



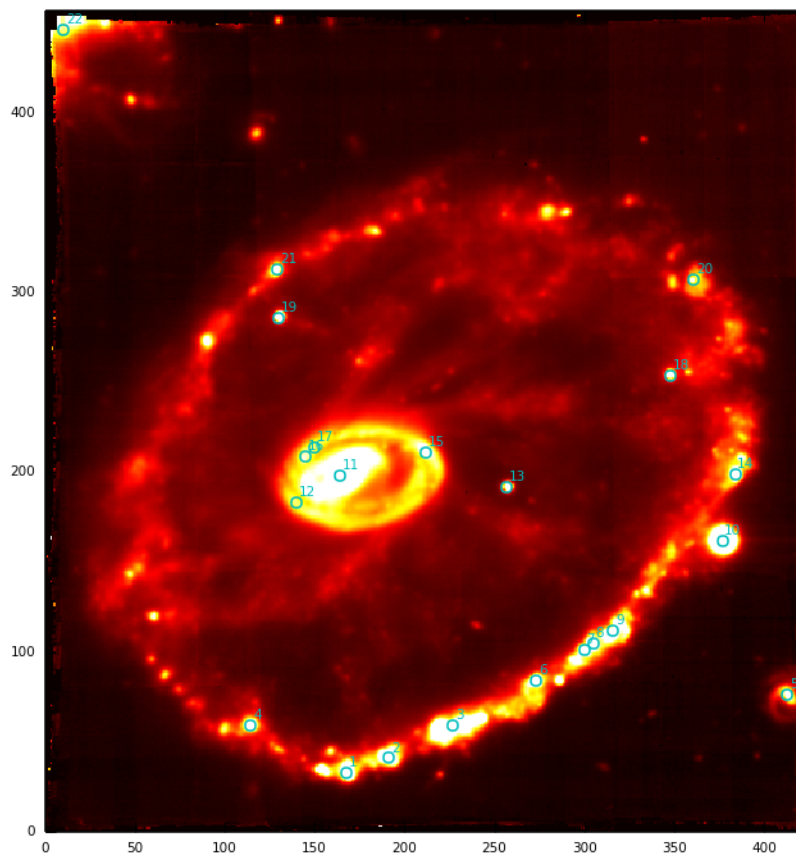
ESO-VLT science data processing

Reinhard Hanuschik

ESO

Data Processing and Quality Control Group

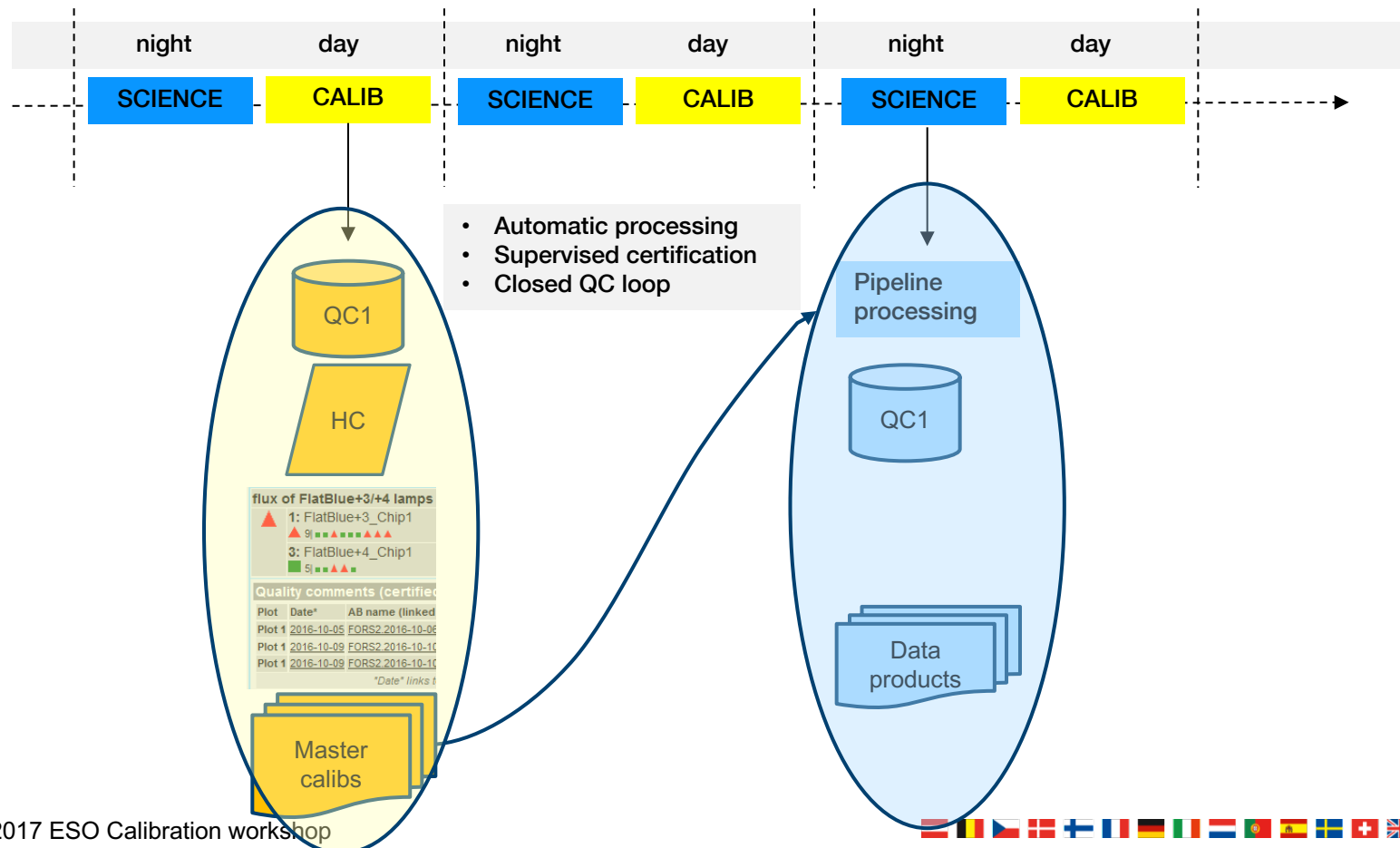
QCG: D. Dobrzycka, W. Hummel, I. Percheron, B. Wolff



Data products: CALIB + SCIENCE

Two pillars of QC services at ESO:

- ◆ CALIB data processing, QC1 parameters, scores, trending
- ◆ SCIENCE data processing



Calibration data processing I

- ◆ Automatic pipelines for all 17 operational instruments
- ◆ QC parameters extracted & stored in database
- ◆ Scores help deciding about quality and certification
- ◆ certification

UVES product quality monitor (date: 2015-02-26)

This is the product quality monitor, with an overview of the processing status of all processing jobs (ABs) and the quality of the products. Those scores which are relevant for instrument health (marked 'HC') are propagated to the HC monitor. The other scores are related to pipeline processing and product quality. Click on the score report to see score details and other information for exploring data quality and trending. For standard DFOS operations, only calibrations are processed, no science data.

last update: 2015-02-27 21:31:33 (UT), machine: muc02 57 ABs in SDFO_AB_DIR browser_refresh: off (every 60 sec | stop | on); tool_refresh: on (every 60 sec + runtime)

number of ABs (all | success | failed | created): 57 | 56 | 1 | 0 scored: 56; result: 1/426

