BATCH : BATCH 107

LESSON : **AWS**

DATE : 01.03.2023

SUBJECT : **Redshift-Elasticache**

techproeducation

techproeducation

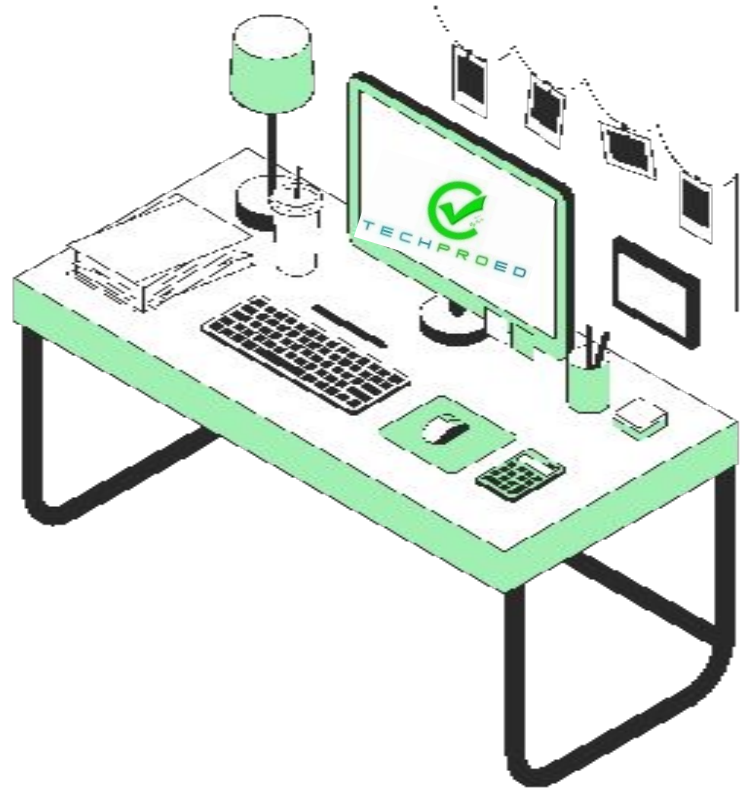
techproeducation

techproeducation

techproedu

Redshift-Elasticache

amazon
REDSHIFT



Redshift

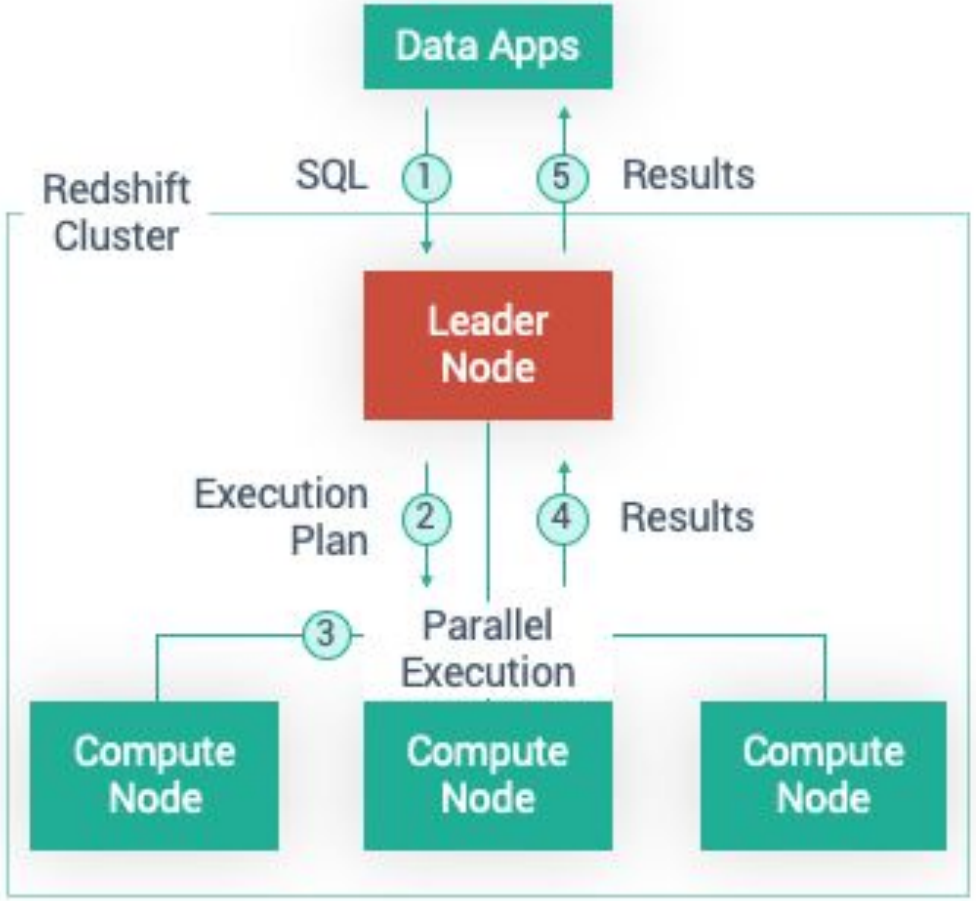
- ✔ Amazon Redshift is a fully managed, cloud-based, petabyte-scale **data warehouse** service by Amazon Web Services (AWS).
- ✔ Amazon Redshift is an efficient solution to **collect and store** all your data to **analyze**. It gives you fast querying capabilities over structured data using familiar SQL-based clients and business intelligence (BI) tools.
- ✔ Amazon Redshift also includes Amazon Redshift Spectrum, allowing you to directly run SQL queries in Amazon S3 data lakes.

Redshift

Database	Data Warehouse
	
OLTP (Online Transaction Processing)	OLAP (Online Analytical Processing)
Used for data storing and managing	Used for data analyzing

Redshift

Amazon Redshift Architecture: The Life of a Query



The 5 Steps to Process a Query

1. The cluster receives a query coming from a data app and parses the SQL in the leader node.
2. The leader node creates an execution plan that breaks a query down into a discrete sequence of steps.
3. The leader node distributes the work of executing the steps in parallel across the compute nodes.
4. The compute nodes send the results back to the leader node to merge data into a single result.
