

Carbon Dioxide Flux Data Computing and Data Warehouses Using Grid Techniques

Jims Hsu

ASGC, Taiwan

The purpose of the present study is to develop a framework which can be used by researchers who are dealing with carbon dioxide flux data management and computing.

Carbon cycle is recognized to have strong relationship to climate change and understanding the net carbon flux is an essential component of climate change research. Researchers around the world measured the carbon exchanges between biosphere and atmosphere utilizing the carbon dioxide flux observation tower that is one of important research method. At present, more and more countries, scientists join this global collaboration. In the FLUXNET, there are over 500 tower sites operating on a long-term and continuous basis in the world. Scientists face a challenge of dealing with a huge amount of data.

We apply a data warehouse and data analysis framework assisting carbon flux research group on the basis of EGEE infrastructure. It is helpful for this research to provide better environment for huge sensor data management and computation model with flexible user control, a workflow engine was integrated with gLite by the GAP.

We use flux data collected from Taiwan Chilan flux towers to test the framework. The result shows that it can be used to solve data management and computing issues such as complicate process of computing, the lack of software, time consuming on data handling and cleaning and real time data management. In addition, we choose the approach to use metadata as data content description and separate metadata from raw data storage has found a benefit to solve the archiving difficulty of the large volume of flux raw data and can be shared through the Grid platform.