REAL-TIME TELEMETRY DATA ARCHIVAL AND DISTRIBUTION

Shawn McFarr and Paul Friedman

Loral instrumentation
15378 Avenue of Science
San Diego, CA, 92128-3407

ABSTRACT

High-performance telemetry systems traditionally store prime and processed data on disk drives attached to a host computer. Bandwidth performance of host minicomputer and disk drives limit the amount of data archived to aggregate rates of a few hundred kilobytes per second. Over the years, several approaches have been used to increase performance from pre-recorded analog tape, but real-time storage still required a large host and expensive proprietary parallel disk technology. The advent of distributed architecture system networks divorced the front-end telemetry processor from direct 'DMA' connections to the host. Today's technology moves data storage to the front end for the highest performance and outward to the network for less demanding archival rates.

This paper explores several schemes and implementations for increased digital data archival performance in a distributed architecture Telemetry Ground Station. It goes on to discuss the variety of industry-standard devices and media available for storage at tens of megabytes per second on Redundant Arrays of Inexpensive Disks (RAID) to slower but much less expensive optical and streaming tape drives on both the front end and network computing resources. But storage is half the task; networks serve many users requiring archived data access. The paper will also show how the sophistication of today's modern Graphical User Interface (GUI) eases data distribution for Telemetry Ground Station engineers and analysts.

INTRODUCTION

Vehicles both under test and in operation often acquire data from a variety of sensors, buses and computers for telemetering, storage and eventual evaluation. On the craft, sensors acquire data that may be conditioned and then sampled and encoded (arranged or commutated) with data from tens or even thousands of other sources. The data may form a continuous stream, as in PCM (Pulse Code Modulation); or may be placed into packets, as in CCSDS. This stream is transmitted to the ground for immediate evaluation or stored on board the vehicle. The media for storing telemetered data offers a host of possibilities. Data can be kept in the original encoded analog formats; decommutated, parallelized into words and stored as frames of data; decommutated to the word level, stored as data along with its defining tag; or sorted in real time and stored as files of individual parameters.

Why store preprocessed or unprocessed data? To reduce the chance of losing often irreplaceable information, therefore analog tape recordings of the original data stream are usually made at a point as close as practicable to the source. Unfortunately, retrieving data in this format is time consuming, and the media is bulky, very expensive and prone to deterioration over long periods of storage. Copying the original media for distribution results in loss of fidelity, and the equipment to replay and decommutate data for multiple dispersed users is prohibitively expensive. Hence a desire for a readily accessible less expensive storage system.

HISTORY

Over the last decade, commercial storage devices grew larger and faster. Telemetry stations' real-time computers were able to select only a subset of the data stream for real-time quick-look analysis and storage (Figure 1).

