New Folder Name Data Acquisition Requirements

LIGO PROJECT
CALIFORNIA INSTITUTE OF TECHNOLOGY

Internal Technical Note	**CDS_N_005**	**draft**	August 10, 1995
	number	*version*	*date*

Outline of LIGO Data Acquisition Requirements

subject

Stan Whitcomb
Volker Schmidt
author(s)

draft distribution: stan;
final document will be distributed to: W. Althouse, A. Abramovici, J. Camp, D. Durance, S. Kawamura, A. Kuhnert, F. Raab, R. Savage, L. Sievers, R. Spero, K. Thorne, R. Vogt, S. Whitcomb, M. ZuckerP. Fritschel, Y. Hefetz, D. Shoemaker, R. Weiss
distribution

This is the first draft of a new document

This is an internal working note of the LIGO Project. It must not be quoted in publications.

file N'005'draft

Contents

List of Figures

1 Purpose of the document

This is a first attempt to outline the data acquisition requirements of LIGO

The document gives an estimate of data rates for one interferometer

It also sketches some basic concepts for data acquisition and data archiving.

It is mainly based on two older notes written in 1992 by Stan Whitcomb.

2 Estimated LIGO Data Rates

The rate at which an interferometer generates data is a key design parameter for its Control and Data System. The following discussion is intended to give a strawman scenario for the handling of the LIGO data and order-of-magnitude estimates for the LIGO data rates. It is essential to keep in mind that the design of the CDS must be reached through a detailed study of the interferometer and the detection strategies.

The handling and recording of gravitational wave detector data is dictated by the unique nature of the interferometers and their data. Each interferometer has a single main output, a voltage which varies as a function of time and represents the difference in arm lengths as measured by the interferometer. This signal will be sampled (at a rate related to the expected signal frequencies) to produce a continuous time series. This time series will be filtered using a set of templates designed to enhance the signal from predicted gravitational wave sources, and it will be examined to look for unknown waveforms. Coincidences among all three interferometers will be used to discriminate against spurious signals. Since much of this analysis will be carried out "off-line", and since the arrival times of the gravitational wave signals are both unpredictable and rare, this portion of the data stream must be recorded on a continuous basis.

The output of the interferometer may be sensitive to a large number of phenomena other than gravitational waves. (Obvious examples are seismic events, laser frequency noise, or fluctuations in the interferometer alignment.) The current plan is to monitor as many of these parameters as possible to help identify and veto events attributable to such non-gravitational-wave sources. The full data stream from these monitors would be extremely large if we were to attempt to record all of the possibly significant signals. Fortunately, most of these signals are expected to consist of random noise with occasional non-gaussian events which may be significant for the interferometer. An example would be a seismometer, which would normally produce a very low level of noise, with occasional large spikes (earthquakes); it would probably not be important to record the continuous noise, but would be important to record periods of time around the spikes.

Thus, we have divided the total data set into two categories:

- the main interferometer output and a small number of auxiliary signal which are required to understand the state of the interferometer or for frequent vetoes of spurious events; these must be recorded on a continuous basis.
- those signals which are of interest only when an anomalous event occurs; this much larger set of signals will only be recorded in response to a trigger generated somewhere within the data stream itself.

Approximate numbers of channels, bandwidths and resolutions for these two classes of signals are given below for a single interferometer.

2.1 Continuously Recorded Signals The continuously recorded signals include the main interferometer output and those auxiliary signals which are identified to be necessary either to understand the interferometer performance (an example might be the dc offset in a servo loop) or which are found to be of greatest use in vetoing spurious events (an example might be the feedback signal for locking one of the auxiliary path lengths in an interferometer).

Table 1 gives an estimate of the number of signals which must be recorded on a continuous basis. These signals will be archived for off-line data analysis.

Number of channels	Sample Rate	Number of Bits	Data Rate (kwords/s)
4	20 kHz	16	
50	200 Hz	16	
200	2 Hz	8	
Total: 254			Total: 90.8

Table 1 : Continuously Recorded Signals

The data rates of table 1 correspond to a daily net data rate of 7.845120 Gwords/day per interferometer.

