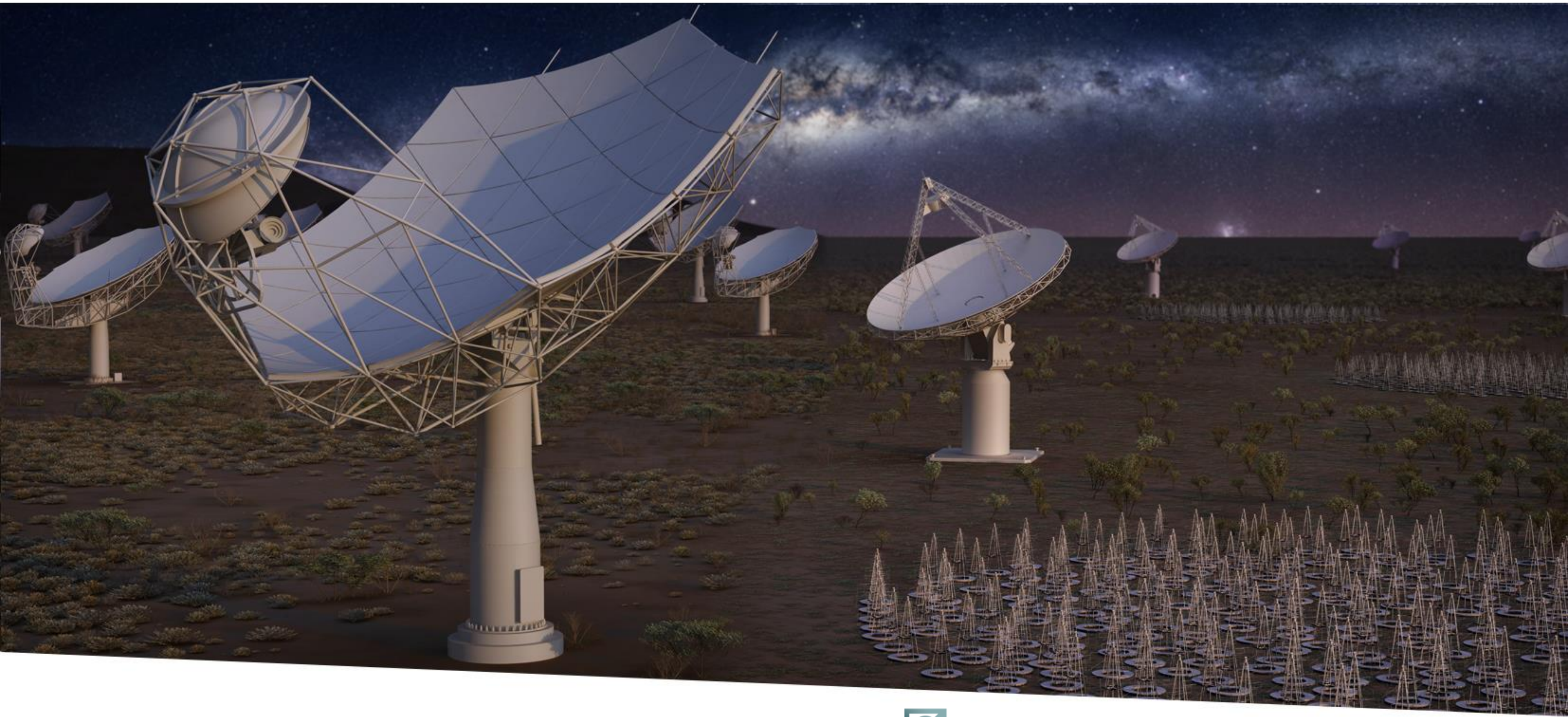
The Square Kilometre Array

ENGAGE SKA PORTUGAL



SQUARE KILOMETRE ARRAY

Exploring the Universe with the world's largest radio telescope

tice.pt

PÓLO DE COMPETITIVIDADE DAS TECNOLOGIAS

Domingos Barbosa

28 Nov 2018

SKA– Key Science Drivers: The history of the Universe

Cosmic Dawn
Galaxies)

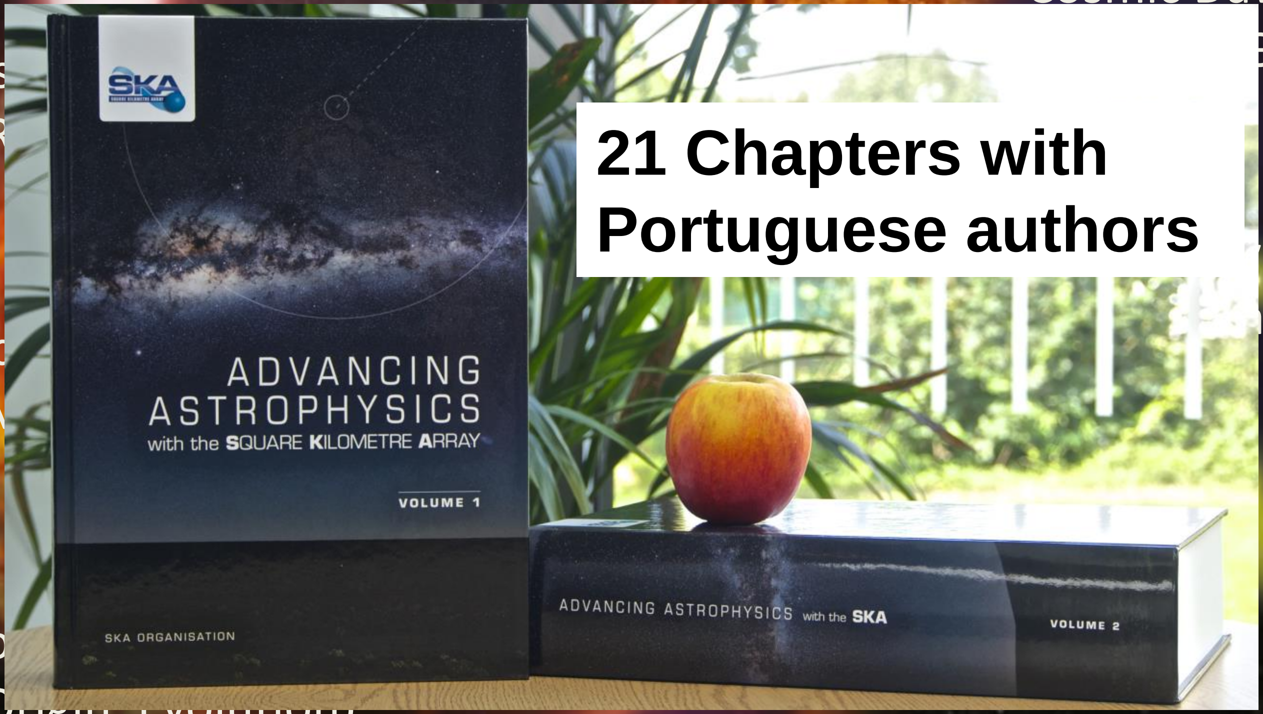
Tes
(Strong R

olution
xies z~2-3)

Crad
(Planets, M

Structure)

Co
(Origin, Evolution)



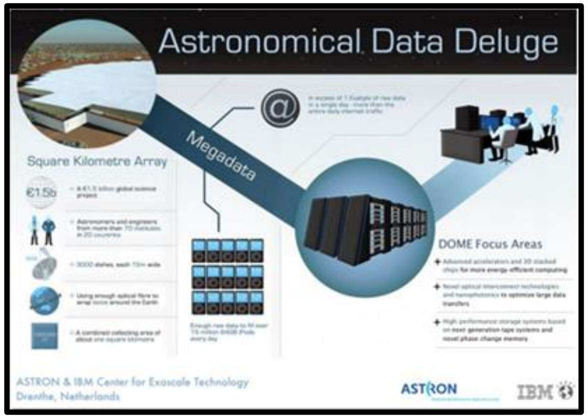
21 Chapters with Portuguese authors

Exploration of the Unknown

Broadest range of science of any facility, worldwide

SKA: driving innovation, creating impact

Infrastructure:
- Long-distance, high-capacity fibre networks
- High-performance computing and data storage
- Green computing



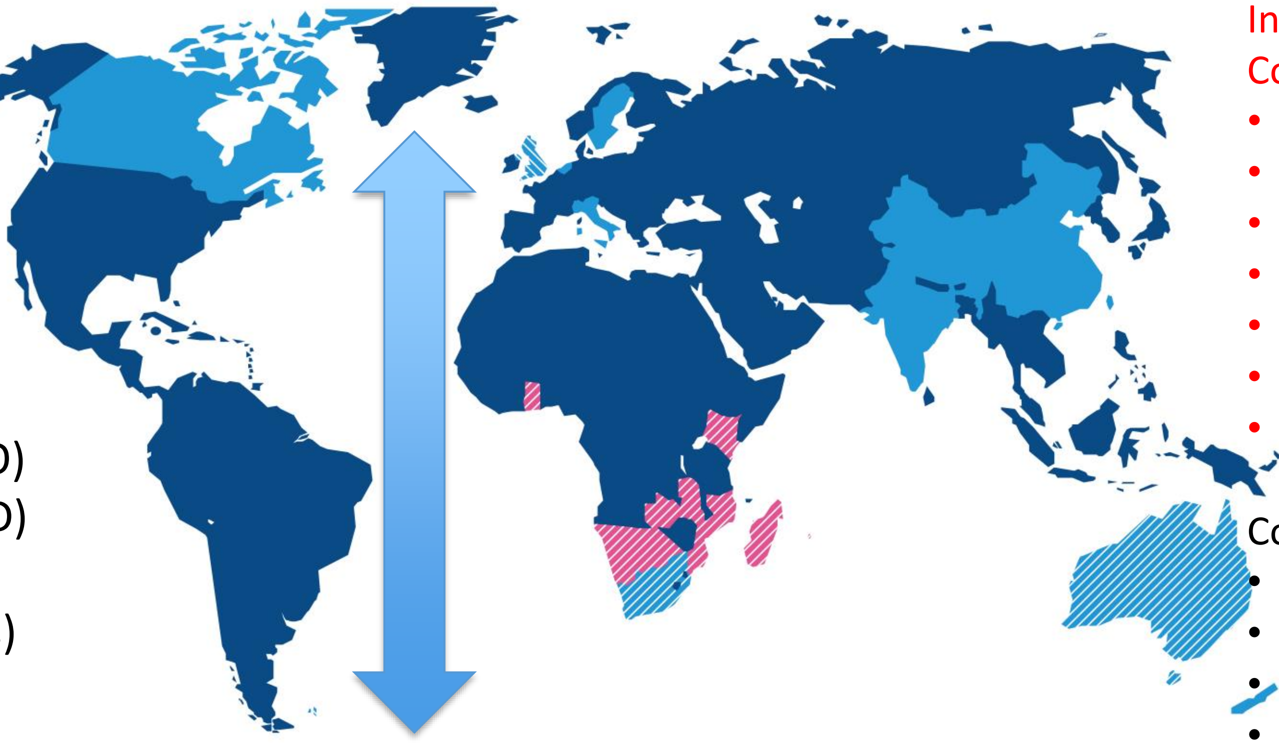
Software:
- Imaging techniques and data visualisation
- Machine learning and Artificial Intelligence
- Data Mining

Past global impacts from Big Science, Research Infrastructures and astronomy:
- Internet, WWW, WiFi, global navigation systems, medicine
  → enormous contribution to global GDP and social wellbeing

Future huge potential impact from SKA and its innovation:
- eg Internet of Things; data analytics; visualisation; 'storage-processing'

SKA Organisation: 10 countries, more to join

Australia (DoI&S)
Canada (NRC-HIA)
China (MOST)
India (DAE)
Italy (INAF)
Netherlands (NWO)
New Zealand (MED)
South Africa (DST)
Sweden (Chalmers)
UK (BEIS/STFC)
Portugal (FCT)



Interested Countries:
- Germany
- France
- Portugal
- Spain
- Switzerland
- Japan
- Korea

Contacts:
- USA
- Malta
- Mexico
- Brazil
- Ireland
- Russia

Full members

SKA Headquarters host country

SKA Phase 1 and Phase 2 host countries

African partner countries (non-member SKA Phase 2 host countries)

This map is intended for reference only and is not meant to represent legal borders

Square Kilometre Array

3 sites; 2 telescopes + HQ

1 Observatory

Design Phase: ~ €200M; 600 scientists+engineers

Phase 1

Construction: 2019 – 2024

Construction cost cap: €674.1M (inflation-adjusted)

Operations cost: under development

MeerKat integrated

Observatory Development Programme (€20M/year planned)

SKA Regional centres out of scope of centrally-funded SKAO.

