

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



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

Demand



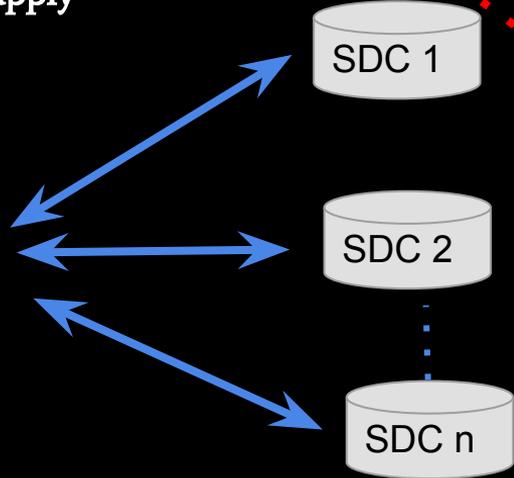
Visualization
Software
(Aladin)

Supply

SDC 1

SDC 2

SDC n



- Real time, Interactive
- Desktop feel but no copy



kapteyn astronomical
institute



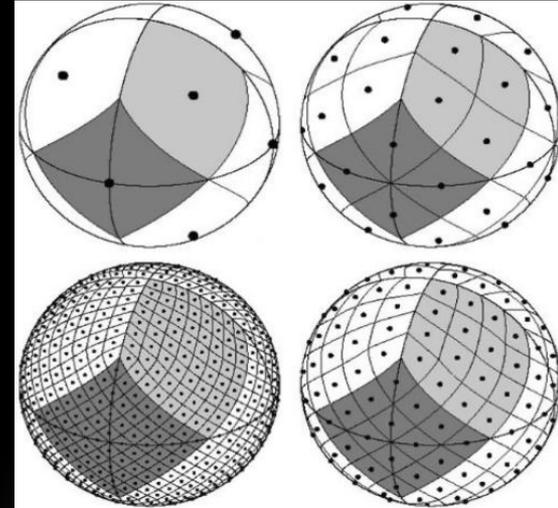
OmegaCEN

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

