AENEAS WP4

Analysis of Global SKA Data Transport and Optimal European Storage Topologies

WP4 Tasks

- Task 4.1: Evaluation of existing data transfer protocols, storage sub-systems and applications

Partners: Chalmers (lead), GÉANT Ltd, Jülich, INAF, UMAN Stakeholders: CSIRO, SANReN, IT

- Task 4.2: Inventory of the storage and network capabilities of existing and planned European Facilities for SKA

Partners: INAF (lead), GÉANT Ltd Stakeholders: ASTRON, IT

- Task 4.3: Optimized design and cost model for a distributed ESDC data topology with world connectivity

Partners: GÉANT Ltd (lead), INAF, Chalmers, Jülich, UMAN Stakeholders: AARNet, CSIRO, SANReN, IT

- Task 4.4: Proof of Concept Activities supporting the design of data access and transport within Europe and from the Host countries to Europe

Partners: GÉANT Ltd (lead), Chalmers, UCAM, UMAN Stakeholders: AARNet, CSIRO, JIV-ERIC, SANReN, IT

Current Milestones & Deliverables

Milestone number	Milestone name	Related work package(s)	Due date (in month)	Means of verification	Notes
7	Storage sub-systems evaluation	WP4	8 Aug17	Technical note written	Tech note in progress
10	Joint Milestone (WP4) on data moving applications & tools	WP3 WP4	9 Sep 17	Technical note written	Doc describing The LHC Grid Environment
11	List of possible regional site locations	WP2 WP4	9 Sep 17	List of possible sites established	NREN inventory done Network & Storage questions done
19	Data transfer test South African site to European site	WP4	13 Jan 18	Technical note written	Note written Tests continue
20	Joint Milestone (WP4) on SKA Sci DMZ recommendations	WP3 WP4	14 Feb 18	Internal memo	3 sites contacted
21	Best practice recommendations Data moving applications, protocols and storage	WP3 WP4	14 Feb 18	D 4.1 written	D4.1 finished Mar 18 Best practice recommendations
22	Specification for SKA Science DMZ	WP3 WP4	14 Feb 18	Specification document written	Work in prog

Joint Milestone
Related WP
WP4 Lead
Deliverable

- Deliverable D4.1 Complete

  Best practice recommendations Data moving applications, protocols and storage

Highlights 1:

MS 20 & MS 22 DMZ Pilot of Network and Storage Questions from MS11

Infrastructure Questions

Network Connectivity and Wide-area-network-facing Devices

1. What is the bandwidth of the access link between the site and the NREN in Gigabits/s?
   a. 1G
   b. multiple 1G
   c. 10 G
   d. multiple 10 G
   e. 40G
   f. **100G (2018)**
   g. Other (Please state)

2. By 2020 / 2022 what will be the expected bandwidth of the access link between the site and the NREN in Gigabits/s?
   a. 10 G
   b. multiple 10 G
   c. 40G
   d. 100G
   e. **current bandwidth to the NREN is limited by demand (it is currently sufficient to meet requirements). In case of increasing demand, we have capacity to increase our uplink by multiples of 100 Gbit/s.**

3. Does the site run a local storage system?
   a. **Yes**
   b. No

4. Does the site currently have designated servers for data transfers (i.e. Data Transfer Nodes (DTN))?
   a. **Yes**
   b. No

Storage Units

14. Please indicate the features of the storage system in your data center that offers the best performance e.g. The Brand and Model and year of purchase
    We have multiple storage systems at SURFsara, with different systems offering best performance based on different performance indicators (throughput, IOPS, network, read, write, combined). It is not clear what defines a "system" here, whether it is a node or a service as a whole.

15. How many replicas do you maintain of raw data?
    Depends on storage system. Typically 1 to 3 copies, for some storage systems the number of replicas can be configured based on user requirements.

