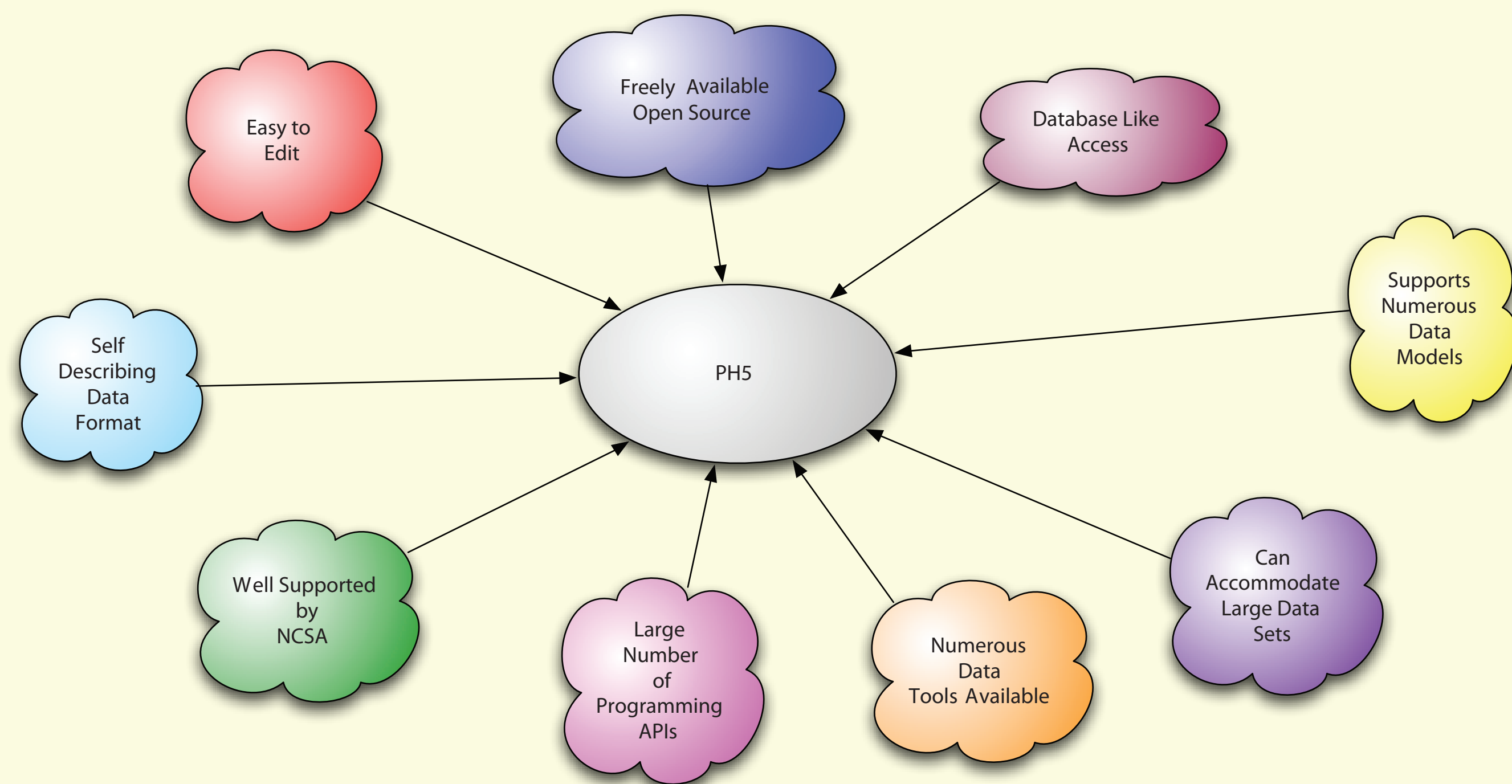


Archiving and Translating Seismic Data using HDF5

Derick Hess, Steve Azevedo, Bruce Beaudoin
IRIS PASSCAL, Socorro NM

Problem

Why HDF5?

