

Databases AWS :

Amazon's database service, relational database service, is called RDS

It has 2 main features :

1. Multi AZ - For disaster recovery
2. Read Replicas : For performance

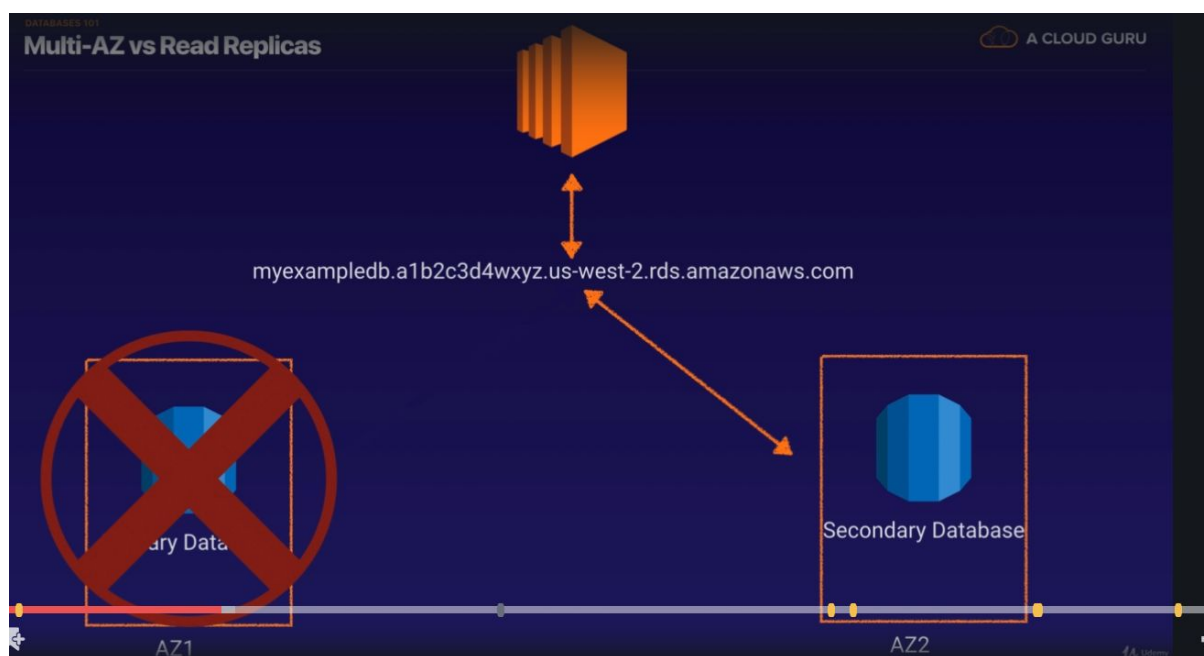
There are exam questions on whether you should be using multi AZ or Read replicas

.

If you want to spread your DB across multiple Availability zones, then you want Multi AZ,

If you wanna try and boost the DB performance, then you go for read replica.

Multi AZ vs Read replica :

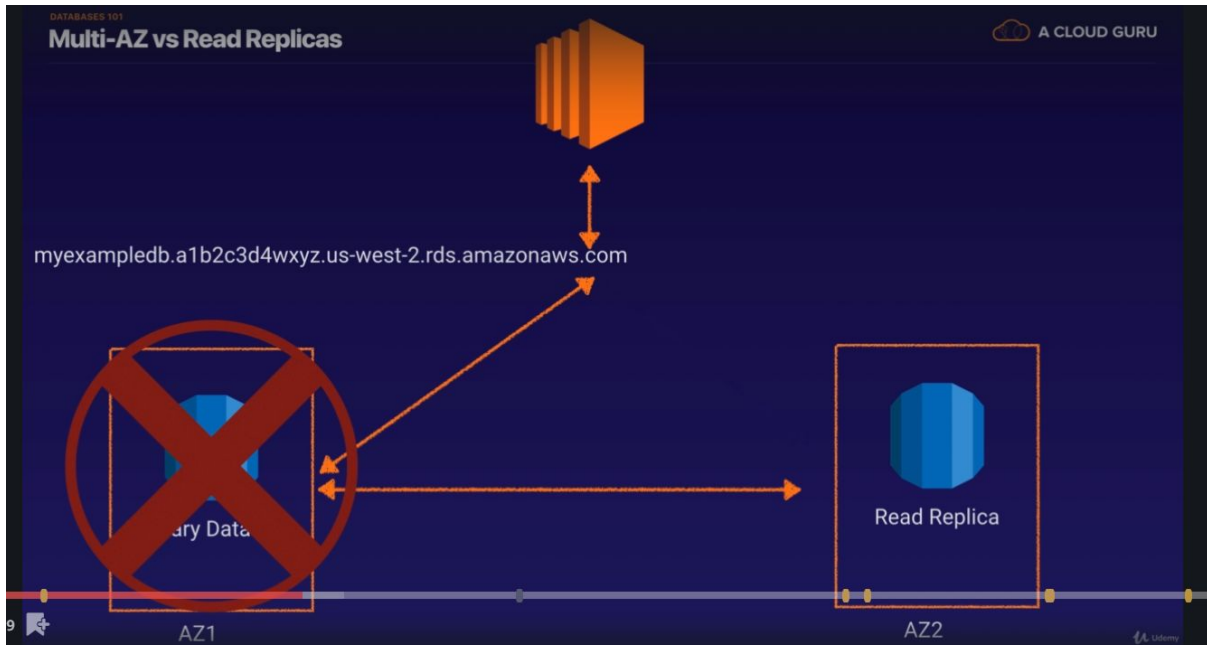


Above is the picture for Multi AZ.

