
SKA
FRANCE
SQUARE KILOMETRE ARRAY
SKA
data processing, storage, distribution
SPDO / Swinburne Astronomy Productions

What are SKA data and compute needs?

Some key numbers

Source: Image Swinburn University

- The ultimate in aperture and collecting area!
  - 3,000 Antennas spread over 5,000km
  - 1,000,000 square meters of collecting area
  - ExaFLOP performance, Terabit data rates, Exabytes per day of output
  - Power consumption of 60MW at the core

- Exa What? – 10^18 or 1,000,000,000,000,000,000
  = 1 billion of billion dpuble precisions (64bits) operations per second

What about computing?

- Technical challenges
  - beyond our capability to currently build!
  - will challenge limits on compute performance, data storage, data transport, power consumption and software systems

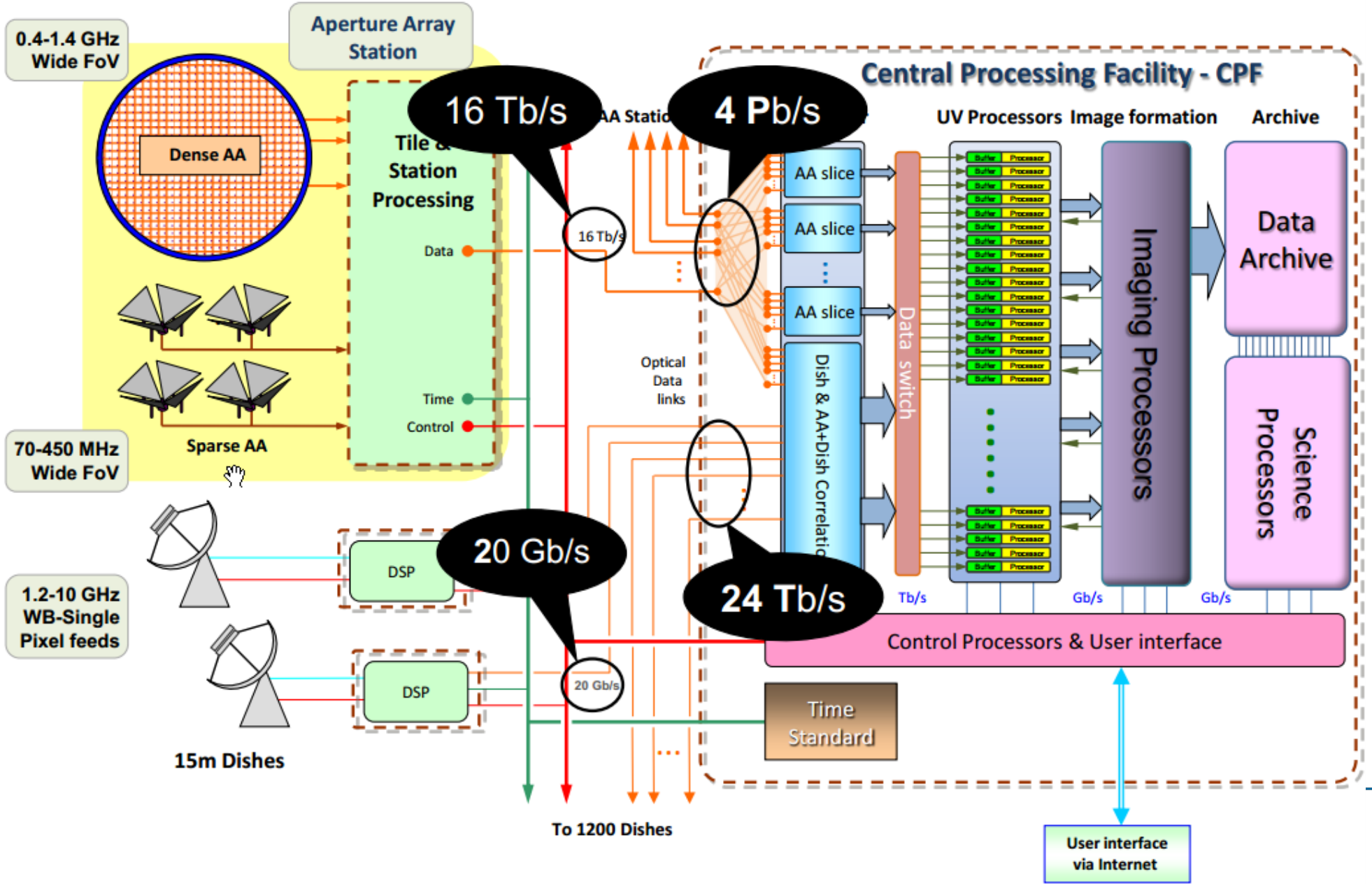
- Cisco (Brett Biddington – Space Initiatives Manager)
  - **“The Mother of All Networks”**

