Mark 5 Disk-Based Gbps VLBI Data System

Alan R. Whitney
MIT Haystack Observatory
Westford, MA, USA
awhitney@haystack.mit.edu

Abstract

The Mark 5 system is being developed at MIT Haystack Observatory as the first Gbps VLBI data system based on magnetic-disk technology. Incorporating primarily low-cost PC-based components, the Mark 5 system supports data rates up to 1024 Mbps recording to an array of inexpensive removable IDE/ATA disks. The system is being developed with support from BKG, EVN, KVN, MPI, NASA, NRAO and USNO; prototype Mark 5 units are now in routine use at several stations and correlators, with ~25 units operating as of fall 2002.

Introduction

The Mark 5 system is being developed as the first Gbps VLBI data system based on magnetic-disk technology. Incorporating primarily low-cost PC-based components, the Mark 5 system supports data rates up to 1024 Mbps, recording to an array of inexpensive removable IDE/ATA disks. It is expected that disk-based VLBI systems will replace current magnetic-tape systems over the next few years.

The goals of the Mark 5 system are:

- Low cost
- Based primarily on unmodified COTS components
- Modular, easily upgradeable
- Robust operation, low maintenance cost
- Easy transportability
- Conformance to VSI specification
- Compatibility with existing VLBI systems during transition
- Flexibility to support e-VLBI
- Minimum of 1 Gbps data rate
- 24-hour unattended operation at 1 Gbps

All but the last goal are clearly achievable with today's technology; 24-hour unattended operation at 1 Gbps is expected to arrive naturally within ~2-3 years with continued development in disk technology.

Why Disks?

Though both magnetic-disk technology and magnetic-tape technology have made great strides over the past few years, the pace of magnetic-disk development has been no less than spectacular, far exceeding even disk-industry projections. Figure 1 displays a comparison of disk and tape prices over the past several years, showing that disk prices (on a $/GB basis) continue downward in a still-accelerating trend. Current (spring 2002) consumer IDE disk costs are ~$1.5US/GB and falling rapidly; current Mark4/VLBA tape prices are ~$2US/GB and remaining steady. By ~2005-2006, industry projections suggest the price of disks will fall to ~$0.5/GB. Similarly, current single-disk capacities are ~200 GB and rising; by ~2005-2006, single-disk capacities are expected to rise to 500-1000 GB! A single Mark 5 system with sixteen 700 GB disk drives will record continuously 1024 Mbps for 24-hours unattended.

In addition to falling prices and increasing capacity, disks have several other advantages:

- Readily available inexpensive consumer product;
  continually improving in price/performance with standard electrical interface
- Self contained drive mechanism, so host system can be inexpensive
- Technology improvements independent of electrical interface
- Rapid random access to any data
- Essentially instant synchronization on playback to correlator (no media-wasting early starts needed)
- No headstacks to wear out or replace – ever!

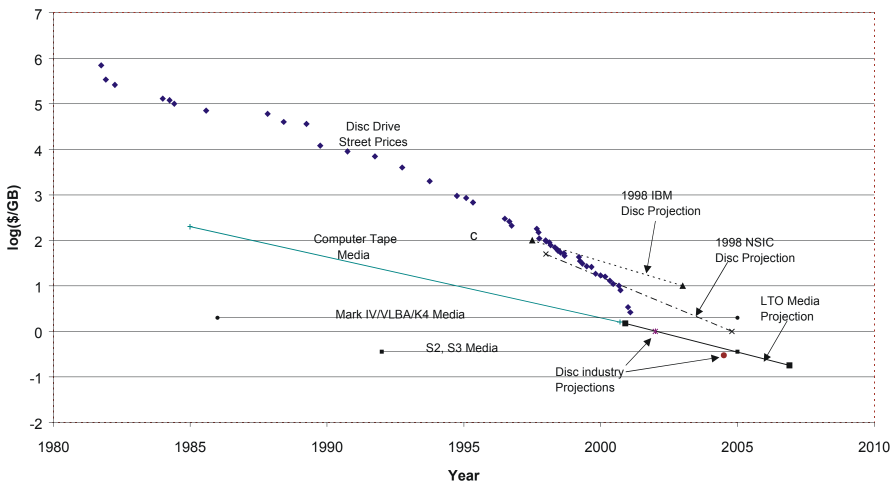

Figure 1: Disk and tape prices vs. time

Mark 5 Development Program

Based on the success of a 512 Mbps Mark 5 demonstration unit [Ref 1] in early 2001 (Figure 2; developed and demonstrated in 3 months time!), Haystack Observatory is developing an operational 1 Gbps Mark 5 system with support from BKG, KVN, MPI, NASA, JIVE, NRAO and USNO.