In this case, a DNS address is configured for a DB. There are 2 instances configured , one is primary and the other secondary . *If the primary one fails, AWS automatically update the ip address of the secondary instance.*

So failover is automatic with Multi AZ.

Read Replicas :



With Read replicas, the DNS is configured with 2 primary instance. Whenever data is written into that primary instance, it is copied into the secondary instance as well, which is a READ REPLICA. So, the read replica is exact copy of the primary DB instance. However, if the primary instance fails, there is no automatic routing to the secondary instance.

There is no automatic failover.
You have to go and create a new url. Configure your EC2 to point to the new DB.

Useful where :

Suppose there is a wordpress website where lots and lots of people are using it to read blogs. So there are lots of read requests on your DB. So you can configure your EC2 to read half the requests from the primary DB instance and the other half from the read replica.


You can have 5 copies of READ replica.

Data Warehousing :

DATABASES 101
What is Data Warehousing? A CLOUD GURU

Used for business intelligence. Tools like Cognos, Jaspersoft, SQL Server Reporting Services, Oracle Hyperion, SAP NetWeaver.

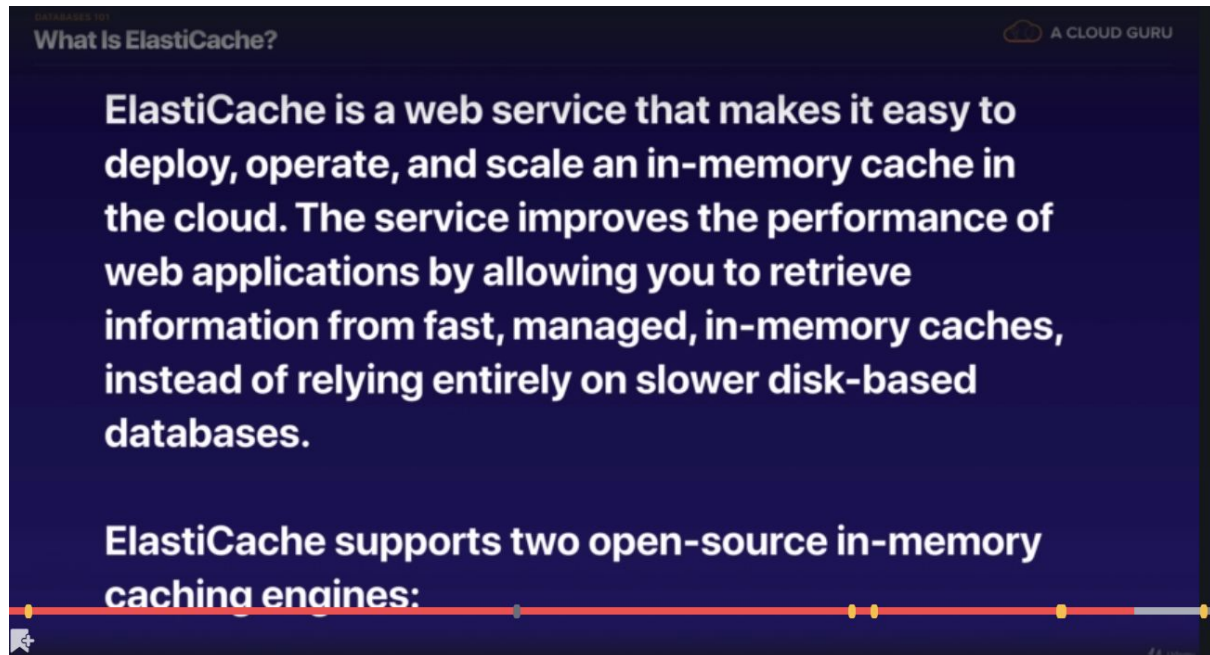
Used to pull in very large and complex data sets. Usually used by management to do queries on data (such as current performance vs targets etc)

The image is a composite of two parts. On the left, there is a 3D icon of a database cylinder with a gear, symbolizing data processing or warehousing. On the right, there is a photograph of a large warehouse interior with high ceilings, blue metal shelving units, and numerous cardboard boxes stacked on the shelves, representing a physical data warehouse.

There is a difference between Online Transaction Processing (OLTP) and Online Analytics Processing (OLAP). Transaction processing involves less queries, like reading records etc, but analytics involves lots of operations, so there are different DBs suited for transaction processing and for analytics, which require complicated queries.

Amazon provides its own Data warehouse solution, which is **RedShift**.

ElastiCache:



AWS service which provides an in-memory cache, allows you to retrieve information from fast, managed, in-memory caches instead of relying entirely on slower disk-based databases.

In memory cache is a cache which can be used to store the results of some queries, like the most commonly used queries whose results do not change very often, so that when those queries are executed, it is not required to read them from the DB, instead their results can be fetched from the memory cache.

