Distributed Visualization -

Building and Using HiPS in Distributed Data Storage

Sweta Singh (MSc. Student)
Supervisors: E.A. Valentijn, A. Belikov, H. Buddelmeijer

IVOA 2019, Groningen

kapteyn astronomical institute

OmegaCEN

Introduction

- Euclid Space Mission (ESA-M)

  - Extra galactic surveys (visible, infrared, spectroscopy)
- 10's of PB data
- Collaborators all over the world
- Multiple Science Data Centers
- Incremental Overall Progress Map - Data available, status of - processing, data release etc.

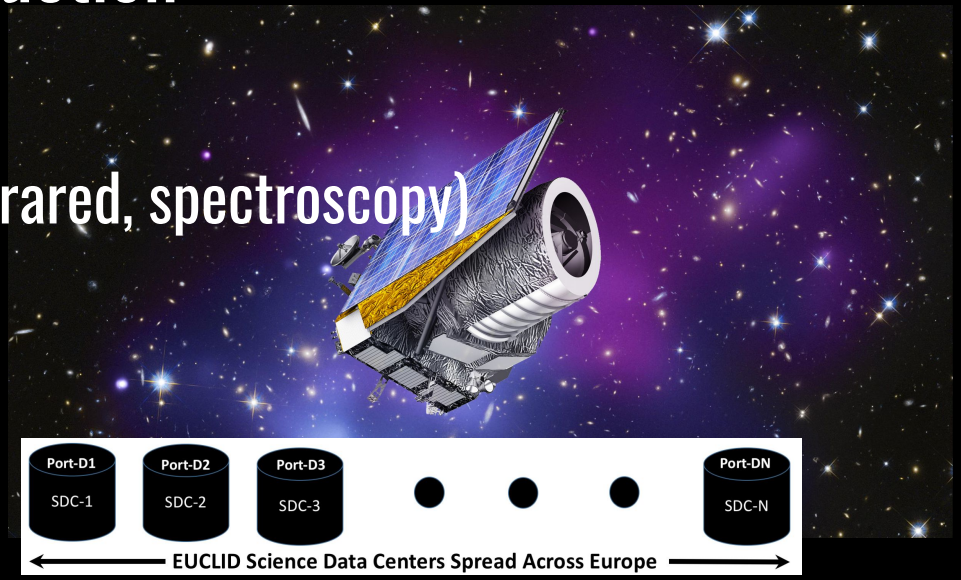


image source: ESA website

kapteyn astronomical institute

OmegaCEN

Challenges

- Need a Distributed Visualization framework
- SDCs (heterogeneity)
- Network between SDCs - from 10s to several 100s Mbps
- Data availability/Progress according to Observation Plan
- Scalable, Flexible, Future proof framework

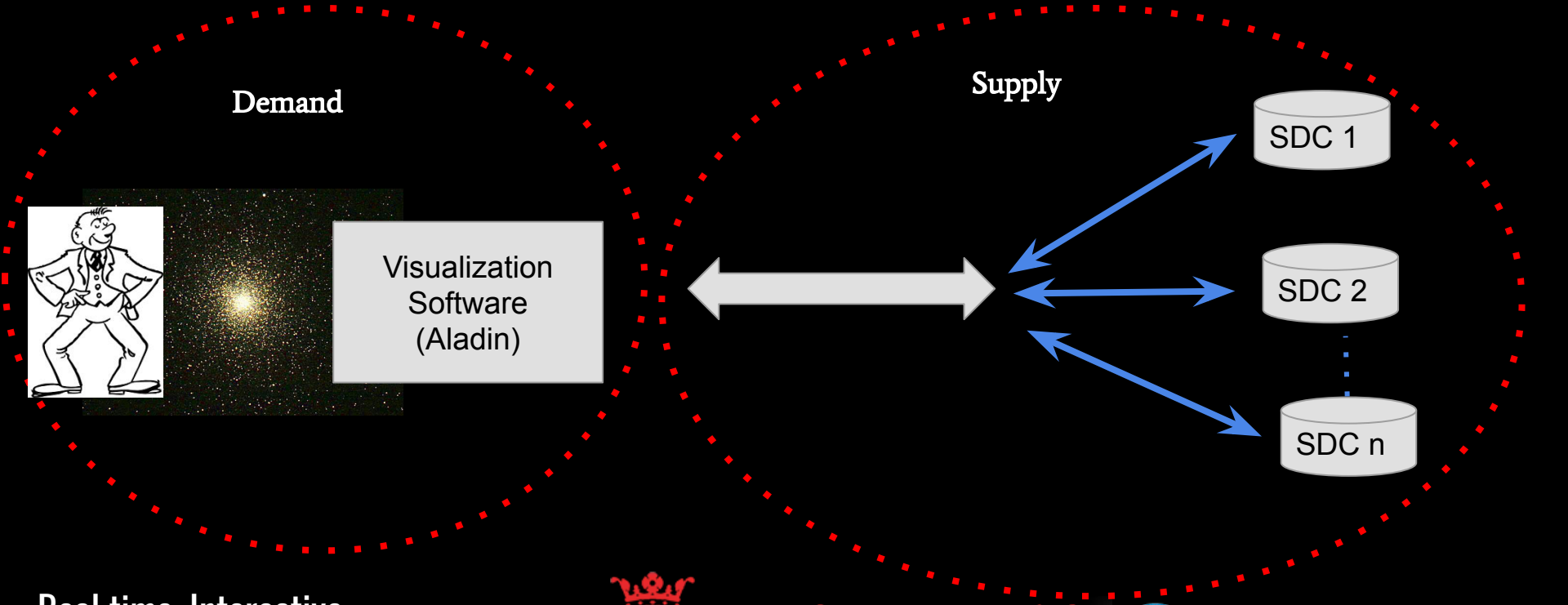
Two prong approach → reduce the data size (optional) &

