Big data exploration -
A hierarchical visualisation solution for cubic surveys

Görlitz – ADASS Oct 2025

Pierre Fernique
Mark Allen, Matthieu Baumman, Thomas Boch,
Caroline Bot, François-Xavier Pineau

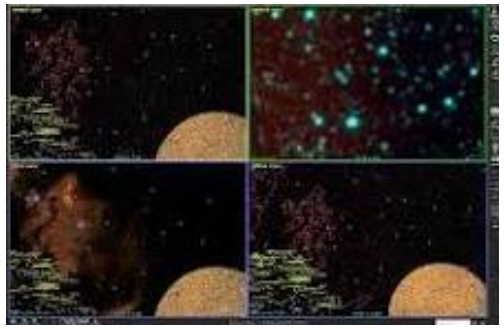
CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

Observatoire astronomique de Strasbourg | ObAS

cnrs

Université de Strasbourg

First challenge: data access/visualisation



Very large image, cube and catalogue data are coming!

- PB-scale data is expected from space missions (e.g. Euclid), and Observatories (e.g. LSST, SKA, ESO +)

***"Access & Visualisation"* involves many different use-cases:**

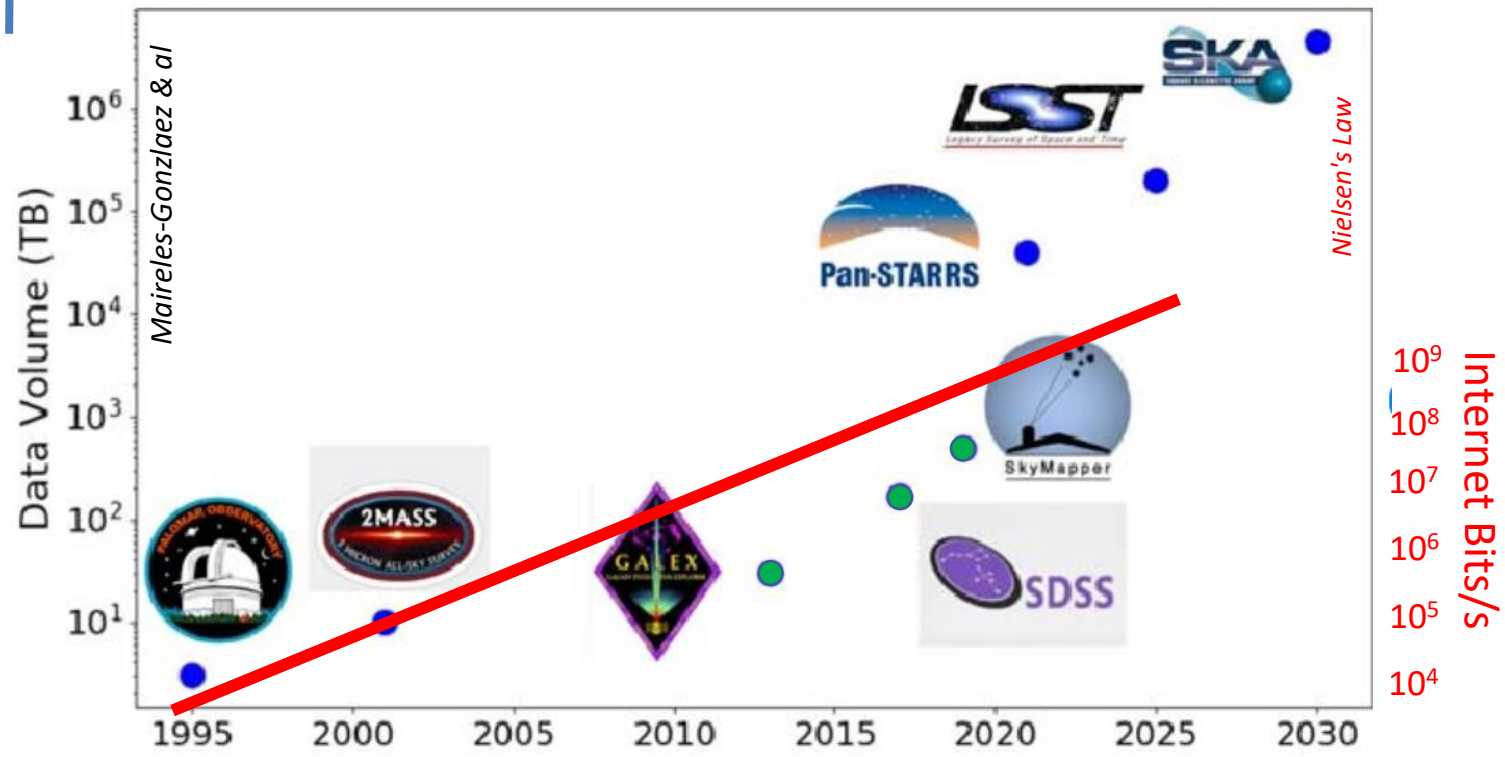
- Exploration of the data on all scales *: Full-sky to individual 'pixels'*
- Detailed analysis, extraction..
- Combination, overlays, colour maps, …

Emphasis in this presentation: ***Visualisation for data discovery***

The dilemma: local or remote data access?

VT100 *TX* *PC* *Web* *Applet* *peer2peer* *Mobile* *VM/cloud/grid* *Notebook* *Remote sessions*

Remote | Download + Local | (Migration +) Remote ?



The key constraint: the **ratio** between **data size** and **network capacity**.

In 2025 – again **too big to be easily moved**!

How to explore/visualise data remotely?

- **Solutions**? Remote session (platform), Remote display (e.g. CARTA), Jupyter notebook, …
- Two major constraints:
  - **Dynamics** - What are the possible interactions? Display reaction time, zoom in/out – the latency can not be removed
  - **Scalability** ? How many simultaneous users are possible? (where is the load (CPU&RAM)? On the server side only?)

- **Alternative: Hierarchical (pre)views**
  - Client - requests only the data that is needed for the current view
  - Server - provides very fast access (data is pre-calculated and organized)

=> **HiPS** : Hierarchical Progressive Survey

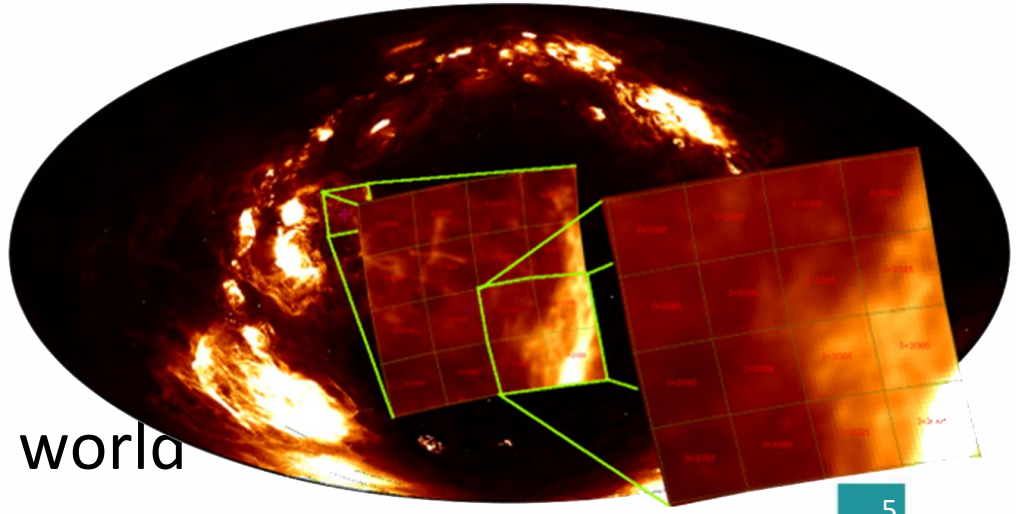
CDS / IVOA approach

HiPS – What is it?

