N94-23942

LONG TERM TRENDING OF ENGINEERING DATA FOR THE HUBBLE SPACE TELESCOPE

Ross M. Cox

Computer Sciences Corporation
Seabrook, Maryland, USA

A major goal in spacecraft engineering analysis is the detection of component failures before the fact. Trending is the process of monitoring subsystem states to discern unusual behaviors. This involves reducing vast amounts of data about a component or subsystem into a form that helps humans discern underlying patterns and correlations. A long term trending system has been developed for the Hubble Space Telescope. Besides processing the data for 988 distinct telemetry measurements each day, it produces plots of 477 important parameters for the entire 24 hours. Daily updates to the trend files also produce 339 thirty day trend plots each month. The total system combines command procedures to control the execution of the C-based data processing program, user-written FORTRAN routines, and commercial off-the-shelf plotting software. This paper includes a discussion the performance of the trending system and of its limitations.

Keywords: trend, processing, software, analysis, plotting, spacecraft

1. INTRODUCTION

Many spacecraft have been designed with expected life times of 3 to 5 years. Even for such short term missions, both the quantity and quality of personnel available for spacecraft engineering operations decline as the mission ages. Budget constraints, individual career advancements, and contractor changes all influence this decline. As mission lifetimes now extend to decades, the need to transmit knowledge concerning the behavior of hundreds of spacecraft components becomes critical. Trending is a useful tool in this transfer. It allows future subsystem engineers (SEs) to have access to the data their predecessors deemed pertinent to proper operation of the spacecraft. The establishment of valid trends also helps an SE become familiar with the way the spacecraft actually operates on-orbit, which is not always as originally designed.

While most SEs acknowledge the need for trending, each has a different idea of the best way to analyze a spacecraft. One way to differentiate between different types of spacecraft analyses is on the basis of where the processing takes place. In this paper, *Online Analysis* is assumed to occur on the same system that takes in realtime data. The realtime support schedule is presumed to be the sole driver for performance of this analysis. *Offline Analysis* is defined as that which goes on despite the availability of the spacecraft.

The duration of the analysis provides another distinction between different types. *Realtime Analysis*, performed during critical activity periods has a duration measured in seconds or minutes. The term realtime is derived from the operations of low earth orbiters, when available commanding time as limited. The concept can be extended to include geosynchronous operations as applied to critical operations such as orbital maneuvers or high precision pointing operations. The basic purpose of realtime analysis is to verify that all expected events occurred. Nominally for prevention of unexpected or hazardous conditions and to quickly detect anomalies, a major drawback to realtime analysis is the many operations proceeding concurrently. The detection of more subtle problems that arise is difficult.

Near-Realtime Analysis, as the name implies, occurs shortly after a realtime contact. Its time scale is from minutes to hours, and its purpose is usually to investigate further some unexpected state or condition recently observed in one or more realtime contacts. It allows for greater correlation between different data points, since it is carried out in a less hurried environment.

Trending is the aspect of offline analysis used to reduce the quantity of data saved to characterize component or subsystem performance. A big problem encountered in trending concerns the volume of data to be processed. Trending can therefore be broken down into short and long terms. Short term implies data collected for hours or days at a time. Long term reflects data collected for weeks, months, and years. The short term trends can normally process every point received from the spacecraft, whereas data reduction strategies must be applied for long term trends to retain information within data storage limits.

This paper will describe issues encountered in the design and implementation of the Long Term Trending System (LOTTS) used by the Hubble Space Telescope (HST). First an overview of the spacecraft, telemetry, and data flow defines the scope of the task. A brief description of the hardware and software environments is followed by discussion of the internal LOTTS processing. Lastly, restrictions and performance considerations are presented.

2. OVERVIEW

A primary goal in creating LOTTS was to provide the SEs with a useful tool that would help in producing required plots for Mission Operations Contractor (MOC) monthly reports. An underlying theme to the development was to design a system the user could easily modify to meet changing mission needs. Before characterizing the LOTTS, the spacecraft and ground system components with which it interacts will be described.

2.1 Spacecraft Subsystems

The HST spacecraft has seven major subsystems: Data Management, Electrical Power, Instruments and Communications, Optical Telescope Assembly, Science Instrument Command and Data Handling (which has four elements, one for each instrument onboard), Pointing Control, and Safing. Each subsystem imposes unique requirements for trending in terms of the input data rate, frequency of desired output, and complexity of the processing involved. Most inputs to LOTTS come from telemetry.

