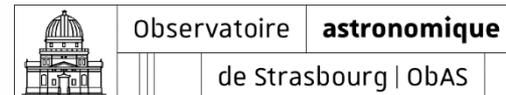


Le CDS, techniquement, comment ça marche ?

Journées ASOV 2022

Pierre Fernique
avec toute l'équipe du CDS



□ Le plan



CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG



- Une mission
- Des personnes
- Une infrastructure
- Un fonctionnement évolutif
- Au service de la communauté scientifique et au-delà

Une mission inscrite dans la durée



« Centre de Données astronomiques de Strasbourg »
créé par l'INAG (maintenant INSU) en 1972

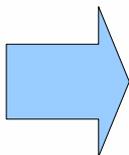


*“Collecter des données « pertinentes » sur les objets astronomiques,
sous forme électronique ; les améliorer en les évaluant de manière
critique et en les combinant ; distribuer les résultats à la communauté
internationale. ...”*

La charte du CDS

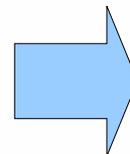
Quelles données ?

Images



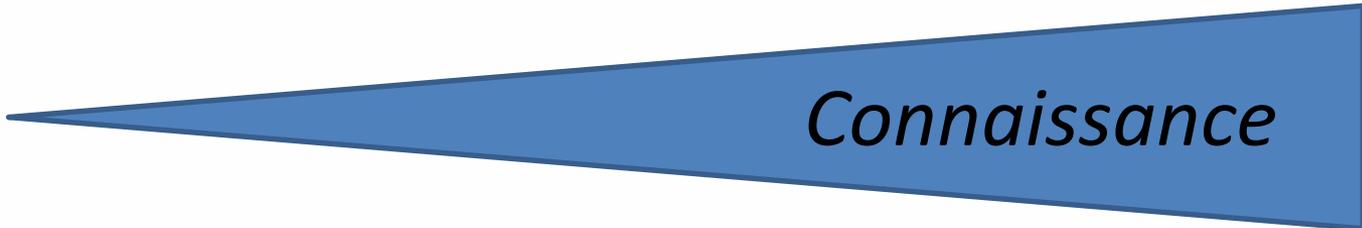
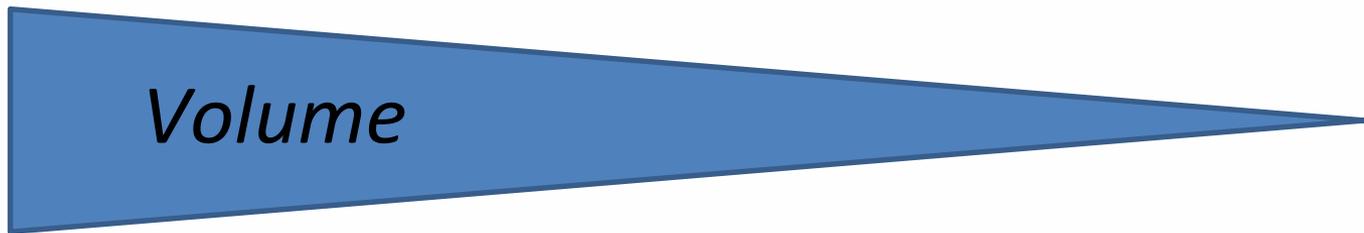
catalogues

MAIN ID	TYPE	RA	DEC	COO	COO	C	PERA	PEREC
M101	Galaxy	14 14 45.5	+54 47 00.0					
NGC 5457	Galaxy	14 14 18.55	+54 25 21.9					
PGC 7141410	Galaxy	14 14 45.72	+54 44 25.6					
M 101	SPC	14 14 35.38	+54 28 14.9	6.74	5.71	10	14.13	-1.41
NGC 5458	Galaxy	14 13 18.00	+54 23 36.0	18000	18000	171		
M 101	Star	14 18 43.97	+54 24 50.3	23.26	20.24	0	-54.9	37.9
NGC 5459	Galaxy	14 13 36.33	+54 20 14.0	3000	3000	90		
NGC 5460	Galaxy	14 13 21.89	+54 40 06.0	3000	3000	40		
M 101	Star	14 14 28.69	+54 16 54.0	3000	3000	171		
M 101	Star	14 14 28.00	+54 36 36.0	37000	19500	9		
M 101	Star	14 12 24.90	+54 42 40.0	30000	30000	157		
M 101	Star	14 12 22.00	+54 42 42.0	65000	57000	0		
M 101	Star	14 15 13.40	+54 40 22.6	3300	3300	90		
M 101	Star	14 15 52.90	+54 55 16.0	20000	20000	120		
M 101	Star	14 16 38.39	+54 07 19.9	300	300	10		
M 101	Star	14 15 37.69	+54 48 05.4	300	300	20		
M 101	Star	14 16 02.75	+54 53 16.3	300	300	100		
M 101	Star	14 13 35.95	+54 26 40.6	6900	2900	90		
M 101	Star	14 11 30.69	+54 32 39.6	43.56	39.54	0	-9.3	-14.9
M 101	Star	14 13 36.09	+54 25 51.6	41.62	34.91	0	-6.2	1.8
M 101	Star	14 14 27.64	+54 17 06.5	36.78	31.93	90	-7.6	-6.8
M 101	Star	14 15 37.96	+54 54 17.6	39.6	38.93	90	-12	-7.1
M 101	Star	14 16 36.70	+54 48 40.4	84.74	41.74	0	21.5	8.6
M 101	Star	14 14 55.99	+54 17 36.6	94.8	39.62	90	-9.1	-7.5
M 101	Star	14 15 02.97	+54 13 24.8	38.55	27.93	90	39.5	-33.2
M 101	Star	14 15 46.48	+54 40 07.7	34.77	31.4	90	-20.7	-13.2
M 101	Star	14 15 39.82	+54 29 45.2	44.1	37.69	0	15.4	-43.0
M 101	Star	14 14 40.00	+54 22 00.0					
M 101	Star	14 13 37.90	+54 24 14.0					
M 101	Star	14 14 15.70	+54 25 23.0					
M 101	Star	14 14 45.70	+54 46 15.5					
M 101	Star	14 14 56.60	+54 56 54.5					
M 101	Star	14 12 21.00	+54 32 26.0					
M 101	Star	14 14 48.00	+54 53 46.0	300	300	107		
M 101	Star	14 12 37.00	+54 29 50.0	300	300	94		
M 101	Star	14 13 40.40	+54 26 37.0	300	300	9		
M 101	Star	14 14 51.60	+54 28 17.0	300	300	18		