So, when a query is executed, AWS first checks **Elasticache** and gets the info from there which takes a huge load off the DB.

You can get a lot of questions, that your DB is overloaded, and would ask for options :

- Using Elasticache is one option
- Read replicas another option

ElastiCache supports 2 opensource cache engines :

- Memcached
- Redis.

Basic TIPS for RDS:

- RDS consists of online transaction processing OLTP
 - Comes in 6 different flavors :
 - SQL.
 - MYSql
 - PostGreSQL
 - Oracle
 - Aurora
 - MariaDB
 - DynaMO DB - AWS noSQL soln
 - Redshift : AWS data warehousing
 - ElastiCache :
 - Memcached
 - Redis
 - AWS in memory cache.
 - Speed up the performance of existing databases or identical queries.
 - RDS runs on virtual machines, you cannot login to these Operating systems however, like EC2.
 - Patching of the RDS operating systems and DB is Amazon's responsibility.
 - RDS is not serverless.
 - Aurora serverless is serverless
-

RDS, Backups, Multi-AZ and Read Replicas

There are 2 types of backups for RDS :

1. Automated Backups
2. Database Snapshots

Automated Backups

They allow you to recover your database to any point in time within a '*Retention Period*'. The retention period can be between 1 and 35 days. Automated backups take a full daily snapshot and store transaction logs throughout the day.

When you do a recovery, AWS will choose the most recent daily backup, and then apply transaction logs relevant to that day. This allows you to do a point in time recovery down to a second, within that retention period.

- Automated backups are enabled by default.

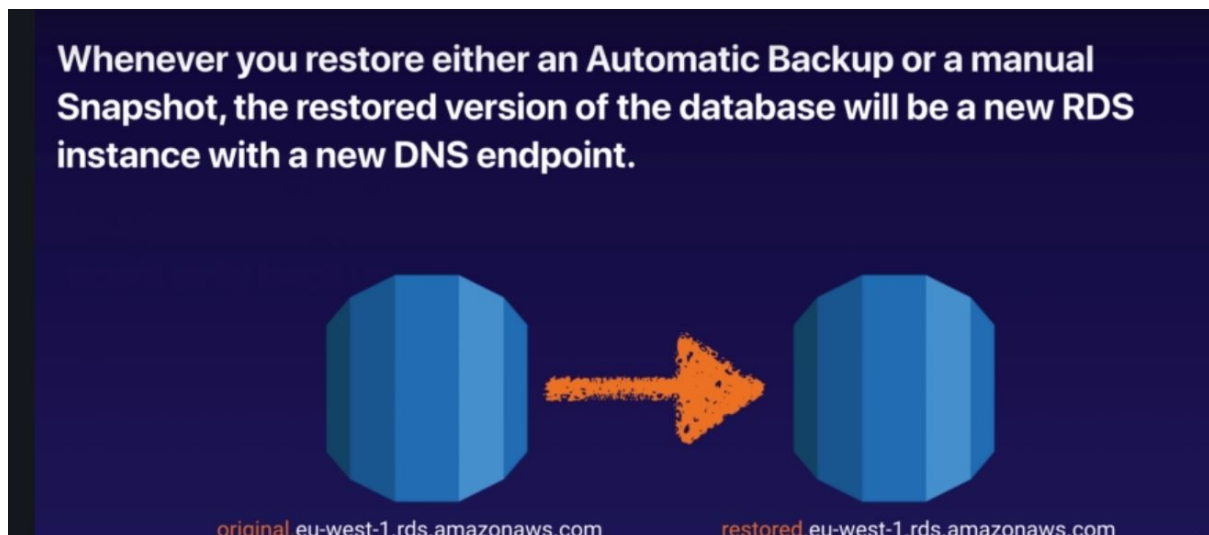
- The backup data is stored in S3, and you get free storage space equal to the size of your DB. If your DB size is 10GB, you will get 10GB worth of storage.
- Backups are defined within a defined window. During the backup window, storage I/O may be suspended while your data is being backed up and you may experience elevated latency.

Database Snapshots:

They are done manually, they are user initiated. They are stored even after you delete the original RDS instance, unlike the automated backups.

- When you go to delete the RDS instance, it will ask if you want to create a final snapshot before you delete it.

Whenever you restore either an Automated backup or a manual snapshot, the restored version of the DB will be a new RDS instance with a new DNS endpoint.

