Reprocessing of D3 CFHT field

Ph.Gris
LPC Clermont-Ferrand

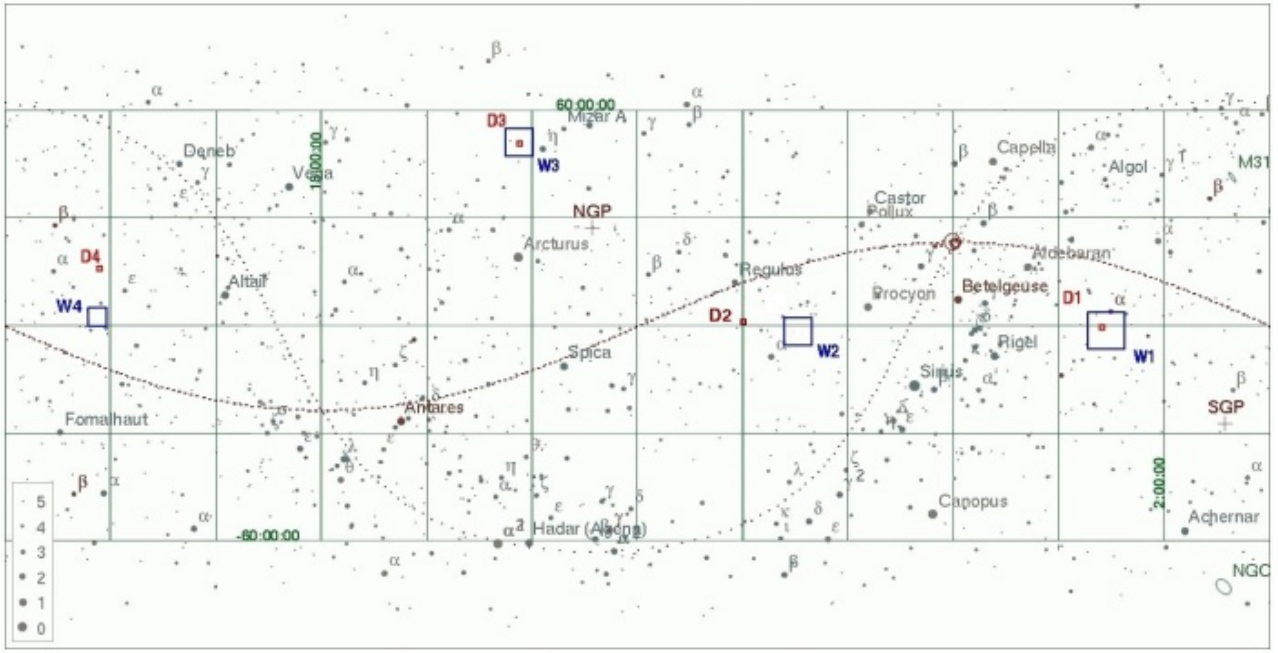
- The subject
- Status of the reprocessing
- Problems encountered
- Next steps

Reprocessing of the D3 CFHT field

- **DESC Task Force (https://github.com/DarkEnergyScienceCollaboration/ReprocessingTaskForce)**
- **Processing with LSST stack: make the job with the same release**
- **Goals:**
  - **technical: exercise the production chain / point at potential problems / effect of data size**
  - **science: quality of the production (performance in astrometry and photometry) for each step of the processing.**
- **Science case: extract the light curves of the Supernovae collected by CFHT**
  - **are we able to extract these light curves from the stack ? Are they of sufficient quality to include the corresponding SN in the Hubble diagram?**

CFHT D3 field

- One of the fourth deep fileds of CFHT
- 1x1 deg2 – RA= 14:19:27 (214.8625 deg) DEC= +52:40:56 (52.6822 deg)
- 121 SN candidates detected



Reprocessing: the steps

Raw CFHT image

Corrected CFHT Image (bias, dark, flat…)

Stack L1

Process CCD
- Calexp Images (Calibrated exposure)
- Sources catalogue

Simultaneous Astrometry
- Calexp Images
- Sources catalogue

Coaddition
- Coadditioned image
- Sources catalogue

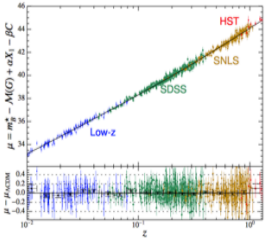
Soustraction
- Sustracted images
- Transient catalogue ?

Stack L2

Lightcurves
- Detection SN Ia?

Stack L3

Cosmology
- Hubble Diagram



Courtesy of A. Ciulli

Reprocessing: the steps

Raw CFHT image

Corrected CFHT Image (bias, dark, flat…)

Stack L1

Process CCD
- Calexp Images (Calibrated exposure)
- Sources catalogue

Simultaneous Astrometry
- Calexp Images
- Sources catalogue

Coaddition
- Coadditioned image
- Sources catalogue

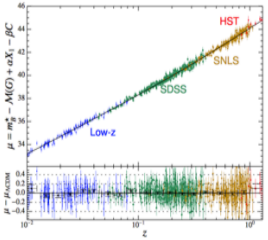
Soustraction
- Sustracted images
- Transient catalogue ?