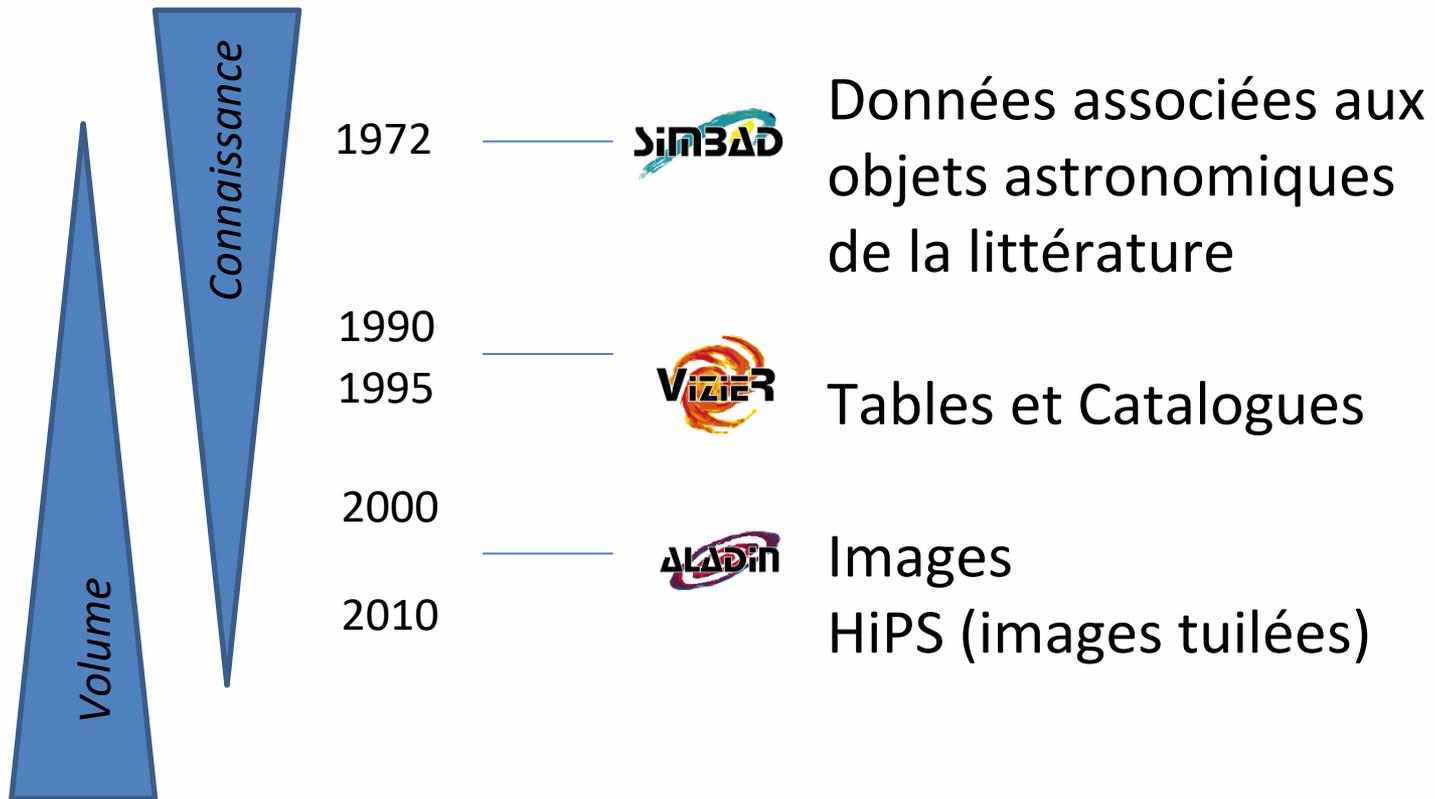


objets

Observation data (J2000 epoch)	
Constellation	Ursa Major
Right ascension	14 ^h 03 ^m 12 ^s
Declination	+54° 20' 57"
Redshift	0.000804
Helio radial velocity	241 ± 2 km/s
Distance	20.9 ± 1.8 Mly (6.4 ± 0.5 Mpc)
Apparent magnitude (V)	7.86
Characteristics	
Type	SAB(rs)cd
Number of stars	1 trillion (10 ¹²)
Size	~170,000 ly in diameter ^[1]
Apparent size (V)	28' 8 × 26' 9
Other designations	
Messier 101, M101, NGC 5457, UGC 8981, PGC 50063, Arp 26	
References: [2][3][4][5][6][7]	



Une réponse qui évolue en fonction des moyens techniques



□ La méthode ?

Une même méthode depuis **50 ans** :

1. Sélectionner des données (issues des journaux et des relevés)
2. Les homogénéiser dans **des modèles de données propres au CDS** (objet astronomique, catalogue/table, HiPS)
3. Les décrire scientifiquement, de manière cohérente
=> les vérifier, les valoriser
4. Les redistribuer



□ CDS figures: 3-500-16-2^E6

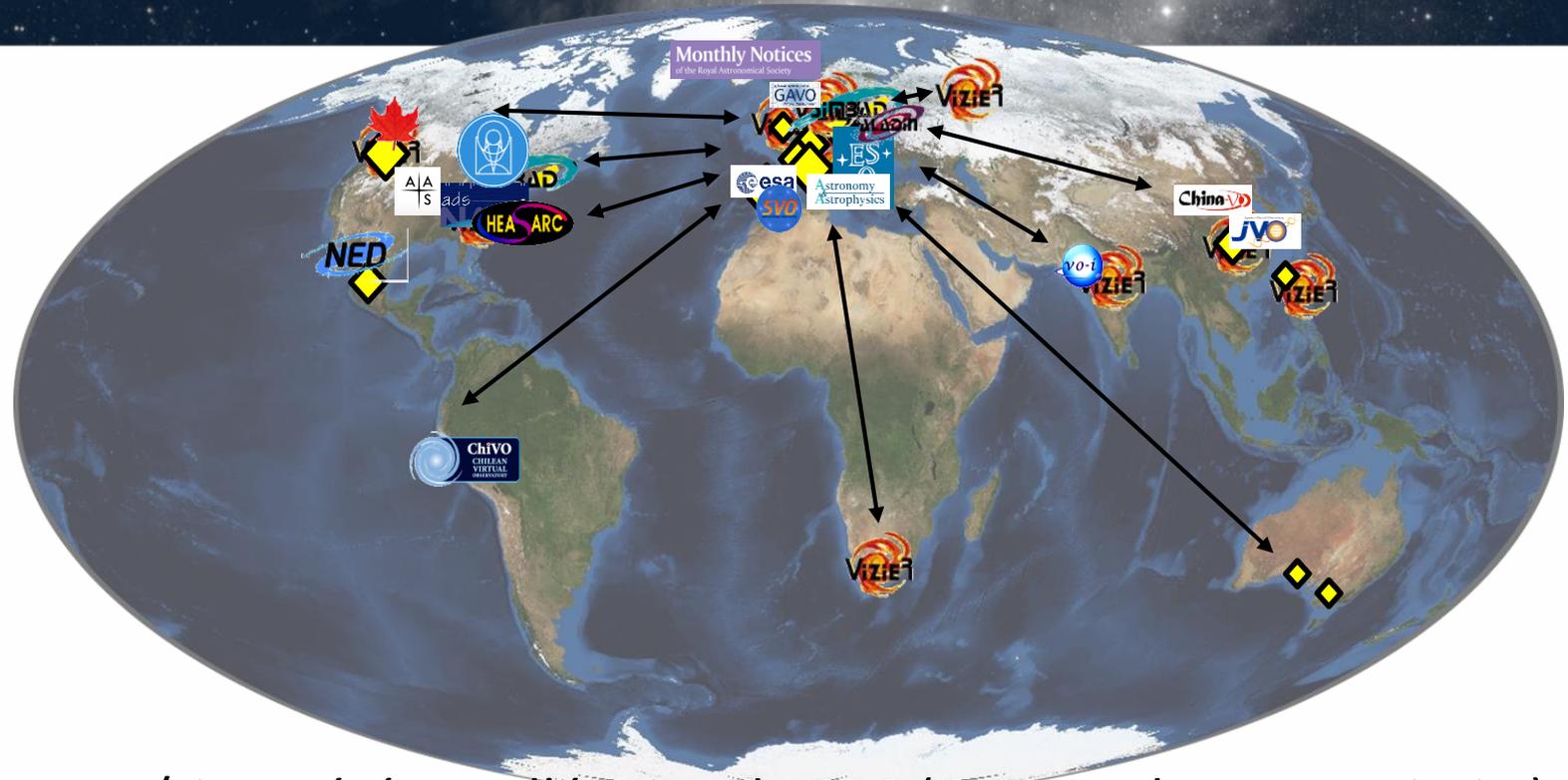


- **3** services principaux: Simbad, VizieR & Aladin
- **500** To de données
- **16** sites/serveurs opérationnels sur 5 continents (7 serveurs locaux + 9 miroirs)
- **2 millions** de requêtes par jour

