Japan Geoscience Union Meeting 2014
(28 April - 02 May 2014 at Pacifico YOKOHAMA, Kanagawa, Japan)

MGI37-05 Room:413 Time:April 29 10:00-10:15

High performance data processing for detection of bipolar waveforms from KAGUYA/WFC-L using the NICT Science Cloud

DAISUKE, Yagi[1*] ; MURATA, Ken T.[2] ; KASAHARA, Yoshiya[1] ; GOTO, Yoshitaka[1]

[1]Kanazawa Univ, [2]National Institute of Information and Communications Technology

Lunar orbiter named KAGUYA was launched in September, 2007, and was operated until June, 2009. The WFC (waveform capture) onboard KAGUYA measures plasma waves below 1MHz around the moon. The WFC-L, one of subsystems of WFC, is a waveform receiver measuring waveform from 100 Hz to 100 kHz with its sampling frequency at 250 kHz. Characteristic bipolar waveforms which can be classified into some patterns were observed by the WFC-L. We developed an automatic detection algorithm to pick up these bipolar waveforms , but it takes huge computation time because the total amount of the WFC-L data is about 190 GB.

In the present study, we introduced the Science Cloud system served by National Institute of Information and Communications (NICT) in order to improve the performance of trial and error process for the development of detection algorithm. The NICT Science Cloud is a cloud system built for scientific research and data service especially for big data science. We utilized parallel data processing under the work flow control implemented in the NICT Science Cloud. We report the performance of the NICT Science Cloud in the present paper.

In order to define an appropriate workflow to the data processing servers, Pwrake (Parallel Workflow extension for Rake) was introduced as a task scheduler. Pwrake is extended for file sharing systems from Rake which is a build tool described by Ruby language. It is possible to assign tasks to each node and to perform parallel data processing by describing the contents of processing, the node to be used, and the number of cores.

We confirmed that total processing time reduced down to 1/140 times compared with a case of 1 node and 1 core, when we used 10 nodes and 24 cores. Because of the effect of hyper-thread technology, processing speed is not proportional to the number of the resources. By utilizing the system, it is expected that a higher-precision detection algorithm can be developed efficiently. As further works, development of more intelligent detection algorithm as well as evaluation of the performance using much larger resources of the NICT Science Cloud will be necessary.

Keywords: Lunar Orbiter KAGUYA, Waveform Capture, NICT Science Cloud, parallel processing