→ enable distributed visualization

kapteyn astronomical institute

OmegaCEN

Visualization Framework



- Real time, Interactive
- Desktop feel but no copy

kapteyn astronomical institute

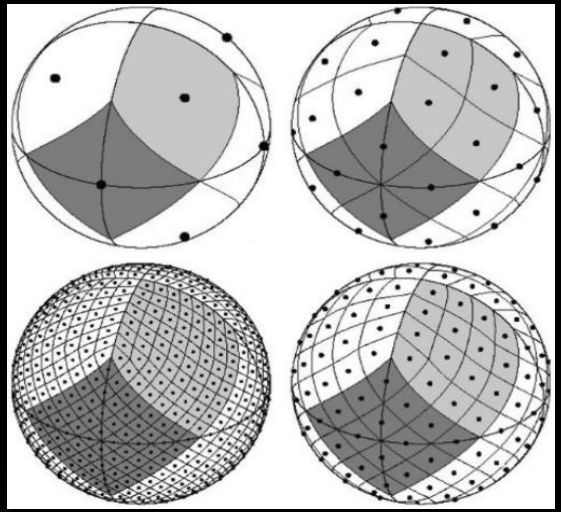
OmegaCEN

HiPS Survey

HiPS is a hierarchical tiling mechanism developed by CDS (P.Fernique et al, A&A 578, 114, 2015)

It supports multi-layer visualization

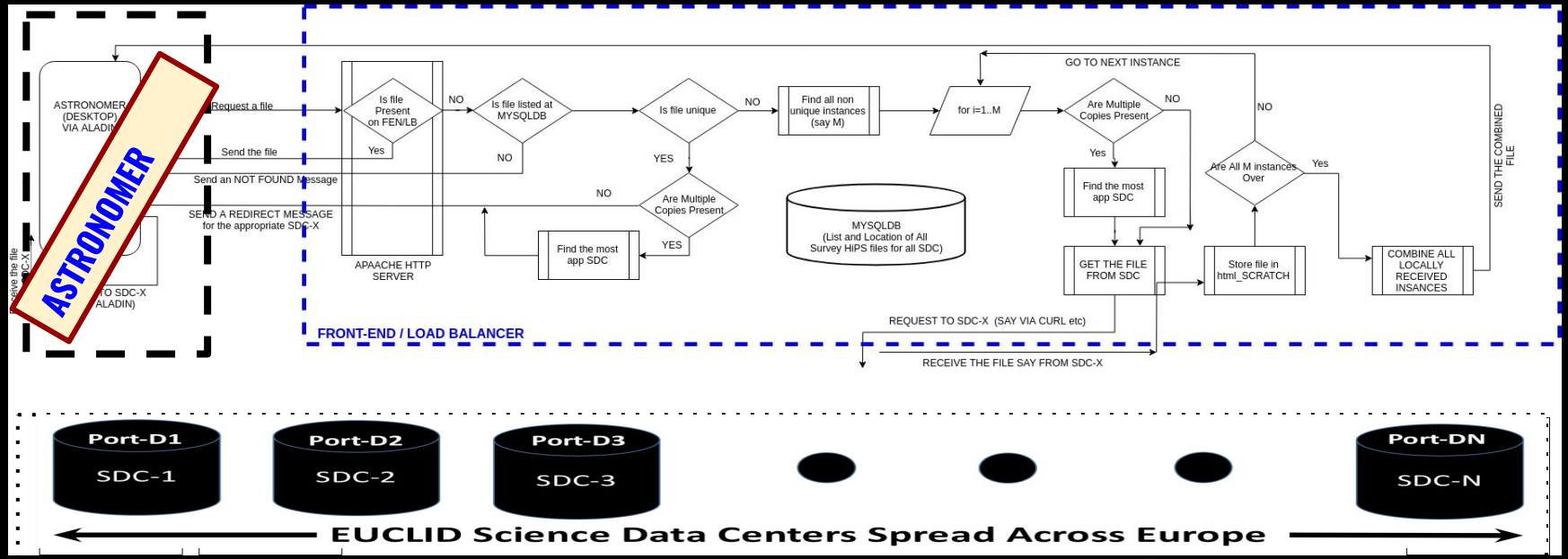
- HiPS is the defacto standard for survey maps
- HiPS is http compliant