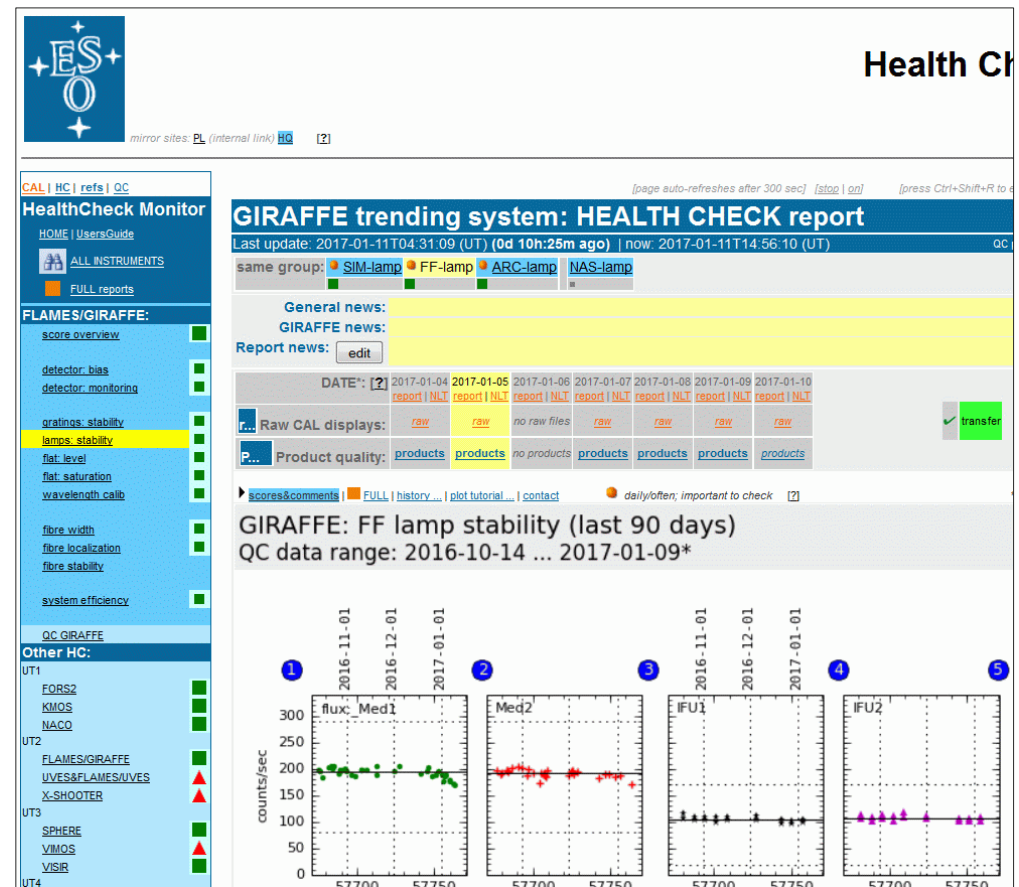
Showing 1 to 11 of 11 ABs (filtered from 57 total ABs)

BQS AB NAME	INDEX	COMPL	AB LOG	RECIPE	RAW_TYPE	SETUP	AB STATUS	P LOG	T_EXEC	QC REPORT	SCORE	CERTIF
UVES_2015-02-27T10:46:39.756_tpl_ab	CAL024	compl	OK	uves_cal_mflat	EFLAT_ECH_BLUE	1_1_1pt/225KHz /g_DICHR#1_FREE_CD#1_346.0_0.40	OK	PLOG	0.1+0.3	QC COVER	✓ (0/2)	AUTO
UVES_2015-02-27T11:01:59.236_tpl_ab	CAL028	compl	OK	uves_cal_mflat	EFLAT_ECH_BLUE	1_1_1pt/225KHz /g_DICHR#2_FREE_CD#2_437.0_0.40	OK	PLOG	0.1+0.3	QC COVER	✓ (0/2)	AUTO
UVES_2015-02-27T11:24:22.382_tpl_ab	CAL032	compl	OK	uves_cal_mflat	EFLAT_ECH_RED	1_1_2pts/225KHz /g_DICHR#1_FREE_564.0_0.30	OK	PLOG	0.2+0.5	QC COVER	✓ (0/2)	AUTO
UVES_2015-02-27T11:39:19.792_tpl_ab	CAL036	compl	OK	uves_cal_mflat	EFLAT_ECH_RED	1_1_2pts/225KHz /g_DICHR#2_FREE_860.0_0.30	OK	PLOG	0.2+0.5	QC COVER	✓ (0/2)	AUTO
UVES_2015-02-27T11:55:46.887_tpl_ab	CAL039	compl	OK	uves_cal_mflat	DFLAT_ECH_BLUE	1_1_1pt/225KHz /g_DICHR#1_FREE_CD#1_346.0_0.40	OK	PLOG	0.1+0.3	QC COVER	✓ (0/5)	AUTO
UVES_2015-02-27T15:25:32.987_tpl_ab	CAL046	compl	OK	uves_cal_mflat	EFLAT_ECH_BLUE	1_1_1pt/225KHz /g_DICHR#2_FREE_CD#2_437.0_1.00	OK	PLOG	0.1+0.3	QC COVER	⚠ (1/4)	OK only 1 flat
UVES_2015-02-27T15:29:56.301_tpl_ab	CAL047	incompl	incompl	uves_cal_mflat	EFLAT_ECH_BLUE	1_1_1pt/225KHz /g_DICHR#2_FREE_CD#2_390.0_1.00	NOK	PLOG	0	missing calib. not used for science.
UVES_2015-02-27T18:40:31.363_tpl_ab	CAL054	compl	OK	uves_cal_mflat	EFLAT_ECH_BLUE	1_1_1pt/225KHz /g_DICHR#2_FREE_CD#2_437.0_0.40	OK	PLOG	0.1+0.3	QC COVER	✓ (0/2)	AUTO
UVES_2015-02-27T19:11:18.806_tpl_ab	CAL055	compl	OK	uves_cal_mflat	EFLAT_ECH_RED	1_1_2pts/225KHz /g_DICHR#2_FREE_860.0_0.30	OK	PLOG	0.8+0.5	QC COVER	✓ (0/2)	AUTO
UVES_2015-02-27T19:33:15.553_tpl_ab	CAL056	compl	OK	uves_cal_mflat	EFLAT_ECH_BLUE	1_1_1pt/225KHz /g_DICHR#1_FREE_CD#1_346.0_0.40	OK	PLOG	0.4+0.3	QC COVER	✓ (0/2)	AUTO
UVES_2015-02-27T19:49:27.873_tpl_ab	CAL057	compl	OK	uves_cal_mflat	EFLAT_ECH_RED	1_1_2pts/225KHz /g_DICHR#1_FREE_564.0_0.30	OK	PLOG	0.8+0.5	QC COVER	✓ (0/2)	AUTO

Calibration data processing II

- ◆ Instrument trending: Health Check Monitor
- ◆ Monitor performance in context
- ◆ Compare to specs, reference values

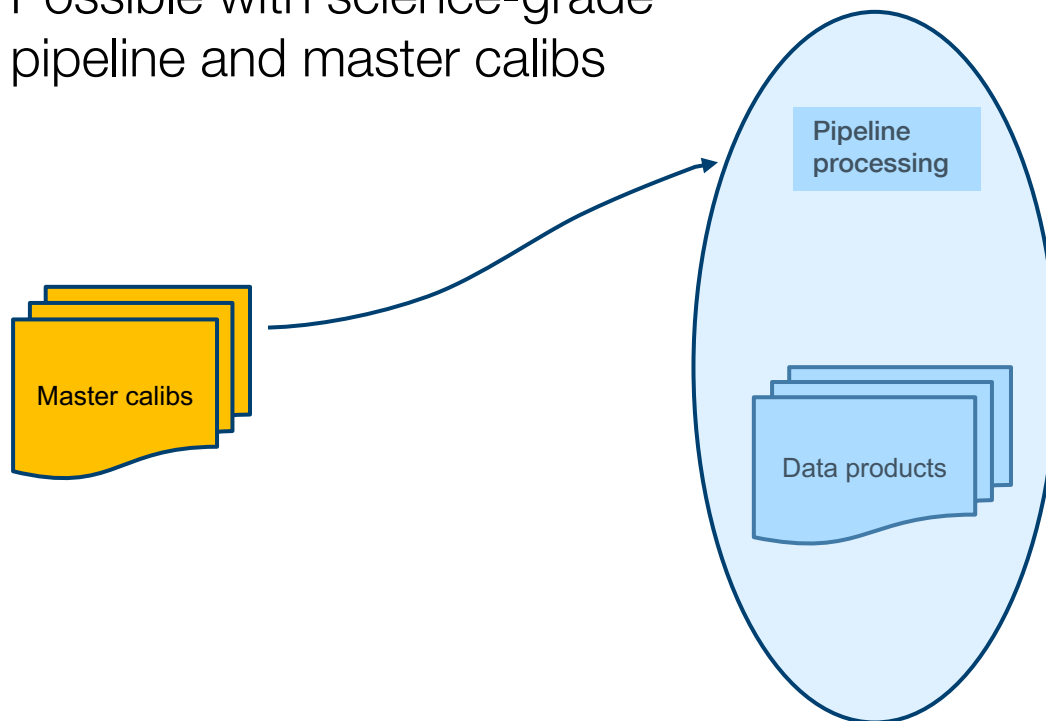
→ Complete knowledge about ins+atm effects stored in master calibration files



