

Minutes EXPR_eS meeting.
April 25th-26th 2007

Attendants: Marcin Okon, PSNC
Dominik Stoklosa, PSNC
Mark Kettenis, JIVE
Nico Kruithof, JIVE
Ruud Oerlemans, JIVE
Yurii Pidopryhora, JIVE
Arpad Szomoru, JIVE
Charles Yun, JIVE
Damien Marchal, UvA

1 Presentations

- Mark opens the meeting and gives an overview of EXPR_eS, FABRIC and SCARIE.
- Marcin gives a presentation on the progress PSNC has made on the Workflow Manager
- Ruud gives a presentation on the current status of the correlator core
- Dominik presents the Workflow Manager
- Nico presents the design of the multi processor software correlator.
- Arpad gave a presentation on the first year review of the EXPR_eS project
- Damien gave a presentation of starplane
 - The supercomputer DAS-3 distributed over 5 locations has switchable light paths and the capability to route them within seconds

2 General

Marcin mentions that the goal of the project is to have an *e*-VLBI experiment with:

- 4 telescopes
- 2-4 hours experiment
- 128 Mb/s

Arpad confirms that this is still the goal.

Convenient telescopes are Torun, Westerborg, Medicina. The last telescope is not decided upon. The experiment will use 3 file servers. Two provided by PSNC and one by JIVE.

3 Grid related

Workflow Manager

The Workflow Manager is a tool which is used:

- to design experiment workflow
- assign file servers and grid resources (correlation)
- operator of an experiment can modify the parameters of the experiment.

The WFM also communicates with the VLBI broker that handles the further processing of the experiment.

- The Workflow Manager will not contain a VEX-parser. It will parse the CCF file, and generate this file using vex2ccf (an utility written in c++ by Ruud).
 - Dominik will investigate two possibilities:
 - integration of the conversion program in the Workflow Manager (Java wrapper for C) or

- a web service that acts as a wrapper around the conversion program.
 - Ruud and Dominik will work on a convenient way to interact with the conversion program, since it currently needs command line input.
- In the Workflow Manager the user will be able to change:
 - the integration time
 - the number of frequency channels
- The icon of a telescope changes depending on the type of input that is used.
- The Workflow Manager will have a text-editor in which the CCF can directly be edited
- The telescopes can be denoted by their two character abbreviation.
- There is no need to receive an acknowledgment in the Workflow Manager from the telescopes before the experiment starts.
- PSNC says that, a monitoring tool for the progress of the correlation is out of the scope of the project for now, but could be added later.

VLBI Grid Broker

The VLBI broker manages the correlation of the experiment. It starts the translation nodes and starts a new correlation job as soon as data from all telescopes for a certain time slice is available. When the correlation process ends, the broker moves all output files to a central place where JIVE can retrieve the output using grid FTP.

- Starting a new correlation job:
 - CCF changes for each coarse time slice
 - filenames of the input and output files
 - start time
 - duration
 - Starting a correlation job takes at least one minute, so smaller time slices are not useful
 - There is no communication possible between the broker and the correlation job: this could speed up the correlation by removing the start up time.
- Deployment of the software correlator on the grid
 - The sources will be compiled on the grid. Necessary libraries have to be installed in advance.
 - Currently used libraries are: fftw, gsl and MPICH2. These libraries do not form a problem.
- The broker will move the output data from a correlation process (a file for each coarse time slice) to a location where JIVE is able to retrieve it from using GSI-FTP. This is the official output. JIVE will analyze the data further.
- PSNC will experiment with different sizes of the coarse time slices

4 Correlator related

Correlator control file

- JIVE will have a close look at the CCF parameters in the coming week and write down a specification of all the parameters with
 - the name
 - a small text that can be used as a caption
 - whether the parameter is possibly subject to change
 - valid and realistic ranges for all the parameters
- JIVE will also specify what inter parameter dependencies are, so that the Workflow Manager

- can check them
- We will continue using key-value pairs and not XML

Delay model

- The delay model will be generated using a pre-installed library (calc-10) on the grid.
- PSNC will try to install the library

Translation nodes

A translation node (also called the “file server”) gets data from the telescopes. It will time slice the data and store each time slice in a file. After a time slice is ready, the translation node sends a message to the VLBI broker. It was also decided to encode some information in the file name of each data chunk.