□ Utiliser au mieux les expertises de chacun

- **Astronomes**
 - Missions : recherche, enseignement, tâches de services
 - Diversité d'expertise scientifique
 - Contributions au sein du CDS et de la communauté
- **Informaticiens**
 - Découpage par services indépendants
 - Multi-casquettes : développeur, architecte software, bases de données, ingénieur système, support
- **Documentalistes**
 - Polyvalence : astronomie, informatique, information scientifique
 - Chaîne documentaire complexe -> données à forte valeur ajoutée

□ L'implantation du CDS

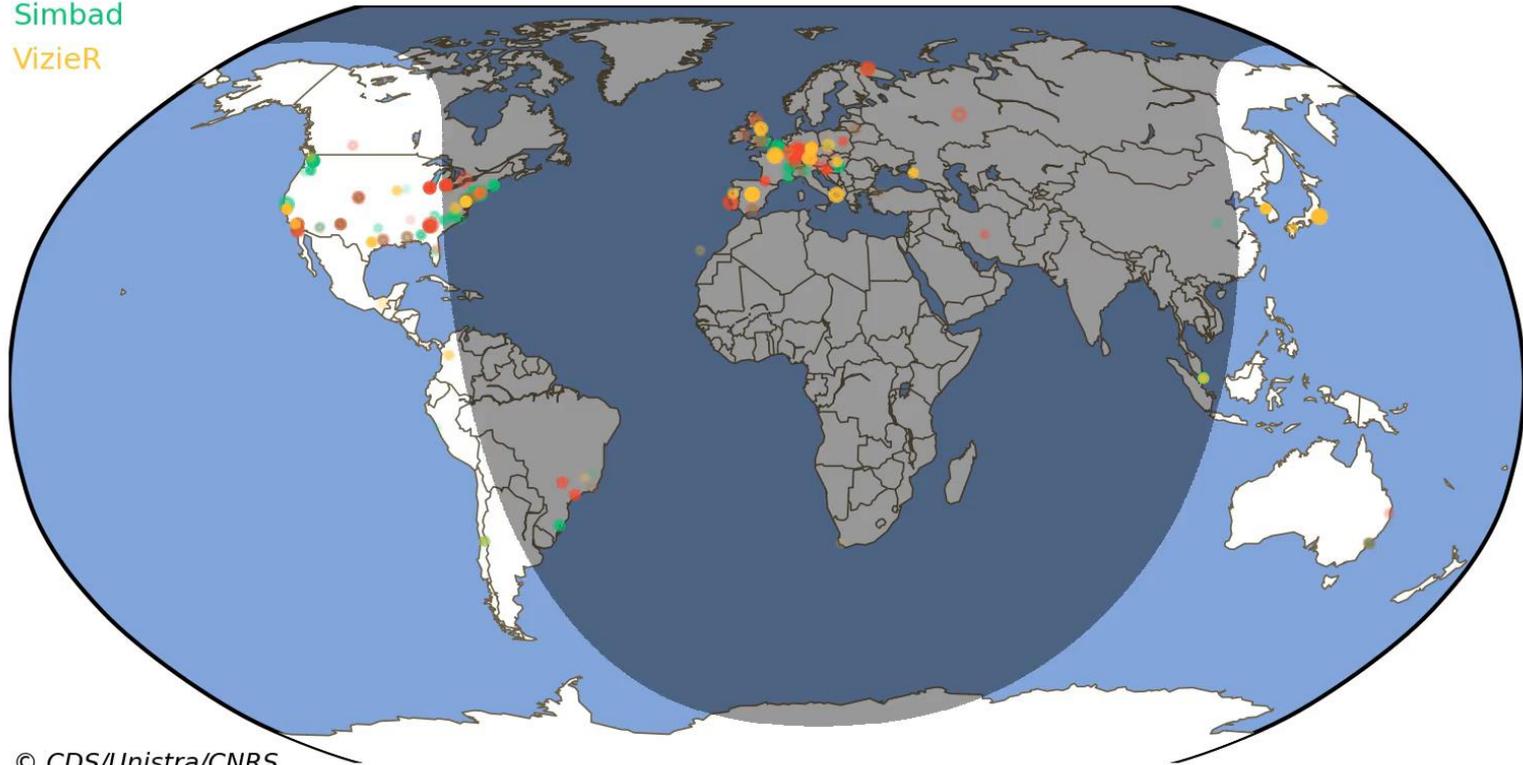


16 serveurs/sites gérés par l'équipe du CDS (serveurs locaux + miroirs)
+ **Dizaine** d'instituts partenaires/de centre de données/d'éditeurs...
qui abondent leurs services avec les données du CDS et **récioproquement**

Au service de la communauté scientifique et au-delà

Aladin
Simbad
VizieR

2022-02-10 00:00:00 UTC+01



□ Les infrastructures

- But : **haut niveau de disponibilité (~99.5%)**
Mais sans astreinte des personnels
(Note: 1600h annuelles de travail => 17% du temps d'une personne)
- 2 lignes de conduite:
 - **Duplication** totale (données, services, infra) dans au moins 2 lieux géographiques distincts
 - **Partage des compétences** systèmes et logiciels (a minima)



□ Continuité de service

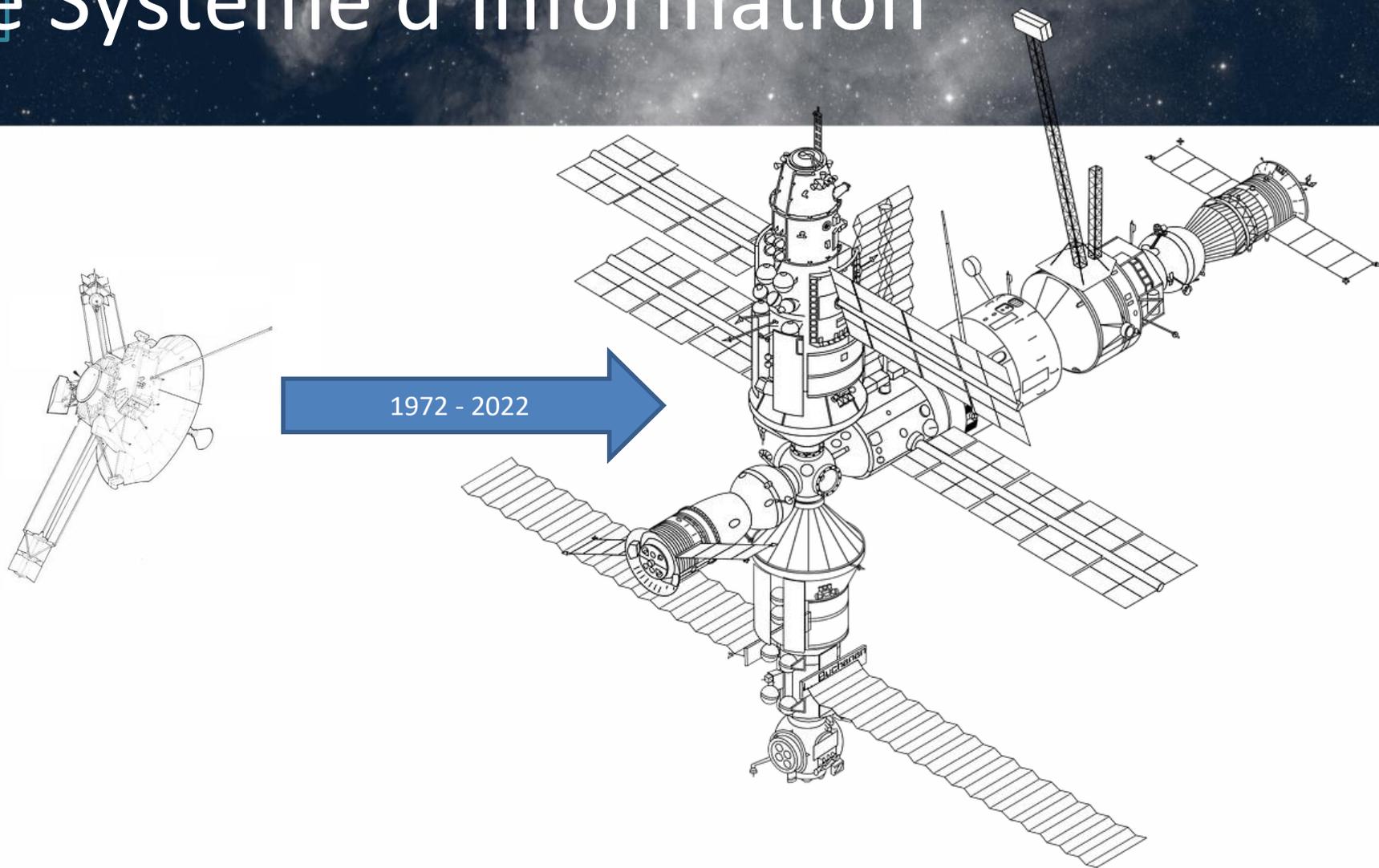
Plan de reprise après sinistre (PRA)

