SKA Design at ICRAR & Industry Involvement

Markus Dolensky

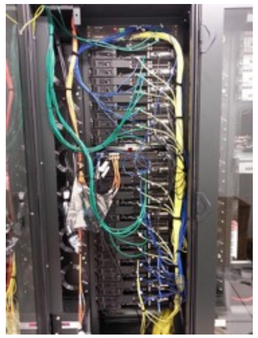
International Centre for Radio Astronomy Research

Curtin University

THE UNIVERSITY OF WESTERN AUSTRALIA

Government of Western Australia
Department of the Premier and Cabinet
Office of Science

SKA Signal & Data Path along Design Elements



LFAA
DISH

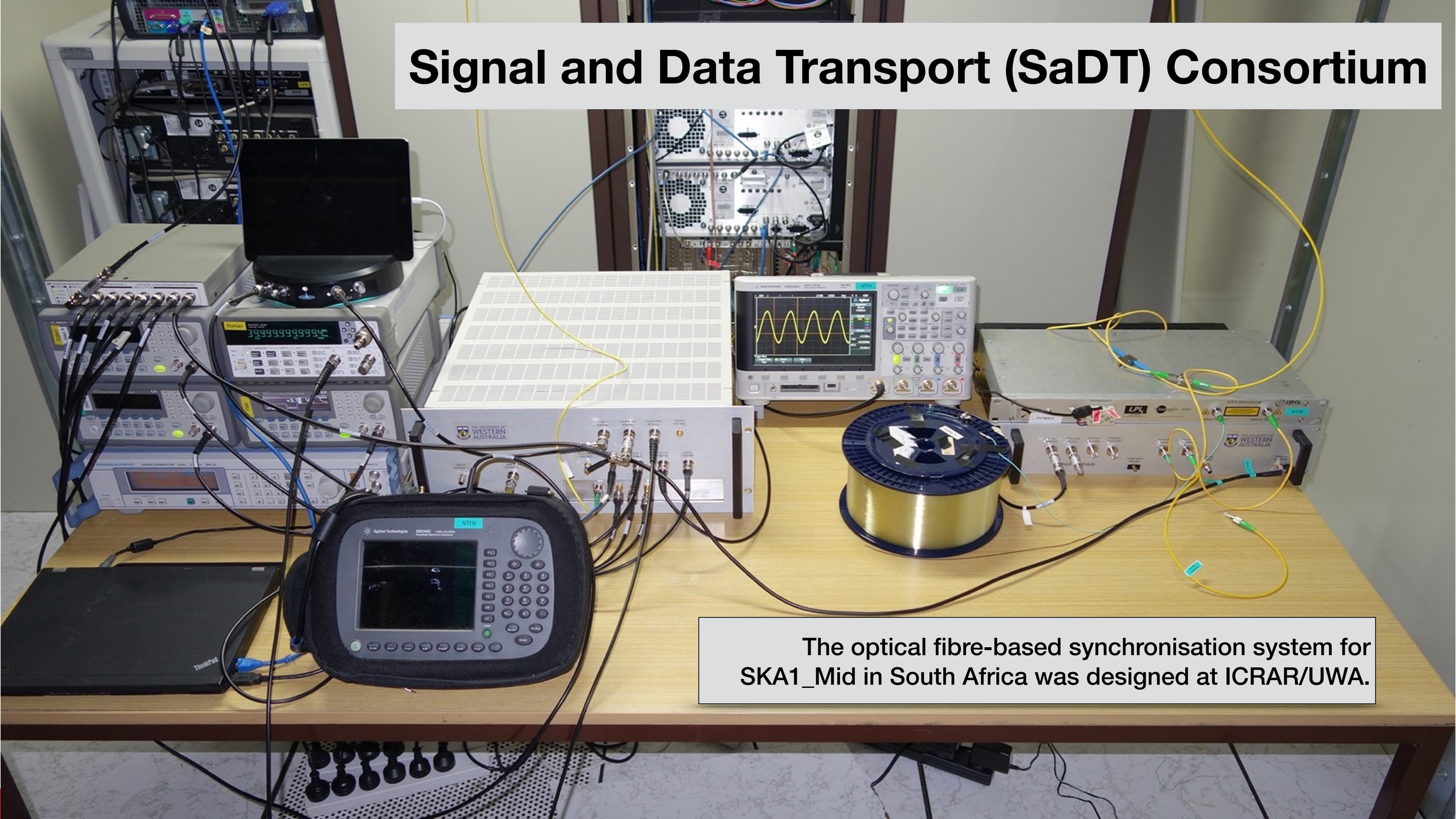
Correlator
CSP

Data Intensive
Astronomy

Science Data
Processor

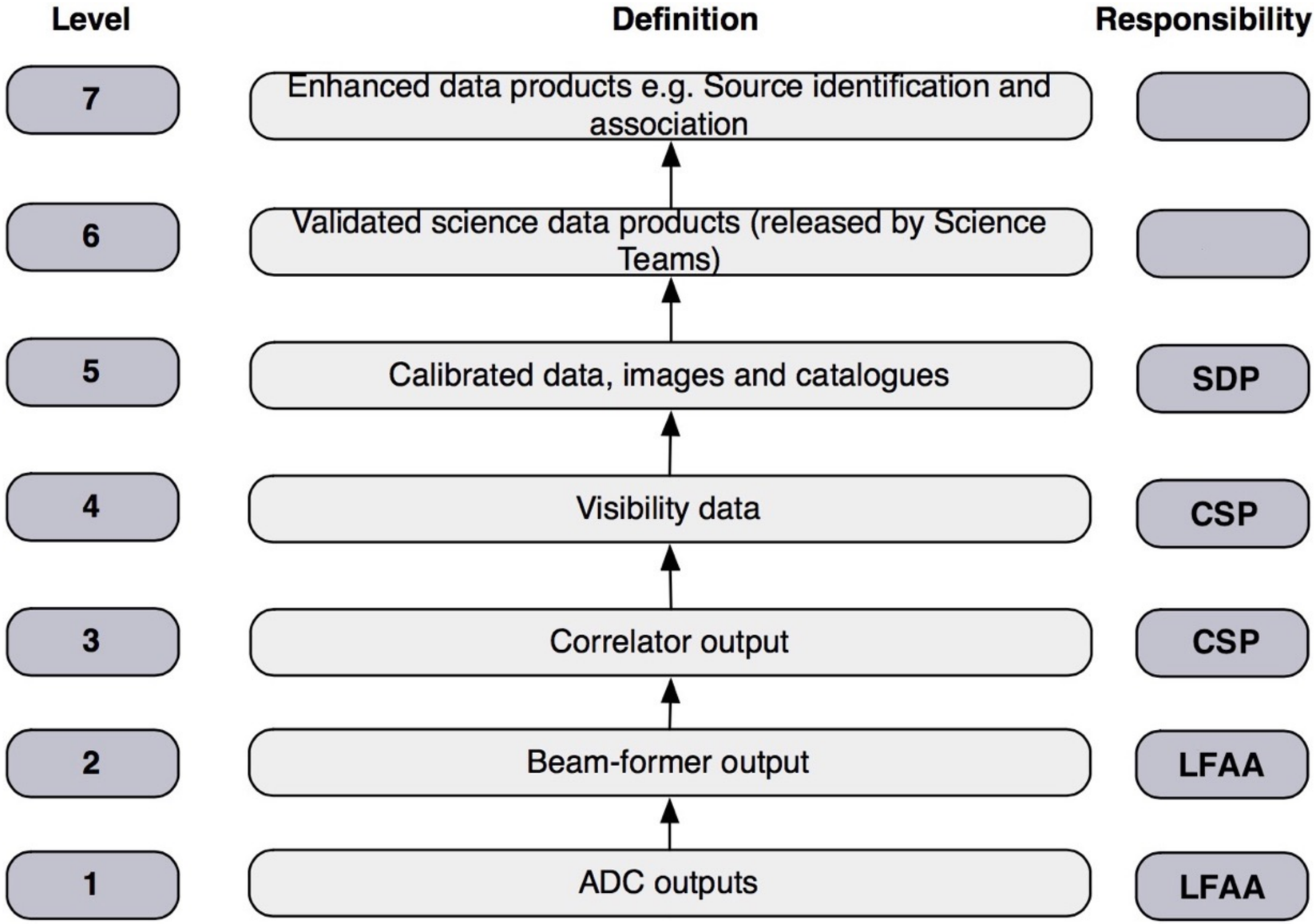
Long-Term
Storage

Signal and Data Transport (SaDT) Consortium



The optical fibre-based synchronisation system for SKA1_Mid in South Africa was designed at ICRAR/UWA.

Data Product Levels



The SDP Consortium

- 23 Partners in 11 countries
  - Australia, Canada, China, France, Germany, Portugal, Netherlands, New Zealand, South Africa, Spain, United Kingdom (UCam Lead)
- 9 Work Areas: at ICRAR 22 FTE over 3 years
  - leads data layer WP, contributes to ARCH, COMP, LMC, MGT, SE
- 5 Supercomputing Centres
  - Pawsey (Perth), HPCS (Cambridge), CHPC (Cape Town), BSC (Barcelona), Forschungszentrum Jülich
- Pre-Construction is self-funded by member countries

Project Rialto

Full federal grant title:
“Finalising the SDP Design and Preparing for SKA Construction”.

Internal project name: Rialto

as in Ponte di Rialto in Venice, because it …

- bridges extended precon with next SKA phase,
- integrates DALiuGE with ASKAPsoft,
- spans data flow design to SKA Regional Centres.