We need to provide at least: name of the station, start and end time.

- The translation node will be developed by JIVE.
- When the translation node is started by the VLBI broker, it will receive:
 - A start time
 - A maximal length of a coarse time slice in bytes
 - A location to retrieve the data from
 - For a telescope this is an ip-address
 - For data from a mark5: an ip-address and a byte offset
 - From a hard drive: an ip-address and a filename or an URL
 - A directory to put the time slices in
 - The address of the broker, to be able to send a message if a time slice is complete
- PSNC will propose a type and syntax for the messages for
 - starting a translation node and
 - notifying the broker that a coarse time slice is available. This message will contain the filename of the data.

The messages will probably be a soap messages. A useful C library would be gsoap.

- The file servers can either be located at JIVE, the telescopes or at PSNC. It is important that the network bandwidth from the telescopes to the translation nodes is sufficient. Moreover, sufficient data storage should be available before the beginning of the experiment. After the time slices are stored, the process is not time critical anymore.
- The translation node starts the incoming data transfer
- The translation nodes do only time slicing (no channel slicing) because it is easier and faster
- A time slice can be deleted from the file server as soon as the grid has successfully correlated the time slice.

5 Networking issues

- JIVE and the PSNC networking team have different visions on networking solutions for FABRIC. The PSNC networking team proposes to use premium IP. JIVE is wondering whether this is the right solution since the guaranteed bandwidth obtained are much lower than needed for the project. Moreover, the telescopes already have dedicated light paths to JIVE and are probably not willing to pay for other connectivity as well.
- Network information requested by PSNC
 - The JIVE network is currently changing. The requested information is available in a couple of weeks and Mark will send the form back.
 - It is probably possible to install a monitoring tool at the end points (on the Mark5's at the telescopes) to test the networking performance, but it is questionable whether it will be

