1) Vizier
2) Content
3) Usage
4) Devs: assoc. data

VizieR Staff and contributors:

Astronomers: P.Ocvirk, C. Bot,
Engineers: G.Landais, T.Boch, F.X.Pineau,
Documentalists: P.Vannier, E.Perret, S.Guéhenneux, M.Brouty.

Non-CDS: L.Michel, J.Y.Hangouet, T.Keller   (Strasbourg Observatory)

What is Vizier ? - I

The vizier **service** provides unified access to a very large collection of astronomical catalogues & associated data:

- Reference catalogues & surveys of astronomical sources at all wavelengths (e.g. SDSS, 2MASS, UCAC, WISE)

- Tables from papers published in the major astronomical journals

- A large variety of astronomical data: astronomical sources , spectra, polarization data, but also models (evolutionary, populations, synthetic spectra, …), statistical analyses, compilations, …

What is Vizier ? - II

Vizier can be queried:
- By catalog name
- By wavelength, Mission name, object type / process, ...
- By keyword (galaxies, quasars, ISM, ...)
- By column description
- By position, through ALL catalogs

The M word

The metadata describes the data:

- Type of data (catalog, spectra, light curves etc...)

- Units

- Wavelengths / filters / instrument

- Magnitude system, Coordinates system

… with <u>homogenized descriptions (metadata)</u>

> Goal: **Discoverability**, **reusability**, platform for higher level services (phot. Viewer, associated data, ...)

VizieR in statistics

Statistics in input

- A regular increasing volumetry

  - 15,366 catalogues, 32,991 tables, 19,2G records (95 % coming fom large surveys)

  - 1,270 ingested catalogue in period sep 2015 – oct 2016

New catalogues ingested

VizieR in statistics
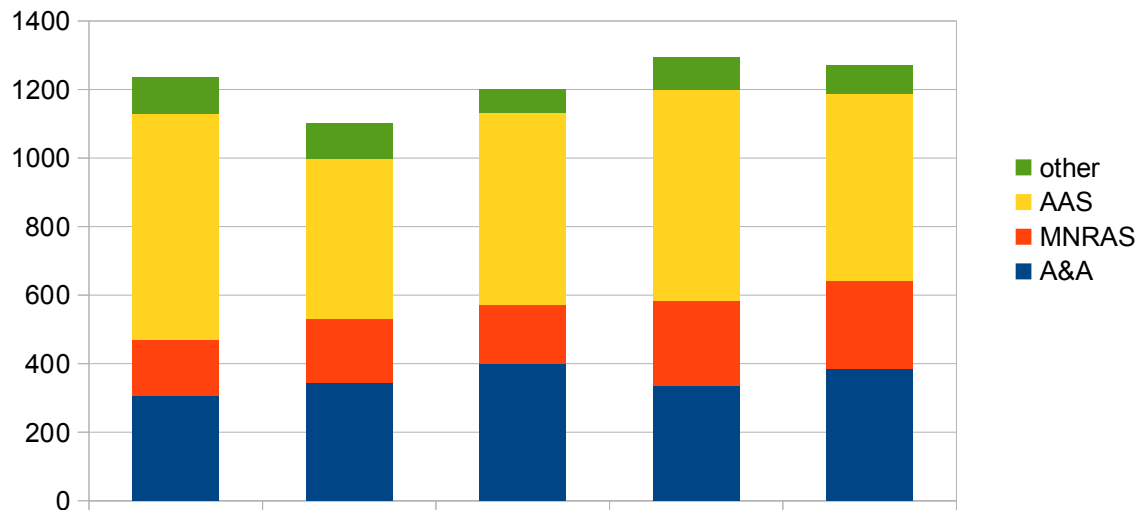
Datasets regulary updated

- ATNF pulsar catalogue (Radio)

- Occultation lights curves (logs of observations)

- (In the pipeline) CFHT, HST :
  observation logs provided by the CADC (Canadian Astronomy Data Center)
  a new VO-pipeline based on the standard TAP

Large surveys

3 next slides courtesy of T. Boch

Large catalogues: what's new?

