

Les données Eros II



Jna

Jean-Noël Albert – LAL / IJClab

Juin 2020

Eros II

90 millions d'étoiles

2 millions d'images

7 années de campagne

2000 nuits d'observation



La Voie lactée à La Silla (Chili).

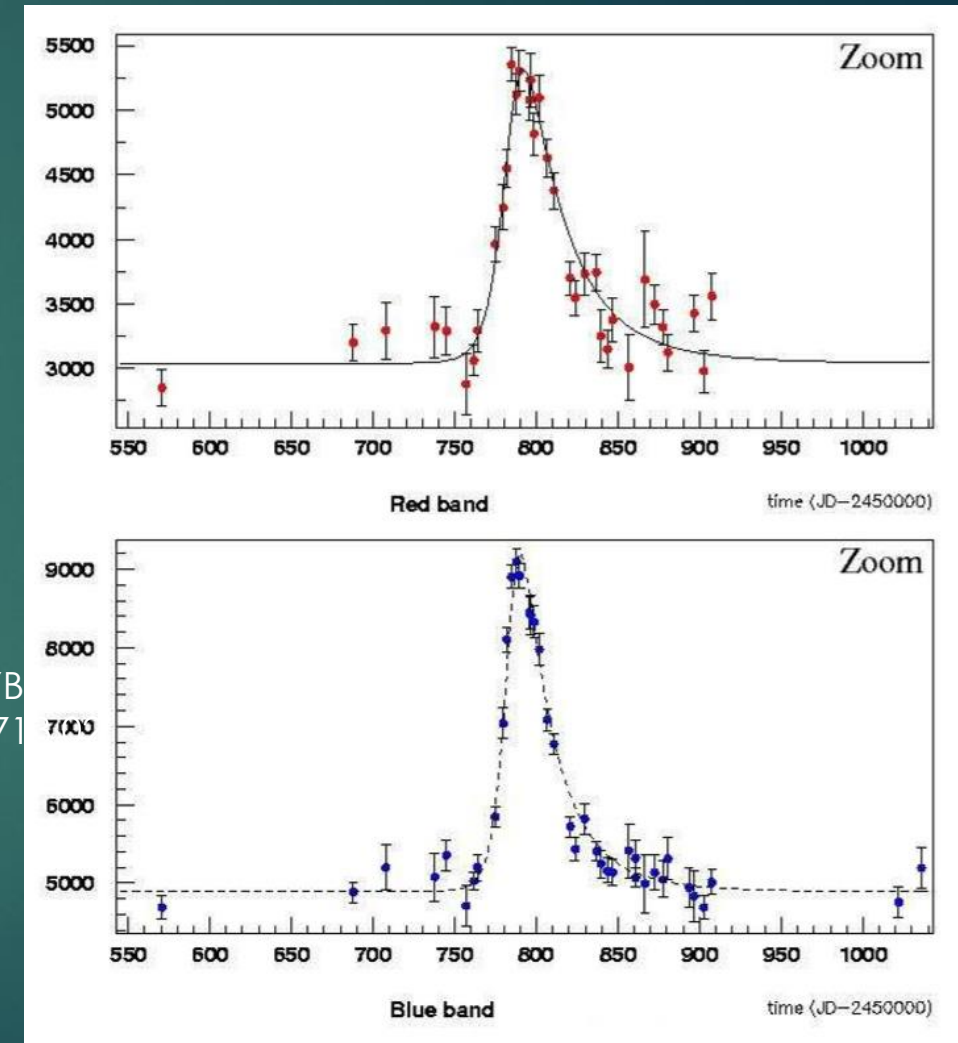
En haut à droite, le Grand Nuage de Magellan,
en bas à droite, le Petit Nuage (ESO/Z.Bardon)