2.2 Telemetry

There are approximately 7000 different engineering measurements that may be placed into the telemetry stream. Each measurement has a User Friendly Mnemonic (UFM) of up to eight characters. For instance, DTR2MTRC identifies tape recorder 2 motor current. The UFMs remove the need to remember minor frame and word locations. The ground system database handles the conversions from UFM to minor frame/word/bit. The telemetry is transmitted in at least one of seventeen data formats. The data rate is 4 or 32 kilobits per second (kbps). The telemetry matrix major frame is 1200 minor frames by 200 words for 32 kbps and 250 minor frames by 125 words in 4 kbps.

2.3 Data Flow from Spacecraft to LOTTS

Data from the spacecraft, is processed through the Tracking and Data Relay Satellite System (TDRSS). It is received at the Space Telescope Operations Control Center (STOCC) at Goddard Space Flight Center (GSFC) located in Greenbelt, Maryland. For realtime operations, data then flows to the Payload Operations Real Time System (PORTS), which produces realtime history files. The PORTS Applications Software System (PASS) accepts the Near RealTime data (seconds to minutes time lag) from PORTS. After an engineering tape recorder dump, the recorded data, PORTS history files, and the NRT data are merged to form the Astrometry and Engineering Data Products (AEDP).

2.4 Engineering Subsets

PASS creates engineering subset files as part of the AEDP based on the contents of the Telemetry Subset Definition (TSD) file[1]. The subsets are written in changes-only format. This means that data is only written to the file if its value has changed since the previous minor frame. This serves to greatly reduce the data processing. The particular subset used by LOTTS contains nearly 1000 telemetry points selected by the subsystem engineers. These data are converted to engineering units and time tagged.

3. HARDWARE ENVIRONMENT

The LOTTS operates on computer systems residing in the Engineering Support Systems(ESS) room at GSFC. The hardware consists of a dedicated Digital Equipment Corporation (DEC) VAX 3100/30 workstation attached to an A/B switch. This switch allows connection into either the network of other ESS 3100 workstations or a cluster of DEC mini-computers that serve as applications processors (AP5/AP7). These reside in the Data Operations Control Center. With the switch in the "A" position, the workstation has access to the AP5/AP7 cluster. This is the normal daytime configuration. It allows SEs with accounts on AP7 to access the trend files or binary data files directly to perform special analyses. The data base and PASS products reside on the cluster. In the "B" position, the workstation can access an LPS20 laser printer to produce hardcopy.

The specifications on the hardware are: 1)the workstation operates at eight mips, has sixteen megabytes of RAM and 1 million blocks (1 block = 512 bytes) of storage, 2) the DEC cluster computers consist of one model 6510 and one model 8650. The 6510 is a 6 mips machine with 60 megabytes of RAM. The 8650 is a 12 mips machine with 64 megabytes of RAM. These two mini-computers share 31 gigabytes of storage. The LPS20 printer has a 1 megabyte onboard memory, processes 20 pages per minute, and has a turbo chip set to improve graphics processing time.

4. SOFTWARE ENVIRONMENT

Since trending is a routine event, it is well suited to a batch environment, requiring minimal operator intervention. LOTTS operates under the VMS 5.4 operating system. Digital Command Language (DCL) command procedures control the operator interface. The VAXC programming language is used for the programs that read the engineering subsets and update trend files. User programs are written in VAX FORTRAN. The plotting software uses a commercial-off-the-shelf product, IDL, developed by Research Systems, Inc.

5. LOTTS INTERNAL DATA FLOW

The data inputs to the LOTTS are the engineering subsets. AEDP creates them on the AP5/AP7 cluster. Trending begins with data technicians monitoring the creation of these files. PASS produces these files, as possible, throughout the day. When an entire day's files appear, the operator executes procedures that copy the engineering subsets to the trending workstation and remove any old data from the cluster. Next, three jobs start up in a batch queue. The first breaks the subsets into a separate 24 hour file for each telemetry point. The second job creates user-defined derived parameters using the output of the first job. The third job creates daily plots (see Figure 1).

