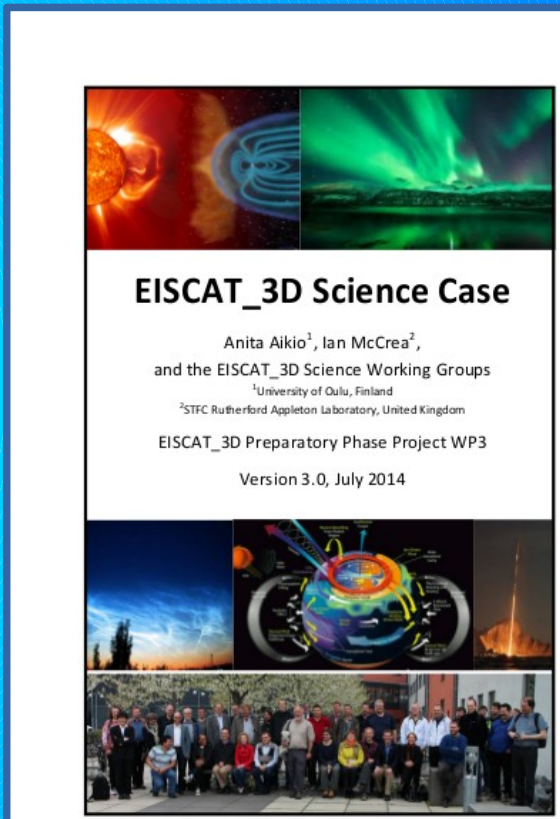
Roadmap for incoherent scatter radar operations

Ingemar Häggström
EISCAT HQ

Why ISR observations?

- ISR technique is one of the most powerful methods for detailed measurements of the ionosphere
  - Electron density
  - Electron and ion temperatures
  - Bulk flow
- Observing connection between Sun and Earth
  - Solar-terrestrial physics
  - Space weather



EISCAT_3D Science Case

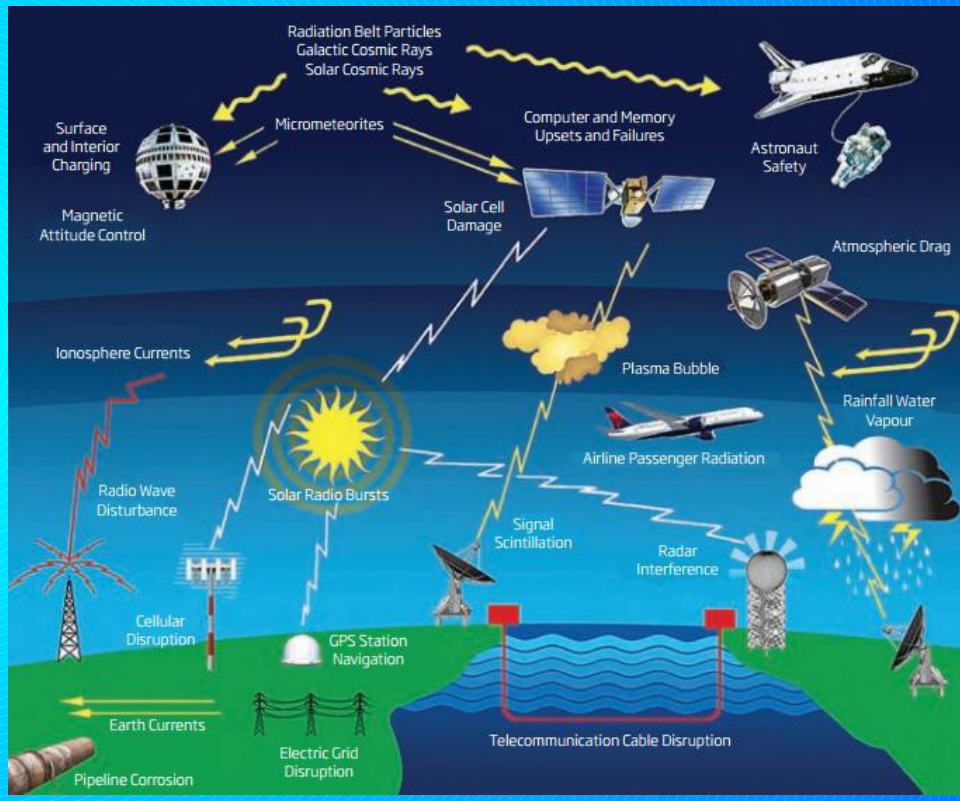
Anita Aikio[1], Ian McCrea[2],
and the EISCAT_3D Science Working Groups
[1]University of Oulu, Finland
[2]STFC Rutherford Appleton Laboratory, United Kingdom

