CSC 369 – Introduction to Distributed Computing

1. CSC 369 - Introduction to Distributed Computing

2. credit units 4  contact hours 6

3. Course Coordinator: Lubomir Stanchev

4. Textbook:
   • Karau, Konwinski, Wendell, and Zaharia, Learning Spark, Lightning-fast data analysis
   • Mahmoud Parsian, Data Algorithms Recipes for Scaling up with Hadoop and Spark

5. a. Course Description: Introduction to distributed computing paradigms and cloud computing. Modern distributed computing infrastructures. Problem-solving in a distributed computing environment. 3 lectures, 1 laboratory.

   b. Prerequisite: CSC/ 203; and one of the following: STAT 301, STAT 312, STAT 321 or STAT 350.

   c. Required/Elective/Selective Elective for CPE, CSC, SE

<table>
<thead>
<tr>
<th></th>
<th>CSC</th>
<th>CPE</th>
<th>SE</th>
<th>CDSM/DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Elective</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective Elective</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

6. a. Course Goals/Outcomes
The student will be able to:
   • Apply principles and techniques of distributed computing to the development of software to solve computational problems in a distributed environment.
   • Analyze the applicability of distributed computing techniques to an existing non-distributed solution of a computational problem and transform (reorganize) the non-distributed solution into a suitable distributed solution.
   • Explain how modern distributed computing and cloud computing infrastructures are organized.
   • Assess the results of solving a computational problem in a distributed environment.

   b. How Student Outcomes addressed
   (“B” = Basic level, “I” = Intermediate level, “A” = Advanced level)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSC</td>
<td>I</td>
<td></td>
<td></td>
<td>I</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE/CPE</td>
<td>I</td>
<td>A</td>
<td></td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. Major Topics Covered: (number of lecture hours each)
• Introduction to distributed computing and cloud computing (6)
• Building distributed algorithms using MapReduce (simple problems) (6)
• Building distributed algorithms using MapReduce (data analysis) (3)
• Building distributed algorithms using MapReduce (text processing) (3)
• Building distributed algorithms using Spark (simple problems) (6)
• Building distributed algorithms using Spark (data analysis) (3)
• Building distributed algorithms using Spark (text processing) (3)