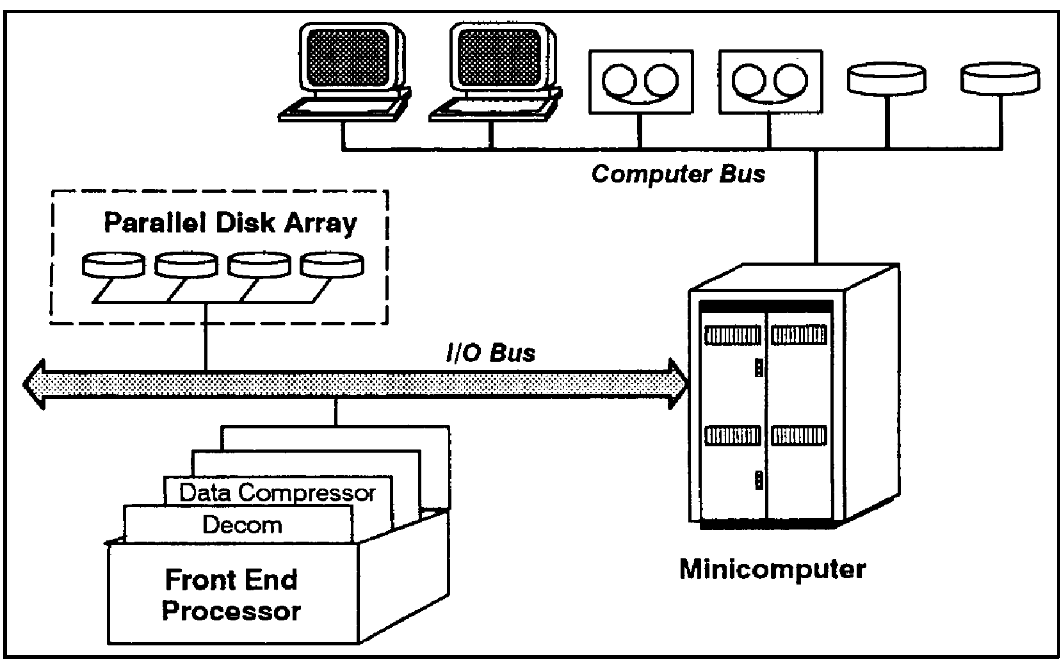

Figure 1. Traditional Host-Based Telemetry Station

The amount that could be stored was severely limited by the then current technology of the device and host computer. Consider storing the data contained in a 1 Mb/s PCM stream after complete decommutation. If the average data word is 10 bits long, the resulting storage rate is 100,000 words/second or 400,000 to 600,000 bytes/second;[Note] the most a super minicomputer could archive to a standard disk drive. Because the capacity of that drive might be 280 MB at most, it filled in 12 minutes. Higher data rates or multiple PCM streams were not practical. Several schemes extended these computer and disk rate limitations. A parallel transfer disk built using multiple heads increased data transfer to a single drive. Techniques employing multiple disks in parallel increased throughput, but at a high cost. Several multi-disk striping scenarios developed (Figure 2), one employing several standard disk controllers on the same computer. Because the computer's bus is several times faster than the drive, each controller's buffer is filled in a round-robin fashion. As the buffer of one controller fills, others are emptying to their respective disk drives. Data integrity is maintained as long as a buffer is written to disk before the next one arrives. Playback is the reverse process.

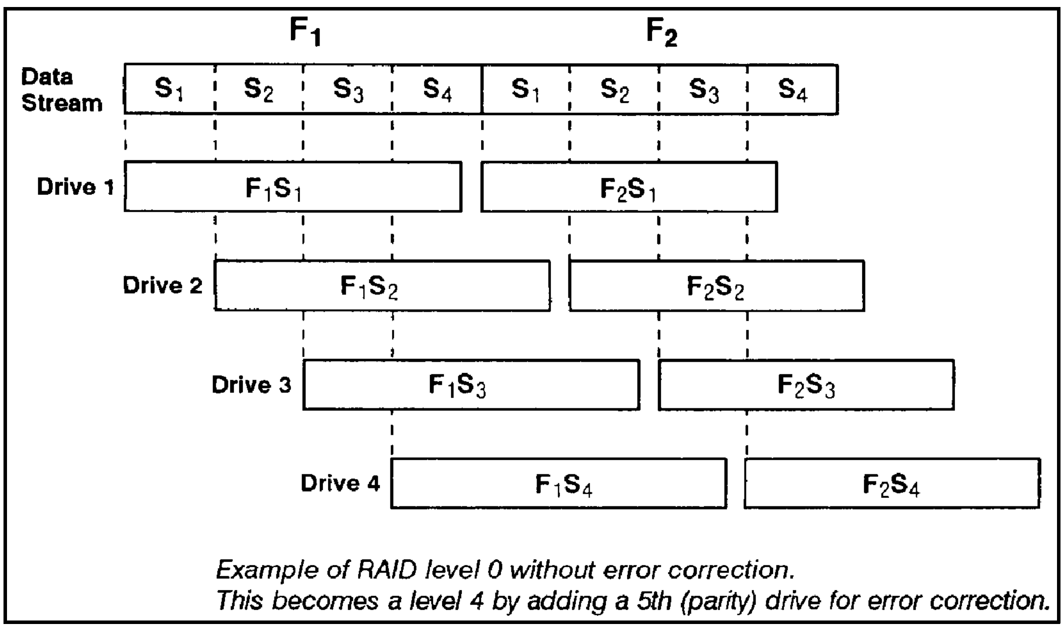

Figure 2. Parallel Disk Array Technology

A separate controller enhance the striping technique. Unfortunately, as the number of parallel drives increased, so did the chances that single drive failure could bring down the entire system. Error correction information was either added via a dedicated drive

Note:
The expansion of data is based on storing a two-byte identification tag with each data work, which is stored as two bytes for fixed format or four bytes for floating point format.

or encoded with each segment stored so that the original data could be reconstructed in the event of a failure. As the aggregate data rate of the telemetry decommutation system and the parallel disks exceeded that of the minicomputer's own bus, a separate front end bus evolved to unburden the minicomputer. Unfortunately, the cost of parallel disks and special front end buses limited deployment.

