The 5^{th} IVS General Meeting Proceedings, 2008, p. 119–123

The EVN Mark IV Data Processor at JIVE and e–VLBI Developments in the EVN

Robert Campbell, Arpad Szomoru

Joint Institute for VLBI in Europe, Netherlands

Abstract. We review the capabilities of the European VLBI Network (EVN) Mark IV Data Processor at JIVE, briefly touching on operations for EVN/global (astronomical) observations. The greatest strides in the past two years have come in establishing an operational real-time e–VLBI facility. The e–EVN has observed 17 proposal-driven real-time science experiments since 2006, using a 6-station array. We can now reliably sustain 512 Mbps data rates, and we anticipate more stations joining the e–EVN in 2008. Tests since the summer of 2007 have achieved real-time fringes at rates of 920–978 Mbps on 3–5 station arrays over a single Gbps connection via packet dropping, and also real-time fringes from Shanghai and Australian stations correlated at JIVE. The e–VLBI development has included work on a grid-based software correlator currently used for processing (non-real-time) the ftp fringe-tests in EVN sessions.

1. Current Capabilities

The EVN Mark IV Data Processor can correlate simultaneously up to 16 stations with 16 channels (8 dual-pol subbands (SB)) per station, each having a maximum sampling rate of 32 Msamples/s, for a maximum of 1 Gbps per station. We can currently correlate/provide: (*i*) Mark 5A or Mark 5B recordings (5B playback via 5A+) in either Mark4 or VLBA format with 1- or 2-bit sampling; (*ii*) full Stokes polarization output; (*iii*) up to 2048 frequency points per baseline/SB/pol; (*iv*) full-correlator integration times $t_{\rm int}$ down to 0.25 s; (*v*) oversampling at 2 or 4 times the Nyquist frequency, providing $BW_{\rm SB}$ down to 500 kHz; (*vi*) multi-pass correlation, (experiments requiring enhanced spectral-resolution per SB, or having $N_{\rm sta} > 16$); and (*vii*) real-time e–VLBI operation. Development of recirculation and Mark 5B playback is underway.

The total correlator capacity can be expressed as:

$$N_{\rm sta}^2 \cdot N_{\rm SB} \cdot N_{\rm pol} \cdot N_{\rm frq} \leq 131072\,. \quad (1)$$

Here, $N_{\rm frq}$ is the number of frequency points per baseline/SB/pol, $N_{\rm SB}$ is the number of subbands, and $N_{\rm pol}$ is the number of polarizations to correlate. The value of $N_{\rm sta}$ is "granular" in multiples of 4 (*e.g.,* for 5–8 stations, $N_{\rm sta}$=8).

Independent of equation 1, the maximum number of input channels ($N_{\rm SB}\cdot N_{{\rm pol}_{\|}}$) is 16, and the minimum and maximum $N_{\rm frq}$ are 16 and 2048. You can evaluate whether your experiment adheres to equation (1) interactively with the EVN calculator (via the Users' Guide on the EVN web site www.evlbi.org).

Recirculation time-shares the correlator chips for observations that don't use the maximum $BW_{\rm SB}$ (16 MHz). This effectively increases the correlator capacity for some types of experiments, by multiplying the right-hand side of equation (1) by a factor $\mathcal{R} = 16\,{\rm MHz}/BW_{\rm SB}$ up to a maximum $\mathcal{R}$ of 8. However, the maximum $N_{\rm frq}$ would remain 2048. The most significant benefit of recirculation would be increased spectral resolution for narrow-band global spectral-line observations.

The minimum $t_{\rm int}$ for a configuration using the whole correlator is now 1/4 s, and some configurations that use less than half the correlator can achieve $t_{\rm int}$ of 1/8 s. With recirculation, the minimum $t_{\rm int}$ would increase by a factor of $\mathcal{R}$ from its nominal value. Short integration times, coupled with the spectral resolution afforded by large $N_{\rm frq}$, allow mapping wider fields of view, via reduced bandwidth- and time-smearing in the u-v plane. The EVN calculator evaluates field-of-view limitations due to each of these smearing effects.

The standard distribution product is IDI FITS files (typically for further analysis in AIPS). One potential drawback to wide-field correlations (with the short $t_{\rm int}$ and large $N_{\rm frq}$ they require) is the rapid growth of the size of the output FITS files — reaching about 7–10 GB per correlator pass per hour of observation. The current record for the total size of output FITS files for a single experiment stands at 1028.7 GB.

