BNL-95005-2011-AB

BIG DATA SYSTEMS AT THE ARM ARCHIVE I: INTERACTIVE VISUALIZATION OPPORTUNITIES

Giri Palanisamy, Oak Ridge National Laboratory
Raymond McCord, Oak Ridge National Laboratory
James Mather, Pacific Northwest National Laboratory
Pavlos Kollias, McGill University
Ieng Jo, McGill University
Pete Eby, Oak Ridge National Laboratory
Karen Gibson, Oak Ridge National Laboratory
Eric Stephan, Pacific Northwest National Laboratory
Michael Jensen, Brookhaven National Laboratory

For presentation at
The Second Science Team Meeting of the
Atmospheric System Research (ASR) Program,
San Antonio, TX
March 28-April 1, 2011

Environmental Sciences Department/Atmospheric Sciences Division
Brookhaven National Laboratory

U.S. Department of Energy
Office of Science

ABSTRACT

A new ARM infrastructure resource provided by the ARM Archive includes a new system that will be dedicated to visualization and software development based on the very large data volumes from new ARM instrumentation (e.g., scanning radars and lidars). As part of the "Big Data System" (BDS) design, the ARM Archive is currently setting up a system with various interactive visualization tools. The registered archive users will be able to request access to this system and perform various visualization and small-scale data extraction tasks. This system will have radar software such as TITAN, IRIS, and IDV, and data processing and visualization tools such as IDL, MATLAB, and NetCDF libraries. In addition to the interactive visualization capabilities, this system will also be used to produce a very large number of pre-computed plots for large volumes of radar data. These plots will be used in various ARM web pages and Archive user interfaces as part of the data discovery and data ordering processes. This system will also have programming languages and their libraries such as C, C++, Perl, Java, Python, and ScientificPython libraries. The ARM Integrated Software Development Environment (ISDE) will be installed as part of the script development feature. The users will be able to use a large collection (10s of TB) of locally stored data or request data through Archive user interfaces for transfer to this system for further processing. This poster will also explain hardware and use policies and operational scheme for this system.