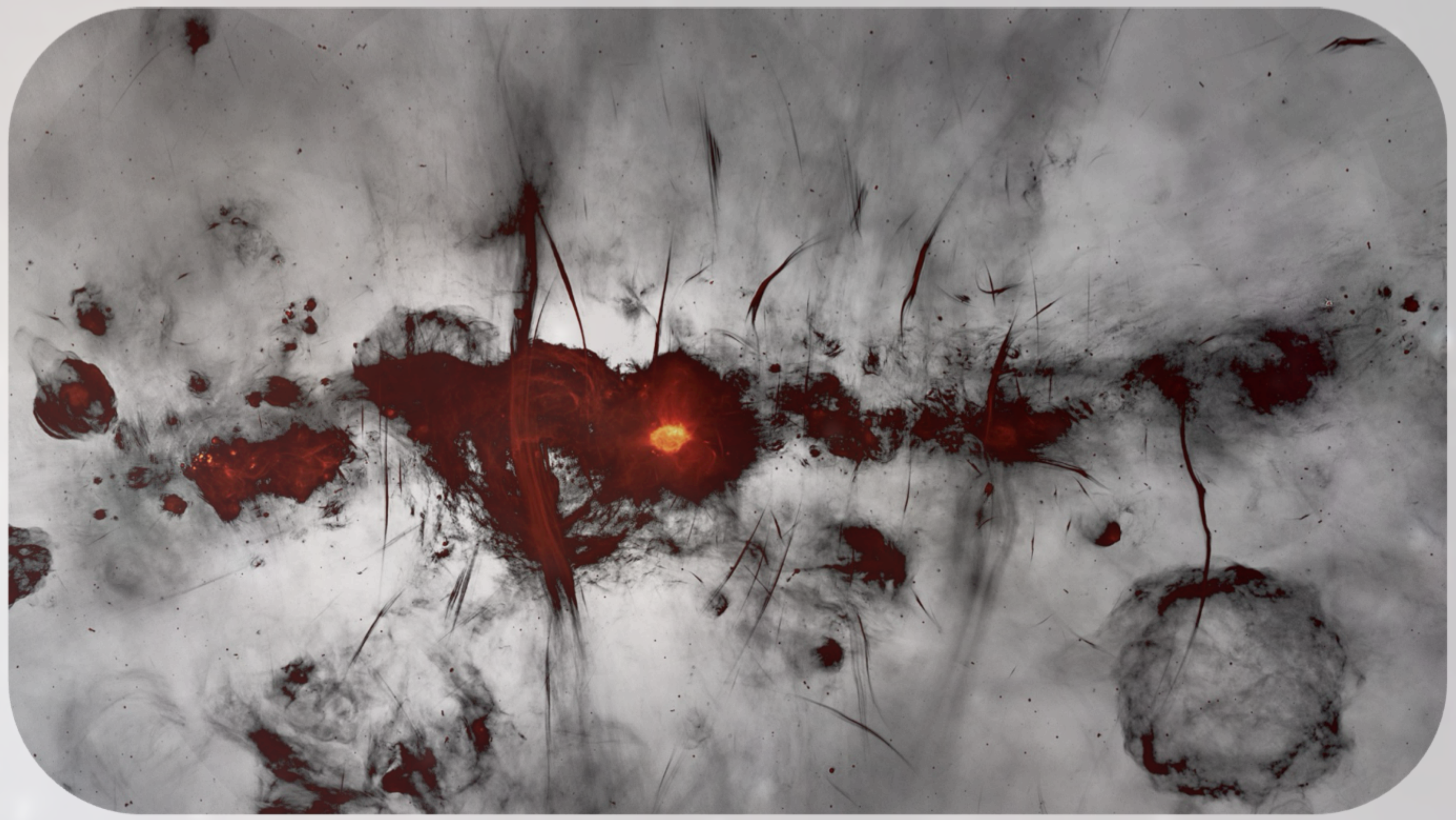
SRCNet (v0.1) for the Square Kilometre Array

James Walder (UKSRC / STFC, RAL)
(With inputs from many people from the SRCNet community)

XRootD and FTS Workshop (Abingdon)
9–13 September 2024

SKA

- The Square Kilometre Array (SKA) is a next-generation radio telescope aimed at being the largest and most powerful ever built
- SKAO (Square Kilometre Array Observatory) is the **governing body** and **Intergovernmental Organisation** responsible for managing the global collaboration, scientific goals, and technical aspects of the SKA project
  - Global project involving over 10 countries, with construction taking place in Australia (SKA-Low) and South Africa(SKA-Mid).
  - > 50 year lifetime
- Two Complementary Telescopes:
  - SKA-Mid (350 MHz – 15.4 GHz): 197 dish antennas
    - located in the Karoo Desert, South Africa. Maximum baseline of O(150) km.
  - SKA-Low (50 MHz – 350 MHz): Over 131,000 dipole antennas
    - grouped into 512 stations in Western Australia, covering a maximum distance of O(60+) km.



Credit: I. Heywood, SARAO

- The SKA Regional Centres (SRCs):
  - SRCs will receive data from the SKAO and act as the scientific archive for SKA data.
  - Global Distribution: SRCs will be located across the world, creating a global network that provides the computational power, data storage, and tools needed for international collaboration.

Science goals

- Diverse and ambitious set of science goals enabling discoveries across a wide range of fields in astronomy, cosmology, and fundamental physics
  - Use of precursor and pathfinder projects to develop and prototype, with own scientific aims

Aspect	Pathfinder Projects	Precursor Projects
Definition	Existing radio telescopes testing technologies relevant to SKA	Prototypes located at SKA sites in South Africa and Australia
Goal	Develop scientific techniques, instrumentation, and methodologies	Conduct early science and engineering directly leading to full SKA
Locations	Worldwide (Europe, Australia, South Africa, etc.)	Located at the SKA sites in South Africa (SKA-Mid) and Australia (SKA-Low)
Notable Projects	LOFAR, ASKAP, MWA, MeerKAT, VLBI	MeerKAT, ASKAP, MWA
Contribution to SKA	Provide foundational data, technology testing, and science exploration	Test SKA-specific designs, contribute to early science, integrate with SKA

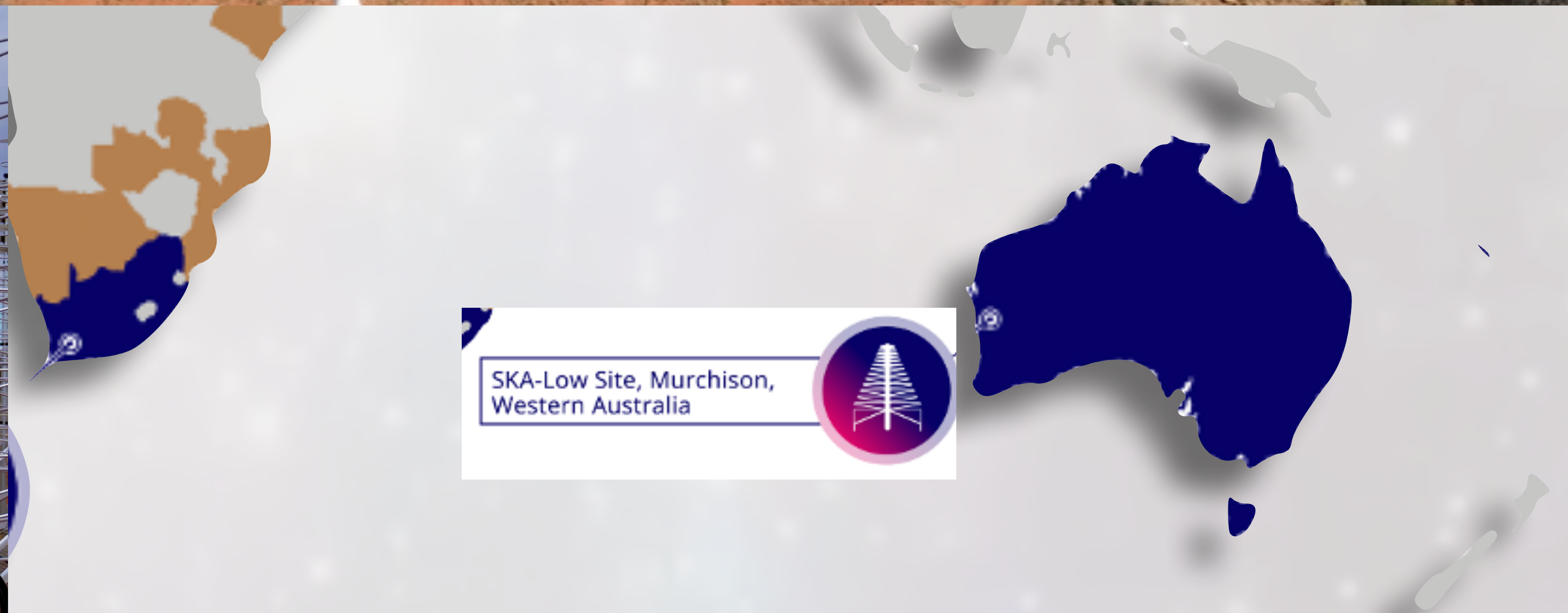
Science Goal	Focus
Cosmic Dawn and Epoch of Reionization	First stars, galaxies, and reionization
Galaxy Evolution and Dark Energy	Large-scale structure, galaxy formation, and dark energy
Cosmic Magnetism	Mapping magnetic fields across the universe
Pulsars and Extreme Gravity	Testing general relativity and extreme gravity with pulsars
Fast Radio Bursts (FRBs)	Understanding the origin and mechanisms of FRBs
Search for Extraterrestrial Life (SETI)	Detecting signals or biosignatures from other civilizations
Cradle of Life	Studying planetary systems and conditions for life
Probing the Dark Ages	Observing the universe before the first stars
Transient Events	Real-time monitoring of cosmic explosions and supernovae
Gravitational Waves and Black Holes	Indirect detection of gravitational waves via pulsar timing arrays

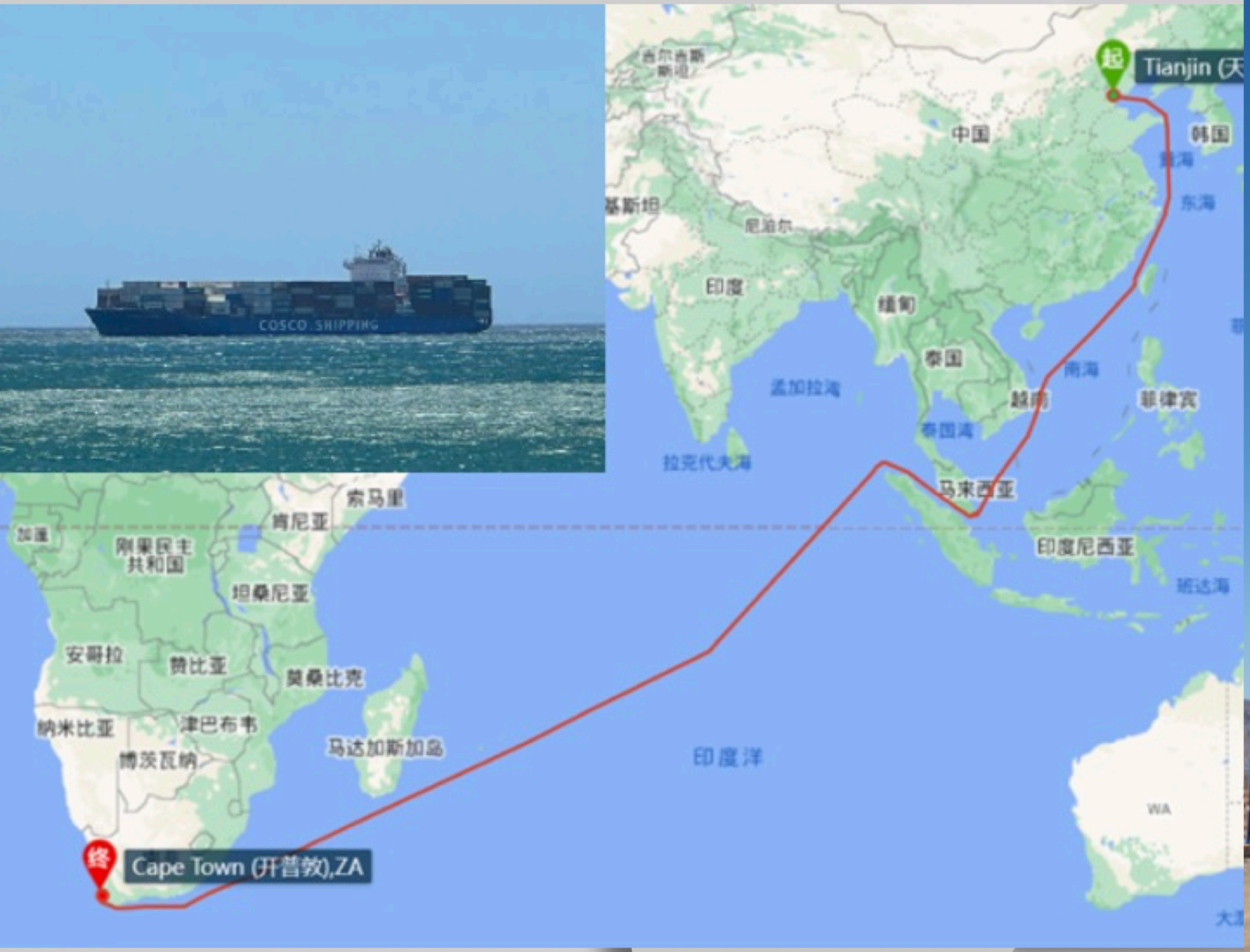
- Don't need to be radio astronomy expert to exploit science data outputs
  - Multi-wavelength data increasingly needed for transformational science
- Property data access and embargo periods for projects

