

---

# **Square Kilometre Array eInfrastructure: Requirements, Planning, Future Directions**

**Duncan Hall  
SPDO Software and Computing  
ESFRI @ EGEE 2009**

SKA – in a nutshell

How do radio telescope arrays work?

What are the SKA's prime characteristics?

Real-time data: pushing the HPC envelope

How much data do we need to store?

Where are we at?

What is the future direction?

Summary

**A global project:  
75+ institutes in 19 countries**

**Similarities to CERN?**

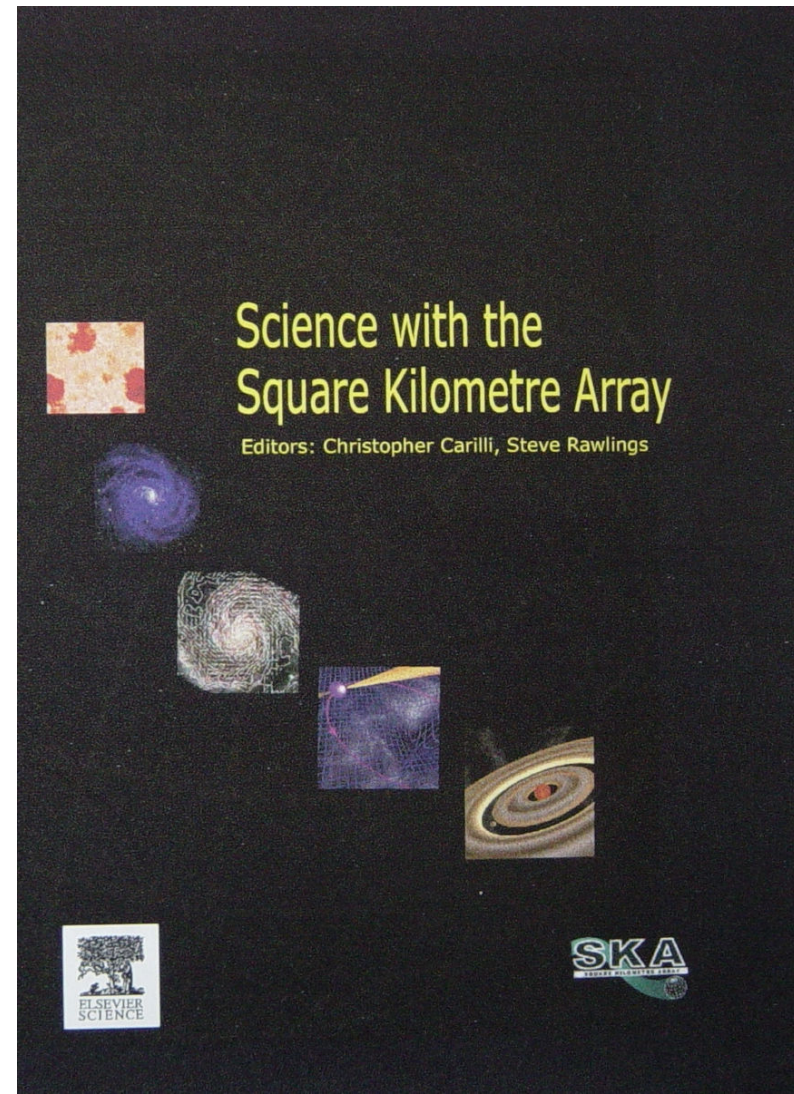
## ■ Origins

- Cosmology and galaxy evolution
- Galaxies, dark matter and dark energy
- Probing the “Dark Ages”
- Formation of the first stars
- Cradle of life
- Search for signs of life

## ■ Fundamental Forces

- Strong-field tests of general relativity
- Was Einstein correct?
- Origin and evolution of cosmic magnetism
- Where does magnetism come from?

## ■ Exploration of the Unknown



---

How do radio telescope arrays work?



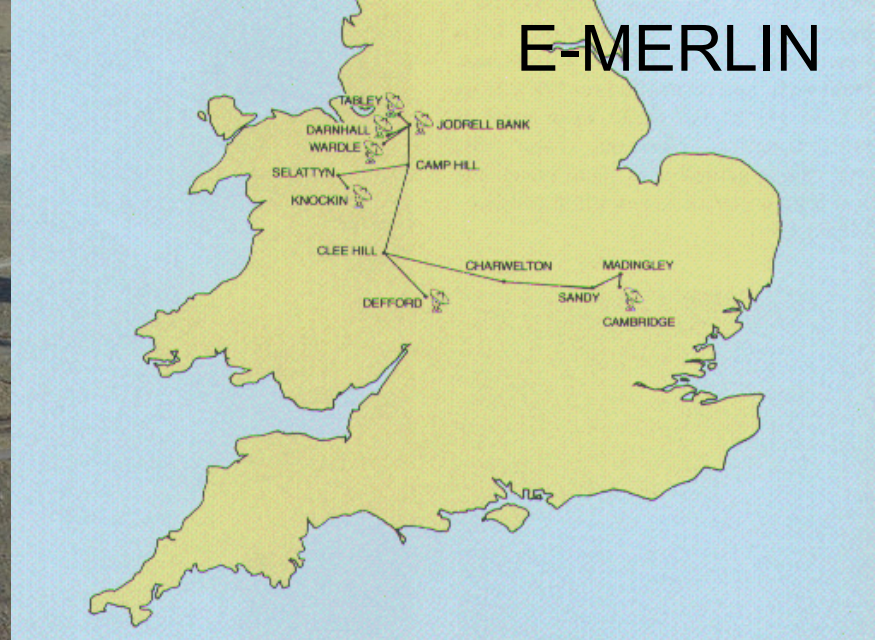
76 m: 3,000+ tonnes







VLA (USA)