Microlentilles gravitationnelles

- ▶ Recherche de microlentilles gravitationnelles
 - ▶ Objets obscures massifs dans le halo galactique
 - ▶ Variations du flux de la lumière d'étoiles lointaines
 - ▶ Observation des mêmes zones sur une longue durée
- ▶ Exemple d'une courbe de lumière
- ▶ Exemple du fichier des mesures

```
$ head lm0153n22431.time
```

```
# star: erosid      MagR  ErrMR  XR      YR      MagB  ErrMB  XB      YB
#   lm0153n22431 19.133  0.124  478.78  684.37  19.238  0.070  431.84  71
#
#  date      MagR  ErMagR  MagB  ErMagB
348.81416  19.224  0.037  19.330  0.025
366.79045  99.999  9.999  19.337  0.035
373.68669  19.115  0.051  19.299  0.032
376.84666  19.195  0.058  19.287  0.036
380.86332  19.210  0.052  99.999  9.999
381.70421  19.186  0.081  99.999  9.999
```



Champs et Programmes

- ▶ Division du ciel en grandes zones : les « programmes scientifiques »
 - ▶ Grand et Petit Nuages de Magellan
 - ▶ Centre galactique
 - ▶ Bras spiraux de la galaxie
 - ▶ ...

- ▶ Chaque zone est divisée en plusieurs « champs » représentant la zone vue par le télescope lors d'un pointage
 - ▶ Grand Nuage : 88 champs
 - ▶ Petit Nuage : 13 champs
 - ▶ Centre galactique : 141 champs
 - ▶ Bras spiraux : 36 champs