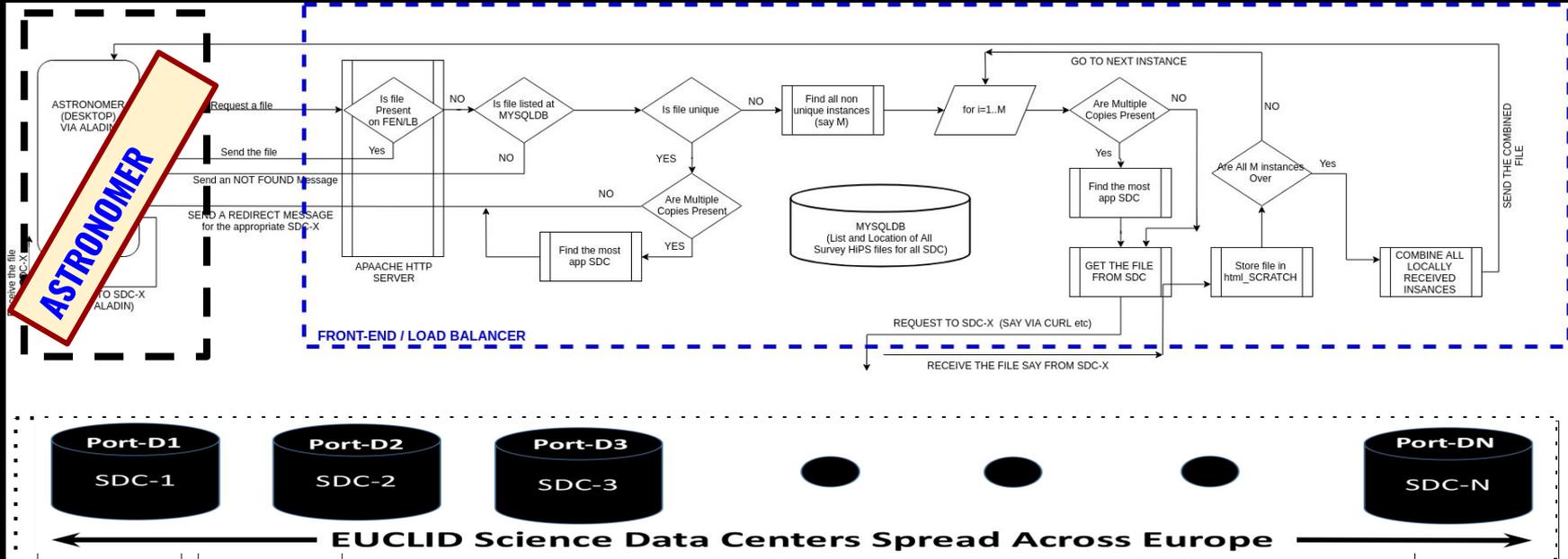


kapteyn astronomical
institute



OmegaCEN

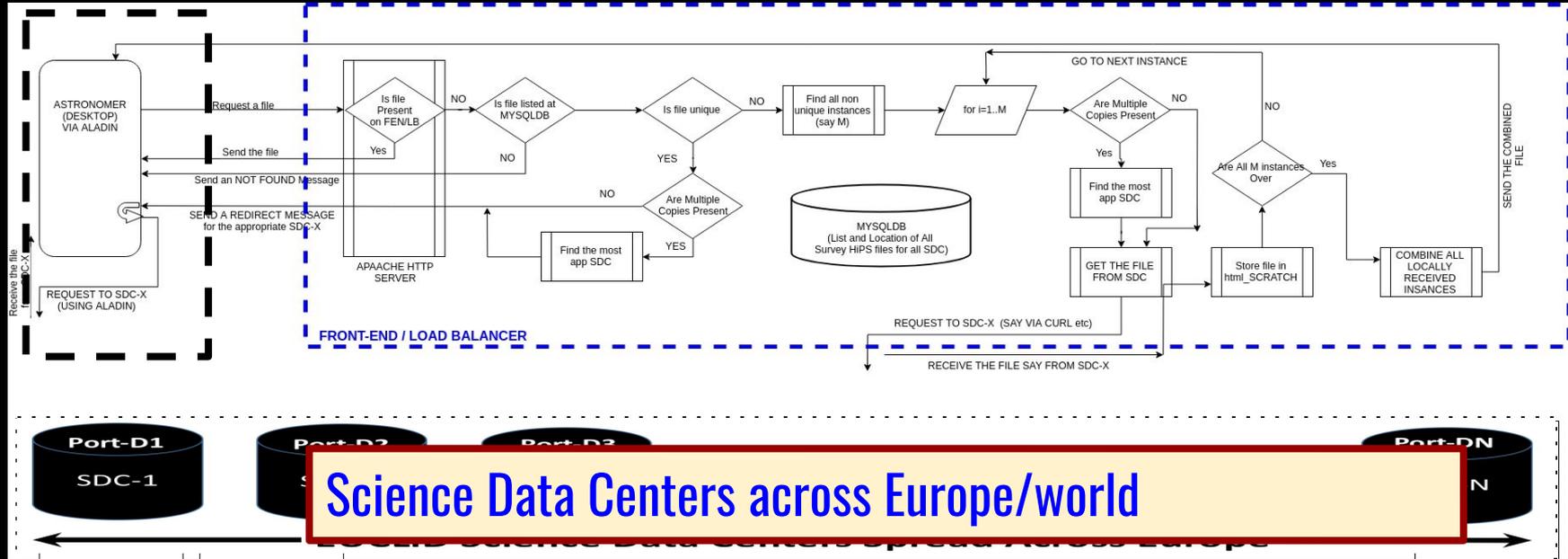
Visualization Framework with LAMP Stack



kapteyn astronomical
institute



Visualization Framework with LAMP Stack