5.1 Creating 24 hour UFM files

Each engineering subset file has the date-time of the first time tag it contains embedded in its name. This file contains variable length records because of the changes-only nature of the data. The first record is a header from which the start time, stop time, and telemetry format of the file are extracted. Each format change causes the current file to close and a new one to open. Every UFM has an associated Measurement Tag Value (MTV), an index into the alphabetical list of UFMs. When writing a subset file, AEDP inserts a minor frame time, followed by an MTV/value pair for each point that has changed in that particular minor frame.

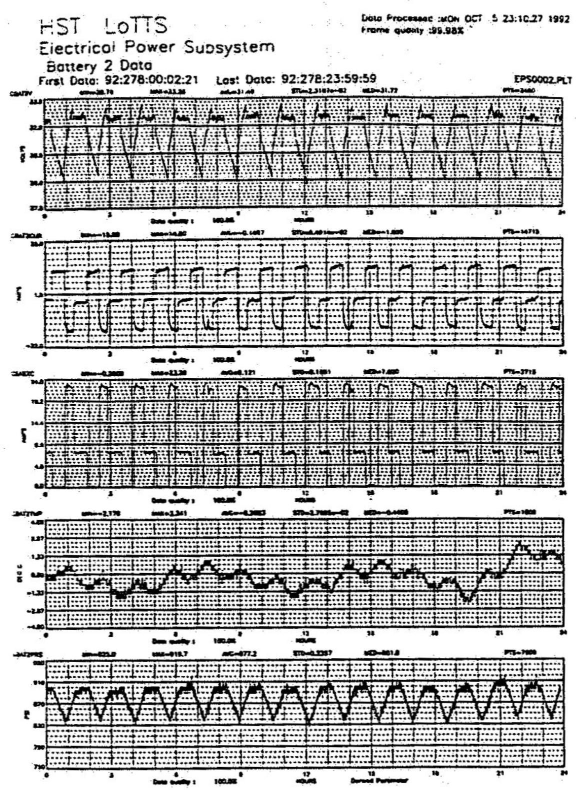

Figure 1 - Typical daily plot produced from LOTTS. It shows, from the top battery voltage, battery current, solar array current, battery temperature and battery pressure for battery two.

The subsets represent large files containing all data for all requested points. The 24 hour UFM process breaks these into many smaller files, each with all data for a single point, over an entire day. These products receive the name of the UFMs for easy identification. They are called bin files, as they are output in a flat binary format. Each bin file has 1536 byte records containing data and time tags. The first record is a header containing status data including the UFM, MTV, data start and stop times, date and time when the data was processed by

LOTTS, number of minor frames read, number of frames with bad quality, the total number of data points read, number of points that failed conversion, data type for this parameter, a descriptor to indicate any special processing of this point (See section 7), and the minimum, maximum and time-based average for the point. This information is later used to help in processing the file or as display data on a daily plot.

Other processing done by the 24 hour UFM process includes marking data dropouts. Each minor frame contains a quality flag. When this flag shows a problem with the data, the program places a large negative value in the file for the given time. This is necessary because the changes-only data produces ambiguity between constant data and no data. The large negative value allows the plotting software to induce a pen-up for dropouts. Note that problem data is not involved with statistics collection.

Error logs from each run show the time locations of data dropouts. These are used by the SE to determine if a data spike was the cause of a data hit or a component anomaly. A final feature of the 24 hour UFM process is the creation of a temporary statistics file. This is used by another program to update the trend files.

5.2 Derived Parameters

The second major processing function is to establish derived parameters. These increase the versatility of the LOTTS by allowing data to be presented in simpler forms. Trending applies to derived parameters just as to normal telemetry. Bin files are created and the temporary statistics file is updated.

Derived parameters are useful for several functions. One form of derived parameters is non-telemetered algebraic combinations of two or more telemetry values. A simple example is the creation of a power parameter using voltage and current telemetry. Derived parameters are also useful when the ground database contains an incorrect calibration curve. Databases should be under the strictest configuration control. In this case, updates will occur only after errors accumulate to the point of warranting a new version. With derived parameters, the SE creates the proper calibration routine until the new curves are implemented. Derived parameters are useful for coordinate transformations and unit conversions as well.

One very important use of derived parameters is made in the treatment of context dependent telemetry, where one telemetry point controls the validity of another, such as a component's telemetry taking on meaningless values when idle. Another example is the SE wanting to see the data only during a part of the orbit such as terminator crossings.

