



Science, injection, and data quality segments for Virgo

Marie Anne Bizouard[†], Leone Bosi[‡], Shourov Chatterji[◊], Gary Hemming^{*}, Nicolas Leroy[†], Didier Verkindt[‡]

[‡] LAPP - CNRS/IN2P3 - Université de Savoie, Annecy, France,

^{*} EGO, S Stefano a Macerata - Cascina, Italy

[†] LAL - CNRS/IN2P3 - Université Paris Sud, Orsay, France

[‡] INFN, Perugia, Italy,

[◊] INFN, Roma, Italy

VIR-XXX-08

Issue: 1

Date : January 14, 2008

Abstract: This note gives the definition and some information on the Virgo detector status and data quality segments which are stored in the Virgo Database (VDB).

VIRGO Collaboration

EGO - Via E. Amaldi - I-56021 S. Stefano a Macerata, Cascina (Pisa)

Secretariat: Telephone (39) 050 752 511 - FAX (39) 050 752 550 - e-mail perus@ego-gw.it

Contents

1	Overview	2
2	Definitions	2
2.1	Science Mode	2
2.2	Injection Mode	4
2.3	Locked Mode	4
2.4	Data Quality	5
2.4.1	Definition	5
2.4.2	Naming convention	5
2.4.3	Categories	6
3	Database	6
4	Online Lists	7
5	Data Quality upload procedure and validation	7
6	Exchange of DQ flags with LSC	7

1 Overview

The purpose of this document is to standardize the definition, naming conventions, and storage of science mode segments, injection segments, and data quality flags for use in Virgo data analysis.

The definition of science mode segments will be static and unchangeable. Any corrections to science mode will instead be handled as data quality segment. Unlike science mode segments, data quality segments can change, which will be handled by having different versions. New versions are self contained and completely replace previous versions. The official location of data quality segments will be in the Virgo DataBase (VDB). Quality flags will have a standard naming convention. Injection segments will be handled the same as data quality.

2 Definitions

This section provides definitions for the "Science Mode", "Injection" and "Data Quality" segments. For any of such segments, the start time is an integer GPS time included in the segment and the stop time is an integer GPS time excluded from the segment, as shown in Figure 1.

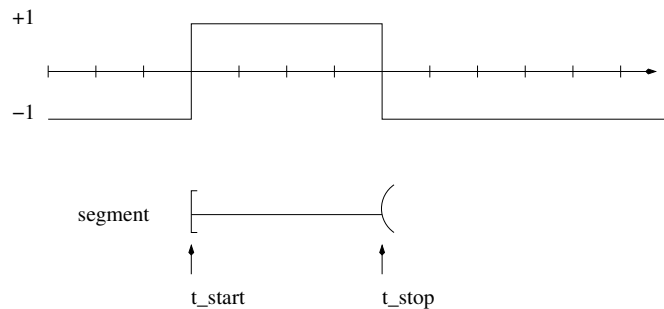


Figure 1: The definition of the segment's boundaries.

2.1 Science Mode

Science mode segments are the baseline segment list for all data analysis.

We desire to recover as reliably and accurately as possible the science mode segments as marked by the operator in the Virgo control room at the time the data was taken.

Even if the raw frame files are the primary reference, we define science mode segments by the value of `Qc_Moni_ScienceMode` as recorded in the trend data frame files, that proved to be derived online from the raw data, to be reliable and to be faster to read. A value of +1 corresponds to science mode data, and a value of -1 corresponds to non science mode data.

A channel is provided online by Automation server: `AlpMain_LOCK_STEP_STATUS`. Its value corresponds to the current step of the interferometer's lock. If, no delay problem is encountered by the AlpMain server or the QcMoni server, the channel `Qc_Moni_ScienceMode` goes to -1 as soon as `AlpMain_LOCK_STEP_STATUS` is not 12, and delay is always below 1 second with respect to the ITF unlock. When the operator enters adjusting mode, the value -1 is put in the `Qc_Moni_ScienceMode` at this time, that means with a delay of about 4 seconds (DAQ latency). To take into account such delays a data quality flag will be created that excludes the last seconds of Science Mode segments.

In the event of missing frame data or just missing `Qc_Moni_ScienceMode` data, the science mode status is defined as the last available science mode status. The only exception is if the detector leaves science mode during the period of missing data. In this case, the science mode segment ends at the beginning of the period of missing data (see Figure 2).

