Chapter 3
Amazon Elastic Compute Cloud (Amazon EC2) and Amazon Elastic Block Store (Amazon EBS)

THE AWS CERTIFIED SOLUTIONS ARCHITECT ASSOCIATE EXAM OBJECTIVES COVERED IN THIS CHAPTER MAY INCLUDE, BUT ARE NOT LIMITED TO, THE FOLLOWING:

Domain 1.0: Designing highly available, cost-efficient, fault-tolerant, scalable systems

✓**1.1 Identify and recognize cloud architecture considerations, such as fundamental components and effective designs.**

Content may include the following:

- How to design cloud services
- Planning and design
- Monitoring and logging

Domain 2.0: Implementation/Deployment

✓**2.1 Identify the appropriate techniques and methods using Amazon EC2, Amazon Simple Storage Service (Amazon S3), AWS Elastic Beanstalk, AWS CloudFormation, AWS OpsWorks, Amazon Virtual Private Cloud (Amazon VPC), and AWS Identity and Access Management (IAM) to code and implement a cloud solution.**

Content may include the following:

- Configure an Amazon Machine Image (AMI)
- Configure services to support compliance requirements in the cloud
- Launch instances across the AWS global infrastructure

Domain 3.0: Data Security

✓**3.2 Recognize critical disaster recovery techniques and their implementation.**

Content may include the following:

- Disaster recovery
- Amazon EB

Introduction

In this chapter, you learn how Amazon Elastic Compute Cloud (Amazon EC2) and Amazon Elastic Block Store (Amazon EBS) provide the basic elements of compute and block-level storage to run your workloads on AWS. It focuses on key topics you need to understand for the exam, including:

- How instance types and Amazon Machine Images (AMIs) define the capabilities of instances you launch on the cloud
- How to securely access your instances running on the cloud
- How to protect your instances with virtual firewalls called security groups
- How to have your instances configure themselves for unattended launch
- How to monitor and manage your instances on the cloud
- How to change the capabilities of an existing instance
- The payment options available for the best mix of affordability and flexibility
- How tenancy options and placement groups provide options to optimize compliance and performance
- How instance stores differ from Amazon EBS volumes and when they are effective
- What types of volumes are available through Amazon EBS
- How to protect your data on Amazon EBS

Amazon Elastic Compute Cloud (Amazon EC2)

Amazon EC2 is AWS primary web service that provides resizable compute capacity in the cloud.

Compute Basics

Compute refers to the amount of computational power required to fulfill your workload. If your workload is very small, such as a website that receives few visitors, then your compute needs are very small. A large workload, such as screening ten million compounds against a common cancer target, might require a great deal of compute. The amount of compute you need might change drastically over time.

Amazon EC2 allows you to acquire compute through the launching of virtual servers called *instances*. When you launch an instance, you can make use of the compute as you wish, just as you would with an on-premises server. Because you are paying for the computing power of the instance, you are charged per hour while the instance is running. When you stop the instance, you are no longer charged.

There are two concepts that are key to launching instances on AWS: (1) the amount of virtual hardware dedicated to the instance and (2) the software loaded on the instance. These two dimensions of new instances are controlled, respectively, by the instance type and the AMI.

Instance Types

The instance type defines the virtual hardware supporting an Amazon EC2 instance. There are dozens of instance types available, varying in the following dimensions:

- Virtual CPUs (vCPUs)
- Memory
- Storage (size and type)
- Network performance

Instance types are grouped into families based on the ratio of these values to each other. For instance, the m4 family provides a balance of compute, memory, and network resources, and it is a good choice for many applications. Within each family there are several choices that scale up linearly in size. Figure 3.1 shows the four instance sizes in the m4 family. Note that the ratio of vCPUs to memory is constant as the sizes scale linearly. The hourly price for each size scales linearly as well. For example, an m4.xlarge instance costs twice as much as the m4.large instance.

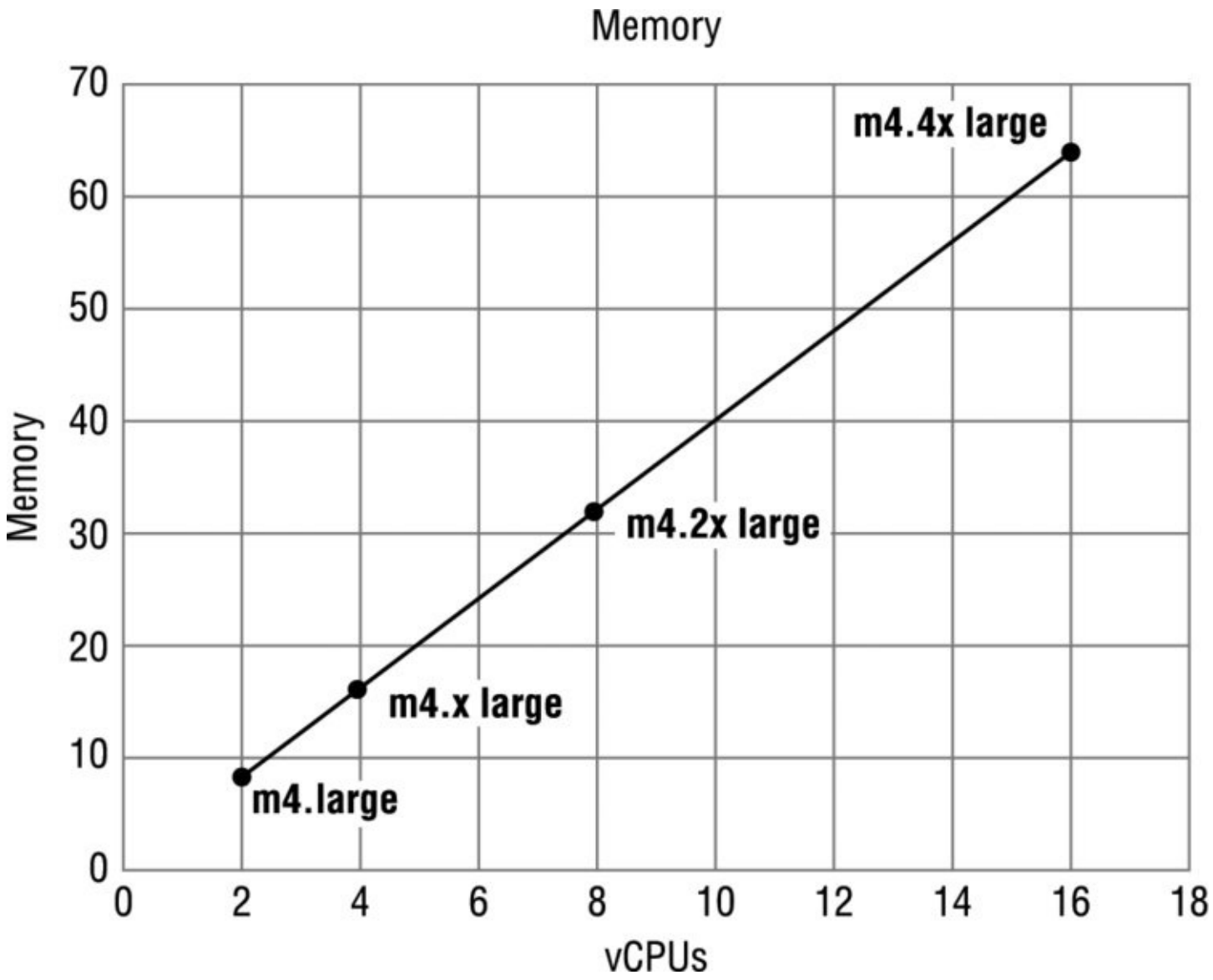

FIGURE 3.1 Memory and vCPUs for the m4 instance family

Different instance type families tilt the ratio to accommodate different types of workloads, but they all exhibit this linear scale up behavior within the family. Table 3.1 lists some of the families available.

