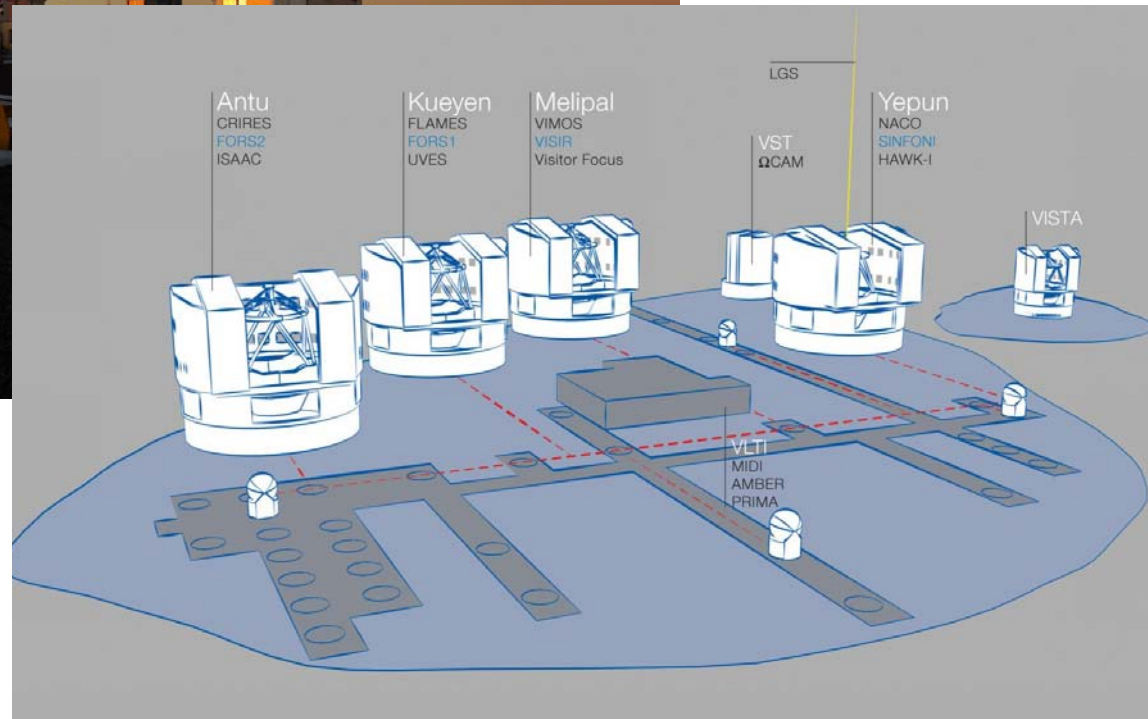
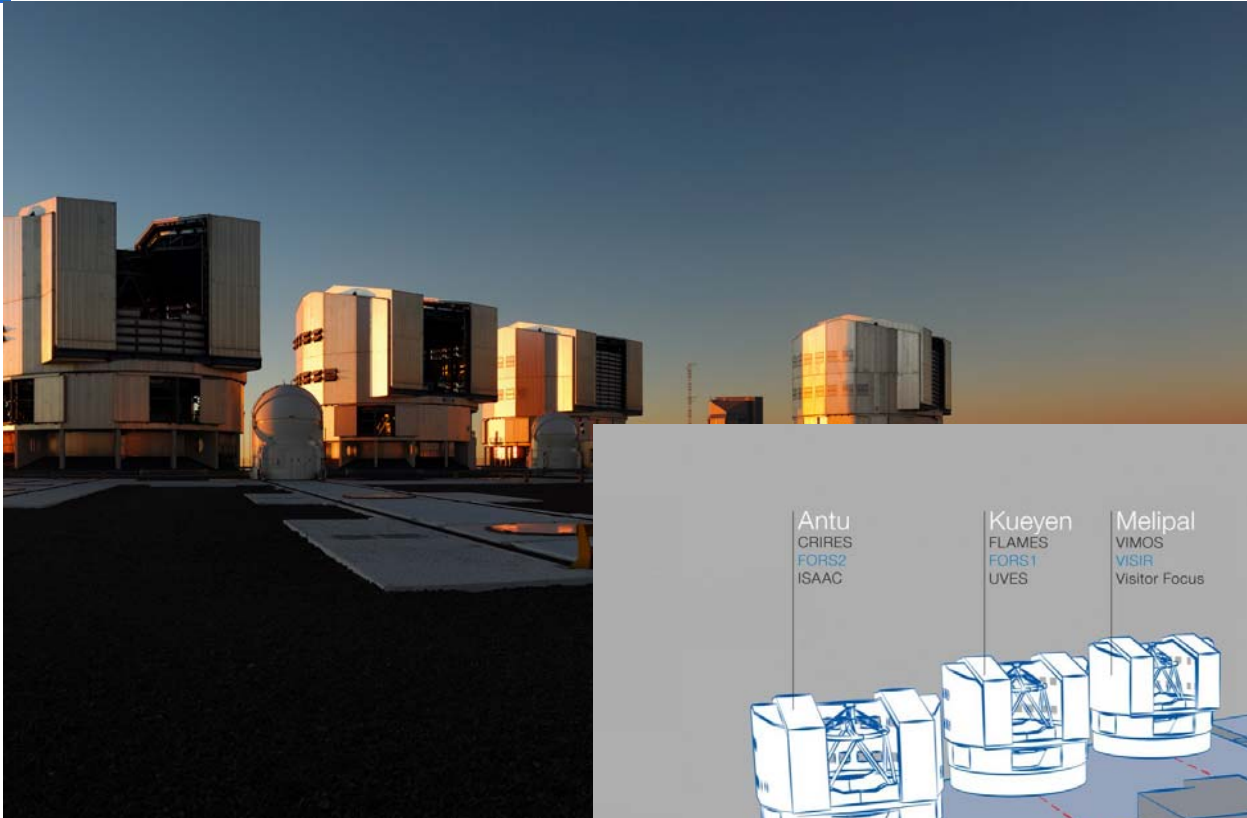




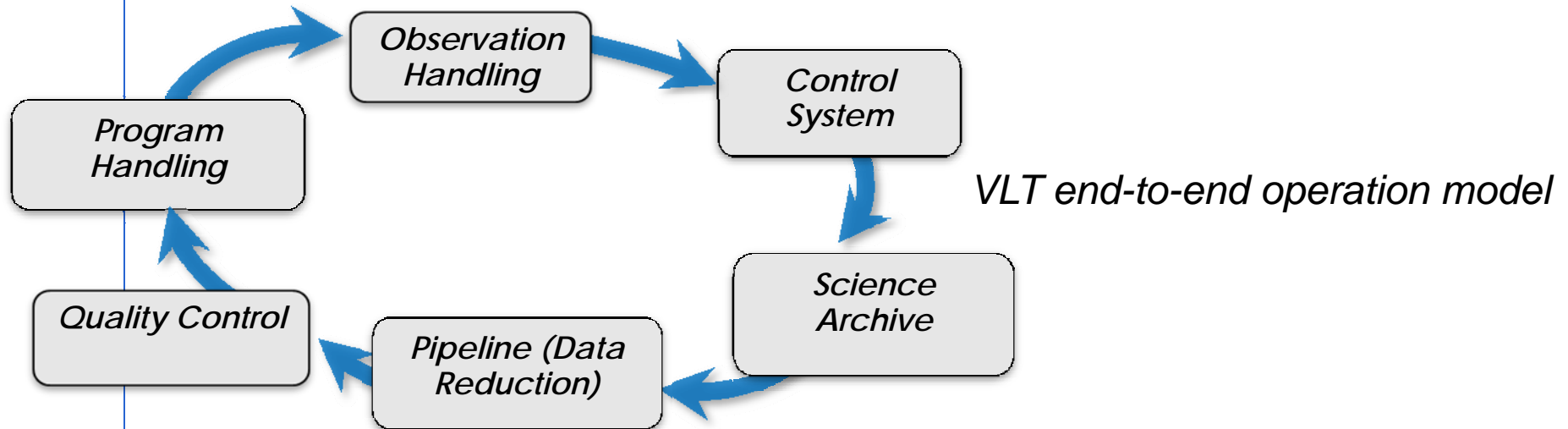
The ESO Data Flow System

Michèle Péron
Software Development Division
ESO

.... what has changed since its inception in 1995 and what has remained the same



Talk Outline



Introduction

Submission, Evaluation and Scheduling of Observing Proposals

Preparation, Scheduling and Execution of Observations

Pipeline (Data Reduction)

Data Archiving and Distribution

The Data Flow System for an ELT



The inception of the ESO Data Flow System

The development of the VLT end-to-end operation model and the requirements analysis of the software started in the fall of 1995 using the Object Modeling Technique (OMT) developed by Rumbaugh around 1991.

A first prototype of the system was verified and validated during the NTT big-bang in 1997. The first release of the system was used during the the VLT first light in 1998.

Few central observation concepts: Observation Block, Reduction Block

Few Design choices: thin interface to control software, instrument independent applications

Since that time, the DFS has evolved to accommodate:

Changes and improvements of the operation model

New instruments (data volume & complexity)

User requests for better services

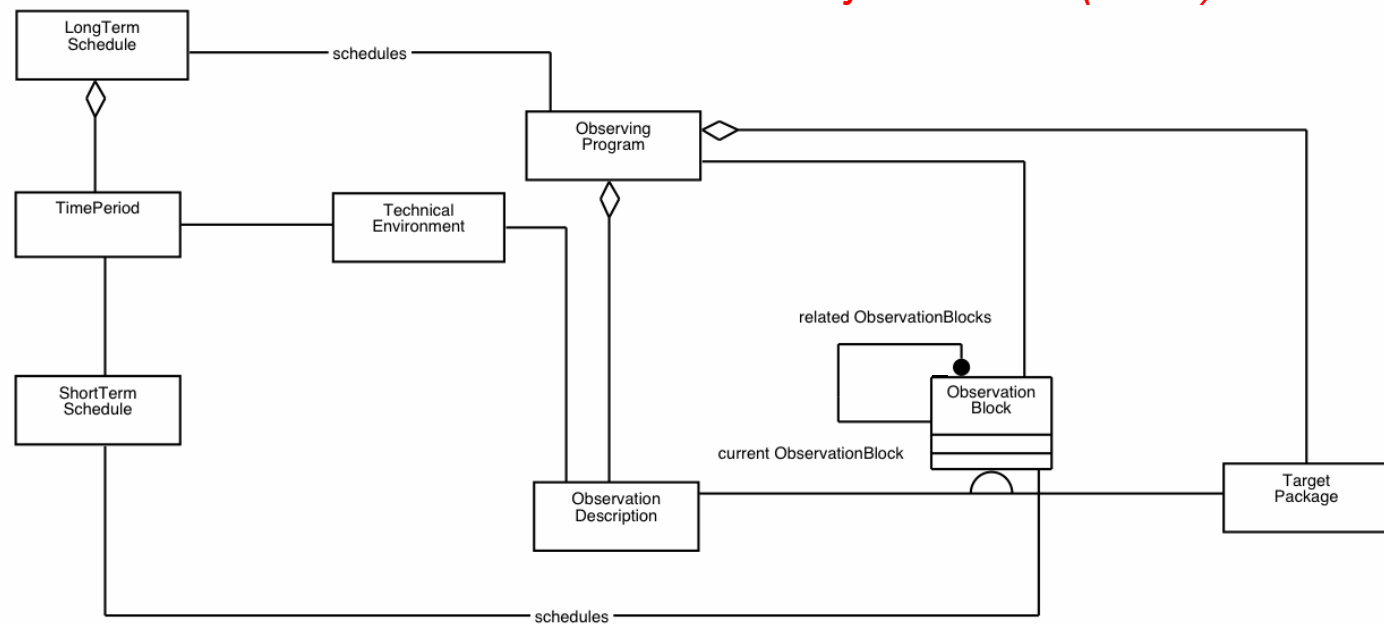
New technologies

The ESO Data Flow System

Formal model describing the system which handles the flow of science data associated with the operation of the Observatory. Focus is on conceptual rather than implementation issues.

