Alan R. Whitney et al.: Data Acquisition and Transport - Looking 2010 and Beyond, IVS 2004 General Meeting Proceedings, p.85–89
http://ivscc.gsfc.nasa.gov/publications/gm2004/whitney2

Data Acquisition and Transport - Looking 2010 and Beyond

Alan R. Whitney [1], *Hans F. Hinteregger* [1], *Tetsuro Kondo* [2], *Yashuhiro Koyama* [2]

1) *MIT Haystack Observatory*

2) *Kashima Space Reserch Center, CRL*

Contact author: Alan R. Whitney, e-mail: awhitney@haystack.mit.edu

Abstract

In contrast to the first ~30 years of VLBI development, where highly specialized equipment for VLBI data acquisition was designed and built at great cost, the last few years are being driven more and more by taking advantage of rapidly developing technology in the computer and networking industry. This trend is only likely to accelerate, and VLBI must position itself to take maximum advantage of these technologies. Already, the transition from magnetic tapes to magnetic disks has been very rapid, and disks will almost certainly be the mainstay of VLBI data acquisition for the next few years. However, development of e-VLBI continues to be rapid and will accelerate. Already, international e-VLBI links of more than 500 Mbps have been demonstrated, with speeds in excess of 1-10 Gbps surely achievable in the near future. The advantages of real-time and near-real-time VLBI made possible by e-VLBI are significant, but major potential stumbling blocks are 'last-mile' connectivity for many telescopes and potentially high networking costs. The future of VLBI data acquisition and transport in the light of current and projected developments will be examined.

1. Introduction

Since the beginnings of VLBI in the late 1960s, a large segment of VLBI instrumentation development has been directed towards the development of specialized data systems to meet the extreme demands of both bandwidth and total data quantity that VLBI requires. Until recently, these systems have primarily been focused on magnetic tape as the only practical media. Beginning with modest data rates of <1 Mbps in 1967 recorded on 1/2" open-reel 2400-ft computer tapes at 800 bpi/track, several generations of specialized magnetic-tape-based data systems have evolved. In the past few years, however, with the rapid advance and broad commercialization of the computer industry, it is increasingly possible to take advantage of relatively inexpensive commodity technology as a basis for very high-performance VLBI data systems. An example is the recent development of magnetic-disk-based systems, such as the K5 [Ref 1] and Mark 5 [Ref 2] systems operating at 1024 Mbps, available for a tiny fraction of the cost of the tape-based systems they replace.

Electronic transmission of VLBI data, dubbed 'e-VLBI', is not new. As early as the mid-1970s, an experiment using satellite data transmission was successfully executed [Ref 3], but was not economically sustainable. In the late 1970s some VLBI data were transmitted via standard voice telephone lines at 1200 bps and successfully correlated [Ref 4], but this data transport mode was practical only for short tests. In the mid-1990s, dedicated fiber networks spanning limited geographic areas in Japan were successfully put into operation at data rates of 256 Mbps [Ref 5]. However, only in the last few years, with the advent of a global grid of high-speed fiber connections, has the possibility of transmitting a substantial amount of VLBI data become a practical possibility.

In this paper, we will examine the state-of-the-art in VLBI data systems and try to project to both the demands and the available technology in 2010 and beyond to meet these demands in the most cost effective and efficient manner.

2. VLBI Data Transport Demands in 2010

For continuing significant improvement in the geodetic VLBI technique over the next decade, it is clear that several things must happen. Among the likely changes:

- A global dedicated network of ~12-25 identical geodetic-VLBI stations, probably co-located with other space geodetic techniques (SLR, GPS among them), will be established.
- For both economic and scientific reasons, the VLBI antennas at these locations will be small and fast-moving. Antenna diameters in the range of 8-20 meters are a reasonable guess.
- In order to counter the burgeoning RFI problems encountered at some of todays VLBI stations, particularly at S-band, it is likely that a shift to higher observing frequencies will take place, perhaps to 8/15 GHz or maybe higher. Systems may even be available that are able to observe over extremely wide single bands from, say, 5 to 20 GHz.
- Due to smaller antennas and likely weaker source strengths at higher observing frequencies, observing bandwidths must be widened to re-gain the necessary SNR for precise geodetic observing. A reasonable guess is that aggregate data bandwidths that must be captured for correlation will be at least a few GHz wide, resulting in corresponding data rates of at least several Gbps/station.
- The duty cycle of observing will substantially increase at these dedicated stations, perhaps up to 50-75% averaged over a week.
- In order to operate efficiently and economically, these stations will be largely unattended by operators and must be remotely managed.