Data products: SCIENCE

Two pillars of QC services at ESO:

- ◆ CALIB data processing, QC1 parameters, scores, trending
- ◆ SCIENCE data processing:
 - ◆ Possible with science-grade pipeline and master calibs



Traditional data reduction

- ◆ “Traditional” way towards data products in ground-based astronomy:
 - ◆ Get your raw data, reduce them by yourself, optimized for project
 - ◆ Cumbersome, long learning curve, “data tsunami”



1. Raw data

science photons
plus atmosphere
plus instrument



2. Processed data

cleaning artefacts
extract signal

3. Analyze data: information → knowledge

New ways: IDPs

◆ More efficient:

- ◆ do it centrally
- ◆ offer to whole community (PI, anyone)



1. Raw data

2. Processed data



3. Analyze data:
information → knowledge

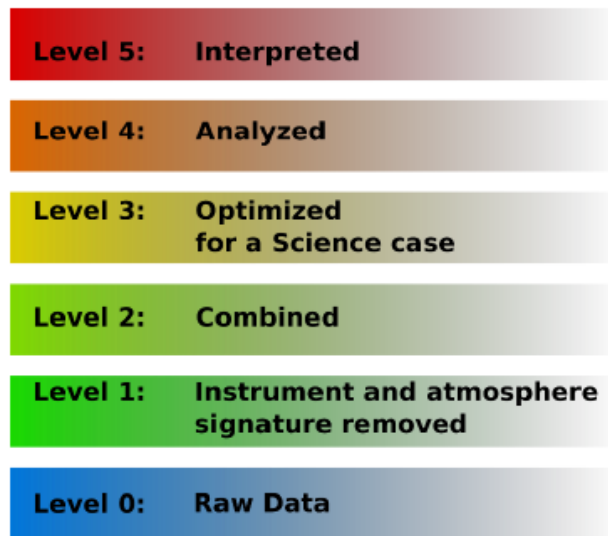
Science data products at ESO

- ◆ IDPs – Internal Data Products:
science data processed in-house at ESO

- ◆ Up to science-grade level:
 - ◆ Instrumental signature removed, atm. effects removed if possible
 - ◆ Physical units (wavelength, flux)
 - ◆ Errors
 - ◆ Ready for analysis

- ◆ Reduction strategy:
 - ◆ Standard reduction (reproduceable and understood)
 - ◆ General strategy (not fine-tuned to specific case)
 - ◆ Reduction efforts done once, centralized, in mass production, with quality control

Science data products at ESO



IDPs:

- processed in-house
- complete by INS
- 100s of programmes per INS
- general-purpose

External data products:

- processed by PIs/data centres
- from surveys or large programmes
- specialized & optimized
- complementary to IDPs

IDP production

- ◆ Needed for the IDP process:
 - ◆ science-grade pipeline, review/certification
(→ *Science Data Products Group*)
 - ◆ master calibrations in archive (covering all instrument aspects, calibration information complete)
 - ◆ QC scheme (QC1 database, scores)
 - ◆ preview plots
 - ◆ data product standard (for 1D spectra)
(→ *Archive Science Group*)

Instruments covered

◆ IDPs:

- ◆ not possible for all VLT instruments (pipelines)
- ◆ since 2013: products from spectroscopic VLT instruments
- ◆ So far 5 projects by QCG:
 - UVES (120,000 spectra);
 - XSHOOTER (45,000 spectra);
 - GIRAFFE (1,200,000 spectra);
 - MUSE (4500 data cubes)
 - coming soon: MUSE_DEEP (200 deep datacubes)
- ◆ also (non-VLT): HARPS, FEROS
- ◆ cover entire history of respective instrument
(UVES: 16 years; GIRAFFE: 13 yr; XSHOOTER: 6 yr; MUSE: 2+ yr)
- ◆ new data added in monthly batches

Phoenix: workflow tool

◆ Important:

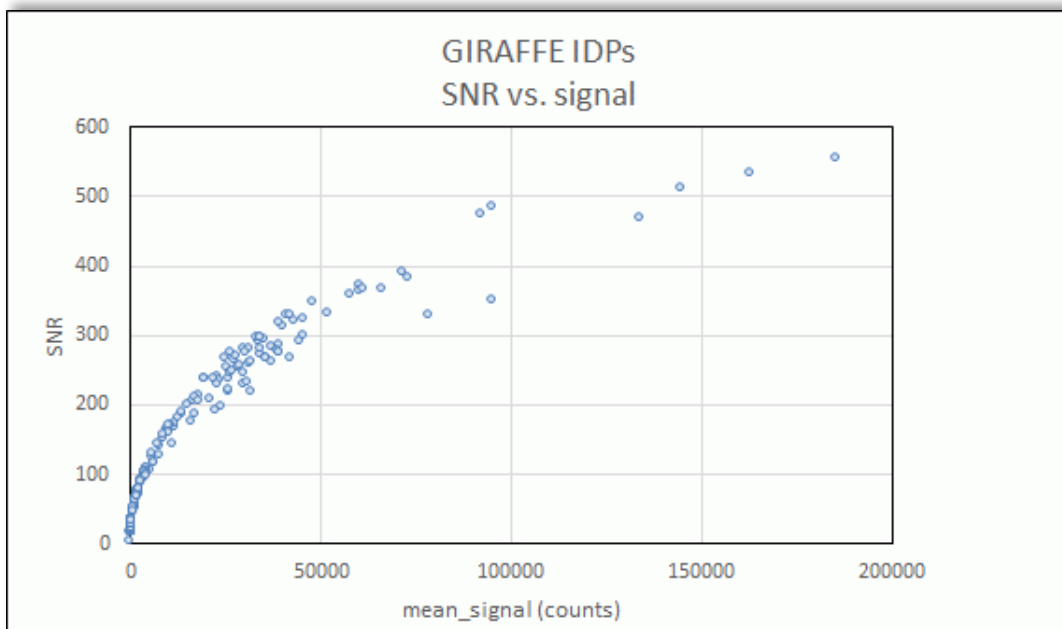
- ◆ Individual issues with calibrations can be ignored; reviewed at QC time → we build on the historical certification
- ◆ Hence we can re-use the certified master calibrations blindly: Important condition for efficient mass-production
- ◆ done with a workflow tool: 'phoenix'

◆ Also:

- ◆ need to check association quality (CAL \leftrightarrow SCI)
- ◆ QC checks for issues with SCI (e.g. saturation)

Process quality control

- ◆ Process QC (as opposed to data QC): we check statistically
 - ◆ SNR \leftrightarrow counts (photon noise dominates error budget)
 - ◆ Association quality (proximity to calibrations)
- ◆ Individually: saturation, number of cosmics
- ◆ Flags: negative counts, anomalies (like no SKY available)



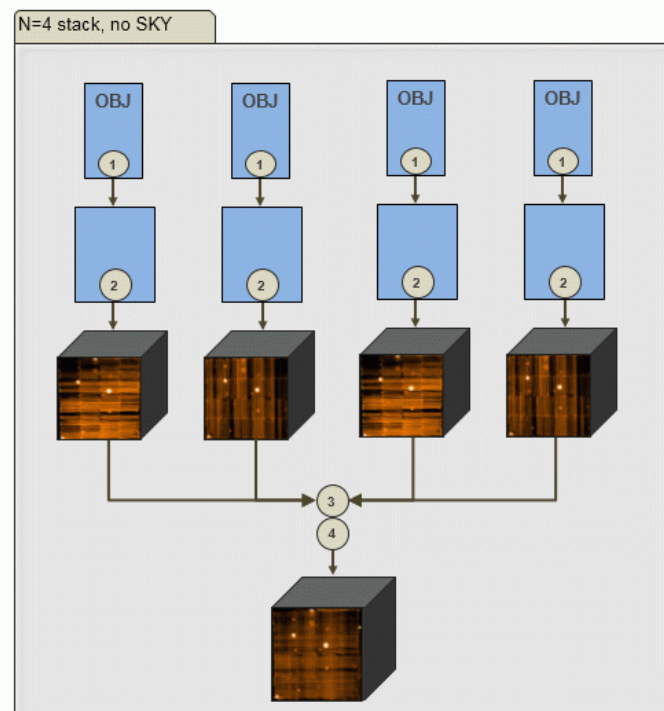
Example: MUSE

- ◆ IDP projects #4 and #5: biggest & most complex so far
- ◆ MUSE: 24 identical spectrographs (4Kx4K each)
 - ◆ Field of view: 1'x1' (WFM)
 - ◆ Spectral resolution: $R=3000$
 - ◆ Spatial resolution: 0.2"/pixel (WFM)
- ◆ MUSE data
 - ◆ Big files (1 GB, 384 mio px): ESO going "big data"
 - ◆ 90,000 spectra in one shot
 - ◆ Challenging in terms of data volume and complexity
 - ◆ Part of community not ready for reducing these data "by hand"