Observation Block (OB): smallest observational unit, with a set of correlated exposures and one target.

Object Model (OMT)



The ESO Data Flow System (cont)

Observing Proposals are invited twice a year

The OPC evaluates and grades the proposals

Successful proposals are scheduled

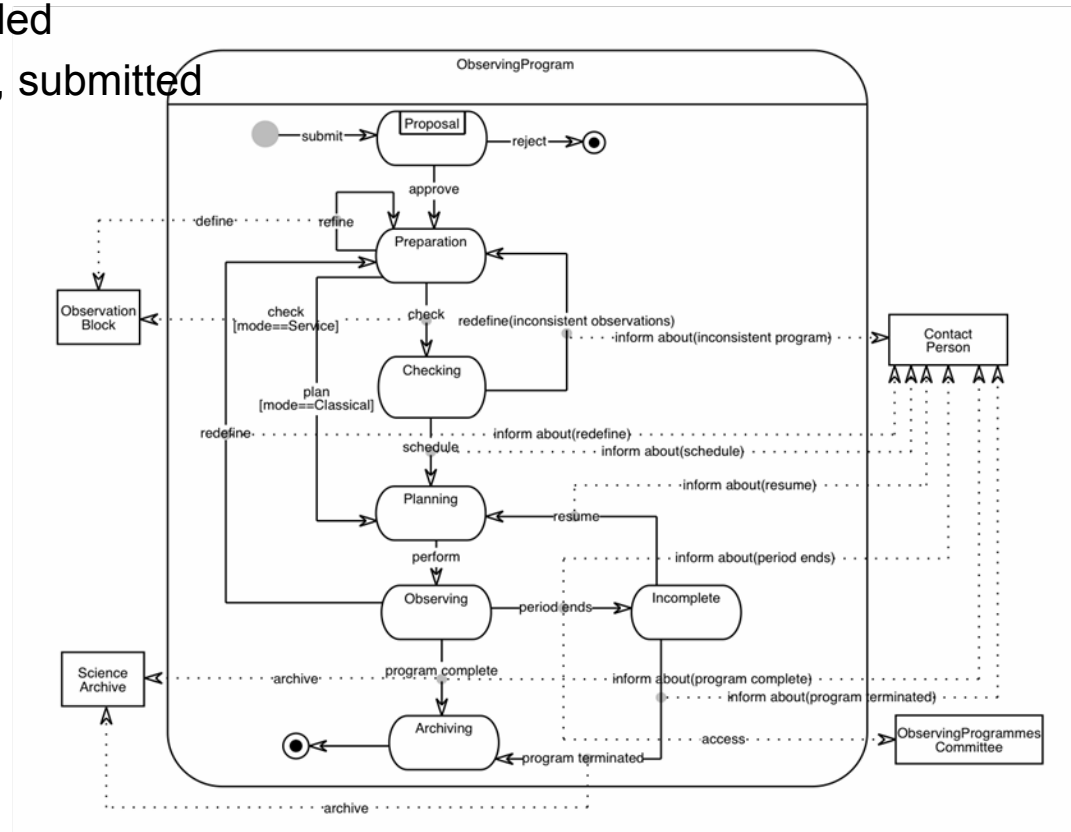
Observation Blocks are prepared, submitted to ESO and validated

Observation Blocks are executed

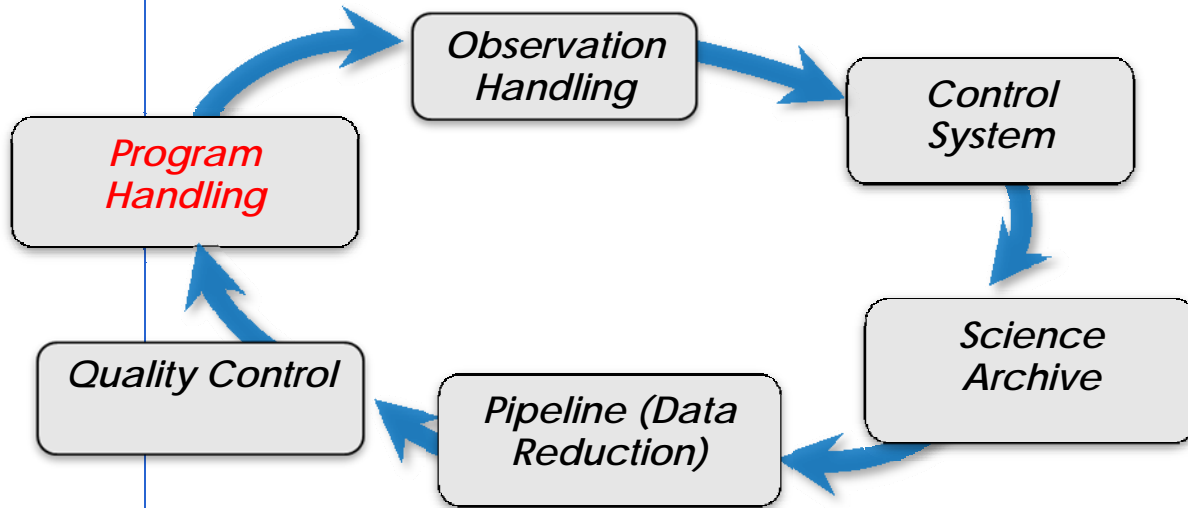
Resulting data is archived

Resulting data is processed for the purpose of quality control

Dynamic Model (OMT)



Program Handling



Introduction

Submission, Evaluation and Scheduling of Observing Proposals

Preparation, Scheduling and Execution of Observations

Pipeline (Data Reduction)

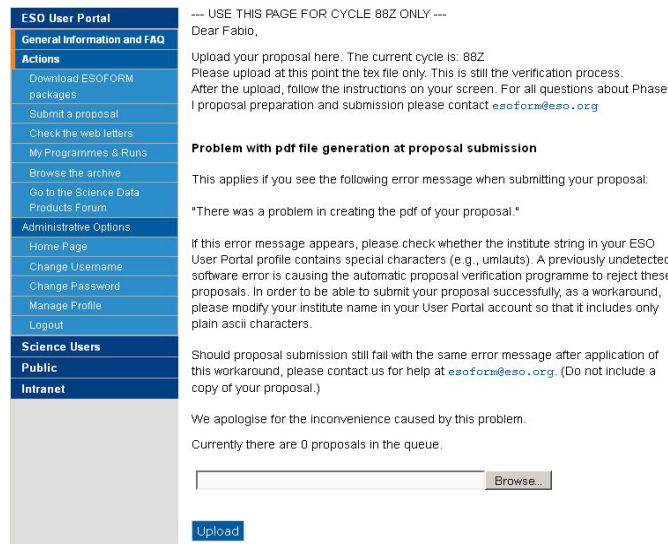
Data Archiving and Distribution

The Data Flow System for an E-ELT

Program Handling (Proposal Submission)

1998

- Latex packages are downloaded from an ftp server
- Users fill in the Latex form
- Users submit the form to ESO per email
- The system parses the LATEX form and returns errors in an email
- If all is fine, email is sent to request submission of pictures



ESO User Portal

General Information and FAQ

Actions

- Download ESOFORM packages
- Submit a proposal
- Check the web letters
- My Programmes & Runs
- Browse the archive
- Go to the Science Data Products Forum
- Administrative Options
- Home Page
- Change Username
- Change Password
- Manage Profile
- Logout

Science Users

Public

Intranet

--- USE THIS PAGE FOR CYCLE 88Z ONLY ---

Dear Fabio,

Upload your proposal here. The current cycle is: 88Z.
Please upload at this point the tex file only. This is still the verification process.
After the upload, follow the instructions on your screen. For all questions about Phase I proposal preparation and submission please contact esoform@eso.org

Problem with pdf file generation at proposal submission

This applies if you see the following error message when submitting your proposal:

"There was a problem in creating the pdf of your proposal."

If this error message appears, please check whether the institute string in your ESO User Portal profile contains special characters (e.g., umlauts). A previously undetected software error is causing the automatic proposal verification programme to reject these proposals. In order to be able to submit your proposal successfully, as a workaround, please modify your institute name in your User Portal account so that it includes only plain ascii characters.

Should proposal submission still fail with the same error message after application of this workaround, please contact us for help at esoform@eso.org. (Do not include a copy of your proposal.)

We apologise for the inconvenience caused by this problem.

Currently there are 0 proposals in the queue.