TABLE 3.1 Sample Instance Type Families

Family	
c4	***Compute optimized***—For workloads requiring significant processing
r3	***Memory optimized***—For memory-intensive workloads
i2	***Storage optimized***—For workloads requiring high amounts of fast SSD storage
g2	***GPU-based instances***—Intended for graphics and general-purpose GPU compute workloads

In response to customer demand and to take advantage of new processor technology, AWS occasionally introduces new instance families. Check the AWS website for the current list.

Another variable to consider when choosing an instance type is network performance. For most instance types, AWS publishes a relative measure of network performance: low, moderate, or high. Some instance types specify a network performance of 10 Gbps. The

network performance increases within a family as the instance type grows.

For workloads requiring greater network performance, many instance types support *enhanced networking*. Enhanced networking reduces the impact of virtualization on network performance by enabling a capability called Single Root I/O Virtualization (SR-IOV). This results in more Packets Per Second (PPS), lower latency, and less jitter. At the time of this writing, there are instance types that support enhanced networking in the C3, C4, D2, I2, M4, and R3 families (consult the AWS documentation for a current list). Enabling enhanced networking on an instance involves ensuring the correct drivers are installed and modifying an instance attribute. Enhanced networking is available only for instances launched in an Amazon Virtual Private Cloud (Amazon VPC), which is discussed in Chapter 4, "Amazon Virtual Private Cloud (Amazon VPC)."

Amazon Machine Images (AMIs)

The *Amazon Machine Image (AMI)* defines the initial software that will be on an instance when it is launched. An AMI defines every aspect of the software state at instance launch, including:

- The Operating System (OS) and its configuration
- The initial state of any patches
- Application or system software

All AMIs are based on x86 OSs, either Linux or Windows.

There are four sources of AMIs:

- ***Published by AWS***—AWS publishes AMIs with versions of many different OSs, both Linux and Windows. These include multiple distributions of Linux (including Ubuntu, Red Hat, and Amazon's own distribution) and Windows 2008 and Windows 2012. Launching an instance based on one of these AMIs will result in the default OS settings, similar to installing an OS from the standard OS ISO image. As with any OS installation, you should immediately apply all appropriate patches upon launch.
- ***The AWS Marketplace***—AWS Marketplace is an online store that helps customers find, buy, and immediately start using the software and services that run on Amazon EC2. Many AWS partners have made their software available in the AWS Marketplace. This provides two benefits: the customer does not need to install the software, and the license agreement is appropriate for the cloud. Instances launched from an AWS Marketplace AMI incur the standard hourly cost of the instance type plus an additional per-hour charge for the additional software (some open-source AWS Marketplace packages have no additional software charge).
- ***Generated from Existing Instances***—An AMI can be created from an existing Amazon EC2 instance. This is a very common source of AMIs. Customers launch an instance from a published AMI, and then the instance is configured to meet all the customer's corporate standards for updates, management, security, and so on. An AMI is then generated from the configured instance and used to generate all instances of that OS. In this way, all new instances follow the corporate standard and it is more difficult for individual projects to launch non-conforming instances.

- ***Uploaded Virtual Servers***—Using AWS VM Import/Export service, customers can create images from various virtualization formats, including raw, VHD, VMDK, and OVA. The current list of supported OSs (Linux and Windows) can be found in the AWS documentation. It is incumbent on the customers to remain compliant with the licensing terms of their OS vendor.

Securely Using an Instance

Once launched, instances can be managed over the Internet. AWS has several services and features to ensure that this management can be done simply and securely.

Addressing an Instance

There are several ways that an instance may be addressed over the web upon creation:

- ***Public Domain Name System (DNS) Name***—When you launch an instance, AWS creates a DNS name that can be used to access the instance. This DNS name is generated automatically and cannot be specified by the customer. The name can be found in the Description tab of the AWS Management Console or via the Command Line Interface (CLI) or Application Programming Interface (API). This DNS name persists only while the instance is running and cannot be transferred to another instance.
- ***Public IP***—A launched instance may also have a public IP address assigned. This IP address is assigned from the addresses reserved by AWS and cannot be specified. This IP address is unique on the Internet, persists only while the instance is running, and cannot be transferred to another instance.
- ***Elastic IP***—An elastic IP address is an address unique on the Internet that you reserve independently and associate with an Amazon EC2 instance. While similar to a public IP, there are some key differences. This IP address persists until the customer releases it and is not tied to the lifetime or state of an individual instance. Because it can be transferred to a replacement instance in the event of an instance failure, it is a public address that can be shared externally without coupling clients to a particular instance.

 Private IP addresses and Elastic Network Interfaces (ENIs) are additional methods of addressing instances that are available in the context of an Amazon VPC. These are discussed in Chapter 4.

Initial Access

Amazon EC2 uses public-key cryptography to encrypt and decrypt login information. Public-key cryptography uses a public key to encrypt a piece of data and an associated private key to decrypt the data. These two keys together are called a *key pair*. Key pairs can be created through the AWS Management Console, CLI, or API, or customers can upload their own key pairs. AWS stores the public key, and the private key is kept by the customer. The private key is essential to acquiring secure access to an instance for the first time.

 Store your private keys securely. When Amazon EC2 launches a Linux instance, the public key is stored in the `~/.ssh/authorized_keys` file on the instance and an initial user is created. The initial user can vary depending on the OS. For example, the Amazon Linux distribution initial user is ec2-user. Initial access to the instance is obtained by using the ec2-user and the private key to log in via SSH. At this point, you can configure other users and enroll in a directory such as LDAP.

When launching a Windows instance, Amazon EC2 generates a random password for the local administrator account and encrypts the password using the public key. Initial access to the instance is obtained by decrypting the password with the private key, either in the console or through the API. The decrypted password can be used to log in to the instance with the local administrator account via RDP. At this point, you can create other local users and/or connect to an Active Directory domain.

 It is a best practice to change the initial local administrator password.

Virtual Firewall Protection

AWS allows you to control traffic in and out of your instances through virtual firewalls called *security groups*. Security groups allow you to control traffic based on port, protocol, and source/destination. Security groups have different capabilities depending on whether they are associated with an Amazon VPC or Amazon EC2-Classic. Table 3.2 compares these different capabilities (Amazon VPC is discussed in Chapter 4).

TABLE 3.2 Different Security Groups

Type of Security Group	Capabilities
EC2-Classic Security Groups	Control outgoing instance traffic
VPC Security Groups	Control outgoing and incoming instance traffic

Security groups are associated with instances when they are launched. Every instance must have at least one security group but can have more.

A security group is default deny; that is, it does not allow any traffic that is not explicitly allowed by a security group rule. A rule is defined by the three attributes in Table 3.3. When an instance is associated with multiple security groups, the rules are aggregated and all traffic allowed by each of the individual groups is allowed. For example, if security group A allows RDP traffic from 72.58.0.0/16 and security group B allows HTTP and HTTPS traffic from 0.0.0.0/0 and your instance is associated with both groups, then both the RDP and HTTP/S traffic will be allowed in to your instance.

TABLE 3.3 Security Group Rule Attributes

Attribute	Meaning
Port	The port number affected by this rule. For instance, port 80 for HTTP traffic.
Protocol	The communications standard for the traffic affected by this rule.
Source/Destination	Identifies the other end of the communication, the source for incoming traffic rules, or the destination for outgoing traffic rules. The source/destination can be defined in two ways: ***CIDR block***—An x.x.x.x/x style definition that defines a specific range of IP addresses. ***Security group***—Includes any instance that is associated with the given security group. This helps prevent coupling security group rules with specific IP addresses.

A security group is a stateful firewall; that is, an outgoing message is remembered so that the response is allowed through the security group without an explicit inbound rule being required.

Security groups are applied at the instance level, as opposed to a traditional on-premises firewall that protects at the perimeter. The effect of this is that instead of having to breach a single perimeter to access all the instances in your security group, an attacker would have to breach the security group repeatedly for each individual instance.

The Lifecycle of Instances

Amazon EC2 has several features and services that facilitate the management of Amazon EC2 instances over their entire lifecycle.

Launching

There are several additional services that are useful when launching new Amazon EC2 instances.

