Distributed Correlators for Interferometry in Space

Raj Thilak Rajan  
National center for Radio Astronomy,  
ASTRON, Dwingeloo, NL  
rajan@astron.nl  
and  
Circuits and Systems, Faculty of ECMCS,  
TU Delft, Delft, NL

Mark Bentum  
National center for Radio Astronomy,  
ASTRON, Dwingeloo, NL  
bentum@astron.nl  
and  
Faculty of Electrical Engineering,  
University of Twente, Enschede, The Netherlands

Andre Gunst  
National center for Radio Astronomy,  
ASTRON, Dwingeloo, NL  
gunst@astron.nl

Abstract—New and interesting science drivers have triggered a renewed interest in radio astronomy at ultra long wavelengths. However, at longer wavelengths (beyond 10 meters) ground-based radio astronomy is severely limited by Earth’s ionosphere, in addition to man-made interferences and solar flares. An unequivocal solution to the problem is to establish a space-based observatory for ultra low frequency (0.3MHz-30MHz) observations. In ground-based radio astronomy, interferometers comprising of widely spaced antennas are employed to enhance the sensitivity and angular resolution of the observations. The signals received from the antennas are pre-processed, phase corrected independently and then cross correlated with one another using a centralized correlator to estimate the coherence function. However, a space-based array, in addition to several other obstacles, presents new challenges for both communication and processing. In this paper, we discuss various conventional correlator architectures, such as $X_F$, $F_X$ and $HFX$. In addition, the importance of a distributed correlator is emphasized for a space-based array, in particular Frequency distributed correlator. We compute transmission, reception and processing requirements for both centralized and distributed architecture. Finally, as a demonstration, we present 2 projects where these signal processing estimates are applied.

Table of Contents

1 Introduction ............................................. 1  
2 Correlation ........................................... 2  
3 Correlator Architectures .............................. 4  
4 Case Studies ......................................... 5  
5 Conclusion ............................................ 7  
6 Acknowledgements ..................................... 7  
References ............................................. 8  
Biography ............................................... 9

1. Introduction

Radio astronomy research in the last decades has contributed significantly to our understanding of the formation and evolution of the universe. At this moment the entire sky has been observed in detail in almost every band of the electromagnetic spectrum. However, there is still one major exception: the Ultra-Long Wavelength (ULW) radio band. This band is well suited for studying the early cosmos at high hydrogen redshifts, the so-called dark ages, extragalactic surveys, (extra) solar planetary bursts, transient radio sources, high energy particle physics and also in space research such as space weather [1]. While Earth based telescopes such as LOFAR (operating at 30-90 MHz, 120-240MHz [2]) and LWA (operating at 10-88 MHz [3]) cater to low frequencies, the spectrum below 10MHz has not been investigated due to two main reasons. The first reason is the ionosphere. Due to ionospheric scintillation (below 30MHz) and its opaqueness (below 15MHz, depending on the ionospheric conditions), Earth-bound radio astronomy observations in the ULW band are either severely limited in sensitivity and spatial resolution or entirely impossible. The second reason is the high impact of man-made interference and solar flares [4]. The only solution to open the last radio band for astronomy is building a radio telescope in space. Several concept studies and workshops have been started in the past [5], [6], [7], [8], however, until today no real ULW instrument is in operation yet. In the light of emerging technologies, new initiatives such as a radio telescope in space using very small satellites (e.g. nano-satellites). A nano-satellite typically has a mass of 10 Kg or less and is largely constructed from Commercial Off-The-Shelf (COTS) components. A consortium of universities and companies is currently investigating the feasibility of this concept. The future radio telescope has been named Orbiting Low Frequency Array (OLFAR) [9] [10]. OLFAR consist of an aperture synthesis interferometric array implemented with a swarm of nano-satellites ($\geq$ 10), in which each satellite carries one element of the array. The swarm will be deployed in a suitable orbit that provides the radio quietness required for the scientific observations. The satellite swarm concept consists in a system made up of simple autonomous units, which perform small tasks that contribute to the completion of a common system goal. In this paper the distributed data processing will be addressed in detail. In Figure 1 a schematic overview of OLFAR is presented. To limit the datastream to Earth, data processing is performed in space. After gathering the data, the data processing consists of correlation and/or beamforming, and time averaging. Several implementations are possible for each of the digital processing tasks, however in this paper we focus on distributed correlation. In Section 2 various conventional correlator architectures will be studied and in Section 3 the implementation choices of the correlator architecture will be discussed. As an illustration, the implementation of the digital signal processing part in the projects DARIS [11] and OLFAR [9] will be presented in Section 4. We summarize with some conclusions on the study.

