The EISCAT_3D project:

Nordic FAIR Data Challenge

Introduction

- NeIC
- EISCAT
- EISCAT_3D
- E3DDS
- EISCAT_3D FAIR Data

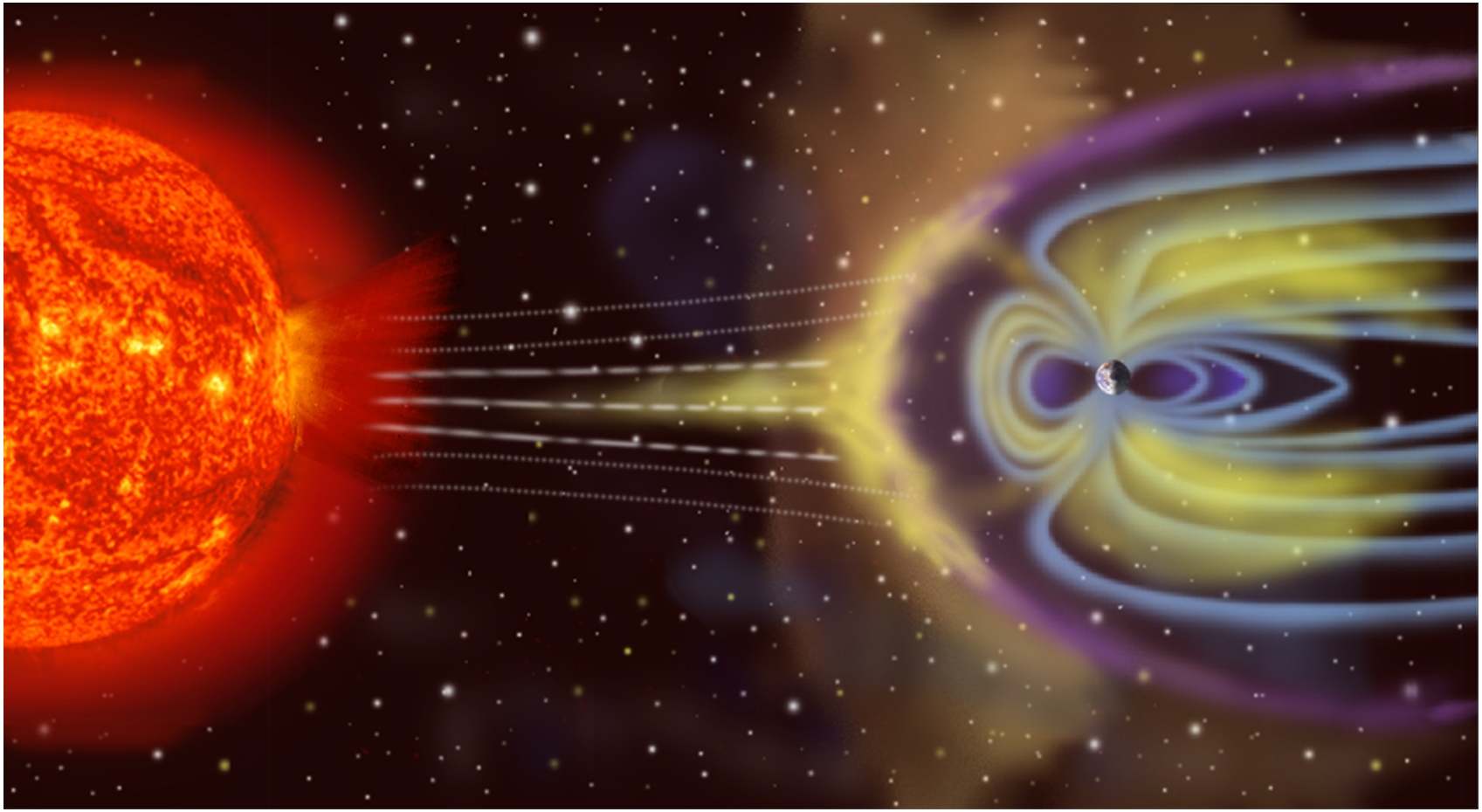
Nordic e-Infrastructure Collaboration (NeIC)

“The NeIC facilitates the development and operation of advanced IT tools and services in areas of importance to Nordic researchers.”

