Report from the
FDSN Archive for Continuous Data
at the IRIS DMC

2009 FDSN Meeting
11-16 January 2009
Capetown, South Africa
by
Tim Ahern, IRIS DMS
(prepared 8 January 2009)

Introduction:
A significant shift has occurred since the FDSN meetings in 2005 and 2007. In 2005 14 FDSN networks were sending data to the FDSN archive in real time. In 2007 that number had grown to 31 networks with an additional member (Canada) sending data with a latency of 24 hours. In 2008 the number of real time FDSN networks sending data to the FDSN Archive at IRIS grew to 38. Data from the majority of FDSN networks routinely flow into the FDSN Archive located at the IRIS DMC. At the IRIS DMC, data received in real time are made available in real time via a SeedLink server. After 30 hours the real time data are archived in DMC mass storage systems. SPYDER® and FARM Products are routinely generated and made available through tools such as WILBER, the Data Handling Interface (DHI) and jWeed within a few hours of an event. All request tools at the DMC can get to all data the DMC has archived within 30 hours after recording.

The DMC continues to ship large amounts of seismological data to researchers around the world. This report will summarize current activity at the FDSN Archive for continuous data at the IRIS DMC in Seattle, Washington.

In the following summary the FDSN membership is considered to be the 65 organizations or networks from 52 countries. It is believed that all graphics capture the dynamic membership of the FDSN correctly.

Real Time Data Flow
Real time data streaming is the most common data reception method at the IRIS DMC at this time. In general delayed transmission of datasets is only common for data from portable instrument deployments. Typically more than 1800 stations contribute data in real time to the IRIS DMC. The following map shows the locations of these real time stations.
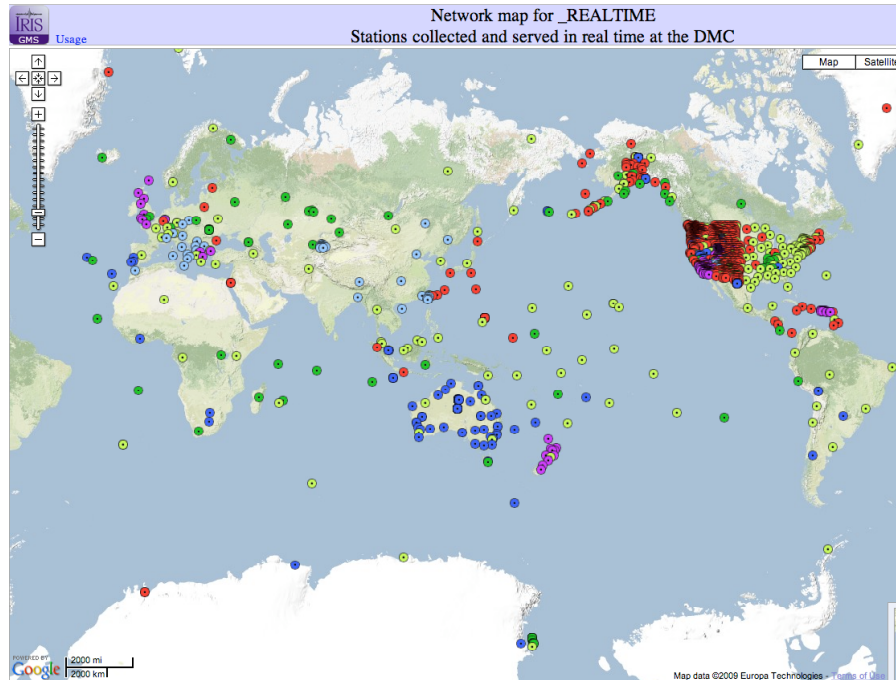
Figure 1. Data from roughly 1800 (1797on January 7, 2008) stations are now received in real time at the IRIS DMC. These are dominated by seismic stations in the US, Europe and Australia/New Zealand but the global extent of the coverage is impressive. We believe this is the largest concentration of real time high quality seismic data of any data center in the world.
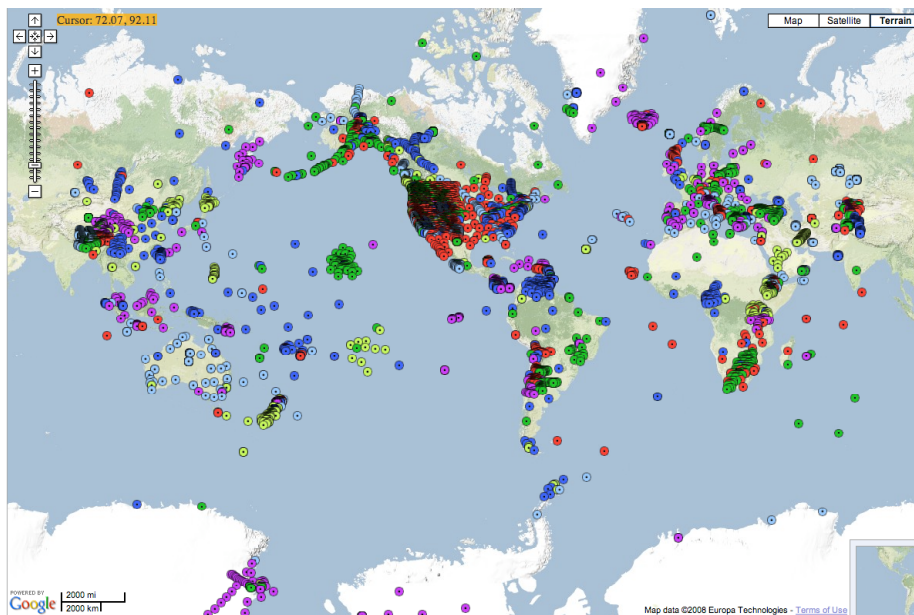