\footnotesize
1 978-1-4577-0557-1/12/$26.00 ©2013 IEEE.  
2 IEEEPC Paper #2245, Updated August 12, 2013
2. Correlation

In this paper, the key focus is to calculate three fundamental parameters data in, data out and processing for the DSP unit (Figure 1) for each satellite. The first order processing model estimates processing for correlation as the number of real $N_{\text{bit}} \times N_{\text{bit}}$ multipliers required for each mode, neglecting addition as a second order computation. Memory is another crucial factor, but is dependent on the implementation style and hence will be ignored in this work. The signal processing system in all scenarios is assumed to run on the hardware and hence will be ignored in this work. The measurement of the same plane wave at the two antennae. The number of cross-correlation products increases as $O(N^2)$ for $N$ antennas and the expectation operator is applied over a period of integration time $\tau_{\text{int}}$. There are 3 ways to implement a correlator. The first option is using the traditional correlator model $XF$ i.e. cross correlation first and Fourier transform later. The second alternative is the more recent $FX$ correlator model which measures the cross-power spectrum between two antenna signals. While $XF$ architecture is beneficial because bandwidth can be traded for spectral resolution, $FX$ architecture reduces processing requirements and offers scalability when the number of antennas is large. Finally, a combination of the $XF$ and the $FX$ architecture, yields an $HFX$ correlator, where the data is first broken down into sub-bands and then each sub-band is analyzed by an $XF$ correlator as complex multiplication (17), Fast Fourier Transform (FFT) (19) and Poly-phase Filter Bank (PFB) (20) are calculated in Appendix and will be invoked often.

### Correlation

Radio astronomers calculate the Fourier transform of the measured coherence function to make maps of the sky. Let $x_i(t)$ and $x_j(t)$ be two time varying signals received at spatial positions labeled $i$ and $j$, then the coherence function $\zeta_{ij}(\tau)$ is the cross correlation product between and is given as

$$\zeta_{ij}(\tau_{\text{int}}) = \langle x_i(t) x_j^*(t-(\tau_{ij} + \tau_{\text{int}})) \rangle$$

where $\langle \rangle$ is the expectation operator, the superscript $(*)$ indicates conjugation and $\tau_{ij}$ is the light travel time between observation of the same plane wave at the two antennae. The number of cross-correlation products increases as $O(N^2)$ for $N$ antennas and the expectation operator is applied over a period of integration time $\tau_{\text{int}}$. There are 3 ways to implement a correlator. The first option is using the traditional correlator model $XF$ i.e. cross correlation first and Fourier transform later. The second alternative is the more recent $FX$ correlator which measures the cross-power spectrum between two antenna signals. While $XF$ architecture is beneficial because bandwidth can be traded for spectral resolution, $FX$ architecture reduces processing requirements and offers scalability when the number of antennas is large. Finally, a combination of the $XF$ and the $FX$ architecture, yields an $HFX$ correlator, where the data is first broken down into sub-bands and then each sub-band is analyzed by an $XF$ correlator.

### XF Correlator

The conventional method to directly measure the cross-correlation function as indicated in (1) which forms the basis for the $XF$ or ‘Lag’ correlator [13]. The signal $x_j(t)$ is delayed and correlated (X) with $x_i(t)$ to produce $\zeta_{ij}$ as a function of $\tau$, which is later Fourier transformed (F) to produce the baseline cross power spectra. The cross correlation operation is a function of lags $(\tau)$, which could be varied in quantization of $n\Delta \tau$, where $-N_{\text{lags}}/2 \leq n < N_{\text{lags}}/2$, where for a given observation bandwidth $\Delta f_{\text{o}}$, $\Delta \tau$ is limited by $\Delta \tau_{\text{min}} = 1/(2 \times \Delta f_{\text{o}})$. Consider a single slice $X F_{ij}$ catering to baseline between antennae $\{i, j\}$ and containing $N = N_{\text{lag}}$ correlation segments as shown in Figure 2. The input signal $x_i$ and the phase compensated $x_j$ are the inputs to the correlator slice $X F_{ij}$. The signal $x_j$ is delayed by $\Delta \tau$, multiplied and accumulated at every sample clock before propagating to the next segment. An $XF$ correlation system will comprise of $N_{\text{lag}}$ such slices with a total input bandwidth of $N_{\text{lag}} \Delta f_{\text{o}} N_{\text{bit}}$. For $N_{\text{lag}}$ such slices, the total number of

