Cloud Computing Final Project

SIT233/SIT706 Cloud Computing (*updated 19th April 2024*)

I. OVERVIEW

This assessment is a multi-graded task that requires students to deploy a highly availability WordPress application in the cloud and explore the alternative services that can be used to achieve the same result.

You will use the **AWS Learner Labs** to work on this final project. You will deploy the required infrastructure for this application via AWS console, and then automate the deployment with *Infrastructure as Code* (for **D** and **HD** grades). **Note:** *You will be provided with an invitation link to join AWS Learner Labs. If you didn't receive the invitation, please contact the Unit Chair.*

You will also research and design a solution using suitable services based on the requirements specified and present your solution as a *video presentation* (for both **SIT233** and **SIT706** students targeting **D** and **HD** grades), on the platform of your preference (e.g., DeakinAir, YouTube, etc). Please include the URL for your presentation in a separate section at the end of the report with the section title: **"Presentation"**.

NOTE: Please indicate clearly whether you are a SIT233 or SIT706 student by including this sentence at the start of your Abstract section in your report:

"This document forms part of the assessment for [SIT233 or SIT706 - choose one] Cloud Computing".

II. FORMATTING REQUIREMENTS

Your submission to OnTrack must meet the following basic requirements:

- Formatted using the **IEEE Style Template** (*Note: It is not required to create your report using LaTeX, but the report must be as per the IEEE Template. You can choose to use it in single-column or double-column mode*). The IEEE LaTeX and Microsoft Word RTF templates can be found at: https://www.ieee.org/conferences/publishing/templates.html.

- If you choose to use Overleaf, you can use these templates:
  - Single-column: https://www.overleaf.com/latex/templates/one-column-ieee-journal-article/ykctxqgxptrs
  - Double-column: https://www.overleaf.com/latex/templates/ieee-conference-template/grfzhhncsfqn

- External sources are appropriately cited and referenced with the **IEEE Referencing Style**

- Report length (page numbers) may vary depending on the scope you cover (and your target grade). There is no hard requirement on the page limit. Although the page length does not necessarily reflect quality, but you will able to cover more with a longer page length and more comprehensively achieve the criteria specified in the rubric.

- Converted to **Portable Document Format (PDF)** format

III. TASK DETAILS

A. Phase 1: Deploy WordPress Application *(SIT233 and SIT706 students)*

In Phase 1, you will login to the AWS console and deploy a highly availability WordPress application that is hosted on AWS EC2 and RDS. The design diagram is shown below. (**Note**: the ap-southeast region may not be available in your account. In that case, you can use the default provided region.)

9 Challenge Lab
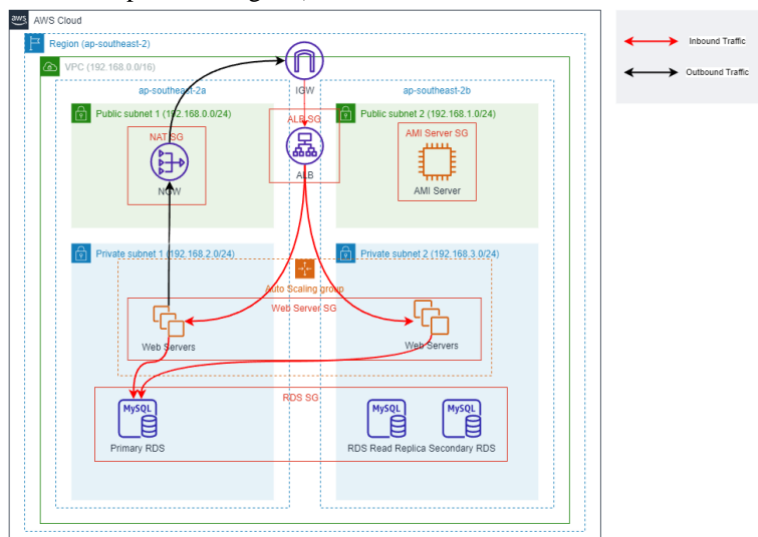
1. **VPC**

A customized VPC with private and public subnets and appropriate routing to NAT and IGW is used to host this application.