Bootstrapping A great benefit of the cloud is the ability to script virtual hardware management in a manner that is not possible with on-premises hardware. In order to realize the value of this, there has to be some way to configure instances and install applications programmatically when an instance is launched. The process of providing code to be run on an instance at launch is called *bootstrapping*.

One of the parameters when an instance is launched is a string value called *UserData*. This string is passed to the operating system to be executed as part of the launch process the first time the instance is booted. On Linux instances this can be shell script, and on Windows instances this can be a batch style script or a PowerShell script. The script can perform tasks such as:

- Applying patches and updates to the OS
- Enrolling in a directory service
- Installing application software

- Copying a longer script or program from storage to be run on the instance
- Installing Chef or Puppet and assigning the instance a role so the configuration management software can configure the instance

 UserData is stored with the instance and is not encrypted, so it is important to not include any secrets such as passwords or keys in the UserData.

VM Import/Export In addition to importing virtual instances as AMIs, VM Import/Export enables you to easily import Virtual Machines (VMs) from your existing environment as an Amazon EC2 instance and export them back to your on-premises environment. You can only export previously imported Amazon EC2 instances. Instances launched within AWS from AMIs cannot be exported.

Instance Metadata *Instance metadata* is data about your instance that you can use to configure or manage the running instance. This is unique in that it is a mechanism to obtain AWS properties of the instance from within the OS without making a call to the AWS API. An HTTP call to `http://169.254.169.254/latest/meta-data/` will return the top node of the instance metadata tree. Instance metadata includes a wide variety of attributes, including:

- The associated security groups
- The instance ID
- The instance type
- The AMI used to launch the instance

This only begins to scratch the surface of the information available in the metadata. Consult the AWS documentation for a full list.

Managing Instances

When the number of instances in your account starts to climb, it can become difficult to keep track of them. Tags can help you manage not just your Amazon EC2 instances, but also many of your AWS Cloud services. Tags are key/value pairs you can associate with your instance or other service. Tags can be used to identify attributes of an instance like project, environment (dev, test, and so on), billable department, and so forth. You can apply up to 10 tags per instance. Table 3.4 shows some tag suggestions.

TABLE 3.4 Sample Tags

Key	Value
Project	TimeEntry
Environment	Production
BillingCode	4004

Monitoring Instances

AWS offers a service called Amazon CloudWatch that provides monitoring and alerting for

Amazon EC2 instances, and also other AWS infrastructure. Amazon CloudWatch is discussed in detail in Chapter 5, "Elastic Load Balancing, Amazon CloudWatch, and Auto Scaling."

Modifying an Instance

There are several aspects of an instance that can be modified after launch.

Instance Type The ability to change the instance type of an instance contributes greatly to the agility of running workloads in the cloud. Instead of committing to a certain hardware configuration months before a workload is launched, the workload can be launched using a best estimate for the instance type. If the compute needs prove to be higher or lower than expected, the instances can be changed to a different size more appropriate to the workload.

Instances can be resized using the AWS Management Console, CLI, or API. To resize an instance, set the state to Stopped. Choose the "Change Instance Type" function in the tool of your choice (the instance type is listed as an Instance Setting in the console and an Instance Attribute in the CLI) and select the desired instance type. Restart the instance and the process is complete.

Security Groups If an instance is running in an Amazon VPC (discussed in Chapter 4), you can change which security groups are associated with an instance while the instance is running. For instances outside of an Amazon VPC (called EC2-Classic), the association of the security groups cannot be changed after launch.

Termination Protection

When an Amazon EC2 instance is no longer needed, the state can be set to Terminated and the instance will be shut down and removed from the AWS infrastructure. In order to prevent termination via the AWS Management Console, CLI, or API, *termination protection* can be enabled for an instance. While enabled, calls to terminate the instance will fail until termination protection is disabled. This helps to prevent accidental termination through human error.

Note that this just protects from termination calls from the AWS Management Console, CLI, or API. It does not prevent termination triggered by an OS shutdown command, termination from an Auto Scaling group (discussed in Chapter 5), or termination of a Spot Instance due to Spot price changes (discussed in the next section).

Options

There are several additional options available in Amazon EC2 to improve cost optimization, security, and performance that are important to know for the exam.

Pricing Options

You are charged for Amazon EC2 instances for each hour that they are in a running state, but the amount you are charged per hour can vary based on three pricing options: On-Demand Instances, Reserved Instances, and Spot Instances.

On-Demand Instances The price per hour for each instance type published on the AWS website represents the price for *On-Demand Instances*. This is the most flexible pricing option, as it requires no up-front commitment, and the customer has control over when the

instance is launched and when it is terminated. It is the least cost effective of the three pricing options per compute hour, but its flexibility allows customers to save by provisioning a variable level of compute for unpredictable workloads.

Reserved Instances The *Reserved Instance* pricing option enables customers to make capacity reservations for predictable workloads. By using Reserved Instances for these workloads, customers can save up to 75 percent over the on-demand hourly rate. When purchasing a reservation, the customer specifies the instance type and Availability Zone for that Reserved Instance and achieves a lower effective hourly price for that instance for the duration of the reservation. An additional benefit is that capacity in the AWS data centers is reserved for that customer. There are two factors that determine the cost of the reservation: the term commitment and the payment option.

The *term commitment* is the duration of the reservation and can be either one or three years. The longer the commitment, the bigger the discount.

There are three different payment options for Reserved Instances:

- ***All Upfront***—Pay for the entire reservation up front. There is no monthly charge for the customer during the term.
- ***Partial Upfront***—Pay a portion of the reservation charge up front and the rest in monthly installments for the duration of the term.
- ***No Upfront***—Pay the entire reservation charge in monthly installments for the duration of the term.

The amount of the discount is greater the more the customer pays up front.

For example, let's look at the effect of an all upfront, three-year reservation on the effective hourly cost of an m4.2xlarge instance. The cost of running one instance continuously for three years (or 26,280 hours) at both pricing options is shown in Table 3.5.

TABLE 3.5 Reserved Instance Pricing Example

Pricing Option	Effective Hourly Cost	Total Three-Year Cost
On-Demand	$0.479/hour	$0.479/hour* 26280 hours = $12588.12
Three-Year All Upfront Reservation	$4694/26280 hours = $0.1786/hour	$4694
Savings		63%


This example uses the published prices at the time of this writing. AWS has lowered prices many times to date, so check the AWS website for current pricing information.

When your computing needs change, you can modify your Reserved Instances and continue to benefit from your capacity reservation. Modification does not change the remaining term of your Reserved Instances; their end dates remain the same. There is no fee, and you do not

receive any new bills or invoices. Modification is separate from purchasing and does not affect how you use, purchase, or sell Reserved Instances. You can modify your whole reservation, or just a subset, in one or more of the following ways:

- Switch Availability Zones within the same region.
- Change between EC2-VPC and EC2-Classic.
- Change the instance type within the same instance family (Linux instances only).

Spot Instances For workloads that are not time critical and are tolerant of interruption, *Spot Instances* offer the greatest discount. With Spot Instances, customers specify the price they are willing to pay for a certain instance type. When the customer's bid price is above the current Spot price, the customer will receive the requested instance(s). These instances will operate like all other Amazon EC2 instances, and the customer will only pay the Spot price for the hours that instance(s) run. The instances will run until:

- The customer terminates them.
- The Spot price goes above the customer's bid price.
- There is not enough unused capacity to meet the demand for Spot Instances.

If Amazon EC2 needs to terminate a Spot Instance, the instance will receive a termination notice providing a two-minute warning prior to Amazon EC2 terminating the instance.

Because of the possibility of interruption, Spot Instances should only be used for workloads tolerant of interruption. This could include analytics, financial modeling, big data, media encoding, scientific computing, and testing.

Architectures with Different Pricing Models For the exam, it's important to know how to take advantage of the different pricing models to create a cost-efficient architecture. Such an architecture may include different pricing models within the same workload. For instance, a website that averages 5,000 visits a day, but ramps up to 20,000 visits a day during periodic peaks, may purchase two Reserved Instances to handle the average traffic, but depend on On-Demand Instances to fulfill compute needs during the peak times. Figure 3.2 shows such an architecture.