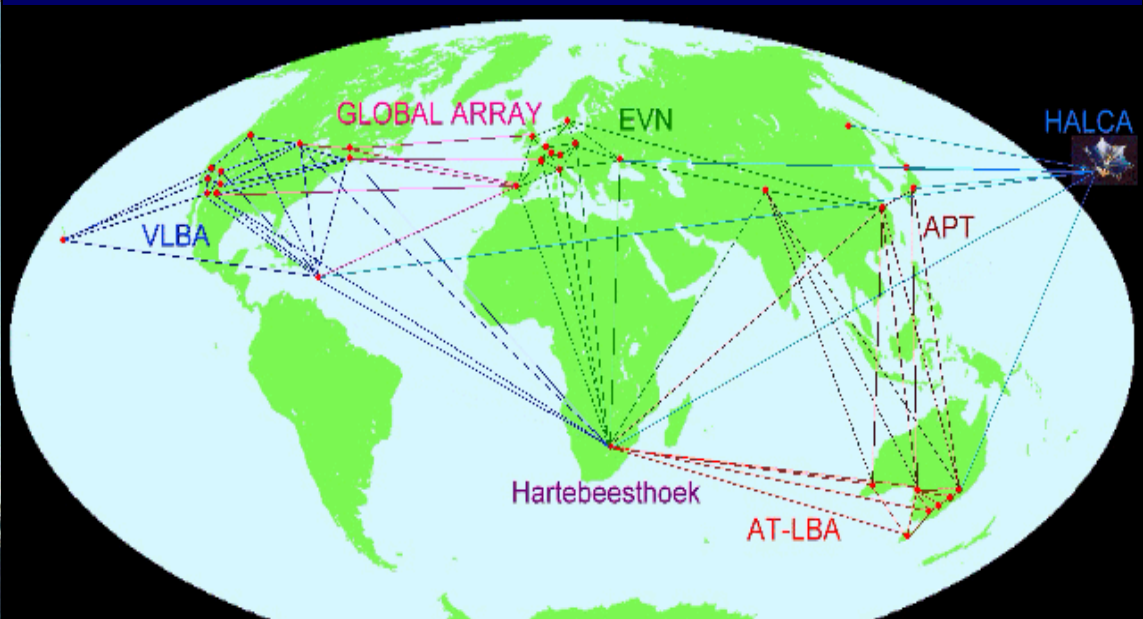


E-MERLIN

# Array Telescopes



Australia Telescope



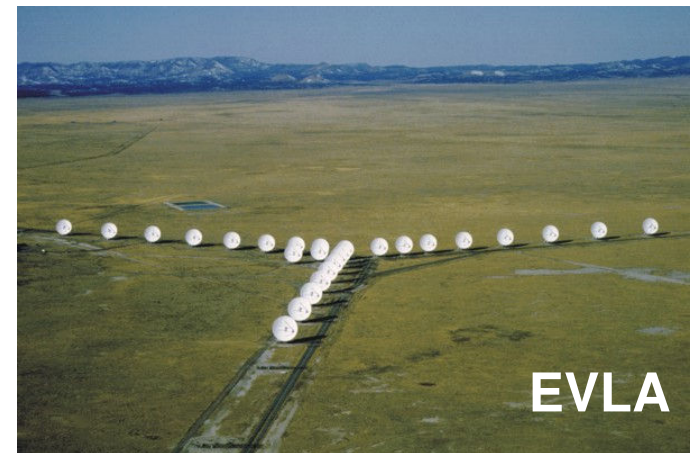
---

What are the SKA's prime characteristics?