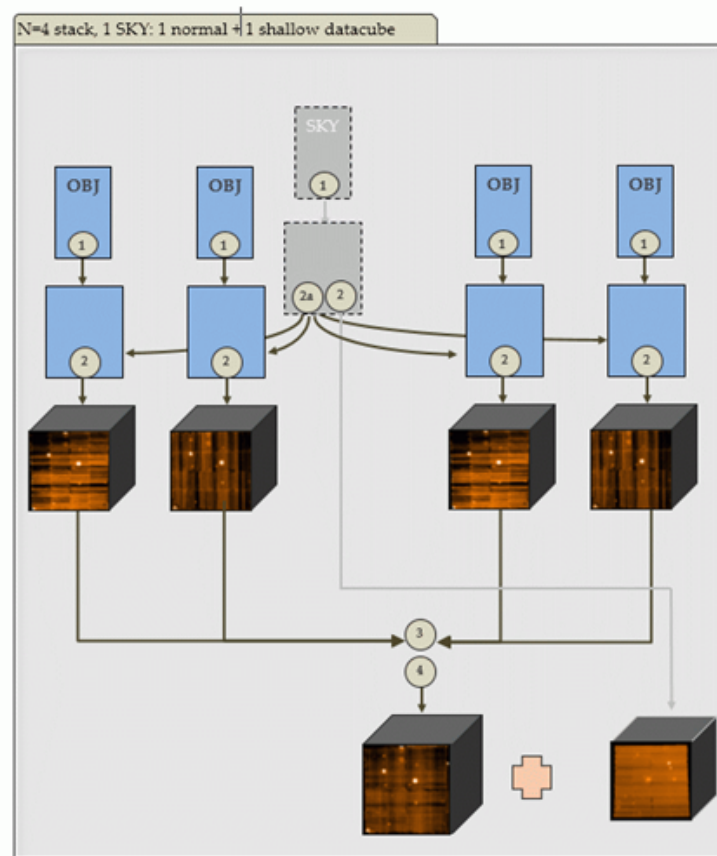
MUSE IDPs

- ◆ MUSE science data processing: requires careful orchestration
 - ◆ up to 5 pipeline recipes
 - ◆ up to 6 steps
- ◆ Typical dataset of 4 raw OBJ:
 - ◆ needs 24 cores, 100 GB memory
 - ◆ takes 2 hrs
 - ◆ delivers 1 OBJ datacube ≥ 3 GB size
 - ◆ flux-calibrated, sky-subtracted: science-ready (still some SKY residuals)



MUSE IDPs

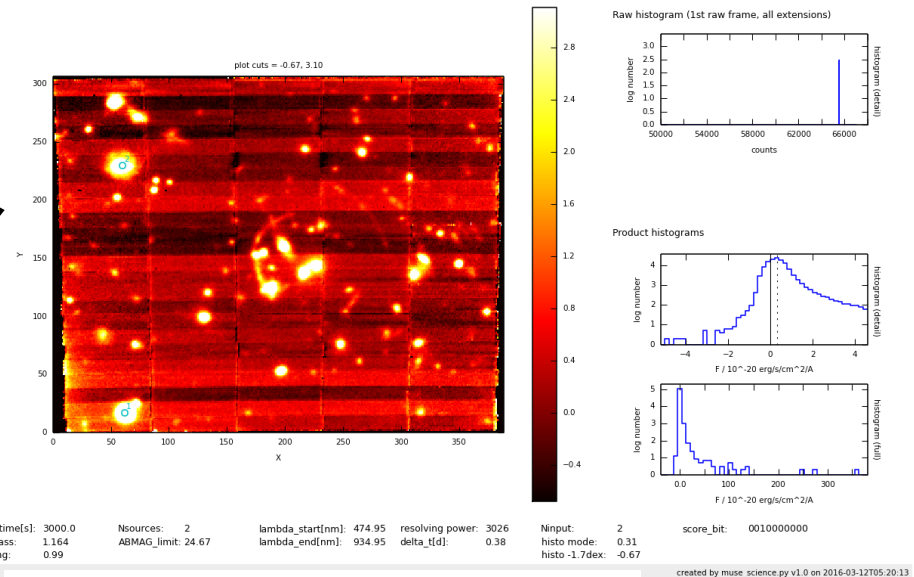
- ◆ OBJ with SKY:
 - ◆ delivers 1 OBJ datacube
 - ◆ plus 1 “shallow” datacube from SKY
- ◆ Archived, downloadable by any user
(proprietary rights respected)



MUSE IDPs

MUSE data product preview r.MUSE.2015-01-10T00:52:40.578_tpl_0000.fits Date: 2015-01-09
 product: DATA_CUBE_COMBINED setup: WFM-NOAO-N
 OBS_ID: 1114634 RUN_ID: 094.A-01411A
 MJD_OBS: 57032.03658076 OB_name: WFM-RCS0224_2
 DPR_TYPE: OBJECT target: WFM-RCS0224-1

- Quality control:
 - QC parameters on single and combined products
 - scores
 - QC reports
 - certification
 - visual inspection



Score report [HELP](#)

MUSE.2014-11-07T07:42:57.317_pst.ab
 RAW_TYPE: PIXTABLE_OBJECT
 TPL_ID: MUSE_wfm-noao_obs_genericoffset
 TPL_NAME: Observation in WFM with user defined offsets

setup: WFM-NOAO-N
 time range: 2014-05-11 ... 2015-05-06

[AB](#) | [ALOG](#) | [PLOG](#) | [QC1 plotter](#) | [factsheet](#)

Point your mouse on QC1 parameter name for short documentation.
 Grey squares: unscored parameter.

NAXIS1	NAXIS2	NAXIS3	num_sat	lshift_max	qc_deltatime	TPL_NEXP
■	■	■	■	■	■	■

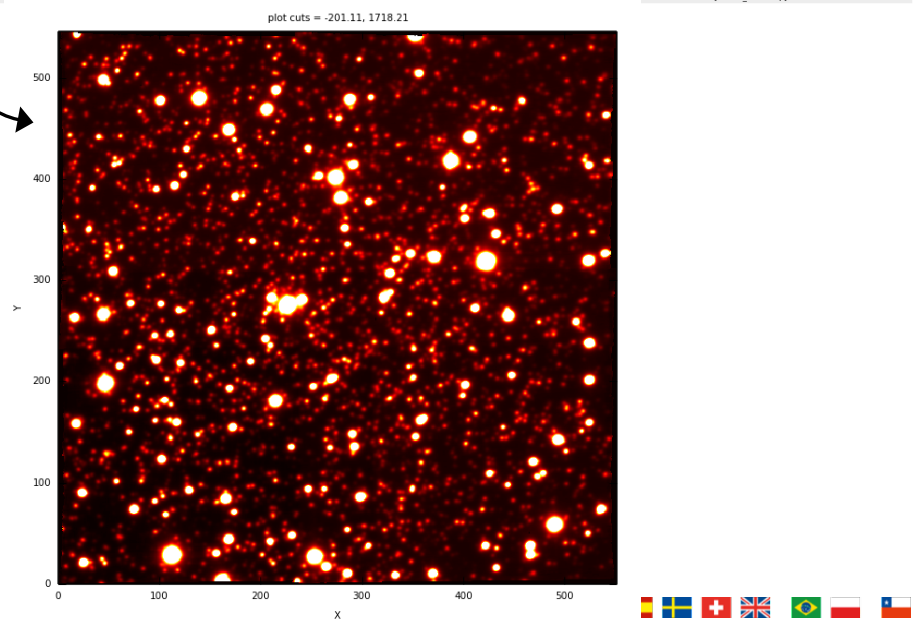
HC plot(s): non configured [explore >>](#)