Figure 2: Mark 5 demonstration system

The Mark 5 system is being developed in two stages:

1. Mark 5A: The Mark 5A system is intended as a direct replacement for a Mark 4 or VLBA magnetic-tape transport at either a station or a correlators. It records 8, 16, 32 or 64 tracks from a Mark4/VLBA formatter, and plays back in the same Mark4/VLBA format. As such, the Mark 5A is a direct replacement for a Mark4 tape unit at 1024 Mbps and VLBA tape unit at 512 Mbps. The Mark 5A system is in operation at several antennas and correlators, with ~25 Mark 5A prototype systems deployed by late 2002.
2. Mark 5B: The Mark 5B is VSI-compliant [Refs 2, 3] system with capability up to 1024 Mbps; no external formatter is necessary. The system will also support several backwards-compatibility modes with existing Mark4/VLBA correlator systems. The Mark 5B is expected to be deployed in late 2003.

Figure 3 shows a photograph of the prototype Mark 5A system with its two removable ‘8-pack’ disk modules. A Mark 5A system may be upgraded to a Mark 5B system simply by replacing a Input/Output PCI board in the host PC.

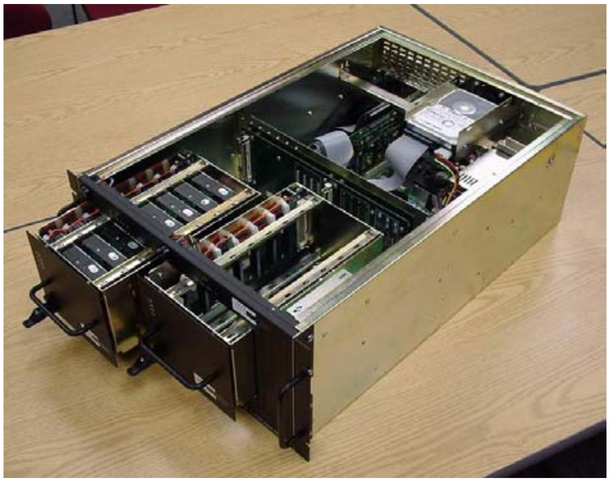

Figure 3: Prototype Mark 5A VLBI data system

Triangle of Connectivity

The Mark 5 is based on a standard PC platform and using a combination of COTS and custom-designed interface cards.

The heart of the system is a 'StreamStor' disk interface card from Conduant Corporation that is specially designed for high-speed real-time data-collection and playback. The StreamStor card supports three physical interfaces in a 'triangle of connectivity' as shown is Figure 4:

1. <u>Data Port/FPDP</u>: This port is present as a 32-bit card-top bus which supports the industry-standard 'Front-Panel Data Port' interface specification [Ref 4]. This is a two-way port through which high-speed real-time data may be either input or output. All 32-bits of the FPDP bus are always active.
2. <u>Disk array</u>: This port supports up to 16 standard IDE disks for reading or writing, arranged in 8 master/slave pairs.
3. <u>PCI bus</u>: This is the standard connection to the host PC platform; the StreamStor card supports a 64-bit/66MHz bus, though it is backwards compatible with standard 32-bit/33MHz buses.

The 'triangle of connectivity' shows that data may be moved in either direction between any two of the three ports. The StreamStor card supports a maximum sustained data transfer rate of up to ~1200 Mbps between any two ports, though only one connection path may be exercised at a time and the maximum data rate for VLBI usage is anticipated to be ~1024 Mbps. The path exercised for traditional VLBI observations is between the FPDP bus and the disk array; note that in this mode, the VLBI data never touch the PCI bus, so the speed of the PC platform is largely irrelevant. Of course, the path between the disk array and the PCI bus allows the PC to read and verify VLBI data written to the disks via the FPDP port. The direct connection of FPDP bus to PCI bus will be used in upcoming e-VLBI experiments where data are transferred directly to a high-speed network and are not recorded locally. An on-board 512 MB buffer provides the necessary 'elasticity' between the three connection nodes to support full real-time operation.