-   Create a VPC as shown in the design diagram above
-   Create a NAT gateway in the public subnet 1
-   Create the appropriate Rout Tables for private and public subnet, and associate them with the subnets

2. **RDS**

A MySQL RDS with the following configurations:

-   DB engine version: MySQL 8.0.35
-   Template: Free tier
-   Public access: No
-   Security Group should only allow necessary traffic from the Web Server
-   Create a new subnet group with private subnets only
-   After the RDS is created, enable Multi-AZ from the console
-   Create a read-replica in different AZ than the primary RDS   Module 9 Guided Lab

3. **ELB**   Module 9 Guided Lab

An Application Load Balancer is used to distribute the traffic between the backend Web Servers.

-   Create an Application Load Balancer and place it in both public subnet 1 and public subnet 2
-   Create a HTTP listener and forward the traffic to a target group
-   Create an ELB SG to allow necessary traffic from the internet

4. **S3**

Create an S3 bucket with public access disabled (**Note:** You may not be able to disable public access in your Learner Labs account. In that case, leave it as enabled.)

5. **EC2 and WordPress**

Launch the AMI server with the following configurations:

-   Amazon Machine Image: Amazon Linux 2 AMI (HVM), SSD Volume Type
-   Instance type: t2.micro
-   Create a new key pair for this EC2
-   Security Groups should only allow necessary traffic types for the web server to work publicly
-   Attach an Instance profile created in the IAM role section
-   VPC and subnet: as show in the design diagram
-   Apply the following Advanced Details – User data:

```bash
#!/bin/bash
yum -y update
yum -y install php httpd mysql
PHP_VERSION=`php -v | head -n 1 | awk '{print $2}' | awk -F "." '{print $1}'`
while [  ${PHP_VERSION} -ne 7 ]
do
amazon-linux-extras install php7.4 -y
PHP_VERSION=`php -v | head -n 1 | awk '{print $2}' | awk -F "." '{print $1}'`
done
yum -y install php-mbstring php-xml
wget http://wordpress.org/latest.tar.gz -P /tmp/
tar zxvf /tmp/latest.tar.gz -C /tmp
cp -r /tmp/wordpress/* /var/www/html/
chown apache:apache -R /var/www/html
systemctl enable httpd.service
systemctl start httpd.service
```

After your EC2 is launched, wait for a few minutes until the Status Check to become "2/2" and the instance state as "running". After that, browser to http://{EC2-Public-IP} in the browser, confirm that you can see the WordPress initial setting screen.

Configure the WordPress website
- Access your WordPress website in the browser, and start configuration
- Fill in the credentials you used to create the RDS
- Run the installation and make sure your database connects successfully

Module 9 Guided Lab

Now you have an "AMI Server" instance with WordPress configured to connect to the RDS. Create an AMI from this server, which will be used to create the AMI for Auto Scaling Groups.

- Create an EC2 instance in the public subnet 2 with the following configurations:
  o Amazon Machine Image: Amazon Linux 2 AMI (HVM), SSD Volume Type
  o Instance type: t2.micro
  o Create a new key pair for this EC2
  o Security Groups should only allow necessary traffic types for the web server to work publicly
  o Attach an Instance profile created in the IAM role section
- On the EC2, configure the same WordPress settings

6. **S3 Integration**

The WordPress site to an S3 bucket for storing media files using WordPress plugin - **Offload Media by Acowebs**

- Search for "OFFLOAD MEDIA" in Plugins > Add New
- After installing, 'activate' the plugin and go to settings.
- To configure this plugin, you would need AWS Access Key and AWS Secret Key which can be found by pressing "Details" button in AWS Academy.
- Add your bucket details to the plugin configuration.
- Upload a photo under media and make sure the file path is pointing to S3
- After the AMI is ready, create a Launch Template with the following configurations:
- Use the AMI your just created
- Instance type: t2.micro
- Use the key pair your just created
- Create a new Security Group that only allows necessary traffic from the ELB

7. **Auto Scaling Group** *Module 9 Guided Lab*

You need to define a scaling policy for your auto-scaling group with at least the following rules:

- The minimum number of servers is 1
- The maximum number of servers is 3
- Desired capacity of 1
- Configure a simple scaling policy to:
  o Scale out when the average CPU utilisation of your ELB target group is above 70%
  o Scale in when the average CPU utilisation of your ELB target group is below 25%
- The ASG should launch instances into the private subnets.