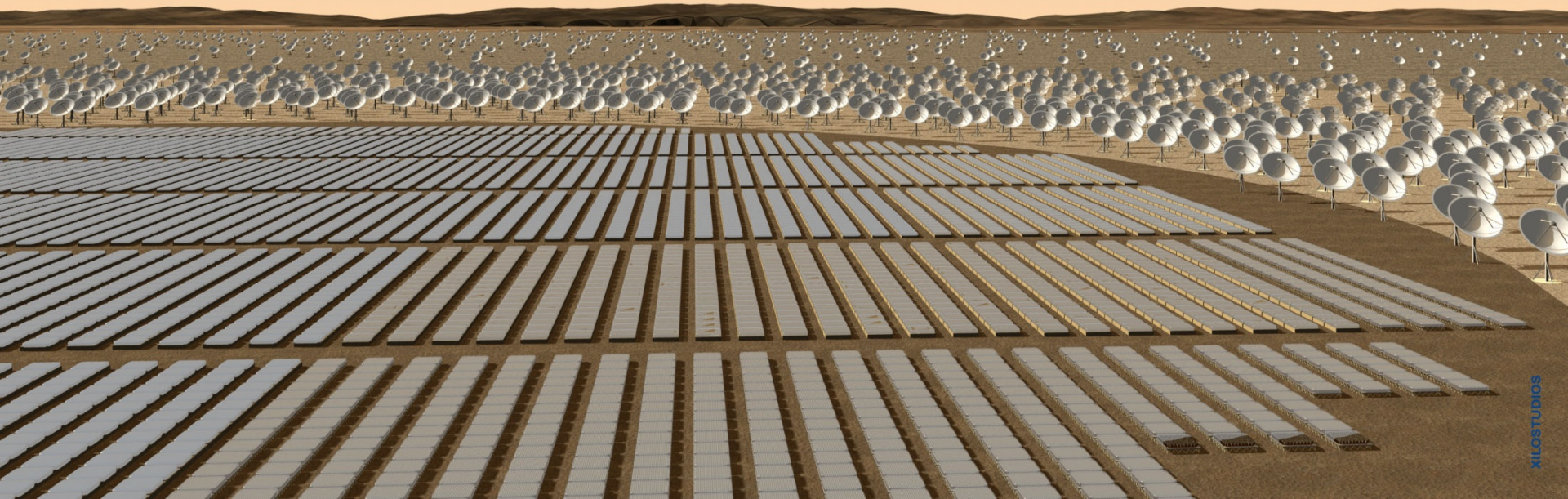
# SKA: prime characteristics

1. **More collecting area:  $\sim 1\text{km}^2$** 
  - Detect and image hydrogen in the early universe
  - Sensitivity  $\sim 50 \times$  EVLA, LOFAR
2. **Bigger field of view**
  - Fast surveying capability over the whole sky
  - Survey speed  $\sim 1,000,000 \times$  EVLA
3. **Wide ranges of frequencies**
  - Low : 70-300 MHz
  - Mid: 300 MHz-10 GHz
  - High: 10-25+ GHz
4. **Large physical extent :  $\sim 3,000+$  km**
  - Detailed imaging of compact objects
  - Astrometry with  $\sim 0.001$  arc second angular resolution



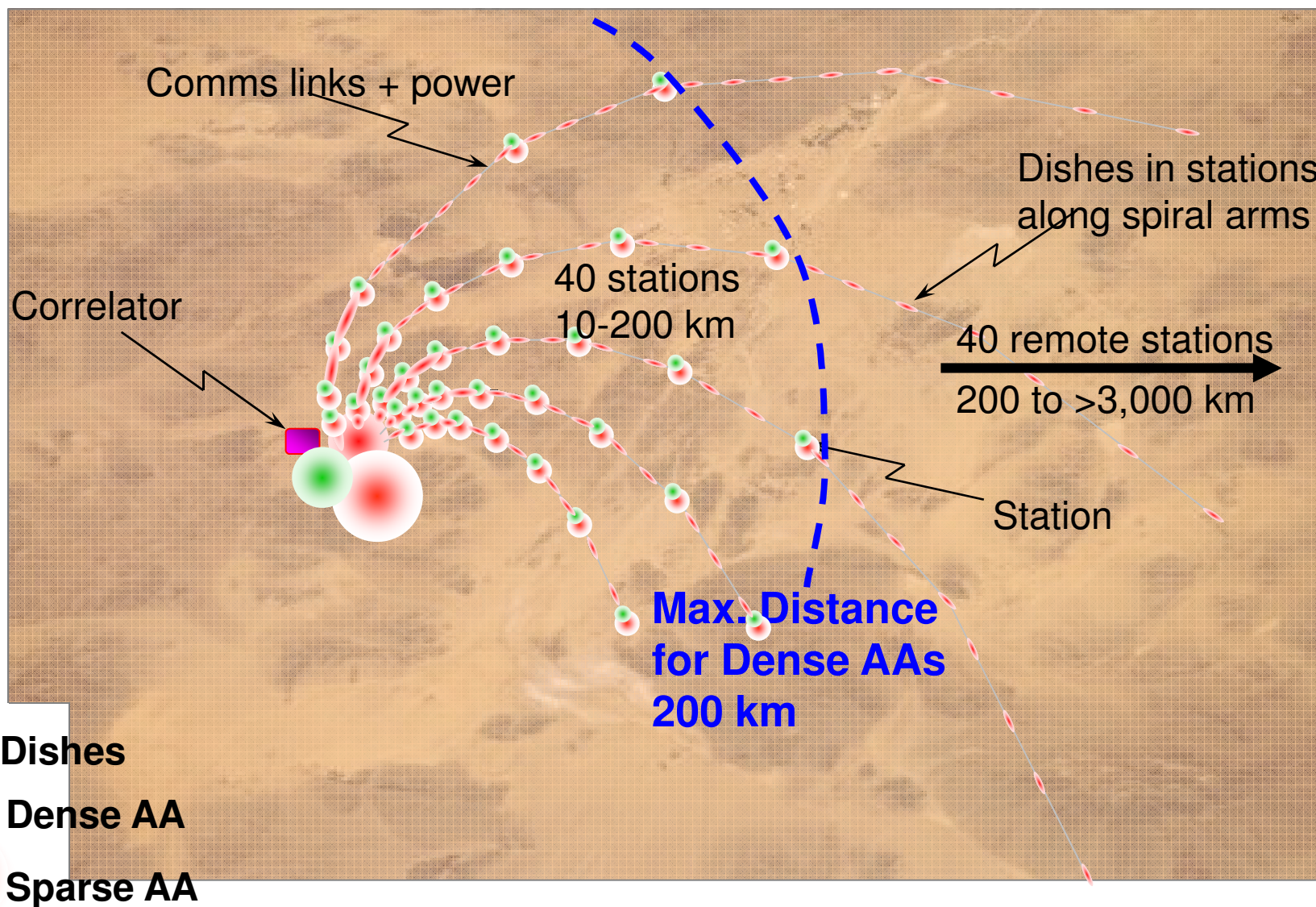
# Artist's impression

- 1,500 dishes (~15m diameter) in a 5 km core
- Additional 1,500 dishes from 5 km to ~3,000+ km
- Aperture arrays (AA) in a core
- Signal processing: (1) beam forming
- Optical fibre connection to (2) correlator
- Optical fibre to remote (3) High Performance Computer



