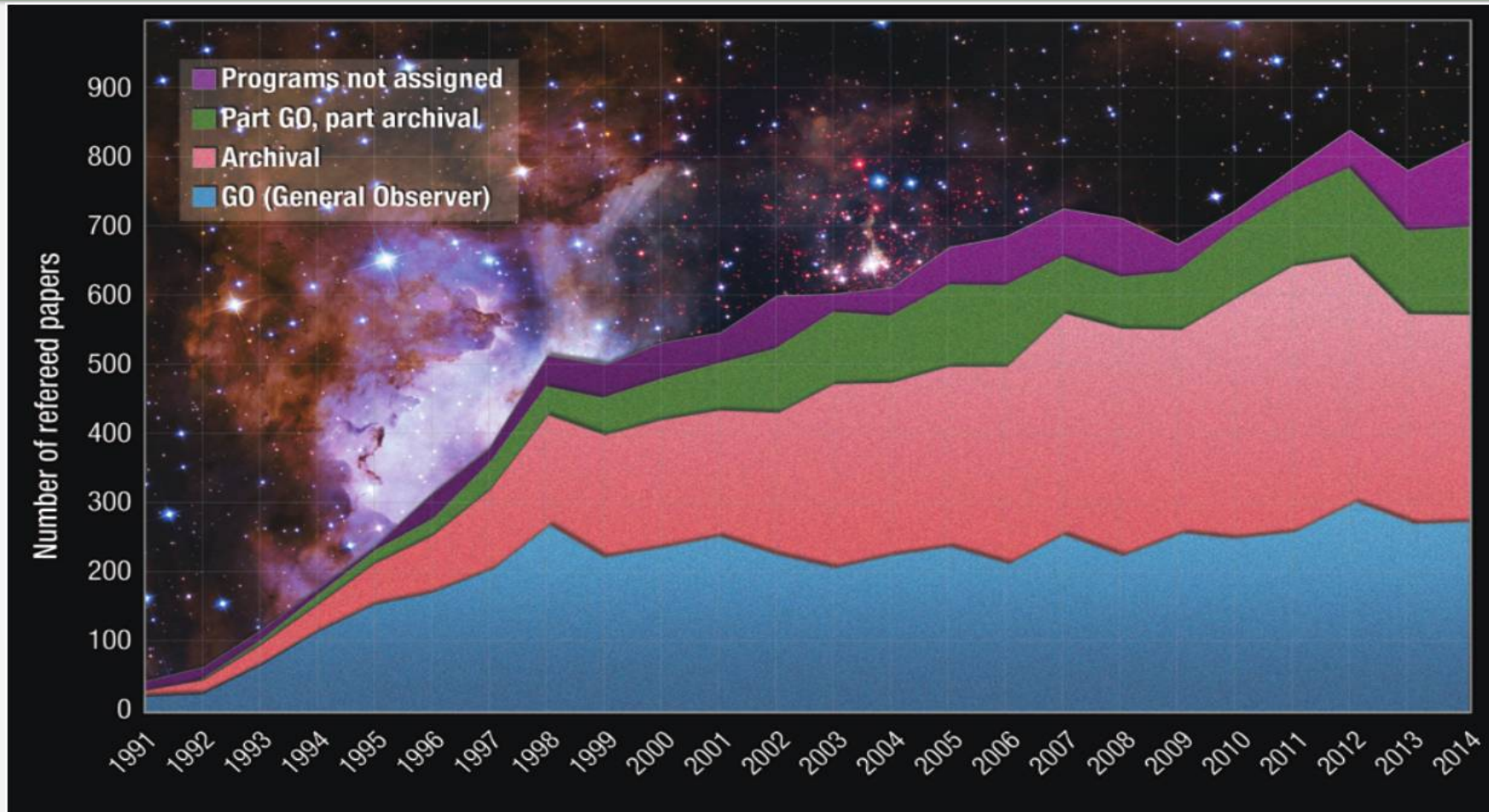# TNG Archive - IA2

Cristina Knapic
03/03/2017 – TNG Meeting - Padova

Archive importance

HST Newsletter: "*At the present time, approximately **half of the refereed publications** based on Hubble observations are derived purely **from archival data**, and, every year, this number is slightly higher than the number of publications based on new observations. .... the Hubble Archive has become a goldmine for the astronomical community....*"

