16a

PNL-10197

Airborne Multisensor Pod System (AMPS) Data Access Concept

Jiing-Yih J. Jou

**W. D. Meitzler
Project Manager**

October 1994

**Prepared for the U.S. Department of Energy
Under Contract DE-AC06-76RLO 1830**

**Pacific Northwest Laboratory
Operated by the U.S. Department of Energy
by Battelle Memorial Institute**

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST LABORATORY
operated by
BATTELLE MEMORIAL INSTITUTE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC06-76RLO 1830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831;
prices available from (615) 576-8401. FTS 626-8401.

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161.

Airborne Multisensor Pod System

Data Access Concept

Jiing-Yih J. Jou

Wayne D. Meitzler,
Project Manager

November 1994

Prepared for
the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

Pacific Northwest Laboratory
Richland, Washington 99352

1.0 Introduction

The purpose of this document is to provide the conceptual design for accessing data collected under the Airborne Multisensor Pod System (AMPS) program and for accessing data about that data. The purpose of the AMPS program is to provide multisensor data for data fusion research as it applies to arms control treaty verification, non-proliferation surveillance, and related programs. Secondary objectives include sensor development and technology demonstration. The AMPS program was established by the U.S. Department of Energy (DOE), Office of Arms Control and Non-proliferation (DOE/AN) and is integrated into the overall DOE AN-10.1 technology development program. The AMPS program interacts with other technology programs of DOE/NN-20.^(a)

The AMPS program participants include five DOE laboratories. Four will be responsible for collecting and supplying data to AMPS data researchers:

- Lawrence Livermore National Laboratory (LLNL)
- Remote Sensing Laboratory (RSL)
- Sandia National Laboratories (SNL)
- Savannah River Technology Center (SRTC).

The Pacific Northwest Laboratory^(b) (PNL), has been assigned data management responsibility for the program, which is defined as the process of planning, acquiring, organizing, qualifying and disseminating data. Data ranges from single sensor readings, to high resolution images, to video. The potential volume of data to be collected under the AMPS program will require considerable management effort.

The AMPS program specifies a distributed approach for the sharing and management of AMPS data. This approach shares the burden of data distribution across the DOE laboratories, which will be responsible for AMPS instruments development and implementation and the data those instruments produce. This approach will also provide the laboratories with direct control over the distribution of the data for which they are responsible.

(a) For a description of the AMPS program, see "Airborne Multisensor Pod System (AMPS) Program Plan," DOE NN-20, U.S. Department of Energy, Washington, D.C.

(b) Pacific Northwest Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under Contract DE-AC06-76RLO 1830.

2.0 Goal

The goal of the AMPS data access development is to provide an easy-to-use, seamless environment for accessing AMPS data and information about that data. This means that the user community will not need to concern itself with the physical location of the data or the data catalog. The data access concept is based on one tool for accessing data being used by all researchers and data suppliers. This tool will be built using government-owned or available low-cost commercial products. AMPS data providers will use the tool to publish their data availability and other parameters concerning their data. AMPS data researchers will use the tool to inquire about AMPS data availability and to submit requests for particular datasets. Finally, the capability to access data will make use of existing government computing networks and software infrastructures.

2.1 Data Catalog

The data catalog for the AMPS data access system will provide functions for two basic user groups: data researchers and data suppliers.

2.1.1 Data Researcher Functions

The AMPS data catalog will provide data researchers with basic information on available AMPS data and characteristics of the data. It will be an essential tool for quickly determining what information is available. The data catalog will be accessed through the Internet and will allow data researchers to obtain the most current information at any time on available data. The functions of the data catalog will be available to the data researcher and include the following:

- **data browsing** - allows the AMPS data researchers to search on specific criteria through the catalog system to locate the data of interest.
- **data dictionary viewing** - allows the researchers to obtain information on the organization and structure of the data and its associated attributes.
- **data ordering** - allows users to obtain data directly over the network or indirectly by supported distribution media.
- **online help** - provides necessary information to help data researchers learn how to use the cataloging system.

2.1.2 Data Supplier Functions

Data supplier functions will include the capabilities to add and maintain data catalog. Specific supplier functions will include the following:

- adding a catalog entry
- deleting a catalog entry
- updating a catalog entry
- maintaining logs on who accesses catalog entries, and when accessing takes place
- synchronizing the cataloging system among data suppliers.

2.2 Data Request

The AMPS data access system will provide functions that will allow users to submit requests for AMPS data directly over a network. Through the cataloging system, a user will be able to identify the datasets of interest simply by marking them and specifying the media desired. The data access system will be able to process the request and distribute formal requests to each data supplier site for the data that is requested. The data request activities will be automatically entered into the status logs for administrative reference.

