

Welcome to the eleventh lesson of AWS Solutions Architect Associate level—Disaster Recovery.



	By the end of the lesson you will be able to:
•	Difference between traditional and Cloud Disaster Recovery practices
•	Discuss Disaster Recovery and its importance
•	DR methodologies
•	Recovery Time Objective and Recovery Point Objective

By the end of this lesson you'll be able to:

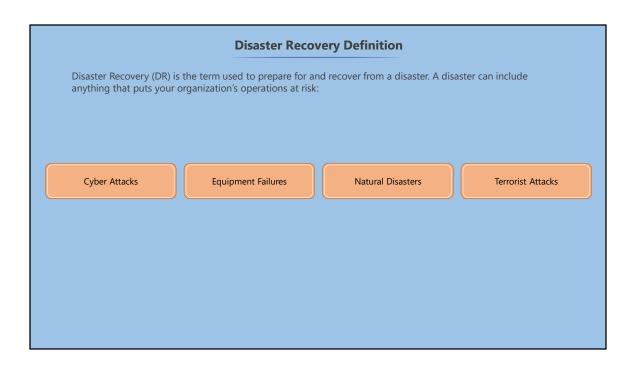
- Understand what Disaster Recovery is and why it is important
- Learn the differences between traditional and Cloud disaster recovery practices
- Understand what Recovery Time Objective and Recovery Point Objective are
- Grasp the concepts surrounding various DR methodologies



Overview of Disaster Recovery	

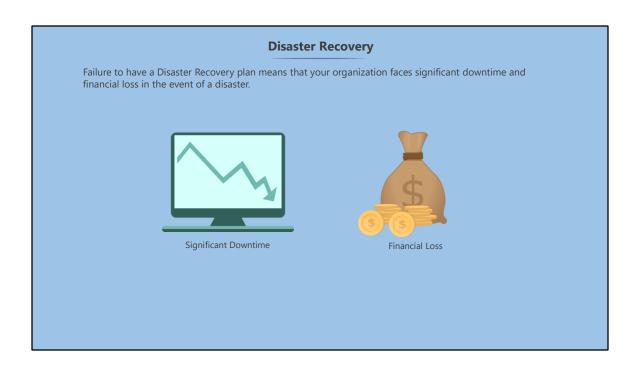
In this section you'll learn about the features of disaster recovery and its features.





Disaster recovery (DR) is the term used to prepare for and recover from a disaster. A disaster can include anything that puts your organization's operations at risk—cyber-attacks, equipment failures, natural disasters such as hurricanes or earthquakes, or even terrorist attacks.

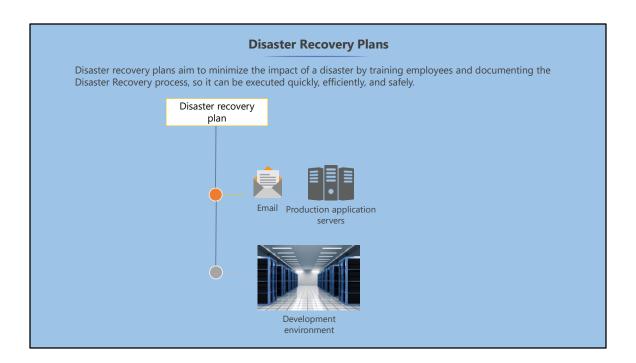




Failure to have a Disaster Recovery plan means that your organization faces significant downtime and financial loss in the event of a disaster.

Even small things such as an air conditioning failure at a data center needs a Disaster Recovery plan to avoid significant losses.





Disaster recovery plans aim to minimize the impact of a disaster by training employees and documenting the Disaster Recovery process, so it can be executed quickly, efficiently, and safely.

The most business critical systems are made online first, for example email, production application servers, and so on. Next in line are the less critical systems, for example the development environment.



#### **Traditional Disaster Recovery**

Following the traditional methods of Disaster Recovery, companies duplicate their environments to another data center in a location remote to the primary site.

The remote data center infrastructure has to be purchased, installed, and maintained so it is ready whenever it is required.

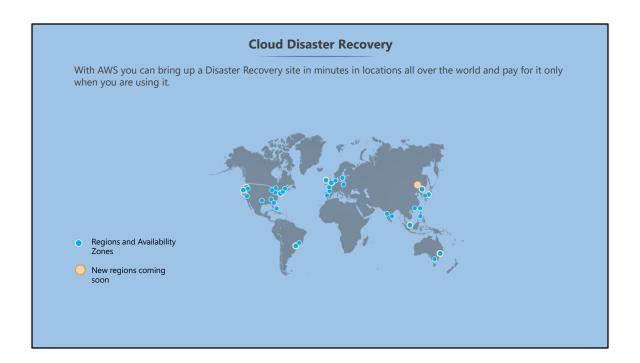
The Disaster Recovery site is often underutilized or over-provisioned, and is nothing but a wastage of money.