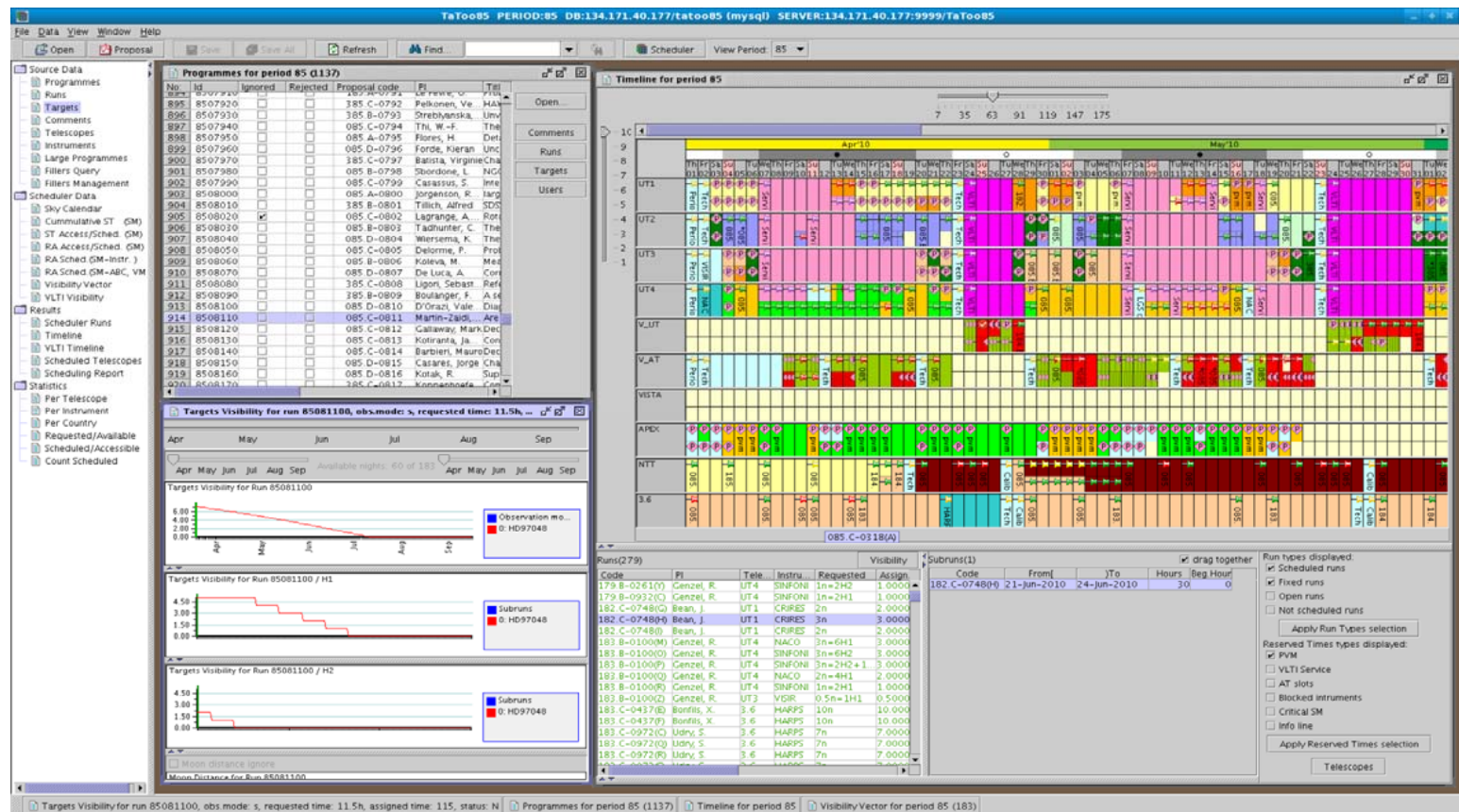
2011

- Users log-in into the User Portal and download the Latex package
- Users fill in the Latex form
- Users upload the form to ESO through a WEB interface
- Pictures can also be uploaded through the WEB.
- A PDF file is generated by the system and checked by users
- Users submit the proposal.

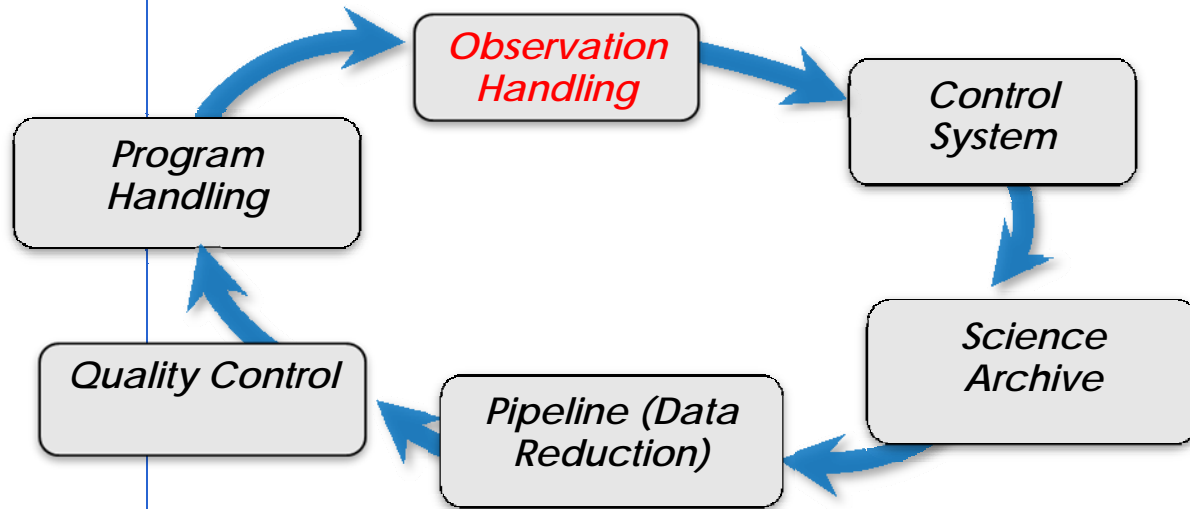
Program Handling (Long-Term Scheduling)

Scheduling of Observations for an observing period of 6 months.

GUI and a constraint programming engine taking in account the constraints of the recommended programs.



Observation Handling



Introduction

Submission, Evaluation and Scheduling of Observing Proposals

Preparation, Scheduling and Execution of Observations

Pipeline (Data Reduction)

Data Archiving and Distribution

The Data Flow System for an E-ELT

Observation Handling (OB preparation)

File Utilities Reports Help

User ID Period 60 Programme ID 60.A-0010 Instrument FORS

Observation Descriptions Observation Blocks Targets

FORs: observationDescription1 60.A-0010: observationBlock1 targetPackage1

Name observationDescription1 Name observationBlock1 Name targetPackage1

Instrument FORs Program ID 60.A-0010 Ra 01 01 03.000

User Type Observation Dec 02 02 03.000

Comments Created 1969-12-31T23:59:59 Last mod. 1969-12-31T23:59:59 Equinox J2000

Templates User Priority 1 Obs. Mode Visitor Prop.mot. Ra 0

Choose template... Detach Status Partial Calibration Req. 0

OK Verify Cancel

OK Cancel

Time Intervals

User Comments

Observer Comments

Target Acq. Templates Choose...

Constraint Set Management

OK Cancel

P2PP (1998)

Observation Handling (OB Preparation)

UVES_5325

Obs. Description Target Constraint Set Time Intervals

Obs. Description

OD Name: UVES Obs.

User Comments:

Instrument Comments:

Execution Time: 00:09:55.000

Recalculate

TemplateType: acquisition

Template:

- UVES_blue_acq_ims1
- UVES_blue_acq_slit
- UVES_blue_acq_slitrm
- UVES_dic1_acq_ims1
- UVES_dic1_acq_slit
- UVES_dic1_acq_slitrm
- UVES_dic2_acq_ims1

Add Duplicate Delete

UVES_blue_acq_ims1	1	UVES_blue_obs_exp	1
RA blind offset	0.0	Blue Readout Mode	225kHz, 1x1, low
DEC blind offset	0.0	Blue Exposure Time	10
Get Guide Star from	CATALOGUE	No. of Exp.	1
Guide star RA	0	Source Type	POINT
Guide star DEC	0	Number of offsets	1
Image Slicer	SLIC#2	X offset in arcsec.	0
Acq. Pre-Slit Filter	FREE	Y offset in arcsec.	0
Depolarizer	OFF	Blue Mode Central Wlgt	346
		Blue Slit Width	0.6

P2PP (2011)



Observation Handling (OB Preparation)

Survey Telescopes (i.e., VISTA, VST) brought in new ways of observing

One program might span several years and including hundreds of OBs.

Scheduling containers allow astronomers to express more complex strategies by creating additional abstraction on top of individual OBs that allow expressing dependencies between them.

The screenshot shows the P2PP 3.2rc2 software interface. The main window displays a tree view of observation containers and blocks. The 'Observing Runs' tab is active, showing a list of containers and their associated blocks. The table below summarizes the data shown in the interface.

Name	Priority	Contrib. to Group	Abs. Time Intervals	Earliest After Prev.	Latest After Prev.
60.A-9253(N)/SM/VIRCAM					
Group Container	1				
Group OB1	✓	20	0		
Group OB2	✓	12	0		
Group OB3	✓	11	0		
Group OB4	✓	40	0		
Group OB5	✓	10	0		
Time Link Container	1				
TimeLink OB1	✓		0		
TimeLink OB2	✓			001d 00:00	003d 00:00
TimeLink OB3	✓			001d 00:00	006d 00:00
TimeLink OB4	✓			005d 00:00	007d 00:00
TimeLink OB5	✓			003d 00:00	007d 00:00
Concatenation Container	1				
Concatenation OB1	✓		0		
Concatenation OB2	✓		0		
Concatenation OB3	✓		0		
Concatenation OB4	✓		0		
Concatenation OB5	✓		0		
60.A-9253(P)/SM/XSHOOTER					
60.A-9252(B)/SM/SUSI2					

The status bar at the bottom indicates 'p2pp server is reachable'.



Observation Handling (OB Execution)

Large number of OBs of short duration, with execution dependencies expressed in scheduling containers.

- Ranking engine suggests the next OB to be executed, taking in account weather condition, visibility constraints, user priority, group score as well as the observing run rank.