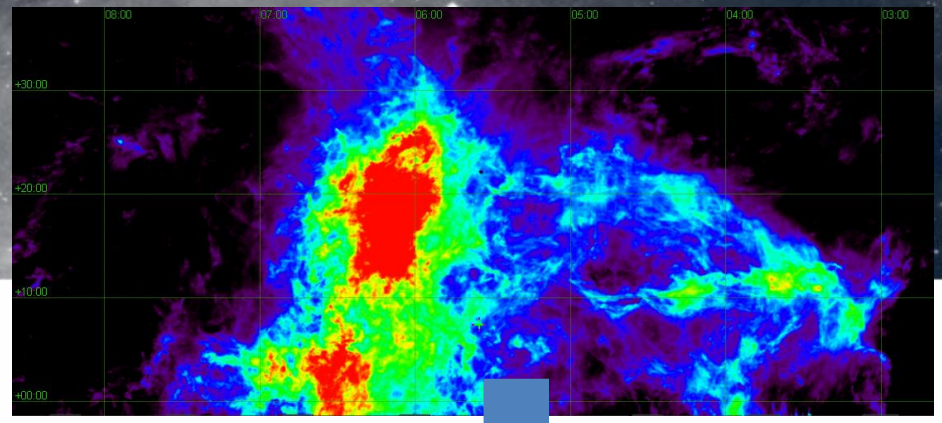
- The **Hierarchical Progressive Survey** method
- Described in **2015A&A…578A.114F**
- **Standardized by IVOA** in 2017

- Makes a sky survey accessible, visualisable and even manipulable, whatever the size of the survey, the quality of the network and the computing power available to the astronomer.

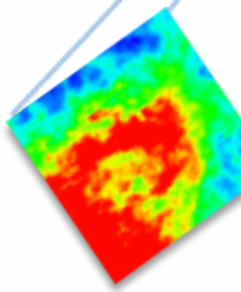
- **Implemented** by scientific, amateur and public visualisation tools and portals:
  Aladin Desktop, Aladin Lite, ESAsky, ESO portal, WWT, Firefly, DIGISTAR, RSACosmos, Stellarium, Astrobrowser,ypiAladin, glue...
- **1400 surveys/missions** already hipsilized, available through **25 HiPS nodes** all over the world



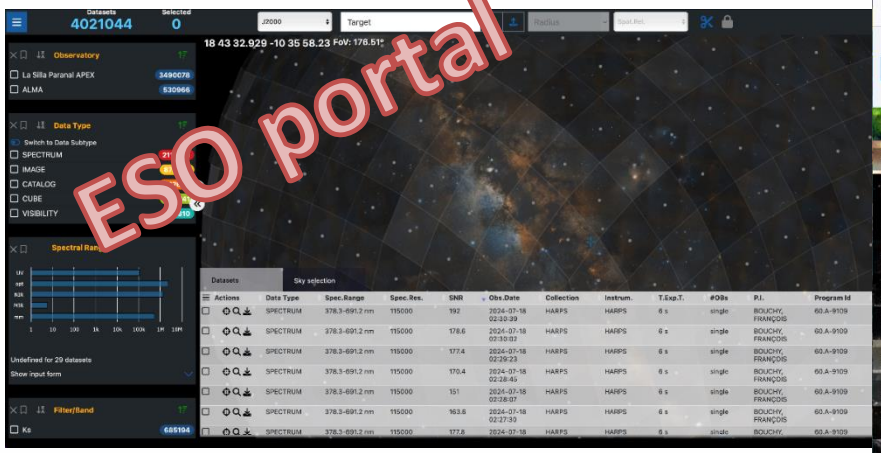
HiPS principles

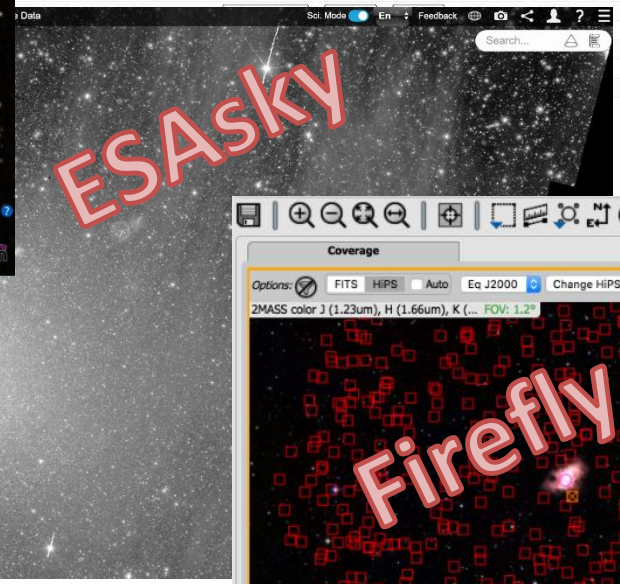


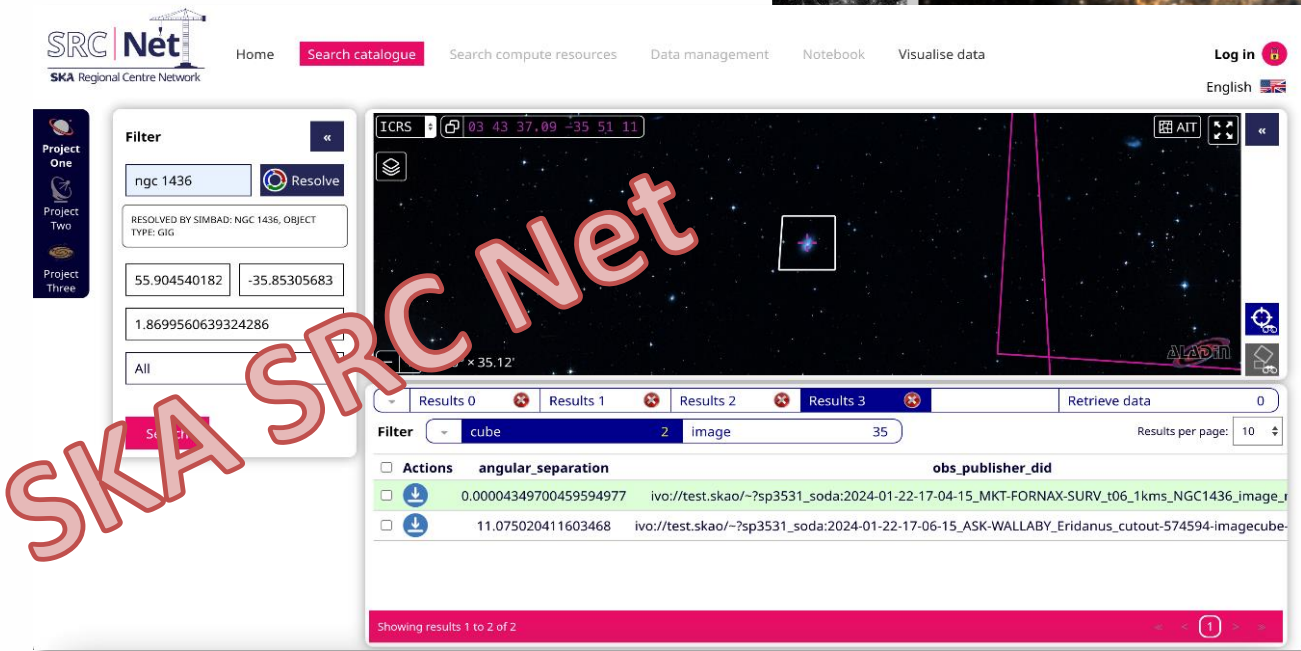
- A collection of astronomical images stored as **hierarchical multi-resolution** tiles using the **HEALPix** sky tessellation.
- The creation of a **HiPS** consists of the generation of the **mosaic**, the **partitioning**, and then the generation of the **tree** of tiles that constitute the final HiPS.
- The resulting **tiles** are image files of the same size **grouped in directories** following a hierarchy described in the IVOA standard.