SKA SDP Pre-Construction at ICRAR

Focus:

1. Execution Framework
2. Preservation System

Furthermore:

Architecture, Compute Platform, PM, SE, LMC

SDP Preservation - Storage Volume

	LOW	MID	
Long term storage volume	110	340	PB
Medium Performance Buffer	30	100	PB
Science Data Volume (HPSOs)	55	170	PB
8 TB drives LTS (0.11+0.34 EB)	13750	42500	Unit
26 TB drives LTS	4200	13000	Unit

=> need high resilience against bit errors

Preservation Design

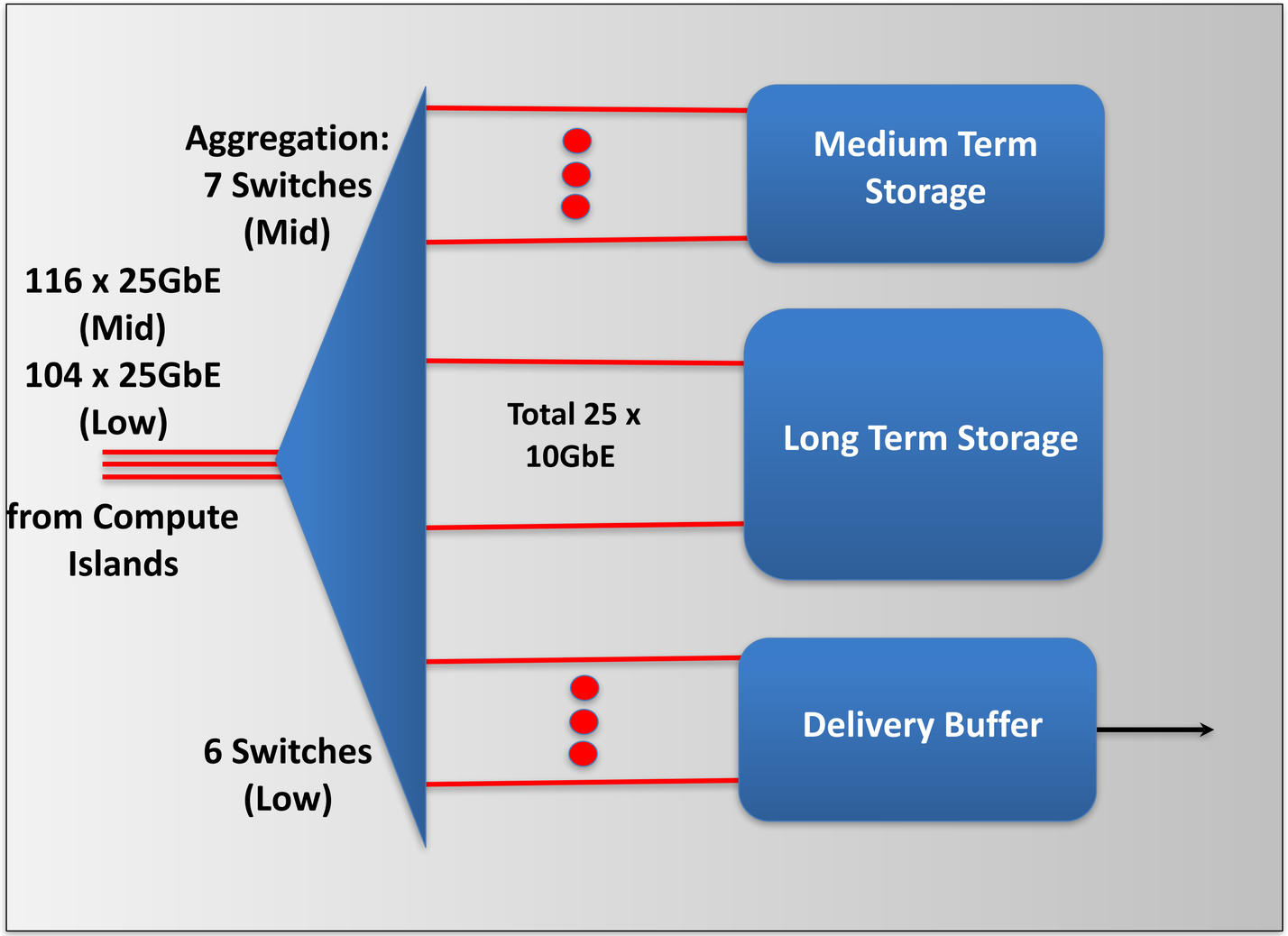
- medium and long term storage system
- for intermediate and final science data products
- capacity: 0.13 and 0.45 Exabyte for MTS and LTS
- ingest bandwidth: 200 Gbps
- populates science product catalogue
- packages, indexes, stores and backs up data products
- is extensible, interoperable, resilient and scalable

- no archive: (external) users have no direct access

Science Data Products & IVOA Support

	IVOA Standard Protocols and Model Candidates		
Data Product Type	**Discovery**	**Access**	**Data Model**
Image cube	TAP, SIAv2, DataLink	SODA	ObsCore, NDimCubeDM
UV-grid	TAP, SIAv2, DataLink	SODA	ObsCore, NDimCubeDM
Calibrated visibilities	TAP, SIAv2, DataLink	SODA	ObsCore, NDimCubeDM
GSM/LSM catalogue	TAP, DataLink	TAP	STC, CharDM, PhotDM
Pulsar timing solution	TAP, DataLink	TAP	ObsCore, TimeSeriesDM
Pulsar & transient cand.	TAP, DataLink	TAP	(PSRFITS)
Science product catalog.	TAP, DataLink	TAP	RegTAP, STC, CatalogDM

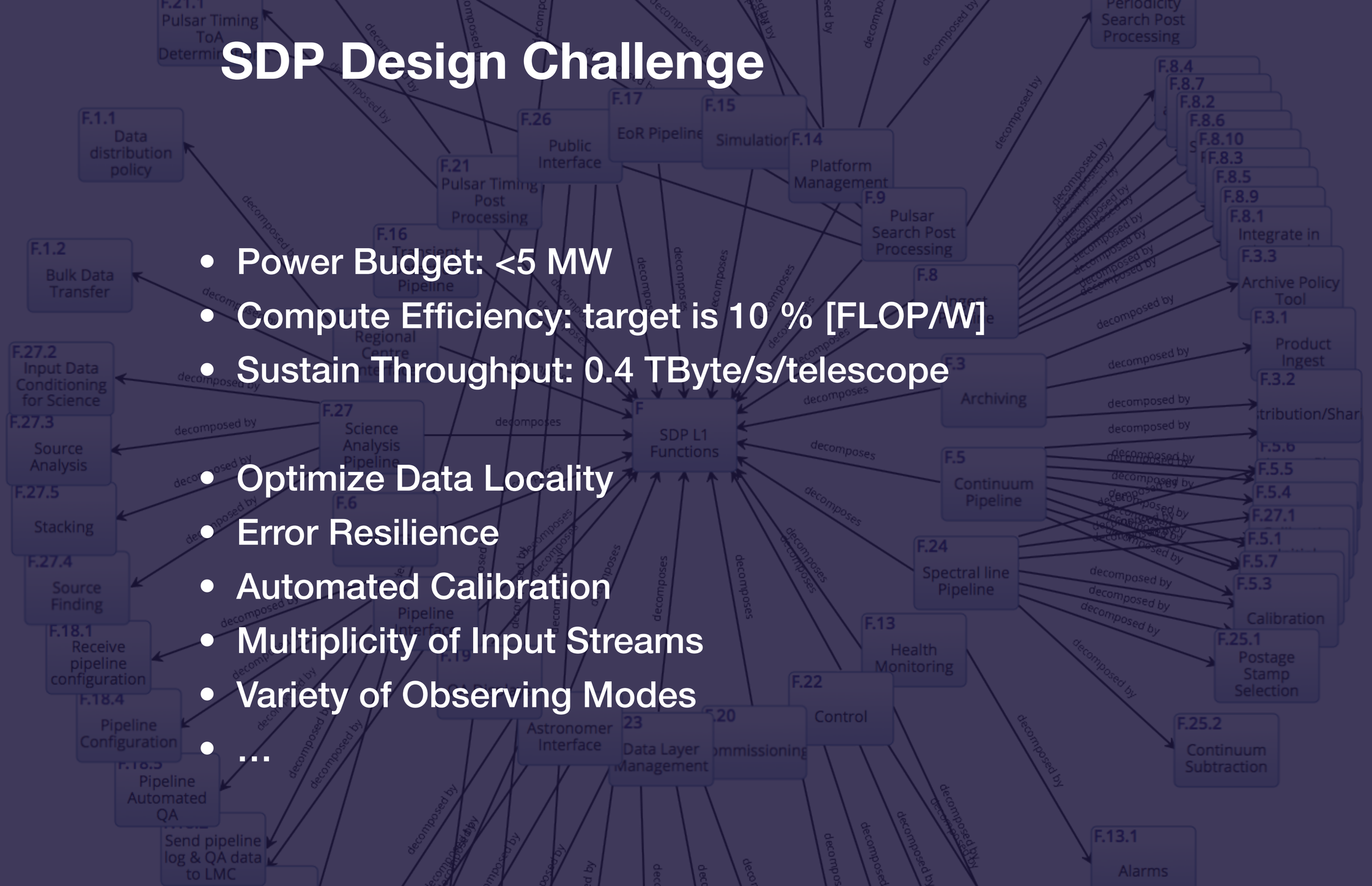
Preservation & Delivery Interconnect



SDP Design Challenge

- Power Budget: <5 MW
- Compute Efficiency: target is 10 % [FLOP/W]
- Sustain Throughput: 0.4 TByte/s/telescope