Locations and participating nations


SKA-Mid Site,
Karoo, South Africa
SKA-Low Site, Murchison,
Western Australia





SKA-Low Site, Murchison,
Western Australia


COSCO SHIPPING
Tianjin (天
中国
韩国
黄海
东海
印度
缅甸
泰国
南海
越南
菲律宾
孟加拉湾
泰国湾
马来西亚
印度尼西亚
拉克代夫海
印度洋
肯尼亚
索马里
坦桑尼亚
安哥拉
赞比亚
莫桑比克
纳米比亚
津巴布韦
博茨瓦纳
马达加斯加岛
Cape Town (开普敦),ZA


17.01.2024

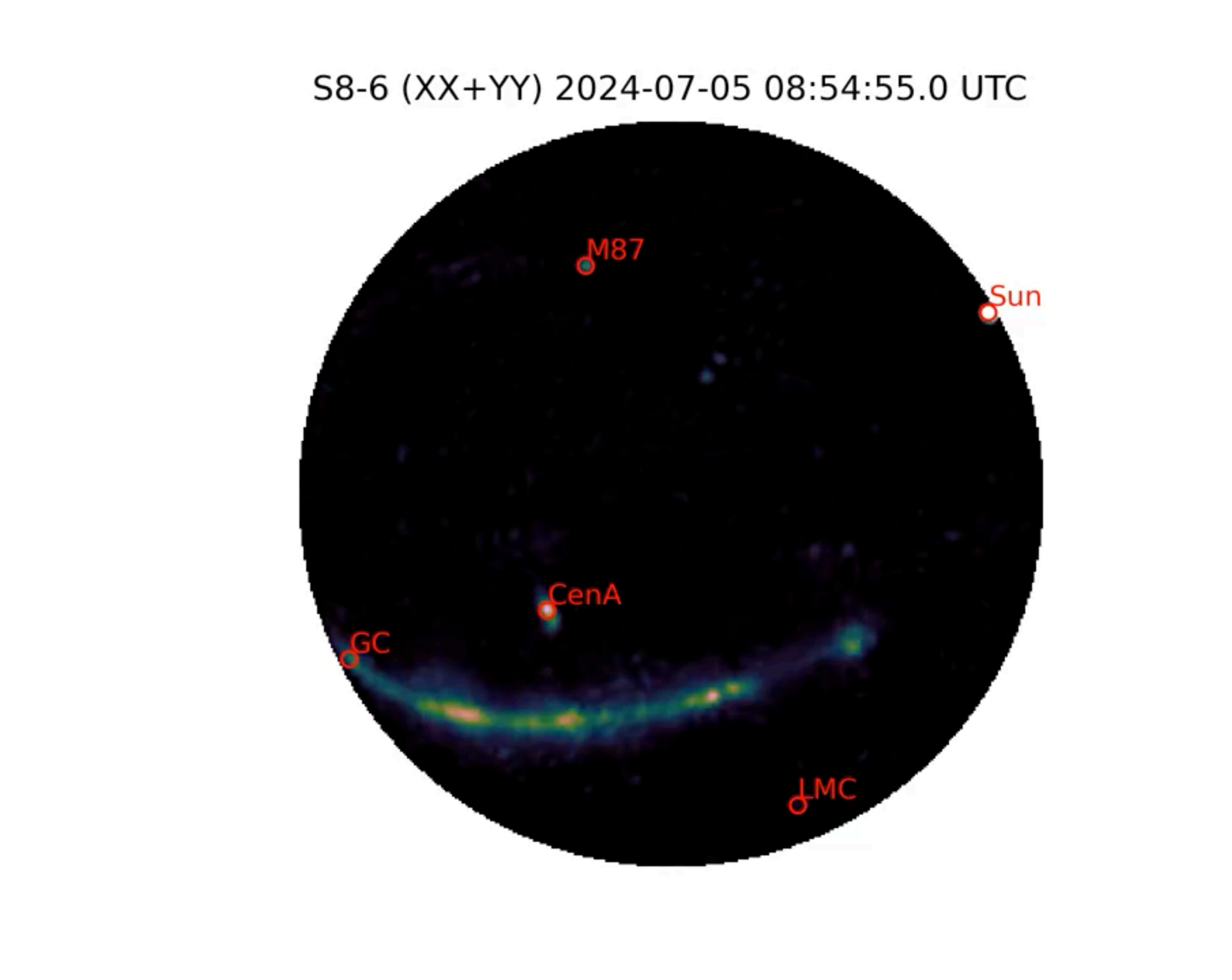
First Light!

- This is the first image and video from observations using **one complete SKA-Low station**, known as S-8, produced only **18 months after the start of construction** activities on site, and five months after the first antenna was installed.

- The completion of a station means not only assembling and installing the **256 antennas**, but also integrating them with all the computing systems behind them.

- The video shows a **24-hour observation**, with the Milky Way rising and passing overhead during the night time hours.

  Some other bright radio sources are marked, including the galaxies Centaurus A and M87, and the Sun is also visible during the day.



https://www.youtube.com/watch?v=zakuQ1-QrGg

Construction Timeline

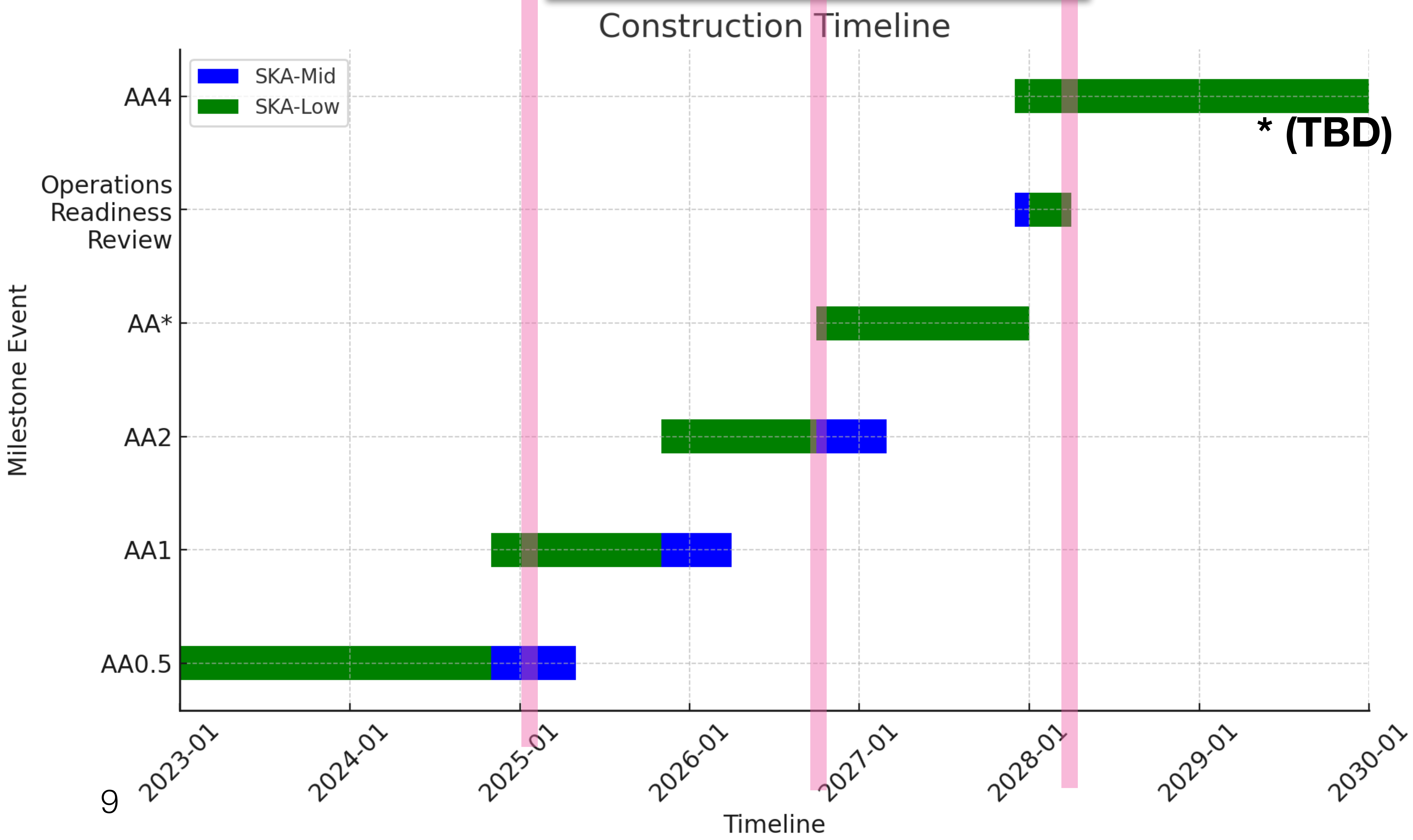
- Construction Strategy
- Target: build the SKA Baseline Design
  (197 Mid dishes; 512 Low stations: AA4)
  - Not all funding yet secured for full AA4 components, following Staged Delivery Plan (AA*)
- Develop the earliest possible working demonstration of the architecture and supply chain (AA0.5).
- Then maintain a continuously working and expanding facility that demonstrates the full performance capabilities of the SKA Design.

Milestone event (earliest)	SKA-Mid (end date)	SKA-Low (end date)
AA0.5 (4 dishes, 6 stations)	2025 May	2024 Nov
AA1 (8 dishes, 18 stations)	2026 Apr	2025 Nov
AA2 (64 dishes, 64 stations)	2027 Mar	2026 Oct
AA* (144 dishes, 307 stations)	2027 Dec	2028 Jan
Operations Readiness Review	2028 Apr	2028 Apr
AA4 (197 dishes, 512 stations)	TBD	TBD

Pre science Verification
- Commissioning (+ Assembly, Integration and Verification) primary activity
- SRCs not needed to support AA0.5/AA1 commissioning
- Opportunity for testing (data, transfer, access, pipelines)!

Cycle 0
- "Proper" shared risk projects
- Teams, proprietary periods, visualisation, ADP creation etc

Science Verification
- Data immediately public
- Full dress rehearsal!
- Some SRCNet resources for analysis would be an advantage
- Observed as trickle but also in dedicated blocks
- (+ Commissioning etc ongoing)



Data Flows

* Data rates approximate

SKA expected data rates*

***these numbers should be used as a guide only - email Shari.Breen@skao.int for further information about ongoing work**

- Numbers refer to data to be delivered to the science community via the SRCNet (i.e. not data used internally for commissioning etc.)

Milestone	Year	Primary activity	Estimated data rate	
			Low	Mid
AA2 • 64 Mid dishes • 64 Low stations	2026 - 2027	**Science Verification** - observed in dedicated ~week long blocks + single observations interspersed throughout. A higher rate of raw data products will be included at this stage.	1.5 PB/week^ 20 Gbps	2 PB/week^ 27 Gbps
AA* • 144 Mid dishes • 307 Low stations	2027 - 2029	**Science Verification** - observed in dedicated ~week long blocks + single observations interspersed throughout. A higher rate of raw data products will be included at this stage.	5 PB/week^ 66 Gbps	9 PB/week^ 119 Gbps
AA* • 144 Mid dishes • 307 Low stations	2029 +	**Operations** - Observation cycles, starting with shared risk observing, building to successful science observations ~90% of the time	173 PB/year 44 Gbps	280 PB/year 72 Gbps
Target is to deliver the SKA Baseline Design but the details of this transition between AA* and AA4 are TBD				
AA4 • 197 Mid dishes • 512 Low stations	2030 +	**Operations** - full SKA baseline design	216 PB/year 55 Gbps	400 PB/year 100 Gbps

^Data rates refer to dedicated Science Verification observing weeks, not an average over a year

SRCNet

- The need for a network of SKA Regional Centres (SRCs) formed around ~ 2018:
  - Distributed compute, storage and expertise to store, process and disseminate data to the communities

- SRCNet is a global, distributed infrastructure designed to:
  - **Store and manage SKA data**:
    - Centralised data management is impractical given the size of the SKA data output, so SRCNet provides regional centres to manage and store data close to its users.
  - **Enable scientific collaboration**: Astronomers worldwide can access, process, and analyze SKA data in a collaborative environment, regardless of their location.
  - **Provide computational resources**: Given the vast amount of data, powerful high-performance computing (HPC)resources are needed. SRCNet provides this by distributing the workload across a global network of SRCs.
  - **Support scientific tools and pipelines**: SRCNet offers standardised tools and software pipelines for researchers to work on SKA data, including supporting advanced data analysis, visualisation, and machine learning techniques.

