

The FDSN Archive at the IRIS Data Management Center

Tim Ahern (*)

Overview

In August of 1990, the IRIS Data Management Center (IRIS DMC) was designated as the first FDSN Data Center. Since that time the IRIS DMC has also come to be known as the FDSN Archive. As the FDSN Archive, all member networks, through their own local Network Data Centers, routinely send data to the IRIS DMC. At the present time we have received data from seven member networks. These include: CNSN (Canada); CDSN (China); GEOSCOPE (France); Graefenberg (Germany); IRIS GSN (U.S.A.); MEDNET (Italy); POSEIDON (Japan).

In addition to acting as the FDSN Archive, the IRIS DMC is responsible for archiving and distributing data from several sources. The Global Seismographic Network (GSN) and the Program for Array Seismic Studies of the Continental Lithosphere (PASSCAL) and the Joint Seismic Program (JSP) are three IRIS programs that presently send data to the IRIS DMS. The DMC also archives data from some regional networks in the United States and in the Former Soviet Union (FSU).

The DMC is one component of a larger Data Management System (DMS) consisting of several components including two Data Collection Centers (DCCs), one Data Management Center (DMC), a DMC Host at the University of Washington, a Wave form Quality Center (WQC), and the Moscow Data Analysis Center (MDC) in Russia.

The two DCCs are responsible for data quality control and reformatting of data into the FDSN SEED format. The IRIS/USGS DCC in Albuquerque is operated by the USGS and provides data from several networks including the DWWSSN, SRO, ASRO, CDSN, TERRAscope, and IRIS/USGS GSN networks. The USGS DCC has also made historical data, beginning in 1972, available in SEED format and the DMC is presently archiving and distributing these data. The IRIS/IDA DCC is operated by the University of California at San Diego and sends data from stations in the IRIS/IDA component of the IRIS GSN to the DMC in Seattle.

The DMC is situated in Seattle, Washington and is hosted by the University of Washington. The DMC receives data from the IRIS/USGS DCC, the IRIS/IDA DCC, from FDSN Data Centers, and from Regional Networks. The DMC archives these data in a Data Base Management System (DIRTS) and stores the wave form data in a 8.6 terabyte mass storage system. Data access tools are developed at the DMC to make access to the data simple for seismological researchers.

The WQC is situated at Harvard University. The WQC uses IRIS and other data in the routine production of the CMT. As problems with data are discovered, the WQC forwards these data problems to the IRIS DMS. Beginning in 1993 the WQC is now using the Data Problem Reporting (DPR) mechanism established by the DMS.

The University of Washington host of the DMC is actively involved in several projects. The development and operation of the GOPHER real time data retrieval system is done by the University

(*) IRIS DMC, 1408 NE 45th Street, Seattle WA 98105, U.S.A.

of Washington. Active projects of the UW researchers involve analyzing many wave forms from the DMC archive and performing quality control on the SEED volumes produced by the DMC. The University of Washington also uses the DPR mechanism.

The MDC supports the activities of seismologists from the Former Soviet Union (FSU). Data requests by FSU researchers are handled by the MDC in coordination with the IRIS DMC. In exchange for financial support, the MDC is developing software modules that will eventually be incorporated within the operational routines of the IRIS DMC. A high priority during this first year of operation was to improve the communication lines within Russia for use by US and FSU seismologists. A program to convert hypocenter information into the WMO New Telegraphic Format has been written and is currently being tested.

The IRIS DMC wave form archive

The IRIS DMC presently has data from the following networks archived and available.

| | |
|--------------------|--|
| ASRO | (US Geological Survey) |
| CAUCUSUS Network | (IRIS JSP) |
| CDSN | (US Geological Survey) |
| CNSN | (Geological Survey of Canada) |
| DWWSSN | (US Geological Survey) |
| GEOSCOPE | (Institute of Physics of the Globe, Paris) |
| GRAEFENBERG | (Erlangen, Germany) |
| IRIS/IDA | (IRIS GSN) |
| IRIS/USGS | (IRIS GSN) |
| Khirgizian Network | (IRIS JSP) |
| MEDNET | (Istituto Nazionale di Geofisica) |
| POSEIDON | (Earthquake Research Institute) |
| SRO | (US Geological Survey) |
| TERRAscope | (Caltech and US Geological Survey) |
| WSRN | (University of Washington) |

Data from all of the above networks is available to the user community with the exception of the KNET, Caucasus, and WRSN networks. Final testing remains for those networks.