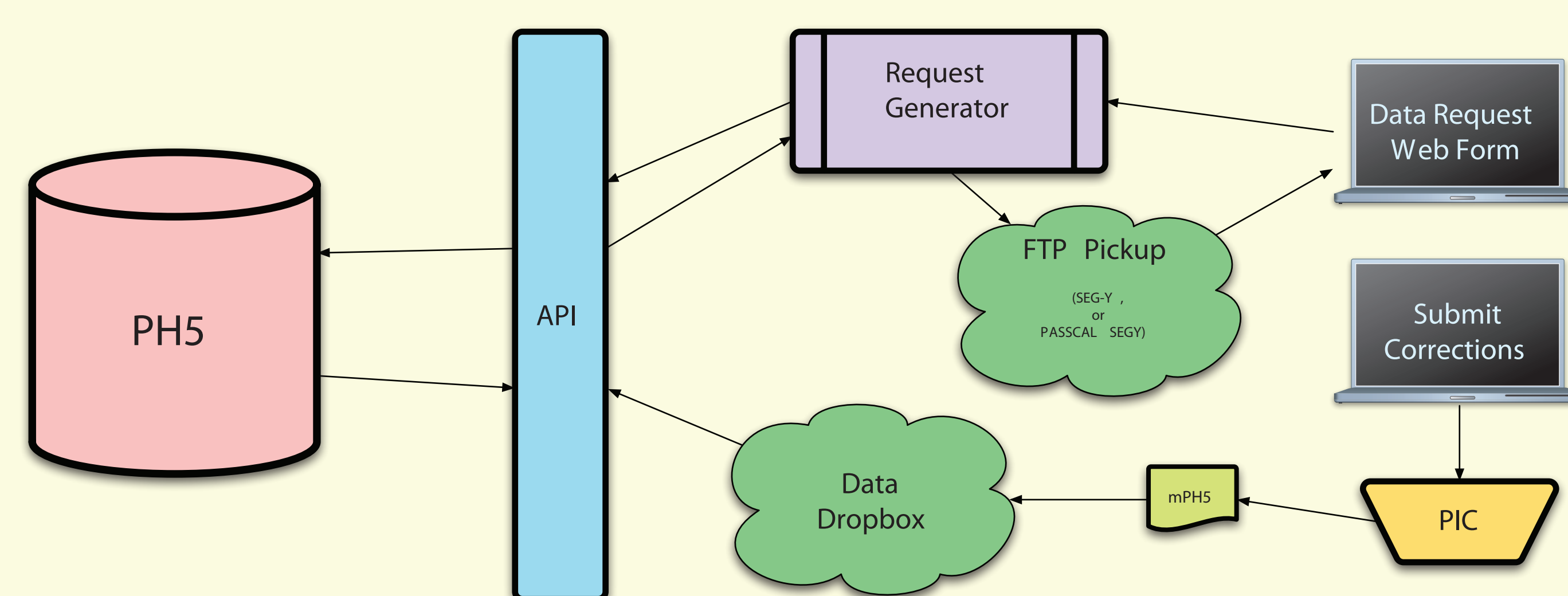


Persons receiving funding for experiments that use IRIS/PASSCAL (Incorporated Research Institutions for Seismology/Program for Array Seismic Studies of the Continental Lithosphere) facilities and equipment are required to submit the data collected to the IRIS DMC (Data Management Center) for archiving. One of the services we perform at IRIS/PASSCAL is assisting principle investigators to archive active source seismic experiments as well as provide data to them in the commonly used analysis format: SEG-Y. Historically data from these experiments was sent to the archive in SEG-Y format but because of the static nature of SEG-Y it has been found to be much more efficient to use HDF5 as an archival and intermediary format.

PH5 is PASSCAL's file organization of HDF5 used for active source data. The extensibility and portability of HDF5 allows the PH5 format to evolve and operate on a variety of platforms and interfaces. Storing data in HDF5 format also facilitates the interactive production SEG-Y gathers, as well as other common seismic data formats based on varying and flexible data request parameters.

To make PH5 more flexible it has now been updated to separate the meta data from the time series data in order to get gains in performance as well as ease of use. By separating the meta data we are able to provide a better web based interface for data requests. The recently developed web interface generates a unique web form and pre-populates much of it based on the meta-data provided to it from the PH5 file. The data requester then can intuitively select the extraction parameters as well as data subsets they wish to receive. The web interface then passes this on to the PH5 processing tools to generate the requested SEG-Y data.

Request and Correction Data Flow



Design

Meta-data Loader:

The task of the meta-data loader is to receive the experiment information from the menu system and read in the corresponding meta-data PH5 file. After reading in all of the meta-data it prepares the data to hand off to the form generator.