Science Data Centers across Europe/world

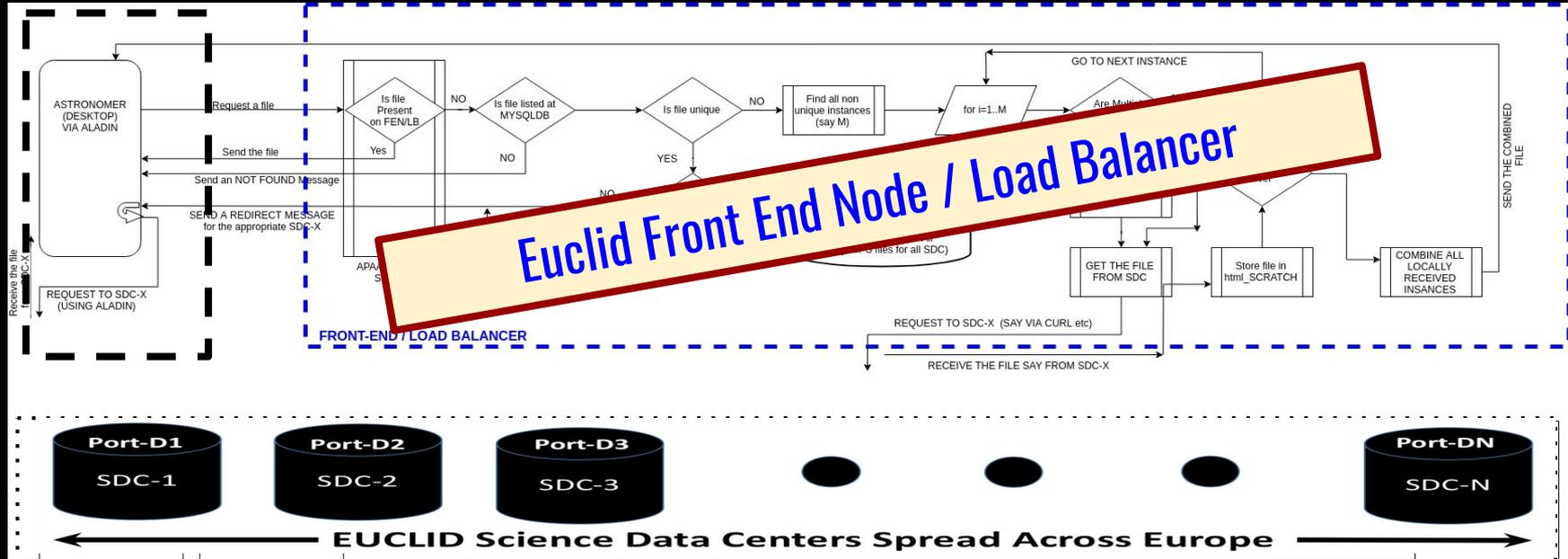


kapteyn astronomical
institute



OmegaCEN

Visualization Framework with LAMP Stack

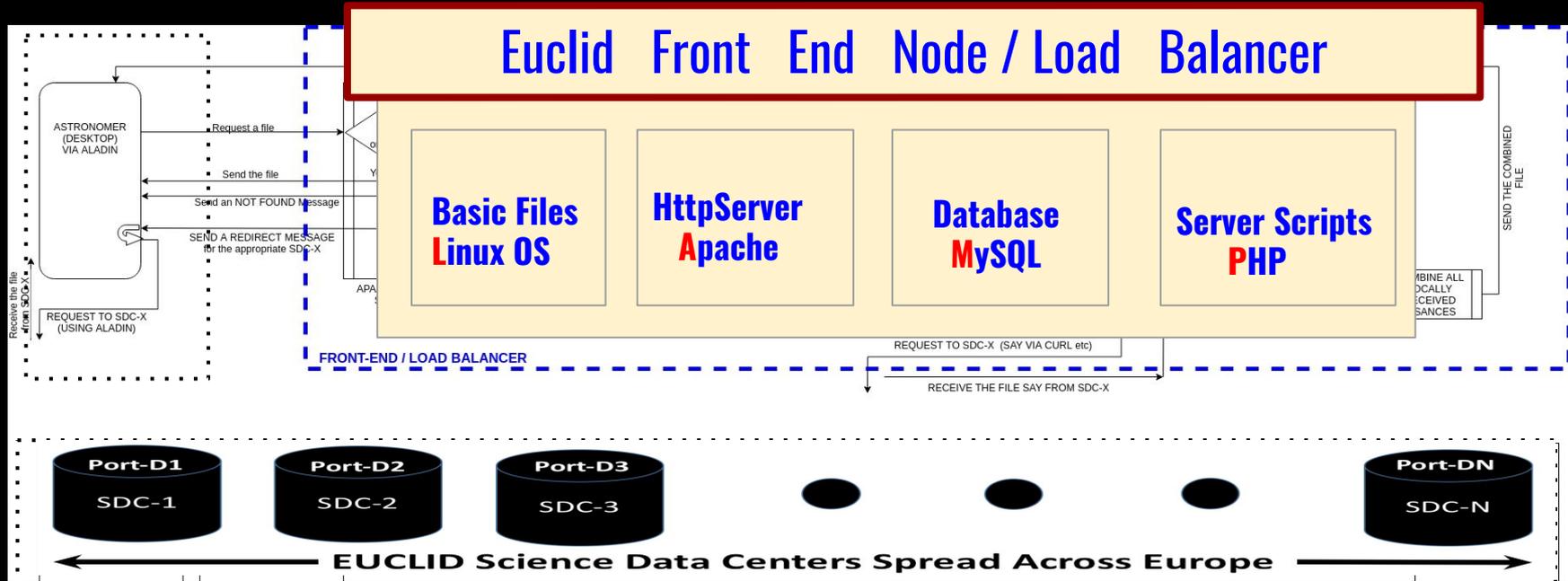


