Reprinted from the preprint volume of the 15th International Conference on IIPS for Meteorology, Oceanography, and Hydrology, 10–15 January 1999, Dallas TX, by the AMS Boston MA

BNL-66881

P1.4

METADATA SYSTEMS FOR ARM'S SGP AND NSA SITES

Kathy J. Doty *
Brookhaven National Laboratory,
Upton, New York

1. INTRODUCTION

The Atmospheric Radiation Measurement (ARM) Program (Stokes and Schwartz, 1994), funded by the Environmental Sciences Division of the U.S. Department of Energy (DOE), is a major global change research program whose focus is to better understand the influences of human activity on the earth's climate.

The ARM program has set up instruments at three primary measuring sites around the world. These sites are the Southern Great Plains (SGP) of the United States, the Tropical Western Pacific (TWP) and the North Slope of Alaska (NSA). The sites were selected to provide opportunities to observe a wide range of climatologically important meteorological conditions. Scientists will be able to use the parametric data gathered at these sites to study the effects and interactions of sunlight, energy transfer, and clouds on climate.

Metadata is important to the ARM program because it provides information about the scientific utility of the parametric data gathered at the ARM sites and also provides information that will enhance the operations and assist in the maintenance of the sites and the site data systems.

Members of the ARM Data and Science Integration Team (DSIT) and the ARM infrastructure have worked together to design and implement a MetaData System (MDS) currently in use at the SGP site, and are now in the process of designing an MDS for the NSA site. We are also working on the design of an MDS for the TWP site which will contain a subset of the metadata being managed by the TWP Operations Management System (OMS). This paper primarily presents the details of the SGP MDS. The NSA and TWP MDS designs will be kept as similar to that of the SGP MDS as possible so that interaction and comparisons between all MDS' will be possible in the future.

2. THE MDS DATABASE

The MDS is based on a relational database management system (RDBMS) and is running on a SUN Sparc 20 workstation under the Solaris 2.4 operating system. We are currently using Empress RDBMS as the database software. As shown in Figure 1, included in the MDS database are: entries produced by instrument data processing circuits (IDPCs) that deal directly with an instrument or an instrument data stream status (collection, ingest, instrument, and quality control modules); entries made by the site operators as they perform duties related to site operations; entries made by anyone in the ARM community via an e-mail message which is forwarded to the MDS' entry processing software. Site operations specific information includes hourly weather observations, instrument preventative and corrective maintenance reports, reports on the surface conditions at the site's facilities, weather alert status reports, entries of a general nature entered by site operations personnel, and summaries of preventative maintenance procedures followed when technicians are performing preventative maintenance activities. In addition, the database contains information on site personnel, facilities, and instrumentation.

3. METADATA INPUT

Various methods exist for populating the MDS. These methods include the transfer of ASCII files to a "data doorstep" on the site data system (SDS) for loading into the MDS, an event-driven (e-mail) entry process and a web-based data entry interface.

3.1 *Transfer of ASCII Files*

Site operations personnel enter instrument corrective maintenance specific, instrument preventative maintenance specific and extended facility surface conditions reports into laptop PCs while in the field. These reports are entered into databases designed using FileMaker Pro database software. The entries are later downloaded to ASCII files, transferred to the SDS data doorstep and loaded into the MDS. The transfer and load procedures have been automated using ftp and the Unix "cron" utility.

3.2 *Event-Driven Entry Process*

The event driven entry process handles the entering of incoming e-mail into the MDS. It does this by using the forwarding capabilities of the *sendmail* program (Costales, Allman and Rickert, 1993) to deliver e-mail to a program run as a specific user on the data system. The program is triggered to run as incoming e-mail arrives. This is the method of entry used by the automatic IDPCs and members of the general ARM community including those on the SDS development team.

