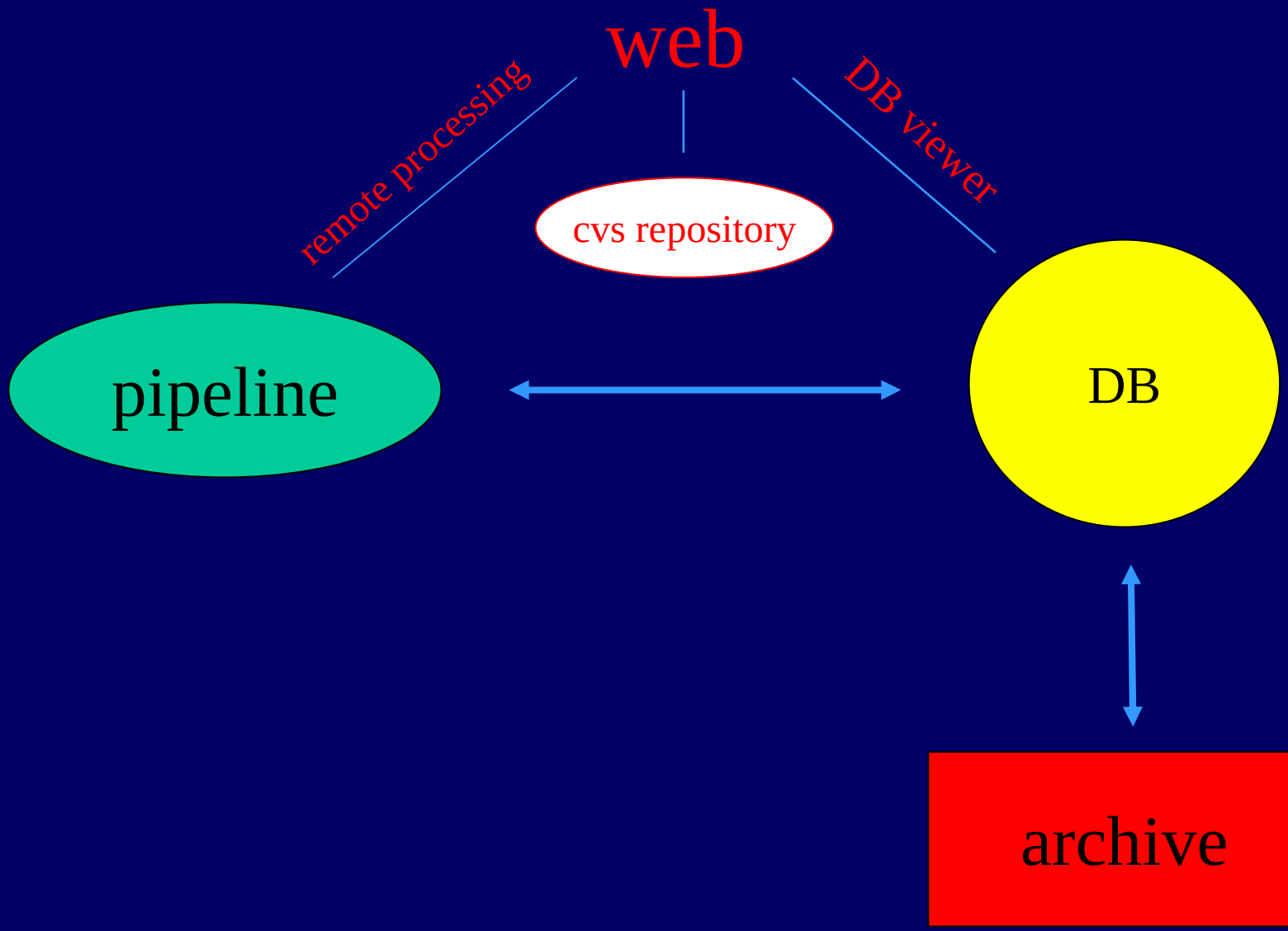
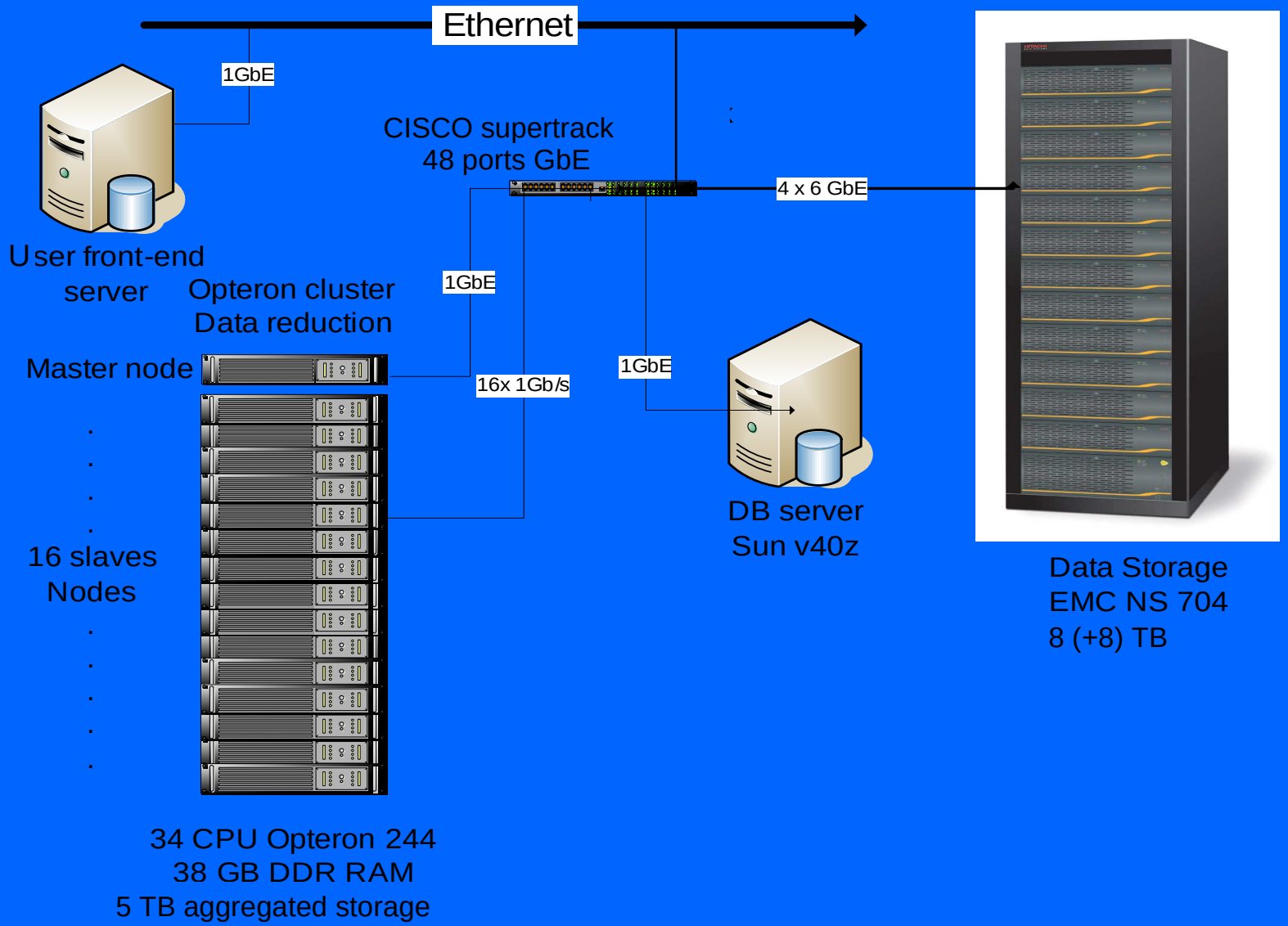
First ASTRO-WISE Course
INAF Napoli, 23-26 June 2008