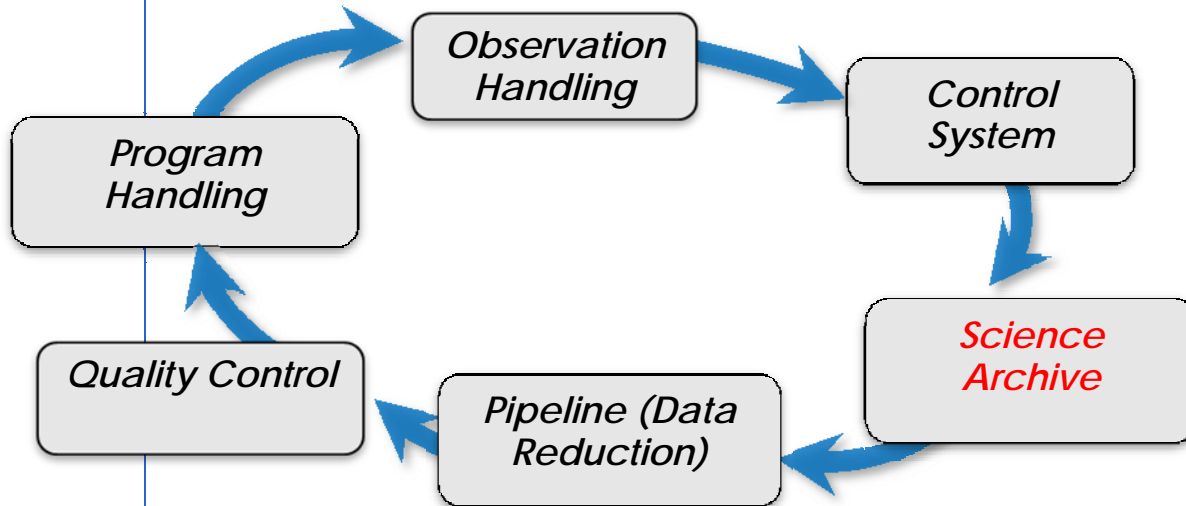
The screenshot displays the ORANG DB server interface (version 6996) with the following components:

- Left Panel:** Configuration settings for observation scheduling.
 - Work:** WORK 2.2m, Work UT1+UT3, Refresh Queue-Lists button.
 - UT Start-Time:** 2010-02-03T07:55:08, UT Start-Time to Current-Time button.
 - Duration in Hours:** All Night, Maximum Orang Rows: 200.
 - Weather-Conditions:** Seeing 600nm Max. (2), Curr. (2), Wind-Direction (90), Sky-Transparency (Photometric).
 - Visibility-Constraints:** Air-Mass(%Tolerance) (30), Sun Max. Elevation (Deg.) (-18), Moon-Elevation (Deg.) (0), FLU(%Tolerance) (0), Moon Distance(%Tolerance) (0), Zenith Avoidance (checked), Exec at UT Start-Time Only (unchecked).
 - Ranking-Algorithm:** Algorithm: VISTA Survey v0.03, Rank button, Re-Load and Rank button.
- Right Panel:** Observation list and ranking details.
 - Observable OB (8):** Table with columns: Rank score, OB ID, Status, Container, Container ID, Sched Id, and group name.
 - Filtered rows: 8**
 - Rank Justification for 729936:** Table showing contribution and execution details for OB 729936.

Rank score	OB ID	Status	Container	Container ID	Sched Id	Group
1	729936	+	G	729929	25048710	grp_0
2	729933	+	G	729929	25048710	grp_0
3	729930	+	G	729929	25048710	grp_0
4	735653	+	G	735652	25048710	con_0
4	735656	+	G	735652	25048710	con_0
4	735659	+	G	735652	25048710	con_0
4	735662	+	G	735652	25048710	con_0
4	734945	+	T	734944	25052161	tlk_er

OB ID	Status	Contrib.	Execution	Started time
729930	+	10.00 %		
729933	+	20.00 %		
729936	+	30.00 %	00:04:28.000	
729939	X	40.00 %	00:04:28.000	2010-01-16T01:2...

Science Archive



Introduction

Submission, Evaluation and Scheduling of Observing Proposals

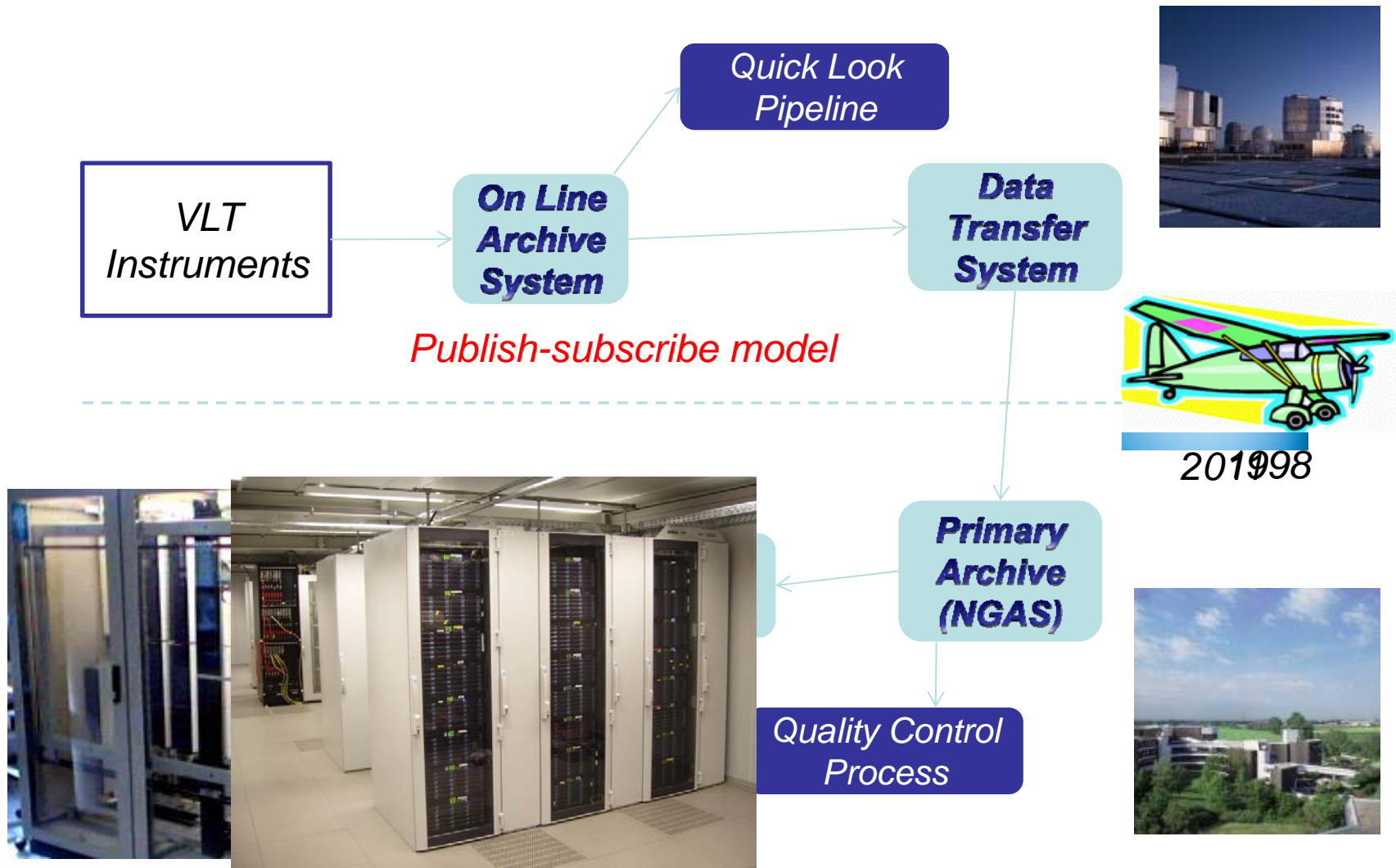
Preparation, Scheduling and Execution of Observations

Pipeline (Data Reduction)

Data Archiving and Distribution

The Data Flow System for an E-ELT

Data Flow Back End






Data Transfer and Distribution

Since middle of 2008 all VLT/VLTI data are transferred to ESO Garching through the network

Highly optimized utilization of high-latency network

File Transfer can be flexibly prioritized

This new system has enabled “more” Quality Control to take place in Garching



Data Transfer Monitor

Data Flow Statistics ♦ Network Transfer Monitor ♦ Network Bandwidth Usage ♦ Network Transmission Time Stats ♦ File Search ♦ Main Page

Network File Transfer

Showing first 3026 of 3026 files with total size 22,565.4 MB

Last query: 2009-03-05 22:04:02 UTC

Define query

Creation date	<input type="text" value="2009-03-04"/>	Date is defined to start at 12:00 UTC
fileID pattern	<input type="text" value="None"/>	Enter one or more matching strings, separated
transfer status	<input type="text" value="all"/>	
file type	<input type="text" value="fits"/>	
# of files to be shown	<input type="text" value="all"/>	

Submit

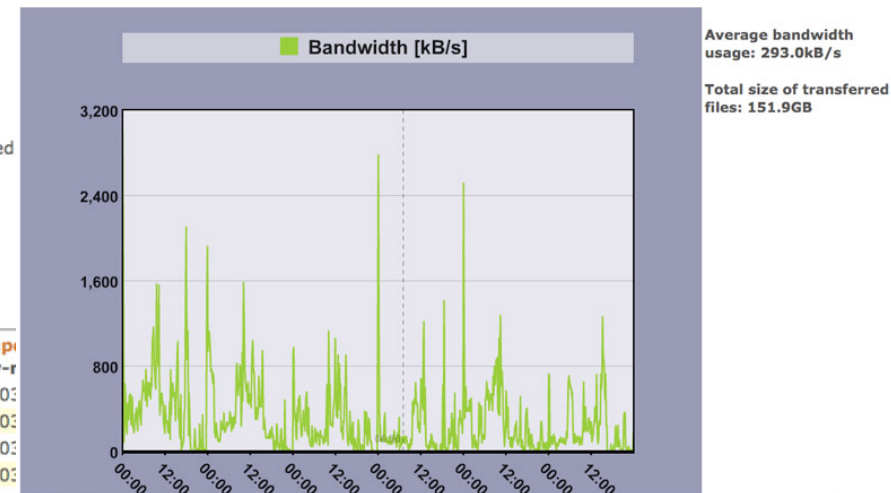
file id	status	size [MB]	priority	waiting Time [hh:mm:ss]	last up [yyyy-mm-dd]
VIMOS.2009-03-05T02:07:49.834	✓	8.56	20	05:01:50	2009-03-05
VIMOS.2009-03-05T02:07:49.835	✓	8.51	20	05:01:51	2009-03-05
VIMOS.2009-03-05T02:07:49.854	✓	8.79	20	05:01:49	2009-03-05
VIMOS.2009-03-05T02:07:49.855	✓	8.91	20	05:01:49	2009-03-05