- **Pool Competencies**
  - Map skills, identify and prioritise needs.
- **Share Resources**
  - Launch pilot projects to establish ambition levels
- **Secure Long-Term Funding**
  - Map funding sources, adapt to unaligned funding periods
- **Strengthen Stakeholder Dialogue**
  - Map stakeholders and partnerships

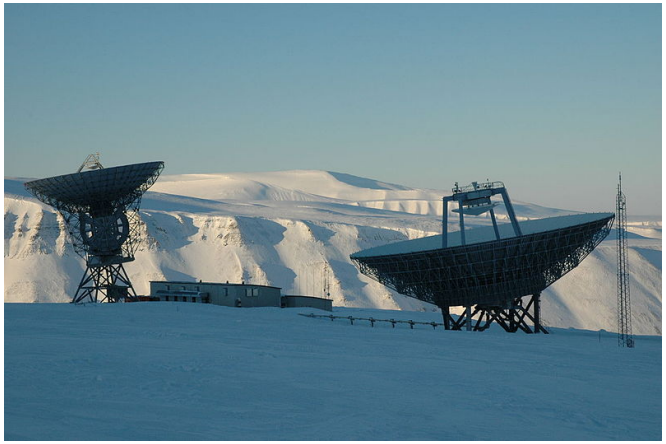
Nordic e-Infrastructure Collaboration (NeIC)

Solar wind and magnetosphere



Direct and indirect Sun-Earth connection through radiation, charged particles and magnetic fields
EISCAT: Studying interactions in the auroral ionosphere and magnetospheric cusp regions

Current EISCAT radars



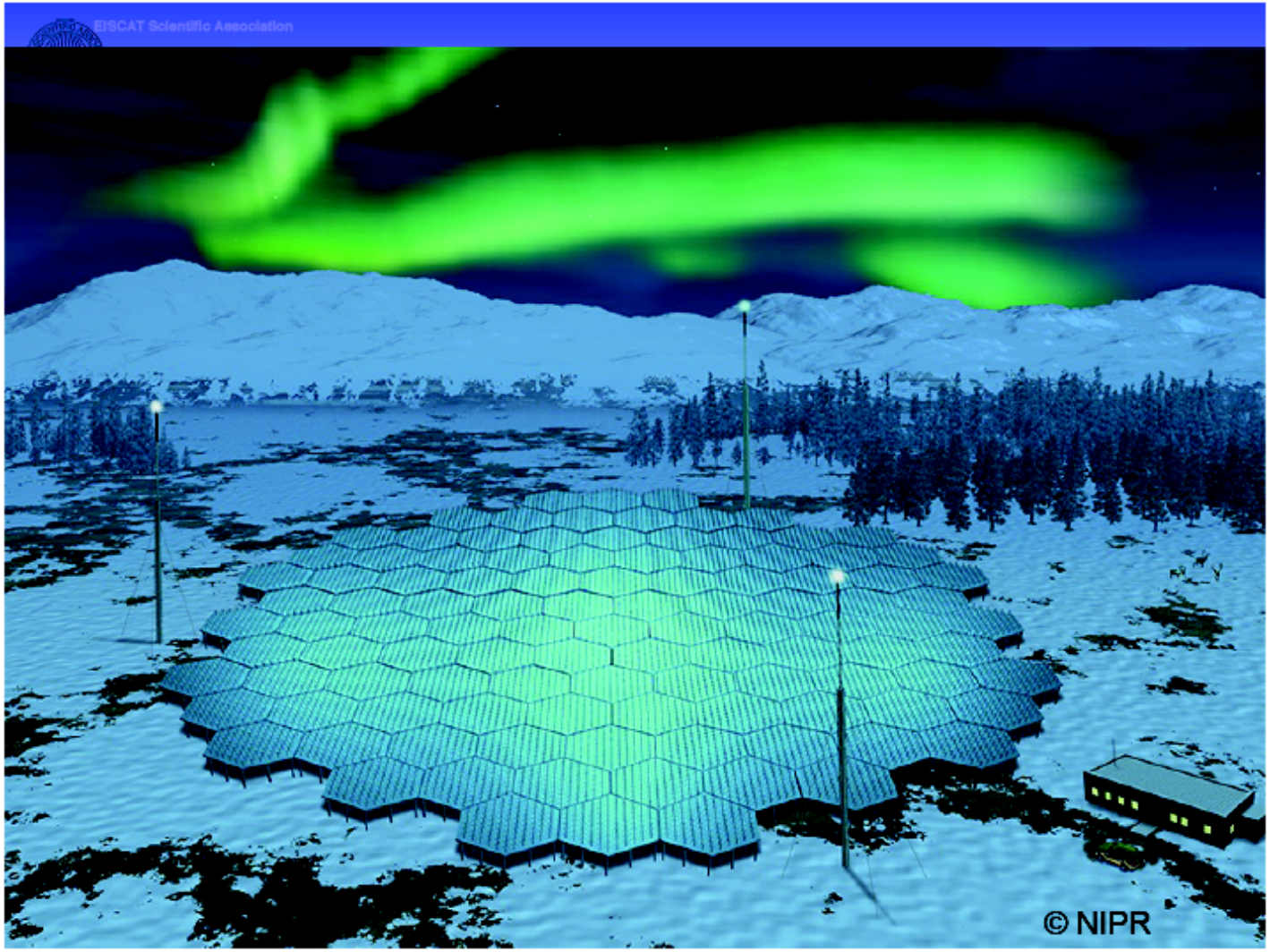
- Campaign operation, fixed schedules.
- Oldest system since 1981.
- **Dataset since 1981 $<$ 100 TB.**