My Naive mapping between LHC and SKA entities

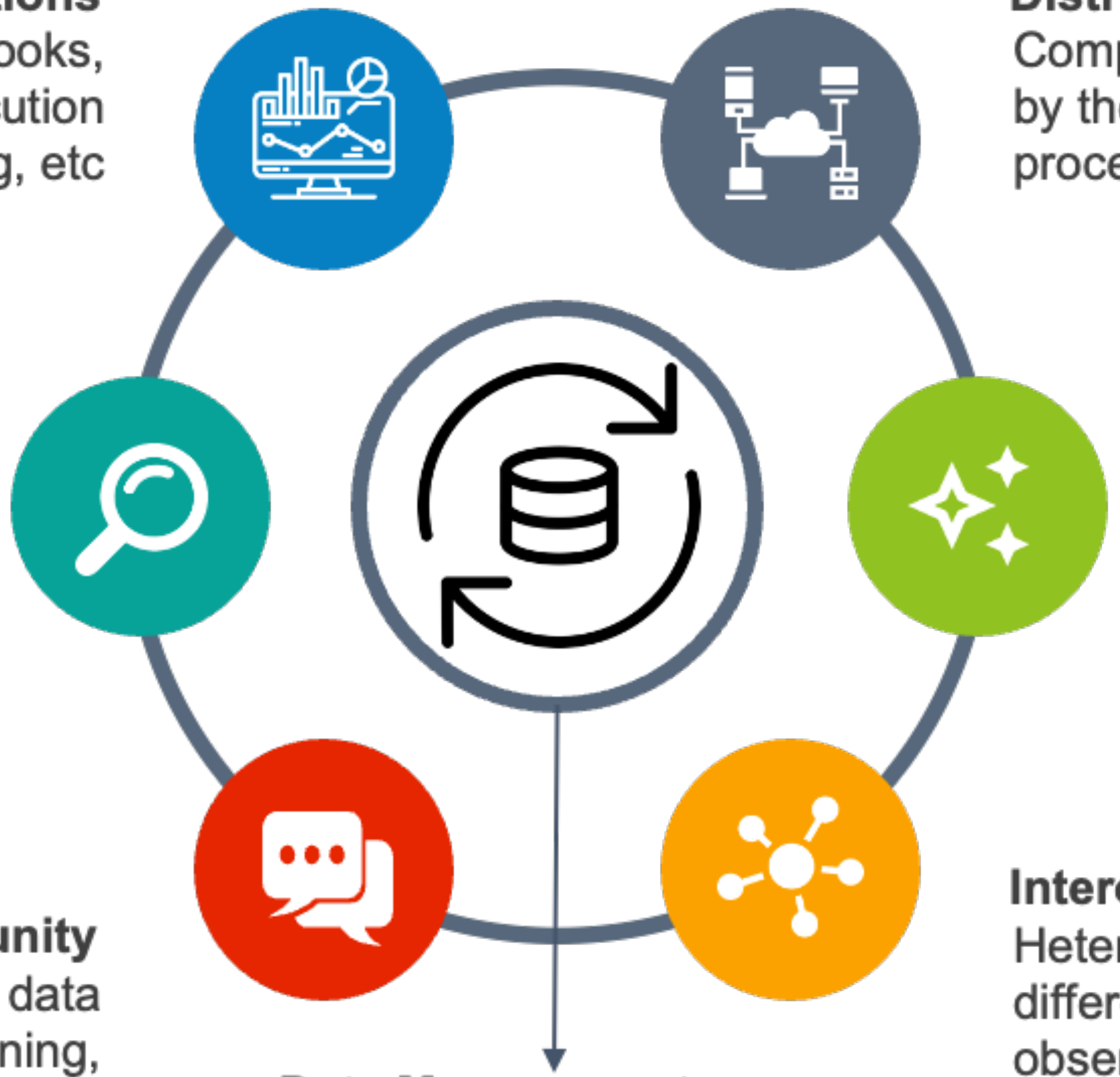
Don't take too literally

LHC/WLCG	SRCNet
Cern + Experiments	~ SKAO + SKA-MID + SKA-LOW
WLCG	~SRCNet
GridPP/ …	~UKSRC / …

Region Center Capabilities Blueprint

Science Enabling Applications
Analysis Tools, Notebooks, Workflows execution Machine Learning, etc

Distributed Data Processing
Computing capabilities provided by the SRCNet to allow data processing



Data Discovery
Discovery of SKA data from the SRCNet, local or remote, transparently to the user

Visualization
Advanced visualizers for SKA data and data from other observatories

Support to Science Community
Support community on SKA data use, SRC services use, Training, Project Impact Dissemination

Interoperability
Heterogeneous SKA data from different SRCs and other observatories

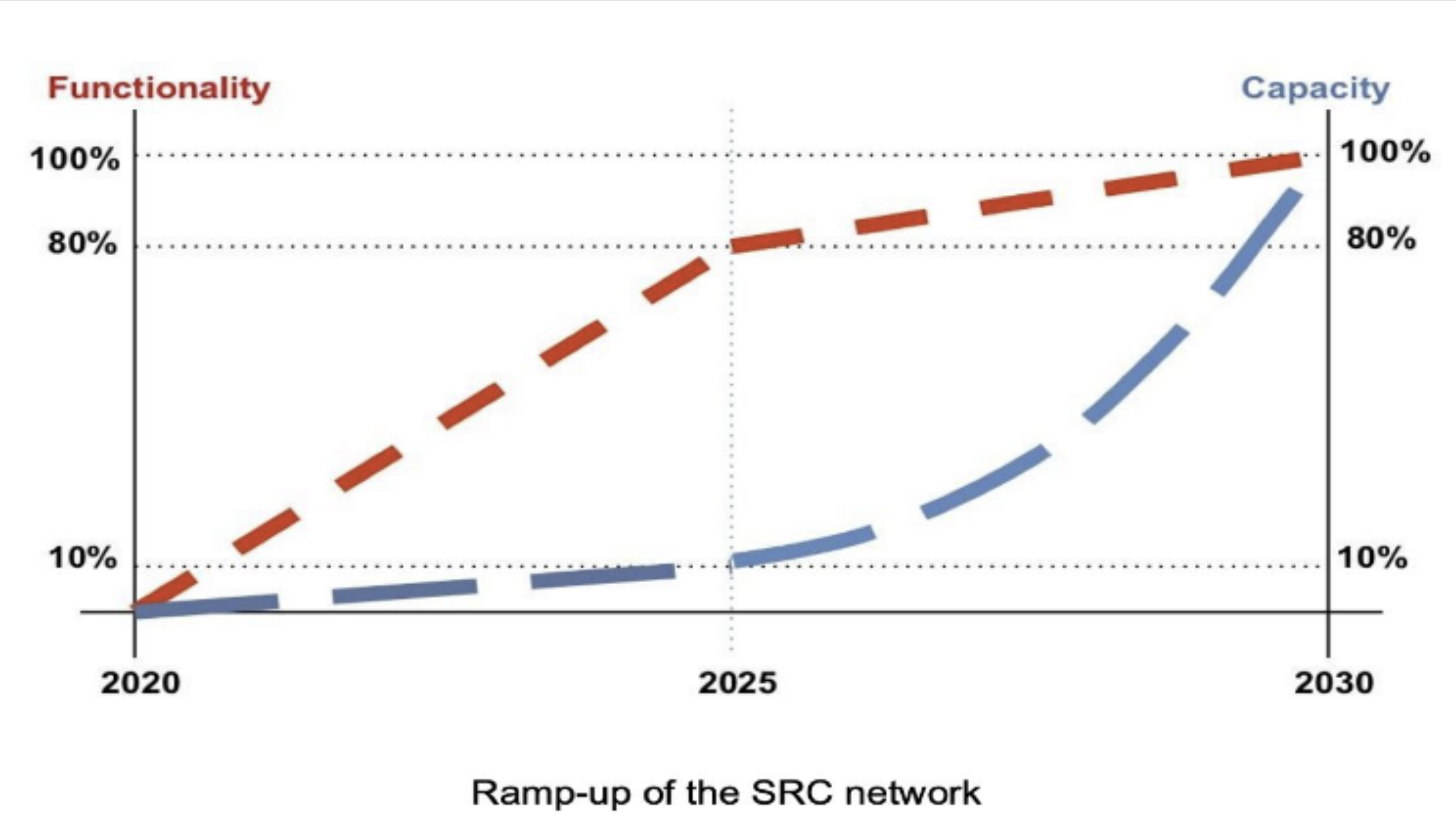
Data Management
Dissemination of Data to SRCs and Distributed Data Storage

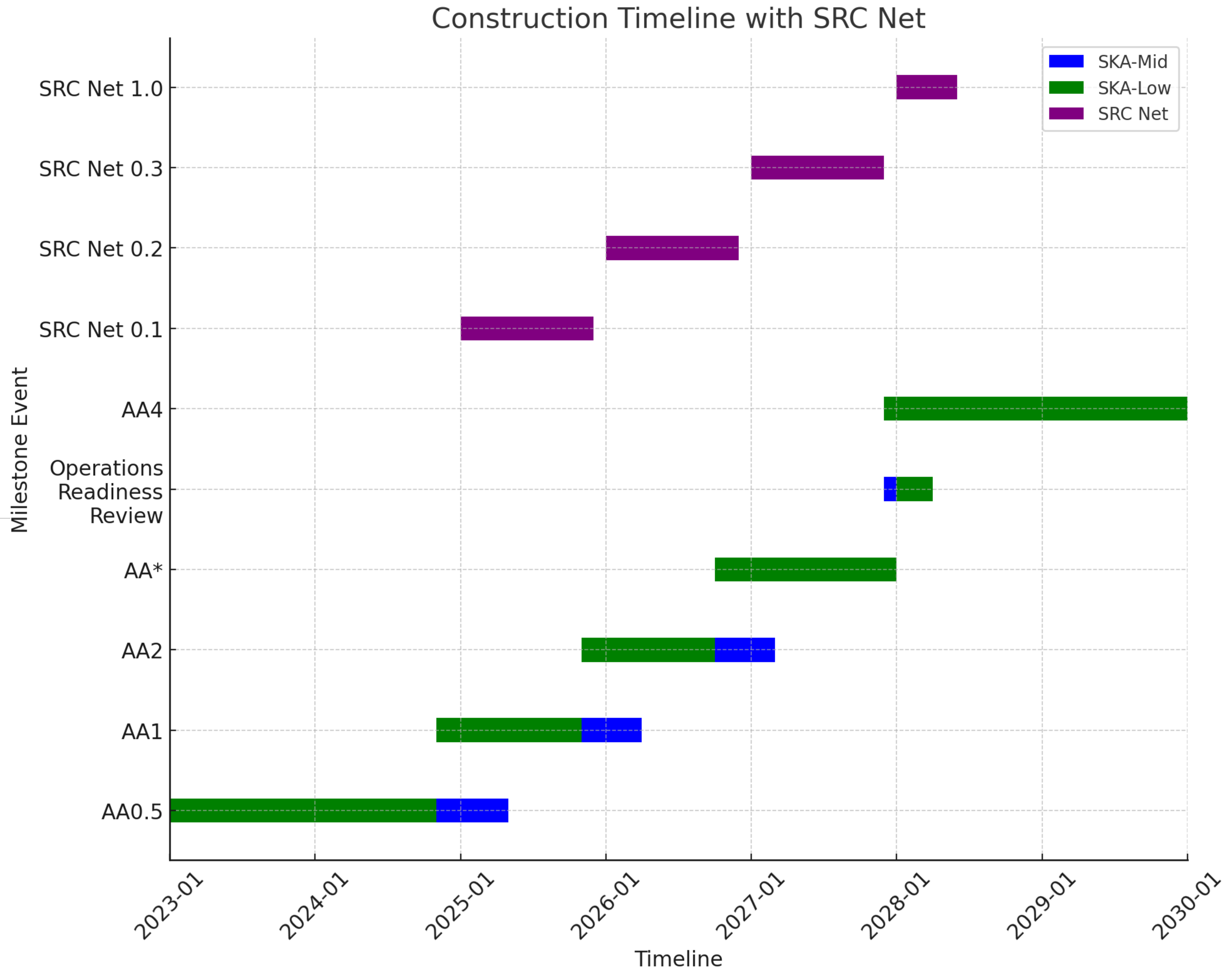
SRCNet Timeline

- SRCNet timeline:
  - SRCNet v0.1 (2025) :
    First major milestone in SRCNet project

- Built to show the architecture and test how it works
- Internal only - no user-facing activities
- Exclusive storage to use in testing
- Compute to use during testing campaigns
  (may be backfilled when idle)
- Learn how to deploy and operate the services