3.3 *Operator Entry Interface*

The operator entry interface is a web-based data entry tool used by the site operators to make entries into the MDS. The web entry interface is password protected and is restricted to site operations personnel use. The types of entries made via the web-based interface are

* *Corresponding author address:* Kathy J. Doty, Brookhaven National Laboratory, Department of Applied Science, Building 490D, Upton, New York 11973-5000; E-mail: doty@bnl.gov

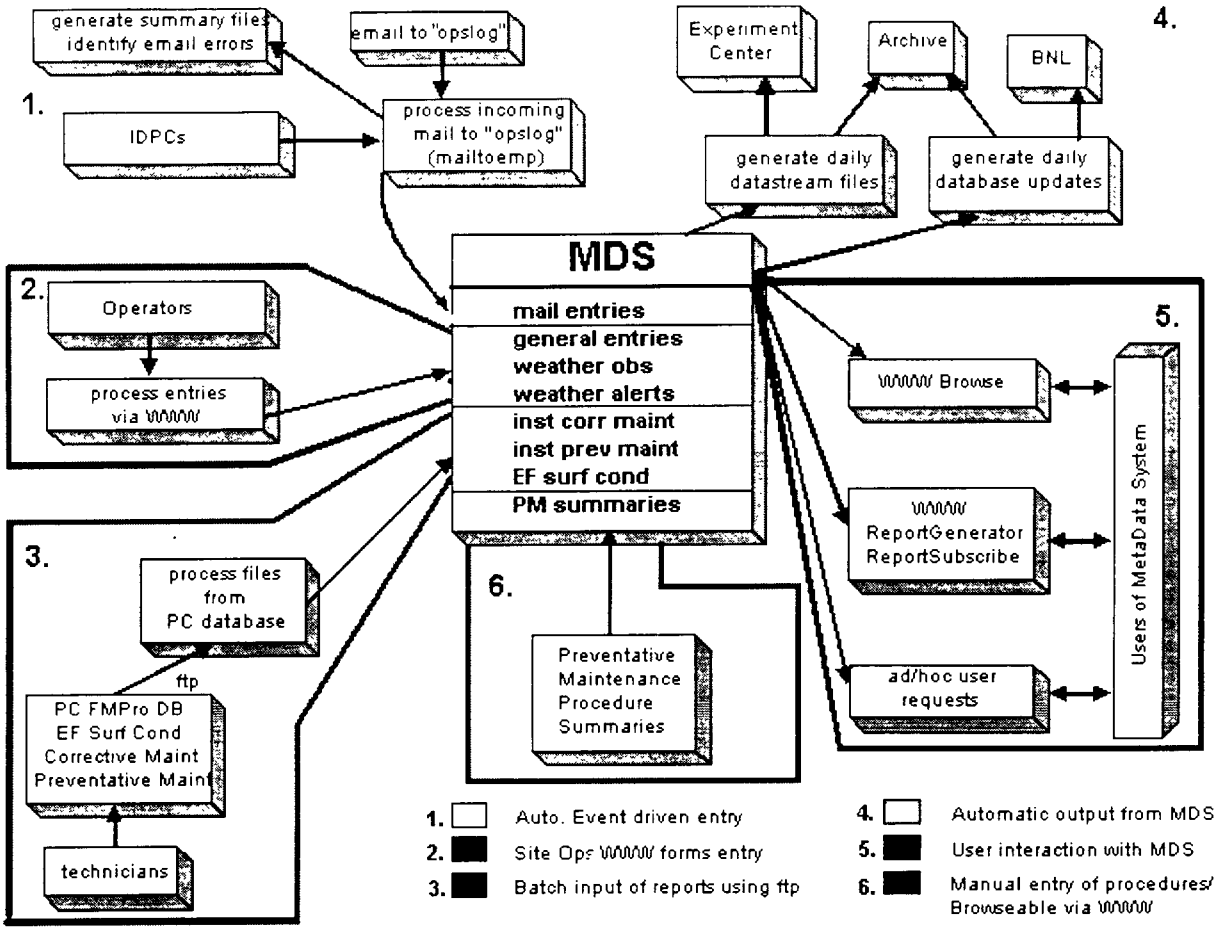

Figure 1. MDS Database Contents and Metadata Flow