Figure 2. All Stations in the IRIS DMC Archive. The IRIS DMC has data from 10,419 stations available as of December 2008. This is up from 8,512 at the time of the 2007 FDSN meeting. This includes stations operating as early as 1970. The various temporary deployments of programs like PASSCAL and SEIS-UK are clearly visible.

FDSN Data Flow into the IRIS DMC

Data are now flowing into the FDSN archive in real time from 38 of the 65 FDSN networks and from centers in 28 FDSN countries. In most cases not all stations operated by a network flow to the FDSN archive, usually just stations that are a part of the FDSN network or which provide improved global coverage and are shared as part of a bi-lateral agreement between the organization and IRIS. FDSN networks (and countries) sending data in real time include:

- Australia,
- Bangladesh,
- Belgium,
- N. California Seismic Network, Berkeley,
- S. California Seismic Network, Caltech,
- Canada (delayed 24 hours),
- China (delayed 30 minutes),
- Czech Republic,
- Geoscope, France
- GEOFON, Germany
- GRSN, Germany
- Great Britain (GB and UK networks)
- Greece, Thessaloniki
- Greece, National Observatory Athens
- GTSN, USA
- Kazakhstan,
- Kyrgyzstan,
- IRIS GSN, USA
- Israel
- Japan, Pacific 21
- Japan, JMA
- Malaysia,
- MedNet, Italy
- Mexico
- New Zealand,
- Netherlands, ORFEUS (networks NL and NA),
- Poland,
- Portugal,
- Romania,
- Russia,
- Switzerland,
- Taiwan, BATS,
- ANSS/USGS (Networks US, LB and CU)
- PRSN and PRSMN, Puerto Rico.
- United Nations (selected stations of the IMS)
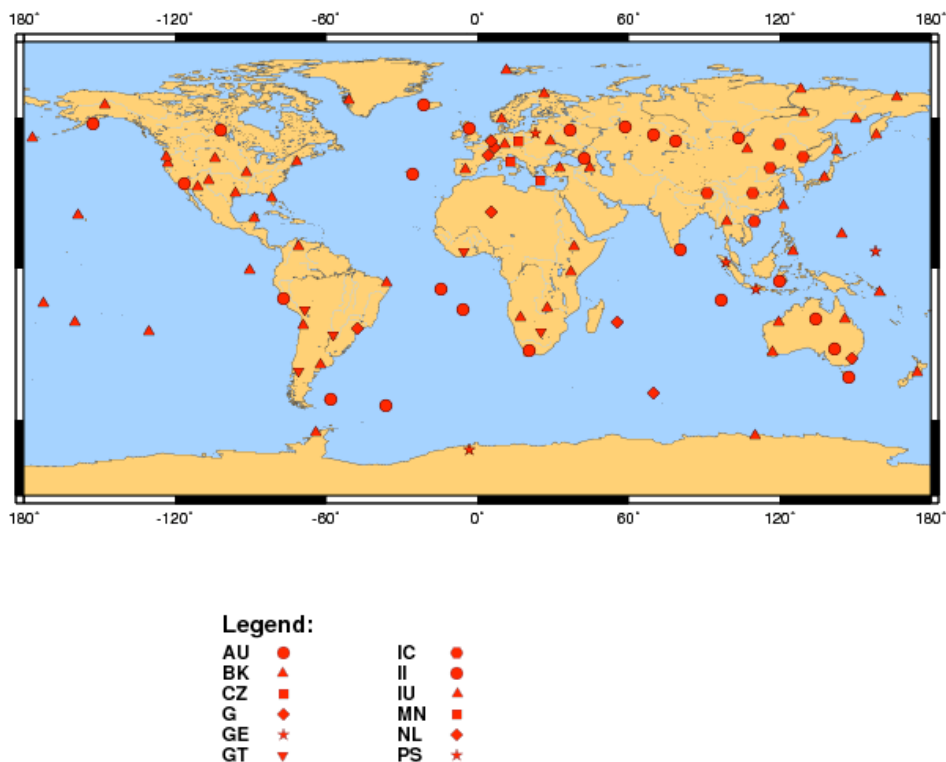
Real Time FDSN Stations 12-Jun