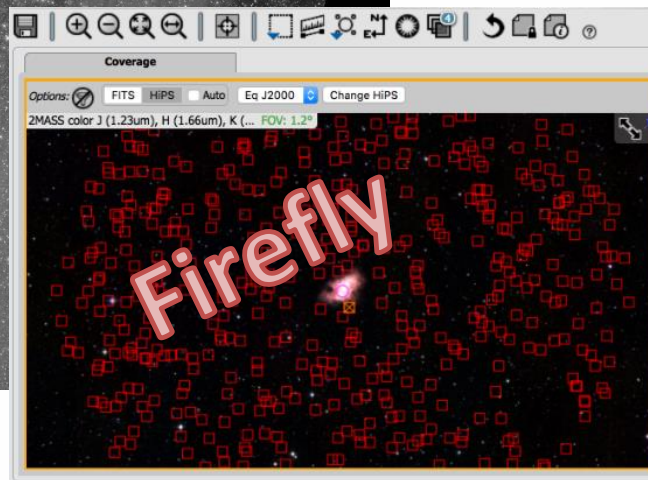


Browse the HiPS sky already provided in hundreds of web “portals”







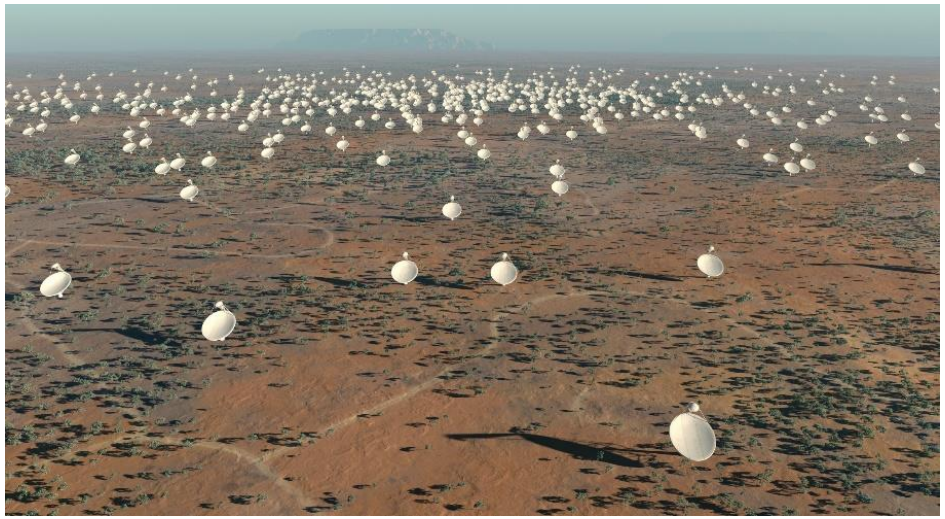


Next challenge: Cubic Surveys

Soon an **avalanche of cubic data** (SKA =700PB of data/year in the form of cubes of several hundred GB, or even much more).

Necessity to invent new solutions.

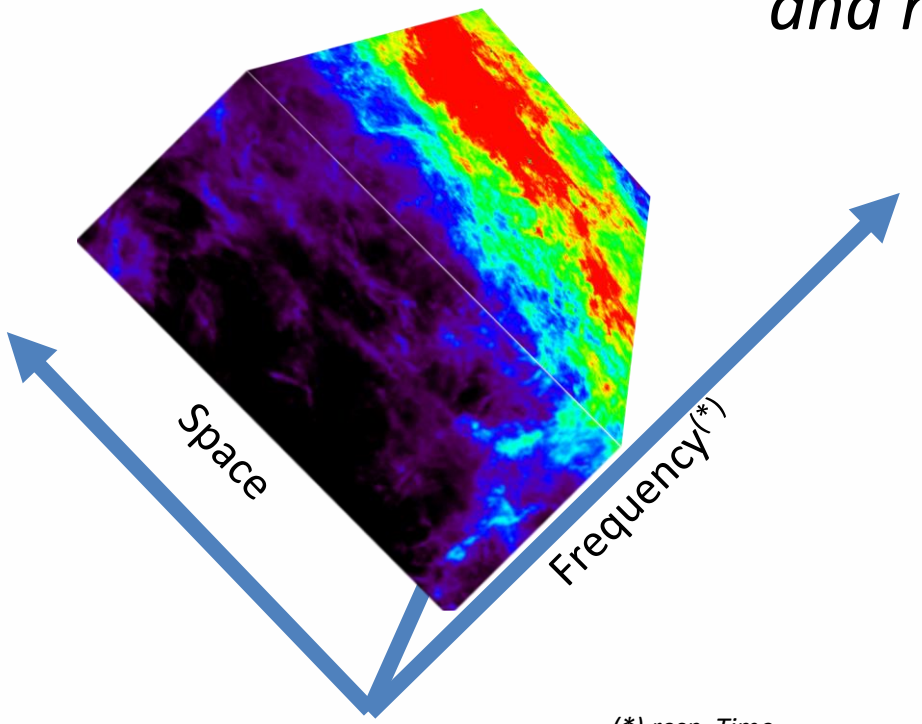
3 years *of discussions, tests, studies and developments supported by the* ***CDS*** *in its contribution to the* ***SKA SRCNet*** *+ help & contributions from IVOA members & other contributors*

