Database Systems
CMPT 308

Big Data Paper Summary - 100 points

Goal

To learn about some important Big Data ideas and techniques. We've been spending our semester on relational database systems. There are others, some especially designed for dealing with more data than relational systems can traditionally handle in a reasonable amount of time. I want you to have some background in this area, and these are some of the most influential papers about those systems.

Instructions

Select one of the papers from the "Big Data Paper Summary" list in the projects section of our class web site. Read it. Sleep with it under your pillow. Read it again. Then read *A Comparison of Approaches to Large-Scale Data Analysis*. Finally, watch Michael Stonbraker's ICDE 2015 talk about his "10 Year Test of Time" paper award. Ponder life, the universe, and the database wisdom you've absorbed.

Deliverables

Analyze and summarize the papers and the talk in presentation form as follows:

- Slide 0 - paper titles and bibliographic data, your name, the date
- Slide 1- the main idea of the paper you chose
- Slide 2 - how that idea is implemented
- Slide 3 - your analysis of that idea and its implementation
- Slide 4 - the main ideas of the comparison paper
- Slide 5 - how those ideas are implemented
- Slide 6 - your analysis of those ideas and their implementations
- Slide 7 - comparison of the ideas and implementations of the two papers
- Slide 8 - the main ideas of the Stonebraker talk
- Slide 9 - advantages and disadvantages of the main idea of the chosen paper in the context of the comparison paper and the Stonebraker talk

Resources

- The papers on our web site, of course, and…
- Chapter 20 in our text

Submitting

Submit your work as a PDF. (Only PDF files will be accepted.) Push it to your GitHub repository **before** the date at which it is due (see syllabus). Remember, neatness counts. So does accuracy, spelling, grammar, correctness, and essentially everything else. It only has to be perfect.