8. **Create the infrastructure in this phase with CloudFormation**

You should break the infrastructure into two CloudFormation templates since they share different lifecycles:

- Networking related services - VPC, NAT, IGW, Route Tables etc.
- Application infrastructure – EC2, Auto Scaling Group, RDS etc.

Note: Do not delete the resources that you created manually, keep those running for marking, test your CloudFormation template to deploy another set of resources in a different VPC.

9. **Testing** *Application Load Balancer*
- Test you can access your WordPress page via the ALB
- Terminate the Web Server and check if the ASG creates a new one *Auto Scaling Group*
- Check you can connect to the Web Server via the Session Manager
- Upload a photo to the WordPress media library and make sure it has been uploaded to the S3 bucket
- Check your RDS configuration and make sure Multi-AZ is enabled, stop your RDS instance and observe the change of Secondary Zone (there might be a few minutes of delay to display the change in console)

Module 9 Guided Lab

B. Phase 2: Project Implementation and Demonstration (SIT706 students only)

In Phase 2, you will propose and design an alternative solution that meets the business scenario given below. High Distinction students will be expected to demonstrate an implementation of the proposed solution.

Business Scenario

You have a client that is migrating their shopping website onto AWS. As part of the migration, the company wants to refactor the application. The following problems/requirements has been identified by the company:

- The company anticipates that the demand for its application will continue to increase, as it has been doubling annually. While the exact extent of growth is uncertain, the company expects this trend to persist for the next a few years at minimum and wants the architecture to be able to accommodate this expansion.
- The application must have lowest possible latency. Data from each user's shopping cart needs to be highly available. User session data must be available even if the user is disconnected and reconnects.
- The company aims to implement a serverless/event-driven approach as much as feasible.
- The company aims to utilize managed cloud services as much as possible in order to reduce the requirement for internal systems administration.
- The current database tier of the application utilizes a relational database, which is relatively slow and expensive to operate. Due to its simple table structure, the company is interested in investigating more cost-efficient alternatives.
- All data in the application need to be encrypted in transit and at rest.
- As the application is expanding to other countries, it has been observed that response times in countries outside of Australia have been relatively slow. The company aims to improve global response times.
- The application design must meet a recovery point objective (RPO) of 15 minutes and a recovery time objective (RTO) of 1 hour.

Architecture Design

You are tasked with creating and documenting a design that addresses the requirements mentioned above. The design document you will produce should comprise of the following elements:

- A visual representation of the architecture, highlighting the cloud services employed and the connections between them. You can use tools such as draw.io to create the diagram
- A comprehensive explanation and rationale for the services selected for the design. For instance, if you choose to use Elastic Container Service (ECS), you will need to explain its function and how it supports the overall design.
- Detailed elaboration of how the business scenario is addressed by the proposed services
- Provide reasons for why a particular service was selected over others (for example, the use of EC2 instead of Lambda or ECS)

IV. PROJECT REPORT STRUCTURE

It is recommended you use the following report structure. You can add new sections/sub-sections as you see fit, depending on the breadth and depth of your project.