TIME CORRELATION

Correlating the time of measurement acquisition with the data can consume a large segment of both storage space and bandwidth. One extreme tags each measurement with either a minor time (least significant portion) only placing the entire time record periodically perhaps once per telemetry frame or once per *n*th measurement. A more economic solution for synchronously acquired data is to insert time periodically; for example, at the end of each frame. The time associated with the acquisition of a particular measurement can be interpolated from its position in the telemetry frame. Time tagging of parameters is required if aperiotic data is stored. An example of asynchronous information is processed data such as the output of"deltaslope" algorithms.

Data retrieval from disk offers a challenge because measurements must be continuously metered programmatically on display workstations. Asynchronously stored data complicates playback even further. Perhaps the most accurate—if not the most cumbersome—reproduction method is to replay the analog tape.

STORAGE INTERFACE

The minicomputer, front end I/O bus and parallel disk arrays all employed proprietary architectures, which means that devices were expensive and unique to a vendor. Realizing the added expense of proprietary architectures, computer users forced the industry to develop architectural standards; i.e., what has evolved into today's "open systems." The prevailing mid-performance mass storage bus is the Small Computer System Interface (SCSI). SCSI is available in upward-compatible formats, including: the original 5 MHz 8-bit parallel interface, SCSI II FAST at twice the speed, and a double-wide version (SCSI II F/W) capable of a transfer rate of 20 MB/s. SCSI bus adapters are available in all forms for everything from personal computers to intelligent peripherals, workstations, and mainframes. The SCSI variants appear to have displaced other mid-range formats including SMD (Storage Module Drive), ESDI (Enhanced Small Device Interface), and IPI (Intelligent Peripheral Interface). At the very high end, HIPPI (High Performance Parallel Interface) and FCS (Fiber Channel Standard) buses are replacing proprietary technology.

STORAGE DEVICES

Mass storage media have many competing technologies and formats. Two major categories will be covered: rotating memories (disk) and tape. Disk offers relatively rapid random access to data, while tape media are inexpensive for archival purposes. Table 1 compares many of the most popular formats available today. Many other formats exist; some have passed into extinction, others are relegated to specific niches, and some are just appearing.

Table 1. Summary of Mass Storage Device Characteristics

Device	Media	Bus	Capacity	Write Rate
Disk:				
5¼″ Hard Disk	magnetic	SCSI II F	3 GB	10 MB/s
3½″ Hard Disk	magnetic	SCSI II F	1.2 GB	10 MB/s
RAID	9 3½″ SCSI	SCSI II F/W	9.6 GB	20 MB/s
Array of Parallel Disk Arrays	9 5½″ ESDI	32-bit Parallel	36 GB	50 MB/s
Optical Disk	5¼″	SCSI II	650 MB	500 KB/s
Floppy 3½″	diskette	SCSI	1.44 MB	125 KB/s

Device	Media	Bus	Capacity	Write Rate
Tape:				
½″ Reel	2,400 ft. tape	SCSI	180 MB	750 KB/s
¼″ Cartridge	Data Cartridge	SCSI	150 MB	90 KB/s
4 mm Cartridge	Digital Audio Tape	SCSI	2 GB	180 KB/s
8 mm Cartridge	Video 8	SCSI	5 GB	500 KB/s
3480 Cartridge	3480 - ½″ tape	SCSI	200 MB	3 MB/s
VHS	T-120 Cassette	SCSI	10.5 GB	4 MB/s
DCSRi 2	Proprietary	8-bit parallel	47.5 GB	13.8 MB/s
ID1/2179	D1 cassette	8-bit parallel	95 GB	32 MB/s
1″ Reel (Analog)	9,200 ft. tape	Analog	460 MB	12 MB/s

Note: The data rates quoted by vendors are often reached for a short duration in telemetry data storage. The continuous nature of telemetry data storage will reduce these peak values to account for such factors as head movement, and waiting for the disk rotating to bring the start of sector under the head.

Disk — only a decade ago, disk drives used multiple platters over a foot in diameter and stored but a few megabytes. Today's technology stores 100 MB on a 1½" diameter drive. For scientific applications, 5¼" and 3½" formats store several Gigabytes each, achieving a 20 MB/s transfer rate.

