LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
- LIGO -
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Technical Note **LIGO-T990109-00 – D** 11/15/99

Reduced Data Set Generation Overview

Rolf Bork, Alex Ivanov and Daniel Sigg

Distribution of this draft:

all

This is an internal working note
of the LIGO Project.

LIGO Hanford Observatory
P.O. Box 1970 S9-02
Richland, WA 99352
Phone (509) 372-8106
FAX (509) 372-8137
E-mail: info@ligo.caltech.edu

LIGO Livingston Observatory
19100 LIGO Lane
Livingston, LA 70754
Phone (504) 686-3100
FAX (504) 686-7189
E-mail: info@ligo.caltech.edu

California Institute of Technology
LIGO Project - MS 51-33
Pasadena CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project - MS NW17-161
Cambridge, MA 01239
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

www: http://www.ligo.caltech.edu/

1 SUMMARY

The reduced data set generator is combination of hardware and software which is installed at both LIGO observatories to enable researchers to obtain their own reduced data stream. The basic features are:

- Data are obtained from the frame broadcaster (~ 1–2 sec delay),
- Date can be reduced in number of channels and in rate,
- Data is written to disk and/or tape,
- Peak data rates of 6 MB/s are supported as well as sustained rates of up to 1MB/s,
- Data volumes of up to 1 TB are supported as well as continuous operations over several months,
- Multiple requests of individual researchers can be processed simultaneously,
- The reduced data set generator operates in batch mode and can automatically recover from system ‘hick-ups’ and reboots,
- A graphical user interface allows the user to assemble and submit batch scripts which describe the reduced data set.

2 SYSTEM OVERVIEW

A hardware overview presented in Fig. 1. Data from the detector is collected by the Frame Broadcaster through the data acquisition system. Data is framed in one second intervals and then broadcasted over gigabit ethernet. The machine responsible for generating the reduced data sets listens to this data stream in ‘real-time’, selects the channels of interests, reduces them if necessary (see Section 3), writes them to an internal disk array and if enough data is collected to tape. Both the input and the output file format is Frames.

The tapes are 50GB (uncompressed) Sony AIT-2. The tape drive is able to write at speeds up 6MB/s (theoretical). The tape robot supports 2 drives and a tape library of about a dozen tapes.