In some cases, missing data segments could be long enough that the detector could leave and returns to science mode during a single instance of missing data. Thus, we should manually double check the science mode status

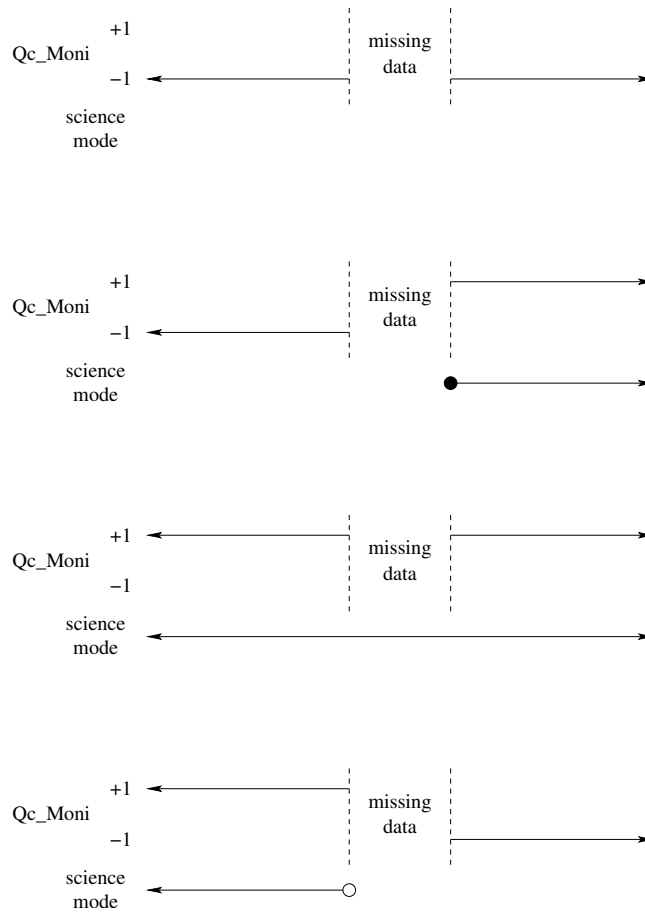


Figure 2: The definition of science mode segments in the case of missing Qc_Moni data.

during periods of missing data. If it is found that data is incorrectly marked as science mode, this will be handled by a specific data quality flag.

More generally, the definition of science mode is static, so there is no version number. Modifications to the definition of science mode should instead be made with data quality flags (see Section 2.4).

Science mode segments (and other information from the trend data frame files) can be ASCII files produced for instance by the package *SeqList*. The segments written in the ASCII file are then inserted into the Virgo DataBase by using the package *VDB* (see section 4).

Adopted definition of the ScienceMode and DQ flags in the VDB:

```
science mode: +1    not science mode: 0    otherwise: -1
bad quality:  +1    good quality:      0    otherwise: -1
```

The following information should be stored for science mode segments (duration of segments is computed by VDB):

1. run tag (WSR9, VSR1, etc.)
2. segment number (1, 2, 3, etc.)
3. GPS start time (inclusive)
4. GPS stop time (exclusive)
5. state (-1, 0, +1)

When output as an ASCII text list, science mode segments lists should consist of the following whitespace separate columns:

1. segment number (e.g. 1, 2, 3, ...)
2. GPS start time
3. GPS stop time
4. duration (optional)

Science mode segment number are not stored, but are generated by the Virgo DataBase at the time segment lists are written.

2.2 Injection Mode

Injection mode segments identify periods of time when simulated burst or inspiral signals were physically injected into the detector.

A flag is produced by one of the Alp server and stored in the raw and trend data frame files: `Alp_Cali_EventInj_STATUS`. It is equal to +1 during injection and -1 otherwise. The missing data policy will be similar to the one defined for science mode segments.

Note that injection mode segments and science mode segments are not mutually exclusive.

Injection mode segments will be specified in the same way as data quality flags (see Section 2.4).

The list of injection details, including the time, SNR, and other parameters of each injection will also be maintained in the database, but independent of the segments considered here.

2.3 Locked Mode

It may also be useful to define segments corresponding to when the detector is locked. For example, these could be useful segments for running online analysis in order to support commissioning activities, particularly between science runs. These segments may also be interesting to analyze when following-up interesting candidates identified with other detectors, potentially including externally triggered events such as gamma-ray bursts or supernovae.

Using the same approach as for "Science Mode" segments, we can use the channel `Alp_Main_LOCK_STEP_STATUS`. Corresponding flag in Virgo DataBase will be +1 when `Alp_Main_LOCK_STEP_STATUS = 12` and 0 otherwise. It will be -1 when the channel is not available.

It has been discovered that, in some cases, due to AlpMain server problems, the interferometer is not locked while *Alp_Main_LOCK_STEP_STATUS* = 12. A specific quality flag will be produced, based on the light power in the recycling cavity Pr_B5_DC channel, to tag those events instead of modifying the lock segments defined by *Alp_Main_LOCK_STEP_STATUS*.

2.4 Data Quality

2.4.1 Definition

The purpose of data quality segments are to account for periods of poor data quality or known problems that affect science mode segments. The list below gives a few examples of the types of segments that should be marked by data quality flags.

- incorrect science mode
- unmarked injections
- missing data
- bad h reconstruction
- the last N seconds of a lock segment
- airplane or helicopter overflight
- power glitch

Data quality segments are used in conjunction with science mode segments to further restrict the segment list. Data quality segments only remove time from science mode. there is no way to add time to science mode. For this reason, a redefinition of science mode in order to recover additional data should only be pursued in exceptional circumstances when there is a strong mistake in the segments listed in the database.