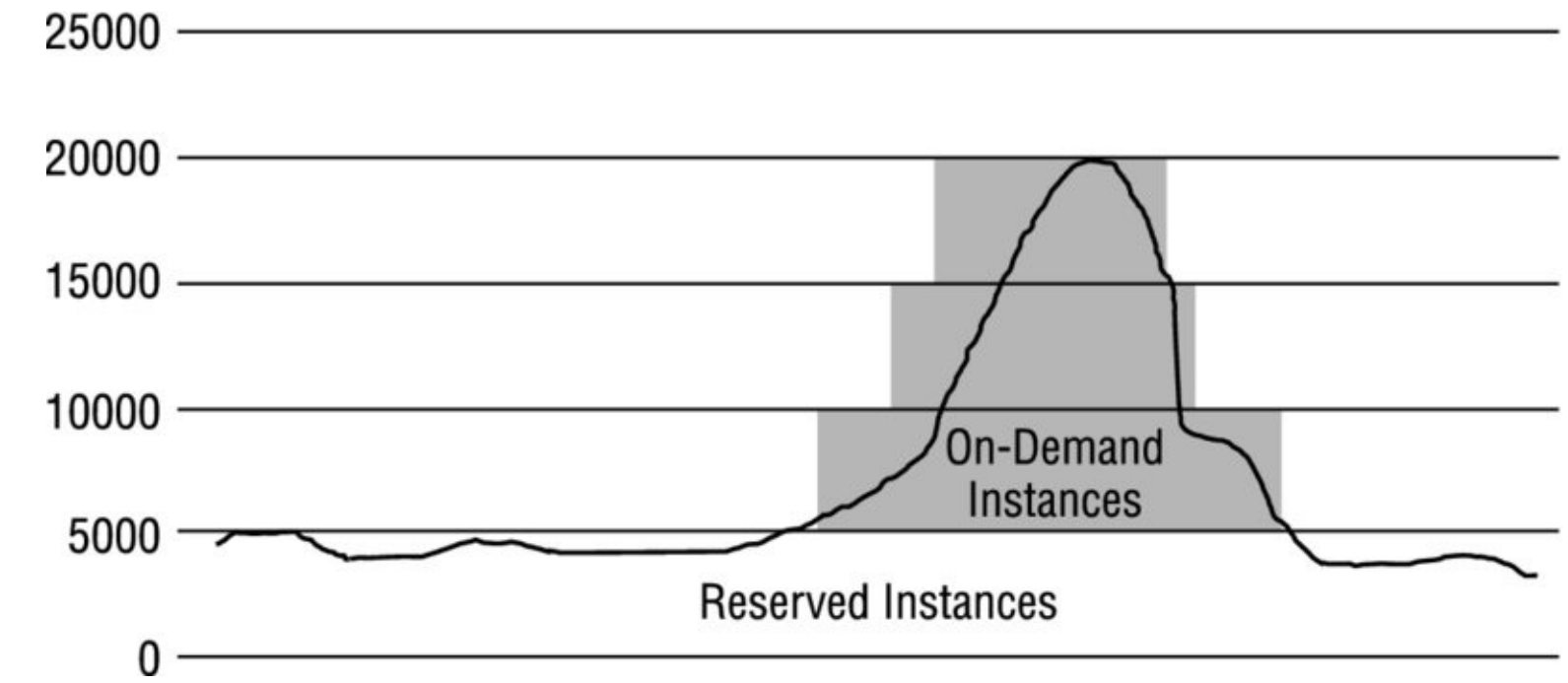

FIGURE 3.2 A workload using a mix of On-Demand and Reserved Instances

Tenancy Options

There are several *tenancy* options for Amazon EC2 instances that can help customers achieve security and compliance goals.

Shared Tenancy Shared tenancy is the default tenancy model for all Amazon EC2 instances, regardless of instance type, pricing model, and so forth. Shared tenancy means that a single host machine may house instances from different customers. As AWS does not use overprovisioning and fully isolates instances from other instances on the same host, this is a secure tenancy model.

Dedicated Instances Dedicated Instances run on hardware that's dedicated to a single customer. As a customer runs more Dedicated Instances, more underlying hardware may be dedicated to their account. Other instances in the account (those not designated as dedicated) will run on shared tenancy and will be isolated at the hardware level from the Dedicated Instances in the account.

Dedicated Host An Amazon EC2 Dedicated Host is a physical server with Amazon EC2 instance capacity fully dedicated to a single customer's use. Dedicated Hosts can help you address licensing requirements and reduce costs by allowing you to use your existing server-bound software licenses. The customer has complete control over which specific host runs an instance at launch. This differs from Dedicated Instances in that a Dedicated Instance can launch on any hardware that has been dedicated to the account.

Placement Groups

A *placement group* is a logical grouping of instances within a single Availability Zone. Placement groups enable applications to participate in a low-latency, 10 Gbps network. Placement groups are recommended for applications that benefit from low network latency, high network throughput, or both. Remember that this represents network connectivity between instances. To fully use this network performance for your placement group, choose an instance type that supports enhanced networking and 10 Gbps network performance.

Instance Stores

An *instance store* (sometimes referred to as *ephemeral storage*) provides temporary block-level storage for your instance. This storage is located on disks that are physically attached to the host computer. An instance store is ideal for temporary storage of information that changes frequently, such as buffers, caches, scratch data, and other temporary content, or for data that is replicated across a fleet of instances, such as a load-balanced pool of web servers.

The size and type of instance stores available with an Amazon EC2 instance depend on the instance type. At this writing, storage available with various instance types ranges from no instance stores up to 24 2 TB instance stores. The instance type also determines the type of hardware for the instance store volumes. While some provide Hard Disk Drive (HDD) instance stores, other instance types use Solid State Drives (SSDs) to deliver very high random I/O performance.

Instance stores are included in the cost of an Amazon EC2 instance, so they are a very cost-effective solution for appropriate workloads. The key aspect of instance stores is that they are temporary. Data in the instance store is lost when:

- The underlying disk drive fails.
- The instance stops (the data will persist if an instance reboots).
- The instance terminates.

Therefore, do not rely on instance stores for valuable, long-term data. Instead, build a degree of redundancy via RAID or use a file system that supports redundancy and fault tolerance such as Hadoop's HDFS. Back up the data to more durable data storage solutions such as Amazon Simple Storage Service (Amazon S3) or Amazon EBS often enough to meet recovery point objectives.

Amazon Elastic Block Store (Amazon EBS)

While instance stores are an economical way to fulfill appropriate workloads, their limited persistence makes them ill-suited for many other workloads. For workloads requiring more durable block storage, Amazon provides Amazon EBS.

Elastic Block Store Basics

Amazon EBS provides persistent block-level storage volumes for use with Amazon EC2 instances. Each Amazon EBS volume is automatically replicated within its Availability Zone to protect you from component failure, offering high availability and durability. Amazon EBS volumes are available in a variety of types that differ in performance characteristics and price. Multiple Amazon EBS volumes can be attached to a single Amazon EC2 instance, although a volume can only be attached to a single instance at a time.

Types of Amazon EBS Volumes

Amazon EBS volumes are available in several different types. Types vary in areas such as underlying hardware, performance, and cost. It is important to know the properties of the different types so you can specify the most cost-efficient type that meets a workload's performance demands on the exam.

Magnetic Volumes

Magnetic volumes have the lowest performance characteristics of all Amazon EBS volume types. As such, they cost the lowest per gigabyte. They are an excellent, cost-effective solution for appropriate workloads.

A magnetic Amazon EBS volume can range in size from 1 GB to 1 TB and will average 100 *IOPS*, but has the ability to burst to hundreds of IOPS. They are best suited for:

- Workloads where data is accessed infrequently
- Sequential reads
- Situations where low-cost storage is a requirement

Magnetic volumes are billed based on the amount of data space provisioned, regardless of how much data you actually store on the volume.