EISCAT_3D Preparatory Phase Project WP3

Version 3.0, July 2014

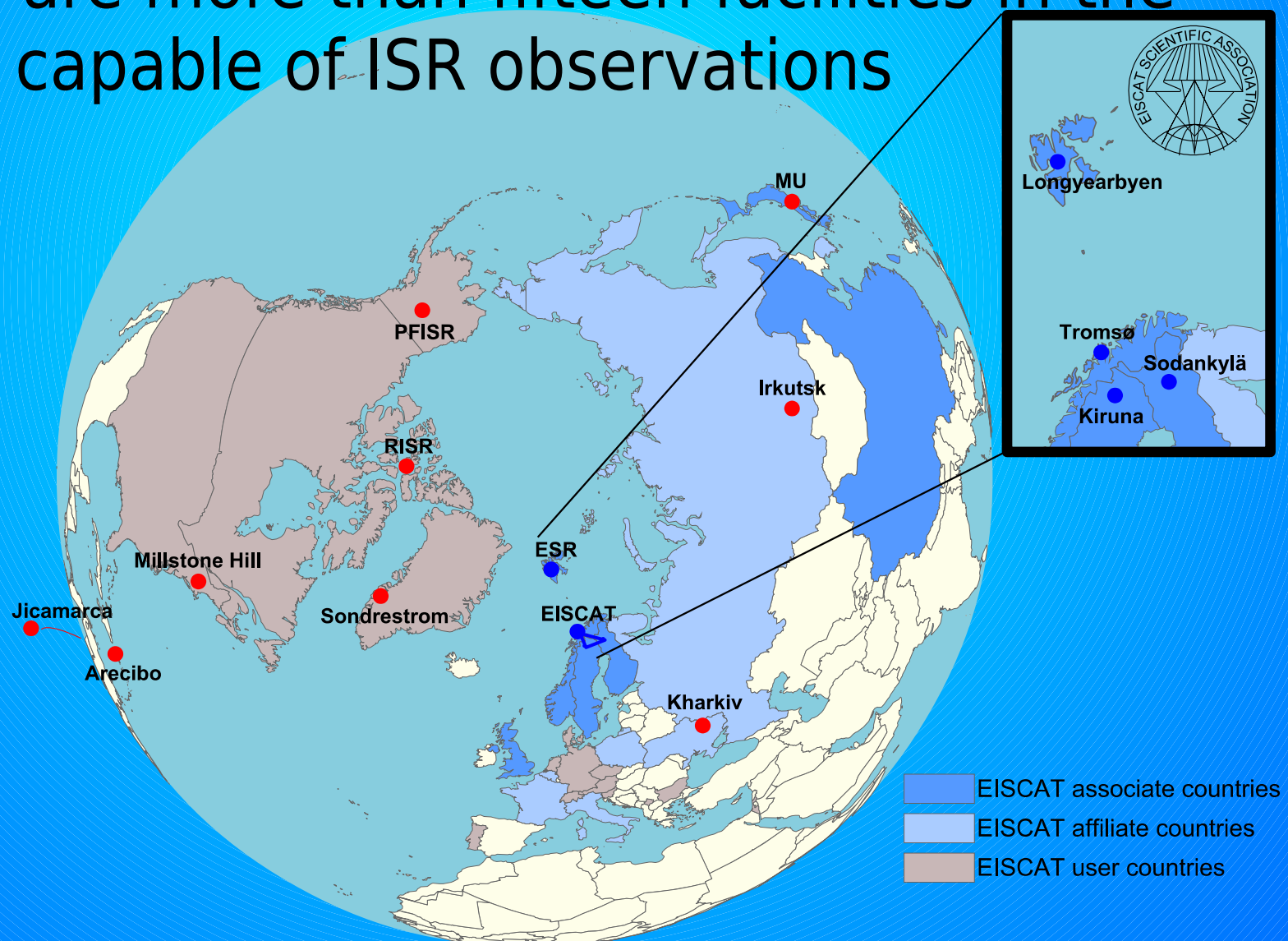
Trends in ISR research

- Generally increasing interest in environmental sciences
  - international collaborations
  - Global coverage
- Space weather prediction
  - Improved dependence on modelling of effects from space weather events
- Continuous observations
  - Moving from campaigns towards observatories
- Larger data volumes and quicker data access requirements



ISR systems in the world

- There are more than fifteen facilities in the world capable of ISR observations



Some ISR facilities

EISCAT
UHF

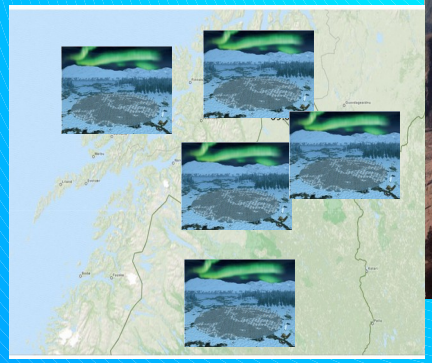
EISCAT Svalbard
Radar

Millstone Hill

Arecibo

PFISR

EISCAT
VHF



EISCAT_3D

Jicamarca

Sondrestrom

RISR

The need for a roadmap

- Large variations in equipment and organisation
  - Harmonisation needed for efficient collaboration
- Increasingly more data intensive
  - Common strategy for data formats and access
  - Must be in line with capabilities and standards set by e-infrastructure providers
- EISCAT_3D to be constructed “soon”
  - Defining system for the future
  - Ready for future standards

