Geodetic VLBI Correlation at the Vienna Scientific Cluster

J. Gruber, J. Böhm, J. McCallum

Abstract Geodetic VLBI correlation is a new challenge in the current activities at the research area Advanced Geodesy at Technische Universität Wien (TU Wien). We have implemented the Distributed FX (DiFX) software correlator and Haystack Observatory Postprocessing System (HOPS) on the Vienna Scientific Cluster 3 (VSC-3), which is a supercomputer located at TU Wien. We provide information about VLBI correlation-related activities in Vienna and we present the VSC-3 by showing some technical aspects of this high performance computer system. We have carried out a performance test to show the most efficient processing setup for DiFX on the VSC-3 and results of successfully correlated data, which includes DiFX processing and HOPS processing. Finally, we list some future challenges and tasks and comment on VGOS correlation at the VSC-3.

Keywords VLBI Correlation, Fringe Fitting, VSC, VieVS, VGOS

1 Introduction

VLBI correlation is referred to the process which determines the difference in arrival times at two stations by comparing the recorded bit streams. After the signals at each antenna site are collected and recorded, VLBI correlation initiates the analysis of a VLBI experiment. Basically a VLBI experiment can be split up into a work stream of several tasks. First the VLBI observations have to be scheduled. On the basis of this schedule the antennas observe data. The recorded bit streams are correlated and then fringe fitted to generate the fundamental VLBI observable which is used in the VLBI analysis to estimate station positions, quasar positions, Earth Orientation Parameters (EOPs) and other important geodetic parameters. So far, at our institute we are capable to schedule VLBI observations and to estimate geodetic parameters with our homemade Vienna VLBI and Satellite Software VieVS (Böhm et al., 2012). With the installation of a VLBI correlation infrastructure at our institute we have established an additional element in the VLBI processing chain of scheduling, observing, correlating and analyzing. Feedback between scheduling and correlation and analysis will help us to improve and refine the individual tasks. Furthermore, with own correlation infrastructure we are more independent to process VLBI projects and we can contribute with a wider field of capabilities to the VLBI community.

With the development of the Distributed FX-style Correlator (DiFX) (Deller, 2007), many hardware correlators have been replaced. The software is designed to run in a multi core environment parallelization to process the large amount of input data. We have installed DiFX on the Vienna Scientific Cluster (VSC) to realize VLBI software correlation at our institute.

Jakob Gruber · Johannes Böhm
TU Wien, Karlsplatz 13, AT-1040 Wien, Austria

Jamie McCallum
UTAS, Churchill Ave, Hobart TAS 7005, Australia

2 Correlation and Fringe Fitting capabilities of the VSC-3

The Vienna Scientific Cluster (VSC) is a supercomputer located in Vienna at the Arsenal TU Wien building. It consists of several cluster systems that have been designed to satisfy the demand for High Performance Computing (HPC) of a consortium of Austrian universities. The latest system is the VSC-3 which was installed in summer 2014 (see Fig. 1). At this time it was ranked 85th in the TOP500 list. The hardware capabilities which are intended for processing large amount of data and for intensive i/o workloads make the VSC-3 to an appropriate multiprocessor computing environment for the correlation of the large amount of VLBI data.

Hardware capabilities of the VSC-3

The VSC-3 was installed by ClusterVision and consists of 2020 nodes, each equipped with 2 processor. The processors are Intel Xeon E5-2650v2 processors from the Ivy Bridge-EP family with 8x2.60 GHz and 20 MB SmartCache. Details of the configuration of such a compute node are given in Tab. 1. The nodes are internally connected with an Intel QDR-80 dual-link high-speed InfiniBand fabric which features very high throughput and very low latency. Energy-efficient cooling is provided by the mineral-oil based CarnotJet System of "Green Revolution Cooling". The high-performance BeeGFS parallel Filesystem, developed for intensive i/o workloads provides the data storage facility on the VSC-3. To specify which nodes have to be used in the correlation process we make use of the SLURM workload manager which organizes the scheduled processing jobs.

Table 1: Configuration table of the compute nodes of the VSC-3.