General-Purpose SSD

General-purpose SSD volumes offer cost-effective storage that is ideal for a broad range of workloads. They deliver strong performance at a moderate price point that is suitable for a wide range of workloads.

A general-purpose SSD volume can range in size from 1 GB to 16 TB and provides a baseline performance of three IOPS per gigabyte provisioned, capping at 10,000 IOPS. For instance, if you provision a 1 TB volume, you can expect a baseline performance of 3,000 IOPS. A 5 TB volume will not provide a 15,000 IOPS baseline, as it would hit the cap at 10,000 IOPS.

General-purpose SSD volumes under 1 TB also feature the ability to burst to up to 3,000

IOPS for extended periods of time. For instance, if you have a 500 GB volume you can expect a baseline of 1,500 IOPS. Whenever you are not using these IOPS, they are accumulated as I/O credits. When your volume then has heavy traffic, it will use the I/O credits at a rate of up to 3,000 IOPS until they are depleted. At that point, your performance reverts to 1,500 IOPS. At 1 TB, the baseline performance of the volume is already at 3,000 IOPS, so bursting behavior does not apply.

General-purpose SSD volumes are billed based on the amount of data space provisioned, regardless of how much data you actually store on the volume. They are suited for a wide range of workloads where the very highest disk performance is not critical, such as:

- System boot volumes
- Small- to medium-sized databases
- Development and test environments

Provisioned IOPS SSD

Provisioned IOPS SSD volumes are designed to meet the needs of I/O-intensive workloads, particularly database workloads that are sensitive to storage performance and consistency in random access I/O throughput. While they are the most expensive Amazon EBS volume type per gigabyte, they provide the highest performance of any Amazon EBS volume type in a predictable manner.

A Provisioned IOPS SSD volume can range in size from 4 GB to 16 TB. When you provision a Provisioned IOPS SSD volume, you specify not just the size, but also the desired number of IOPS, up to the lower of the maximum of 30 times the number of GB of the volume, or 20,000 IOPS. You can stripe multiple volumes together in a RAID 0 configuration for larger size and greater performance. Amazon EBS delivers within 10 percent of the provisioned IOPS performance 99.9 percent of the time over a given year.

Pricing is based on the size of the volume and the amount of IOPS reserved. The cost per gigabyte is slightly more than that of general-purpose SSD volumes and is applied based on the size of the volume, not the amount of the volume used to store data. An additional monthly fee is applied based on the number of IOPS provisioned, whether they are consumed or not.

Provisioned IOPS SSD volumes provide predictable, high performance and are well suited for:

- Critical business applications that require sustained IOPS performance
- Large database workloads

Table 3.6 compares these Amazon EBS volume types.

TABLE 3.6 EBS Volume Type Comparison

Characteristic	General-Purpose SSD	Provisioned IOPS SSD	Magnetic
Use cases	• System boot volumes • Virtual desktops • Small-to-medium sized databases • Development and test environments	• Critical business applications that require sustained IOPS performance or more than 10,000 IOPS or 160MB of throughput per volume • Large database workloads	• Cold workloads where data is infrequently accessed • Scenarios where the lowest storage cost is important
Volume size	1 GiB–16TiB	4 GiB–16TiB	1 GiB–1TiB
Maximum throughput	160MB	320MB	40–90MB
IOPS performance	Baseline performance of 3 IOPS/GiB (up to 10,000 IOPS) with the ability to burst to 3,000 IOPS for volumes under 1,000 GiB	Consistently performs at provisioned level, up to 20,000 IOPS maximum	Averages 100 IOPS, with the ability to burst to hundreds of IOPS

At the time of this writing, AWS released two new HDD volume types: Throughput-Optimized HDD and Cold HDD. Over time, it is expected that these new types will eclipse the current magnetic volume type, fulfilling the needs of any workload requiring HDD performance.

Throughput-Optimized HDD volumes are low-cost HDD volumes designed for frequent-access, throughput-intensive workloads such as big data, data warehouses, and log processing. Volumes can be up to 16 TB with a maximum IOPS of 500 and maximum throughput of 500 MB/s. These volumes are significantly less expensive than general-purpose SSD volumes.

Cold HDD volumes are designed for less frequently accessed workloads, such as colder data requiring fewer scans per day. Volumes can be up to 16 TB with a maximum IOPS of 250 and maximum throughput of 250 MB/s. These volumes are significantly less expensive than Throughput-Optimized HDD volumes.

Amazon EBS-Optimized Instances

When using any volume type other than magnetic and Amazon EBS I/O is of consequence, it is important to use Amazon EBS-optimized instances to ensure that the Amazon EC2 instance is prepared to take advantage of the I/O of the Amazon EBS volume. An Amazon

EBS-optimized instance uses an optimized configuration stack and provides additional, dedicated capacity for Amazon EBS I/O. This optimization provides the best performance for your Amazon EBS volumes by minimizing contention between Amazon EBS I/O and other traffic from your instance. When you select Amazon EBS-optimized for an instance, you pay an additional hourly charge for that instance. Check the AWS documentation to confirm which instance types are available as Amazon EBS-optimized instance.

Protecting Data

Over the lifecycle of an Amazon EBS volume, there are several practices and services that you should know about when taking the exam.

Backup/Recovery (Snapshots)

You can back up the data on your Amazon EBS volumes, regardless of volume type, by taking point-in-time snapshots. Snapshots are incremental backups, which means that only the blocks on the device that have changed since your most recent snapshot are saved.

Taking Snapshots You can take snapshots in many ways:

- Through the AWS Management Console
- Through the CLI
- Through the API
- By setting up a schedule of regular snapshots

Data for the snapshot is stored using Amazon S3 technology. The action of taking a snapshot is free. You pay only the storage costs for the snapshot data.

When you request a snapshot, the point-in-time snapshot is created immediately and the volume may continue to be used, but the snapshot may remain in pending status until all the modified blocks have been transferred to Amazon S3.

It's important to know that while snapshots are stored using Amazon S3 technology, they are stored in AWS-controlled storage and not in your account's Amazon S3 buckets. This means you cannot manipulate them like other Amazon S3 objects. Rather, you must use the Amazon EBS snapshot features to manage them. Snapshots are constrained to the region in which they are created, meaning you can use them to create new volumes only in the same region. If you need to restore a snapshot in a different region, you can copy a snapshot to another region.

Creating a Volume from a Snapshot To use a snapshot, you create a new Amazon EBS volume from the snapshot. When you do this, the volume is created immediately but the data is loaded lazily. This means that the volume can be accessed upon creation, and if the data being requested has not yet been restored, it will be restored upon first request. Because of this, it is a best practice to initialize a volume created from a snapshot by accessing all the blocks in the volume.

Snapshots can also be used to increase the size of an Amazon EBS volume. To increase the size of an Amazon EBS volume, take a snapshot of the volume, then create a new volume of the desired size from the snapshot. Replace the original volume with the new volume.

Recovering Volumes

Because Amazon EBS volumes persist beyond the lifetime of an instance, it is possible to recover data if an instance fails. If an Amazon EBS-backed instance fails and there is data on the boot drive, it is relatively straightforward to detach the volume from the instance. Unless the `DeleteOnTermination` flag for the volume has been set to false, the volume should be detached before the instance is terminated. The volume can then be attached as a data volume to another instance and the data read and recovered.

Encryption Options

Many workloads have requirements that data be encrypted at rest, either because of compliance regulations or internal corporate standards. Amazon EBS offers native encryption on all volume types.

When you launch an encrypted Amazon EBS volume, Amazon uses the *AWS Key Management Service (KMS)* to handle key management. A new master key will be created unless you select a master key that you created separately in the service. Your data and associated keys are encrypted using the industry-standard AES-256 algorithm. The encryption occurs on the servers that host Amazon EC2 instances, so the data is actually encrypted in transit between the host and the storage media and also on the media. (Consult the AWS documentation for a list of instance types that support Amazon EBS encryption.) Encryption is transparent, so all data access is the same as unencrypted volumes, and you can expect the same IOPS performance on encrypted volumes as you would with unencrypted volumes, with a minimal effect on latency. Snapshots that are taken from encrypted volumes are automatically encrypted, as are volumes that are created from encrypted snapshots.

