BEAM

Blacknest Data System (BDS)

BDS TapeDigitiser Import– 2.0.0 – 2012-06-18

Table of Contents

1. Introduction

The BdsImportTapeDigitiserData program is one of a suit of command line programs provided to allow the import of seismic data to the BDS system. The BdsImportTapeDigitiserData program allows import of TapeDigitiser data and is a client of the BdsServer and connects through the DataAddAccess API in order to import the requested data.

2. Tape Data

The original data comes from 24 track analogue magnetic tapes where the data is encoded using analogue FM techniques which a separate channel per track. There are two heads A and B each with 12 channels. The heads are offset by 1 track width to form the 24 track recording and playback. The tapes are recorded with different FM centre frequencies and differing recording tape speeds depending on the Seismic array and time period.

The 21 data channels contain data from different stations/channels. Sometimes these get swapped around from the standard settings.

2.1. Error Tracks

There are normally two error tracks, often on channels 6 and 18. These are recorded with just the FM carrier frequency. They are used to reduce the effects of tape speed fluctuations on recording and playback machines. As all the signals are FM encoded tape speed fluctuations result in amplitude fluctuations of the actual seismic signals after FM decoding. So to reduce this effect the demodulated FM error signals are subtracted from each of their respective heads data channels demodulated FM values. This helps

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk

significantly reduce the effects of tape speed fluctuation, however it does not handle the signals phase variation.

2.2. Time code tracks

Normally track 24 contains a time code signal. Sometimes this is on track 12 and sometimes there are more than one track containing a time code. Most tapes use a VELA time code format while some use a Hutchins time code format. There are other formats as well. These are FM modulated as per the seismic data tracks and the error tracks are used to reduce errors.

Time code tracks can also be inverted, especially Hutchins time code tracks.

2.2.1. VELA Time Code Tracks

VELA time code tracks contain a stream of pulse width modulated pulses, one per second. The time code is formatted as a frame of 60 time code pulses and defines the time, to the minute, of the first pulse of the time code frame. They always contain the year day number, the hour and the minute. Some tapes also have the least significant digit of the year (0 to 9) stored in the frame.

There are some issues with VELA timecode tracks:

1. They are subject to the tape degradation like the rest of the data tracks which result in low and fluctuating FM signal levels often disappearing altogether and occasional with large noise blocks.
2. As track 24 is at the tape edge it can be damaged quite often by the tape drive mechanics. This can result in severe loss of signal for long periods of time.
3. Noise on the track can result in bit errors leading to incorrect time stamps being recovered.
4. There are issues at the start of a new year. The VELA encoder has to be manually reset at the start of a year. Sometimes this was not done for a few days after the end of the year. In these cases a year day beyond the end of a year is often seen. There are other isses that can occur at this point as well.
5. The year or at least the first 3 digits of the year have to be obtained from tape labels or other information.

2.2.2. Hutchins Time Code Tracks

Hutchins time code tracks are similar to VELA time code tracks and were used before the VELA time code was introduced. They also contain a stream of pulse width modulated pulses, one per second. The time code is formatted as a frame of 60 time code pulses and defines the time, to the minute, of the first pulse of the time code frame. They only contain the time in hours and minutes. They have the same issues as the VELA time code tracks but require the setting of the time to the day from tape labels or other sources as no year day is stored.

3. Tape Digitisation

The analogue tapes are digitised by the TapeDigitiser system. This is documented at:
https://portal.beam.ltd.uk/support/blacknest/files/tapeDigitiser

The TapeDigitiser system employs digital signal processing techniques, on a multiple processor system, to

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk

decode the FM data and decode the time code tracks. It also provides for user provided and system generated MetaData and Quality control and overall management of the digitisation process.

The tape playback machine can be set to run at different speeds to quicken digitisation time.

For each tapes digitisation it generates a directory containing a set of binary data files, one for each contiguous digitisation (data-*.bs) and an ASCII job information file jobInfo.tdi. Normally just one binary data file is present but sometimes, due to tape head cleaning or other issues, multiple data files for different segments of the tape are present. These will often overlap in time periods.

The data files contain a header containing information on the digitisation followed by a set of fixed size and sample number blocks of data. Each block contains a header with a checksum and some MetaData including the start and end time stamps and the FM signal levels averaged out over the block.

The job information file jobInfo.tdi contains overall information on the job, information from tape labels and statistics including tape quality measures. As the TapeDigiser was enhanced over the years extra fields have been added to the jobInfo.tdi file.