- Meeting these challenges will drive improvements across whole IT industry

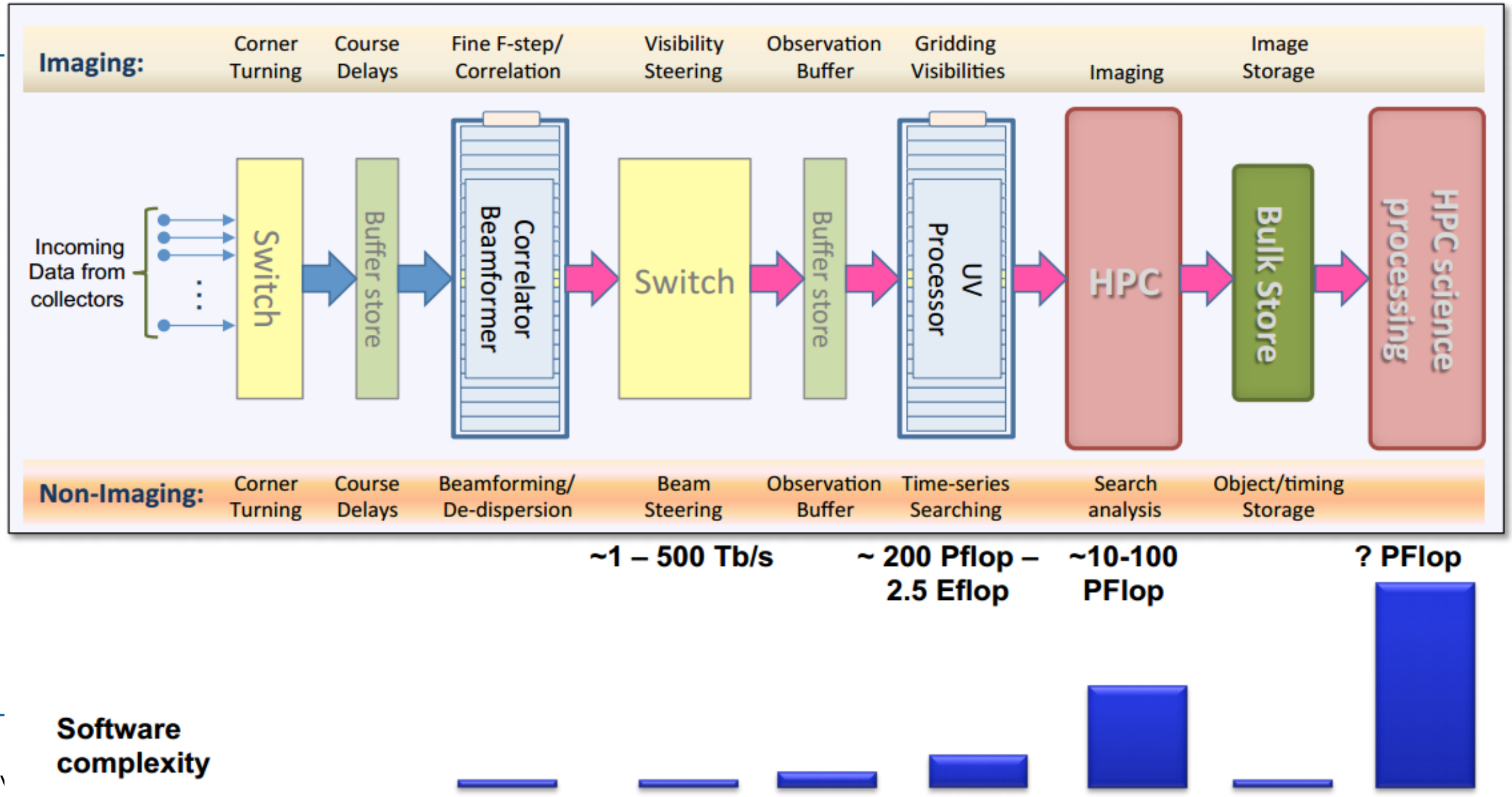
SKA FRANCE
SQUARE KILOMETRE ARRAY

SKA Data Flow

SKA2 wide area data flow



Processing pipeline



SKA raw data
14 Exabytes/day

Books 700M km Earth-Jupiter *per day*

280,000 years HDTV
1.4B DVDs
per day

Data Scale (log)