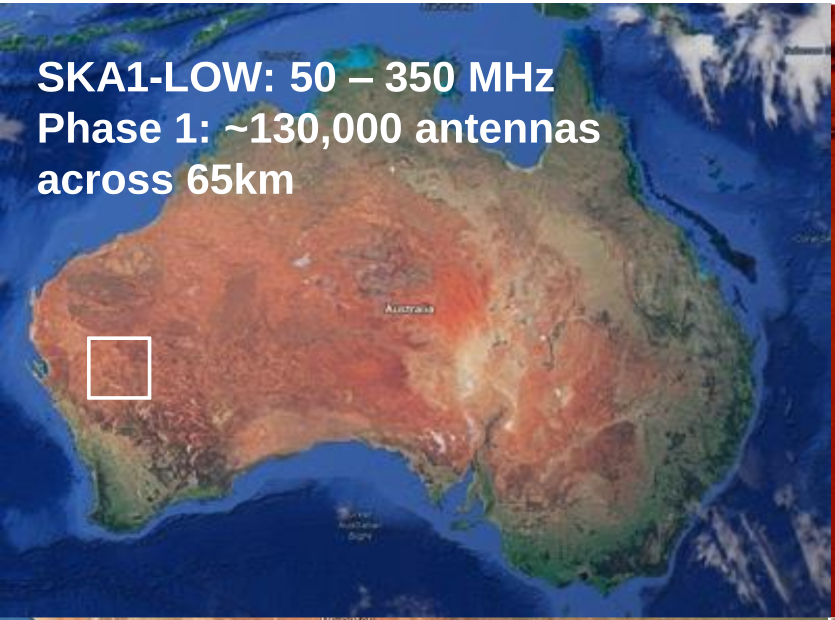
Phase 2: start mid-2020s

~2000 dishes across 3500km of Southern Africa

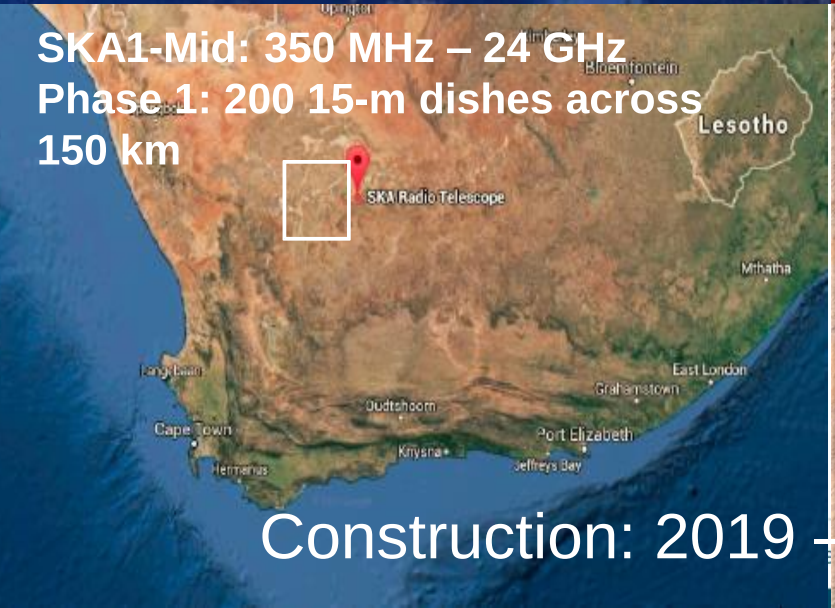
Major expansion of SKA1-Low across Western Australia

SKA: HQ in UK; telescopes in AUS & RSA

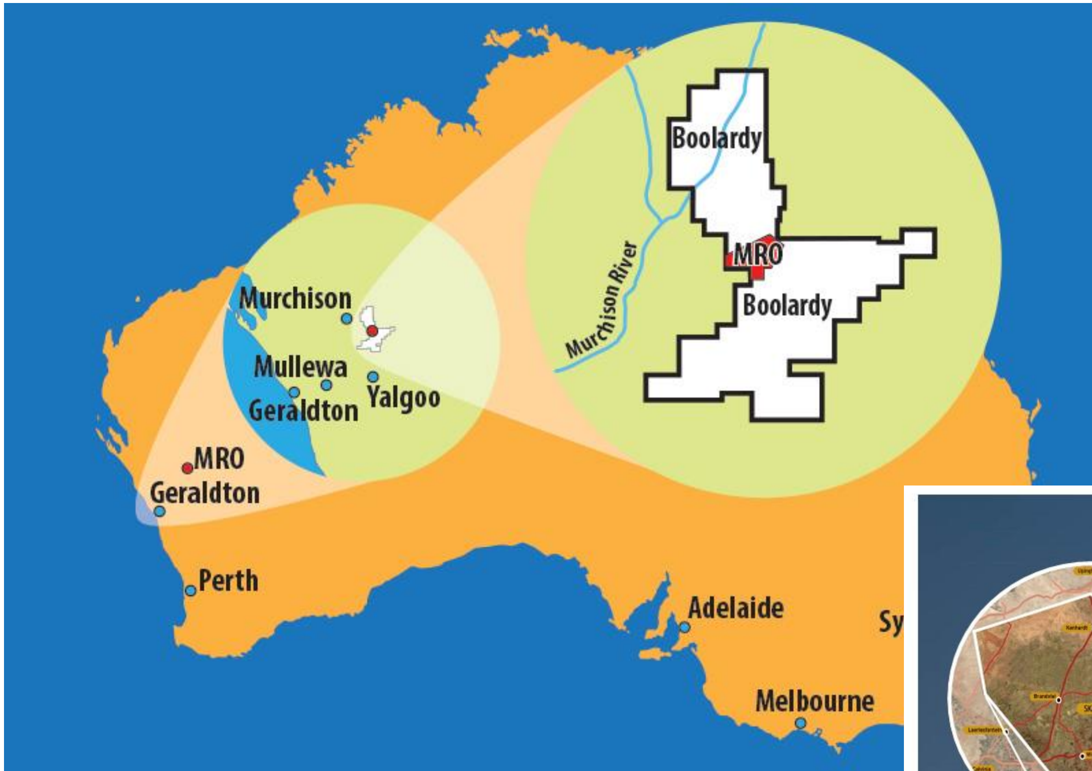
SKA1-LOW: 50 – 350 MHz
Phase 1: ~130,000 antennas across 65km

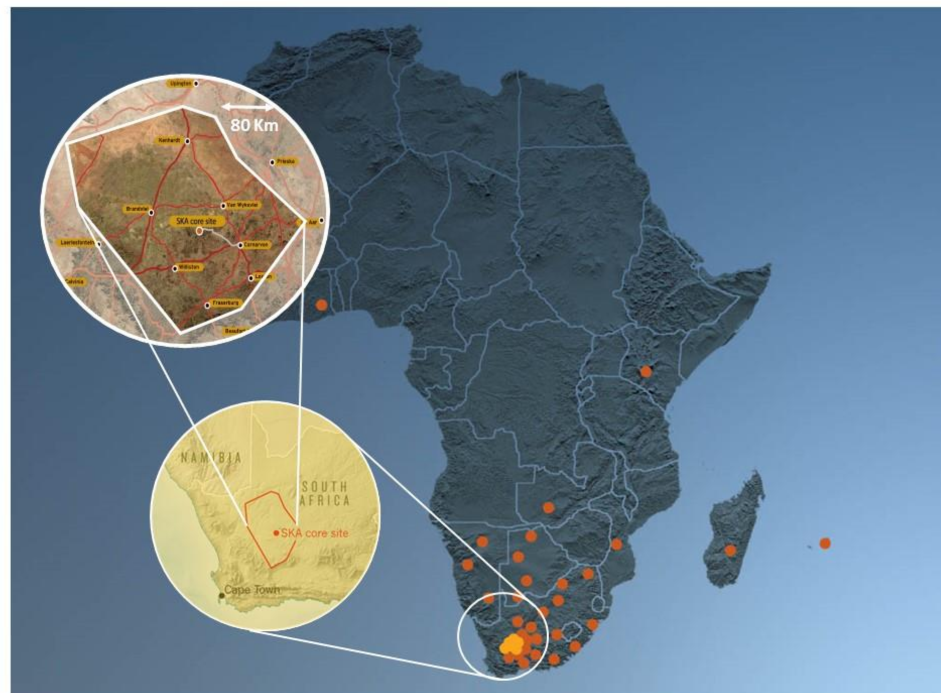


SKA1-Mid: 350 MHz – 24 GHz
Phase 1: 200 15-m dishes across 150 km



Construction: 2019 – 2024; Cost cap: €675M


Boolardy
MRO
Boolardy
Murchison River
Murchison
Mullewa
Geraldton
Yalgoo
MRO
Geraldton
Perth
Adelaide
Melbourne


80 Km
SKA core site
NAMIBIA
SOUTH AFRICA
SKA core site
Cape Town


SKA
SQUARE KILOMETRE ARRAY
www.skatelescope.org
SWINBURNE ASTRONOMY PRODUCTIONS

Status

International Design Teams

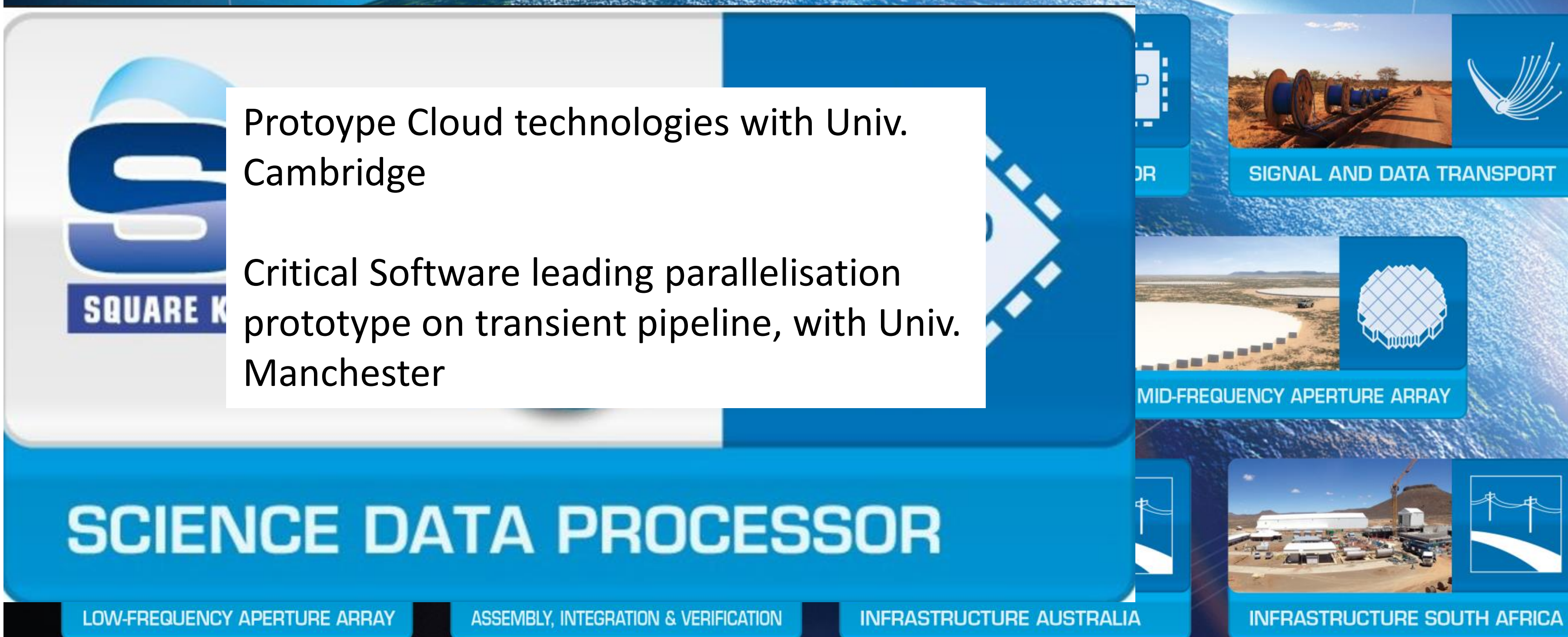
ENGAGE SKA leads local infrastructure design.