Form Generator:

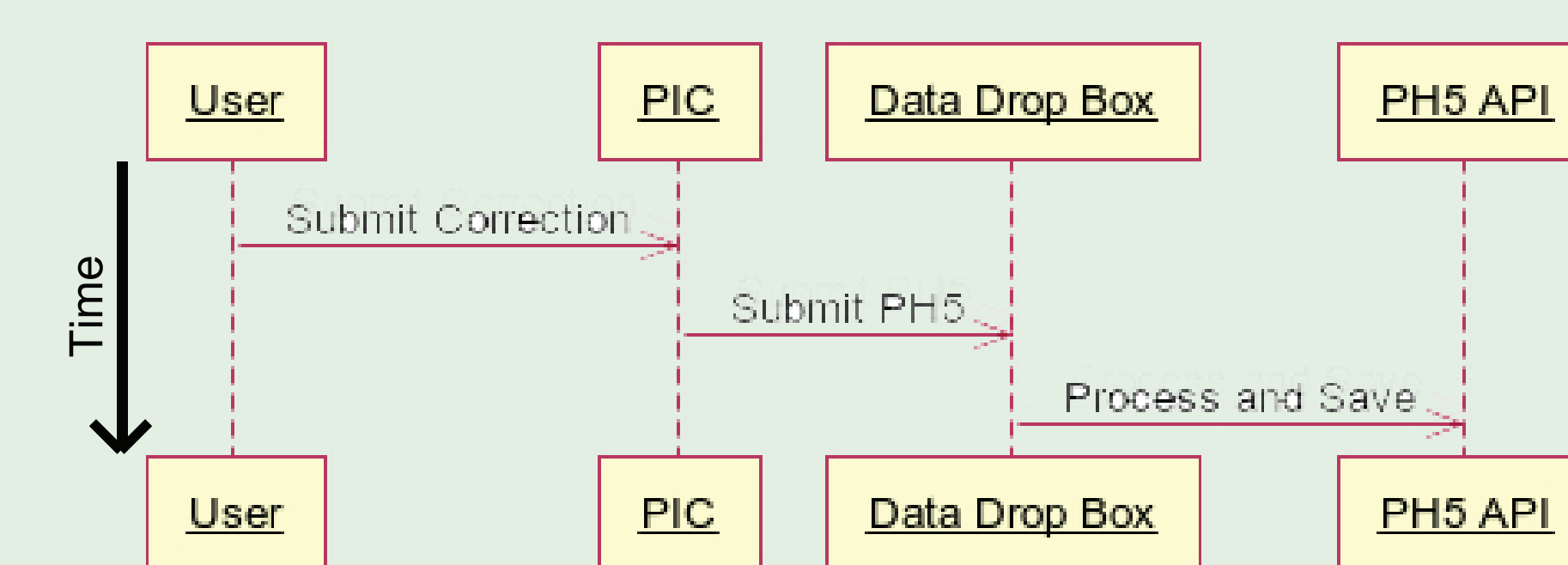
The form generator reads in all the meta-data types from the meta-data loader and create a form based on the meta-data provided, and pre-populate some form fields.

Form Handler:

