Homework 2

Submit Assignment

Due Mar 12 by 11:59pm **Points** 100 **Submitting** a file upload
Available Feb 20 at 12am - Mar 12 at 11:59pm 21 days

Optimization

Pick **any one of the optimizations** described in the optimizations catalog paper (also covered in the class): Operator Reordering, Redundancy Elimination, Batching, Fusion, Fission, Operator Separation, Load-balancing, Load-shedding, Algorithm Selection
and implement the same benchmark application described. Try to reproduce the graph from the paper.

Guidelines

1. People with a virtual machine or laptop may be limited in their choice of optimizations (e.g. you will not be able to deal with placement).
2. Please make appropriate assumptions:
a. Define an appropriate workload (what data you will use), how you will stream it etc.
b. Define how you will measure the performance, such as throughput, selectivity etc.
Please make these assumptions clear in your writeups. Please also include a plot of the performance graph you are trying to reproduce.

Deliverables

Please put the code into a single file (or directory).
Tar/gz or zip the contents of your homework in to a single file.
Please upload your zipped submission to courseworks. Do not include binaries or jar files. If you have data files, please just include a small sample of it.
Please include a short write-up explaining your assumptions and implementation (1-2 pages).

Course Chat