Working closely with NCRA-TIFR India

ank, Manchester, UK
r countries of the SKA

International Design Teams

- Project Management and System Engineering based at Jodrell Bank, Manchester, UK
- >600 scientists & engineers in institutes and industry in Member countries of the SKA



Protoype Cloud technologies with Univ. Cambridge

Critical Software leading parallelisation prototype on transient pipeline, with Univ. Manchester

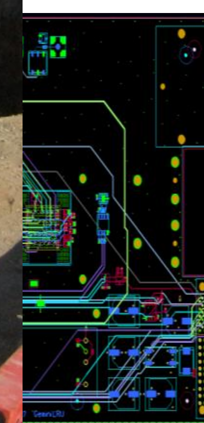
International Design Teams

- Project Manageme
- >600 scientists & e

Technical Progress

1st dish prototype complete in China in Oct 17
Chinese (CETC54)/German (MTM) design

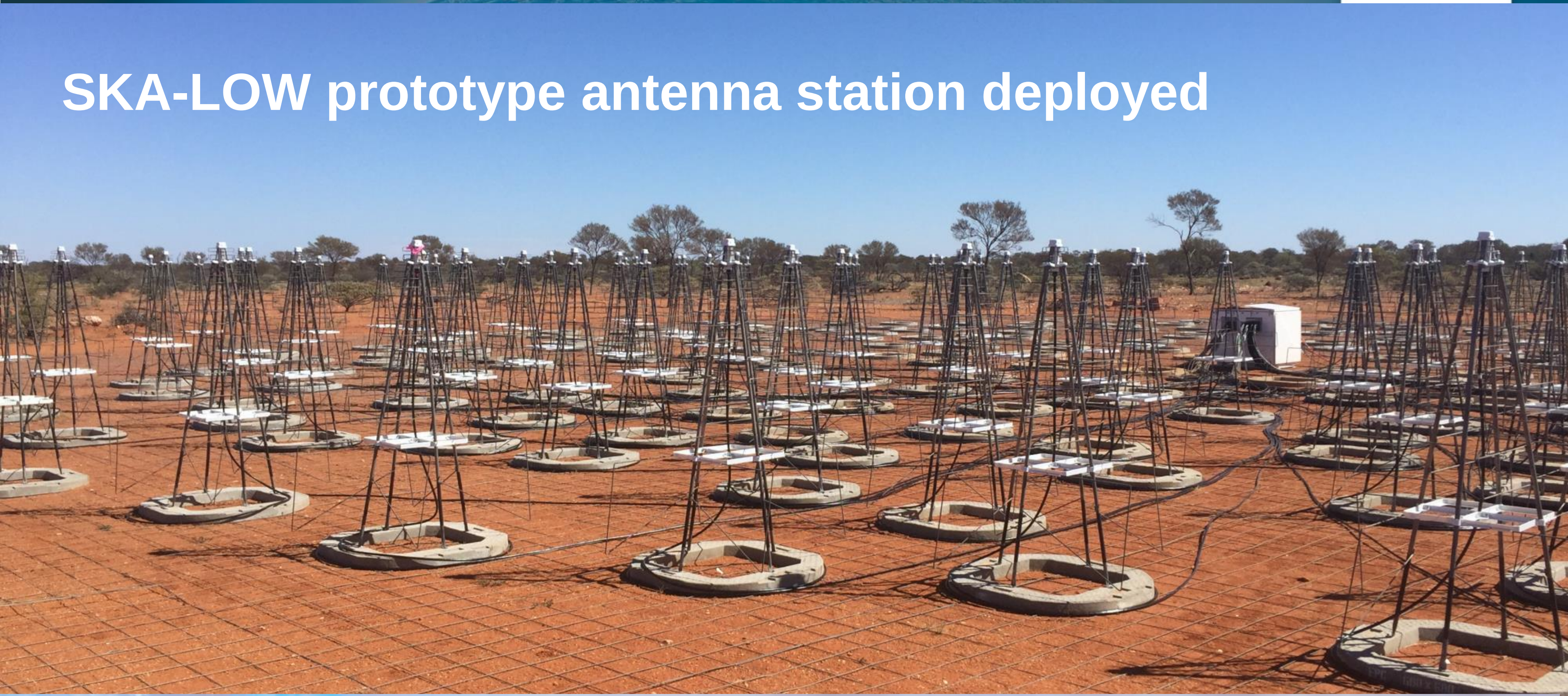
Surface I
350 micr



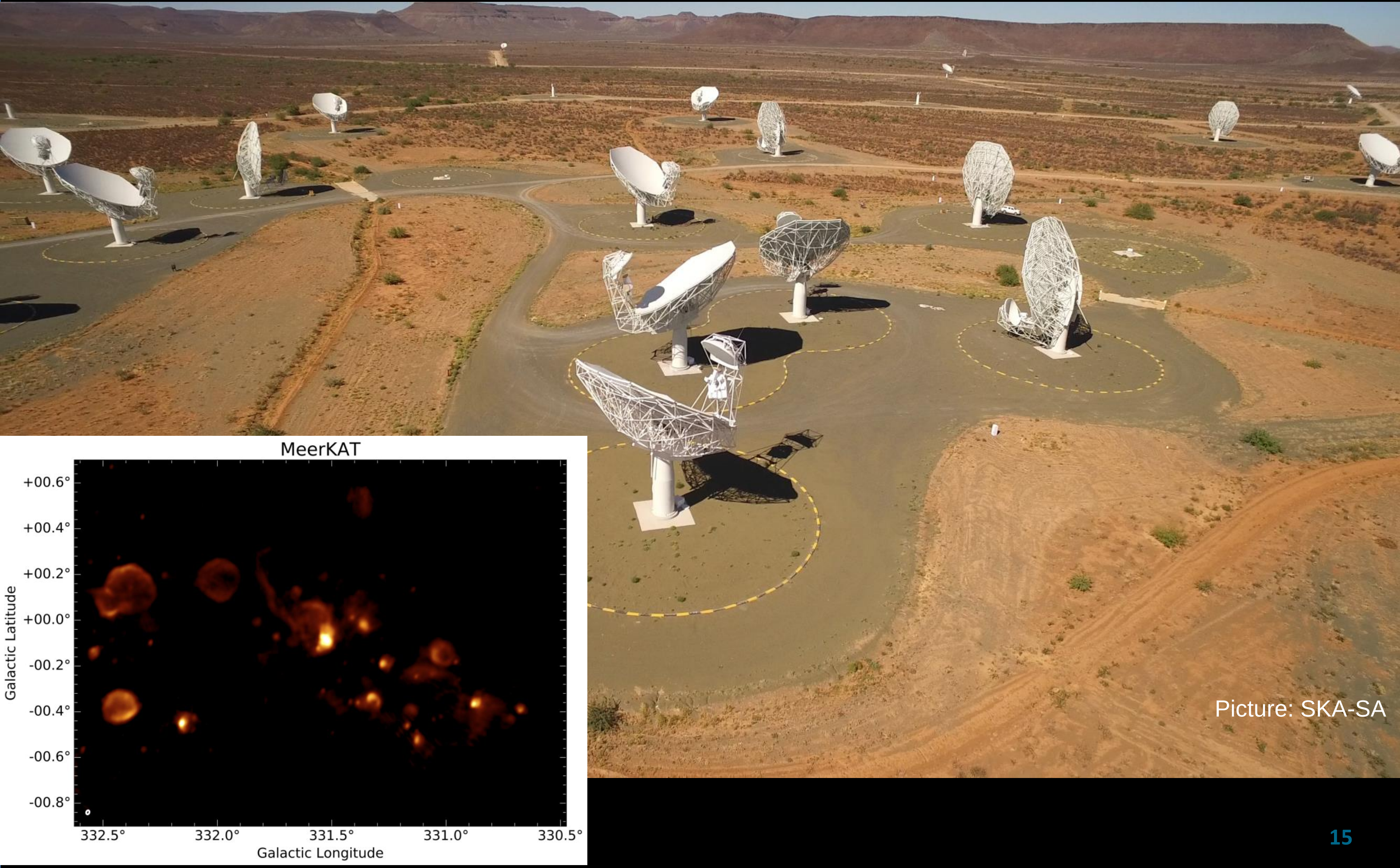
E

largest

SKA-LOW prototype antenna station deployed



MeerKAT: 49/64 antennas now built



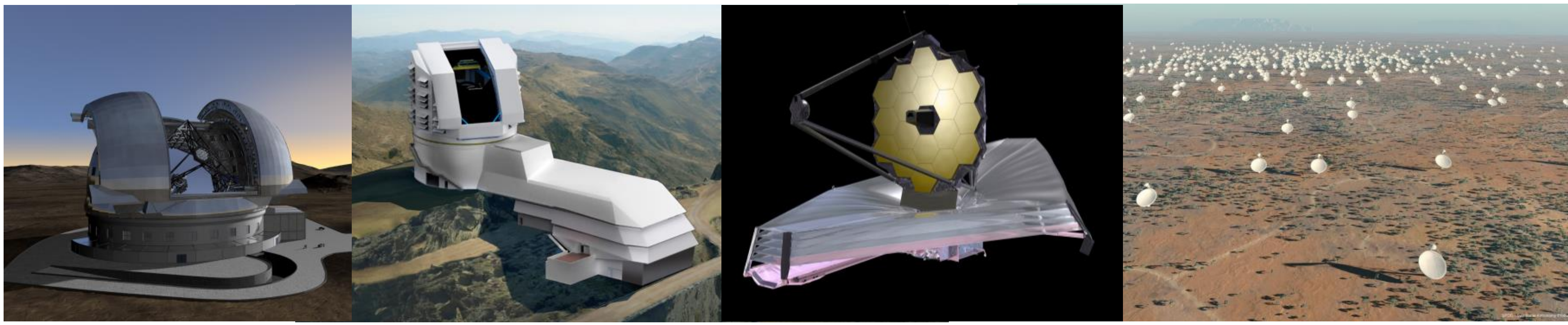
Picture: SKA-SA

Impact of Radio Astronomy

- Cheap low-noise-amplifiers → telecommunications
- Radio imaging algorithms → medical tomography, NMR imaging, finger-print detection, speed cameras!
- Very Long Baseline Interferometry measures rotational parameters of Earth →GPS and navigation: civil and military applications.
- NEOs: using radar astronomy to determine orbits of Near Earth Objects
- Wifi: digital processing techniques led directly to development of WiFi. CSIRO owns patent; ~2 billion devices sold; >$600M income

Data Intensive Astronomy

Data Intensive Astronomy



- Science is increasingly driven by large data sets
- Massive data collections and large scientific collaborations
- Most science extraction is based on the archived data
- Current instruments already producing petascale datasets

New science infrastructures will produce exascale data!

• Data flow

Global internet traffic ~360 Tb/s

(Cisco: 2016)

SKA1-LOW

~2 Pb/ s

SKA SQUARE KILOMETRE ARRAY
LOW-FREQUENCY APERTURE ARRAY

7.2 Tb/ s

SKA1-MID

8.8 Tb/ s

SKA SQUARE KILOMETRE ARRAY
SIGNAL AND DATA TRANSPORT

SKA SQUARE KILOMETRE ARRAY CSP
CENTRAL SIGNAL PROCESSOR

~50 PFlops

~5 Tb/ s

SKA SQUARE KILOMETRE ARRAY SDP
SCIENCE DATA PROCESSOR

~250 PFlops

~300 PB/ yr

SKA: A Leading Big Data Challenge for 2020 decade

Antennas

Digital Signal Processing (DSP)

Transfer antennas to DSP
2020: 5,000 PBytes/day
2030: 100,000 PBytes/day

Over 10's to 1000's kms

To Process in HPC
2020: 50 PBytes/day
2030: 10,000 PBytes/day

Over 10's to 1000's kms

High Performance Computing Facility (HPC)