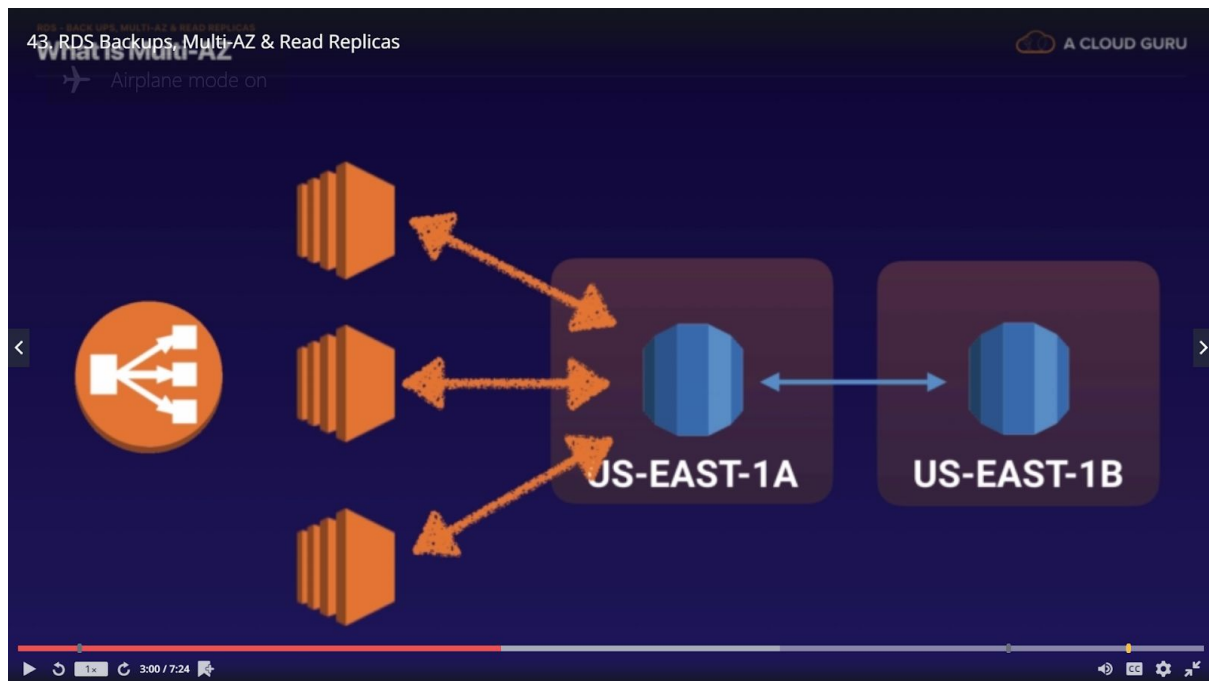


Encryption at Rest :

is supported for MySQL, Oracle, SQLServer, PostgreSQL, MariaDB and Aurora.

Encryption is done using the AWS KMS / Key management service. Once your RDS instance is encrypted, the data stored at rest in the underlying storage is encrypted, as are its automated backups, read replicas and snapshots.

Multi AZ :



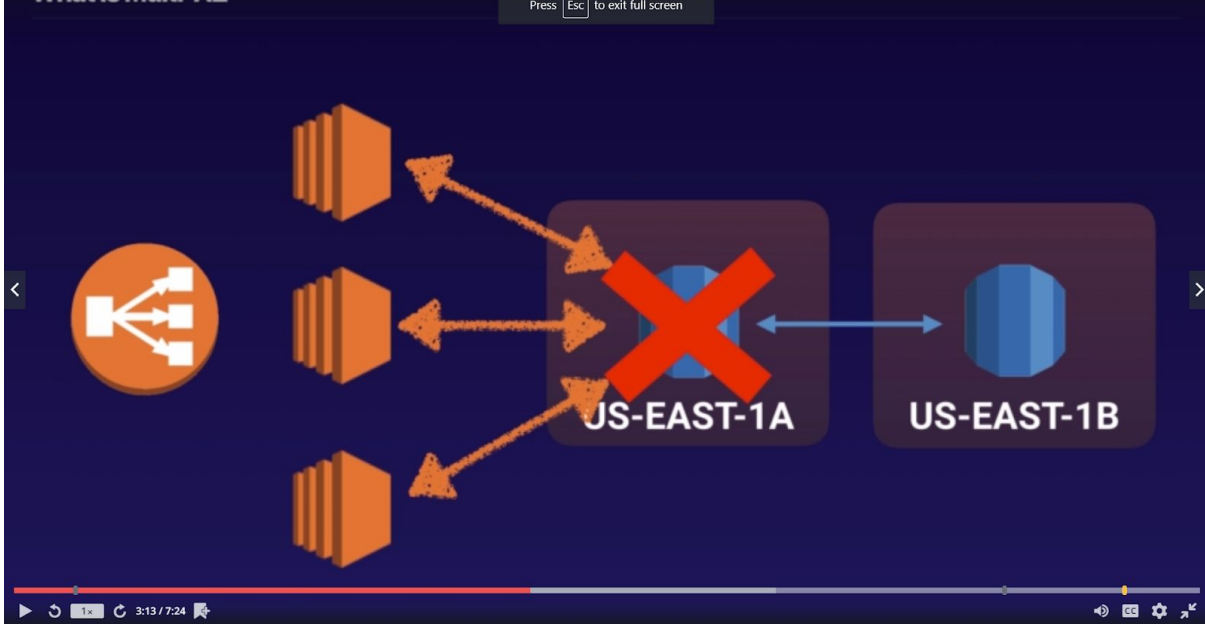
-- Lets say there is a load balancer here routing traffic among 3 EC2 instances, which are connected to DB in US-EAST 1A. With **MultiAZ**, there is a copy of the DB in another AZ, here - US-EAST 1B. All the data.

It synchronously updates all changes to the DB in US-East 1B as well.

-- So there are 2 copies of you DB.

-- If you were to lose the AZ / DB in AZ 1, AWS is going to update the DNS settings and the RDS is automatically going to failover / point to the RDS instance in the second AZ, so the system does not crash.

-- Multi Az can be selected during creation of RDS instance.