The TapeDigitiser employs a DSP based time code decoder. This operates on the VELA track and in later versions the Hutchins track as well. Due to the real-time requirements of the TapeDigitiser the time code decoder, although sophisticated, is relatively simple in its operation. It decodes the time code track searching for valid time code frames. It validates each time code bit and the decoded time code frame in comparison with previous time code times. If there is a mismatch the time code is ignored. Each data block is marked with its start time and end time based on a sample interpolated value from the last valid time code decoded.

Due to tape speed fluctuations and missing time code data the blocks time stamps can jitter and sometimes jump. Each block contains roughly 12 seconds of data. The time code decoder assumes that time “jitters” of less than 2 seconds constitute contiguous data and anything above is a jump in time. Jumps in time code be due to a tape head cleaning or other issue during recording. Note that there can be jumps backward in time. If the time code track goes bad and the tape slows down for a while, the the time code decoder which is working on sample based interpolation can be ahead of the real tapes time. In this case when more than two consecutive valid time code stamps are received, there may be a backwards jump in time. Also the original time code generator will have its time synchronised occasionally.

In newer versions of the TapeDigitiser, the parameter timecodeType is set to “Vela”, “Hutchins” or “Manual” in the jobInfo.tdi file to indicate the type of timecode processing used.

Note that at all times the TapeDigitiser keeps all of the data. So the timestamps can be regenerated later and indeed the error channel processing can be re-applied if required. The resulting data files are effectively sampled at 100 Hz, quite a lot higher than necessary, allowing for future data processing with good accuracy.

4. TapeView

The tape digitisation process is carried out by non-seismic trained personnel. The TapeView program allows the TapeDigitiser data to be viewed and checked. It also offers additional processing.

1. It allows the MetaData stored in the jobInfo.tdi file to be modified. A backup of the jobInfo.tdi file is always made when this occurs to keep the original information intact.
2. It allows digitisation sessions to be marked as to be ignored if there content is not useful.

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk

3. It allows the time code channel to be re-processed by a better DSP phase locked loop based time decoder. The data files time-stamps are updated based on the results of this.
4. It allows a section of a digitisation to be marked with a start and an end cursor based on sample number. This limits the time code reprocessing to just this region of the tape and is used to restrict the data import to this section of data.

The time code preprocessing in TapeView uses a PLL based system. This synchronises to the 1 second time code pulses and offers better time code accuracy even when the time code track is bad. The system may not be able to successively decode a time code frame, but it is likely to be able to synchronise to the 1 second time code pulses. Thus it can maintain an interpolated time based on the time code pulses for quite a long time with high accuracy following tape speed fluctuations.

TapeView allows the time code type to be set to VelaStd, VelaPll, Hutchins and Manual. In Manual mode it reprocesses the time based on the sample rate. This can be used when no useful time code track is present. Its accuracy will be poor, but sufficient to store data to the correct hour or or better in the day. The mode, VelaStd, leaves the timecode track alone so that the original TapeDigitiser processed timecode track time stamps are used. VelaPll re-processes the timecode track using the newer PLL based system as does the Hutchins mode. All of the time code re-processing requires the start time of the data to be entered. This is normally take from the tape label MetaData stored in the jobInfo.tdi file by the TapeDigitisation operator. It can be manually entered though for re-processing if required.

TapeView also allows the time code track to be inverted prior to processing and also has a flag “keepGoing”. The KeepGoing flag tells the time code preprocessor to keep going even if 4 hours have passed since the last valid time code.

The TapeView program stores its time code reprocessing information in the jobInfo.tdi file after backing up the old one. This includes the time code type, start time and cursors etc. It also calculates tape quality figures for the time code re-processing and the data based on FM signal levels and stores this information in the jobInfo.tdi file.

TapeView does not modify the TapeDigitiser data files only the jobInfo.tdi file.

5. BDS Import

The TapeDigitiser data files are imported into the BDS using the BdsImportTapeDigitiserData command line program. This program parses the jobInfo.tdi file for information and imports the data from the data files. All of the data and MetaData is imported. The BdsImportTapeDigitiserData program uses the information added by the TapeView program to ignore sessions and re-process the time-code etc. The following information is used:

- Sessions marked as bad are ignored.
- Data session start and end cursors are used to define the periods for time code re-processing and import.
- The time code type, start time, invert and keepGoing parameters are used for another time-code reprocessing run unless the timecode type is VelaStd.