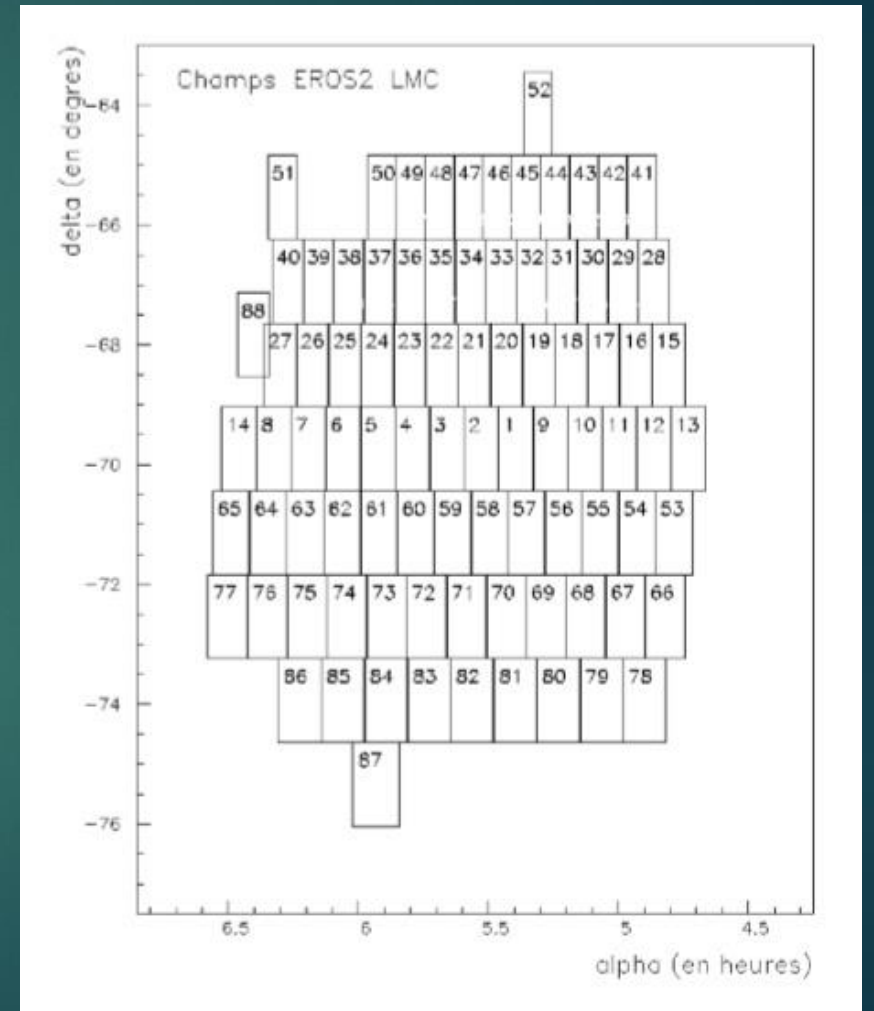
Les Champs du Grand Nuage



Les Nuages de Magellan

Le Petit

Le Grand



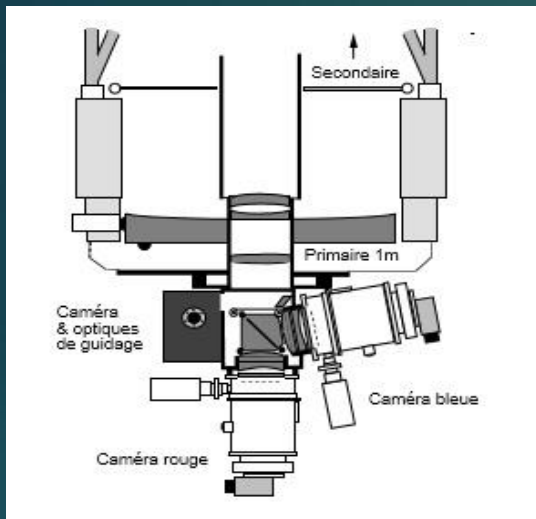
P. Tisserand – fg 4.6b p82

Images et CCD

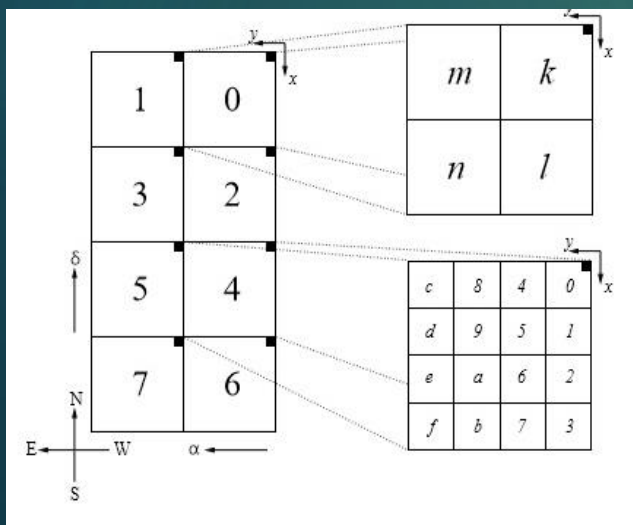
- ▶ Un télescope
- ▶ Un prisme divisant la lumière du télescope en deux flux : rouge et bleu
- ▶ Deux caméras de 8 CCD de grande taille : 2048 x 2048 pixels 16 bits
 - ▶ Chaque caméra reçoit la lumière des mêmes étoiles de la même zone au même moment.
- ▶ 16 images de 8 MB capturées lors d'une même prise de vue
 - ▶ Une moyenne de 60 champs par nuit
 - ▶ Près de 1 000 images par nuit
 - ▶ A peu près 8 GB