- Multi AZ allows you to have an exact replica of your production DB in another Availability zone, and that is done synchronously through synchronous replication.
- AWS handles the replication, so when your DB is written to, this will automatically be synchronized to the standby DB.
- In the event of a planned DB maintenance, DB instance failure or Availability zone failure, Amazon RDS will automatically failover to the standby so that DB operations can resume quickly without administrative intervention.
- You do not need to change anything in case of a failover of AZ/ instance. You need to use the same DNS endpoint, AWS will itself change the IP to point to the failover RDS instance, to restore connectivity.

43. RDS Backups, Multi-AZ & Read Replicas

What is Multi-AZ

Press Esc to exit full screen

Multi-AZ is for Disaster Recovery only



It is not primarily used for improving performance. For performance improvement, you need Read Replicas.

4:11 / 7:24


43. RDS Backups, Multi-AZ & Read Replicas

Multi-AZ Support

Press Esc to exit full screen

Multi-AZ is available for the following databases

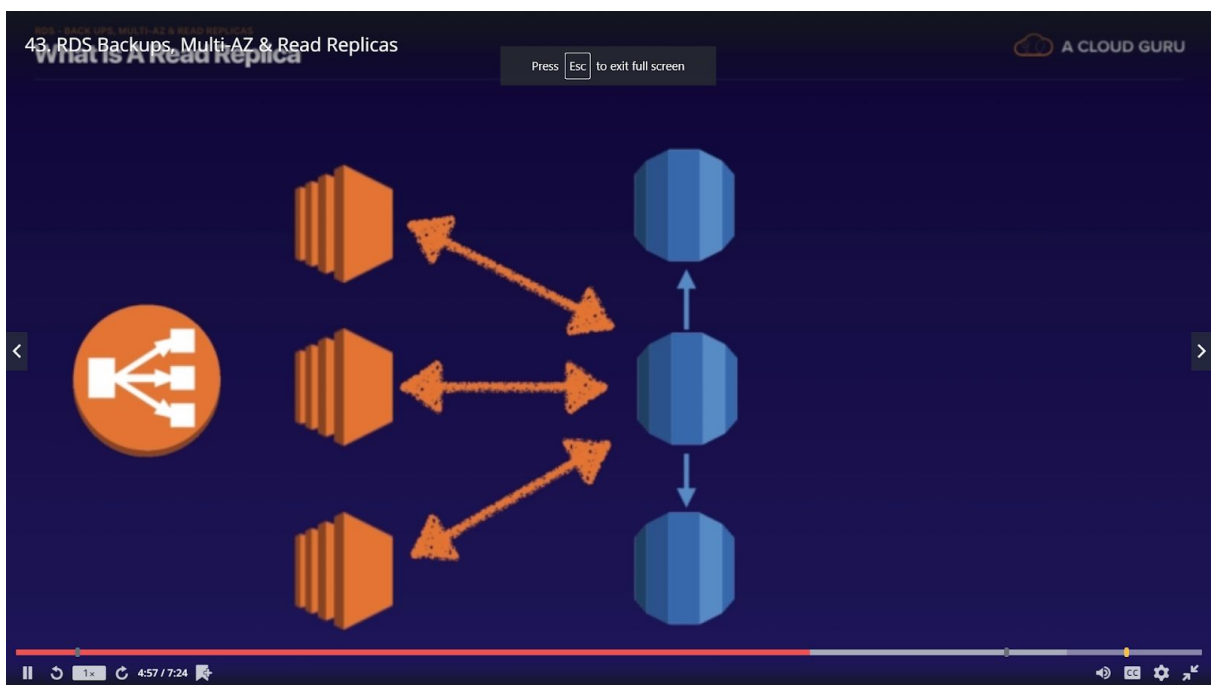
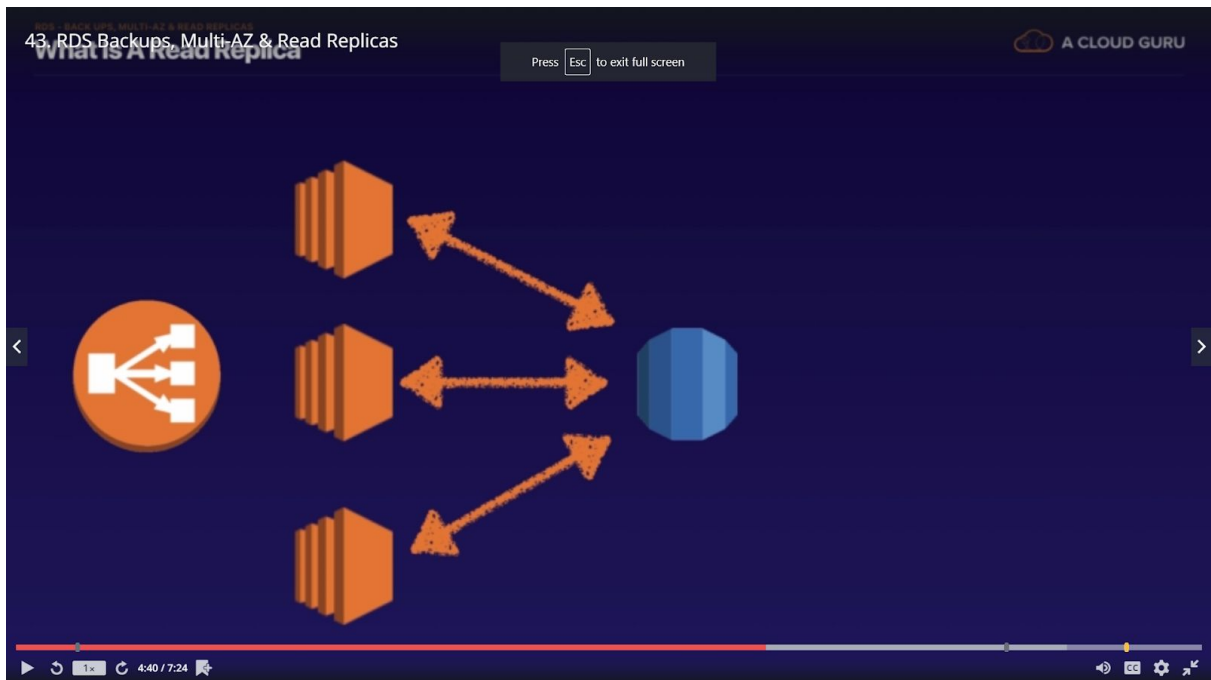
- SQL Server
- Oracle
- MySQL Server
- PostgreSQL
- MariaDB