Bandwidth Usage from 2009-02-01 00:00 UTC to 2009-02-07 00:00 UTC

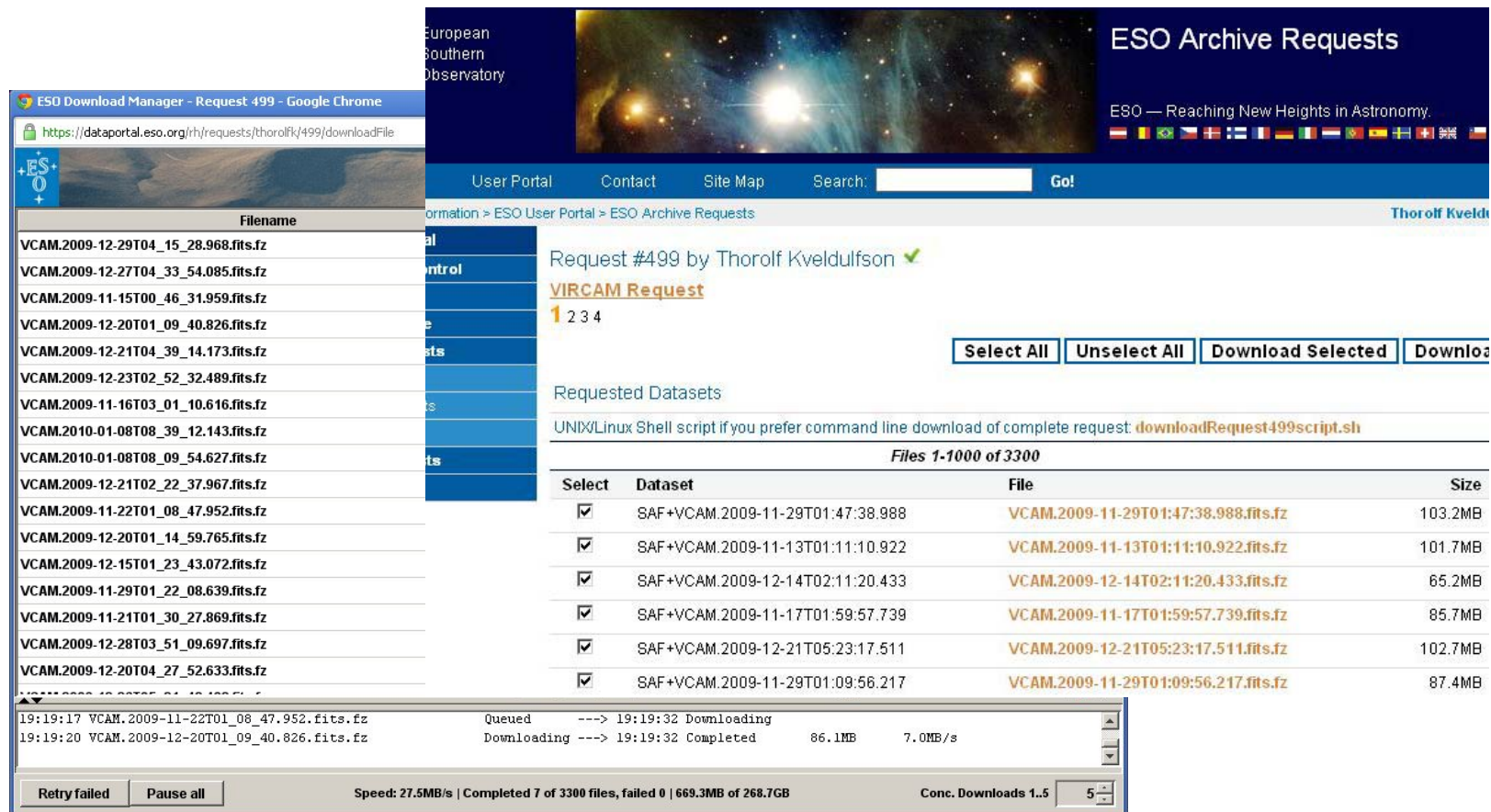
Show data from: (valid format: yyyy-mm-dd hh:mm)
to:

Submit



Data Distribution (request Handler)

Code re-use from CADC/ALMA

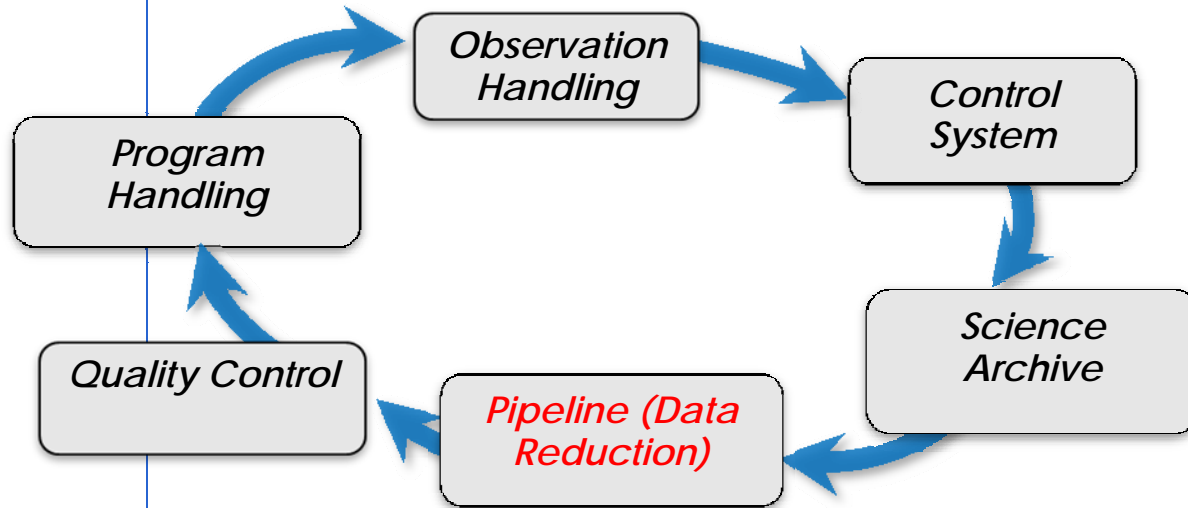


The screenshot displays the ESO Archive Requests web interface. The top navigation bar includes links for User Portal, Contact, Site Map, and a search bar. The main content area shows 'Request #499 by Thorolf Kveldulfson' with a 'VIRCAM Request' status. Below this, there are buttons for 'Select All', 'Unselect All', 'Download Selected', and 'Download'. A table lists 'Requested Datasets' with columns for 'Select', 'Dataset', 'File', and 'Size'. The table shows several files selected for download. At the bottom, a progress bar indicates the download status, showing a speed of 27.5MB/s and 7 of 3300 files completed.

Select	Dataset	File	Size
<input checked="" type="checkbox"/>	SAF+VCAM.2009-11-29T01:47:38.988	VCAM.2009-11-29T01:47:38.988.fits.gz	103.2MB
<input checked="" type="checkbox"/>	SAF+VCAM.2009-11-13T01:11:10.922	VCAM.2009-11-13T01:11:10.922.fits.gz	101.7MB
<input checked="" type="checkbox"/>	SAF+VCAM.2009-12-14T02:11:20.433	VCAM.2009-12-14T02:11:20.433.fits.gz	65.2MB
<input checked="" type="checkbox"/>	SAF+VCAM.2009-11-17T01:59:57.739	VCAM.2009-11-17T01:59:57.739.fits.gz	85.7MB
<input checked="" type="checkbox"/>	SAF+VCAM.2009-12-21T05:23:17.511	VCAM.2009-12-21T05:23:17.511.fits.gz	102.7MB
<input checked="" type="checkbox"/>	SAF+VCAM.2009-11-29T01:09:56.217	VCAM.2009-11-29T01:09:56.217.fits.gz	87.4MB

Nathalie Fourniol: News about ESO Archive services
Ignacio Vera: hFits: from storing metadata to publishing ESO data

Pipeline (Data Reduction)



Introduction

Submission, Evaluation and Scheduling of Observing Proposals

Preparation, Scheduling and Execution of Observations

Pipeline (Data Reduction)

Data Archiving and Distribution

The Data Flow System for an E-ELT

Data Reduction at the Telescope

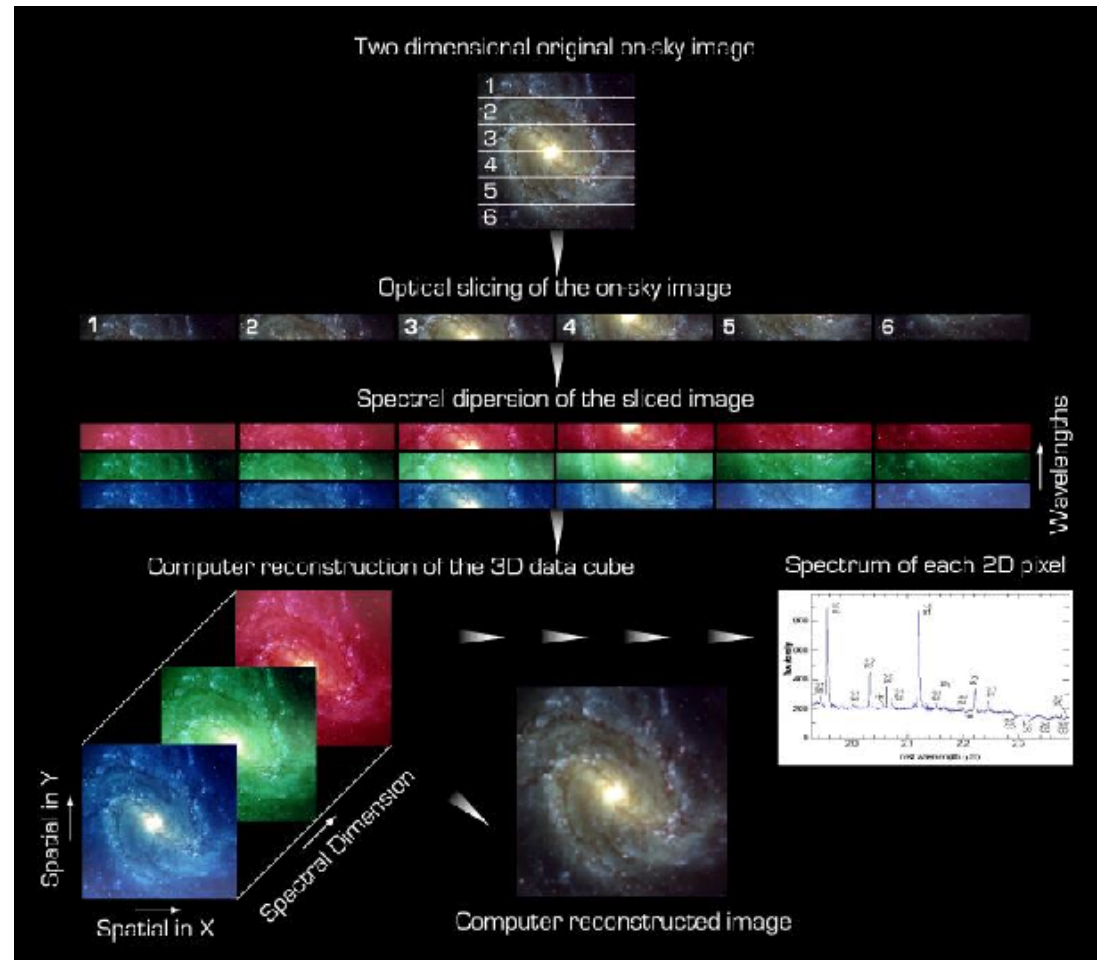
Required to control the health of the instruments and check the quality of the observations

