




## SDP Memo 100: Guide to the SDP Cost Model

Document number..... SKA-TEL-SDP-0000201  
 Document Type..... MEMO  
 Revision..... 01  
 Author..... F. Graser  
 Release Date..... 2019-03-29  
 Document Classification..... Unrestricted

Lead Author	Designation	Affiliation
Ferdl Graser	SDP Systems Engineer	SARAO
Signature & Date:		

# SDP Memo Disclaimer

The SDP memos are designed to allow the quick recording of investigations and research done by members of the SDP. They are also designed to raise questions about parts of the SDP design or SDP process. The contents of a memo may be the opinion of the author, not the whole of the SDP.

## Table of Contents

<b>1. Introduction</b>	<b>4</b>
<b>2. Sheets</b>	<b>4</b>
2.1. Header	4
2.2. SAFe SDP Construction Summary	4
2.3. SDP Opex Summary	4
2.4. SDP Construction Cashf	4
2.5. Assumptions & Inputs	5
2.6. Hardware (Totals)	5
2.7. Software	5
2.8. System level tasks	6
2.9. ART resource profile	6
2.10. Compute Rack hardware	6
2.11. operations phase cost	6
2.12. Buffer hardware	6
2.13. SDP Infrastructure	6
2.14. Interconnect System	7
2.15. Hierarchical Storage	7
2.16. System Specification Summary	7
2.17. OST (112) Summary	7
2.18. phased deployment scenarios	7
2.19. hw refresh compare	8
2.20. Ingest Nodes	8
<b>3. References</b>	<b>9</b>
3.1. Reference Documents	9

## List of Figures

## List of Tables

## List of Abbreviations

PI	Program Increment
SDP	Science Data Processor
WBS	Work Breakdown Structure

# 1. Introduction

This document is a guide on how to use the SDP Cost Model spreadsheet [RD01]. This guide does not duplicate the information contained in the SDP Cost Basis of Estimate [RD02], the SKA Construction Cost Estimation Guidance [RD03] and the SKA Operations Cost Estimation Guidance [RD04].

The guide uses specific row, column and cell references (i.e. D6) and therefore this guide is specific to revision 5 (15-03-2019) of the SDP Cost Model spreadsheet [RD01]. The row, column and cell references may be different in older or newer revisions of [RD01].

Text in *italics* refers to sheet names, column names or row names in the SDP Cost Model spreadsheet [RD01].

## 2. Sheets

### 2.1. Header

The header sheet contains the relevant document and change control information.

### 2.2. SAFe SDP Construction Summary

Summary of construction (CAPEX) costs according to the Cost Guidance [RD03]. Costs are aggregated according to the WBS for SDP and the prescribed columns (D-G). Hardware costs are mapped from the *Hardware (Totals)* sheet. Agile Release Train costs are mapped from the *Software* sheet and the *System level tasks* sheet.

### 2.3. SDP Opex Summary

Summary of operations phase costs (OPEX) according to the Operations Cost Guidance [RD04]. Costs are aggregated according to the WBS for SDP and the prescribed columns (D-F). Hardware costs are mapped from the *Hardware (Totals)* sheet. Agile Release Train costs are mapped from the *Software* sheet. Non-labour costs (column F) are aggregated from the *operations phase cost* sheet.

The *Lumped Costs* for the full system refresh after 5 years is calculated manually adjusting the *SKA1 SDP final deployment start date* (B11) on the *Assumptions & Inputs* sheet and then copying the value from the *Hardware (Totals)* sheet, e.g. to calculate the *SDP Compute Hardware* (full system refresh) value for the year 2030 follow these steps:

1. Change *SKA1 SDP final deployment start date* (B11) on the *Assumptions & Inputs* sheet to 2030-02-01 (add 5 years).
2. Copy the value in D6 (*SDP Compute Hardware*) on the *Hardware (Totals)* sheet and paste into J9 on the *SDP Opex Summary* sheet. Repeat this for the LOW value (copy C6 to J13)
3. Change *SKA1 SDP final deployment start date* (B11) on the *Assumptions & Inputs* sheet to back to its previous value (2025-02-01).

### 2.4. SDP Construction Cashf

Summary of construction (CAPEX) cashflow according to the Cost Guidance [RD03]. Costs are aggregated according to the WBS for SDP. Hardware costs are mapped from the *Hardware (Totals)* sheet according to the 3 hardware deployments (AA1, AA3 & AA4) for SDP.

Agile Release Train costs are mapped from the *Software* sheet and the *System level tasks* sheet. The cost for each month is calculated by looking up the required effort (number of people) for a particular PI on the *ART resource profile* sheet. The PI to month mapping is given in row 28 on the *SDP Construction Cashf* sheet.