Set up of the SRC Operations Group, with limited scope

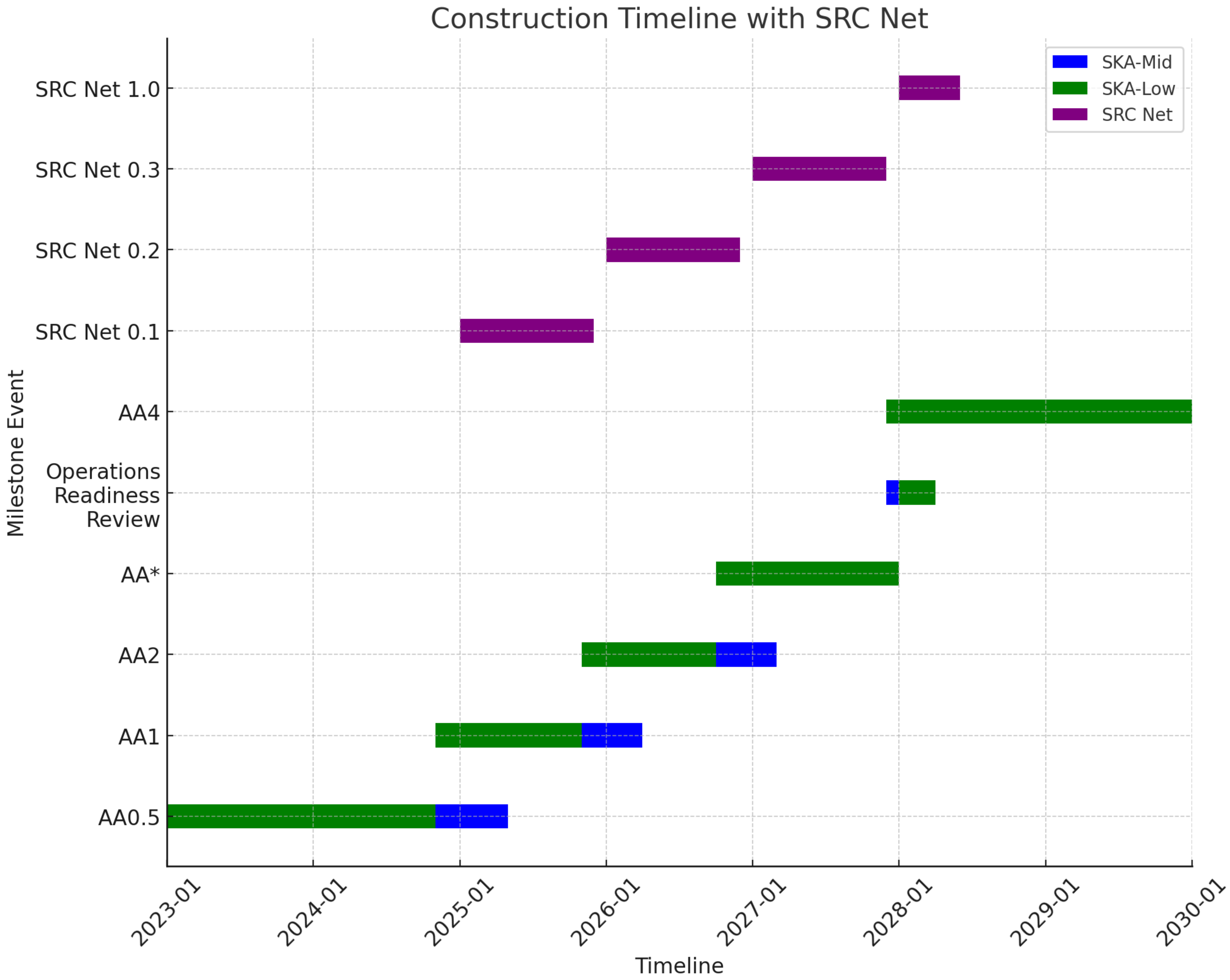




SRCNet Timeline

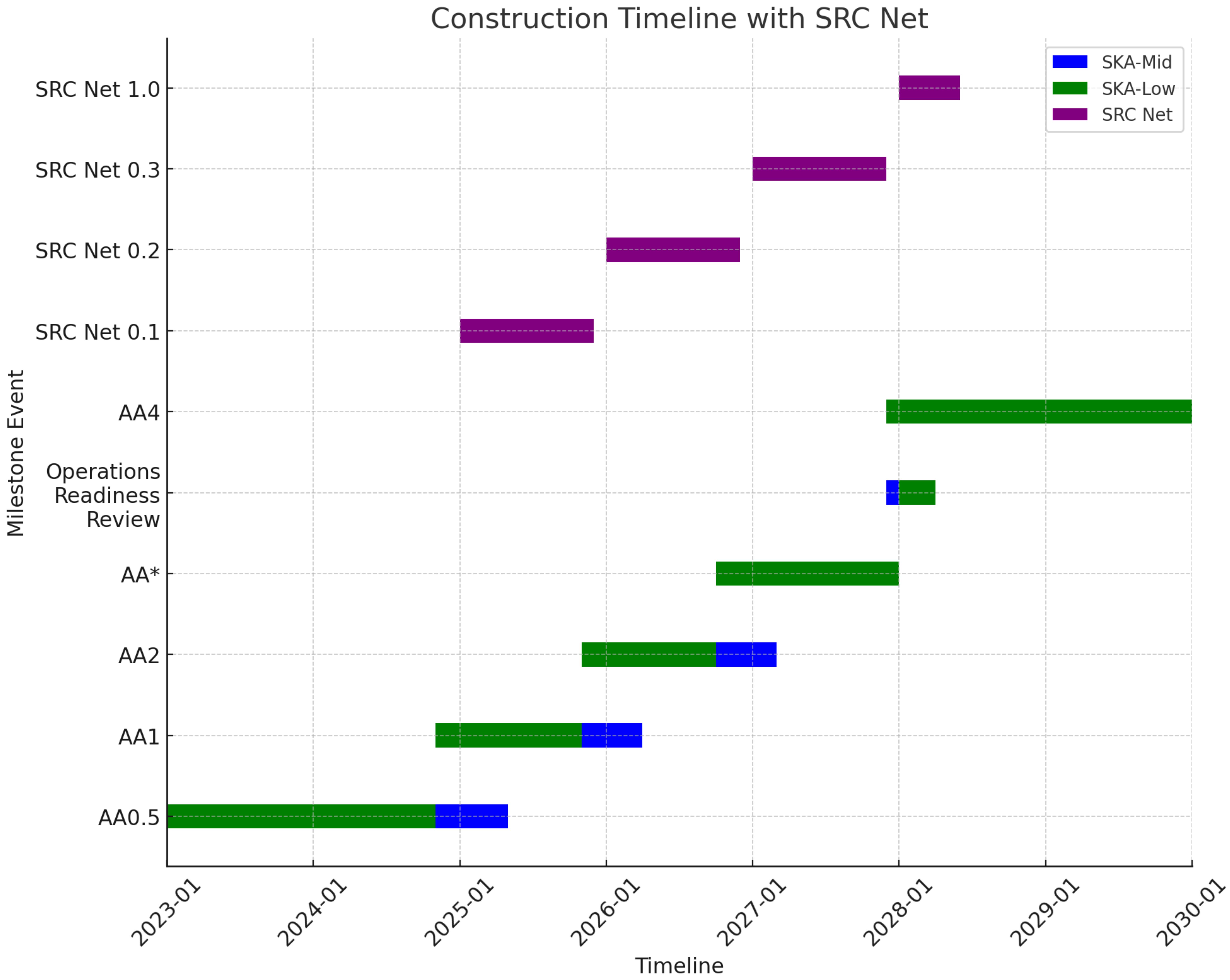
- SRCNet timeline:

- SRCNet v0.2 (2026)
  - Increased Federation
  - Science verification starts
  - Data ingestion from telescopes



SRCNet Timeline

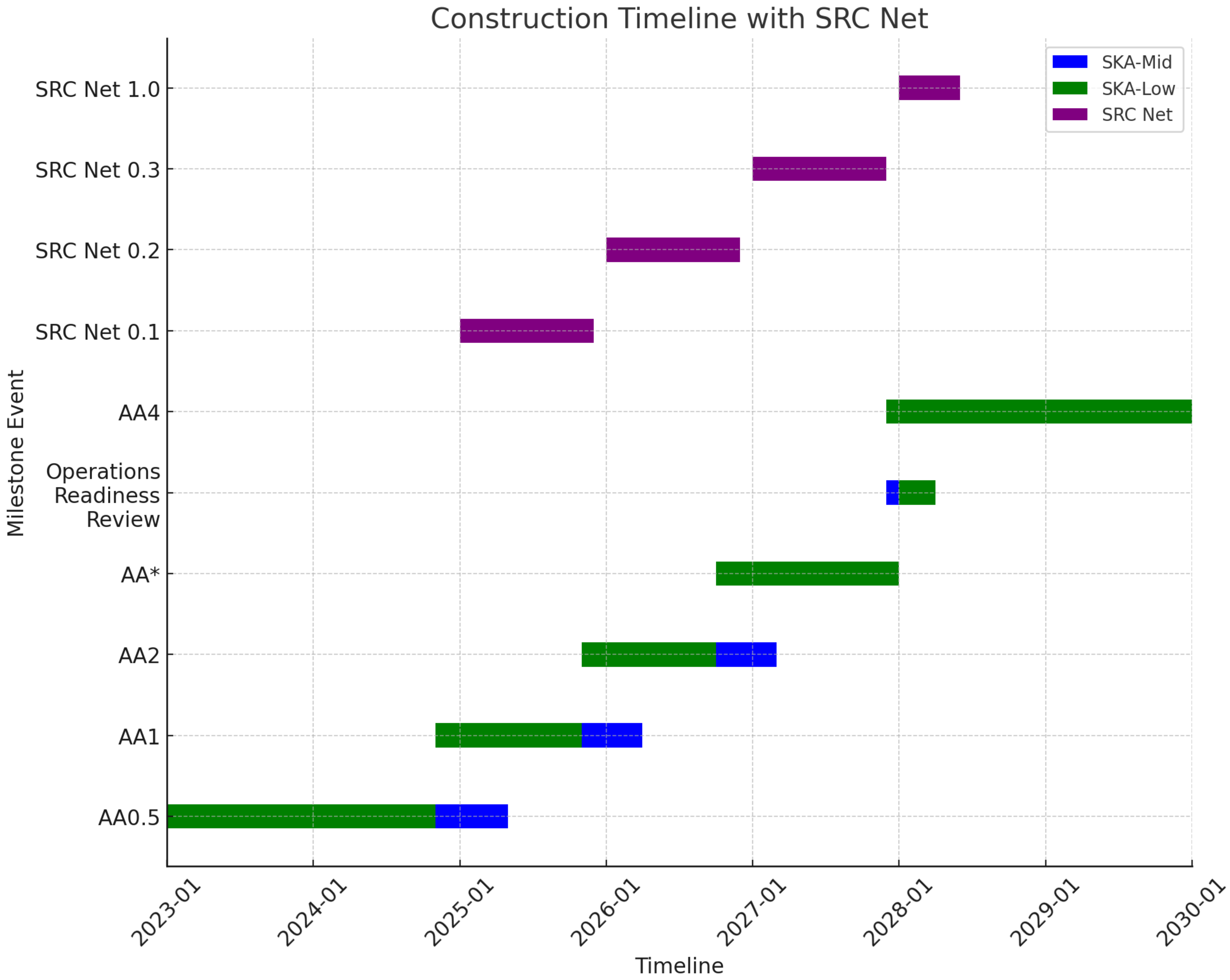
- SRCNet timeline:



- SRCNet v0.3 (end 2026)
  - Increased science verification
  - Science user testing

SRCNet Timeline

- SRCNet timeline:



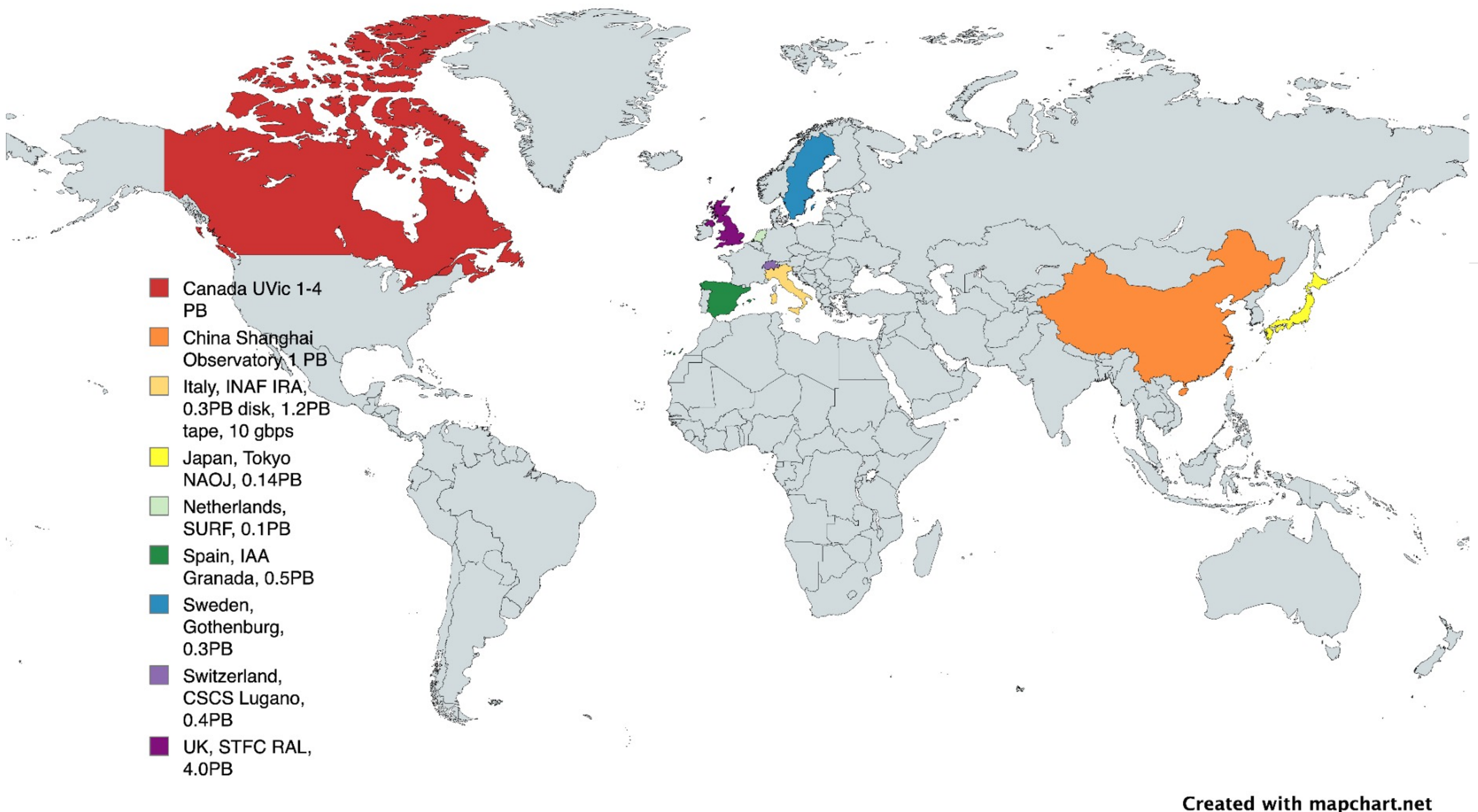
- SRCNet v1.0 (~2028+)
  - Full operations

