Mark 5 and e-VLBI Status Update

Alan Whitney
MIT Haystack Observatory

Mark 5A Data System

- Direct plug-compatible replacement for 64-track Mark4 or VLBA tape drives
- Record/Playback at rates up to 1024 Mbps
- Two independent ‘8-pack’ disk modules per system can be used in ‘ping-pong’ fashion
- Will record 8, 16, 32 or 64 tracks from Mark4 formatter (1024 Mbps max) or VLBA formatter (512 Mbps max)
- Parity bits are stripped before recording; re-inserted on playback

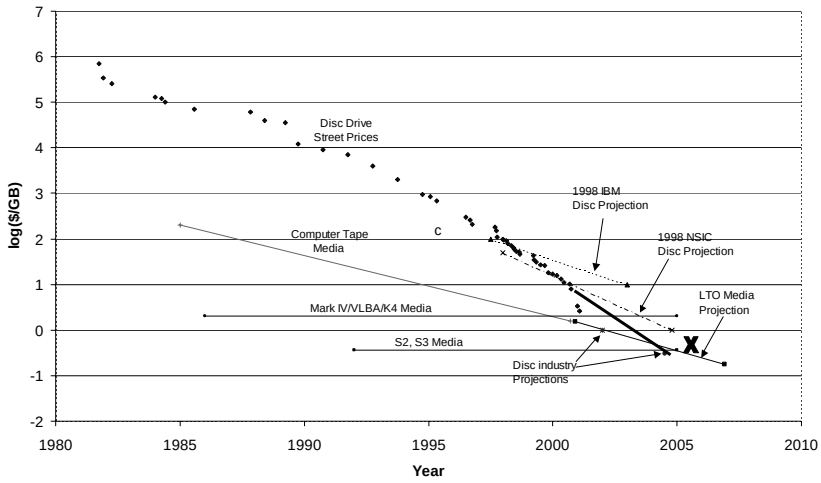
Current Mark 5A Status

- ~100 Mark 5 systems deployed;
  except for VLBA, ~95% displacement of Mark4/VLBA tape systems
- ~1000 Mark 5A '8-pack' disk modules deployed (>1.5 PB of storage!);
  growing rapidly!
- Many stations are now entirely Mark 5A, except for data which goes to VLBA correlator
- 1 Gbps experiments are now routine for both geodesy and astronomy
- VLBA has converted ~3 stations to Mark 5; hopes to complete remainder of stations by end 2005

Disk-Media Status

- Hard disk price vs capacity/performance continues to drop
  - Now approaching ~$0.50/GB and will continue to drop
    (Mark 4/VLBA tape is ~$2.00/GB)
- 250 GB disks now common –
  8-pack of 250GB disks comparable to ~7 VLBA/Mark 4 tapes
- 400 GB disks are available –
  8-pack of 400GB disks comparable to ~11 VLBA/Mark 4 tapes
- 700 GB disks expected by end 2005 –
  8-pack of 700GB disks comparable to ~19 VLBA/Mark 4 tapes;
  ~24 hours @ 1 Gbps unattended!

Tape vs. Disc Price Comparison



Disk-Media Reliability

- Based on statistics collected at Haystack, average disk drive failure rate is ~0.5% per year
- Failure rate of Hitachi 250GB has been higher than expected
- Conduant has qualified drives from Maxtor, WD and Seagate
  - 400GB – Seagate
  - 250GB – Western Digital RAID Edition (high-reliability)
  - 250GB – Maxtor Maxline III (high-reliability)
- Conduant will be shipping these disks in assembled Mark 5 disk modules
- Disk reliability at high altitudes is under investigation

High-Reliability PATA Drives

High Reliability PATA Drives			Hybrid.xls	DLS	16-Apr-05
Manufacturer	W. Digital	Maxtor	Maxtor	Maxtor	Maxtor
	Raid Edition	MaxLine II (5400rpm)	MaxLine Plus II	MaXLine III	MaXLine III
Model	WD2500SB	5A250J0-QV	7Y250P0	7B250R0	7B300R0
Capacity	250	250	250	250	300
Buffer Size	8 MB	2 MB	8 MB	16 MB	16 MB
Shock	250 G	300 G	300 G	300 G	300 G
Altitude	10,000 ft			10,000 ft	10,000 ft
Noise	28 dBA	25 dB	27 dB	27 dB	27 dB
Linear Vibration	0.75G			0.86G	0.86G
Rand. Vibration	0.004 g2/Hz				
Rot. Vibration				12.5 rad/s2	12.5 rad/s2
Start/Stop		>50,000	>50,000	50,000	50,000
Design Life		5 years	5 years		
Warranty	3 years	5 years	5 years	5 years	5 years
MTBF, hours	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
Price	$177.46	$162.48	$169.50	$173.68	$287.80
Price/GB	$0.89	$0.83	$0.86	$0.87	$1.11