Each data quality flag may have three possible states.

1. The data has been tested and the flag does not apply state = 0
2. The data has been tested and the flag does apply state = +1
3. The data has not been tested state = -1

The third state is represented by a value of -1 (flag not computed), and is the default state of a data quality flag.

The LSC uses 0 and 1 to indicate that the flag should not or should be applied, and does not provide a value when the data has not yet been tested. To minimize difficulties in sharing segments, any segments list produced by Virgo will not include the -1 values.

It may be useful to define multiple thresholds of some data quality flags. For example *WIND_OVER_30KPH*, *WIND_OVER_40KPH*, *WIND_OVER_50KPH*, etc... Similarly, there could be different frequency bands. These will be implemented as separate flags.

The newest version of a data quality flag supercedes all previous versions. By default, the database will return the most recent version unless a user explicitly requests an earlier version.

2.4.2 Naming convention

Run tag will be in a separate table instead of being included in the flag's name. The VDB interface can make this transparent. The DQ flag name inserted in VDB have all capital letters and “_” should be preferred to “-”. Example: *SSFS_CORR*. In case of DQ flags which depend on the h(t) processing (for instance *HREC_MISS*

to flag h(t) missing periods), the flag name must contain a reference to the h(t) processing version number. Example: HREC_MISS_VH1 for h(t) v1 stream. There is no other restriction except that it must be unique and as much as possible meaningful.

When output as an ASCII text file, data quality segment lists should consist of the following whitespace separate columns:

1. data quality flag name
2. run tag (e.g. VSR1)
3. version number (e.g. 1, 2, ...)
4. segment GPS start time (inclusive)
5. segment GPS stop time (exclusive)
6. state (e.g. -1, 0, +1)

Data quality GPS times are integer numbers.

2.4.3 Categories

Depending of the type of usage and the pertinence (mainly efficiency and dead time) of a data quality flag, it belongs to a given category:

- **Category I** : flag should be applied on data segments before events search processing
- **Category II** : flag is applied on the triggers. It corresponds to obvious malfunctioning of the detector and the deadtime remains small (few %).
- **Category III**: flag is applied on the triggers. The deatime may be large and the interest of the DQ may depend on the analysis.

Generally, the use of the DQ flags depends on the type of analysis. Thus, a category I flag for an analysis can be a category II flag for an other. However each DQ flag uploaded in the data base has a "recommended" category, determined according to the studies performed by the DQ group.

3 Database

The Virgo Data Base (VDB) is the official repository of science mode, injection mode, and data quality segments.

A web interface to the VDB is available at <http://virgo.pg.infn.it/VDB>. The following general conventions apply to all segment lists:

- Start times are rounded down to the nearest integer GPS second.
- Stop times are rounded up to the nearest integer GPS seconds.
- Segments are inclusive of the specified start time (see Figure 1).
- Segments are exclusive of the specified stop time.

4 Online Lists

For online analysis, it is desirable to have science and injection mode segment lists available with the minimum possible latency in a form that is easily parsed by automated scripts.

Those lists will be ASCII files posted in a location that is publicly accessible using utilities such as wget and curl. This location (or a second one) must be easily accessible from computing nodes on the EGO, IN2P3, CNAF, and LSC computing centers. This information should follow the definitions provided in the VDB documentation.

A possible solution is to the following: Using the process MdTrend or cronjobs using rsync, the trend data frame files will be read every N minutes by the SegList process which will produce text files "science_list.txt" and "injection_list.txt". Those files will be read by VDBtoolkit every N minutes to update online the Virgo DataBase. Currently trend data frame files are produced every 30 minutes. This implies $N \leq 30$.

For each list, the current segment will be tagged "in progress" if it is a +1 segment.

5 Data Quality upload procedure and validation

Any user will be allowed to upload a quality flag segments list in the Virgo Data Base. The person proposing the quality flag will be responsible for demonstrating its usefulness. The threshold for usefulness will be set high enough to prevent overwhelming the database, but low enough to error on the side of caution. To validate this quality flag, the following procedure will be followed:

1. Any new quality flag is presented on a weekly teleconference to discuss data quality issues.
2. A short description about the detector problem the DQ flag tackles has to be provided. This description should also explain how the DQ is produced (channel, threshold, ...)
3. If there is sufficient interest, then the quality flag segments list is uploaded into VDB. This list is tagged as "unstable".
4. Deadtime is computed with respect to ScienceMode segments
5. Efficiency is then studied with respect to some analysis triggers (MBTA, Qonline, etc...). Some figures of merit such as efficiency, use percentage should be provided.
6. Sanity checks on the segments list will be done by the validator. If the proposed DQ flag is identical to an already defined DQ, it will be abandoned.
7. A category will be proposed depending on the problem and the dead time. The recommended category is defined in the file *VSR1_dqinfo.html* that is saved in CVS.
8. When all cross-checks are performed, the validator will decide after a while to move the DQ flag to the "stable" state and attribute a version number (run tag + version number).

6 Exchange of DQ flags with LSC