La double caméra

P. Tisserand – fg 4.1 pg 75



P. Tisserand – fg 4.2 pg 76



Jna



Jna

Images brutes et réduites

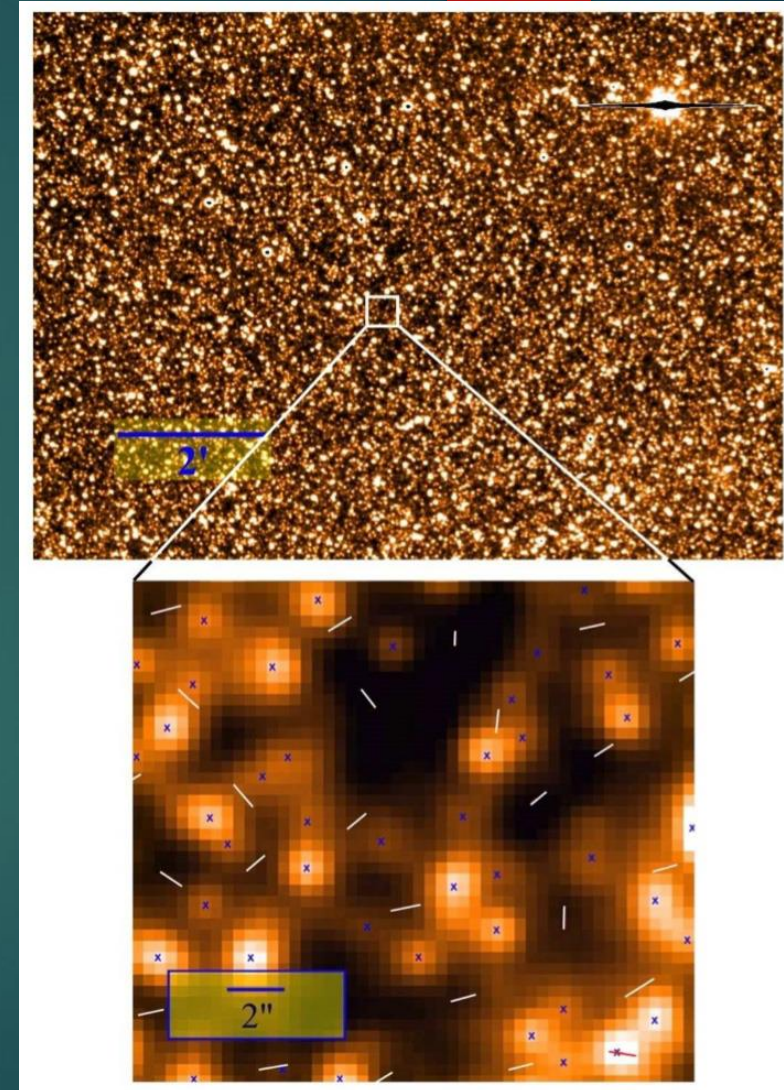
- ▶ Les images issues des CCD sont des images « brutes »
- ▶ Elles sont « normalisées » grâce à des images « de calibration » pour former les images « réduites » utilisées dans les analyses.
- ▶ La réduction est réalisée directement sur le site.
- ▶ Il y a donc une image « brute » pour chaque image « réduite ».
- ▶ Images brutes, réduites et de calibration envoyées par DLT.

Transfert des images

- ▶ Les images brutes sont sauvées sur des DLT « brutes »,
Les images réduites sont sauvées sur des DLT « réduites »
- ▶ Les DLT sont envoyées à Saclay
 - ▶ Les DLT brutes restent à Saclay
 - ▶ Les DLT réduites continuent vers le CC à Lyon
- ▶ Les images réduites sont recopiées sur des 3490
 - ▶ L'emplacement des fichiers et les principales caractéristiques des images sont enregistrées dans la base de données Oracle de l'expérience
 - ▶ La gestion des cartouches et du placement des fichiers est assuré par un environnement développé autour de TCL.
- ▶ Plusieurs réorganisations ont été réalisées autour du HPSS, des streamers modernes et enfin de Irods