As of March 27, 1994 the total archive contained 447 gigabytes of GSN/FDSN data in 131 212 station day files. All of these data have been archived since the beginning of 1992 and the wave forms were transferred to the mass storage system, the IRIS DIRTS data base was also populated with the descriptive information about the stations and channels as well as with information from the Hypocenter Data Files (HDF) from the NEIC.

Figure 1 shows the rate of growth in the IRIS DMC archive for GSN/FDSN data.

The IRIS DMC Archive is dominated by data from the IRIS networks. Figure 2 more clearly indicates the amount of data that is archived from the other networks of the FDSN.

SPYDER – System to Provide You Data from Earthquakes Rapidly – near real time access to data

The SPYDER system was developed by Steve Malone at the University of Washington through a subaward from the IRIS DMS. It is now maintained by Dr. Malone and Sandy Stromme as part of the UW DMC Host activity. Since its inception the system has expanded significantly. This system

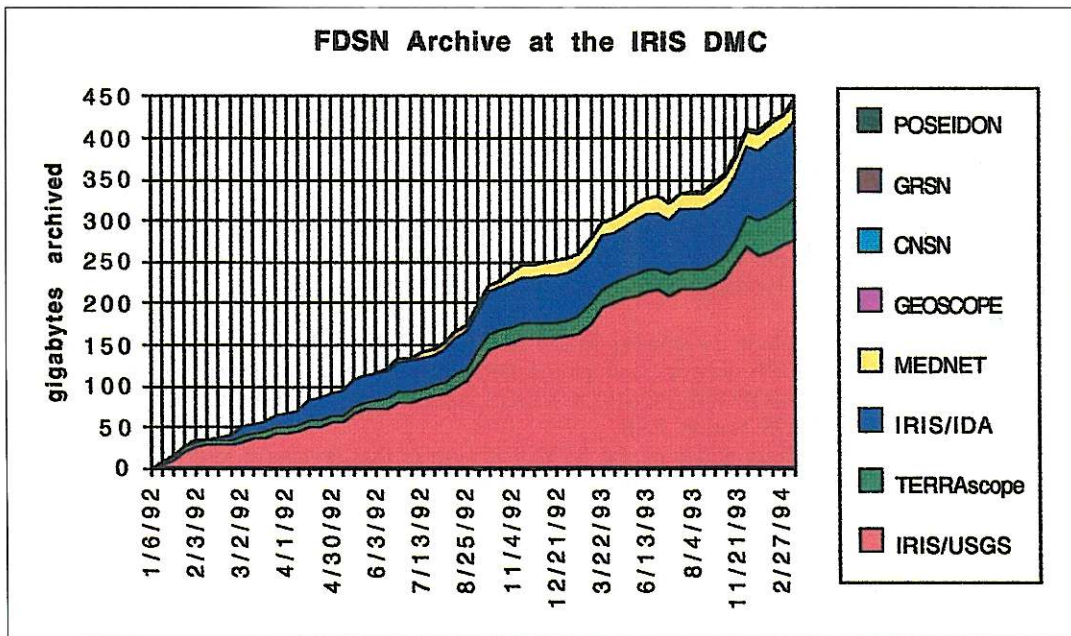


Fig. 1

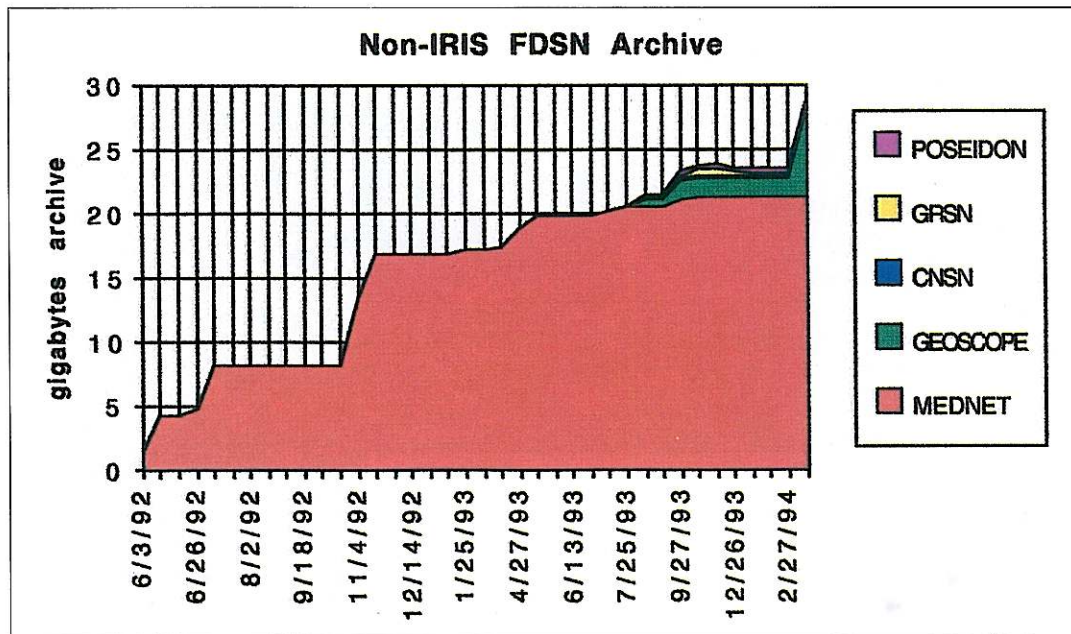


Fig. 2

now has nodes installed at six locations outside the continental United States including nodes at FDSN members in Japan (POSEIDON), the Netherlands (ORFEUS), Australian National University, Institute of Geological and Nuclear Sciences in Wellington and in Potsdam. The goal of IRIS is to place systems that are located geographically close to IRIS stations and therefore reduce telephone access problems. The data are returned to the originating SPYDER location using the INTERNET. The system has been working very well in Tokyo, Utrecht, Australia, New Zealand, Hawaii, Caltech and the DMC. The system was also recently installed in Potsdam, Germany in conjunction with the GEOFONE program.

Another advantage of distributing SPYDER systems is that the time needed to recover data is greatly reduced. There are presently up to 34 stations that return data in near real time to the IRIS DMC GOPHER system. It typically takes one to two hours to return data from all stations after an earthquake.

The 34 stations now include 24 stations in the IRIS/USGS network, 5 from the IRIS/IDA network, 4 from the GEOSCOPE network and station DPC in the Czech Republic. Some stations are joint stations such as the GEOFONE/POSEIDON/IRIS station PMG, the GEOSCOPE/IRIS station KIP, and the MEDNET/IRIS station PAB.

Figure 3 shows the various nodes of the SPYDER system that presently return data to FDSN data centers and the seismological community in general.

