ACRES'
INTERFACE
CONTROL DOCUMENT
ICD #1

ACRES & TERSS
Archive Data Format (ADF / TERSS), Version 2.2

9 December 1999

ACKNOWLEDGMENTS:

- Stephen Clift, CSIRO, for having conceived and designed the tape format
- Stuart Swan & David Edwards, CSIRO, for having written the first version of the tape format description

CHANGE RECORD

Issue	Date	Pages	Description of Change	Release
0.9	?	All	Initial 'TERSS' format document	
1.0	20/12/1999	All	Included MSS and IRS data Included sub-sample file	

TABLE OF CONTENTS

INTRODUCTION

This document describes a tape format used by ACRES Data Acquisition Facility (DAF) in Alice Springs and the Tasmanian Earth Resources Satellite Station (TERSS) in Hobart for raw data storage of satellite data.

Although the format is generic enough to be used with a variety of tape drive technologies, the ADF implementation is currently using Digital Linear Tapes (DLTs). This format description is therefore explicitly discussing (and may therefore appear limited to) the DLT. More specifically, the tape drives used at the DAF and TERSS ground stations are DLT7000s.

PURPOSE OF THE DOCUMENT

This is an Interface Control Document (ICD). The intent is to produce a description detailed enough for any developer of a system or software module to make use of and interface in an optimum way to the device/module/format subject of this description.

DATA RECORDING

The formatting of the data on tape is done by the "Archiving" function at the ground station. This is the process of writing the data onto tape. It takes place after the satellite data has been recorded onto a disk array in real time, and will always commence after the pass has ended. Since the tape drive usually has a much lower data rate capability, (than what corresponds to real time data rate) the process of writing the dataset to tape takes up to several times longer than was the duration of the pass.

TAPE FILES

The archive tapes produced are digital 0.5 in cassettes written on Quantum DLT7000 tape drives, with a nominal data capacity of 35 Gbyte per tape. Individual satellite passes can produce datasets varying in size from several hundred Mbytes up to about 10 Gbytes. The archiving software tries to make efficient use of the tape capacity, with a few constraints. As many complete datasets as will fit may be archived on a single tape. Datasets that are too large to fit on a single tape start on a new tape (they are never appended to a tape with data already on it) and extend over as many complete tapes as necessary to produce a multi-volume dataset. Unused space on the last tape of a multi-volume dataset may be used to store any complete datasets that will fit.

A satellite dataset comprises a number of files. The main file contains the satellite data; the others contain header information or system log data. Each tape record in the satellite data file comprises a header containing time and status information, followed by a block of satellite data.

The sequence of satellite data on tape is the same as in the downlink. In the case the downlink format needs to be reconstituted, the bytes shall be output in the order they appear in the tape record and the most significant bit of each byte shall be output first.

The tape format is shown in a block diagram (Figure 1).