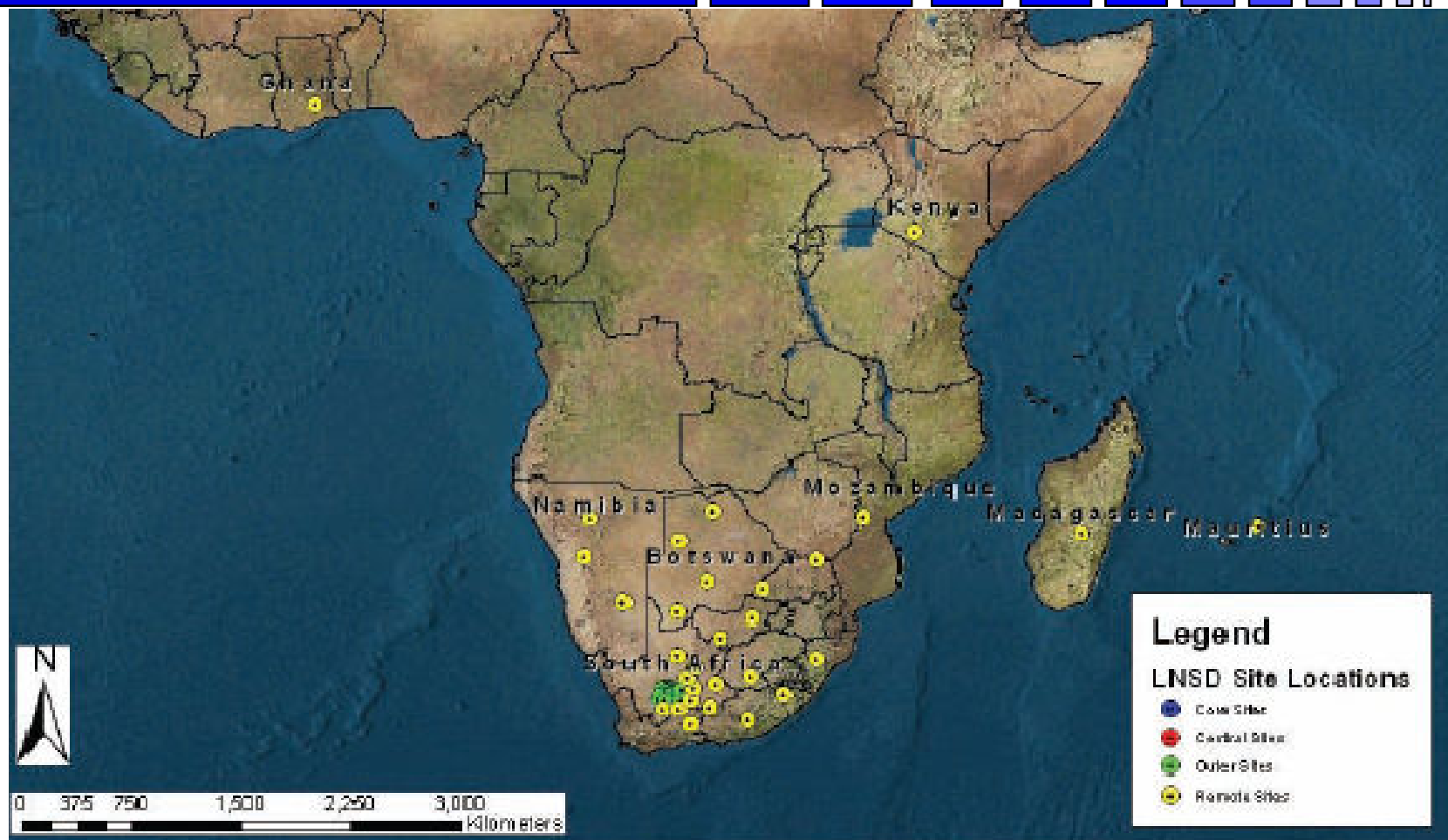


# One possible configuration





# South Africa + 7 countries



# Australia + New Zealand



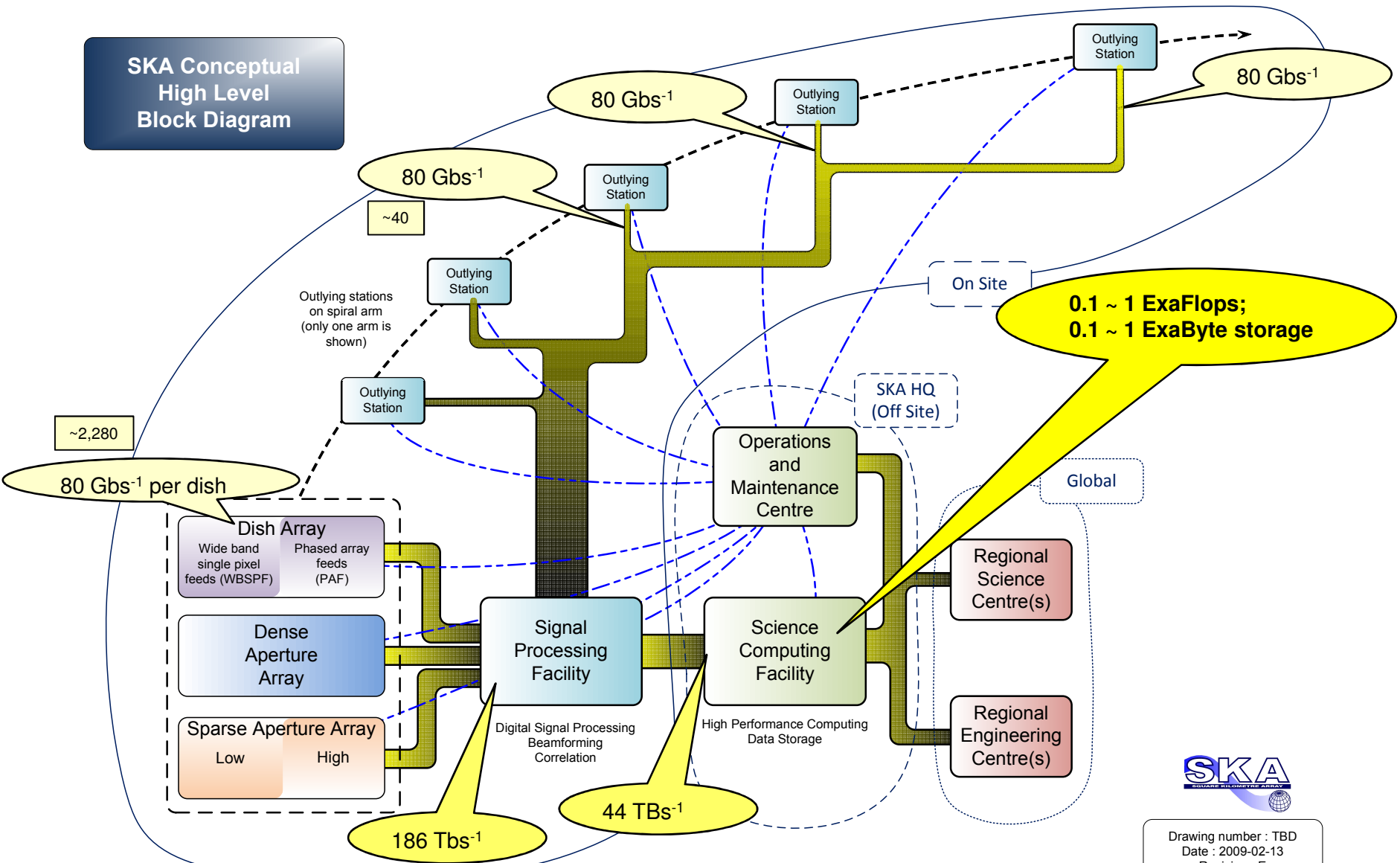
---

Real-time data:  
pushing the HPC envelope ...