NASA via Wikimedia Commons, EISCAT

European **I**ncoherent **Scat**ter Scientific Association (**EISCAT**)

`https://www.eiscat.se/`

EISCAT_3D Project



109 Arrays of 91 antennas. VHF 233 MHz. Solid state amplifiers.
Up to 100 simultaneous beams. 60° zenith angle

https://www.eiscat.se/about/eiscat3d7/

EISCAT_3D Project

Phase 1: One core station, two remote receiver sites.
Remote control and data access

EISCAT_3D Project

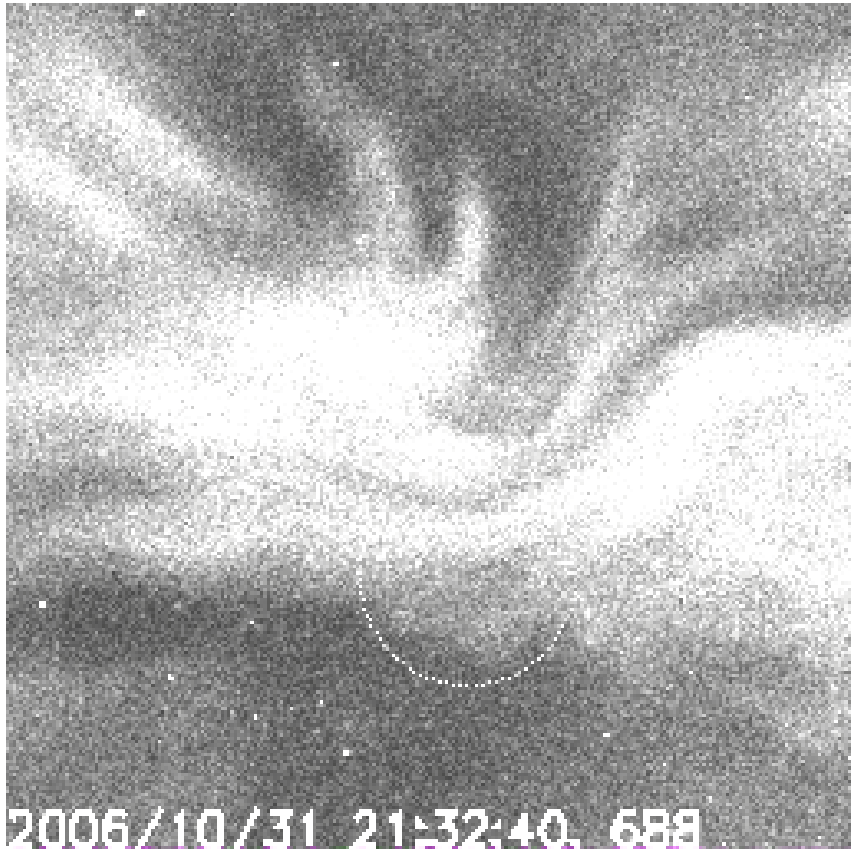
https://heinselslug.smugmug.com/Professional/EISCAT/2017-06-Ramfjordmoen/