Legend:

AU	●	IC	●
BK	▲	II	●
CZ	■	IU	▲
G	◆	MN	■
GE	★	NL	◆
GT	▼	PS	★

Figure 3. Real Time Data Flow from FDSN Stations. This figure shows the stations providing data to the FDSN archive in real time. These are as specified in the _FDSN virtual network.

The FDSN holdings at the IRIS DMC have continued to increase. The holdings now exceed 8.7 terabytes, up from 6.2 terabytes at the 2007 FDSN meeting. While not the most voluminous data set they are extremely important due to their wide geographic distribution and high data quality.
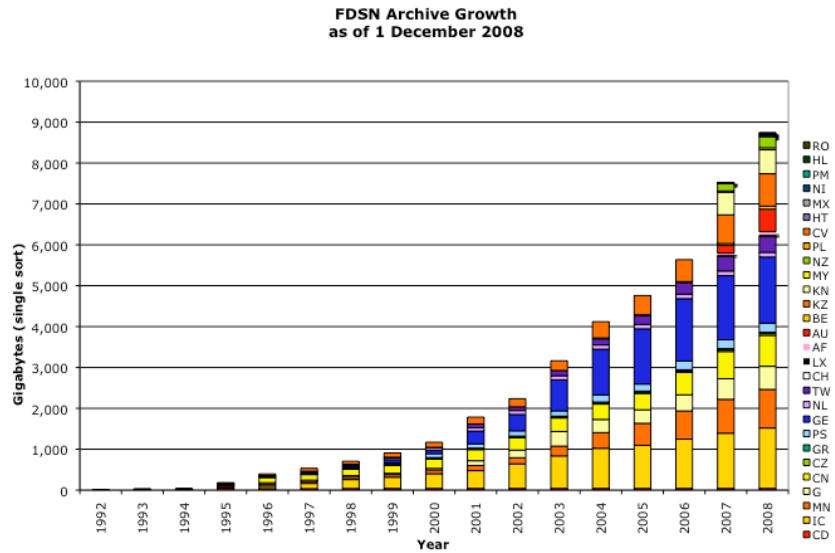
FDSN Archive Growth
as of 1 December 2008