16. Do you use a replica manager?
    (not clear what this means)

17. What is the total capacity in TBytes?
    (unclear what to answer, different capacities per system, most systems are expanded on user demands, somewere between 25TB and 25PB)

18. What type of disks (SATA, SAS, SSD)?
    a. SATA
    b. SAS
    c. SSD
    d. **All types available.**

19. Number of disks in the storage Unit
    a. 4
    b. 8
    c. 12
    d. 24
    e. 48
    f. 96
    g. **Different configurations applicable.**

20. What type of RAID is used?
    a. Raid0
    b. Raid 10
    c. Raid 5
    d. Raid 6
    e. **Different configurations applicable.**

Network or Distributed File System

23. Please indicate the features of the network file system in the data center that offers the best performance.
    Depends on criteria for "best performance" sequential single- or multi-threaded, parallel scale-up figures, read or write, IOPS, etcetera.

24. What is the number of nodes of the Network File System?
    a. 2
    b. 4
    c. 8
    d. 16
    e. 32
    f. 64
    g. 128
    h. **Differs per system, but usually sufficient to meet user demands.**

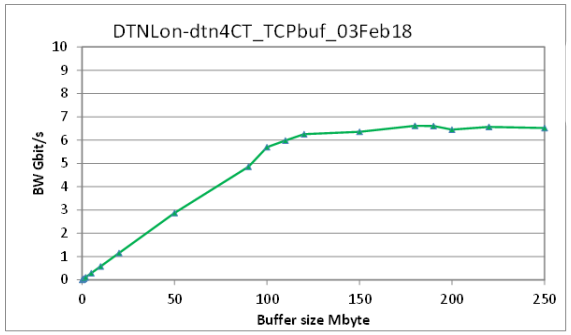
25. What is the type of file system?
    a. AFS
    b. GPFS
    c. Lustre
    d. Ceph
    e. HDFS
    f. BeeGFS
    g. GlusterFS
    h. Infinit
    i. **Differs per system**

26. What type of network is used between the File System nodes?
    a. Eth 1G
    b. Eth 10 G
    c. Eth 40
    d. Eth 100 G
    e. Infiniband 20 G
    f. Infiniband 40- 56G
    g. Fiber Channel
    h. **Differs per system**

Highlights 2:

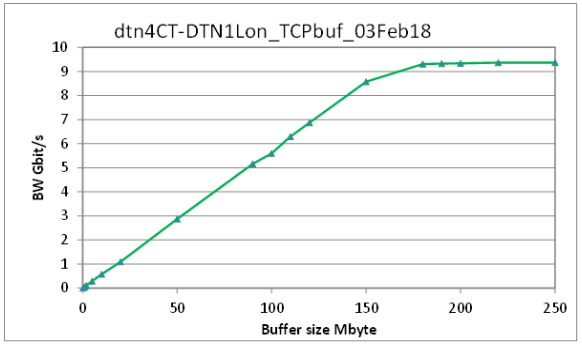
MS 19 Network Performance Between Europe and South Africa

