AENEAS WP4
Network Design for SKA

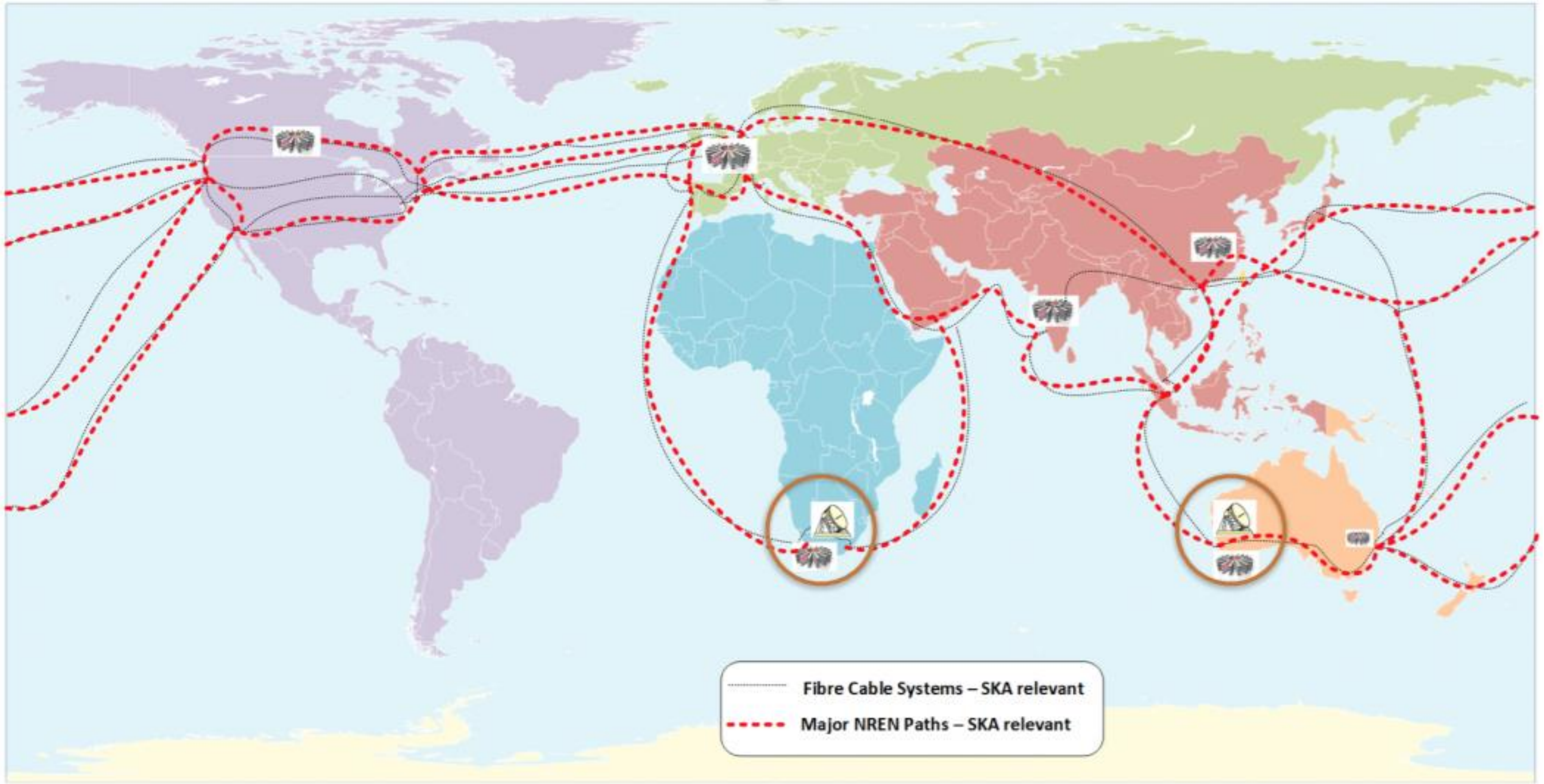
D4.3 “Architecture and cost model for the European ESRC network”

D4.4 “Architecture and cost model for a World-wide network for SKA”

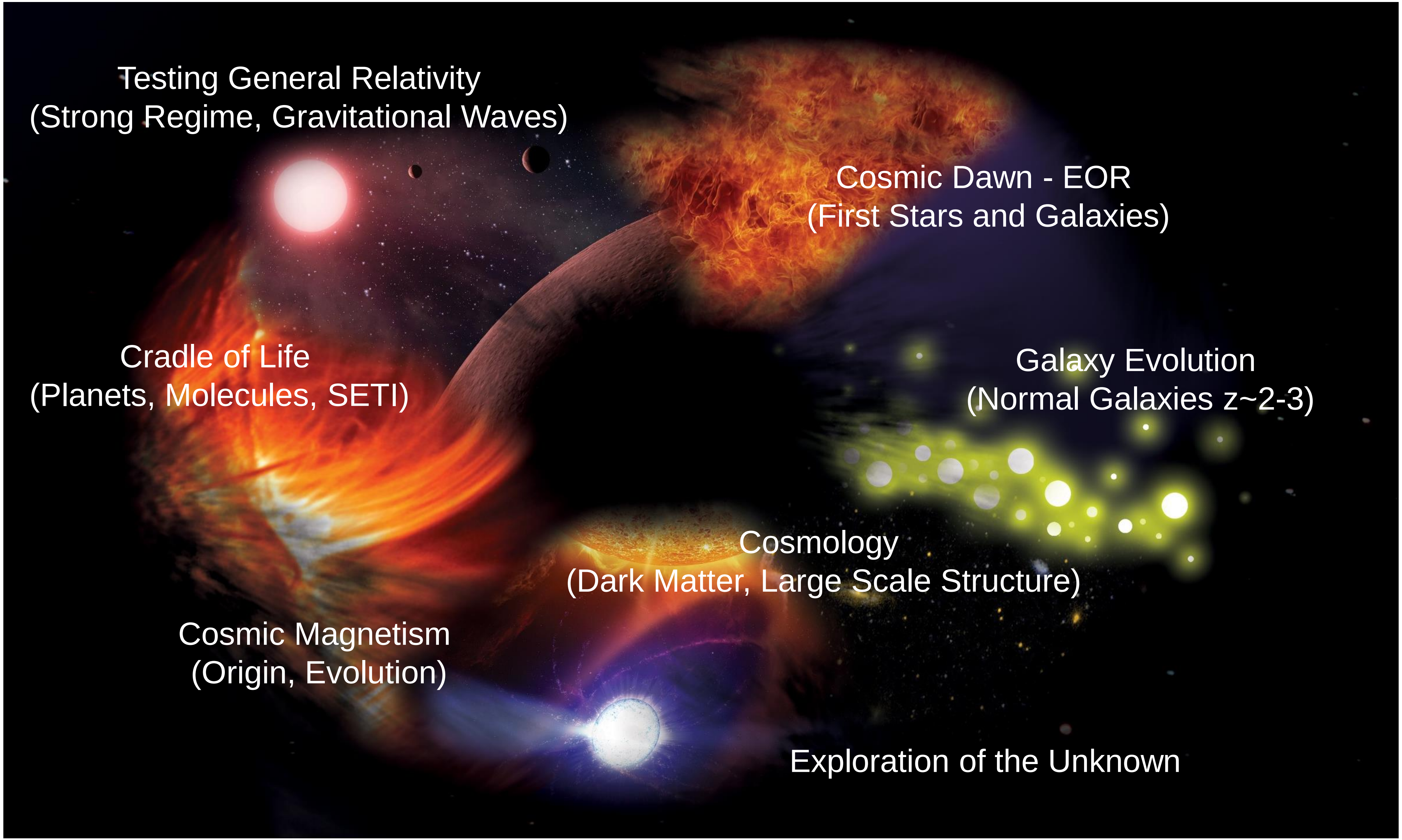
Richard Hughes-Jones
GÉANT Association

Data Flow in SKA

Observatory Data Products flow from Science Data Processors
in Perth and Cape Town to SKA Regional Centres around the globe