Stack L2

Lightcurves
- Detection SN Ia?

Stack L3

Cosmology
- Hubble Diagram



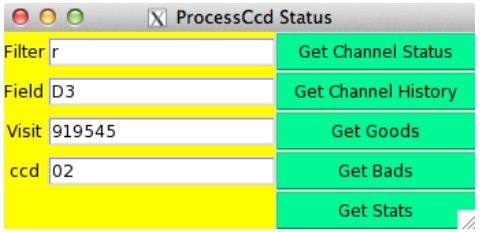
Courtesy of A. Ciulli

D3 : calexps production status

- First step of L2: production of "Calibrated Exposure":
  - Calexp : A fully-qualified LSST image, which includes a science pixel array, and concomitant data including a quality mask and a variance array, in addition to a PSF characterization and metadata (including calibration metadata) about the image (https://confluence.lsstcorp.org/display/LSWUG/DMS+Glossary#DMSGlossary-C)
  - calibrations (astrometry and photometry) (reference catalog: SDSS)
  - production of catalogs (sources) (stars and galaxies)
- Calexps to be produced (D. Boutigny)
  - u : 153 x 36 = 5508
  - g : 542 x 36 = 19512
  - r : 645 x 36 = 23220
  - i : 713 x 36 = 25668
  - z : 682 x 36 = 25582

D3 : Calexps production status

- filter u,g,r,i,z : > 99.5 % of the production completed
- To access the status of the production
  - python script accessible from ccage: python /sps/lsst/dev/lsstprod/CFHT/D3/utils/Production_Status/processCcd/processCcdgui.py &



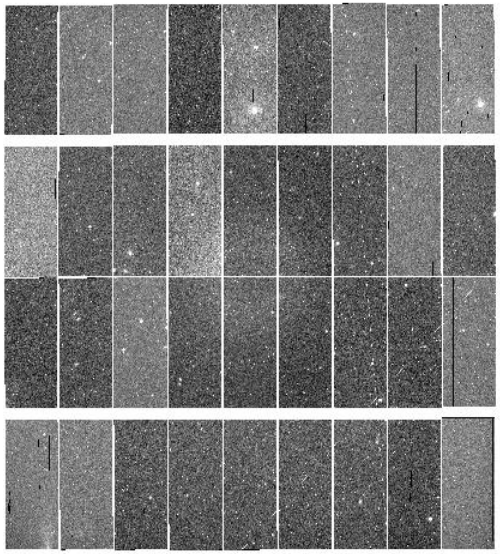
Number of events processed: 5508 / 5508 Nimages= 153
Number of events ok : 5484 frac = 0.996
Number of events not ok : 24 frac = 0.004
Fatal errors :
24 cannot convert float NaN to integer : 5 event(s)
Unable to match sources : 8 event(s)
Unknown : 1 event(s)
No viable PSF candidates survive : 2 event(s)
No sources remaining in match list after magnitude : 6 event(s) No input matches : 2 event(s)

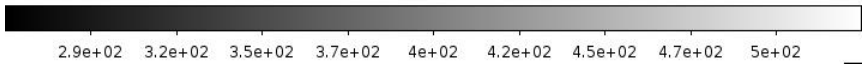
  - from the wiki page of the DESC reprocessing TF: https://github.com/DarkEnergyScienceCollaboration/ReprocessingTaskForce/wiki/CFHT-D3-reprocessing-@ccin2p3
- Calexps: 8.5 Tb, catalogue: 668 Gb (fits)

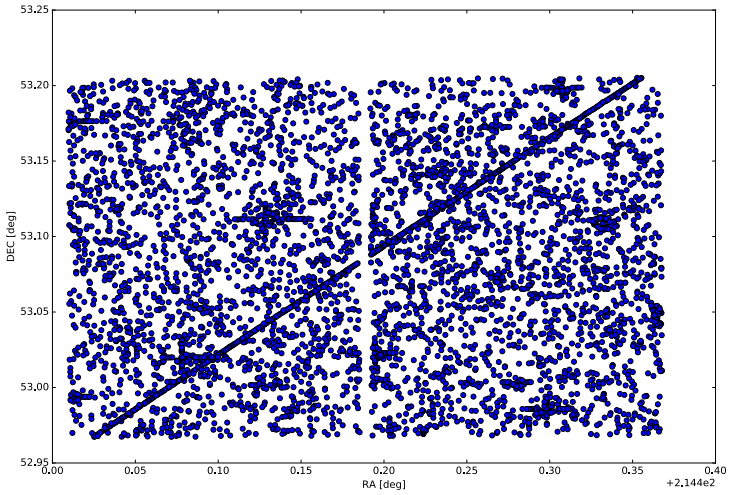
D3 : calexps production status

- Two major problems encountered (up to now):
  - attributed to afs problems:
    - on these workers: impossible to setup the release
    - systematic crash of the job submitted
    - this problem is transient, random and difficult to diagnose
  - Memory problems:
    - at the deblender level
    - identification: “bad::alloc” in error messages
    - appear in images crossed by KFO (Pierre Astier)

May be partly solved by increasing job memory (up to 14 Gb !)