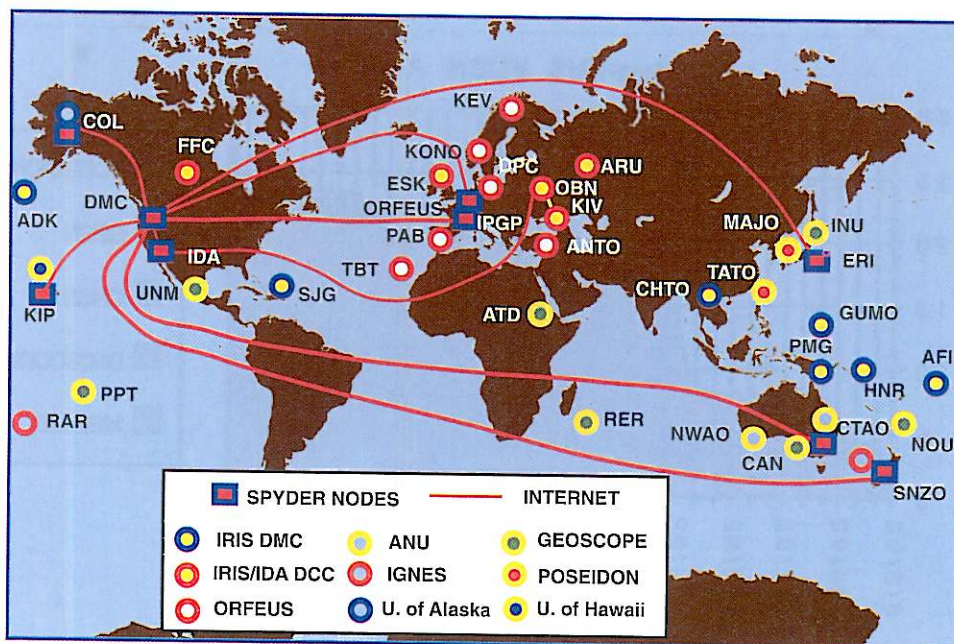


Fig. 3

Software activities

The IRIS DMC has long been developing tools related to:

- SEED format support
 - RDSEED, the FDSN SEED reading program
 - POD, a SEED writing utility
 - VERSEED, a SEED format verifier
 - EVAL_RESP, a SEED response computation engine.
- Data base management software for seismic data
 - DIRTS – a network DBMS closely coupled to the SEED format.
- User data request tools
 - BREQ_FAST, a batch method of making requests by specifying station-channel-time windows.
 - RUMBLE, a batch method of requesting data specifying events
 - XRETRIEVE, an interactive X window based data request tool
 - XTRACT, an interactive X window GUI SQL generation interface
 - SPROUT, an SQL interface to DIRTS.
- SPYDER, System to Provide You Data from Earthquakes Rapidly
This software has been made available to all FDSN Data Centers and can be used to access Quanterra Data Loggers such as those operated by GEOFONE, IRIS, and other institutions. The access method combines INTERNET data transfer between data centers and high speed modem connections between the SPYDER nodes and the Quanterra stations.
- Analysis Software
The IRIS DMS will soon support the Programmable Interactive Toolkit for Seismological Analysis (PITSA) software for both SUN workstations and IBM PC computers.

FARM – FTP access to event volumes

The IRIS DMC continues to support a large SCSI disk farm that is attached to the main user interface computer, dmc. These disks are attached to the anonymous ftp area of dmc. We are in the process of building complete SEED volumes for all events of moment magnitude greater than or equal to 5.8 and for all events of moment magnitude greater than or equal to 5.5 if the event depth is 100 km or greater. The windows are the same as the windows on the NEIC CD-ROM with the exception that the window length for broadband and higher frequency triggered channels is an hour after the estimated P arrival time at the station. SEED volumes are kept to sizes of about 20 megabytes or less so that they can be easily transferred over the INTERNET. It is our belief that we can improve service to the community by pre-assembling event volumes like this at the DMC. This will ultimately reduce the work load on the DMC as well. 1993 saw 754 shipments of FARM SEED volumes since the spring of 1993.

General data distribution statistics

This year has seen the user load on the DMC grow exponentially. From 1989-1992 the DMC physically shipped a total of 2103 items to users all over the world. During 1993 the DMC shipped a total of 3790 items. These include shipment of programs, documentation, assembled data products, FARM products and customized user requests for data. During the period 1989-1992 the

DMC shipped 1088 data volumes. In 1993 the DMC shipped 3043 data volumes exceeding the number shipped in the previous four years by nearly a factor of three. Figure 4 shows the total number of physical shipments from the DMC for the past 5 years. This includes shipments of data, software and documentation.