kapteyn astronomical institute

OmegaCEN

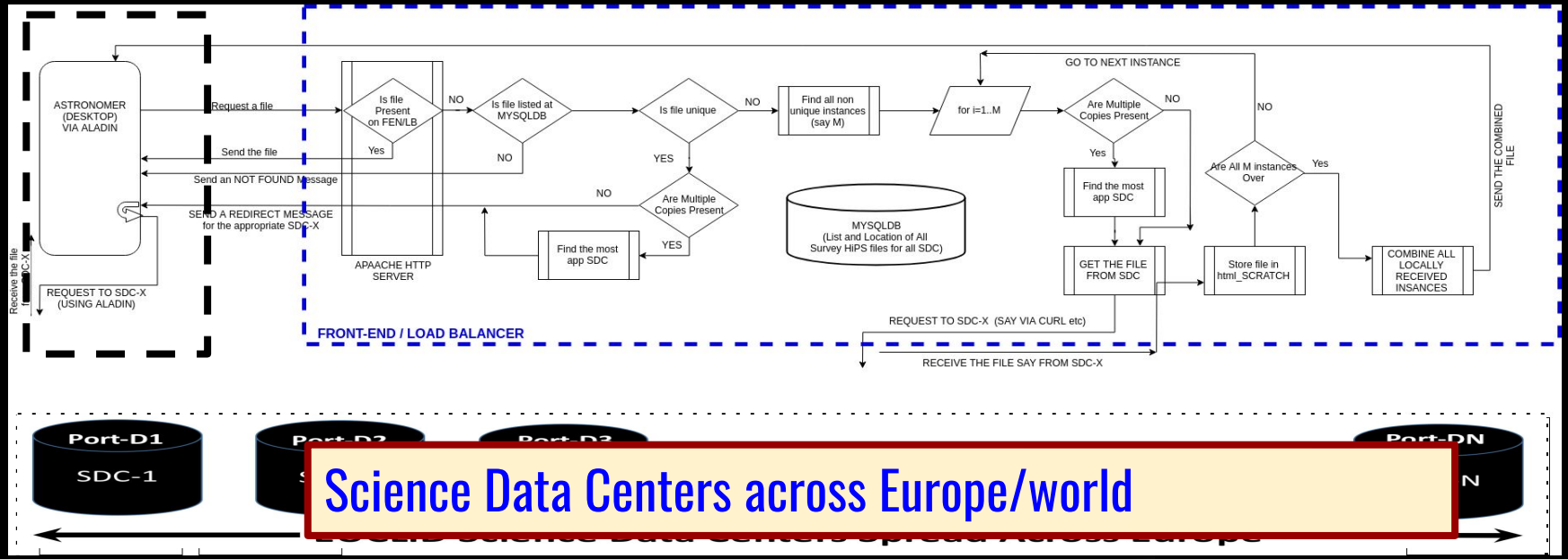
Visualization Framework with LAMP Stack



kapteyn astronomical institute

OmegaCEN

Visualization Framework with LAMP Stack

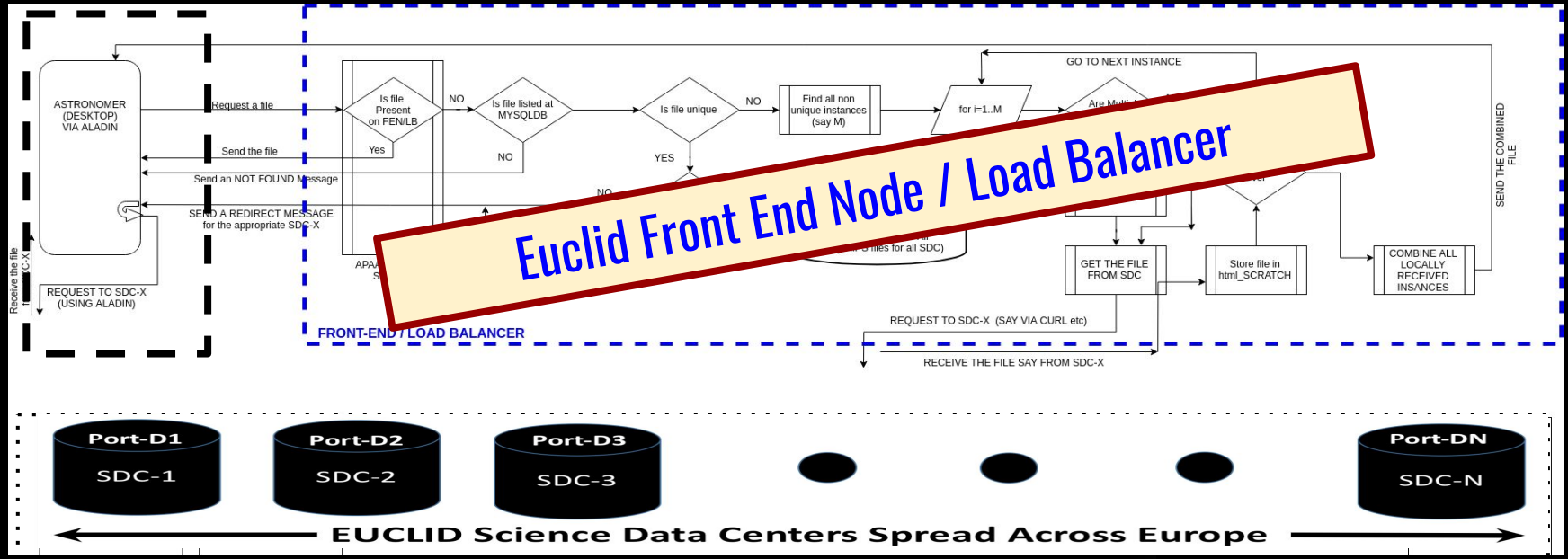


Science Data Centers across Europe/world

kapteyn astronomical institute

OmegaCEN

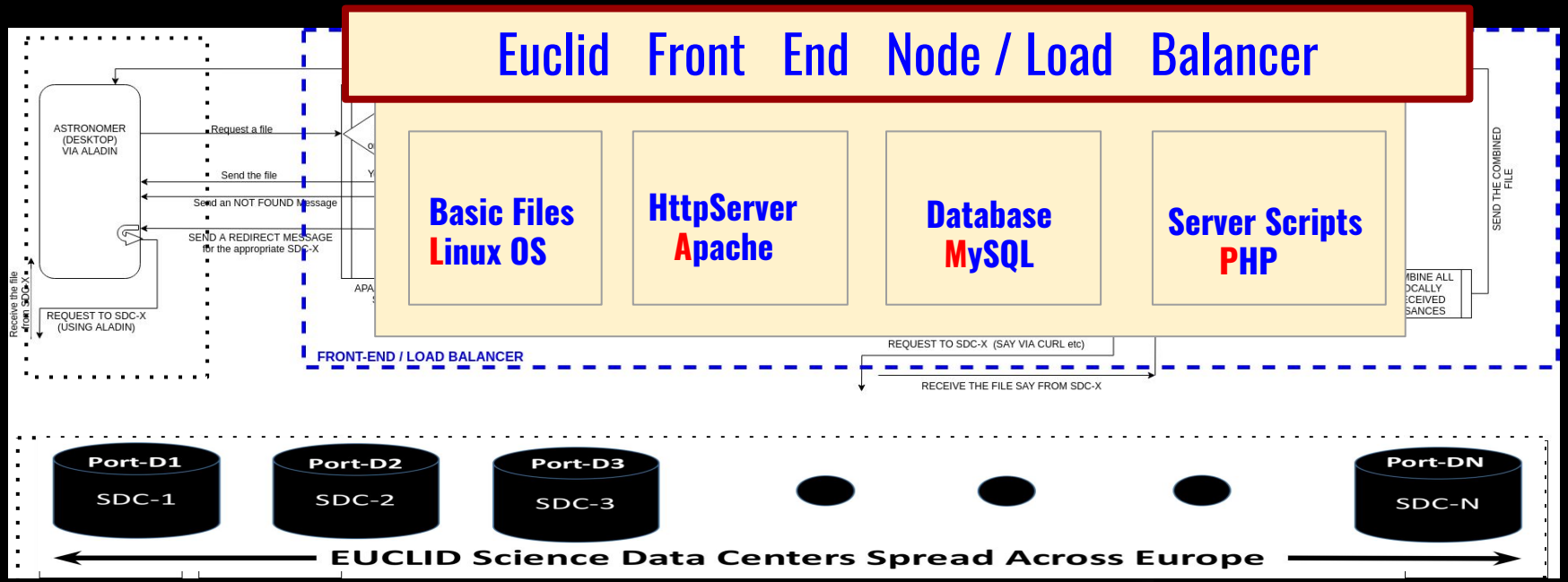
Visualization Framework with LAMP Stack



kapteyn astronomical institute

OmegaCEN

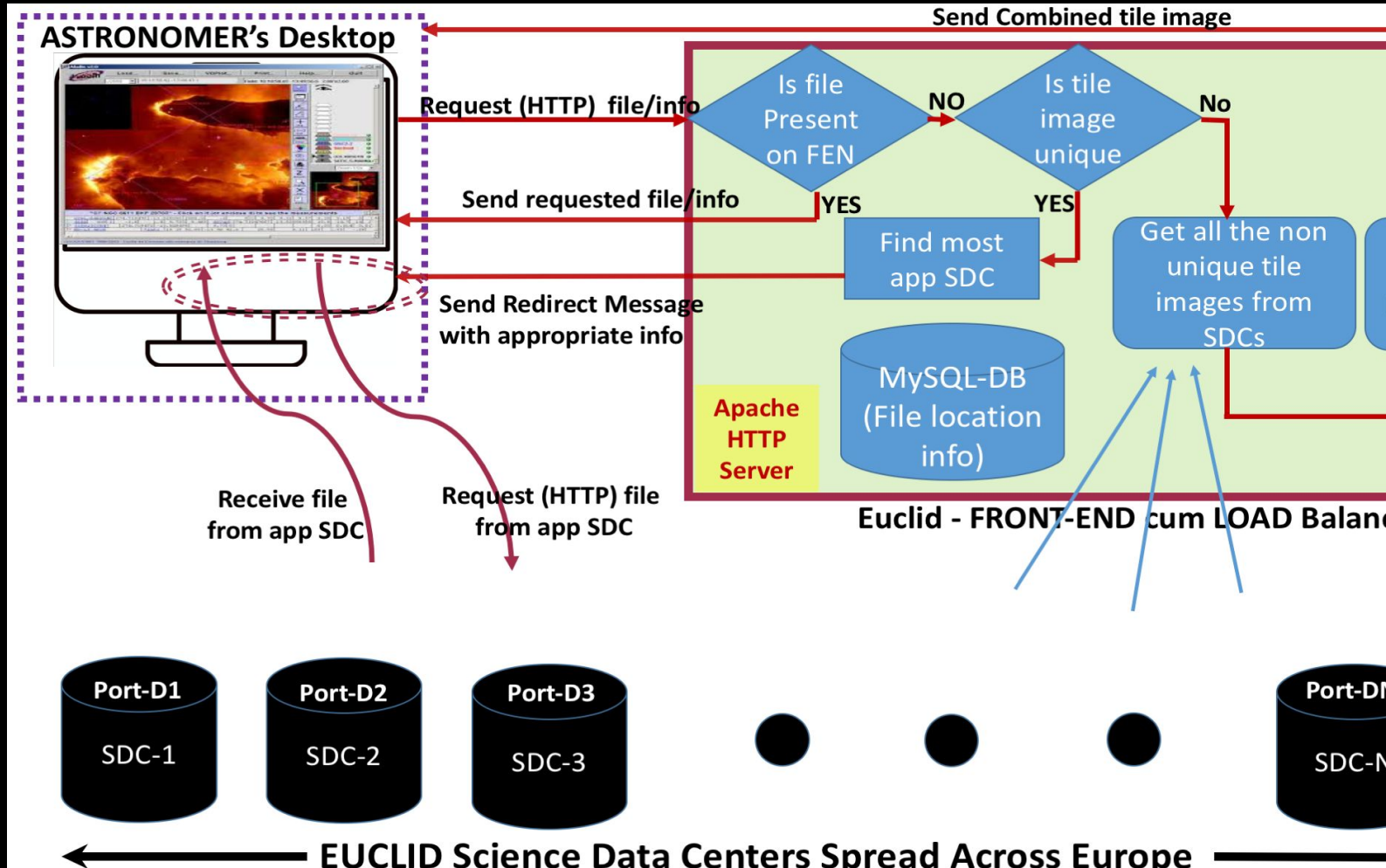