The AW environment (AWe)



Hardware


Ethernet
1GbE
User front-end server
CISCO supertrack
48 ports GbE
4 x 6 GbE
1GbE
Opteron cluster
Data reduction
Master node
16x 1Gb/s
1GbE
16 slaves
Nodes
DB server
Sun v40z
Data Storage
EMC NS 704
8 (+8) TB
34 CPU Opteron 244
38 GB DDR RAM
5 TB aggregated storage

(16+1) Opteron beowulf

17 TYAN Transport TX28 (B2880) server systems based on TYAN Thunder K8S (S2880) platform featuring two 64-bit AMD Opteron processors 244 + 128-bit DDR dual channel memory controller.

The system has five expansion slots of which 2 independent PCI-X buses, can support up to 12 GB of Registered DDR2700 memory and has two integrated Gigabit Ethernet controller connected to the PCI-X bus. Total RAM is 38 GB and total aggregate disk space is 5TB. The connectivity inside the cluster is provided by a 24 copper ports Gigabit switch. This will be upgraded to a 48 GbE switch to allow the connectivity to the data storage.

The system runs Scientific Linux created at CERN which is based on Red Hat and OSCAR.

SUN V40z equipped with two Opteron CPUs, 4 GB of DDR RAM and 500 GB of U320SCSI 10000 rpm disks.

- Commercial solution in order to guarantee the necessary reliability, availability, performance and maintenance.
- The system is a NS 704 from EMC, equipped with four NAS heads, two storage controllers populated with 12.5 TB (raw) of fibre channel disks.
- Storage capacity can be scaled up to 130 TB with a very high performance in data delivery (the storage is connected to the Beowulf cluster through 24 Gigabit Ethernet connections divided in four trunked channels).
- An upgrade (from 8 to 16 TB of effective capacity) is foreseen later this year.

AW Architectural Design Document (ADD)

Requirements/concepts:

- Pipeline performance: at least 1 Mpix/sec ⇒ parallel processing
- DB driven operations
- Multi-site operations (processing and administration) → federated DB (Oracle 10g)
- Extendibility
- On-the-fly reprocessing
- Multi-instrument support
- VO/Grid compatibility

AW development framework:

- Modular programming
- Object Oriented programming (→ python): > awe
- CVS: central repository of the code (+ documentation) implemented at the different nodes

AW pipeline

Data reduction is normally performed step by step through python command lines, using either:

- the task.execute() command [single CCDs]
- the dpu.run method [distributed processing, 1 CCD = 1 CPU]