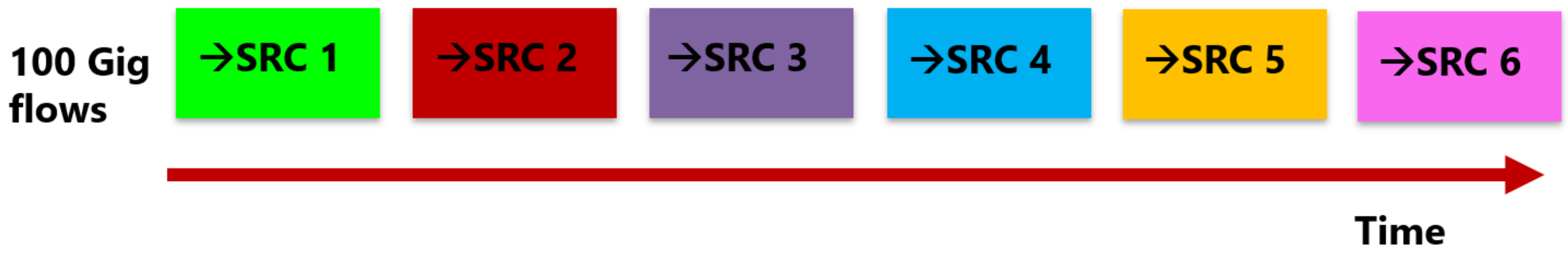
Key Science Drivers



The Compute Architecture for the Different Observatory Data Products

- Pulsars — Independent parallel – most like LHC
- Visibilities image data cubes — All image data for one sky section needs to be in one place – high bandwidth storage – memory – cpu power
- Radio Weak Gravitational Lensing — Uses gridded visibilities – needs high memory
- Epoch Of Re-ionisation — Power spectra – closely coupled supercomputer – cell cubes influence each other
- Cosmic Magnetism — Faraday rotation measure synthesis can be done as independent parallel computations
- Object detection / classification — Requires large numbers of GPUs – like HL-LHC

Models of Data Flows from the Telescopes to a SRC

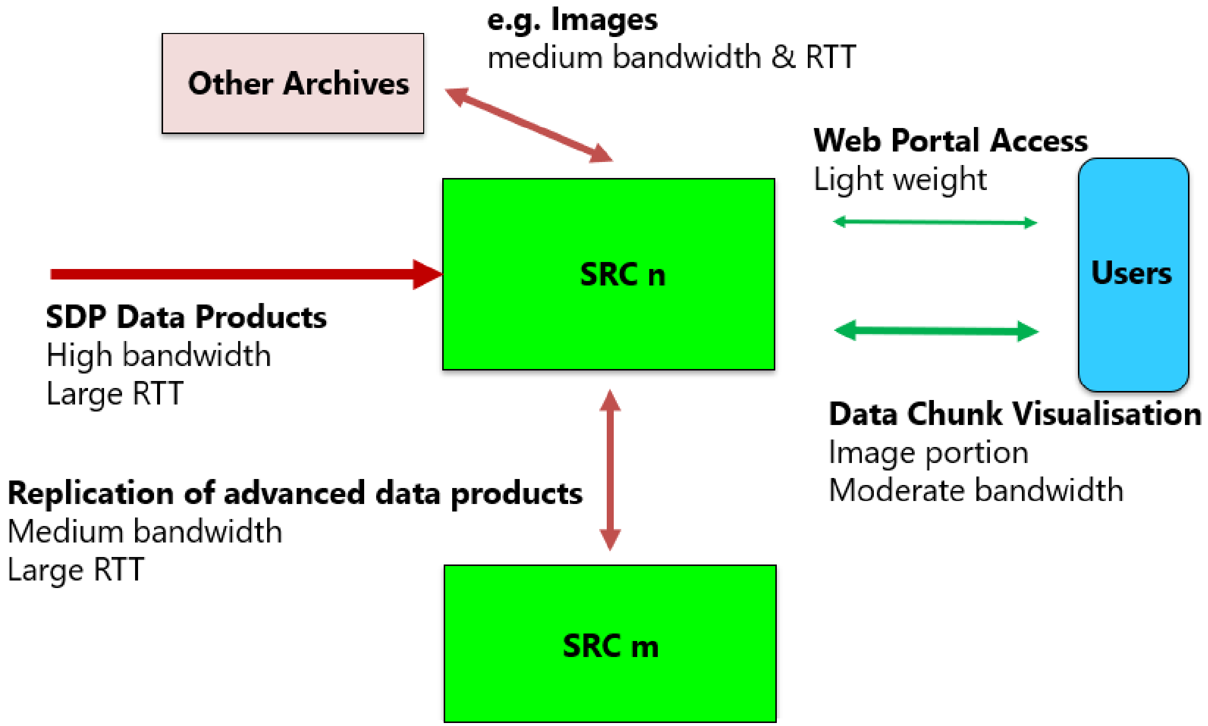


- ODP sent to each SRC in turn.
- Requires clean 100Gbit/s paths to each SRC.
- Worst case the SRC would require a 200 Gbits/s link.

- Several ODPs flow to different SRC taking place in parallel.
- Operationally more realistic.
- Makes efficient use of the network to each SRC.
- WP4 demonstrated stable 28 Gbit/s flows – TCP limit.

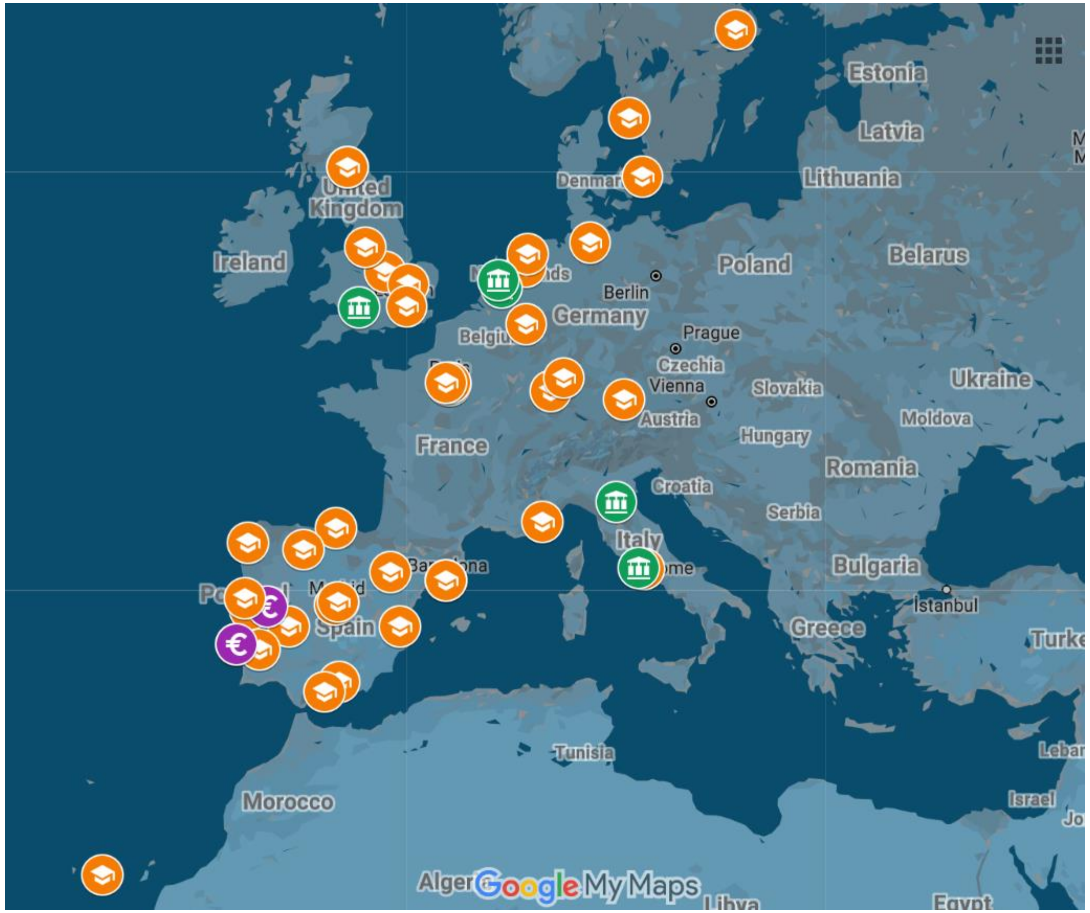
The SDP push model gives the advantage of scheduling the use of the bandwidth on the telescope access link .