IA2 Main Goals

Centro Italiano Archivi Astronomici (IA2) main goals are :

- archiving systems
- safety
- data curation and preservation
- distribution over several geographical distributed sites
- providing services and tools (TWiki, work-flow, etc..)
- data publication in the VO

 of *Astronomical Data*

IA2 manages data of several PROJECTS.
Mainly they come from:

- TELESCOPES (raw; INAF ground based)
- SURVEYS (raw and/or calibrated)
- SIMULATIONS (ITVO)
- EU Projects (GENIUS, VIALACTEA, INDIGO)

IA2 is an INAF e-Infrastructure
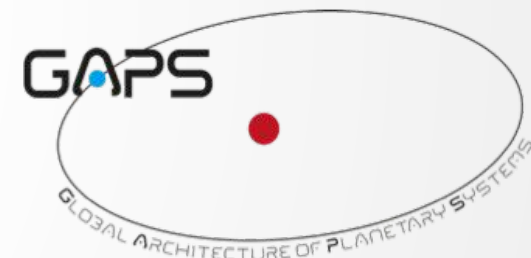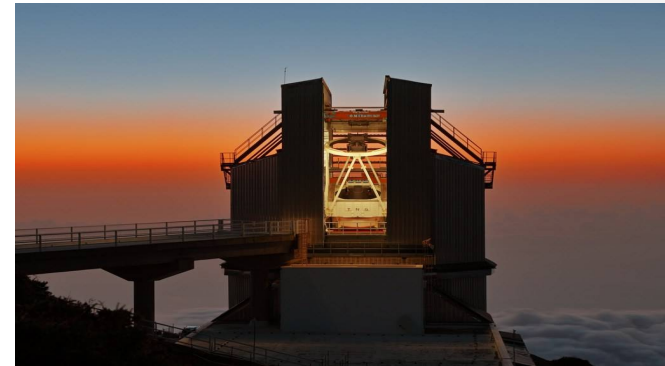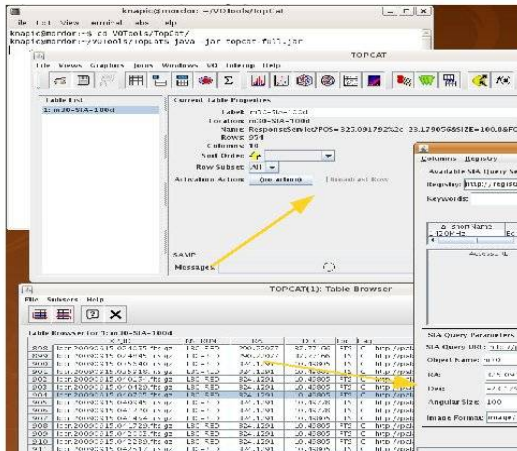
IA2 components applied to TNG Archive

- IA2
  - Main Goals
  - Hosted Services
    - NADIR
    - Workflow (Yabi)
    - Collaborative tools (Twiki, OwnCloud...)
  - Observatory Archives (**TNG**, LBT, Asiago, Radio, ExoClimate, Serra La Nave....)
  - VO compliant services (SIAP, SSAP, TAP, Cone..)
  - Some examples
  - Projects hosted @ IA2 (**GAPS**)
- Services in debug
  - Single Sign-On
  - VOSpace, user space
- Future plans