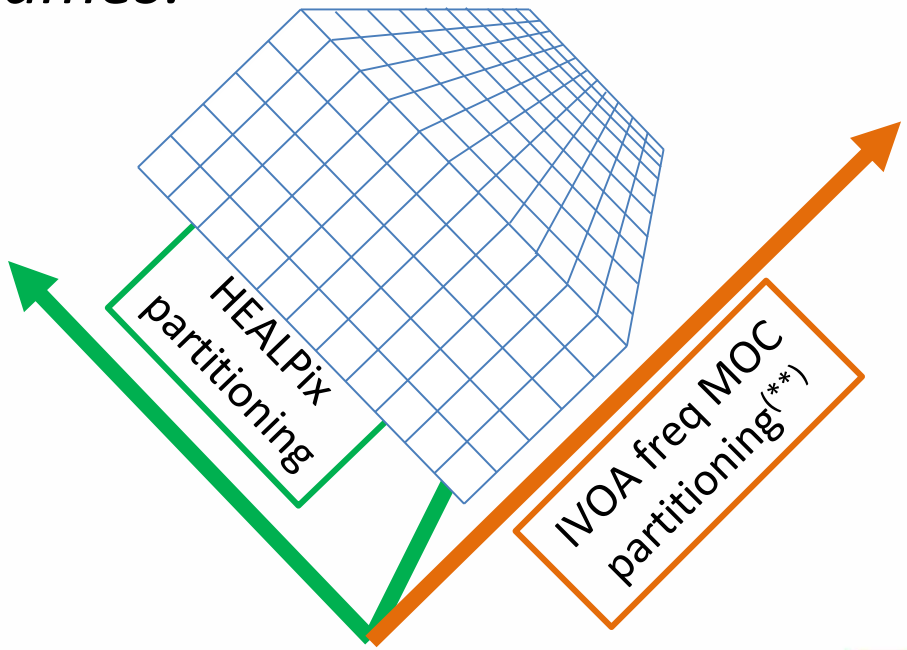


HiPS3D idea

*Extend the current HiPS towards a **HiPS 3D***

i.e. *use a **hierarchical partition in each physical dimension***
and based on absolute physical units
and reference frames.



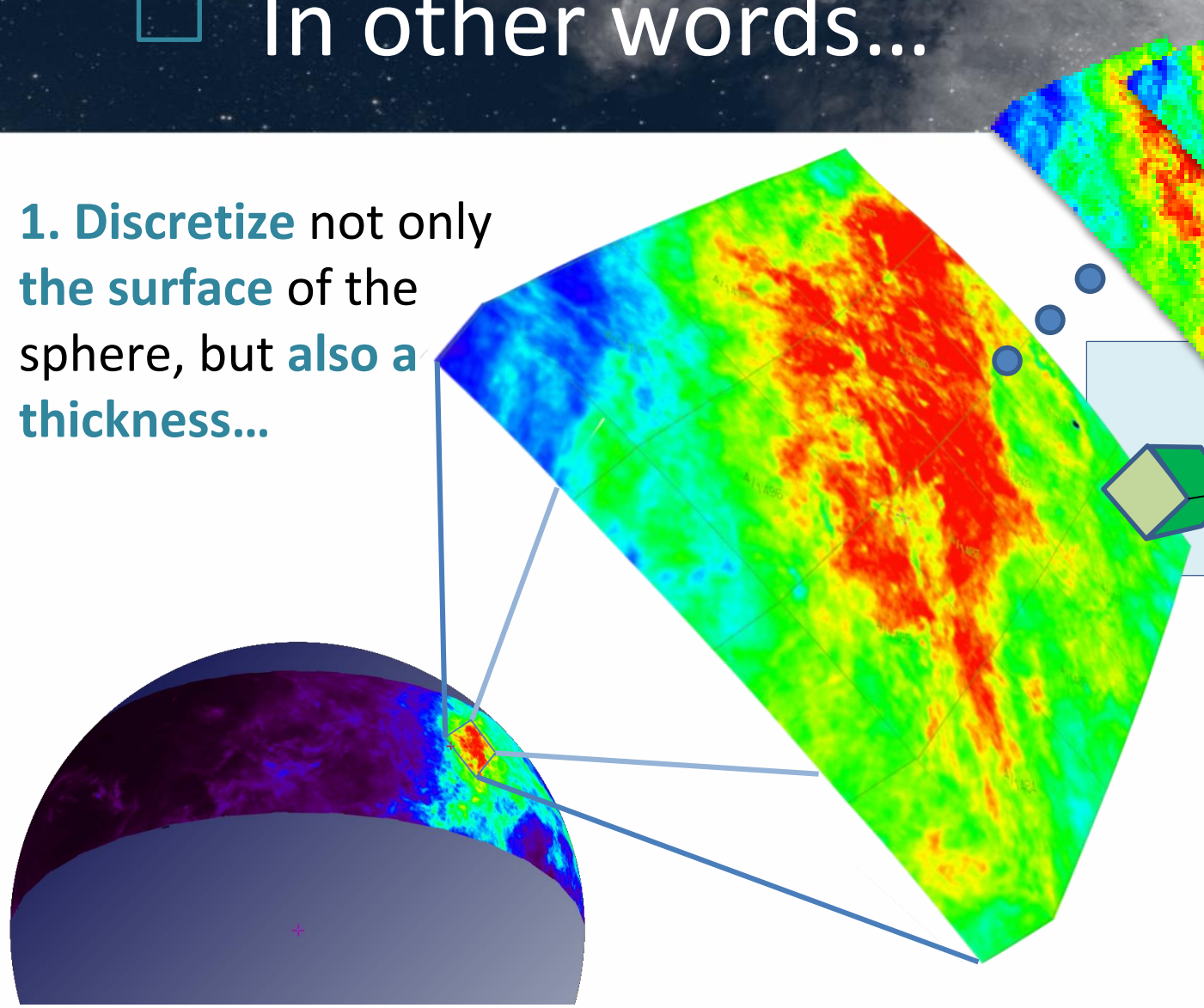


(*) resp. Time

(**) based on IVOA MOC 2.0 & frequency extensions

In other words…

1. Discretize not only **the surface** of the sphere, but **also a thickness…**



2. And do this at **several resolutions**

Each 3D pixel in the pyramid has a unique HiPS3D address:

(OrderS/NpixS_OrderF/NpixF)