The number of seismograms shipped also continues to increase. It is this number that has a great impact on how actively the mass storage system is accessed. Figure 5 provides some indication of the total number of seismograms and total volume of customized SEED data the DMC has shipped for the past several years. Figure 5 only contains statistics on the customized SEED volumes we build. It does not include PASSCAL shipments, assembled data products and FARM shipments. The total amount of data shipped from the IRIS DMC in 1993 was 275 gigabytes, far more data than was recorded by all FDSN networks in 1993.

Figure 6 shows the total number of shipments sent to the various FDSN networks.

Shipments to FDSN members receive priority in request servicing. After shipments to IRIS members, shipments to FDSN members are the most numerous. In total, the FDSN received 418 shipments in 1993, 345 of which were data shipments. During this same period IRIS members received 3372 shipments, 2698 of which were for data. FDSN activity represents about 11% of the total output of the IRIS DMC.

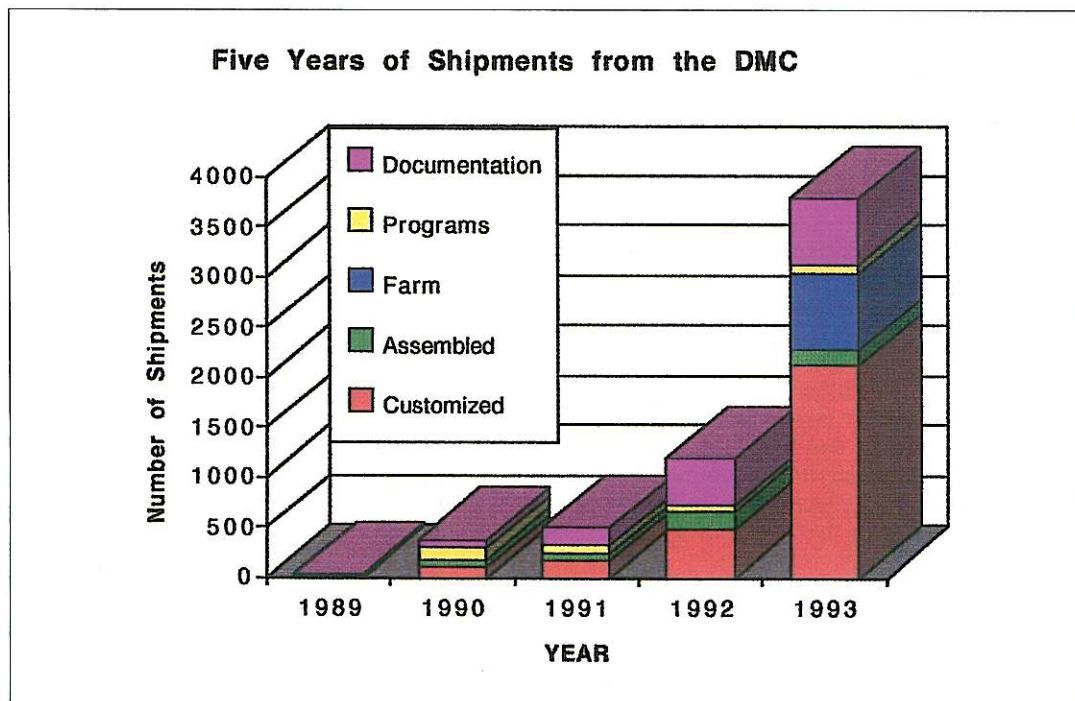


Fig. 4

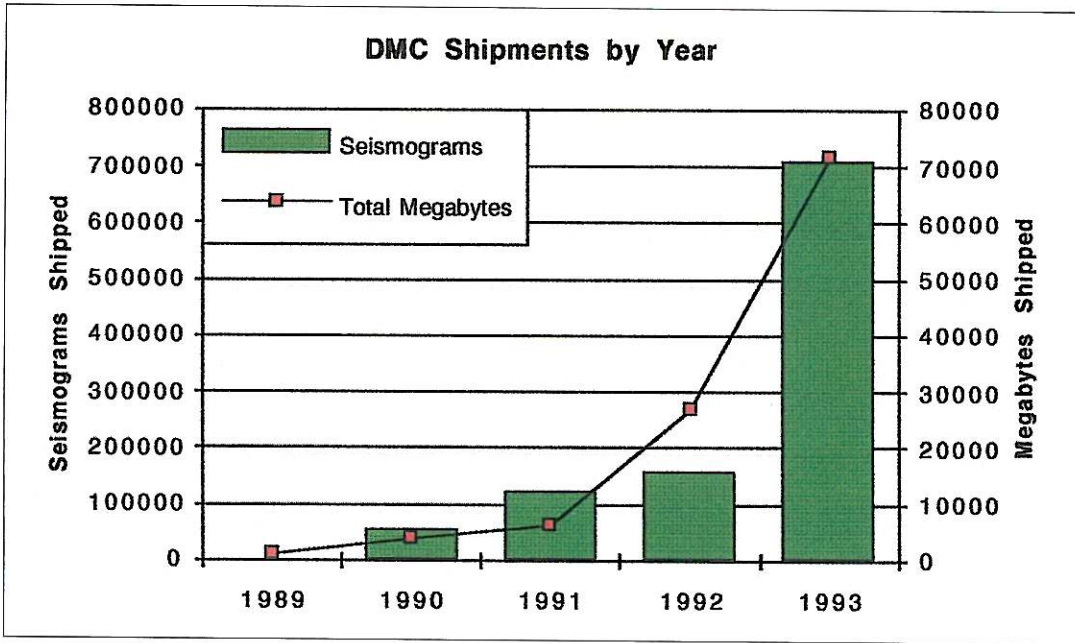


Fig. 5

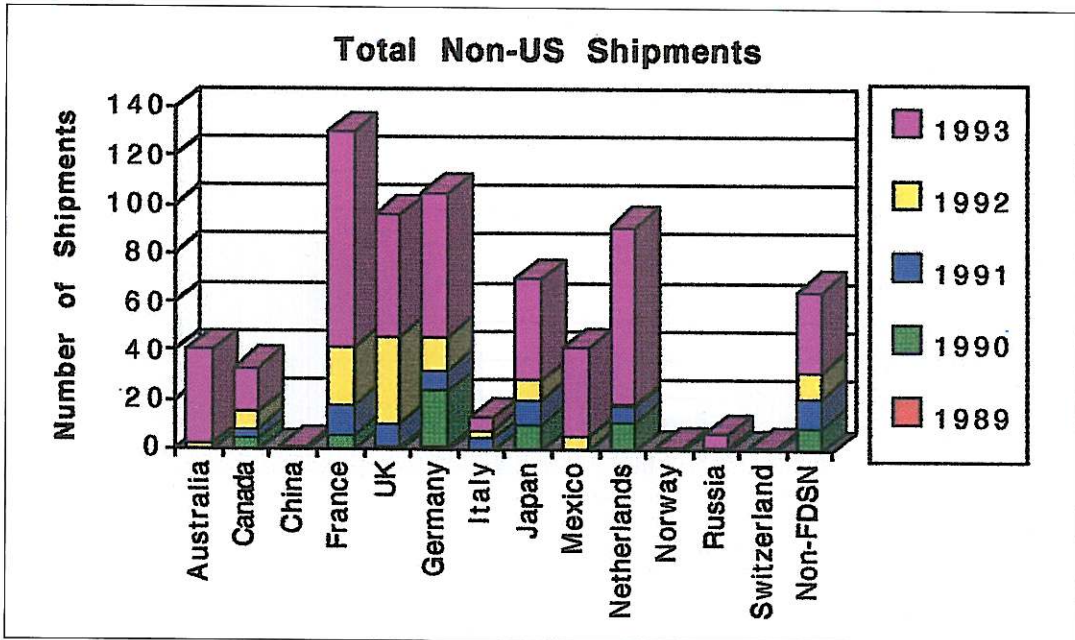


Fig. 6

Electronic accesses

The DMC is seeing an ever increasing amount of electronic activity by users. First and foremost is the activity in the area of the IRIS DMC Electronic Bulletin Board and of SPYDER. Figure 7 shows how the general level of usage has increased for these two common methods of accessing the IRIS DMC. There were nearly 20 000 accesses of the IRIS DMC through either the Electronic Bulletin Board or the SPYDER system in 1993. Many of these accesses came from members of the FDSN.

