

Construction of spatio-temporal data mining system for time-series satellite imagery using Hadoop

NISHIMAE, Kou¹ ; MIYOSHI, Tomoya¹ ; HONDA, Rie^{1*}

¹Kochi University

A large number of spatio-temporal data have been stored in various fields of science, such as remote sensing, numerical simulation, and astronomical observation, in which data often appears as time-series images. To extract spatio-temporal knowledge from spatio-temporal data including time-series images, spatio-temporal cross section relevant to a target task has to be extracted from a mass of data. Since these data are stored as a large number of files, utilization of distributed processing framework such as Hadoop or Gfarm is promising.

We constructed distributed data mining system for time-series satellite images using 53 nodes (3 masters and 50 slaves at maximum) of iMac and Hadoop which enables distributed file system and distributed processing using MapReduce. We evaluated the scalability and performance of the system for the task extracting time-series data from a large number of images carefully and found that partitioning the images into optimum numbers and reducing the data between map phase and reduce phase is essential.

The system was then applied to two different tasks focusing on time-series data analysis extracted from satellite imagery: statistical modeling of seasonal changes in vegetation index and spatio-temporal correlation analysis of weather satellite images. The tasks were successfully implemented on the system and the computational time was decreased in inverse proportion to the number of slave nodes, thus usefulness of the distributed system to spatio-temporal data mining for time-series images.

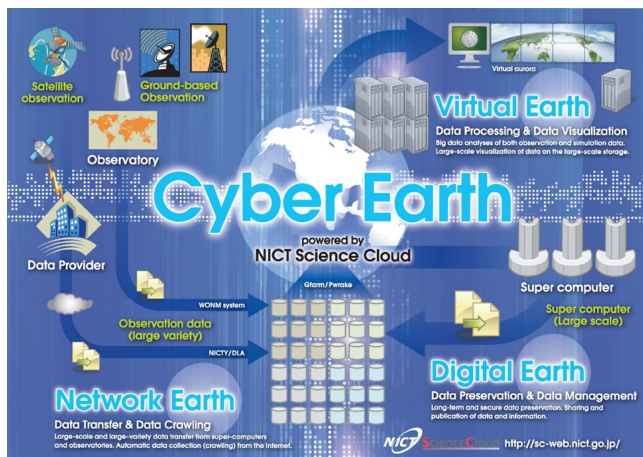
Keywords: distributed processing, Hadoop, MapReduce, data mining, spatio-temporal, satellite imagery

Basic Technologies, Integrated Systems and Applications of the NICT Science Cloud

MURATA, Ken T.^{1*}; WATANABE, Hidenobu¹; UKAWA, Kentaro²; MURANAGA, Kazuya²; YUTAKA, Suzuki²; KASAI, Yasuko¹; SATOH, Shinsuke¹; ISHII, Shoken¹; YAMAMOTO, Kazunori¹; NAGATSUMA, Tsutomu¹; TSUGAWA, Takuya¹; NISHIOKA, Michi¹; GUO, Zhihong³; KUROSAWA, Takashi⁴; MIZUHARA, Takamichi⁵

¹National Institute of Information and Communications Technology, ²Systems Engineering Consultants Co., LTD., ³SURIGIKEN Co.,Ltd., ⁴Hitachi Solutions East Japan, Ltd., ⁵CLEALINKTECHNOLOGY Co.,Ltd.

This paper is to propose a cloud system for science, which has been developed at NICT (National Institute of Information and Communications Technology), Japan. The NICT science cloud is an open cloud system for scientists who are going to carry out their informatics studies for their own science. The NICT science cloud is not for simple uses. Many functions are expected to the science cloud; such as data standardization, data collection and crawling, large and distributed data storage system, security and reliability, database and meta-database, data stewardship, long-term data preservation, data rescue and preservation, data mining, parallel processing, data publication and provision, semantic web, 3D and 4D visualization, out-reach and in-reach, and capacity buildings.



Global spectral crustal model

TENZER, Robert^{1*} ; CHEN, Wenjin¹

¹School of Geodesy and Geomatics, Wuhan University

We compile the harmonic coefficients, which describe the Earth crustal density structure with a spectral resolution complete to degree/order 180. These coefficients can be used in gravimetric studies of the Earth lithosphere structure, isostasy, crustal loading, sedimentary basins and related topics. The global spectral crustal model is separated into 9 specific layers of the topography, bathymetry, polar ice sheets, sediments (3-layers) and consolidated crust (3-layers). The harmonic coefficients describe uniformly the geometry and density (or density contrast) distribution within each crustal component. The topographic and bathymetric coefficients are generated from the topographic/bathymetric model ETOPO1 and the global geoid model GOCO03s. A uniform density model is adopted for the topography. The ocean density distribution is approximated by the depth-dependent seawater density model. The ETOPO1 topographic and the DTM2006.0 ice thickness data are used to generate the ice coefficients, while assuming a uniform density of the glacial ice. The geometry and density distribution within sediments is described by the 3 stratigraphic layers of a laterally varying density model, and the same structure is used to describe the density distribution within the consolidated crust down to the Moho interface. The sediment and consolidated crust coefficients are generated from the global crustal model CRUST1.0. The density contrasts of the ocean, ice, sediments and remaining crustal structures are taken relative to the reference crustal density.

Keywords: crust, density, gravimetric forward modeling, harmonic analysis