QC report(s):

Score data: [details](#)

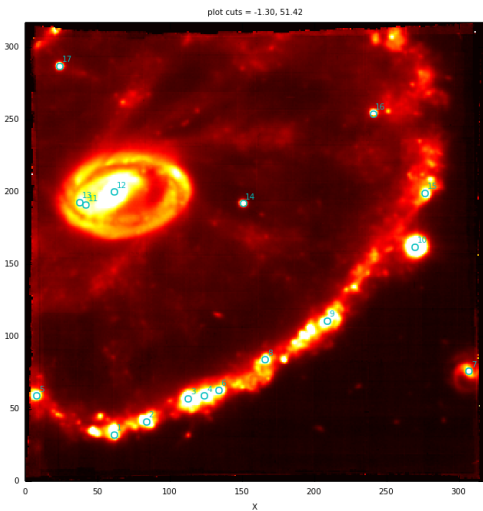
score result: 1/7 best: 0/7

powered by QC [scoreQC v2.1.3]

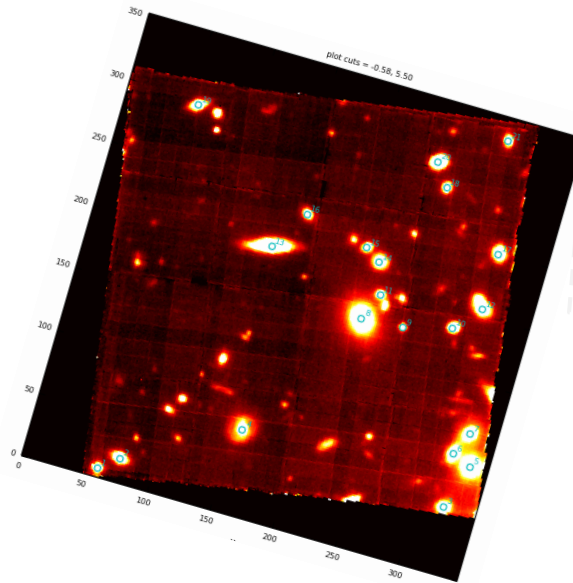


MUSE

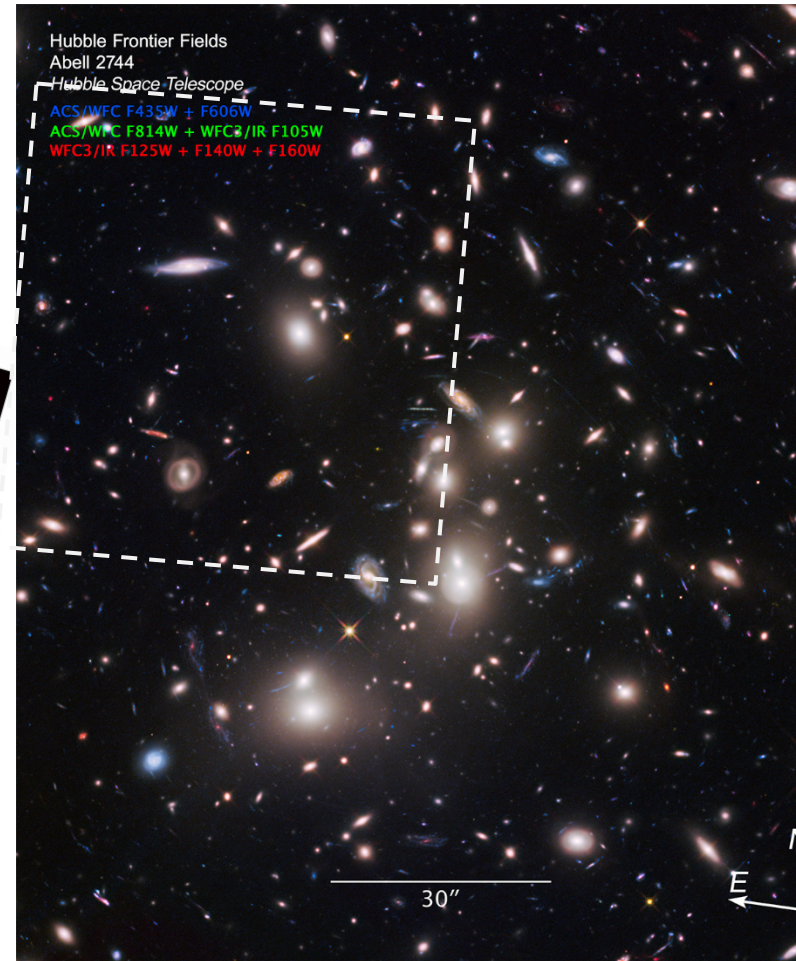
- ◆ Processed + ingested:
 - ◆ Science Verif. #1 and #2
 - ◆ Start of ops (2014-09) until now
- ◆ IDPs at the OB level (single pointings), like for all other IDPs



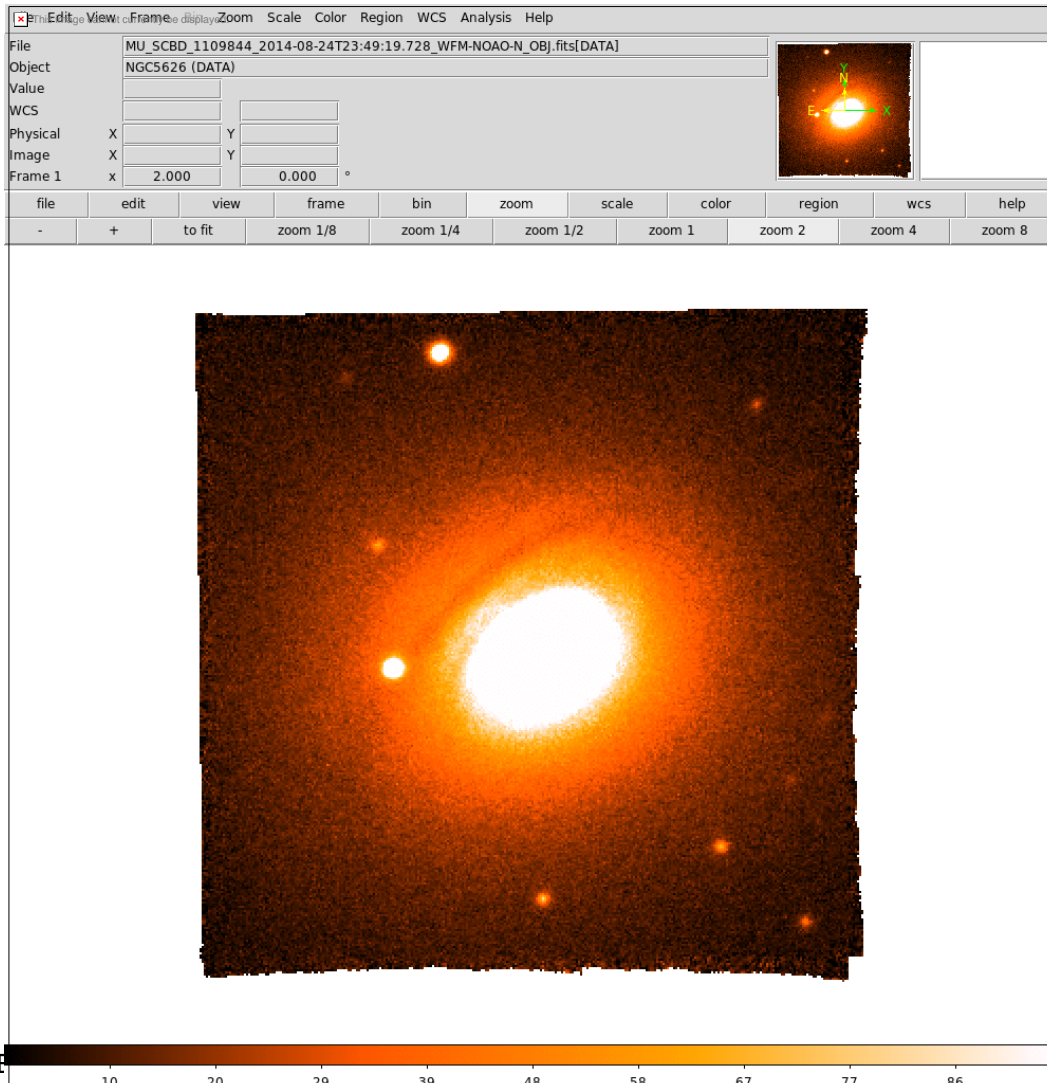
Cartwheel galaxy.
35 min, N=4



Abell 2744.
60 min, N=2



◆ Data quality and content: amazing ...

