Design of a VGOS Software Correlator Based on GPUs

Voytsekh Ken, Vladimir Mishin, Dmitry Pavlov, Nadezda Sokolova, Igor Surkis

Abstract The design of a new FX software correlator intended for the new small antenna VLBI network was started at IAA RAS in 2012. The main principles, topology, hardware, and software solutions for the design of a six-station, near-real-time correlator with the ability to process 16 Gbps data are presented. Benchmark test results are given. First fringes with wideband observations were obtained.

Keywords Correlator, GPU, VLBI, VGOS

1 Introduction

The design goal is to realize a near-real-time, six-station correlator to process a data stream of up to 16 Gb/s from each observatory. VLBI data are recorded on four frequency bands at a bandwidth of up to 1024 MHz for each circular polarization or up to 512 MHz for two linear polarizations using 2-bit sampling. The input data format is VDIF. The correlator computes cross-spectra at a resolution of 4,096 spectral channels and extracts 32-phase calibration tones in each frequency band for each station in near-real time. A two-station prototype was designed in 2013 [1].

2 Why GPU?

The most significant distinctive feature of our correlator from existing VLBI software correlators is the usage of graphical processing units (GPUs) for the main computations such as bit repacking, Fourier transformation, doppler tracking, spectra multiplication, and phase calibration signal extraction. The choice of the GPUs was dictated by the fact that a GPU contains hundreds of ALUs, which are able to simultaneously operate the same instructions. Besides the global memory with large latency, a GPU also contains shared, texture, and constant memories, which are of cache-type. This allows different threads to access the same data almost simultaneously. Thus, usage of GPUs should sufficiently decrease the amount of the central processing units. A corrrelator design using traditional processors requires approximately 3,000 cores or 375 eight-core CPUs.

Another key feature is that data transformation from bit-type to float-type is perfomed during the GPU's kernel execution in the GPU's DRAM. It allows a decrease in traffic between all modules of the cluster of up to 32 times.

3 High-perfomance Computing Cluster

The correlator's hardware is based on hybrid-blade-server technology. The hardware of the current single-baseline correlator prototype contains five blade servers, which are installed into one chassis (Figure 1). Each blade server contains two Intel Xeon E5 CPUs and two Nvidia Tesla M2090 GPUs based on Fermi technology. The cluster also includes power supply and power distribution units, two cache servers (which are similar to blade servers but having 256 GB RAM and 2x10 GbE optic input), one head (master) server, data storage with a capacity of 20 TB, an infiniband

Institute of Applied Astronomy, Russian Academy of Sciences

data commutator, and a fiber optic commutator. The cluster components are mounted in one cabinet.

Fig. 1 HPC cluster with five blade servers for a two-station correlator prototype at IAA RAS.

In 2014 the cluster will be upgraded to 40-blade servers, eight cache servers, and 75 TB data storage. Also, the existing ten GPUs will be changed to Tesla K20 (Kepler technology). These components require four cabinets.

4 Correlator Topology

The correlator's topology consists of head, station, and correlation software modules. The head module controls all interblock processes and collects the results. Each station module processes a data stream from one station and provides phase calibration signal extraction, data synchronization, delay tracking, and bit repacking. Each correlation module provides cross- and autocorrelation spectra computing for all the stations. Figure 2 illustrates the principle of the station module (SM). SM decodes a VDIF input stream and transfers to delay tracking block.

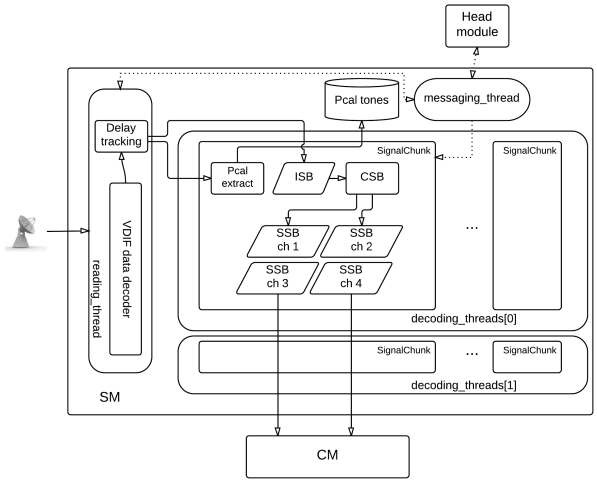

Fig. 2 Station module algorithm.

