

TEAM PROJECT

Twitter Data Analytics



Team Project Phase 2 Live Test

Top Q1 Teams

PanDa	73834.43
Interstellar	59690.21
MacUserNoXX	56227.48
s21survive	53671.11
AyoJin	47940.98

Q1H

AceCC	74316.76
Hello	74051.99
AyoJin	72166.82
PanDa	71107.69
CloudCrew	67603.62

Q1M

Congrats to PanDa, AyoJin for top performance for both Q1H and Q1M tests.



Team Project Phase 2 Live Test

Top Q2 Teams

MacUserNoXX	25051.42
cloudcompewpew	18817.51
CloudCrew	17325.59
Chong	16574.74
rapidbug	14579.25

Q2H

Chong	39271.64
TheQuinjet	25557.03
MacUserNoXX	25067.37
OnePiece	24954.83
CloudCrew	24616.03

Q2M

Congrats to Chong, CloudCrew, MacUserNoXX for top performance for both HBase and MySQL tests.



Team Project Phase 2 Live Test

Top Q3 Teams

cloudcompewpew	12421.05
CloudCrew	5145.9
Tuixue	3219.06
Chong	2997.7
rapidbug	2776.49

Q3H

Chong	10527.79
Cumulonimbus	7483.56
OnePiece	6722.71
rapidbug	6042.26
TheQuinjet	5365.66

Q3M

Congrats to Chong, rapidbug 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**
 - Perform ETL on your beloved GCP and Azure

Team Project - Phase 3

- RPS targets have been changed (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
- Monitor cloud metrics
- Please deploy resources in the same region

Team Project - Phase 3

- Teams should **NOT** use any EC2 instances or EBS volumes
- Use AWS Managed Services (ECS Fargate and EKS for web tier, and managed databases as storage tier)
- Rule of thumb:
 - See resources in the EC2 dashboard besides managed node groups? NOT ALLOWED
<https://console.aws.amazon.com/ec2/v2/home?region=us-east-1#Instances:>
<https://console.aws.amazon.com/ec2/v2/home?region=us-east-1#Volumes:>
 - 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

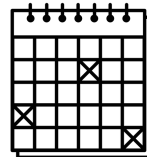
Team Project General Hints

- 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 2 Q1, Q2, Q3	Monday 03/29/2021 00:00:00 ET	Q3 Early Bird Bonus: Sunday 04/04/2021 23:59:59 ET Phase2 Due: Sunday 04/18/2021 15:59:59 ET	
Phase 2 Live Test (Hbase AND MySQL) Q1, Q2, Q3	Sunday 04/18/2021 16:00:00 ET	Sunday 04/18/2021 23:59:59 ET	Tuesday 04/20/2021 23:59:59 ET (upload PDF report and verify your submission)
Phase 3 Q1, Q2, Q3 (Managed services)	Monday 04/19/2021 00:00:00 ET	Sunday 05/02/2021 15:59:59 ET	
Phase 3 Live Test Q1, Q2, Q3 (Managed services)	Sunday 05/02/2021 17:00:00 ET	Sunday 05/02/2021 23:59:59 ET	Tuesday 05/04/2021 23:59:59 ET

Upcoming Team Project Deadlines

- Team Project: Phase 2
 - Code and report due:
 - Tuesday, April 20, 2021 11:59 PM ET
- Team Project: Phase 3
 - Live-test:
 - Sunday, May 2th, 2021 3:59 PM ET
 - Code and report due:
 - Tuesday, May 4th, 2021 11:59 PM ET

Best Wishes and Have Fun!