

ASSIGNMENT 5: PARALLEL PROCESSING WITH MAPREDUCE AND BIG DATA APPLICATIONS.

Chapter 7: Parallel Processing with Map Reduce

Instructor: Houssein Dhayne (houssein.dhayne@net.usj.edu.lb)

ASSIGNMENT OVERVIEW:

The objective of this assignment is to delve into the concepts of parallel processing using MapReduce and explore its applications in the selected topic. Additionally, students will be tasked with developing a MapReduce algorithm to solve an analytical problem within their chosen field of study.

TASK

Review MapReduce Concepts:

Revisit the concepts covered in the chapter on Parallel Processing with MapReduce.

Understand the Map and Reduce programs, key-pair data structure, program structure, and the execution of MapReduce programs despite node failures.

Understand MapReduce Applications:

Explore real-world applications of MapReduce, examining instances such as Google search algorithms and seek out additional use cases from different companies or industries.

Develop a MapReduce Algorithm:

Identify an analytical problem relevant to your chosen domain.

Formulate a MapReduce algorithm to solve the identified problem.

Clearly define the input, output, map function, and reduce function for your algorithm.

Apply MapReduce to Chosen Domain:

Discuss how the MapReduce algorithm you developed can be applied to the specific characteristics and requirements of your chosen domain.

Explore the potential benefits and challenges of implementing MapReduce in your domain.

Case Study Development:

Develop a hypothetical case study or scenario showcasing the practical application of MapReduce in your chosen domain.

Illustrate how the MapReduce algorithm solves the analytical problem identified earlier.

Challenges and Future Trends:

Identify potential challenges associated with implementing MapReduce in your domain.

Discuss future trends and advancements in MapReduce technology that could further benefit your chosen domain.

MAPREDUCE ALGORITHM IMPLEMENTATION:

MapReduce Code:

Provide the actual MapReduce code for the algorithm you developed.

Use a programming language of your choice (e.g., Java).

Explain Code Logic:

Provide a detailed explanation of the logic behind your Map and Reduce functions.

Clarify how the code addresses the identified analytical problem.

FORMAT AND SUBMISSION:

The assignment should be well-organized and clearly written.

Utilize appropriate headings and subheadings to structure your assignment.

Include diagrams, charts, or graphs to illustrate key points.

Submit the MapReduce code as a separate document or within the assignment document.

GRADING CRITERIA:

Develops a clear hypothetical case study showcasing MapReduce application.

Connects the case study with real-world applications in the chosen domain.

Offers thoughtful insights into the interdisciplinary nature of Big Data analysis.

Discusses how MapReduce bridges gaps between data analysis and decision-making.

Clearly and fully presents the developed MapReduce algorithm.

Proficiently implements Map and Reduce functions.

Successfully applies the algorithm to solve the identified analytical problem.