Summary

Compute is the amount of computational power required to fulfill your workload. Amazon EC2 is the primary service for providing compute to customers.

The instance type defines the virtual hardware supporting the instance. Available instance types vary in vCPUs, memory, storage, and network performance to address nearly any workload.

An AMI defines the initial software state of the instance, both OS and applications. There are four sources of AMIs: AWS published generic OSs, partner-published AMIs in the AWS Marketplace with software packages preinstalled, customer-generated AMIs from existing Amazon EC2 instances, and uploaded AMIs from virtual servers.

Instances can be addressed by public DNS name, public IP address, or elastic IP address. To access a newly launched Linux instance, use the private half of the key pair to connect to the instance via SSH. To access a newly created Windows instance, use the private half of the key pair to decrypt the randomly initialized local administrator password.

Network traffic in and out of an instance can be controlled by a virtual firewall called a security group. A security group allows rules that block traffic based on direction, port, protocol, and source/destination address.

Bootstrapping allows you to run a script to initialize your instance with OS configurations and applications. This feature allows instances to configure themselves upon launch. Once an instance is launched, you can change its instance type or, for Amazon VPC instances, the security groups with which it is associated.

The three pricing options for instances are On-Demand, Reserved Instance, and Spot. On-Demand has the highest per hour cost, requiring no up-front commitment and giving you complete control over the lifetime of the instance. Reserved Instances require a commitment and provide a reduced overall cost over the lifetime of the reservation. Spot Instances are idle compute capacity that AWS makes available based on bid prices from customers. The savings on the per-hour cost can be significant, but instances can be shut down when the bid price exceeds the customer's current bid.

Instance stores are block storage included with the hourly cost of the instance. The amount and type of storage available varies with the instance type. Instance stores terminate when the associated instance is stopped, so they should only be used for temporary data or in architectures providing redundancy such as Hadoop's HDFS.

Amazon EBS provides durable block storage in several types. Magnetic has the lowest cost per gigabyte and delivers modest performance. General-purpose SSD is cost-effective storage that can provide up to 10,000 IOPS. Provisioned IOPS SSD has the highest cost per gigabyte and is well suited for I/O-intensive workloads sensitive to storage performance. Snapshots are incremental backups of Amazon EBS volumes stored in Amazon S3. Amazon EBS volumes can be encrypted.

Exam Essentials

Know the basics of launching an Amazon ec2 instance. To launch an instance, you must specify an AMI, which defines the software on the instance at launch, and an instance type, which defines the virtual hardware supporting the instance (memory, vCPUs, and so on).

Know what architectures are suited for what Amazon ec2 pricing options. Spot Instances are best suited for workloads that can accommodate interruption. Reserved Instances are best for consistent, long-term compute needs. On-Demand Instances provide flexible compute to respond to scaling needs.

Know how to combine multiple pricing options that result in cost optimization and scalability. On-Demand Instances can be used to scale up a web application running on Reserved Instances in response to a temporary traffic spike. For a workload with several Reserved Instances reading from a queue, it's possible to use Spot Instances to alleviate heavy traffic in a cost-effective way. These are just two of countless examples where a workload may use different pricing options.

Know the benefits of enhanced networking. Enhanced networking enables you to get significantly higher PPS performance, lower network jitter, and lower latencies.

Know the capabilities of vm import/export. VM Import/Export allows you to import existing VMs to AWS as Amazon EC2 instances or AMIs. Amazon EC2 instances that were imported through VM Import/Export can also be exported back to a virtual environment.

Know the methods for accessing an instance over the internet. You can access an Amazon EC2 instance over the web via public IP address, elastic IP address, or public DNS name. There are additional ways to access an instance within an Amazon VPC, including private IP addresses and ENIs.

Know the lifetime of an instance store. Data on an instance store is lost when the instance is stopped or terminated. Instance store data survives an OS reboot.

Know the properties of the Amazon EC2 pricing options. On-Demand Instances require no up-front commitment, can be launched any time, and are billed by the hour. Reserved Instances require an up-front commitment and vary in cost depending on whether they are paid all up front, partially up front, or not up front. Spot Instances are launched when your bid price exceeds the current spot price. Spot Instances will run until the spot price exceeds your bid price, in which case the instance will get a two-minute warning and terminate.

Know what determines network performance. Every instance type is rated for low, moderate, high, or 10 Gbps network performance, with larger instance types generally having higher ratings. Additionally, some instance types offer enhanced networking, which provides additional improvement in network performance.

Know what instance metadata is and how it's obtained. Metadata is information about an Amazon EC2 instance, such as instance ID, instance type, and security groups, that is available from within the instance. It can be obtained through an HTTP call to a specific IP address.

Know how security groups protect instances. Security groups are virtual firewalls controlling traffic in and out of your Amazon EC2 instances. They are deny by default, and you can allow traffic by adding rules specifying traffic direction, port, protocol, and destination address (via Classless Inter-Domain Routing [CIDR] block). They are applied at the instance level, meaning that traffic between instances in the same security group must adhere to the rules of that security group. They are stateful, meaning that an outgoing rule will allow the response without a correlating incoming rule.

Know how to interpret the effect of security groups. When an instance is a member of multiple security groups, the effect is a union of all the rules in all the groups.

Know the different Amazon ebs volume types, their characteristics, and their appropriate workloads. Magnetic volumes provide an average performance of 100 IOPS and can be provisioned up to 1 TB. They are good for cold and infrequently accessed data. General-purpose SSD volumes provide three IOPS/GB up to 10,000 IOPS, with smaller volumes able to burst 3,000 IOPS. They can be provisioned up to 16 TB and are appropriate for dev/test environments, small databases, and so forth. Provisioned IOPS SSD can provide up to 20,000 consistent IOPS for volumes up to 16 TB. They are the best choice for workloads such as large databases executing many transactions.

Know how to encrypt an Amazon ebs volume. Any volume type can be encrypted at launch. Encryption is based on AWS KMS and is transparent to applications on the attached instances.

Understand the concept and process of snapshots. Snapshots provide a point-in-time backup of an Amazon EBS volume and are stored in Amazon S3. Subsequent snapshots are incremental—they only store deltas. When you request a snapshot, the point-in-time snapshot is created immediately and the volume may continue to be used, but the snapshot may remain in pending status until all the modified blocks have been transferred to Amazon S3. Snapshots may be copied between regions.

Know how Amazon ebs-optimized instances affect Amazon ebs performance. In addition to the IOPS that control the performance in and out of the Amazon EBS volume, use Amazon EBS-optimized instances to ensure additional, dedicated capacity for Amazon EBS I/O.

Exercises

For assistance in completing these exercises, refer to these user guides:

- ***Amazon EC2 (Linux)***—http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/concepts.html
- ***Amazon EC2 (Windows)***—http://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/concepts.html
- ***Amazon EBS***—http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html

EXERCISE 3.1

Launch and Connect to a Linux Instance

In this exercise, you will launch a new Linux instance, log in with SSH, and install any security updates.

1. Launch an instance in the Amazon EC2 console.
2. Choose the Amazon Linux AMI.
3. Choose the t2.medium instance type.
4. Launch the instance in either the default VPC or EC2-Classic.
5. Assign the instance a public IP address.
6. Add a tag to the instance of Key: Name, Value: **`Exercise 3.1`**.
7. Create a new security group called **`Cert Book`**.
8. Add a rule to Cert Book allowing SSH access from the IP address of your workstation (`www.WhatsMyIP.org` is a good way to determine your IP address).
9. Launch the instance.
10. When prompted for a key pair, choose a key pair you already have or create a new one and download the private portion.

    Amazon generates a `keyname.pem` file, and you will need a `keyname.ppk` file to connect to the instance via SSH. `Puttygen.exe` is one utility that will create a `.ppk` file from a `.pem` file.
11. SSH into the instance using the public IP address, the user name `ec2-user`, and the `keyname.ppk` file.
12. From the command-line prompt, run `sudo yum update–security -y`.
13. Close the SSH window and terminate the instance.