The greatest value of derived parameters in the LOTTS is that the SEs can modify, create or add algorithms at their discretion without extensive programmer intervention. Basic utilities are provided for reading and writing both ASCII and binary files and performing calibrations. There is single subprogram for each subsystem to handle its derived parameters. The SE simply creates a subroutine to perform the desired calculation and modifies the derived parameter subprogram to call it. This leaves LOTTS programmers to handle the more complex issues and prevents errors in one subsystem from propagating to the others.

5.3 Trending

After the subsets are processed and derived parameters are produced, the trending takes place. The subroutine takes the statistics file, generated by the 24 hour UFM and Derived Parameter processes, and updates the trend files. This occurs quickly since the trend files are indexed by year and day of year. The trend data is stored in 80 byte records and includes: minimum, maximum and average for the given day and year.

5.4 Plotting

Next, the daily plots are produced. A template file defines each plotted page. Concatenating many of these templates allows each SE to customize plot requests as the mission requires. Theoretically, there is no limit to the number of plots placed on a page, but the resolution becomes impaired after 8 in portrait mode (6 in landscape). The templates specify the number of plots per page, plot orientation for the page, UFMs of the points to be plotted, number of coarse and fine grids lines, upper and lower bounds on both axes, plotting symbols and connectivity, and limits. The user has the option of placing statistics for a grid just above the top line. These statistics include minimum, maximum, average, standard deviation, and number of points plotted. Pages can be labeled with major and minor titles as well. The user can select automatic scaling based either on the minimum and maximum data values or n-sigma, where n defined by the user. The plotting software provides default values to all unsupplied parameters except the UFM and page numbers. The same templates are used for trend plots. An example is given in Figure 2.

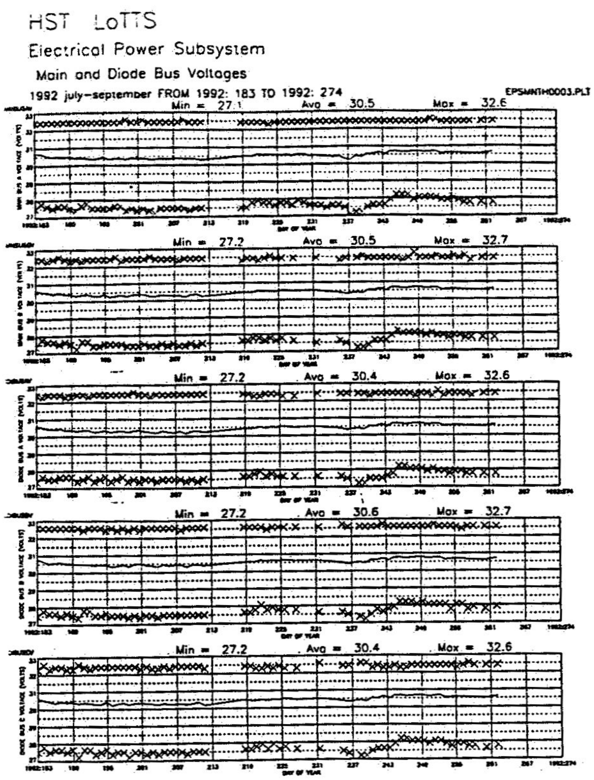

Figure 2 - A sample long term plot. Averages are connected dots, minimums and maximums are shown with unconnected Xs.

6. RESTRICTIONS

It was known from the beginning that certain restrictions would have to be placed on the users of the system. This is because, in general, users have the uncanny ability to consume all resources, in terms of CPU storage, etc., they are allowed. For LOTTS, this meant devising compromises to maximize engineering content while minimizing total processing.

6.1 Number of Telemetry Points Processed

If you ask an SE how much of the available data he or she would like to see, the response is invariably "All of it!". In practice, however, the SE looks at only a portion of the data on a routine basis. For instance, it is not uncommon for data on a backup component to be meaningless when not in use. The backup component is still important, but the trended data is meaningless. The criteria established for allowing mnemonics into LOTTS was to impose a total limit of 100 per subsystem plus 100 for each instrument or 1000 total. The limit relies on the supposition that an engineer, in normal operations, would not use more than 100 points, (approx. 12 pages), for daily or monthly reports. The supposition comes from both past experiences with other spacecraft, and actual monthly report content.

