**CSEE5590-0001/490-0003: Big Data Programming**

**Lesson Plan # 3**

**ICP Feedback and Submission Link :** <https://docs.google.com/forms/d/e/1FAIpQLSe1WigyMHd3SVLiynDleH1Njm0DWp8skiYBfE-ZJx6HhflG0Q/viewform>

**For Online students:** <https://docs.google.com/forms/d/e/1FAIpQLSe1WigyMHd3SVLiynDleH1Njm0DWp8skiYBfE-ZJx6HhflG0Q/viewform>

**Importing Libraries**

|  |
| --- |
| import org.apache.commons.logging.log;  import org.apache.commons.logging.logFactory; |

**Required Libraries (POM.XML)**

https://umkc.box.com/s/at5sr2ms080gt7lh53msce3gtoxm1lgt

**Lesson Title: *Hadoop 2***

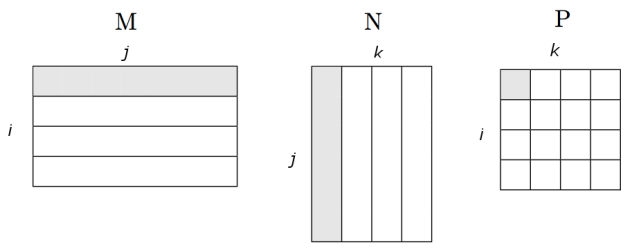
**Lesson Description: *Hadoop Distributed File System (HDFS)/ MapReduce and Big Data Applications***

**Lesson Overview:**

Overview of Hadoop and Map Reduce Paradigm. The Lesson focuses on map reduce coding exercises by actual implementation

**In class exercise:**

1. **Matrix Multiplication in Map Reduce**



Suppose we have a *i x j matrix* M, whose element in row *i* and column *j*will be denoted m_{ij} and a *j x k matrix* N whose element in row j and column k is donated by n_{jk} then the product P = MN will be *i x k* matrix P whose element in row i and column k will be donated by p_{ik}, where P(i,k) = m_{ij} * n_{jk}.

Create a Map-Reduce Program to perform the task of matrix multiplication

Marks will be distributed between logic, implementation and UI (presentation in GITHUB Wiki)

**Programming elements:**

Hadoop MapReduce and HDFS

**Reference:**

<https://lendap.wordpress.com/2015/02/16/matrix-multiplication-with-mapreduce/>

**Source Code:**

<https://umkc.box.com/s/a8rpby47fer2zeqzwdmsgokuroha7act>

**Prerequisites:**

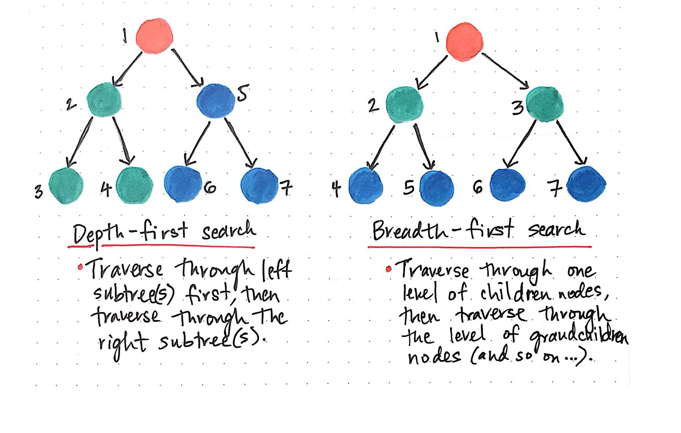
Ensure that Hadoop is installed, configured and is running. More details:

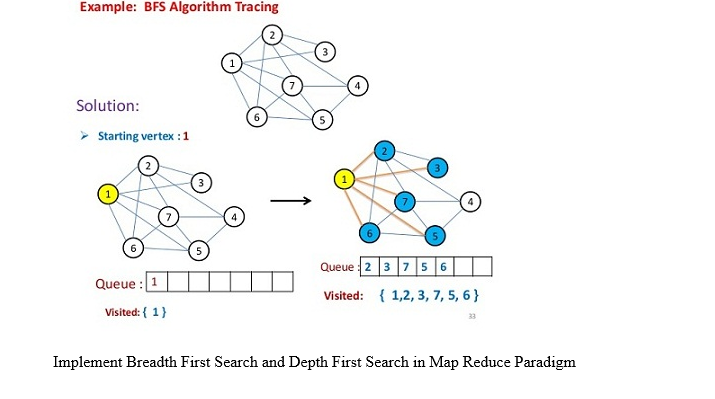
[Single Node Setup](https://hadoop.apache.org/docs/r1.2.1/single_node_setup.html) for first-time users.

[Cluster Setup](https://hadoop.apache.org/docs/r1.2.1/cluster_setup.html) for large, distributed clusters**.**

**Bonus Point:**

**Breadth First Search and Depth First Search in Graph using Map Reduce**





After completion of your ICP fill in the form. Any available TA/instructor will come to you and evaluate ICP

**ICP Guidelines (In Class Students):**

1. ICP Submission is in pairs of two students.
2. Once completed, must be presented to TA or Instructor before the completion of the class
3. Submission after class is considered as late submission. (Check the late submission policy in the syllabus)
4. ICP Code with brief explanation in wiki should be pushed to GitHub (submit your screenshots as well. The screenshot should have both the code and the output)Submit GitHub Link through the Feedback Form (<https://docs.google.com/forms/d/e/1FAIpQLSe1WigyMHd3SVLiynDleH1Njm0DWp8skiYBfE-ZJx6HhflG0Q/viewform>**)**

**Submission Guidelines (for online students):**

1. Submit your source code and documentation to GitHub and represent the work through wiki page properly with detailed explanation (submit your screenshots as well. The screenshot should have both the code and the output)
2. Comment your code appropriately.
3. Submit a brief demo video 2-3 min showing your assignment with a voice over explaining your work through the Submission Link.
4. Use the following Google link to submit your assignment

(ICP Submission Link# <https://docs.google.com/forms/d/e/1FAIpQLSe1WigyMHd3SVLiynDleH1Njm0DWp8skiYBfE-ZJx6HhflG0Q/viewform>

***Cheating, plagiarism, disruptive behavior and other forms of unacceptable conduct are subject to strong sanctions in accordance with university policy. See detailed description of university policy at the following URL:*** [*https://catalog.umkc.edu/special-notices/academic-honesty/*](https://catalog.umkc.edu/special-notices/academic-honesty/)