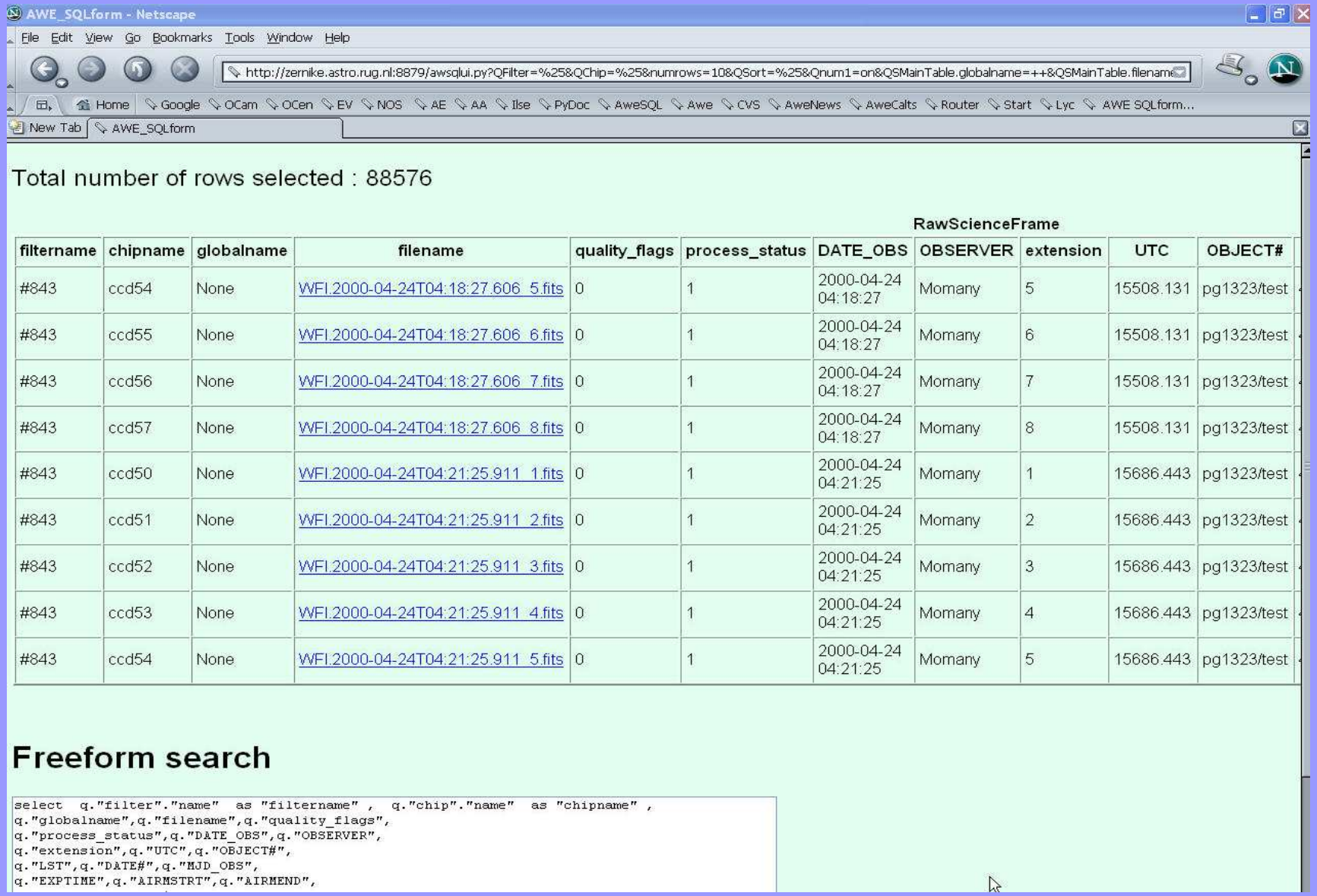
The Pars method allows to change parameters without entering the python (or C/C++) code

When everything is well optimized it is possible to launch more complex automatic python scripts that do several steps at once (even via web interface)

AW DB (Oracle 11g)

- The DB contains metadata (links to the fits files stored in the data storage) or other objects (e.g. catalogs).
- All data can be easily retrieved, reproduced and published.
- The DB (and the DS) is “federated”: the same DB is seen by each AW node and is continuously synchronized. This allows to share files between users at different nodes working in the same project.
- The DB is divided in different “contexts” corresponding to the different survey projects.

The content of the DB can be seen through an SQL query or through a web interface (DB viewer)