Following the traditional methods of Disaster Recovery, companies duplicate their environments to another data center in a location remote to the primary site.

The remote data center infrastructure has to be purchased, installed, and maintained so it is ready whenever it is required.

The Disaster Recovery site is often under-utilized or over-provisioned, and is nothing but a wastage of money.





With AWS you can bring up a Disaster Recovery site in minutes in locations all over the world and pay for it only when you are using it.

AWS allows you to scale immediately and quickly change and optimize your resources during a DR event.



#### **Traditional Versus Cloud**

The following table lists the differences between traditional and Cloud followed Disaster Recovery methodologies:

	Traditional	AWS
Facilities: Data centers, power, and cooling, and so on	User responsibility	AWS responsibility
Security to ensure the physical protection of assets	User responsibility	AWS responsibility
Suitable capacity to scale the environment	User responsibility	AWS responsibility
Support for repairing, replacing, and refreshing the infrastructure	User responsibility	AWS responsibility
Internet service provider contracts to provide bandwidth utilization for your environment under full load	User responsibility	AWS responsibility
Network infrastructure such as firewalls, routers, switches, and load balancers	User responsibility	AWS responsibility
Enough server capacity to run mission critical services	User responsibility	AWS responsibility

The table lists the differences between traditional and Cloud Disaster Recovery methodologies.

The following features are provided by AWS:

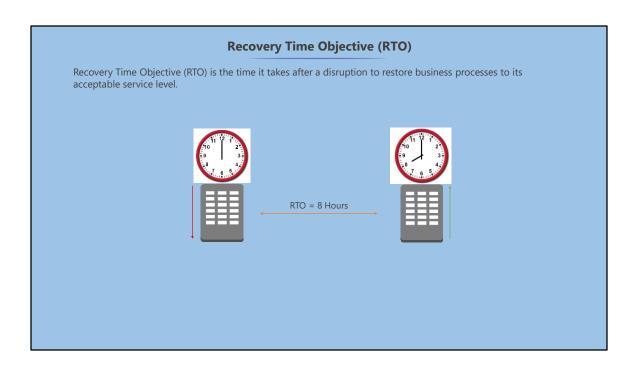
- · Facilities to house the infrastructure, including power and cooling.
- Security to ensure the physical protection of assets.
- Suitable capacity to scale the environment.
- Support for repairing, replacing, and refreshing the infrastructure.
- Contractual agreements with an Internet service provider (ISP) to provide Internet connectivity that can sustain bandwidth utilization for the environment under a full load.
- Network infrastructure such as firewalls, routers, switches, and load balancers.
- Enough server capacity to run all mission-critical services, including storage appliances for the supporting data and servers to run applications and backend services such as user authentication, Domain Name System (DNS).



RTO Versus RPO
Two important considerations with Disaster Recovery are the Recovery Time Objective (RTO) and Recovery Point Objective (RPO).

Two important considerations with Disaster Recovery are the Recovery Time Objective (RTO) and Recovery Point Objective (RPO).

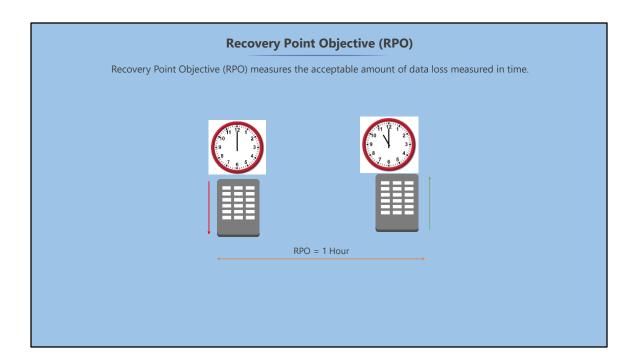




Recovery Time Objective (RTO) is the time it takes after a disruption to restore business processes to its acceptable service level.

For example, if a disaster occurs at 12:00 PM and the RTO is eight hours, the DR process should restore the business process to the acceptable service level by 8:00 PM.