A “typical” aurora...

EISCAT_3D Kick-off, Tromsø, September 2017

Courtesy Thomas Ulich, SGO, Finland

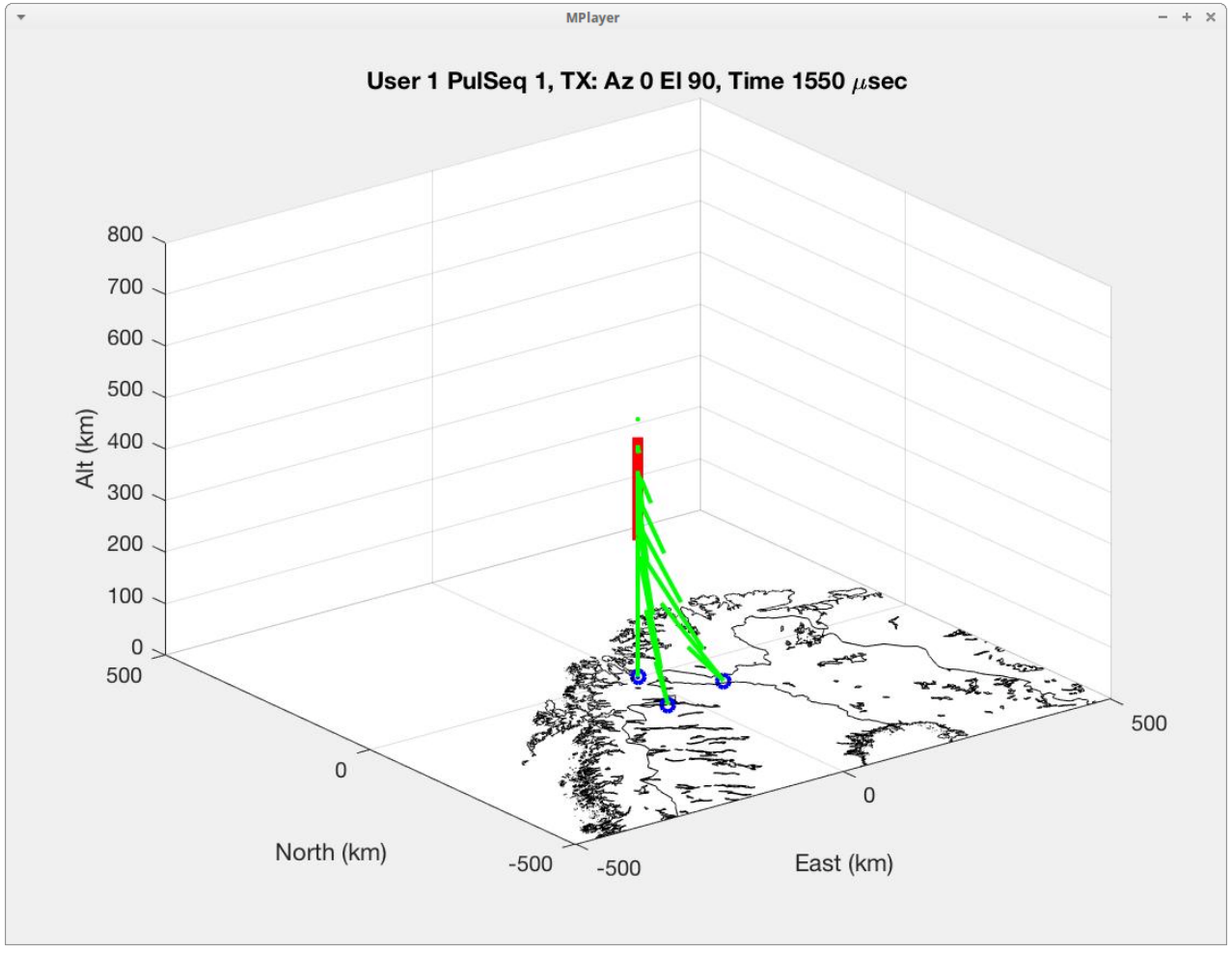
Why EISCAT_3D?