Areas for collaborative efforts

- Identified areas where collaborative efforts suitable and of great value
  - Definition of data levels
  - Data format
  - Access
  - Standard & non-standard operations
  - Training and education
  - Outreach

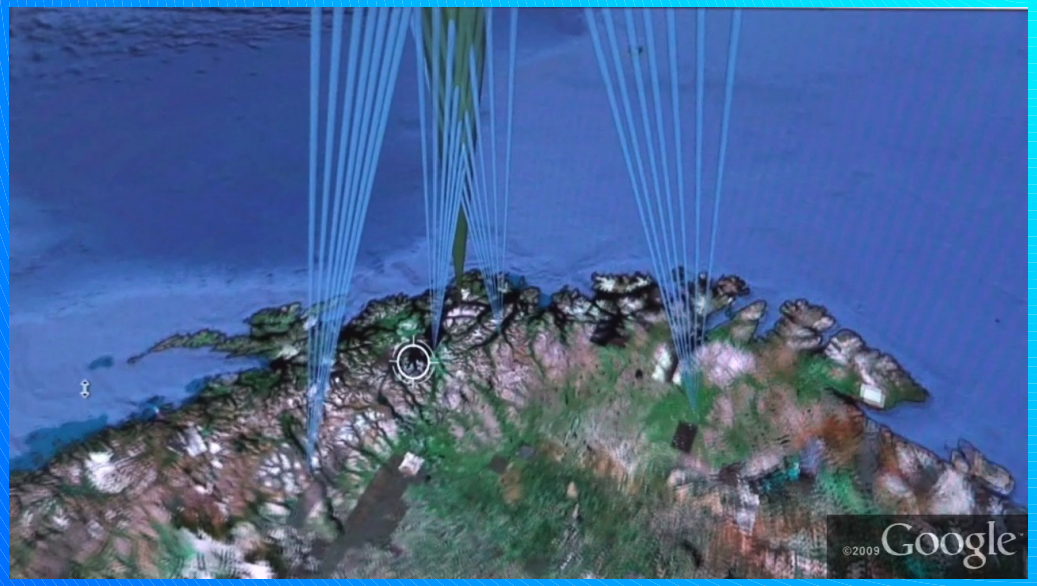
EISCAT_3D

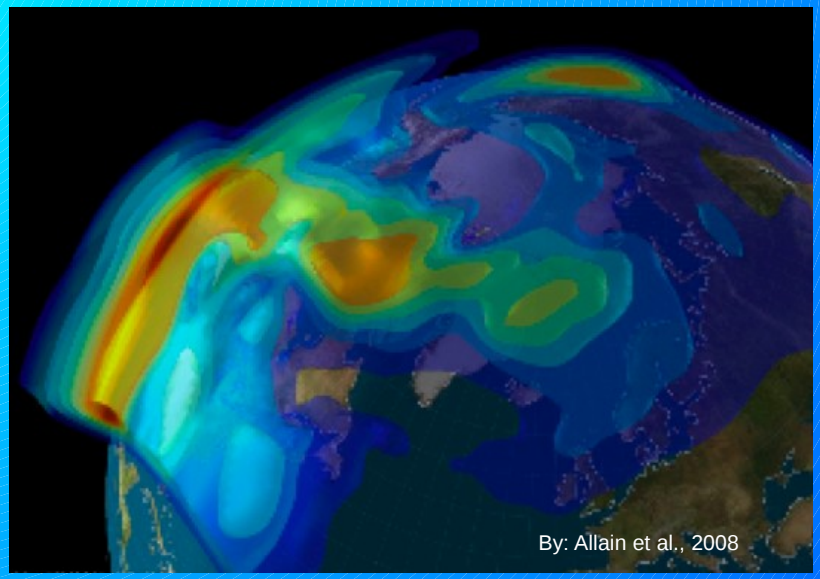
EISCAT_3D will be a volumetric vector-imaging radar for studying the geospace environment

It will support a measurements of the space environment - atmosphere coupling at the southern edges of the polar vortex and the auroral oval.