kapteyn astronomical
institute



OmegaCEN

Visualization Framework with LAMP Stack

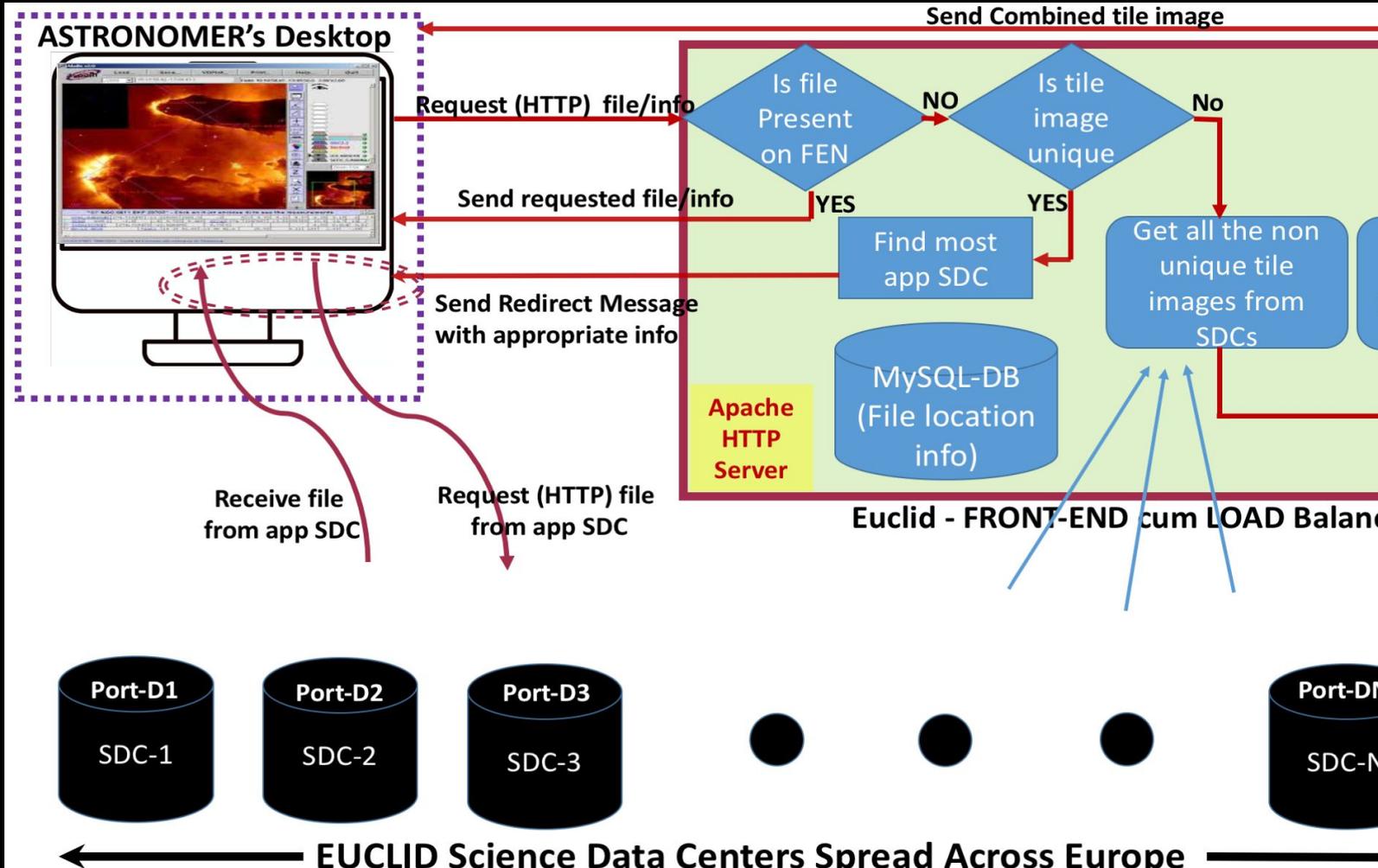


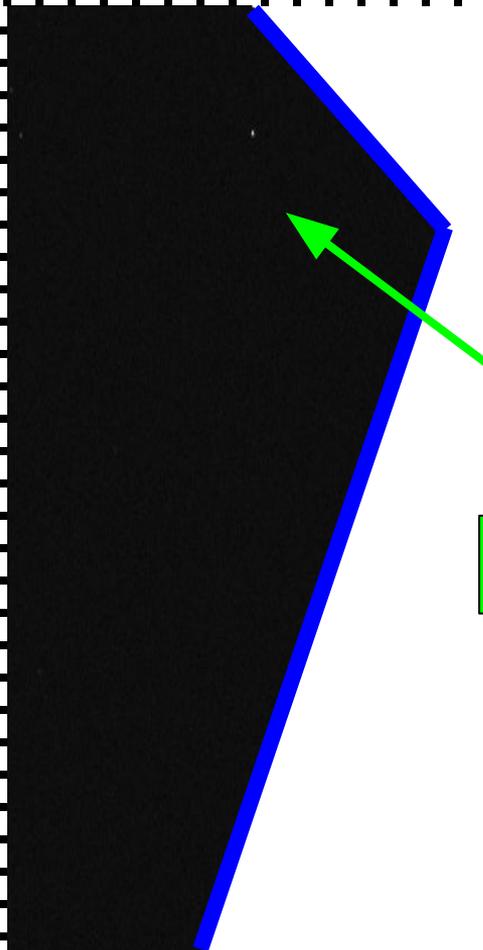
kapteyn astronomical
institute



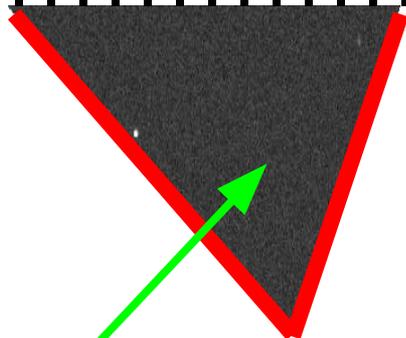