We will examine the possible options for data transport based on this likely scenario.

3. Modes of Operation

Though direct transmission of the data via e-VLBI to a correlator is clearly the most desirable from an efficiency point of view, there are a range of alternative data transport strategies that must be considered. Furthermore, the data transport strategy may evolve with time as technology permits or economics dictates, so the data transport system may need to be designed from the beginning with this evolution in mind. The range of possible data transport options is shown in Table 1, listed in order of least desirable to most desirable.

4. Magnetic Disk

The rapid advance of magnetic-disk technology over the past few years has seen magnetic disks literally burst onto the VLBI stage as performance and capacity rapidly increase and price per GB decrease. Still, we should assess the advantages and disadvantages of magnetic disks and try to project where the technology may be in 2010 and beyond. Some of the advantages and disadvantages are:

Table 1. Possible modes of data transport (in order from least to most desirable)

Situation at station	**Data transport mode**
No data link to correlator	Record, ship to correlator
Nearby high-speed POP with link to correlator	Record, transport to POP, transmit to correlator and record at correlator
'Slow' network connection to correlator	Use disk FIFO buffer - transmit at whatever speed is available; record at correlator. Sustainable observation duty cycle is constrained.
High-speed link at station, but not to all stations	Transmit to correlator and record
High-speed links to all stations	Transmit and correlate in real-time (no recording)

Advantages

- Fast technology advance: In terms of capacity, the technology advancement rate of disks over the past 10 years has outstripped the Moores Law advancement rate of semi-conductor technology by about 50%; this shows some signs of slowing, but many analysts attribute this more to market forces rather than technological advances. Nevertheless, advances in capacity are remarkable even over the past 3 years, moving single-disk capacity from ~120 MB in 2001 to 1 TB today (early 2004). Improvements in single-disk data rates have been much slower, but aggregation of multiple disks into single physical modules to increase aggregate data rates is both practical and economical.
- Rapidly falling prices: Price per GB has been dropping about 50% per year for the past 5 years. Though there is some evidence this rate may be slowing, this is still a much faster decline than any other competing media. Typical disk media costs have now dipped below $1/GB and are expected to drop to ~$0.50/GB over the next year.
- Easy to take advantage of new technology: Because disk drives connect to the system at a standard electrical interface, the technology used in the disk drive is invisible to the user and may advance and change radically with no impact on the design of the data system.
- Very large market: It is obvious that the vast market for disk drives is the driving force in the rapid technology advance and decrease in price. Currently, there are no obvious challengers to the domination of magnetic disks for large mass-market random-access storage systems.

Disadvantages

- Shipping fragility: Compared to magnetic tapes, magnetic disks are more susceptible to shipping damage due to the delicate nature of their internal mechanisms. However, experience has shown that susceptibility to damage is somewhat vendor dependent. It should be pointed out, however, that some of the disk-based VLBI data systems can continue to operate in the face of the loss of one or more drives with only slightly degraded performance. And, of course, not all data transport scenarios that involve local media require shipping the media (see Table 1).

- Limited data rates: Compared to magnetic tape, multiple-disk modules can easily support data rates to 1 Gbps. However, if date rates of the order of 10 Gbps or higher are required, even disks become problematic due to the large number necessary to support these high rates. For example, at least 30 disks of todays commodity variety, operating simultaneously, would be necessary to sustain operation at 10 Gbps; the practicalities and economics of such a situation need to be studied further.

5. Other Types of Recordable Media

Over the years we have seen many promises for alternative types of media, particularly optical-based media, but in practice these have not become contenders for VLBI data storage. For example, recordable optical disks are still more expensive than comparable magnetic-disk storage, especially re-writable optical media; some non-re-writable optical media is comparable in cost to magnetic disk cost, but data densities are rather now and the media costs would be recurring. Additionally, sustainable data rates on optical media are generally lower than for magnetic disks. Optical holographic storage has been long promised but has never become a commercial success.

In terms of recordable media, magnetic disks reign as king and show no hint that they will relinquish their crown over the next decade.