4:25 / 7:24

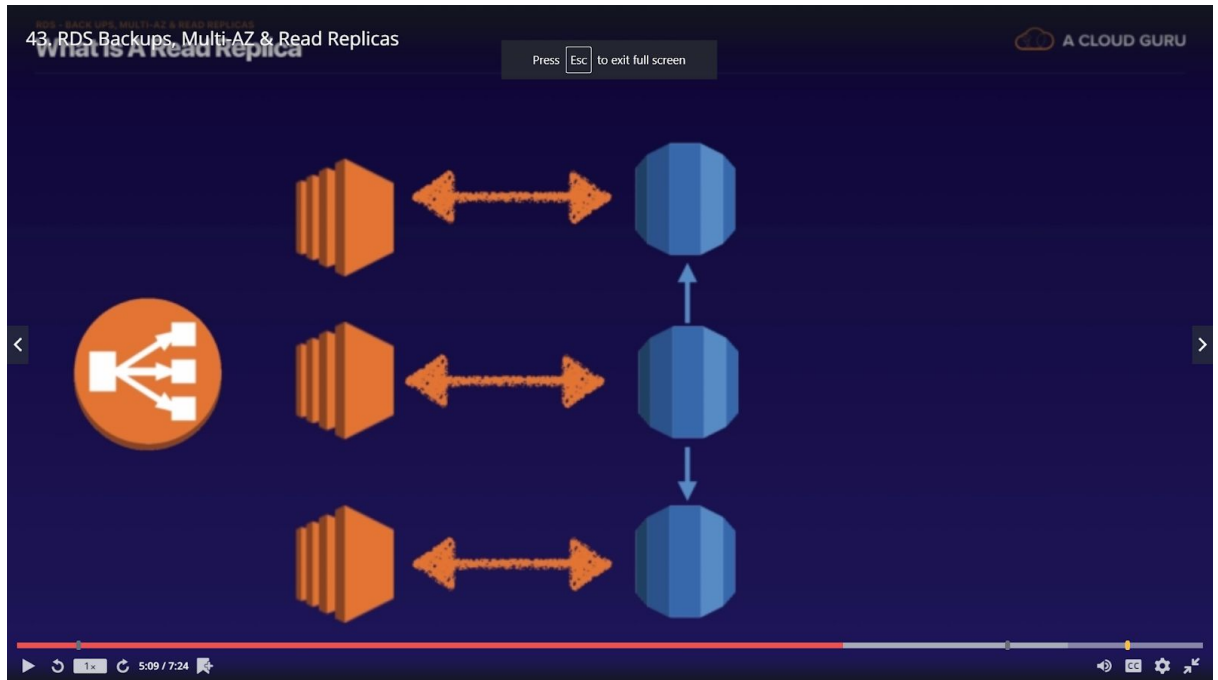
Aurora, by its own architecture, is completely fault tolerant itself.

Read Replicas :