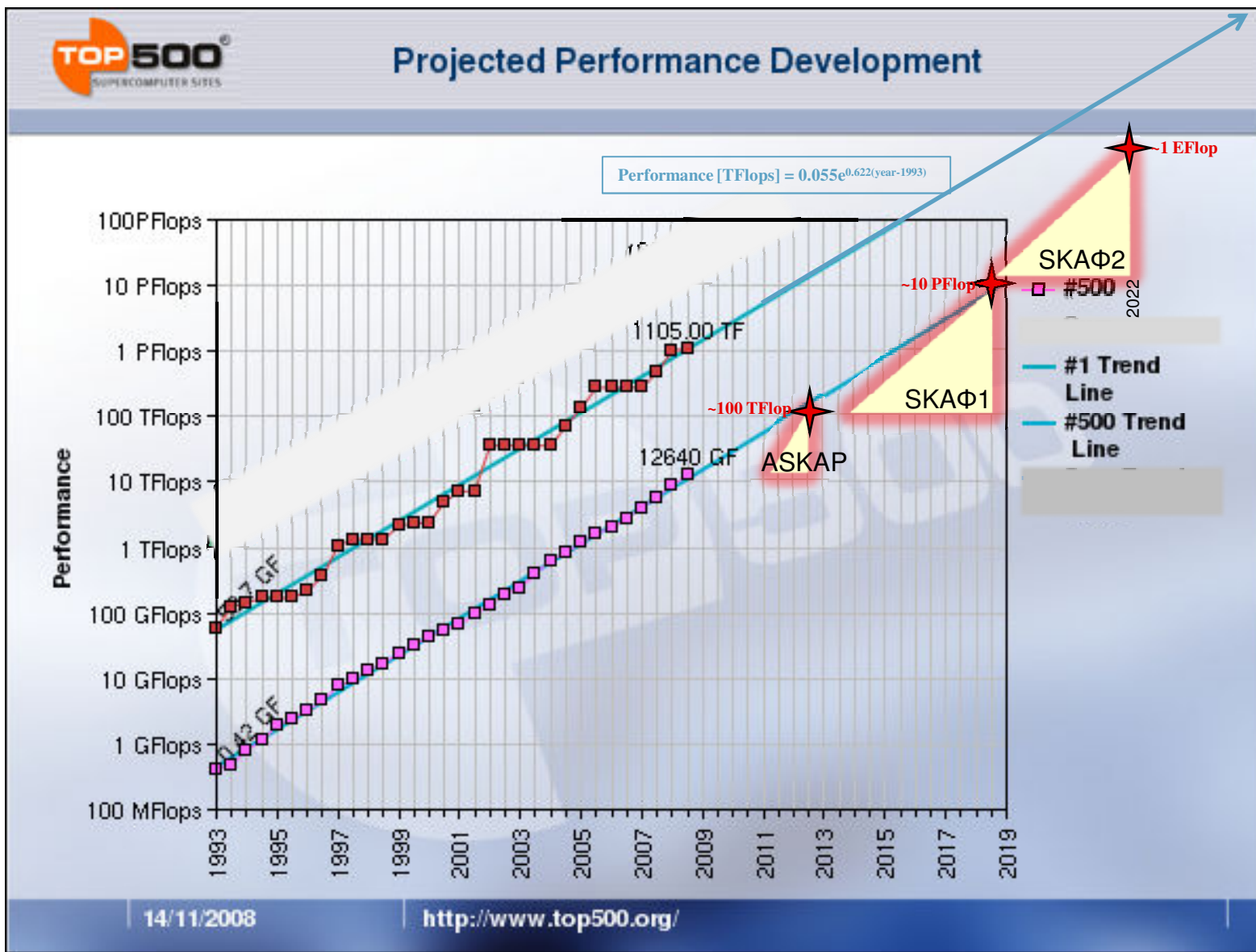


# Φ2 real-time data from dishes

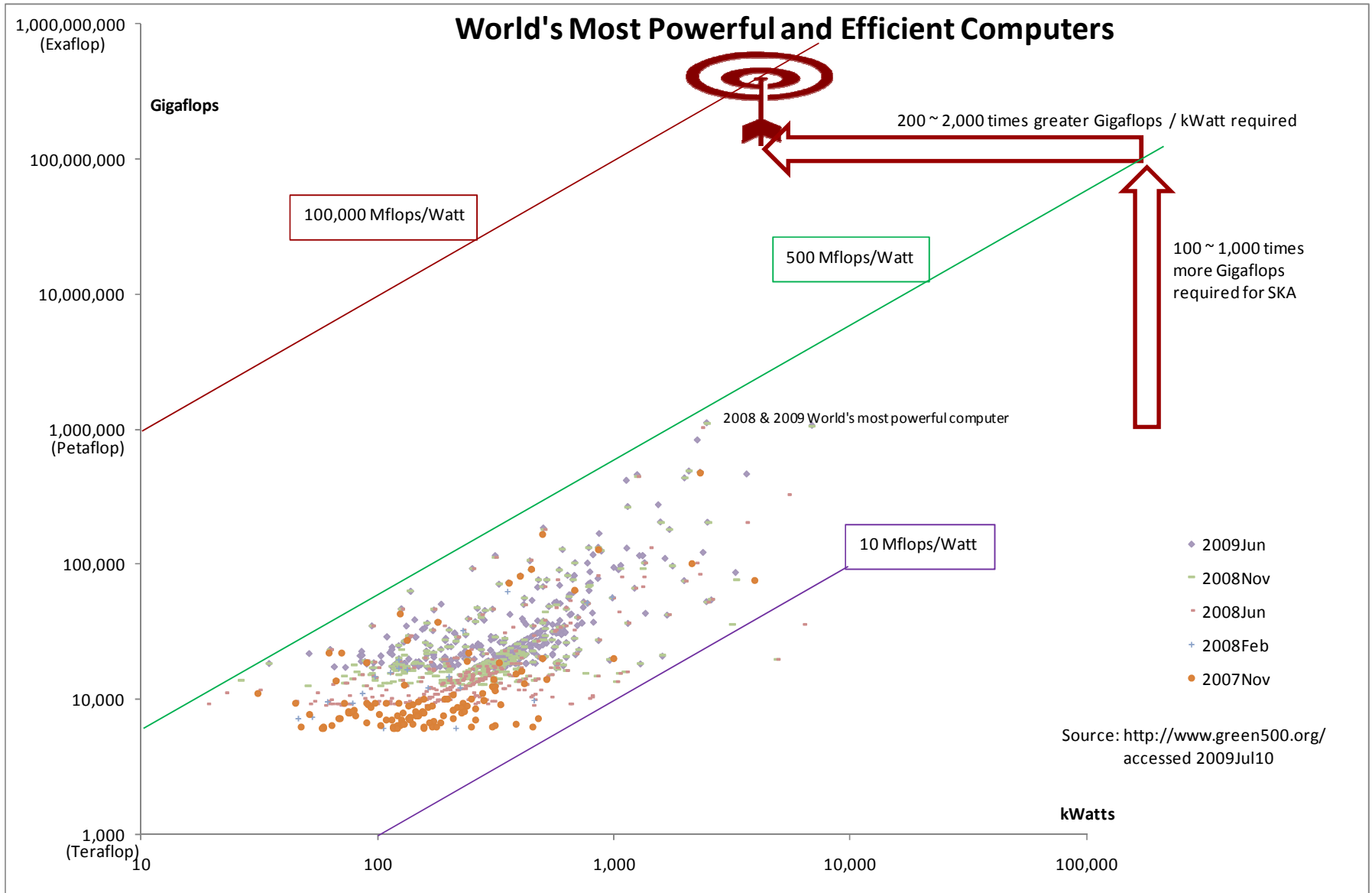