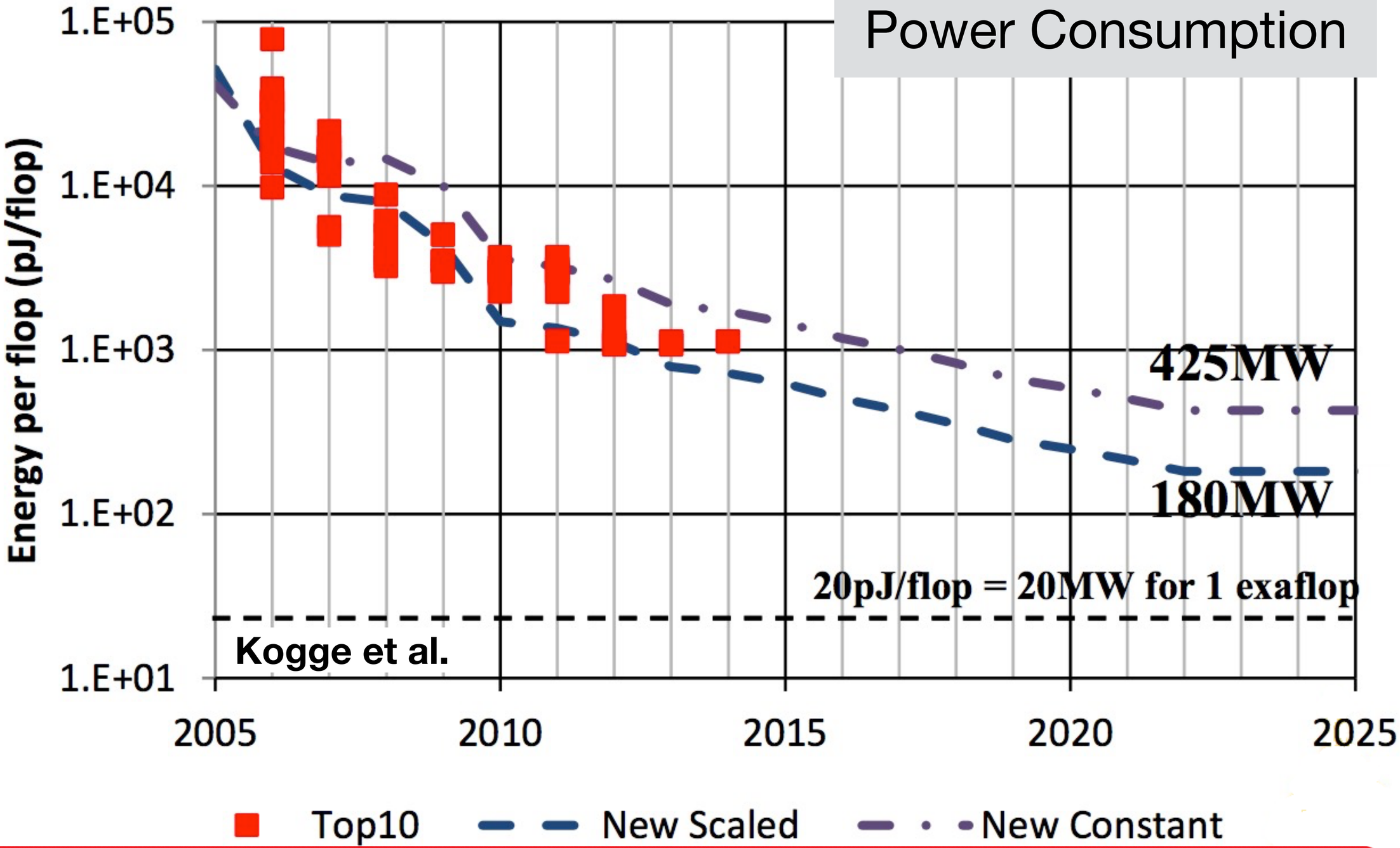
- Optimize Data Locality
- Error Resilience
- Automated Calibration
- Multiplicity of Input Streams
- Variety of Observing Modes
- …



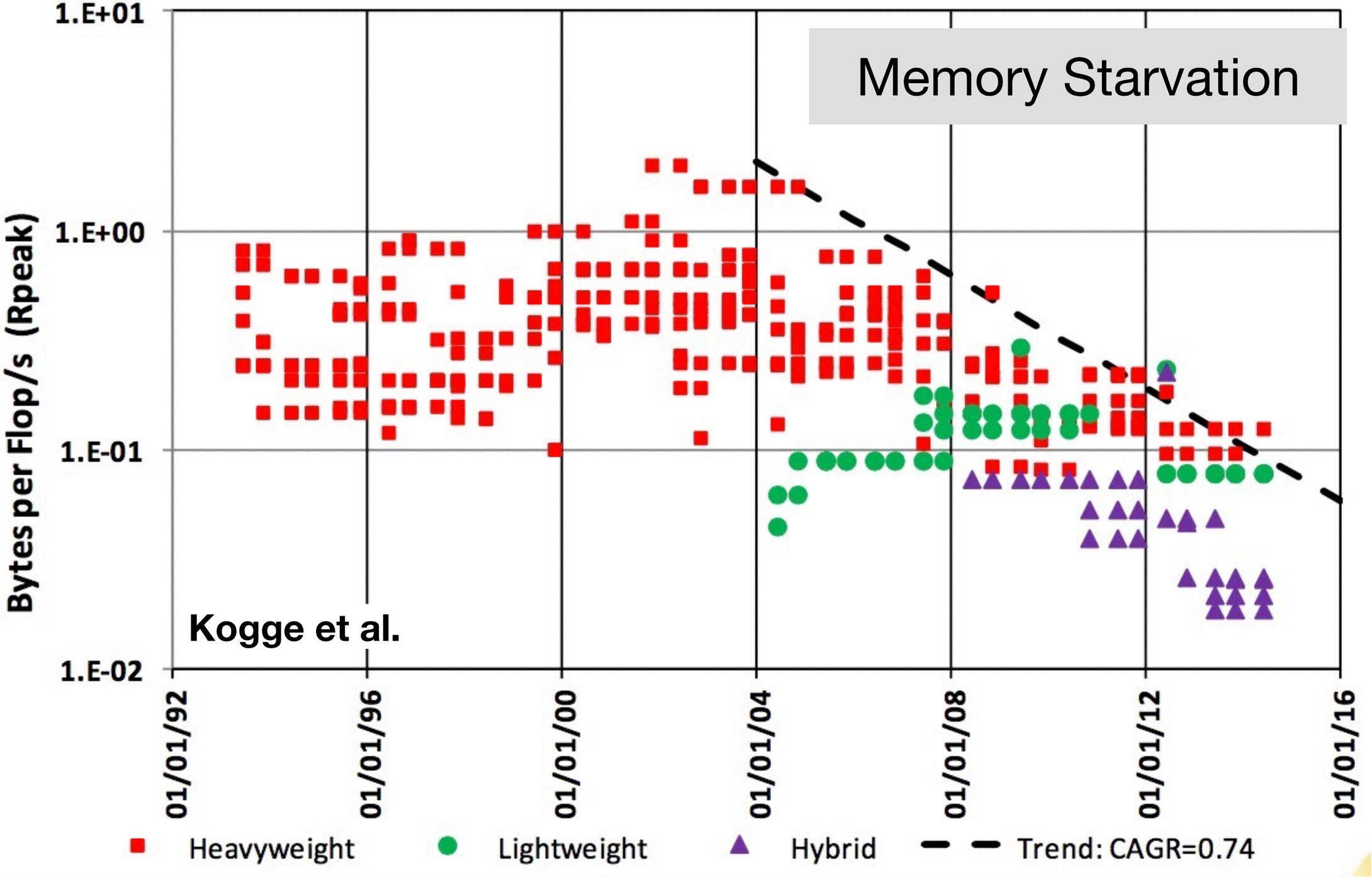
Size after Rebaselining

Telescope	SKA1 Low rebaselined	SKA1 Mid rebaselined
Antennae / Dishes	130000	200
max. Baseline [km]	65	150
Frequency channels	65,536	65,536
Complex Correlations / s	3.8E+10	6.4E+10
Image side length [pix]	16000	20000