5. The leader node addresses any final sorting or aggregation and returns the results to the data app.

Redshift Features 1/3

- Amazon Redshift is a fast, fully managed data warehouse that makes it simple and costeffective
- to analyze all your data using standard SQL and existing Business Intelligence (BI)
- tools.
- RedShift is a SQL-based data WAREHOUSE used for analytics applications.
- RedShift is an Online Analytics Processing (OLAP) type of DB.
- RedShift is ideal for processing large amounts of data for business intelligence.
- Data can be loaded from S3, EMR, DynamoDB, or multiple data sources on remote hosts
- Redshift can handle petabytes worth of data. Redshift is for Data Warehousing
- RedShift uses replication and continuous backups to enhance availability and improve
- durability and can automatically recover from component and node failures.
- Redshift can only run in a 1 AZ (Single AZ)

Redshift Features 2/3

- Redshift can run via a single node or multiple-node (cluster)
- A single node is 160 GB in size
- RedShift is 10x faster than a traditional SQL DB.
- A multi-node is comprised of a leader node and multiple compute nodes
- You are bill per hour for each node (excluding leader node in multi-node)
- You are not billed fort he leader node
- You can have up to 128 compute nodes
- Redshift has two kinds of Node Type; Dense Compute and Dense Storage
- Redshift attempts to backup 3 copies of your data, the original, on compute node and on S3
- Similar data is stored on disk sequentially for faster reads