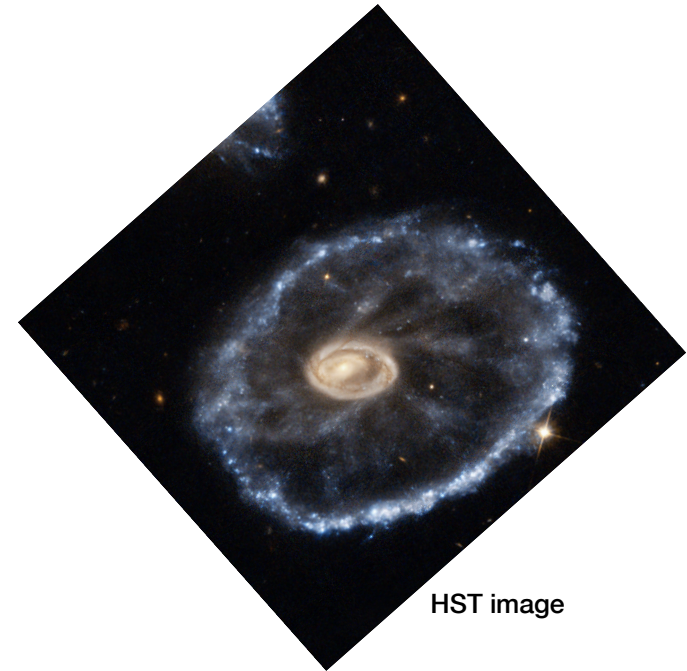
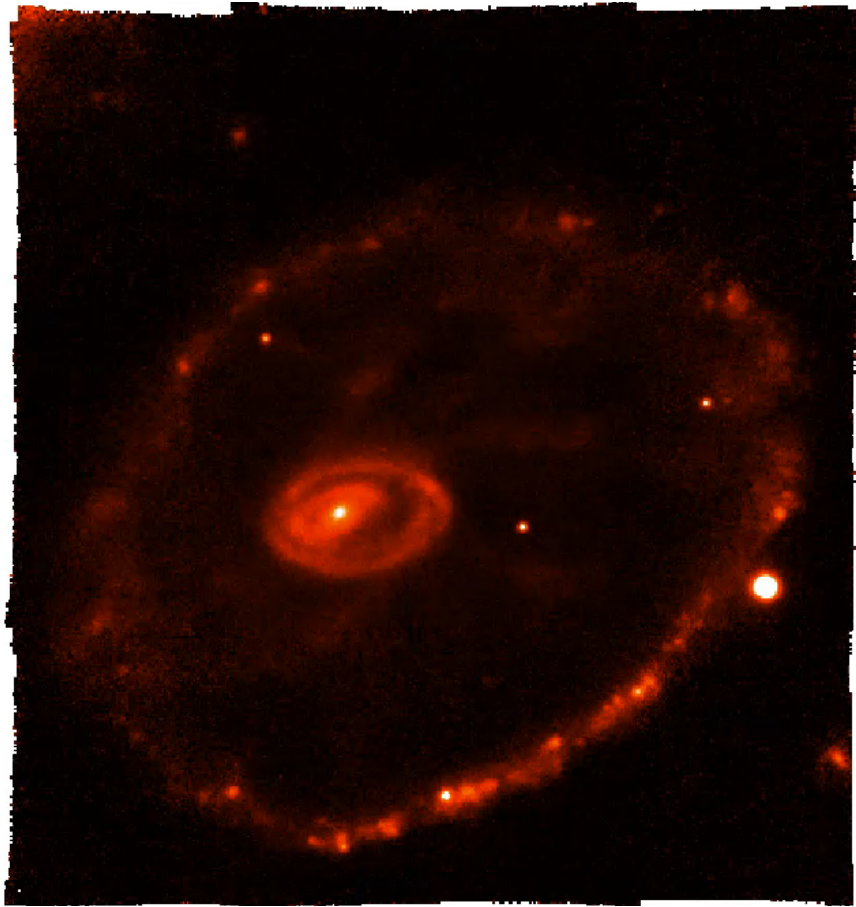


Science Verification run 60.A-9337A:

Velocity maps of emission-line regions in NGC 5626;

17 wavelength bins shown around 6740 Å

MUSE



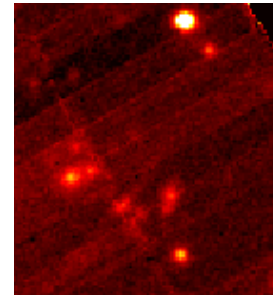
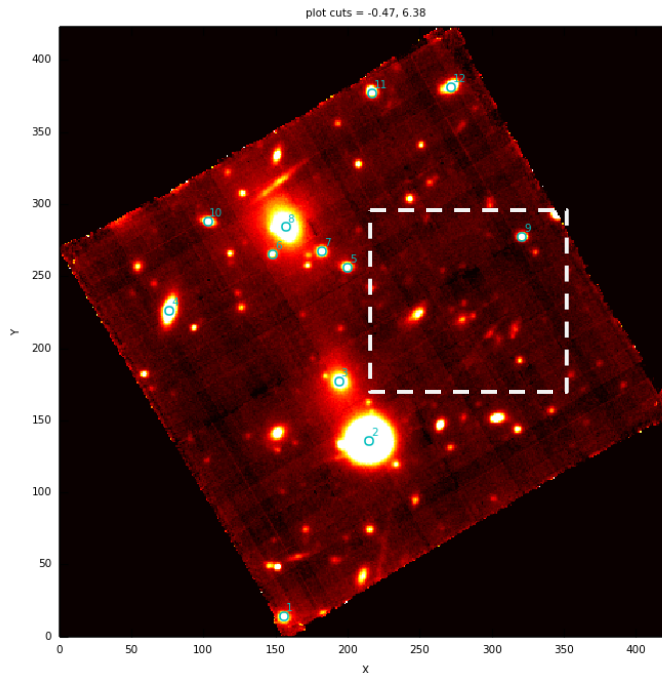
HST image

SV run 60.A-9333A

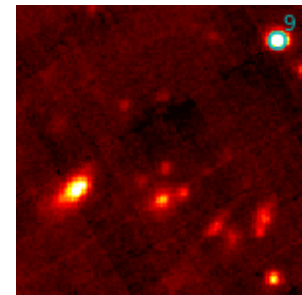
Emission-line regions in
Cartwheel galaxy
(6730-6800 Å; 2 Å/sec)

MUSE: going deep

- ◆ What we have: OB based datacubes
 - ◆ Typical stack of N=2-4, covering ~ 1 hour (operational limit)

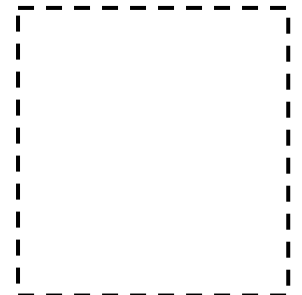


Single exposure,
10 min



One OB, 40 min,
4 exposures

**MUSE OB-
based IDPs**



15 OBs, 11.2 hrs,
58 exposures

**MUSE_DEEP
IDPs**

- ◆ for many programmes: not good enough
 - ◆ > 50% of all MUSE runs: want to go DEEP!
 - ◆ OBs designed with 10 hr or more, in different nights

MUSE_DEEP

- ◆ going DEEP:
 - ◆ multiple visits of same target: multiple OBs
 - ◆ some spread across multiple run_IDs
 - ◆ challenge to find them