EXERCISE 3.2

Launch a Windows Instance with Bootstrapping

In this exercise, you will launch a Windows instance and specify a very simple bootstrap script. You will then confirm that the bootstrap script was executed on the instance.

1. Launch an instance in the Amazon EC2 console.
2. Choose the Microsoft Windows Server 2012 Base AMI.
3. Choose the t2.medium instance type.
4. Launch the instance in either the default VPC or EC2-Classic.
5. Assign the instance a public IP address.
6. In the Advanced Details section, enter the following text as UserData:

```
<script>
md c:\temp
</script>
```

7. Add a tag to the instance of Key: Name, Value: ***Exercise 3.2***.
8. Use the Cert Book security group from Exercise 3.1.
9. Launch the instance.
10. Use the key pair from Exercise 3.1.
11. On the Connect Instance UI, decrypt the administrator password and then download the RDP file to attempt to connect to the instance. Your attempt should fail because the Cert Book security group does not allow RDP access.
12. Open the Cert Book security group and add a rule that allows RDP access from your IP address.
13. Attempt to access the instance via RDP again.
14. Once the RDP session is connected, open Windows Explorer and confirm that the `c:\temp` folder has been created.
15. End the RDP session and terminate the instance.

EXERCISE 3.3

Confirm That Instance Stores Are Lost When an Instance Is Stopped

In this exercise, you will observe that the data on an Amazon EC2 instance store is lost when the instance is stopped.

1. Launch an instance in the Amazon Management Console.
2. Choose the Microsoft Windows Server 2012 Base AMI.
3. Choose the m3.medium instance type.
4. Launch the instance in either the default VPC or EC2-Classic.
5. Assign the instance a public IP address.
6. Add a tag to the instance of Key: Name, Value: **`Exercise 3.3`**.
7. Use the Cert Book security group as updated in Exercise 3.2.
8. Launch the instance.
9. Use the key pair from Exercise 3.1.
10. Decrypt the administrator password login to the instance via RDP.
11. Once the RDP session is connected, open Windows Explorer.
12. Create a new folder named **`z:\temp`**.
13. Log out of the RDP session.
14. In the console, set the state of the instance to Stopped.
15. Once the instance is stopped, start it again.
16. Log back into the instance using RDP.
17. Open Windows Explorer and confirm that the `z:\temp` folder is gone.
18. End the RDP session and terminate the instance.

EXERCISE 3.4

Launch a Spot Instance

In this exercise, you will create a Spot Instance.

1. In the Amazon EC2 console, go to the Spot Request page.
2. Look at the pricing history for m3.medium, especially the recent price.
3. Make a note of the most recent price and Availability Zone.
4. Launch an instance in the Amazon EC2 console.
5. Choose the Amazon Linux AMI.
6. Choose the t2.medium instance type.
7. On the Configure Instance page, request a Spot Instance.
8. Launch the instance in either the Default VPC or EC2-Classic. (Note the Default VPC will define the Availability Zone for the instance.)
9. Assign the instance a public IP address.
10. Request a Spot Instance and enter a bid a few cents above the recorded Spot price.
11. Finish launching the instance.
12. Go back to the Spot Request page.

    Watch your request. If your bid was high enough, you should see it change to Active and an instance ID appear.
13. Find the instance on the instances page of the Amazon EC2 console.

    Note the Lifecycle field in the Description that says Spot.
14. Once the instance is running, terminate it.

EXERCISE 3.5

Access Metadata

In this exercise, you will access the instance metadata from the OS.

1. Launch an instance in the Amazon EC2 console.
2. Choose the Amazon Linux AMI.
3. Choose the t2.medium instance type.
4. Launch the instance in either the default VPC or EC2-Classic.
5. Assign the instance a public IP address.
6. Add a tag to the instance of Key: Name, Value: **`Exercise 3.5`**.
7. Use the Cert Book security group.
8. Launch the instance.
9. Use the key pair from Exercise 3.1.
10. Connect the instance via SSH using the public IP address, the user name `ec2-user`, and the `keyname.ppk` file.
11. At the Linux command prompt, retrieve a list of the available metadata by typing:

    curl http://169.254.169.254/latest/meta-data/
12. To see a value, add the name to the end of the URL. For example, to see the security groups, type:

    curl http://169.254.169.254/latest/meta-data/security-groups
13. Try other values as well. Names that end with a / indicate a longer list of sub-values.
14. Close the SSH window and terminate the instance.

EXERCISE 3.6

Create an Amazon EBS Volume and Show That It Remains After the Instance Is Terminated

In this exercise, you will see how an Amazon EBS volume persists beyond the life of an instance.

1. Launch an instance in the Amazon EC2 console.
2. Choose the Amazon Linux AMI.
3. Choose the t2.medium instance type.
4. Launch the instance in either the default VPC or EC2-Classic.
5. Assign the instance a public IP address.
6. Add a second Amazon EBS volume of size 50 GB. Note that the Root Volume is set to Delete on Termination.
7. Add a tag to the instance of Key: Name, Value: **`Exercise 3.6`**.
8. Use the Cert Book security group from earlier exercises.
9. Launch the instance.
10. Find the two Amazon EBS volumes on the Amazon EBS console. Name them both **`Exercise 3.6`**.
11. Terminate the instance.

Notice that the boot drive is destroyed, but the additional Amazon EBS volume remains and now says Available. Do not delete the Available volume.

EXERCISE 3.7

Take a Snapshot and Restore

This exercise guides you through taking a snapshot and restoring it in three different ways.

1. Find the volume you created in Exercise 3.6 in the Amazon EBS console.
2. Take a snapshot of that volume. Name the snapshot **`Exercise 3.7`**.
3. On the snapshot console, wait for the snapshot to be completed. (As the volume was empty, this should be very quick.)
4. On the snapshot page in the AWS Management Console, choose the new snapshot and select Create Volume.
5. Create the volume with all the defaults.
6. Locate the snapshot again and again choose Create Volume, setting the size of the new volume to 100 GB (taking a snapshot and restoring the snapshot to a new, larger volume is how you address the problem of increasing the size of an existing volume). Locate the snapshot again and choose Copy. Copy the snapshot to another region. Make the description **`Exercise 3.7`**.
7. Go to the other region and wait for the snapshot to become available.
8. Create a volume from the snapshot in the new region. This is how you share an Amazon EBS volume between regions; that is, by taking a snapshot and copying the snapshot.
9. Delete all four volumes.

EXERCISE 3.8

Launch an Encrypted Volume

In this exercise, you will launch an Amazon EC2 instance with an encrypted Amazon EBS volume and store some data on it to confirm that the encryption is transparent to the instance itself.

1. Launch an instance in the Amazon EC2 console.
2. Choose the Microsoft Windows Server 2012 Base AMI.
3. Choose the m3.medium instance type.
4. Launch the instance in either the default VPC or EC2-Classic.
5. Assign the instance a public IP address.
6. On the storage page, add a 50 GB encrypted Amazon EBS volume.
7. Add a tag to the instance of Key: Name, Value: **`Exercise 3.8`**.
8. Use the Cert Book security group as updated in Exercise 3.2.
9. Launch the instance.
10. Choose the key pair from Exercise 3.1.
11. Decrypt the administrator password and log in to the instance using RDP.
12. Once the RDP session is connected, open Notepad.
13. Type some random information into Notepad, save it at `d:\testfile.txt`, and then close Notepad.
14. Find `d:\testfile.txt` in Windows Explorer and open it with Notepad. Confirm that the data is not encrypted in Notepad.
15. Log out.
16. Terminate the instance.

EXERCISE 3.9

Detach a Boot Drive and Reattach to Another Instance

In this exercise, you will practice removing an Amazon EBS volume from a stopped drive and attaching to another instance to recover the data.

1. Launch an instance in the Amazon EC2 console.
2. Choose the Microsoft Windows Server 2012 Base AMI.
3. Choose the t2.medium instance type.
4. Launch the instance in either the default VPC or EC2-Classic.
5. Assign the instance a public IP address.

