*Student name: Le Thuy Hang Nguyen*

*Student ID: 44499159*

*Group: 36*

Priority Based Job Scheduler in Cloud Computing

1. **Introduction**

Cloud computing, a new technology, has been gaining immense popularity over a past few years. It is a distributed, parallel and on-demand available form of computer system resources (It is a form of distributed system in which servers are virtualized), allowing access to a shared pool of configurable computing resources such as servers, applications, networks, storage space and services. Clouds provide services to their users over multiple locations from data centres. Cloud computing achieves coherence and economies of scale from resource sharing. It enables organizations to avoid up-front IT costs, improve maintenance and management ability over their applications and allows IT department to quickly meet unpredictable demands by easy adjustments of resources. The benefits of high-capacity networks, low-cost computers and storage devices has led to the high popularity of cloud computing.

The exponential growth of cloud computing poses new major challenges, network bottlenecks. It is that network gets overwhelmed when they have to deal with a large amount of data. For example, the breakdown of telecom operators due to bursting messages during spring festivals or website crashes on traditional Christmas sales. Cloud computing involves processing a diverse set of tasks and performing complex computations on a huge amount of data amongst multiple nodes. Therefore, the tasks arriving to be processed must be scheduled efficiently so that task delays or data loss can be avoided in a highly clustered network. For that reason, resource allocation and scheduling play crucial roles in cloud systems.

1. **Aim**

The aim of this project is to propose a scheduling algorithm based on priorities of the tasks in order to increase the throughput of the system, improve performance and meet user’s satisfactions. We will use metrics like computational complexity, job parallelism, resource parallelism to determine upon job assignment, which jobs need to be executed prior to others. We will evaluate our algorithm using simulation and compare them with similar algorithms to show the effectiveness of our approach.

1. **Background**

In this project, we will build a cloud job scheduler in a discrete event simulation, which adopt the client-server communication model. The server-side simulator has already been developed. It simulates a distributed system with user specified configurations and connects with the client-side simulator containing one or more scheduling policies/algorithms. In this project, the client-side simulator acts as a job scheduler while the server-side simulator simulates everything else including user (job submissions) and servers (job execution)

1. **Project plan**

The project will be conducted by Group 36 with 3 members:

* Allan Nguyen: System Developer
* Le Thuy Hang Nguyen (me): Data Analyst
* Marcus Palio:

Schedule: the cloud job scheduler will be built up through 5 stages, 1 design stage, 3 development stages and final presentation

* Stage 0: document basic knowledge about distributed systems and cloud systems; the project’s purposes, goals, backgrounds and plans
* Stage 1: design and implement a ‘vanilla’ client – side simulator with a simple job dispatcher
* Stage 2: implement three baseline scheduling algorithms: First-Fit (FF), Best-Fit (BF), and Worst-Fit (WF)
* Stage 3: design and implement a new scheduling algorithm
* Stage 4: GitHub data repository and Wikipedia page

Meetings will be weekly held at 1pm-2pm on Friday at library

Communication channel: WhatsApp group

All documentations and source code of the project are all available online at GitHub repository: [Priority Based Job Scheduler for Cloud Computing (GitHub)](https://github.com/sotiey511/Priority-Based-Job-Scheduler-in-Cloud-Computing)

Reference

Vijindra, & Sudhir, S. (2012). Procedia Engineering: Survey on Scheduling Issues in Cloud Computing