Les images Eros II

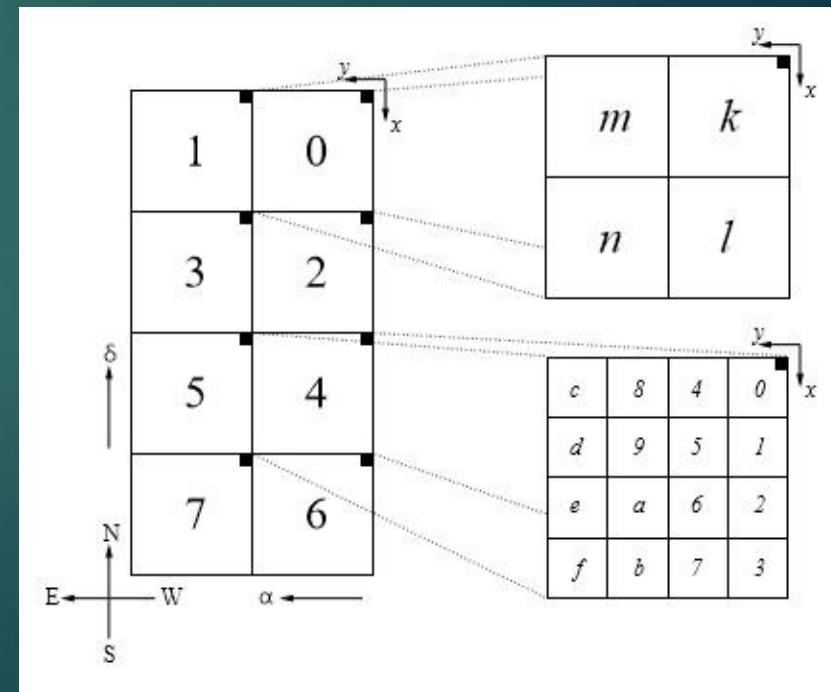
- ▶ 2 027 572 images référencées dans la base de donnée
- ▶ 1 884 001 images effectivement présentes dans Irods, soit 14.3 TB
- ▶ 1 832 463 images réduites, soit 14.0 TB
- ▶ 80 650 images perdues, soit 4.0 %, dont l'ensemble du programme « Naines rouges »
 - ▶ Episodes de crash des cartouches
 - ▶ Erreurs lors des restructurations ?



Quart de CCD

- ▶ Le « quart de CCD » est un sous-ensemble d'un CCD de 1600 x 1600 pixels
 - ▶ Rappel : un CCD = 2048 x 2048 pixels 16 bits
- ▶ Les quarts de CCD sont nommés « k », « l », « m » et « n ».
- ▶ Un quart de CCD : 1600 x 1600 pixels 16 bits, soit 6.2 Mo
 - ▶ Rappel : une image CCD = 8 Mo

P. Tisserand – fg 4.2 pg 76



Traitement des données

- ▶ Une étape sur le site : la réduction des images brutes
- ▶ Plusieurs étapes au Centre de calcul
 - ▶ Création de très bonnes images pour une meilleure détection des étoiles
 - ▶ Détection des étoiles et création de catalogues
 - ▶ Mesure des paramètres des étoiles des différentes images
 - ▶ Création de courbes de lumière ASCII par étoile
 - ▶ Pour réduire la taille des fichiers et les volumes traités, les analyses sont conduites par quart de CCD et par flux de couleur

Production

- ▶ Traitement massif des images pour une version du logiciel d'analyse et un ensemble de paramètres.
 - ▶ Réalisé en batch sur la ferme Anastasie du Centre de calcul
 - ▶ En général par programme et sous-ensembles d'un programme
- ▶ Les références et les suivis sont identifiés par un code de traitement et un numéro de version, correspondant à une « production ».
- ▶ Les « productions » P1 et P5 représentent 85% des catalogues et 75% des mesures.
- ▶ Seule la production P5 a conduit à la création de courbes de lumière.