6. e-VLBI

e-VLBI, particularly with full real-time processing, is the ultimate solution for VLBI data transport; e-VLBI is rapidly coming of age, though it also brings its share of problems as well. Some of the advantages and disadvantages:

Advantages

- Potential for very high data rates: Since increasing data rate translates into increasing sensitivity for geodetic VLBI, the promise of very high data rates for e-VLBI is very enticing. This is particularly attractive and important if observations are done with small antennas. Global networks operating at 10 Gbps/wavelength/fiber are standard commercial practice today, with capability of up to ~100 wavelengths on a single fiber for an aggregate of ~1 Tb/sec/fiber.
- Rapid turnaround for results: e-VLBI, particularly real-time or near-real-time e-VLBI, has the potential for reducing the correlation processing delay to near zero and producing the final analyzed geodetic results in hours. This will help to meet the turnaround requirements specified in the IVS Working Group 2 for IVS Product Specification and Observing Programs.
- Lower or eliminate media costs: This is obvious, of course. In Table 1, each of the succeeding scenarios requires less recording of data at a station and hence a reduction of media cost, down to zero in the full 'real-time data transfer' scenario.
- Fully automated operations: The last three scenarios in Table 1 require no shipping of media and therefore allow for completely unattended operation of a station, helping to lower operational costs.
- Monitor station health: With the capability to transmit data rapidly to the correlator, it is possible to monitor the overall health on a comprehensive basis, as well as assist in the remote diagnosis of problems for assistance in repair.

- Enables distributed correlation: Though not practical today, distributed correlation using arrays of commodity PCs, perhaps geographically dispersed, is a possibility for the 2010 era. With e-VLBI, the data are easily dispersed to many geographically distributed sites. Experiments in distributed software correlation are now taking place in Japan, and though the performance is not now sufficient to accommodate the overall needs of geodetic VLBI, the continued improvement in commodity PC computing power may change the situation by 2010.

Disadvantages

- Last-mile connectivity: Though high-speed global networks are now a reality, the 'last-mile' connection to stations is a potentially expensive issue that must be faced before one can move to full real-time capability.

There is no doubt that e-VLBI, particularly real-time or near-real-time, offers many advantages. One-time last-mile connectivity costs and recurring fiber leasing costs are the primary issues to be resolved. Clearly, co-location of several space-geodetic techniques at each station is advantageous to help reduce these costs.

The cost of fiber data connections can be kept to a minimum by ownership or lease of dark fiber or dark wavelengths on fiber; in these cases, the owner is also responsible for terminal equipment, but is free to upgrade or replace this equipment with no additional costs associated with the fiber itself. In general, total communications-infrastructure costs are much lower for this approach compared to leasing configured services from a service provider.

7. Questions Still to be Answered

There are still a number of questions to be answered and issues to analyze to complete the VLBI2010 analysis of the data transport issue. Among such questions:1) Cost estimates for last-mile connections. 2) What are the technology options for the last-mile terminal equipment? 3) Cost-benefit analysis of various options in Table 1. 4) How can stations such as O'Higgins, which cannot be connected by fiber, be managed (satellite?)? 5) Are current VSI (H,S and E) specifications [Ref 6] adequate?

References

[1] Osaki, H, Advancement of Versatile Scientific Sampling Processor, CRL/Technology Development Center News, Serial No.21, Nov 2002; available http://www2.crl.go.jp/ka/radioastro/tdc/news_21/pdf/tdcnews_21.pdf.

[2] Mark 5 information available at http://web.haystack.mit.edu/mark5/Mark5.htm.

[3] Yen, J. L., K. I. Kellerman, B. Rayher. N. W. Broten, D. N. Fort, S. H. Knowles, W. B. Waltman, and G. W. Swenson, Jr., Real-Time, Very Long Baseline Interferometry Based on the Use of a Communications Satellite, Science, 198, 289-291, 1977.

[4] Levine, J. and A. Whitney, Mark III Real-Time Fringe-Detection System, Radio Interferometry Techniques for Geodesy, NASA Conference Publication 2115, 1979.

[5] Yoshino, T. et al, Keystone Project and VLBI real-time data processing, Proceeding of 4th APT Workshop, Sydney, Australia, 1995.

[6] VLBI Standard Hardware Interface information available at http://web.haystack.edu/vsi/index.html.