Seventh Framework Programme: Research Infrastructures

Contract No. 212243

Preparatory Phase for the Square Kilometre Array

PrepSKA Workpackage 3:

System Preliminary Design Report

Deliverable 2.3
Memo 130
SKA Phase 1: Preliminary System Description

P. Dewdney (SPDO)
J-G bij de Vaate (ASTRON)
K. Cloete (SPDO)
A. Gunst (ASTRON)
D. Hall (SPDO)
R. McCool (SPDO)
N. Roddis (SPDO)
W. Turner (SPDO)

November 2010
**DOCUMENT HISTORY**

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date Of Issue</th>
<th>Engineering Change Number</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>2010-05-25</td>
<td>-</td>
<td>First draft. Also distributed to JGBdV and PA for review and inputs.</td>
</tr>
<tr>
<td>0.7</td>
<td>2010-06-28</td>
<td>-</td>
<td>All the main sections inserted. Recent cost information inserted. Performance graphs developed.</td>
</tr>
<tr>
<td>1.0</td>
<td>2010-07-15</td>
<td>-</td>
<td>Changes to wording. Response to comments.</td>
</tr>
<tr>
<td>1.1</td>
<td>2010-08-20</td>
<td>-</td>
<td>Added sections 8 and 10, editing.</td>
</tr>
<tr>
<td>1.2</td>
<td>2010-08-26</td>
<td>-</td>
<td>Updated following internal SPDO review</td>
</tr>
<tr>
<td>2.0</td>
<td>2010-11-20</td>
<td>-</td>
<td>Removed section on costs, updated the last section, updated performance graphs.</td>
</tr>
<tr>
<td>2.1</td>
<td>2010-11-22</td>
<td>-</td>
<td>Minor formatting updates.</td>
</tr>
</tbody>
</table>

**DOCUMENT SOFTWARE**

<table>
<thead>
<tr>
<th>Package</th>
<th>Version</th>
<th>Filename</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wordprocessor</td>
<td>MsWord</td>
<td>MGT-040.070.010-MR-001-2.1_PrelimSysDefandCost</td>
</tr>
</tbody>
</table>

**ORGANISATION DETAILS**

<table>
<thead>
<tr>
<th>Name</th>
<th>SKA Program Development Office</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical/Postal Address</td>
<td>Jodrell Bank Centre for Astrophysics</td>
</tr>
<tr>
<td></td>
<td>Alan Turing Building</td>
</tr>
<tr>
<td></td>
<td>The University of Manchester</td>
</tr>
<tr>
<td></td>
<td>Oxford Road</td>
</tr>
<tr>
<td></td>
<td>Manchester, UK</td>
</tr>
<tr>
<td></td>
<td>M13 9PL</td>
</tr>
<tr>
<td>Fax.</td>
<td>+44 (0)161 275 4049</td>
</tr>
<tr>
<td>Website</td>
<td><a href="http://www.skatelescope.org">www.skatelescope.org</a></td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

1 **INTRODUCTION** ............................................................................................................................................. 6  
1.1 Purpose of the document ....................................................................................................................................... 6  
1.2 Scope of the document ........................................................................................................................................... 6  

2 **REFERENCES** ...................................................................................................................................................... 6  

3 **GLOBAL ASSUMPTIONS** ................................................................................................................................. 7  
3.1 Technical ................................................................................................................................................................. 7  
3.1.1 Notes on the System ........................................................................................................................................... 7  
3.1.2 Sky Noise ........................................................................................................................................................... 8  

4 **DISH AND SPARSE AA ARRAY CONFIGURATION ASSUMPTIONS** ................................................................. 8  

5 **DISH ARRAY** ....................................................................................................................................................... 10  

6 **SPARSE APERTURE ARRAY** ................................................................................................................................. 12  

7 **IMAGING PROCESSING** ........................................................................................................................................ 14  

8 **NON-IMAGING PROCESSING** ............................................................................................................................... 15  

9 **PHASE 1 SYSTEM SENSITIVITY** .......................................................................................................................... 16  

10 **THE ROUTE FORWARD** .................................................................................................................................... 17  
10.1 The Starting Point .................................................................................................................................................. 18  
10.2 The Design Process leading from Phase 1 and Phase 2 ......................................................................................... 19  
10.2.1 Stepping Through Phase 1 ............................................................................................................................ 23  
10.2.2 Stepping Through Phase 2 ............................................................................................................................ 23  
10.2.3 Stepping Through the Advanced Instrumentation Programme ...................................................................... 24  

11 **CONCLUSION** ....................................................................................................................................................... 25
LIST OF FIGURES

Figure 1: Sky Noise Temperature vs frequency for low radio frequencies.................................................. 8
Figure 2: The distribution of collecting area (AAs and dishes) for the Phase 1 array. Note that the dish core and the AA core will actually be displaced from each other. In SKA Phase 2 the mid region will extend to 180 km......................................................... 9
Figure 3: An example of an inner array configuration conforming to the assumptions in this document.......................................................................................................................... 10
Figure 4: System sensitivity (Ae/Tsys) in m²/K. The blue line (left) is for the Sparse AA with receiver temperature of 150K; the green line, for a receiver temperature of 100K; the red line (right) is for the Dish Array. Note the offset zero........................................................................ 16
Figure 5: Survey Speed Figure of Merit (SSFoM): (Ae/Tsys)² Ωfov in units of 10⁶ m⁴ K⁻² deg². The blue line (left) is for the Sparse AA with receiver temperature of 150K; the green line, for a receiver temperature of 100K; the red line (right) is for the Dish Array............................................................. 17
Figure 6: Part A: A standard ‘V’ diagram showing the phases of a large project. The dotted boxes correspond approximately to those in Figure 7. Part B: A schematic representation of the juxtaposition of the Phase 1, Phase 2 and AIP activities. The arrows between the ‘V’s illustrate critical information flows between the activities. More detail is shown in Figure 8, where SKA-specific activities are expanded. Major milestones are shown at the bottom. .................. 20
Figure 7: A simplified view of the approach to the SKA design and construction. The dotted boxes correspond to activities in the schedule shown in Figure 8.................................................. 21
Figure 8: A preliminary schedule based on the activities shown in Figure 7. The double-headed arrows denote information flow from both user and non-science sources through to preliminary design. Major SKA milestones are shown at the bottom................................................ 22

