An IVS Pilot Study for Distributed Correlation in the VGOS era

A. Bertarini, G. Tuccari, R. Haas, J. McCallum, S. Weston

Abstract In this study we explore the feasibility of distributed correlation as a potential VGOS correlation architecture. Specific challenges are that the amount of data to be transferred to the correlator is going to increase enormously, hence also the correlation time and the capacity of the RAIDs at the correlators to store the data will increase. This study will identify whether distributing the correlation of one experiment among more correlators will help to keep the latency from observation to analysis within a reasonable time; now the guideline set by the IVS from observation to analysis should not be more than two weeks.

Keywords Correlation, VGOS, Geodetic VLBI

1 Introduction

In this study we present the first test made toward the goal of distributed correlation. The general idea of this study is to send segment of data to multiple correlators; one main correlator receives the data at the beginning and the end of the experiment for clock and drift adjustment, prepares the vex and v2d files required for the correlation within DiFX and distributes them to the branch correlators that then correlate their allocated segment of data within the experiment, ideally all using the same DiFX version. After the correlation the branch correlators send the visibilities to the main correlator for fringe fitting and further geodetic processing.

Alessandra Bertarini
Reichert GmbH und Bundesamt für Kartographie und Geodäsie
Bonn, Germany

Gino Tuccari
Istituto di Radioastronomia, Italian National Institute for Astrophysics, Italy

Rüdiger Haas
Chalmers University of Technology, Onsala Space Observatory, SE-439 92 Onsala, Sweden

Jamie McCallum
University of Tasmania, Private Bag 37, 7001 Hobart, Australia

Stuart Weston
Institute for Radio Astronomy & Space Research, AUT University, Auckland, New Zealand

2 Background

The VGOS requirements forcasted for 2020 are summarized in Table 1.

Table 1: VGOS requirements (forecast for 2020) (Petrachenko et al., 2015).

Observation cycle of 1 source	30 s
Data rate	16 Gbps
Session length	24 h/day
Periodicity	7 days/week
Number of sites	24

The expected amount of data per day, that a VGOS correlator will have to handle is about 1000 TB /day hence huge in comparison to the data arriving at the correlator for the legacy system which is about 30 TB/-day for a 10 station 512 Mbps R1 session. Some consequences are as follows.

- The correlators need to buy new RAIDs for storage.
- There is a need of more bandwidth to and from the correlators.
- IVS needs more VGOS correlators.
- Maybe use cloud computing.
- Maybe use distributed correlation.

In this study we begin to explore the distributed correlation.

3 Distributed correlation

For the pilot test presented here, we selected Bonn to be the main correlator and Hobart (Tasmania), Onsala (Sweden) and Workwarth (New Zealand) as branch correlators. We agreed to use the IVS experiment R1.785 because R1.785 was observed at 256 Mbps, and we wanted to start with a low data rate to shorten the electronic data transfer.

Each of the branch correlators received 1 hour of data from the 10 stations involved in R1.785, i.e. Hobart, HartRAO, Ishioka, Katherine, Ny-Ålesund, Onsala, Shanghai, Warkworth, Wettzell and Yarragadee. Using the clock information provided by the main correlator, each branch correlator correlated the assigned hour of data, using identical v2d and vex files provided from the main correlator. After the correlation the branch correlators sent the visibilities to the main correlator for further processing. In parallel, the main correlator correlated all the experiment for regular submission to the data centers. Then the results produced by the branch-correlator processing and the main correlator can be compared and the strategy can be evaluated (still pending).

4 Does distributed correlation work?

We cannot yet give a definitive answer since the project just started and requires more tests, but we can say that

- it works best if all the participating stations transfer their data electronically. If using modules, the stations have to spread the data over multiple modules: one per correlator involved,
- each station should have enough storage in the form of RAIDs or FlexBuff that can be accessed simultaneously by all the correlators,
- the main and branch correlators should have the same DiFX version,
- this study is DiFX-centric, but not all IVS correlators use DiFX. This implies that the main correlator should master all the various file formats used by the various correlator architectures, or the branch correlators should be able to convert the file produced by the main correlator into their control format.
- we are unsure whether this system would work for military-based institutions,
- the correlator reliability drops with the power of the number of the correlators involved. E.g. reliability $\sim \eta^{Ncorr}$, if $\eta = 0.9$ and $N = 5$ then we expect 59 % successful correlation.

5 Future planning

We need to continue with R1.785 analysis to see whether the final results are consistent with those obtained from the correlation performed fully at Bonn. We need to perform more structured tests adjusting the number of correlated hours per branch correlator to be proportional to the number of computing nodes and bandwidth available at the branch correlators. Also we need to check whether the latency from observation to analysis stays within 15 days.

Acknowledgements

We thank S. Casey and R. Hammargren from Onsala, L. La Porta from Bonn correlator, U. Michiko from GSI, C. Plötz from BKG Wettzell, J. Quick from HartRAO, M. Siebert from NMA and B. Xia from Shanghai Astronomical Observatory for their help and assistance during this test.

References

Petrachenko B, Bertarini A, Alef W, Behrend D, Cappallo R, Hase H, Ma C, Niell A, Nothnagel A, Zhang X (2015) VGOS Data Transmission and Correlation Plan. In: K. D. Baver, D. Behrend, K. L. Armstrong (eds.) *International VLBI Service for Geodesy and Astrometry 2014 Annual Report* NASA/TP-2015-217532.