- New organization of the large catalogues team
  - for each big catalogue to be ingested:
    1 astronomer + 1 software engineer + 1 documentalist
  - focus on each other's expertise
  - speed up the overall ingestion process
- Systematic integration of *ESO public surveys* catalogues
  ongoing work
- 4 new large catalogues ingested in the past year

New large catalogues

- **Gaia DR1**

  1+ billion sources

  - unusual process: CDS had data in advance in order to make them ready for the release date

  - *more about Gaia DR1 in CDS services later this afternoon*

- **VPHAS+ DR2**   ES O public survey

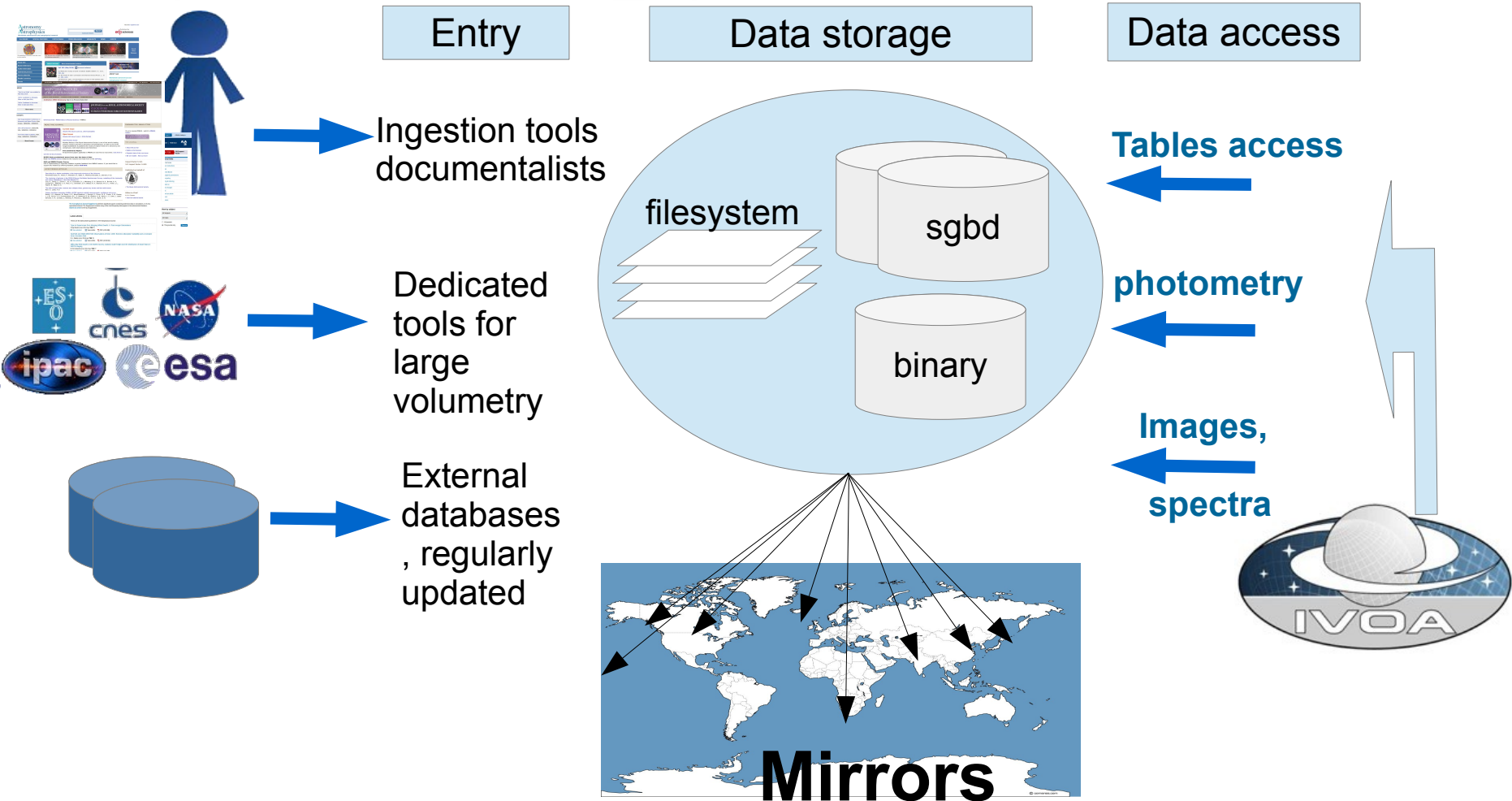 319M rows