LIST OF TABLES

Table 1: Phase 1 Dishes with SPFs ........................................................................................................... 11
Table 2: Key System Performance Factors for Dishes and SPFs .............................................................. 12
Table 3: Phase 1 Sparse Aperture Arrays ............................................................................................... 13
Table 4: Key System Performance Factors for Sparse AAs ................................................................. 14
LIST OF ABBREVIATIONS

AA..............................Aperture Array
AIP...........................Advanced Instrumentation Programme
CoDR..........................Conceptual Design Review
CPU...........................Central Processing Unit
DAA...........................Dense Aperture Array
DRM...........................Design Reference Mission
EoR...........................Epoch of Reionisation
FLOPS.........................Floating Point Operations per second
FoV...........................Field of View
HPC...........................High Performance Computer
Hz.............................Hertz
K...............................Kelvin
Km.............................kilometre
LNA...........................Low Noise Amplifier
m...............................metre
OTPF.........................Observing Time Performance Factor
PAF...........................Phased Array Feed
PrepSKA.....................Preparatory Phase for the SKA
RFI...........................Radio Frequency Interference
s..............................second
SEFD..........................System Equivalent Flux Density
SKA...........................Square Kilometre Array
SKADS......................SKA Design Studies
SPDO.........................SKA Program Development Office
SSEC..........................SKA Science and Engineering Committee
SSFoM......................Survey Speed Figure of Merit
T...............................Temperature
TBD..................To be determined
WBSPF......................Wide Band Single Pixel Feed
1 Introduction

The SPDO held a System Conceptual Design Review (Sys CoDR) in Feb, 2010. The CoDR review panel recommended a focussing of science and technology goals, and the SSEC has responded by defining a reduced set of science goals and technologies for SKA Phase 1. This is summarised in “Concept Design for SKA Phase 1 (SKA1)” [1], in which a Baseline Design is adopted and an Advanced Instrumentation Program (AIP) is described, elements of which could be implemented in Phase 2 of the telescope construction. The baseline design incorporates a Dish Array with single pixel feeds and a low frequency aperture array. The AIP includes phased array feeds (PAFs) for the dishes and mid-frequency dense aperture arrays (DAAs), both of which provide greatly expanded fields of view for fast surveying, and wide-band single-pixel feeds (WBSPFs), with much wider frequency coverage than currently available feeds.

1.1 Purpose of the document

Based on [1], this document has been written to provide a minimal technical foundation for a plausible estimate of performance and cost for Phase 1 SKA. It lays the foundation for further technical development and for obtaining a cost estimate. Because sufficient information on cost is not yet available and will in any case require a more complete system design, an actual cost estimate is not provided.

1.2 Scope of the document

The system sketched in this document not the refined, optimised system that we expect to emerge from a more extensive, systematic study, which the international community will continue to carry out. Such a study will involve an orderly decision-making process which trades-off technical requirements, derived from the user requirements (essentially the Design Reference Mission (DRM) [2]) and other sources, system performance and total cost of ownership. This process is needed to define the most cost-effective Phase 1 sub-set as well as SKA Phase 2, but is beyond the scope of this document.

This document attempts to be compatible with the precepts contained in [1]. Variations are noted in the text.

The process for developing the Phase 1 technical foundation and cost estimate will recognise that SKA Phase 1 is not an end in itself, but rather a “waypoint” on the way to SKA Phase 2. This has implications for the specifications of components, infrastructure, and other items, the design and cost of which will be developed so that a systematic progression from SKA Phase 1 to Phase 2 is possible. A planning outline has previously been described in [6], and is included here in Section 10.

2 References


3 Global Assumptions

3.1 Technical

3.1.1 Notes on the System

The following assumptions have been used for the purposes of this document only. Design and performance assumptions are mainly conservative, and have been minimised to provide as much latitude for further development as possible.

- The two main receptors, the Dish Array and the Sparse Aperture Array, are assumed to occupy spatially separated cores. The cores are serviced by data transport and time transfer sub-systems that lead to a common signal processing facility. Outside the inner region (see Section 4), the two types of receptors are assumed to be grouped together and served by common routing of data transport and time transfer.
- The infrastructure for the signal processing equipment servicing the two receptor arrays is assumed to be shared (power provision, buildings, etc).
- Each Sparse Aperture Array station is assumed to have its coarse beamformers at the station location.
- Separate imaging processors (channelizers, correlators and array beam-formers) are assumed for the two receptor arrays.
- The non-imaging processor is assumed to be shared between the two receptor arrays.
- The HPC is assumed to be located far off site, and to be shared between the two receptors arrays.
• Both receptor arrays and their respective support services are assumed to be operating concurrently and continuously. The Dish Array can use only one feed (frequency range) at a time.