Exa 10^{18} 1'000'000'000'000'000'000

Books 50M km
1/3 way Earth – Sun
1B years reading

20,000 years HDTV, 100M DVDs

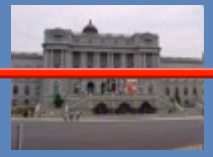
SKA stores
1 Pbyte/day

Peta 10^{15}

Books around equator
1B years reading

20 years HDTV
100,000 DVDs
per day

per day



Tera 10^{12}

Large Library
(50 Km of books)
1000 years reading

One week full HDTV

Giga 10^{9} 1'000'000'000

Small Library
50m of books
1 years reading

10 minutes full HDTV

Mega 10^{6} 1'000'000

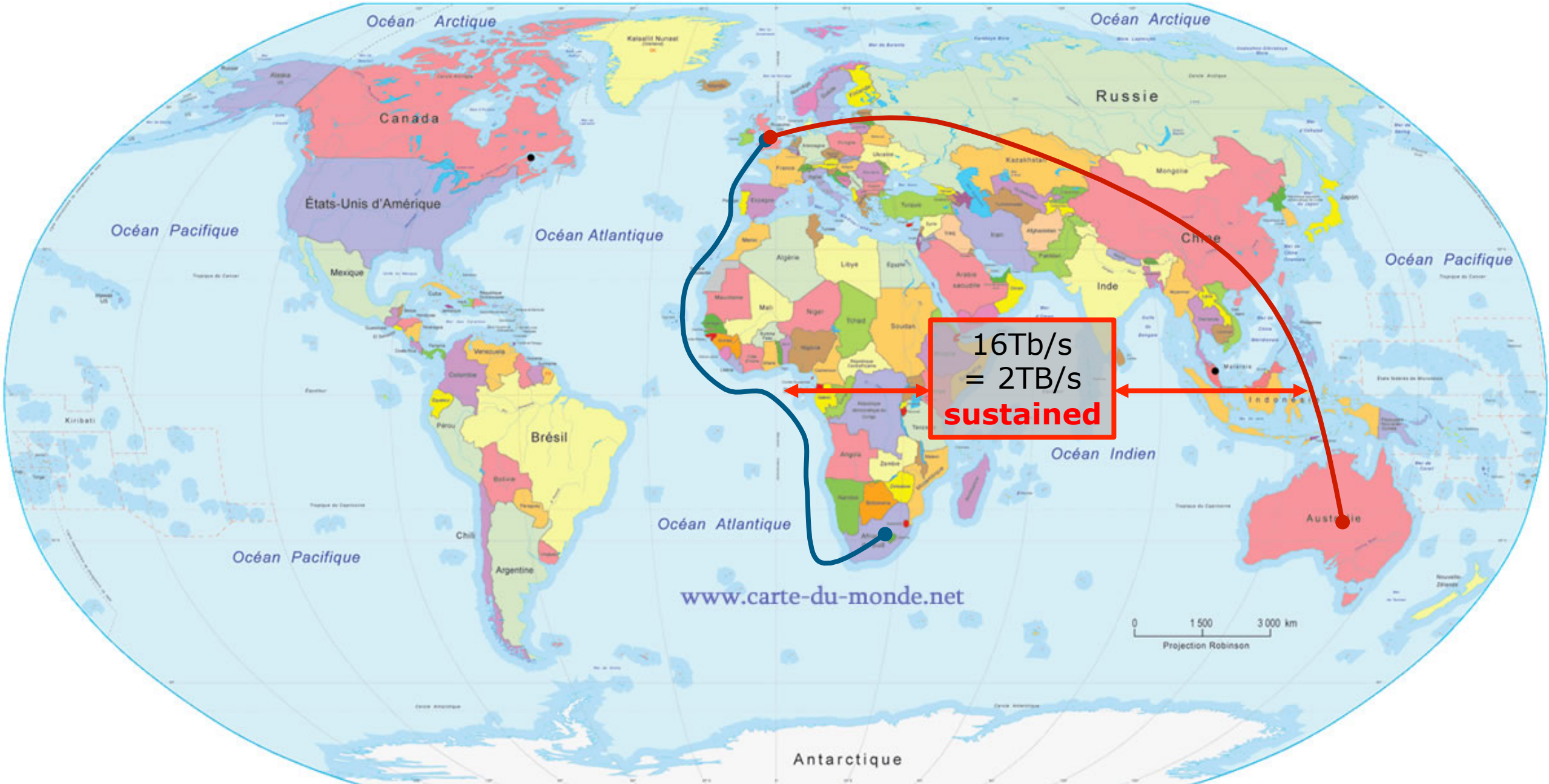
Decent Book (5cm)
8 hours of reading

Kilo 10^{3} 1'000

Half page

Adapted from "Towards an ASTRON IBM Center for Exascale Technology", Dr. Ton Engbersen (IBM Research Zurich), and Dr. Chris Broekema (ASTRON)