During tracking, the signal from the station is being transformed to the reference of the Earth center time scale. During the transformation, samples are inserted/removed (Figure 3) into/from the input stream buffer (ISB). The VDIF decoder is continuously being asked for portions of samples to fill the ISB between the inserted/removed samples, or for the whole ISB at once for the case that there are no inserted/removed samples in the ISB. In parallel, the genuine signal (without added or removed samples) is being copied into a separate ISB, which we call "ISB-Genuine".

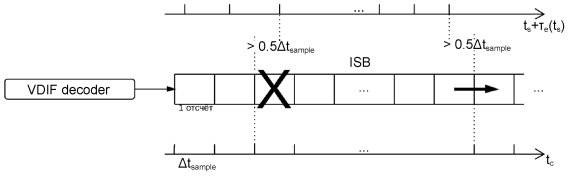

Fig. 3 Delay tracking scheme.

The ISB is separated into channels in the channel separation block (CSB). The CSB essentially repacks bits using the GPU. As a result, the station signal buffers (SSB) are formed and are ready to be transferred to the correlation modules (CM).

ISB-Genuine is used for the pcal extraction operations, which are also perfomed with the GPU. Pcal extraction is implemented using the method proposed by Pogrebenko [2]. If the pcal offset is equal to zero, then the phase tones have an equidistant frequency spacing. The shifted data is divided into frames with a size of doubled pcal tones, and then all the frames are summed. We used a reduction algorithm [3] to accelerate the computation. As a result, white noise from the cosmic data is averaged to zero, and only the phase tone information remains. Then the FFT is run to get the pcal data.

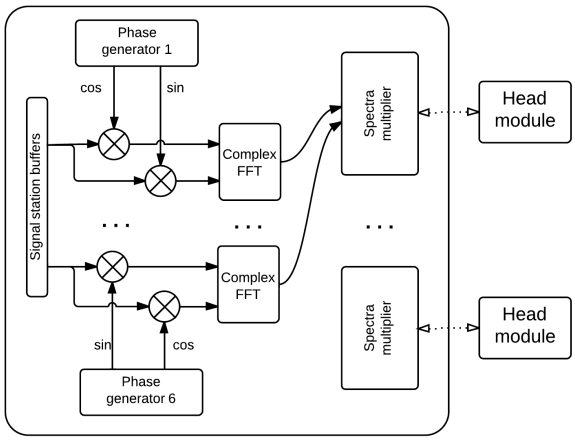

Fig. 4 Correlation module algorithm.

The cross-spectra computation algorithm is shown in Figure 4. First the fringe rotation is produced on the input stream. This operation is perfomed in conjunction with a bit transformation into the float type. Then the FFT operation is completed; and, finally, the obtained data are transferred to the spectra multipliers, where they are multiplied together and averaged within the chosen time period.

When using computation with GPUs, we are able to use several approaches in order to improve the performance by optimization of the operations with memory. Constant memory was used when the precomputed values of the trigonometrical functions for fringe rotation were saved. Different threads had multiple access to these values, so we took sufficient advantage of the memory caching. To sum up the spectra, shared memory was used to decrease the number of requests to global memory.

5 Benchmark Tests

Time and perfomance estimation tests were perfomed in order to check the SM and CM algorithms. The tests were made using Tesla M2090 and Tesla K20x. The pass criterion is to provide the delay tracking and pcal extraction operations in real time for the SM. The input signal is a wideband VDIF data stream written in 2-bit, eight-channel 512-MHz mode.

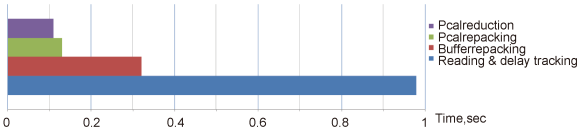

Fig. 5 Station module time benchmark test for one second of a 16 Gbps signal (Tesla M2090).