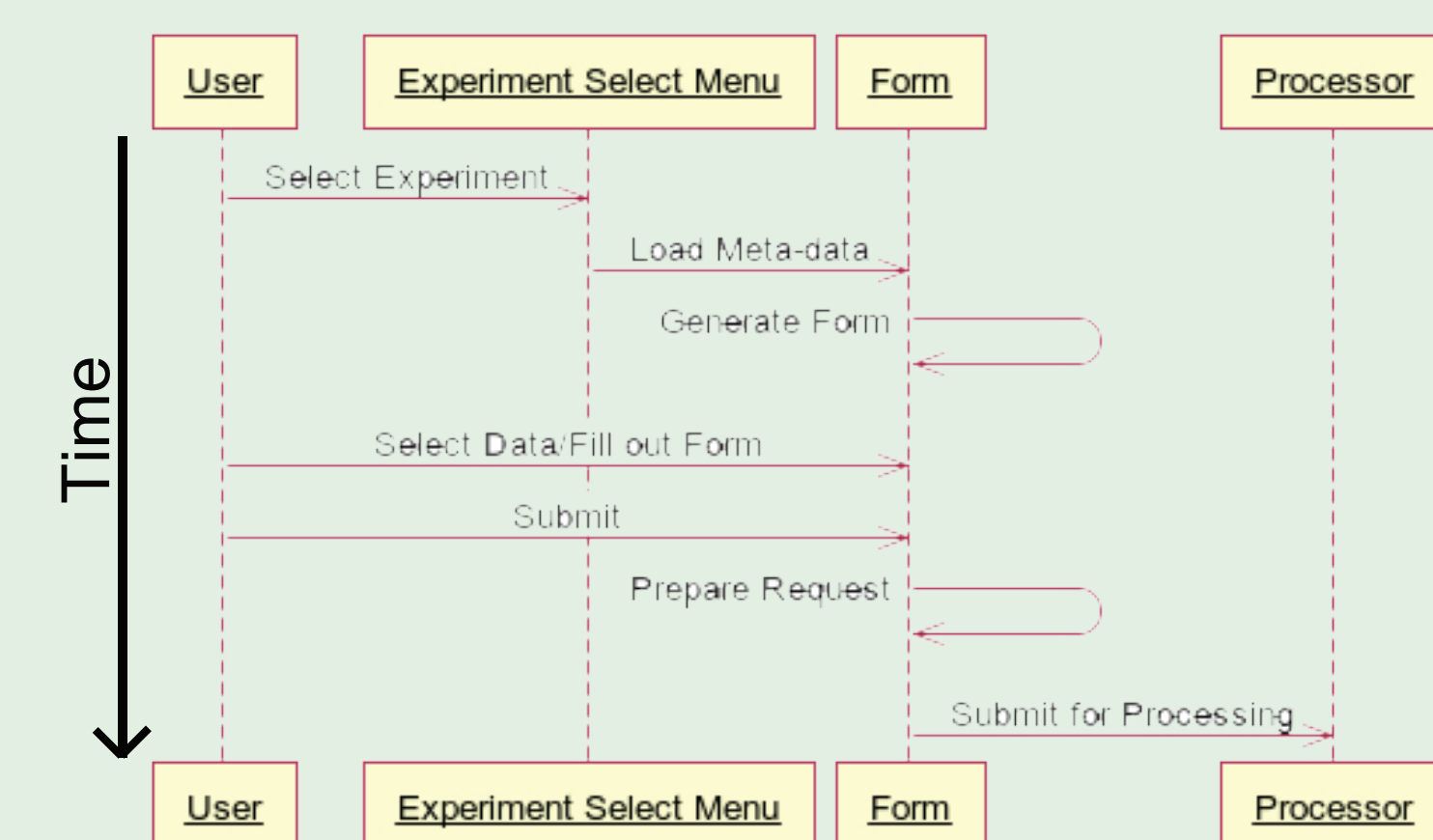
The form handler is called when the form is submitted. After submission the form hands off all the data from the user request to the form handler. It then formats the request and passes it to the processing software.