Majeurs, mineurs et exotique

- ▶ Trois « catégories » de programmes : « majeurs », « mineurs » et « exotiques »...
- ▶ Les **7 majeurs** : programmes entièrement analysés avec des courbes de lumière
 - ▶ Beta Scuti (BS), Centre galactique (CG), Gamma Normae (GN), Gamma Scuti (GS), Grand et Petit nuage de Magellan (LM, SM), Theta Muscae (TM)
 - ▶ 7 programmes sur 33 référencés
 - ▶ 278 champs sur 1594 référencés
 - ▶ 1 763 282 images réduites sur les 1 974 701 référencées ~ 90 % des observations
- ▶ **4 mineurs** : programmes analysés sans courbes de lumière
 - ▶ Tous ces programmes sont marqués « Monte-Carlo », copie de CG et LMC

Image composée

- ▶ Très bonnes images créées à partir d'une petite série de très bonnes images pour améliorer la détection des étoiles
 - ▶ 12-15 images par série
 - ▶ Les images créées sont nommées « images de références »
 - ▶ Elles sont produites par couleur et par quart de CCD
- ▶ 12 652 images de références calibrées astronomiquement
 - ▶ Rappel : il y a 278 champs « principaux » x 8 CCD x 2 caméras x 4 quarts, soit en principe 17 792 qCCD
 - ▶ 14 012 images « c » 84.8 GB
 - ▶ 12 652 images « w » 76.5 GB
 - ▶ 19 866 images « x » 120.0 GB

Catalogues d'étoiles

- ▶ Les catalogues d'étoiles regroupent les références des étoiles détectées
 - ▶ Ces fichiers nommés « fichiers de référence »
 - ▶ Format binaire propre à l'expérience
 - ▶ Un fichier de références par couleur et quart de CCD
- ▶ Les étoiles détectées sont associées dans les deux couleurs pour permettre un traitement indépendant des flux bleus et des flux rouges
- ▶ Un jeu de fichiers de références par « production »
 - ▶ 13 376 références P 5 pour les 7 programmes majeurs
 - ▶ 13 344 références P 5 sauvées dans Irods – 66.9 GB

Suivis

- ▶ Les fichiers de suivis regroupent les mesures réalisées sur les étoiles des différentes images
 - ▶ Format binaire propre à l'expérience
 - ▶ Analyse réalisées par quart de CCD et par couleur, par lot, en batch, au fur et à mesure des arrivées
- ▶ Division en plusieurs blocs pour réduire les mouvements avec les 3490
 - ▶ Chaque bloc correspond à un fichier SV identifié par son numéro de blocs
- ▶ 72 234 fichiers de suivis pour les 7 Majeurs P5 dans Irods
 - ▶ Soit 6.3 TB
- ▶ 13 343 blocs #1 – donc de suivis distincts – soit 75 % des champs « majeurs »