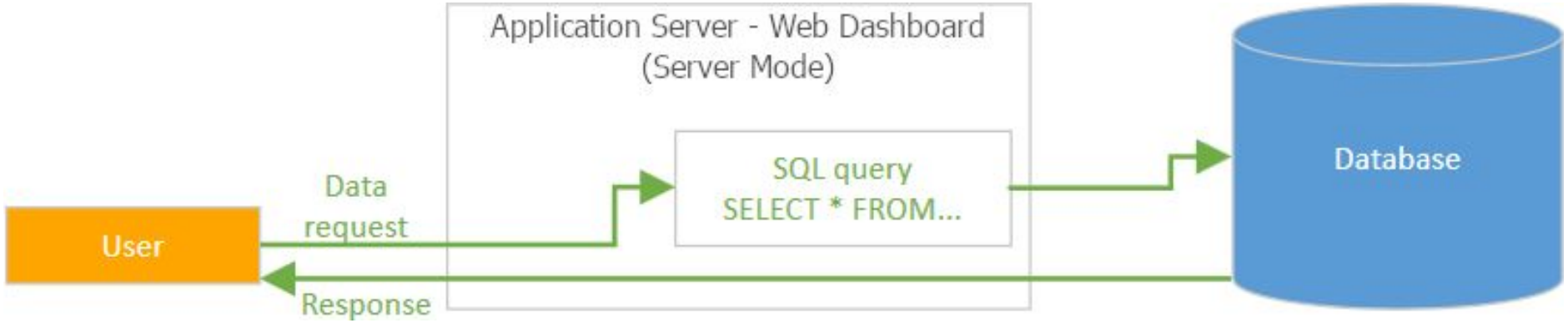
Redshift Features 3/3

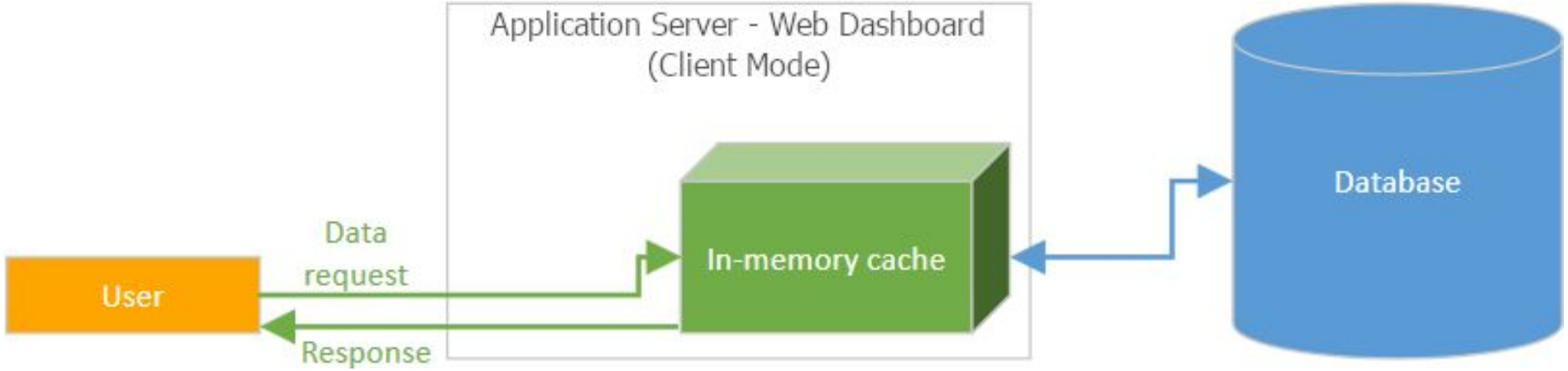
- Redshift database can be encrypted via KMS or CloudHSM
- Backup Retention is default to 1 day and can be increased to a maximum of 35 days
- Redshift can asynchronously back up your snapshot to another region delivered to S3
- Redshift uses Massively Parallel Processing (MPP) to distribute queries and data
- across all loads
- In case of an empty table, when importing Redshift will sample data to create a
- schema
- RedShift uses columnar data storage;
- Data is stored sequentially in columns instead of rows.
- Columnar based DB is ideal for data warehousing and analytics.
- Requires fewer I/Os which greatly enhances performance.