### Table 1. Conventions

<table>
<thead>
<tr>
<th>Notation</th>
<th>Explanation</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{stat}}$</td>
<td>Number of stations/stations/Nodes</td>
<td>N_{\text{stat}} \times N_{\text{pol}}</td>
</tr>
<tr>
<td>$N_{\text{pol}}$</td>
<td>Number of polarizations</td>
<td>$N_{\text{stat}} \times N_{\text{pol}}$</td>
</tr>
<tr>
<td>$N_{\text{sig}}$</td>
<td>Number of signals</td>
<td>$N_{\text{bit}} \times N_{\text{bit}}$</td>
</tr>
<tr>
<td>$N_{\text{bits}}$</td>
<td>Number of bits</td>
<td>$N_{\text{stat}} \times N_{\text{pol}}$</td>
</tr>
<tr>
<td>$N_{\text{vis}}$</td>
<td>Number of visibilities</td>
<td>$N_{\text{bits}} \times N_{\text{bits}}$</td>
</tr>
<tr>
<td>$N_{X}$</td>
<td>Number of real multipliers</td>
<td>$N_{\text{bit}} \times N_{\text{bit}}$</td>
</tr>
<tr>
<td>$N_{\text{bins}}$</td>
<td>Number of subbands</td>
<td>$N_{\text{bins}} \times N_{\text{bins}}$</td>
</tr>
<tr>
<td>$N_{\text{fags}}$</td>
<td>Number of frequency bins</td>
<td>$N_{\text{fags}} \times N_{\text{fags}}$</td>
</tr>
<tr>
<td>$\tau_{\text{int}}$</td>
<td>Correlation - integration time</td>
<td>$N_{\text{fags}} \times N_{\text{fags}}$</td>
</tr>
<tr>
<td>$f_{\text{sys}}$</td>
<td>Observation frequency</td>
<td>$N_{\text{fags}} \times N_{\text{fags}}$</td>
</tr>
<tr>
<td>$\Delta f_{\text{o}}$</td>
<td>Instantaneous bandwidth</td>
<td>$N_{\text{fags}} \times N_{\text{fags}}$</td>
</tr>
<tr>
<td>$\Delta f_{\text{sb}}$</td>
<td>Sub-band bandwidth</td>
<td>$N_{\text{fags}} \times N_{\text{fags}}$</td>
</tr>
</tbody>
</table>
real multiplications per second is given from (1) and (17) is
\[
N_{i}^{xf} / \text{sec} = N_{x} N_{i} \Delta f_{i} = 2 N_{i}^{2} N_{i} \Delta f_{i},
\]
where \(N_{i} = N_{i} \text{stat} N_{i} \text{pol} \). If the system frequency is \(f_{\text{sys}}\) then the number of multipliers is given as
\[
N_{i}^{xf} = 2 N_{i}^{2} N_{i} \Delta f_{i} / f_{\text{sys}}.
\]
An advantage of XF architecture is that the Fourier transform (FFT) operation can either be an online or an off-line process. This means that the correlator operates on the entire bandwidth for all nodes \(N_{\text{stat}}\) and the data generated can be transmitted/stored without the immediate need to Fourier transform. Furthermore, since the FFT is an isomorphic process, the output rate is unaffected as shown in Table 2.

FX Correlator

An alternative to the XF correlator is to directly measure the cross-power spectrum. The term FX correlator was coined by Chikada [14] who built the first such correlator, implementing the reversal of the order of operations compared to the XF architecture. If \(X_{i}(\nu) \) and \(X_{j}(\nu)\) are the real time Fourier transforms of the delay compensated waveform \(x_{i}(t)\) and \(x_{j}(t)\) then using convolution theorem, the correlation function \(\zeta_{ij}(\tau)\) in (1) can be written as a multiplication in the Fourier spectrum [15]
\[
\zeta_{ij}(\tau) \equiv X_{i}(\nu) X_{j}^{*}(\nu) \equiv X_{ij}(\nu).
\]
Note that the cross-power spectrum to the right is a function of frequency and its Fourier transform to the left is a function of lags (\(\tau\)). From implementation perspective, the essence is to transform each input signal \(x_{i}\) into frequency domain (F) and then multiply-accumulate (X) over each spectral bin for all the nodes, to produce the cross-power spectrum and later the visibility function off-line. Unlike the XF correlator, the FX correlator must do a node based Fourier transform online, as indicated in Figure 3. The number of points \(N_{i}\) is given by the spectral resolution intended for the application i.e \(N_{\text{bins}}\). A shift register loads \(N_{\text{bins}}\) samples which is Fourier transformed to produce \(N_{\text{bins}}\) points. The number of points \(N_{\text{bins}}\) can be interpreted as the spectral translation of \(N_{i} \text{lags}\) from \(XF\) correlator in the time domain. The input and output data rates for such a system is given in Table 2. The total number of multipliers (\(N_{i}^{xf}\)) required is the sum of multipliers for node-based FFT (\(N_{i}^{xf}\)) and multipliers for Correlations (\(N_{i}^{corr}\)).

Consider \(N_{i} \text{sig}\) signals inputted to the correlator block, then using a \(N_{\text{bins}}\) - point FFT processor for each signal and referring to eq(19) we have the total number of multipliers required as
\[
N_{i}^{xf} = 2 N_{i}^{2} \zeta_{i} / f_{\text{sys}} \log_{2} N_{\text{bins}}.
\]

In contrast to the XF mode, where each sample is cross-multiplied at input rate, in FX mode each spectral bin is cross-multiplied and accumulated only once for \(N_{\text{bins}}\) samples, thereby reducing the number of computations by a factor \(N_{\text{bins}}\). In other words, the cross-multipliers operate at a rate \((\Delta f_{i} / N_{\text{bins}})\) instead of \(\Delta f_{i}\). From (17) we have, the number of multipliers required for correlation as
\[
N_{i}^{corr} = 2 N_{i}^{2} \zeta_{i} / f_{\text{sys}} \left[N_{i} \text{sig} + \log_{2} N_{\text{bins}}\right].
\]