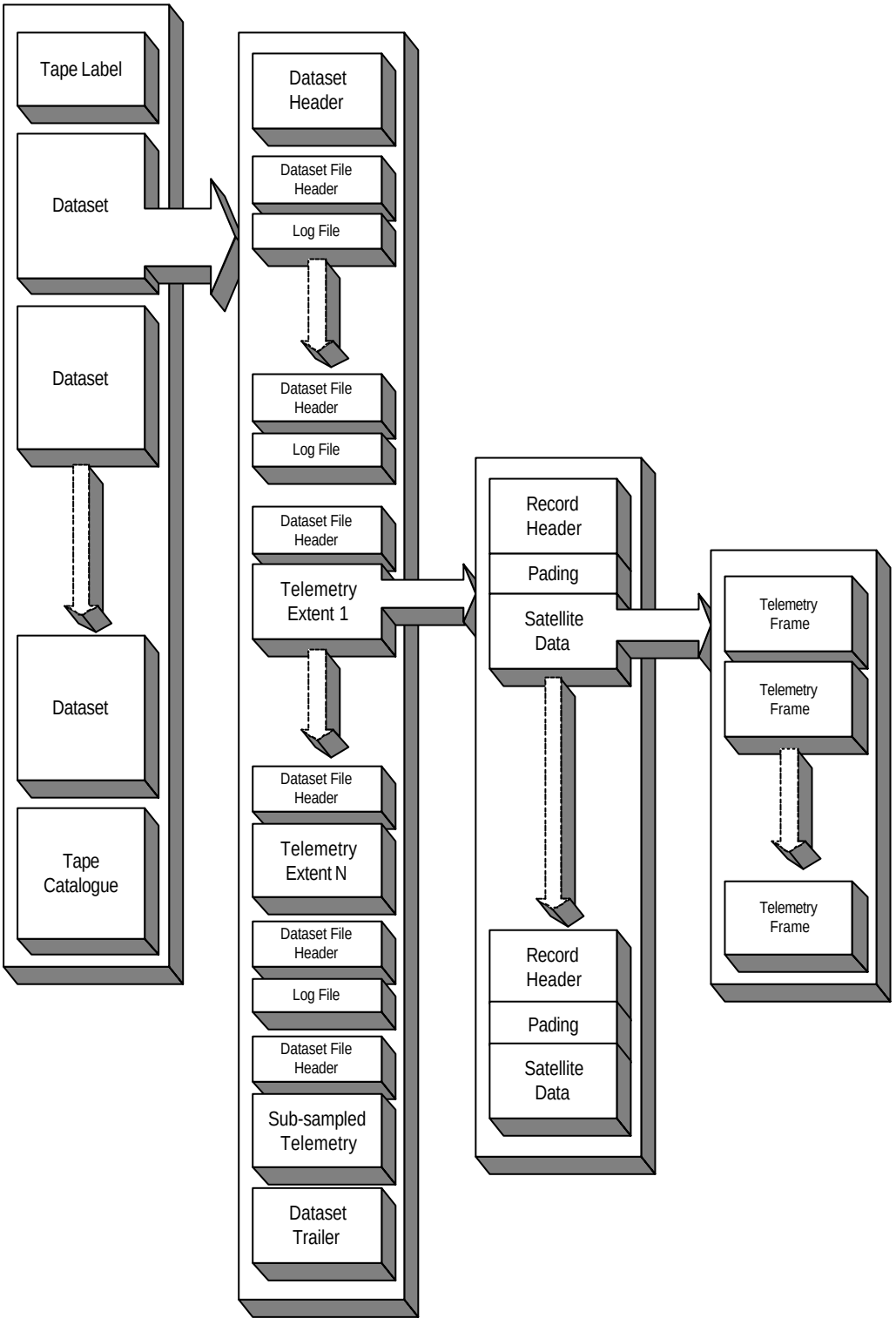

Figure 1 Tape Format Block Diagram

TAPE FORMAT

The archive format of each tape shall consist of one tape label, one or more pieces of a dataset, followed by one tape catalogue. The format of the header files is designed to be flexible and easy for humans to read. Each header file contains a number of records, zero-padded out to (typically) 32kb. A record is a linefeed (0x0A) terminated line. The maximum length of a record will be 512 characters. No non-printing ASCII characters, including the CR character (0xD), will occur within a record. The first line of each header file describes the type of the header file and commences with the character sequence "< TERSS RMS ". Each of the subsequent lines is of the format

"<Identifier>: <Attribute>"

Line lengths, column layouts, etc. will not be assumed. The set of identifiers for each type of header file may change with new revisions of the TERSS system software. Identifiers do not contain embedded colon characters. Identifiers are case sensitive.

Tape Label

The tape label is a 4096 byte file. Each record is ASCII text, must start in column 1 and must be space filled. The first record is the string "< TERSS RMS TAPE LABEL >". This is followed by multiple records that describe the tape.