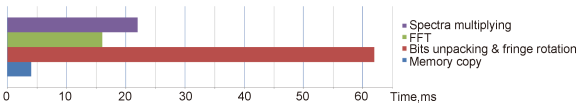

Fig. 6 Correlation module time benchmark test (Tesla M2090).

Figure 5 and Table 1 show the time intervals to perform the required operations for SM. The reading and delay tracking operation as well as other operations are provided by parallel threads so the total time does not exceed the period of one second for one second of data.

Table 1 Required time for station module operations, in seconds.

Operation	Tesla M2090	Tesla K20x
Reading and delay tracking	0.98	0.98
Buffer repacking	0.32	<0.32
Pcal repacking	0.13	0.21
Pcal reduction	0.11	0.19

Benchmark tests were also done for the CM in order to estimate the amount of blade servers needed to

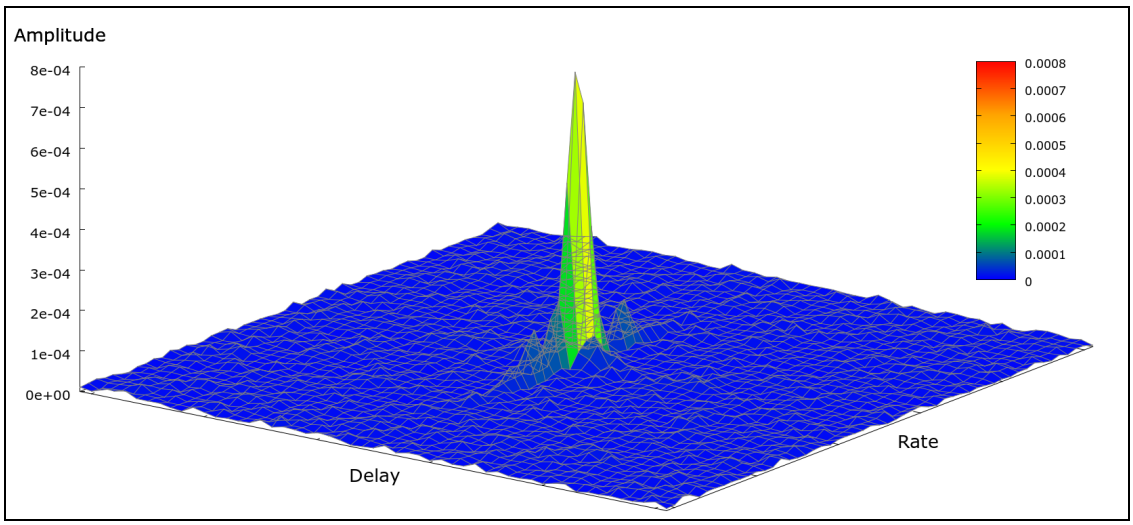

Fig. 7 3D fringe plot for the source 1300+580.

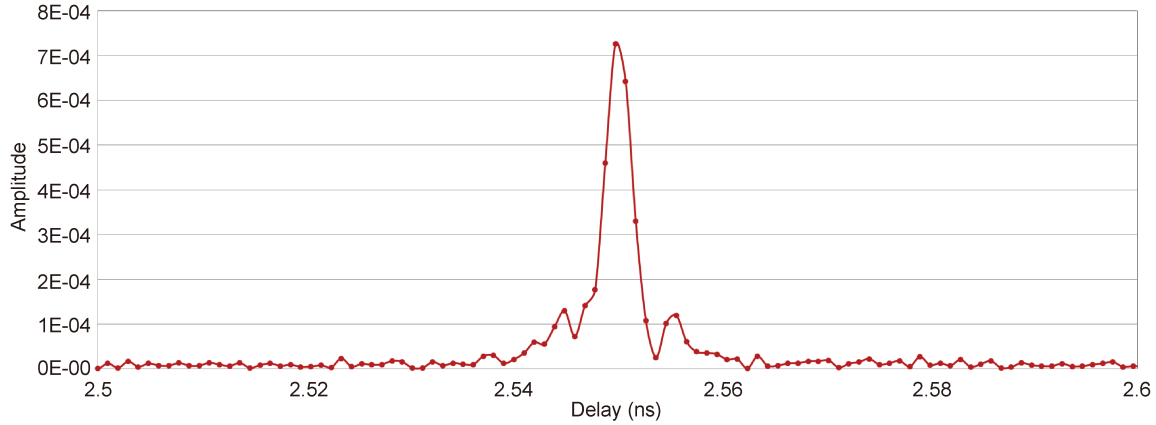

Fig. 8 Fringe plot for the source 1300+580, delay vs. amplitude.