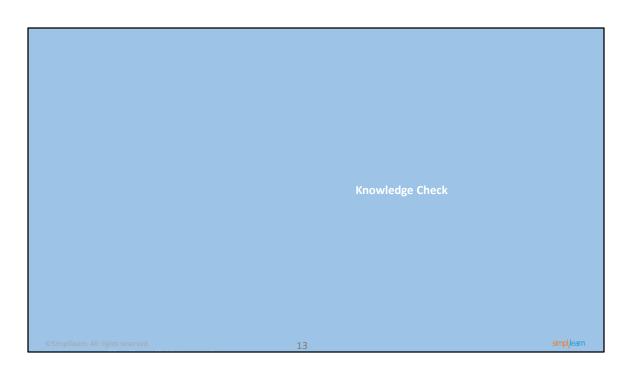




Recovery Point Objective (RPO) measures the acceptable amount of data loss measured in time.

For example, if a disaster occurs at 12:00 PM and the RPO is one hour, the system should recover all data that was in the system before 11:00 AM. Data loss will span only one hour, between 11:00 AM and 12:00 PM.







## Which of the following statements correctly defines Recovery Point Objective?

The time it takes after a disruption to restore a business process to its service level

How to prepare for and recover from a disaster

The acceptable amount of data loss measured in time

The point in the recovery plan that IT needs to start from



# Which of the following statements correctly defines Recovery Point Objective?

The time it takes after a disruption to restore a business process to its service level

How to prepare for and recover from a disaster

The acceptable amount of data loss measured in time

The point in the recovery plan that IT needs to start from

c

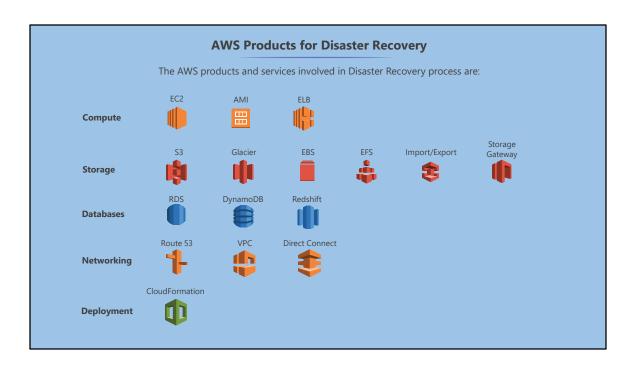
Recovery Point Objective measures the acceptable amount of data loss measured in time.



Using AWS Products for Disaster Recovery
--

In this section you'll learn about the AWS products that provide Disaster Recovery.

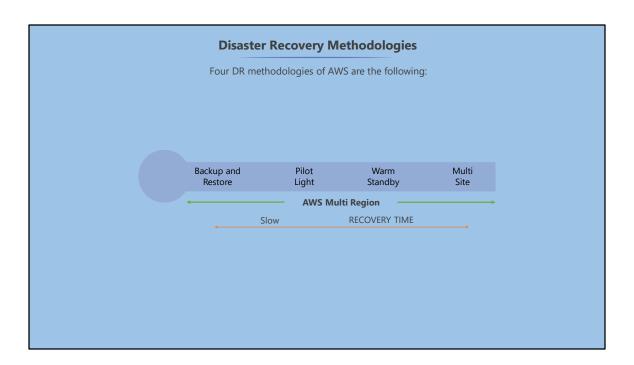




The AWS products and services involved in Disaster Recovery process are:

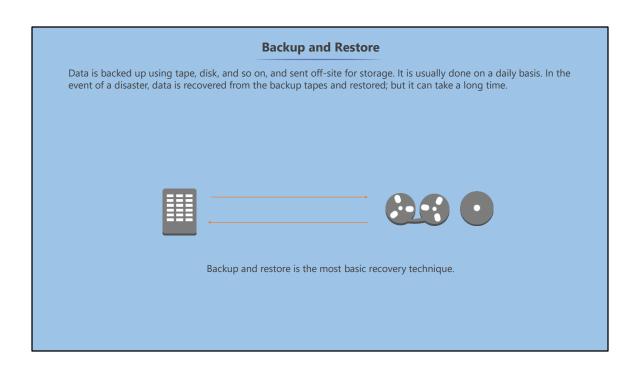
- Compute—EC2, AMIs, and ELB
- Storage—S3, Glacier, EBS, EFS, Import/Export, and Storage Gateway
- · Databases—RDS, DynamoDB, and Redshift
- Networking—Amazon Route 53, VPC, and Direct Connect
- Deployment—CloudFormation





In this section you'll learn about the four DR methodologies that showcase how AWS can be used.
Backup and restore
Pilot light
Warm standby
Multi-site



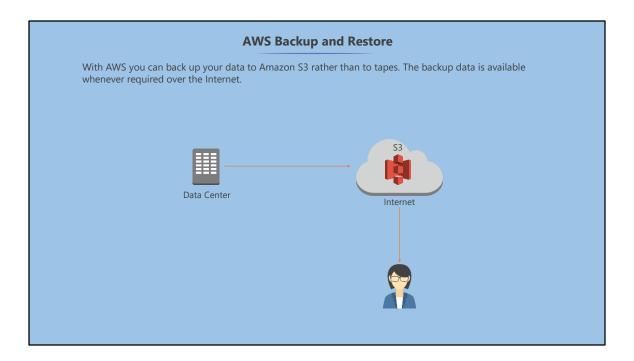


Data is backed up using tape, disk, and so on, and sent off-site for storage. It is usually done on a daily basis.