- **Deux sites locaux** : Observatoire + Unistra DataCenter
 - ⇒ Hyperviseurs VMware réparties (2 + 1)
 - ⇒ Serveurs dédiés dupliqués
 - ⇒ Système de stockage redondant (1.6Pox2)
- **Des miroirs externes** chez nos partenaires : NASA, ESA, ESO, ...
=> Vérification régulièrement qu'au moins un miroir distant peut être utilisé pour reconstruire le service (en cas de catastrophe majeure à l'Obs)

- **Des fiches « survie »**: 37 cartes « panne => solution » que les collègues habilités peuvent mettre en œuvre en cas d'absence de la personne responsable
=> méthode simple et efficace pour les défaillances mineures

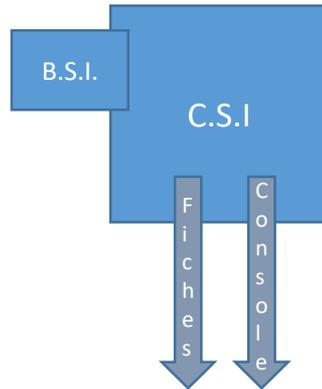



Le Système d'Information



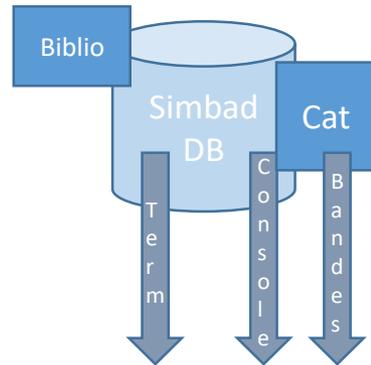
1972 – 1980 – CDS à Strasbourg
2 à 6 personnes

1 IBM 360/65 à
Meudon



1981-1985 – (83 -> + galaxies)
6-12 personnes

1 Univac 1110 à Cronenbourg
réseau dédié– lang PL1



1986-1994 – (91: +biblio, +catalogues)

