CMPT 756 - G100 Systems For Big Data

Instructor(s): George Chow

Calendar Objective/Description:
Systems For Big Data

Instructor's Objectives:
This course will survey a variety of tools essential for developing data analytics that run at scale in modern data centres. Approximation techniques are often required to handle high volumes and velocities of data. Agile development methods allow teams to adjust their products to changes in data variety and customer needs. Latency and throughput can be improved by tailoring implementations to GPUs and wide-scale distribution. You navigate these complex choices guided by the high-level overview of the design space of distributed applications, producing scalable, fault-tolerant, efficient analytic tools.

Prerequisites:
see go.sfu.ca

Topics:
- Agile software development
- Instrumenting and monitoring performance
- Cloud computing
- Acceleration using GPU and cloud resources
- Designing approximation algorithms
- Reliability and consistency in distributed systems
- The design space of distributed applications

Grading:
Tentative grading breakdown: • Assignments (30%) • Term Project (50%) • Quizzes (20%) The Term Project, as well as some assignments, will be a group project. This breakdown is tentative. The grading and assignment policy will be finalized in the first week of class. Students must attain an overall passing grade on the Term Project in order to obtain a clear pass (C- or better).

Recommended Books:

Academic Honesty Statement:
Academic honesty plays a key role in our efforts to maintain a high standard of academic excellence and integrity. Students are advised that ALL acts of intellectual dishonesty will be handled in accordance with the SFU Academic Honesty and Student Conduct Policies ( http://www.sfu.ca/policies/gazette/student.html ).