Chapter 14
Resource Provisioning and Scheduling of Big Data Processing Jobs

Rajni Aron
Sahyadri College of Engineering and Management, India

Deepak Kumar Aggarwal
Concordia University, Canada

ABSTRACT

Cloud Computing has become a buzzword in the IT industry. Cloud Computing which provides inexpensive computing resources on the pay-as-you-go basis is promptly gaining momentum as a substitute for traditional Information Technology (IT) based organizations. Therefore, the increased utilization of Clouds makes an execution of Big Data processing jobs a vital research area. As more and more users have started to store/process their real-time data in Cloud environments, Resource Provisioning and Scheduling of Big Data processing jobs becomes a key element of consideration for efficient execution of Big Data applications. This chapter discusses the fundamental concepts supporting Cloud Computing & Big Data terms and the relationship between them. This chapter will help researchers find the important characteristics of Cloud Resource Management Systems to handle Big Data processing jobs and will also help to select the most suitable technique for processing Big Data jobs in Cloud Computing environment.

INTRODUCTION

The year 2007 witnessed the advent of a new term Cloud Computing which has become a buzzword now in the IT industry. The market now has plenty of Cloud technologies and platforms like Google App Engine, Microsoft Azure, Manjrasoft Aneka and many more to fit in this slot. Cloud Computing which provides inexpensive computing resources on the pay-as-you-go basis is promptly gaining momentum as a substitute for traditional information technology (IT) based organizations. Therefore, the increased utilization of Clouds makes an execution of Big Data processing jobs a vital research area.

As more and more users have started to store/process their real time data in Cloud environments, resource provisioning and scheduling of Big Data processing jobs becomes a key element of consideration for efficient execution of Big Data applications. The base of any real time system is a resource and to manage the resources to handle Big Data jobs in Cloud Computing environment is a very tedious task. An inefficient resource management system can have a direct negative effect on performance cost and indirect effect on functionality of the system. Indeed, some functions provided by the system may become too expensive or may be avoided due to poor performance. Thus, Cloud Computing faces the challenge of resource management especially with respect to choosing resource provisioning strategies and suitable algorithms for particular application. The major components of resource management systems are: resource provisioning, scheduling and load balancing. If any system is able to fulfil the requirements of these three components, the processing of Big Data jobs will become much easier.

This chapter discusses the fundamental concepts supporting Cloud Computing & Big Data terms and the relationship between them. It examines both computing paradigms under varied contexts such as the underlying characteristics, requirements, types, challenges and solutions. Comparison of different cloud service providers is done for various application paradigms. It reflects the essential perceptions behind the resource provisioning in the Cloud and identifies requirements based on user's applications associated with handling real time data. An architecture for dynamic provisioning of resources based on user's requirements to maximize efficiency and analysis of Big Data processing jobs is also discussed. This chapter will help researchers find the important characteristics of cloud resource management systems to handle Big Data processing jobs and will also help to select the most suitable technique for processing Big Data jobs in Cloud computing environment along with significant future research directions.

The remainder of this book chapter is organized as follows: In the section "Relationship between Cloud Computing and Big Data", a relationship between cloud computing and Big Data is presented. The section "Resource Provisioning for Big Data" discusses resource provisioning with Big Data in detail including types of resource provisioning techniques, requirements and related work in this area. Scheduling of Big Data based application is reviewed in the section "Scheduling for Big Data" with types of scheduling, characteristics of scheduling and scheduling in IaaS. The section "Load Balancing for Big Data" presents load balancing techniques for Big Data based applications. The section "Open issues and Challenges" presents future research directions related with resource provisioning, scheduling and load balancing for Big Data related problems in cloud computing environment. The section "Conclusion" discusses the final conclusion of the chapter.

Relationship Between Cloud Computing and Big Data

Cloud Computing is an elaborate and complex distributed computing paradigm providing an easy to use, on-demand network access to a shared pool of configurable computing resources involving networks, servers, storage, applications and services that can be dynamically provisioned and released with least management effort as well as service providers interaction. This cloud model can be visualized as an amalgamation of five essential characteristics and three service models. Cloud Computing services offerings can broadly be put into three major categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). Cloud Computing is primarily concerned about sharing of the resources. Any application in this paradigm comes with three requirements: computation, storage and network. Resources here are virtualized giving us an illusion of infinite capacity as well as hiding the implementation of their multiplexing and sharing.

Related Content

The Paradoxical Effect of Social Media Usage on Developer Creativity and the Moderating Role of Openness to Experience
Yumeng Miao, Rong Duand Veda C. Storey (2021). *Journal of Database Management (pp. 29-45).*
www.irma-international.org/article/the-paradoxical-effect-of-social-media-usage-on-developer-creativity-and-the-moderating-role-of-openness-to-experience/282443

Cost Modeling and Range Estimation for Top-k Retrieval in Relational Databases
Anteneh Ayanso, Paulo B. Goesand Kumar Mehta (2011). *Theoretical and Practical Advances in Information Systems Development: Emerging Trends and Approaches (pp. 295-315).*
www.irma-international.org/chapter/cost-modeling-range-estimation-top/52960

Evaluating XML-Extended OLAP Queries Based on Physical Algebra
Xuepeng Yinand Torben Bach Pedersen (2006). *Journal of Database Management (pp. 85-116).*
www.irma-international.org/article/evaluating-xml-extended-olap-queries/3354

Spatial Joins: Algorithms, Cost Models and Optimization Techniques
Nikos Mamoulis, Yannis Theodoridisand Dimitris Papadias (2005). *Spatial Databases: Technologies, Techniques and Trends (pp. 155-185).*
www.irma-international.org/chapter/spatial-joins-algorithms-cost-models/29663

Symbolic Objects and Symbolic Data Analysis
Héctor Oscar Nigroand Sandra Elizabeth González Císaro (2005). *Encyclopedia of Database Technologies and Applications (pp. 665-670).*
www.irma-international.org/chapter/symbolic-objects-symbolic-data-analysis/11221