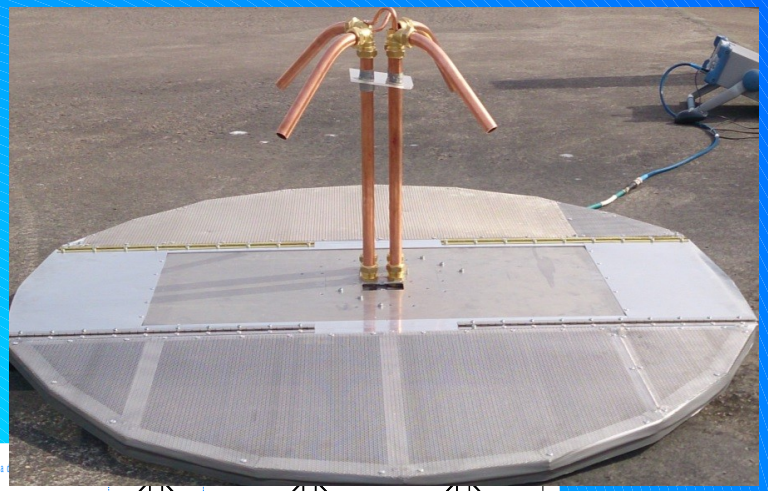
An extremely versatile and largely software-defined instrument

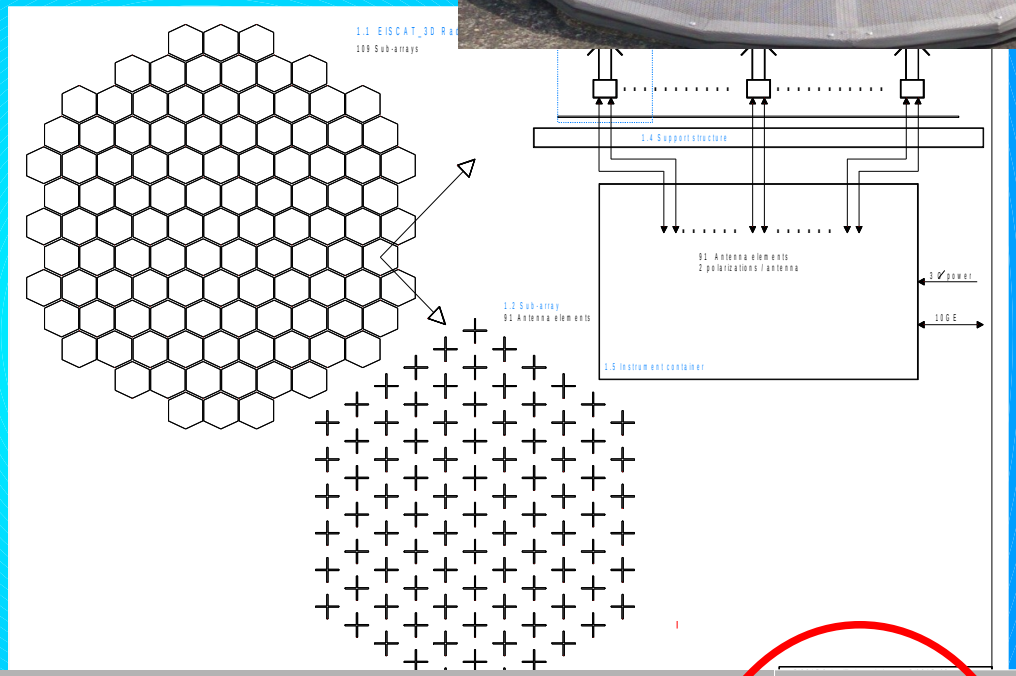
Easy expansion to new fields





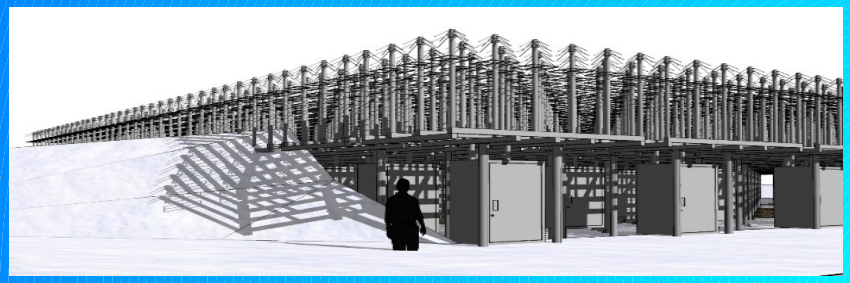
10000 antennas at each radar site!

