Large scale data analytics

Deskripsi Lengkap: https://lib.ui.ac.id/detail?id=20509153&lokasi=lokal

----

Abstrak

This book presents a language integrated query framework for big data. The continuous, rapid growth of data information to volumes of up to terabytes (1,024 gigabytes) or petabytes (1,048,576 gigabytes) means that the need for a system to manage and query information from large scale data sources is becoming more urgent. Currently available frameworks and methodologies are limited in terms of efficiency and querying compatibility between data sources due to the differences in information storage structures. For this research, the authors designed and programmed a framework based on the fundamentals of language integrated query to query existing data sources without the process of data restructuring. A web portal for the framework was also built to enable users to query protein data from the Protein Data Bank (PDB) and implement it on Microsoft Azure, a cloud computing environment known for its reliability, vast computing resources and cost-effectiveness.