Revision:	The format revision of the tape
Tape Format:	The format revision of the tape
Tape Name:	The identifier for this tape.
Created:	ISO time stamp of when this tape label was written.
Site:	Where the tape label was written.
User:	Owner of the process that wrote the tape.
Machine:	The architecture of the machine.
OS:	The operating system and version number.
Tape Type:	The type of tape (may be "unknown")
Drive Type:	The type of drive that created the tape label.
Capacity:	The nominal capacity of this tape

Below is an actual example from a tape:

```
< TERSS RMS TAPE LABEL >
Revision:    2.1
Tape Format: 2.1
Tape Name:   SL0001
Created:     1999-05-24T05:38:10
Site:        ALICE
User:        root (Super-User)
Machine:     IP27
OS:          IRIX64 6.5
Tape Type:   Unknown
Drive Type:  QUANTUM DLT7000 2255
Capacity:    0.0
```

Tape Catalogue

The tape catalogue is the last file on the tape. It is a 32768 byte file, which consists of a header line containing the string "< TERSS RMS TAPE CATALOG >" and a revision statement as in the tape label. This is followed by multiple dataset descriptors, which describe the datasets or partial datasets on the tape. Each dataset descriptor consists of 3 records. Each record is of the same format as the tape label and has the following identifiers.

`Dataset Number:`	The dataset number on this tape (starts at 1).
`Dataset Identifier:`	A dataset identifier (pass name).
`Dataset Files:`	The number of files associated with this dataset.

The Dataset Identifier is a space-separated string of the form

pass_id sat_id orbit_no AOS

where *pass_id* is a unique pass descriptor
sat_id identifies the type of satellite
orbit_no is the orbit number
AOS is the AOS date/time in ISO format
Eg. Dataset Identifier: FRED ERS-1 25082 1994-09-29T13:24:52.397

Datasets

Each dataset or partial dataset starts with a dataset header. This is followed by sequence of file sets, each comprising two tape files. The first file of each pair is a header file that identifies the second file. The second file contains either acquisition log information or satellite telemetry. Several Gbytes of telemetry data are typically acquired for each satellite pass; this is broken into conveniently sized "extents" (files) on tape (currently about 500MB each). Telemetry file records are less than or equal to 64K in length. The dataset concludes with a dataset trailer.

Dataset Header

The format of the dataset header is similar to the tape label header and consists of a header line followed by multiple records. The header line contains the string "< TERSS RMS DATASET HEADER >". The dataset header has a block size of 32768 bytes. The first record in a dataset header contains a revision statement as in the tape label. This denotes the format of the dataset. The identifiers, which are required for this revision, are as follows.

`Created:`	ISO creation date
`Dataset Number:`	The dataset number of this tape (starts at 1).
`Pass Identifier:`	The operator-determined identifier for this satellite pass
`Satellite Name:`	Name of the satellite for this pass.
`Orbit Number:`	Orbit number of the satellite.
`Link ID:`	Link identifier for the satellite (eg LR,LBR,HR,HBR)

`Data Type:`	Type of data received from the satellite (eg SAR)
`Scheduled Start:`	ISO scheduled start time of the pass.
`Scheduled Stop:`	ISO scheduled end time of the pass.
`Scheduled Duration:`	Duration in seconds of the pass.
`AOS:`	ISO time of signal acquisition.
`LOS:`	ISO time of signal loss.
`Comments:`	Operator-determined comments
`Bit Rate:`	Bit Rate of satellite in bits per second (floating point number)
`Megabytes Acquired:`	Number of megabytes acquired between AOS and LOS
`Ephemeris:`	Ephemeris information as sent to antenna controller.
`Tape Name:`	The name/ID of this tape.
`Ephemeris:`	Ephemeris information as sent to antenna controller.
`Tapes Allocated:`	Space separated list of tapes originally allocated to archive this satellite pass. Depending on drive compression and data compressibility, some of these tapes may not be used.
`Tape Number:`	The number of this tape in the dataset sequence (1-origin).
`Tape Name:`	The name (identifier) of this tape.
`Tape Record Size:`	Size of tape records in the telemetry data files (bytes)
`Telemetry Frame Size:`	Satellite's basic telemetry (data block) frame size (bytes)
`Extent Size:`	Approximate size of each telemetry data extent (bytes)