Network Traffic into & out of a SKA Regional Centre



- Assumptions
- 20 Gbits/s at a time from each Telescope
  peak 40 Gbit/s to a site
- At a given time all of a Data Product goes to one site
- 10 Gigabits/s to other ESRC sites
- 10 Gigabits/s to non SKA data Archives
- 10 Gigabits/s to other RDC
- 70 Gigabit/s (peak) to each site.
- Flexible but secure ACL and high performance DMZ
- Say 100 Gigabit/s for SKA to each site in 2024/25.
  This is affordable.
  (c.f. the UK WLCG sites which have 10-40 Gbit/s dedicated links now)

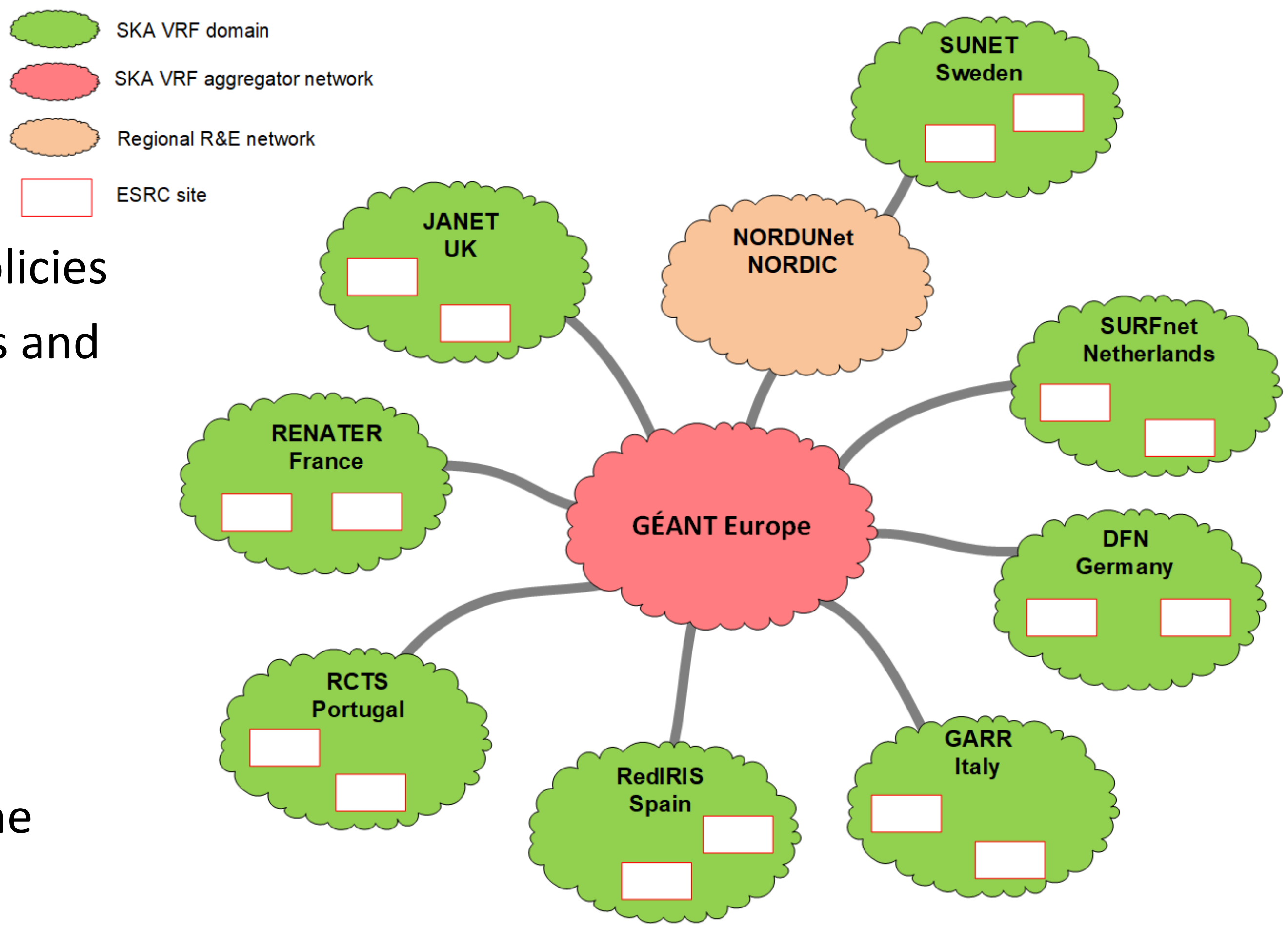
The European ESRC network



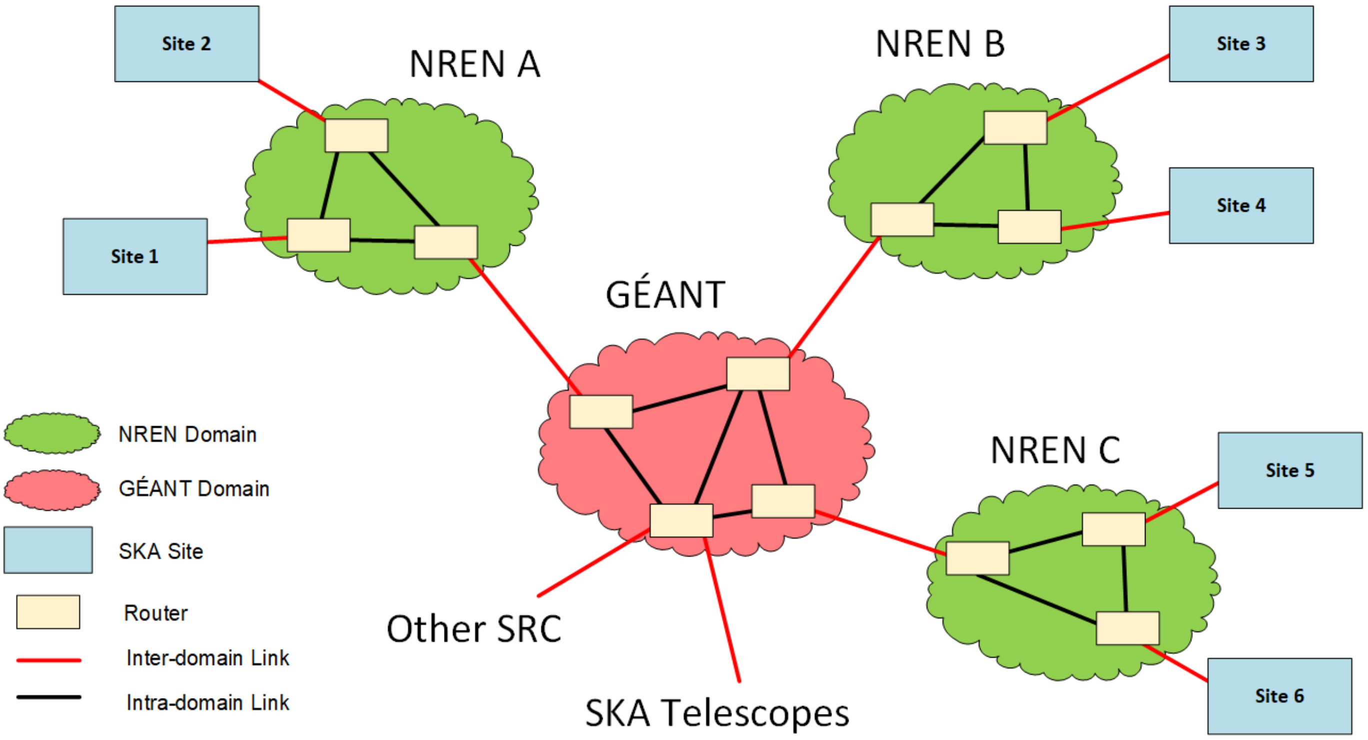
- Order of two sites per country
- Large sites like LCG Tier1s not Tier2s