OmegaCEN

Fully Distributed Visualization

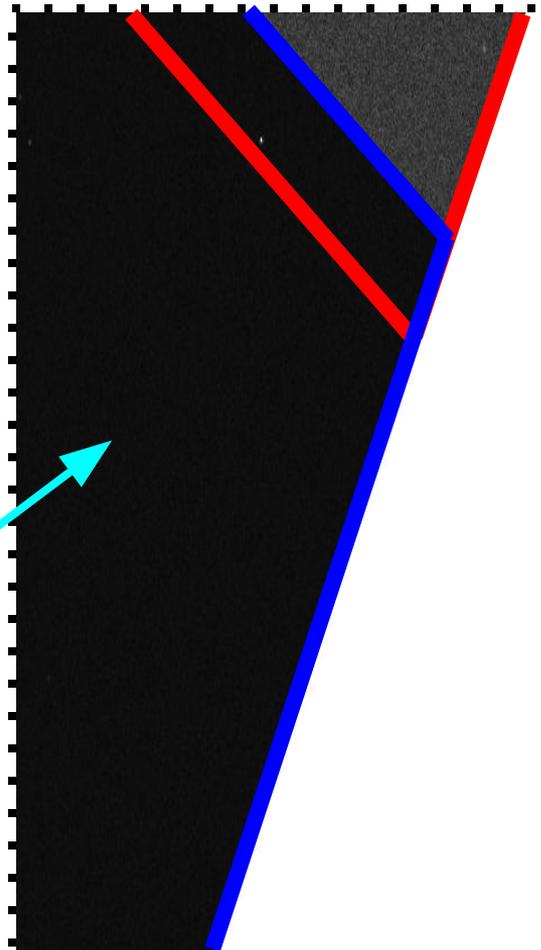




Npix281386.png



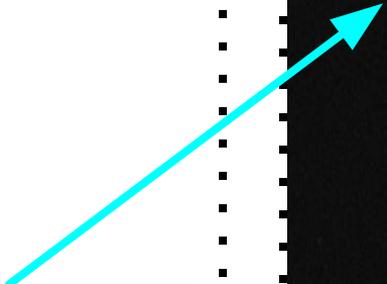
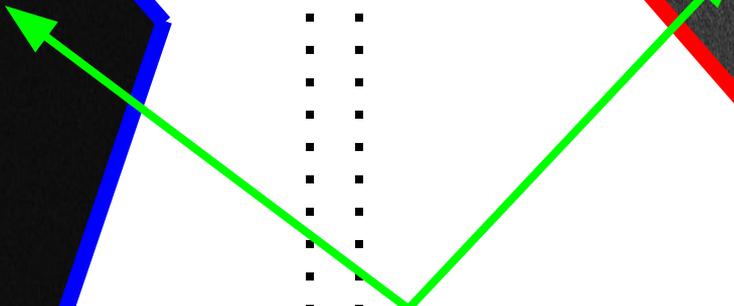
Npix281386.png