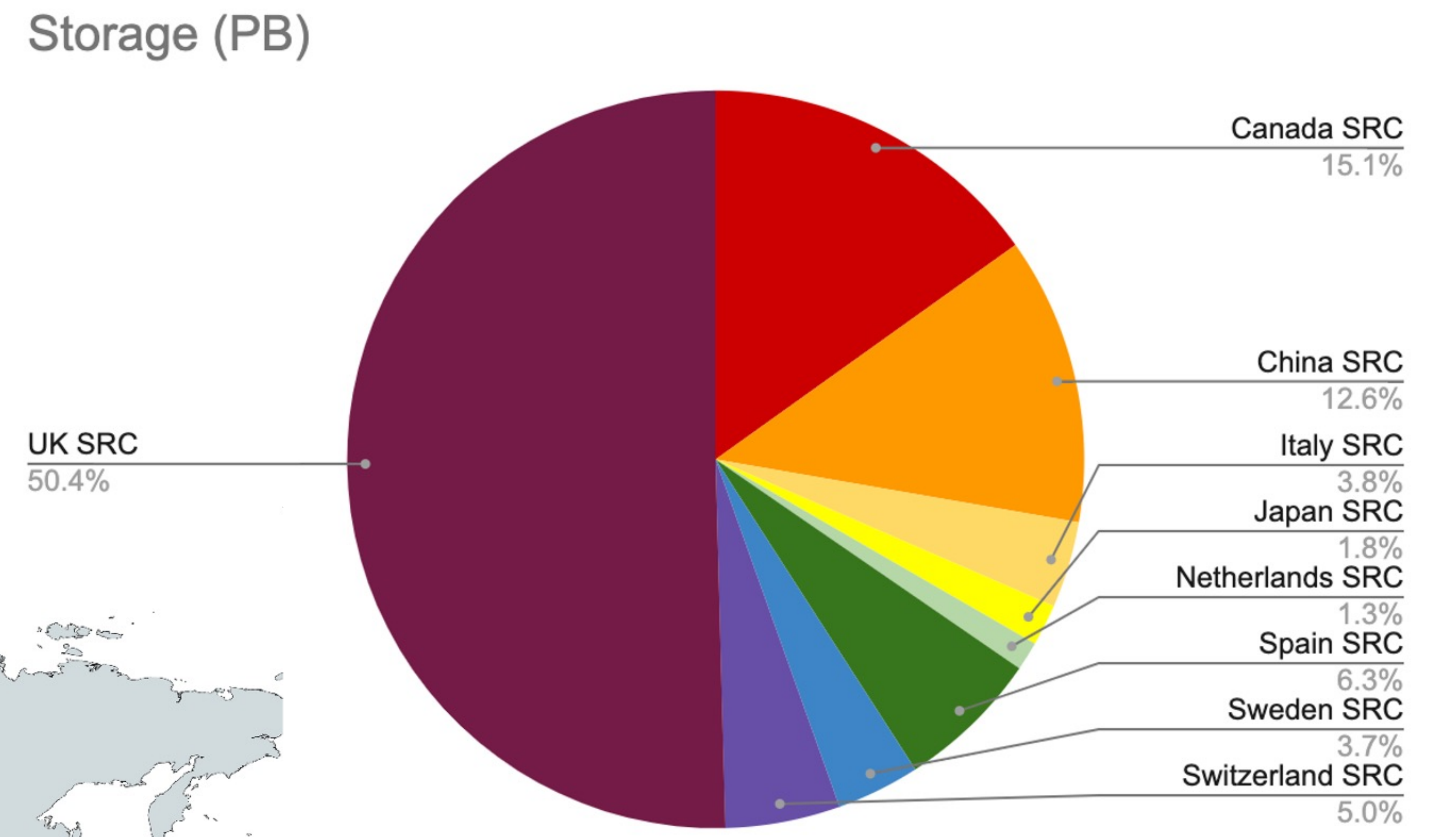
SRCNet v0.1

- Expression of Interest circulated for countries to participate in v0.1.
- 9 resulting sites; providing resources in line with SKA Top-level roadmap requirements

SRCNet0.1 included sites

8 PBytes total storage offered for SRCNet0.1 (c.f original target of 20 PB)

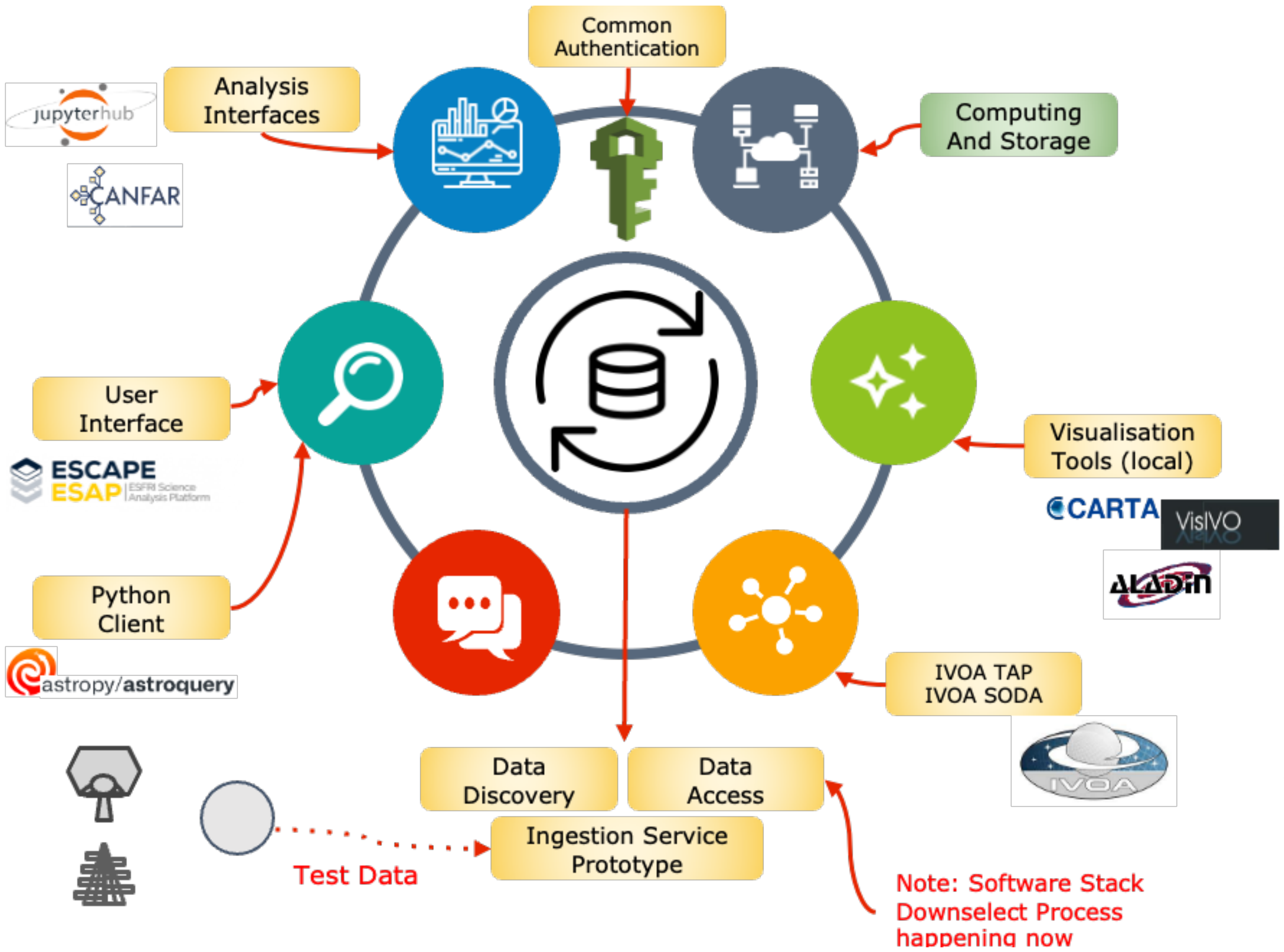




WLCG experience at some sites (Canada, Netherlands, Sweden, Switzerland, UK)

Several new sites and teams will learn by being involved

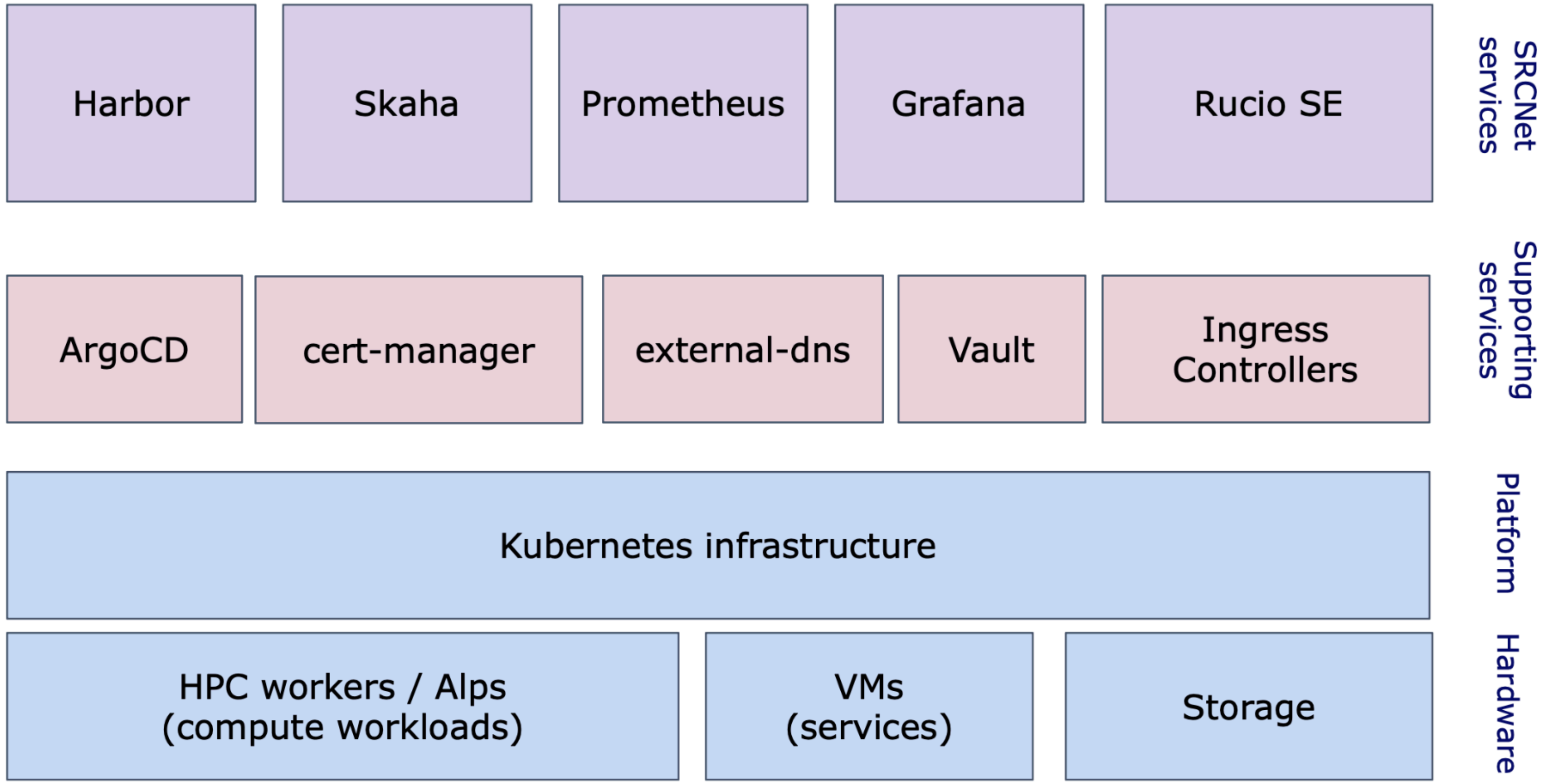
SRCNet v0.1 Deployment



- Common Authentication
  - IAM
- Visualisation Tools (local)
- IVOA Protocols
  - TAP, SODA
- Data Discovery and Access from Data Lake
- Ingestion Service Prototype
- Python Client
  - Astroquery Module
- User Interface
  - ESAP
  - https://esap.srcdev.skao.int/
- Analysis Interfaces
  - JupyterHub
  - CANFAR Science Platform

Deployment Methodology

- Aim for common deployment methods in the SRCNet where possible
  - GitOps style approach recommended (e.g. ArgoCD/FluxCD, k8s);
    - Some difference in deployments due to heterogeneous sites / infrastructure adds complexity. (e.g. for established multi-project sites).
- Target the end goal to decouple site from service and enable deployment / upgrades with minimal complexity and allow central management (within reason …)
- **Tools** from (e.g) HEP, such as **FTS / XRootD** help to **abstract away** site implementation details to allow common designs.



- GitOps - managing the configuration for your infrastructure using git

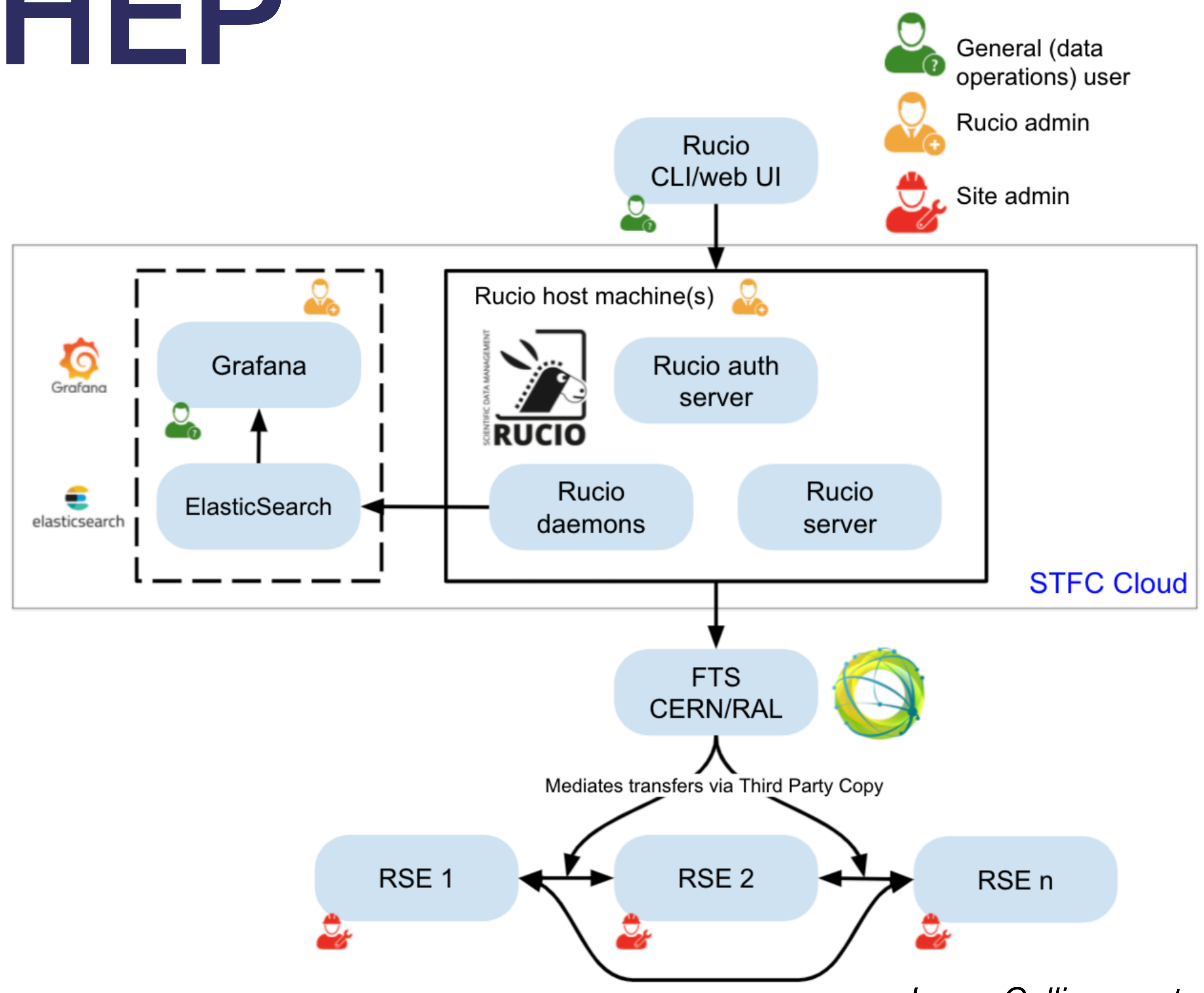
Recommended / Proposed approach

Other sites may have own requirements deployment methods

Pablo Llopis

Synergies with HEP

- Significant prior work undertaken via ESCAPE and with pre-SRCNet v0.1 test deployments.
- DDM assessment in early 2024:
  - Conclusion to use **Rucio** (and associated stack) for SRCNet v0.1
    - Other options, e.g. Storage Inventory considered
    - Currently hosted on (via k8s) STFC-Cloud
  - Abstracted via a **DM API service** for interactions with other service
    - Interoperability with Metadata services and other standards (IVOA-VOSpace)
  - **INDIGO-IAM**
    - No use of X509–VOMS user certificates
    - To migrate shortly from ESCAPE to SKA IAM
  - **FTS**
    - Recent migration from FTS-pilot to SKA instance (See talk from Rose Cooper)
  - Sites using **StoRM, dCache, XRootD**
  - **perfSONAR** to be used for network monitoring and diagnosis



