INAF IA2: data provider experience
including the VIALACTEA Knowledge Base

Marco Molinaro
Cristina Knapic, Andrea Bignamini, Francesco Cepparo,
Sonia Zorba, Robert Butora, Riccardo Smareglia

outline

- IA2 data center
  - data
  - requirements
  - tools & systems
- LBT DA architecture
- TNG & GAPS use case
- VIALACTEA Knowledge Base

IA2 Data Content

- Italian center for Astronomical Archives
  - Ground based observatories
    - TNG, as the main full archive
    - LBT, managing all the consortium distributed infrastructure
    - Asiago Astronomical Observatory
    - Medicina and Noto radio observatories
  - Project dedicated infrastructures
    - GAPS, VIALACTEA
  - Survey data
    - WINGS, VIPERS

IA2 Architecture Requirements

- Common file handling
  - FITS (various flavours)
  - ASCII, GADGET, custom project formats
- Infrastructure needs
  - Data handling/ingestion/distribution
  - Policy handling, User management
  - Resource web deployment/distribution
    - Custom UI and all purpose API (~VO)

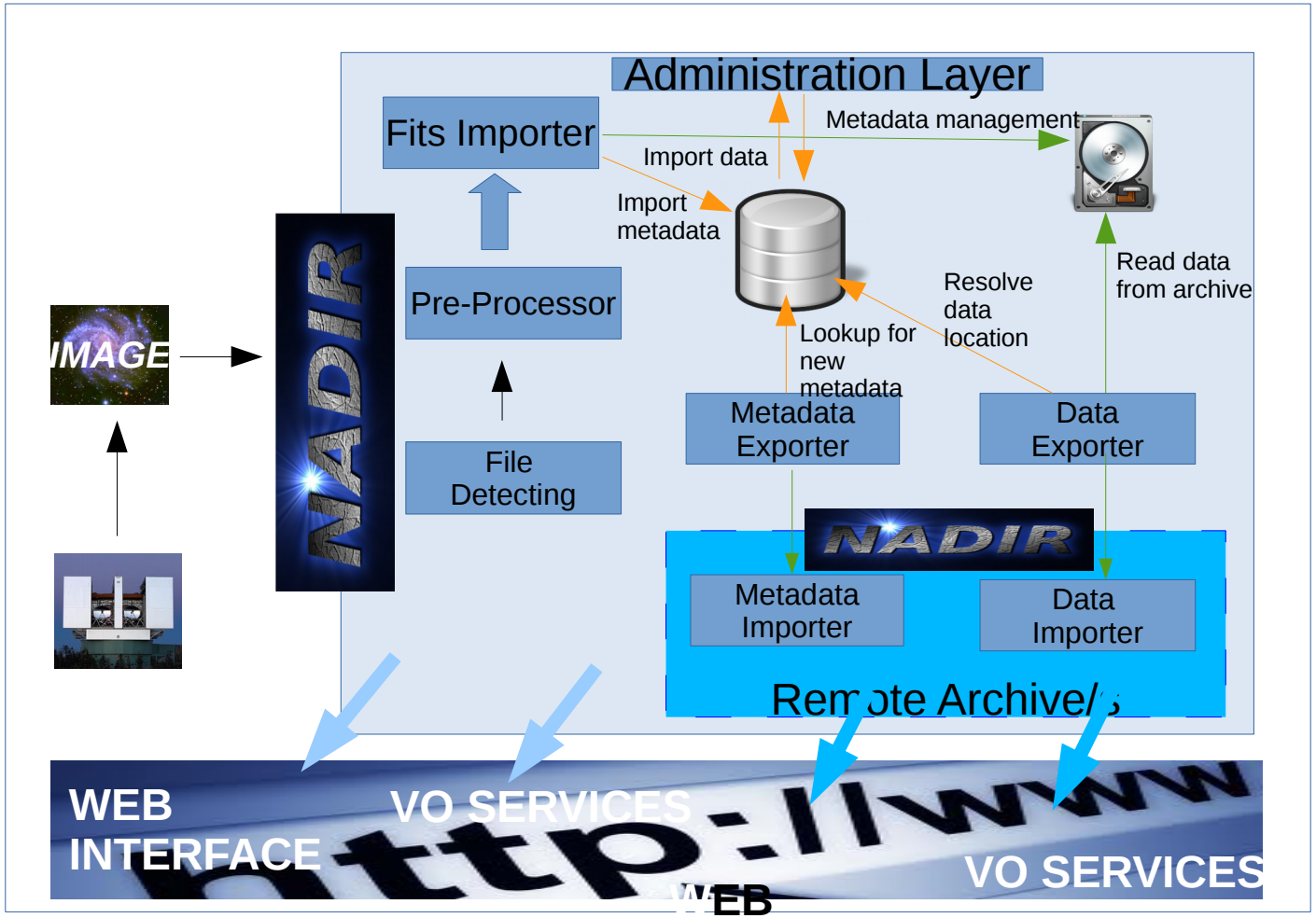
Tools & Systems

- NADIR
  - New Archive Distributed InfrastructuRe
- VO-Dance, IA2TAP, VO-Ball
  - VO inspired resource publishing
  - Under refurbishment
- TS-Man, Portal Generator
  - UI to help configure and generate interfaces
- VAPE
  - VO Aided Publishing for Education
- YABI
  - Workflow manager
- VLKB
  - Independent resource for the VIALACTEA project