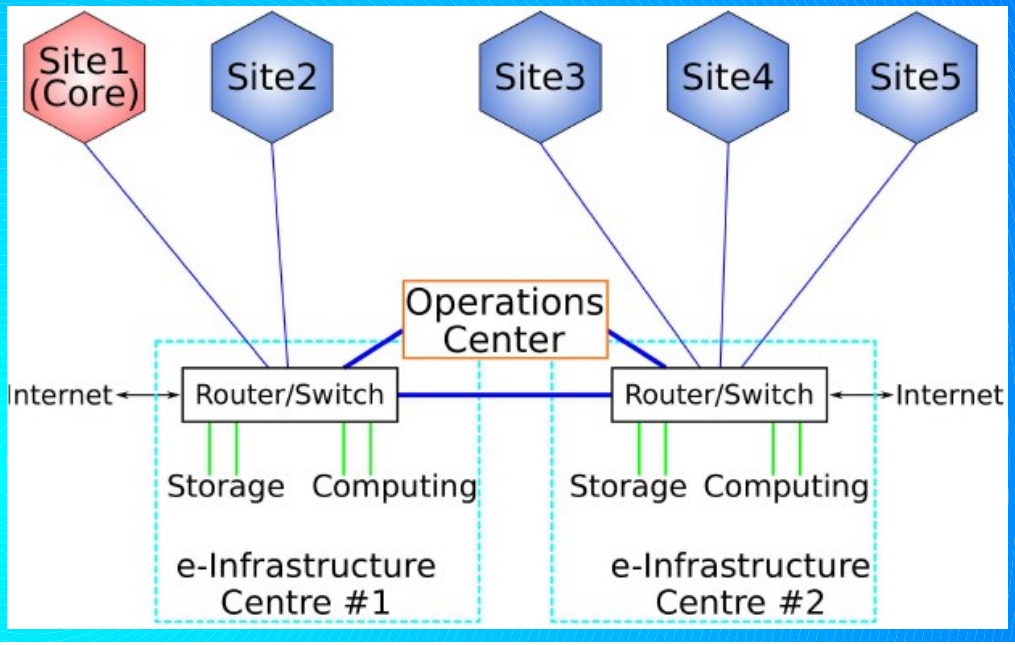




	Transmitters	Antennas	Beams	Sites	Volumes
Mainland	1	1	1	3	1^3
Svalbard	16	2	2	1	$1*2^1$
EISCAT_3D	10000	10000	10*10	5	$10*10^5$

Computing & Storage first 5 years of E3D





	Sites 5x	Operation Centre		Data Centre	
	Real-time	Post-processing		User Access & Archive	
Storage	86 TB RAM 1 PB Disk	20 PB Disk		2 PB Disk	2 x 10 PB Tape (2 PB per year of operation)
Computing	590 Tflops/s (FPGA) 77 Tflops/s (other)	500 Tflops/s		50 Tflops/s	

EISCAT_3D data

- Operations centre
  - 3 sites 54Gb/s
    - 3rd data buffer 20PB
      - 100 days
  - Multistatic analysis
    - 500 Tflop/s

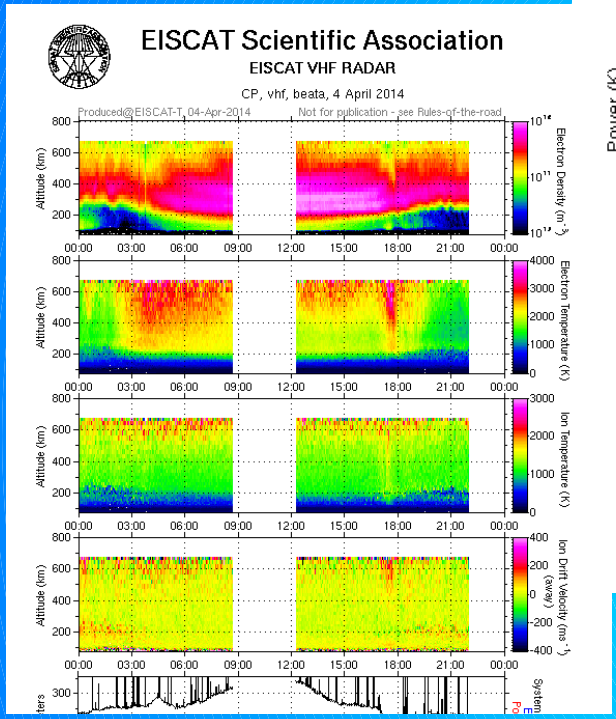
Data levels

- Level 0
  - From ADC, never store
- Level 1, Voltage
  - Reduced BW (filtered, resampled)
  - Sample interval, microseconds
  - Profile interval milliseconds
- Level 2, Spectral data
  - Correlated and integrated time/space)
  - File interval, seconds
- Level 3, Ionospheric parameters
  - integrated (time/site)
  - ~Minute
- Level 4
  - User data, reports (filtered, resampled)