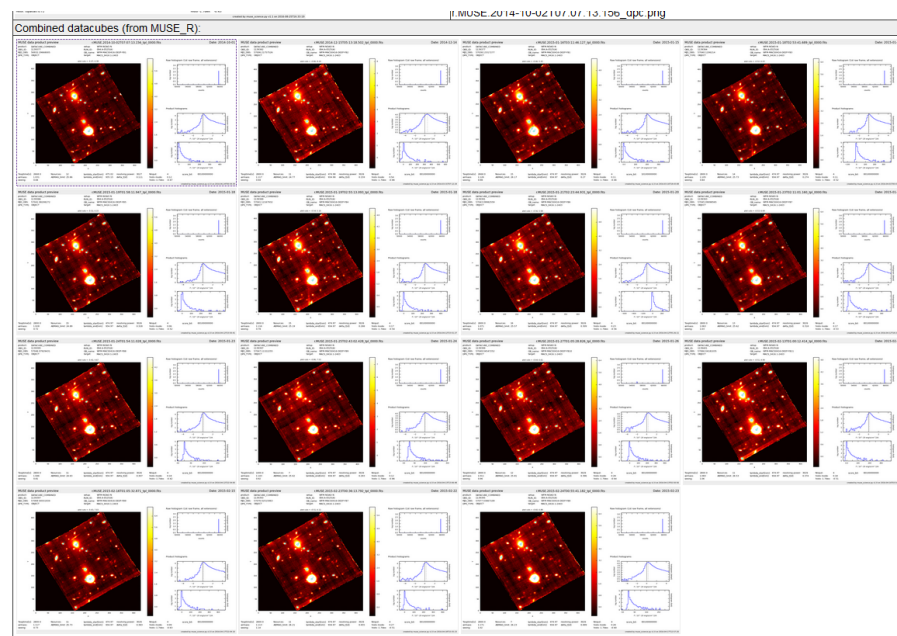
- ◆ stable calibration plan at its best!

- ◆ challenge to process them
 - ◆ 2 stages: OB processing
 - ◆ combine all products from all OBs to final DEEP product

- ◆ full scientific signal available ONLY in deep cubes

Abstract: "... we propose a pilot study to spectrally image ... the HFF Cluster MACSJ0416.1-2403 with MUSE. The resulting datacube will yield robust redshifts for strong emission-line objects out to $z=6.6$, enable emission-line diagnostic, kinematics, and morphology studies ..."

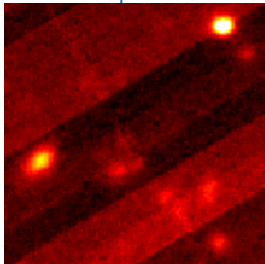
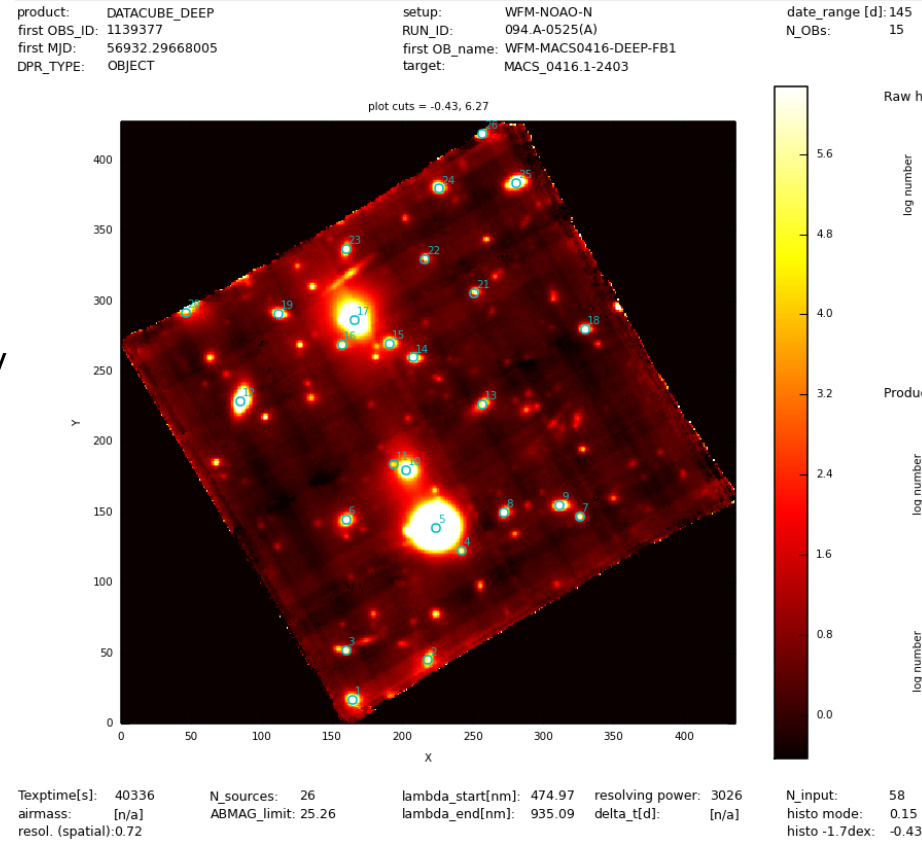
15 OBs spread across a semester; currently 15 individual datacubes; MUSE_DEEP: 1 deep cube



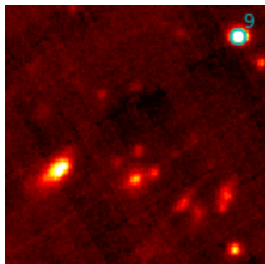
MUSE_DEEP

◆ Example: what DEEP can mean

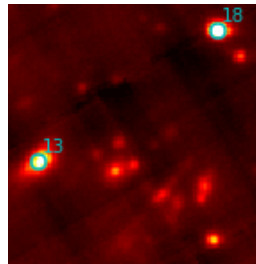
- ◆ 15 OBs combined
 - ◆ 58 input files
 - ◆ 145 days apart
 - ◆ 11.2 hr exposure time
 - ◆ 15.8 hr compute time on 48-core blade, 1 TB memory
 - ◆ Current high-end of IDPs
- ## ◆ 200 such deep cubes
- ◆ from SV+P94+P95, to be available soon