NADIR

- Archival software tool aimed to guarantee high scalability, modularity, robustness, error tracking, tools to monitor the status of the archiving system
- The archiving system is composed of several blocks (devices) in a TANGO Controls System
  - an open source device-oriented controls toolkit for controlling hardware or software devices
  - operating system independent and supports C++, Java and Python for all of the components
  - A device is logically seen as an entity to be controlled with public attributes and private properties
  - Attributes and commands are defined to control and monitor the device, its status and configuration
  - TANGO devices for NADIR are archiving software objects

NADIR


Administration Layer
Fits Importer
Metadata management
Import data
Import metadata
Pre-Processor
Resolve data location
Read data from archive
Lookup for new metadata
IMAGE
NADIR
File Detecting
Metadata Exporter
Data Exporter
NADIR
Metadata Importer
Data Importer
Remote Archive/s
WEB INTERFACE
VO SERVICES
http://www
VO SERVICES
WEB INTERFACE

LBT Distributed Archive

Source point of information:

- LBT telescope site – temporary archive/repository
- Ingestion site for standard FITS files of raw, technical or guiding data
- Limited storage capacity
- Full metadata DB since the first operations.

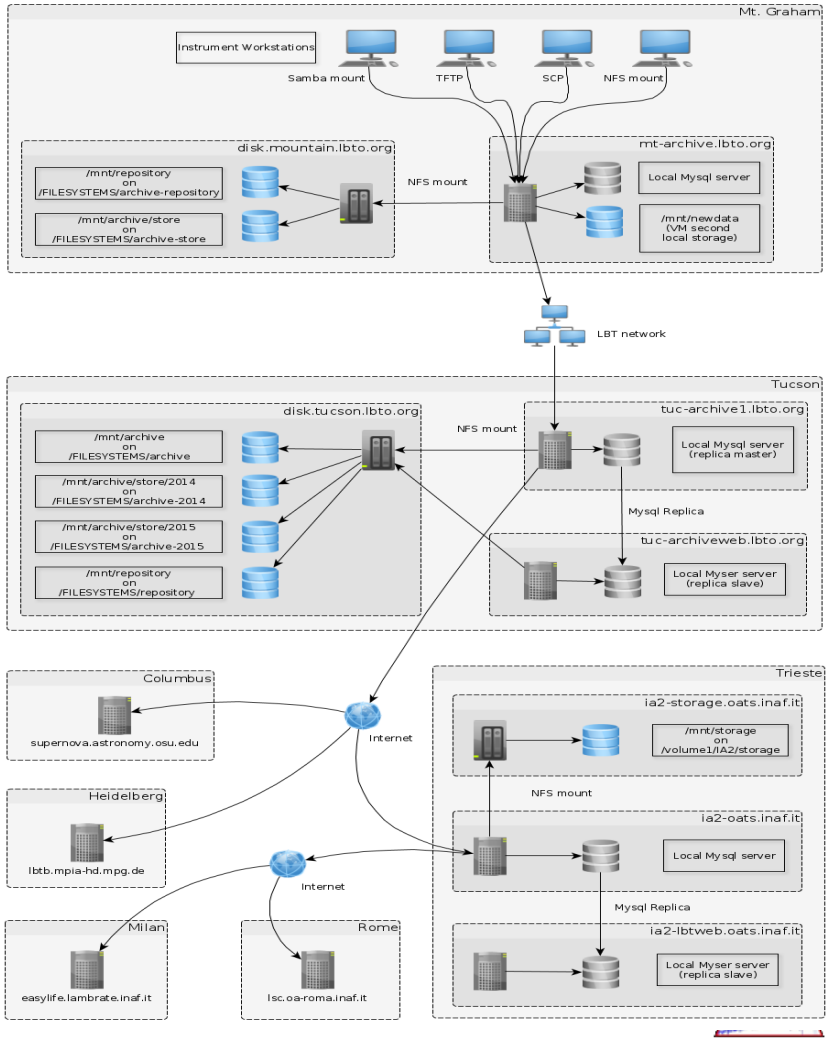
Master archive in Tucson:

- Full archive (metadata and data) for all partners and all data
- In the same place, first data reduction for some data, data quality evaluation foreseen in the near future (interaction with operation team)
- Data distribution to partners both to other LBT distributed Archive sites and to externals (data delivery, not archive)

Distributed sites
(MPIA – Heidelberg; OATs – Trieste)

- Partial archives with proprietary data
- Data reduction with proprietary pipelines

Data delivery to other archives
(Ohio State University – Columbus; INAF – IASF – Lambrate (Mi) ; INAF – OARm - Roma)

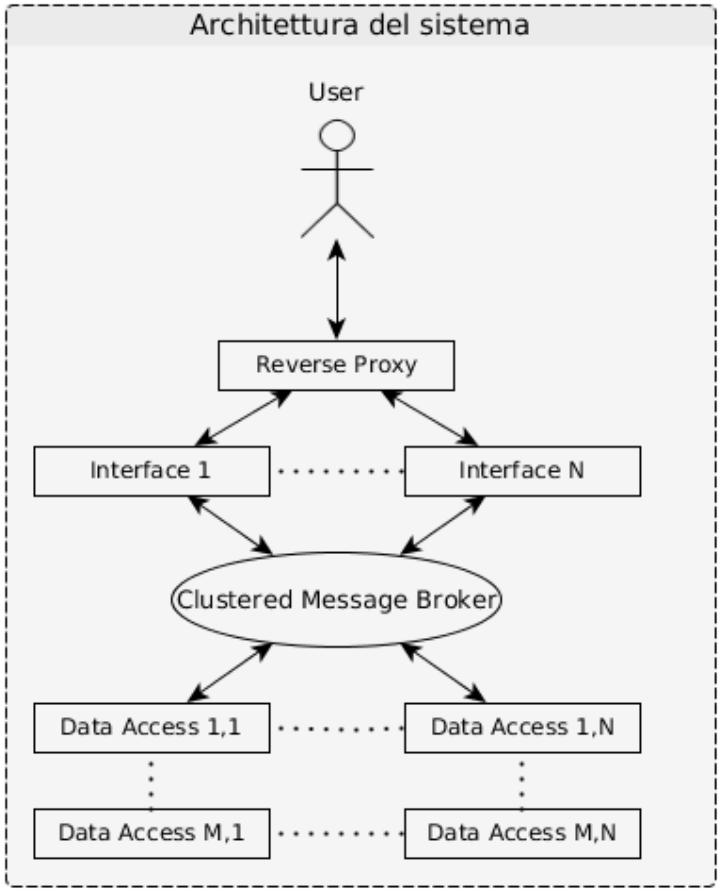


VO inspired tools

- "old" generation
  - VO-Dance
    - Monolithic Java EE
    - Worked on the assumption of "simple" VO services
  - IA2TAP
    - Broke the "simple" scenario
    - Software integration solution

VO inspired tools

- "new" publishing architecture
  - Currently *VO-Ball*
    - Modular
    - Message driven
    - Distributable
    - H&V scalable
    - Language agnostic
      - (as far as AMQP allows)
  - Will supersede VO-Dance, IA2TAP and help the VLKB



Cepparo & al. (SPIE 2016)
A distributed infrastructure for publishing VO services: an implementation (9913-116)