HXF Correlator

The Hybrid XF correlator splits the input bandwidth \(\Delta f_{i}\) into smaller sub-bands using a array of filters or filter banks, making it very similar to the FX implementation. A XF correlation operating at a lower data rate is then applied to each of the sub-bands. The input/output data rates in this case are the same as \(XF\) or the FX correlator module given in Table 2. The number of multipliers required to split the \(N_{i} \text{sig}\) signals of bandwidth \(\Delta f_{i}\) into smaller \(N_{\text{ab}}\) sub-bands of width \(\Delta f_{\text{ab}}\) is given from
\[
N_{i}^{pfb} = 2 N_{i}^{2} \zeta_{i} / f_{\text{sys}} \left[2N_{\text{taps}} \log_{2} N_{\text{ab}}\right],
\]

where \(N_{\text{taps}}\) is the number of taps of the Poly-phase Filter Bank (PFB). In this mode a XF correlator operates on a single sub-band of band-width \(\Delta f_{i} / N_{\text{ab}}\). However since there are \(N_{ab}\) such XF correlator modules, the total number of operations/sec for the entire process remains constant. The total number of multipliers is thus given from (2) and (7) we have
\[
N_{i}^{xf} = N_{i}^{pfb} + N_{i}^{corr} = 2 N_{i}^{2} \zeta_{i} / f_{\text{sys}} \left[N_{i} \text{sig} + 2N_{\text{taps}} \log_{2} N_{\text{ab}}\right].
\]

The hybrid modes effectively splits the large input bandwidth into non-overlapping sub-bands using PFBs and then applies \(XF\) [16]. This enables smaller sub-systems to operate on down sampled rates rather than on the entire bandwidth.

<table>
<thead>
<tr>
<th>No. of channels</th>
<th>bandwidth/channel units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>(N_{\text{sig}}, N_{\text{bins}})</td>
</tr>
<tr>
<td>Output</td>
<td>(N_{\text{lags}}, N_{\text{bins}})</td>
</tr>
</tbody>
</table>

Table 2. Data rates for Lag/ XF , FX and HXF correlator
Earth leading/trailing (orbit Earth). Given a far away deployment location, such as Lunar, each satellite must down-link for correlator architecture design. The choice between the FX and additional benefits, we chose a phase rotation. Hence, given the reduction in computations fractional-sample delay compensation and truly station-based we have the processing factor given as

$$N_{FX} = \left( \frac{N_{sig} N_{bins}}{N_{sig} + \log_2 N_{bins}} \right).$$

As can be seen, the multiplicands in the XF mode are additive in the FX mode, besides the log reduction on the number of frequency bins. Thus, although for lower number of nodes the XF is comparable to FX mode, for large scalable architectures the FX mode is computationally cost effective. A similar reduction can be observed, by comparing the Hybrid mode against the XF mode,

$$N_{FX/fx} = \left( \frac{N_{sig} N_{bins}}{N_{sig} + 2 N_{taps}} \right),$$

which emphasizes the importance of breaking the bandwidth into smaller bands for processing. In addition, the FX and the HFX correlator also offers additional benefits such as fractional-sample delay compensation and truly station-based phase rotation. Hence, given the reduction in computations and additional benefits, we chose a FX correlator as our basis for correlator architecture design. The choice between the FX and HFX correlators depends on detailed parameters such as power consumption and flexibility of implementation.

3. CORRELATOR ARCHITECTURES

Among the three prevalent modes of correlation (XF, FX, Hybrid), the FX mode is assumed for all the below architectures.

Raw Transmission

The simplest scenario for a space-based interferometer is to transmit all the observed raw data to Earth. As illustrated in Figure 4, each satellite has transmission capability back to Earth, which is indicated with label ‘T’ and transmits $2N_{pol} \Delta f_i N_{bits}$ bits/sec. Observe that, with $N_{pol} = 3$, for a nominal instantaneous bandwidth of $\Delta f_i = 1$ MHz for 1 bit correlation, each satellite must down-link 6 Mbits/sec to Earth. Given a far away deployment location, such as Lunar orbit ($\approx 384,000$ km), Earth-Moon L2 ($\approx 378,000$ km), Earth leading/trailing ($\approx 2 \times 10^8$ to $4 \times 10^8$ km), this down-link data rate levies heavy prerequisites on the resources of a small satellite. Hence, the satellite cluster must not only employ onboard pre-processing, but also onboard correlation which is shown to minimize down-link data rate back to Earth.