In the event of a disaster, data is recovered from the backup tapes and restored; but it can take a long time.

In this scenario the RTO is long and the RPO is low-to-high, depending on frequency of backups.

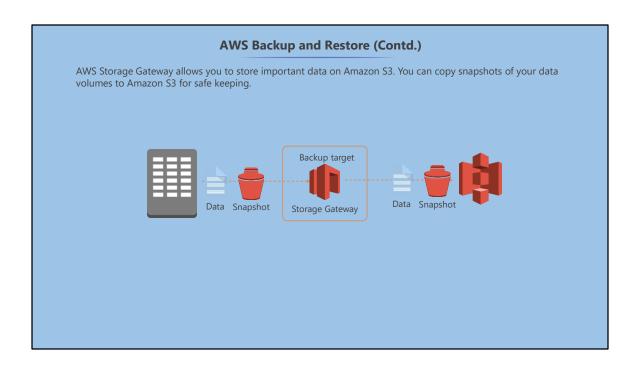




## AWS methods:

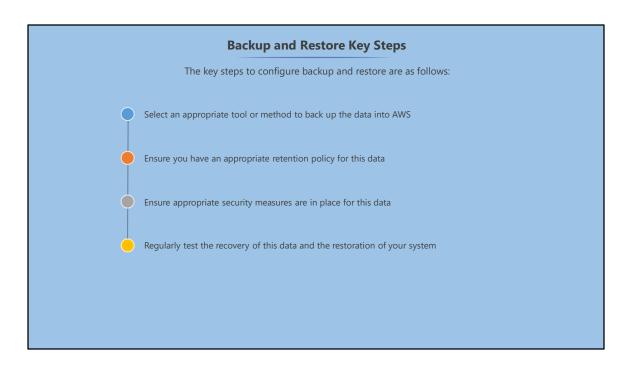
- With AWS you can back up your data to Amazon S3 rather than to tapes. The backup data is available over the Internet, whenever required.
- AWS S3 provides different classes of storage. You can store huge amounts of backups cost-effectively.





AWS Storage Gateway allows you to store important data on Amazon S3. You can copy snapshots of your data volumes to Amazon S3 for safe keeping. You can use Storage Gateway as a backup target for your existing backup software to replace traditional magnetic tapes.





The key steps for backup and restore are as following:

- 1. Select an appropriate tool or method to back up your data into AWS.
- 2. Ensure that you have an appropriate retention policy for this data.
- 3. Ensure that appropriate security measures are in place, including encryption and access policies.
- 4. Regularly test the recovery of this data and the restoration of your system.



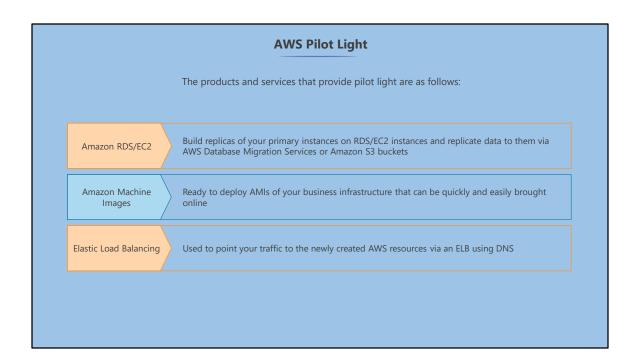
# Pilot Light

Pilot light allows you to run a minimal version of your environments in the cloud. In case of a disaster, you can scale up the standby copies of your core systems up to production capacity. In pilot light, the RTO is medium and the RPO is low to medium depending on the frequency of replication.

Pilot light allows you to run a minimal version of your environments in the cloud. In case of a disaster, you can scale up the standby copies of your core systems up to production capacity.