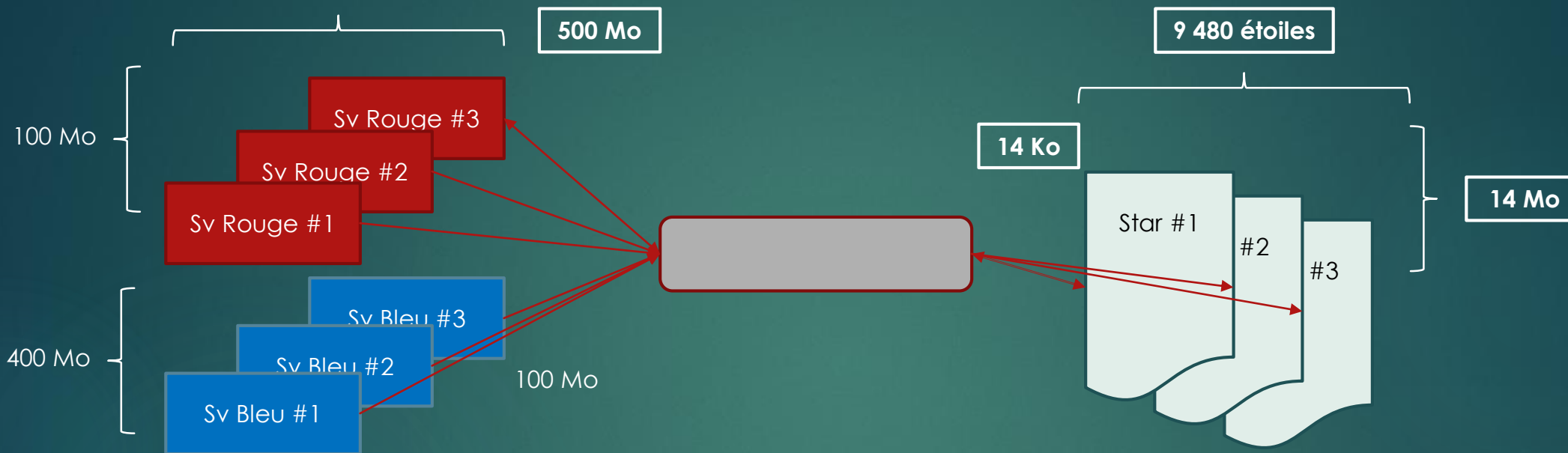
Réorganisation des résultats

- ▶ L'organisation en couleurs, quarts de CCD et blocs de suivis est efficace pour des productions massives
- ▶ Pas très pratique pour des analyses fines sur de petites sélections d'étoiles, surtout si elles sont éloignées
- ▶ La réorganisation consiste à regrouper les mesures de l'ensemble d'une même étoile pour les deux couleurs dans un seul fichier
 - ▶ Ces fichiers sont nommés « courbes de lumière »

Les courbes de lumière

- ▶ Regroupe les 7 années de campagne dans les deux couleurs dans un seul fichier par étoile suivie
- ▶ Format ASCII ne nécessitant pas de librairie de lecture dédiée
 - ▶ Accès universel
- ▶ Ensemble réduit de données : date de mesure, mesures bleues et rouges, erreurs sur les mesures.
- ▶ Fichier de petites tailles : 15-20 KB
- ▶ Environ 1.2 TB non comprimés, soit 450 GB comprimés.

Construction des courbes de lumière



Tous les blocs rouges et tous les blocs bleus pour le même quart de CCD,

Un fichier ASCII par étoile identifiée du quart de CCD contenant toutes les mesures rouges et bleues

Identification et paramètres globaux	# star:	erosid	MagR	ErrMR	XR	YR	MagB	ErrMB	XB	YB
	#	Im0153n22431	19.133	0.124	478.78	684.37	19.238	0.070	431.84	713.02
	#									
Mesures rouges et bleues	#	date	MagR	ErMagR	MagB	ErMagB				
		348.81416	19.224	0.037	19.330	0.025				
		366.79045	99.999	9.999	19.337	0.035				
		373.68669	19.115	0.051	19.299	0.032				
		376.84666	19.195	0.058	19.287	0.036				
		380.86332	19.210	0.052	99.999	9.999				
	381.70421	19.186	0.081	99.999	9.999					

90 millions d'étoiles

Programme	Etoiles bleues	Rouges	Moyennes	Courbes	%
Bs	6 699 360	6 585 525	6 642 443	2 416 924	36%
Cg	92 001 893	92 642 126	92 322 010	43 681 008	47%
Gn	12 871 686	13 018 914	12 945 300	4 307 755	33%
Gs	5 589 593	5 428 873	5 509 233	1 804 578	33%
Lm	68 696 035	70 057 582	69 376 809	28 797 428	42%
Sm	8 708 758	8 887 279	8 798 019	4 065 298	46%
Tm	6 355 664	6 143 296	6 249 480	2 129 271	34%
total	200 922 989	202 763 595	201 843 292	87 202 262	43%

Les nombres d'étoiles bleues et rouges proviennent du programme d'analyse et sont enregistrés dans la base de données au niveau de la description des références.

A suivre : perte de données Eros II

Jean-Noël Albert – LAL / IJClab

- ▶ Le blog Eros / Anastasis : <https://groups.lal.in2p3.fr/erosanastasis/>
- ▶ Le site ErosDb II : <http://eros.lal.in2p3.fr/ErosDB/>
- ▶ Le site historique Eros : <http://eros.in2p3.fr/>