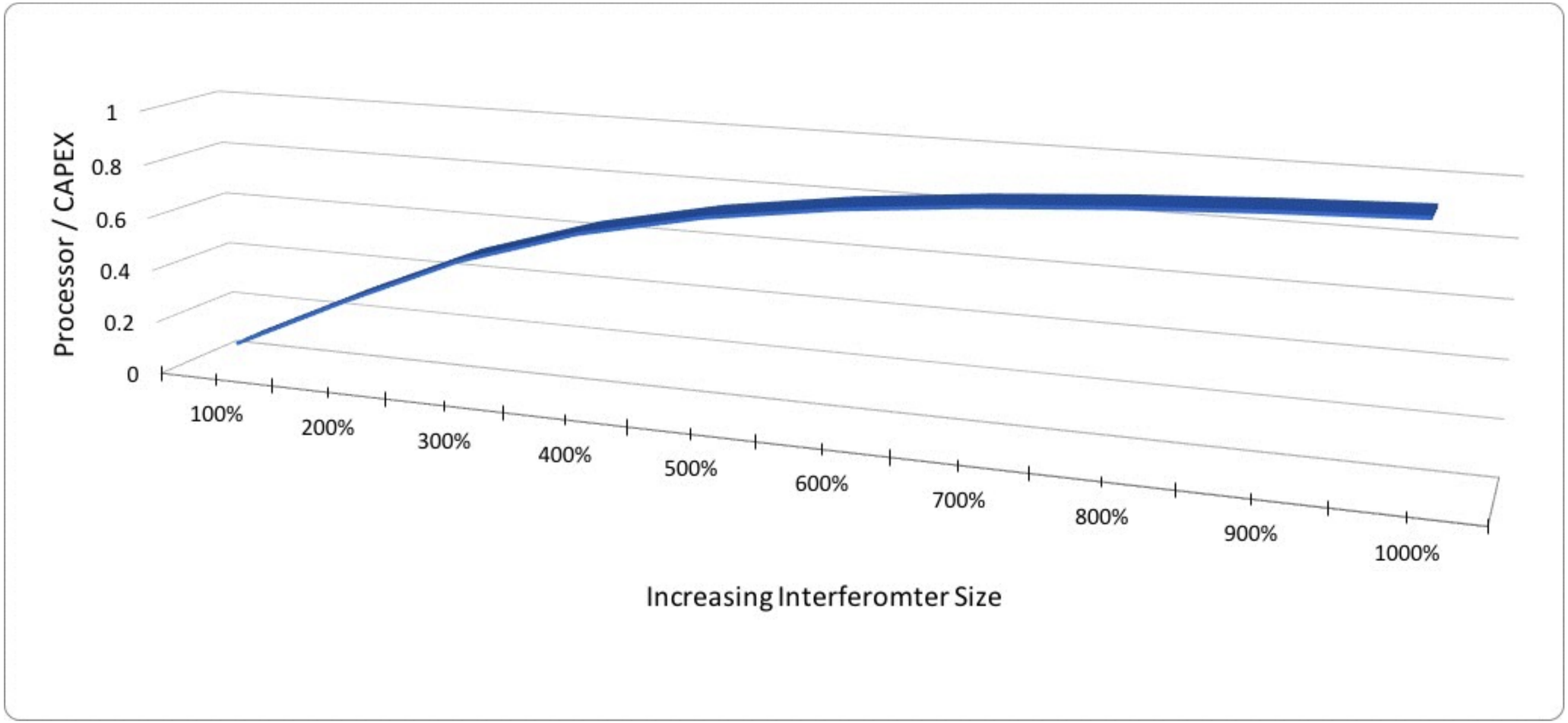
Power Consumption



Memory Starvation

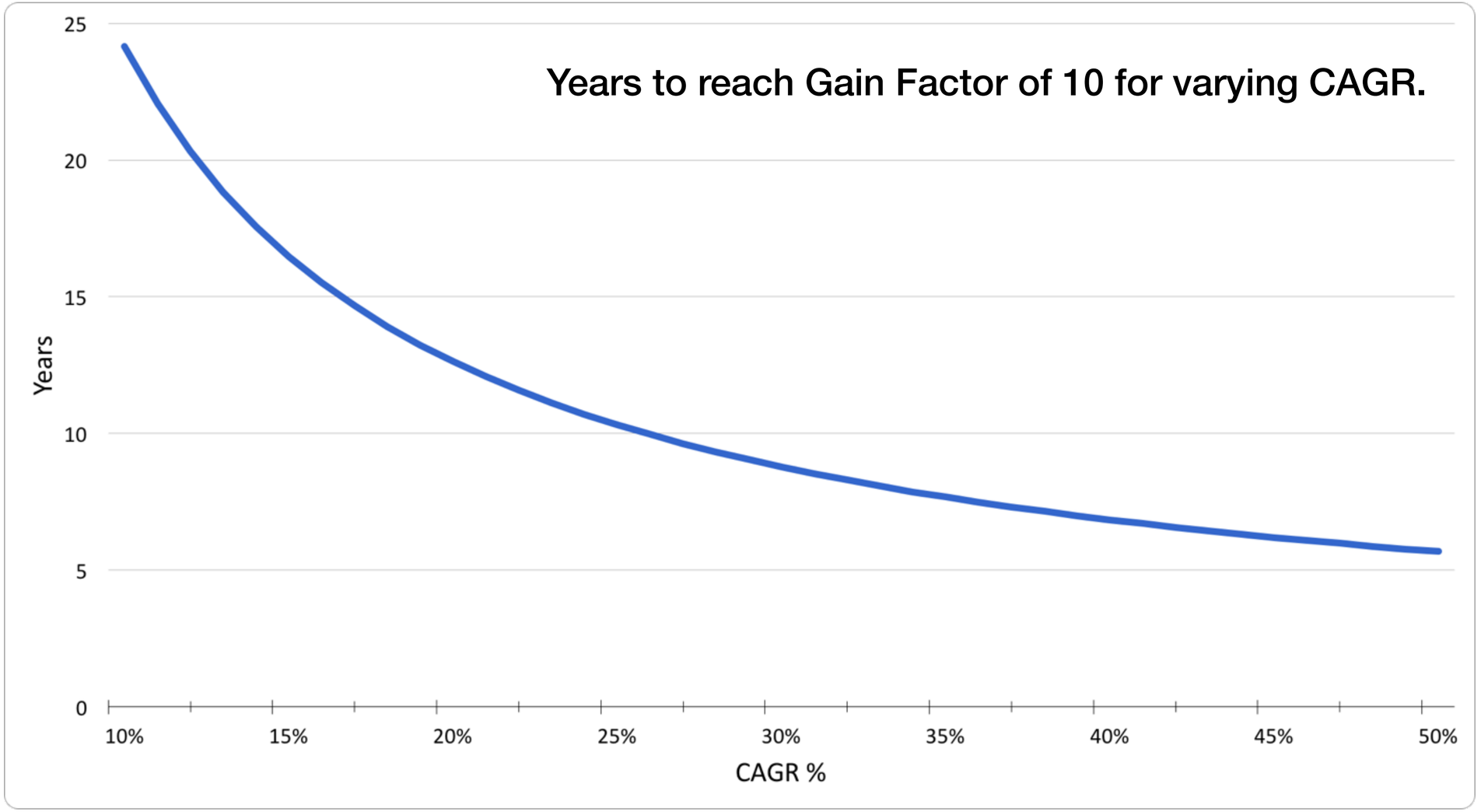


Doubling Interferometer means 8x Processing - $O(n^3)$



Processing System quickly dominates CAPEX when increasing Interferometer Size.

$O(n^3)$ and Impact on Timeline



At 24 month doubling rate (CAGR 41%) it takes 7 years to gain a factor of 10,
The rebaselining downsized by a factor of 10.

Finding Slack in SDP System

- data format: Measurement Set (MS) v2 table system
  v3 initiative by SKAO, NRAO etc.; JPIP/JPEG2000 for vis.
- storage manager: ADIOS (Adaptable I/O System)
  abstract parallel IO interface from application
- cluster management
  e.g. OpenStack for bare metal performance in contrast to VM
- Containerization: e.g. Docker
  machine learning: RFI detection and classification
- scheduling: DAG; NP hard
- algorithms/maths: BDA, facetting
  review data reduction/calibration approach
- going from batch to stream processing

SDP Execution Framework Prototype DALiuGE

DALiu**GE** – **D**ata-**A**ctivated **Liu** (流) **G**raph **E**ngine

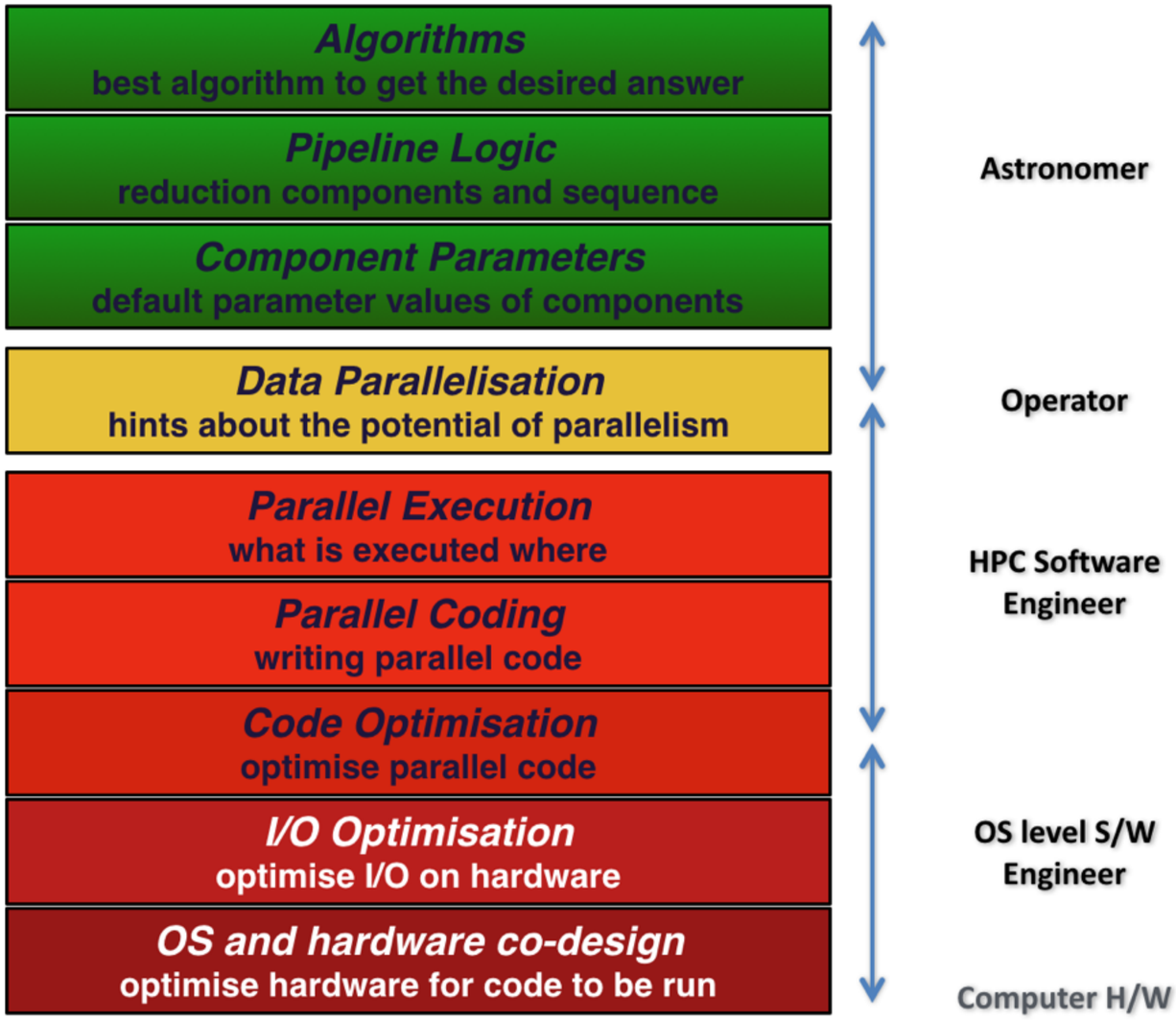
- https://github.com/ICRAR/daliuge
- 20,000 lines of Python code + minimal C library code
- refs
  - http://www.sciencedirect.com/science/article/pii/S2213133716301214
  - https://arxiv.org/abs/1702.07617

Data-centric Approach

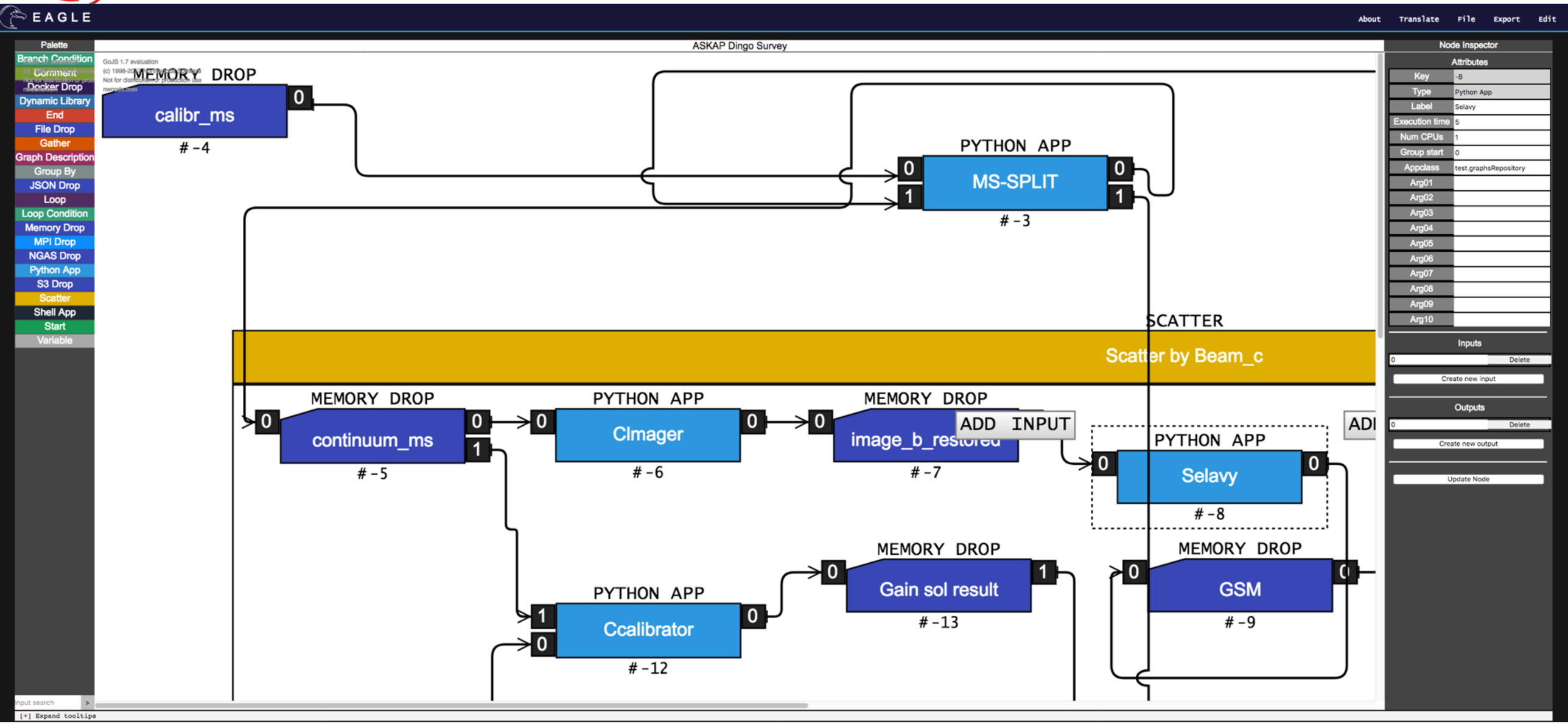
+ control of data locality
+ exploitation of data parallelism
+ works with commodity hardware
+ no global file system required, just a GUID server like Ceph
+ DAG based control flow; not limited by single master node
+ load-balancing
+ fault tolerance: depends on storage manager