IA2 Services

IA2 Services

Service ▲	Protocol ⇕	Number of Accesses				
		Day	Week	Month	Trimester	Year
		Ok	Ok	Ok	Ok	Ok
aaotap	tap	0	0	0	0	4
⊟ Asiago Archive	web	7	49	79	454	1221
		• Downloaded files: 0	• Downloaded files: 187	• Downloaded files: 982	• Downloaded files: 11925	• Downloaded files: 25355
DAbrusco2009QSO	cone	3	28	107	361	1841
DAbrusco2009QSORedshift	cone	2	24	107	352	1869
epntap	tap	38	260	1394	2622	19190
Laurino2010	cone	2	24	97	347	1551
Laurino2011_optUV_QSO	cone	2	23	104	370	2004
⊟ LBT Archive	web	74	502	3972	8812	31619
		• Downloaded files: 24	• Downloaded files: 162	• Downloaded files: 1275	• Downloaded files: 2783	• Downloaded files: 9478
⊞ Owncloud	web	17	35	69	100	188
Planck_ECC	cone	1	18	84	325	1776
Planck_ERCSC_f030	cone	1	20	88	312	1666
Planck_ERCSC_f044	cone	1	20	88	312	1621
Planck_ERCSC_f070	cone	1	20	82	303	1536

Service	Protocol	Day			Week			Month			Trimester			Year		
⊞ Asiago Archive	web	1	0	0	15	0	0	137	0	0	488	0	23	1589	0	102
⊟ LBT Archive	web	72	0	0	627	0	0	2501	0	13	7368	0	13	30772	0	16
		• Downloaded files: 17			• Downloaded files: 161			• Downloaded files: 734			• Downloaded files: 2252			• Downloaded files: 9475		
⊞ Owncloud	web	26			51			92			137			228		
⊟ TNG Archive	web	48	0	0	198	0	3	923	0	8	1987	0	43	7739	0	302
		• Downloaded files: 31			• Downloaded files: 163			• Downloaded files: 19724			• Downloaded files: 38547			• Downloaded files: 134322		

Service	Protocol	Day	Week	Month	Trimester	Year
TNG-NICS	siap			15	376	2842
TNG-OIG	siap	0	0	17	337	1934
⊞ Vape	web	0	0	1	1	32
⊞ Vialactea Cutout	web	0	0	0	0	36167
⊞ Vialactea Search	web	0	0	0	0	34451
VIPERS SSAP	ssap	56	345	2703	3327	5137
vlkb	tap	0	0	0	286	4351