• Some assumptions about the design of the low frequency aperture array have been made. These have been noted in the text.

3.1.2 Sky Noise

The component of system temperature originating as “sky noise” that will be used in this document is given in Figure 1, which also shows the approximation, $T_{\text{sky}} = 60\lambda^{2.55}$ used in [3] at low frequencies. This approximation agrees with an “average” between those given for the Galactic pole and the Galactic Plane, and is valid up to ~450 MHz. For the purposes of determining collecting area, $T_{\text{sky}}$ is assumed to be 500 K at 130 MHz ($z \cong 10$).

![Figure 1: Sky Noise Temperature vs frequency for low radio frequencies.](image)

4 Dish and Sparse AA Array Configuration Assumptions

The configuration description uses the nomenclature of previous documents (core, inner, mid, outer). Phase 1 does not cover the outer region and the mid region is limited to 100 km (compared to 180 km for Phase 2).

The fractional distribution of collecting area is also not the same for Phase 2 as that approved by the SSEC in March 2010, and is shown in Figure 2. Figure 3 is an example of a configuration that conforms to the assumptions outline below.\(^1\)

The following additional assumptions are made for this document. These are simplified to some extent compared with what might actually be needed for Phase 1. For example, the distribution of

\(^1\) This example is drawn from the work of the SKA configuration task force (R. Bolton et al).
dishes and AAs may not be actually the same in the core and inner regions, and the positions of clusters\(^2\) of dishes and AAs along spiral arms will likely have to be adjusted.

**SKA Phase 1 Array Distribution**

![Diagram showing the distribution of dishes and AAs in SKA Phase 1 array]

**Figure 2:** The distribution of collecting area (AAs and dishes) for the Phase 1 array. Note that the dish core and the AA core will actually be displaced from each other. In SKA Phase 2 the mid region will extend to 180 km.

1. The distribution of Sparse AA and dish aperture will be as similar as possible.
2. The configuration of collecting area within the core and inner regions will be a sub-set of locations designed for Phase 2.
3. The mid array is a 3-arm spiral each arm extending 100 km from the centre.
4. There are five receptor locations on each spiral arm, located equidistant from each other and from the edge of inner (This assumption will ultimately have to be adjusted to account for the desired log distribution of collecting area with distance). Each location contains a cluster of dishes and one AA station.
5. 250 dishes total; 175 in the core and inner regions, 75 in the mid region.
6. In the mid region there are 5 dishes and one AA station per cluster.
7. 50 AA stations total; 35 in the core and inner regions, 15 in the mid region.

---

\(^2\) A cluster refers to location on the spiral arm where both dishes and sparse aperture array stations are located close together and are services by common power and infrastructure. The data from each of the individual dishes in the cluster is sent back to the signal processing facility.
5 Dish Array

Table 1 contains the parameters for the Phase 1 Dish Array. The values of two of the parameters (frequency range and channel bandwidth) adopted here differ from [1]:

- For SKA Phase 1 it is assumed that two feeds are available for each antenna. They are based on high-performance corrugated horns with ~2:1 bandwidth ratio, possibly enabling $T_{sys}$ to be as low as 30 K with an aperture efficiency of 70%. These two feeds will not provide coverage above 2 GHz, if positioned in frequency starting at 450 MHz. They could be re-positioned in frequency without much cost impact.

- Channel width has been taken to be that needed for full-field imaging in continuum (~7.5 kHz). Although this is not a science priority, it is a more stringent limit than might be reasonable for HI line observations (~10 km/s). However, this is still much larger than the 1 kHz channel bandwidth than that given in Table 1 of [1]. A channel width of 1 kHz would imply an order of magnitude larger processing load than described here.
Table 1: Phase 1 Dishes with SPFs

<table>
<thead>
<tr>
<th>Aperture</th>
<th>250, 15-m diameter parabolic antennas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Frequency</td>
<td>300 MHz</td>
</tr>
<tr>
<td>Upper Frequency</td>
<td>10 GHz</td>
</tr>
<tr>
<td>Total physical aperture</td>
<td>44179 m²</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Array Configuration</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Core (radius &lt; 0.5 km)</td>
<td>250 antennas</td>
</tr>
<tr>
<td>Inner (0.5 &lt; radius &lt; 2.5 km)</td>
<td>“50% (125 ant.)”</td>
</tr>
<tr>
<td>Mid (2.5 &lt; radius &lt; 100 km)</td>
<td>“20% (50 ant.)”</td>
</tr>
<tr>
<td>Core filling factor</td>
<td>0.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Antenna RF System</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Feed/LNA 1</td>
<td>0.45 – 1.0 GHz</td>
</tr>
<tr>
<td>Feed/LNA 2</td>
<td>1.0 – 2.0 GHz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Digital Outputs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Core-to-Mid antennas</td>
<td></td>
</tr>
<tr>
<td>Number of Sample streams</td>
<td>2</td>
</tr>
<tr>
<td>bits per sample</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signal Transport System</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Radius &lt; 100 km</td>
<td>24 Gbit s⁻¹</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signal Processing System</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlator</td>
<td></td>
</tr>
<tr>
<td>Input data streams</td>
<td>500</td>
</tr>
<tr>
<td>Frequency channels (line limit)</td>
<td>30,000</td>
</tr>
<tr>
<td>Frequency channels (cont. limit)</td>
<td>67,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Science Computing System</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Input data rate</td>
<td>(8 + 336) x 10⁹ Byte s⁻¹ av'ge from correlator (8-Byte complex, core-inner + long)</td>
</tr>
</tbody>
</table>