The DMC places a considerable amount of information in an anonymous ftp area on the main user interface computer. We instituted logging of ftp transfers on this machine in April of 1992. During 21 months we have seen the following amount of anonymous ftp activity in the major categories listed (fig. 8). PROGRAMS includes all software, USER DATA are ftp transfers of customized SEED volumes, EBB is information that users transfer out of the Electronic Bulletin Board, MANUALS are the various manuals IRIS has available by anonymous ftp, and the TUTORIAL category indicates users that have transferred information on how to access data at the IRIS DMC.

We estimate that nearly one half of all data shipments are now done electronically using anonymous ftp.

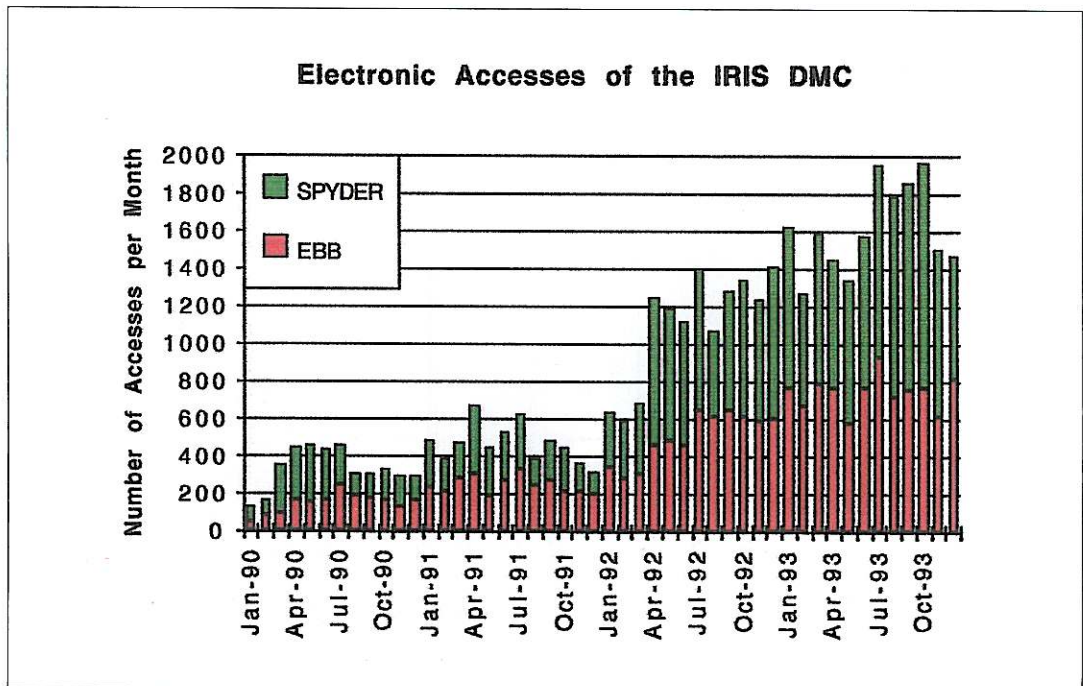


Fig. 7

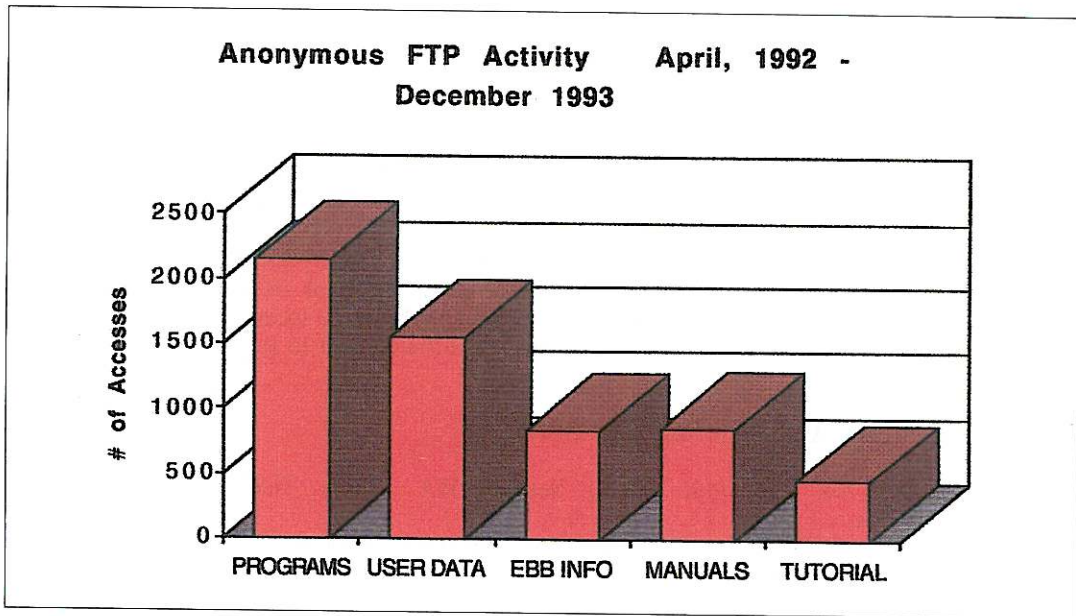


Fig. 8

Documentation

The IRIS DMS produces several manuals to assist users of seismological data. Presently we have the following manuals available.

- 1) SEED Manual – First edition
The first edition was released in 1990 and since that time over 800 manuals have been distributed.
- 2) SEED Manual – Second edition
The second edition was released early in 1993. This manual incorporates all modifications made to the format at the present time.
- 3) SEED Programmer's Toolkit
This manual documents C routines written by Scott Halbert of the USGS and Reinoud Sleeman of ORFEUS. These routines can be used as a library of routines by which individuals can develop SEED reading and writing utilities.
- 4) POD – The IRIS SEED writer