Centralized Correlator

As the name suggests, the centralized correlator model, as shown in Figure 5, proposes to have a single correlation satellite for a cluster of $N_{stat} + 1$ satellites. The $N_{stat}$ observation nodes are primarily responsible for transmitting the data to the centralized correlator node. Each satellite node transmits $N_{pol}$ channels of capacity $2\Delta f_i N_{bits}$ bits/sec to the centralized mother ship, which receives at reception rate $D_{in}$

$$D_{in} = 2\Delta f_i N_{bits} N_{sig} \text{ (bits/sec).}$$

The input data from all satellites is then correlated and the output is then down-linked/transmitted down to Earth. Referring to (6) the total number of $N_{bit} \times N_{bit}$ multipliers needed per node are

$$N_{X}^{cs} = 2N_{sig} \left( \frac{\Delta f_i}{f_{sys}} \right) [N_{sig} + \frac{\log_2 N_{bins}}{N_{pol}}].$$

where $N_{sig} = N_{sig} N_{pol}$. After the correlation processes, the data is temporarily buffered and later transmitted down to Earth at scheduled intervals. Referring to Table 2, the data rate post-correlation is

$$D_{out}^{cs} = \left( \frac{2N_{sig} N_{bins} N_{bits}}{\tau_{int}} \right) \text{ (bits/sec).}$$

Distributed Correlator, Distributed Transmission

Centralized correlation depends heavily on the healthy operation of a single correlation node. Alternatively, the processing can be distributed to all nodes, removing the need for centralized correlation. This means that all antenna nodes take part in correlation. Additionally we also propose that after the processing each node down-links the data independently to Earth based node. Hence every node does both correlation and transmission denoted by ‘XT’ in the Figure 6. Three modes of distributing the signal processing for correlation are [17] (a) Antenna based distribution (b) Time based distribution and (c) Frequency based distribution. In Antenna based distribution, the entire data observed by each satellite is sent over to all other satellites for correlation. This immediately removes Single Point of Failure (SPOST), however increases communication and computation by incorporating high redundancy in the system. Alternatively, in Time based distribution, every satellite is assigned a particular time slot for correlation. This minimizes communication overload of the
network by minimizing the data transmitted. However, this architecture requires high levels of coherent communication and processing between the satellites, which is an additional overload. The third alternative, which offers a truly distributed solution for correlation and optimal communication and processing is frequency distributed correlation.

**Frequency distributed correlation**—In this framework, each node is assigned to correlate a specific sub-band of the complete bandwidth $\Delta f_s$. Hence, after the preliminary signal conditioning and quantization, each node splits the instantaneous bandwidth $\Delta f_s$ into $N_{sb}$ sub-bands, each of bandwidth $\Delta f_{sb}$, such that $\Delta f_s = N_{sb}\Delta f_{sb}$. Each of these $N_{sb}$ is then transmitted to the corresponding node assigned for that sub-band. For even distribution of data, we enforce the number of sub-bands equal to the number of nodes i.e

$$N_{sb} = N_{stat}.$$  \hspace{1cm} (13)

Each satellite node receives $N_{pol}$ channels each of bandwidth $\Delta f_{sb}$ from all other $N_{stat} - 1$ nodes, with single sub-band of bandwidth $f_{sb}$. Thus, the total inter satellite reception rate for each node is given by

$$D_{ins}^{f_{dc}} = \Delta f_{sb}(N_{sb} - 1)N_{pol}N_{bits} \text{(bits/sec/node)}.$$  \hspace{1cm} (14)

Observe that the data transmission is reduced by a factor $N_{sb}$ compared to the centralized correlation. In other words, every node transmits all the sub-bands of its input signal to all other nodes, except for the band it is assigned for correlation. In terms of processing, each satellite correlates only in a single sub-band $\Delta f_{sb}$, and hence referring to (6) the number of $N_{bit} \times N_{bit}$ multipliers per node needed are

$$N_x^{f_{dc}} = 2N_{sig}\left(\frac{\Delta f_{sb}}{f_{sys}}\right)[N_{sig} + \log_2 N_{bins}].$$  \hspace{1cm} (15)

Every node correlates a specific sub-band over a given integration time $\tau_{int}$ and then down-links the correlated data to Earth. From Table 2, the data rate of the final correlated output is given as

$$D_{out}^{f_{dc}} = \left(\frac{N_{sys}^2 N_{bins} N_{bits}}{\tau_{int} N_{stat}}\right) \text{(bits/sec/node)}.$$  \hspace{1cm} (16)

which, as expected, is a factor $N_{sb} = N_{stat}$ lower than the centralized correlator architecture and thereby dividing the communication resources evenly across all the satellite nodes. Evidently, if a single satellite fails, only a single sub-band of the total observation is lost. Furthermore, given communication flexibility, other active satellites in the network can poll and reassign a larger sub-band, to retain sensitivity. Moreover, additional satellites can be added to the network to minimize communication overload and/or maximize instantaneous bandwidth.