2.2 Intermittently Recorded Signals Table 2 indicates the number of signals for each interferometer subsystem which must be monitored by the data acquisition system at each local CDS interface computer (as shown in the overall CDS block diagram). These signals would be digitized and examined by the local computer to search for anomalies (for example, average value exceeding some threshold and indicating a need for adjustment, or a pulse which might appear in a veto signal indicating a false input to the interferometer). If an anomaly is detected, a trigger might be broadcast to the entire data system resulting in the saving of a larger number of data channels for some period of time (perhaps 1–10 seconds) around the anomaly; therefore all of the following signals should be buffered in the local CDS interface (computer) for a time long enough to permit saving of the larger data set for delayed triggers. The thresholds for triggering such a period of extended data storage would have to be adjusted so that these periods would not increase the overall data storage too much (perhaps by a factor of 2).

Subsystem	Number of Signals	Sample Rate	Bits	Data Rate (kwords/s)
Sensitive Component Suspension (6 places)	10 x 6	20 kHz	16	
	5 x 6	2 kHz	16	
	15 x 6	200 Hz	16	
	20 x 6	2 Hz	8	
Subtotal	*300*			*1278.24*
Prestabilized Laser	8	20 kHz	16	
	5	2 kHz	16	
	10	200 Hz	16	
	20	2 Hz	8	
Subtotal	*43*			*172.04*
Mode Cleaner	20	20 kHz	16	
	10	2 kHz	16	
	20	200 Hz	16	
	30	2 Hz	8	
Subtotal	*80*			*424.06*

Table 2 : Intermittently Recorded Signals (Continued) . . .

Subsystem	Number of Signals	Sample Rate	Bits	Data Rate (kwords/s)
Injection Optics	10	20 kHz	16	
	5	2 kHz	16	
	15	200 Hz	16	
	20	2 Hz	8	
Subtotal	*50*			*213.04*
Interferometer Read-out	30	20 kHz	16	
	15	200 Hz	16	
	20	2 Hz	8	
Subtotal	*65*			*603.04*
Auto- Alignment	15	200 Hz	16	
	20	2 Hz	8	
Subtotal	*35*			*3.04*
Auxiliary Monitors	30	2 kHz[1]	16	
	10	200 Hz[2]	16	
	200	2 Hz[3]	8	
Subtotal				*62.04*
Housekeeping	20	2 kHz	16	
	50	200 Hz	16	
	300	2 Hz	8	
Subtotal				*51.6*
total # of channels	*1203*	*maximum total data rate*		*2807.10*

Table 2 : Intermittently Recorded Signals

The ratio between continuously recorded signals and intermittently recorded signals in terms of bandwidth is
90.8 : 2807.1 = 1 : 31

---

[1] 20 kHz in an earlier version!

[2] 2 kHz in an earlier version!

[3] 200Hz in an earlier version!

2.3 State Commands In addition to the data acquisition task, the CDS must also control the operating state of the interferometer. This entails determination of which parts of the system are active, determination of gains and filter states of servo-loops, and so on. Estimates of the number of such parameters which need to be controlled are rather uncertain, but are probably in the range of 500–5000. None of these controls are expected to be time critical.

3 Triggers

The storage of a body of data in response to some trigger (an event in one or more of the interferometer data channels) is one of the major tasks of the data acquisition system.

3.1 Trigger classes Triggers fall in one of two classes: the main trigger signal which is derived by looking at the main gravitational wave signal, and instrumental triggers which are derived from a variety of other sources. Instrumental triggers would be derived from at least three types of sources: internal signals from the interferometer (such as the feedback voltage used to orient a test mass, or fast fluctuations in the laser intensity), external detectors of possible noise sources (an example would be seismic motion detectors), and one or more data-independent triggers (these would be used to occasionally sample the operating state of the interferometer and might be generated by a random number generator).

One important distinction between the two classes of triggers is where they are generated. The main gravitational wave signal is always brought to the central computer and recorded. Thus its trigger signal would probably be generated there. On the other hand, most of the signals used to generate the instrumental triggers would not normally be transmitted to the central computer; these signals would be available only locally and any triggers from them would have to be generated locally and transmitted to the central computer over our network.

A second distinction is in the priority assigned to the different types of triggers. The main gravitational wave trigger must have absolute priority. Gravitational wave events are likely to be rare (at least in the beginning) and each candidate must be saved. The instrumental triggers have a lower priority. Their purpose is to help us understand the operation of the interferometer, to help diagnose and eliminate noise sources and to build a database of instrumental "glitches" against which candidate gravitational wave detections can compared. It seems likely that we will want to have more than one priority for the instrumental triggers and that the priorities will be adjusted from time to time as we learn more about the interferometer's performance. For example, if we find that the interferometer is relatively insensitive to seismic noise but more susceptible to magnetic noise than expected, we would probably want to adjust the trigger priorities accordingly.