## 2.5. Assumptions & Inputs

This sheet contains all the input variables for scaling the quantity of hardware and the hardware quantity and forward projection (Moore's Law) as well as the labour rates, travel rates, assumptions to calculate the quantity and cost of the AA1 and AA3 deployments and operational cost assumptions. The cells in bright green are typically used to adjust the hardware quantities and forward projection.

Any power usage values given or calculated on this sheet are for reference only as the official SDP power figures are calculated in a different spreadsheet (SDP Load List).

## 2.6. Hardware (Totals)

This sheet provides an overview all the hardware CAPEX costs and calculates the contingency factors and costs. *Compute Rack* costs (C8:D13) are calculated from the *Compute Rack hardware*, *Buffer hardware*, *SDP Infrastructure* and *Assumptions & Inputs* sheets. *Interconnect System* costs (C15:D21) are calculated from the *Interconnect System* and *Assumptions & Inputs* sheets. *Management Compute Islands* costs (C23:D23) are calculated from the *Compute Rack hardware* sheet. *SDP Preservation Hardware* costs (C26:D27) are calculated from the *Hierarchical Storage* and *Assumptions & Inputs* sheets.

The cost of the AA1 and AA3 deployments are calculated from the *SDP Compute Hardware* (C8:D23) costs and scaled according to the scaling factors calculated on the *Assumptions & Inputs* sheet.

The costs calculated for *Observatory Support Tools* at the bottom of the sheet are for reference only and they are not included in the *SAFe SDP Construction Summary* sheet.

## 2.7. Software

The *Software* sheet contains all the software development effort (and travel) that will be carried out by the Agile Teams. Costs are calculated based on the labour rates given on the *Assumptions & Inputs* sheet and contingency factors and costs are also calculated. For each software module (column A) the total effort (in person years) is calculated from the *Effort duration* (columns AA & AB) and the individual roles (N:Z), e.g. 1.00 *Intermediate Engineer (Eng:Int)* for 3 years equals 3.0 person years of effort. For each software module (column A) a *Daily rate* (column AC) is calculated as the sum of the daily rate for each role. The cost excluding contingency (column C) is calculated by multiplying the *Daily rate*, *Effort Duration* (in years) and *Work Days Per Year* (on the *Assumptions & Inputs* sheet).

The costs calculated for *Observatory Support Tools* at the bottom of the sheet are for reference only and they are not included in the *SAFe SDP Construction Summary* sheet.

## 2.8. System level tasks

The *System level tasks* sheet contains the effort (and travel) for the Agile Release Train Program Team and System Team. Cost and effort is calculated in the same way as on the *Software* sheet.

## 2.9. ART resource profile

This sheet provides the resource profile for the Agile Release Train and this is used for the cash flow estimate. The *Resource loading* is given per PI (3 month period) for each team category as *number of teams* working during that PI. The *Resource profile* (number of people per PI for each team category) is calculated from the normalised *Resource loading* (*number of teams* teams divided by the average number of teams over the total period), total effort (person years), total number of PIs and number of PIs per year (4).

## 2.10. Compute Rack hardware

This sheet calculated the cost of a compute rack for MID and LOW as well as the cost of the management (*MGMT*) compute rack. The total number of nodes required is calculated from the *Performance (Sustained FLOPS)* required and *Compute Efficiency* (on the *Assumptions & Inputs* sheet), the *Peak Performance per GPU* (forward projected GPU performance) and the number of *GPUs per node*.

Note that the *Storage Node* section at the bottom of the sheet is not used in the compute rack calculation, but rather in the buffer cost calculation. The *Receive and Real Time Node* section at the bottom of the sheet is currently not used anywhere.

Any power usage values given or calculated on this sheet are for reference only as the official SDP power figures are calculated in a different spreadsheet (SDP Load List).

## 2.11. operations phase cost

This sheet contains the effort (and cost) needed to support the hardware and the cost of hardware maintenance supplied by the hardware vendor. The effort (and its cost) is calculated in the same way as on the *Software* sheet. The cost of the hardware maintenance (supplied by the hardware vendor) is calculated from the CAPEX cost of the hardware (given on the *Hardware (Totals)* sheet) and the maintenance cost factor given on the *Assumptions & Inputs* sheet. Although maintenance cost factors are given for 3, 5, and 7 years, the *5 yr maintenance* values are used.

## 2.12. Buffer hardware

This sheet calculated the cost of Buffer Storage (B35 and C35) based on the *Buffer Size* required, the *Price per GB* for various storage media (given on the *Assumptions & Inputs* sheet) and the *Hot/Cold Buffer Ratio*. Performance values given and calculated on this sheet are for reference only and can be used as a sanity check to make sure that the design is feasible, but the performance values don't contribute to the cost of Buffer Storage.

Any power usage values given or calculated on this sheet are for reference only as the official SDP power figures are calculated in a different spreadsheet (SDP Load List).