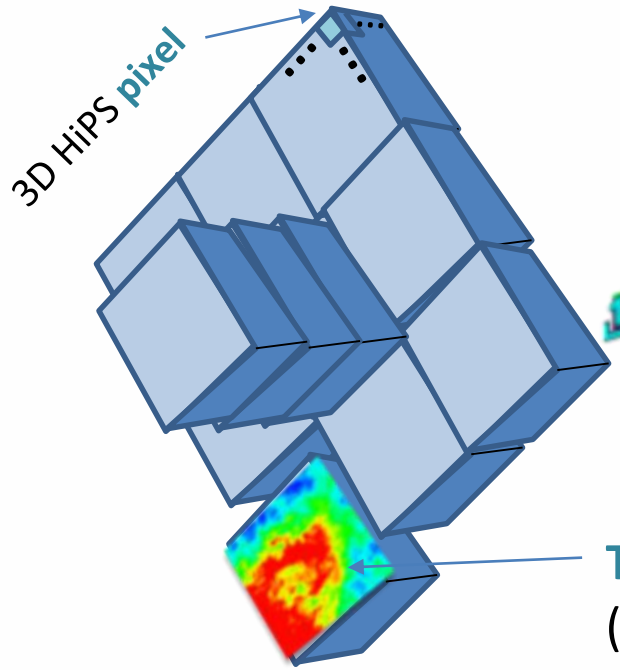
HEALPix partitioning

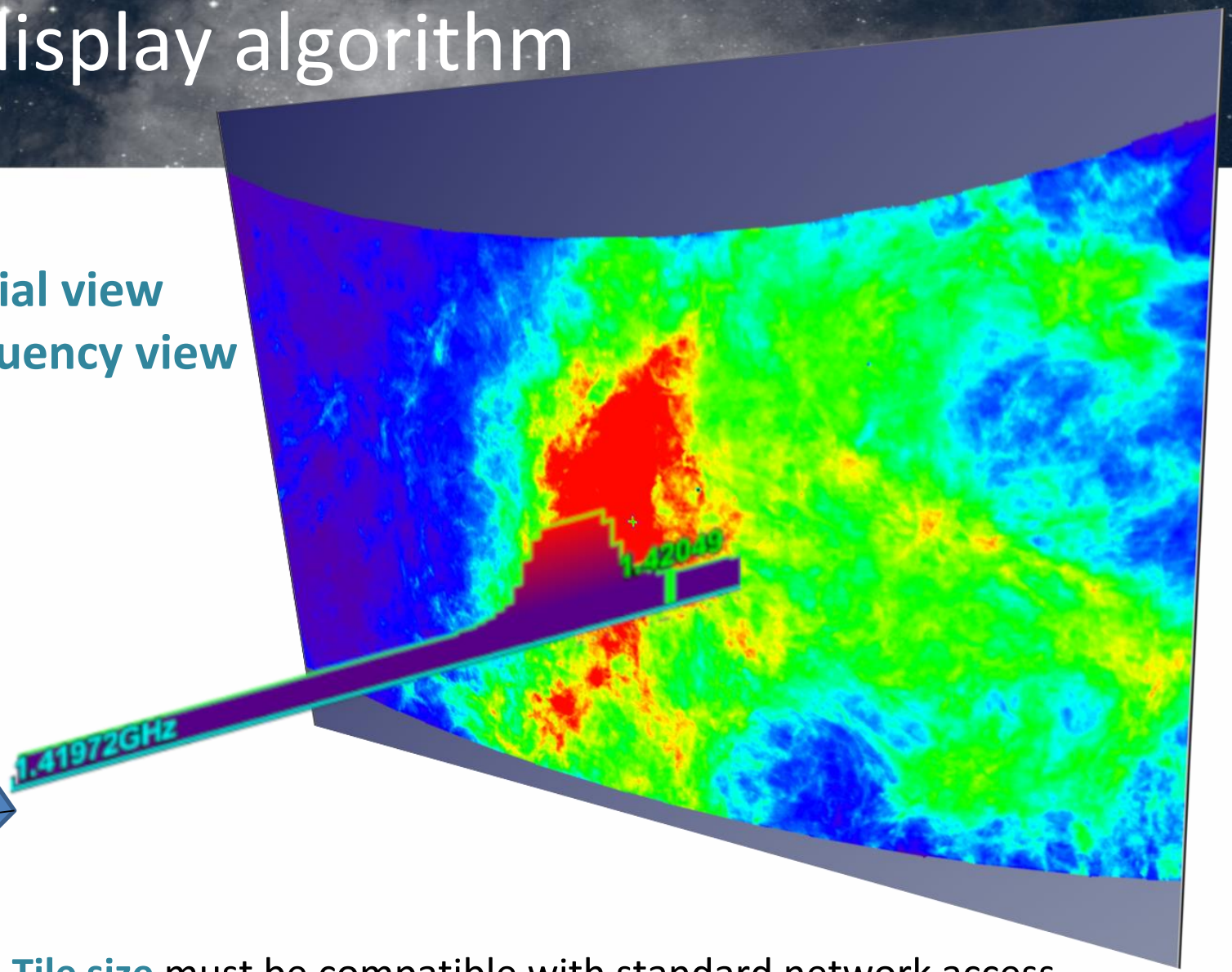
IVOA freq MOC partitioning

HiPS3D client display algorithm

The HiPS3D client loads :

- Only the tiles covering the **spatial view**
- Only the tiles covering the **frequency view**
- at the **appropriate resolution**





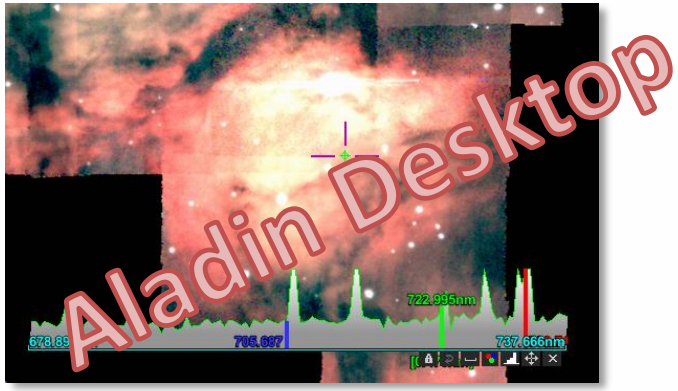
Tile size must be compatible with standard network access (typically 256x256 spatial pixels x 16 channels).

HiPS3D tool collection
(*) in fast progress

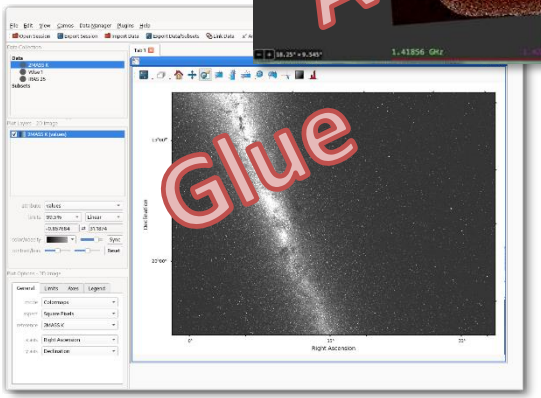
```
INFO  : Starting HipsGen 20/12/22 19:46:21 (based on Aladin v12.023)...
OPTION: in=HalphaNorth
OPTION: out=hips
INFO  : Action => INDEX: Build spatial index (                der directory) + MOC index
INFO  : Action => TILES: Build all true                        tiles + Allsky + MOC
INFO  : Action => PNG: Build all                               ) + Allsky.png
INFO  : Action => CHECKCODE:                                   the check codes (and the size) associated to the target HiPS
INFO  : Action => DETAILS: Ada                                 ex for supporting the "detail table" facility
```

Hipsgen

- **Hipsgen**: HiPS generation (public version)

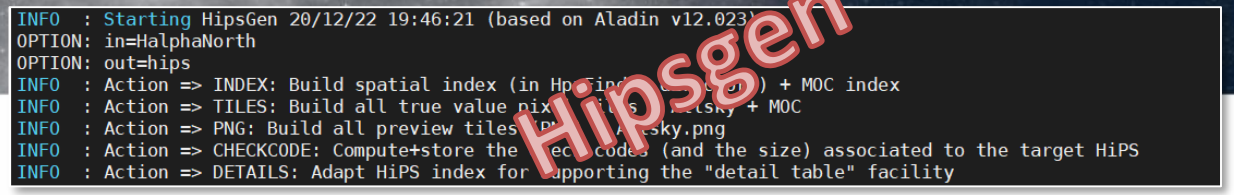


- **Aladin Desktop:** Standalone HiPS visualization (version beta 12.6)
- **Aladin Lite:** Web & jupyter(*) HiPS visualisation (prototype v3.7)
- **Astropy/reproject/Glue**: Generation & visualisation(*) [T.Robitaille]



- **Hips2fits3D** : Server side extraction of any cube at any resolution from a HiPS3D
  -> See T.Boch's ADASS poster

Creating a HiPS3D from your own data!