The main trigger will probably have to be generated by a digital data analysis. This trigger will be generated from analysis of the main gravitational wave signal. In this case there will probably be a set of filters operating to optimize the signal-to-noise ratio for different types of expected gravitational wave signals. In some cases, the filter will actually act on several seconds worth of data. A further advantage of using a digital filter for this particular signal is that the filters are expected to evolve as we learn more about the sources.

Many of the instrumental trigger signals probably could be generated by hardware circuits. Simple triggers, which might include thresholds, DC levels, or simple pulse filters, can certainly be implemented in hardware. These circuits would have to have adjustable thresholds, with anywhere from one to four parameters (levels, time constants, etc.) which themselves should be under computer control.

It is not clear at this time whether it would be better to generate the instrumental trigger signals in hardware or software. The answer might well depend on the availability of suitable commercial circuits or modules.

3.2 Trigger Generation

3.2.1 Main gravitational wave signal will be analyzed a large number of ways to give trigger signals. One could imagine having a bank of DSPs operating in parallel on this signal to pass it through 100 or more different filters, each optimized for a different type of wave. This one signal is clearly a unique case, which will receive special attention.

3.2.2 The instrumental triggers are intended to provide additional data on the performance of the interferometer. These signals will probably require only a few basic types of simple trigger generators, e.g.:

- DC level
- Short pulse
- RMS AC level

3.2.3 Trigger Generation

The slow channels (the ones with a sample rate ≤ 2 Hz) do not produce recording requests (triggers). This class includes the majority of the channels, but only a small part of the data. Most of these are monitors of the state of the system (temperature, pressure, etc.).

The fast signals include most of the internal signals within the interferometer and the external monitoring system for vetoing external disturbances. It can be expected that it could be important to get trigger (recording request) signals from 50-100% of these signals. Flexibility in the system design is crucial, since we will probably want to change the pattern of both the instrumental and main triggers as we accumulate some working experience with the interferometer.

3.3 Trigger Rates The rate of triggers in any particular channel can be treated as a design parameter, to a large extent under our control. By raising the threshold for that trigger, the rate can be reduced to arbitrarily low levels; the penalty for doing so is to lower the sensitivity of the interferometer or to lower the amount of data available to analyze to improve our understanding of the interferometer.

The rate of triggers from the gravitational wave signal would most likely be set by the maximum rate which is tolerable for a threefold coincidence detection. This rate depends on the correlated event rate of the two interferometers at the initial LIGO site. It can be expected to be up to 10 per hour.

One way to set the rate of instrumentally triggered data (averaged over some reasonably long period of time) is to set it equal to the rate of continuously recorded data, i.e. about 90 kBytes per second. Thus, if we recorded all data in a 10 s burst we would do that no more often than once every 300 seconds on average. There is nothing fundamental about this choice; it just seems reasonable that we would want to have some samples of detailed interferometer performance covering 5–10%[4] of its operating time, with an emphasis on "interesting" times.

As for dead time, one should plan that the data buffers should allow a minimum dead time between a pair of main gravitational wave triggers. We would be willing to lose some instrumental triggers, using some priority ranking, but would be very reluctant to accept the loss of any potential gravitational wave signals. Provided the average rate of data triggered by gravitational wave triggers is within the capability of the network and archiving system, ensuring that the memory buffers for the data are large enough to accommodate multiple triggers should be a relatively small cost.

---

[4] this is more than the 1:30 rate suggested earlier! (to be discussed)

4 Instrument Diagnostics

In addition to the data recording mechanisms described so far LIGO requires additional remote instrument diagnostics. This is composed of two types of components:

- fast analogue output from CDS
- fast recording

"Fast" in this context means most likely up to 1Msamples/s. The system has to have the capability to send arbitrary test signals to selectable output channels to the interferometer, and to record data from a selectable set of data channels. In addition this system should be able to read also data from all normal instrument channels (mainly the 20 kHz main and auxiliary data channels). The functionality one wants is a combination of a programmable function generator and fast, remote storage scope.

The number of channels is not yet known. This part of the system has to be very flexible.

Basically the system has to provide:

- the interactive selection of the output channels to be activated
- the interactive creation of test waveforms
- the loading of "canned" waveforms
- the interlock with machine operation (avoiding of dangerous operation)
- the interactive selection of the channel(s) to be recorded
- the interactive setting of the trigger conditions (time or signal) and the time bases
- selection of sampling frequencies and intervals
-

Instrument diagnostic data are normally not be recorded in the general data archives. They will be handled separately.