Mark 5A Software/Firmware Upgrades

- Improved capabilities to deal with bad or slow disks on both write and read
- Ability to recover from inadvertant operational errors using ‘recover’ command
  - recover data from interrupted recording
  - recover from accidental use of ‘sstest’ or ‘WRSpeed’
  - recover data from accidentally erased module (in latest development version of software)
- Support of disk FIFO mode up to 512Mbps
- Implementation of ‘disk state mask’; not being used outside NRAO
- Improved data-buffer handling for support of faster network transfers

Plans for Serial-ATA Support

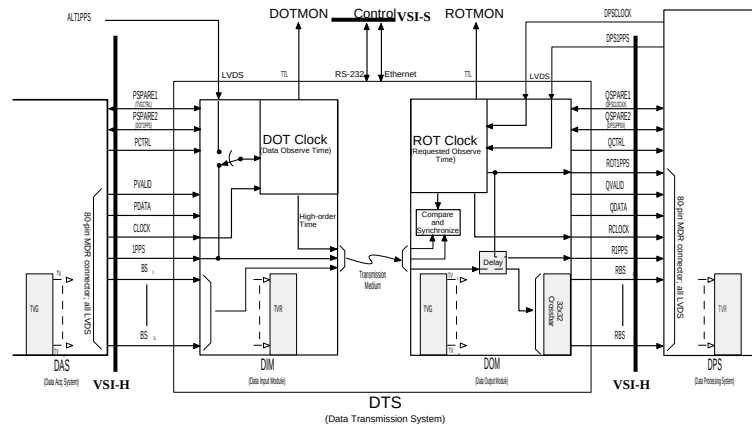
- PATA and SATA disk modules will be interchangeable in Mark 5 systems
- Biggest challenge was finding module SATA connectors with durability
- Requires upgrade kit for current Mark 5 systems - new chassis backplane with separate connectors for PATA and SATA
- New SATA disk modules have been designed
- SATA module should be able to support 2048 Mbps in near future with Mark 5B and NG StreamStor board
- SATA module price will be about same as current module
- Expect SATA modules to be available late 2005

Mark 5B Data System

- Full VSI compatibility
- Same chassis as Mark 5A; uses same disk modules; requires Mark 5B I/O card
- 1024 Mbps record/playback
- Expandable to 2048 Mbps with NG StreamStor card using 64 MHz VSI clock
- Eliminates need for external formatters
- With a 14-BBC Mark4 or VLBA4 system, up to 1792 Mbps can be recorded with two parallel Mark 5B systems
- Requires sampler adapter for Mark 4/VLBA DASs to provide VSI-compatible output [but will not be needed with DBE]
- Station Unit capabilities for connection to Mark 4 correlators is being designed into Mark 5B
- Built-in phase-cal extraction and state counting
- Front-panel status display – 8 tri-color LEDs
- DIM and DOM capabilities are separate FPGA downloads
- FPGA is in-place programmable via software

Development support from Mark 5 development consortium –
BKG, EVN, KVN, JPL, MPI, NASA, NRAO, USNO

Mark 5B supports VSI-H



Notes:
1. Shaded items are for illustrative purposes only.
2. PVALID is optionally transmitted from DIM to DOM.
3. PDATA is optionally transmitted from DIM to DOM.
4. Data delay in DOM is required only for storage-based systems.
5. See text for discussion of use of optional use of P/QSPARE1/2 signals.
6. If DIM/DOM in single box, ALT1PPS/DPSCLOCK/DPS1PPS share single MDR-14 connector.
7. This diagram does not show all functions and options -- see VSI-H specification for details.