TNG Archive 2016 Map of Accesses
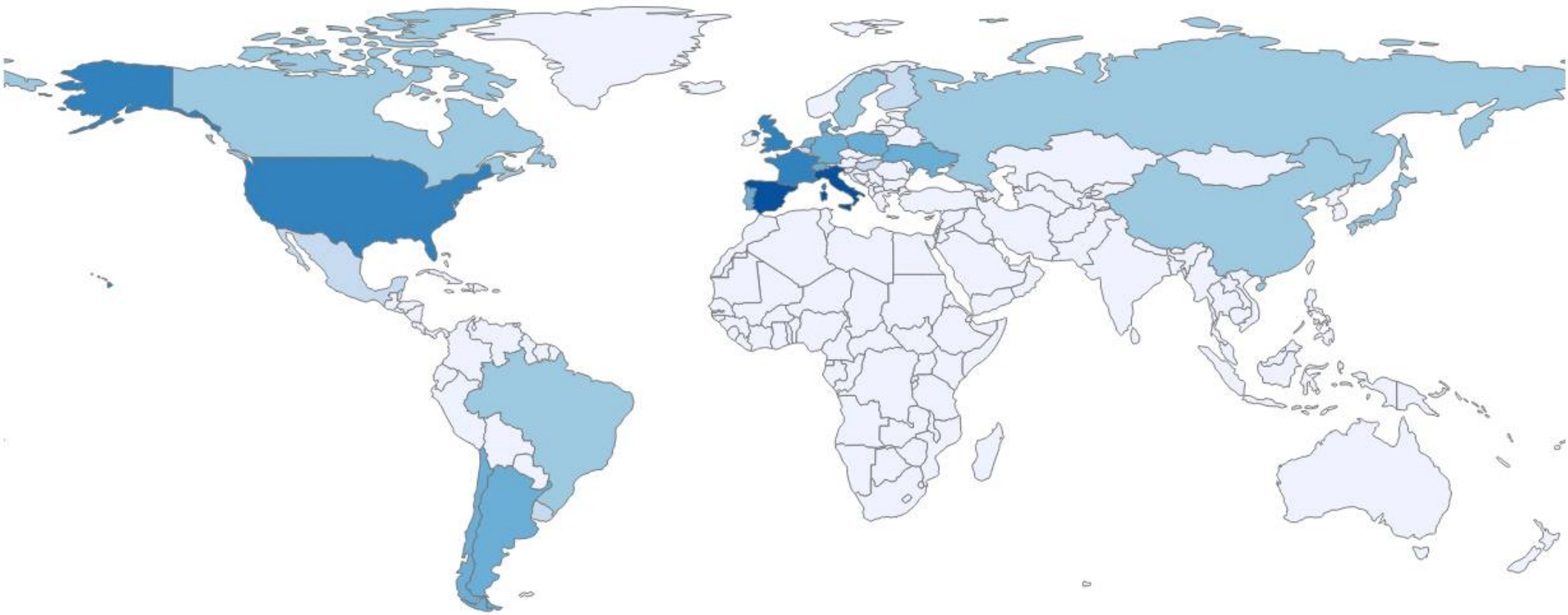
ITA: 2879	NLD: 161	DNK: 71	RUS: 26	URY: 8	FIN: 3	LTU: 1
ESP: 1367	CHE: 148	PRT: 70	SWE: 24	MEX: 6	ROU: 2	IND: 1
GBR: 517	ARG: 94	POL: 68	CHN: 24	BEL: 6	AUS: 2	ARM: 1
USA: 488	DEU: 92	UKR: 61	CAN: 17	ISR: 5	TWN: 1	
FRA: 288	CHL: 76	BRA: 28	JPN: 14	HUN: 5	MUS: 1	