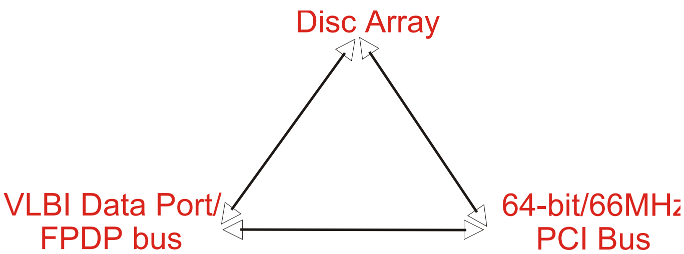

Figure 4: 'Triangle of connectivity' of StreamStor interface card

Mark 5A

The Mark 5A system is designed to be a *direct plug-compatible replacement for a Mark 4 or a VLBA tape recorder at either a field station or a correlator*, except that:

1. The Mark 5A can record and playback 8, 16, 32 or 64 'tracks' of data from a Mark 4 or VLBA formatter[1].
2. The software control of the Mark 5A is based on VSI-S rules and syntax, and is vastly simpler than controlling a Mark 4/VLBA tape transport.

The I/O Panel on the rear of the Mark 5A is shown in Figure 5. It contains the same set of connectors as a Mark 4 or VLBA tape drive so that the Mark 5A is very easily substituted for a tape drive.

Figure 5: Mark 5A I/O panel at rear of unit

The Mark 4 formatter can output up to 64 tracks of data at 18 Mbps/track, which includes the parity overhead, for a total of 1152 Mbps. However, it is not necessary to record the parity bits in the low-error-rate environment of the disks; therefore, the Mark 5A Input/Output Interface Board does do some special processing to remove the parity bits for recording and restore them on playback. A simplified block diagram is shown in Figure 6.

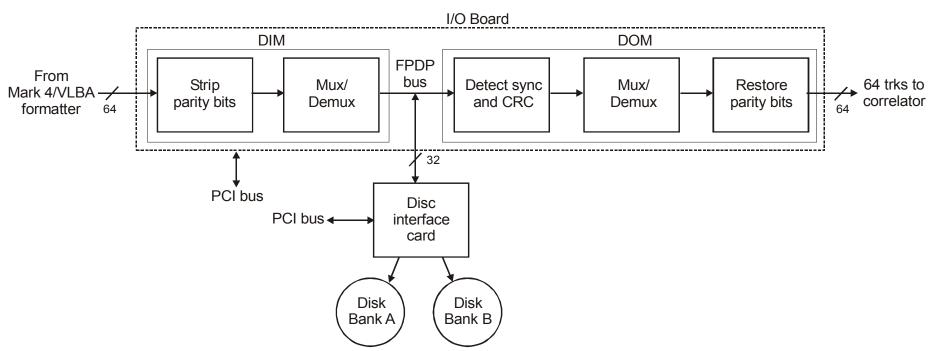

Figure 6: Mark 5A simplified block diagram

Data Input Module (DIM) Section of I/O board (in block diagram in Figure 6):

1. All parity bits are removed.
2. In the 8 and 16-track modes, the data are demultiplexed by a factor of 4 or 2, respectively, and the resulting 32 parallel bit streams are sent directly to the FPDP bus for recording.
3. In 32-track mode (equivalent to Mark 4/VLBA headstack 1 only), the resulting 32 parallel bit streams are sent directly to the FPDP bus for recording.

[1] When recording 8, 16, 32 or 64 'tracks', 100% disk utilization is achieved; other numbers of tracks can be recorded, but with <100% disk utilization. Both Mark 4 and VLBA formatters can be configured to multiplex a single channel to 1, 2 or 4 tracks with 1-bit samples or to 2, 4 or 8 tracks with 2-bit samples, so that normally disk utilization is near 100%.

4. In 64-track mode (equivalent to Mark 4/VLBA headstack 1 and 2), adjacent even and odd track-pairs are interleaved bit-by-bit before being sent to the FPDP bus. Maximum data rate at this point will be 1024 Mbps for Mark 4 and 512 Mbps for VLBA[2].

For compatibility with standard VLBA and Mark 4 formatter modes [Ref 5], 8-track mode records VLBA-equivalent even track numbers in the range 2-16 inclusive. 16-track mode records VLBA-equivalent even track numbers in the range 2-32 inclusive. 32-track mode records VLBA-equivalent tracks 2-33 inclusive from headstack 1. 64-track mode records tracks 2-33 from both headstacks 1 and 2.

The DIM also includes a 'straight-through' mode in which 32 input bit streams are recorded directly on the disk with no processing or multiplexing; this is a useful test mode and may also be used for other applications.

The DIM also can generate several test patterns for system testing and diagnosis, including a standard VSI-H test pattern.