5 Data Recording

The simplest system would be to store all of the data every time any one of the triggers is fired, but this is not necessary if it creates a problem. Certain signals would (for example the all-important gravitational wave trigger) would probably result in storing all of the data; other triggers might only require a smaller selection of channels.

More complex schemes are also possible, and may be desirable. For example, we could have a channel for which we store a small set of data most of the time, but periodically (every tenth time?) a larger set is stored. In another case, we might want to store all data if a high threshold is exceeded, but a smaller set if a lower threshold is passed.

It is quite likely (almost certain) that the data storage scheme will change from the initial concept, as we gain experience with it. It is very important to maintain flexibility.

6 Data Archives

Recorded data are placed in a data archives. Each data record is identified by

- signal name
- start date and time
- sampling frequency
- duration of the time window of the data record (or end date and time)

It is assumed that the sampling frequency for any time window is constant. If a signal is sampled with different frequencies, the signal is stored as several records with each record having constant sampling frequency.

Data can be retrieved from the archives by specifying:

- signal name
- start date and time
- duration of the time window of interest (or end date and time)

The data retrieval is successful if the requested signal was recorded at any time within the specified time interval. The retrieval is unsuccessful if the signal was not recorded during the time interval. The successful retrieval can return more than one record for the same signal.

7 Technical Implementation of Data Acquisition

A uniform approach should be applied to the acquisition of continuously and intermittently acquired data.

The basic principle of the data acquisition scheme is the following:

- data is continuously acquired at the local level
- data is archived upon request only

7.1 Local Data Acquisition Data is stored locally (i.e. near the AD converter) in a circulating buffer which has the following characteristics:

- the sampling frequency must be set to the maximum for the channel in question
- the buffer length must be at least sufficient for data to be taken for the maximum interval and the maximum sampling frequency for the channel in question
- precision time information must be stored together with the data

7.2 Data Archival Data is archived on request: a request for data archiving consists of the following information

- signal name
- start date and time
- sampling frequency
- duration of the time window of the data record (or end date and time)

Data archiving requests can be generated by different sources:

- time triggered
- asynchronously triggered

Time triggered data archiving covers three categories:

- continuous archiving (in this case the time interval between successive triggers must be not longer than the duration of the acquisition window)
- data taking at regular times (in this case the trigger is produced at regular intervals)
- data taking at random times (in this case the trigger is generated at random time intervals)

Asynchronously triggered data taking is typically initiated by the detection of specific patterns in some incoming data stream.

Archiving requests can be generated by any computer within CDS. The request is dispatched by the message broadcast system

7.3 Hardware Requirements The scheme described above requires double data buffers at the local level in order to simultaneously archive and acquire data from continuous signals (this is the worst case). The switching between memories at the local level should be as transparent as possible.

8 Technical Implementation of Instrument Diagnostics

8.1 Hardware requirements

8.1.1 Function Generator modules The output functions are produced by programmable function generator modules. These consist of a combination of on-board memory and one or more digital-to-analogue converter channels. The function to be generated is pre-loaded into the on-board memory. The waveform production is then triggered by an external (hardware) trigger signal. The time base for the waveform generation is to be tied to the overall synchronization signals produced by the central timing system.

8.1.2 Transient recorder modules The signal inputs from the instrument are recorded by transient recorder modules. These consist of a combination of analogue-to-digital converter channels and on-board memory. The analogue signal is converted and stored into the on-board memory when triggered by an external (hardware) trigger signal. The time base used for the recording is to be tied to the overall synchronization signals produced by the timing system.

8.1.3 Multiplexers The instrument diagnostic system requires to types of remotely controlled multiplexers: Input and output multiplexers.

Output[5] multiplexers have one analogue input and n analogue outputs. The routing is controlled by the (digital) routing control information.

Input multiplexers have n analogue inputs and one analogue output. The routing is controlled by the (digital) routing control information.

The technical details of the multiplexers have still t.b.d.

8.1.4 Cabling The instrument diagnostics sketched here require a careful planning of the cable connections. All possible signal candidates have to be wired to input multiplexers; a sufficient number of multiplexers has to be installed permanently.

Signal cables have to be installed to all possible points for receiving fast waveform signals. Care has to be taken not to interfere with the signals used in normal operation of the instrument.

---

[5] output relative to CDS

8.2 Software The instrument diagnostics software does not need to be fully integrated with the "normal" data acquisition software. Its use is nearly exclusively interactive; its results will not be put into the general data archives. On the other hand does the data acquisition software need to have some interactive features as well. The result may well be that the two are identical.