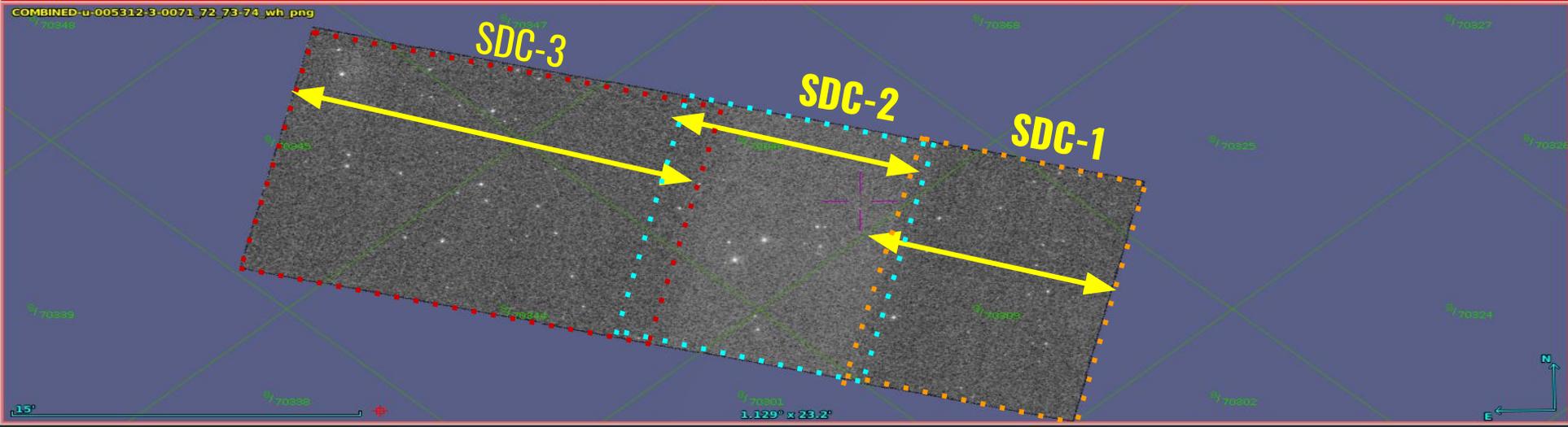
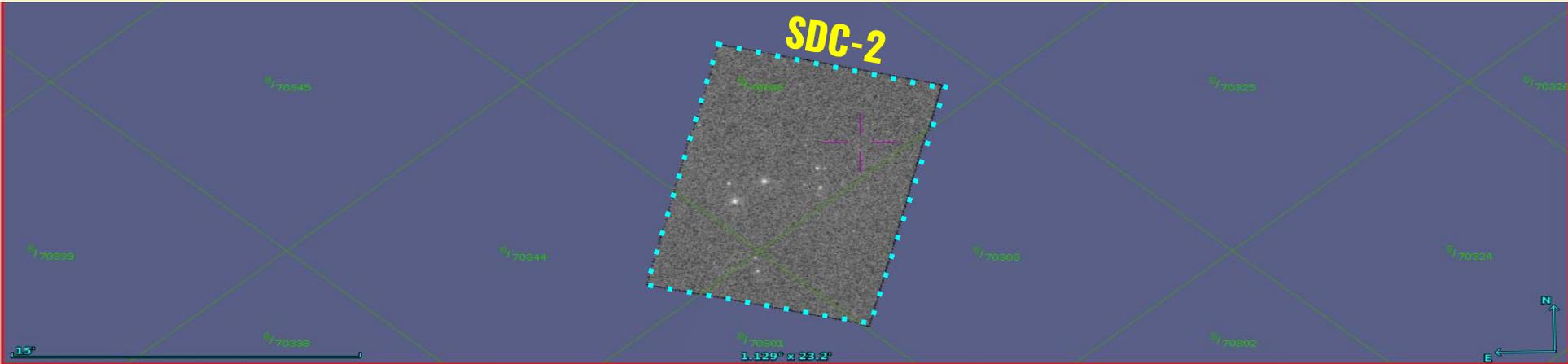
Npix281386.png

Same tile image

Combined tile image



Visualization in practice



Virtual Machines as SDCs

The image displays a virtualization environment setup for Software-Defined Clouds (SDCs). It features the Oracle VM VirtualBox Manager interface on the left, showing a list of virtual machines and their snapshots. The main area shows three running Ubuntu 17.04 VMs, each with a green label in the top right corner: **SDC-1** (hukka3u), **SDC-2** (hukka4u), and **SDC-3** (hukka5u). Each VM screen shows a login prompt for the user 'sweta' with a password field. In the bottom left, a terminal window shows the prompt `[Sun Dec 17-16:34:03-sweta@hukka: ~/REPORT/FIRST_DRAFT]>` and is labeled **FEN/LB** in a green box.

Virtual Machine List:

Name	State	Current State	Snapshot Name	Time
ubuntu1	Powered Off			
ubuntu2 (openssh-server installed)	Powered Off			
hukka3u (cassandra cluster working t...)	Running		ssh setup and login to hukka4u working	12/6/17 10:49 PM (10 days ago)
hukka4u (cassandra cluster working t...)	Running		logins working	12/6/17 11:06 PM (10 days ago)
hukka4u (cassandra cluster working t...)	Running		cassandra cluster working tag 1A	12/7/17 8:08 PM (9 days ago)
hukka5u (cassandra cluster working t...)	Running		cassandra cluster working tag 1B	12/8/17 12:29 AM (9 days ago)

Terminal Output:

```
[Sun Dec 17-16:34:03-sweta@hukka: ~/REPORT/FIRST_DRAFT ]>
```

Virtual Machines as SDCs

From Demo to Real Data

- KiDS DR 4
- 1.5TB (g band)
- On OmegaCen nodes

FEN/L

SDC-3

sweta

Password

ubuntu® 17.04

Right Ctrl

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 hops 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**