- **APASS9**

  611M rows

- **XPM**

Soon available

- SDSS DR12

- Hubble Source Catalog v1 & v2

- KiDS DR2
- Viking DR2
- VVV DR2

public surveys

Query VizieR catalogues

The information system

Information System

| Entry | Data storage | Data access |

Ingestion tools
documentalists

Dedicated
tools for
large
volumetry

External
databases
, regularly
updated

filesystem

sgbd

binary

Tables access

photometry

Images,

spectra

Mirrors

Output statistics

The last year: 380,000 queries/day, 12% coming from mirrors
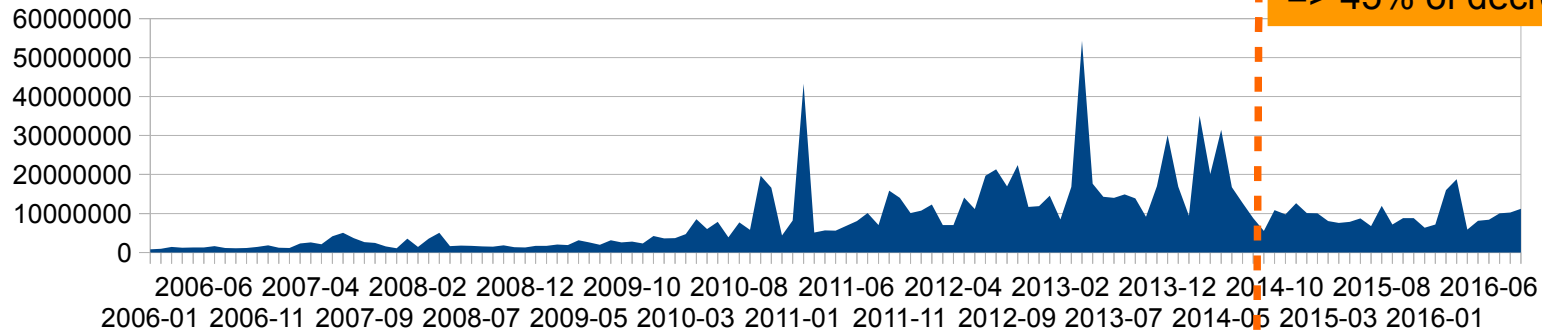
Decrease due to jobs on smart X-Match API instead of direct VizieR queries => 45% of decreased