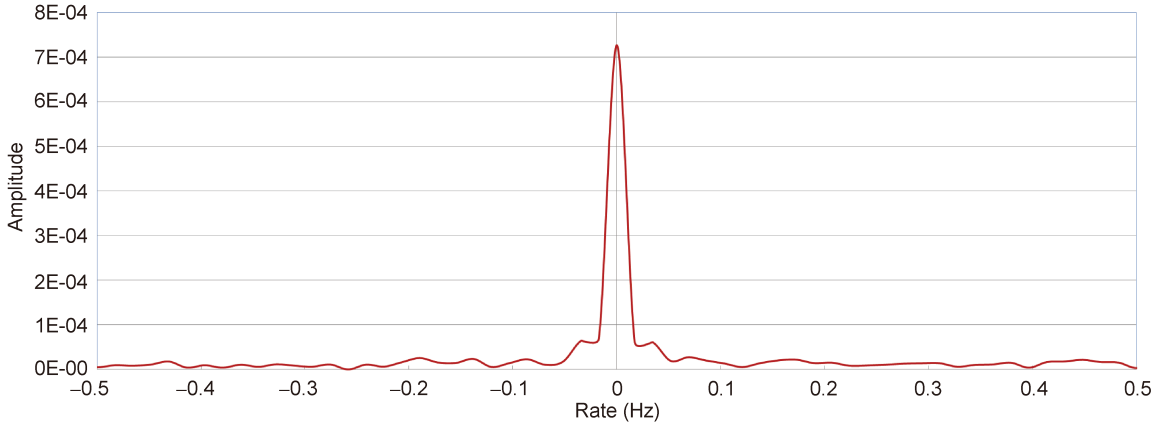

Fig. 9 Fringe plot for the source 1300+580, delay rate vs. amplitude.

process a 16-Gbps data stream in near-real time. Figure 6 illustrates the interval distribution during spectra computation. The bits unpacking and fringe rotation operations are the most intensive, because high-latency DRAM is used. The comparison results between M2090 and K20x are compiled into Table 2. According to the results, the implementation of the algorithms requires eight Fermi M2090 blades or seven Kepler K20x blades for near-real-time processing of one wideband (512 MHz, 2 Gbps) data stream.

Table 2 Required time for correlation module operations, in microseconds.

Operation	Tesla M2090	Tesla K20x
Bits unpacking & fringe rotation	31	23
FFT	8.4	6.5
Spectra multiplication	7.4	6.6

6 First Fringes

A test session, Ru-TEST074, was carried out on September 21, 2012. A scan of the source 1300+580 was chosen to find fringes. The 56-s scan was recorded with the BRoadband Acquisition System (BRAS) prototype on the single baseline Svetloe–Zelenchukskaya. Figures 7–9 illustrate the fringe fitting results. More information about this experiment is presented in [4].

7 Conclusions

The two-station, near-real-time FX-correlator prototype was designed. An HPC cluster was developed and installed at IAA RAS. A six-station correlator is under development. This work will be completed in 2015. All well-parallelized algorithms are perfomed using GPUs. The algorithms require 76 GPUs. Fringes were found for the source 1300+580 with a fringe rate of 1.4 MHz.

References

1. Surkis, I., Ken, V., Kurdubova, Ya., Melnikov, A., Mishin, V., Pavlov, D., Sokolova, N., Shantyr, V., Zimovsky, V. IAA Correlator Center. International VLBI Service for Geodesy and Astrometry 2013 Annual Report. Submitted.
2. Wagner, J., Pogrebenko, S. Fast multi-tone phase calibration signal extraction. http://www.metsahovi.fi/∼jwagner/pcal_extraction_practicals.pdf
3. Harris, M. Optimizing parallel reduction in CUDA. http://developer.download.nvidia.com/assets/cuda/files/reduction.pdf
4. Melnikov, A., et al. First fringes with BRAS on VLBI network “Quasar”. This volume.