Figure 1: VSI-H Functional Block Diagram

VSI6.DRW
ARW 21 Jun 2000

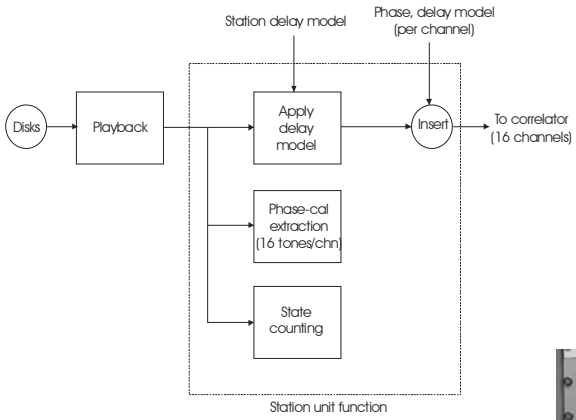
Mark 5B Compatibility

- Mark 5B can play only Mark 5B recordings (VSI format in/out)
- Upgraded Mark 5A will be able to play:
  - All Mark 5A recordings
  - Mark 5B recordings played back in VLBA-track-format mode
- Design of Mark 5A upgrade will start when Mark 5B checkout is complete

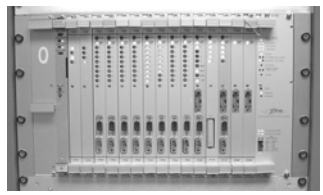
Consequence of Mark 5A upgrade:

Existing Mark 4 correlators with Mark 5A units will be able to process data from both Mark 5A and Mark 5B units during transition period to Mark 5B.

Mark 4 Station Unit emulation



Mark 4 Station Unit



Mark 5B Status

- 3 prototype Mark 5B boards currently under test
- DIM capability ~50% tested; VSI TVG data successfully written to disks; checkout of SU capabilities is slow and tedious
- ~20 Mark 5B interface boards will be distributed to Mark 5 development consortia members – 3Q 2005
- Design will be transferred to Conduant Corp for further replication and sale – mid/late 2005

Mark 5B Interface Boards

- VSI-4 sampler adapter board
  - •Will be placed inside Mark 4 formatter to upgrade to VSI
  - •Uses existing Mark 4 samplers and 1pps generator
  - •2 VSI output connecters at 1024Mbps each (though data rate is restricted to 1792 Mbps by 14 BBC's)
  - •Prototype boards ready for checkout
- VSI-C sampler adapter
  - •Interfaces VLBA samplers to VSI
  - •Designed at Metsahovi
- Correlator Interface Board (CIB)
  - •Interface between Mark 5B and Mark 4 correlator
  - •PCB currently in fabrication
- Upgraded Serial Link boards for Mark 4 correlator
  - •Designed at MPI
  - •Prototypes have been tested; replication to begin soon

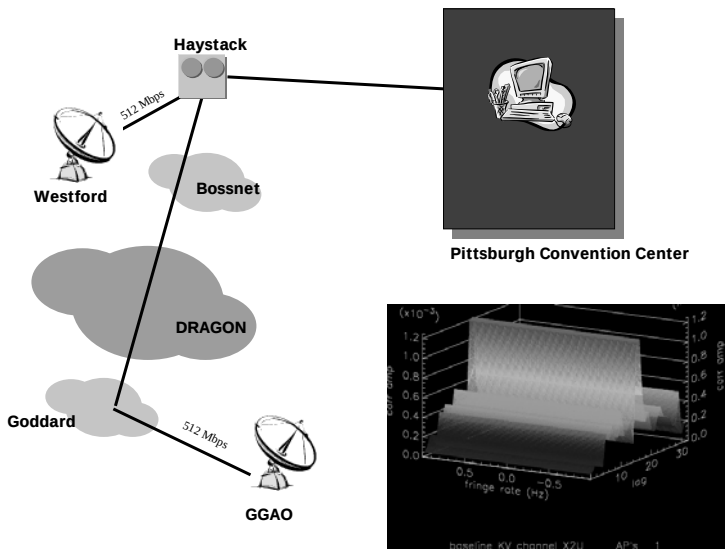
VLBI Standard Interface (VSI)

