

15-319 / 15-619

Cloud Computing

Recitation 15

December 1st 2020

Overview

- **This week's schedule**
 - Phase 3 Live test this Sunday
 - Please join [Gather.town](#) and have fun!
 - Project 4.3 will be released 1 day early on this Sunday, due next Friday.

TEAM PROJECT

Twitter Data Analytics



+



=



Phase 3 Live Test on Sunday 12/06

Time	Task	Description
4:00 pm	DNS Submission	Submit your DNS for the Live Test before the deadline
5:30 pm - 5:31 pm	DNS Validation	Validate your DNS. This is the last chance to update your DNS for the Live Test

Time	Value	Target	Weight
6:00 pm - 6:30 pm	Warm-up (Q1 only)	0	0%
6:30 pm - 7:00 pm	Q1	35000	15%
7:00 pm - 7:30 pm	Q2	12000	25%
7:30 pm - 8:00 pm	Q3	6000	25%
8:00 pm - 8:30 pm	Mixed (Q1,Q2,Q3)	12000/3000/2000	5+5+5 = 15%

All times are ET

Phase 3: Web tier Requirements

Choose one of the following as your web tier.

- ECS+Fargate
 - You only need to pay for the vCPU and memory you specified. See [Fargate Pricing](#)
- EKS+managed node group
 - Any instance type except for T series
 - Besides the normal EC2 + EBS cost, don't forget **\$0.1** cluster management fee. See [EKS Pricing](#)

You need to compare both managed services in the report

You need to provide terraform scripts for ECS

You need to provide eksctl cluster.yaml, k8s deployment & service yaml file for EKS

Phase 3: Storage tier Requirements

Choose from the following as your storage tier.

- Amazon Aurora
- Amazon DynamoDB
- Amazon DynamoDB Accelerator (DAX)
- Amazon DocumentDB
- Amazon ElastiCache
- Amazon Keyspaces (for Apache Cassandra)
- Amazon Neptune
- Amazon Relational Database Service (RDS)

You need to compare at least 2 in your report.

You only need to provide terraform scripts for the storage tier you eventually choose

You can choose multiple managed databases for your storage tier. (You need to run them at the same time for the mixed query, please calculate your budget carefully)

- e.g. DynamoDB for query 2 and Aurora for query 3

Phase 3: Live Requirements

- To get a score for a query, you **MUST** get at least 80% correctness on that query
- Hourly budget is \$1.28, including all possible costs except S3, Network, and Disk I/O
 - Storage cost should be included. e.g. 200GB for the database
 - Aurora/RDS IOPS(Disk I/O) cost will not be included, but it can be expensive for total budget
 - DynamoDB RCU cost should be included
 - DynamoDB secondary index the cost should be included
 - **[HINT]** You should place your web-tier and storage-tier in the **same availability zone** within the **same VPC** to avoid any unnecessary data transfer cost. (Don't forget that eksctl will by default create a new VPC for you.)
- EC2 / EBS is not allowed expect for the ones created by EKS managed groups
- You cannot install any Storage/Caching layer in your web tier.
 - You cannot have MySQL/Redis/Custom caching layer in your container/pod
 - You need to fetch the data from the storage tier for every valid request.

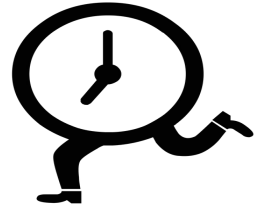
Phase 3: Summary

- RPS targets have been increased →
 - Q1: 35000
 - Q2: 12000
 - Q3: 6000
 - Mixed: 12000/3000/2000
- Terraform and deployment files are required (on Github)

Hints:

- Do mock Live Test by yourselves, including mixed queries by submitting q1, q2, q3 in TPZ at the same time.
- Monitor cloud metrics
- Please deploy resources in the same availability zone

Upcoming Deadlines



- Team Project Phase 3
 - **Live-test:** Sunday, Dec. 6th, 2020 4:00 PM ET
 - **Code and report due:** Tuesday, Dec. 8th, 2020 11:59 PM ET
- Project 4.3
 - Due **Friday**, Dec. 11th, 2020 23:59 ET