Network Architecture for the ESRC

- VRF based overlay on the academic networks
- Isolation of SKA traffic from other users
- Easier for NRENs to implement the routing and policies
- SKA traffic can be engineered to use specific paths and routes to provide the high bandwidth
- Layer 3 routing provides isolation of any network configuration issues and strictly limits broadcast storms
- Layer 3 will re-route traffic as long as there is an alternative network path
- Configuration actions have to be undertaken by the NREN and a Site to join the SKA VRF, which provides an extra layer of security



Forming the VRF - Connecting ESRC Sites to the NREN



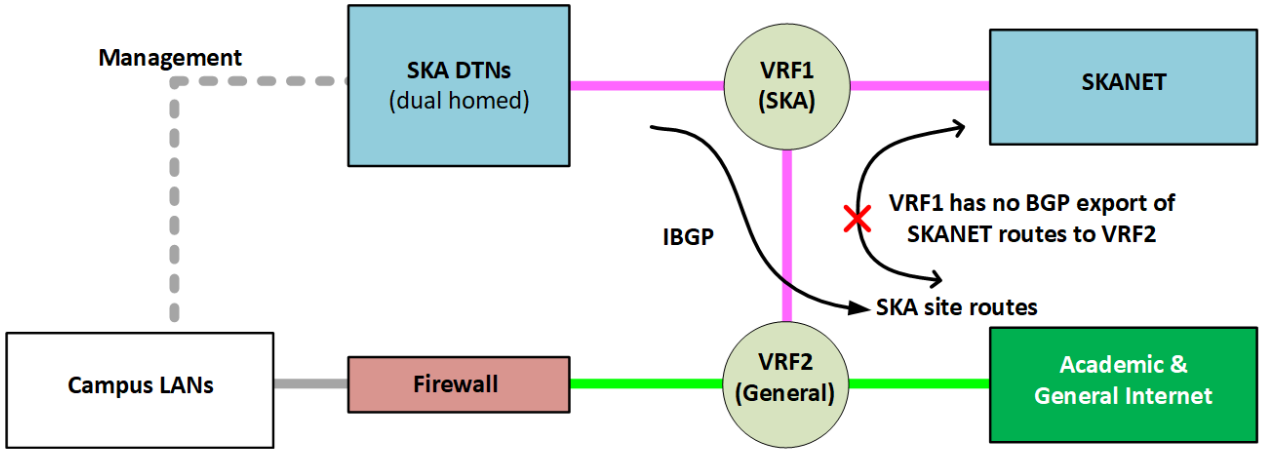
- Site routers connect to VRF in the NREN routers
- NREN backbone forms the VRF in that country.

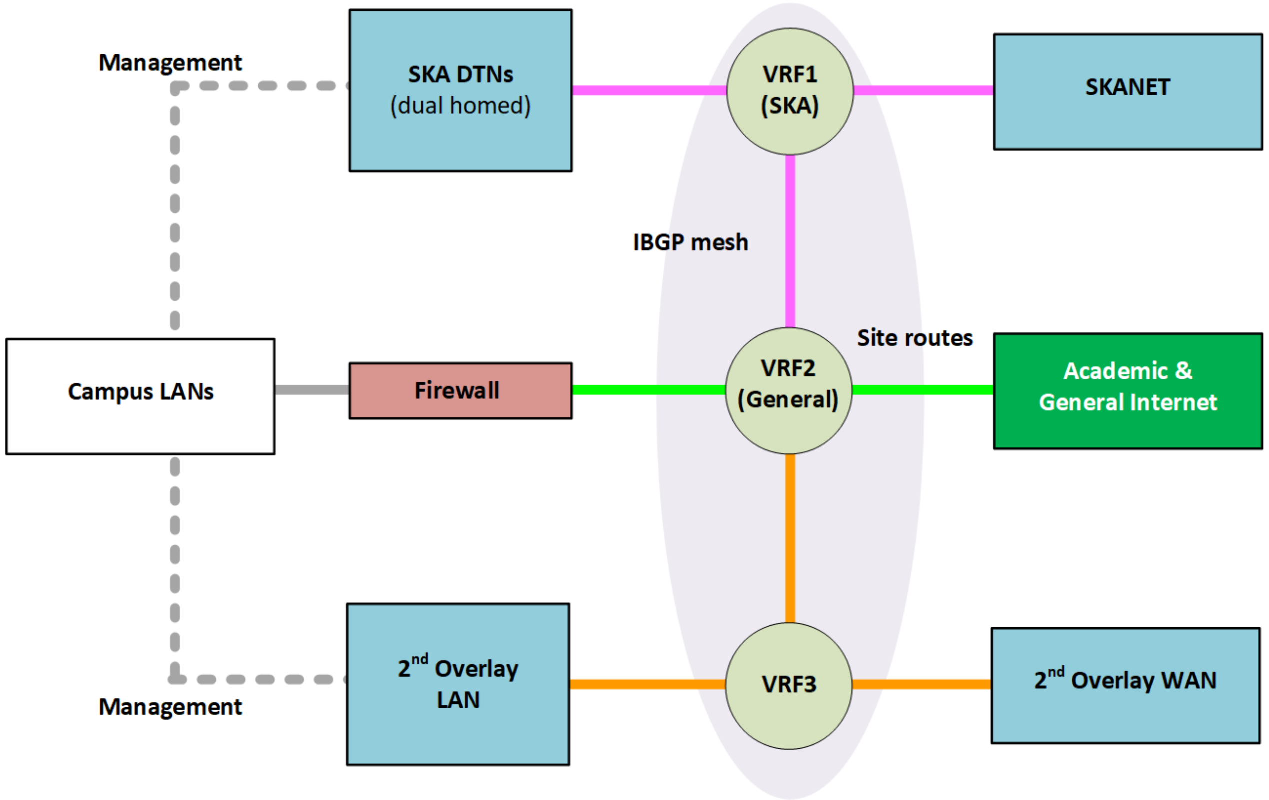
Operational aspects:

- Use NREN access link to peer with routers on the GÉANT backbone
- Sites will need to define and implement local Site Policy and Filtering Requirements.
- Project-wide Acceptable Usage Policy to be defined by SKAO & the whole SKA community.
- NRENs and GÉANT will need to implement an access policy based on the set of accepted ESRC site prefixes.

Network Considerations for SKANET at an ESRC Site

- Need for high performance Data Transport Node hardware
  - Tuned for RTT ~300 ms
  - Network – disk transfer rate ~20 Gbit/s
- Flexible but secure ACLs and high performance DMZ connected to the VRF
- Separate SKA data & Campus LAN traffic
- Uncongested site access link