- idle times when waiting for matching data

Separation of Concerns facilitated by DALiuGE



EAGLE … Editor for the Astronomical Graph Language Environment



Code Reuse - JACAL (**J**oint **A**stronomy imaging and **CAL**ibration)

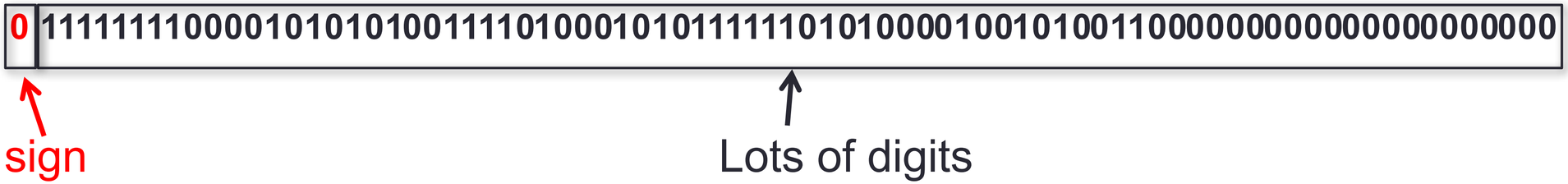
JACAL - a Rialto project deliverable:

- **Investigate reuse of ASKAPsoft in context of SKA1-Low**
- **Integrate with DALiuGE to scale system performance**
- **Resulting source package JACAL (ASKAPsoft & DALiuGE)**
- **Determine workload characteristics/data access patterns of algorithmic components**
- **Reuse MWA archival data and codes:**

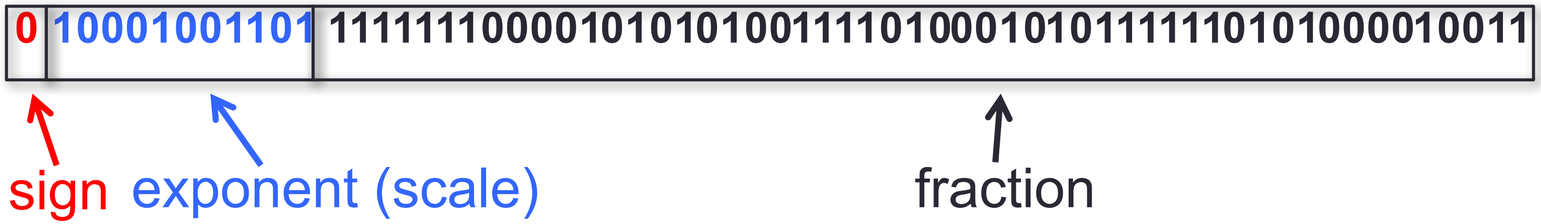
  e.g. continuum imaging and precision spectral line work at low frequencies

Unum Number Format by Gustafson

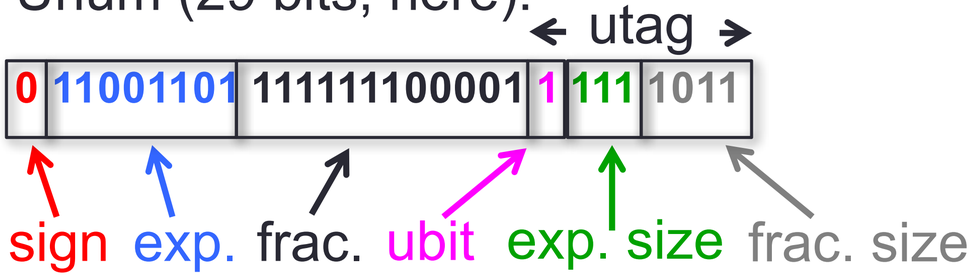
Sign-Magnitude Integer (80 bits):



IEEE Standard Float (64 bits):

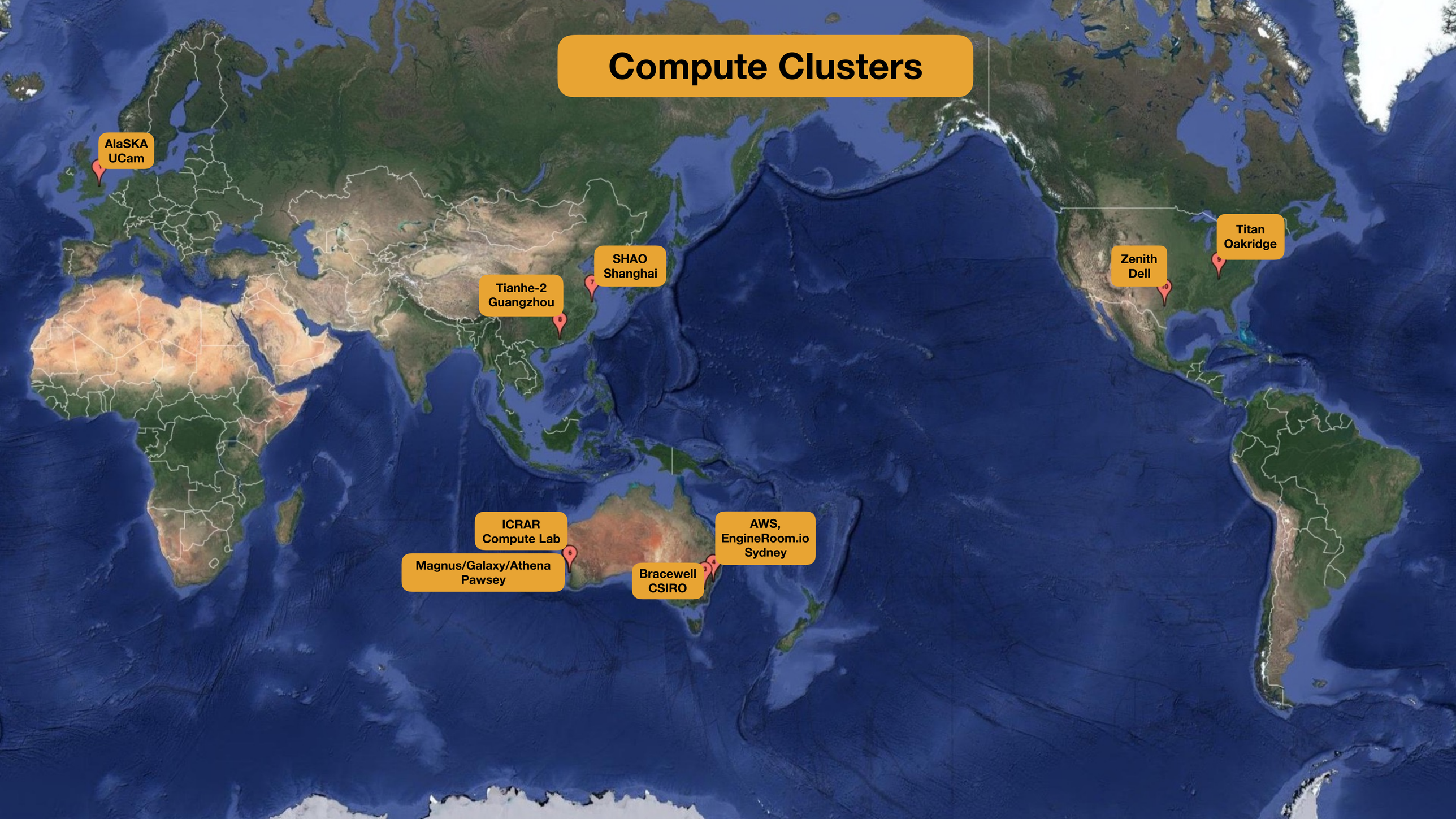


Unum (29 bits, here):