Evolution of the VizieR queries

Evolution of the VizieR queries repartition
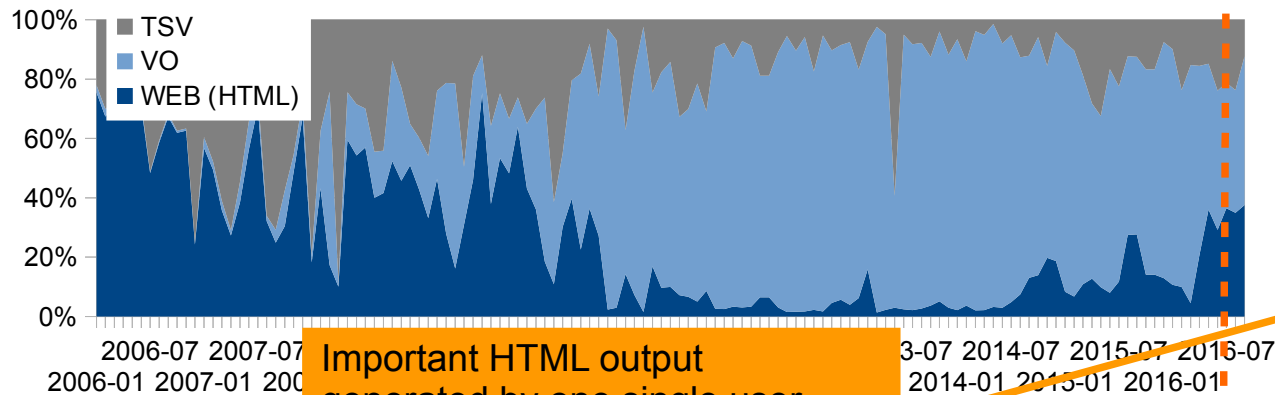
Important HTML output generated by one single user using wget and perl user-agent

Output statistics

Gaia DR1 (released in september 2016) ~3,322,000 queries

TOP 20 of the most popular catalogues

Catalogues (top to bottom): Hipparcos (1997), UKIDSS-DR9 (2012), APASS DR9 (2016), PPMXL (2010), URAT1 (2015), SDSS 9 (2012), PPMX (2008), 2MASS 6X (2006), Tycho-2 (2000), USNO-B1.0 (2003), USNO-A2.0 (1998), 2MASS (2003)

