Hello Cloud Gurus and welcome to this lecture,

which is going to introduce RDS backups and snapshots.

Now when it comes to backing up your data held in RDS,

you've actually got two options.

And they're similar,

but they have some very important differences.

So firstly, we have database snapshots,

which are manual, ad-hoc, and user initiated.

And this provides a snapshot or a point in time copy

of the storage volume attached to your database instance.

And we also have automated backups,

and these are enabled by default.

And with automated backups,

RDS creates daily backups or snapshots

which run during a backup window that you define.

And in addition to this daily backup,

it also generates transaction logs,

which are used to replay transactions

when you come to restore the database.

So let's take a closer look at each of these options

in a little bit more detail,

beginning with automated backups.

Now with automated backups,

you actually get a lot more bang for your buck,

because this gives you the ability

to perform a point-in-time recovery

and recover your database to any point in time

within a retention period of between one and 35 days.

So this gives you the ability to recover

to any point in time within your defined retention period.

So how does it do that?

Well, it performs a full daily backup or snapshot,

and it also stores transaction logs throughout the day.

And then during the recovery process,

when you perform a recovery,

AWS will first select the most recent daily backup.

It will then apply the transaction logs relevant to that day

up to the recovery point that you select.

So imagine that you have an RDS database,

storing details of your customers online purchases.

You have an automated backup scheduled to complete

within a backup window of between 2:00 AM and 5:00 AM.

So RDS will automatically take a daily snapshot

during that backup window,

and the snapshots are stored in S3.

And it's also creating transaction logs as well.

So let's say that somebody accidentally deleted

some of your data and you need to recover the database back

to the state that it was in at 9:00 AM this morning.

Well to do that,

RDS will restore using the latest snapshot,

and it's then going to use the transaction logs

to replay any changes which happened

since the snapshot was taken.

So any changes that happened up to our recovery point,

which was 9:00 AM this morning.

And it's this combination of the daily snapshot

and the transaction log which enables point-in-time recovery

down to the second within your retention period.

And the retention period is of course user-defined.

And it can be anything between one and 35 days.

So with automated backups,

these are stored in S3,

and the manual snapshots are also stored in S3.

You get free storage space equivalent

to the size of your database.

So if you've got an RDS instance with 10 gigs of data,

then you will get 10 gigs worth of storage

for your automated backups.

And everything happens within a backup window

that you define.

However, during the backup window,

your storage I/O may be suspended for just a few seconds

while the backup process initializes itself.

And during this time, you may experience increased latency.

But this is only during the initialization process.

And after that,

your latency on your performance should go back to normal.

So what about snapshots?

Well, these are not automated.

So database snapshots are done manually,

and they are user-initiated.

There is no retention period when you do a manual snapshot.

So they are not deleted even after you delete

the original RDS instance, including any automated backups.

So the database and its backups can be deleted,

and the snapshots will remain.

And snapshots enable you to back up your database

to a known state.

And you can back up your database instance to a known state

as frequently as you wish,

and then restore to that specific state at any time.

So when would you use this?

Well, imagine you were planning

some significant updates to your data.

You might wish to backup to a known state

before you make any changes to your data,

so that you can always restore back to that state

if anything goes wrong.

Now when you come to restore an RDS database,

whether you restore from an automated backup

or from a manual snapshot,

the restored version of the database

will always be a completely new RDS instance

with a new DNS endpoint.

So it's going to have a different endpoint

to your original RDS instance.

So we've got our original endpoint

and the new endpoint for the restored database.

Moving on to encryption,

and you might be wondering,

how does encryption fit in with all of this?

Well, there is a link,

and there are some important features to be aware of

when it comes to encryption and database snapshots.

Encryption at rest is very easy to enable with RDS,

and you can enable it at creation time

by just selecting the encryption option in the console.

And this is something that we're going to see later on

when we can come to create our own RDS instance.

And RDS is completely integrated with KMS.

So encryption is done using KMS,

and it uses the industry standard, AES-256 bit encryption.

Now when we enable encryption on an RDS database,

then RDS is going to encrypt all of the database storage.

So that includes all of the underlying storage associated

with your RDS database,

including any automated backups,

any manual snapshots, any logs,

and read replicas as well.

And if you haven't heard of read replicas,

we will be covering those later on

in this section of the course.

Now encryption can only be enabled

when you first create the database,

and you cannot later on enable encryption

on an unencrypted RDS database instance.

So what do you do

if you have an existing unencrypted database

and you are suddenly asked to encrypt that data?

Well, snapshots can actually help you with that situation,

to help encrypt your database.

So if you've already created an unencrypted database

and you need to encrypt that data,

you can take a snapshot

and that snapshot will also be unencrypted.

And then from that unencrypted snapshot,

you can create an encrypted snapshot,

and then perform a database restore

using that encrypted snapshot.

And in that way, you will get an encrypted database.

So this will be a copy of your original database,

but with everything encrypted.

And it sounds like a pretty long winded process.

So if possible, to avoid that,

make sure you find out if the data needs to be encrypted

before creating your database.

And then you can just enable encryption in the beginning.

So onto my exam tips for backups and snapshots.

And you'll need to understand

the difference between an automated backup

and a database snapshot or a manual snapshot.

So automated backups are enabled by default

and you define the backup window.

It runs a daily backup consisting

of a point-in-time snapshot or backup plus transaction logs,

allowing you to restore to any point in time

within your retention period.

And you can define a retention period

of anything between one and 35 days,

and automated backups can be used to recover your database

to any point in time down to the second

within that retention period.

Alternatively, with database snapshots,

and you might also see them called manual snapshots.

These are user-initiated and ad-hoc.

So they don't happen automatically.

It's a point-in-time snapshot only,

and there is no retention period.

So your snapshots are stored indefinitely,

even if the database and its automated backups are deleted.

And a database snapshot can be used

to backup your database instance to a known state

and restore to that specific state at any time.

For example before making a change

to your production database,

you might take a snapshot

so that you can easily roll back the changes

if something goes wrong.

When it comes to encryption,

you can only enable encryption

when you first create the database.

So you cannot enable it later.

And when encryption is enabled,

this includes all of the underlying storage

of the RDS database, including your automated backups,

any snapshots, logs, and read replicas.

And when you enable encryption,

it uses the AWS key management service

with the industry standard, AES-256 bit encryption.

And finally if you didn't enable encryption

at creation time,

and then you'd later on need to encrypt the data

in your RDS instance,

then you can use snapshot technology to take a snapshot.

And from that snapshot,

create an encrypted snapshot,

and create a new RDS instance from the encrypted snapshot.

So that is the end of this lecture.

If you have any questions, please let me know.

Otherwise, I'll see you in the next lecture.

Thank you.