User Interfaces

- Help configuring and deploying services
  - VAPE: VO Aided Publishing for Education
    - NADIR and VO-Dance based
  - TAP_SCHEMA Manager (TS-Man)
    - GUI to help fill in a TAP_SCHEMA out of a DB
    - Includes UCD helper
  - Portal Generator
    - Piggyback fork of TS-Man
    - Configure/generate a web form front end for archives
      - VOTable, SAMP, UWS capable

Zorba & al. (SPIE 2016)
Aided generation of search interfaces to astronomical archives (9913-157)

A couple of figures

Default Statistics | Ancillary Statistics | Custom Statistics | Geolocation

Grouped by protocol

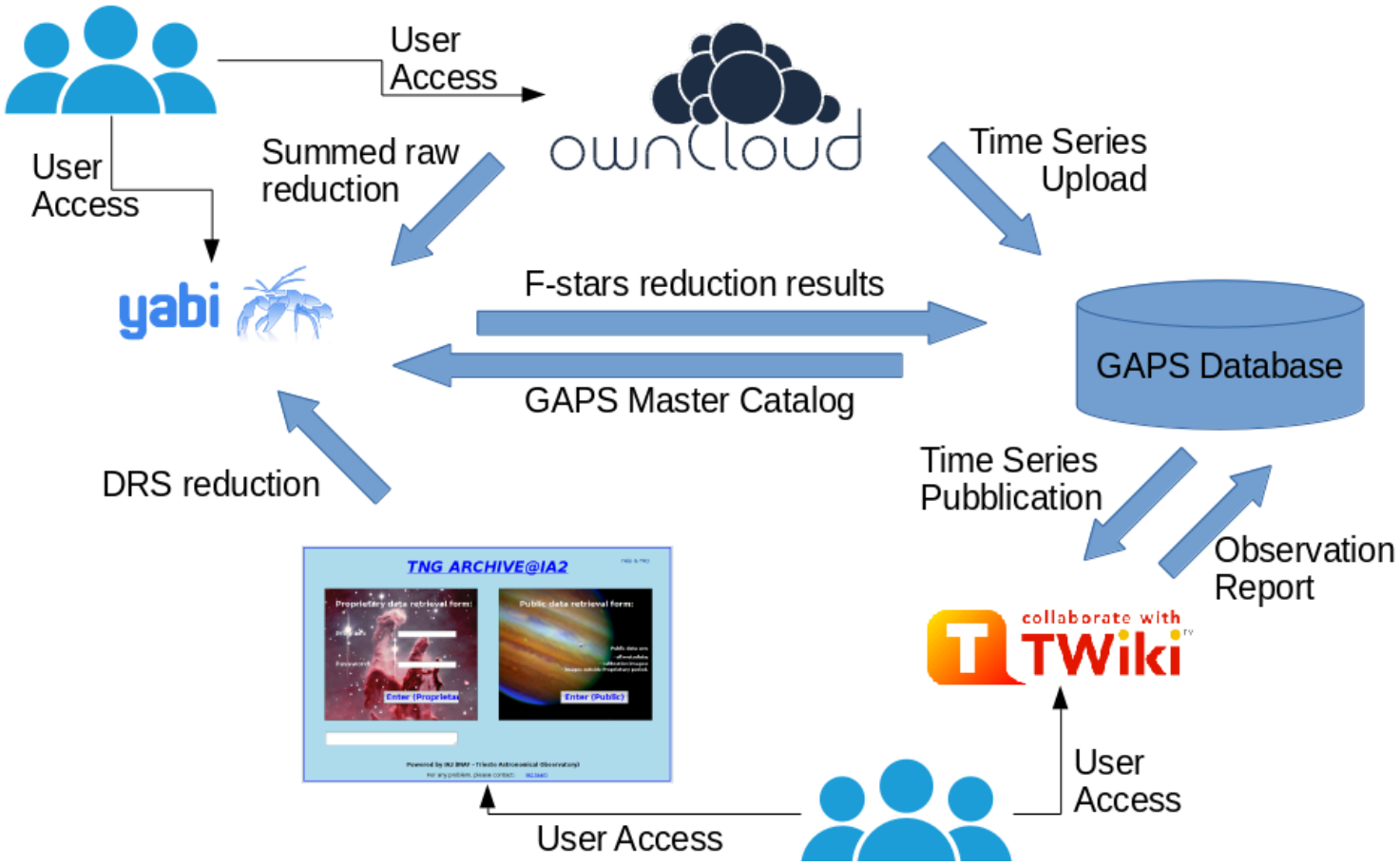
☑ Day ☐ Week ☑ Month ☐ Trimester ☑ Year
☑ Ok ☐ Error ☐ Local

Search...