Motherboard	Supermicro X9DRD-iF
	Intel Patsburg Chipset
	QuickPathInterconnect (QPI) 8.0GT/s
	Dual Xeon Sandybridge (E5 Series)
	Up to 256GB DDR3 1600/1333/1066/800MHz
	Slots: 1 (x16) PCI-E 3.0 and 4 (x8) PCI-E 3.0 slots
	Intel®350 Dual-Port Gigabit Ethernet Controller
	8x SATA2 and 2x SATA3 ports
	Integrated IPMI 2.0 with Dedicated LAN
	Supermicro RSC-RR1U-E8 1U PCI-E Riser Rad
Chassis	SNK-P0047P passive 1U heat sink
	X9 Generation Motherboard
	Indium Foil replaces heatsink paste
	1U PowerSupply 350W
CPU	2 x Intel Xeon IvyBridge-EP E5-2650v2
	2.60GHz
	8 Core - 20MB Cache
	Intel HT Technology - Intel Turbo Boost Technology
	95W TDP (Thermal Design Power)
Memory	8 x 8192MB DDRIII1866 ECC Registered (512Mx8)

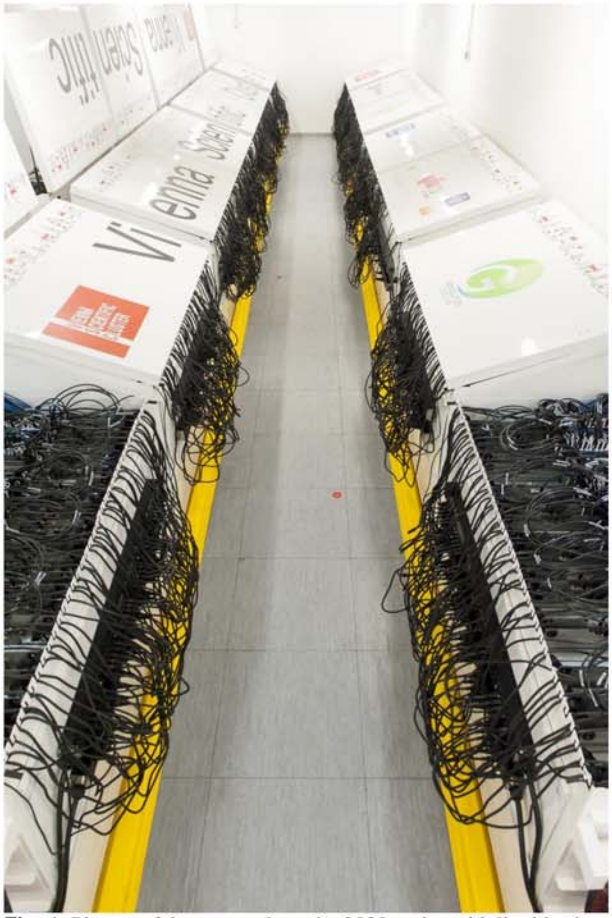

Fig. 1: Picture of the room where the 2020 nodes with liquid submersion cooling of the VSC-3 are stored. Picture taken at VSC Arsenal TU Wien building ©Claudia Blaas-Schenner.

Software correlation capabilities of the VSC-3

We have installed the Distributed FX (DiFX) software correlator version 2.4 on the VSC-3. It was developed at the Swinburne University in Melbourne, Australia by Adam Deller et al. (Deller, 2007). It is a software correlator in FX-style written in C++ and the correlation algorithms are intended to run in multiprocessor computing environments. The fundamental operations performed by the DiFX correlator can be listed as follows: streaming of digitized signals, application of the correlator model, padding the data from 2 bits to 16 bits, alignment of data within +/- 1 sample, performing an FFT, fractional-sample delay correction, complex multiplication and integration, writing output as complex visibilities. In the software correlator architecture data are loaded by data-stream nodes (one per station). These nodes are directed by a master node (FxManager) to send data to the processing elements (core nodes). The processed data are sent to the FxManager node for storage to disk.

After the generation of complex visibilities fringe fitting has to be carried out to obtain the geodetic relevant multi-band delay. This can be achieved by the Haystack Observatory Postprocessing System (HOPS). It is a software package written in C which performs basic fringe fitting, data editing, problem diagnosis, and correlator support functions. HOPS is not specifically intended to run in a multiprocessor computing environment but to have the whole correlation process chain ready at one place we have installed and integrated HOPS in our working environment on the VSC-3.

3 DiFX performance test and efficiency analysis

To evaluate the most efficient processing configuration of DiFX on the VSC-3 we have carried out performance tests. The result will help us to get a clearer understanding of the correlation processing speed on the VSC-3 of VLBI data and it will provide us information to predict the processing time for future sessions. With respect to a shared correlation network of several correlation centers we can also show our correlation capabilities to the VLBI community.

