# Cloud Computing Project Report

# Choice of Front End

## Grizzly or Tomcat zhuolinl :

1. the design of the front end system
2. code location

## Load Balance Choice yinsuc

1 code location

3. The type of instances used and justification

4. The number of instances used and justification

5. Other configuration parameters used

6. The cost per hour for the front end system

7. The total development cost of the front end system

8. The throughput of your front end for the given workload of q1 queries.

# Backend Choice

## Basic design (Mysql or HBase) qianm

Assume we put all data in the backend and get the not so good result

1 The design of the back end system (why hbase)

2. The table structure of the database, justify your design decisions

3. The type of instances used and justification

4. The number of instances used and justification

5. The cost per hour for the back end system

6. The spot cost for all instances used

7. The total development cost of the back end system

# ETL method yinsuc

1 code location

2 The programming model used for the ETL job and justification

3. The type of instances used and justification

4. The number of instances used and justification

5. The spot cost for all instances used

6. The execution time for the entire ETL process

7. The overall cost of the ETL process

8. The number of incomplete ETL runs before your final run

9. Discuss difficulties encountered

10. The size of the resulting database and reasoning

11. The time required to backup the database on S3

12. The size of S3 backup

# Optimize for Throughput and Latency

## Large or Small qianm

data size: backend, we first put all our data in backend and find we actually don’t need so much data

1. The insight from Step 4 to influence optimizations

2. The optimizations utilized and justifications

3. Changes to the overall system design

## Large data choice qianm

The design of the back end system

The table structure of the database, justify your design decisions

The type of instances used and justification

The number of instances used and justification

The cost per hour for the back end system

The spot cost for all instances used

The total development cost of the back end system

## Small data ETL qianm

etl method and code location

## Small data choice zhuolinl

code location

The design of the back end system

The table structure of the database, justi fy your design decisions

## Optimization Algorithm zhuolinl

## Optimization Result yinsuc

5. For the given workload

a. The maximum throughput of the optimized system for q1, q2, q3 & q4

b. The latency of your optimized system for q1, q2, q3 & q4

6. The cost per hour of your optimized system at low and high load

# System Test zhuolinl

The design of the overall system and difficulties encountered

2. The AWS resources utilized to realize your system on A WS

3. For the given workload

a. The maximum throughput of your overall system for q1, q2, q3 & q4

b. The latency of your overall system for q1, q2, q3 & q4

4. The cost per hour of your system at low and high load

# Provision, Load and Prepare for Live Test yinsuc

1. The IP address of your system

2. Final configuration of each part (instance type and number)

3. Estimated per hour cost of the web service during the test period