2. Operations & Data Flow

The Science Operations & Support Group at JIVE provides assistance for all phases of a user's experiment — from proposing/scheduling, through correlation and review of the raw output, to analysis of the resulting FITS data. Refer to the contributions in the most recent EVN symposium [1] or EVGA Working Meeting [2] for further details related to the pre-observation/pre-correlation steps, the post-correlation data review process, the operation and content of the EVN pipeline, and the population and use of the EVN Data Archive.

3. Real-time e–VLBI in the EVN

One of the highlights at JIVE over the past two years has been the funding and commencement of EXPReS (Express Production Real-Time e–VLBI Service). This is a three-year I3 (Integrated Infrastructure Initiative) project funded by the EC (DG-INFSO), with the broad goals of (*i*) enabling an operational real-time e–EVN for astronomical applications and (*ii*) looking beyond the current 1 Gbps recording/correlation environment. The EXPReS web-site www.expres-eu.org provides the most up-to-date information.

By the time of the 4$^{\rm th}$ General Meeting, we had established a 5-station e–EVN array (Westerbork, Jodrell Bank, Cambridge, Onsala, Torun) capable of reliable 128 Mbps operation. In Mar. 2006, Medicina became the sixth station, and in April 2006 we conducted our first successful real-time e–EVN user projects (two experiments spanning 16 h). Tests at the end of May 2006 achieved 256 Mbps from 6 stations. Further work on improving the reliability of the network connections and Mark 5 units at the stations and at JIVE led to the first user e–EVN experiments at 256 Mbps in Oct. 2006. We saw our first real-time e–VLBI fringes from Mh in Mar. 2007 (256 Mbps at K-band). In Jan. 2007, we achieved a sustained 512 Mbps for 1 h on a 5-station array (no Tr); 4 hours of 512 Mbps from all six stations came in Jun. 2007. Efforts in the winter of 2007/8 focused on approaching Gbps rates through packet dropping. In Dec., we maintained 920 Mbps for over an hour on a 5-station array (On did not then have an upgraded mother-board in their Mark 5 unit). By the end of Jan., we could sustain 978 Mbps for over 3 hours on a 3-station array (Mc, On, Tr) by dropping packets such that all track-frame headers remained intact. However, there was no synchronization among the stations in terms of the packet dropping. We are working towards a system in which we can target an entire VEX-channel exclusively for packet dropping.

During the late summer and autumn of 2007 we conducted various tests involving intercontinental data transmission. During the 24th APAN conference in Aug., we obtained real-time fringes on Sh–Mopra and Sh–Western-European baselines at 256 Mbps, with the data from Sh routed overland via Siberia. We also had about 10 minutes of fringes on the Mopra–Darnhall baseline — at 12304 km (NSW, Australia to Cheshire, UK) this may well be the longest real-time e-baseline to date. In Oct., we sustained real-time fringes at 512 Mbps for over 12 h from the Australian stations Parkes–Mopra–ATCA, with data streaming into the EVN correlator at JIVE along three dedicated light-paths across the Pacific. Fig. 1 shows the evolution of data rates and array sizes in e–VLBI science runs and tests through Feb. 2008.

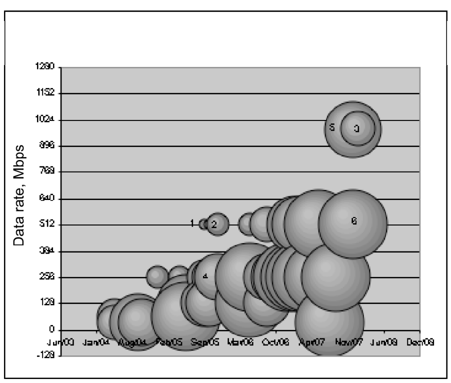

Figure 1. Evolution of e–VLBI data rates and array sizes. The size of the bubbles reflects the number of participating stations

In early Apr. 2008, the 100-m telescope at Effelsberg successfully participated in a real-time test for the first time, producing fringes at 512 Mbps. Inclusion of Ef along with the "original 6" more than doubles the sensitivity of the e–EVN array, clearly making it a much more attractive instrument for astronomers. Also in Apr., we conducted the first user experiments at 512 Mbps (two experiments spanning more than 20 h), including a $12^{h}48^{m}21^{s}$ stretch of continuous correlator output, which established a new record for our longest single job of any sort (e–, disk–, or tape–VLBI).