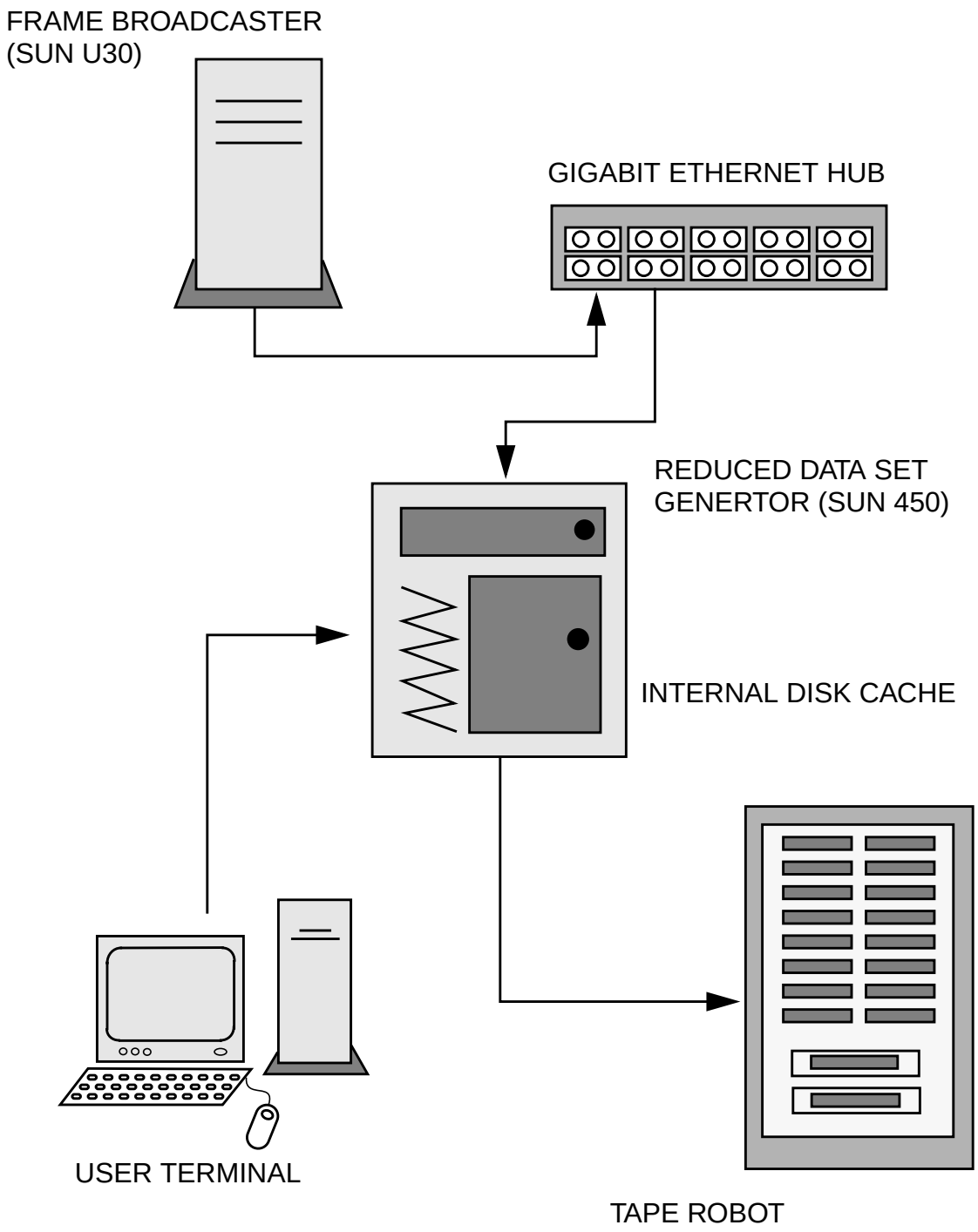

Figure 1: Hardware overview of the reduced data set generator.

3 DATA PREPROCESSING

The reduced data set generator has the capability to preprocess the data before sending it to disk or tape (see Fig. 2). The first stage of the preprocessing is a data collection stage which makes sure that the data segments which are fed into the decimation filters are long enough to yield at least one output data point. The next stage is a time delay filter which can delay the data by an integer number of samples. This stage is followed by the first filter-decimation stage, the down-conversion stage and the second filter-decimation stage. If the down-conversion stage is included the data streams becomes complex. Before the data is partitioned into segments, the time series can be shifted forward in time to compensate for the delay accumulated in the decimation filters and to correct for the calibration. The time shift has to be a multiple of the sampling period after decimation. Since the filter-decimation stages are based on FIR filters, the time delays accumulated within are a multiple of the input sampling period, but no necessarily a multiple of the output sampling rate. To compensate for fractional sample time delays at the output stages, the delay filter at the input can be used for compensation. By default all filter delays are removed automatically and the partitioned data has no remaining time deviation. The output samples are aligned as follows: *(i)* data rates equal to 1 Hz are aligned with the one second GPS clock, *(ii)* data rates of 2^n with $n \geq 1$ are aligned so that every 2^n-th sample is synchronized with the one second GPS clock, and *(iii)* data rates of 2^n with $n < 0$ are aligned so that the GPS time divided by the rate yields an integer. GPS time is counted in seconds starting at Sunday, January 6, 1980, 00:00 UTC.

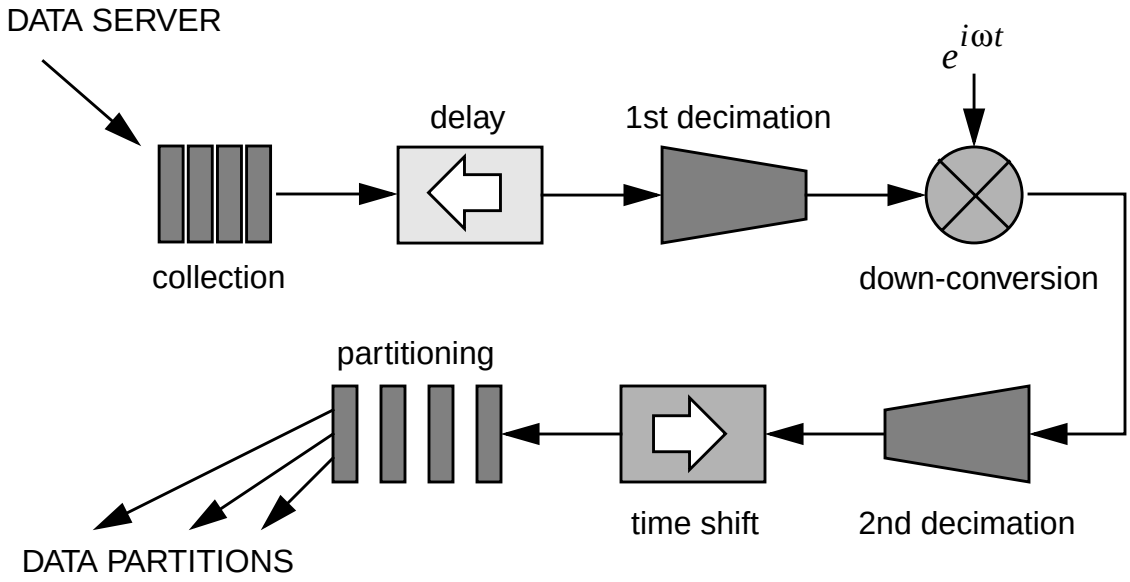

Figure 2: Data preprocessing for the reduced data set generator.