2.3 Data Transfer

Once a data request is made, the action for transferring the data to the requester will be turned over to the data supplier. Two conditions govern how the transfer is to be made:

1. If the requester specifies that the data is to be transferred over a network, the data will be sent directly from the supplier to the requester.
2. If the requester specifies that the data is to be transferred using a storage media, such as tape, the supplier will be notified through the data access system; the supplier will then prepare the data on the storage media and mail it to the requester.

The size of datasets will vary widely, from a few kilobytes to hundreds of megabytes. The data access system will provide the data requester with a practical data transfer policy based on the following:

- whether online transfer is the preferred method for small to medium size datasets
- whether CD-ROM transfer is the preferred method for large size datasets.

Although the AMPS data access system defines the data transfer policy, other transfer methods, such as 8-mm tape, will be allowed if the data supplier is capable of providing an alternative method. It will be the data supplier's responsibility to provide information to the cataloging system about what other distribution media are available so that data requesters know their options.

3.0 Options

It will be possible to distribute data in numerous ways. Table 3.1 lists some possibilities. Each possibility is addressed in the table in terms of the following three major elements in data distribution:

- **Data Catalog** - Data suppliers will list the data that are available in this catalog. Researchers and data suppliers will be able to use it to determine what data are available and their condition.
- **Data Request** - To obtain data listed in the data catalog, researchers and data suppliers will file a request with the supplier responsible for that data.
- **Data Transfer** - Based on data requests from researchers and data suppliers, data suppliers will provide requesters with the data they request.

The shaded areas of Table 3.1 identify the approach recommended for the AMPS program for each data distribution element. The reasons for this recommendation are as follows. Paper-based options were eliminated because they are labor intensive. The use of CD-ROM for the data catalog was eliminated because a new release of the catalog would be needed for each update to keep users current.

In general, the use of networks for the catalog, data request, and data transfer of small to medium datasets are recommended. For a large dataset, the use of CD-ROM is suggested, although transfer over the Internet or use of tape media is not ruled out.

Table 3.1. Data Access Options

Distribution Technique	Data Catalog	Data Request	Data Transfer
CD/Tape	update periodically	not practical	any size
FTP/WAIS/WWW	direct access	direct access	small & medium size
paper	update periodically	postal service	view-only
fax	possible	phone line	small size
E-mail	mail message	mail message	small & medium size
Data in the range of kilobytes is considered small. Data in the range of a few megabytes is considered medium. Data in the range of hundreds of megabytes or more is considered large. FTP is the acronym of File Transfer Protocol over a network. WAIS is the acronym of Wide Area Information System, an Internet protocol. WWW is the acronym of World Wide Web, a hypertext-based Internet protocol.			

4.0 Data Access Approach

The primary method to access AMPS data will be through Internet, using the Mosaic user interface. The network will offer all data access functions including data catalog, data request, and data transfer. The only exception will be data transfer of large datasets, which may use CD-ROM or other mass storage media supported by the data suppliers. In the rare case when the network will not be available or user access will be prohibited, a backup distribution means will be provided.

4.1 AMPS Network Services

A high-level schematic of the network associated with AMPS is shown in Figure 4.1. Most data researchers and data suppliers will access the network directly through the network infrastructure already in place at their facilities. Data researchers without network connections will either arrange to have direct network access or use a dial-up modem to access a facility that does support the network.

The data suppliers at the national laboratories will have accessible nodes on the network. Each node consists of computer equipment in support of AMPS that will provide access to the AMPS data catalog, online data that they provide, and the AMPS data access software. The node will be operated and maintained by the data suppliers's organization with the understanding that it will be kept online at all times. The computer equipment can be shared among data suppliers as long as timely response can be provided for data access by the AMPS user community.

4.2 AMPS Non-network Services

Non-network AMPS services will focus on two areas:

- large dataset transfer
- backup support should the network be unavailable or prohibited.

The AMPS data access concept suggests that the transfer of large datasets should be by CD-ROM. The data transfer activity will be conducted directly by the data supplier in response to a network data request. However, the AMPS data access system may not rule out the possibility of transferring a large dataset over the Internet should the data requester insist.

Backup support will be provided for AMPS users who cannot establish an Internet connection or are prohibited from access. In this case, the AMPS catalog will be made available to those users in both electronic and paper forms generated by the AMPS catalog system. Ordering AMPS datasets will be accomplished by traditional means, such as by fax or mail from the requester directly to the AMPS data supplier. Because of the dynamic nature of the AMPS cataloging system, it is the requester's responsibility to make sure that the AMPS catalog they have in hand is current.

Figure 4.1. AMPS Network Connection Scheme

5.0 System Support

The AMPS data access system requires adequate **networking**, hardware, security measures, and skillful personnel support. These needs are described in the following sections.

5.1 Networking

Because the concept presented in this document is based on networking, an affordable network for the AMPS user community will be vital for success. Internet, the world's largest computer network, represents a viable solution. Internet comprises many communication resources, integrated to form a very large, virtual network available to all major research institutions, government agencies, and universities. Access to the Internet costs as little as the cost of a modem (less than \$200). Users of all major computer operating systems can access Internet resources online.