James Collinson, et. al

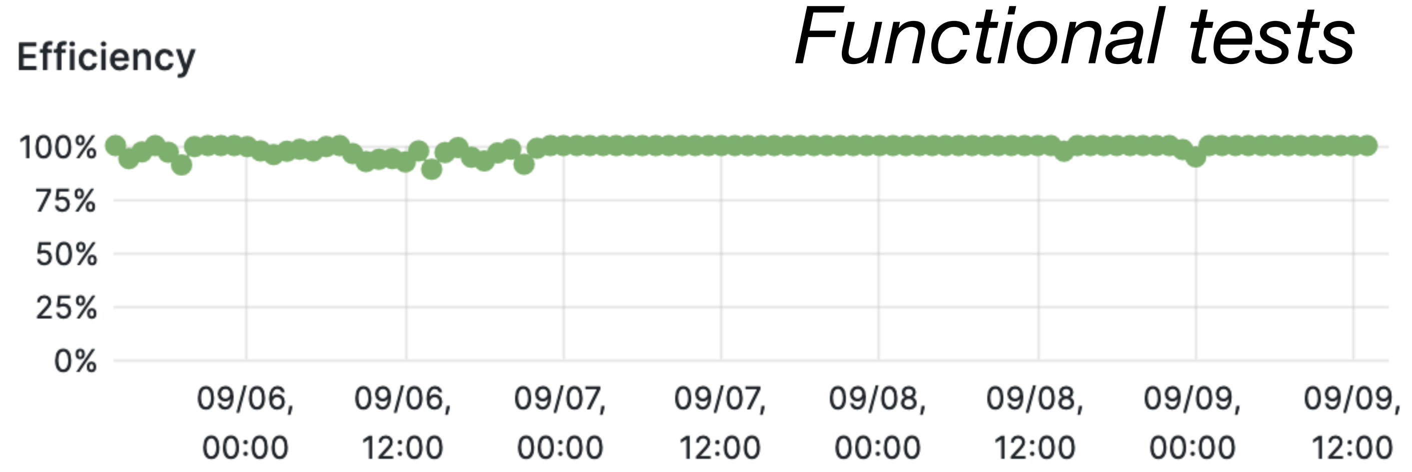
Transfer failure site matrix

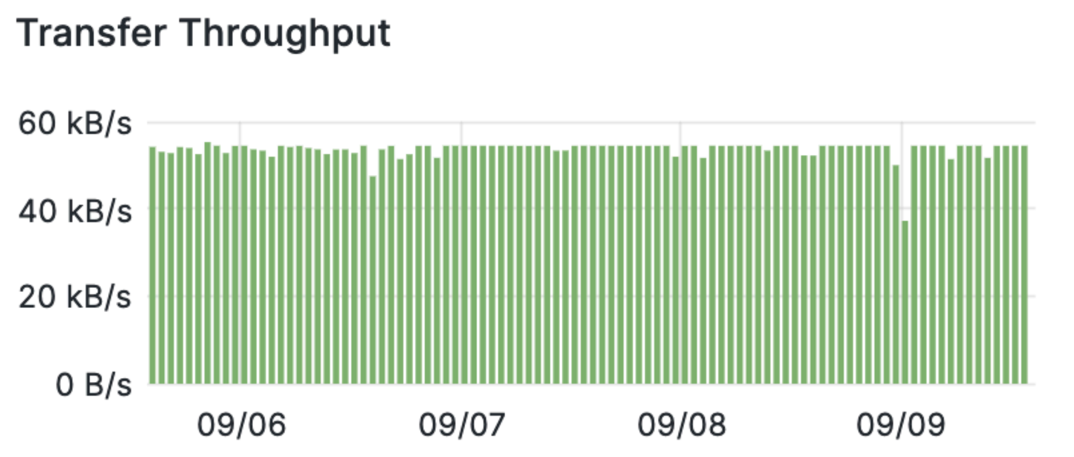
Src\Dst	STFC_STORM_ND	STFC_STORM	SPSRC_STORM	NLSRC_PROD_DC	NLSRC_DCACHE	KRSRC_STORM	JPSRC_STORM	IMPERIAL
_XRD_DEVCEPHFS	0%	0%	0%	0%	0%	100%	0%	0%
STFC_STORM	100%	-	0%	0%	0%	100%	0%	0%
SPSRC_STORM	100%	100%	-	0%	0%	100%	0%	0%
RC_PROD_DCACHE	0%	0%	0%	-	0%	100%	0%	0%
NLSRC_DCACHE	0%	0%	0%	0%	-	100%	0%	0%
JPSRC_STORM	100%	100%	0%	0%	0%	100%	-	0%
IMPERIAL	0%	0%	0%	0%	0%	100%	0%	-
CNAF	100%	100%	0%	0%	0%	100%	0%	0%
CASRC_XRD	0%	0%	0%	0%	0%	100%	0%	0%

Data Lifecycle / Movement Challenges 2025

- SRCNet v0.1 to execute a number of mini Data moving and Data lifecycle challenges.
  - Demonstrate data ingestion (this includes metadata ingestion)
  - Data movement campaigns
  - Data access and usage tests

Functional tests

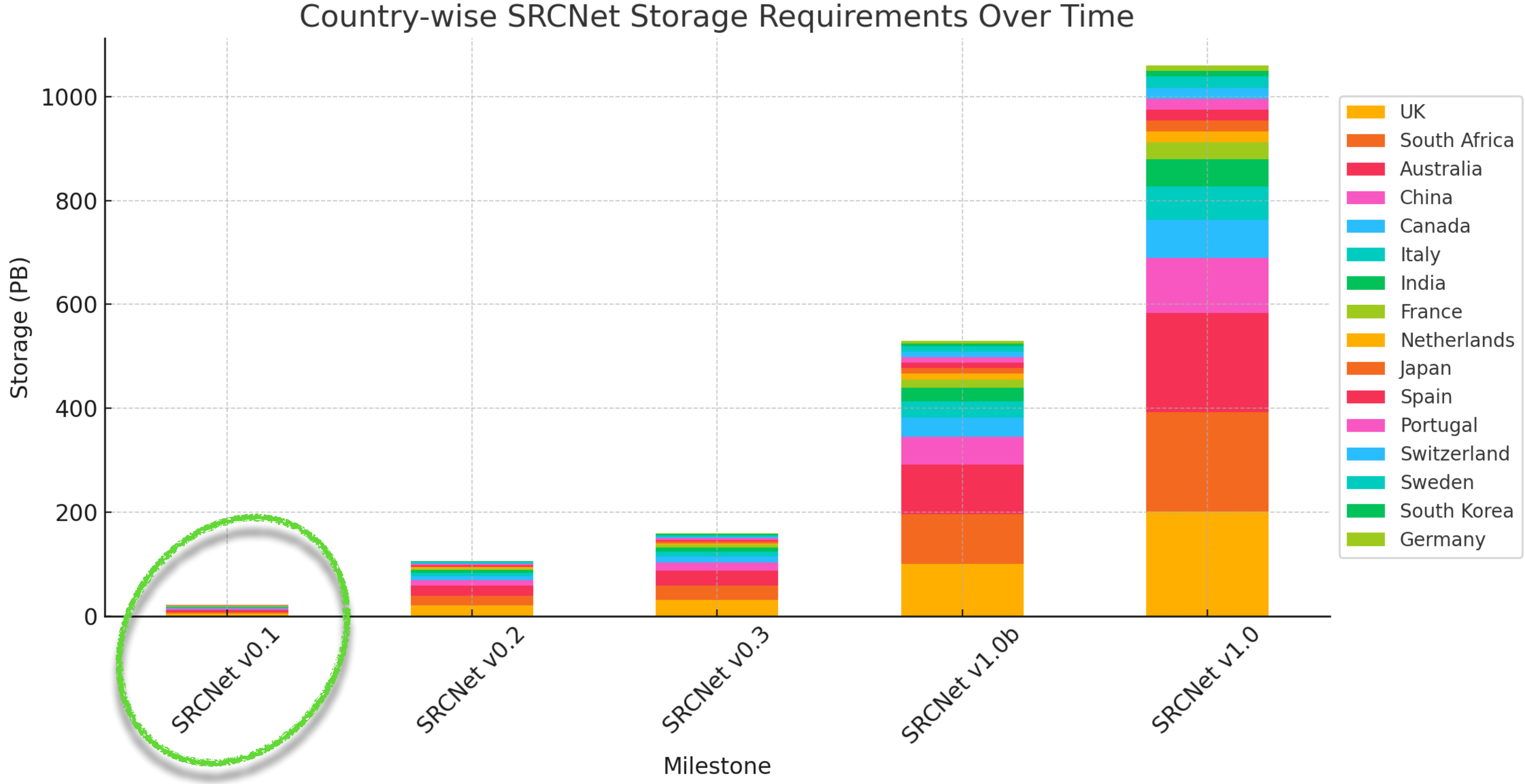




- Public data and simulated data to be used in campaigns
- Dedicated O(week)s during the year. Resources can be assigned when needed

- Work ongoing to define the details of the tests
  - Main requirement to verify prototype architecture
  - Need also to explore scalability to prepare for scale of full operations

- Interest in participating in DC2X challenges as experience grows
  - Experience from a number of SRCNet v0.1 sites from WLCG already.



SKA Top-Level Roadmap requirements timeline for Online storage

Challenges for Data Movement

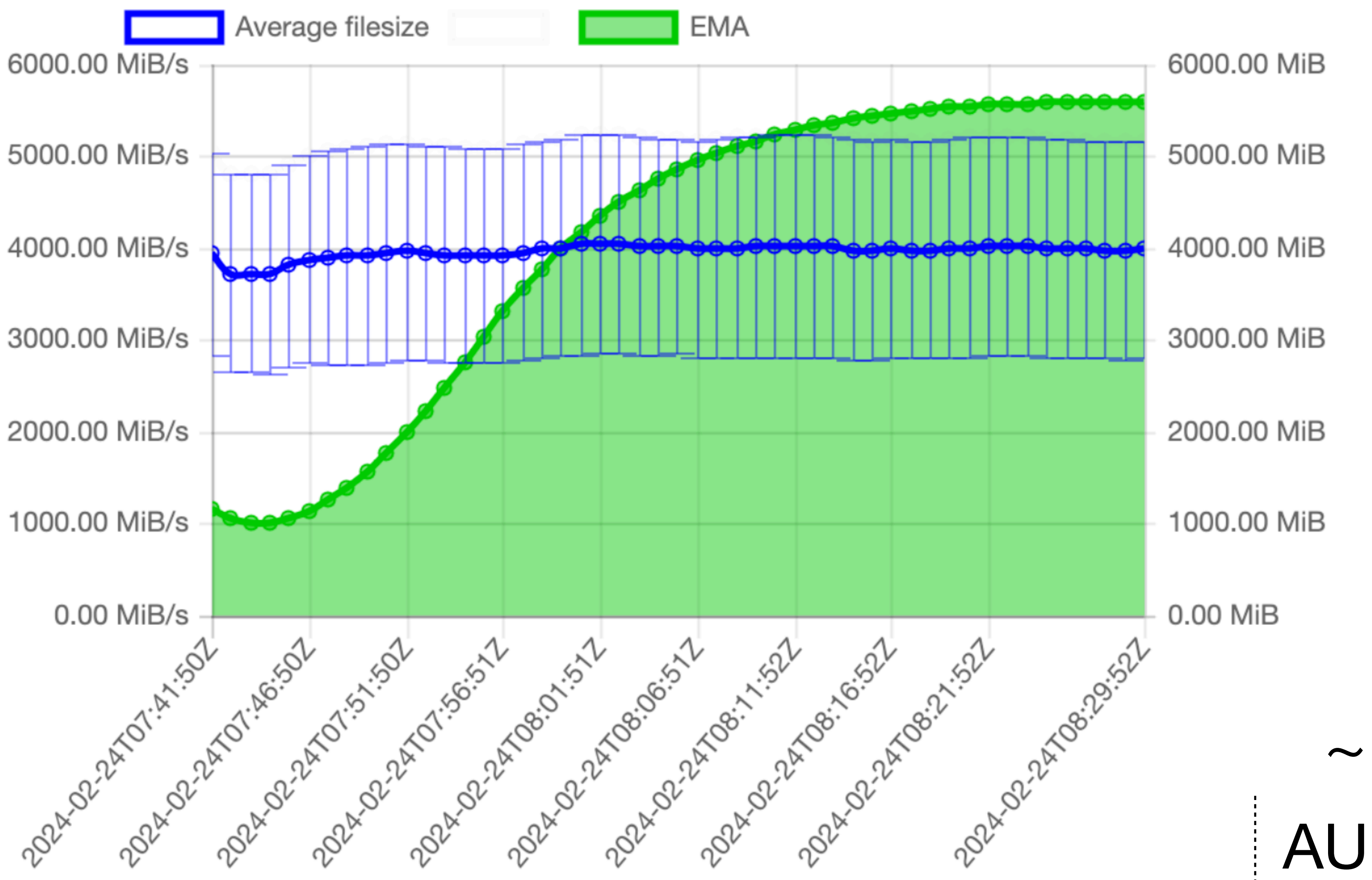
- (Nominal) Target of 100 Gbps from each Telescope site (AA4)
  - O(600) PB data generated / year (50 year lifetime)
- Online storage expected to grow slowly with time
  - Will hold recent and popular data
- Nearline storage to archive old data
  - Noting; Old data may be just as useful as new …

- 'Datasets' may be large O(PB), to be analysed 'together'
- RTT values from Data source (ZA / AU) O(150 – 300) ms
  - High latency
  - Larger file sizes, improved per-file transfer speeds
  - BDP (10 – 100 Gbps) 250MB – 2.5 GB
  - Careful tuning, ~ zero packet loss, parallel streaming (?)
  - Other protocols ? QUIC ?
- Bandwidth in NREN and commodity networks likely to be sufficient by O(2030).