Command example:

```
Java –jar Hipsgen.jar –hips3D \
   in=YourCube.fits out=YourHips3D id=AUTH/C/xxx INDEX TILES PNG
```

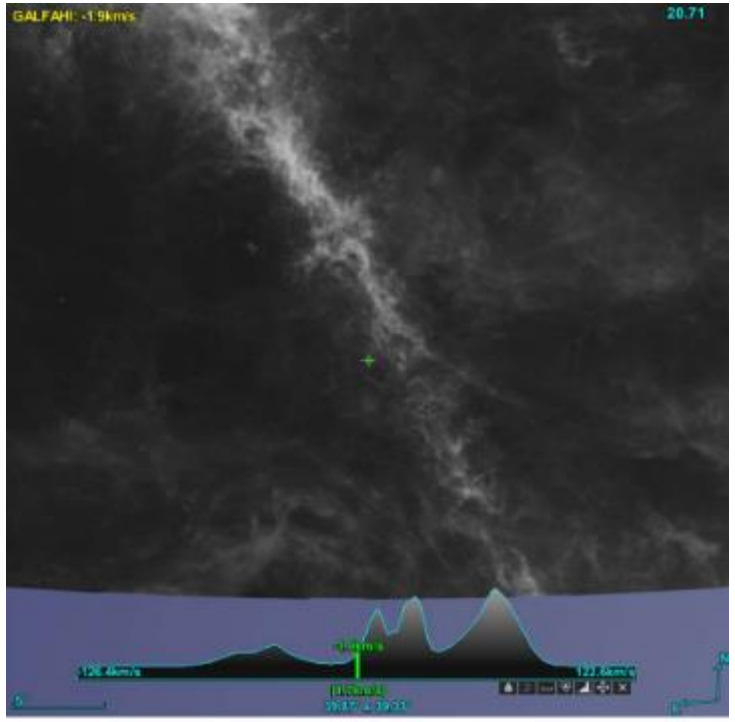
Key options:

- YourCube.fits → FITS cube (or directory of cubes) with spatial & frequency calibration
- YourHips3D → output directory for the HiPS3D product
- AUTH/C/xxx → identifier of your choice
- INDEX TILES PNG → actions to build HiPS3D with FITS + PNG tiles
- -hips3D → enables HiPS3D generation

More options (resolution, overlays, etc.): HiPSgen Manual

Various types of cubic surveys already successfully hipsilised (Hipsgen tool)

- **Spatial mosaics**
  - **GALFAHI**: 225 cubes (512x512x2048 = 225GB)
- **Pointed observations**
  - **MUSE**: 2600 cubes (350x350x3700 = 4TB)
  - **ASKAP**: 4 cubes (11000x11000x144 = 177GB)
- **Frequency mosaics**
  - **MEERKAT**: 3 cubes (5000x5000x2000 = 310GB)
- **Pointed obs. in "space & frequency" mosaic**
  - **ALMA**: 560 cubes (heterogeneous = 1.4TB)
- **Simulations**
  - **SKADC2**: 1 cube (5851x5851x6668 = 850GB)

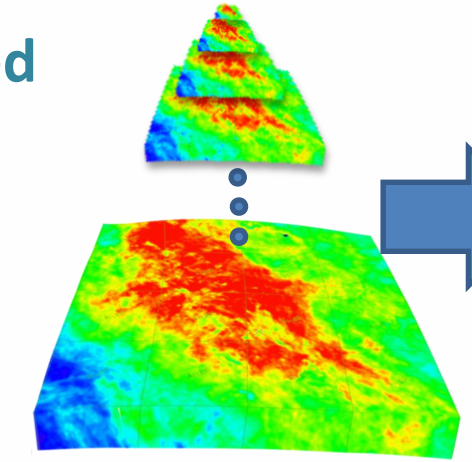


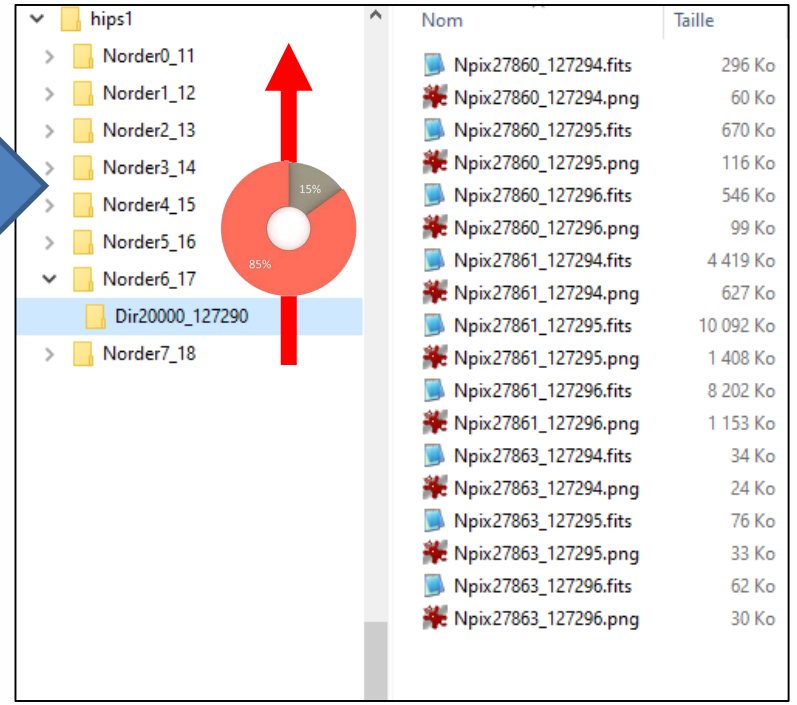
- **Others: LGBSHI, DHIGLS, MANGA, HARMONI…**

https://aladin.cds.unistra.fr/java/HipsList3D.txt

HiPS3D server directory structure

- As regular HiPS, HiPS3D-tiles may be **stored** in a **regular file system**
- **allowing all combinations of resolution** (spatial vs. frequency)(*)
- But for visualisation a single combination is enough
  = **simultaneous reduction in spatial & frequency(*) resolution**





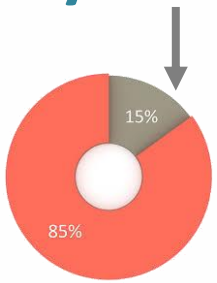
The hierarchy **adds only 15%** to the volume of HiPS
(reduced by a factor of 8 for each sub-order)

(*) resp. time

Insights on the HiPS3D time generation/volume

4 volume factors

1. **Resolution**
   (original vs HiPS)
2. **Compression**
3. **Overlays**
4. **HiPS Pyramid**



Input data	*HiPS processing*	*Generation time*	*Nb. files (tiles)*	*Final volume*
GALFAHI (225 cubes, 225 GB)	FITS tiles	**32min** 5s	140,502	**250** GB
	PNG tiles	+2min 48s	140,502	7.8 GB
	JPEG tiles	+1min 25s	140,502	**4.2** GB (2%)
ALMA (88 cubes, 675 GB)	FITS tiles	**53min**	10,060	**28,3** GB (5%)
	PNG tiles	+1min 21s	10,060	3.3 GB (0,5%)