TNG foreseen Data flow

- Source point of information:
  - TNG 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 Trieste:
  - Full archive (metadata and data) for all raw and reduced data (es. GAPS)
 - In the same place, first data re-reduction for some data
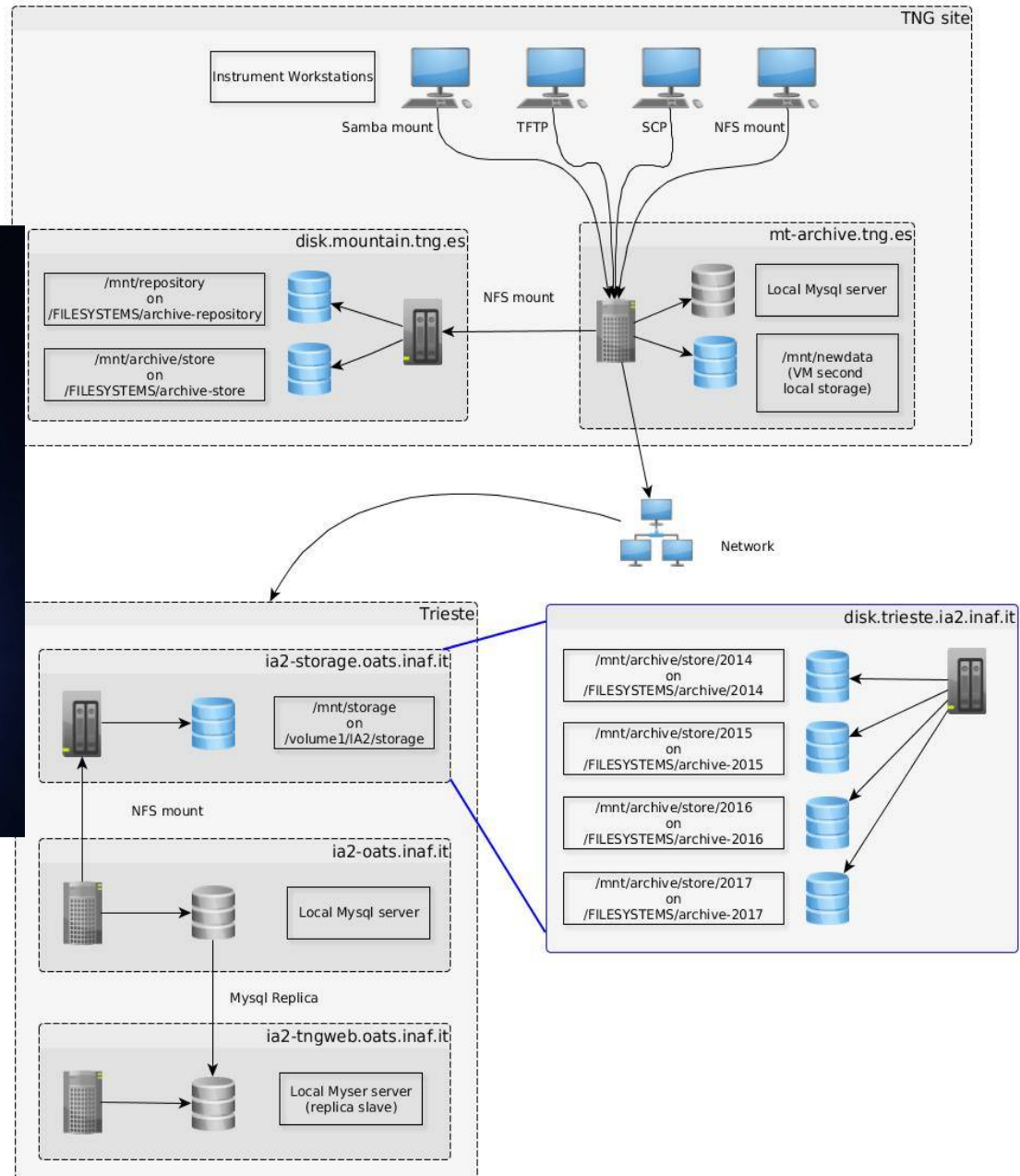
GAPS@IA2
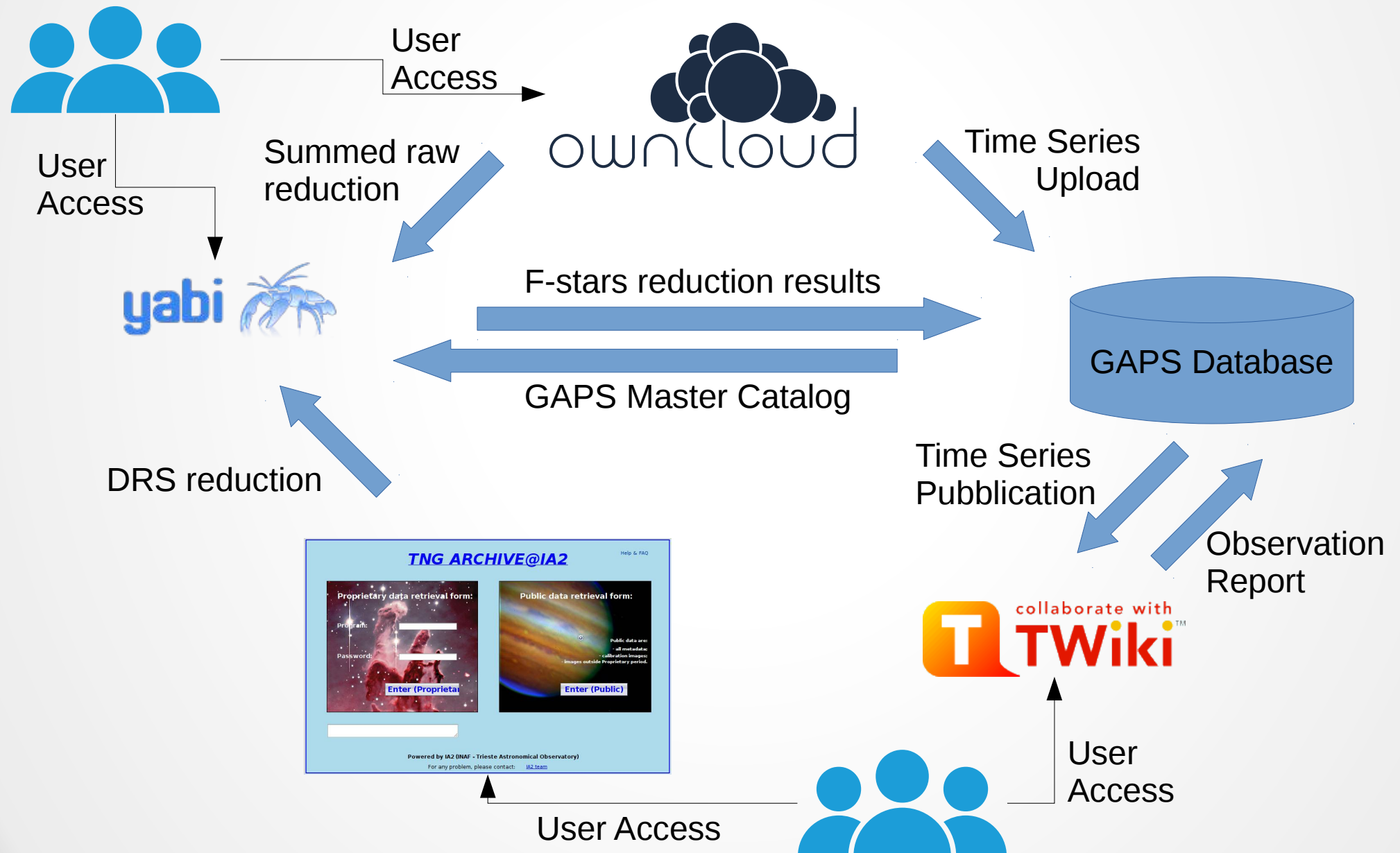
IA2 Workflow system : YABI