London to Cape Town

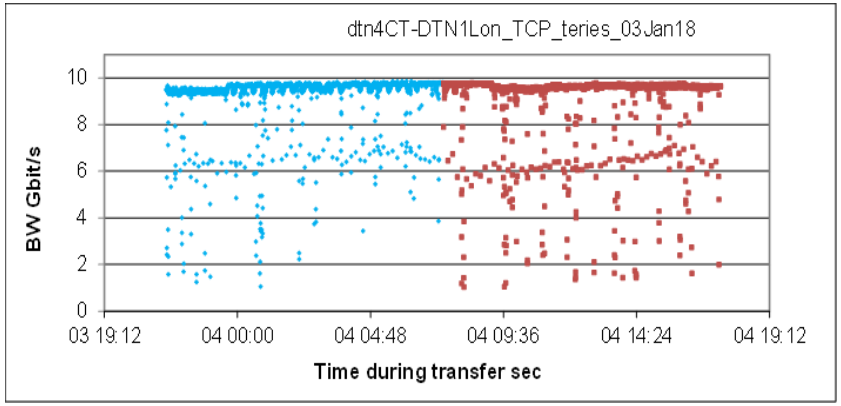


RTT 142 ms
BDP 178 MBytes for 10 Gig

Cape Town to London



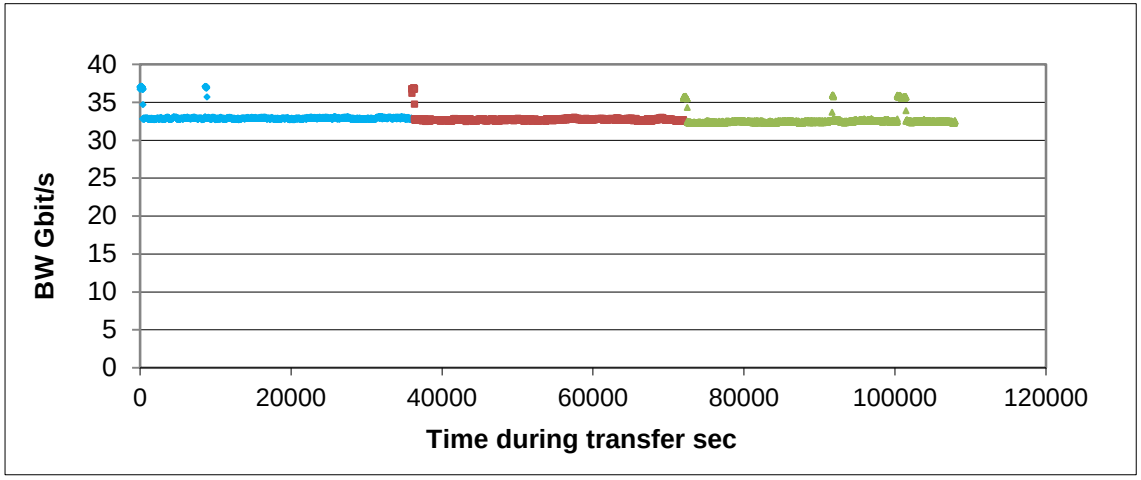
Cape Town to London



Production routed IP path
Achievable TCP throughput over 20 Hrs
Peak 9.5 Gbit/s

Highlights 3:

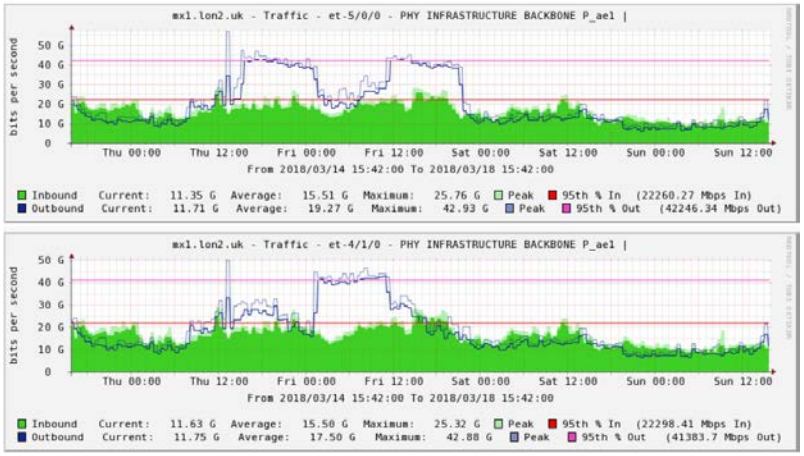
MS 19 TCP Performance Over GÉANT - 30 Gbit/s Single Flow



London to Paris

RTT 7.5 ms
BDP 34.7 MBytes for 35 Gig

TCP throughput over 30 Hrs
32.5 Gbit/s
No TCP segment re-transmissions



Milestones in the next 6 Months

Milestone number	Milestone name	Related work package(s)	Due date (in month)	Means of verification	Notes
25	radio astronomy data over global routes from South Africa to Europe	WP3 WP4	18 Jun 18	WP3 Technical note written	Continuation of MS19
					D4.2 Site Catalogue
27	Joint Milestone (WP4) on demonstration of moving data from observatory sites (SA) to ESDC	WP3 WP4	19 Jul 18	Demonstration completed	
30	Joint Milestone (WP4) on data replica manager	WP3 WP4	21 Sep 18	Internal memo	
31	Specifications for SKA Replica Manager	WP3 WP4	21 Sep 18	Specification document written	
33	Joint Milestone (WP4) on demonstration of moving data from observatory sites (AUS) to ESDC	WP3 WP4	24 Dec 18	Demonstration completed	
35	Data transfer test Australia site to European site	WP4	27 Mar 19	Technical note written	