Visualization Framework with LAMP Stack

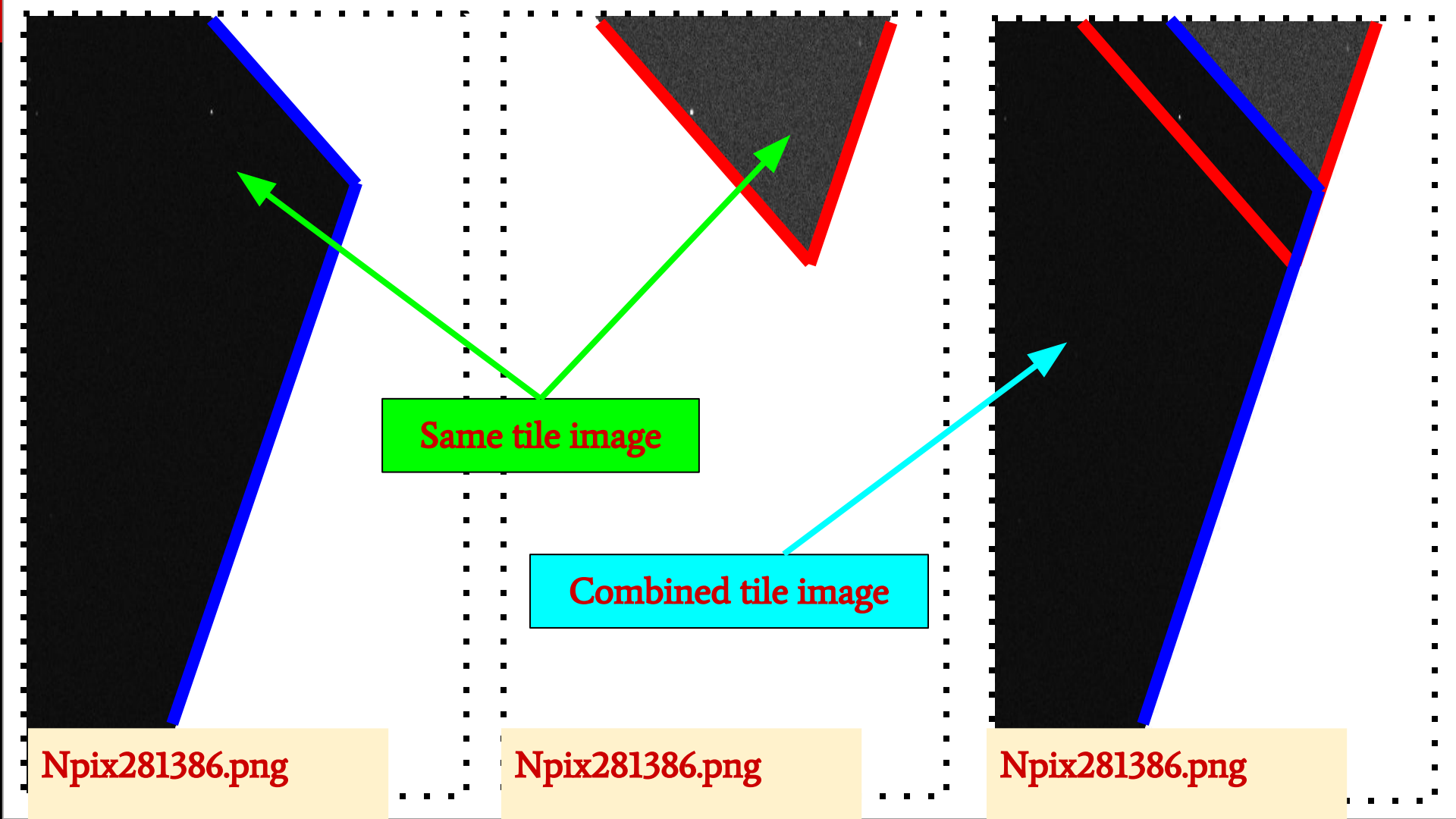


kapteyn astronomical institute

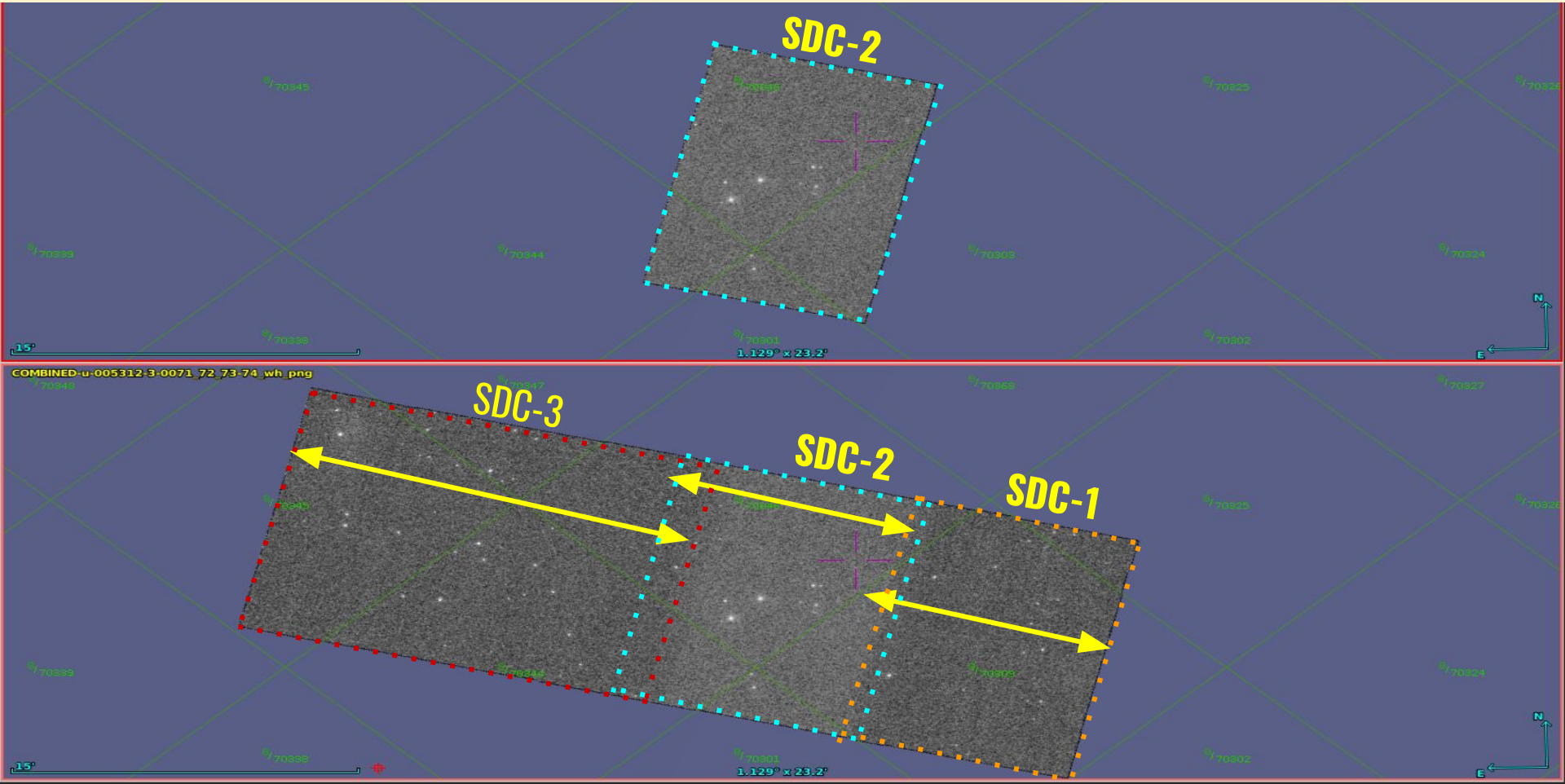
OmegaCEN

Fully Distributed Visualization

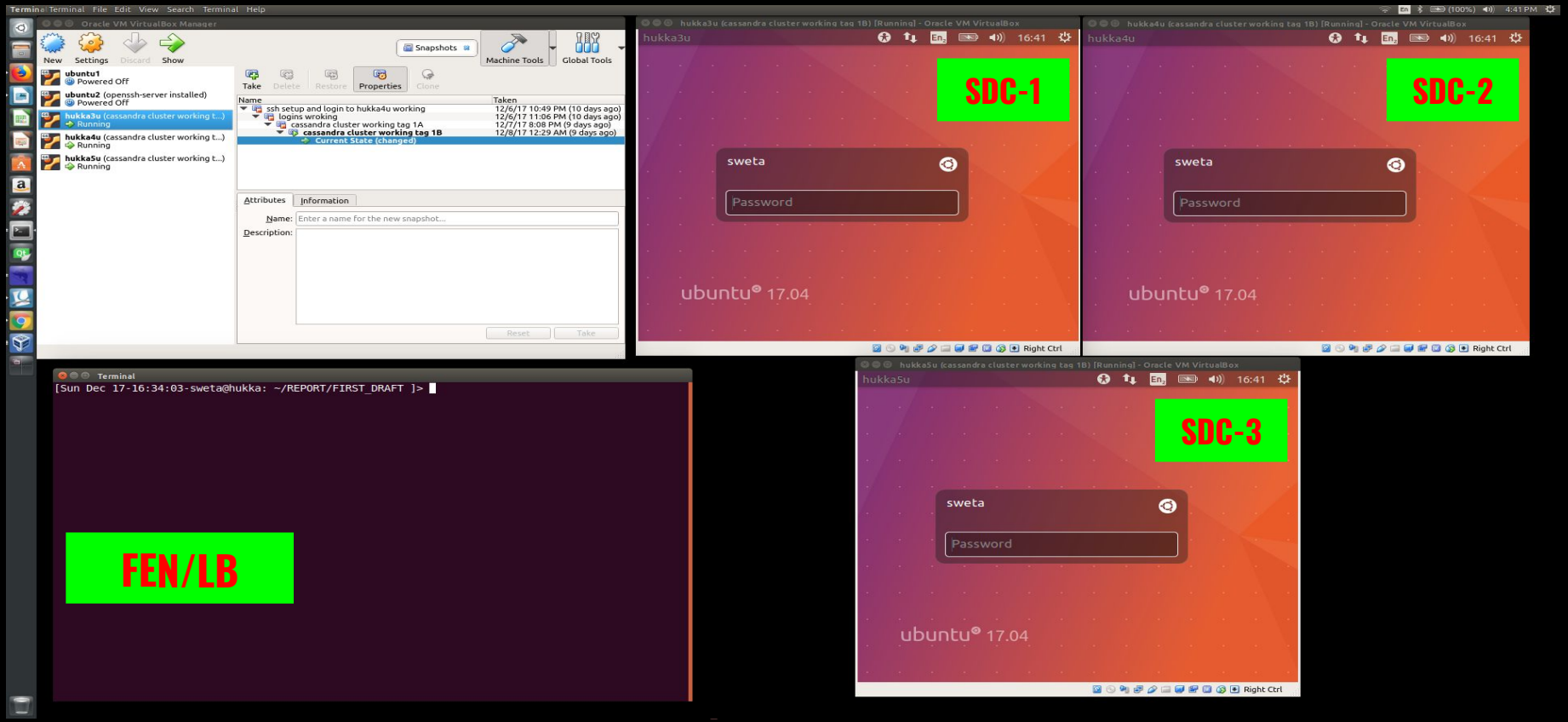



Same tile image
Combined tile image
Npix281386.png
Npix281386.png
Npix281386.png

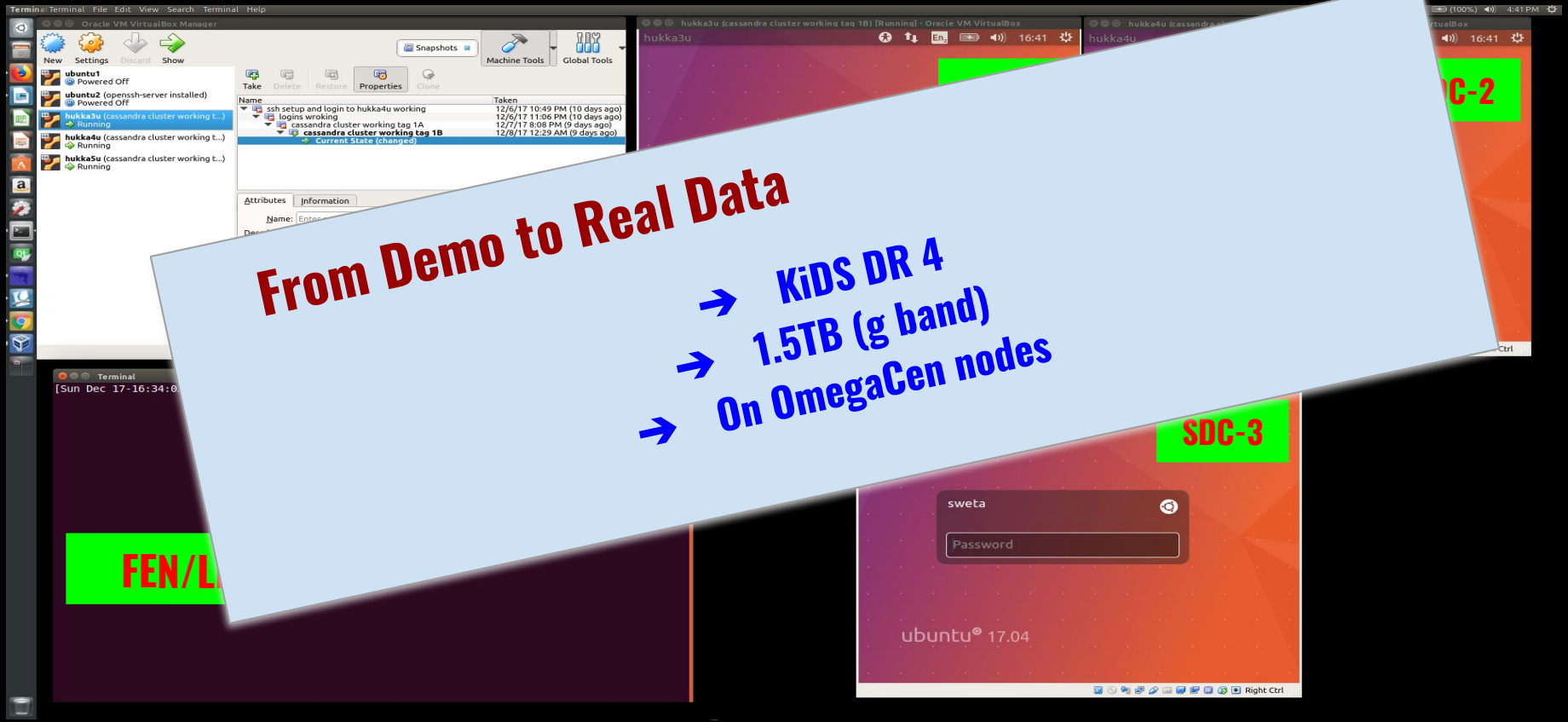
Visualization in practice



Virtual Machines as SDCs



Virtual Machines as SDCs



Important Aspects

- **Discussion Points**