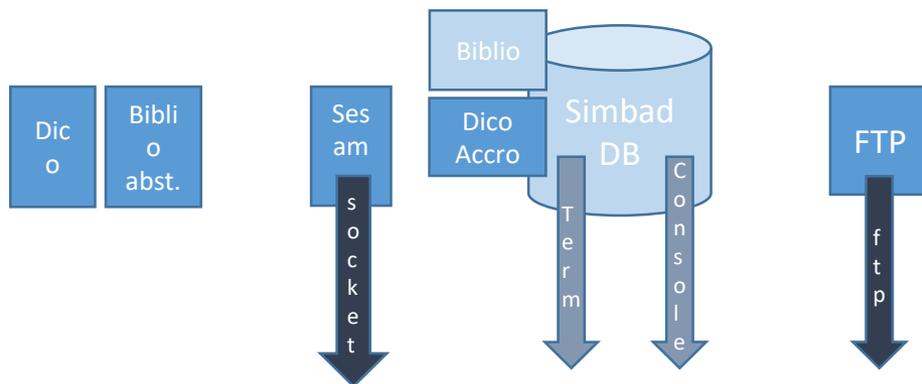
12-20 personnes

1 Vax/VMS, suivi by Dec/Ultrix à Strasbourg

lang C - DECnet,

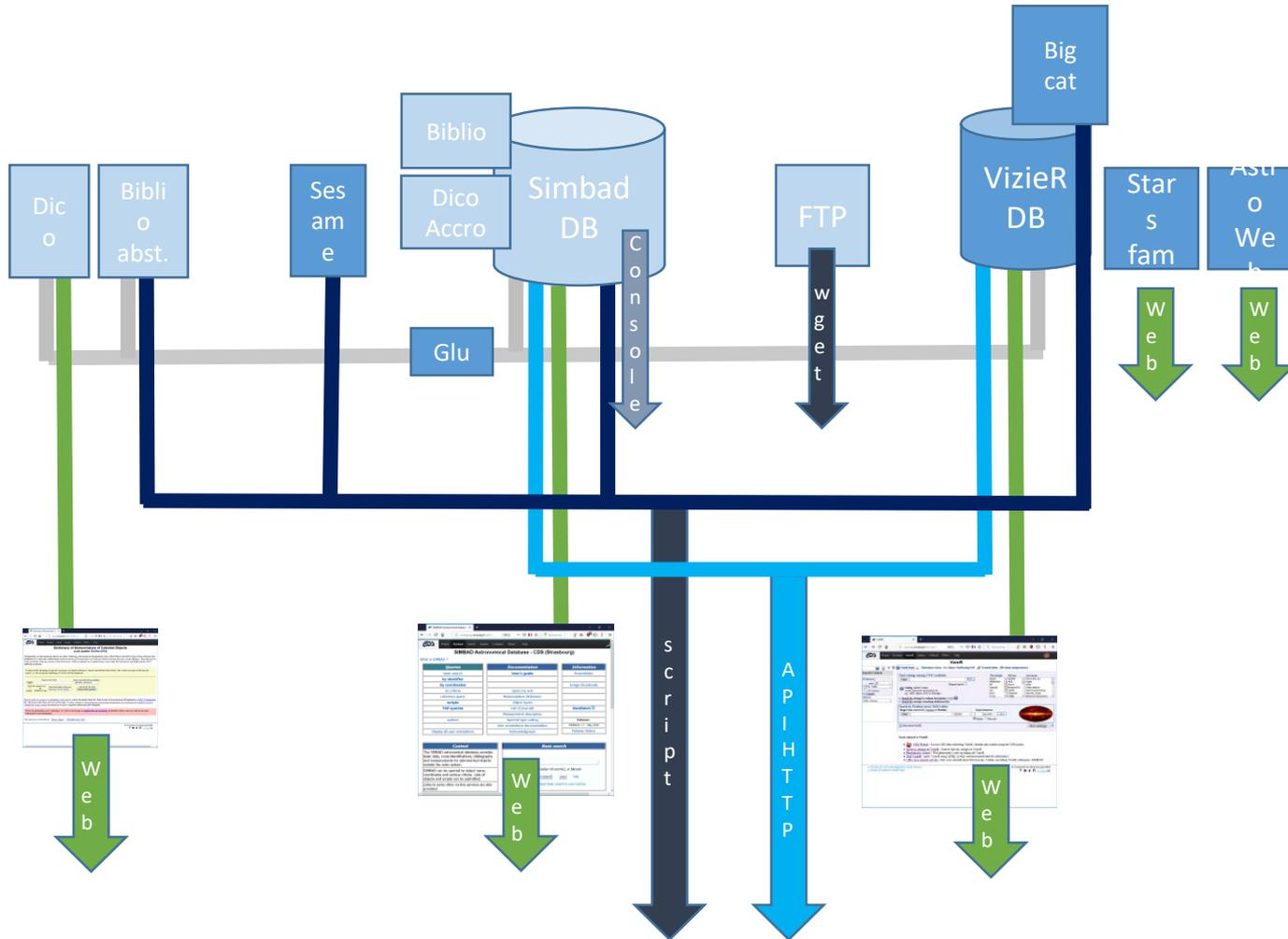


suivi par 3 Suns
lang C, AWK – Internet



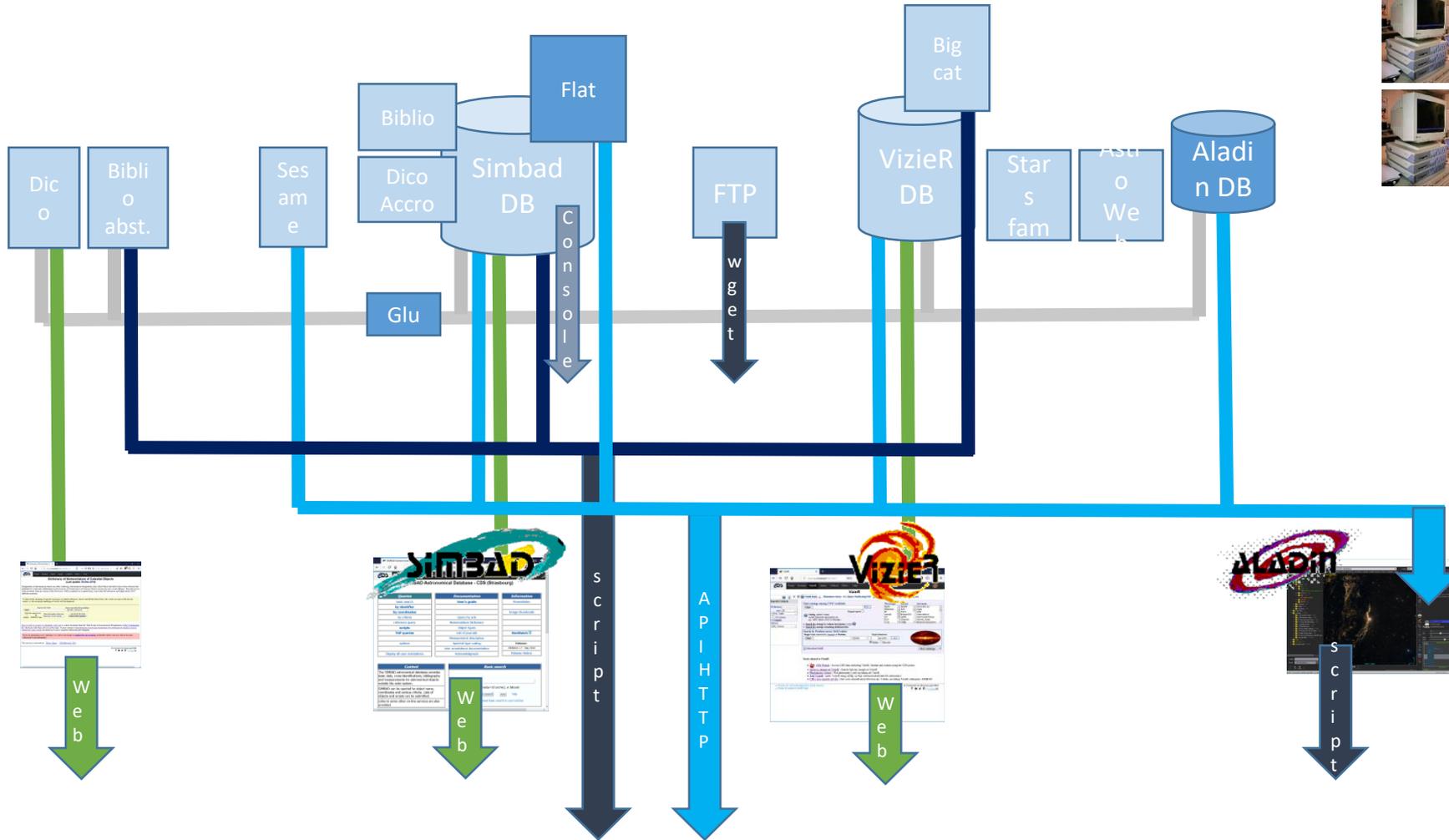
1995-1997 (94: +images, +big catalogs, +indexAstro)
20-25 personnes

5 Suns Obs
Lang C,C++,AWK,perl



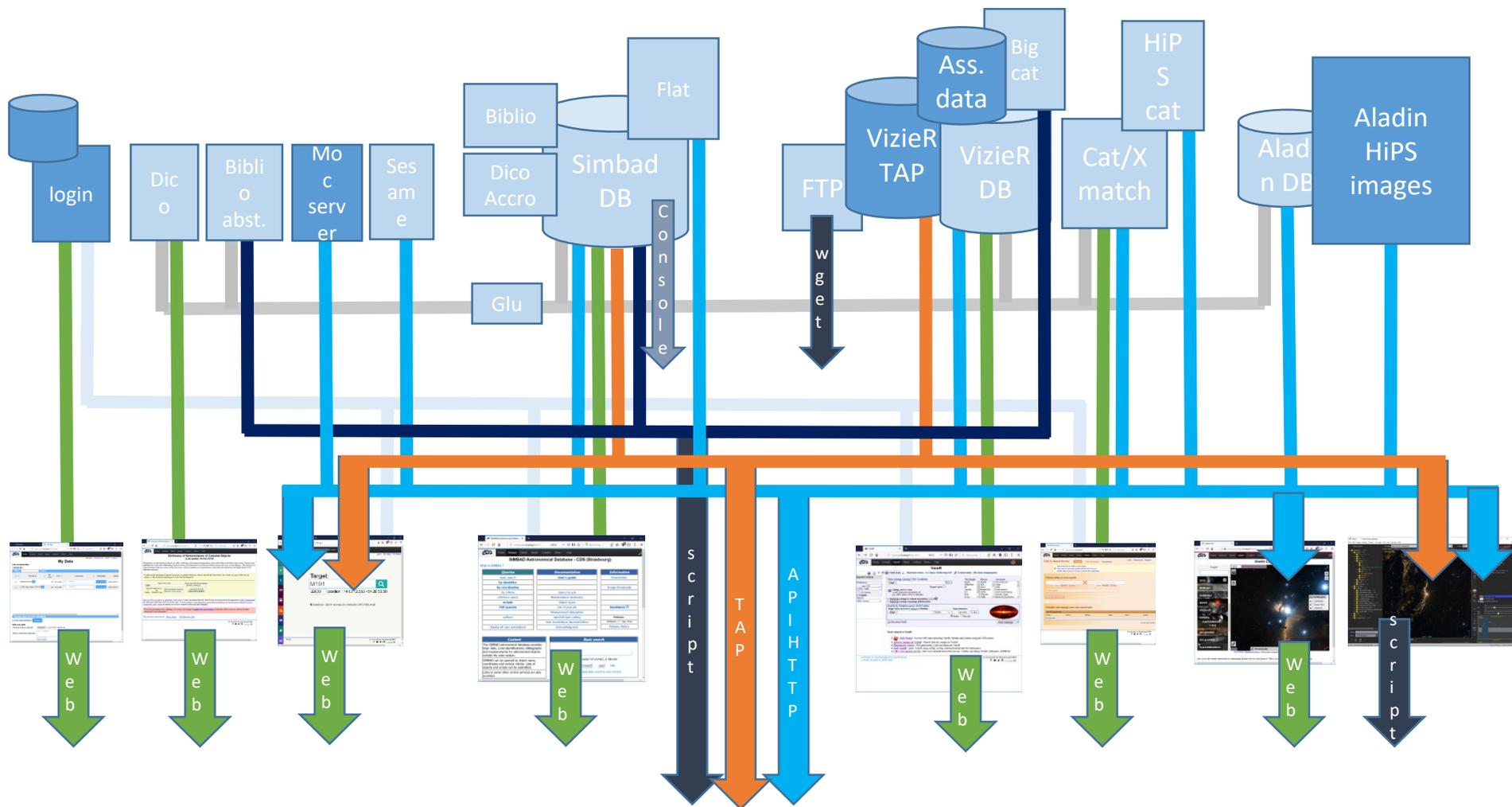
1998-2007 (03: -indexAstro)
25-30 personnes