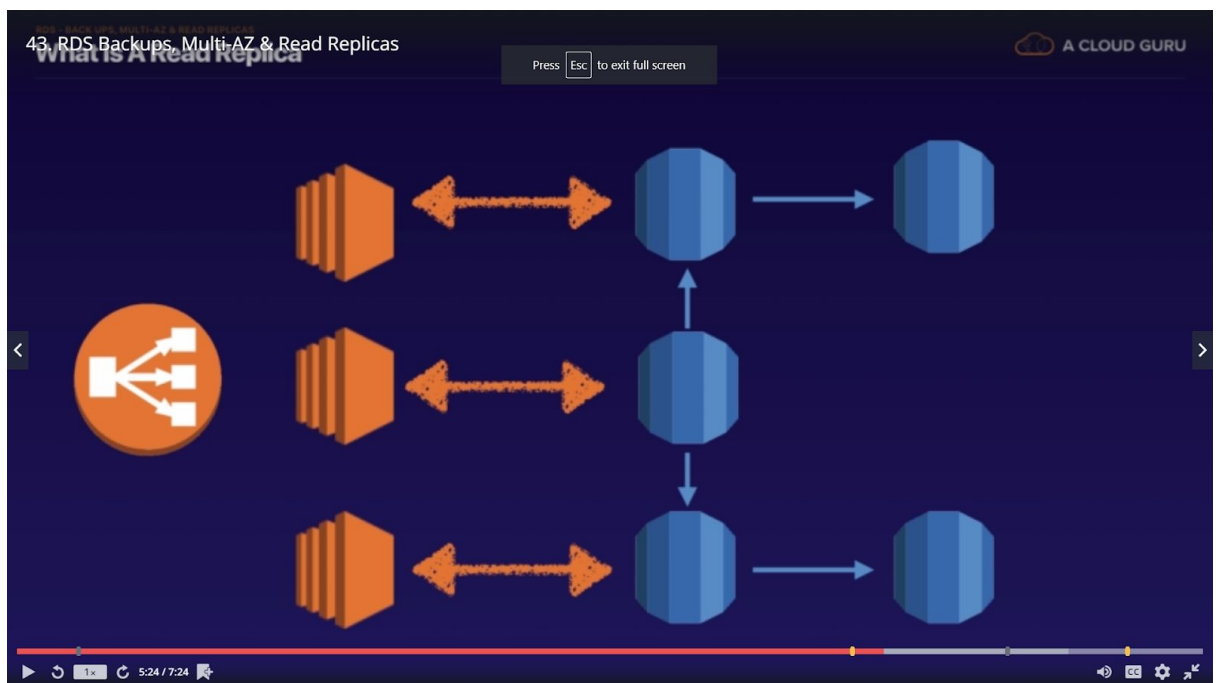


- Above, ELB is distributing load to EC2 instances, and they are then connecting to a DB instance.
- Production DB is then **asynchronously replicating** to multiple copies in the 2nd figure. Now, in the 2nd figure, all the EC2 instance are reading and writing from the same DB instance.

- **You can actually architect this so that different EC2 instances read from different read replicas, and write to a single DB, and those writes will then be replicated out.**



- You can also have Read replicas of Read replicas. Copies of read replicas :



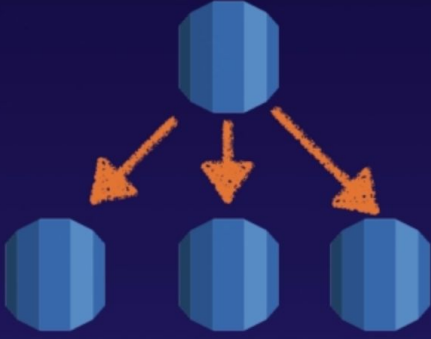
- Read replicas can also be promoted to be their own standalone instance.

RDS - BACK UPS, MULTI-AZ & READ REPLICAS

What Is A Read Replica

A CLOUD GURU

Read replicas allow you to have a read-only copy of your production database. This is achieved by using Asynchronous replication from the primary RDS instance to the read replica. You use read replicas primarily for very read-heavy database workloads.



```
graph TD; A[Primary RDS Instance] --> B[Read Replica 1]; A --> C[Read Replica 2]; A --> D[Read Replica 3];
```

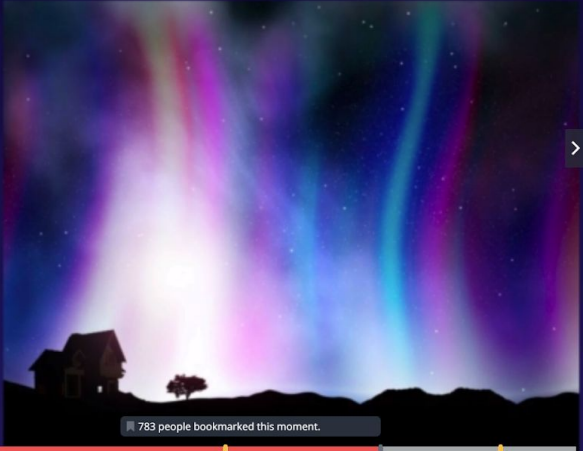
Ways to improve DB performance :

1. Use read replicas for read heavy Db workloads.
2. ElastiCache also an option for in - memory cache.

43. RDS Backups, Multi-AZ & Read Replicas

Read Replicas are available for the following databases

- MySQL Server
- PostgreSQL
- MariaDB
- Oracle
- Aurora



783 people bookmarked this moment.

TIPS Read Replicas :

- Used for scaling, not for DR.
- Must have automated backups turned on to deploy a Read replica.

- You can have 5 read replica copies of any DB.
- You can have read replicas of read replica (but watch out for latency).
- ***Each Read Replica will have its own DNS Endpoint.***
- You can now have ***Read replicas that also have Multi AZ.***
- You can create read Replicas of multi AZ source databases.
- Read replicas can be promoted to be their own databases. This breaks the replication.
- You can also have a Read replica in a second region.

44. RDS Backups, Multi-AZ & Read Replicas - Lab

Exam Tips

A CLOUD GURU

There are two different types of Backups for RDS:

- Automated Backups
- Database Snapshots

6:40 / 8:03

RDS - BACKUPS, MULTI-AZ & READ REPLICAS

Exam Tips

A CLOUD GURU

Read Replicas

- Can be Multi-AZ.
- Used to increase performance.
- Must have backups turned on.
- Can be in different regions.
- Can be MySQL, PostgreSQL, MariaDB, Oracle, Aurora.
- Can be promoted to master, this will break the Read Replica

Udemy

MultiAZ

- Used For DR.
- You can force a failover from one AZ to another by rebooting the RDS instance.