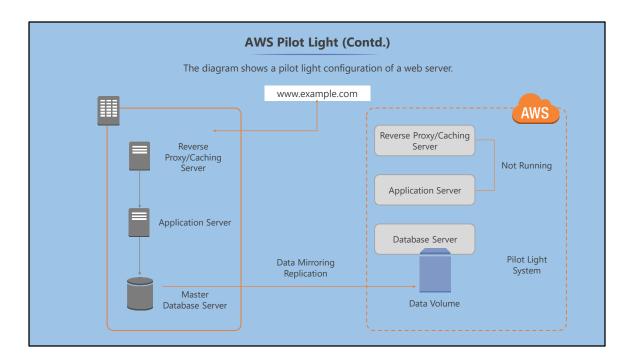
In pilot light the RTO is medium and the RPO is low-to-medium depending on the frequency of replication.





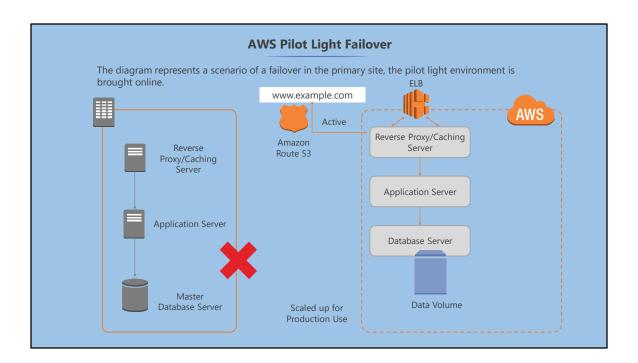
- Amazon RDS/EC2: Create replicas of your primary instances on RDS/EC2 instances and replicate data to them via AWS DMS or Amazon S3 buckets. When required you can scale them to larger instance types and move your production operations to them.
- Amazon Machine Images: You can have ready-to-deploy AMIs of your business infrastructure that can be quickly and easily brought online, for example application servers or web servers.
- Elastic Load Balancing: You can use Elastic Load Balancing for web servers to point your traffic to the newly created AWS resources.





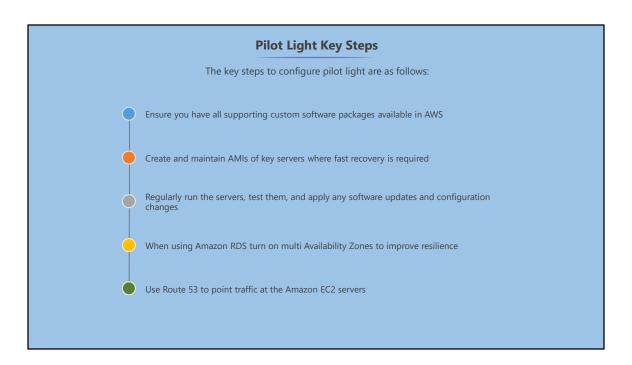
The diagram shows a pilot light configuration of a web server. In the left hand side there is a corporate data center. There is the caching server, application server, and the database server. The database server is replicated to AWS where a small database server is running. However, in AWS the caching server and the application server are not currently running.





If there is a Disaster Recovery required in the data center, you can fail over to the AWS environment, where the database server is already up and running. You can scale it to a larger instance as required. And bring the application server and the caching server online either via AMIs or by just starting them up. You can use ELB to push traffic to the new resources.





The key steps to configure pilot light are listed as follows:

- 1. Ensure you have all supporting custom software packages available in AWS
- 2. Create and maintain AMIs of key servers where fast recovery is required
- 3. Regularly run the servers, test them, and apply any software updates and configuration changes
- 4. When using Amazon RDS turn on multi Availability Zones to improve resilience
- 5. Use Route 53 to point traffic at the Amazon EC2 servers



#### **Warm Standby**

It extends the idea of pilot light by running a scaled down version of a fully functional environment always running in the Cloud.

The method decreases recovery time as the critical systems are already operational. The warm standby site can either be on instances ready for production or on lower sized instances that can be scaled as required.

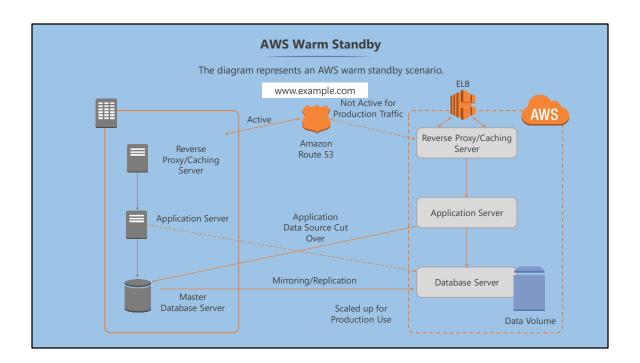
The RTO is low to medium and the RPO is low.

It extends the idea of pilot light by running a scaled down version of a fully functional environment always running in the Cloud.

The method decreases recovery time as the critical systems are already operational. The warm standby site can either be on instances ready for production or on lower sized instances that can be scaled as required.