A FOI is being implemented in the stack

D3 : calexps production status

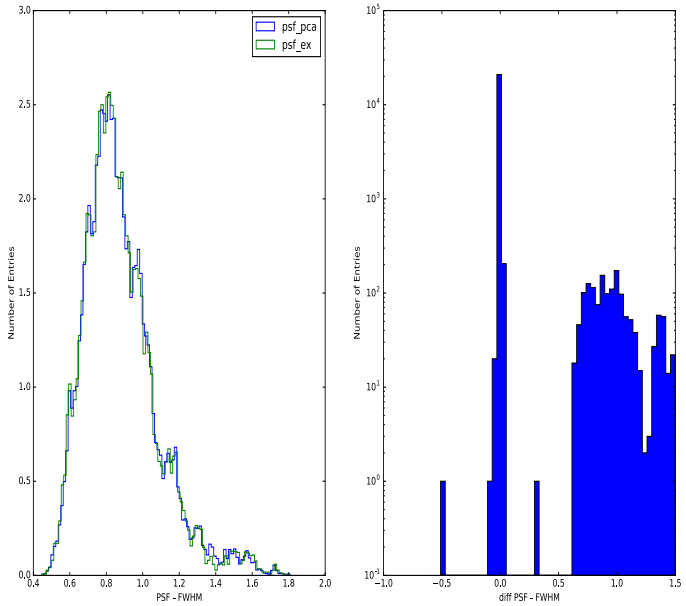
- **Few remarks concerning the processing:**
  - **we would like to use the same release to perform the full L2 processing (let us be optimistic)**
  - **production is made in a rather non professional way (job submissions)**
  - **check of the production: reading of log files -> not optimal**
  - **no ingest in databases for the moment -> will be done at some point (qserv?)**

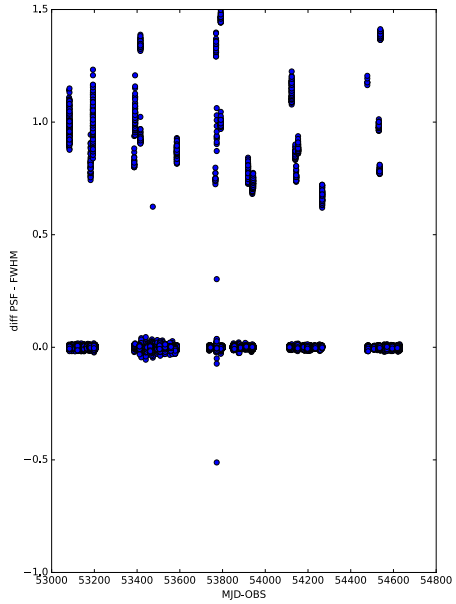
Tools ?

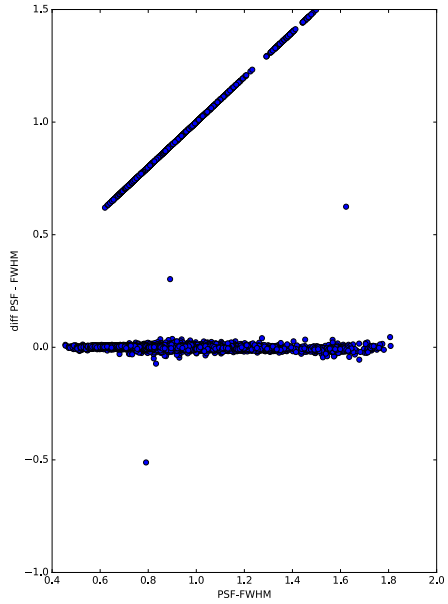
- **Validation of the production:**
  - **estimation of the stack performance (astrometry, photometry)**
  - **package validate_drp : scores and precisions-> see Alexandre's talk during SN session.**
  - **This validation may help at pointing stack shortcomings.**

D3 reprocessing status

- Calexp production u,g,r,i,z validation in progress
- BUT:
  - it seems that the default algorithm to estimate the PSF (psf_pca) and the shape of the objects is not optimal (Dominique B. found some problems when selecting objects – on the basis of their "extendedness" – at the calibration level)
  - production of the r filter with another algorithm for the PSF (psf_ex)







Next steps

- **Evaluate the impact of the new PSF estimation**
- **(Probably) new production u,g,i,z with psf_ex**
- **Simultaneous astrometry (jointcal) on calexps**
- **Coaddition with and without simultaneous astrometry**
- **Qualification of the coaddition step (validate_drp, with modifications)**
- **Forced photometry**