- Joint effort by astronomy (GVWG) and geodetic (IVS) communities
- **The purpose of VSI is to define a standard interface to and from a VLBI Data Transmission System (DTS) that allows heterogeneous DTS's to be interfaced to both data-acquisition and correlator systems with a minimum of effort**
- Focuses on those functions independent of DTS technology
- VSI-H – complete (Aug 2000)
  - Data and control interfaces for recording and playback, including connectors and pinouts
  - Electrical and timing specs
  - Won national award in Japan in 2001!
- VSI-S – complete (Feb 2003)
  - Communications model
  - Application Protocols
  - Command/Response Syntax
  - Suggested base command set
- VSI-E (e-VLBI) specification nearly complete
- VSI specs available at http://www.haystack.edu/vsi
- VSI is now widely accepted worldwide with a substantial number of VSI-H compliant components available
- Haystack Observatory has led the VSI effort

Recent e-VLBI Developments

- August 2004:
  - Haystack link link upgraded to 2.5 Gbps
  - Real-time fringes at 128 Mbps, Westford and GGAO antennas, Haystack Correlator
- September 2004:
  - Real-time fringes at 512 Mbps (new world record!), Westford and GGAO antennas, Haystack Correlator
- November 2004
  - Real-time e-VLBI demonstration at SC2004
  - Use DRAGON optically-switched light paths
- February 2005
  - Real-time fringes Westford-Onsala at 256Mbps
  - Used optically-switched light paths over part of route
- October 2004 – present
  - Regular transfers from Kashima (~300GB per experiment; ~200 Mbps)
- Starting April 2005
  - First transfer from Tsukuba (~240GB; ~240Mbps)
  - Preparing for CONT05 (15 days continuously at 256 Mbps)

Real-time e-VLBI SC2004 Demo



(Theoretical) Antenna/Correlator Connectivity

(geodetic antennas in red)

- JIVE Correlator (6 x 1 Gbps)
- Haystack (2.5 Gbps)
- Kashima, Japan (1 Gbps)
- Tsukuba, Japan (1 Gbps)
- GGAO, MD (1 Gbps)
- Onsala, Sweden (1 Gbps)
- Torun, Poland (1 Gbps)
- Westerbork, The Netherlands (1 Gbps)
- Westford, MA (2 Gbps)
- Jodrell Bank (1 Gbps?)
- Arecibo, PR (155 Mbps)
- Wettzell, Germany (~30 Mbps)
- Kokee Park, HA (nominally ~30 Mbps, but problems)
- TIGO (~2 Mbps)

In progress:

- Hobart – agreement reached to install high-speed fiber
- NyAlesund – work in progress to provide ~200Mbps link to NASA/GSFC

Challenges

- Network bottlenecks well below advertised rates
- Performance of transport protocols
  - untuned TCP stacks, fundamental limits of regular TCP
- Throughput limitations of COTS hardware
  - Disk-I/O - Network
- Complexity of e-VLBI experiments
  - e-VLBI experiments currently require significant network expertise to conduct
- Time-varying nature of network
- Define standard formats for transfer of data and control information between different VLBI systems
- 'Last-mile' connectivity to telescopes
  - Most telescopes are deliberately placed in remote areas
  - Extensive initiatives in Europe and Japan to connect; U.S. is lagging

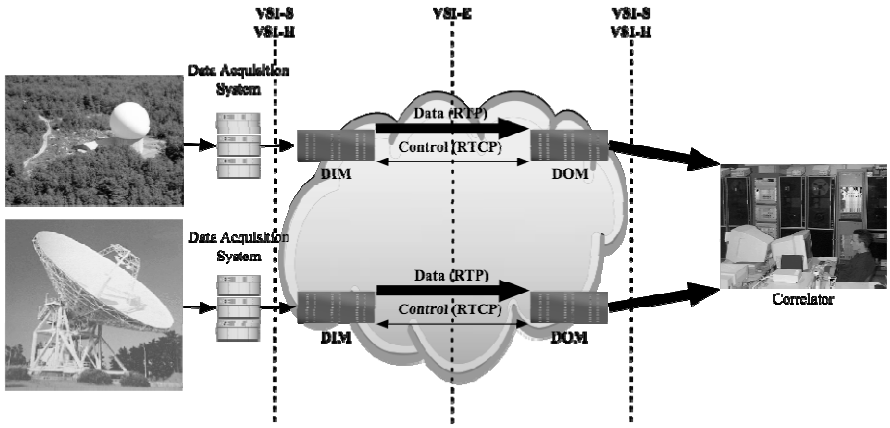
Current Projects at Haystack Observatory