ASK 3×3 degrees, EISCAT Tromsø site 31 Oct 2006
Courtesy Hanna Sundberg (formerly Dahlgren), KTH, Stockholm,
Now at Swedish Defence Research Agency

Instrument now at EISCAT Svalbard site `http://ask1.esr.eiscat.no/`

EISCAT_3D Operation

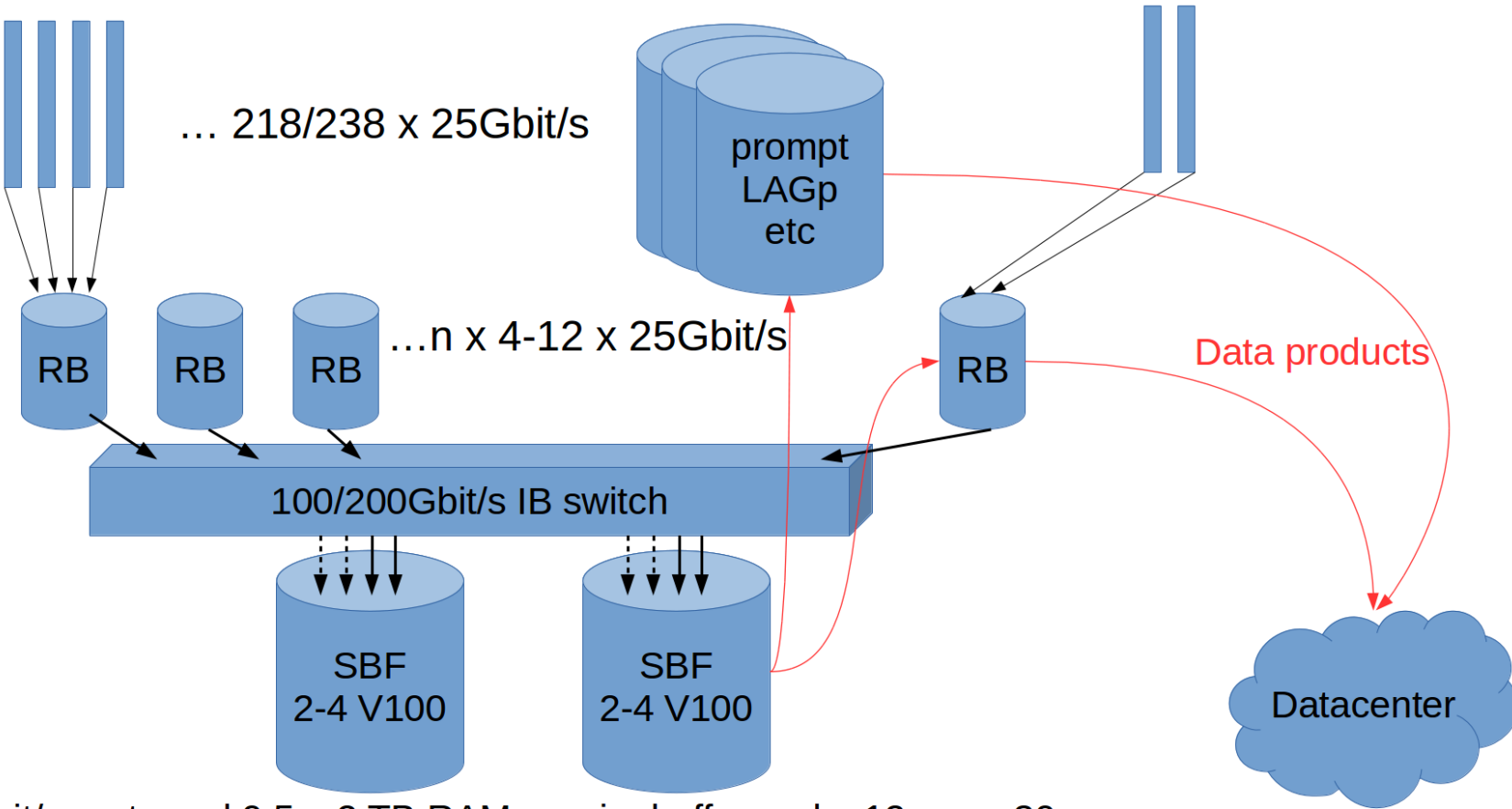


Volume imaging. Maximum data rate after beamforming $>$ 50 Gb/s

`https://www.eiscat.se/about/eiscat3d/eiscat_3d-operation-illustration/`

EISCAT_3D Online Computing

GPU beamforming with Ringbuffer



4-12 25Gbit/s ports and 0.5 – 2 TB RAM per ringbuffer node, 12 < n < 30
Loss of a single RB node is “acceptable degradation”
Second stage Beam Former computing on GPU

See: `https://wiki.neic.no/wiki/EISCAT_3D_Data_Solutions#Deliverable_1`

EISCAT Data

- EISCAT_3D is a project of the EISCAT Scientific Association, therefore:
- Governed by the EISCAT rules [1]
- EISCAT data policy governed by blue book [2](2015)
- Low-level data from each experiment embargoed for defined periods (typically 1 year for EISCAT member carrying out experiment, 3 years within EISCAT membership) .
- Data should be archived at two redundant Data Centres.
- Analysis of data either close to Data Centres or “spare” on-site computing.

[1] https://www.eiscat.se/scientist/document/governing-rules/

[2] Page 39 onwards of https://www.eiscat.se/wp-content/uploads/2017/06/BlueBook_Edition2015.pdf

EISCAT Data Levels

Level	Type	Produced by	Storage	Format	Rate
1a	Ring buffer data	1st stage beam former	4 months*	UDP stream/ HDF5	$\leq$ 4 **Tb/s** [3]
1b	Beam-formed data	2nd stage beam former	4 months*	HDF5	**64 Gb/s**
2	Time integrated correlated data	All sites	Archived	HDF5	