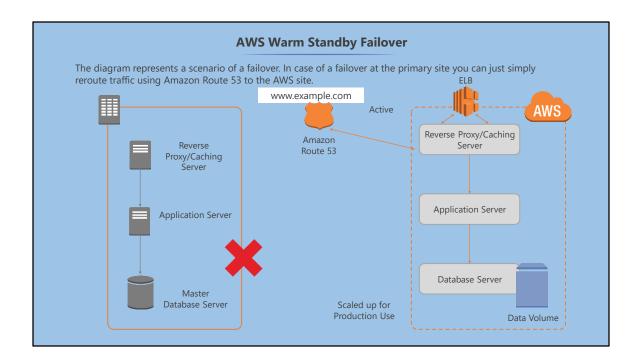
The RTO is low to medium and the RPO is low.





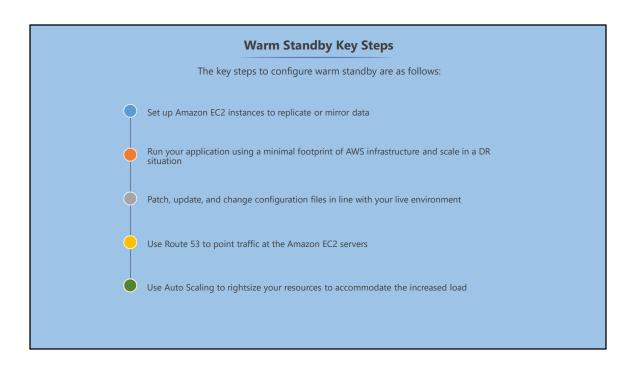
The diagram represents an AWS warm standby scenario. In the left hand side there is the corporate data center, the caching server, application server, and the database server. The database server is replicated to AWS. However, in warm standby the caching server and the application server are already up and running and can communicate with the resources on the other side.





In the case of failover at the primary site you can just reroute traffic using Amazon Route 53 to the AWS site. This method involves minimum downtime for users. If required, you can scale up your instances to larger instance types.





The key steps to configure warm standby are as follows:

Set up Amazon EC2 instances to replicate or mirror data

Run your application using a minimal footprint of AWS infrastructure and scale in a DR situation

Patch, update, and change configuration files in line with your live environment Use Route 53 to point traffic at the Amazon EC2 servers

Use Auto Scaling to right-size your resources to accommodate the increased load



#### **Multi-Site**

In the multisite scenario the infrastructure runs both in AWS and in existing on-premise sites in an active-active configuration.

Users can use both sites. You can direct the user traffic via Route 53 weighted routing. In this case the RTO is short and the RPO is very low.

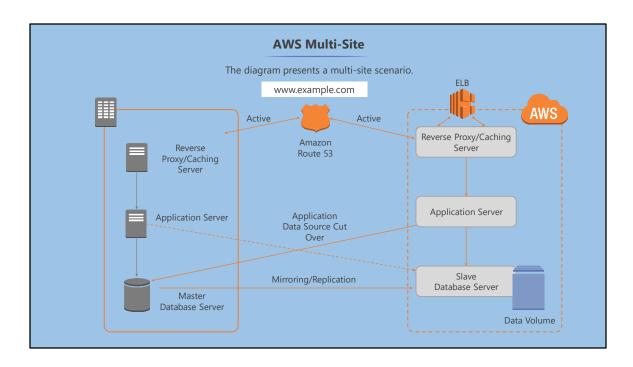
In the multisite scenario the infrastructure runs both in AWS and in existing onpremise sites in an active-active configuration.

Users can use both sites. You can direct the user traffic via Route 53 weighted routing .

In this case the RTO is short and the RPO is very low.

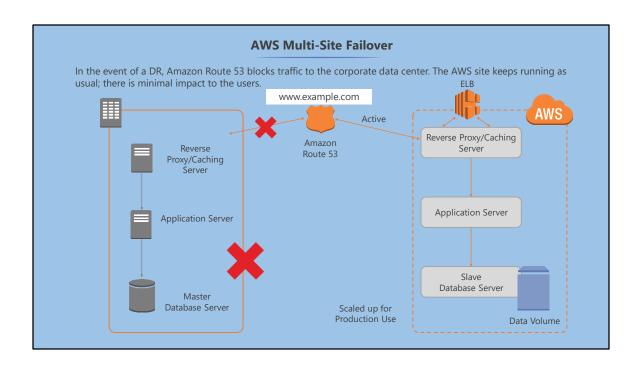
It uses the same methods as pilot light and warm standby, except that the AWS environment is full sized in this case.