*1.0 GHz x 2 (Nyquist) x 4 (bits) x 1.25 (encoding) x 1.2 (oversample) x 2 (streams) = 24 Gbit s⁻¹.
* For imaging, the channel bandwidth is assumed to be given by \( \Delta \nu / \nu \ll 20 / \theta_{\text{fov}} \), where \( \Delta \nu \) is the channel bandwidth, \( \nu \) is the centre frequency, \( \theta_i \) is the synthesised beamwidth, and \( \theta_{\text{fov}} \) is the width of the FoV. The integration time is assumed to be \( \Delta t = 20 / \theta_{\text{fov}} \times 1.37 \times 10^4 \text{ s} \), where \( \Delta t \) is the integration time, and \( 1.37 \times 10^4 \) is the number of sidereal seconds in a radian. The \( \ll \) is taken to be a factor of 10. Also, if the FoV is taken to be at approximately the half-power point of the aperture, \( 20 / \theta_{\text{fov}} \) is approximately \( d / D_{\text{max}} \), where \( d \) is the aperture diameter and \( D_{\text{max}} \) is the maximum baseline.
** Assuming that ~half of the baselines are core-inner only. Overall compute requirement may be a serious underestimate. This figure assumes 100% efficiency, and it is known that HPCs typically operate at much lower efficiency. Moreover, recent information indicates that \( 10^8 \) flops / float will be needed for wide-field, high dynamic range imaging. Thus the actual compute requirement may be more than an order of magnitude greater.
Table 2 contains the key system performance factors for the dishes equipped with single-pixel feeds described in Table 1.

<table>
<thead>
<tr>
<th>Table 2: Key System Performance Factors for Dishes and SPFs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antenna/Feed Efficiency</td>
</tr>
<tr>
<td>Minimum Elevation Angle</td>
</tr>
<tr>
<td>Average T_{sys} in low band feed</td>
</tr>
<tr>
<td>Average T_{sys} in high band feed</td>
</tr>
<tr>
<td>System Equiv. Flux Density (each antenna, Stokes I)</td>
</tr>
<tr>
<td>System Equiv. Flux Density (250 antennas, Stokes I)</td>
</tr>
<tr>
<td>∆S (1σ) for 1.0 GHz bandwidth**</td>
</tr>
<tr>
<td>Field-of-View @ 30 cm wavelength (1 GHz)***</td>
</tr>
<tr>
<td>Equivalent “pill-box” Field-of-view</td>
</tr>
<tr>
<td>Field-of-View @ 450 MHz</td>
</tr>
<tr>
<td>A_e/T_{sys} (all antennas)</td>
</tr>
<tr>
<td>773</td>
</tr>
<tr>
<td>SSFoM @ \lambda = 30 cm (1 GHz), T_{sys} = 30K</td>
</tr>
<tr>
<td>250 Antennas @ &lt;100 km radius</td>
</tr>
<tr>
<td>SSFoM @ 450 MHz, T_{sys} = 40K</td>
</tr>
<tr>
<td>250 Antennas @ &lt;100 km radius</td>
</tr>
<tr>
<td>Imaging Dynamic Range****</td>
</tr>
<tr>
<td>Spectral Dynamic Range</td>
</tr>
</tbody>
</table>

* SEFD = 2 kbolts T_{sys} 1026 / A_e Jy.
** ∆S = SEFD / (2 B/1/2 Jy s^{-1/2} for Stokes I.
*** This is the FoV used for the science performance evaluation. Assumes illumination tapers to zero at the edge of the dish and follows I(x) = 1 – 2r^2/3, where r is the distance from the dish centre. This results in a first sidelobe level of ~-23 dB.
**** Ratio taken between the flux of a typical L-band source in field of 100 mJy and 1σ noise level after a 1000-hr integration on a single field.

6 Sparse Aperture Array

Table 3 contains the parameters for the Sparse Aperture Array (AA). There is one difference between the parameters adopted in this document and those described in [1]:

- The number of elements (11,200 per station) in the array is used as a starting point in this document. This provides sensitivity (A_e/T_{sys}) of 1515 m^2/K at 130 MHz (z_{HI} ≈ 10) at the zenith, rather than 2000 m^2/K as described in [1].

In addition:

- The station diameter, which is not stated in [1], is assumed to be 180 m. This leads to an almost filled core (80%, compared with the maximum possible of 91% (\pi/(2\sqrt{3})) for circular stations) and to dense AA performance below ~115 MHz. For practical reasons, this may mean that the distribution of stations for AAs in the core be somewhat larger than for dishes.
- Reference [3] is used in determining most array parameters and performance.
- The receiver temperature for the AA elements is assumed to be either 150K across the band. This value is similar to that obtained in the present LOFAR high-band antennas.

---

[^3]: Private communication, Aperture Array Verification Program. This is the basis for developing a Phase 1 aperture array consistent with the cost allocation given in [1].
• The term “Sparse Aperture Array” has been retained, even though the elements are close enough together to become dense at low frequencies. For most of the bandwidth the arrays are sparse.

A discussion of low frequency options for Phase 1 is contained in [8]. This document contains descriptions of potential element designs and other design aspects that could serve as a starting point for continuing the development of Sparse AAs for Phase 1.