- Custodial storage (i.e tape) should not be in the originating country
  - Telescopes will hold a 'copy of last resort'.

- Data transfers between SRC Nodes and within regional centres also to be modelled.

Cern -> 100Gbps Host (RAL) [Memory testing]

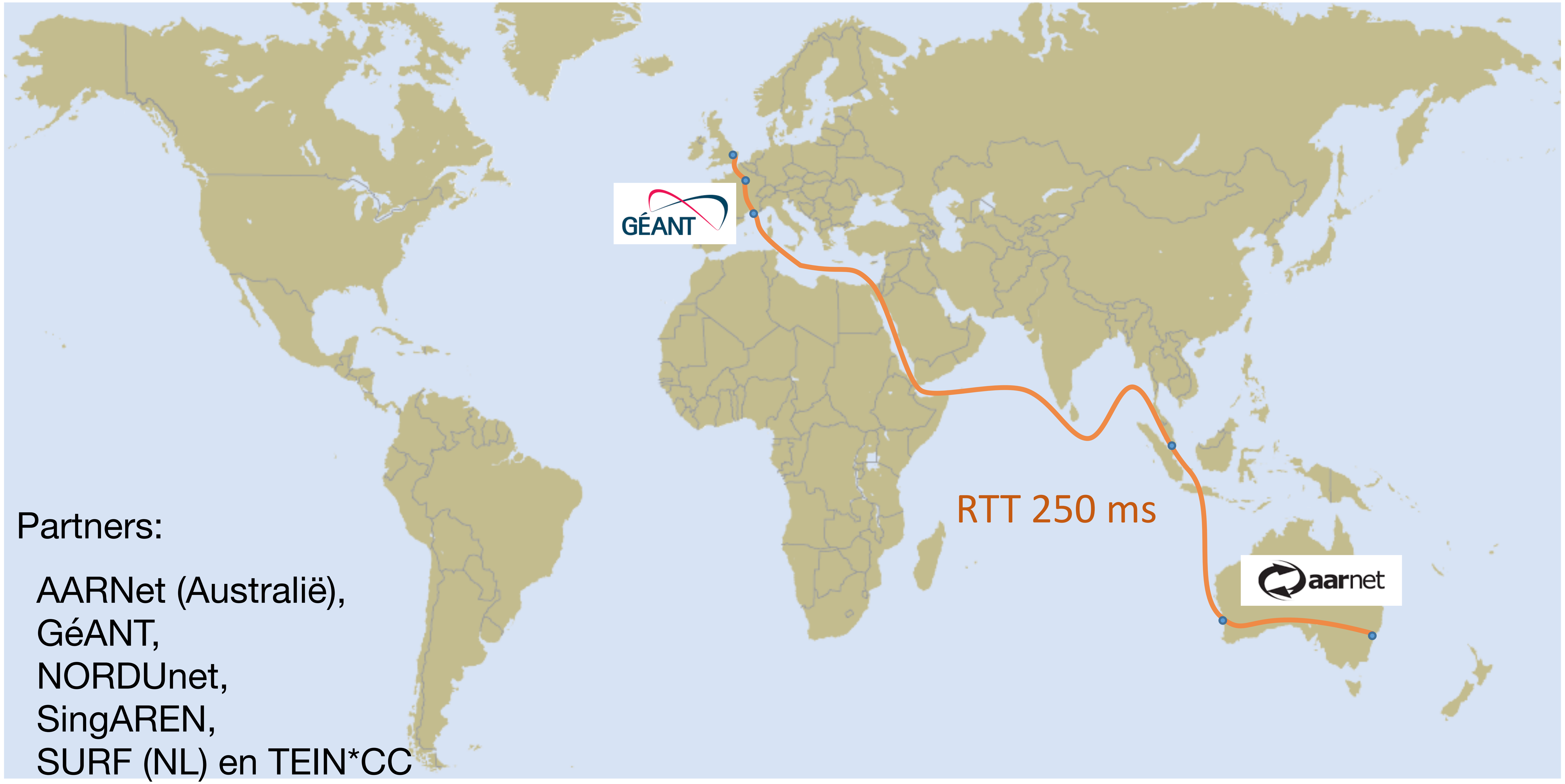


~ theoretical RTT [ms]

	AU	CN	CA	IN	JP	Europe
ZA	280	220	280	200	270	160
AU		120	220	180	100	250
CN			180	90	50	200
CA				220	150	100
IN					200	170
JP						230

The 100 Gigabit Path between AARNet Canberra and London using CAE-1

Example flow



Partners:

AARNet (Australië),
GéANT,
NORDUnet,
SingAREN,
SURF (NL) en TEIN*CC

Thanks to Karl Meyer

Low-level testing

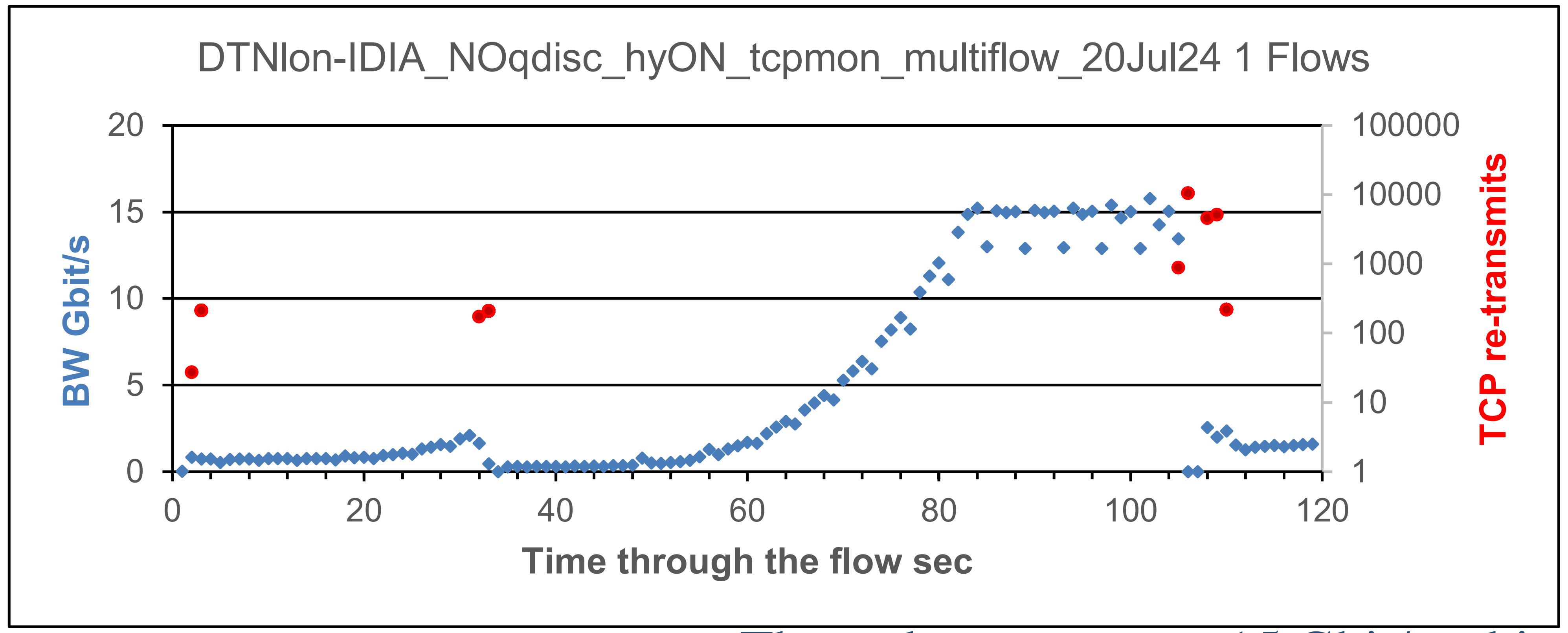
- Activities ongoing on low-level testing
  - Understand of tuning of core parameters on flow throughput
- Delay Bandwidth Product for 20 Gbit/s 1125 MBytes
  - Packet pacing;
  - hystart (None, ON)

•*Aggregate throughput for multiple concurrent TCP flows for 60 sec with tcpmon.*
•*TCP buffer size scan set to Autotune*
•*Set tcpmon affinity to the "good" cores on AARnetCBR & CamDTN1*
•*Queue discipline fq with maxrate 30 Gbit*

•Bandwidth
- Single flow ave 15.3 Gbit/s after slow start
- 15 Gbit/s after 7 sec
- Scaling good 1 flow to 4 flows 52.5 Gbit/s

flows	1	2	3	4
BW Gbit/s	14	23.2	42	52.5

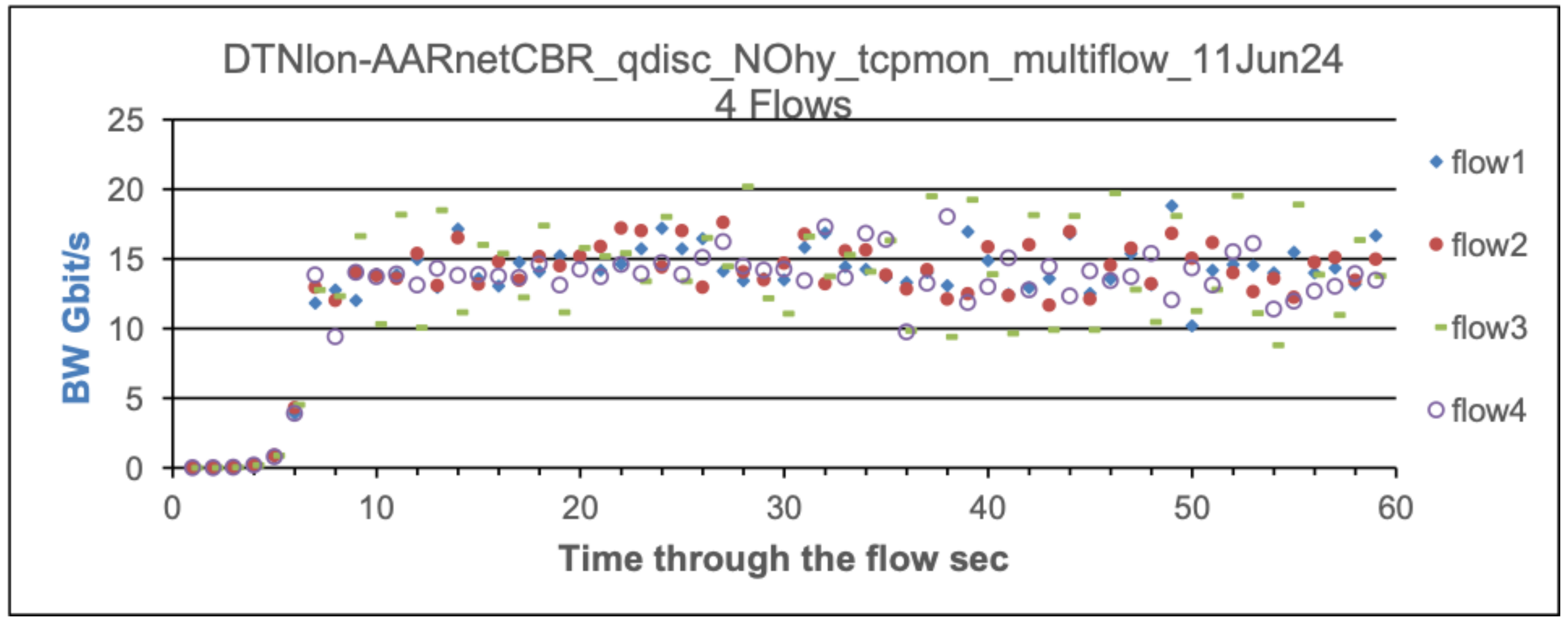
No Packet Pacing Standard cubic hystart ON



•Throughput grows to 15 Gbit/s taking 50s
Although tuned line-speed packet bursts cause loss in the network