- Standardization
  - VSI-E Draft VSI-E standard distributed in January 2004
  - Reference implementation released in October 2004
- Network interfacing equipment for e-VLBI
  - Mark 5 VLBI data system
- Network Monitoring
  - Evaluation, development and deployment of monitoring systems
- Transport Protocols
  - Testing and evaluation in collaboration with NASA GSFC and Caltech
- Intelligent Applications
  - Automation of e-VLBI transfers an ongoing process
  - Development of optimization-based algorithms for intelligent applications ongoing (EGAE)
  - Intelligent optically-switched networks (DRAGON)
- Production e-VLBI
  - Put e-VLBI into routine use – progressing in limited venues

VSI-E

- Characteristics:
  - Based on standard RTP/RTCP high-level protocols
  - Allows choice of IP transport protocols (TCP-IP, UDP, FAST, etc.)
  - Scalable Implementation; supports up to 100Gbps
  - Ability to transport individual data-channel streams as individual packet streams
  - Ability to make use of multicasting to transport data and/or control information in an efficient manner
- Status
  - Draft VSI-E specification available
  - could be used in the future for support of distributed correlation
  - Prototype VSI-E prototype implementation Nov 2004

VSI-E Architecture



Mark 5 Updates for e-VLBI

- Improved buffer handling for higher-speed transfer between PCI bus and NICs
- Testing of e-VLBI on high-performance motherboards
- Use of dual-GigE connections to achieve 1024Mbps network transfer (to date in lab only)

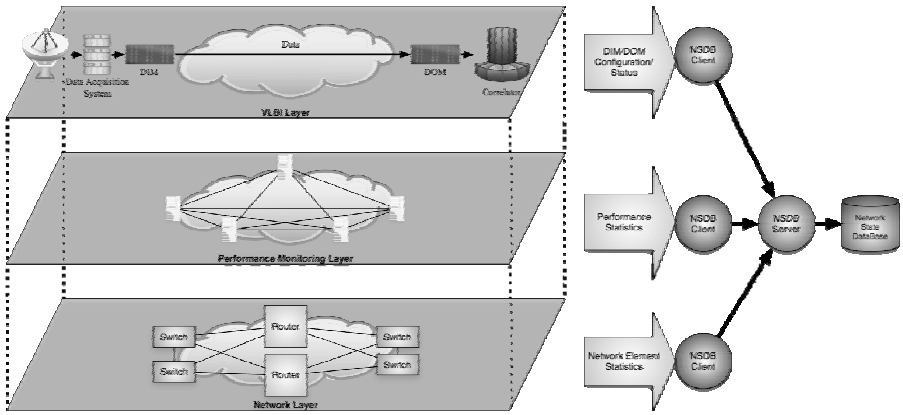
e-VLBI Network Monitoring

- Use of centralized/integrated network monitoring helped to enable identification of bottleneck (hardware fault)
- Automated monitoring allows view of network throughput variation over time
  - Highlights route changes, network outages
- Automated monitoring also helps to highlight any throughput issues at end points:
  - E.g. Network Inteface Card failures, Untuned TCP Stacks
- Integrated monitoring provides overall view of network behavior at a glance

Network State DataBase (NSDB)

- Tool to keep track of state of e-VLBI state:
  - Network performance
  - Configuration of end systems
  - State of end systems
- Integrates and builds on standard monitoring tools to provide a single, coherent view of e-VLBI network state:
  - Maintain continuous state monitoring of entire e-VLBI system
  - Essential for being able to identify issues with network/end system configuration
  - Diagnose at-a-glance (cf. current practice)

