Summer 2019

Performance Optimization of Big Data Workflows in Distributed Environments

Daqing Yun

Follow this and additional works at: https://digitalcommons.harrisburgu.edu/cisc_pti
Title: Performance Optimization of Big Data Workflows in Distributed Environments (Tentative)

Author: Daqing Yun - dyun@harrisburgu.edu

Difficulty: Hard

Specialization: Computer and Network Security

If other, please specify:

Most Appropriate Course: Project II

Brief Description: Formulate performance optimization problems of big data workflows with various objects; prove complexities of these problems; design and evaluate proposed heuristics using simulations

Number of students needed: 1

Outcomes and Deliverable: Simulation code + project report

Skills Required: Programming skills in C/C++ or other languages; Algorithm design skills;

Available Resources: Testbeds; code base; reading references

Program Goal:
CISC 1.1: Mathematical Analysis, CISC 1.2: Sound Reasoning, CISC 1.3: Develop Solution, CISC 1.4: Deploy Solution CISC 2.2: Software Platform, CISC 2.4 Data Structure, CISC 2.5 Analysis of Algorithms CISC 3.2: Explore New Design CISC 4.1: Written Communication, CISC 4.2: Oral Communications

Student Learning Outcomes:
1a: The student should be able to analyze a problem in a manner that facilitates the design of its solution. 1b: The student should be able to apply relevant principles of computing during their analysis of a problem. 2b: Student is able to develop a software solution from a formal design specification. 3a: Student will be able write in a standardized format in order to organize their thoughts and deconstruct their ideas at a level appropriate for the desired audience. 3b: Student will be able to verbally communicate effectively with an advisor, group of colleagues or an audience to express a thought or idea at a level appropriate for the desired audience.