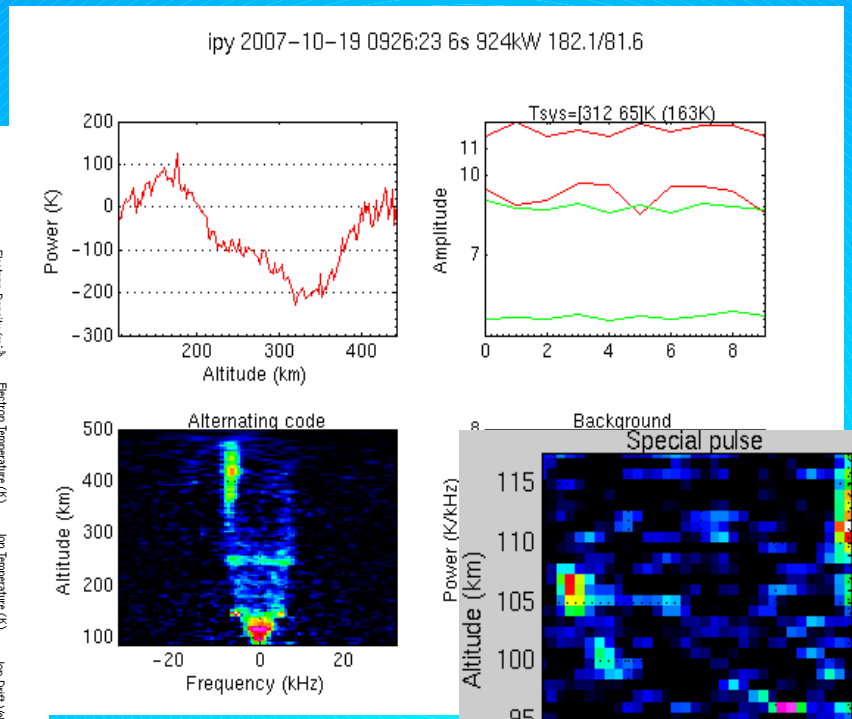
Research with EISCAT

- Historic
  - Campaign/event based
  - Space Physics
- EISCAT_3D
  - Radar to run continuously
  - Environmental science
  - Volumetric data
  - Data mining will be the main source for scientific work

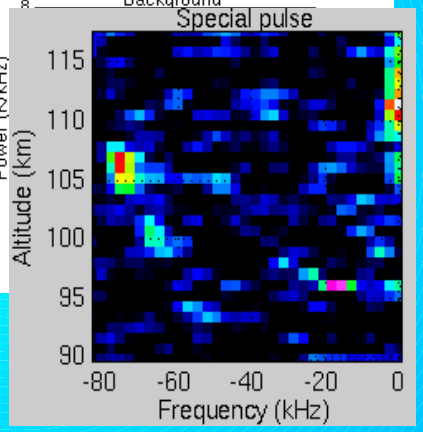
Significant computational resources needed



Level 3



Level 2

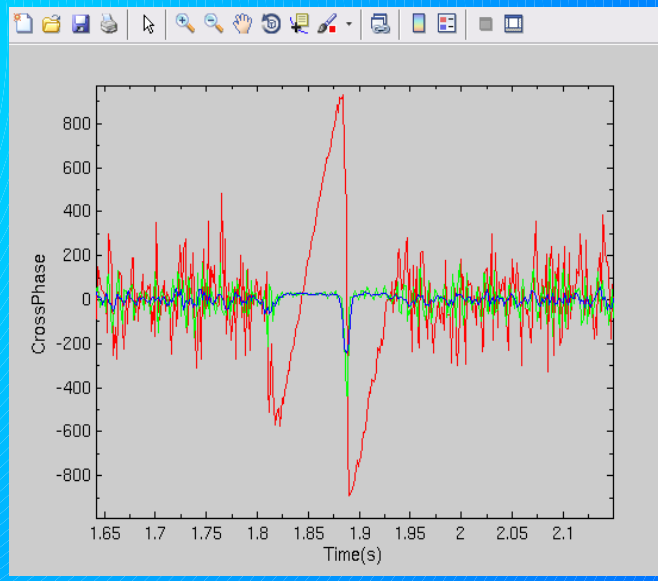


Spatial scales

- Global/regional/km/m/cm

Time scales

- Solar cycle/ y/d/h/m/s/ms-/µs



Level 1

- Metadata search
- Data search & visualization
  - Volumetric data

EGI Engage
EISCAT_3D competence centre

- Participants
  - EISCAT 14M
  - SNIC, Sweden 7M
  - CSC, Finland 7M
- Own funding
  - Uninett Sigma, Norway
  - NeIC, Nordic eInfrastructure Collaboration

EISCAT_3D CC Objectives

- To build common e-Infrastructure to meet the requirements of a big scientific data system such as EISCAT_3D data system
- To demonstrate the developed e-Infrastructure can support the EISCAT science community in their acquisition, curation, access to and processing of the data
- To train data scientists who can explore new approaches to solve problems via new data-centric way of conceptualising, organising and carrying out research activities