<table>
<thead>
<tr>
<th>Table 3: Phase 1 Sparse Aperture Arrays</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sparse Aperture Array</strong></td>
</tr>
<tr>
<td><strong>Lower Frequency</strong></td>
</tr>
<tr>
<td><strong>Upper Frequency</strong></td>
</tr>
<tr>
<td><strong>Number of antennas</strong></td>
</tr>
<tr>
<td><strong>Total physical aperture</strong></td>
</tr>
<tr>
<td><strong>Area per antenna</strong></td>
</tr>
<tr>
<td><strong>Dense/Sparse Transition $\lambda$</strong></td>
</tr>
</tbody>
</table>

**Array Configuration**
- Core (radius < 0.5 km) ~50% (25 stations) Fractional total number of AA stations.
- Inner (1 < radius < 2.5 km) ~20% (10 stations)
- Mid (2.5 < radius < 180 km) ~30% (15 stations) In clusters of 5 stations (Total of 15 clusters)
- Core filling factor 0.81
- Number of antennas 11,200
- Instantaneous bandwidth per beam 380 MHz Assures full bandwidth is available (70-450 MHz)

**Digital Outputs**
- Sample streams 960 Max - sub-bands
- bits per sample 4

**Signal Transport System**
- Optical fibre to signal processor
- Data rate per station
- Optimized layout, buried fibre, highly multiplexed.

**Signal Processing System**
- Coarse frequency channels 3040 Bandwidth / 125kHz
- Fine Frequency channels*** 3.8x10$^3$ Bandwidth / 1kHz
- Complex Correlations 9.3x10$^{12}$ (50'/2)baselines x (480) bms x 4 pol'n prod's x 3.8 x 10$^6$ chans
- Core-Inner Dump Time 49 s
- Minimum Dump Time 1.2 s Longest Baseline, taken as 200 km.

**Science Computing System****
- Input data rate (75 + 3100) x 10$^9$ Byte s$^{-1}$ av’ge from correlator (8-Byte complex, core-inner + long)
- Imaging Processing Required (core + inner) 189 Tflops @ 10$^7$ flops / float (EVLA Memo 24) (core-inner, 50% of baselines)
- Imaging Processing Required (non-core-inner baselines) 7750 Tflops @ 10$^8$ flops / float (minimum dump time, 50% of baselines)
- Archive 1 Exabyte Limited by cost in 2015 timeframe.

* $\lambda_{ds} = (3\pi d_{station}^2/4 \cdot 11200)^{1/2}$.
** 380 MHz x 2 (Nyquist) x 4 (bits) x 960 (streams) x 1.25 (coding) = 3.65Tbit s$^{-1}$, assuming that channelization and beamforming are carried out at the stations.
***Assuming that the entire frequency range is channelized at 1 kHz resolution. This may be an overestimate since this narrow channel bandwidth would only be needed at the low end of the band.
****See Table 1 for assumptions and cautionary notes for this section.
Table 4 contains the key system performance factors for the sparse aperture arrays described in Table 3.

<table>
<thead>
<tr>
<th>Table 4: Key System Performance Factors for Sparse AAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antenna Efficiency</td>
</tr>
<tr>
<td>Projection loss</td>
</tr>
<tr>
<td>Minimum Elevation Angle</td>
</tr>
<tr>
<td>$T_{\text{rcvr}}$</td>
</tr>
<tr>
<td>$T_{\text{sys}}$ at 70 MHz*</td>
</tr>
<tr>
<td>130 MHz</td>
</tr>
<tr>
<td>300 MHz</td>
</tr>
<tr>
<td>450 MHz</td>
</tr>
<tr>
<td>Station Sys. Equiv. Flux Density (SEFD)</td>
</tr>
<tr>
<td>130 MHz</td>
</tr>
<tr>
<td>300 MHz</td>
</tr>
<tr>
<td>450 MHz</td>
</tr>
<tr>
<td>Field-of-View @ 70 MHz**</td>
</tr>
<tr>
<td>130 MHz</td>
</tr>
<tr>
<td>300 MHz</td>
</tr>
<tr>
<td>450 MHz</td>
</tr>
<tr>
<td>$A_e/T_{\text{sys}}$ @ 70 MHz, bore sight, all antennas</td>
</tr>
<tr>
<td>130 MHz</td>
</tr>
<tr>
<td>300 MHz</td>
</tr>
<tr>
<td>450 MHz</td>
</tr>
<tr>
<td>SSFoM @ 70 MHz</td>
</tr>
<tr>
<td>130 MHz</td>
</tr>
<tr>
<td>300 MHz</td>
</tr>
<tr>
<td>450 MHz</td>
</tr>
</tbody>
</table>

* $T_{\text{sys}} = T_{\text{rcvr}} + 60\lambda^{2.55}$

** $\text{FoV}_{\text{beam}} = \pi(1.3\lambda/d_{\text{station}})^2/4$

*** $A_e = \lambda^2/3$ at frequencies above sparse/dense transition.

7 Imaging Processing

The image processing part of the telescope “sticks out” as a major cost and performance driver. The system described here is based on the “traditional” model, using commercially available computing equipment to carry out all the processing beyond the correlator output. While this may not ultimately turn out to be a viable model, it serves here to illustrate that very large computing systems will be required to provide images over the very wide fields-of-view and large numbers of frequency channels in this representation of SKA Phase 1. Table 1 and Table 3 indicate that image processing computers capable of ~10 Petaflops will be needed for SKA Phase 1. This is based on calibration and imaging algorithms that can yield good results with $10^5$ floating point operations per input floating point number (flops per float) from the correlator. For Phase 1 imaging requirements, especially at frequencies below ~1.4 GHz, current algorithms require approximately $10^5$ flops per float [7].