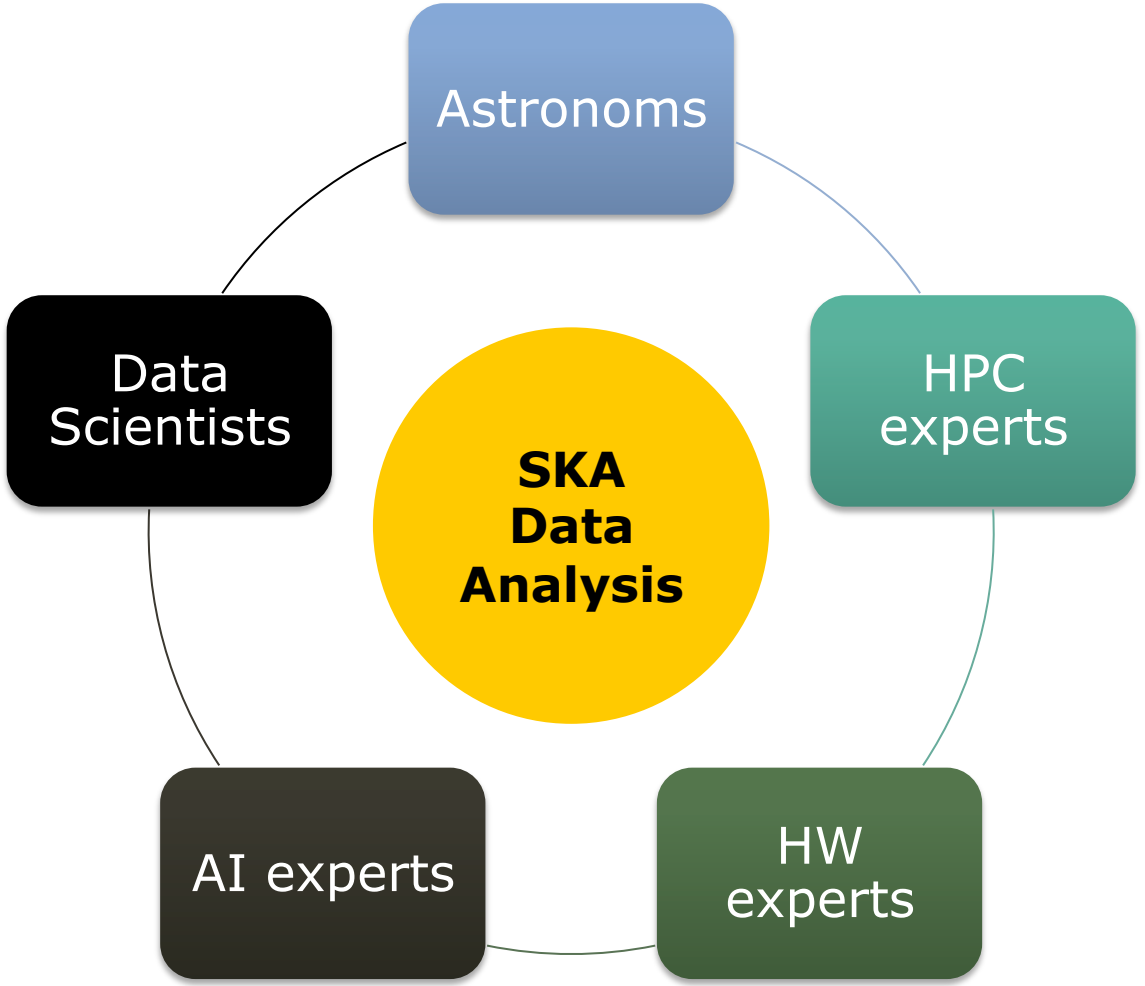
Data transfer



How to address the needs

Proposed computing systems

- Integrated HPC systems
  - Possibly useful for analysis
  - Not affordable for earlier stage problems
- Commodity x86 coupled with high performance networking (IB)
  - Also not possible to (affordably) achieve goals with this technology!
- GPU's
  - Becoming more mainstream
  - Non-standard programming
  - Questions over future gains
- FPGA's
  - Harder to program
  - Better power profile
  - Deeper pipeline
  - Better energy characteristics



Reduce Data volume
Co-design approach at any stage

700 people
400 HPC dedicated

1900
patents

6%
75M € EUR
50M € in HPC

Factory
In France