Must be done
automatically and
in quasi real-time

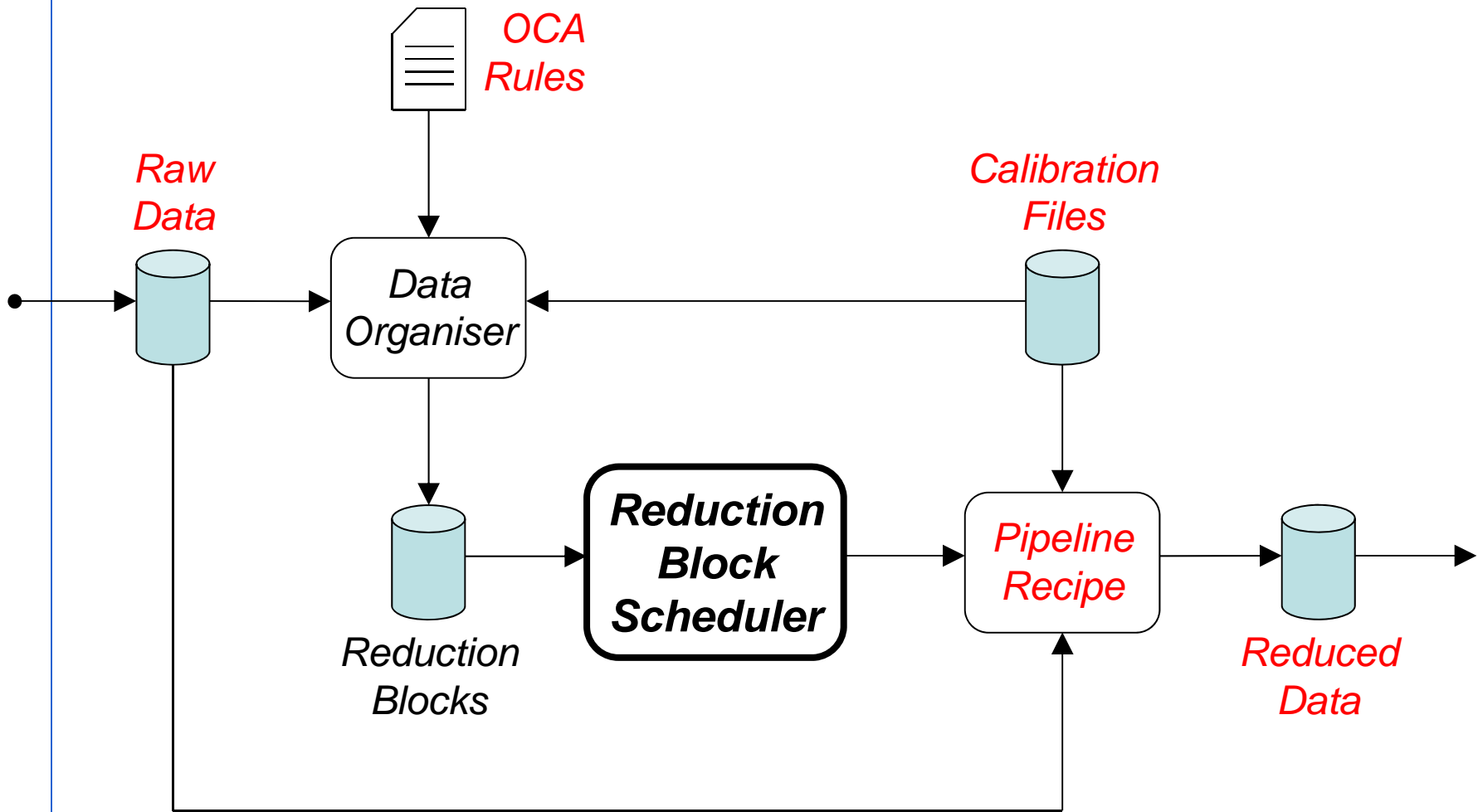
Large amount of data
(few hundreds of
Gigabytes per night)

Multi-core hardware and parallel processing

Complex instruments and complex reduction algorithms



Pipeline Infrastructure

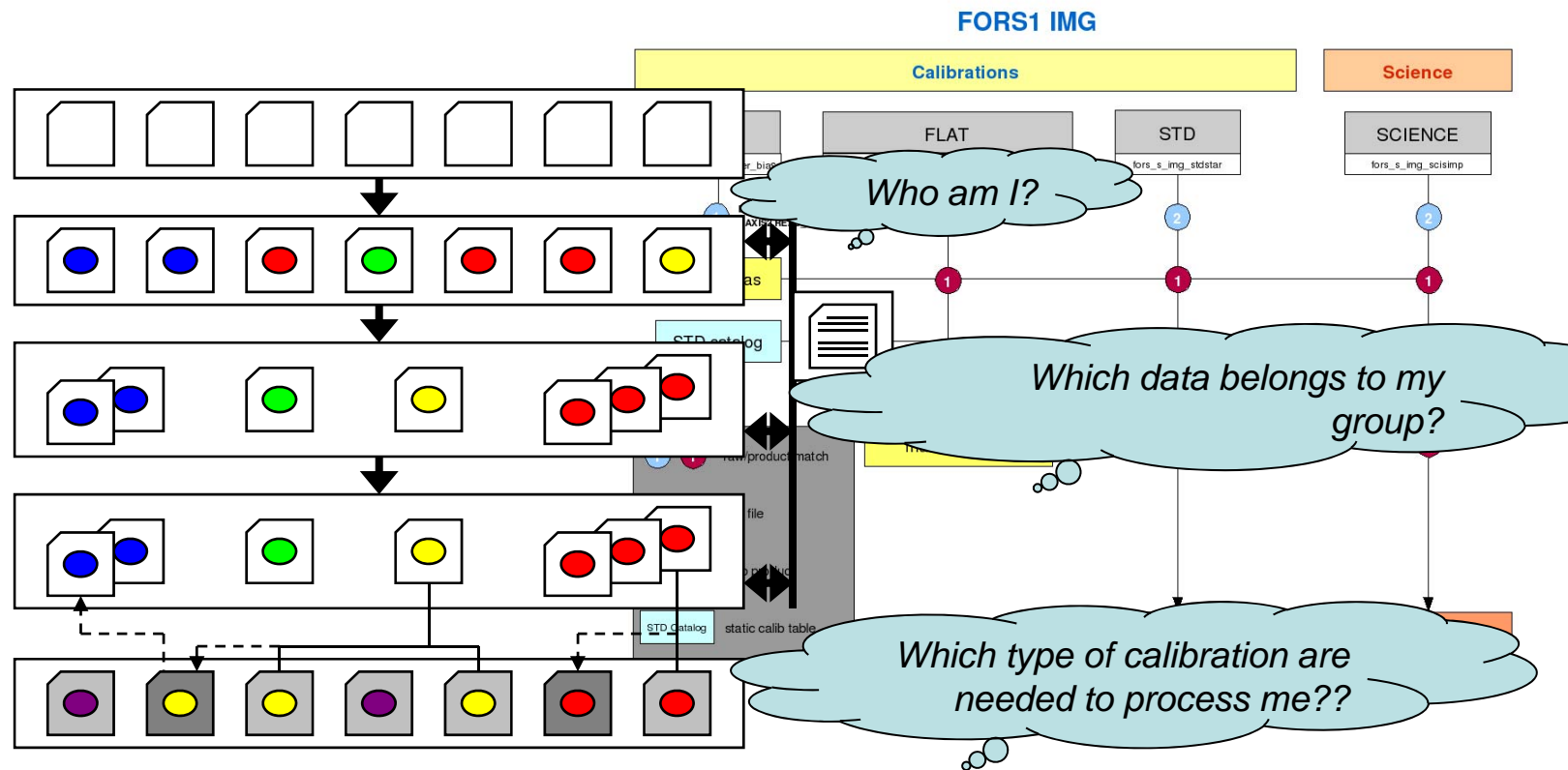


Data Reduction at the Telescope (cont)

Automatic Data Organization (available in 1998 in C++, re-engineered in Java in 2005)

Based on a flexible rule-engine & a domain-specific language

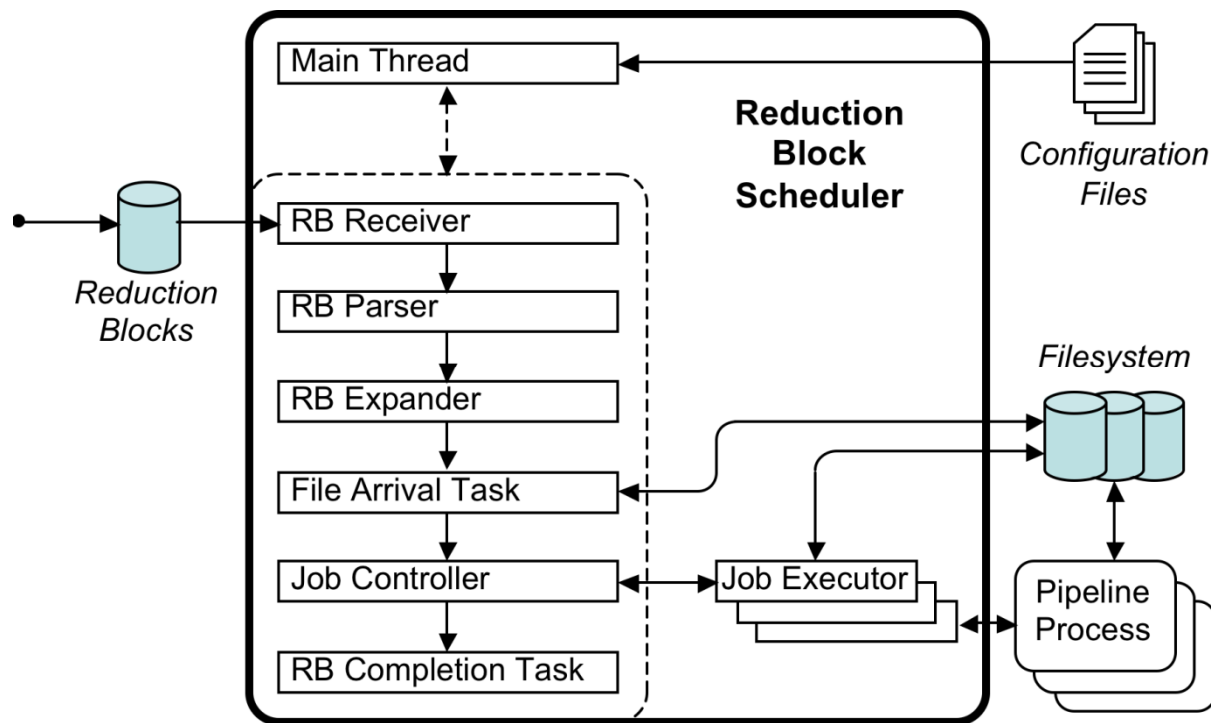
Creates Reduction Blocks (contains all information for reducing a set of related data)