All of these are over-ridable by command line flags.

Note that the BdsImportTapeDigitiserData will import all of the data within the start and end cursors for

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk

each session not marked as ignored. It does not do anything special with some of the TapeDigitisation MetaData. For example tracks marked as bad in the TapeDigitiser will be imported and no Notes set. However the information is stored in the BDS Sensor data MetaData and can be viewed.

6. Usage

The BdsImportTapeDigitiserData programs can be run by any user that has permissions to add data to the BDS system. The BdsImportTapeDigitiserData program accepts the following command line options:

Core Options	
-help	Help on command line parameters
-verbose	Be verbose with comments
-host <hostname>	BDS Server host name
-user <user:password>	The BDS user id and password
-network	Network organisation
-array	The Array the set of data is from, if an array.
-source	The data source (TapeDigitiser)
-channels <station1:chan1,station2:chan2>	The list of station/channel names for each data channel. This is needed for data formats that are not self describing such as TapeDigitiser.
-description <description>	Set the file description
-addWarning <warning>	Add an importWarningUser to the notes system and data file meta data. Multiple addWarning's can be used. Format is: "errorno,errorstring,fileName,startTime,endTime,network:station:channel:source,description". Note that the network:station:channel:source field will be filled in automatically (using the array as the station and a null channel) if this parameter is null.
-warnings	Display warnings
-dryRun	Perform all data validation tests but don't actually import the data
Validation Options	
-ignoreMissingBlocks	Allow import although validation fails with missing blocks of data. Missing blocks of data are where there is a gap between the end time stamp of one block and the start time stamp of the next block.
-ignoreTimeBackwards	Allow import although validation fails with blocks going backwards in time.

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk

-ignoreFilenameTime	Allow import although validation fails with file name times that do not match block timestamps
-ignoreMetaData	Allow import although validation fails where Database meta-data does not match files meta-data or is not present
-ignoreCorruptions	Fix block corruptions (short/long or otherwise corrupted blocks) by ignoring them (BDRS, WRA-40 and WRA-64)
-ignoreAll	Allow import although validation fails
-ignoreSessions <0,1,2,3,...>	For BdsImportTapeDigitiserDataTapeDigitiserData only. Provides a list of TapeDigitiser “sessions” to be ignored during import.
-includeSessions <0,1,2,3,...>	For BdsImportTapeDigitiserDataTapeDigitiserData only. Provides a list of TapeDigitiser “sessions” to be imported during import. All others, or those stated in -ignoreSessions will be ignored.
-ignoreVelaReprocessingErrors	For BdsImportTapeDigitiserDataTapeDigitiserData only. If the VELA reprocessing on a session fails then import the original session file and save a warning.
-numBlocks <numBlocks>	Limit the number of blocks to import (for testing)
-doNotReprocess	Do not reprocess the TapeDigitiser data with bdsTapeDigitiserFile
-timecode <type>	The timecode type (Manual, Vela, Hutchins). Default is Vela.
-timecodeChannel <n>	The timecode channel (1 – 24). Default is to use channel specified in the jobInfo.tdi file.
-timecodeStartTime <time>	The timecodes start time. Default is to use channel specified in the jobInfo.tdi file.
-timecodeInvert	Invert the timecode

The BdsImportTapeDigitiserData program will read the BDS_HOST environment variable at start-up. This variable, if set, defines the default BdsServer host name to contact. The default is “localhost” if this is not set.

The data is always input as synchronously sampled data.

The user will need to supply the directory name where the TapeDigitiser data is stored, the network, the array, the source and list of station/channel names for each of the files data channels.

The BdsImportTapeDigitiserData programs will first validate the data. The following validations will be performed:

- Generally check for data file corruption. This looks at block headers, makes sure that time stamps

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk

are in time order and in range.

- Check that the blocks time-stamps are continuous. If they are not there are missing blocks of data. This error can be ignored.
- Check that the sample rate is correct.
- Check that there is meta-data information in the BDS system for the channels being imported. This error can be ignored.

The "-ignore*" flags of the BdsImportTapeDigitiserData programs allows certain validation failures to be allowed. If the validation passes the data is uploaded into the BDS system. If it fails an error message is reported to the user.

There are some example scripts of various Data Imports in /usr/bds/import directory. This code is quite rudimentary.