3a	Physical parameters	All sites	Archived	HDF5	
3b	3D-voxel parameters	Operations centre	Archived	HDF5	$\approx$ 1 **Gb/s**
4	Derived geophysical parameters	Users	Users	Publications etc	

- The EISCAT_3D Data Centres will receive, serve and archive all data at levels 2 and 3.
- Data used in research should be given Persistent Identifiers (PIDs) according to a common standard such as DOI, DataCite, or similar, to be unambiguously citable in publications.
- A 4 months period is selected as this is the estimated time required to perform a “real-time” analysis on low-level data.
- A portion of the level 1 data will also be archived permanently, on the order of 1% of the level 1 data rate, e.g. one beam per site and/or bandwidth-limited data.

[3] At full 52 Msample/s from each FSRU

EISCAT FAIR Data

Strong recommendation to follow ENVRI-FAIR principles
FAIR = **F**indable. **A**ccessible, **I**nteroperable, **R**eusable

- Will use PIDs or similar.
- Data stored in standard format (e.g. HDF5).
- Data model using standard vocabulary.
- Standard file and metadata catalogues.
- Standard data management system e.g. from LHC projects.
- Metadata always accessible and associated with provenance.
- AAI (EGI checkin implemented).

References

Open Information Linking for Environmental science:

`http://oil-e.net/ontology/`

ENVRI Reference Model:

`https://wiki.envri.eu/display/EC/ENVRI+Reference+Model`

EISCAT_3D Data Model:

`https://drive.google.com/open?id=0B83vmItCJRAYMktBdEoxdmlyMjA`

Thank you
Questions?

https://wiki.neic.no/wiki/EISCAT_3D_Data_Solutions
https://www.eiscat.se/eiscat3d/

Produced with LaTeX
Updated April 4, 2019