Joint Milestone
Related WP
WP4 Lead
Deliverable

Deliverable D 4.2

Site Catalogue of the Storage and Network Capabilities

- Boundary Conditions
  - Due in June 2018
  - The questionnaire to potential supplier sites has not gone out yet
  - AENEAS will have a EC review in September

- The Network and Storage Capability questions have been written
  - Use Qaire as a guide for structured interview –telcon/skype
- WP4 will create a minimal draft for D4.2
  - Based on the sites contacted for the DMZ pilot
- The document will be private to AENEAS

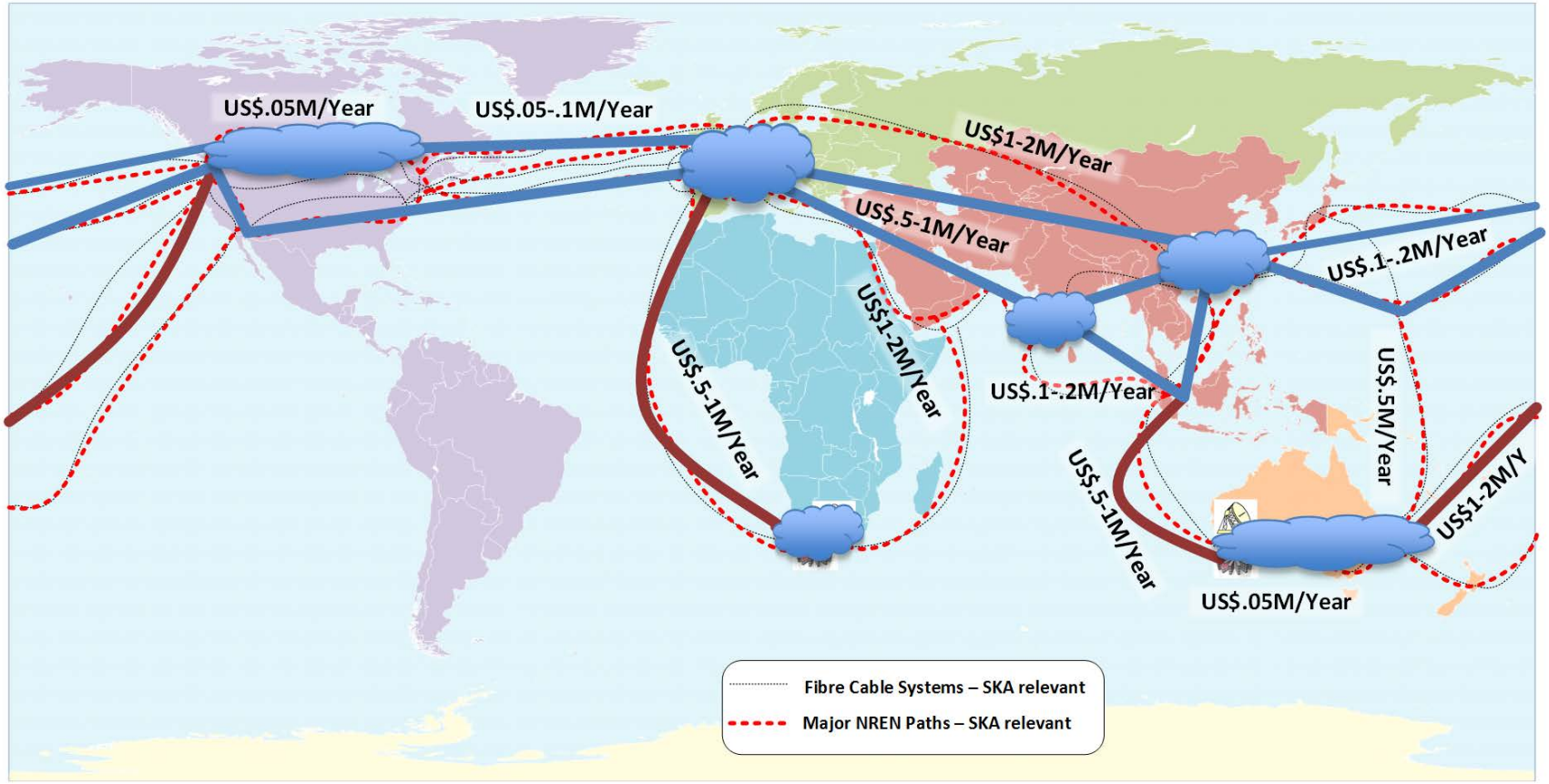
Opportunities and Challenges

- Deliverable D4.3 – Mar 19
  - Architecture and cost model for European ESDN network.
- Deliverable D4.4 – Aug 19
  - Architecture and cost model for World-wide network for SKA .

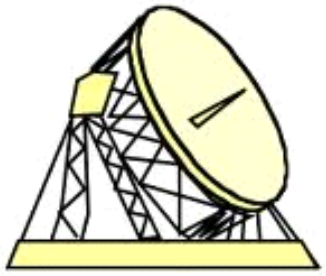
- The Data Placement and Compute models lead to the Network Requirements
- NREN Funding
  - Funding agency and strategy different in the European countries
  - Cost of an access link general depends on the site location

Task 4.3: Global Overlay & Dedicated links

- Not a commitment just some thoughts
- L3 VPNs linked over the academic network
- Dedicated links from both telescopes
- 1 PetaByte/day pushed by SDP from each Telescope → 100 Gigabit/s