- Yabi is a 3-tier application stack to provide users with an intuitive, easy to use, abstraction of compute and data environments. Developed at the Centre for Comparative Genomics and Murdoch University, Yabi has been deployed across a diverse set of scientific disciplines and high performance computing environments.

- Yabi is a workflow system that allows accredited users to run part or full HARPN pipeline

- Yabi has a few key features:
  - simplified web based access to High Performance Computing
  - easy tool addition and maintenance
  - handling of disparate compute and storage resouces ie. PBSPro, SGE, Torque, SSH, SFTP, Amazon S3, Swift
  - easy and powerful workflow creation environment

- Reference: https://ccgapps.com.au/yabi/
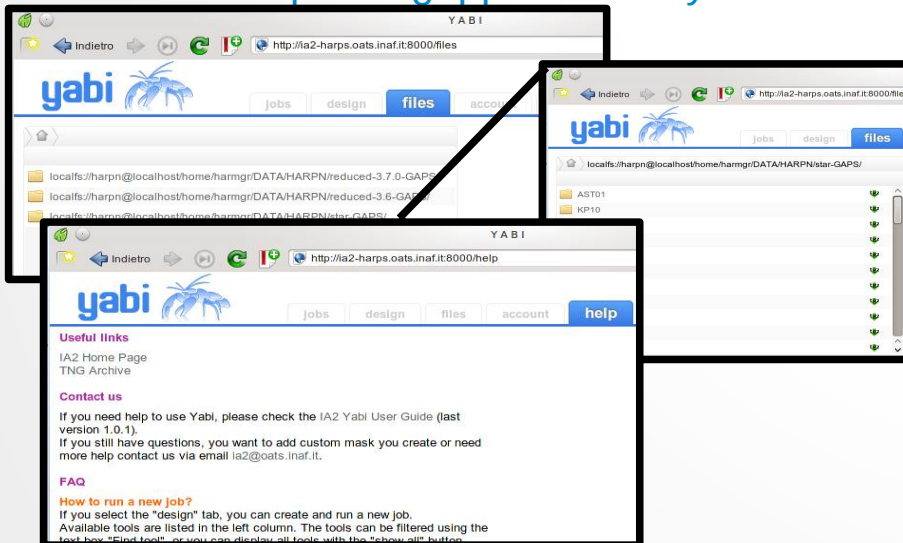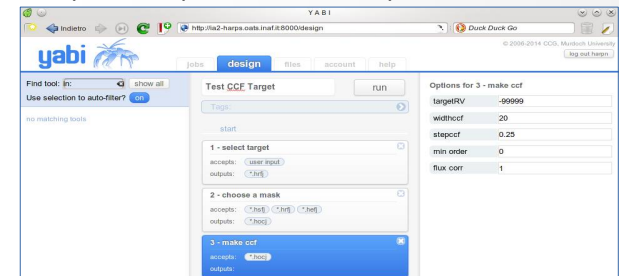
Yabi is running since 2014-03-23 at IA2 with 105 total number of users. Statistics:
Total size of workflow outputs 778 GB

| Description	Nr. Workflows	Nr. Jobs
Total	4019	11000
Last Year	1200	3217
Last Month	117	298
Last week	5	15

GAPS@IA2 - HARPN-DRS

- HARPN-DRS pipeline installed at IA2 ⟵ Multiple version support!
  - Version 3.6 installed in January 2013
  - Version 3.7 installed in February 2014
- Raw data are automatically reduced in Trieste and saved on the Yabi storage (available 12-16 hours after observation)
  - Data reduced at TNG → TNG archive
  - Data reduced in Trieste → Yabi
- Data before 2014-02-14 reduced with both version, while after 2014-02-14 reduced only with 3.7 version
  - Version 3.6 is no more maintained
- In the Yabi interface directories GAPS users can find:
  - Reduced data subdivided by observation night
  - F-Star folders

SSO activities

Authentication (SAML based)
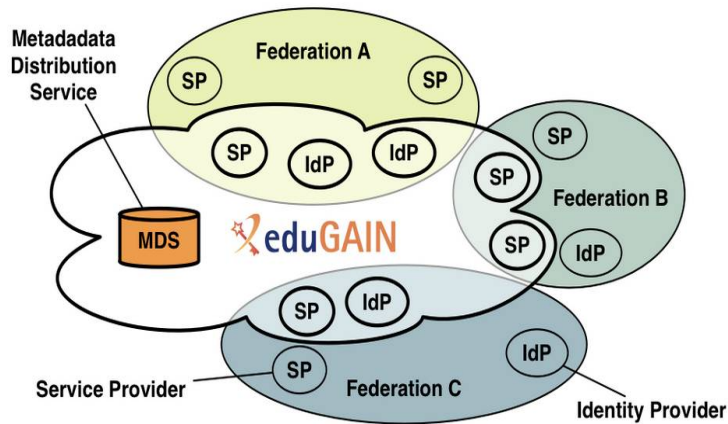
Authorization Internet2 application based / VO compliant based

Future plans of interoperability:

What about RAP????