6. Add a tag to the instance of Key: Name, Value: **`Exercise 3.9 Source`**.
7. Use the Cert Book security group from earlier exercises.
8. Launch the instance with the key pair from Exercise 3.1.
9. Launch a second instance in the Amazon EC2 Console.
10. Choose the Microsoft Windows Server 2012 Base AMI.
11. Choose the t2.medium instance type.
12. Launch the instance in either the default VPC or EC2-Classic.
13. Assign the instance a public IP address.
14. Add a tag to the instance of Key: Name, Value: **`Exercise 3.9 Destination`**.
15. Use the Cert Book security group from earlier exercises.
16. Launch the instance with the key pair you used in Exercise 3.1.
17. Once both instances are running, stop the first instance (Source). Make a note of the instance ID.
18. Go to the Amazon EBS page in the Amazon EC2 console and find the volume attached to the Source instance via the instance ID. Detach the instance.
19. When the volume becomes Available, attach the instance to the second instance (Destination).
20. Log in to the Destination instance via RDP using the administrator account.
21. Open a command window (`cmd.exe`).
22. At the command prompt, type the following commands:

```
C:\Users\Administrator >diskpart
DISKPART>select disk 1
DISKPART>online disk
DISKPART>exit
C:\Users\Administrator>dir e:
```

    The volume removed from the stopped source drive can now be read as the E: drive on the destination instance, so its data can be retrieved.

23. Terminate all the instances and ensure the volumes are deleted in the process.

Review Questions

1. Your web application needs four instances to support steady traffic nearly all of the time. On the last day of each month, the traffic triples. What is a cost-effective way to handle this traffic pattern?
   A. Run 12 Reserved Instances all of the time.
   B. Run four On-Demand Instances constantly, then add eight more On-Demand Instances on the last day of each month.
   C. Run four Reserved Instances constantly, then add eight On-Demand Instances on the last day of each month.
   D. Run four On-Demand Instances constantly, then add eight Reserved Instances on the last day of each month.
2. Your order-processing application processes orders extracted from a queue with two Reserved Instances processing 10 orders/minute. If an order fails during processing, then it is returned to the queue without penalty. Due to a weekend sale, the queues have several hundred orders backed up. While the backup is not catastrophic, you would like to drain it so that customers get their confirmation emails faster. What is a cost-effective way to drain the queue for orders?
   A. Create more queues.
   B. Deploy additional Spot Instances to assist in processing the orders.
   C. Deploy additional Reserved Instances to assist in processing the orders.
   D. Deploy additional On-Demand Instances to assist in processing the orders.
3. Which of the following must be specified when launching a new Amazon Elastic Compute Cloud (Amazon EC2) Windows instance? (Choose 2 answers)
   A. The Amazon EC2 instance ID
   B. Password for the administrator account
   C. Amazon EC2 instance type
   D. Amazon Machine Image (AMI)
4. You have purchased an m3.xlarge Linux Reserved instance in us-east-1a. In which ways can you modify this reservation? (Choose 2 answers)
   A. Change it into two m3.large instances.
   B. Change it to a Windows instance.
   C. Move it to us-east-1b.
   D. Change it to an m4.xlarge.
5. Your instance is associated with two security groups. The first allows Remote Desktop Protocol (RDP) access over port 3389 from Classless Inter-Domain Routing (CIDR) block 72.14.0.0/16. The second allows HTTP access over port 80 from CIDR block

0.0.0.0/0. What traffic can reach your instance?

A. RDP and HTTP access from CIDR block 0.0.0.0/0
B. No traffic is allowed.
C. RDP and HTTP traffic from 72.14.0.0/16
D. RDP traffic over port 3389 from 72.14.0.0/16 and HTTP traffic over port 80 from 0.0.00/0

6. Which of the following are features of enhanced networking? (Choose 3 answers)

A. More Packets Per Second (PPS)
B. Lower latency
C. Multiple network interfaces
D. Border Gateway Protocol (BGP) routing
E. Less jitter

7. You are creating a High-Performance Computing (HPC) cluster and need very low latency and high bandwidth between instances. What combination of the following will allow this? (Choose 3 answers)

A. Use an instance type with 10 Gbps network performance.
B. Put the instances in a placement group.
C. Use Dedicated Instances.
D. Enable enhanced networking on the instances.
E. Use Reserved Instances.

8. Which Amazon Elastic Compute Cloud (Amazon EC2) feature ensures that your instances will not share a physical host with instances from any other AWS customer?

A. Amazon Virtual Private Cloud (VPC)
B. Placement groups
C. Dedicated Instances
D. Reserved Instances

9. Which of the following are true of instance stores? (Choose 2 answers)

A. Automatic backups
B. Data is lost when the instance stops.
C. Very high IOPS
D. Charge is based on the total amount of storage provisioned.

10. Which of the following are features of Amazon Elastic Block Store (Amazon EBS)? (Choose 2 answers)

A. Data stored on Amazon EBS is automatically replicated within an Availability Zone.

B. Amazon EBS data is automatically backed up to tape.

C. Amazon EBS volumes can be encrypted transparently to workloads on the attached instance.

D. Data on an Amazon EBS volume is lost when the attached instance is stopped.

11. You need to take a snapshot of an Amazon Elastic Block Store (Amazon EBS) volume. How long will the volume be unavailable?

A. It depends on the provisioned size of the volume.

B. The volume will be available immediately.

C. It depends on the amount of data stored on the volume.

D. It depends on whether the attached instance is an Amazon EBS-optimized instance.

12. You are restoring an Amazon Elastic Block Store (Amazon EBS) volume from a snapshot. How long will it be before the data is available?

A. It depends on the provisioned size of the volume.

B. The data will be available immediately.

C. It depends on the amount of data stored on the volume.

D. It depends on whether the attached instance is an Amazon EBS-optimized instance.

13. You have a workload that requires 15,000 consistent IOPS for data that must be durable. What combination of the following steps do you need? (Choose 2 answers)

A. Use an Amazon Elastic Block Store (Amazon EBS)-optimized instance.

B. Use an instance store.

C. Use a Provisioned IOPS SSD volume.

D. Use a magnetic volume.

14. Which of the following can be accomplished through bootstrapping?

A. Install the most current security updates.

B. Install the current version of the application.

C. Configure Operating System (OS) services.

D. All of the above.

15. How can you connect to a new Linux instance using SSH?

A. Decrypt the root password.

B. Using a certificate

C. Using the private half of the instance's key pair

D. Using Multi-Factor Authentication (MFA)

16. VM Import/Export can import existing virtual machines as: (Choose 2 answers)

A. Amazon Elastic Block Store (Amazon EBS) volumes

B. Amazon Elastic Compute Cloud (Amazon EC2) instances

C. Amazon Machine Images (AMIs)

D. Security groups

17. Which of the following can be used to address an Amazon Elastic Compute Cloud (Amazon EC2) instance over the web? (Choose 2 answers)

A. Windows machine name

B. Public DNS name

C. Amazon EC2 instance ID

D. Elastic IP address

18. Using the correctly decrypted Administrator password and RDP, you cannot log in to a Windows instance you just launched. Which of the following is a possible reason?

A. There is no security group rule that allows RDP access over port 3389 from your IP address.

B. The instance is a Reserved Instance.

C. The instance is not using enhanced networking.

D. The instance is not an Amazon EBS-optimized instance.

19. You have a workload that requires 1 TB of durable block storage at 1,500 IOPS during normal use. Every night there is an Extract, Transform, Load (ETL) task that requires 3,000 IOPS for 15 minutes. What is the most appropriate volume type for this workload?

A. Use a Provisioned IOPS SSD volume at 3,000 IOPS.

B. Use an instance store.

C. Use a general-purpose SSD volume.

D. Use a magnetic volume.

20. How are you billed for elastic IP addresses?

A. Hourly when they are associated with an instance

B. Hourly when they are not associated with an instance

C. Based on the data that flows through them

D. Based on the instance type to which they are attached