Figure 4. The FDSN archive now totals roughly 8.7 terabytes (single sorted) with data from 28 FDSN networks not including US data from US operated networks (BK, CI, CU, GT, IU, II, LB, US, PR).
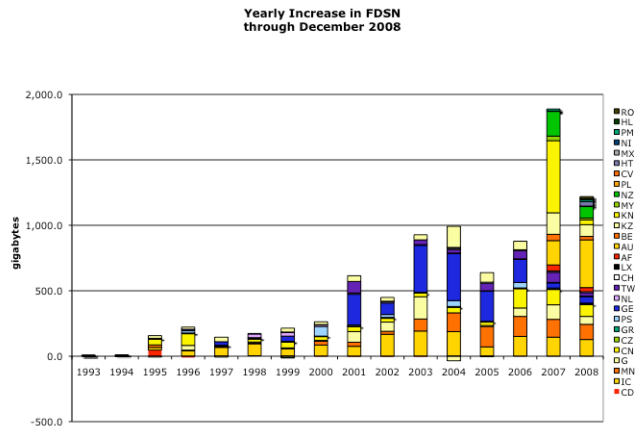

Yearly Increase in FDSN
through December 2008

Figure 5. The above figure shows the net increase of FDSN data entering the FDSN archive by year. Data from IRIS GSN Networks of IU and II are not included. Data from US operated networks are also omitted (BK, CI, CU, GT, LB, US, PR) the figures for 2008 are valid through 1 December 2008. The large increase from KN stations reflects the fact that Kyrgyzstan became part of the FDSN in 2007, not

that that much new data was received that year. Negative values reflected in the above chart are a result of removing data from the archive, usually in preparation for receiving retransmissions of the data. In general the FDSN Archive is receiving about 1 terabyte of data from non-US sources.
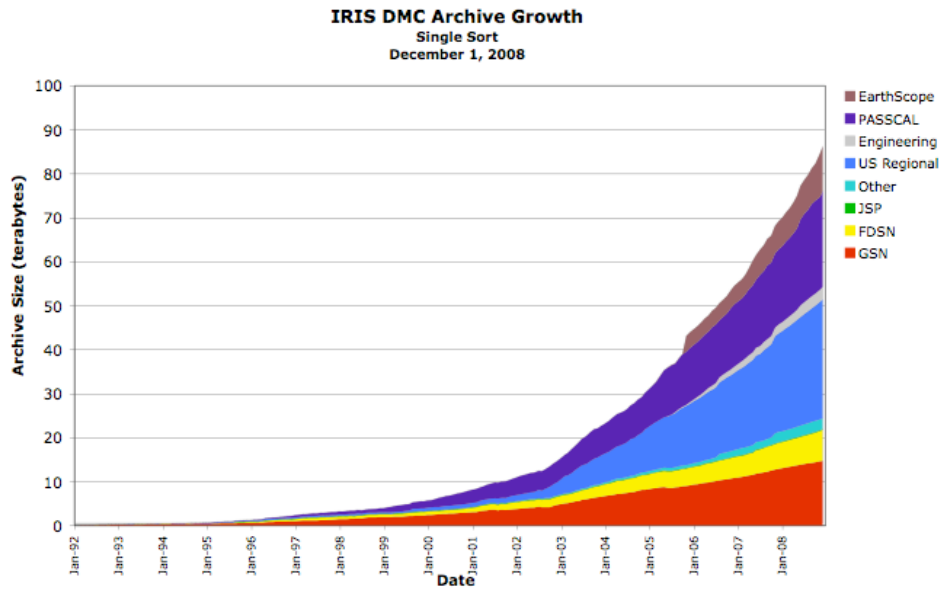

Figure 6. The above figure shows the total amount of data archived at the IRIS DMC. As of the beginning of December 2008 the IRIS DMC had a total of 86.3 terabytes in the archive, up from 61 terabytes at the 2007 FDSN meeting. The non-IRIS FDSN contributions are shown in yellow. The data from the IC network is included in the GSN value, not the FDSN value in the above chart explaining the discrepancy between Figure 4 and 5.

Data Shipments from the Archive (Excluding Real Time and DHI Data Shipments)