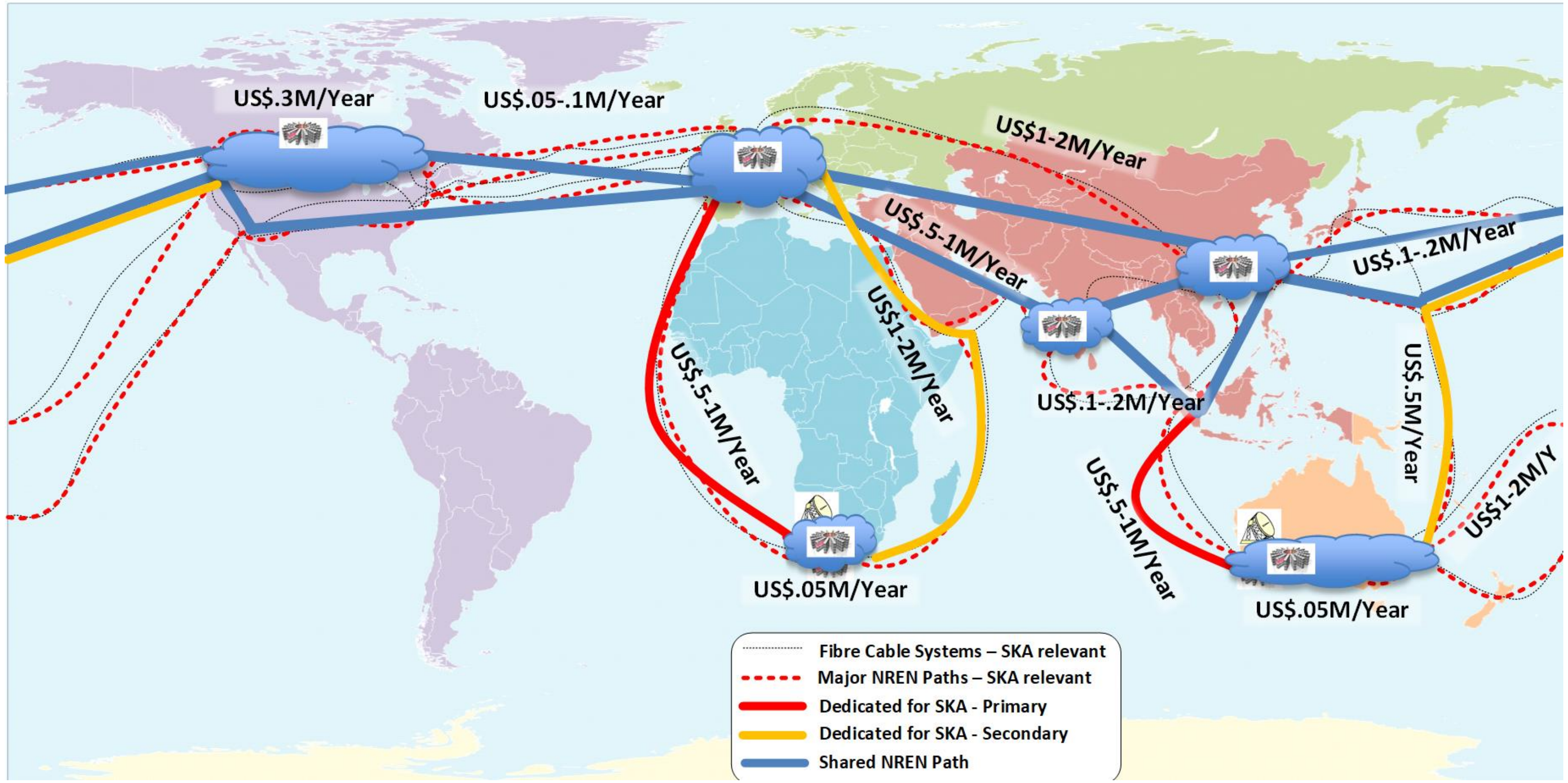




World-wide Network for SKA

Global Network & Paths of Interest to SKA

- Dedicated Primary (red lines) & Backup links (yellow lines) from both telescopes
- Use of the shared academic network (blue lines).
- 1 PetaByte/day pushed by SDP from each Telescope → 100 Gigabit/s
- Costs based on 10 to 15 year IRU per 100 Gbit circuit projected to 2020 prices