**Distributed Correlator, Centralized Transmission**

Providing down-linking capability to all satellite nodes is an expensive task and in addition consumes valuable observation time. Alternatively, this can be avoided by adding a Transmitting node to the cloud of existing observation nodes $N_{stat}$. In this mode, aptly called Distributed Correlator- Centralized Transmission, the processing is distributed, but the down-link transmission is centralized. The total number of nodes are thus

$$N_{stat}^{out} = N_{stat} + 1.$$  \hspace{1cm} Since the processing is still distributed, the equation for inter-satellite communication (14) and distributed processing given in (15) still hold. As shown in Figure 7, the centralized down-link node collects inter network processed data and down-links the processed data to Earth. The total data it receives is given from (10) and the down-link data rate is (12). Along similar lines, more down link satellites can be added to the cluster.

**4. CASE STUDIES**

To illustrate the distributed and centralized architectures discussed in the previous section, we investigate two feasibility studies, namely DARIS and OLFAR.

**DARIS**

DARIS is Distributed Aperture Array for Radio Astronomy in Space, a feasibility study to investigate a distributed aperture array in space for radio astronomy observations. The DARIS cluster comprises of 9 Nodes observing the cosmos in the ultra low frequency spectrum of $1 - 10$ MHz. Table 3 lists the first order requirements of the DARIS cluster. DARIS is an array of less than 10 nodes and is a feasibility study of a space-based array using Commercial Off-The-Shelf (COTS) components. Hence, although a distributed architecture is preferred, technology (in particular inter satellite communication) limits us to a centralized solution [11].

**Node level Signal Processing for DARIS**—Figure 8, shows a Node level Signal Processing (NSP) unit, which will be a part of all satellites in the DARIS cluster. Each of the observational satellite node will have $N_{sys} = 3$ separate data paths to the Node level Signal Processing (NSP) unit. Since the total observational bandwidth is too low, Direct Digital
Conversion (DDC) is employed. This eliminates the need for mixing stages and IF to baseband conversion. Thus, the entire observational bandwidth of $\Delta f_o = 10$ MHz input signal from each dipole is conditioned and directly sampled using a $N_{\text{bits}}$ bits Analog to Digital Converter (ADC) at the Nyquist rate of $2\Delta f_o$. A coarsely Poly phase Filter Bank (PFB) [18] is used to selectively choose the desired instantaneous bandwidth of $\Delta f_i = \Delta f_o / N_{\text{bins}} = 1$ MHz [19]. The PFB is essentially a commutator, followed by an array of Finite Impulse Response (FIR) filters and an FFT block. This windowing prior using the FIR filters prior to the FFT block significantly improves the out-of-band rejection ratio of the obtained spectrum. Furthermore, RFI mitigation techniques can be employed to eliminate interference and reducing the total number of bits to $N_{\text{bits}} \leq N_{\text{bits}}'$. The requirements of the dynamic range and hence $N_{\text{bits}}'$ will depend heavily on the deployment location of the satellites cluster [4]. However, we would assume a bare minimum of $N_{\text{bits}} = 1$ bits. The total amount of $N_{\text{bit}} \times N_{\text{bit}}$ multipliers for implementing the DSP for DARIS is given from (20) as $2\Delta f_o (2N_{\text{taps}} + \log_2 N_{\text{bins}}) / f_{\text{sys}} = 71$, for $N_{\text{taps}} = 16$ and $f_{\text{sys}} = 10$ MHz. The output from each signal processing data path will be $2\Delta f_i N_{\text{bits}} = 2$ Mbits/sec. For $N_{\text{pol}} = 3$ data paths, the total output from each satellite is $2N_{\text{pol}}\Delta f_i N_{\text{bits}} = 6$ Mbits/sec, which is sent to the intra satellite communication layer for transport. The centralized mother ship receives $2N_{\text{stat}} N_{\text{pol}} \Delta f_i N_{\text{bits}} = 48$ Mbits/sec, where $N_{\text{stat}} = 8$ excluding the mother ship.

Centralized Processing and Down-link

As discussed earlier, the DARIS cluster will employ a centralized correlator-centralized down-link architecture. The centralized mother ship is also an observation satellite. Hence for correlating $N_{\text{sig}} = N_{\text{pol}}(N_{\text{stat}} + 1) = 27$ signal paths, the total number of processing units is given from (11) as $N_{\text{daris}} X = 2N_{\text{sig}} \Delta f_i [N_{\text{sig}} + \log_2 N_{\text{bins}}] / f_{\text{sys}} = 163 N_{\text{bit}}$ multipliers, given $f_{\text{sys}} = 10$ MHz. The centralized mother ship then down-links data at the rate given by (12), as $D_{\text{daris}} \text{out} = 2N_{\text{pol}} N_{\text{stat}} N_{\text{bit}} \tau_{\text{int}} = 1.46$ Mbits/sec.

OLFAR

The OLFAR (Orbiting Low Frequency Array for Radio astronomy) project aims to design and develop a detailed system concept for an un-tethered swarm of more than 10 scalable autonomous nano satellites in space (well above the ionosphere) to be used as a scientific instrument for ultra low frequency observations (Table 3). The large number of such satellites spread over large distances will collectively synthesize an aperture dish of diameter 100 kilometers. OLFAR is very similar in design to DARIS, however aims to observe a wider spectrum of $\Delta f_o \approx 30$ MHz and will consist of a
larger number of satellites. For the sake of illustration, we will consider a nominal $N_{stat} = 10$ satellites in this design.