Sequence Diagram

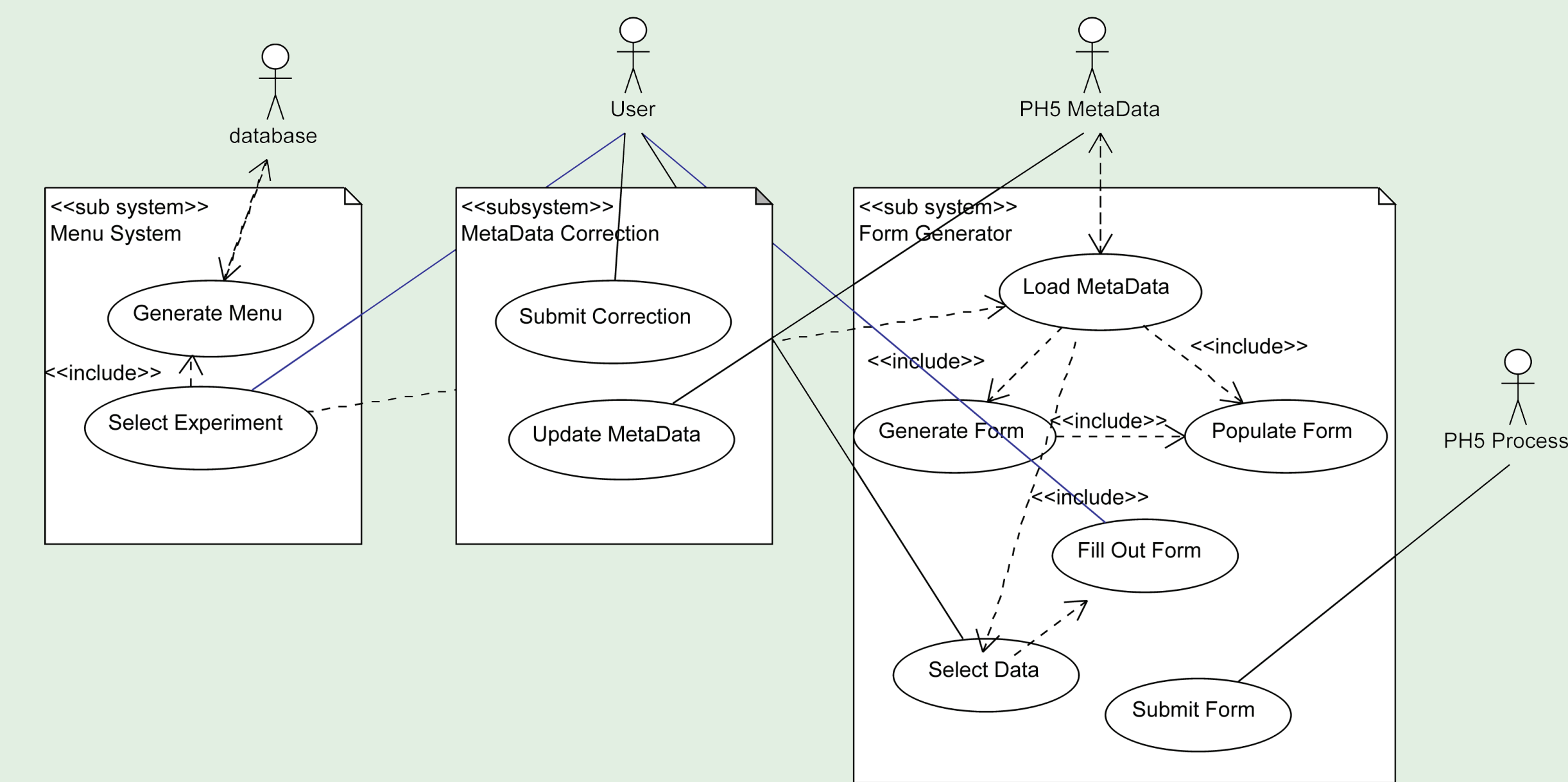
MetaData Correction



Form Interaction



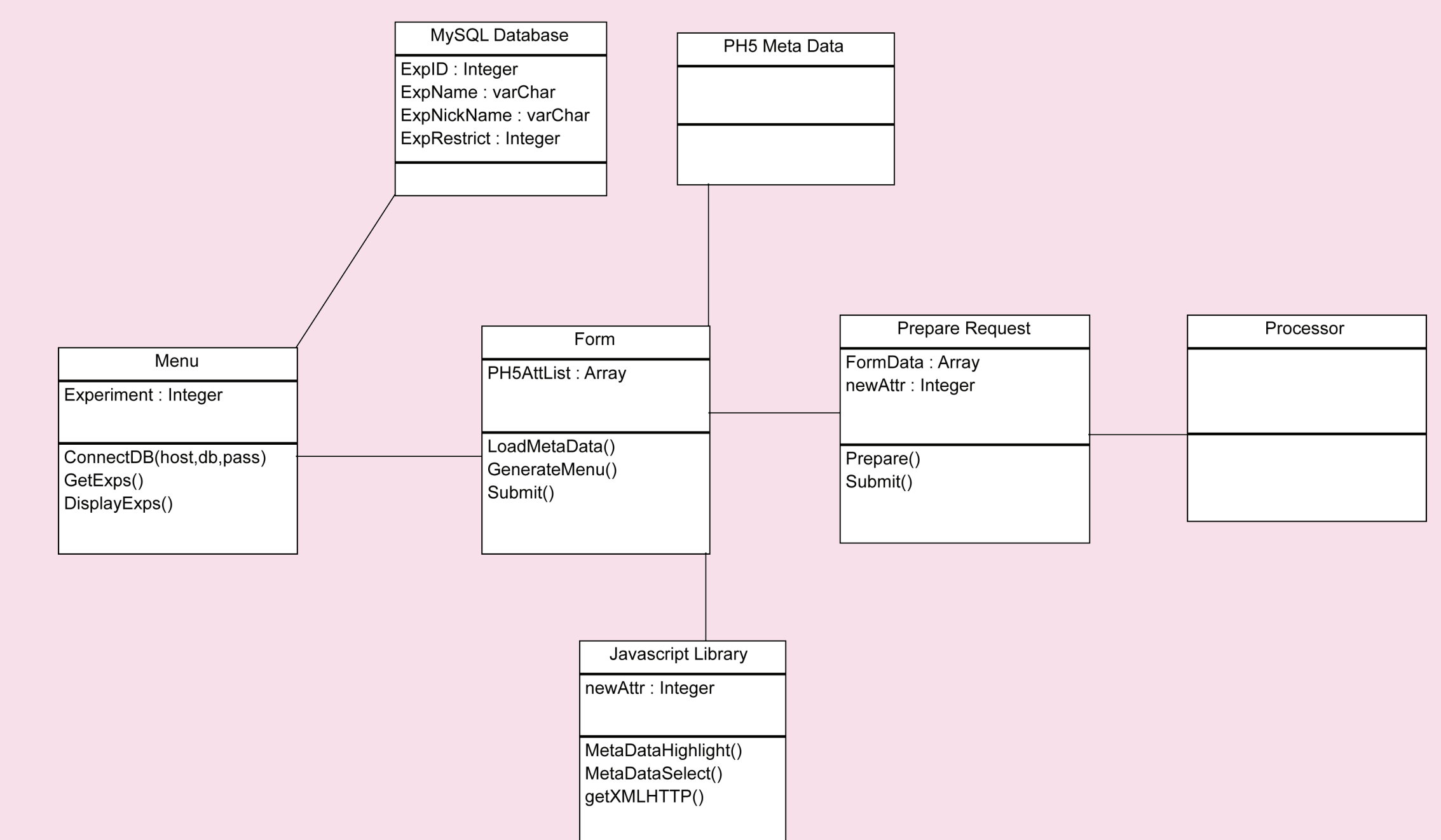
Use Case



Solution

The separation of meta-data into it's own PH5 file has increased usability and productivity in two ways. The first is in the ease of updating and changing the meta-data. Corrections can now be done quickly and efficiently. The second benefit is in the usability of the data request service. Many requestors are graduate students and may be unfamiliar with the intricacies of the file formats. The new form system makes requesting data and data subsets from an experiment much more intuitive. The user now can click on the needed data to add it into the request form reducing the need to manually type in the request information.

System Architecture



Request Form

Please use the information in the right frame to correctly request data.

Data Request Form for UNRESTRICTED Experiment:
07-012 - Mt St Helens Temporary Array - 2005 - MSH_TRANSECT

These files may include pdf's, kmz's (google map), and other format files of interest.
*Restricted data requires prior permission to access it.
*Restricted data will be encrypted prior to being put on the DMC's ftp site.

The fields below will be used to track your request and are all required.
Hover over input fields below for help.

Name:
Institution:
Your E-MAIL:
LABEL:

The fields below specify your request. An asterisk (*) indicates required field.

Data Type: about PASSCAL SEG-Y about Standard SEG-Y Standard SEG-Y
Component:
Decimation (useful for high sample rate data):