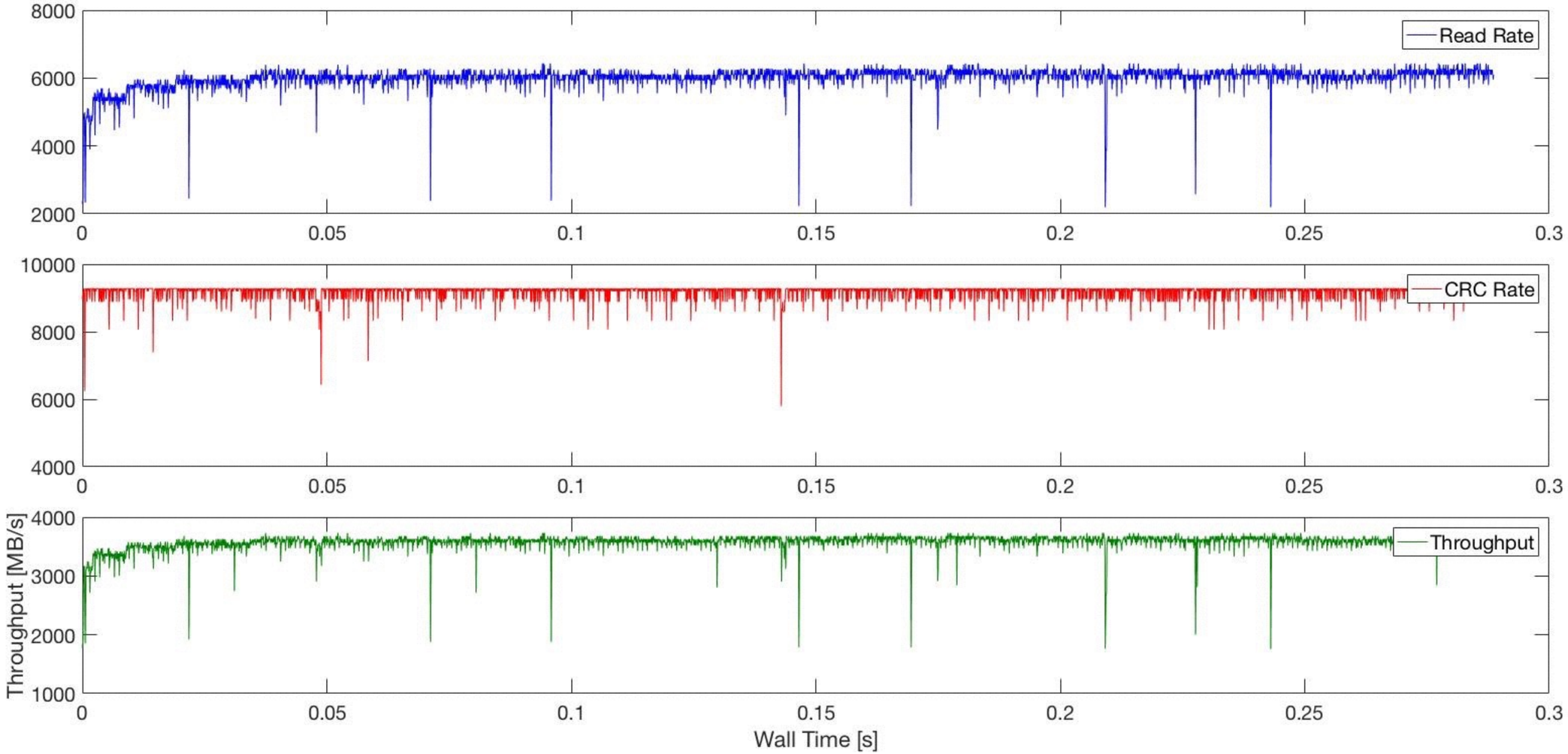


Self-descriptive "utag" bits track and manage uncertainty, exponent size, and fraction size

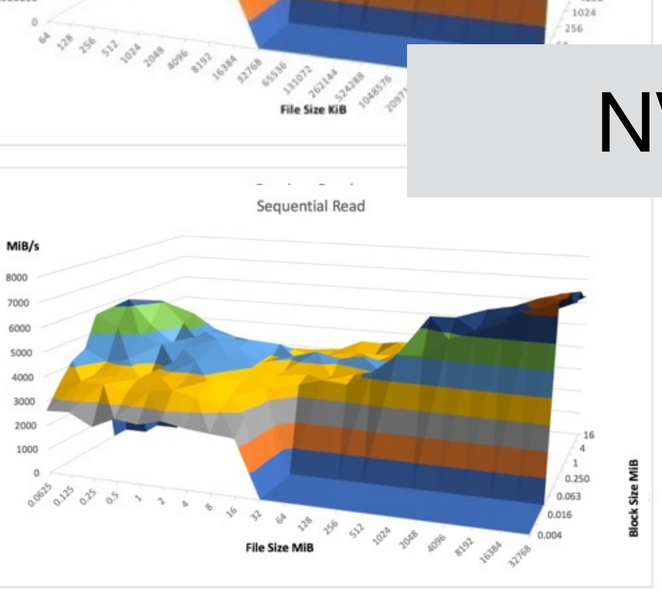
more compact, accurate and efficient

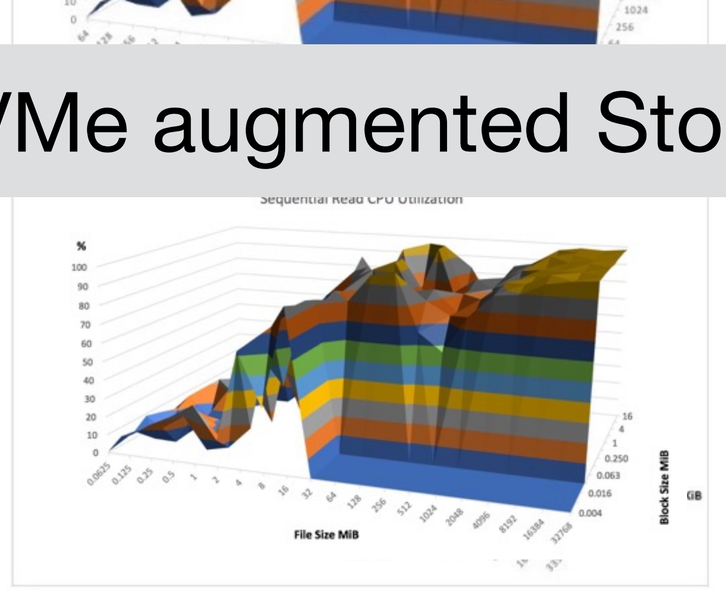

Compute Clusters
AlaSKA
UCam
SHAO
Shanghai
Tianhe-2
Guangzhou
Zenith
Dell
Titan
Oakridge
ICRAR
Compute Lab
Magnus/Galaxy/Athena
Pawsey
Bracewell
CSIRO
AWS,
EngineRoom.io
Sydney

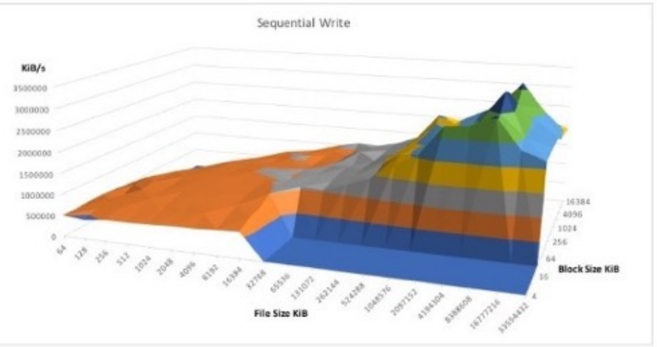
CRC Calc Rate on Xeon Phi (KNL)