Data Output Module (DOM) Section of I/O board (block diagram in Figure 6):

The DOM must reverse the actions of the DIM to exactly reproduce the original 'track' data streams. In all but the 'straight-through' mode, the board must first undo any multiplexing/de-multipexing to reconstruct parity-free 'track' data streams. Embedded Mark4/VLBA sync words and CRC characters are then detected to determine data framing, parity is restored, and the fully reconstructed track data are sent to the output.

For normal VLBI operation, the DOM is instructed whether the data are VLBI or 'straight-through'. If VLBI mode, the board automatically recognizes whether data is in 8, 16, 32 or 64-track mode and reconstructs track data accordingly.

Data Clocking

The DIM is driven by data and clock from the Mark 4 or VLBA formatter at a maximum frequency of 18 MHz. In turn, the DIM drives the FPDP bus with data and clock at an average frequency 8/9 (32-track case) or 16/9 (64-track case) times the formatter clock frequency.
The DOM has two clock modes:

1. In <u>Record/Bypass</u> mode, the DIM passes the 'track' clock over the FPDP bus to the DOM.
2. In <u>Playback</u> mode, the DOM may clock the output data either according to an on-board digitally-synthesized clock (0-40MHz, ~23mHz resolution) or from an external user-supplied clock in the range 0-40 MHz. The playback data rate is completely independent of the record data rate, though the normal maximum 'track' playback data rate is ~18 Mbps/track for compatibility with existing correlators. The DOM reads disk data at whatever rate is necessary to sustain the requested 'track' data rate.

Data format on Disk

The format of the data recorded to disk is such that barrel-rolling and/or data modulation are not necessary. The StreamStor card accumulates 32-bit 'words' (corresponding to the 32-bit wide FPDP bus) into 64 kB 'chunks and writes these chunks sequentially in a 'round-robin' fashion to the disk array. Neither 'barrel-rolling' or 'data-modulation' need to be applied by the formatter since these functions are specifically designed to overcome limitations with magnetic tape.

Operating modes

For normal operation with disks, the Mark 5 system has three basic operating modes:

1. <u>Idle</u> – When the system is 'idle' (i.e. not recording or playing back), the data entering the DIM passes through the FPDP bus and is re-created (with a delay of a few clocks cycles) at the DOM. This is also known as 'pass-through' mode and allows multiple Mark 5 systems to be daisy-chained together to increase overall data capacity, as shown in Figure 7..
2. <u>Record</u> – This mode is identical to 'idle' mode' except that the disks record the data on the FPDP bus.

[2] Actual maximum data rate for VLBA is slightly higher (516.096 Mbps to be exact) due to non-data-replacement nature of the format.

3. Playback – Data pre-recorded on disk is played back to the FPDP bus and then to the DOM. In this mode, the playback rate is controlled by a clock provided by the DOM (either on-board or external) and is independent of the record clock rate.

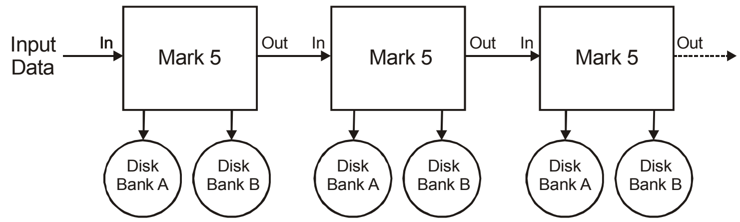

Figure 7: Recording with cascaded Mark 5 systems

When recording, the Mark 5 adds a new scan (recording) after the end of the last scan, just like tape. However, scans may be deleted one-by-one from the last scan towards the first scan and the disk space reused.

Additional features

Several additional features exist in the Mark 5 system to ensure data integrity and continuity:

1. The Mark 5A chassis holds two '8-pack' modules of disks. Each module can support up to 1024 Mbps and each module can operate independently (but not simultaneously). A 'module switching' mode is being implemented so that recording is automatically switched from one module to another when the first is filled; this allows continuous data to be taken over module boundaries. Modules are 'hot-swappable'.
2. If a disk becomes slow or fails during recording, the recording load is dynamically adjusted among the disks so that no data are lost.
3. If a disk fails or is missing during playback, the Mark 5 will fill any data gaps with a user-specified data pattern that can be detected and cause the data to be invalidated at the correlators.
4. '8-pack' disk modules can be managed just like tapes with little or no modification to present management procedures. Each '8-pack' disk module contains a permanent electronically-readable Volume Serial Number (VSN) just like Mark 4/VLBA tape. Current tape labeling schemes can be directly transferred to 8-pack disk modules.