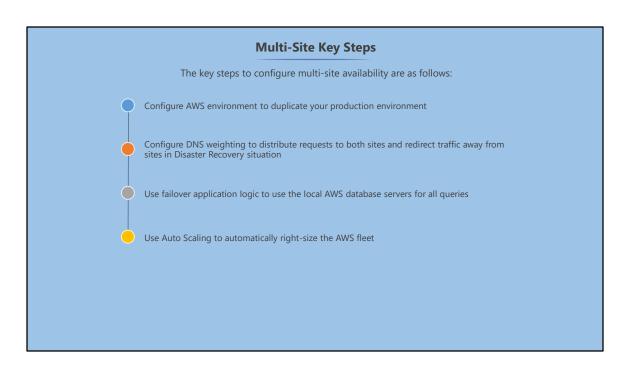
The diagram presents a multi-site scenario. Every server is up and running in both the corporate data center and in AWS. Amazon Route 53 weighted routing is being used to split user traffic equally between the two sites.





In the event of a DR, Amazon Route 53 blocks traffic to the corporate data center. The AWS site keeps running as usual; there is minimal impact to the users.

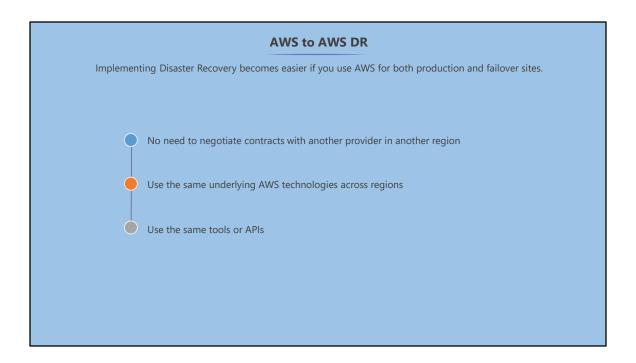




The key steps for configuring multi-site are as follows:

- 1. Configure AWS environment to duplicate your production environment
- 2. Configure DNS weighting to distribute requests to both sites and redirect traffic away from sites in Disaster Recovery situation
- 3. Use failover application logic to use the local AWS database servers for all queries
- 4. Use Auto Scaling to automatically right-size the AWS fleet

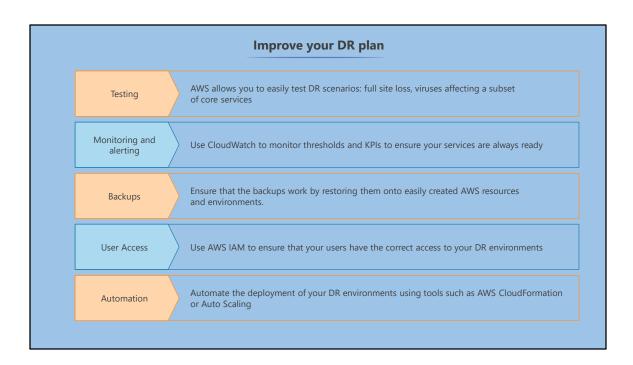




If you are already using AWS, then setting up DR is easier. The following reasons explain why:

- You don't need to negotiate contracts with another provider in another region
- You can use the same underlying AWS technologies across regions
- You can use the same tools or APIs





The following points explain how you can use AWS to improve your DR plan:

- Testing: AWS allows you to easily test DR scenarios: full site loss, viruses affecting a subset
  - of core services
- Monitoring and alerting: Use CloudWatch to monitor thresholds and KPIs to ensure your services are always ready
- Backups: Ensure that the backups work by restoring them onto easily created AWS resources
- User Access: Use AWS IAM to ensure that your users have the correct access to your DR environments
- Automation: Automate the deployment of your DR environments using tools such as AWS CloudFormation or Auto Scaling



Knowledge Check



Which Disaster Recovery methodology offers the lowest RTO?
Backup and restore
Pilot light
Warm standby
Multi-site



Which Disaster Recovery methodology offers the lowest RTO?
Backup and restore
Pilot light
Warm standby
Multi-site
d
Multi-site offers the lowest RTO as infrastructure runs both in AWS and in existing on-premise sites in an active-active configuration.



Practice Assignment: Disaster Recovery	
Calculate RTO and RPO	

In this section you'll apply what you've learned to calculate to calculate RTO and RPO.



## Disaster Recovery Assignment



Your company is concerned about the RTO and RPO of their accounting database, which is used to store important accounting data and run reports.

Management wants to have an RTO < 1 hours and RPO < 15 minutes.

You need to calculate the maximum RTO and RPO for the accounting database to see if it meets the proposed SLA and if it doesn't, suggest an infrastructure solution using AWS.