HPC Processing
2023: 250 PFlop
2033+: 25 EFlop

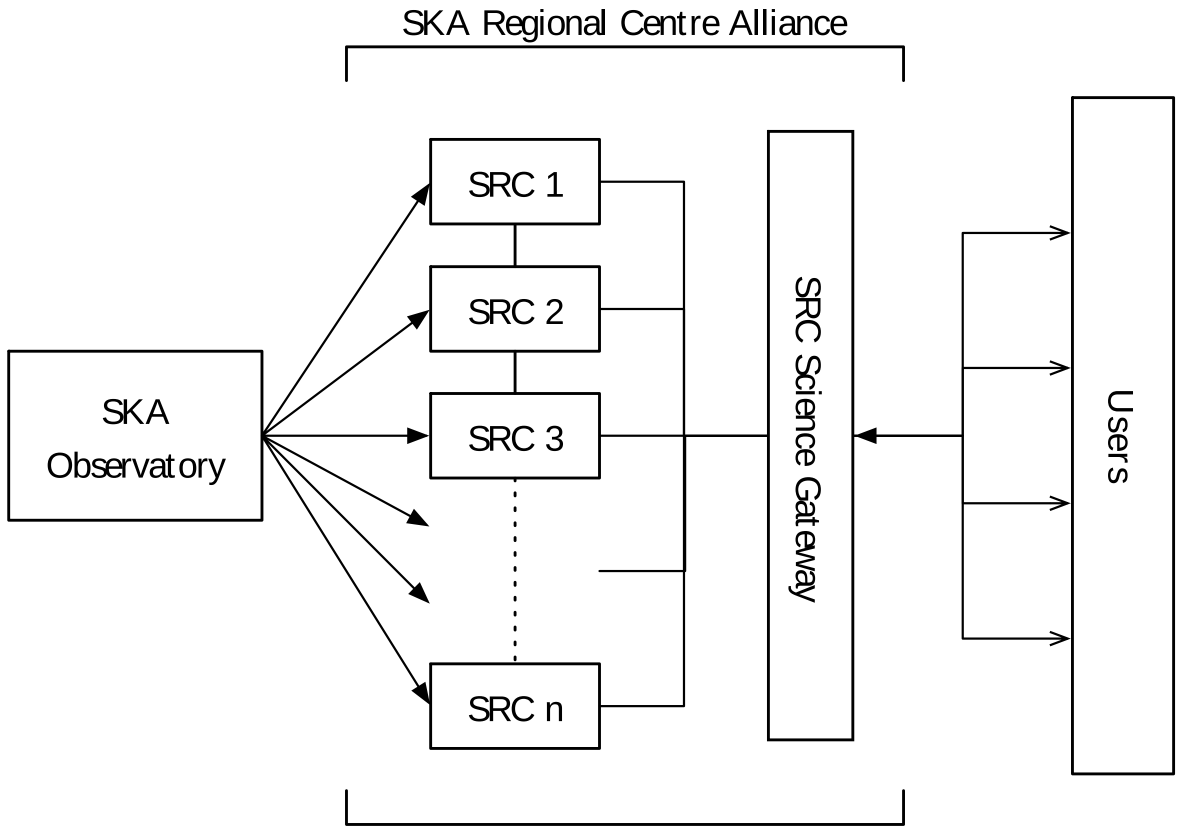
SKA Big Data

- Science is increasingly driven by large data sets
- Massive data collections and large scientific collaborations
- Most science extraction is based on the archived data
- Current instruments already producing petascale datasets

New science infrastructures will produce exascale data!

Future SKA Science Archive

- A collaborative model for SKA Regional Centres



Archive Estimates SKA1

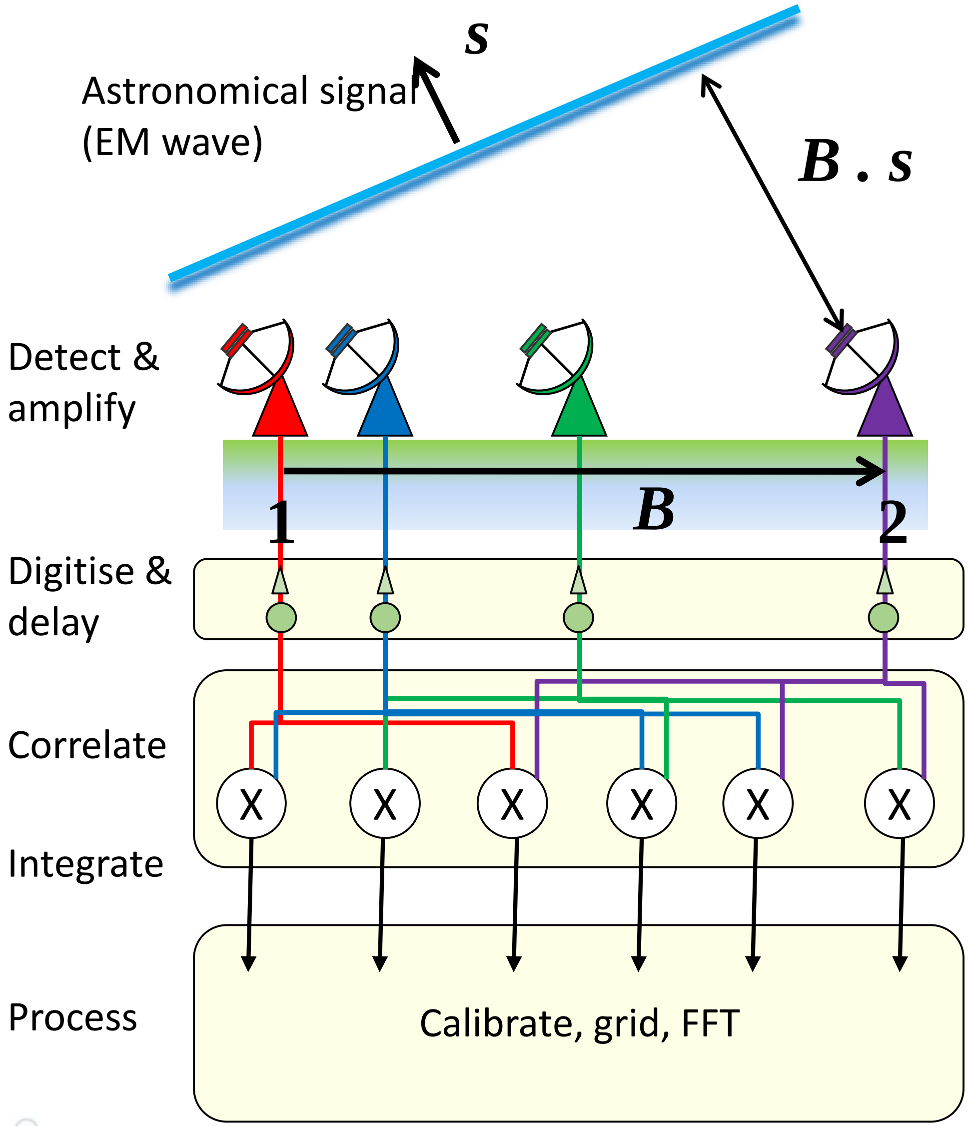
Archive Growth Rate: ~ 45 – 120 **Pbytes/yr**
Larger including discovery space: ~300 Pbytes/yr

Data Products

- Standard products
  - calibrated multi-dimensional sky images;
  - time-series data
  - catalogued data for discrete sources (global sky model) and pulsar candidates
- No derived products from science teams

Further Processing and Science Extraction at Regional Centres

Science Data Processor (SDP)



- Visibility:

$$V(\boldsymbol{B}) = E_1 E_2^*$$
$$= I(\boldsymbol{s}) \exp(\, i\, \omega\, \boldsymbol{B.s}/c\,)$$