**Node level Signal Processing for OLFAR**

Each observational satellite will be equipped with a NSP unit for preprocessing before correlation, which is shown in Figure 9. Along similar lines to the DARIS NSP, the input signal will be digitized, RFI mitigated (not shown in Figure 9) and filtered using a PFB to obtain the instantaneous bandwidth $\Delta f_s = 1$ MHz. In reality, OLFAR aims to achieve a larger instantaneous bandwidth $\geq 1$ MHz and thereby improving the sensitivity of the instrument. A second fine PFB is used to further split $\Delta f_s$ into $N_{sb}$ sub-bands, each of bandwidth $\Delta f_{sb} = \Delta f_s/N_{sb}$. Furthermore, for even distribution of data, we enforce the number of sub-bands to be equal to the number of stations, $N_{sb} = N_{stat} = 10$ and hence $\Delta f_{sb} = 100$ KHz. The intra-satellite communication layer then transmits $N_{sb} = 1$ sub-bands to all other satellites. The number of multipliers required for the coarse PFB is similar to the DARIS PFB, which is 71 multipliers running at $f_{sys} = 10$ MHz and for the Fine PFB, we have $2\Delta f_{i}s[N_{sig} + \log_2 N_{bins}]/f_{sys} = 7.1$ multipliers, with $N_{taps} = 16$ in both cases. Each satellite then transmits $2N_{pol}(N_{sb} - 1)\Delta f_i = 5.4$ Mbits/sec to every other satellite in the network.

**Distributed correlation and Down-link**

The OLFAR network would employ Distributed correlation-Distributed down-link architecture, where by all the satellites will correlate data from a specific sub-band assigned to them. Using (15), DARIS NSP will employ $2N_{pol}\Delta f_{sb}[N_{sig} + \log_2 N_{bins}]/f_{sys} = 10$ multipliers, where $N_{sig} = N_{stat} N_{pol} = 30$. Observe that this requirement is an order smaller compared to the DARIS cluster of 9 nodes. Furthermore, from (16), each of the satellites will down-link the data at a factor of $N_{stat} = N_{sb}$ lower compared to the DARIS central mother ship, with $D_{out}^{defl} = 2N_{sig}^2 N_{bins} N_{bits}/(N_{stat}^7 int) = 146$ Kbits/sec/node.

### 5. Conclusion

A first order model of computational requirements for a correlator were presented, including an overview of conventional terrestrial correlator architectures such as $XF$, $FX$ and $HX$. For space-based interferometry, centralized and distributed architectures were discussed, with emphasis on Frequency Distributed Correlation which is seen to be the optimal choice for a space-based array, distributing the downlink and processing evenly among all the satellite nodes, in addition to removing Single Point of Failure. As a computational illustration, the signal processing units of DARIS and OLFAR projects discussed, where centralized and distributed solutions are presented respectively.

**APPENDIX**

**Complex Multiplications**

The product of any 2 complex numbers; say $(x_1 + iy_1)$ and $(x_2 + iy_2)$ is given by $(x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$. Hence cross correlation of 2 complex signals involves 4 multiply accumulates (MAC). However for auto-correlation this scales down to 2 multiplications and 1 addition. Thus, we have

1. Complex MAC = 4 Real MAC for cross-correlations
2. Complex MAC = 2 Real MAC for auto-correlations

Hence, the total number of real $N_{mb}$ multipliers required of for correlation of $N_{sig}$ with $N_{sig}$ is

$$N_{X}^{ff} = 4N_{sig}(N_{sig} - 1) + 2N_{sig}$$

$$= 2N_{sig}^2$$

(17)

**Fourier Transform**

The Discrete Fourier Transform (DFT) of a finite duration sequence $x(n) \leq n \leq N - 1$ is given by

$$X(k) = \sum_{n=0}^{N-1} x(n)W^{nk}$$

where $W = e^{-j(2\pi/N)}$, where $W^{nk}$ is periodic in $N$. As seen from the equation for an N-point DFT the number of multiplications is $N^2$. Fast Fourier Transform or FFT is an alternative efficient algorithm to compute DFT. Using the straight forward pipelined Radix-2 implementation, the DFT equation can be broken down to $N/2$ stages of 4 multiplexes each. For a given input signal of bandwidth $\Delta f$, the FFT provides $N_{bins}$ coefficients at the rate $\Delta f/N_{bins}$. Hence the total number of real $N_{mb} \times N_{mb}$ multiplications per second required is given as

$$N_{X}^{ff} = 4(\Delta f/N_{bins})(N_{bins}/2) \log_2 N_{bins}$$

$$= 2\Delta f \log_2 N_{bins}$$

(19)

**Poly-phase Filter Bank**

A poly phase filter bank for a single node, single polarization consists of a FIR filter and a FFT module.