REDSHIFT :

Redshift is AWS Data warehousing solution.

Redshift can be configured as follows :

1. Single Node : 160 gb
2. Multi Node :
 - a. Leader Node : manages client connections and receives queries
 - b. Compute node : store data and perform queries and computations. Up to 128 compute nodes.

Tips :

1. Redshift is used for business intelligence.
2. Only available in 1 Az.
3. Backups :
 - a. **Backups** for Reshift are enabled by **default** with a **1 day** retention period.
 - b. Maximum retention period is 35 days.
 - c. Redshift always attempts to maintain at least 3 copies of your data :
The original and replica on the compute nodes and a backup in Amazon S3.
 - d. Redshift can also asynchronously replicate your snapshots to S3 in another region for disaster recovery.

Aurora :

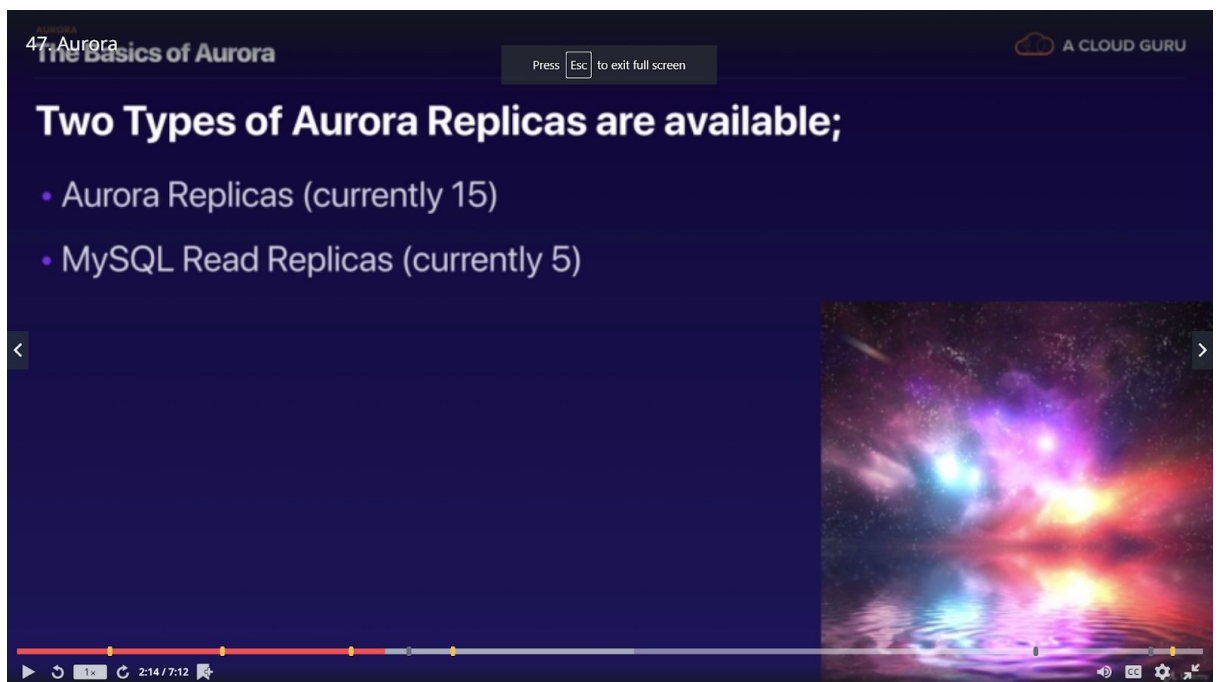
Things to know about Aurora :

- Starts with 10 GB, scales with 10 gb increments to 64 TB.
- Compute resources can scale up to 32 vCpus and 244GB of memory.
- 2 copies of your data is contained in each AZ, with minimum of 3 AZs. 6 copies of your data.

Scaling Aurora :

It is designed to transparently handle the loss of up to 2 copies of data without affecting the database write availability and up to 3 copies of data without affecting the database read availability.

Storage is self healing. Data blocks and disks are continuously scanned for errors and repaired automatically.



Feature	Amazon Aurora Replicas	MySQL Replicas
Number of replicas	Up to 15	Up to 5
Replication type	Asynchronous (milliseconds)	Asynchronous (seconds)
Performance impact on primary	Low	High
Act as failover target	Yes (no data loss)	Yes (potentially minutes of data loss)
Automated failover	Yes	No
Support for user-defined replication delay	No	Yes
Support for different data or schema vs. primary	No	Yes

Backups with Aurora :

- Automated backups are always enabled. Backups do not affect DB performance.
- You can also take snapshots with Aurora. This also does not affect performance.
- You can also share Aurora snapshots with other AWS accounts.
- 2 types of Replicas are available :
 - MySQL Replicas and Aurora Replicas
 - Automated failover is only available with Aurora replicas.
- You can do a migration from MySQL to Aurora by creating an Aurora read replica from MySQL DB instance, and then promoting the Aurora instance.