

Assignment 1

CSCI 5410 (Serverless Data Processing)

Date Given: May 15, 2020

Due Date: May 25, 2020 at 11:59 pm

Late Submissions are not accepted, and will result in scoring "0" in the assignment.

To avoid any additional charges for resource consumption - Delete the AWS S3 storage, and AWS MySQL after fulfilling the assignment submission requirements

Objective:

This assignment covers some basic concepts of cloud computing and services. The primary objective of this assignment is to introduce you to the cloud computing platform and perform a cloud computing literature review.

Plagiarism Policy:

- This assignment is an individual task. Collaboration of any type amounts to a violation of the academic integrity policy and will be reported to the AIO.
- Content cannot be copied verbatim from any source(s). Please understand the concept and write in your own words. In addition, cite the actual source. Failing to do so will be considered as plagiarism and/or cheating.
- The Dalhousie Academic Integrity policy applies to all material submitted as part of this course. Please understand the policy, which is available at: https://www.dal.ca/dept/university_secretariat/academic-integrity.html

Grading Scheme:

- A. Research: 20 points
 - a. Completeness of tasks: 5 points, (if all parts are completed a,b,c,d)
 - b. Research quality: 10 points, (if the context is captured correctly)
 - c. Quality of document: 5 points (sentence formation, categorization, grammatical errors etc.)
- B. AWS S3 storage experiment: 10 points
 - a. Completeness of tasks: 5 points (if all parts are completed)
 - b. Submission: 5 points (if submission requirement is fulfilled)
- C. AWS RDS database service experiment: 10 points
 - a. Completeness of tasks: 5 points (if all parts are completed)
 - b. Submission: 5 points (if submission requirement is fulfilled)

Tasks:

This assignment has three parts. Part A has a small reading task, and part B, part C have small programming tasks:

Part A. Read the attached paper “A1(5410)_Reading.pdf”, and write the central idea discussed in the paper.

- It can be approximately 1 page summary, and should be written in your own words. The summary should include - (a) what the authors have presented in the paper, (b) if any specific issue is addressed, (c) if any experiments or studies performed, (d) analysis or findings made by the authors.

Part A - Submission requirement: A pdf file with the summary

Part B. AWS S3 Storage experiment:

Using AWS Educate account or AWS starter account, perform the following
take screenshots at every step:

- a. Create a S3 bucket from management console. Once it is done, create a text file (empty file) in your computer and rename it with your “First Name”. e.g. “Alice.txt”.
- b. Explore AWS SDK for Python - and using python program, upload the file on the S3 bucket you created.
- c. Create a second bucket in AWS S3 using python, and programmatically change the access permission, “disable public access”. In addition, programmatically change the ACL write option to “no” for bucket owner.
- d. Try to move (using program/script) the file from 1st bucket to 2nd bucket.
- e. Create a flowchart using draw.io/ word or any similar tool to show the steps that you have performed in this experiment.

Note: If you are not a Python programmer, you can select Java as your programming environment for this question.

Part B - Submission requirement: A pdf file with the (i) flowchart, (ii) a paragraph on your overall observation, (iii) screenshots of the S3 buckets and operations (capture all steps) (iv) copy-paste the program script in the pdf. (Separate source code submission is NOT required)

Part C. AWS RDS database service experiment:

Using AWS Educate account or AWS starter account, perform the following:
take screenshots at every step

- a. Using AWS RDS - Create MySQL DB instance
- b. Create a single table database with two fields (userID, and Password)

- c. Write a Python program or Java program to insert an ID and Password into the database.
- d. In the same program, you need to write another block/function/method to retrieve password based on the given ID.
- e. You need to use python-mysql connector or JDBC driver to connect to the database.
- f. The password must be encrypted before inserting to the database. To encrypt password, use the lookup table provided, however, first store the table in S3 bucket, and access from there.

[Hint: Consider each password contains only lower case English alphabet ('a' to 'z'). In the given lookup table, 1st column represents the original alphabet, and the 2nd column represents its "replacement". E.g. if password contains "a", you need to replace it with "xt".]

Part C - Submission requirement: A pdf file with the (i) screenshots of the S3 bucket with Lookup table, (ii) screenshots of the RDBMS, (iii) also copy-paste program script in the pdf. (Separate source code submission is NOT required), (iv) output – displays password before encryption, after encryption, and after decryption.