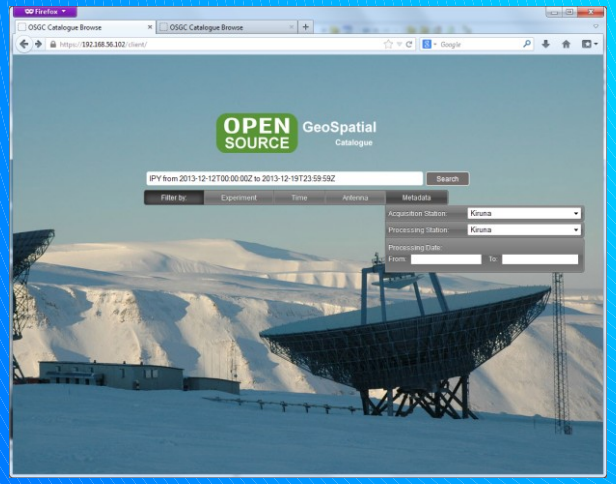
EISCAT_3D CC Tasks

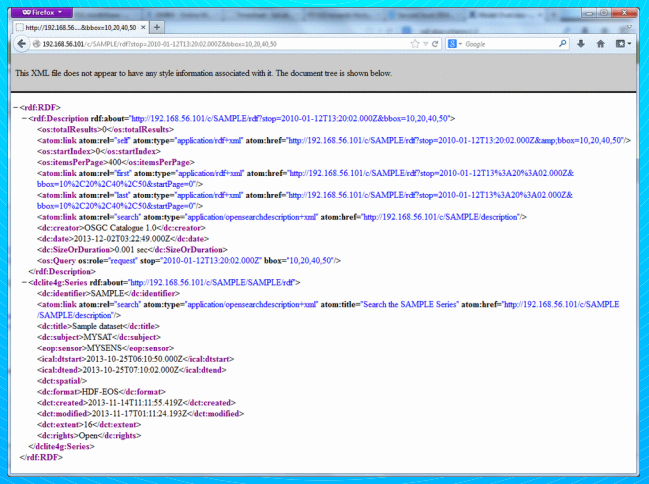
- User Support and Training
- Deploy the portal as a production system
  - For human and machine
- Basic reanalysis within the portal
  - Metadata update
- Use level 3 data as metadata
  - Link the levels
- Exploitation

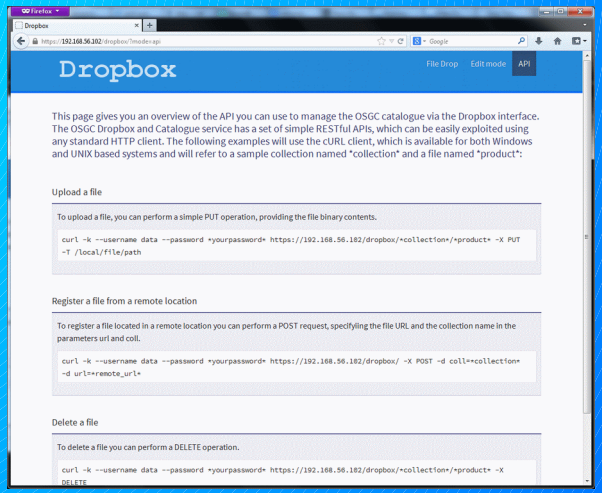
OpenSource Geospatial Catalogue

PaaS is based on the **OpenSource GeoSpatial Catalogue**, which is an:

- OpenSource implementation of OpenSearch and OGC OpenSearch GeoSpatial Extension Catalogue systems.
- Released under GPLv3 on SourceForge (https://sourceforge.net/projects/osgcat/)
- Built in PHP/PostgreSQL

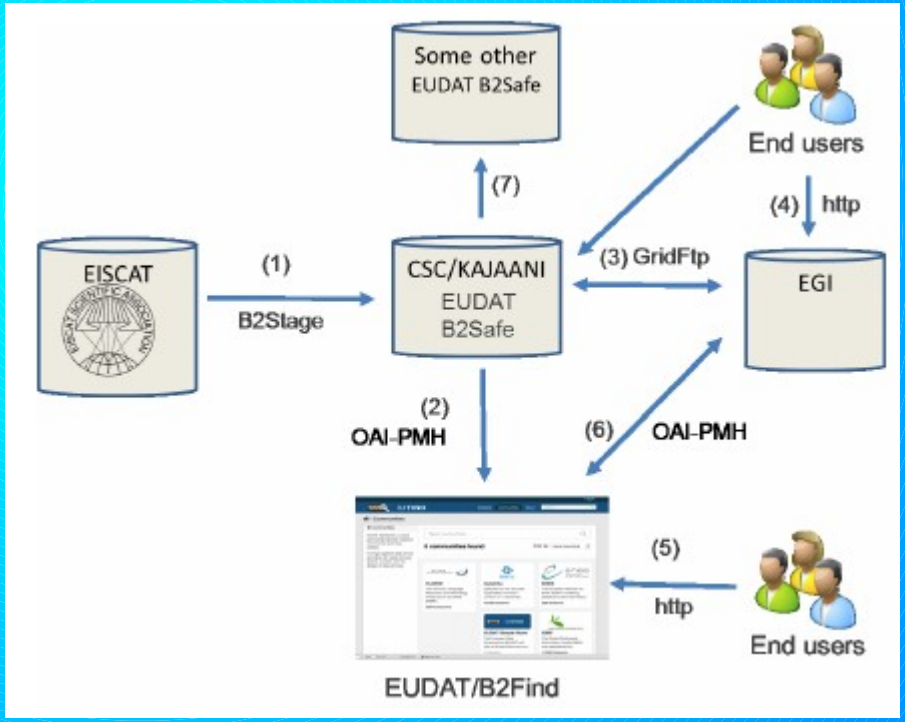






EGI-ENVRI Portal

- Some metadata
  - Radar parameters
  - Open search
- Data storages
  - EISCAT
  - Cloud
  - EUDAT
- Web interface
  - EGI
  - B2Find