The IRIS DMC has seen a significant increase in the number of shipments in 2008. At the present time we are estimating 650,000 customized requests for data will be serviced 2008. This is more than double the number service in 2007.
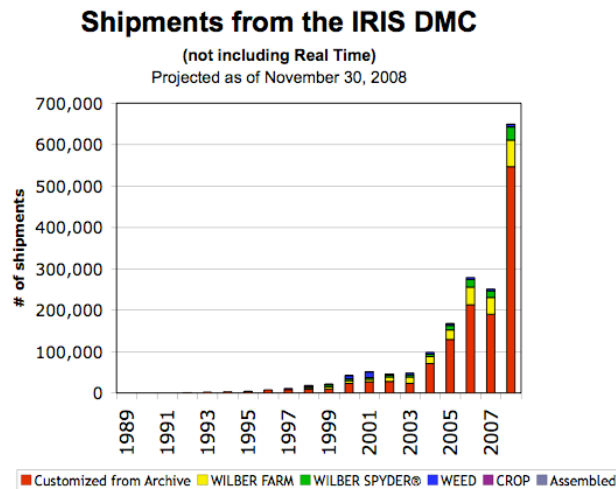

Shipments from the IRIS DMC
(not including Real Time)
Projected as of November 30, 2008

Figure 7. This figure shows the number of shipments from the Archive, and the SPYDER® and FARM products through WILBER that we project will be shipped this year. A significant fraction of this increase is FDSN members outside the United States.

Of the 650,000, approximately 140,000 are being shipped outside the United States with France generating the most requests other than the US as shown in Figure 9.
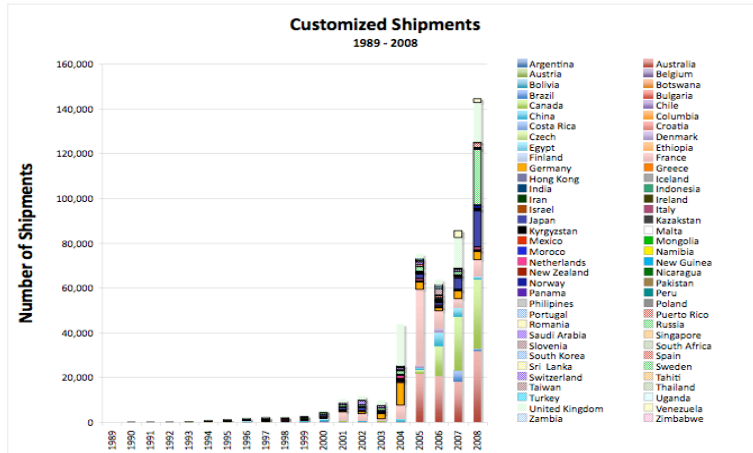
Figure 8. This figure shows the number of shipments flowing outside the United States. We shipped about 140,000 customized requests for data outside the US in 2008. This contains information through the beginning of December 2008.
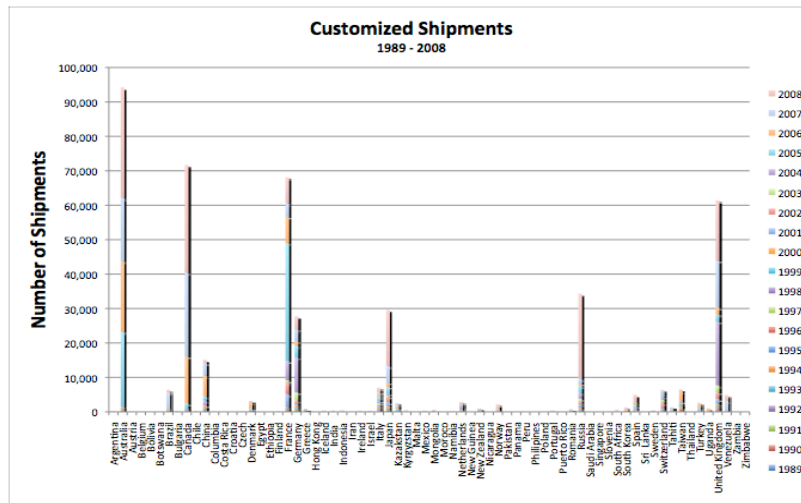

Figure 9. This figure shows the number of customized requests sent outside the US for various countries from 1989 through the beginning of December 2008. The countries making the most use of the FDSN archive at the DMC is Australia, followed by Canada, France, the United Kingdom and Russia. The column labeled Misc contains requests where the country of the requester could not be determined and is likely dominated by US requests.