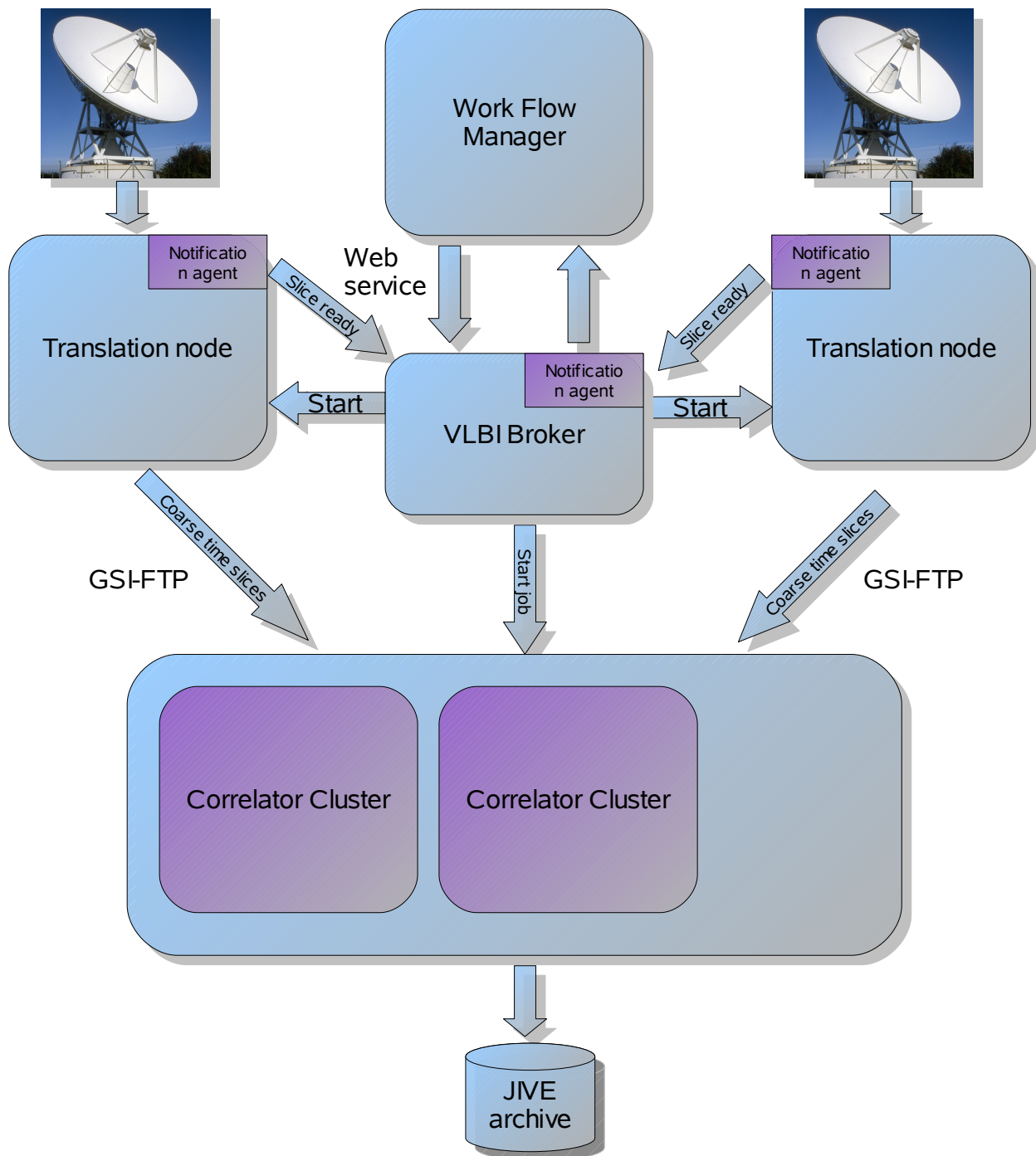
- possible to get a dedicated machine there.
- JIVE will also have some network monitoring tools installed.