For **SIT233** students, you may use the following structure:

1) **Title:** Provide a suitable title you see fit (be creative, otherwise you can simply use *"SIT233 or SIT706 Final Project Report"* as your title.
2) **Abstract:** Provide a brief overview of your project report.
3) **Introduction:** Introducing your project, including its objectives and motivations.
4) **Design Diagram:** Present the design requirements of the project with appropriate use of diagrams.
5) **Implementation:** Present how you implemented the project in detail with description of screenshots of the system you built along the way, including the commands/scripts used, and the final outcomes.
6) **Discussion and Reflections:** Discuss the final outcomes of the system (what worked and what didn't), and your reflections on how the project can be implemented and beneficial in real-world businesses.
7) **Conclusion:** Summarise and conclude your project report.
8) **References:** Include authoritative reference sources here.

For **SIT706** students, add two new sections to the above structure before the Conclusion section:

7) **Business Scenario Overview:** Discuss the business scenario you have been presented and how you would provide/propose a solution based on the scenario.

8) **Architecture Design:** Present your architecture design diagram here. Present the technical implementation in this section if you are aiming for HD.

V. ASSESSMENT

You will be graded according to the **Marking Rubric** on the last page of this task sheet.

VI. SUBMISSION

Please submit your document in PDF to OnTrack by the specified due date.

*** Please see the next page for Marking Rubric ***

Marking Rubric	Pass	Credit	Distinction	High Distinction
Format	- The submitted report is formatted using the IEEE Conference template - All figures/tables are appropriately labeled - The submitted report is in PDF format	All **Pass** requirements plus: - PDF fonts compiled into document - Images/figures in vector format	All **Credit** requirements plus: - Formatting is clean with no words/tables wrapping beyond the edge of a column/page	All **Distinction** requirements
Structure	- The submitted document is properly structured in the format of an IEEE paper - The topic is divided into a simple but clear structure using the section headings detailed in Section IV - The content is complete but some sections are lacking required detail - A reference list is provided and correct	All **Pass** requirements plus: - Clear use of sub-sections as required to clearly delineate different aspects of the project - All sections have a sufficient level of detail - The reference list is professionally structured and complete	All **Credit** requirements plus: - Suitable references have been located and used for all claims made by the student	All **Distinction** requirements
Technical Content – Phase 1	- VPC and subnets are correctly configured with correct routes and internet access - WordPress website is configured and can be accessed from either a public IP address or the ELB public DNS - RDS is configured and working correctly	All **Pass** requirements plus: - Website is accessible via the ELB public DNS - Auto-scaling group is configured and working correctly - IAM role is created and attached to EC2 - Photos can be uploaded to a private S3 bucket successfully - RDS is created in the correct subnets with Multi-AZ and Read Replica - Web servers are created in the correct subnets	All **Credit** requirements plus: - Security Groups are configured following least privilege - Auto Scaling Policy is configured correctly - S3 bucket is configured as private with the correct bucket policy - CloudFormation template is created for infrastructure provisioning, and captures most of the technical requirements and a working	All **Distinction** requirements plus: - CloudFormation template is created for infrastructure provisioning, and captures all of the technical requirements and a working

Marking Rubric	Pass	Credit	Distinction	High Distinction
Technical Content – Phase 2 **(SIT706 students only)**	• Basic proposed solution for the business scenario	All **Pass** requirements plus: • A basic architectural design of the proposed solution	All **Credit** requirements plus: • Most of the requirements from the business scenario are captured in the design • Good understanding of the services used • Good architectural design of the proposed solution	All **Distinction** requirements plus: • All of the requirements from the business scenario are captured in the design • In-depth understanding of the services used • Excellent proposed solution and architectural design • A good implementation to demonstrate the proposed solution for the business scenario • Discussions on the real-world implications and applications of the proposed solution
Presentation	• No video presentation required	• No video presentation required	Presentation required: • Engaging presentation, with good explanation of the solution design and reasonable understanding of the services used	All **Distinction** requirements plus: • <mark>Very engaging presentation</mark>, with <mark>excellent explanatio</mark>n of the solution design and in depth understanding of the services used
Language	• Basic language and grammatical skills	All **Pass** requirements plus: • Good grammatical structure and flow of argument	All **Credit** requirements plus: • A document suitable for reading by a professional audience	All **Distinction** requirements plus: • <mark>An excellent report suitable for reading by an professional and academic audience</mark>