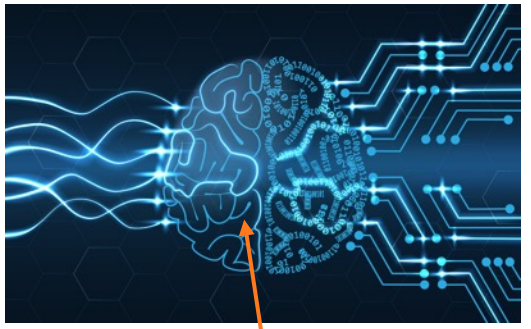
R&D centres
Paris, Grenoble
Sophia,
Bangalore

Infrastructure (mobull)
- Data Center design
- Mobile Data Center
- Water-Cooling

Expertise & Services
- Consulting & Architecture
- Cloud HPC (public & private)
- Big Data Analytics & management
- Artificial Intelligence & Deep Learning
- Applications & Performance
- 24/7 SLA
- On site services

Software (bullx SCS, Extreme Factory)
- Open, scalable, reliable SW
- Development Environment
- Linux, OpenMPI, Lustre, Slurm
- Administration & monitoring

Supercomputers (bullx)
- Full range development from ASICs to boards, blades, racks
- Direct Liquid Cooling, PUE <1.1
- SMP, Cluster & MPP systems




Thanks