5.1.1 Public Domain Protocols

The AMPS data access concept is expected to make use of the public domain protocols Wide Area Information Server (WAIS) and/or World Wide Web (WWW). The WWW is the latest Internet technology available. It uses hypertext technology to link related topics. The WAIS is a distributed text searching system based on the drafted ANSI Z39.50 standard. It allows computer users to find and access resources on the network without regard to their location. The AMPS data access specification and data access design will provide additional details on the network implementation specifications.

5.1.2 Network Data Transfers

AMPS users will be able to acquire small to medium datasets directly over the network. The AMPS data access system may impose limitations on the size of data that can be sent over the network.

5.2 Computer Media

All AMPS data that can be converted to digital form will be available on computer media. Computer media will be the primary means of transferring large AMPS datasets. The AMPS data access concept dictates the CD-ROM as the preferred data transfer media for large datasets. However, AMPS data may be supplied on non-CD-ROM media if the data supplier supports it. The following are the computer media to be used:

- **CD-ROM** - CD-ROM is the preferred media for the distribution of AMPS data when distribution by means of the network is not possible or too costly. CD-ROM is gaining wide acceptance among commercial data and software suppliers. In relation to the magnetic media commonly used, CD-ROM has a much longer shelf-life, costs less, and provides faster data access using random access methods.

- **Tape** - Magnetic tape media (8-mm or 9 track tape) is the approach historically used for distributing large amounts of data. Some AMPS instrument suppliers already use this medium. Distribution of AMPS data in this form is recommended for special situations where the data supplier has the capacity to produce the tape.
- **Other Media** - Distribution of AMPS data on other media will be confined to those media that cannot be digitized to a level that replicates the data. Cases warranting further investigation for this purpose include high-resolution camera film and video tape. A final decision on which will be used will be made during the data access specification phase.

5.3 Security

The AMPS requirement for data distribution is to provide AMPS data to DOE data researchers. For data that is not classified or sensitive, the security offered through Internet and computer passwords will be adequate for basic AMPS security requirements. For classified and sensitive data, a different distribution method will be needed. This document does not attempt to address this issue.

5.4 Data Access Support Staff

AMPS data access will require staff for operation and support. An AMPS data access manager will be necessary for the following functions:

- monitoring data access requests
- maintaining the AMPS data access system.

Data suppliers will require staff resources to accomplish the following:

- support computer equipment on the network nodes
- enter instrument data into the AMPS data access system
- prepare and distribute CD-ROMs or other media.

5.5 Data Access Equipment

For data researchers, the following equipment will be needed for AMPS data access:

- a DOS or UNIX-based operating environment
- a computer capable of directly accessing the Internet

- a CD-ROM or tape drive to read AMPS distributed data
- an adequate on-line data storage medium.

For data suppliers, the following equipment will be needed:

- a computer as a node on the Internet network
- on-line mass storage for their data repository
- a writable CD-ROM or 8-mm drive to produce data for distribution.

6.0 Conclusion

This concept paper recommends that AMPS data access be achieved using a computer network approach for all services, except for large dataset transfers that would be performed by mail using CD-ROM or 8-mm tape. Implementation of this approach would accomplish the following:

- Make use of existing protocols available on the Internet.
- Provide a cataloging system for the data suppliers and the researchers to perform access functions.
- Have the data suppliers distribute their data over the network (small and medium datasets) or CD-ROM or 8-mm tape (large datasets).
- Provide a backup mechanism in the case that the researchers cannot access Internet or are prohibited from using it.

Although the current program direction is to use a distributed approach to handle AMPS data, the AMPS Integration Team has concluded that a centralized system is a better approach from both technical and cost-effectiveness points of view. The concepts and architecture described in this document could also be used in a centralized approach.

Distribution

**No. of
Copies**

OFFSITE

- 12 DOE/Office of Scientific and
Technical Information

Michael F. O'Connell
Office of Arms Control and
Nonproliferation
U.S. Department of Energy
1000 Independence Avenue
Washington, DC 20585

Ralph Hastings
NN-20, GA-033
Office of Research and Development
U.S. Department of Energy
1000 Independence Avenue
Washington, DC 20585

**No. of
Copies**

ONSITE

- 22 Pacific Northwest Laboratory

K. K. Anderson	K7-34
M. J. Burford	K7-28
J. C. Cordero	K7-28
R. A. Gimenez	K7-28
J. A. Hucks	K7-28
R. W. Jordan	K7-28
J. J. Jou	K7-22
T. F. Lundeen	K6-84
W. D. Meitzler	K7-22
R. B. Melton	K7-02
J. W. Nageley	K7-10
B. A. Roberts	K6-48
T. L. Tran	K7-28
H. A. Undem	K6-48
M. S. Vorce	K1-96
R. A. Warner	K6-48
Publishing Coordination	
Technical Report Files (5)	