Below is an actual example from a tape:

```
< TERSS RMS DATASET HEADER >
Revision:             2.1
Created:              1999-05-24T05:38:42
Dataset Number:       1
Pass Identifier:      SPOT-1.37114
Satellite Name:       SPOT-1
Orbit Number:         37114
Link ID:
Data Type:            SPOT-1 HRV
Scheduled Start:      1999-05-16T00:43:38
Scheduled Stop:       1999-05-16T00:43:54
Scheduled Duration:   16
AOS:                  1999-05-16T00:43:38
LOS:                  1999-05-16T00:43:54
Comments:             ads
```

```
Bit Rate:              49372400
Megabytes Acquired:    145.8
Ephemeris:             -
Tapes Allocated:       1
Tape Number:           1
Tape Name:             SL0001
Tape Record Size:      56320
Telemetry Frame Size:  18564
Extent Size:           500000000
```

The ISO format for dates and times is CCYY-MM-DD[Tt]HH:MM:SS[Ff.]FFF, where [] denotes a choice of several characters, CC is the century, MM is the month, DD is the day of the month, HH is hours, MM is minutes, SS is seconds and FFF is milliseconds. For example,

1994-09-29T13:24:52F397

Dataset File Header

The dataset file header is a 32768-byte file, which contains as its first record "`< TERSS RMS DATASET FILE HEADER >`". As in the dataset header this is followed by a revision statement, which determines the format of the remaining identifiers and attributes. All records are space filled.

`Created:`	ISO date and time when the file was written to tape.
`Pass Identifier:`	Operator-determined identifier for this satellite pass.
`Estimated Size:`	Size of following file in bytes.(approximate for telemetry files)
`Extent Number:`	Zero for log files, 1 upward for satellite data files.
`Data Type:`	Text description of the type of data in the following file.

Below are two actual examples from a tape:

```
< TERSS RMS DATASET FILE HEADER >
Revision:        2.1
Created:         1999-05-24T05:38:45
Pass Identifier: SPOT-1.37114
Estimated Size:  147000000
Extent Number:   1
Data Type:       Telemetry (SPOT-1 HRV)
```

```
< TERSS RMS DATASET FILE HEADER >
Revision:        2.1
Created:         1999-05-24T05:39:21
Pass Identifier: SPOT-1.37114
Estimated Size:  6792
Extent Number:   0
Data Type:       ref
```

```
< TERSS RMS DATASET FILE HEADER >
Revision:        2.1
Created:         1999-05-24T05:39:26
Pass Identifier: SPOT-1.37114
Estimated Size:  1010
Extent Number:   0
Data Type:       card7

< TERSS RMS DATASET FILE HEADER >
Revision:        2.1
Created:         1999-05-24T05:39:31
Pass Identifier: SPOT-1.37114
Estimated Size:  1039224
Extent Number:   0
Data Type:       sub
```

Log File Format

Each log file is written out using 32768 byte blocks. The format of each is specific to the log file however it should be in an ASCII, human readable form. The last record of a log file is zero-padded out to 32kb as necessary.

Telemetry Data File Format

The format of each satellite telemetry file is satellite specific because the notion of a fundamental telemetry frame varies from satellite to satellite. The record size is as specified in the telemetry data record header and should take into account any problems associated with reading the data on other systems.

Note that a design limitation of the format means that for satellites with variable length telemetry frames (ie Landsat TM) there can be no more that one frame per tape record. As TM has such a large (major) telemetry frame it spans several tape-records and thus does not present a problem.

Satellite Data Record Format

Each satellite record contains 200 bytes of header information. This is followed by zero or more padding bytes, followed by the satellite data stream. The first byte of the first satellite frame in the data stream part of the record came from a particular demodulator frame. The demodulator status information in the header is the status for that demodulator frame, and the offsets detailed below are relative to the start of that demodulator frame. The format of the header is as follows:

Satellite Telemetry Data Record Header Format

Byte numbers are inclusive. All numbers are unsigned MSB based unless otherwise stated.