  - Variable network Bandwidth ? - FEN/LB (optimal redirection), independent parallel transfer between SDCs & User, stateless nature of http.
  - Network failure at SDC ? - copy at other SDC used, low order hips on FEN.
  - Headers - jpeg , png
  - Data Cubes
  - All code/configuration etc. would be placed on Gitlab.
- **Multiple FEN/LB**
- **Distributed nosql databases like Apache Cassandra**
- **Nginx on SDCs for improved performance**
- **Reverse proxy configuration as an alternative topology**

OmegaCEN

Conclusion & future steps

- ➔ Successfully developed and demonstrated distributed visualization framework for very large surveys
- ➔ Our framework works on heterogenous SDCs.
- ➔ Data size reduction using png instead of pure fits for hips survey generation has been explored
- ➔ Applicable to big, collaborative project like SKA
- ➔ Extremely Important IVOA Role - especially in context of Big projects
- ➔ It is being implemented on OmegaCen server nodes as SDCs
- ➔ Performance and Monitoring using ELK stack is in progress
- ➔ Optimisation - caching, key value stores(cdb), distributed db(Apache Cassandra)
- ➔ We are on our way to implement the framework for Aladinlite, as it works on the desktop version

kapteyn astronomical institute

OmegaCEN

Suggestions/Feedback

- Combining distributed mandatory and recommended files using Aladin
  - Mandatory: properties
  - Recommended: index.html, moc.fits, allsky
- Is it possible to specify the quality png resolution ?