- Resolution determined by maximum baseline

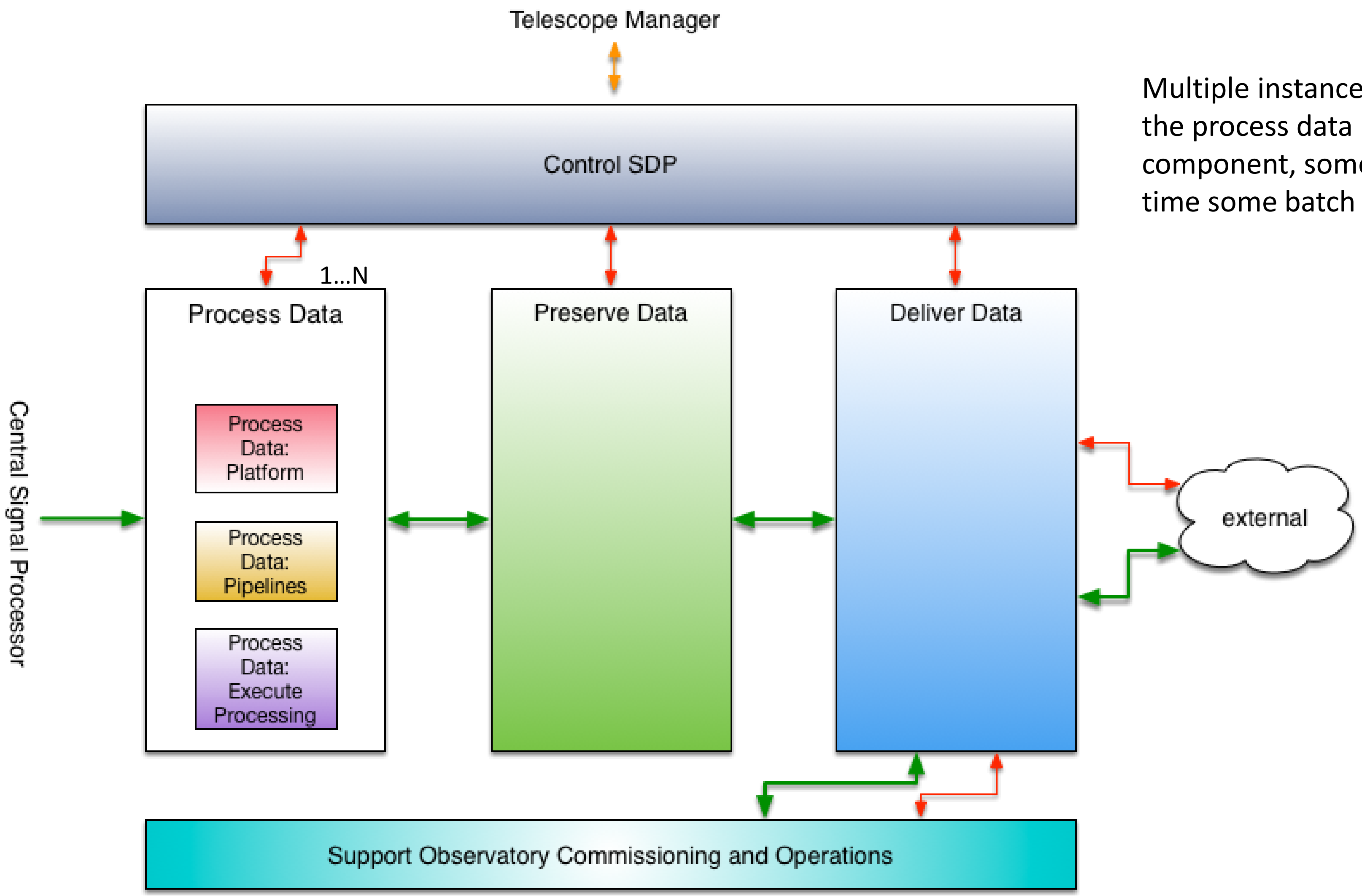
  $\theta_{max} \sim \lambda / B_{max}$

- Field of View (FoV) determined by the size of each dish

  $\theta_{dish} \sim \lambda / D$

SKY Image

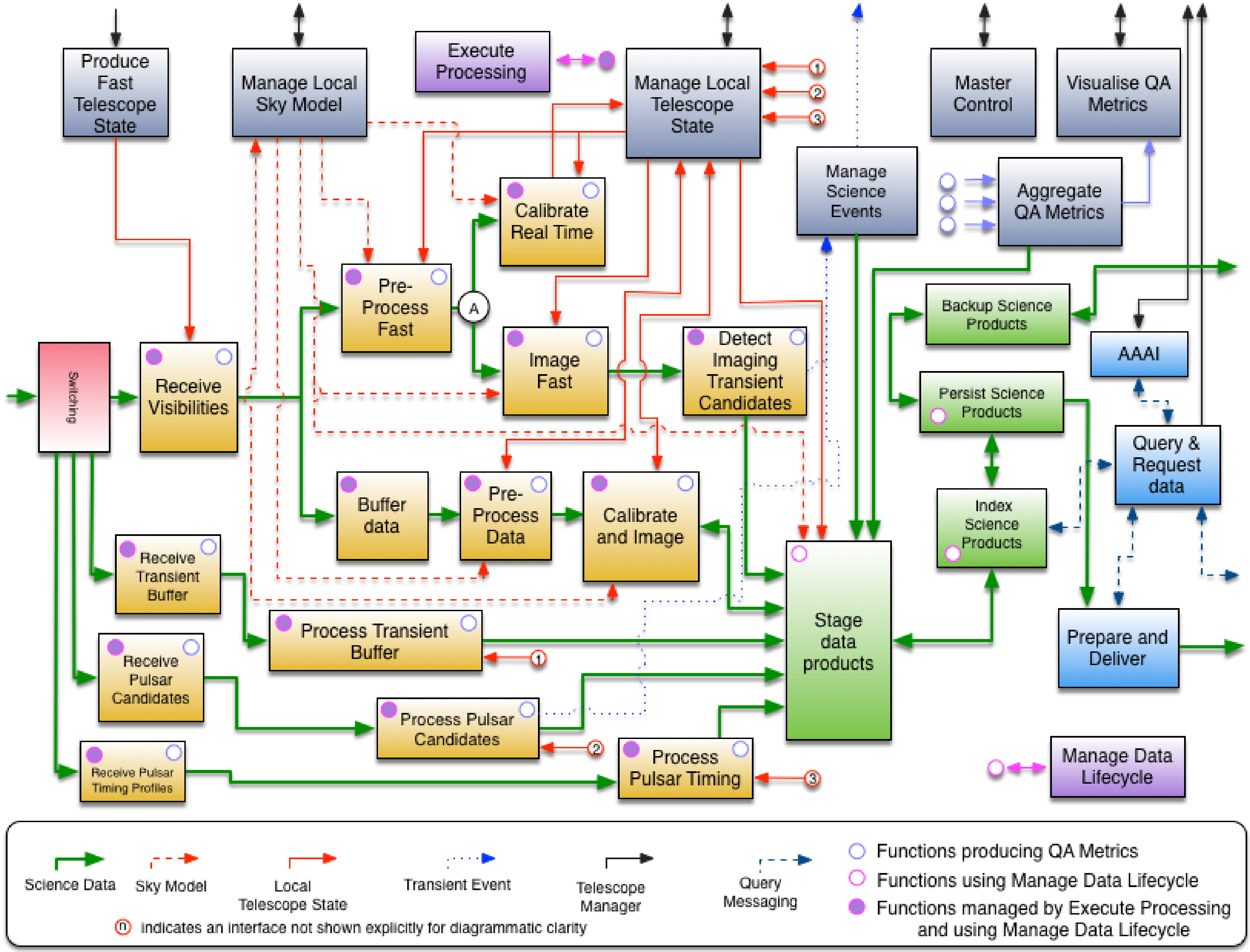
The SDP System



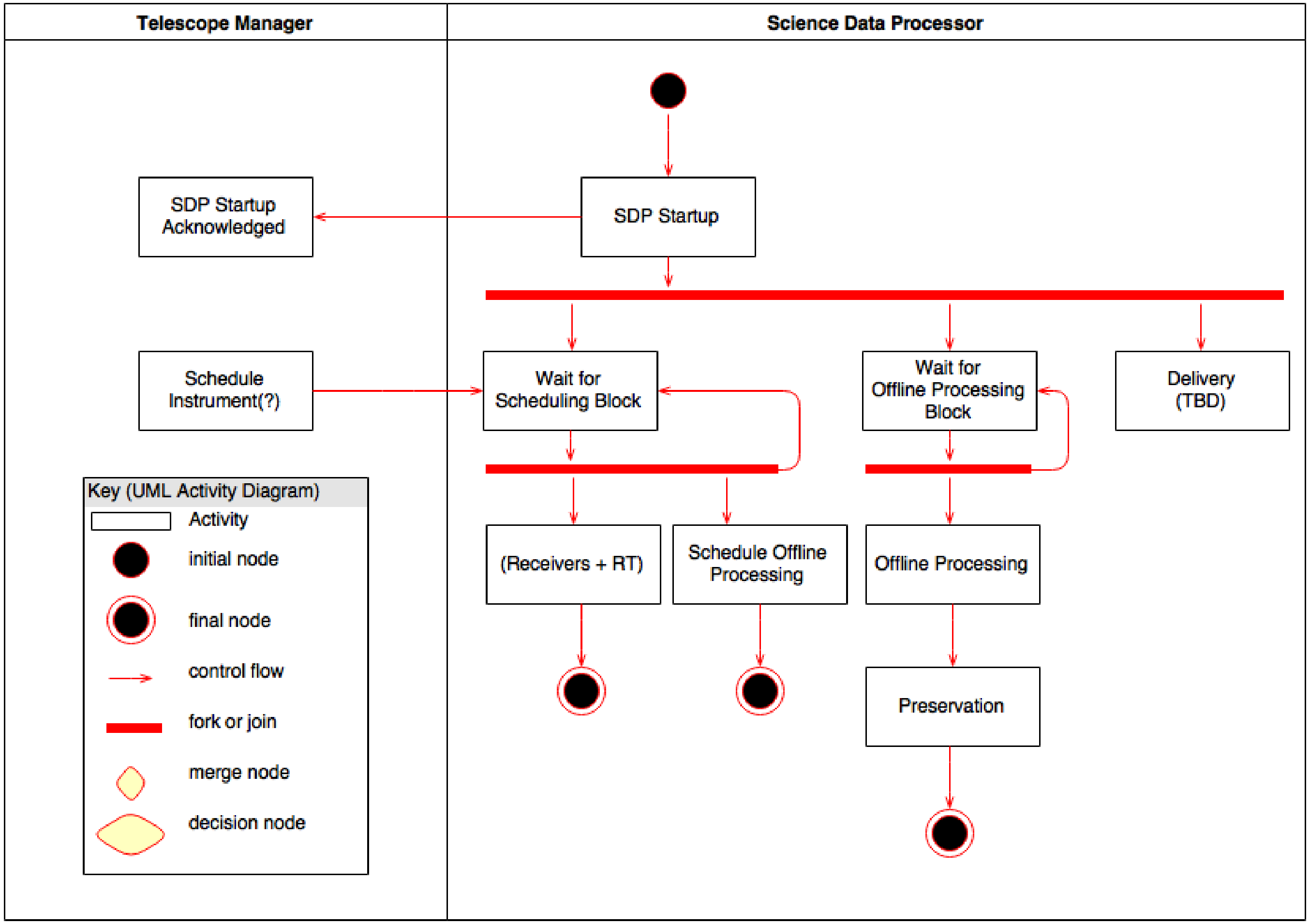
Multiple instances of the process data component, some real time some batch

Illustrative Computing Requirements

- ~250 PetaFLOP system
- ~200 PetaByte/s aggregate BW to fast working memory
- ~80 PetaByte Storage
- ~1 TeraByte/s sustained write to storage
- ~10 TeraByte/s sustained read from storage
- ~ 10000 FLOPS/byte read from storage
- ~2 Bytes/Flop memory bandwidth


Produce Fast Telescope State
Manage Local Sky Model
Execute Processing
Manage Local Telescope State
Master Control
Visualise QA Metrics
Manage Science Events
Aggregate QA Metrics
Calibrate Real Time
Pre-Process Fast
A
Image Fast
Detect Imaging Transient Candidates
Backup Science Products
AAAI
Persist Science Products
Query & Request data
Switching
Receive Visibilities
Buffer data
Pre-Process Data
Calibrate and Image
Index Science Products
Stage data products
Receive Transient Buffer
Process Transient Buffer
Prepare and Deliver
Receive Pulsar Candidates
Process Pulsar Candidates
Receive Pulsar Timing Profiles
Process Pulsar Timing
Manage Data Lifecycle
Science Data
Sky Model
Local Telescope State
Transient Event
Telescope Manager
Query Messaging
Functions producing QA Metrics
Functions using Manage Data Lifecycle
Functions managed by Execute Processing and using Manage Data Lifecycle
n indicates an interface not shown explicitly for diagrammatic clarity

What does SDP do?



SDP is coupled to rest of the telescope

Try to make the coupling as loose as possible, but some time critical aspects

For each observation:

- Controlled by a scheduling block
- Run a Real time (RT) process to ingest data and feed back information
- Schedule a batch processing for later
- Must manage resources to SDP keeps up on timesacle of approximately a week

SDP Architecture and Architectural Drivers

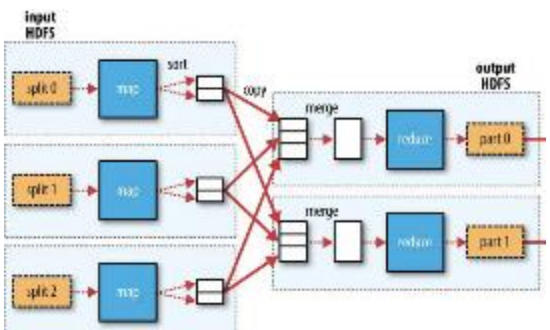
- How do we achieve this functionality
- Architecture to deliver at scale
- Some key considerations
  - SDP is required to evolve over time to support changing work flows required to deliver the science
  - Different Science experiments have different computational costs and resource demands
    - Use for load balancing
    - Expect evolution of system to more demanding experiments during the operational lifetime

How do we get performance and manage data volume?

Approach: Build on BigData Concepts

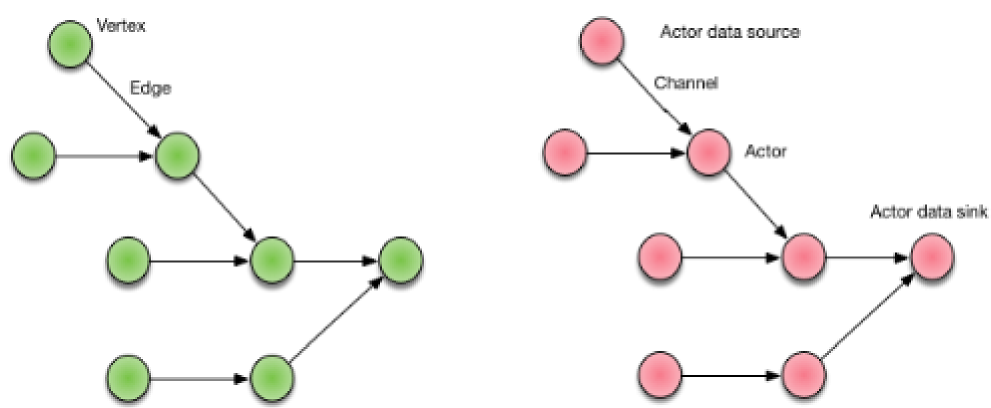
"data driven" → graph-based processing approach receiving a lot of attention