Scientific council 2016, VizieR

13
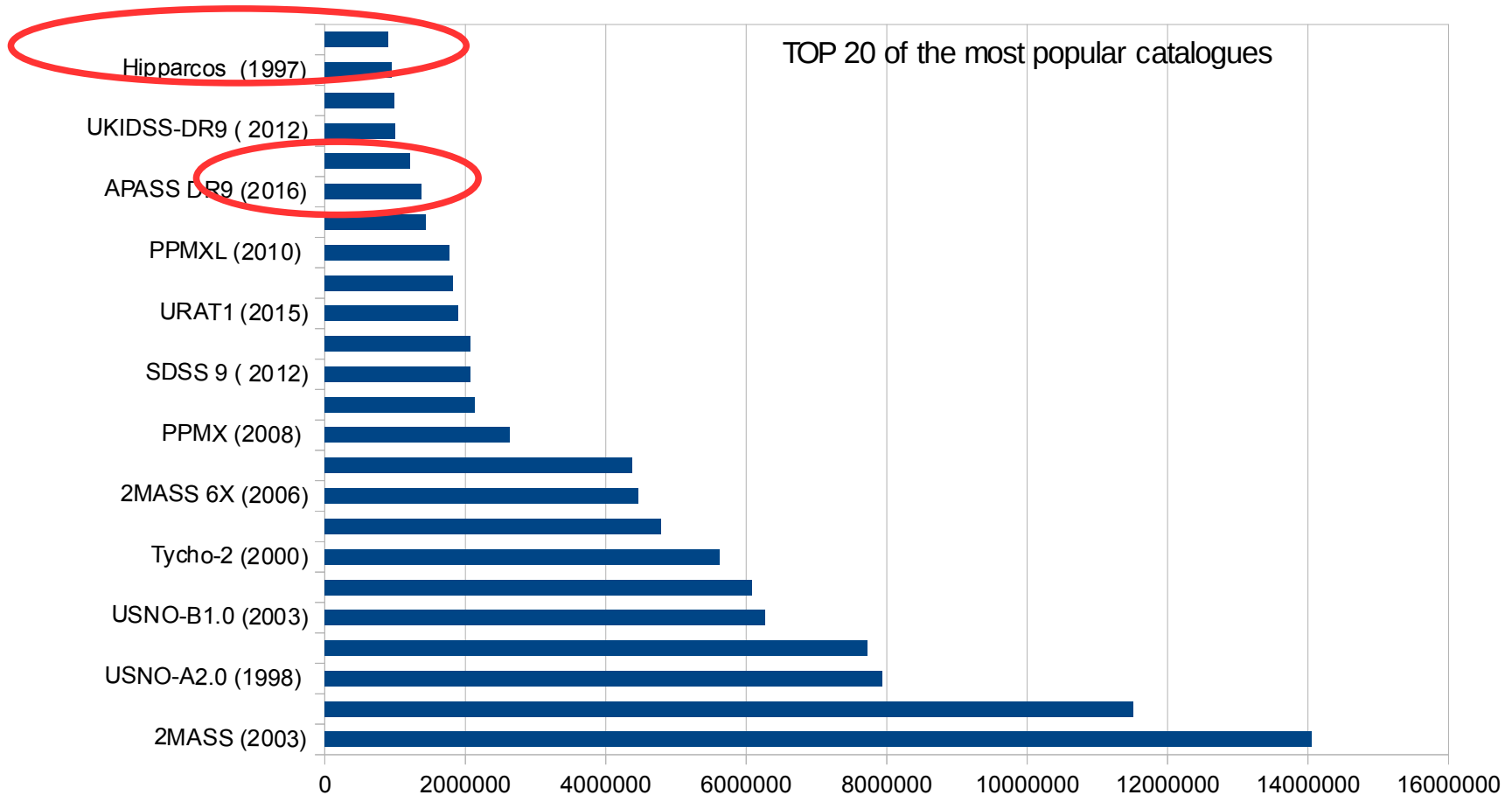
Work accomplished

- Improvement of the logs observations pipeline

  in particular a new VO-pipeline based on the standard TAP for the logs observation provided by the CADC (Canadian Astronomy Data Center) : CFHT, HST

- VizieR updates resulting of Gaia ingestion

  - Work with original column name in VizieR (including TAP queries)

  - Updates into the internal TAP engine

- Indexation of the spectra, images in VizieR

Works in progress

- (still in developpement) XML Extraction tool using the XML distributed by ApJ journals (Thomas Delacour)

- Finalise the CFHT/HST pipeline : the process could be reused for others pipelines

- DOI attribution for the VizieR catalogues

The associated data

Indexation of the associated data (images, spectra, time-series)

Improving the associated data visibility with indexation and meta-data based on the ObsCore Data model (VO).

Services ready from end to end and partially deployed

- Development began in 2014 after the AAS meeting

- build on the Saada database generator (L.Michel, Observatoire de Strasbourg)

- Discovery services are online with dedicated web application and VO access

- Ingestion pipeline is ready, but **not in action**

The increase of the ingestion pipeline workload

- Meta-data generation shared between authors and the CDS staff

- Trial period is needed for CDS & authors

- Evaluation of the additional workload at CDS

The associated data

The contents

Ingestion of 64 VizieR catalogues having FITS

- 18,400 images from 33 catalogues

- 2,470,000 spectra from 34 catalogues

- 2 surveys:

  - LAMOST : ~2,2M spectra (Chinese Optical Telescope)

  - CoRoT: ~170,000 time-series (Exoplanet detection CNES)

The softwares/services

- Graphical and console GUI sofwares for metadata mapping

- Data access services in a dedicated web-application and through the VO

Ingestions tools

- **Tools to make the mapping with ObsCore**
- **Semi-automatized detection operated with Saada**

Authors

CDS documentalists

VizieR Associated data web page

VizieR associated data through the VO

SSA, Samp

SIA, Samp, ObsTAP

ObsTAP, Samp

2017....

- Deploy the associated data pipeline

- Renewal of the acceptation of the DSA label (Data Seal of Approval)

What is VizieR ?