Methodology

In contrast to common benchmarking approaches of DiFX, in which fake eVLBI data streams are used to eliminate disk i/o as a bottle neck, we see our performance test in a more practical way. We want to test how DiFX performs in our working environment on the VSC-3 with real data and in a real processing environment. We do not avoid i/o issues because we want to determine the impact of the communication between the processing nodes and the hard drive. This means, i/o issues are part of this performance test, because they are part of processing real VLBI data. With this test it is difficult to ascertain the true performance of the DiFX source code but we will obtain information about processing performance of a real VLBI experiment. For more details on a benchmarking methodology and benchmarking results, see Phillips (2009).

The DiFX processing consists of several programs of the DiFX software package. A shell script is used to start DiFX and to run all the required functions. There are head operations which have to be executed just once in beginning of the whole DiFX processing, such as the generation of the machine and threads file and the generation of the .input and .calc files. Then there are processes which have to be executed once per scan such as the calcif2 program. All those functions are not intended for multiprocessor environments. Any change of the node number will not influence the processing time of this functions. The core function of the DiFX processing is "mpifxcorr" which actually carries out the correlation task. Mpifxcorr makes heavy use of the Intel Integrated Performance Primitives (IPP) library for optimization and uses MPI (Message Passing Interface) for parallelization. It is designed to run in multiprocessor environments. In this performance test we change the number of nodes to estimate the influence of number of nodes on the processing speed of this function. Furthermore we process the same data with the same cluster configuration at different times to reveal the possible impact of current workload on the shared user storage area.

Test Setup

In this test we used several scans of the session ds317 observed by the AUSTRAL network. The participating stations are Hobart 12 m, Kathereine 12 m, Yarragadee 12 m and Hartrao 15 m with a recording rate of 1 Gbps. The integration time in the DiFX processing configuration was set to 1 s and the spectral resolution of visibilities produced was set to 0.0625 MHz and was not changed for the entire performance test.

Results

The reading and writing speed on the high-performance BeeGFS parallel filesystem on the VSC-3 depends strongly on the current user workload. Our results show that the processing speed varies on average within 15 sec for the same processing setup (see Fig. 2). This means that the processing time of the same scan with the same configuration can fluctuate about 15 sec due to user workload on the storage area. Another impact on processing time is due to number of nodes used in the processing configuration. Using the same scans as used in the test shown in Fig. 2 to evaluate the impact of the number of nodes, we see that the processing time almost halves when changing from 1 to 2 nodes (see Fig. 3). By further increasing number of nodes the processing time stops decreasing at approximately 8 nodes. At this time the DiFX performance becomes data-limited rather than CPU-limited. This means that, at some point (in our case at 8 nodes), obtaining data from the data source (network socket) and transmitting it across the local network to processing nodes will no longer occur quickly enough. The decrease of processing time from 5 to 8 nodes is small in comparison to the decrease from 1 to 5 nodes. On average the decrease in processing time changing from 5 to 8 nodes is about two seconds. We think it would be more efficient to work with 5 nodes and use the remaining three nodes for processing scans in parallel. This strategy can be realized with the SLURM job array, which makes it possible to process several scans in parallel. In this test we executed each scan after another to evaluate the number of nodes at which the DiFX performance becomes data-limited, but in practice we use the SLURM job array to process several scans in parallel. For example if we want to make use of 100 nodes on the VSC-3, we will process 20 scans in parallel. Each scan will be processed by 5 nodes.

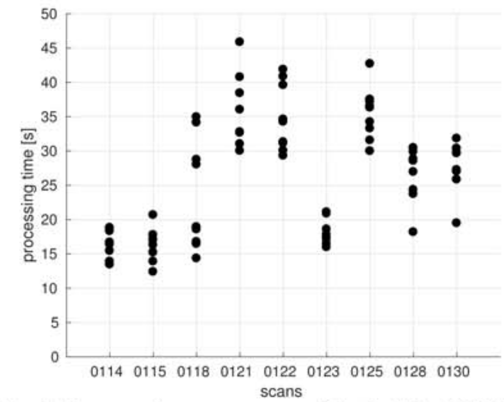

Fig. 2: Impact of current user workload of the VSC-3 storage area on processing time. The dots per scan represent the processing time of several DiFX runs for the same scan with the same processing configuration.