Single, 10 min

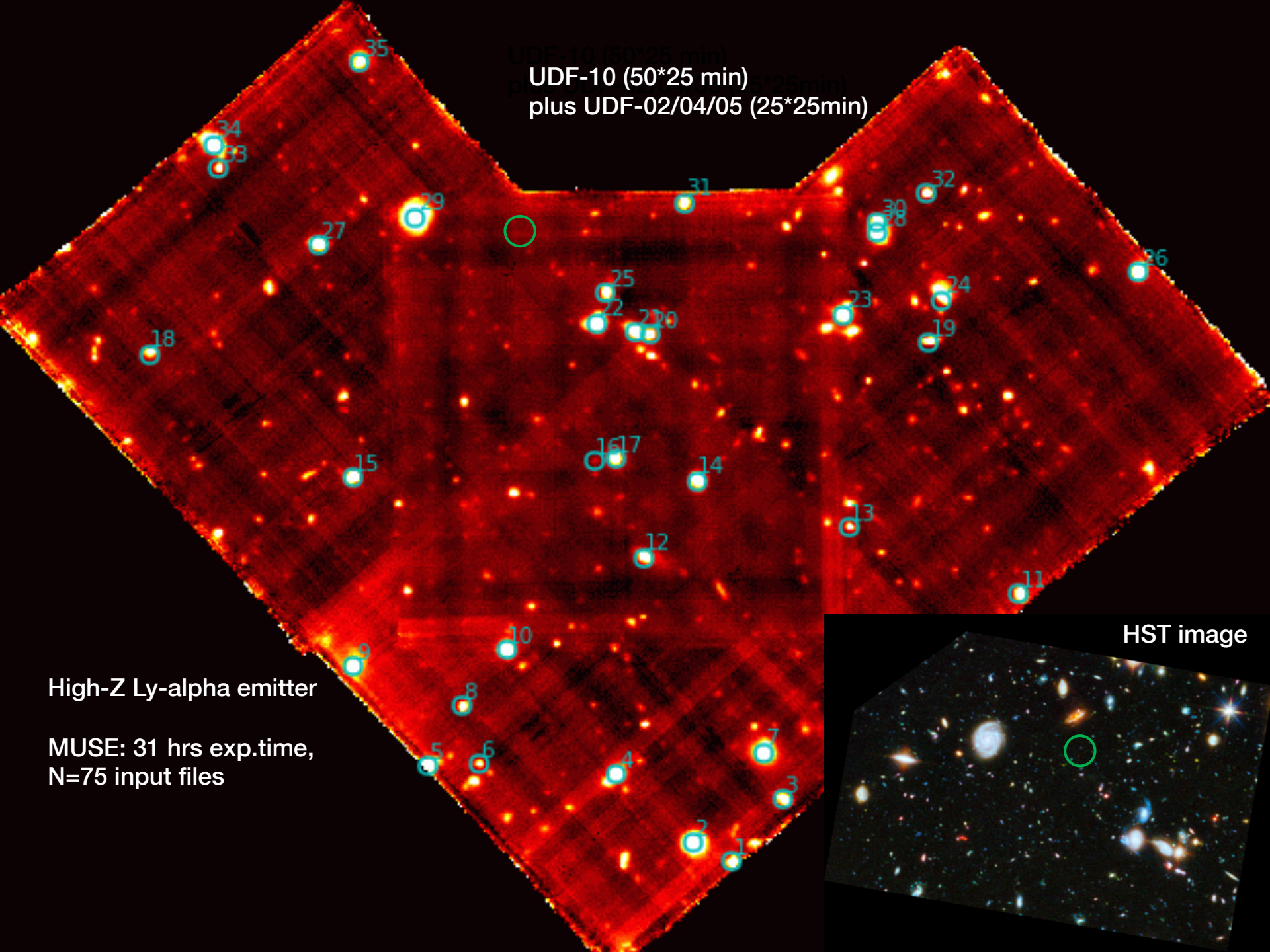


One OB, 40 min, 4 exposures



15 OBs, 11.2 hrs, 58 exposures

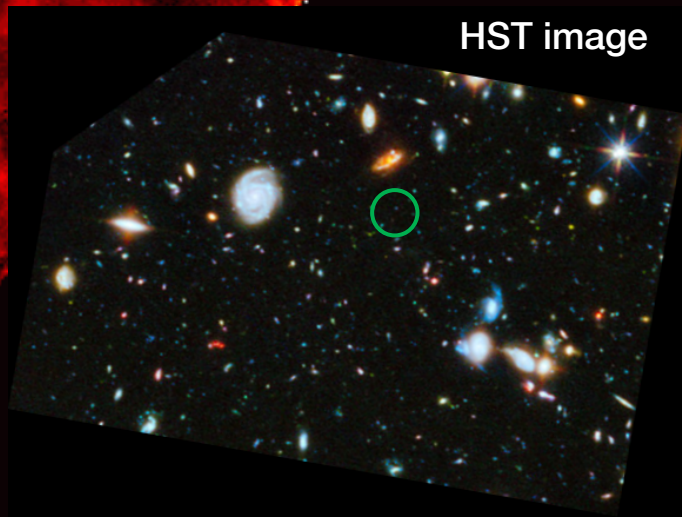
UDF-10 (50*25 min)
plus UDF-02/04/05 (25*25min)



High-Z Ly-alpha emitter

MUSE: 31 hrs exp.time,
N=75 input files

HST image



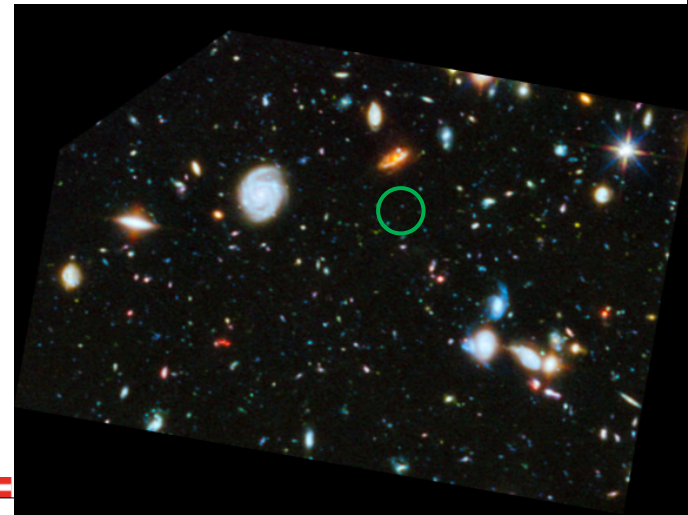


UDF-10 (50*25 min)
plus UDF-02/04/05 (25*25min)



High-Z Ly-alpha emitter

Wavelength scan 669nm+

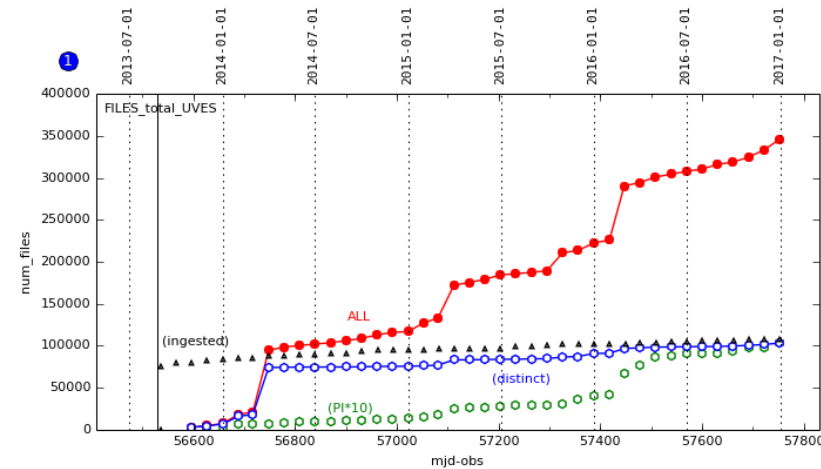


IDPs: downloads

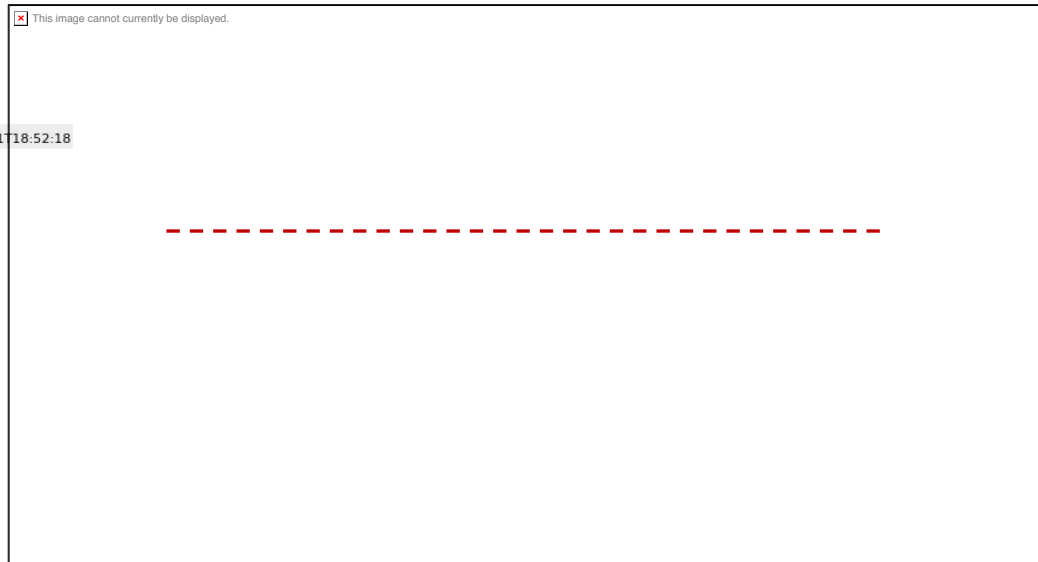
◆ IDPs are used: most popular data products at ESO

IDPs: files&requests (FULL range, 1291 days, close-up)
Data range: 2013-07-01 ... 2016-12-30*

UVES IDP downloads



← Total downloads (~350,000)
← available IDPs (~105,000)
← PI downloads *10 (~10,000)



powered by QC: www.eso.org/HC created by trendPlotter v3.6.2 on 2017-01-11 18:52:18

IDP downloads normalized:
fractional downloads vs. time
since release



Thanks!