Inspired by Hadoop but for our complex data flow

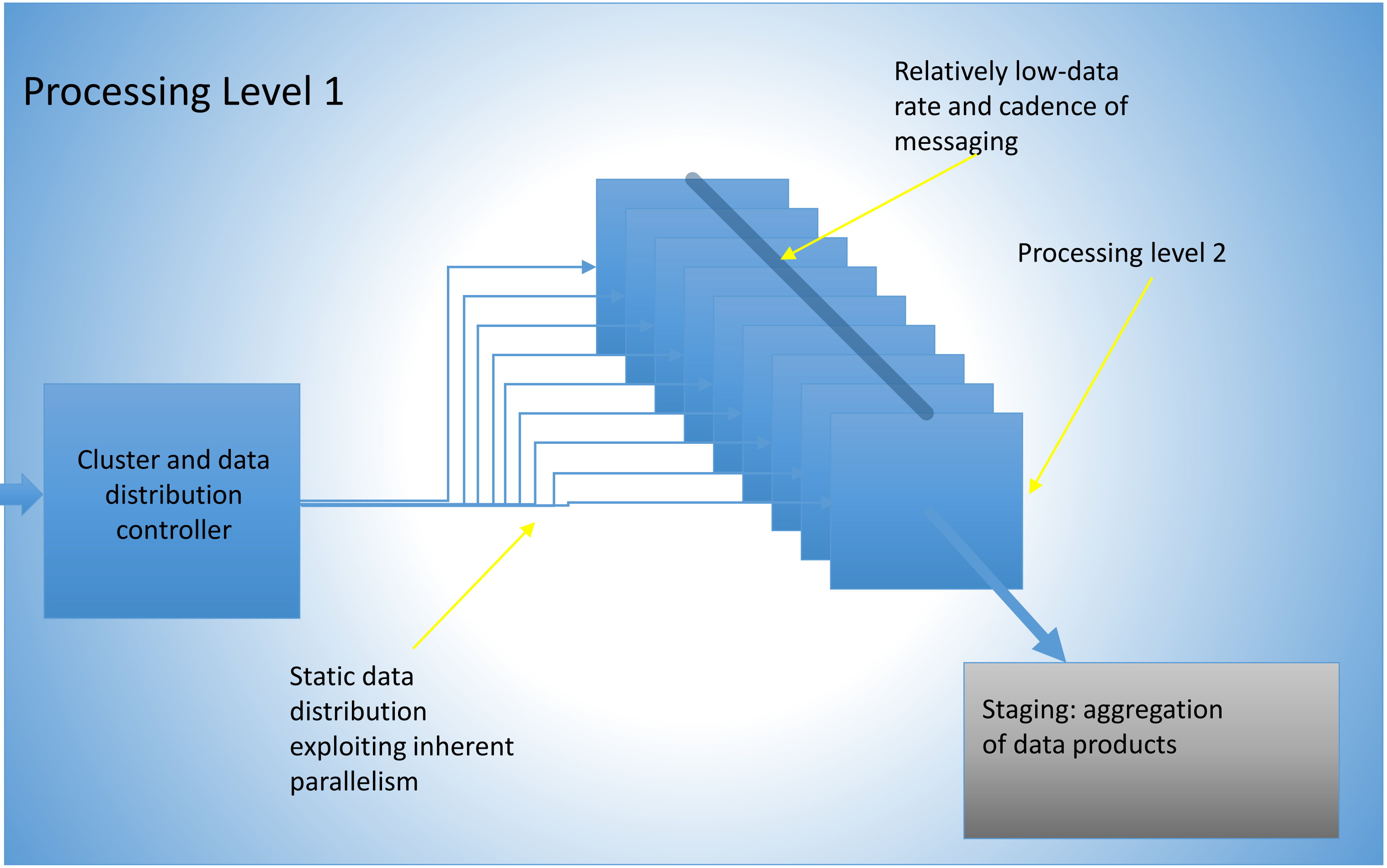


Hadoop

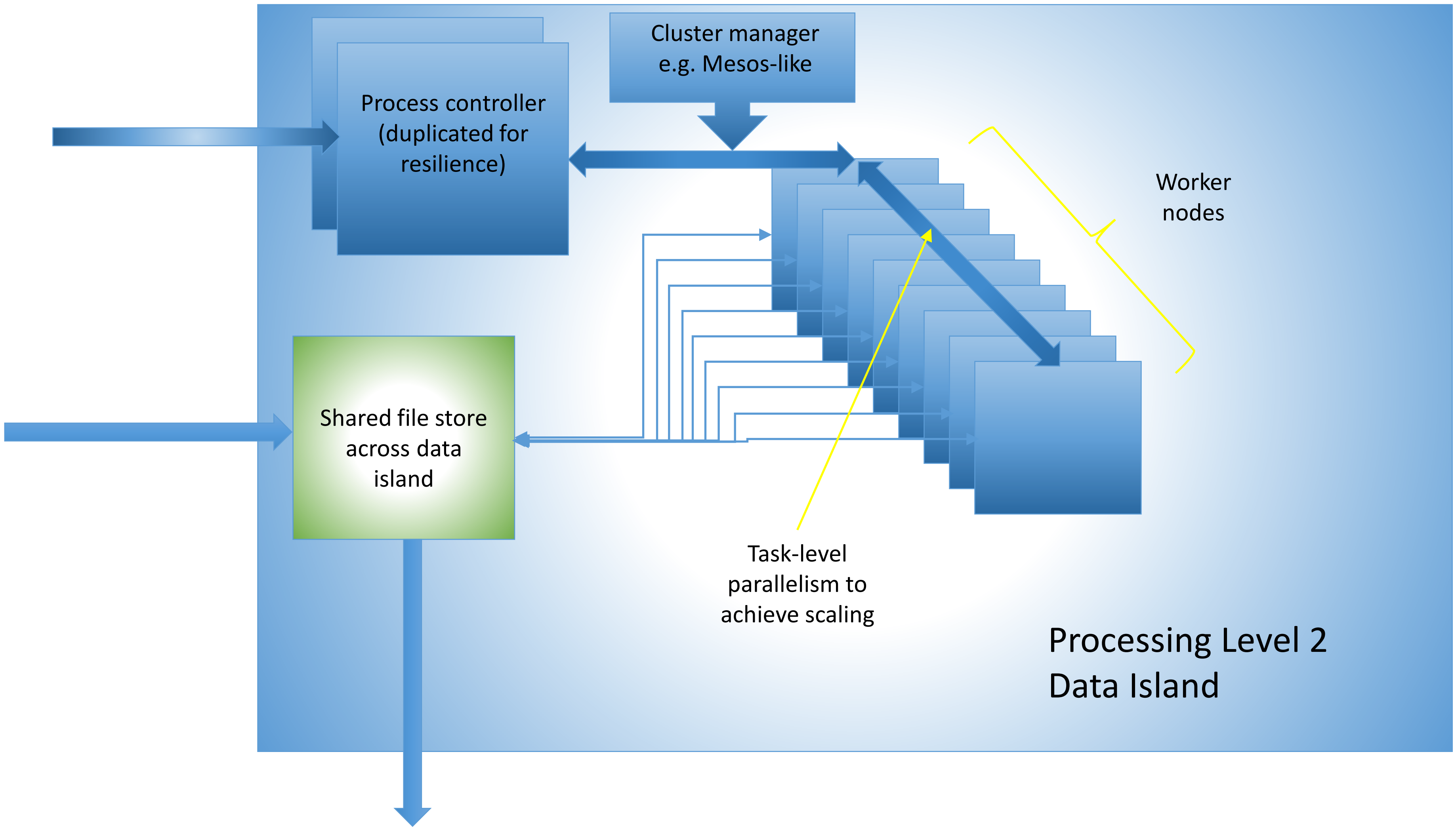
Graph-based approach



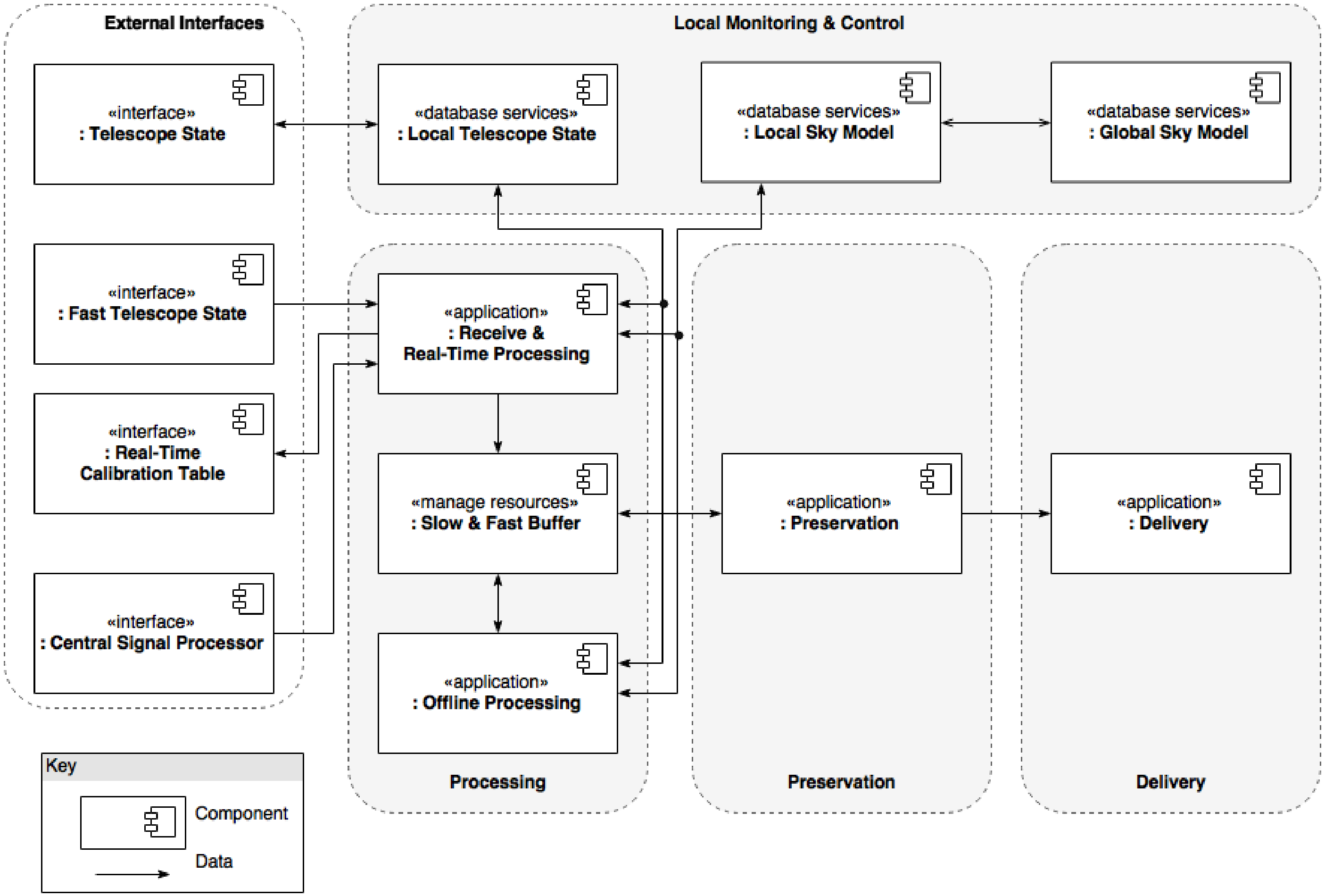
Execution Engine: hierarchical



Execution Engine: hierarchical



Data flow for an observation



What do we need

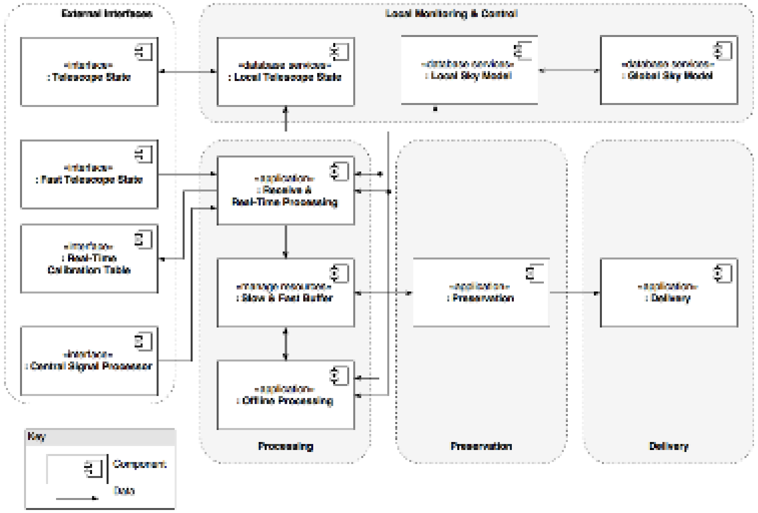
- Processing Level 1:
  - Custom framework to provide scale out
- Processing Level 2:
  - Many similarities to Big-Data frameworks
  - Need to support multiple execution frameworks
    - Streaming processing – possibly Apache STORM
    - Work flows with relatively large task size or where data-driven load balancing is straight forward – possibly development of prototype code
    - ”Big Data” workflows with load balancing, resilience – possibly something like Apache SPARK
  - Framework manages tasks
    - Tasks need to be pure functions
    - Most tasks domain specific, but need supporting infrastructure
      - Processing libraries
      - Interfaces to the execution framework
      - Performance tuning on target hardware

What do we need

- Skills and work – e.g. for SPARK
  - Need to develop for High Throughput
    - High Throughput data analytics framework (HTDAF)
    - New data models
    - Links to external processes
    - Memory management between framework and processes
  - Likely to have significant applicability
  - Shared file system needs to be very intelligent about data placement / management or by tightly coupled with the HTDAF
  - Integration of in-memory file system / object store systems with the cluster file system / object store
  - Software development in data analytics, distributed processing systems, JVM, …

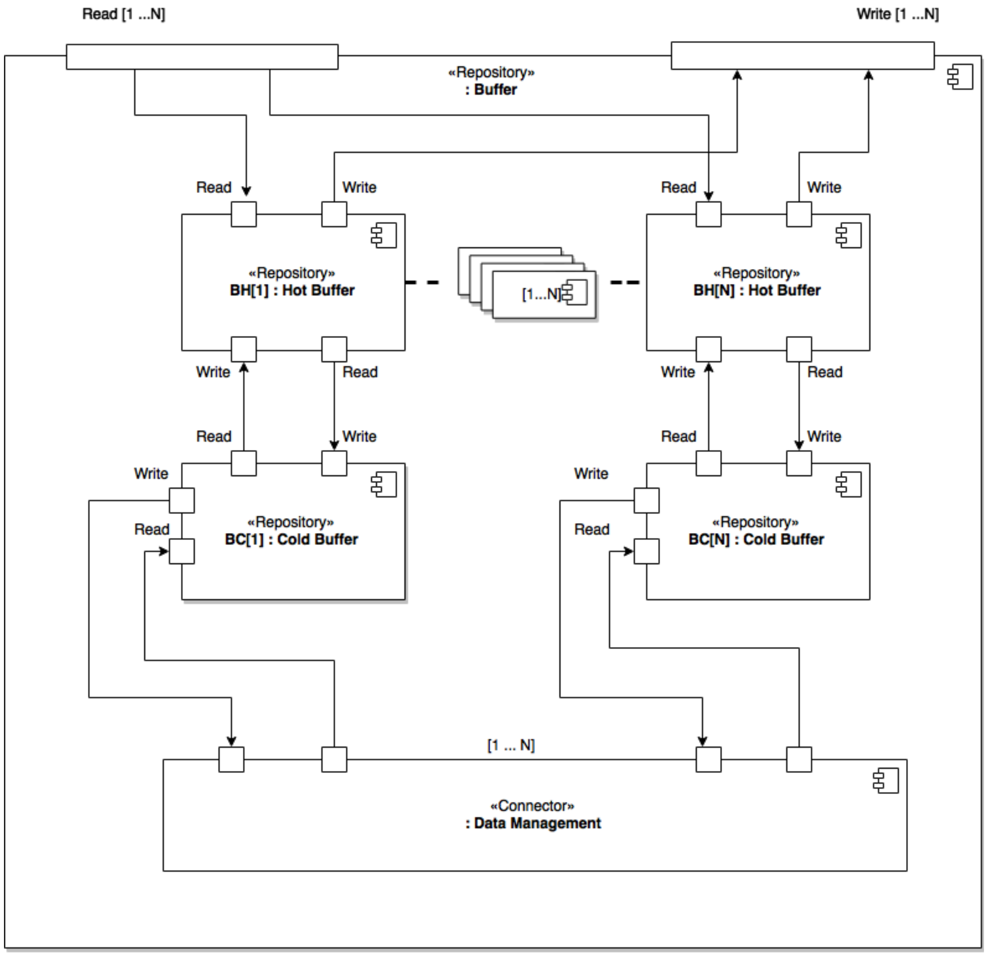
Database services

- Tasks running within processing framework need metadata and to communicate updates / results
  - High performance distributed database
  - Prototyping using REDIS
  - Needs to interface to processing requirements
  - Schemas