Data Reduction at the Telescope (cont)

Reduction Block Scheduler (available in 1998 in C++, re-engineered in Java in 2007)

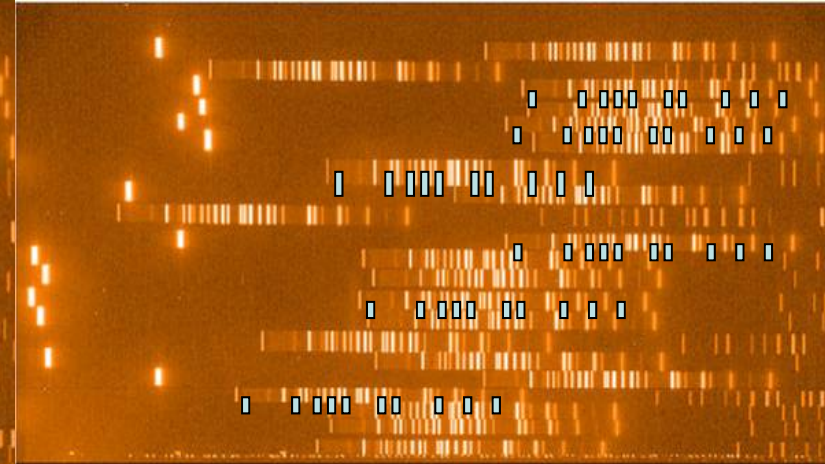
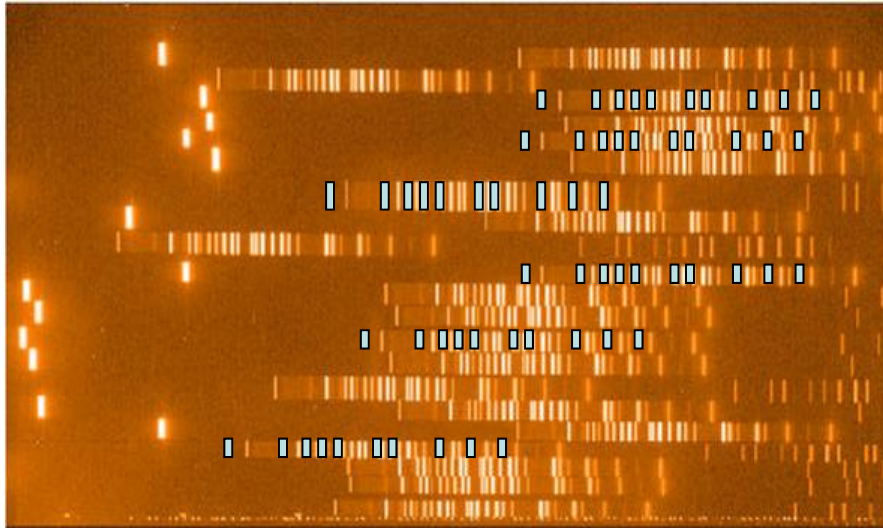
- Multi-threaded application
- Takes in account dependencies between Reduction Blocks



Pipeline Algorithms – New approaches

Wavelength calibration of a MOS exposure using first guess model to find reference lines

EARTHQUAKE!



New approaches (such as pattern-matching) are needed



In Memoriam Carlo Izzo □

Data Reduction at Home (Reflex)

Components | Data

Search

☐ Search repository

Search Reset

Components

Projects

Disciplines

Statistics

0 results found.

UVES Workflow For Point Source Echelle Data (v. 4.9.0)

Workflow Instructions

To run this workflow on the demo data:

- Turn on highlighting. Choose "Tools"-> "Animate at Runtime" from top menu and set it to "1".
- Press the "Run" button OR cntrl-R to start the workflow.

To run on a different data set:

- Click on ROOT_DATA_DIR and set as appropriate. All subdirectories of RAWDATA_DIR will be searched for data.
- If desired, change END_PRODUCTS_DIR.
- Press the "Run" button OR cntrl-R to start the workflow.

To monitor the progress of the workflow in more detail:

- Open "Window" -> "Runtime Window" in top menu before starting the workflow.

Input:

- ROOT_DATA_DIR: /data/reflex_data
- RAWDATA_DIR: \$ROOT_DATA_DIR/reflex_input/UVes

Working Directories:

- BOOKKEEPING_DIR: \$ROOT_DATA_DIR/reflex_book_keeping/UVes
- LOGS_DIR: \$ROOT_DATA_DIR/reflex_logs/UVes
- TMP_PRODUCTS_DIR: \$ROOT_DATA_DIR/reflex_tmp_products/UVes
- END_PRODUCTS_DIR: \$ROOT_DATA_DIR/reflex_end_products

Output:

If END_PRODUCTS_DIR or ROOT_DATA_DIR is changed using the Browse button, the leading file: has to be removed manually!

Global Parameters

= actor with interactive option

- FITS_VIEWER: fv
- ESORexArgs: --suppress-prefix=TRUE
- EraseDirs: false
- END_PRODUCT_SUBDIR: 2010-12-14T11:12:20/Science_DataSet_2
- GLOBAL_TIMESTAMP: 2010-12-14T11:12:20

fits viewer to use for the inspection of input/output products

esorex arguments

Change "EraseDirs" to true to erase BOOKKEEPING_DIR, TMP_PRODUCTS_DIR and LOGS_DIR each time the workflow is run (Lazy Mode won't work anymore)