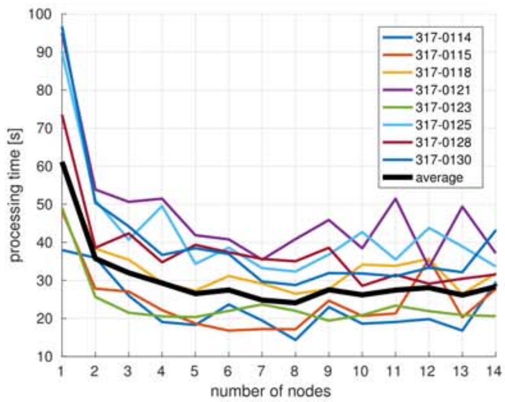

Fig. 3: Impact of number of nodes on the processing time. The colored lines represent change in processing time by changing number of nodes for each scan.

4 Experiments correlated at the VSC-3

Besides some test data we have successfully correlated the dynamic scheduling session ds317, which has been observed on 11/12/2016. We used this data also for the performance test. It was a four-antenna session of the AUSTRAL observing network Hobart 12 m, Katherine 12 m, Yarragadee 12 m and Hartrao 15 m. The session lasted 50 hours and the raw input of recorded data of the VSC-3 DiFX correlator amounts to 22.5 TB. The result of the fringe fitting process are shown in Fig. 4.

5 VGOS correlation in Vienna

The "Observing Plan" of the VLBI Global Observing System (VGOS) shows increasing requirements for the correlation and data transmission of VLBI data, see Petrachenko (2014). Due to an increasing number of observing sites per session and broadband observations more VLBI data will be generated and higher correlation capabilities are required. In view of a distributed correlation strategy, which means that data for a session will be spread to many correlation sites, the VSC-3 would have the resources to process vast amount of VGOS data as part of a shared correlation network. The VSC-3 is connected to the Global Research and Education Network with multiple 10 Gbps which would enable the transmission of recorded data.

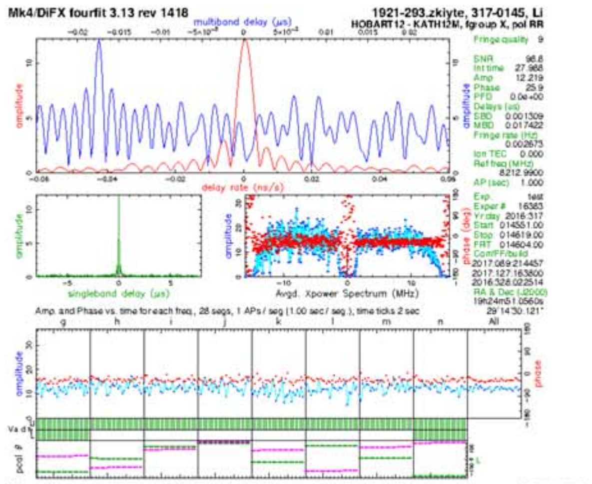

Fig. 4: Results of fringe fitting with the fourfit program of HOPS.

Fig. 5: The Vienna GEO correlator is located in Vienna in Europe and is connected to the Global Research and Education Network with multiple 10 Gbps links.

6 Conclusions and Outlook

We have installed DiFX and HOPS on the VSC-3 which allows us to correlate and fringe fit VLBI data. We determined the most efficient node configuration for a common four station AUSTRAL experiment and we showed the influence of the user workload on the storage area. We have successfully correlated the dynamic scheduling session ds317 and we plan to correlate several sessions per year of the AUSTRAL VLBI array in future. We are working on a refined process chain to process data automatically and we are developing a tool to convert the HOPS/MK4 output database to VGOS database to be able to feed our correlation results into VieVS. This implies the development of a tool in VieVS to detect and correct for ambiguities in the fringe fitted data.

References

Böhm J, Böhm S, Nilsson T, Pany A, Plank L, Spicakova H, Teke K, Schuh H (2012) The New Vienna VLBI Software VieVS. *Springer Berlin Heidelberg*, 1007–1011.

Deller A T, Tingay S J, Bailes M, West C (2007) DiFX: A Software Correlator for Very Long Baseline Interferometry Using Multiprocessor Computing Environments. *PASP*, 119:318–336.

Petrachenko B, Behrend D, Hase H, Ma C, Niell A, Nothnagel A, Zhang X (2014) VGOS Observing Plan. In: *Proc. IVS 2014 General Meeting*.

Phillips C, Deller A T (2009) Performance testing of the DiFX correlator. In: *Proc. 8th e-VLBI Workshop EXPReS09*.