- IA2 IdP is based on a replica of the Bologna LDAP that provide Digital Identities to INAF (IDEM);
- IA2 integrate the LDAP with non IDEM identities that use to one or more IA2 Telescopes;

Courtesy of F. Tinarelli

IA2 Archives – new approach

TNG credential creation in Archive

Based on accepted Proposal Principal Investigator references, **TNG administrators** could add new users to the TNG archive Identity provider and grouping management system (Grouper)

TNG Users Manager - Admin Panel

Multiple users

CSV file

[Choose File] No file chosen

[Process file]

Valid CSV example

```
"PROGRAM_X","Name Surname","test@test.com",1,"The Address","12.0","A","Call","magazzu@tng.iac.es"
```

Single user

⦿Search existing user ◯Insert a new user

User [type e-mail or user name]

You can search by username, but the selection will use the e-mail, so, if multiple users share the same e-mail all of them will be added into the selected program.

TNG credential creation in Archive

Single user

◉ Search existing user ◯ Insert a new user

User | cristina.

You can search by username, but the selection will use the e-mail, so, if multiple users share the same e-mail all of them will be added into the selected program.

bernasconi.cristina (cristi...

cristina.bernasconi (cristi...

cristina.botta (botta@oats...

cristina.knapic (knapic@o...

cristina.leto (leto@asdc.a...

TNG Users Manager - Admin Panel

Email	test@test.com	
Uid	mario.rossi	Same as email
Initial password	i3cs016f	Generate a random one
Name	Mario	
Surname	Rossi	

☐ Write also into the PostgreSQL database

◯ Select existent program ◉ Create new program

Program name | TestProg1

TNG Users Manager - Admin Pan...

Email	knapic@oats.inaf.it
Uid	knapic.cristina, cristina.knapic

ⓘ User already exists: user information will be updated only in PostgreSQL database

Name	Cristina
Surname	Knapic

☑ Write also into the PostgreSQL database

Password	u5ebndqc	Gener...
Address		
Schedule time	0.0	
Priority		
Call		
Support e-mail	magazzu@tng.iac.es	

◉ Select existent program ◯ Create new program

Program name | CAT15A_S

Summary

New account will be added to LDAP ou=custom_users,dc=oats,dc=inaf,dc=it

uid	mail	cn	sn	userPassword
mario.rossi	test@test.com	Mario Rossi	Rossi	i3cs016f

New group will be created on Grouper

Group folder	Group name
OATS/TNG_GROUPS	TestProg1

The account will be associated to the group in Grouper

Group folder	Group name		Uid	E-mail
OATS/TNG_GROUPS	TestProg1	⇄	mario.rossi	test@test.com

Following data will be added to PostgreSQL database

name	address	e_mail	usrname	password	program	sch_time	priority	call	support
Mario Rossi		test@test.com	TestProg1	i3cs016f	TestProg1	10.0	2	A	avet@tng.iac.es

Confirmation e-mail

A confirmation e-mail will be send to the address ✉ zorba@oats.inaf.it

ⓘ Mail sender is in testing mode

✔ Confirm ✖ Reject

Future plans

Estratto dal "OdG Coll Dir del 5 ottobre 16_allegato 2) verb. 32.pdf"

Iniziative sul Calcolo (II)

- Revisione di IA2
  - E' in previsione una revisione di IA2 con l'introduzione di un "layer alto" per l'accesso ai dati science ready da parte di una comunità larga.

- Improve the services offered (link to proposal submission?);
- Increase the number of services offered (possibility to **re-process** data? **Pipelines**?);
- Increase the number of supported data models;
- Grant access to all the offered services via SSO;
- Adopt a VOSpace compliant service to allow **user space** utilization;
- **Improve the user experience in order to stimulate the Astronomical Data life cycle**.

Needs

- Feedback from USERS!!
- Citation and references to IA2;
- Continuous interaction with stakeholders: users!!

Conclusions

- Target of IA2:
  - Offer Archival services and maintain them
  - To be ready for Big Data

- Target of the archive user:
  - Use archive services
  - Have his own User Space where to make science!
  - Reporting problems
  - Suggest upgrade/necessities