8 serveurs Obs
+ miroirs externes
Lang C,java,AWK

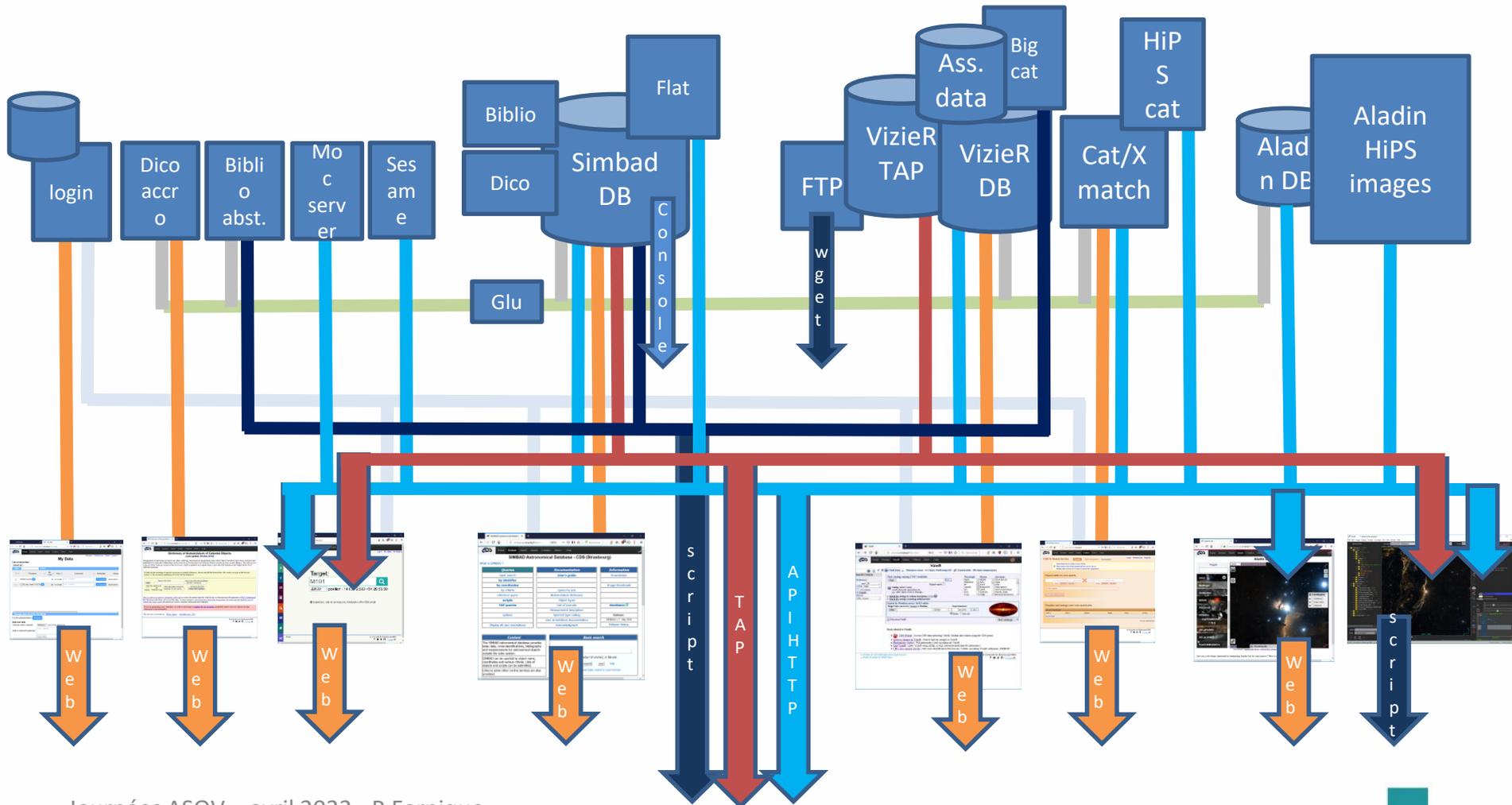


2009-2019 (12:-biblio, 14: +Big Images Surveys, 16: +AssocData)
35-40 personnes

12 serveurs Obs + miroirs
lang: +python



Le Système d'Information du CDS: en fait une vingtaine de composantes



□ Gérer les évolutions



- **Inflexions lentes** => Beaucoup « d'énergie cinétique » 
- Dictées par le choix des **innovations** retenues, généralement précédées par un prototypage sur une plate-forme de test 
- Pilotées par une connaissance précise du SI basée sur :
 - **L'expérience** des personnels
 - **Le suivi** de l'utilisation et du contenu 
 - **La supervision** opérationnelle
- En gardant à l'esprit:
 - **Rationalisation** des composants du système
 - **Suppression** ou **évolution** des composants obsolètes 
 - **Redondance complète** dans des lieux géographiques distincts (dans la mesure du possible)

□ Innovations

- Adaptation en fonction de l'évolution du contexte:

- **Science:**

- Collaborations scientifiques internationales
- Interactions inter-équipe

- **Technique:** Web, Qbox -> HEALPix, virtualisation

- **Traitement des données :**

- quantité augmente, complexité
- Best Practices

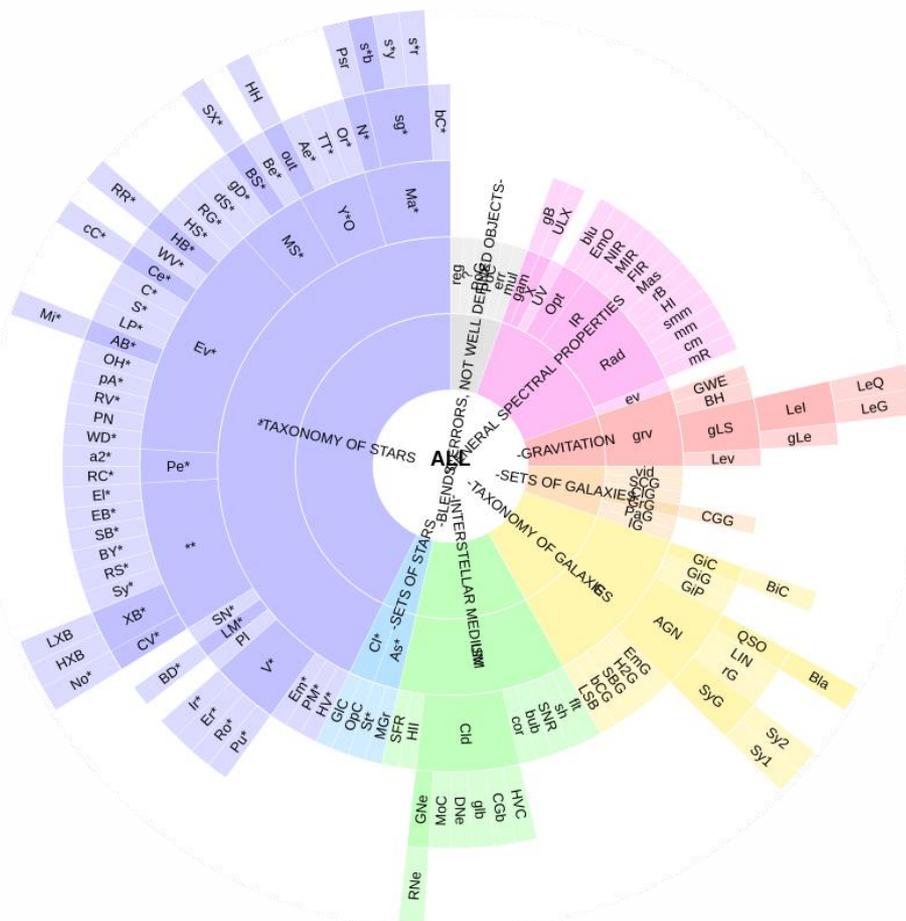
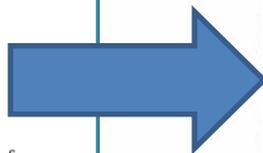