Field	Bytes	Description
1	0-3	'Magic Number' (0xE914AD33)
2	4-5	Revision number of the satellite record header format
3	6-7	Reserved
		The following two fields are only filled in tapes made on the old TERSS system. They are zero filled for contemporary DIS ingested data. See Appendix 1
4	8-9	Demodulator status information format revision
5	10-39	Demodulator status information (Redundant)
6	40-47	Ground station acquisition timestamp data converted to seconds since 1970 in unsigned 64-bit fixed-point format. The first four bytes represent seconds since 1970-01-01T00:00:00 (i.e. Unix time). The following four bytes represent fractional seconds, with an implied decimal point to the left of the MSB.
7	48-51	Offset in bits of the MSB of the first byte of the satellite telemetry data from the MSB of the first byte in the associated DIS frame.
8	52	Reserved.
9	53	The (left) bit shift required to byte align the MSB of the first byte of the satellite telemetry data (the value in bytes 48-51 modulo 8).
10	54-55	Reserved.
11	56-59	Number of satellite frames in a record.
12	60-63	Number of records required to make a satellite frame.
13	64-67	Sequence number of this record in a multi-record satellite frame.
14	68-71	Size of this record in bytes.
15	72-75	Byte offset from beginning of this record to start of satellite data.
16	76-79	Bit mask of frame validity.
17	80-83	Bit error count.
18	84-87	Number of bits tested for errors.
19	88-91	Size of current satellite frame(s) in bytes.
20	92	XOR mask for DTU/ACRES compatibility
21	93	Reserved.
22	94-95	The current extent number in this dataset.
23	96-99	The record count within the current file.
24	100-104	The record count within the current dataset.
25	105-199	Reserved.

An example hex dump from SPOT (line numbers are in hexadecimal notation, each line of hex is followed by its decimal equivalent):

```
dafdis1 129# od -D SPOT-1.40009.ref | head -25
0000000 e914 ad33  0004 0000  0001 5fd7  5e01 0009
        3910446387 0000262144 0000090071 1577123849
0000010 0000 0000  0085 0000  0000 0000  0000 0000
        0000000000 0008716288 0000000000 0000000000
0000020 0000 0000  0000 0000  384b 01c4  0700 7c00
        0000000000 0000000000 0944439748 0117472256
0000030 0000 e177  0002 0000  0000 0003  0000 0000
        0000057719 0000131072 0000000003 0000000000
0000040 0000 0000  0000 dc00  0000 0274  0000 0007
        0000000000 0000056320 0000000628 0000000007
0000050 0000 0000  0000 0000  0000 4884  ff00 0001
        0000000000 0000000000 0000018564 4278190081
0000060 0000 0001  0000 0001  0000 0000  0000 0000
        0000000001 0000000001 0000000000 0000000000
0000070 0000 0000  0000 0000  0000 0000  0000 0000
        0000000000 0000000000 0000000000 0000000000
*
* data starts here | with the spot sync word…
0000270 0000 0000  0030 0f03  30ff 303f  0c33 cfcf
        0000000000 0003149571 0822030399 0204722127
0000280 0f33 3ffc  6c6c ff09  0033 ffc0  f03f fcf3
        0255016956 1819082505 0003407808 4030725363
```

The ‘magic number’ is a unique identifier used to determine if the record is a satellite data record. The revision number details the contents of the header information. Newer revisions will utilise the reserved bytes to add extra information.

The ground station timestamp is the ingest time of the start of the first telemetry frame in the record. For frames that span multiple records this time is repeated in each block that the frame spans.

The number of satellite frames in a record represents the number of satellite frames required to complete a tape record. If a record contains only a partial satellite frame, (ie TM or MSS) this value will be 0. The number of records required to make a satellite frame will be 0 except where a satellite frame is larger than the record size. The sequence number is used when a satellite frames spans more than one record. If one or more satellite frames fit within a record the sequence number will be 0.