Global Network Flows to a SRC



SKA1-Low
SKA1-Mid

SRC 1

SRC 2

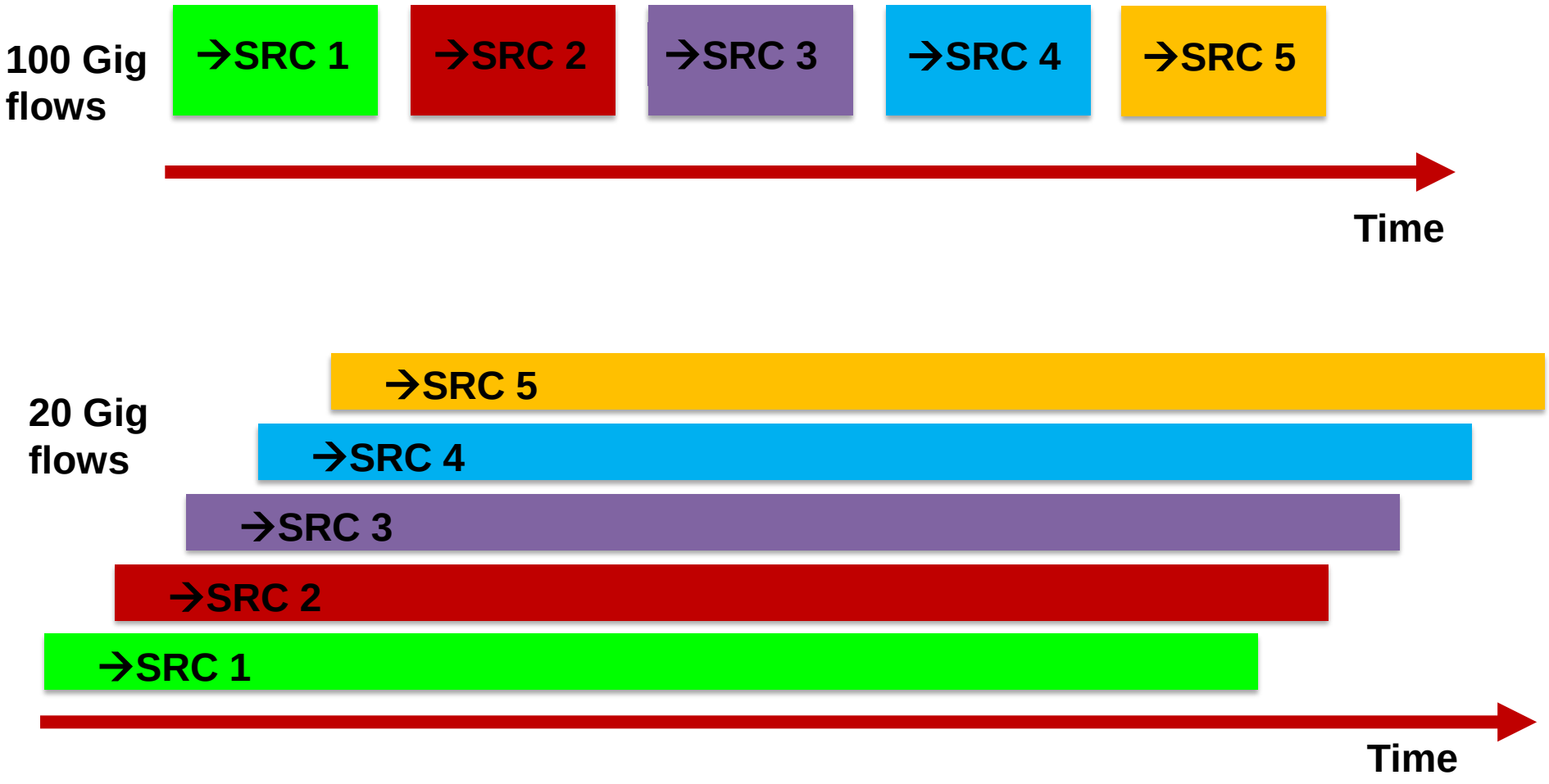
SRC 3

SRC 4

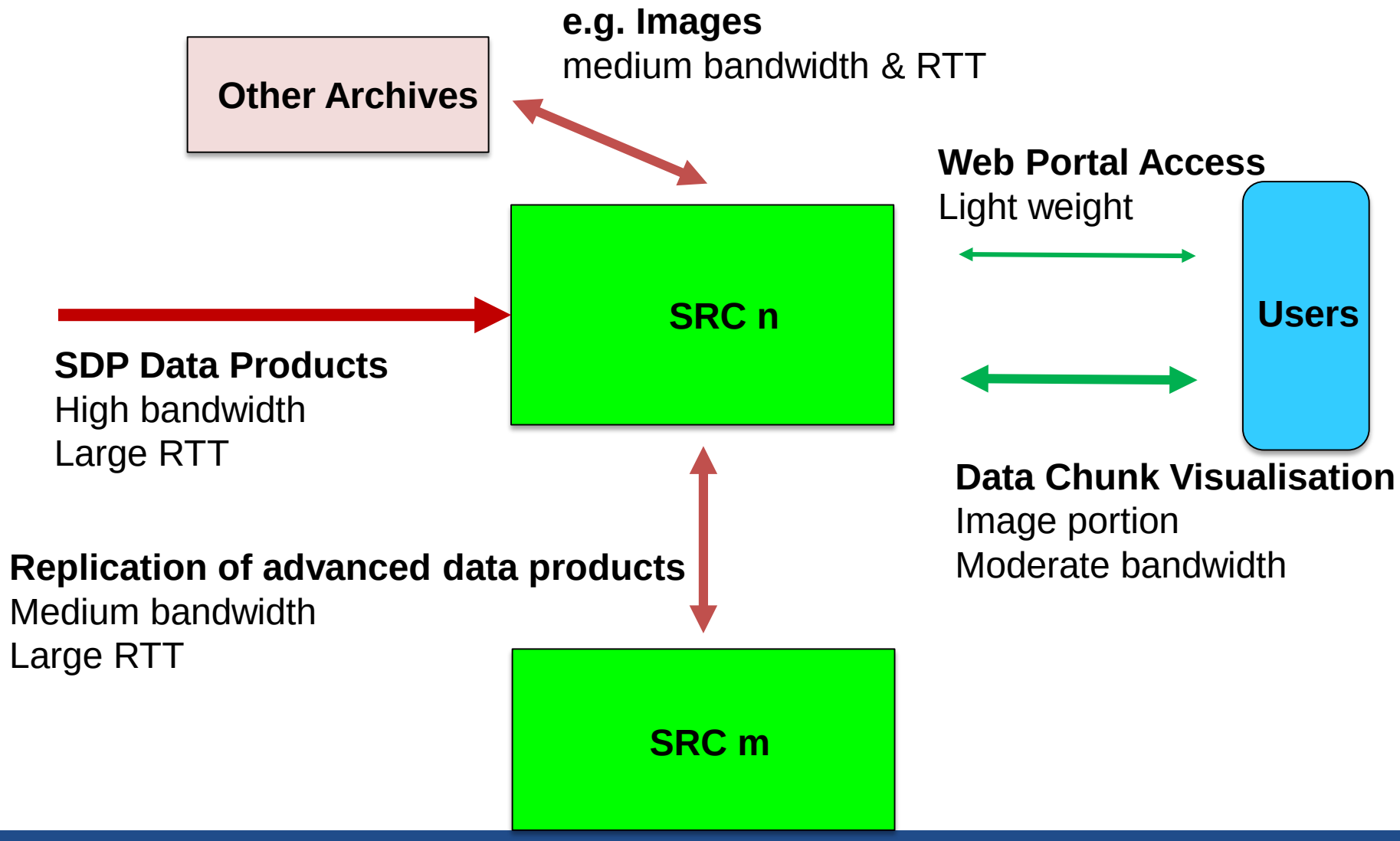
SRC 5

One copy at SDP One replica at a SRC

Task 4.3: Models of Data Flows to a SRC



Task 4.3: Network Connectivity to the ESDC



A Network Requirements Forward Look

Year	SKA Activity	Precursor Data	AENEAS PoC Network transfers Data moving tools	AENEAS ESDC PoC and operations	SKA commissioning and early science	SKA Data Challenge	Data from SKA Telescopes
2017		64 MeerKAT dishes			-	-	-
2018	CDR complete	1 – 10G MeerKAT → ASTRON 1 – 10G MWA → ASTRON ASKAP 10G?	10G SA → Lon, JBO 100G Aus → Lon, JBO 100G JBO, Lon, Par	1-10G Euo sites SA data → Europe	-	-	-
2019	IGO operation Approve / start c						
2020	C0 30 March						
2021	MID 1 dish						
2022	MID 8 dish Low 4 stations						