MetaData List

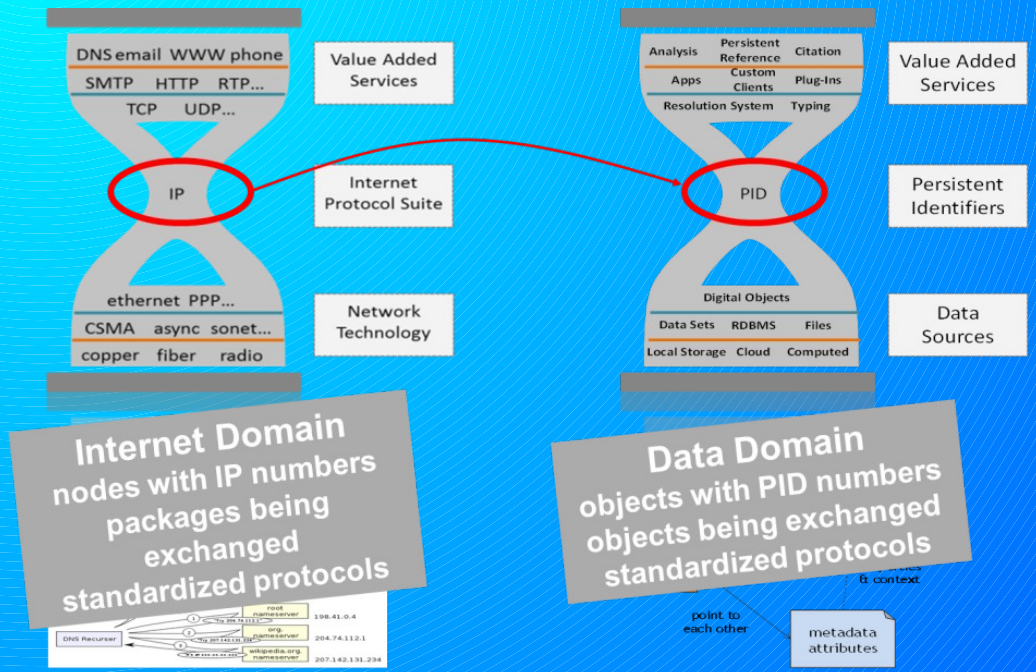
- Complete list
  - Radar parameters (headers in the data files)
  - Experiment configurations (scripts in info dir)
  - Analysis setups (simulations of exps in matlab)
  - Schedule db (experimentors...)
  - SQL db (Owner(s), data sizes)
  - Log files
- E3D
  - Need to identify
    - what is missing
    - the structures of the metadata (volumetric data)

Access

- Combination of
  - EISCAT SQL db info (Country)
    - Now we do 'whois' on the IPs
  - Certificates
    - EISCAT affiliate membership
      - University groups, single users
    - Better logging

Data identifiers

- Identifiers
  - To link the different layers of data
- Tied to time (UTC/TAI)
  - Environmental data
  - Plasma Physics (heating)
    - Event based (like LHC) but in natural environment!
  - Different timescales
    - IP like ID?
      - RDA analogy!



Work done

- Kickoff 11 March
- ~biweekly telecons
- NEIC conference 5 May
- Portal
  - Human/machine interface?
    - (Liferay)
    - Metadata setup
  - Search engine
  - Analyser
  - Packer for download (hdf5)

Thank you!

Search Big Data

To search the big data by specifying the 'patterns' at different levels, to limit the data sizes. For that one need PIDs, to be able to jump between the levels. Identifiers are probably not only time and location.

The data at the higher levels are the metadata for the lower levels.

The patterns are features in 0-4 dimensions.

For zero it means simple values searches. Higher orders, demands operations to be done on the data before comparing with the patterns.

Examples

Give me level 1 data where:

0-D on level3 data: Find Te > 3000K

1-D on level3 data: Find Te > 3000K lasting less than 5 seconds

2-D on level2 data: Find asymmetric spectra with 95% confidence levels

3-D on level2 data: Find asymmetric spectra restriced to 10km altitudes

and so on.

Mainly features following a certain behaviour, in Visual domain

Workflows

Scalable as EISCAT_3D is growing = Robust

Workflows to redo the a going between layers.

Workflows to do unusual operations on data:

Nonstationary spectra, crosspectra, etc.

Workflows for user defined operations on the data.

Visualise big data

Volumetric data, 4D data.

There are soft, volume filling, targets with gradients and possibly higher degrees of variation within the volume.

Hard targets have more distinct borders, normally smaller than the measurement volume.

There is also anisotropic data, with different values at different directions (drifts, non-Maxwellian) giving different spectral responses.