The estimates of computing power in this document are necessarily crude and do not attempt to take into account any specific algorithms or calibration schemes. They also assume that computers
can be operated at 100% efficiency. Thus, they may be underestimates in some respects and overestimates in other respects.

In the era of Phase 1, high-performance computers (HPC) will take the form of millions of processors operating in parallel. The standard CPU-style processors are also likely to be assisted by specialized hardware (accelerated processing) to speed up some types of processing. The challenge for the SKA will be to adapt the algorithms to operate on the type of architectures that will be available in this class. Moreover, current code, originally developed for small numbers of processors, is unlikely to be easily adapted to these architectures, and new code written for them may not be as portable. The other challenge is the very high cost of machines, code development, and support for the machines, including their power consumptions, of which current estimates are of the order of 10s of megawatts. In addition, the radio astronomy imaging problem is unlikely to influence the design of the next generation HPC architectures.

Clearly the imaging problem is high parallelisable. For the Sparse Aperture Arrays each element of the beam-channel product is independent to first order, and for dishes each channel is also quasi-independent. However, more sophisticated algorithms are likely to require cross-access to data streams. Complex calibration schemes and developing algorithms for seamlessly stitching together large fields-of-view are potential examples. The solution will require a careful analysis of requirements and functions, starting at the correlator.

A combination of HPC machines, bespoke hardware solutions, and software is likely to be required. For example, the best solution could be a large array of much smaller computers, rather than a full-blown HPC architecture. This may require a sacrifice in imaging processing flexibility in order to reach the throughput required.

8 Non-Imaging Processing

Pulsar observations have been identified as key science for Phase 1. Observations fall into two main categories: timing of a set of known, suitable pulsars, and searching/surveying for pulsars that have not yet been identified. Pulsar searches require much more backend equipment and/or processing than timing, although timing requires well-understood on-axis polarisation parameters. Searching is likely to be done over a frequency range from ~400 MHz to ~3 GHz (higher would be better for some searches), while timing is likely to be done at the higher frequencies (2-3 GHz).

Other time-domain science (e.g. transient detection) is not key science for Phase 1, and will be carried out only if it can be done with minimum additional cost or effort.

Pulsar processing must be included in the Phase 1 system design. At this stage the requirements for pulsar searching and timing are still being assembled in detail (e.g. in each frequency band - range of dispersion measures, range of periods, range of acceleration parameters, etc.). Overall searching through-put requirements must also be defined. Although some design development has been done and cost estimates made, these cannot yet be traced to science requirements.
9 Phase 1 System Sensitivity

Figure 4 and Figure 5 show the system sensitivity over the entire frequency range, based on the parameters given in Tables 1-4. An additional curve is given for a receiver temperature of 100 K instead of the more conservative 150 K. Inspection of these figures shows that the frequency of peak sensitivity and the sensitivity at high frequencies is improved with Trcvr = 100 K, as would be expected. However, since no AA element/low-noise amplifier design for this bandwidth ratio (6.4:1) has yet been developed or tested, it is difficult to predict performance at the high-frequency end. Ultimately it could be better than 100 K (see [8]).

The peak sensitivity for the Sparse AAs occurs at about 115 MHz, the frequency at which the array becomes dense. The transition is assumed to occur when the effective collecting area of a single element is equal to the actual element packing density, determined in this case by packing 11,200 antenna elements into a 180-m diameter station. In practice, the transition will be smoothed by a gradual relaxation of coupling effects between elements. At frequencies less than the transition, the collecting area is constant but the increasing sky temperature decreases $A_e/T_{sys}$. At higher frequencies the effective collecting area of an element goes as $\lambda^2$.

The dishes show two different values of $A_e/T_{sys}$ for the two frequency ranges given. The reason is that a higher system temperature, by 10K, has been assumed for the lower band feed. This is to account for the probable increase of spillover in this band. Note that although these graphs show the Dish Array operating to 3 GHz, the array described in Table 1 operates only to 2 GHz. The reason is explained in Section 5.

In Figure 4, for the Sparse AA component, the values are for the peak value (bore-sight) in a single beam pointed at the zenith with no taper. An average reduction in sensitivity for use at elevation angles above 30 degrees will be 30-35% because of foreshortening. Tapering the beam to reduce...
sidelobes will reduce the sensitivity by an additional 30-50% for typical tapers [3]. However, tapering may not be needed for all types of observations.

![Survey Speed Figure of Merit (SSFoM): \((Ae/T_{sys})^2 \Omega_{FoV}\) in units of \(10^6 \text{ m}^4 \text{ K}^{-2} \text{ deg}^2\).](image)

The blue line (left) is for the Sparse AA with receiver temperature of 150K; the green line, for a receiver temperature of 100K; the red line (right) is for the Dish Array.

The dish sensitivity values are for the peak value (boresight) in the beam. Taper has already been accounted for in the aperture efficiency.

In Figure 5, AA foreshortening loss will increase the beam-size in one dimension. Formally this will increase the FoV, but reduce the effective area. The net effect on the SSFoM is a reduction, because the effective area enters as the square in the expression for the SSFoM. A similar effect occurs for taper loss.

10 The Route Forward

Synopsis

This section is based on [6], a discussion paper which outlines the plan to develop the technical concept for Phase 1, outlined in [1], while maintaining the vision of SKA Phase 2.