2023	MID 32 dish SDP AA2 Mid AIV AA1 Low 18 stations Low AIV AA1						
2024	MID 96 dish SPC MID ready Mid AIV AA2 Low 64 stations SPC Low ready Low AIV AA2						...overy space")
2025	MID 133 dish MeeKAT integration SDP AA3 Full imaging SDP AA4 Mid AIV AA3 Mid AIV AA4 Low 256 stations Low AIV AA3			Several 10s G between ESDC sites Several 10s G between SKA Regional Centres			Low 100G Mid 100G (including "discovery space")
2026	Low 512 stations Low AIV AA4			Several 10s G between ESDC sites Several 10s G between SKA Regional Centres			Low 100G Mid 100G (including "discovery space")
2027	AIV AA4 complete			Several 10s G between ESDC sites Several 10s G between SKA Regional Centres			Low 100G Mid 100G (including "discovery space")

2018-2020 1 – 20Gigabit science data globally & into Europe
100Gigabit tests

2020-2022 Construction phase with data challenges expected
A few 10s Gigabits globally

2022 – 2024 Commissioning / Early Science / data challenges
100 gigabit staring to be used out of each telescope.

2025 onwards
100 Gigabit from each telescope globally

Task 4.3: Network Model within the ESDC

- Assumptions
  - Take 7 sites over Europe
  - 20 Gbits/s at a time from each Telescope
    peak 40 Gbit/s to a site
  - At a given time:
    all of the information from 1 section of a data product goes to one site
  - 10 Gigabits/s to other ESDC sites
  - 10 Gigabits/s to non SKA data Archives
  - 10 Gigabits/s to other RDC

- 70 Gigabit/s (peak) to each site.

- Say 100 Gigabit/s to each site for SKA in 2024/25. This is affordable.
  (c.f. the UK WLCG sites which have 10-40 Gbit/s dedicated links now)

Questions ?

Thanks to Richard Hughes-Jones