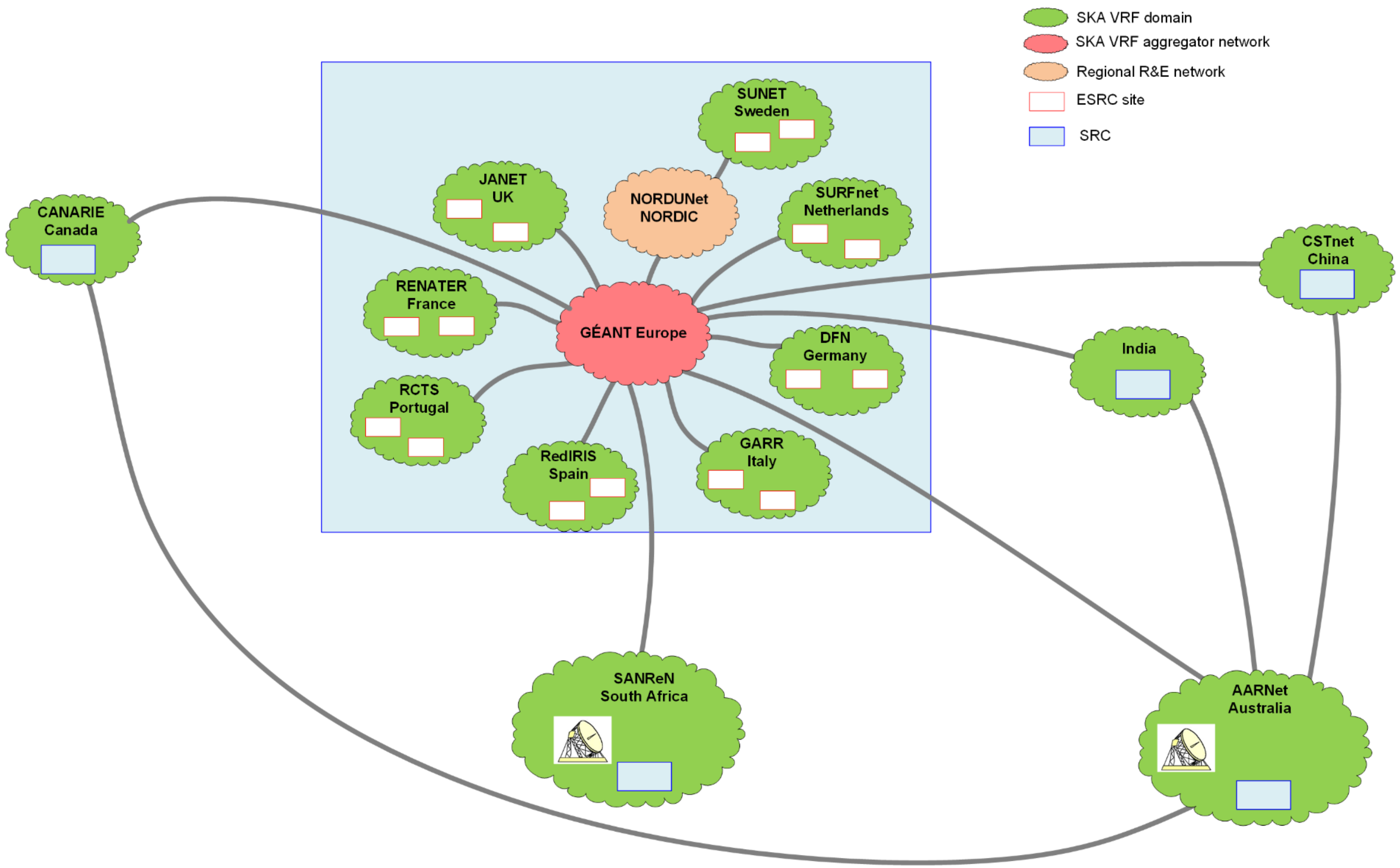
Budgetary OPex Costs

- **Primary links USD 2M per year**
- **Backup links USD 4M per year**

Now a new affordable path
Singapore – Europe direct

Global Network Architecture for SKA

- Global VRF based overlay with peering linked over the shared academic network

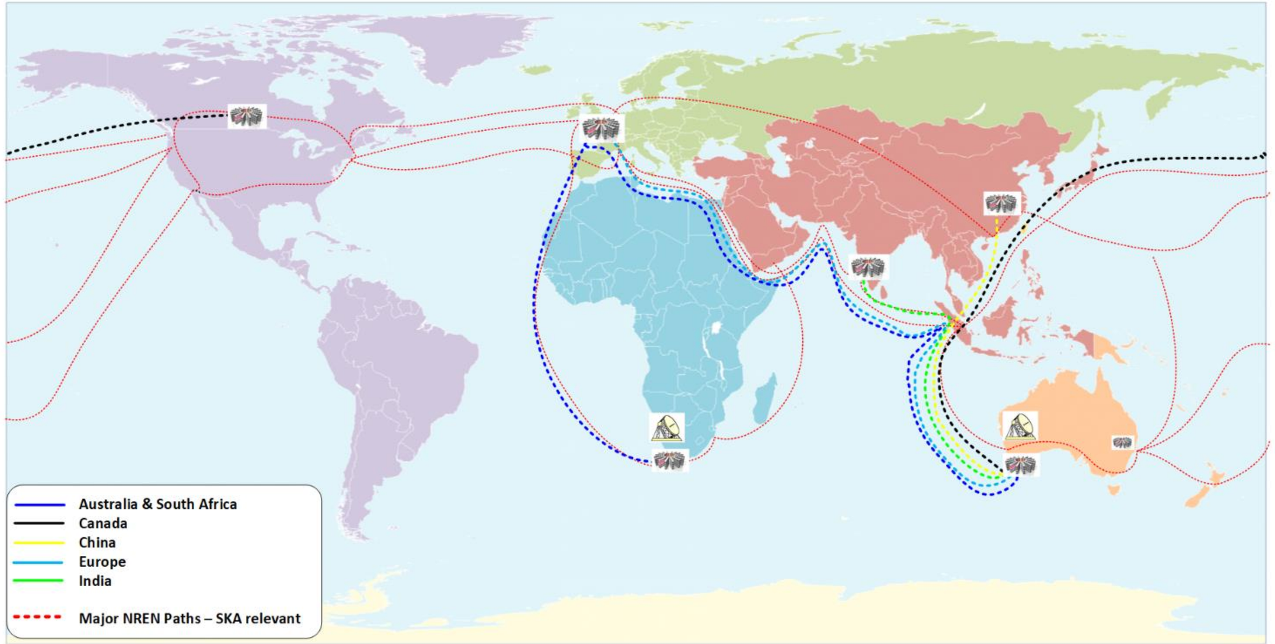


Global Paths of the Data Flows Pushed to the SRC – 1 Replica

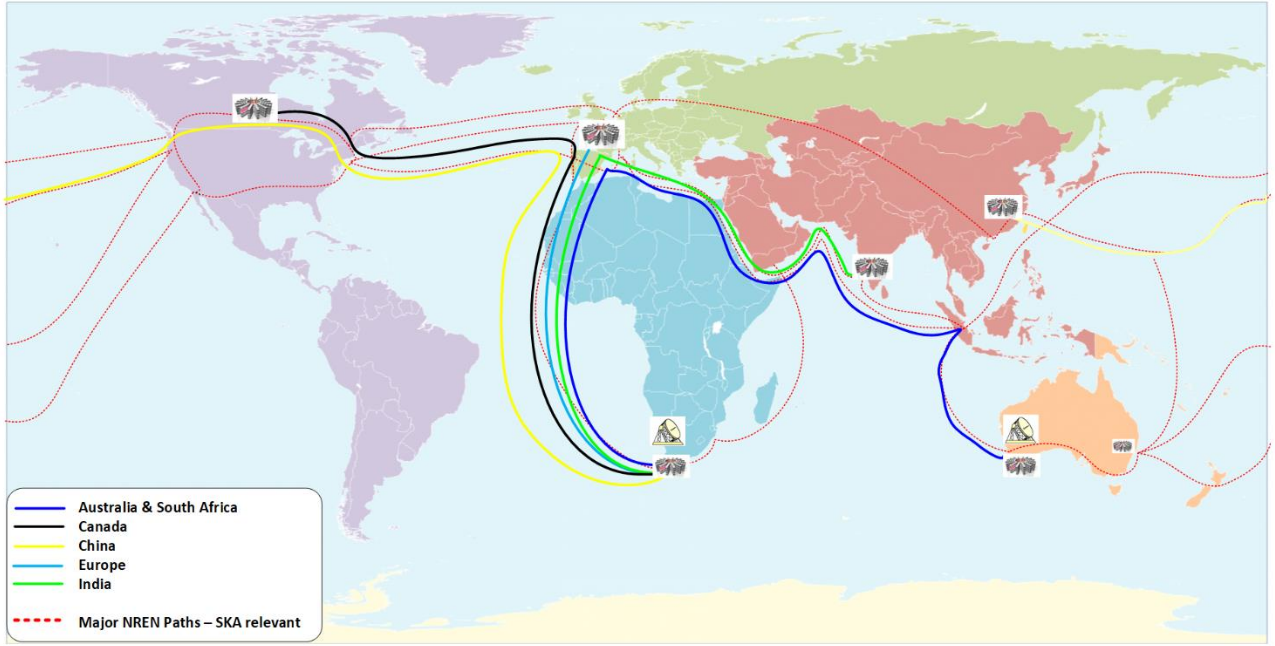
- Five flows on the submarine cable from Perth to Singapore .
- Then join the general purpose routed IP academic network.
- Single flows on the routes to Canada, China and India, Australia is local, and
  two 20 Gbit/s flows would be carried to London to reach SRCs in Europe and South Africa.

- Five flows on the submarine cable from Cape Town to London.
- Then join the general purpose routed IP academic network.
- Different submarine cables used to reach India and Australia, Europe is local, and
  two 20 Gbit/s flows cross the Atlantic to SRC in Canada and China.