RAID (Redundant Array of Inexpensive [or Independent] Disks) is a term coined at the University of California at Berkeley for Parallel Disk Arrays. Originally, five levels were defined by architecture or application, such as mirroring for backup, high performances, error correction, and transaction processing. Manufacturers are now refining these definitions and expanding the number of levels. A RAID not only increases data transfer rate but appears to the computer as a single volume, permitting larger contiguous files.

Floppy drives, where the writing head touches the media, transfer data very slowly and store relatively little data.

Optical disk drive technology is limited to a fraction of the transfer rate of disk drives. As with magnetic disk drives, however, access is random. In addition, the low media cost makes it ideal for archival purposes. Two media types are available: WORM (Write Once Read Many) and erasable. Today's ISO standard drive format encodes 330 MB per side.

Tape — in various formats and carriers is also relatively inexpensive; however, data access is serial.

One half inch reel-to-reel tape, found originally on mainframes and minicomputers, was supplanted by various cartridge formats for use on workstations and file servers.

Cartridge tape is available in a variety of formats, including computer vendor specific (3480 and TK50) industry standard (¼") and standard audio/video formats applied for digital recording (4 mm, T-120, and 8 mm).

Cartridge media, such as optical disk and tape, are excellent media for long-term storage. **Robotic** systems store tens to hundreds of cartridges in "jukeboxes" for retrieval and insertion into a single drive by a mechanical arm. A hierarchy of media may be assembled to move data from fast retrieval disk to on-line tape to archived cartridges. In this way, many terabytes may be accessed on-line.

ARCHITECTURE ALTERNATIVES

There are several ways of achieving telemetry storage requirements:

- Traditional DMA approach to a host computer
- Tightly-coupled Telemetry Front End (TFE) Storage
- TFE Data Server (TDS)

Traditional DMA Approach to a Host Computer

In the traditional DMA approach, the TFE uses a DMA controller to store the raw or processed data to a host computer. The choice of the host computer is widely open in this approach. It can start from a low-end Sun workstation with an Sbus DMA adapter and end with a highly parallel supercomputer with high-performance computer interfaces such as: HIPPI, Fiber Channel, etc.

The single most important advantage of this approach is the fact that it requires minimal software changes to the already existing post-processing applications. The slow data rates on the host computer storage devices are the primary disadvantage of this method. Many current operating systems lack a real-time file system that is finely tuned to store data at the maximum possible rates of current storage devices such as RAIDs. The primary example of such file systems is the UNIX file system, which uses an index node (inode), an array of pointers to file blocks. The file block allocation is not accomplished on a contiguous disk block. Therefore, disk block access times vary based on their location on the different cylinders and tracks. A raw UNIX file system is a better alternative for achieving higher data storage rates because the entire disk can be treated as one inode and the overhead of disk block accesses is reduced greatly. The storage management software can then overlay a contiguous file system on the raw disk device. The UNIX kernel itself uses a raw partition for doing process/program swapping. This points up another disadvantage of using the host computer for storage; that is, the lack of real-time scheduling in either a UNIX or a VMS operating system cannot guarantee a sustained storage rate. The primary overhead of using such operating systems is the amount of time required to do context switching in a real-time operating-system. This time is 6 μs (based on Wind River Systems' VxWorks 5.0 Benchmark Report using a 68040, zero wait state memory) whereas in real-time proprietary UNIX implementations, context switching is on the order of 100 μs. The primary reason for a small context switch time is that the amount of per-process/task resources required to be saved and restored is less than the UNIX or VMS process control block data structures. Secondly, the memory model in a real-time operating system does not use a demand paging architecture, whereas the UNIX and VMS operating systems have to use a paging system. The context switch time for such a paging system can result in having to write pages of a process to a swap space on the disk, resulting in even lower storage data rates of a PCM stream.

Tightly-Coupled TFE Storage

The TFE Bus—whether VME or a proprietary bus—is used directly to store the PCM data into a storage medium. The advantage of this approach is in the higher sustained storage rates. The major disadvantage is in the post-processing application software changes required to interface to the TFE stored files. The Client/Server model, using Remote Procedure Call (RPC) libraries, provides similar functionality to the use of the host operating system calls for opening, reading, writing, seeking, and closing files. A post-processing program can use the open RPC routine to open a file on the TFE storage medium and then use the read RPC routine for reading raw or processed data into an application buffer for post processing. This method does not require copying of the TFE file onto the host storage medium; therefore, no storage space is used on the host computer. Coupling storage tightly with the telemetry hardware can have an operational advantage resulting from the fact that telemetry decommutation is shared with the storage and post-processing functions. The function of storage and post-processing services can be combined as a data server functionality in a separate environment.