Packet Pacing qdisc fq 30gbit NO hystart

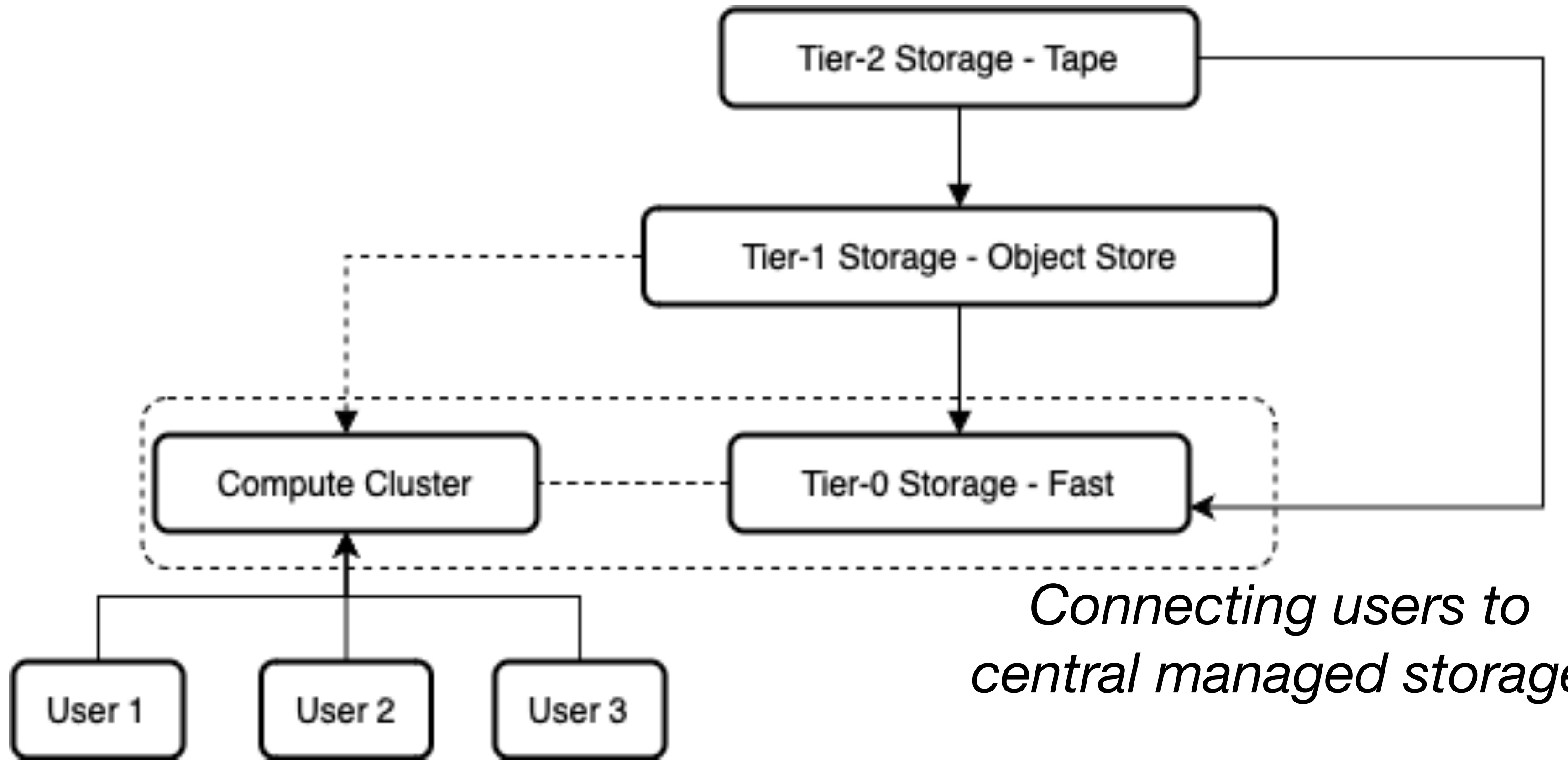
Individual flows well behaved



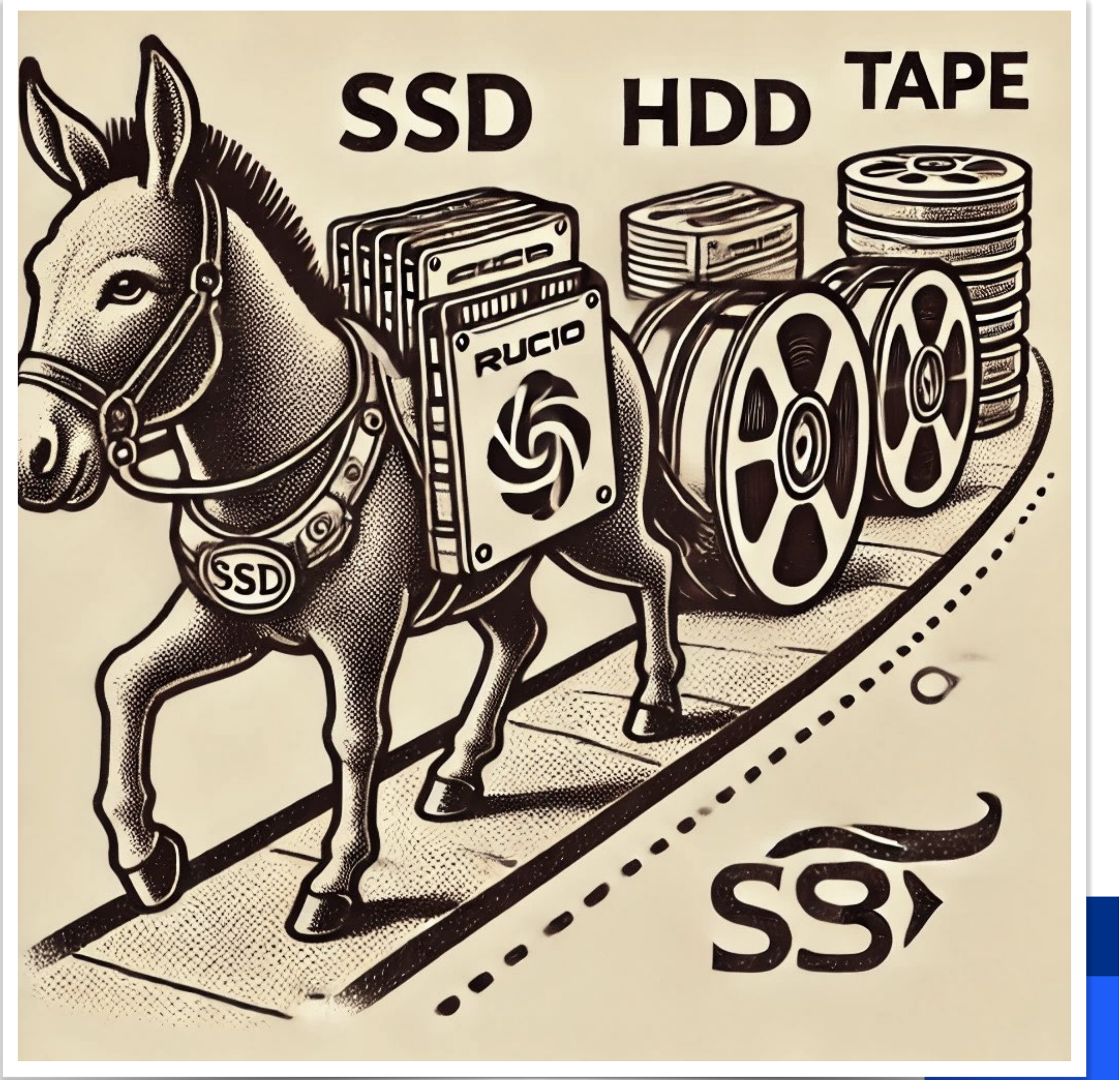
Slower climb in rate due to pacing & hystart putting cubic into congestion avoidance
Use of pacing reduces the number ot re-transmitted segments

Open Questions: Data placement / access

- Many Astronomy codes assume POSIX-like filesystems
  - Scalability to EB scale with POSIX not likely …
- S3 / object store for scalability
  - Already included within some designs (e.g. Vault with CANFAR)
  - Benchmarking tests with S3 fuse ongoing.
  - TPC S3 <–> S3 ?
    - Operational experience with XRootD S3 plugin(s)?
    - FTS streaming, …
      - How to design Object Stores into Rucio / FTS / SEs for scalable and interoperable solutions
- Data placement for Compute workloads to be refined:
  - Staging may be too costly for size of datasets
  - R/O to RSE storage; Management of permissions/ groups / embargoed data
  - Direct streaming to local / remote storage (performance?)
  - Pre-placement strategies and data reductions (e.g. cutouts)

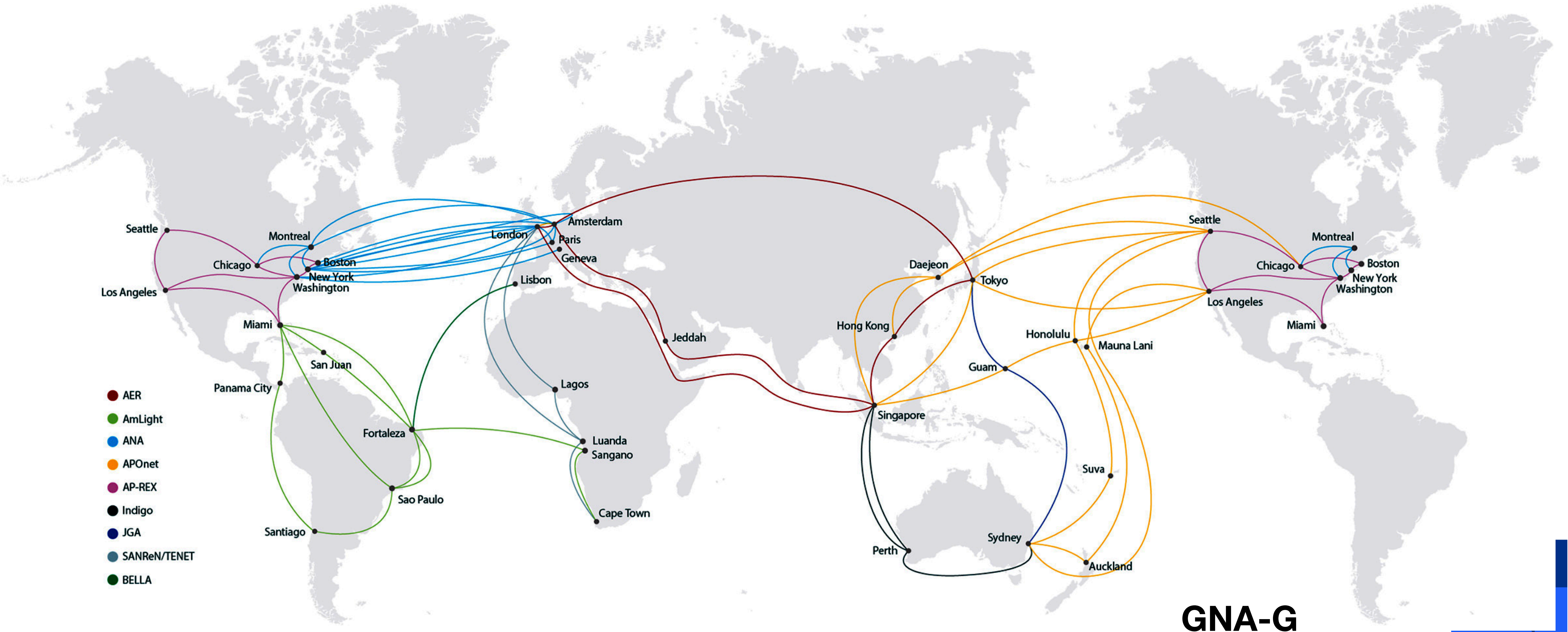


Connecting users to central managed storage



Open Questions (2)

- My personal (biased) thoughts:
  - Not necessarily limited to FTS / XRootD
- Data files will be larger (TBD)
- Round-trip times are larger (especially from Data source to primary sites).
  - Dealing with failures without total file retransmission
  - Partial file transfers (.e.g Cutouts)?
- In-flight checksumming
- Per-file throughput speeds
- Overall throughput
- Tokens:
  - File path hierarchy considerations, common namespaces



GNA-G

Summary

- The SKA project is a monumental effort combining global expertise, computational power, and scientific ambition to unlock new discoveries in the universe
  - Keen to share and engage with other communities to design for the next level of data intensive science
  - **Community engagement and participation**

- **Rucio / FTS / XRootD** (+other SE technologies) for **initial DM verification**
  - To understand POSIX access vs S3 (object store) and tensions between application code and storage scalability
  - **Embargoed proprietary data**, dynamic group structures, public data access and multi-wavelength analysis add much complexity
  - User experience will be through Science Gateways (aka analysis facilities).

- **O(500–700) PB** of data ingestion each year (from AA4) from SKA-MID and SKA-LOW sites (combined
  - **50 year** expected lifespan of SKA
  - Required excellence in User Experience for interactive analysis.
  - Challenges on the **scale of HL-LHC**
- Aiming to deploy services via GitOps style approaches

- **SRCNet v0.1** marks a pivotal shift from development, to prototyping and operational readiness
  - Data lifecycle tests to demonstrate the architecture and test functionality
  - Understanding of data sizes / distribution and high-latency data movement

Backup

www.uksrc.org

Joint mailing list: STFC UK SKA Observatory Science Committee
UKSKA-SCIENCECOMMUNITY@JISCMAIL.AC.UK

uk-ska-regional-centre-uksrc

@UK_SKARC

UNIVERSITY OF CAMBRIDGE

THE UNIVERSITY of EDINBURGH

Durham University

University of Hertfordshire UH

MANCHESTER 1824 The University of Manchester

UCL

UKRI Science and Technology Facilities Council | Scientific Computing


Country-wise SRCNet Storage Requirements Over Time
Storage (PB)
0
200
400
600
800
1000
SRCNet v0.1
SRCNet v0.2
SRCNet v0.3
SRCNet v1.0b
SRCNet v1.0
Milestone
UK
South Africa
Australia
China
Canada
Italy
India
France
Netherlands
Japan
Spain
Portugal
Switzerland
Sweden
South Korea
Germany


Country-wise SRCNet Computing Requirements Over Time
Computing (PFLOPS)
0
5
10
15
20
25
30
35
SRCNet v0.1
SRCNet v0.2
SRCNet v0.3
SRCNet v1.0b
SRCNet v1.0
Milestone
UK
South Africa
Australia
China
Canada
Italy
India
France
Netherlands
Japan
Spain
Portugal
Switzerland
Sweden
South Korea
Germany

Architecture Layers