Data provider & astronomers papers

tables

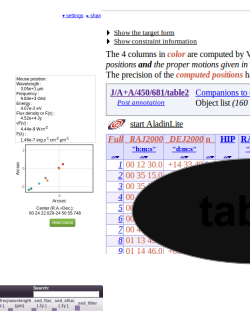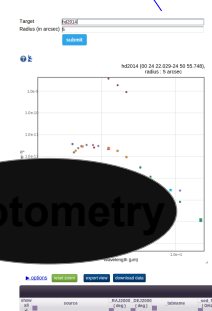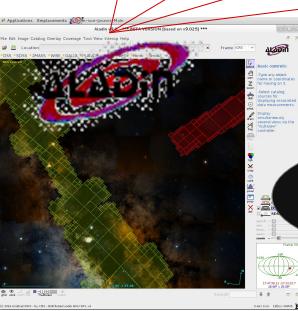*Images, spectra, light-curves*

FITS

Hips	MOC	Tables DB	ReadMe	Standardized tables		FITS	ObsCore DB

Hips server	MOC server	VizieR	FTP	Saada-VizieR
		TAP, SCS		SIA, SSA, ObsTAP
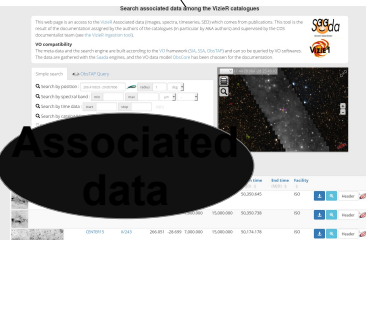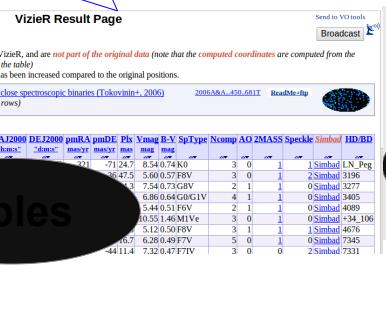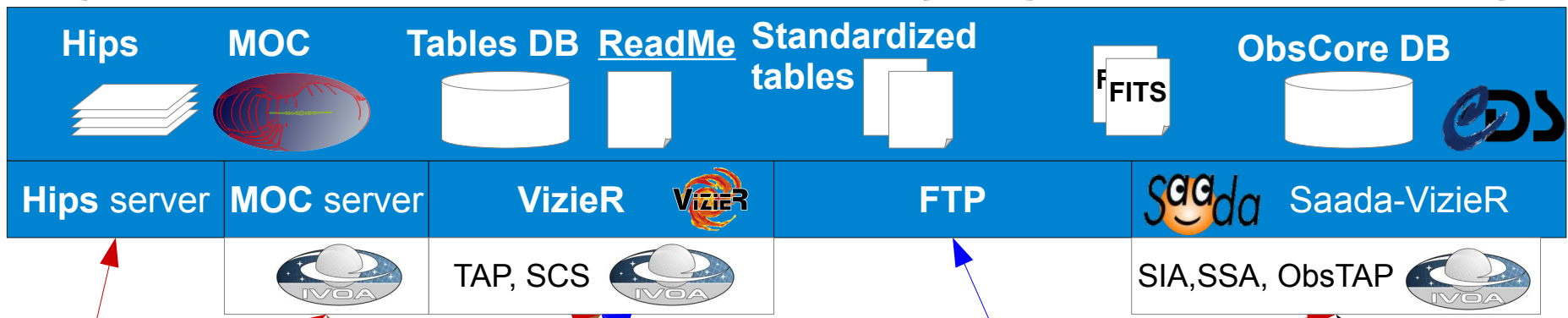
Xmatch

photometry

tables

Associated data

	09/2013	09/2014	09/2015	09/2016
Catalogs	11,601	12,802	14,096	15,366
Tables	24,912	27,450	30,058	32,991
Total of records	14.2G	16.8G	17.4G	19.2G
Total of records coming from big catalogues (binary format)	13.8G	16.1G	16.7G	18.3G

1270 new catalogues ingested from oct 2015 to sept 2016
95% of records come from large catalogues.