Done on a 64 thread machine, disk access by NFS

Let’s imagine a very large cube survey…

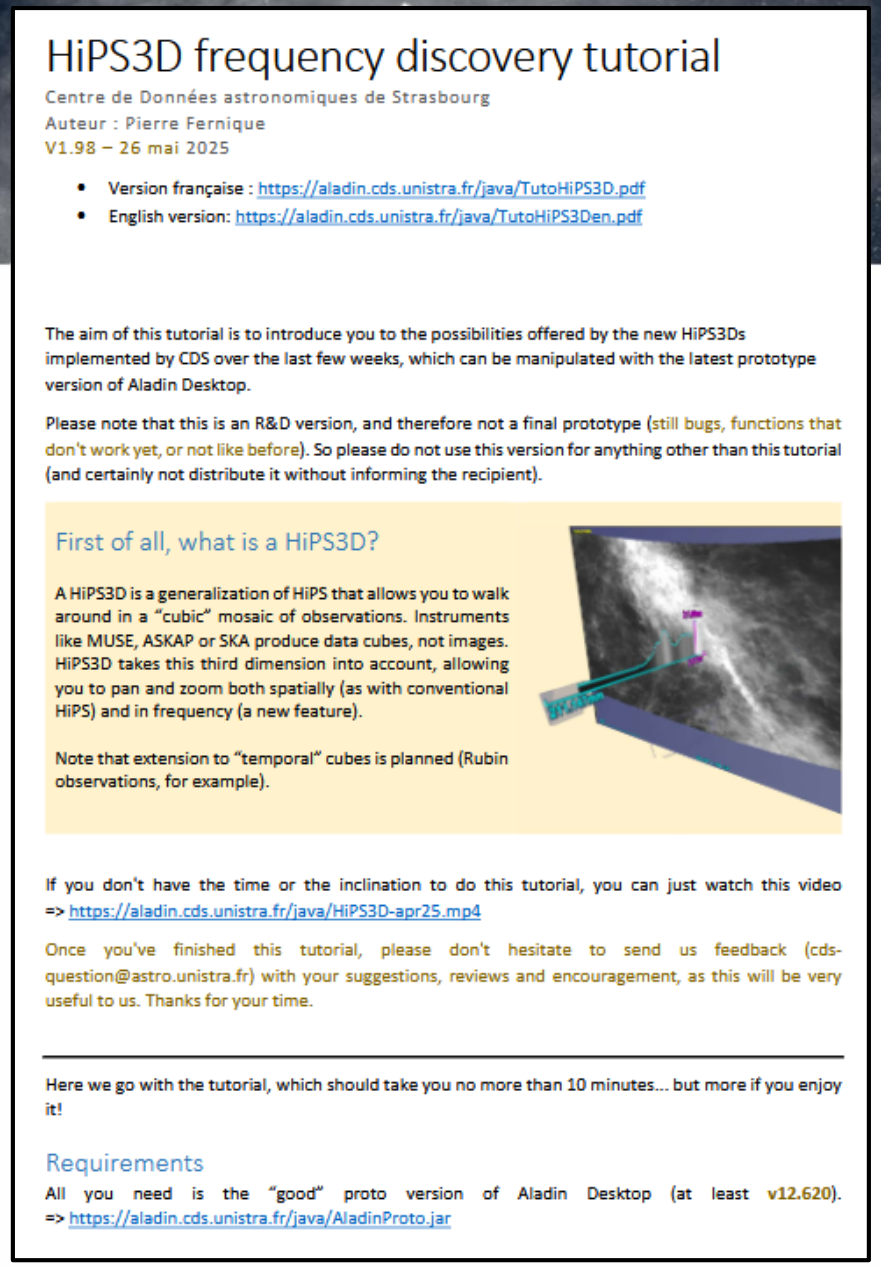
- **Several PB large** - e.g.SKA.
- By reducing the resolution by **2 HiPS orders** and **keeping only compressed tiles,** the additional storage cost will be **0.2%**.
- Furthermore, **let us imagine an “aggregated” HiPS3D** so that each SKA node generates only its own part **without having to move the original cubes** to generate this HiPS, and thus the final HiPS is simply obtained dynamically from all sites.
- At the end, with a few additionnal TB, **we will be able to visualise PB cubes surveys from any portals**, even from a mobile device.

…not a dream, **just the reality**

Thanks – Question ?

ADASS – Görlitz – Nov 2025 – P. Fernique

Manual, Technical information, Tutorial, Demos, all in one link



HiPS3D frequency discovery tutorial

Centre de Données astronomiques de Strasbourg
Auteur : Pierre Fernique
V1.98 – 26 mai 2025

- Version française : https://aladin.cds.unistra.fr/java/TutoHiPS3D.pdf
- English version: https://aladin.cds.unistra.fr/java/TutoHiPS3Den.pdf

The aim of this tutorial is to introduce you to the possibilities offered by the new HiPS3Ds implemented by CDS over the last few weeks, which can be manipulated with the latest prototype version of Aladin Desktop.

Please note that this is an R&D version, and therefore not a final prototype (still bugs, functions that don't work yet, or not like before). So please do not use this version for anything other than this tutorial (and certainly not distribute it without informing the recipient).

First of all, what is a HiPS3D?

A HiPS3D is a generalization of HiPS that allows you to walk around in a "cubic" mosaic of observations. Instruments like MUSE, ASKAP or SKA produce data cubes, not images. HiPS3D takes this third dimension into account, allowing you to pan and zoom both spatially (as with conventional HiPS) and in frequency (a new feature).

Note that extension to "temporal" cubes is planned (Rubin observations, for example).

If you don't have the time or the inclination to do this tutorial, you can just watch this video => https://aladin.cds.unistra.fr/java/HiPS3D-apr25.mp4