 Apply Time Correction to Texan Data
Request By Shot IDs By Time Range Das Serial Number
Hover over input fields below for help. [Click here for detailed help.](#)

*Shot IDs (comma separated list): 2002254, 200852
Array IDs (comma separated list): 002
*Length (secs): Offset (s):

University of Washington, USGS
The purpose is to determine velocity model of shallow edifice of Mount St. Helens using earthquakes in new dome. The layout is a ~4 km-long transect beginning at the base of the old dome and extending north to the pumice plain. Stations 1978 and 1968 were also deployed on the west glacier for observational purposes.

Shot Information and IDs:

shot_id	time	lat	long	elev (m)	size (code)	dept (m)
2002254	20051232123515.000	44.196500	-122.189000	2200.000	2.600	0.000
2003005	20051232123515.000	44.196500	-122.189000	2200.000	1.700	0.000
2004582	20051232123515.000	44.196500	-122.193500	2200.000	1.900	0.000
2005235	20051232123515.000	44.194000	-122.191500	2200.000	2.600	0.000
2006221	20051232123515.000	44.194700	-122.188700	2200.000	2.600	0.000
2012345	20051232123515.000	44.196700	-122.188500	2200.000	2.400	0.000

Array: 001

station	das	lat	long	elev (m)	component
1979	11979	44.205159	-122.188619	1783.000	1
1956	11956	44.206045	-122.188635	1782.000	1
1974	11974	44.206953	-122.18895	1759.000	1

Sample Rate: 100 sps
Deployment Time: 2005123122:00
Pickup Time: 2005123412:00
Components: 1 => 2, 2 => B, 3 => E

Array: 002

station	das	lat	long	elev (m)	component
1938	11938	44.205159	-122.193500	1928.000	1
1968	11968	44.198333	-122.194500	1928.000	1