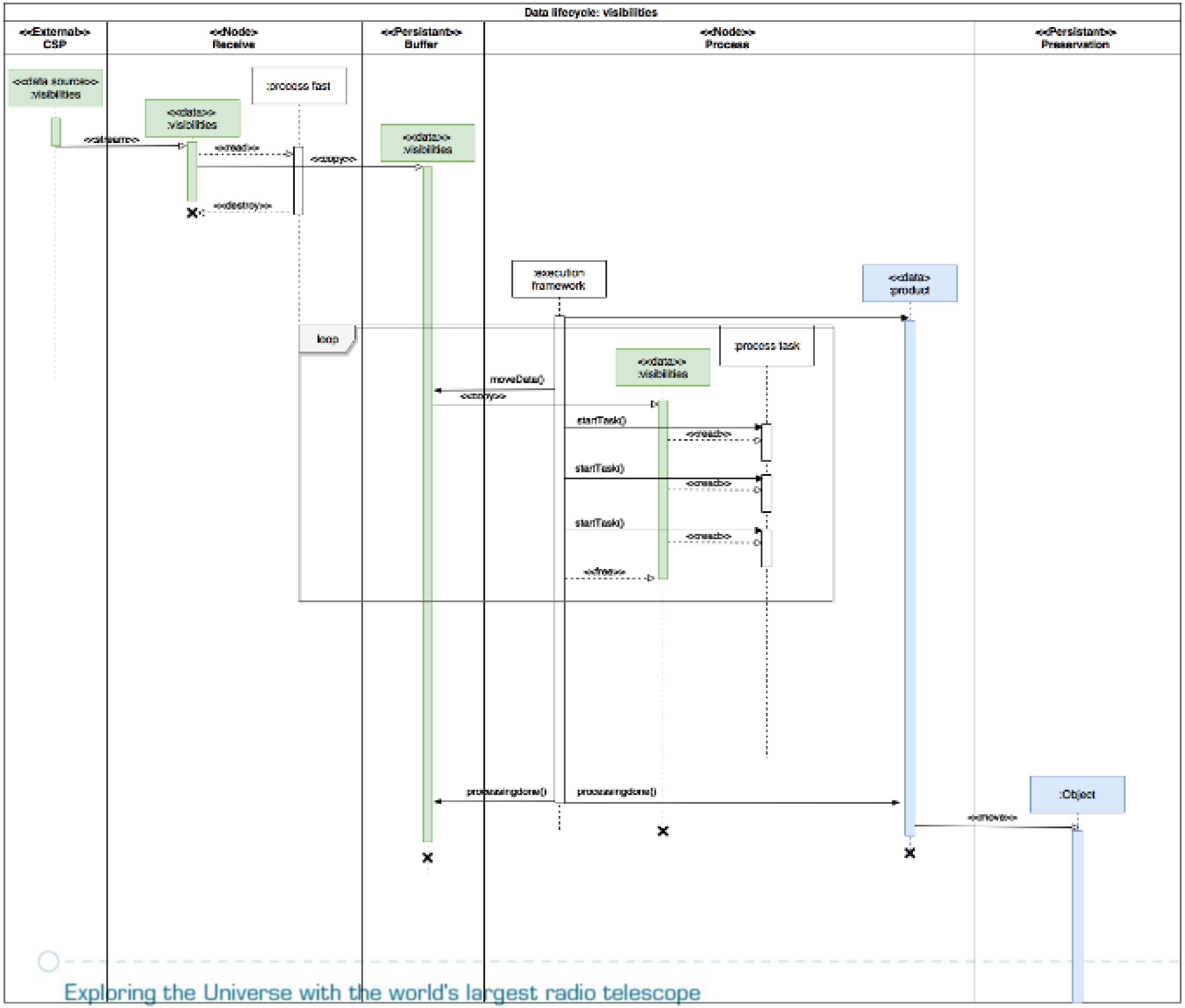


File / object store

- Multiple file / object stores, one per "data island"
  - Integration into execution framework
  - Resource allocation integrated with cluster management system
  - Complex resource management
  - Migration / interface to long-term persistent store likely to be separate Hierarchical Storage Manager



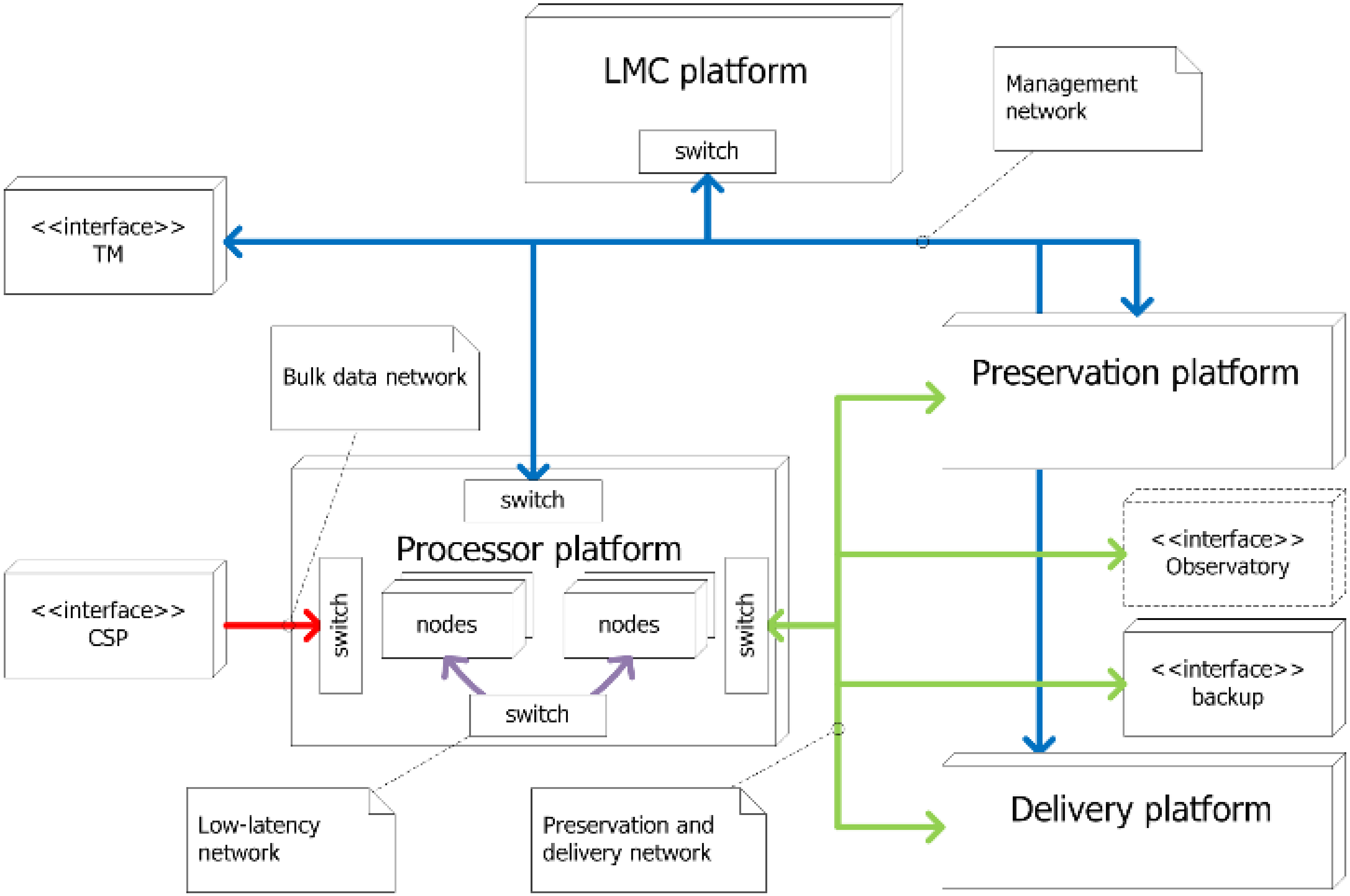
Data lifecycle management



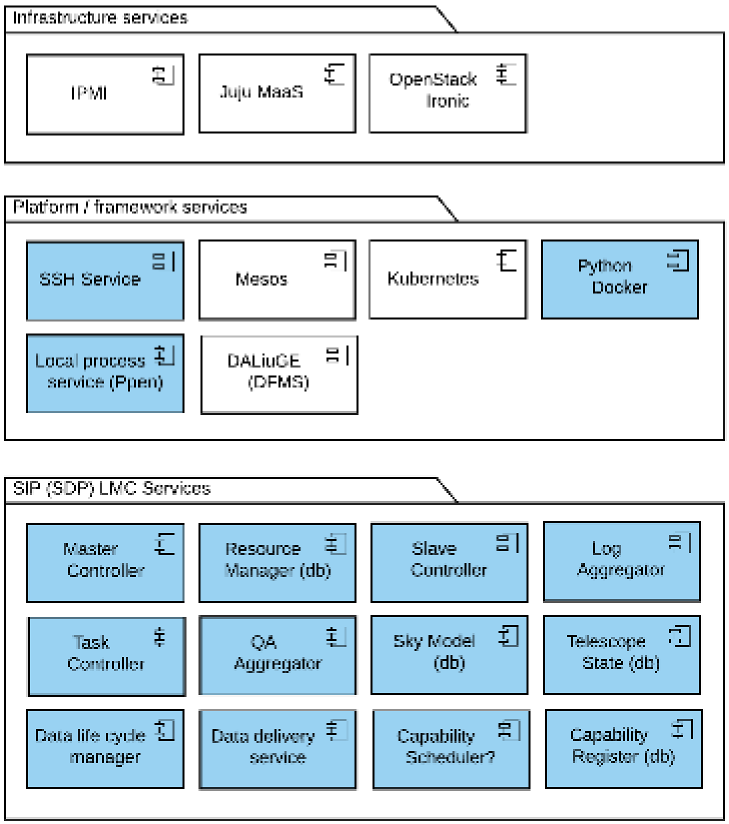
Manage data lifecycle across the cluster and through time

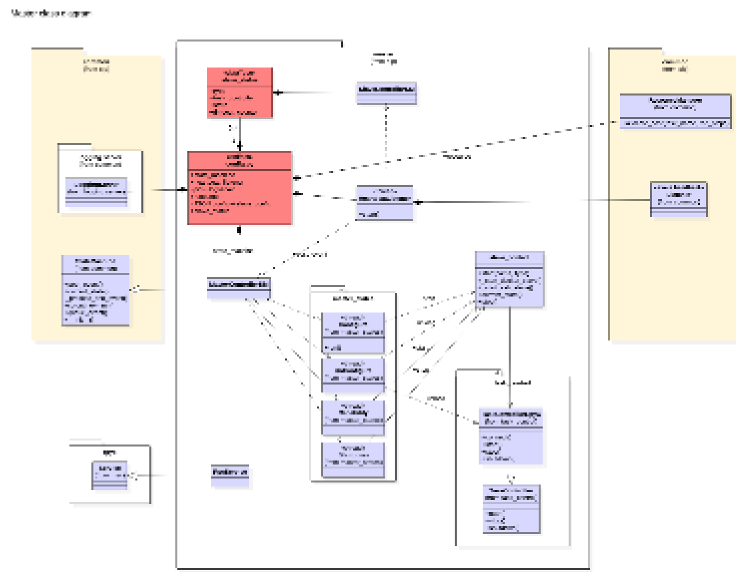
- Strong link to overall processing and observing schedule
- Performance is key
- Determine initial data placements
- Manage data movement and migration
- Local high performance obtained by light-weight layer (e.g. policies) on a file system