The data byte offset indicates the beginning of the satellite data within the current record.

The bit error count indicates the number of bit errors found whilst testing the satellite data in the current tape record or satellite frame, whichever is larger. The total number of bits tested is contained in the bit errors tested counter. Thus the BER (bit error rate) can simply be calculated by dividing field 17 by 18.

Each data byte read from the tape will have to be exclusively ORed (XORed) with the value of the Processing byte found in the Satellite Data Block header before being output as serial data.

The validity bit mask lists the validity of the frames within the record. The least significant bit represents the first satellite frame. 1 represents a valid frame, 0 an invalid frame.

ERS-1 Tape Block Format

Amount of padding = 56 bytes (256 bytes of header information)
Number of frames (formats) within one tape block = 7.
Size of each frame (format) = 7424 bytes.
Size of each block = 51Kb (52224 bytes).

RadarSat-1 Tape Block Format

Amount of padding = 90 bytes (290 bytes of header information)
Number of frames (formats) within one tape block = 202.
Size of each frame (format) = 323 bytes.
Size of each block = 64Kb (65536 bytes).

SPOT Data File Format

Amount of padding = 428 bytes (628 bytes of header information)
Number of frames within one tape block = 3 HRV1 & HRV2 interleaved.
Size of each frame = 9282 x 2 = 18564 bytes.
Size of each block = 55Kb (56320 bytes).

LandSat MSS Data File Format

Landsat MSS scan-line records consist of 184320 6-bit words (1105920 bits) which is equivalent to exactly 138240 (8-bit) bytes. Thus writing one frame across three 46Kb-tape blocks makes the most efficient use of space. As for TM the first two blocks will be filled, the third one containing the required filler.
Amount of padding = 2472 bytes (third frame only).
Number of tape blocks for one frame = 3.
Size of each frame = 138240 bytes.
Size of each block = 46Kb (47104 bytes).

Landsat TM Data File Format

Landsat TM scan-line records vary in length from line to line and the average length increases slowly with time as the TM instrument ages. As of May 1996, the average length is about 767kb, which is well outside the original line-length variation specification
To keep tape records within the nominal 64k limit, complete scan lines are packed into groups of as many 64kbyte tape records as are necessary to hold the data (typically 12). Padding is zero for all records except the last, where the padding is adjusted so the last byte of the scan line is also the last byte of the record.

Landsat-7 ETM Data File Format

Amount of padding = 855 bytes (1056 bytes of header information)
Number of frames within one tape block = 31 CADU's 1& 2 interleaved.
Size of each frame = 1040 x 2 = 2080 bytes.
Size of each block = 64Kb (65536 bytes).
NOTE: Any switches in polarity (I & Q) will not be detected and corrected until the next 'frame' resulting in the offending frame being omitted.

IRS LISS Data File Format

Amount of padding = 781 bytes (981 bytes of header information)
Number of frames within one tape block = 3.
Size of each frame = 21861 7-bit words rounded up to 19129 bytes.
Size of each block = 57Kb (58368 bytes).

IRS PAN Data File Format

The PAN sensor data is split into two 6252 6-bit word frames which are transmitted independently on the I and Q channels.
Amount of padding = 876 bytes (1076 bytes of header information)
Number of frames within one tape block = 6 PAN 1& 2 interleaved.
Size of each frame = 6252 6-bit words x 2 = 9378 bytes.
Size of each block = 56Kb (57344 bytes).

Sub-Sample File

The sub-sample file contains the sub-sampled telemetry data extracted and used for catalogue generation. Thus not all datasets will contain this data.

In general the sub-sample file contains lines of sub-sampled imagery with auxiliary data appended to the start. This file is simply written to DLT with a blocking factor of 62Kb with no regard to it's internal structure.