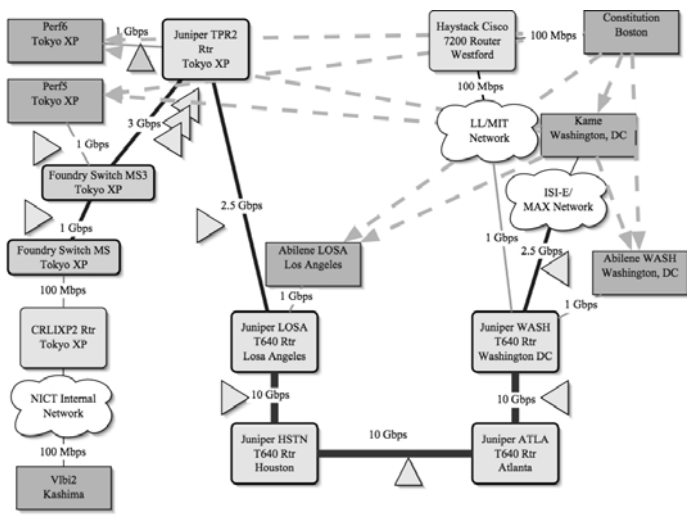
NSDB Architecture



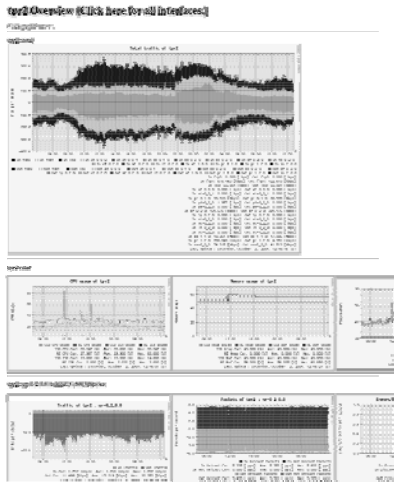
e-VLBI Weather Map Web Page

(Haystack to Kashima)

http://web.haystack.mit.edu/e-vlbi/evlbi.html



Network Layer Statistics



New Application-Layer Protocols for e-VLBI

- Based on observed usage statistics of networks such as Abilene, it is clear there is much unused capacity
- New protocols are being developed which are tailored to e-VLBI characteristics; for example:
  - Can tolerate some loss of data (perhaps 1% or so) in many cases
  - Can tolerate delay in transmission of data in many cases
- 'Experiment-Guided Adaptive Endpoint' (EGAE) strategy being developed at Haystack Observatory under 3-year NSF grant:
  - Will 'scavenge' and use 'secondary' bandwidth
  - 'Less than best effort' service will not interfere with high-priority users

EGAE Progress

- Prototype EGAE software now undergoing shakeout
- Currently ramping up to support routine non-real-time e-VLBI data transfers soon to support regular e-VLBI transfers from Japan, Hawaii, Germany
  - 'Production' e-VLBI facility established at Haystack to support routine transfers
  - Now transfer >1TB per month – expected to ramp up through 2005

Production e-VLBI

- Progress has been slower than we thought it would be
- 'Production' e-VLBI facility established at Haystack to support routine transfers
- Currently, data are being transferred from Kashima and Tsukuba to Haystack on a regular basis
  - Now transfer >1TB per month
  - Hope to support CONT05 in Sep 05
  - Major bottleneck is data-format translation from K5 to Mark 5
- UT1 Intensive transfers (Kokee-Wettzell)
  - Wettzell: regular UT1 Intensive transfers now is place to Haystack
    - Reliably achieve ~30Mbps data rate
    - Destination soon to be moved to Arlington, VA, where data will be physically picked up and carried to USNO for correlation
  - Kokee: many difficulties
    - Highly variable and unreliable network conditions; have on occasion achieved as high as 80 Mbps transfer rate
    - 'Last-mile' connection under control of PMRF; want money to continue

e-VLBI Technical Working Group

- Established at this e-VLBI workshop as group of technical experts, David Lapsley chair
  - On hold until David Lapsley replacement is on-board
  - Hope to re-invigorate at July e-VLBI workshop in Sydney
- Objectives
  - Evaluate e-VLBI/VSI-E hardware/software/procedures
  - Implement standardized global e-VLBI network performance/monitoring tools
  - Provide expert assistance to e-VLBI users
- ~2 members from each major e-VLBI geographical area

Future Directions

- Further EGAE and VSI-E development and deployment
- Improved IP protocols for e-VLBI
- Optically-switched networks for highly provisioned high-data-rate pipes
- Solving 'last mile' problem to U.S. telescopes
- Distributed correlation using clusters and/or highly distributed PC's
- Extending to higher bandwidths
  - Haystack has Astronomy NSF grant to push for 4Gbps/station
  - Preparing NSF proposal to extend to 16Gbps/station using new digital-filter and recording technology
- Continuing to move e-VLBI into routine practice on a global basis