To realise the full SKA science case, SKA Phase 1 will be extended in terms of collecting area, baselines and performance, in particular for wide-field imaging, to SKA Phase 2. There will be up to ten times more total collecting area (depending on frequency), additional antenna stations at longer baselines, upgraded Central Signal Processing systems to accommodate the additional dataflow, and scaled up Software & Computing systems. A system description for Phase 2 can be found in [9].
In conjunction with the detailed design and pre-construction work, an Advanced Instrumentation Program (AIP) will be executed to further develop the new technologies (see [1] and descriptions below). These technologies will be assessed in terms of science impact, cost and technical readiness, and deployed in Phase 2 if shown to be feasible and cost-effective.

The phased construction of the SKA enables the project to make maximum use of advances in technology. Depending on the results of the AIP, the Phase 2 Baseline Design can be optimised for wide-field imaging by the inclusion of Phased Array Feeds on the dishes or a mid-frequency Aperture Array (AA-mid), and/or for ultra-wide band observations by the inclusion of Ultra-Wide Band Feeds on the dishes. These modifications will have a large system impact and will require a thorough impact analysis prior to deployment.

A limited number of Phase 1 subsystems will be obsolete by the construction of Phase 2, and will be decommissioned accordingly.

10.1 The Starting Point

- Phase 1 has been broadly defined in [1]. It contains:
  - a selection of the full set of science goals;
  - a description of Baseline Technology for the full SKA, consisting of dishes with single-pixel feeds, and sparse aperture arrays. Initially in Phase 1 the dishes will cover frequencies from 450 to 3000 MHz and the sparse aperture arrays will cover frequencies from 70 to 450 MHz.

- Phase 1 is considered as a “step” along the way to constructing Phase 2. However, because the full technology complement for Phase 2 will not be well defined (see below) a flexible plan for carrying out the development from Phase 1 to Phase 2 is required.

- The AIP will continue to develop innovative technology for:
  - Phased Array Feeds on the dishes (PAFs),
  - Mid-frequency Aperture Arrays (AA-mid), and
  - Ultra-wideband single pixel feeds on the dishes.

- As described in [1], the development of these technologies will run in parallel with the continued design and roll-out of Phase 1 until early 2016, when a decision will be made on whether they will be incorporated into Phase 2.
  - Note that by 2016, Phase 1 will be partially rolled out, and major changes in Phase 1 design will not be possible. However, it may be possible to equip Phase 1 with some AIP technology if the design and funding allows.
  - An assessment of whether PAFs can be incorporated into Phase 1 will be made in 2014.

- Full consideration of the cost and performance of the system described in [1] is likely to result in changes to the design, but the receptor technology will not change.
10.2 The Design Process leading from Phase 1 and Phase 2

The conceptual route to the final SKA system can be summarised as follows:

1. From the Phase 1 science requirements selected in [1] and further developed in a Phase 1 subset of the Design Reference Mission (DRM), the details of Phase 1 technical requirements will be derived. Initial Phase 2 system technical requirements will be developed, based on the science requirements contained in the DRM [2].

2. This will be followed by a full preliminary system design for Phase 1, informed by early preliminary designs for the Phase 2 options, so as to maximise the probability of extensibility to Phase 2.

3. The full detailed design for Phase 1 is the third major step. In conjunction with this, the development of the AIP components will be continued and incorporated into the final SKA design when sufficiently mature and cost-effective.

Although this overall process is inherently complex, it is closely linked to standard system engineering design procedures, shown diagrammatically in Figure 6. This figure shows the standard design steps interlinked at key points and approximately how they relate to the overall SKA project schedule. After User and System Requirements are obtained for Phase 1 and Phase 2, Phase 1 design/development becomes the pacing process. The AIP work must also be on track to meet the technology decision points in 2014 and 2016. Phase 2 system design work will be carried out at a slower pace aimed at a Preliminary Design following the completion of the AIP.

Figure 7 shows the process in more detail. The coloured boxes and labels in both Figure 6 and Figure 7 show the relationship of the process to standard system design steps. Note the four potential receptor technology options for Phase 2: 1) Baseline Technology, 2) Baseline Technology plus Phased Array Feeds, 3) Baseline Technology plus Dense Aperture Arrays, 4) Baseline Technology plus Wide-band Single-Pixel Feeds. The fraction of total collecting area and the frequency ranges for each technology in all options are to be determined as part of the process.

All three tracks in Figure 7 must be carried out and managed in parallel. This is essential so that sufficient design information is available from the initial stages of the Phase 2 track to maximise the likelihood that Phase 1 retains, as much as possible, compatibility with an extension to any of the Phase 2 options. This will require system design work for all four options.

Feedback Loops in the Preliminary Design Process

System design ultimately requires trading performance against cost. This is an iterative process. In Figure 7 the iterations are confined to the Preliminary Design part of the design process. In the feedback loop, the system scope will be adjusted, which means that not all of the requirements can
Figure 6: Part A: A standard 'V' diagram showing the phases of a large project. The dotted boxes correspond approximately to those in Figure 7. Part B: A schematic representation of the juxtaposition of the Phase 1, Phase 2 and AIP activities. The arrows between the 'V's illustrate critical information flows between the activities. More detail is shown in Figure 8, where SKA-specific activities are expanded. Major milestones are shown at the bottom.
Figure 7: A simplified view of the approach to the SKA design and construction. The dotted boxes correspond to activities in the schedule shown in Figure 8.
Figure 8: A preliminary schedule based on the activities shown in Figure 7. The double-headed arrows denote information flow from both user and non-science sources through to preliminary design. Major SKA milestones are shown at the bottom.
be met. This process will be tracked so that a clear record of the reasons for changes in scope is available.

A preliminary schedule is shown in Figure 8, including the three major Phase 1 project phases: The Preparatory Phase; the Preconstruction Phase; and the Construction, Verification, Commissioning, Integration and First Science Phase. Figure 8 also shows linkages and dependencies in the form of information flow from one stage of the process to the next.