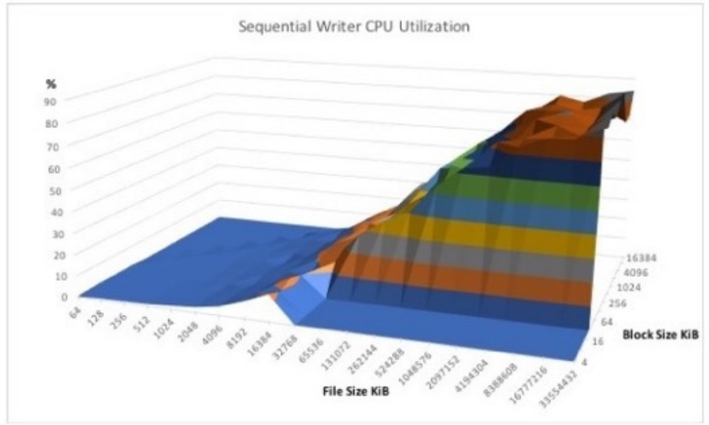


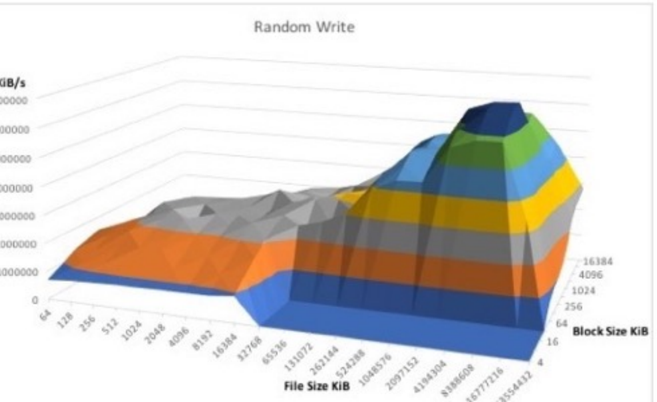
NVMe augmented Storage Node Performance

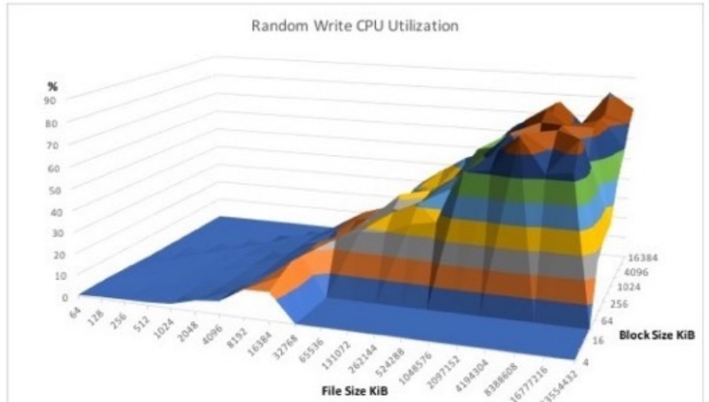


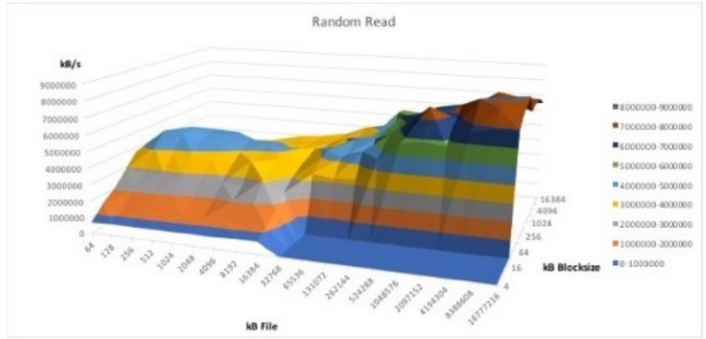


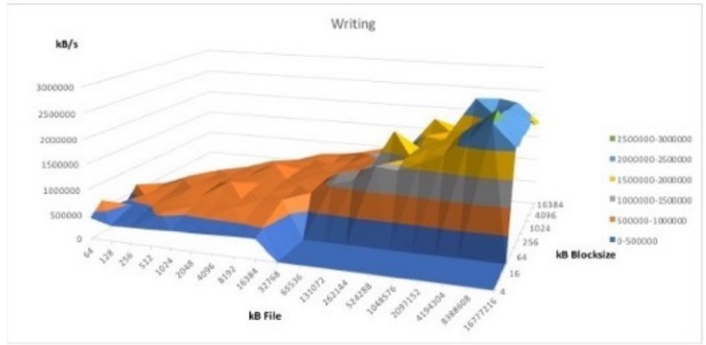


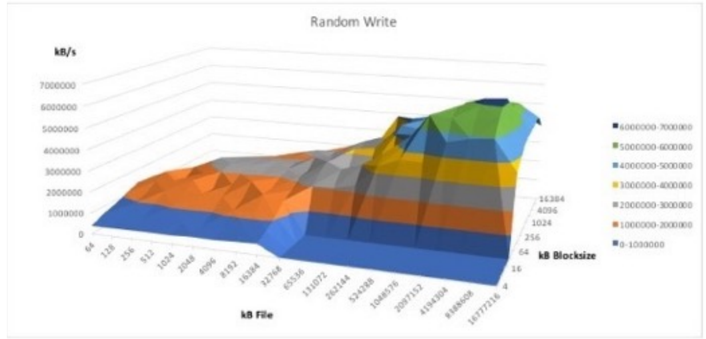




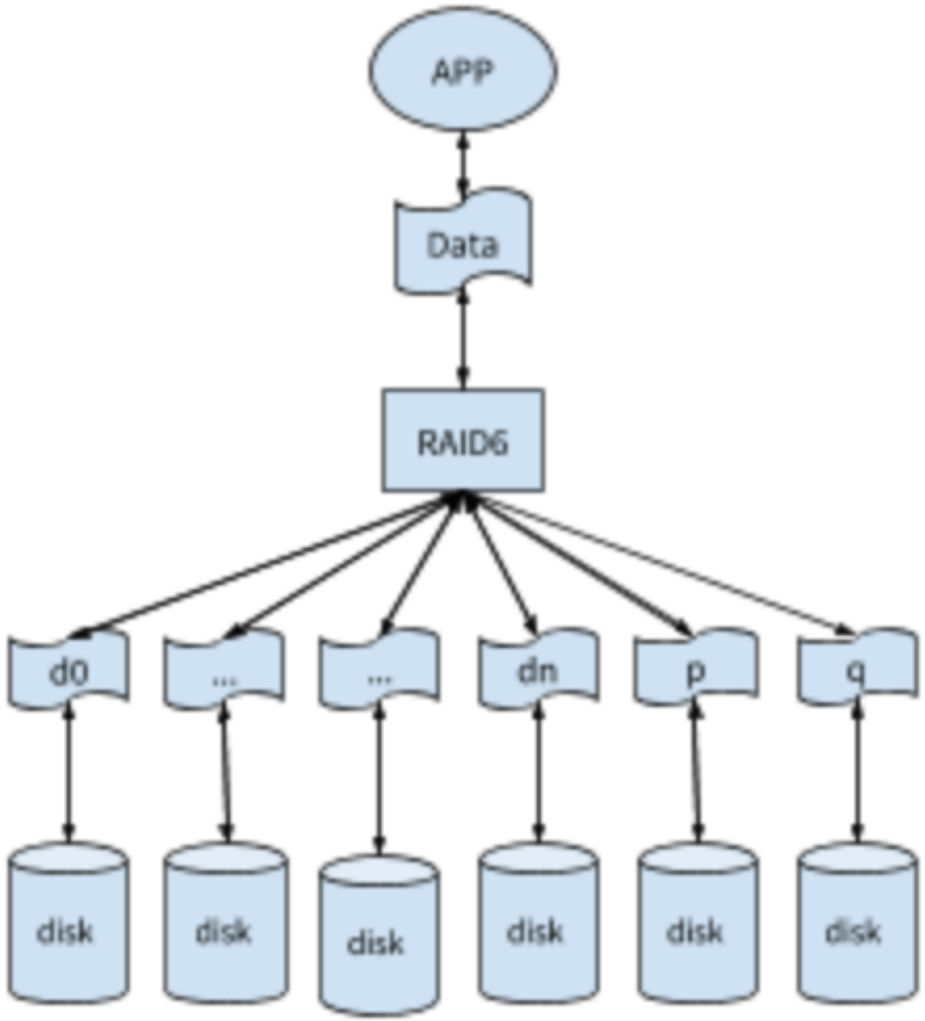




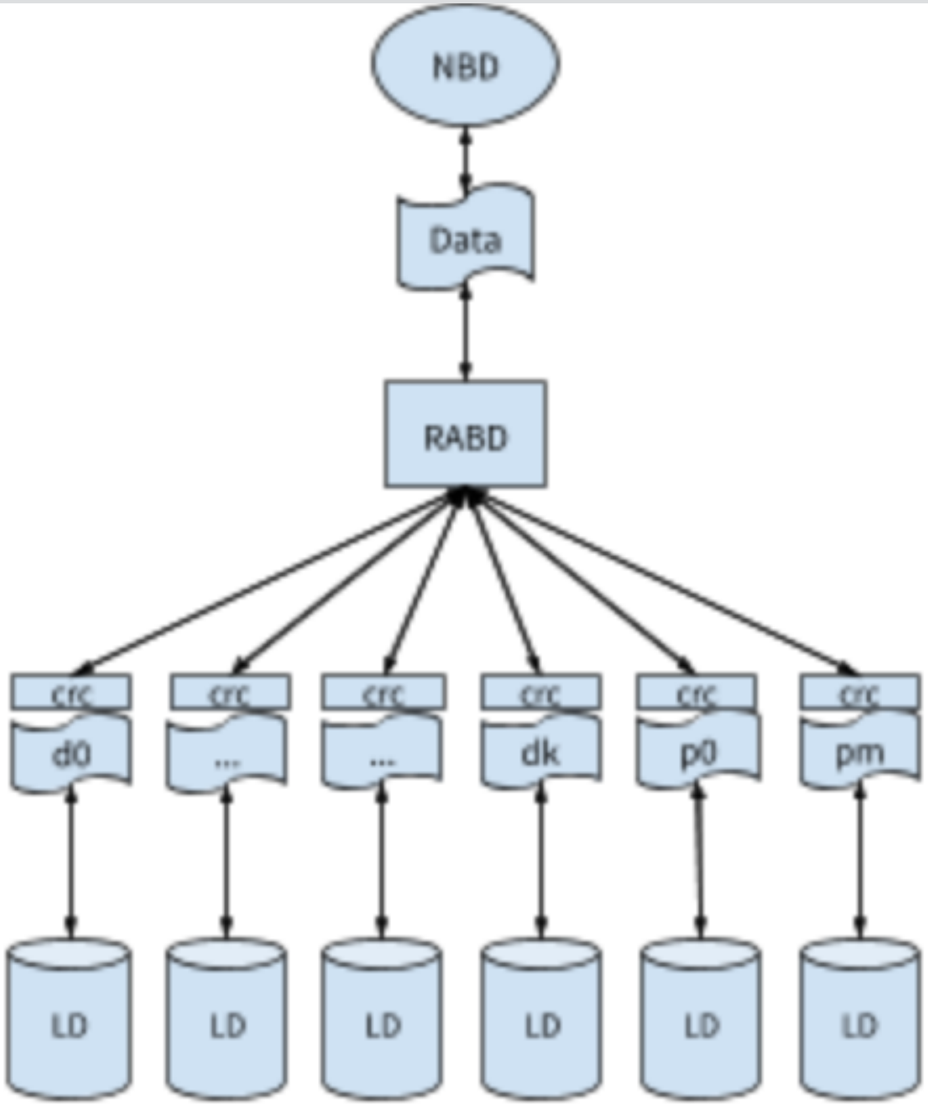




Radio Astronomy Block Device

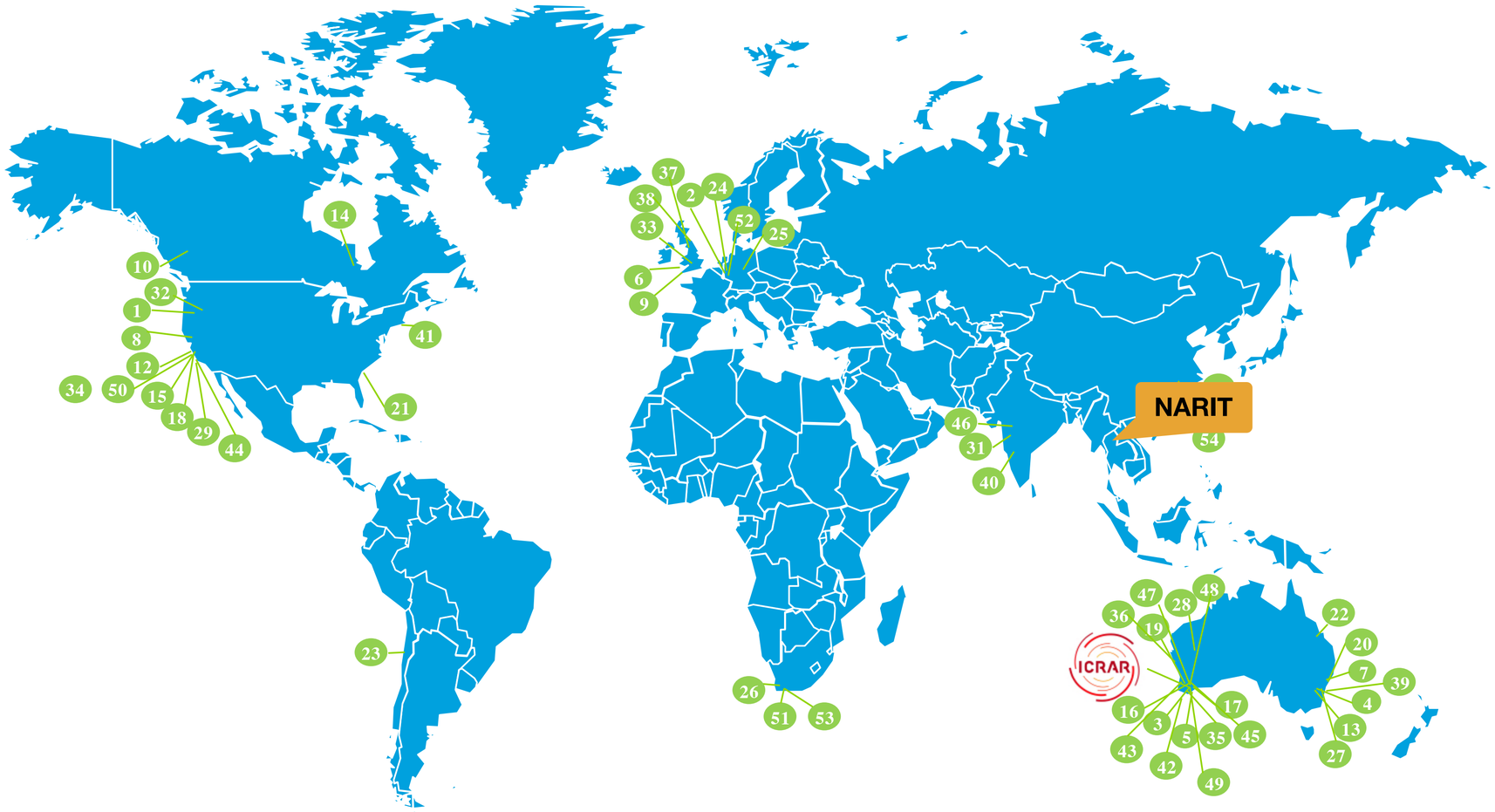


Maximum of 2 parity (p and q) no matter the amount of data



RABD can be configured to store up to 127 parity blocks with its data with a maximum recommended configuration being 96 data blocks with 32 parity across 128 logical devices.