Fits → HiPS(png) using Aladin/hipsgen

- Combining same tile number images
  - Linux Commands e.g. convert
  - Python libraries
  - More feedback is welcome

kapteyn astronomical institute

OmegaCEN

Thank You !

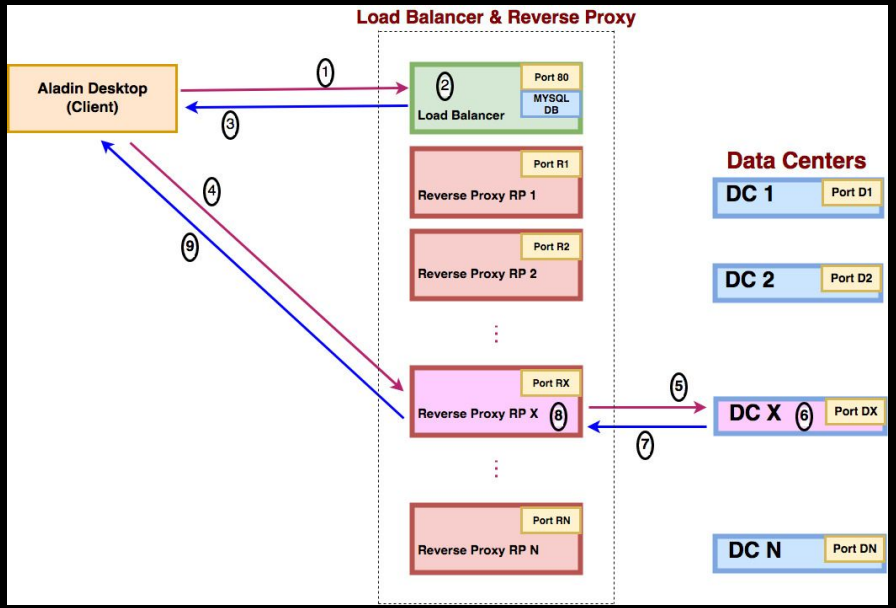
kapteyn astronomical institute

OmegaCEN

Architecture

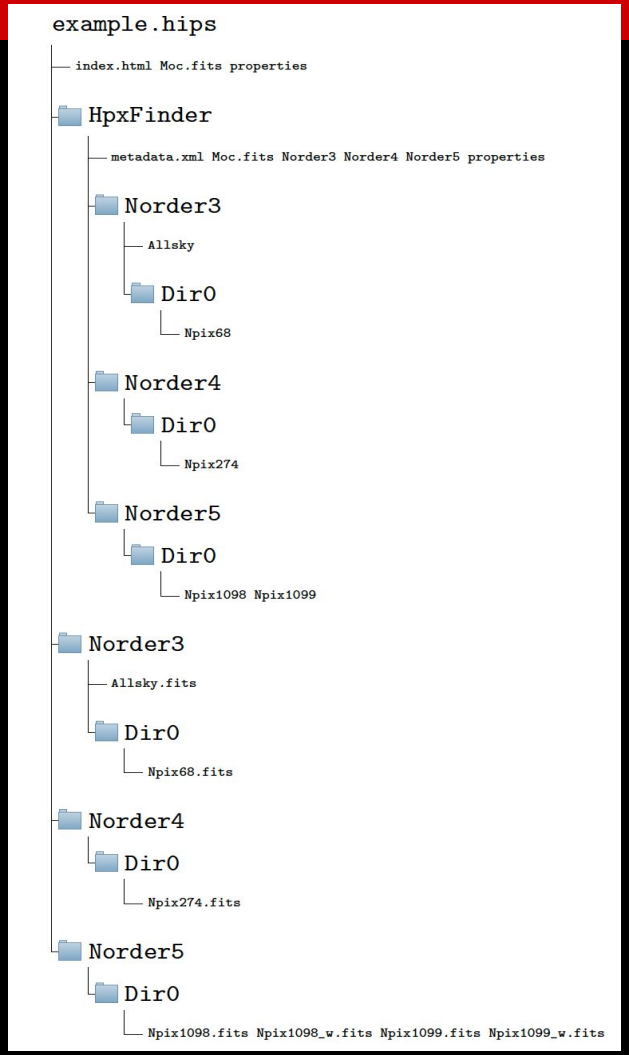
Distributed Network Mode

Reverse Proxy Network Mode



kapteyn astronomical institute

OmegaCEN



HiPS Survey

- HiPS is http compliant - allows it to be accessed via http server
- Simple Hierarchical Tree structure with directories and files

kapteyn astronomical institute

OmegaCEN

LAMP Stack

1. Linux → open source(free), reliable(virus free)
2. Apache → most popular, open source, reliable, secure, fast, http,
3. MySQL → simple, sql, open source
4. PHP → open source, server side(code executed on server side), scripting language, communicate with MySQL

kapteyn astronomical institute

OmegaCEN