SKA Conceptual High Level Block Diagram



# Pushing the Flops envelope



# Power efficiency challenges:

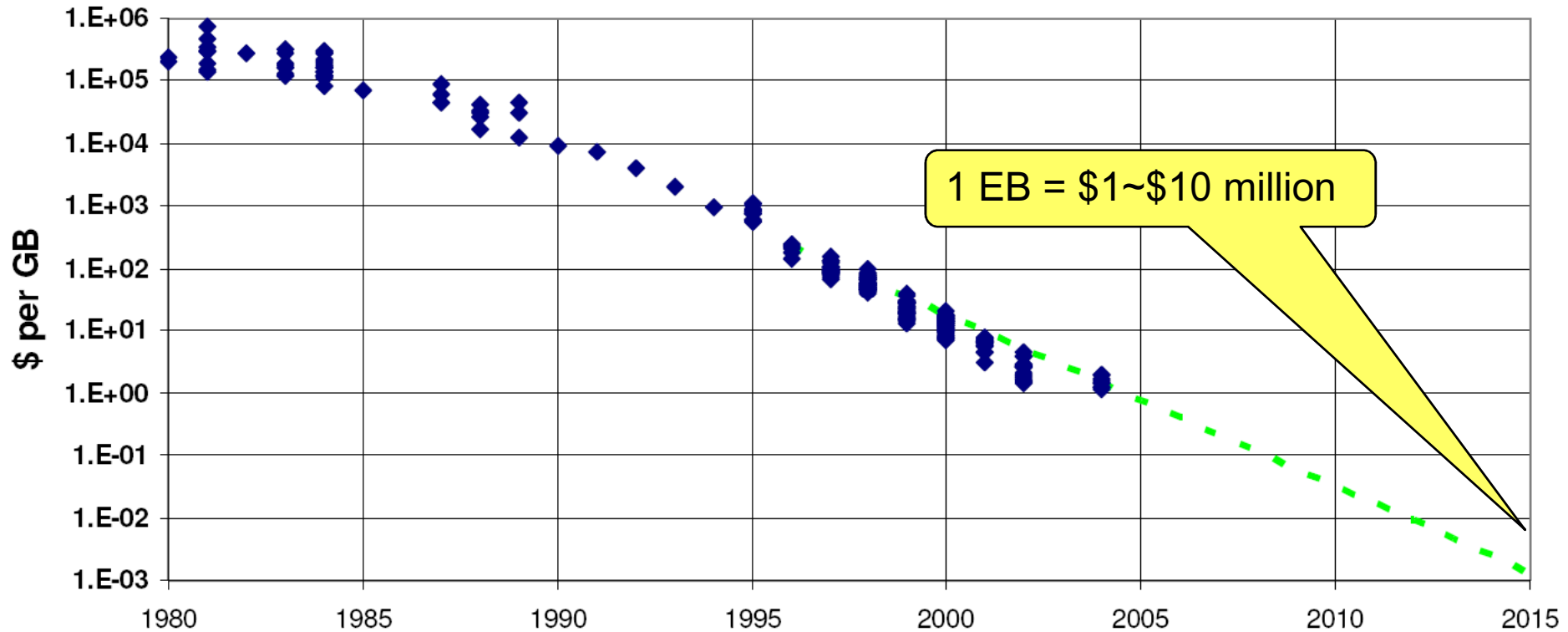




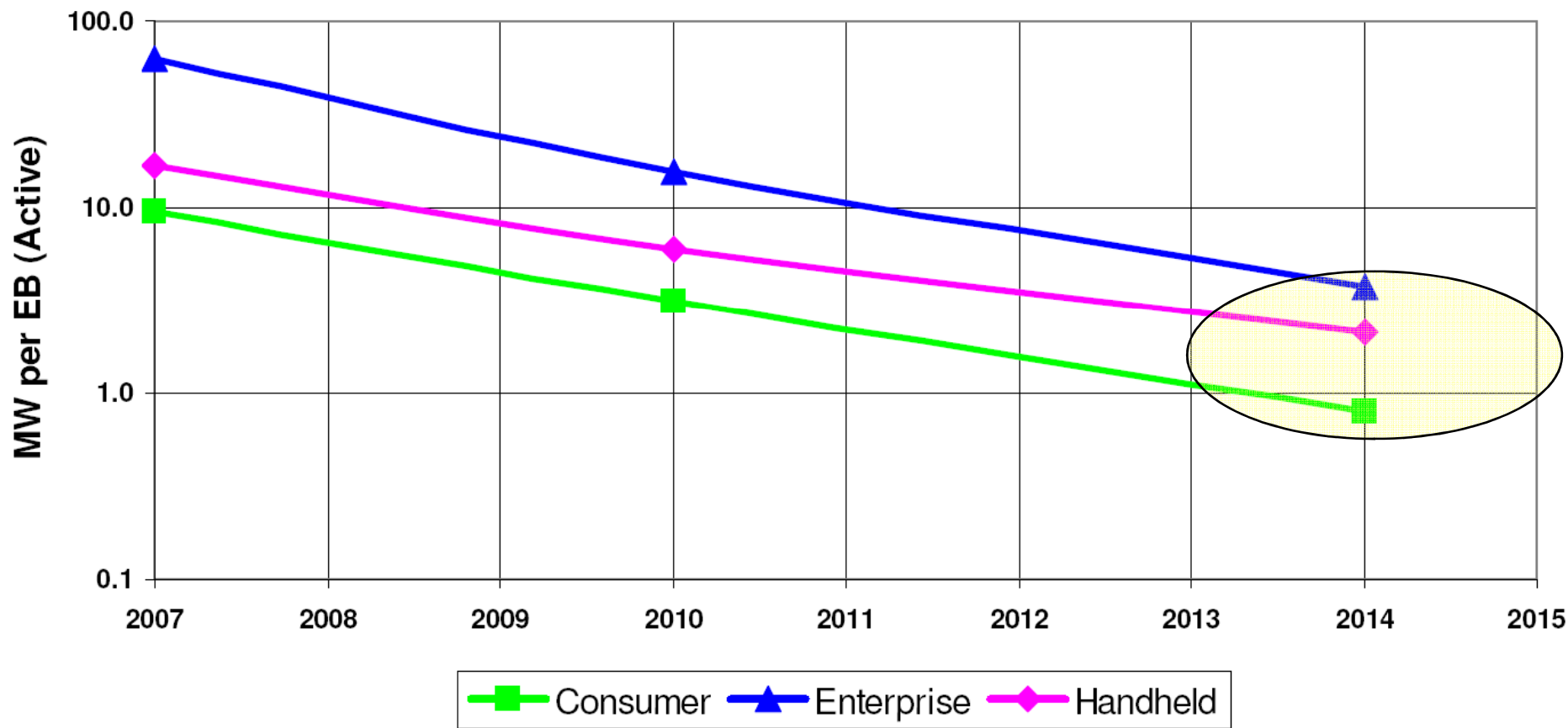
---

How much data do we need to store?