Connections to the currently participating e–EVN stations are at least 1 Gbps, and all but On and Ef now have dedicated dark fibre or a light-path. Besides the route over Siberia, Sh also has a light-path available across the Pacific at 622 Mbps. Arecibo (PR) has a 512 Mbps connection into Florida. See the EXPReS wiki (www.jive.nl/dokuwiki/doku.php?id=expres:expres) under NA4 for a network diagram showing how these and other stations route their data into the EVN correlator at JIVE.

Coming into the EVN correlator, we currently have a 5 Gbps IP-switched connection to SURFNET (the Dutch NREN), plus eight 1 Gbps light-paths, and 1 Gbps dark-fibre from Wb. To provide more control and independence, we also now have our own central switch and IP space. We have upgraded our Mark 5 units by replacing mother boards (dual 3.2 GHz Xeon CPUs, 2 GB of RAM), and boosting the power supply to 650 W. We are also migrating to DEBIAN linux for the OS. The increased processing capacity permits running modified Mark 5A code that incorporates use of UDP transfer protocols in conjunction with targeted packet dropping tactics, enabling real-time data rates above 512 Mbps.

3.1. Real-time e–VLBI Astronomy

The EVN has offered proposal-driven real-time e–VLBI observations since 2006. There have been 17 successful observations from 28 proposals. The key distinction of a real-time e–EVN experiment is the extremely reduced latency in receiving correlation results; final FITS files are available within a day of the end of observations. For sources that vary on short time-scales — flaring X-ray binaries, gamma-ray bursts, just-exploded supernovae — such short-latency high-resolution VLBI results are vital to adapting observing tactics based on the source's behavior.

e–EVN observing sessions are currently scheduled for 24 h on pre-arranged dates, about once per month. There have been paired dates only 2–3 days apart to permit following source changes on shorter time-scales. A new class of "triggered" proposal enables e–EVN observations of a pre-selected source when its behavior prior to an e–EVN session shows that it has entered an interesting state. The proposals themselves are due at the usual thrice-annual EVN deadlines, but the trigger request can be made up to 24 h before the start of an e–EVN session. The more frequent e–EVN sessions are also beneficial for astrometric monitoring. Work continues towards a more fully dynamic, on-demand observing capability at non-scheduled times.

4. Software Correlation at JIVE

We have been inserting ftp scans into the standard monitoring experiments during EVN sessions for some time. We originally processed those ftp scans using the NICT Software Correlator fx_cor [3], running on an 8-node dual-processor PC cluster, where independent baselines are assigned to different nodes/processors. An automatic-ftp feature added to the Field System has been exercised since the Nov. 2006 EVN sessions. This automatically copies a specified portion of a scan from the Mark 5 disk to a linux file and ftp's it directly to the cluster at JIVE, where the arrival of new data is detected, correlation performed, and results posted to a web page available to the stations. NICT has provided responsive support to enable correlation of new recording modes that they would not normally handle. The ftp fringe tests themselves have been very successful in identifying problems early enough to allow stations to repair them before user experiments would have been affected.

We now process the ftp fringe tests on the software correlator we are developing as part of the EXPReS. This addresses workflow management and correlation algorithms that can be applied to a real-time GRID-based correlator. A workflow manager application prototype, to route data to cluster computers such that the network keeps up with the incoming data, has been developed at Poznań Supercomputing and Networking Center. We are adapting the correlator algorithms developed for observing the descent of the Huygens probe onto Titan [4] to wide-field applications. The resulting XF software correlator has handled the EVN ftp fringe tests since June 2007. It currently runs on a single node, quad-processor CPU, and can correlate 1 second of data from an 8-station array with 8×8 MHz SBs, providing four polarizations per SB and 1024 frequency points per baseline/SB/pol, in $\sim$100 s.

References

[1] Campbell, R.M. Developments at the EVN Mk IV Data Processor at JIVE. Proceedings of the 8th EVN Symposium. A.R. Marecki (ed.), 2007, 396.

[2] Campbell, R.M., A. Szomoru. Recent Developments at the EVN Mk IV Data Processor at JIVE. Proceedings of the 18th European VLBI for Geodesy and Astrometry Working Meeting. J. Böhm, A. Pany, H. Schuh (eds.), 2007, 15.

[3] Kondo, T., Y. Koyama, H. Takeuchi, et al. Current Status of the K5 Software Correlator. IVS NICT-TDC News, v. 25, 23, 2004.

[4] Pogrebenko, S.V., L.I. Gurvits, R.M. Campbell, et al. VLBI Tracking of the Huygens Probe in the Atmosphere of Titan. Proc. of Planetary Probe Atmospheric Entry and Descent Trajectory Analysis & Science. D. Atkinson, A. Suleman (eds.), C-4.6, 2003.