4 USER MODEL

An schematic layout of the software is shown in Fig. 3. The data is obtained through a DMT shared memory partition. A batch processor which is controlled by a queue of data set descriptors feeds the appropriate data channels through a preprocessing stage and then writes it to a disk cache. From there the data is transferred to tape when ready. The user interacts with the system by building a data set descriptor and submitting it to the work queue. Data set descriptors use the LIGO lightweight format(?).

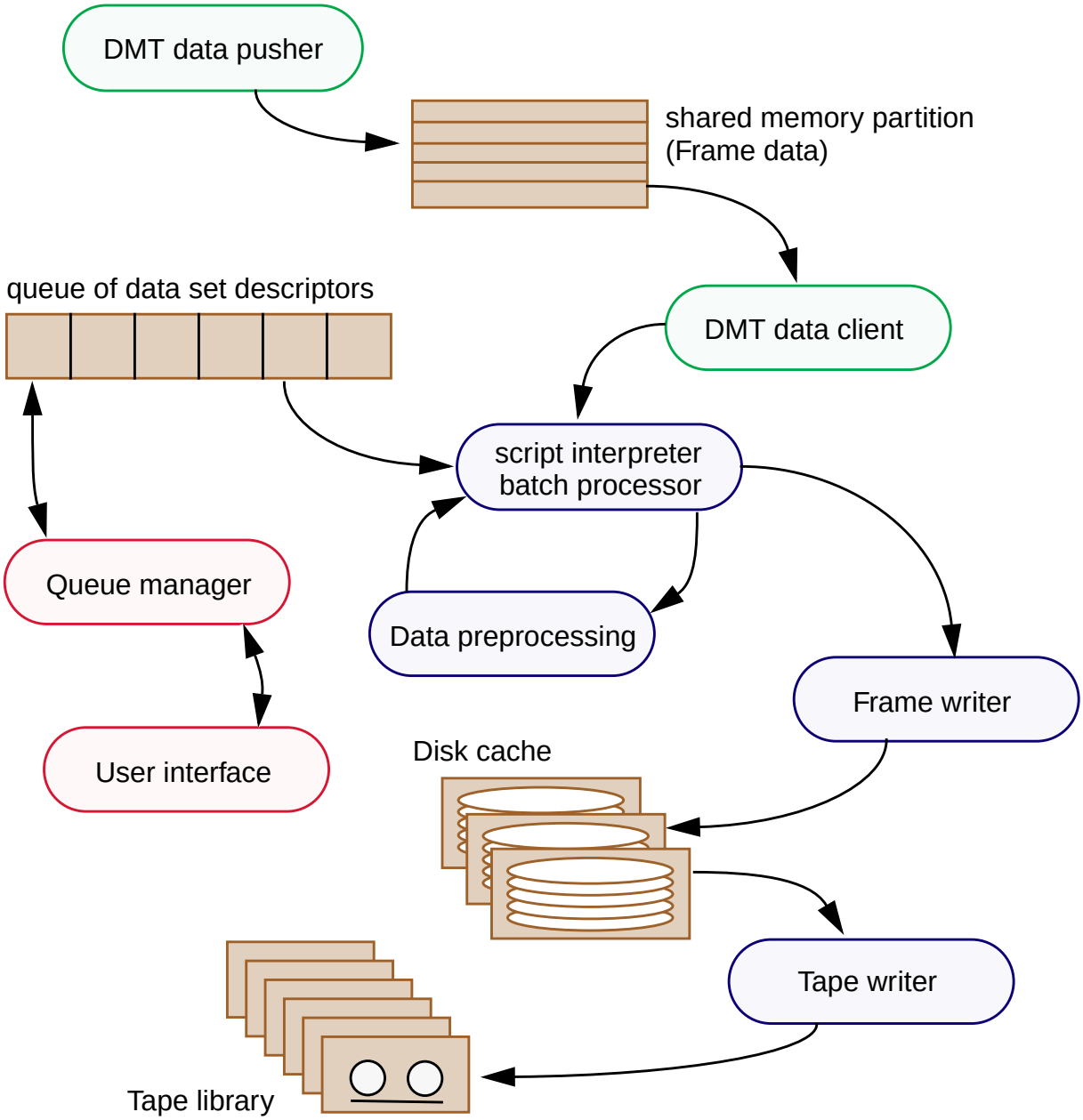

Figure 3: Schematic view of the software layout for the reduced data set generator.