The benefit of imposing a limit is that it forces each SE to seriously consider those points they really need and prevents them from simply asking for everything, much of which they would never use. Note that the limit applies only on the total number of points, not points per subsystem, so one group may be plotting or trending more than 100 points.

An important consideration when imposing limits, is that specific mnemonics must be easily added to or deleted from the total, such as when a primary fails or interest rises in a previously untrended point. LOTTS accommodates this through requesting changes to the PASS controlled telemetry subset definition file.

6.2 Culling Algorithm

Some data are telemetered at 40 hz. If one of these points were monitored every second over 24 hours it would produce over 3.4 million time-value pairs. Plotting such large files can take hours, with no real benefit, since the number of points is greater than the printer's resoultion. The problem is how to throw out data without loosing information content. The solution used in LOTTS, termed culling, is to use a time window of averaging. The user specifies time period over which the minimum, maximum, and average will be found for each offending UFM. The bin files will contain just these three points for each time period. Using this method, plots of over one million points reduce to near eight thousand points, with no visible information loss on the plots. Currently, less than ten points require this treatment, but the plot software is spared the need to plot over 9 million points a day.

6.3 Filtering Scheme

Another problem of using plots to analyze data is what to do with the spikes. They may be an indication of an imminent failure, but most of the time the problem is in the data. Data spikes make automatically scaled plots take on unreasonable ranges and wreck statistical calculations.

LOTTS solves this by removing the spikes from the plotted data, but providing a listing of their times and values. This is called *filtering*. There are two kind of filtering rate and absolute. In rate filtering, the SE defines a delta amount by which a point can change and still be plotted. Any changes greater than the delta will not be written to the plot file. Absolute filtering disallows any values over or under the user-defined range.

7. PERFORMANCE

7.1 Processing Time

The engineering subsets are available within 3 days (usually 1), of realtime. The delay is largely due to waiting for the tape recorder to be dumped. It then takes 2-4 hours to create the bin files and derived parameters. 1-1.5 hours are required to create plot files. And 2-2.5 hours are required before hard copies of the daily plots are in hand. This totals roughly the time of an 8 hour shift. Monthly plot creation takes 2-3 hours.

7.2 Short Term Storage

The normal telemetry data consists a mixture of 32 and 4 kbps data. The designed mixture was 20% 32 kbps and 80% 4 kbps. Due to mission changes since launch, the mixture is now closer to 80% 32 kbps and 20% 4 kbps. The size of engineering subsets for a given day ranges from 45,000 blocks to 120,000 blocks. A rule of thumb to process the data completely is that three times the space used by the subsets will be required. LOTTS programs and permanently stored files require 200,000 blocks. This limits the amount of on-line storage of bin files to about three days[2].

7.3 Long Term Storage

A final performance issue on LOTTS is due to size of the trend files themselves. Currently, these files take up around 67,000 blocks. Storing 1000 sets of statistics, per day, for 15 years will eventually consume over the 1,000,000 allocated to LOTTS. As stated earlier, present operations allow the bin files to remain on the system for three days. Allowing for a doubling of LOTTS and user defined software, 400,000 more blocks will be used up. For these longer term trends it is not necessary to keep each days statistics. The solution is to compress the daily trend files into monthly statistical form. For LOTTS, daily files for the last five years will be stored on-line and older data will be stored by month.

8. CONCLUSIONS

This paper has addressed various aspects of the LOTTS as applied on HST. First, trending is performed as a routine activity that does not interfere with critical operations. Also, the trending and daily products are readily accessible to the subsystem engineers. The design incorporates the ability to change with mission needs. Several methods were presented to show that large volumes of data are successfully reduced, without loss of information content, for both analytical and storage reasons. Long range storage solutions have shown that the LOTTS can last throughout the mission lifetime.

9. ACKNOWLEDGEMENTS

The author wishes to thank the following: Charlie Bengston for all his work on the initial design and prototyping, Kim Hawkins for her many refinements, Betty Colhoun for her critiques and suggestions throughout the development, all the subsystem engineers for their many ideas and the data technicians who make it so.

10. REFERENCES

1. Computer Sciences Corporation, Hubble Space Telescope Astrometry and Engineering Data Processing System to Post Observation Data Processing System Interface Control Document, ST-ICD-46, February, 1986.

2. Computer Science Corporation, Long Term Trending User's Guide, K. Hawkins and B. Colhoun, October 1992.