Exemple: Types d'objets Simbad



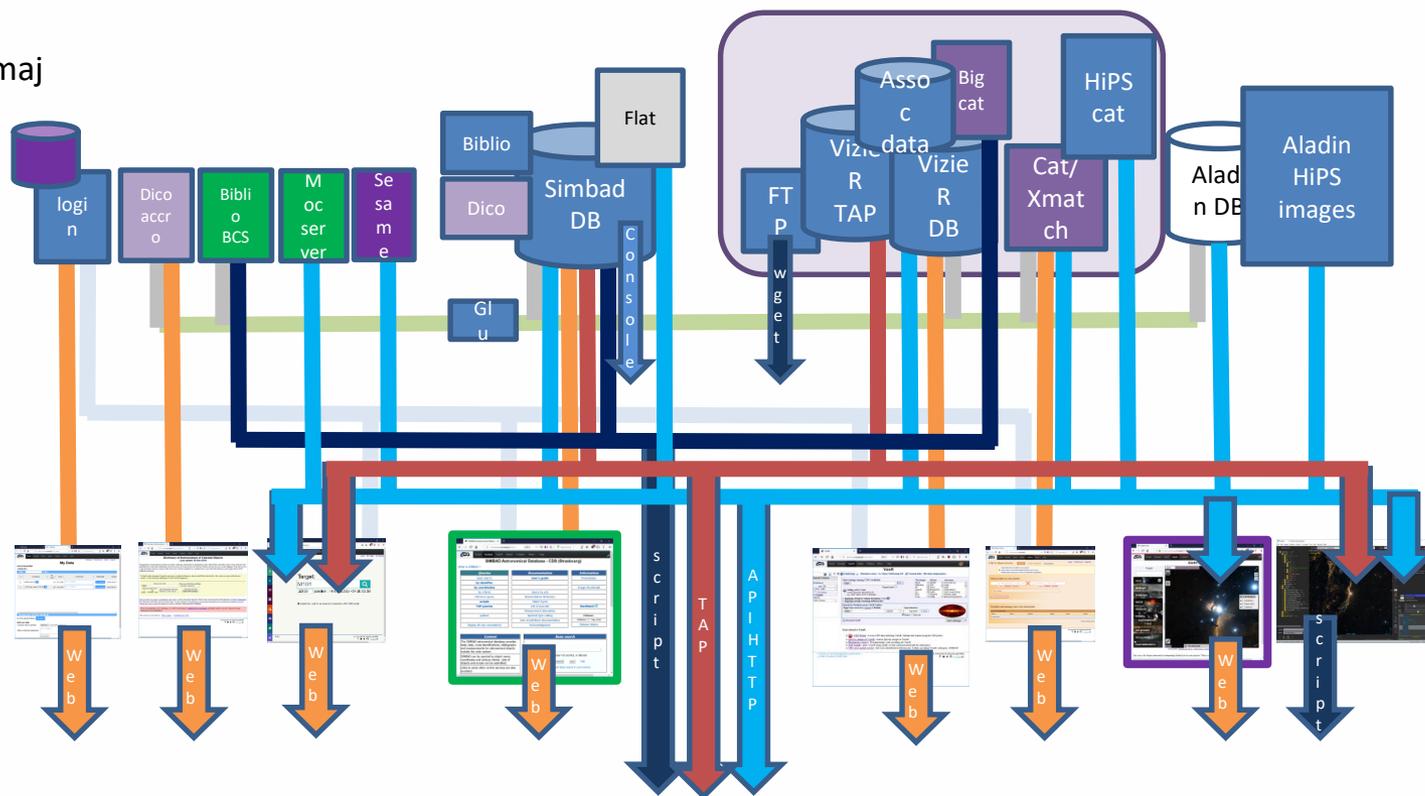
Table 1. Object type classification: extracts from the object type table illustrating examples of the four levels of the classification scheme

Level	Standard name	Short name	Extended Explanation
...			
1	Star	*	Star
2	*inCl	*iC	Star in Cluster
2	*inNeb	*iN	Star in Nebula
2	*inAssoc	*iA	Star in Association
2	*in**	*i*	Star in double system
2	V*?	V*?	Star suspected of Variability
2	Pec*	Pe*	Peculiar Star
3	HB*	HB*	Horizontal Branch Star
3	YSO	Y*O	Young Stellar Object
3	Em*	Em*	Emission-line Star
4	Be*	Be*	Be Star
...			
1	Galaxy	G	Galaxy
2	PartofG	PoG	Part of a Galaxy
2	GinCl	GiC	Galaxy in Cluster of Galaxies
2	GinGroup	GiG	Galaxy in Group of Galaxies
2	GinPair	GiP	Galaxy in Pair of Galaxies
2	High_z_G	H_zG	Galaxy with high redshift
...			
2	AGN	AGN	Active Galaxy Nucleus
3	LINER	LIN	LINER-type Active Galaxy Nucleus
3	Seyfert	SyG	Seyfert Galaxy
4	Seyfert_1	Sy1	Seyfert 1 Galaxy
4	Seyfert_2	Sy2	Seyfert 2 Galaxy
3	Blazar	Bla	Blazar
4	BLLac	BLL	BL Lac - type object
4	OVV	OVV	Optically Violently Variable
3	QSO	QSO	Quasar



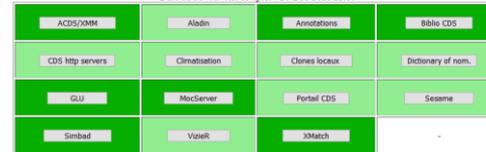
Évolution des composantes du SI

- Renouvelé ou maj
- Supprimé
- En cours
- Sans modif.