Industrial & Academic Collaboration Partners



Map Ref. No.	Collaboration Partner
1	Amazon
2	ASTRON
3	Astronomy WA
4	Australian NZ SKA coordination committee (ANZSCC)
5	Balance Utility Solutions
6	BIOS IT
7	CAASTRO
8	California Institute of Technology
9	Cambridge University
10	Canadian Astronomical Data Centre (CADC)
11	Chinese Academy of Science
12	CISCO
13	CSIRO
14	CSP work package consortium (Led by National Research Council of Canada)
15	Data Direct Networks
16	Future engineering and communications
17	IBM Australia
18	INTEL
19	IVEC
20	Jet Cut
21	John Hopkins University
22	Kaelus
23	Large synoptic survey telescope Chile
24	LFAA work package consortium (Led by ASTRON)
25	Max Planck Institute for Nuclear Physics
26	MeerKAT/SKA RSA
27	Motion technologies
28	MWA consortium
29	NASA Jet Propulsion Laboratory
30	National Astronomical Observatory of China
31	National Centre for Radio astrophysics (NCRA)
32	Nvidia
33	Oxford University
34	PANSTARRS
35	Pawsey Centre
36	Poseidon Scientific Instruments (Now owned by Raytheon)
37	PrepSKA committee
38	SKA Project Development Office
39	Questacon
40	Raman Research Institute RRI
41	Raytheon
42	Romteck
43	Scitech
44	SGI
45	SPICE
46	Suzlon Energy
47	Systemic
48	Thoughtworks
49	Tooling Solutions
50	University of California, Davis
51	University of Cape Town
52	University of Groningen
53	University of Western Cape
54	VLBI team in Shanghai and India

Industrial Co-Design

Company	Area of Collaboration
Amazon Web Services	On demand compute and storage services AstroCompute: computing in the Cloud grant scheme
BT	Data analysis and data mining
DDN	Provider of high performance storage solutions
Intel	Access to core accelerators and silicon based, high-performance storage products
Kakadu Software	JPEG2000 based solutions
Nyriad	GPU acceleration, compression and erasure coding, Radio Astronomy Block Device
SGI	Compute cluster and storage design
Systemic	System and software engineering (SysML)
ThinkBottomUp	Distributed Database engine
Thoughtworks	S/w development environment and methodology

Opportunities

- co-design
- technical review
- exchange of product roadmaps
- product evaluation
- catch up meetings

Developing Roadmaps to (SKA) Regional Centres

EU

Advanced European Network of E-infrastructures for Astronomy with the SKA

Funded: 15.07.16

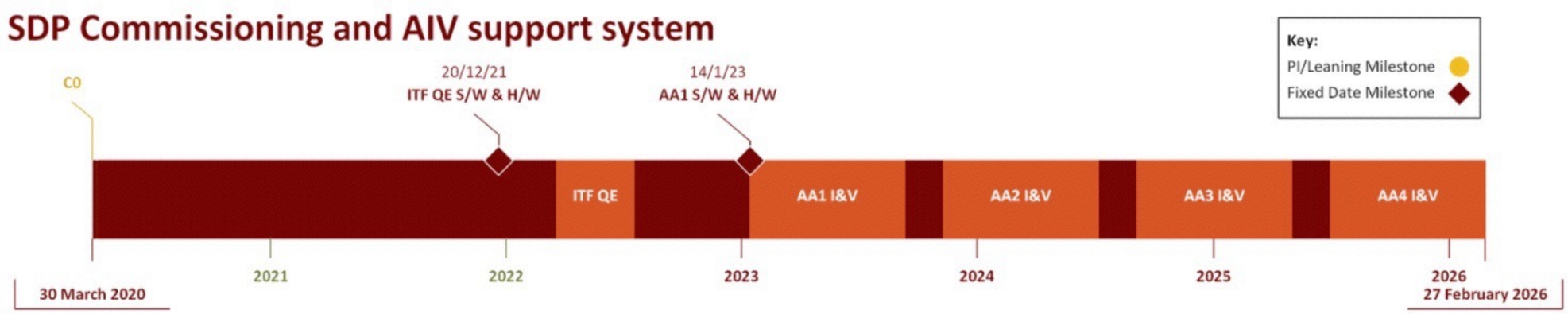
Australia

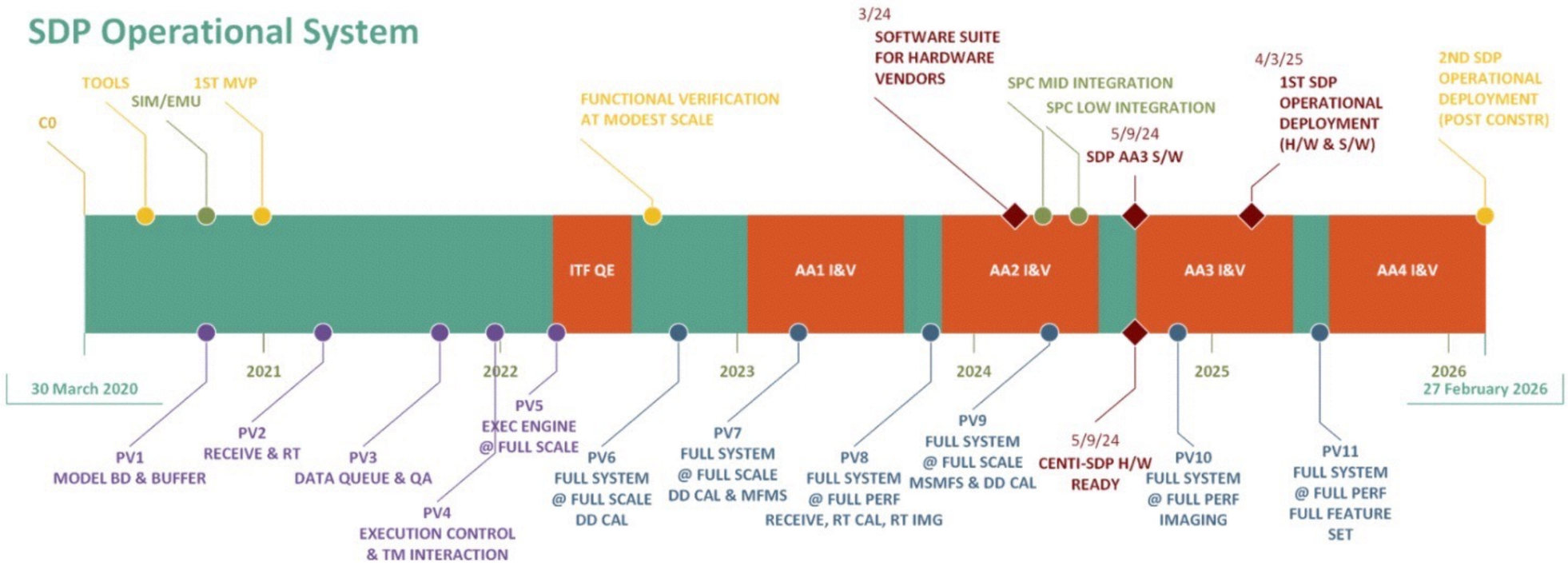
Exascale Research Infrastructure for Data in Asian-Pacific astroNomy Using the SKA

Launched 10.04.17

http://eridanus.net.au

SDP Rollout


SDP Commissioning and AIV support system
Key:
PI/Leaning Milestone
Fixed Date Milestone
C0
20/12/21
ITF QE S/W & H/W
14/1/23
AA1 S/W & H/W
ITF QE
AA1 I&V
AA2 I&V
AA3 I&V
AA4 I&V
2021
2022
2023
2024
2025
2026
30 March 2020
27 February 2026


SDP Operational System
C0
TOOLS
SIM/EMU
1ST MVP
FUNCTIONAL VERIFICATION
AT MODEST SCALE
3/24
SOFTWARE SUITE
FOR HARDWARE
VENDORS
SPC MID INTEGRATION
SPC LOW INTEGRATION
5/9/24
SDP AA3 S/W
4/3/25
1ST SDP
OPERATIONAL
DEPLOYMENT
(H/W & S/W)
2ND SDP
OPERATIONAL
DEPLOYMENT
(POST CONSTR)
ITF QE
AA1 I&V
AA2 I&V
AA3 I&V
AA4 I&V
30 March 2020
2021
2022
2023
2024
2025
2026
27 February 2026
PV1
MODEL BD & BUFFER
PV2
RECEIVE & RT
PV3
DATA QUEUE & QA
PV4
EXECUTION CONTROL
& TM INTERACTION
PV5
EXEC ENGINE
@ FULL SCALE
PV6
FULL SYSTEM
@ FULL SCALE
DD CAL
PV7
FULL SYSTEM
@ FULL SCALE
DD CAL & MFMS
PV8
FULL SYSTEM
@ FULL PERF
RECEIVE, RT CAL, RT IMG
PV9
FULL SYSTEM
@ FULL SCALE
MSMFS & DD CAL
5/9/24
CENTI-SDP H/W
READY
PV10
FULL SYSTEM
@ FULL PERF
IMAGING
PV11
FULL SYSTEM
@ FULL PERF
FULL FEATURE
SET