Currently only the following datasets include a sub-sample file:

- SPOT HRV
- Landsat TM
- Landsat-7 ETM

Dataset Trailer

The dataset trailer is the final file in a dataset and has a similar format to the dataset header. It contains as its first record "`< TERSS RMS DATASET TRAILER >`". As in the dataset header this is followed by an revision statement, which determines the format of the remaining identifiers and attributes.
All records are space filled.

Created:	ISO date and time when the file was written to tape.
Pass Identifier:	Operator-determined identifier for this satellite pass.

Below is an actual example from a tape:

```
< TERSS RMS DATASET TRAILER >
Revision:        3.1
Created:         1999-05-24T05:39:44
Pass Identifier: SPOT-1.37114
```

APPENDIX 1: DEMODULATOR STATUS INFORMATION

Zero filled for data ingested through the DIS
The demodulator used for the original TERSS system inserted its own status information into the satellite data stream. Satellite data was blocked by the demodulator into 4K frames, which contain 16 bytes of demodulator status data. During tape generation the demodulator status data was stripped out of the telemetry stream and placed in the tape record headers.

Space is left in the tape record headers for 15 x 16-bit words of demodulator status information, although only the first 8 are used. Each word is written out with the most significant byte first (big endian). The 8 words are the demodulator "overhead" words. The remaining 7 words are reserved for future use. The format of the demodulator overhead words is as follows (bit 0 is the LSB, bit 15 the MSB):

Word 0 - Sync Word (B7C2 in hex)

Word 1 - Frame Counter (16 bits)

Word 2 - Internal Status/Alarm Information

[0-3]	Buffer/MUX FIFO overflow count.
[4]	Buffer FIFO Full - a 1 indicates a FIFO overflow.
[5]	Signal Lock
[6]	HSI bus status (not particularly meaningful)
[7]	MUX FIFO Full - a 1 indicates a fault in the demodulator interface.
[8]	XILINX Programmable bit (Default 0)
[9]	Timeout (not used in TERSS)
[10]	Spare
[11]	Spare
[12]	Test mode indicator
[13-15]	Demodulator symbol rate indicator

Word 3 - OMEGA Clock Information

[0]	1 msec strobe
[1]	1's of msec
[2]	2's of msec
[3]	4's of msec
[4]	8's of msec
[5]	1's of 10's of msec
[6]	2's of 10's of msec
[7]	4's of 10's of msec
[8]	8's of 10's of msec
[9]	1's of 100's of msec
[10]	2's of 100's of msec
[11]	4's of 100's of msec
[12]	8's of 100's of msec
[13]	1's of seconds
[14]	2's of seconds
[15]	4's of seconds

Word 4 - OMEGA Clock Information

[0]	1 second strobe
[1]	8's of seconds
[2]	1's of 10's of seconds
[3]	2's of 10's of seconds
[4]	4's of 10's of seconds
[5]	1's of mins
[6]	2's of mins
[7]	4's of mins
[8]	8's of mins
[9]	1's of 10's of mins
[10]	2's of 10's of mins
[11]	4's of 10's of mins
[12]	1's of hours
[13]	2's of hours
[14]	4's of hours
[15]	8's of hours

Word 5 - OMEGA Clock Information

[0]	1's of 10's of hours
[1]	2's of 10's of hours
[2]	1's of days
[3]	2's of days
[4]	4's of days
[5]	8's of days
[6]	1's of 10's of days
[7]	2's of 10's of days
[8]	4's of 10's of days
[9]	8's of 10's of days
[10]	1's of 100's of days
[11]	2's of 100's of days
[12]	±10 msec accuracy
[13]	±100 msec accuracy
[14]	±500 msec accuracy
[15]	±1 sec accuracy

Word 6 - Uncommitted Bus

LSB	User input A byte (rear panel connector PD14)
MSB	User input B byte (rear panel connector PD13)

Word 7 - Uncommitted Bus

LSB	User input CA byte (rear panel connector PD16)
MSB	User input D byte (rear panel connector PD15)