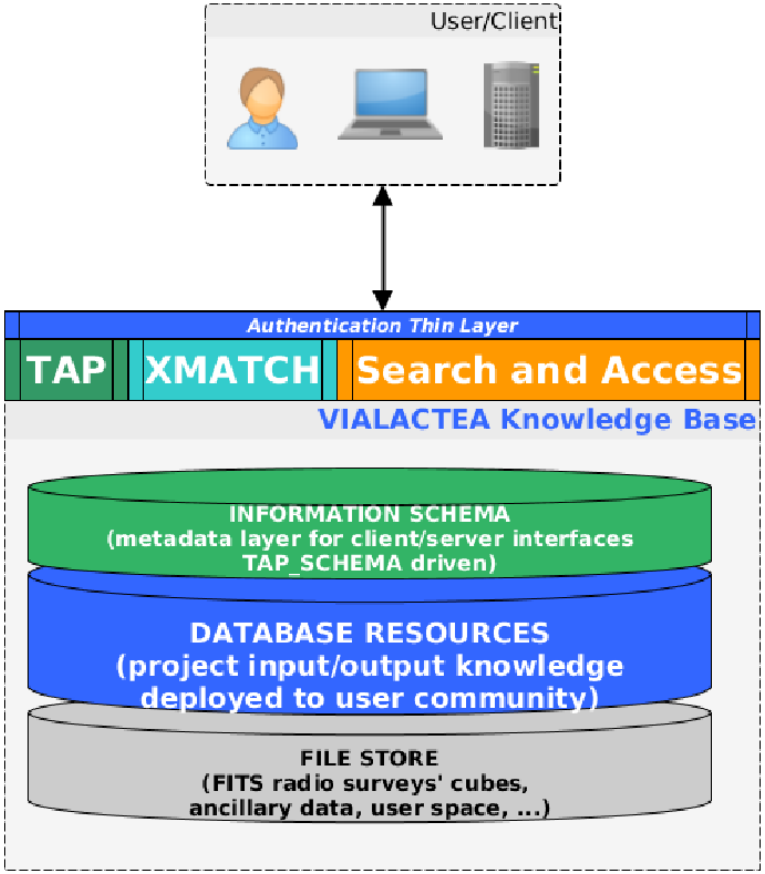
Protocol	Services	Number of Accesses		
		Day	Month	Year
		Ok	Ok	Ok
cone	27	213	3511	67013
siap	7	36	644	27133
ssap	1	11	156	2380
tap	4	38	1979	36228
web	7	108	3570	122570

Disclaimer: I'm cheating, I've hidden errors and local calls

TNG & GAPS architecture



VIALACTEA Knowledge Base



- (VLKB) On top of
  - 2D radio continuum images
  - 3D radio velocity cubes
  - 3D extinction maps
  - Compact sources catalogues
  - Diffuse structure catalogues
  - SED numerical models
  - Derived velocity profiles
- Galactic sky frame based
- Search & access scenario

If needed: Via Lactea = (the) Milky Way

VLKB Search&Access

- Features
  - TAP interface on top of all database content
  - Positional discovery on circle and boxed region
  - Velocity range and survey filtering
  - 3-axes cutout and/or merge
- Will see (near future)
  - Tessellation based cross-match between compact and diffuse objects
  - VO-Space solution for initial and post-processed use data

Molinaro & al. (SPIE 2016)
VIALACTEA knowledge base: homogenizing access to Milky Way data (9913-17)

VIALACTEA – The Milky Way as a star formation engine - Conference
http://vialactea2016.iaps.inaf.it/

...and moving on...

- Authentication & Authorization
  - GMS & Grouper on LDAP solutions
- DOI for data resources
  - Data preservation

Thank you!

Marco, Cristina, Andrea, Francesco,
Sonia, Robert, Riccardo