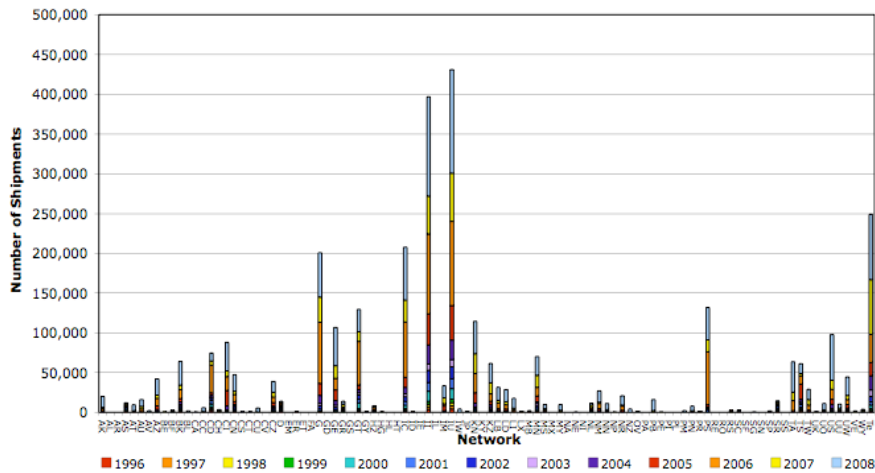**Shipments by Network Code**
as of 1 December 2008

Figure 10. The above figure shows how requests vary by Network Code. The ten most frequently requested networks are IU, II, temporary networks [X?,Y?,Z?], IC, G, PS, GT, KN, GE, and US.

Shipments via all methods (Archive, DHI and Real Time)

Just as the IRIS DMC is now receiving data in near real time through automated techniques, we are also beginning to support distribution of data via streaming mechanisms. We currently only support SeedLink as a streaming protocol. Historically we also supported LISS and the DHI can also effectively serve data in near real time.

The DMC also released a new version of the autoDRM software that provides access to data in the BUD system. The IRIS DMC shows availability from more seismic channels than any other system that is monitored by Waves4U (http://www.seismo.ethz.ch/waves4u/) at ETH.

The most effective way the DMC now has to measure output data rate is the number of bytes we send to end users either from
1. Customized requests to the archive
2. Real time distribution such as through SeedLink or LISS and
3. Data distribution through the DHI.

DMC Shipments by Request Type
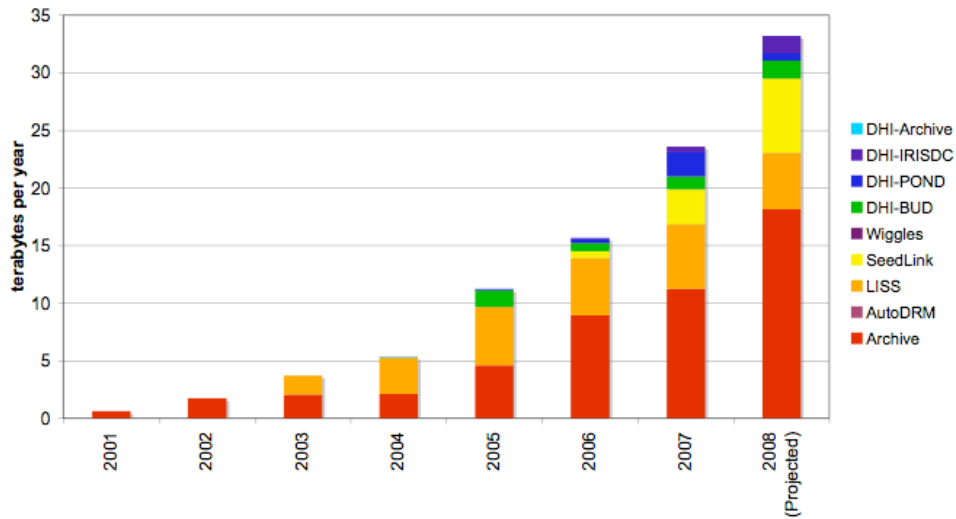Estimated November 30, 2008

Figure 11. This figure shows the number of gigabytes of data shipped conventionally from the archive in red and the amount of data shipped through the two real time methods of (LISS and SeedLink) and DHI based techniques. The DMC discontinued support for LISS in the fall of 2008.

The IRIS DMC will ship a total of more than 34 terabytes of waveform data this year if shipments continue at the current rate.