Accounting database backup details:

- Backed up to tape at 10PM each night
- Tapes are rotated at 9AM each morning and taken offsite
- Retrieval time for the tapes is 4 hours
- It takes 2 hours to restore the database from a backup

Your company is concerned about the RTO and RPO of their accounting database, which is used to store important accounting data and run reports.

Management wants to have an RTO < 1 hours and RPO < 15 minutes.

You need to calculate the maximum RTO and RPO for the accounting database to see if it meets the proposed SLA and if it doesn't, suggest an infrastructure solution using AWS.

Accounting database backup details:

- Backed up to tape at 10PM each night
- Tapes are rotated at 9AM each morning and taken offsite
- Retrieval time for the tapes is 4 hours
- It takes 2 hours to restore the database from a backup



- Disaster Recover (DR) is the term used to prepare for and recover from a disaster.
- With traditional Disaster Recovery methodologies the Disaster Recovery site remains often under-utilized or over-provisioned and normally just wastes money.
- Two important considerations with Disaster Recovery are the Recovery Time Objective (RTO) and Recovery Point Objective (RPO).
- Four DR scenarios that showcase how AWS can be used are: Backup and Restore, Pilot light, Warm standby, and Multi-site

- Disaster Recover (DR) is the term used to prepare for and recover from a disaster.
- With traditional Disaster Recovery methodologies the Disaster Recovery site remains often under-utilized or over-provisioned and normally just wastes money.
- Two important considerations with Disaster Recovery are the Recovery Time Objective (RTO) and Recovery Point Objective (RPO).
- Four DR scenarios that showcase how AWS can be used are: backup and restore, pilot light, warm standby, and multi-site.



Knowledge Check

- 1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



When using AWS for disaster recovery, who is responsible for the facilities housing the infrastructure?

Your organization

AWS

Both AWS and your organization

A third party data center provider

- 1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



When using AWS for disaster recovery, who is responsible for the facilities housing the infrastructure?

Your organization

AWS

Both AWS and your organization

A third party data center provider

b

AWS is responsible for facilities to house the infrastructure, as well as security of the physical assets, capacity, support, network, and ISP contracts.

- 1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



If yo 2 at 12	our systems fail at 12PM and your SLA says that you need all systems to be back online 2AM. What is your RTO?
12 hours	
0 hours	
24 hours	
36 hours	

- 1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



If your systems fail at 12PM and your SLA says that you need all systems to be back online at 12AM. What is your RTO?
12 hours
0 hours
24 hours
36 hours
a  The RTO (Recovery Time Objective) is the time after a disruption to restore services to acceptable levels, which in this
example is 12 hours.

- Amazon Web Services (AWS) is a secure cloud services platform that offers cloudbased infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



Pilot Light disaster recovery methodology means 3
Running a scaled down version of a fully functional environment
Running your environment both in AWS and in on premise sites in an active-active configuration
Running backups and storing on tapes offsite
Running a minimal version of your environments in the cloud.

- 1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



Pilot Light disaster recovery methodology means \_\_\_\_\_\_.

Running a scaled down version of a fully functional environment

Running your environment both in AWS and in on premise sites in an active-active configuration

Running backups and storing on tapes offsite

Running a minimal version of your environments in the cloud.

d

Pilot Light means running a minimal version of your environments in the cloud that can be quickly scaled up and launched in the event of a disaster.

- Amazon Web Services (AWS) is a secure cloud services platform that offers cloudbased infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



Which of the following is NOT a key benefit of using AWS for disaster recovery?

Reduced Cost

Ability to fully test disaster recovery plans

Increased RPO and RTO

Option to scale your DR environment on the fly

- 1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



Which of the following is NOT a key benefit of using AWS for disaster recovery?

Reduced Cost
Ability to fully test disaster recovery plans
Increased RPO and RTO
Option to scale your DR environment on the fly

C
Using AWS for DR will reduce your RTO and RPO in most cases

- Amazon Web Services (AWS) is a secure cloud services platform that offers cloudbased infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



:	Which AWS service allows you to replace traditional magnetic tape backups?
	AWS Direct Connect
	AWS Storage Gateway
	AWS WorkSpaces
	AWS Lambda

- 1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- 3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.



Which AWS service allows you to replace traditional magnetic tape backups?

AWS Direct Connect

AWS Storage Gateway

AWS WorkSpaces

AWS Lambda

b

AWS Storage Gateway is used as a backup target for your existing backup software to replace traditional magnetic tape backups.

- Amazon Web Services (AWS) is a secure cloud services platform that offers cloudbased infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
- 2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
- A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
- 4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
- 5. AWS has various cloud-based products to help your business grow.