6.1. Recommended import options

The recommended standard import options for Blackest TapeDigitiser data are:

"-ignoreMissingBlocks -ignoreTimeBackwards"

These will handle general import without losing any import data.

7. Errors and Warnings

Each of the import programs perform a validation of the data prior to import. If there are any failures the errors are printed on stderr. The import programs have the ability to "fix" certain errors and to ignore certain errors if the appropriate flags (-ignore...) are used. When an import issue is ignored or fixed a warning is issued instead of an error. This is added to the BDS Notes system and also stored the the BDS sensor data files MetaData. Normally no warnings are printed on stderr unless the -warnings option is used.

The warnings are stored in the BDS Notes database system and in the seismic data file;s MetaData. They are listed in the "importWarning[0-9]*" and "importWarningUser[0-9]*" info fields (The [0-9]* is a number). You can add to the "importWarningUser[0-9]*" fields using the -addWarning flag which is useful to add any information based on pre-processing done to the data files.

The system also adds the file pathnames of the imported files to the "importFile[0-9]*" fields in the BDS seismic data file.

The syntax of the user added warnings is a comma separated list of parameters. These are checked to validity. The syntax is as follows:

ErrorNumber, ErrorString, FileName, StartTime, EndTime, Network:Station:Channel:Source, Decsription

When these are output on stderr, they are prepended with either "Error, " or "Warning, " as appropriate. Note that the Network:Station:Channel:Source field will be filled in automatically (using the array as the station and a null channel) if this parameter is null.

8. Validation Errors

The current Validation Warning/Errors include:

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk

Number	*Error/Warning*	*Description*
19	ErrorValidateReorder	If the -reorder flag is specified then data blocks will be re-ordered into time-stamp order. A warning will be issued which describes the re-ordering performed in the form of an ASCII string. The format of this string is a list of block ranges and the re-ordered position. For example "0-10:30" means blocks 0 through 10 have been moved to position 30.
18	ErrorValidateDuplicate	If any duplicate blocks are detected (blocks havingan identical time span) then this error is issued. The -deleteDuplicate blocks option will delete any duplicate blocks making sure that the blocks are truely identical in data as well as time span.
13	ErrorValidateMissingBlocks	Each blocks start and end times are compared. If there is a gap in time where there is no data then the Missing Blocks Error is issued.
14	ErrorValidateTimeBackwards	If a blocks start time is before the previous blocks end time then this error is issued.
15	ErrorValidateFilenameTime	The date/time encoded in the filename does not match the time stamp of the first block of data.
16	ErrorValidateMetaData	There is in-complete MetaData for the data in the MetaData database on the BdsServer
17	ErrorValidateFix	If the -ignoreCorruptions flag is used an the system “fixes” corruptions then this warning is issued for every block that has been fixed.
27	ErrorDataQuality	Data quality issues such as bad time code tracks etc.

9. Data Quality Measures

The TapeDigitiser measures and stores various data quality statistics and user settings. All of this information is available in the MetaData stored in the BDS Sensor data files. Some of the more important data is presented in the BDS Notes system during import. The main quality measures stored this way include:

QualitySignalLevel	The overall FM signal level. Low levels indicate dirty tape heads or a bad tape. Data will have noise on it and possibly glitches.
QualitySignalLevelVela	The FM signal level on the time code track. If low this indicates a creased tape edge, dirty tape heads or a bad tape. Time code decoding will be suspect and can thus have a degree of jitter.

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk

QualitySpeedVariation	The tape speed fluctuation. If high this will affect the phase of data signals and general signal quality. Time codes are more likely to have jitter.
QualityVela	The fraction of time codes that have been successfully decoded. Low levels below 70% indicate that the time codes could have jitter and jumps.

If values are the above are outside set levels then appropriate warning Notes will be added to the BDS system for the data.

10. Return Value

The programs will return a status value of 0 if all was Ok. They will return a non zero value on error together with a message output on stderr. BdsError numbers are listed in the BDS user manual.

11. Further Information

For further information please look at the BDS system documentation at:
https://portal.beam.ltd.uk/support/blacknest.

Directors: Dr T.K. Barnaby, M.J. Thomas
Registered Number: 2415065
Registered office: Beam Ltd, Northavon Business Centre, Dean Road, Yate, Bristol BS37 5NH, UK.
VAT Number: 520-2521-05
Web: www.beam.ltd.uk
Email: info@beam.ltd.uk