TFE Data Server Architecture

The Data Server architecture is based on providing high-performance I/O coupled with a high-speed network interface to provide data services to the post-processing engineers who use powerful graphics workstations. The Telemetry Data Server (TDS) can be set up such that the ground station TFE can decommutate/process data and route data to the TDS via a high speed LAN interface. The data is then stored in a high-speed RAID storage system. TDS can also provide concurrent playback operation while a current storage session is in progress. This is accomplished by archiving the current file to medium-speed SCSI drives in the background while acquiring real-time data. This hierarchical storage design allows multiple users to play back different archived files while a recording session is in progress. Having a dedicated environment for post processing also means the ability to run algorithms on the stored data in the data server. This is accomplished by providing dedicated processors in the data server to do post processing. The processors can be easily programmed to provide data reports and data tapes for the post processing end users. The Data Server also supports the Network File System (NFS), which is the industry standard for a transparent file system that can be accessed from a PC, a UNIX host, or a VMS host. NFS files appear as regular files on a local machine; these files can be accessed transparently without changes to the application software.

TECHNIQUES IN HIGH-SPEED STORAGE

There are several issues to consider when designing a storage solution:

- Telemetry Input Buffer Size
- Storage Block Size
- Storage Device Command Processing
- High-Speed File System

Telemetry Input Buffer Size

The telemetry data must be stored at decommutation rates in ping-pong buffers so that one buffer is stored in the storage device while the telemetry data is saved in the secondary buffer. The buffer size must be large enough to accommodate higher storage rates for the storage medium. This buffer size is fully adjustable based on different storage devices and interfaces.

Storage Block Size

The storage interfaces support an adjustable block size based on different media such as RAIDs, IPI drives, and 8-mm tape drives. The block size transfers can directly affect the storage data rate; therefore the ability to adjust the block size can guarantee sustained transfer rates based on the storage device specifications.

Storage Device Command Processing

The command processing time between host adapters and storage devices can greatly affect data transfer times. The ability to do command chaining reduces the command processing time because one transaction provides multiple commands that can be executed in a queue. This requires a certain level of synchronization as to when data is available for the command queue, but the benefit of reducing command transaction overhead time is greater than the synchronization complexity.

High-Speed File System

In a high-speed file system, the simple technique is to structure files in contiguous blocks so that the data is not scattered across cylinder boundaries. This method is optimized for high-speed transfer rates. Files are pre-allocated before a recording session starts. Once a file is closed, directory information is updated.

A GRAPHICAL USER INTERFACE FOR DATA ARCHIVAL

The graphical user interface (GUI) to the SCSI storage subsystem on the System 500 Model 550 is an example of a user interface that simplifies the process of setting up and controlling data storage. One SCSI storage subsystem includes up to seven SCSI disk or tape units. The operator's interface to the storage subsystem is a family of Motif windows that contain programmable fields with toggle buttons and pop-up menus that allow the operator to define all the conditions for recording, playback, disk backup to tape, and disk restore from tape. The operator executes tape or disk control commands by using the mouse or the keyboard to "press" push buttons in Command Windows onscreen. Menu bar commands allow the operator to manipulate SCSI storage subsystem setup files; initialize the hardware or the display; post information about the current SCSI storage subsystem hardware, software, and data base; and open other windows.

Figure 3 shows examples of the Motif windows used to set up and control recording to a SCSI disk. One screen, the Main Window, contains all of the fields required to set up the conditions for recording, including the name of the disk file, the data format, the event that triggers recording start/stop, the conditions for recording time with the data, the user-programmable distribution code that designates the parameters to be directed to disk, and the size of the disk file. A Command Window contains push buttons used to start and stop recording when the "Manual Start/Stop" recording method is selected. Recording start/stop may also be triggered by the appearance of designated parameters on the TFE's real-time bus. When the parameter trigger start/stop method is selected, the operator uses listpick menus to select the trigger parameters from a complete list of the parameters in the data base. A dynamically updated status window reports the status of the current operation.