This is set automatically

This is set automatically

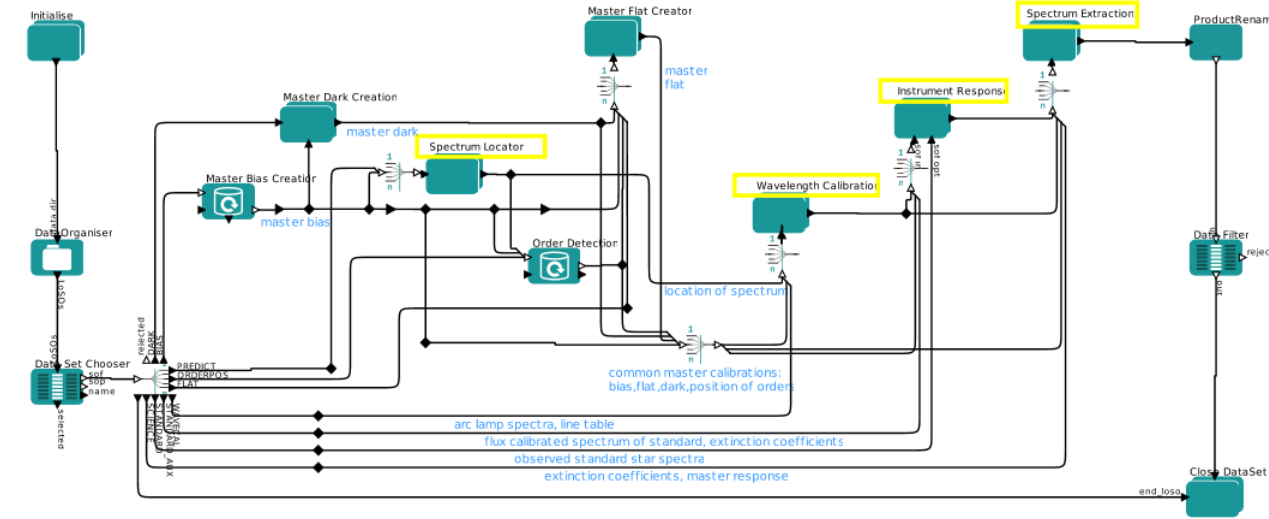
Step 1:
Data Organisation
and Selection

Step 2:
Creation of Master
Calibration Files

Step 3:
Wavelength and Response
Calibration

Step 4:
Spectrum
Extraction

Step 5:
Output
Organisation



Initialise

Data Organiser

Data Set Chooser

Master Dark Creation

Master Bias Creator

Spectrum Locator

Master Flat Creator

Order Detection

Wavelength Calibration

Instrument Response

Spectrum Extraction

Product Renamer

Data Filter

Close DataSet

rejected

end_logs

common master calibrations:
bias, flat, dark, position of order

arc lamp spectra, line table


flux calibrated spectrum of standard, extinction coefficients

observed standard star spectra

extinction coefficients, master response

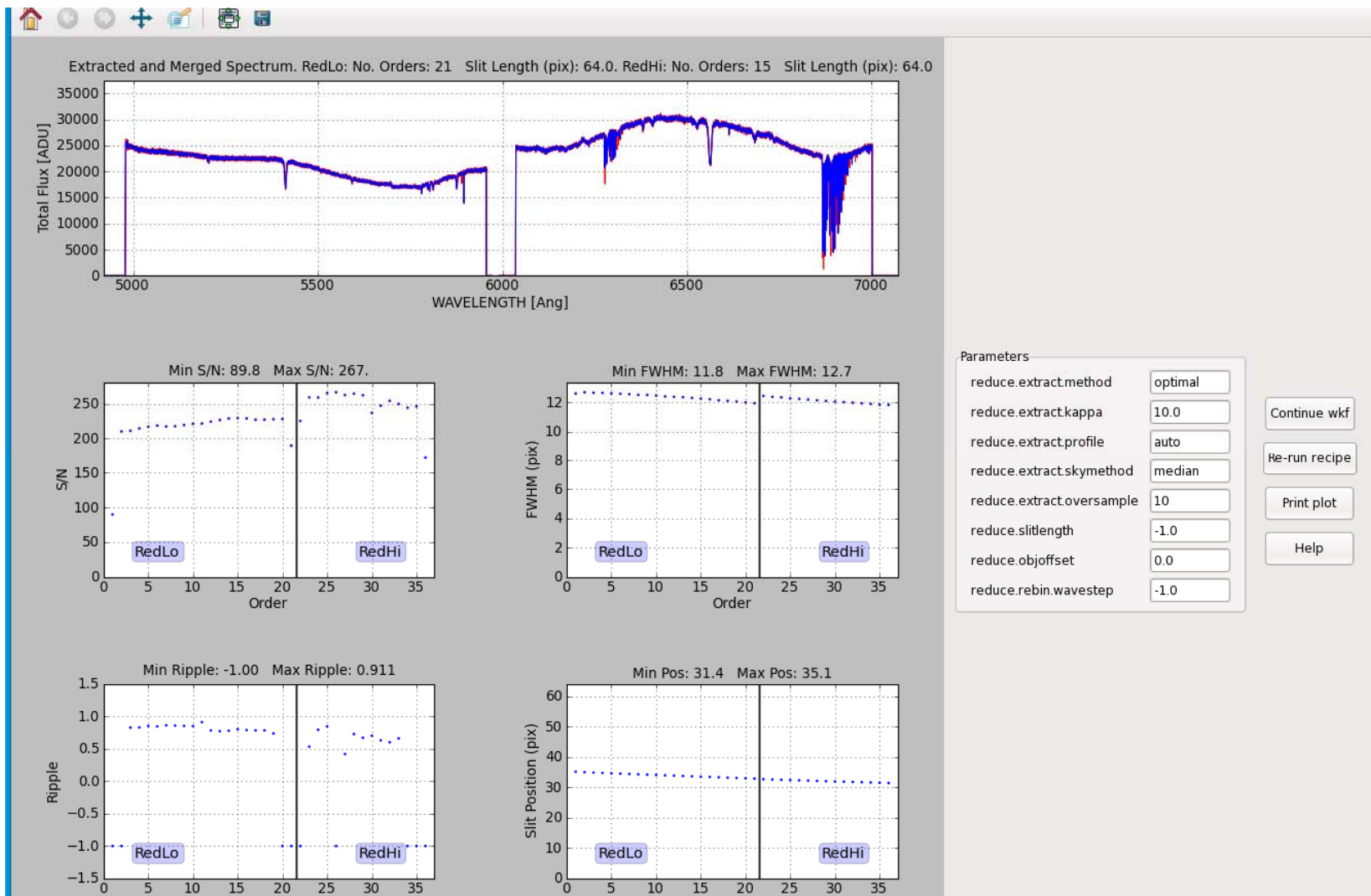
ADASS 2011 – Michèle Peron

25



Data Reduction at home (Reflex)

Demo: Ballester et al.



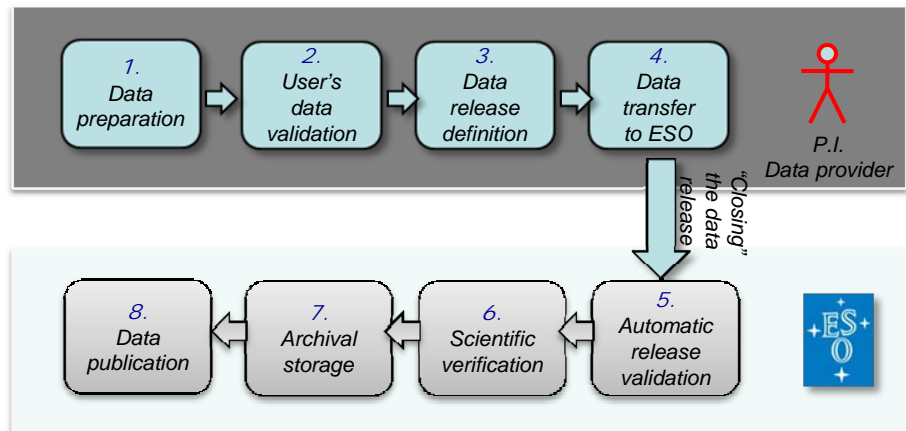
Phase 3: Handling Survey Data Products

J. Retzlaff, M. Arnaboldi, V. Forchí, P. Nunes, S. Zampieri,
T. Bierwirth, M. Peron, M. Romaniello, J. Lockhart, D.
Suchar (ESO)

Phase 3 denotes the process in which principal investigators of ESO observing programmes return their reduced data products to ESO for storage in the ESO archive and subsequent data publication to the scientific community.

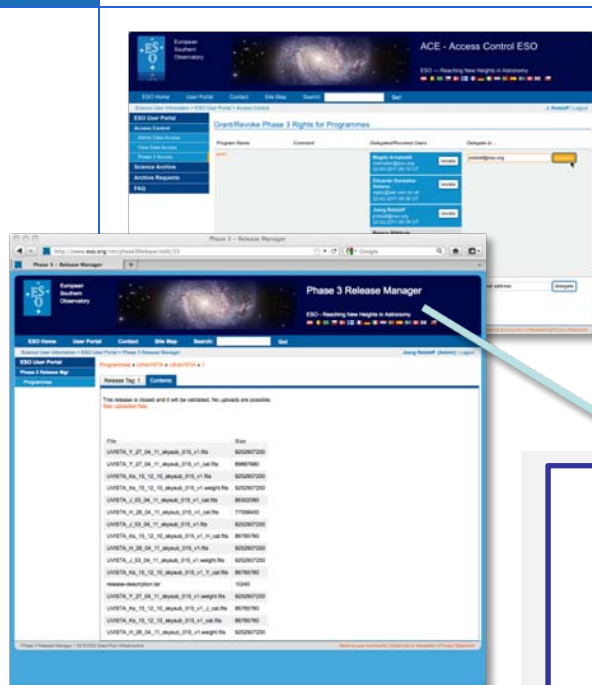
The new Phase 3 infrastructure supports the reception, validation and publication of data products from the public survey projects and large programmes to the ESO Science Archive Facility.

Phase 3 Process and Responsibilities



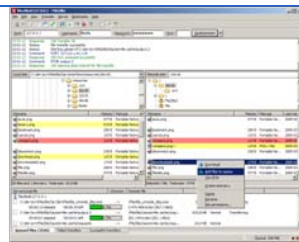
<http://www.eso.org/sci/observing/phase3.html>

Phase 3 data flow & infrastructure



The release manager is a web application that allows the P.I. to define data collections and releases and to manage the Phase 3 delegation to co-investigators.

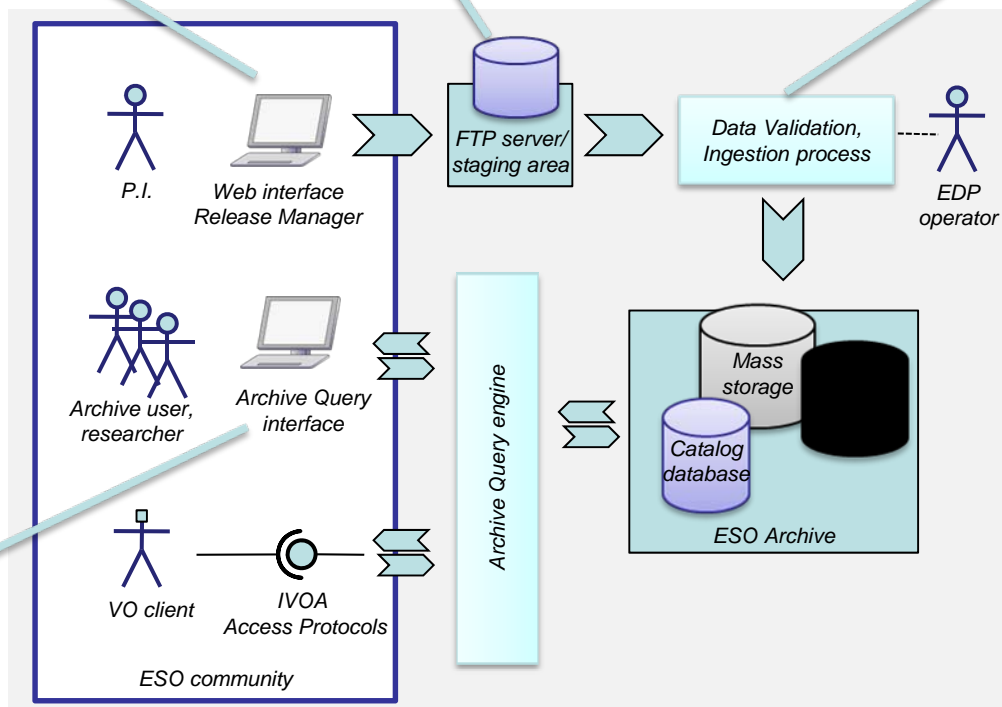
<http://archive.eso.org/wdb/wdb/adp/phase3/main/form>



The release validator is a command-line application that helps to verify the data standard and validity of the header keywords against predefined rules.



The data is transferred by the PI/Co-I via FTP to the dedicated staging area.



Interfaces between the Phase 3 data flow, its users and the ESO Science Archive Facility.

Start of operations: 10 March 2011

Data Flow System for the ELT

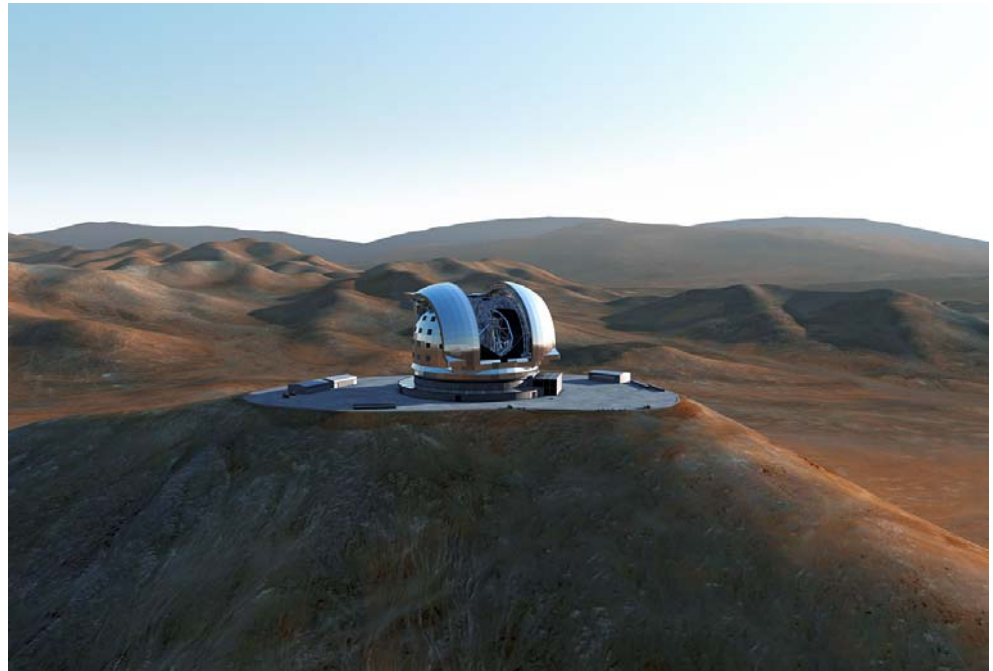
End-To-End Operational Model similar to VLT.

Instruments will have the same level of complexity as the VLT ones.

There will not be a DFS-VLT and a DFS-ELT but a DFS

Will be an evolution of the current system or a revolution?

Underlying Data Model has grown organically and might require complete re-engineering



THANK YOU

... to all those who have contributed to the DFS over the past years:

Software Development Division



Data Management and Operations Division

The Observing Programme Office

Science Operations Department of the Observatory

