Abstract

Geographic Information Systems (GIS) platforms are used to implement and deal with the massive spatial data, especially image data. Therefore, these platforms require high storage capacity and high computational power to process them. This paper aims of processing of Geodata using Distributed processing Frameworks. Large volume of Geodata cannot be processed using desktop GIS tools such as QGIS, ArcGIS, GRASS, OpenJUMP etc. Therefore, to handle and process on these types of large data, use of Hadoop Distributed processing framework needs to be deployed. GeoProcessing is a GIS operation used to manipulate spatial data. It is one of the original proposal in GIS development. Almost every GIS application is represented by a GeoProcessing Workflow. This paper explains the GeoProcessing Workflow for processing of image data. Also explains Hadoop Distributed File System (HDFS), MapReduce Programming Model and Yet Another Resource Negotiator (YARN) architecture, useful in large spatial data handling and analysis at fast rate.


Central Africa Regional Program for the Environment: http://carpe.umd.edu/geospatial/satellite_imagery_resources.php


The Apache Software Foundation: http://hadoop.apache.org/


Hadoop YARN: http://hortonworks.com/hadoop/yarn/

Hadoop at Yahoo!: https://developer.yahoo.com/hadoop/


ISO Standards: www.iso.org/iso/catalogue_detail.htm

OGC Standards: www.opengeospatial.org/standards

Index Terms

Computer Science

Distributed Systems
Keywords

GIS  Geoprocessing Workflow  Distributed System  Hadoop  YARN  MapReduce  Classic.