**FIR Filter**—The number of multiplications for a single FIR filter is equal to the number of taps ($N_{taps}$). For an input bandwidth $\Delta f$, the number of real multiplications required are $\Delta f N_{taps}$. For a complex signal the coefficients will also be complex, hence from (17) we have the total number of real multiplications for a complex signal as $4\Delta f N_{taps}$.

**Filter Bank**—Referring to (19), we have the total number of real multiplications required for a poly phase filter bank as

$$N_{X}^{pol} = 4N_{taps} + 2\Delta f (\log_2 N_{bins})$$

$$= 2\Delta f (2N_{taps} + (\log_2 N_{bins}))$$

(20)

**6. Acknowledgements**

The authors would like to thank all members of the OLFAR and DARIS projects for discussions, in particular Noah Saks for discussions on Correlators. This research was funded in part by two projects namely, the STW-OLFAR (Contract Number: 10556) within the ASSYS perspectief program and secondly the ESA-DARIS Contract “Feasibility of Very Large Effective Receiving Antenna Aperture in Space” (Contract Number 22108/08/NL/ST).
REFERENCES


BIography

Raj Thilak Rajan (S’10) received the B.Sc (with distinction) and M.Sc (with distinction) in Electronics science from University of Pune, India in 2004 and 2006 respectively. He is presently with the Digital and Embedded Signal Processing (DESP) R&D group at ASTRON, The Netherlands (since December 2008) and also a PhD candidate at Circuits And Systems (CAS), Faculty of EEMCS, TU-Delft, The Netherlands (since April 2010). Previously, he worked at Whirlpool (India), Politecnio di Bari (Italy) under MIUR and INFN fellowship, and was a visiting researcher at CERN (Switzerland) for the Compact Muon Solenoid (CMS) experiment in the Large Hadron Collider (LHC) project. His research interest lies in signal processing, algorithms and implementation.

Mark Bentum (S’92-M’95-SM’10) was born in Smilde, The Netherlands, in 1967. He received the M.Sc. degree in electrical engineering (with honors) from the University of Twente, Enschede, The Netherlands, in August 1991. In December 1995 he received the Ph.D. degree for his thesis "Interactive Visualization of Volume Data" also from the University of Twente. From December 1995 to June 1996 he was a research assistant at the University of Twente in the field of signal processing for mobile telecommunications and medical data processing. In June 1996 he joined the Netherlands Foundation for Research in Astronomy (ASTRON). He was in various positions at ASTRON. In 2005 he was involved in the eSMA project in Hawaii to correlate the Dutch JCMT mm-telescope with the Submillimeter Array (SMA) of Harvard University. From 2005 to 2008 he was responsible for the construction of the first software radio telescope in the world, LOFAR (Low Frequency Array). In 2008 he became an Associate Professor in the Telecommunication Engineering Group at the University of Twente. He now works in digital signal processing and digital system design. Dr. Bentum is a Senior Member of the IEEE, the Dutch Electronics and Radio Society NERG, the Dutch Royal Institute of Engineers KIVI NIRIA, and the Dutch Pattern Recognition Society, Secretary of the Dutch URSI committee, Secretary of the Foundation of Scientific Activities of the Dutch URSI Committee, Executive Committee member of the IEEE Benelux Communications and Vehicular Technology Chapter, and has acted as a reviewer for various conferences and journals.

Andre Gunst was born in Havelte, The Netherlands, in 1972. He received the M.Sc degree in electrical engineering from the University of Twente, Enschede, The Netherlands, in June 1999. In the same year, he joined ASTRON as a digital system engineer with the R&D Department. Since 2004 he has worked extensively on the development of station systems for the LOw Frequency ARray (LOFAR) project and since 2006, he leads the LOFAR technical team. From 2010 onwards, he assumed a dual role, working for both ASTRON and the Square Kilometer Array (SKA) project. SKA will be the largest radio telescope on Earth, built in Australia or South-Africa. He is an Aperture Array Domain Specialist for the SKA office located in Manchester. In addition, he also leads the digital system development of several projects in ASTRON. His research interests include (digital) system design and digital signal processing.

Albert-Jan Boonstra was born in The Netherlands in 1961. He received the B.Sc. and M.Sc. degrees in applied physics from Groningen University, Groningen, the Netherlands, in 1984 and 1987, respectively. In 2005 he received the PhD degree for his thesis Radio frequency interference mitigation in radio astronomy from the Delft University of Technology, Delft, The Netherlands. He was with the Laboratory for Space Research, Groningen, from 1987 to 1991, where he was involved in developing the short wavelength spectrometer (SWS) for the infrared space observatory satellite (ISO). In 1992, he joined ASTRON, the Netherlands Foundation for Research in Astronomy, initially at the Radio Observatory Westerbork, Westerbork, The Netherlands. He is currently with the ASTRON R&D Department, Dwingeloo, The Netherlands, where he heads the DSP group. His research interests lie in the area of signal processing, specifically RFI mitigation by digital filtering.