AWS
ElastiCache


TECHPROED

Elasticache

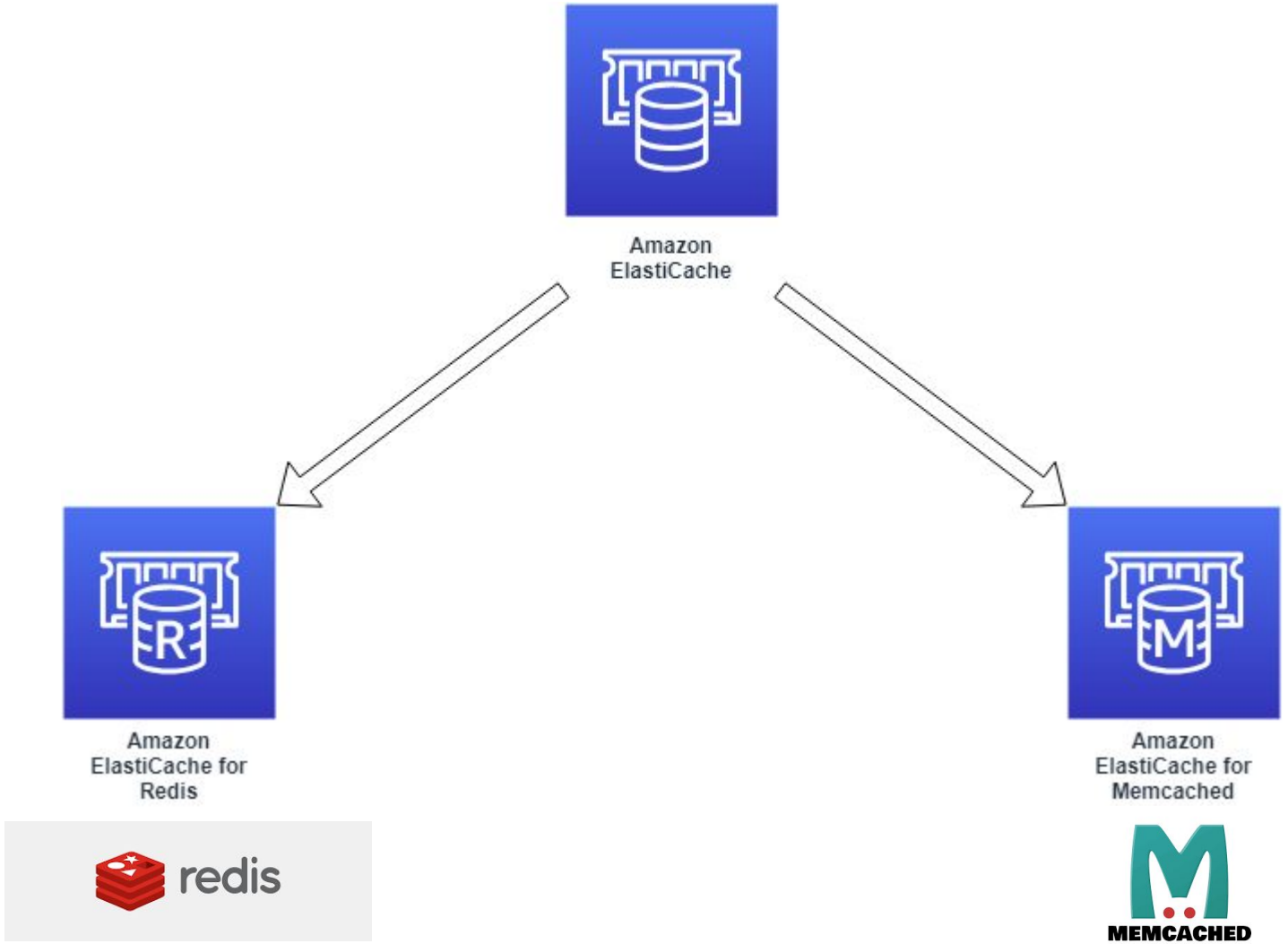




Elasticache

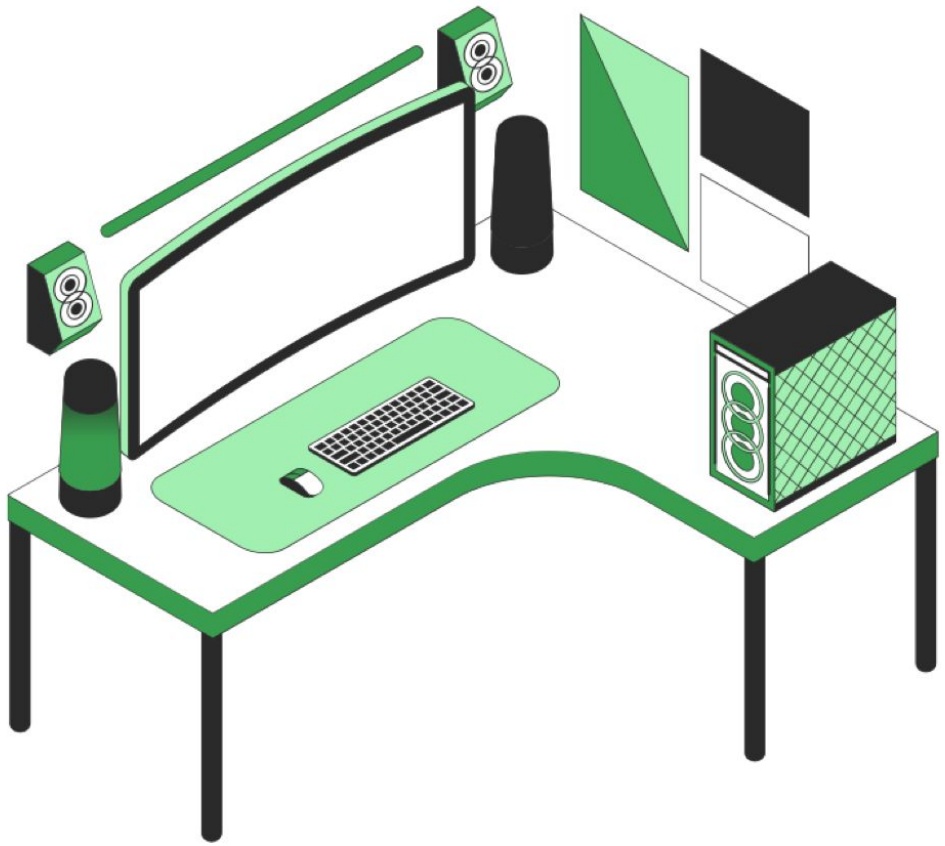
- ✔ Amazon Elasticache makes it easy to set up, manage, and scale a distributed in-memory data store or cache environment in the cloud. It provides a high-performance, scalable, and cost-effective caching solution.

Elasticache



Redis vs Memcached

Features	MEMCACHED	REDIS
★ Sub-millisecond latency ★ Developer ease of use ★ Data partitioning ★ Support for a broad set of programming languages ★ Cache clearing ★ Scaling	★ Yes ★ Yes ★ Yes ★ Yes ★ Yes ★ Yes	★ Yes ★ Yes ★ Yes ★ Yes ★ Yes ★ Yes
★ Advanced data structures	-	★ Yes
★ Multithreaded architecture	★ Yes	-
★ Snapshots	-	★ Yes
★ Replication	-	★ Yes
★ Transactions	-	★ Yes
★ Pub/Sub	-	★ Yes
★ Lua scripting	-	★ Yes
★ Geospatial support	-	★ Yes



Do you have any questions?

Send it to us! We hope you learned something new.

TECHPROED