

15-319 / 15-619

Cloud Computing

Recitation 13

November 17th 2020

TEAM PROJECT

Twitter Data Analytics



+



=



Team Project Phase 2 Live Test

Top Q1 Teams

Jedi	63328.90
KuchBhiRakhde	47426.90
CloudComputing	42163.60
DoNutDisturb	42117.80
CloudProject	39899.30

Q1H

Jedi	65109.10
CCTeam	62083.10
Ayanami	59449.20
KuchBhiRakhde	54572.15
LastButNotLeast	54258.99

Q1M

Congrats to **Jedi** and **KuchBhiRakhde** for top performance for both HBase and MySQL tests.



Team Project Phase 2 Live Test

Top Q2 Teams

Jedi	49701.53
This is baseline	20234.02
Ogg3	18803.50
CC too simple	10742.13
LoadDestroyer	9428.81

Q2H

Jedi	48915.94
Ayanami	26526.30
Ogg3	19761.60
This is baseline	14567.00
A++	10794.80

Q2M

Congrats to **Jedi**, **This is baseline** and **Ogg3** for top performance for both HBase and MySQL tests.



Team Project Phase 2 Live Test

Top Q3 Teams

Jedi	36804.30
LoadDestroyer	1904.69
This is baseline	1863.56
Bonaparte	1813.53
Ayanami	1429.07

Q3H

Jedi	37062.60
Ogg3	6012.80
Ayanami	4286.30
CloudComputing	3424.70
ThorNotFoundError	2676.96

Q3M

Congrats to **Jedi and Ayanami** for top performance for both HBase and MySQL tests.



Team Project - Phase 3

- Use only **AWS managed services** for all queries
- Development budget: \$120
 - 100% Penalty for over-budget: >\$150
- Live test:
 - Per-hour-budget: **\$1.28 (included in \$100)**
- Perform ETL on your beloved GCP and Azure

Cloud Managed Services

- Managed services remove the burden from having to operate the provisioned cloud resources.
- Management tools such as monitoring, patching, security, backup are offered as part of the service.

Team Project - Phase 3

- RPS targets have **not** been changed →
 - Q1: 32000
 - Q2: 10000
 - Q3: 1500
- For web-tier, you could use either ECS or EKS.
- For storage-tier, you could use any managed services we gives in the writeup.
- Rule of thumb:
 - If you want to start any resources on EC2 dashboard, **stop**
 - If you are doing `sudo apt install mysql-server`, **stop**
- Teams should explore the managed services provided by AWS to come up with a solution
- Teams are required to use Terraform (unless Terraform does not have support for your particular managed service)

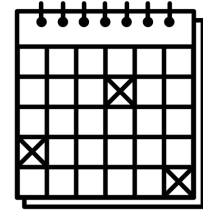
Team Project General Hints

- No EC2 instances and EBS volumes in the live test!
 - Nonetheless, you can use those to do verification or comparison to the hosted service you built before in the development process.
 - Note that ECS and EKS are backed up by EC2 instances, so it's perfectly normal to see EC2 instances in dashboard after you fire up your containers.
- Don't use T-family instances for ECS or EKS, as their performance are not consistent.

Team Project General Hints

- One option would be to split the services into web-tier and storage-tier and choose different managed services.
 - If so, the compatibility of these two services should be taken into account.
- Consider the different characteristics of queries to decide what kind of managed services to use.
- High performance/cost ratio is valued.
 - Try your best to achieve the highest possible ratio.

Team Project Time Table



Phase (and query due)	Start	Deadlines	Code and Report Due
Phase 1 <ul style="list-style-type: none"> Q1, Q2 	Monday 10/12/2020 00:00:00 ET	Q1 Checkpoint: Sunday 10/18/2020 23:59:59 ET Q1: Sunday 10/25/2020 23:59:59 ET Q2 Checkpoint: Monday 10/25/2020 23:59:59 ET Q2: Sunday 11/01/2020 23:59:59 ET	Tuesday 11/03/2020 23:59:59 ET
Phase 2 <ul style="list-style-type: none"> Q1, Q2, Q3 	Monday 11/02/2020 00:00:00 ET	Sunday 11/15/2020 18:00:00 ET	
Phase 2 Live Test (Hbase AND MySQL) <ul style="list-style-type: none"> Q1, Q2, Q3 	Sunday 11/15/2020 18:00:00 ET	Sunday 11/15/2020 23:59:59 ET	Tuesday 11/17/2020 23:59:59 ET
Phase 3 <ul style="list-style-type: none"> Q1, Q2, Q3 (Managed services) 	Monday 11/16/2020 00:00:00 ET	Sunday 12/06/2020 15:59:59 ET	
Phase 3 Live Test <ul style="list-style-type: none"> Q1, Q2, Q3 (Managed services) 	Sunday 12/06/2020 18:00:00 ET	Sunday 12/06/2020 23:59:59 ET	Tuesday 12/08/2020 23:59:59 ET

Upcoming Deadlines

- Team Project: Phase 2
 - Code and report due:
 - Tuesday, Nov 17, 2020 11:59 PM ET
- Project 4.2: Machine Learning on the Cloud
 - Due Sunday, Nov 22, 2020, 11:59 PM ET
- Team Project: Phase 3
 - Live-test:
 - Sunday, Dec 06, 2020 6:00 PM ET
 - Code and report due:
 - Tuesday, Dec 08, 2020 11:59 PM ET

Good Luck and Have Fun!