SKA1-LOW Australia

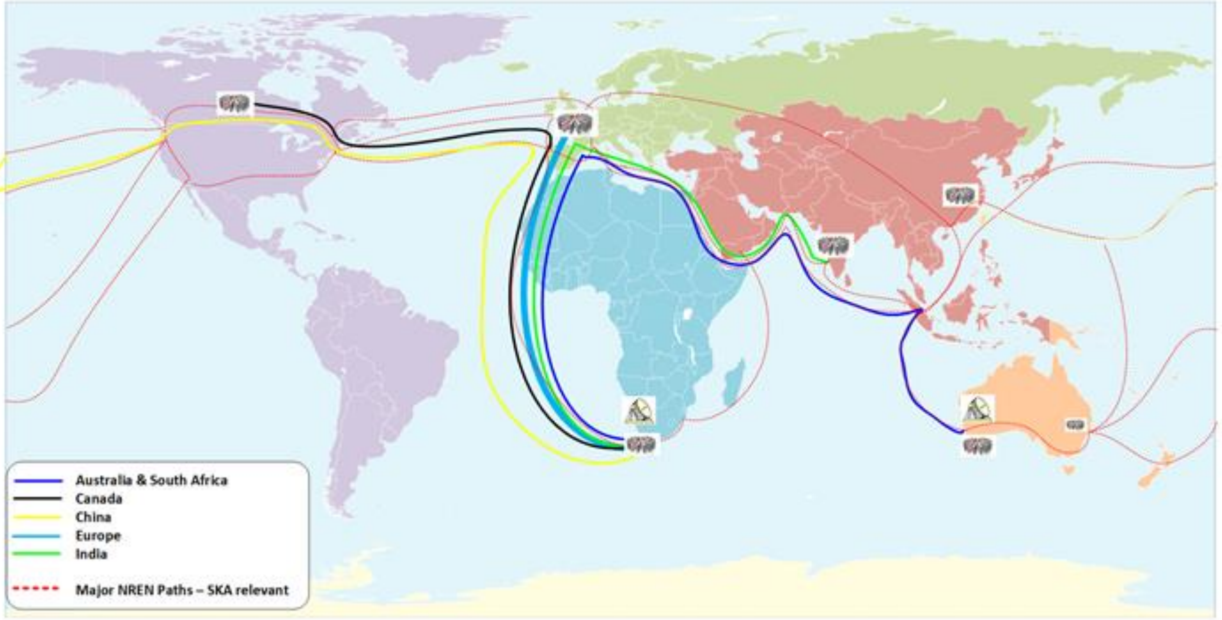


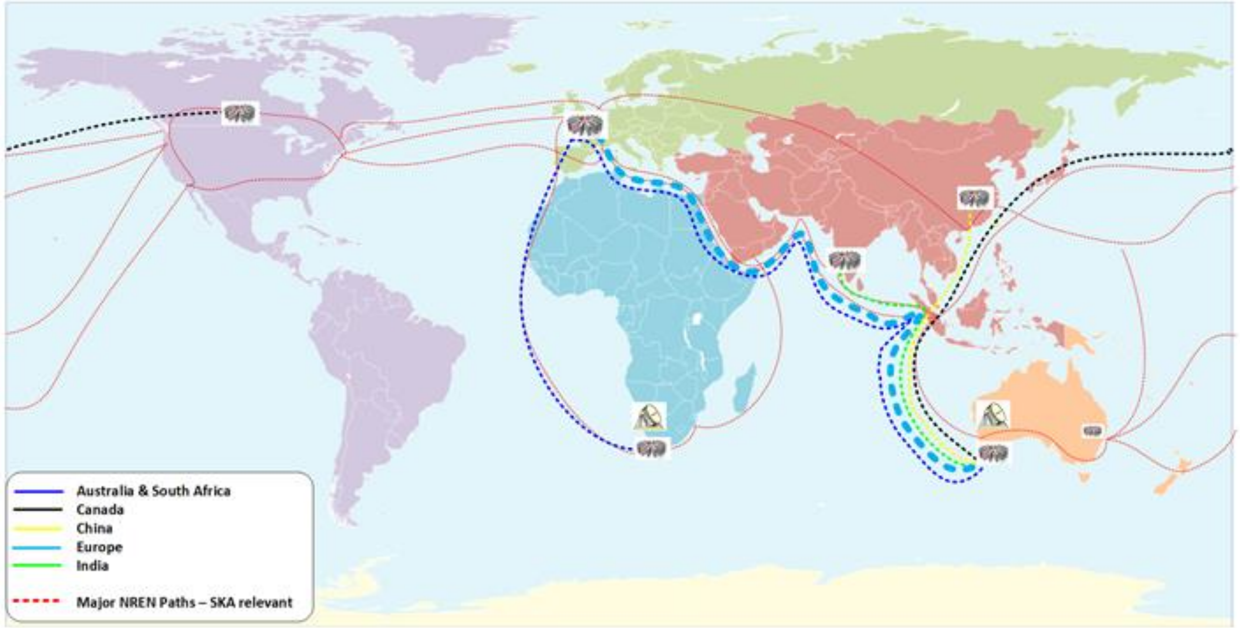
SKA1-MID South Africa



Global Data Flows if ESRC has a Full Copy of SKA Data

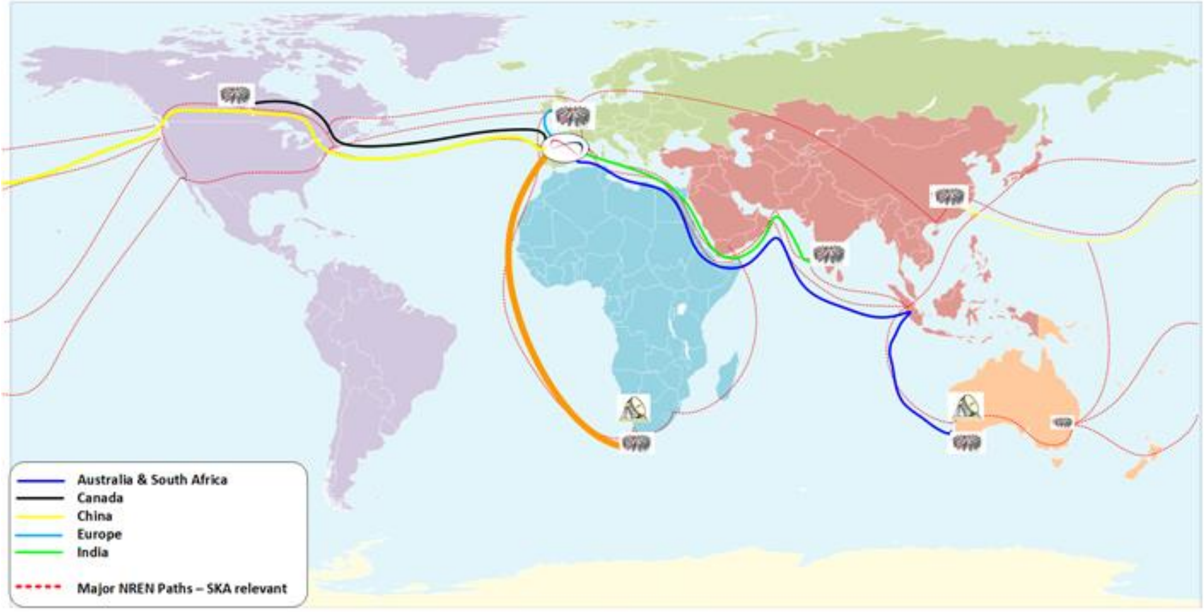
- One lambda carries the ~20 Gbit/s flows to the global SRCs like the 1 replica scenario
- Another 100 Gbit/s supplies the other Data Products to ESRC
- Europe funds extra intercontinental OPEX of USD 1.7M/year.
- Extra load on SDP buffer systems.
- Consideration to be given to the cost of provision of diverse backup paths.

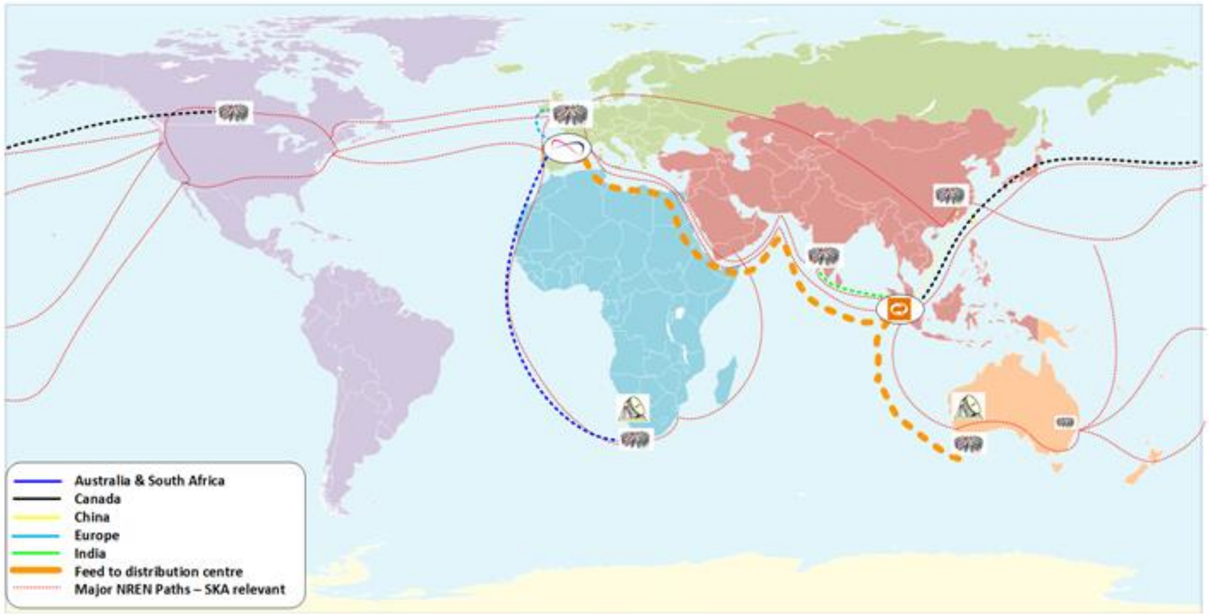




Distribution Centres Send Data to the SRCs

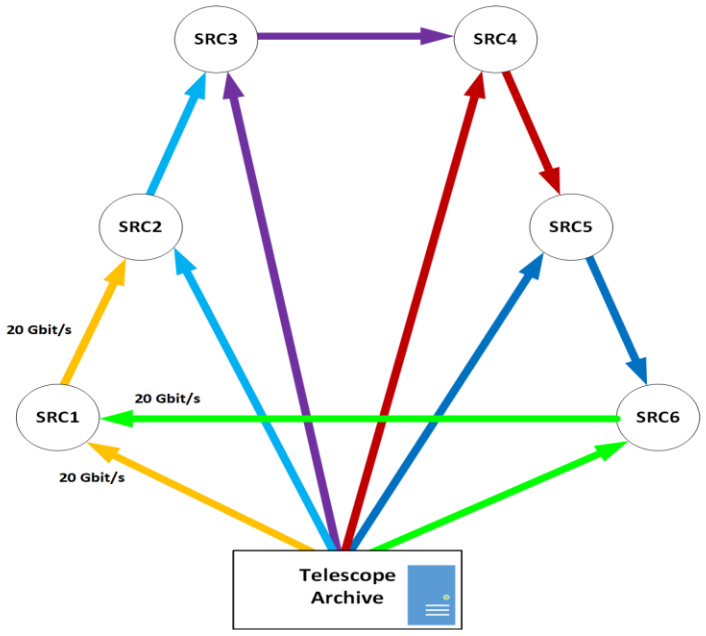
- Observatory Data Products sent to SKA Distribution Centres located at the remote ends of the expensive Submarine Cable links.
- These centres forward the data to the SRCs as required.
- London good location for SKA1-mid Singapore for SKA1-low
- Two days data from a telescope is ~2 PBytes of storage
- Need a medium size data centre, similar to a WLCG Tier 2 site
  - expensive 24/7 cost with power, staffing, etc. estimate ~ 0.3M Euro/year/centre