Software Control

Control of the Mark 5A system is through a software program on the host PC machine operating under a Linux OS. The control syntax and rules are taken from VSI-S, though the commands are mostly not standard VSI-S commands. Control be from a local or remote terminal or over programmatically over a network. Though there are many modes and many test and status commands and queries, the fundamental operation of the Mark 5A is quite simple. The Mark 5A system is supported by the NASA/GSFC Field System for control at stations.

Deployment and Availability

As of late 2002 approximately 25 Mark 5A systems will be operating at various stations and correlators in the world. Mark 5A systems may be ordered from Conduant Corporation for ~$17,500 each without 8-pack modules or disks. Those who wish to assemble and test Mark 5 system themselves can do so for ~$12-13K each.

Mark 5B

The Mark 5B will be fully VSI-H and VSI-S compatible, allowing more compatibility among various VSI-compatible VLBI data systems. A new Mark 5B Input/Output Board will be designed to replace the Mark 5A Input/Output Board. Among the features of Mark 5B will be:

1. Internal data formatting eliminates need for external formatter.
2. Simple input and output interface, each on a single 80-pin connector.
3. 32-bit wide input and output data channel.
4. Any 1, 2, 4, 8, 16 or 32 input bits streams may be selected for recording with 100% disk-capacity efficiency.
5. Easy synchronization at a correlator with an external second tick.

The Mark 5B system is expected to be available in late 2003.

Compatibility Considerations

The Mark 5 system is being designed for extensive forward and backwards compatibility with existing VLBI systems. For example, data may be recorded with a VSI-compatible interface and re-played into any Mark4/VLBA correlator. Conversely, data may be recorded from a Mark4/VLBA system and re-played into any VSI-compatible correlator. In addition, it is expected that existing interfaces to S2 recorders can be easily adapted to record on Mark 5B, which can then be re-played into either a VSI-compatible or Mark4/VLBA correlator.

This inter-compatibility among various systems will allow a much broader and flexible use of existing VLBI facilities throughout the world.

e-VLBI Support

The Mark 5 system allows easy connection of a VLBI data system to a high-speed network connection. Because the Mark 5 system is based on a standard PC platform, any standard network connection is supported.

Depending on the availability of high-speed network connections, this can be accomplished in at least two ways:

1. Direct Station to Correlator: If network connections allow, data may be transferred in real-time at up to 1 Gbps from Station to Correlator, either for immediate real-time correlation or buffering to disk at the Correlator.
2. Station Disk to Correlator Disk: If network connections are not sufficient to allow real-time transmission of data to the Correlator for processing, data may be recorded locally to disk at the Station, then transferred to disk at the Correlator at leisure for later correlation.

Depending on available network facilities, either entire experiments or small portions of experiments may be transmitted electronically. The latter may be particularly useful for verifying fringes in advance of important experiments.

Possible Future Enhancements

Among the future enhancements being studied for Mark 5:

1. Support for interchangeability between 8-pack modules containing parallel-ATA disk drives and 8-pack modules containing serial-ATA disk drives.
2. Support for an inexpensive ‘expansion’ chassis that will support additional 8-pack disk modules.
3. Full compatibility with standard Linux file systems.

Summary

The Mark 5 system promises to move VLBI data systems to new levels of high-performance and low-cost by leveraging the enormous investments of the computer industry in high-speed data technology. Within only a very short time, the possibilities to economically expand VLBI observing programs by large factors appear to be within reach.

References

1. “Second interim report on COTS-VLBI project” by Alan R. Whitney, 8 Mar 2001, Mark 5 memo 3, available at http://web.haystack.mit.edu/mark5/Mark5.htm.
2. “VLBI Standard Hardware Interface Specification –VSI-H”, Revision 1.0, 7 August 2000, available at http://dopey.haystack.edu/vsi/index.html.
3. “VLBI Standard Software Interface Specification – VSI-S”, under development.
4. FPDP Specification, available at http://www.fpdp.com/.
5. “Mark IIIA/IV/VLBA Tape Formats, Recording Modes and Compatibility - Rev 1.2” by Alan R. Whitney, 28 Sep 2000, Mark 4 memo 230, available at ftp://dopey.haystack.edu/pub/mark4/memos/index.html.

Note: The Mark 5 web site at http://web.haystack.mit.edu/mark5/Mark5.htm contains much additional information about the Mark 5 system.