The Motif interface also includes windows for viewing and manipulating disk and tape files. Figure 4 shows an example of the windows used to manage and view SCSI disk files. The File Command window provides programmable fields for defining the attributes of disk files, such as the file name, file size, and write protection status. Push buttons allow the operator to execute file management commands such as File Protect, Copy, Rename, Truncate, Create, and Delete. A Disk List window provides two list types, a disk directory and a detailed information profile on a selected disk file. The File View window displays the contents of a disk file in ASCII form. The operator selects the file name, the disk location at which the file begins, and the size of the desired segment. A scroll bar allows the operator to scroll through the file

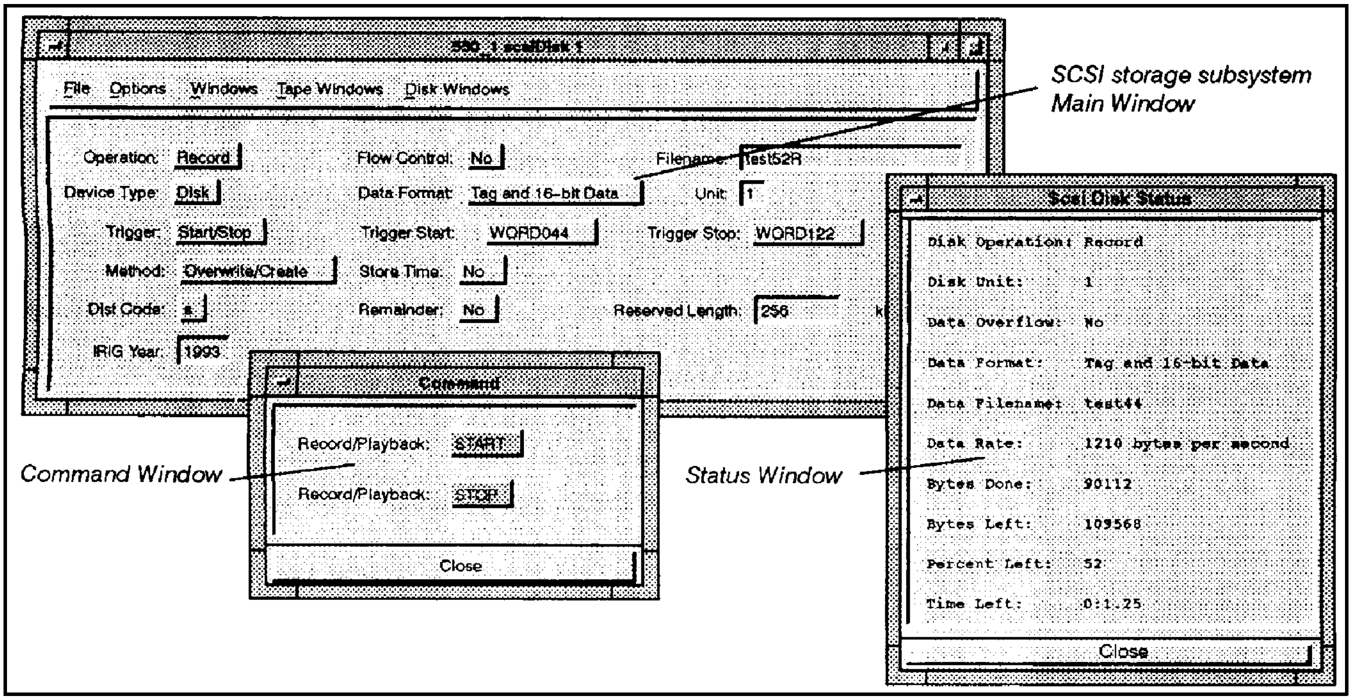

Figure 3. Examples of Windows Used to Set Up and Control Recording to Disk

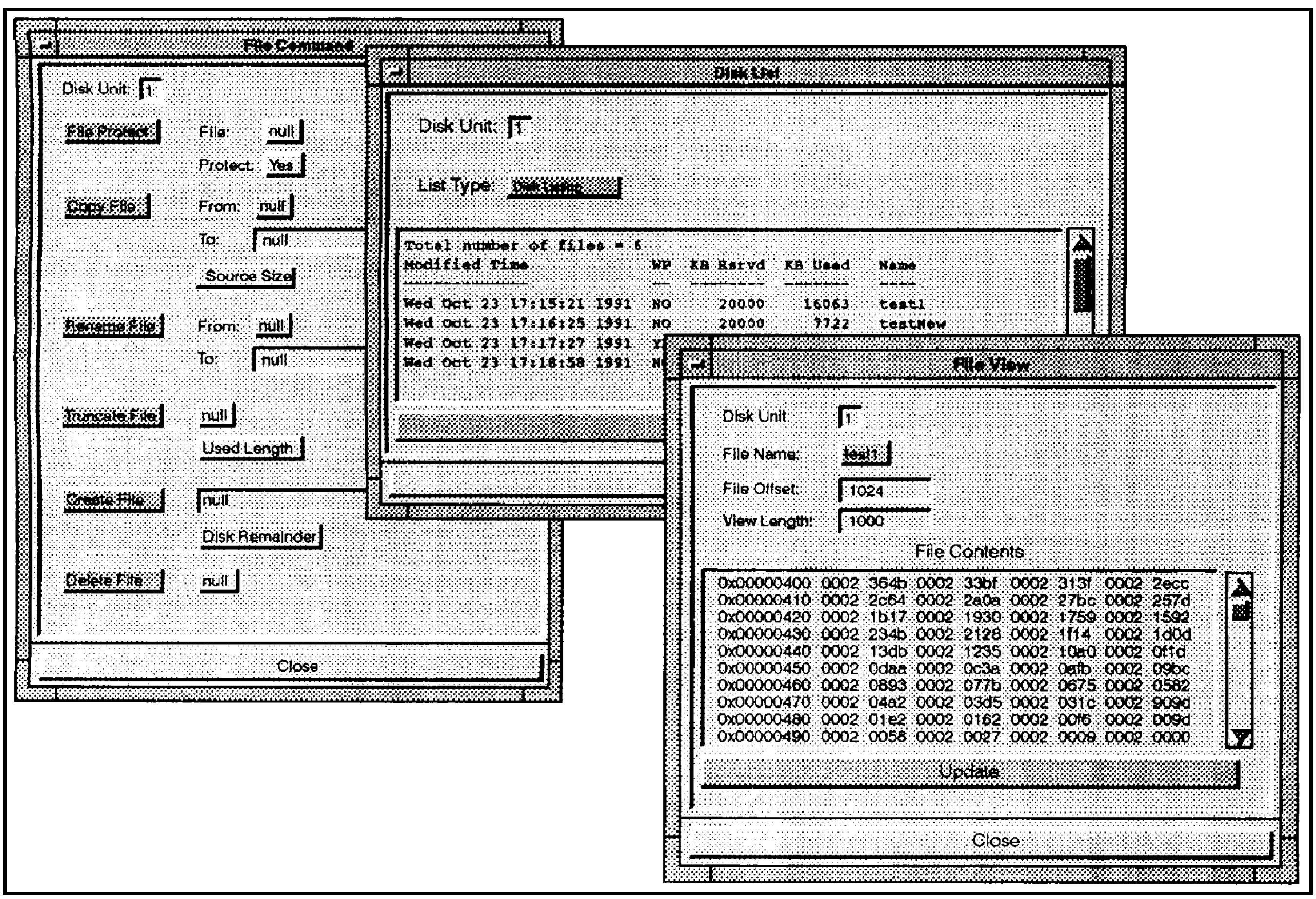

Figure 4. Windows Used to Manipulate, List, and View Disk Files

contents. By selecting different file names, offsets, and view lengths and then pressing the Update bar, the operator may examine one file or file segment after another without exiting the window.

CONCLUSION

Current computer standards and technology provide a host of solutions for real-time data storage in a telemetry station. A variety of disk alternatives offer data storage from 3 GB at 10 MB/s to 36 GB at 50 MB/sec on a single volume. Though the tape media is relatively inexpensive, the various formats lack the ability to quickly access random data, as is possible in disk formats. The highest performance tape can store 95 GB at 32 MB/s on a professional video tape cartridge. These high transfer rates preclude the use of general-purpose workstations or host computers. The storage device must be tightly coupled to the Telemetry Front End (TFE); however, computing resources on high-speed networks attached to the TFE can access selected data using the Network File System (NFS). Modern Graphical User Interfaces can ease the control and manipulation of archived data while presenting meaningful status information.

ACKNOWLEDGEMENTS

Bette Jo Welch

Linda Richards