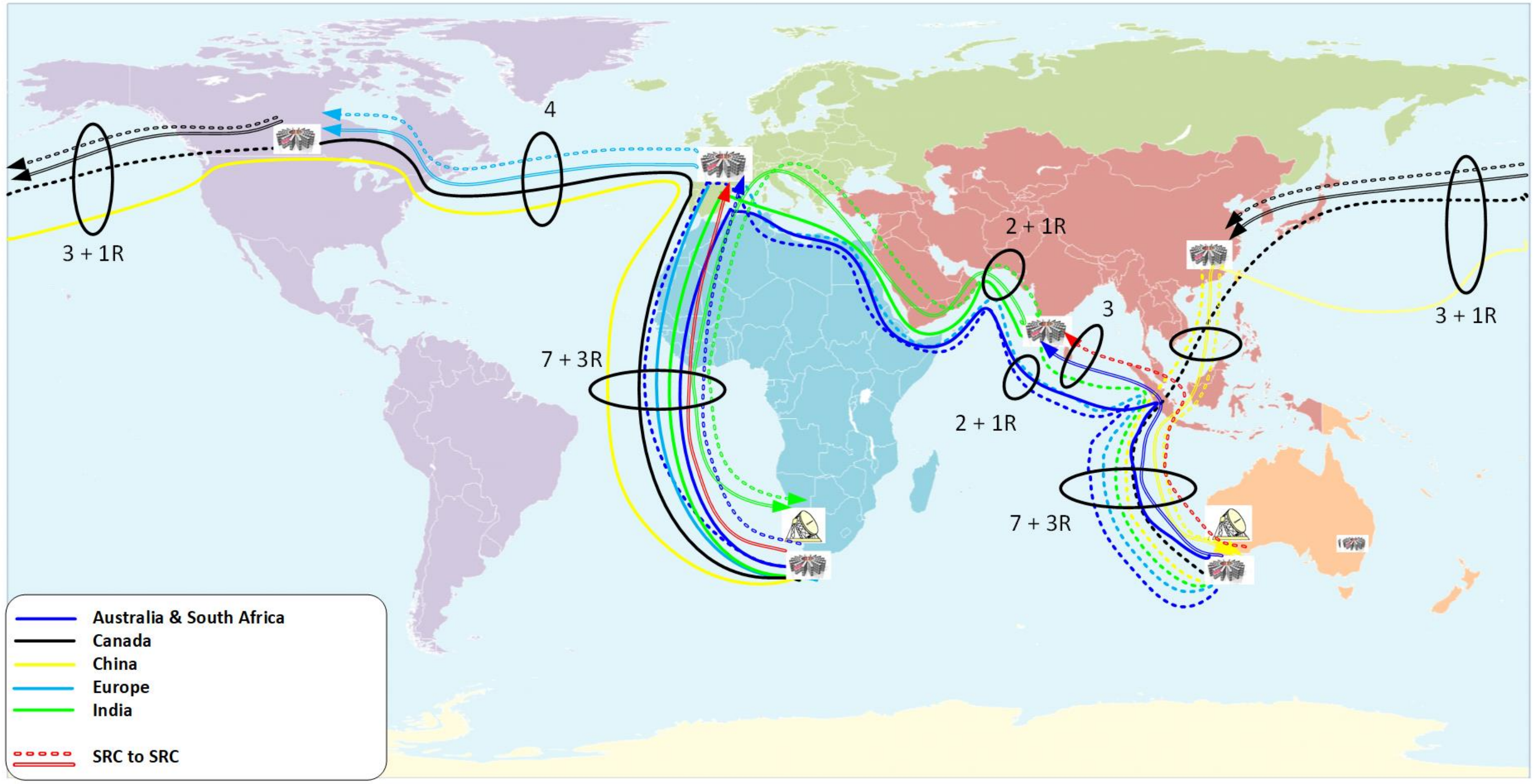




Global Data Flows if the SRC Re-distribute data – 2 Replicas

- Each SRC accepts its fraction of the Observatory Data Products and re-distributes to another SRC.
- SRC has 20 Gbit/s flow from the telescope & a second continuous 20 Gbit/s flow from another SRC.
- Each SRC sends out a 20 Gbit/s flow.
- Makes substantial use of the shared academic network which would imply charges to the SKA community.
- Cost to SKA community Very approx. ~ 0.8 M USD/year not allowing for the extra BW from the telescopes.





Summary

WP4 Conclusions, Costs and Recommendations

Richard Hughes-Jones

Recommendations: The Global SKA Network

- The global SKA network be constructed as an overlay network formed from a set of VRFs.
- These VRFs are connected with suitable inter-continental paths to enable the peering.
- That SKAO funds the use of dedicated 100 Gbit/s paths out of each telescope on the submarine cable routes to the Northern hemisphere.
  The budgetary OPex cost of these two primary paths is USD 2M per year projected to 2024 prices.
- These primary dedicated 100 Gbit/s paths out of each telescope be commissioned and in service to assist with:
  - The hardware and science commissioning of the telescopes.
  - The establishment of the SRCs and the SKA data challenges.
- The purchase of dedicated secondary 100 Gbit/s backup paths from each telescope be considered only when SKA is fully operational – probably about 2028.
- The SKA community gives careful consideration to the use of one replica of the Observatory Data Products external to the SDP archives.

Recommendations: ESRC Connectivity

- The SKA network in Europe be constructed as an overlay network formed from a set of VRFs each one covering a country and aggregated by a VRF on the GÉANT backbone.
- In discussion with the NRENs sites provision a suitable access link for SKA data
- ESRC sites and the Science Data Centres establish DMZs for SKA data transfers:
  - Ensuring a loss free end site network with no bottlenecks
  - Suitable access control policies for SKA data flows
  - Support for both high volume data transfers and public access to certain datasets
- The European SKA community make provision for funding the cost of the SKA VRF network estimated at
  CAPex of about €50k and an annual OPex of about €15k for each 100 Gbit/s path over the network

Recommendations: General Implementation & Operation

- That GÉANT & global NRENs work with the SKA community during the pre and implementation phase of the SRCs
  - Provide technical collaboration for SKA in the work of ESCAPE & DOMA
  - To assist in determining the sites, the bandwidth and other network requirements.
  - Coordinate requests for information and costs to the NRENs
  - Help with the Operations planning

- That GÉANT & global NRENs work with the SKAO / SRCSC to form a SKA-NREN forum.
  - The forum should cover technical, governance and operational aspects
  - Include European and global astronomy networking & infrastructure members.
  - The forum should enable the exchanges of roadmaps, requirements, and facilitate operational aspects.
  - Worked very well for VLBI and WLCG


8


4
3

Integrating Global Distributed Computing Sites (AENEAS) & SDP

- Need for state of the art middleware for:
  - Replica management.
  - Orchestration of the data file streaming.
  - Workflow management
  - E.g. Rucio & FTS

- The SDP design needs to be integrated into the distributed computing environment, e.g.
  - Read-write loads on hot/cold buffers or archive
  - Linking SDP & global file catalogues
  - Defining the policy interface

- Realistic flow demonstrations are an excellent infrastructure candidate for SKA Data Challenges