10.2.1 Stepping Through Phase 1

- In Figure 7, the Phase 1 (Concept, Definition box) begins with defining the Phase 1 science requirements, extracted from the applicable subset of the DRM. Technical requirements are assembled from both science (user) and non-science sources.

- Using information from a preliminary Phase 2 analysis, a Phase 1 preliminary system design that considers extensibility to Phase 2 will be carried out.
  - Part of the Phase 1 design process must include a categorisation of elements/sub-systems in the following way: 1) those too valuable to abandon must be designed/constructed for Phase 2, 2) those with short lifetimes or that impede Phase 2 installation must be removed or abandoned from the built Phase 1 system, 3) software, algorithms and much of the underlying hardware that will be continuously developed and upgraded for at least a decade. Examples of category 1 could be dishes; category 2, correlator and/or non-visibility processor; category 3, image processing software.

- Anticipating the build-out to Phase 2 in this way will consume resources (time and funds); the cost of extensibility to Phase 2 will have to be balanced against the cost of Phase 1 performance.
  - Examples: planning for additional fibre routing from intermediate baseline stations for either PAFs or Dense Aperture Arrays (DAAs); provision of additional power for dishes for PAFs; opening up an additional core for DAAs; provision of additional power for the entire site.

  - As a result, it must be accepted that Phase 1 will cost more than if it were the final product, and that some things may be discarded afterwards or be re-done.

- Phase 1 will be cost-capped. Science/cost/performance trades will be carried out, using the methods outlined in [4] and [5]. Phase 1 scope will be adjusted to fit the cost and to satisfy as many of the initial requirements as possible, including aspects related to extension to Phase 2.

- This leads to a detailed system design for Phase 1. Figure 7 and Figure 8 show the continuation of the process through detailed design, construction and operations for Phase 1, linked eventually to the build-out to Phase 2.

10.2.2 Stepping Through Phase 2

- Figure 7 (centre column) shows the process by which the Phase 2 options are initially evaluated. For Phase 2 the entire DRM is the source of science requirements.

- To inform the Phase 1 design process, it will be necessary to draft initial technical requirements for the four system options for Phase 2 (Baseline Technology alone + the three AIP options in combination with the Baseline Technology), and carry out early preliminary system designs for
each of the options. A first-order optimisation of each of the four options will be carried out to provide input to the Phase 1 design on what may have to be supported (see Figure 7).

- This optimisation is mainly driven by science requirements, but must also take into account cost and performance information, recognising that only very rough estimates of capital and operating costs will be available. Projections of the science performance of the three options would be evaluated using “observing time performance factors (OTPF)”, as described in [5]. The goal is not to define every aspect of the Phase 2 system design, but to provide sufficient guidance on what must be anticipated in the design of Phase 1.

- As shown in the definition stage for Phase 2 in Figure 7, Phase 2 requirements will continue to be developed in parallel with design work being done on Phase 1. This will be necessary to prepare for the decision on technology selection in 2016.

- The preliminary design phase shown in Figure 7 considers all the mature technology options in a system design in which the use of the technologies is fully optimised. In particular, the fractions of collecting area, the frequency ranges and the array configuration for each technology will be optimised to maximise science performance under cost constraints. This will be an iterative trade-off process.

  - More detail on actual Phase 1 system performance and cost will be available from the initial rollout of Phase 1, which will use Baseline Technology. This information flow is shown in Figure 7 as a connection from the Phase 1 to the Phase 2 track.

  - Figure 8 shows the preliminary design phase running for about three years. This length of time will be needed to include one or more AIP technologies in the system design. If Baseline Technology is adopted for Phase 2, preliminary design is likely to be shorter.

- Detailed design for Phase 2 will be required after this selection is made.

- In much less detail, Figure 7 shows the process continuing to the build-out and operation of Phase 2.

### 10.2.3 Stepping Through the Advanced Instrumentation Programme

- The AIP will be designed to build up the level of maturity of each of the three technologies, as well as anticipate their utilisation in Phase 2.

- The three technologies (PAFs, DAAs and WBSPFs) will require different approaches in detail, but they will share the following general steps (see Figure 7):

  - Analyse the Phase 2 requirements so as to maximise the potential of the new technology to enhance system performance, achieve more of the initial system requirements and/or reduce cost, as compared with the baseline.

  - Develop a preliminary design as if the AIP technology were to be used at the element or subsystem level of the system.

  - Carry out a verification program to test the level of achievement of requirements, to develop performance/cost models for the AIP technology, and to ascertain remaining risk.

- The expected outcomes of the verification programmes are hard data on the following:

  - Potential in-system performance, projected from tests, and measured against requirements.
- Good estimates of volume manufacturing, deployment and maintenance costs;
- Verified operational models in the intended physical environment, including power, temperature control, support, etc.
- Verified calibration models needed to achieve performance, including estimates of limitations placed by systematic and other residual errors on performance;
- Risk assessments.

- The decision to include the technology for Phase 2 will turn on the level of achievement of these outcomes and the remaining risk. This will be made in 2016 for Phase 2, but an assessment of the possibility of utilising PAFs on dishes will be made separately in 2014 for Phase 1.

11 Conclusion

A well resourced project could meet the Phase 1 science performance goals described in [1] on the schedule outlined in Figure 8, and continue on to build the full SKA later. The high-level plan described in Section 10 is a means of systematically approaching the design, trade-off and costing work needed to accomplish this result.