Total number of rows selected : 88576

							RawScienceFrame			
filtername	chipname	globalname	filename	quality_flags	process_status	DATE_OBS	OBSERVER	extension	UTC	OBJECT#
#843	ccd54	None	WFI.2000-04-24T04:18:27.606_5.fits	0	1	2000-04-24 04:18:27	Momany	5	15508.131	pg1323/test
#843	ccd55	None	WFI.2000-04-24T04:18:27.606_6.fits	0	1	2000-04-24 04:18:27	Momany	6	15508.131	pg1323/test
#843	ccd56	None	WFI.2000-04-24T04:18:27.606_7.fits	0	1	2000-04-24 04:18:27	Momany	7	15508.131	pg1323/test
#843	ccd57	None	WFI.2000-04-24T04:18:27.606_8.fits	0	1	2000-04-24 04:18:27	Momany	8	15508.131	pg1323/test
#843	ccd50	None	WFI.2000-04-24T04:21:25.911_1.fits	0	1	2000-04-24 04:21:25	Momany	1	15686.443	pg1323/test
#843	ccd51	None	WFI.2000-04-24T04:21:25.911_2.fits	0	1	2000-04-24 04:21:25	Momany	2	15686.443	pg1323/test
#843	ccd52	None	WFI.2000-04-24T04:21:25.911_3.fits	0	1	2000-04-24 04:21:25	Momany	3	15686.443	pg1323/test
#843	ccd53	None	WFI.2000-04-24T04:21:25.911_4.fits	0	1	2000-04-24 04:21:25	Momany	4	15686.443	pg1323/test
#843	ccd54	None	WFI.2000-04-24T04:21:25.911_5.fits	0	1	2000-04-24 04:21:25	Momany	5	15686.443	pg1323/test

Freeform search

```
select  q."filter"."name"  as "filtername" ,  q."chip"."name"  as "chipname" ,
q."globalname",q."filename",q."quality_flags",
q."process_status",q."DATE_OBS",q."OBSERVER",
q."extension",q."UTC",q."OBJECT#",
q."LST",q."DATE#",q."MJD_OBS",
q."EXPTIME",q."AIRMSTRT",q."AIRMEND",
```

What will (likely) not be seen in this course

1) Ingestion of the data

2) Visualization

ObsViewer, MEF, PSF_Anisotropy, DBviewer compatibility with Aladin & Topcat, PLASTIC connectivity

3) Analysis tools

GalFit, GalPhot, PhotRed, VODIA, Mdia

Recent developments (1)

New web services & tools

1) "Extreme data lineage" [Dependency Cutout]

From a single entry in a source list, it is possible to obtain "on the flight" a series of stamps with that particular source in all the frames that have contributed to create the source list, from the BIAS frame to the coadded image.

2) Image visualization

Tiff images of mosaics are produced on the flight, including the possibility to create 3-colour images.

3) New instruments: a standard protocol to introduce new instruments has been defined. Currently interfaces exist for the following instruments:
WFI(ESO2.2m), WFC(INT2.5m), MDM8K(MDM2.4m), LOFAR, WSRT, SUP(SUBARU), LBC-BLUE(LBT), ISAAC(VLT-UT1), ACS(HST).

Recent developments (2)

4) Calibration timestamp editor (CALTS)

5) Ingestion of reduced data or catalogs

It is now possible to ingest not only raw data but also reduced frames and source lists.

6) N-dim visualization & PLASTIC (PLatform for Astronomical Tool InterConnection)

PLASTIC is available in AWe to connect different applications (e.g. Aladin, Topcat, …)

7) "Federated data storage": now it is fully operative (tested here)

8) GRID: DPU task can be started in EGEE GRID (not yet tested here).

9) "Target processor" [web]

To process images from the web on one of the AW nodes.

AW documentation

http://www.astro-wise.org/

PYDOC

CVS

MANUAL

HOW-TOs

The people

AW personnel (hired with AW funds)

A. Volpicelli (2003-2005) ► code implementation

F. Getman (since 2003)
- ► DB administrator
- ► AWe maintenance
- ► code implementation

Science staff and others

M. Capaccioli (PI, since 2002)

R. Silvotti (contact person, since 2002)

J.M. Alcalá (since 2002)

A. Grado (since 2002)

M. Pavlov (2002-2005)

E. Puddu (since 2002)

M. Radovich (since 2002)

S. Leccia (since 2005)

A. Romano (since 2006)

M. Dall'Ora (since 2006)

The ASTRO-WISE node in Naples

Thanks to:

- All the Groningen AW team

- All the OAC AW team
  (without Fedor NO course !)

[Monica]

THE END

Or the beginning ?

if you find bugs, if you have suggestions

on how to improve the documentation

or the algorithms (or the next course), etc. ...

please send me by e-mail your comments !