# Disk storage: annual 50% cost reduction



# Power for EB-size disk looks reasonable





# 10 PetaByte tape robot at CERN



500-GB tapes switched to 1-TB models – an upgrade that took a year of continuous load/read/load/write/discard operations, running in the interstices between the data centre’s higher-priority tasks

---

Where are we at?

## PrepSKA 2008-2011

The Preparatory Phase for the SKA is being funded by the European Commission's 7<sup>th</sup> Framework Program

€5.5M EC funding for 3 years + €17M contributed funding from partners (still growing)

€150M SKA-related R&D around the world

Coordinated by the Science and Technology Facilities Council (UK)

Coordinated by the SKA Program Development Office in Manchester

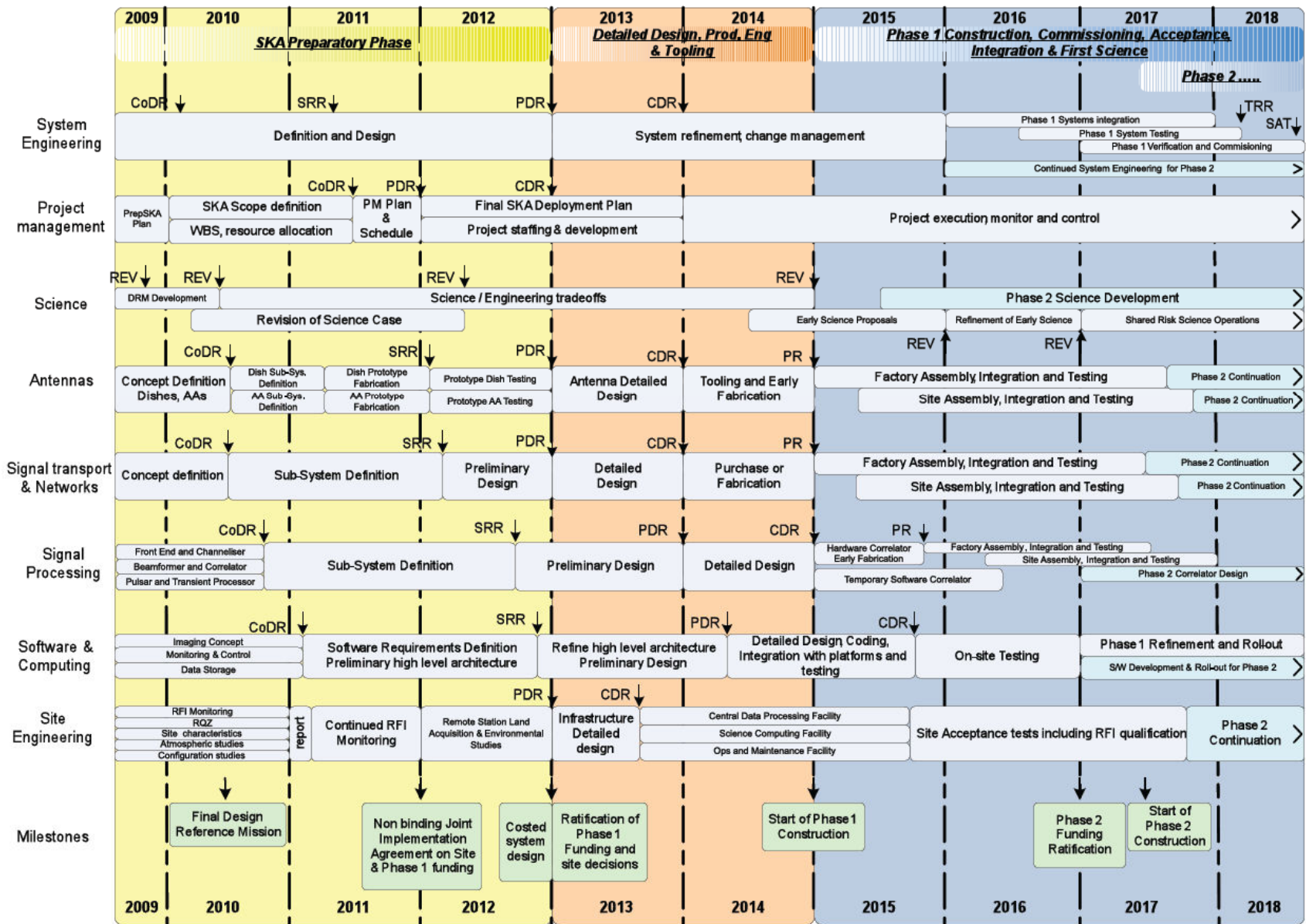
- System Definition
- Dishes, feeds, receivers
- Aperture arrays
- Signal transport
- Signal processing
- Software
- High performance computers
- Data storage
- Power requirements

---

What is the future direction?



# Schedule to 2018



# Phased construction

## ■ Phase 1:

- 2013 – 2018 construction
- 10-20% of the collecting area

## ■ Phase 2:

- 2018 – 2022 construction
- Full array at low and mid frequencies

## ■ Phase 3:

- 2023+ construction
- High frequencies

# Summary

---

## ■ Required eInfrastructure in host region

- Distributed 100's Gb fibre optic signal transport and networks
- Distributed 10's MW power for collecting dishes and stations
- Central Signal Processing Facility: specialised hardware
- Fibre to Science Processing Facility

## ■ Other eInfrastructure: could be anywhere

- Science Processing Facility – ExaFlop “production” environment
- Tiered hierarchy (1-3 ?) of ExaByte storage, archive and distribution
- Software applications development and maintenance environments
- Power for computing and (-H)VAC: 10's MW

## ■ Current state of planning

- Preparatory costing and high level design / architecture development