COMPUTER ENGINEERING DEPARTMENT

Graduation Project Proposal Form

CMPE 405 /CMSE 405 /BLGM 405

|  |  |
| --- | --- |
| Instructor Name: | Dogu Arifler |
| Project Title: | Redundancy Systems for Warehouse-Scale Computing |
| Number of team members: | 2 or 3 |
| Semester & Year: | Spring 2024 |
| Type of Project (HW/SW): | Systems |

# Project proposal should be in accordance with ABET requirements, which are stated as: “The curriculum must include a culminating major engineering design experience that 1) incorporates appropriate engineering standards and multiple constraints, and 2) is based on the knowledge and skills acquired in earlier course work.”

# 1. Project overview

Availability of warehouse-scale computing is critical for building large-scale Web services. Responsiveness of these services requires software-based latency variability mitigation techniques. To this end, a dispatcher can be employed to create multiple copies (also called replicas) of the jobs and send these replicas to different servers. The goal of the project is to develop software and/or models for such redundancy systems. The students will learn how warehouse-scale computing systems, such as those of Google, work.

**2. Detailed specification of the project**

A possible strategy to implement the project can be as follows: as soon as the first replica completes service on some server, all the remaining replicas are removed from the system immediately. Such an approach has been demonstrated to reduce the mean and the tail of response time. If models are developed, they should be validated via computer simulations or real testbeds.

# 3. Standards that should be used in the project

Standards and papers should be investigated as part of the project.

# 4. Constraints

Job replica sizes should not be assumed to be independent and identically distributed, but instead must be correlated at the servers in order to represent realistic situations.

# 5. Tasks to be completed by students

# Stages and tasks must be identified by the students during the first semester (CMPE 405) of the project.

# 6. Tools Required/Expected to be used in the development of the project

Students must be proficient in C or C++ or Java or Python. Students can use any of the aforementioned programming languages for implementation/modeling/simulation. Completion of CMPE 344 is required, and completion of CMPE 474 or knowledge of performance analysis and statistical methods is a plus.

# 7. Resources Required (HW/SW/Data, etc.)

# Reasonably powerful laptop computers.

# Complete list of Project Deliverables Expected at the End of Project Completion

|  |  |
| --- | --- |
| Deliverable | Description |
| Software | Solution to reduce the mean and the tail of response time |
| Model/Simulator | A model to reduce the mean and the tail of response time accompanied by proper validation |
|  |  |
|  |  |
|  |  |

(add more rows as needed)

# 9. Other related information

See the following link for warehouse-scale computing:

<https://research.google/pubs/the-datacenter-as-a-computer-an-introduction-to-the-design-of-warehouse-scale-machines/>

The following textbook is useful:

M. Harchol-Balter, *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*, Cambridge University Press, 2013.