6 Time line

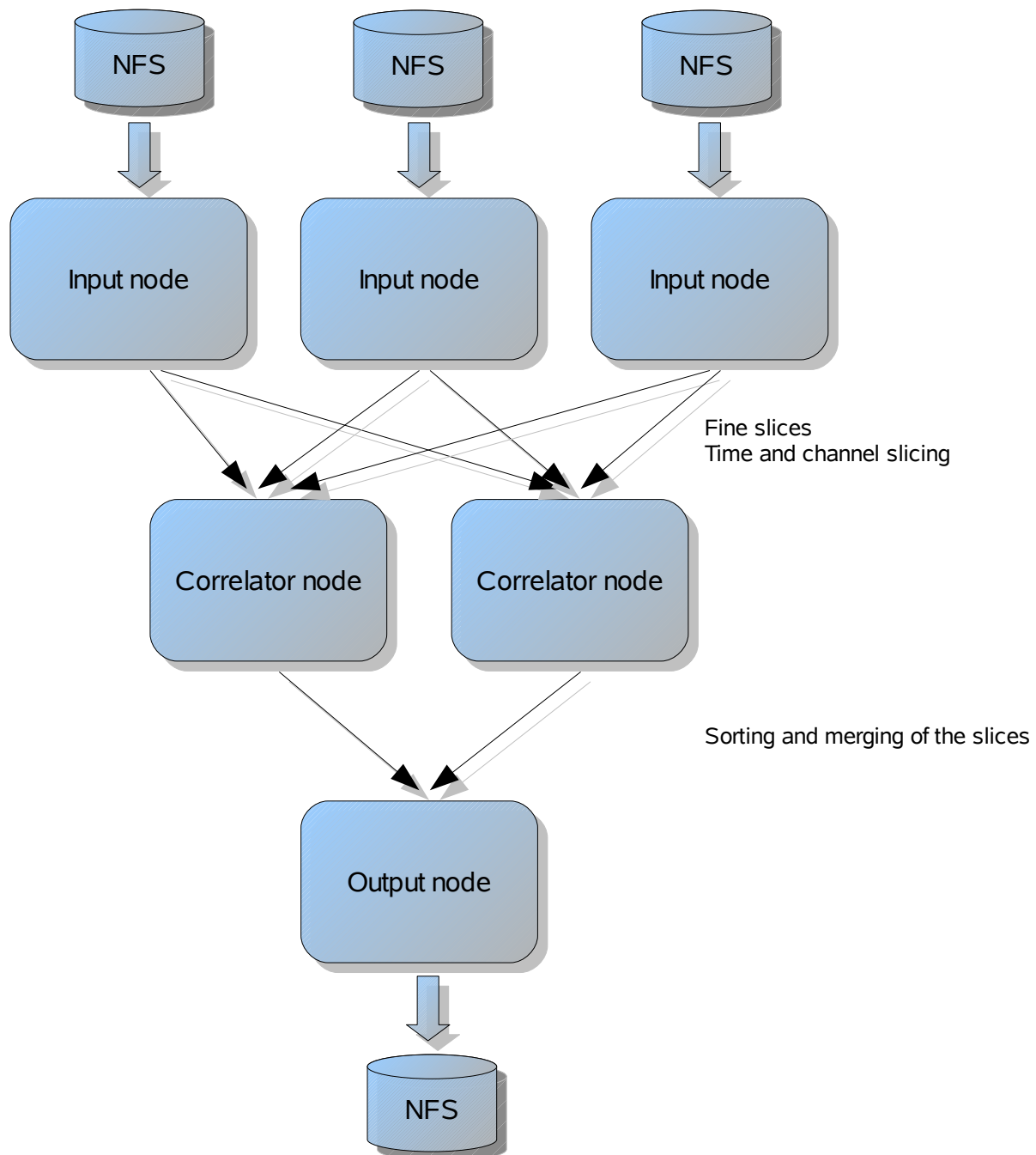
- PSNC is concerned that they are unable to develop and implement their part of the system by the end of August, which was assigned as the deadline. The problem lies in the complexity of the modules that have to be developed and also many interaction points with JIVE part, which will not be completed by August.
- At the end of August, the PSNC networking department will start working for EXPReS.
- For JIVE, the software correlator is not due until the end of december, which makes it hard for PSNC to test their software.
- Charles suggests to mention the delays in the monthly EXPReS reports and to Norbert Meyer.
- JIVE will send a new preliminary software correlator to PSNC. This allows PSNC to do some more benchmarks with the code.

7 Action items

- JIVE
 - list of parameters for the conversion of VEX to CCF
 - list of parameters for CCF
 - send vex2ccf application
 - send delay generation program
 - send a stable version of the software correlator as soon as possible, with large data files to test the parallelization of the code
- PSNC
 - propose an interface for the translation node
 - the networking people at PSNC will contact JIVE
 - test the maximum data rate to retrieve data from NFS (also multiple concurrent data streams)



- The WFM starts the correlation process. The user specifies a VEX file, which the WFM transfers it to a CCF file. WFM notifies VLBI Broker about new experiment, which is responsible for managing the further process.
- The VLBI broker starts the translation nodes, which open a connection to a telescope.
- The translation nodes send a message back to the broker when a coarse time slice is available.
- When a coarse time slice is available from all telescopes, the broker transfers the data to the grid and starts a job for that time slice.
- After the correlation the output of all coarse time slices is transferred to a central location from which JIVE can retrieve the data using grid FTP.



- For each coarse time slice, a new job is started on a cluster.
- The input nodes can access the data through NFS. The input nodes extract the channels. A time slice of one channel is sent to a correlator node.
- The correlator node performs the correlation for the time slice and the channel.
- The output of the correlation is sent to the output nodes, that sorts the output and writes it to file.