Chapter 77

Streaming Remote Sensing Data Processing for the Future Smart Cities:
State of the Art and Future Challenges

Xihuang Sun
Chinese Academy of Science, China

Yan Ma
Chinese Academy of Science, China

Peng Liu
Chinese Academy of Science, China

Dingsheng Liu
Chinese Academy of Science, China

Yechao Sun
China Centre for Resources Satellite Data and Application, China

ABSTRACT

The explosion of data and the increase in processing complexity, together with the increasing needs of real-time processing and concurrent data access, make remote sensing data streaming processing a wide research area to study. This paper introduces current situation of remote sensing data processing and how timely remote sensing data processing can help build future smart cities. Current research on remote sensing data streaming is also introduced where the three typical and open-source stream processing frameworks are introduced. This paper also discusses some design concerns for remote sensing data streaming processing systems, such as data model and transmission, system model, programming interfaces, storage management, availability, etc. Finally, this research specifically addresses some of the challenges of remote sensing data streaming processing, such as scalability, fault tolerance, consistency, load balancing and throughput.

1. INTRODUCTION

With the fast development of satellite technology, satellite reception capability and the decrease in the costs of computer memory, remote sensing (RS) data processing has grown rapidly. The RS data transmitted in Institute of Remote Sensing and Digital Earth (RADI) reached 412.97GB/day in 2012 (RADI, 2012) and continues to grow rapidly. Furthermore, the RS data in EOSDIS grew to 8.5TBper day in 2012 and its total archived dataset reached 10PB (EOSDIS, 2012).

In addition to the data explosion, the needs of real-time processing and concurrent access of RS data processing has also increased as more time-strict applications are applied.RS data processing plays an important role in disaster monitoring, climate prediction, and many other city-related applications. Real-time processing of RS data can encourage the development of smart disaster management, pre-disaster risk assessment, and post-disaster risk assessment big data applications with more effective solutions and fewer energy costs. Increasing the speed of RS data processing of massive data and real-time analysis will benefit a lot in building future smart cities.

RS data are the earth-observing data continuously obtained from space borne and airborne sensors and other data acquisition measurements. As RS data is mainly transmitted as data streams from sensors but stored as data files for offline processing, research based on data files cannot achieve thorough real-time processing since they cannot fill the gap between data transmission and data processing. In fact, problems such as mixed programming across different hardware, IO bottlenecks caused by external disk storage, and coordination between distributed datacenters still cannot be fixed perfectly. Streaming RS data processing treats RS data as continuous data streams and completes data transmission and computing in memory. It may reduce the processing units by eliminating the disk storage of intermediate data and build a pipeline where sub part of processing can make interactions overlap in time. RS data streaming processing brings no extra burden to processing hardware and can break the parallel IO bottleneck with memory.

In this paper we introduce current research in RS data processing. Then we summarize mainstream data streaming processing frameworks. We also discuss aspects of building RS data streaming processing frameworks. Finally we discuss the challenges of future RS data streaming processing, such as scalability, fault tolerance, consistency, load balancing and throughput.

2. RELATED WORKS

2.1. Current Works in RS Data Processing

To boost the speed of RS data processing, many research have been studied based on different architectures. These research mainly focus on offline RS data processing. Algorithms and processing patterns (Hailiang & LU, 2010; Jun-jun & Gang, 2010; Ke-fei, 2010; LIU et al., 2014) take the advantage of different hardware architectures such as GPU and FPGA. High performance clusters (Wei, Liu, & Wang, 2014; Xue-ping, 2013) enhance the RS data processing performance by using parallel computing. Distributed computing environments (Sekiguchi et al., 2008) can increase the data-storage capacity and expand the diversity of processing services.

There are also works based on cloud computing and RS big data computing. G-Hadoop (Wang et al., 2013) applied MapReduce across distributed data centers. Almeer(2012) employed Hadoop MapReduce framework to implement parallel processing of RS images and built an experimental 112-core high-performance cloud computing system. Wang (2014) estimates the statistical characteristics of RS big data in the wavelet transform domain. Task scheduling strategies (W. Zhang, Wang, Ma, & Liu, 2014) across data centers were also studied. Wang (2015) discusses some research issues about processing distributed internet of things data in clouds.

Apart from these works, Research on RS data streaming processing mainly focuses on data model and data querying. StreamInsight (Kazemitabar, Demiryurek, Ali, Akdogan, & Shahabi, 2010) support

Related Content

Soil Pollution by Chromium in Sub-Saharan Africa
 (2023). *Global Industrial Impacts of Heavy Metal Pollution in Sub-Saharan Africa (pp. 178-195).*
www.irma-international.org/chapter/soil-pollution-by-chromium-in-sub-saharan-africa/328148

SiRCub, A Novel Approach to Recognize Agricultural Crops Using Supervised Classification
Jordi Creus Tomàs, Fabio Augusto Faria, Júlio César Dalla Mora Esquerdo, Alexandre Camargo
Coutinhoand Claudia Bauzer Medeiros (2019). *Environmental Information Systems: Concepts,
Methodologies, Tools, and Applications (pp. 1129-1147).*
www.irma-international.org/chapter/sircub-a-novel-approach-to-recognize-agricultural-crops-using-supervised-classification/212986

Mathematical Models Used for Hydrological Floodplain Modeling
Carmen Mafteiand Konstantinos Papatheodorou (2015). *Extreme Weather and Impacts of Climate Change
on Water Resources in the Dobrogea Region (pp. 240-283).*
www.irma-international.org/chapter/mathematical-models-used-for-hydrological-floodplain-modeling/131532

Geospatial Influence in Science Mapping
Carlos Granell-Canutand Estefanía Aguilar-Moreno (2019). *Advanced Methodologies and Technologies in
Engineering and Environmental Science (pp. 349-361).*
www.irma-international.org/chapter/geospatial-influence-in-science-mapping/211883

General Perspectives on Electromyography Signal Features and Classifiers Used for Control of
Human Arm Prosthetics
Faruk Ortes, Derya Karabulutand Yunus Ziya Arslan (2019). *Advanced Methodologies and Technologies in
Engineering and Environmental Science (pp. 1-17).*
www.irma-international.org/chapter/general-perspectives-on-electromyography-signal-features-and-classifiers-used-for-control-of-human-arm-prosthetics/211855