□ Supervision opérationnelle

- En synergie étroite/efficace avec **l'équipe infra de Observatoire** de Strasb. (2 ingénieurs)
- **Supervision** :
 - Globale => GluSupervisor
 - Détaillée => spécifique à chaque service (nagios...), technique (hardware) => zabbix
- **Tableau de bord**:
 - => Outil maison pour remonter l'ensemble des statistiques de tous les services (contenu, activité, audience, profils et disponibilité)

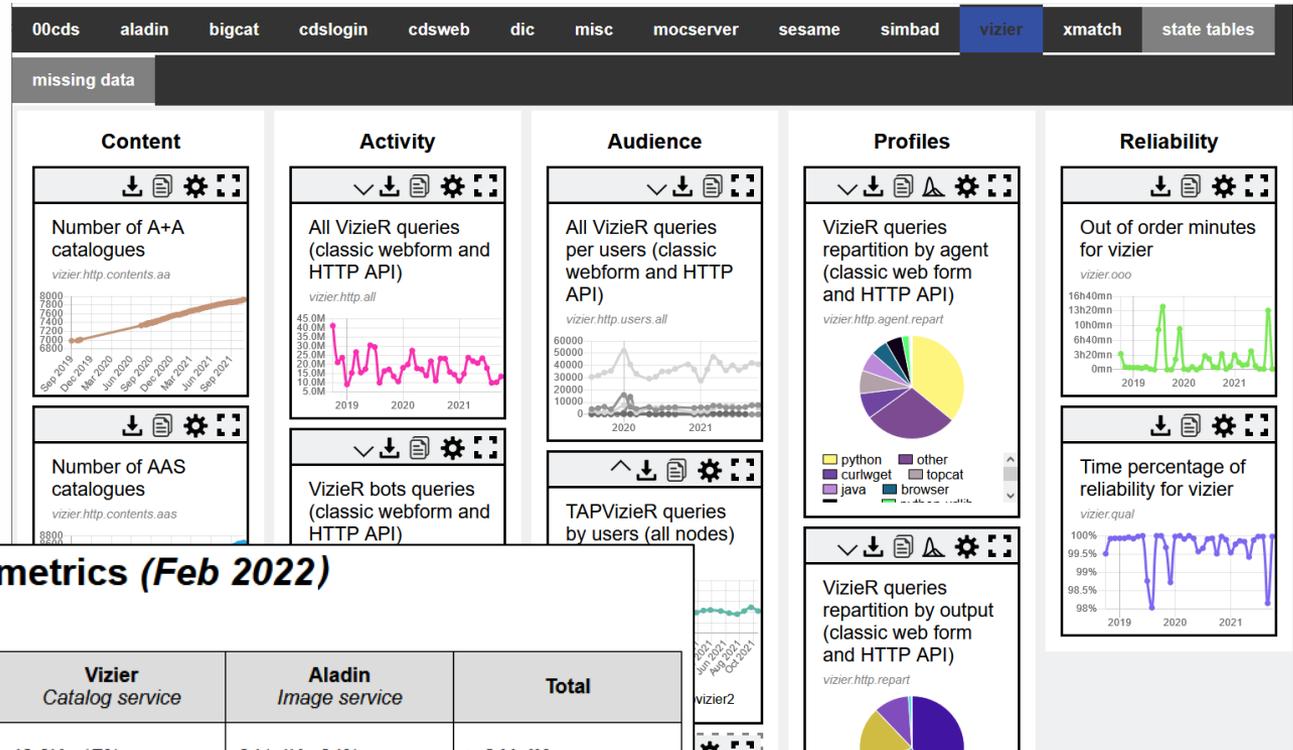


ACDS/OMM	Abelin	Annotations	Bible CDS
CDS http servers	Climatisation	Clones locaux	Dictionary of nom...
GLU	MacServer	Partiel CDS	Sesame
Simbad	VizR	XMatch	-



Tableau de bord CDS

Les services
du CDS en
un coup
d'œil



CDS main metrics (Feb 2022)
compared to Feb 2021

Main services	Simbad <i>Astro.object DB</i>	Vizier <i>Catalog service</i>	Aladin <i>Image service</i>	Total
users/month	154.4K +4% Adv.users: 1454	43.0K +17%	341.4K +34%	> 341.4K
queries/day	552.0K +65%	512.6K +7%	1.3M +22%	2.3M
load/day	11.6GB +182%	no data	122.6GB +5%	> 134.2GB
data volume	29.8GB +15%	52.8TB	406.8TB +2%	459.6TB
data content	13.1M obj +8%	21.9K cats +5%	986 HIPS +0.9%	
reliability	98.31% -0.6%	100.00% +0.2%	99.94% -0.04%	> 98.31%

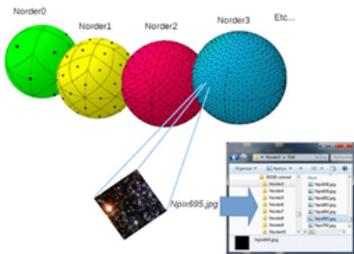
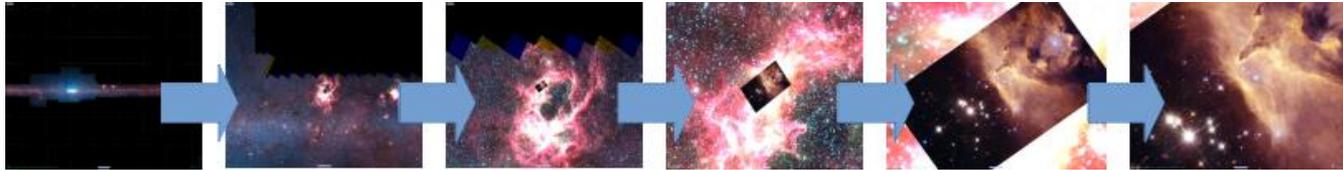
□ Et le VO dans tout ça ?



□ Les moyens d'une synergie globale

- Des partenariats bi ou multi-parties depuis des décennies: NED (bibcode), ADC (catalog), ADS (Simbad), ...
- L'IVOA depuis 2001:
 - => Une **synergie élargie** d'une richesse inestimable
 - => Une bibliothèque de **formats et protocoles d'échanges** : CS, VOTable, UCD, SAMP, HiPS, MOC, TAP...
 - Font désormais partie intégrante de nos outils et méthodes de travail.
 - Deux cas de figure:
 - => **Adoption des standards** existants (ex: CS, TAP, ...)
 - => ou **proposition d'un nouveau standard** pour une solution CDS existante que l'IVOA souhaite endosser (ex: UCD, HiPS, ...)

Exemple de proposition du nouveau standard HiPS



Start of the story at CDS
by A.Oberto, P.Ferrière, T.Boch & Al.

Proto in CDS Aladin Desktop
ADASS talk

Mizar/CNES client

CDS Aladin lite

IVOA HiPS 1.0

IVOA standardization
HiPS network
ESAsky, JUDO2, ...
A&A article

ADASS talk
HiPS cube, cat, ...

- 1034 HiPS
- 2483 instances (masters + mirrors)
- 370TB of HiPS
- around 800K tile queries per day

HiPS nodes

- 20 HiPS nodes (stable compared to last year)
CFA, WFAU, CASDA, PADC, IPAC, ANU, LEIDEN, IRAP, SSC, CDS, AMIGA, SVO, IAS, ESAC, JAXA, CADC, HEASARC, China-VO...

+ dozens of undeclared HiPS nodes (private, project oriented...)

www.sternwarte.uni-erlangen.de
www.aobobservatory.org
www.eao.hawaii.edu
www.atnf.csiro.au
mail.astro.puc.cl
lofar.stw.leidenuniv.nl
gpc.idc.cab.inta-csic.es
nova.mpe.mpg.de
etelchally.net
datalab.noao.edu
darts.iias.java.jp
archive-new.noao.edu
ada.physics.usyd.edu.au:8021
192.168.56.180
192.168.56.100-80
192.168.2.45

IVOA Garching - Oct 2019 - P. Ferrière

Dozens of HiPS clients covering various niches

- Available
 - Aladin Desktop (CDS/Java)
 - MIZAR (JS/WebGL)
 - Aladin Lite (CDS/JS)
 - ESAsky, ESOportal, JUDO2, ++ (Aladin Lite based tools)
 - HscMap (Subaru project/WebGL)
 - Stellarium Desktop (C)
 - Stellarium Web (C -> Web assembly)
 - Stellarium mobile (?)
 - Kstar (C++)
 - Firefly (IPAC/JS+java backend)
- Proto
 - WWT China-VO (China-VO/C#)
- In preparation
 - Aladin Lite WebGL (CDS/JS/WebGL/RUST)
 - Digistar (?)

Scientific purpose
Planetary control
Amateurs
Public

IVOA Garching - Oct 2019 - P. Ferrière

□ Questions ? Remarques ?
Commentaires ?