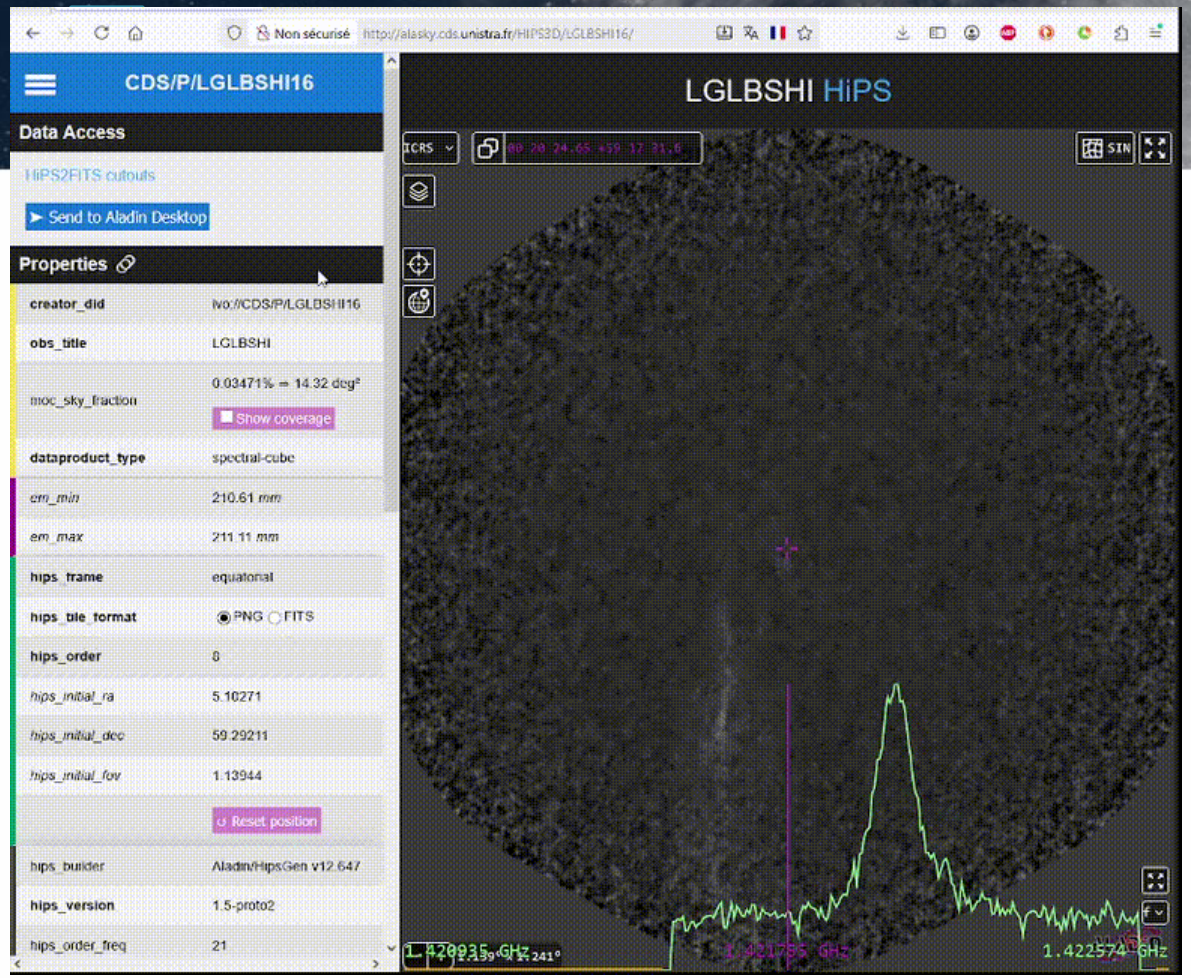
Once you've finished this tutorial, please don't hesitate to send us feedback (cds-question@astro.unistra.fr) with your suggestions, reviews and encouragement, as this will be very useful to us. Thanks for your time.

Here we go with the tutorial, which should take you no more than 10 minutes... but more if you enjoy it!

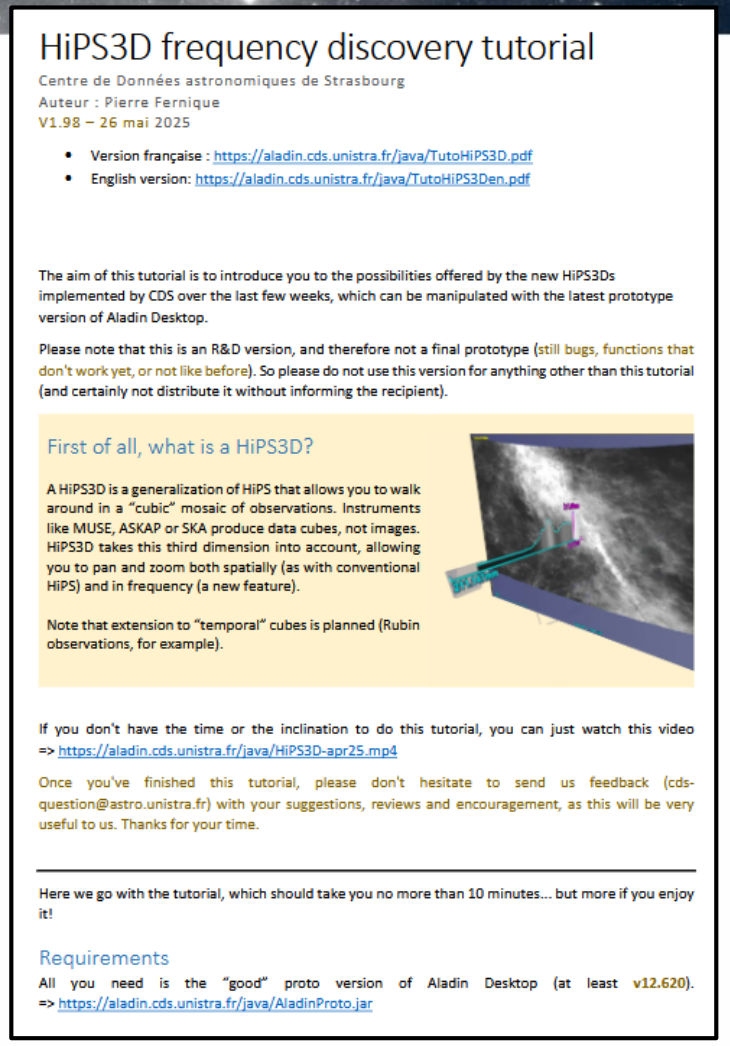
Requirements

All you need is the "good" proto version of Aladin Desktop (at least v12.620).
=> https://aladin.cds.unistra.fr/java/AladinProto.jar

https://aladin.cds.unistra.fr/**java/TutoHiPS3Den.pdf**



https://alasky.cds.unistra.fr/**HIPS3D/LGLBSHI16**



HiPS3D frequency discovery tutorial

Centre de Données astronomiques de Strasbourg
Auteur : Pierre Fernique
V1.98 – 26 mai 2025

- Version française : https://aladin.cds.unistra.fr/java/TutoHiPS3D.pdf
- English version: https://aladin.cds.unistra.fr/java/TutoHiPS3Den.pdf

The aim of this tutorial is to introduce you to the possibilities offered by the new HiPS3Ds implemented by CDS over the last few weeks, which can be manipulated with the latest prototype version of Aladin Desktop.

Please note that this is an R&D version, and therefore not a final prototype (still bugs, functions that don't work yet, or not like before). So please do not use this version for anything other than this tutorial (and certainly not distribute it without informing the recipient).

First of all, what is a HiPS3D?

A HiPS3D is a generalization of HiPS that allows you to walk around in a "cubic" mosaic of observations. Instruments like MUSE, ASKAP or SKA produce data cubes, not images. HiPS3D takes this third dimension into account, allowing you to pan and zoom both spatially (as with conventional HiPS) and in frequency (a new feature).

Note that extension to "temporal" cubes is planned (Rubin observations, for example).

If you don't have the time or the inclination to do this tutorial, you can just watch this video => https://aladin.cds.unistra.fr/java/HiPS3D-apr25.mp4

Once you've finished this tutorial, please don't hesitate to send us feedback (cds-question@astro.unistra.fr) with your suggestions, reviews and encouragement, as this will be very useful to us. Thanks for your time.

Here we go with the tutorial, which should take you no more than 10 minutes... but more if you enjoy it!

Requirements

All you need is the "good" proto version of Aladin Desktop (at least v12.620).
=> https://aladin.cds.unistra.fr/java/AladinProto.jar

https://aladin.cds.unistra.fr/**java/TutoHiPS3Den.pdf**