Hardware Platform



Control and system components





Other system Software

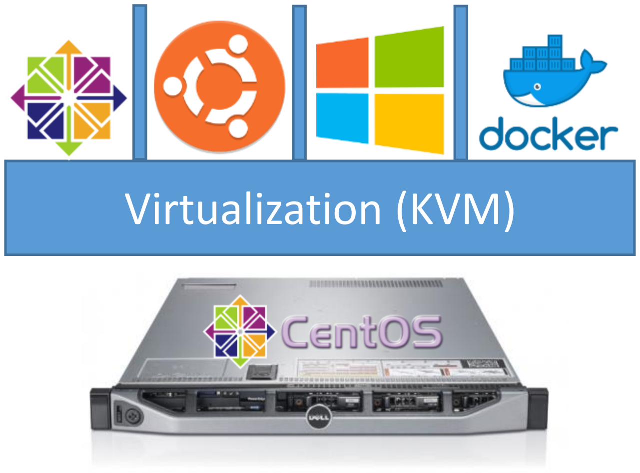
Compute System Software

- **Compute Operating System software**
  - The O/S will be based on a widely-adopted community version of Linux such as Scientific Linux (Fermi Lab, et al) or CentOS (CERN, et al)
- **Middleware**
  - **Messaging Layer** based on TANGO's messaging layer and/or open source messaging layers (like ZeroMQ) is possible with effort
- **Interface to TANGO**
  - TANGO is the selected product for overall monitor and control – interfacing of other components to TANGO will be required for e.g. monitor information from the cluster manager
- **Middleware stack**
  - Based on widely adopted versions of community-supported software that will mature up to deployment, like OpenStack. Effort will be required during the construction period to develop enhancements and SDP-specific customisation. Effort will also be required to maintain enhancements, relevant to the hardware and software requirements. Close collaboration and convergence with large contributors (like CERN) in the open source community could lead to reduced long-term maintenance effort.

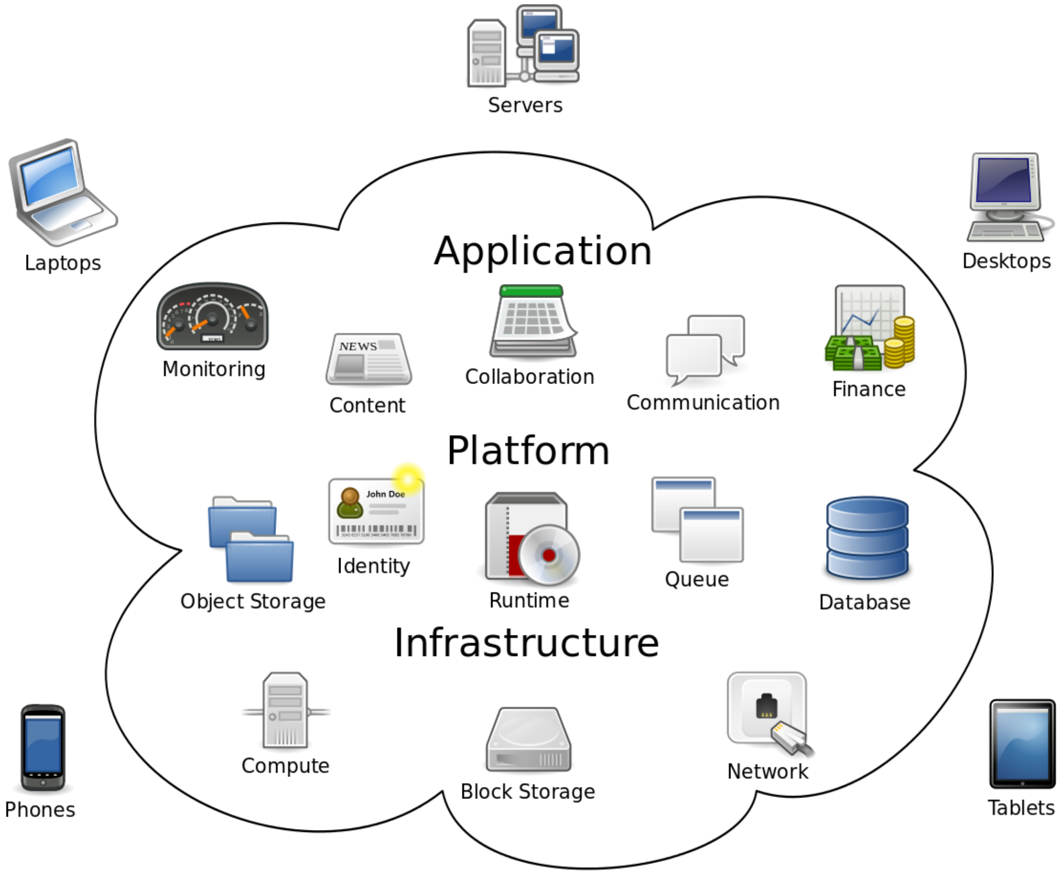
- At Aveiro:

Science Data Processor (SDP)

The Basic principle of Cloud Computing



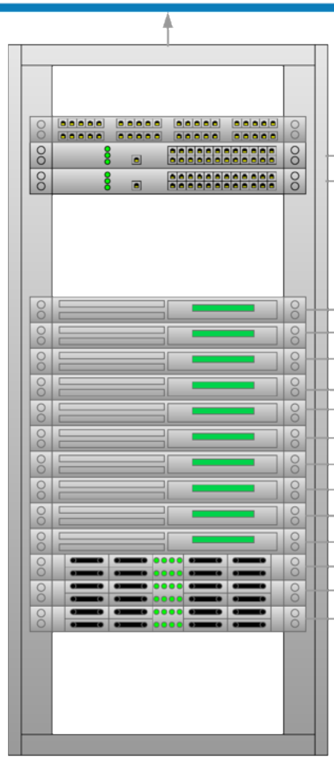
Cloud computing



Prototype of Cloud Infrastructure
"BASED ON ALASKA – SDP SYSTEM"

EngageSKA Cluster

Compute: 232C/464T
Memory: 2.5TB of RAM
SSD: 21.76TB
HDD: 33.60TB
Link: 10Gigabit with redundancy



10Gigabit Switch (Redundancy)
Server: Dell N4032F
Port: 24 10Gigabit

Compute Node 512
Server: Dell PowerEdge R630
CPU: 2x Intel Xeon E5-2650 v4 2.2GHz 12C/24T
RAM: 512 GB
SSD: 2x 400GB in RAID 0
Link: 2x 10Gigabit and 2x 1Gigabit

Compute Node 128
Server: Dell PowerEdge R630
CPU: 2x Intel Xeon E5-2650 v4 2.2GHz 12C/24T
RAM: 128 GB
SSD: 2x 400GB in RAID 0
Link: 2x 10Gigabit and 2x 1Gigabit

Controller Node (Redundancy)
Server: Dell PowerEdge R630
CPU: Intel Xeon E5-2620 v4 2.1GHz 8C/16T
RAM: 64 GB
HDD: 4x 600GB SAS in RAID 4
Link: 2x 10Gigabit and 2x 1Gigabit

Storage Node (SSD)
Server: Dell PowerEdge R630
CPU: Intel Xeon E5-2620 v4 2.1GHz 8C/16T
RAM: 64 GB
SSD: 8x 1.92TB 6Gbps
Link: 2x 10Gigabit and 2x 1Gigabit

Storage Node (SAS)
Server: Dell PowerEdge R630
CPU: Intel Xeon E5-2620 v4 2.1GHz 8C/16T
RAM: 64 GB
HDD: 8x 1.8TB 10K RPM SAS 12Gbps
Link: 2x 10Gigabit and 2x 1Gigabit

Orchestration

vmware

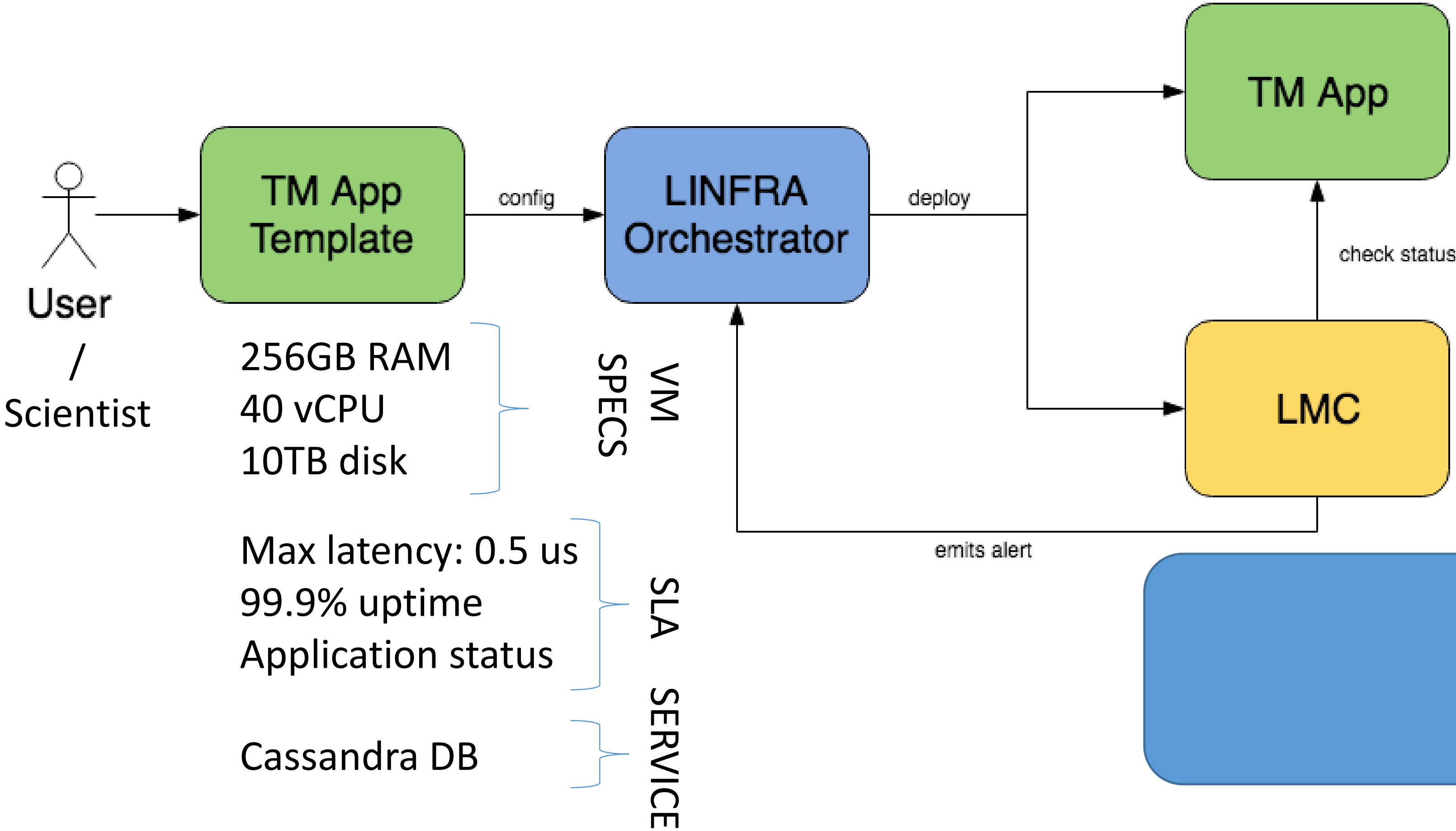
business oriented

openstack™

kubernetes

docker

Orchestration of Cloud Services with Critical Components in SKA



*TM - Telescope Manager

Summary

- Automatic Deployment
  - Orchestration system
- Monitorization (SNMP)
  - Performance
  - State
  - Network
  - Service
- Failover mechanisms
  - Recovery methods
  - Recovery strategies

ENAbling Green E-science for SKA
Capacitation and Sustainability of Portuguese participation in the SKA

ENGAGE SKA PORTUGAL

Radioastronomy as an Open Innovation Living Lab
A National Infrastructure

5 Academic Partners
1 Interface Institutions
7 Industry partners
1 Competitivity Cluster