## 2.13. SDP Infrastructure

This sheet calculates the cost of *Rack Infrastructure* for all racks required.

## 2.14. Interconnect System

This sheet calculates the cost of the Low Latency Network, Bulk Data Network and the Preservation and Delivery Network.

The number of switches (32 port, 100GbE) needed for the *Bulk Data Network* is calculated in the *BDN Calculations* section (F28:J40) according to the network topology (explained in [RD02]). The cost of the switches is calculated from the total number of ports (32 ports per switch) and the cost per port (given on the *Assumptions & Inputs* sheet F35:I39). The cost of *SMF Cables & transceivers* is calculated from the number of SaDT links and the cost per *SMF Cables/Transceivers Pair*. The cost of QSFP cables is calculated from the number of *DownLinks* and cost per *QSFP Cable*.

The cost of the Preservation and Delivery Network is calculated in the *Archival Interconnect System* section (A43:C54). The number of switches (56 port, 25GbE) needed is calculated from the number of Compute Racks (*Compute Rack hardware* sheet B10 & C11) and the number of aggregation switches. The cost of the switches is calculated from the total number of ports (56 ports per switch) and the cost per port (given on the *Assumptions & Inputs* sheet I36). The cost of cables is calculated from the *number of 25GbE Uplinks*, *number of 25GbE DownLinks* and cost per *SFP+ or QSFP Cable*.

The cost of the Low Latency Network is calculated in the *Low-Latency Interconnect System* section (A94:D128). Costs are calculated for *Over-Subscription* ratios of 1:1, 3:1 and 8:1. The *Low-Latency Over-Subscription* ratio is set on the *Assumptions & Inputs* sheet B49. The number of switches needed is calculated from the *Number of nodes*, size of the switches and the network topology (2 level). The total cost of the Low Latency Network is calculated from the number of switches and cables and the cost of the switches and cables.

The section *Per-SDP Inter-Island Interconnect System (Low-Latency Switch)* (row 59-90) is no longer used and can be ignored.

## 2.15. Hierarchical Storage

This sheet calculates the cost of Long Term Storage (*SATA HDD*) and Long Term Staging Storage (*SATA SSD*). The cost of a complete storage pod (with either SATA hard drive or SATA solid state drive media) is calculated and then converted to a cost per PetaByte of storage. Cost projection is done using *Storage density growth rate (CAGR)* projections from industry roadmaps by scaling the capacity per media (*SATA HDD* or *SATA SSD* disk) to the desired future date.

## 2.16. System Specification Summary

This sheet is not used for cost calculation purposes. It provides a summary of the hardware specification as costed.

## 2.17. OST (112) Summary

This sheet provides an estimate of the *Observatory Support Tools* costs. This estimate is not officially part of the SDP cost submission and is for reference only.

## 2.18. phased deployment scenarios

This sheet is used to analyse the cost of phased deployment scenarios for SDP hardware. The data in this sheet is based on values calculated by the other sheets in this spreadsheet. This data was last

updated on 02/06/2017 and therefore is out of date. To generate new data for analysis follow these steps:

1. Change the following parameters on the *Assumptions & Inputs* sheet to the desired value:
  - a. *Performance (Sustained FLOPS)*
  - b. *Buffer size (binary)*
  - c. *SKA1 SDP final deployment start date*
2. A new hardware cost is now calculated and shown in *Hardware cost output (G6)*, copy this value (and other values as needed) into the tables below.
3. Repeat step 1 and 2 in order to populate the analysis tables with the hardware cost for various dates, performance and buffer sizes as needed.
4. Set the parameters in step 1 back to the original values

## **2.19. hw refresh compare**

This sheet is used to analyse the total lifecycle cost of SDP hardware (excl power and labour) for a 3 year and 5 year hardware refresh cycle. The *Hardware refresh cost* for various dates is generated by setting the *SKA1 SDP final deployment start date* (on the *Assumptions & Inputs* sheet) to the desired date and copying the values from shown in column B. Since spares are bought up-front, the cost is taken from the previous or first hardware deployment.

At the bottom of the sheet a cost comparison over a 5 year lifecycle is given comparing a 3 year and 5 year refresh cycle.

## **2.20. Ingest Nodes**

This sheet is not used for cost calculation purposes. The sheets was used to analyse a possible variant of the costed hardware concept design using special Ingest nodes, but this is not part of the SDP baseline and therefore can be ignored.



## 3. References

### 3.1. Reference Documents

- RD01 SKA-TEL-SDP-0000043, SKA1 SDP Cost Model, Rev05
- RD02 SKA-TEL-SDP-0000046, SKA1 SDP Cost Basis of Estimate, Rev05
- RD03 SKA-TEL-SKO-0000845, Construction Cost Estimation Guidance, Rev01
- RD04 SKA-TEL-SKO-0000830, Operations Cost Estimation Guidance, Rev 01