This manual is nearly ready to release. It documents the use of the IRIS SEED writer. Many of the details of the SEED format are taken care of by this utility. Developers of new SEED writers can concentrate on assembling the information needed to build the various SEED blockettes. Documentation is included for an IRIS STEIM compression utility that understands how to include blockettes in SEED data records.

- 5) SPROUT – The user interface to the IRIS DBMS
This manual documents the IRIS DBMS. It not only identifies all the fields stored in the data base but provides many examples of how to use this system. This manual is presently being updated.
- 6) DMC tutorial
This short manual provides some specific examples of using IRIS DMC data access utilities. This manual concentrates on using SPROUT to make data requests.
- 7) FDSN station Book
Starting with contributions from the various FDSN member networks, Kris Skjellerup of the IRIS DMC has nearly completed the FDSN station book. This book includes basic information about each of the FDSN stations, contains color photographs of the location when provided, and contains noise estimates for all stations that had contributed data to the DMC.

Accessing the FDSN Archive at the IRIS DMC

IRIS DMC, 1408 NE 45th Street, Seattle, WA 98105, U.S.A., (206) 547-0393
(206) 547-1093 Fax.

Electronic access is made through the main user interface computer. You can access the IRIS DMC Electronic Bulletin Board by

`rlogin dmc.iris.washington.edu -l bulletin` and use the password board.

The DMC can also be accessed via high speed modem by dialing (206) 547-1294.
The INTERNET address of the above system is 128.95.166.2.