Kaptijn Institute
institute



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

Fits → HiPS(png) using Aladin/hipsgen



kapteyn astronomical
institute



OmegaCEN

Thank You !



kapteyn astronomical
institute

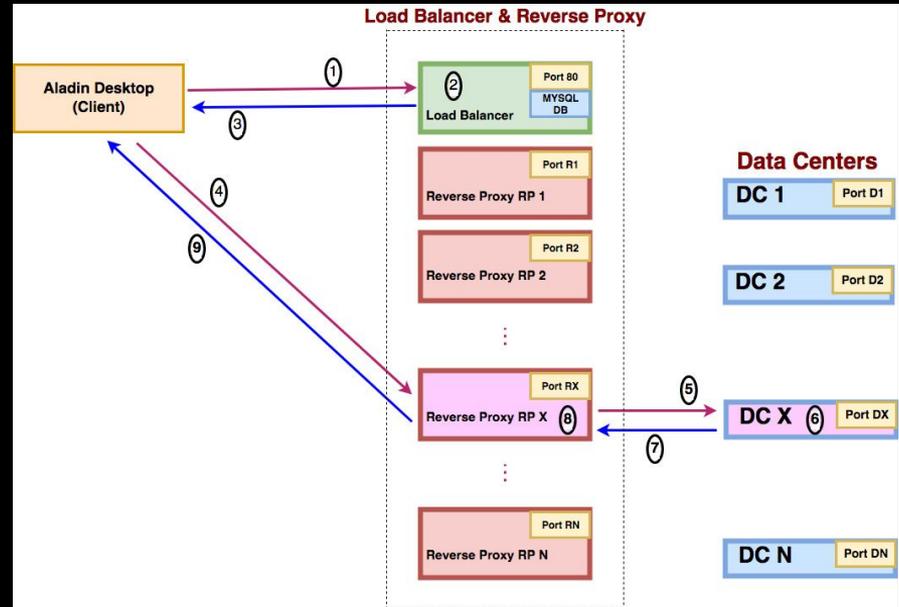


OmegaCEN

Architecture

Distributed Network Mode

Reverse Proxy Network Mode

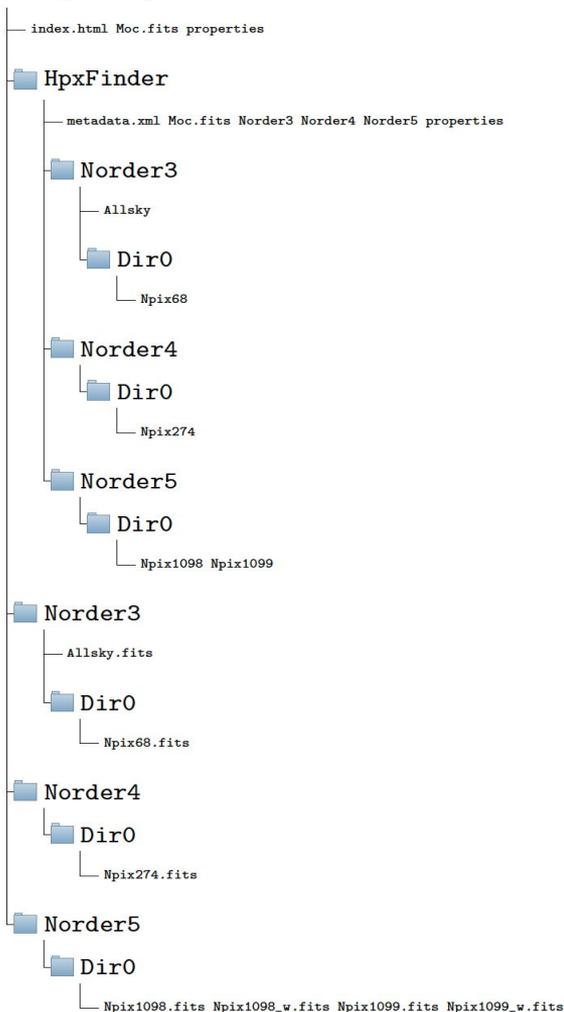


kapteyn astronomical
institute



OmegaCEN

example.hips



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