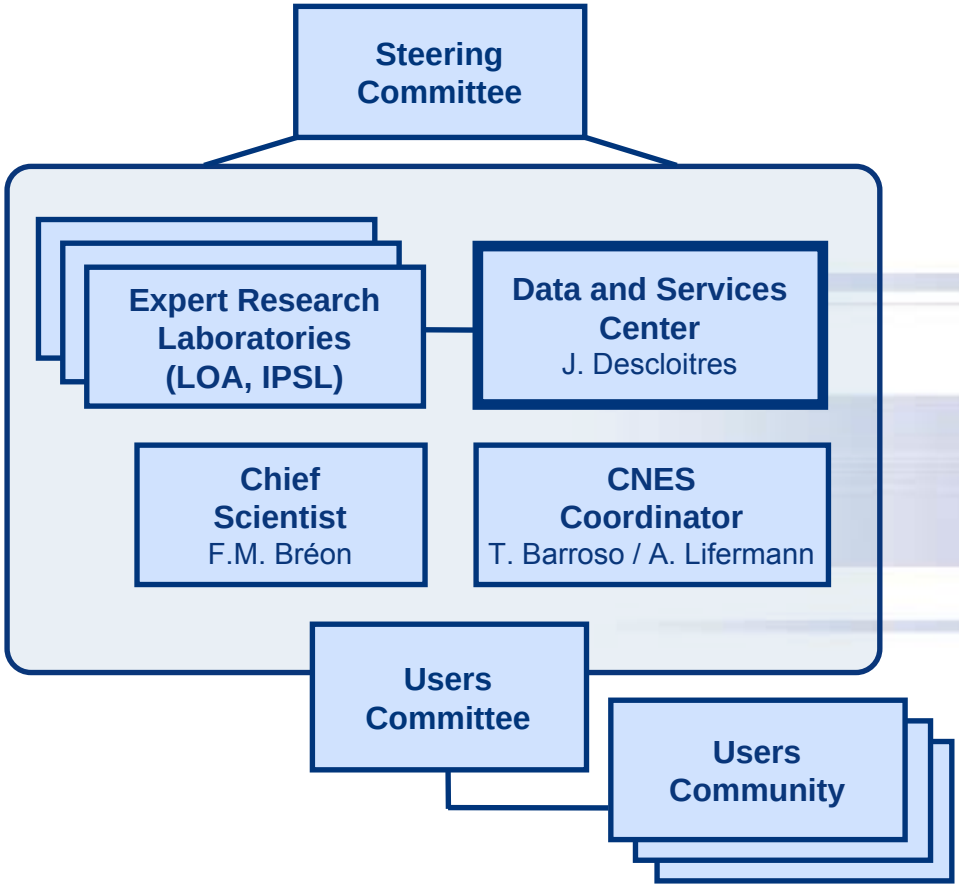
Icare Data and Services Center

on aerosols, clouds and water cycle
Computing facility and user interfaces

Loredana Focsa, Jean-Marc Nicolas, Jacques Descloitres
And the ICARE team

Table of Contents

Icare Thematic Center (1/3)



- ICARE is a structure created in 2003 to coordinate research and observations within a specific thematic perimeter: about **aerosols, clouds, water cycle, radiative transfer in the atmosphere**
- Its key mission is to provide data, products and services to the international scientific community

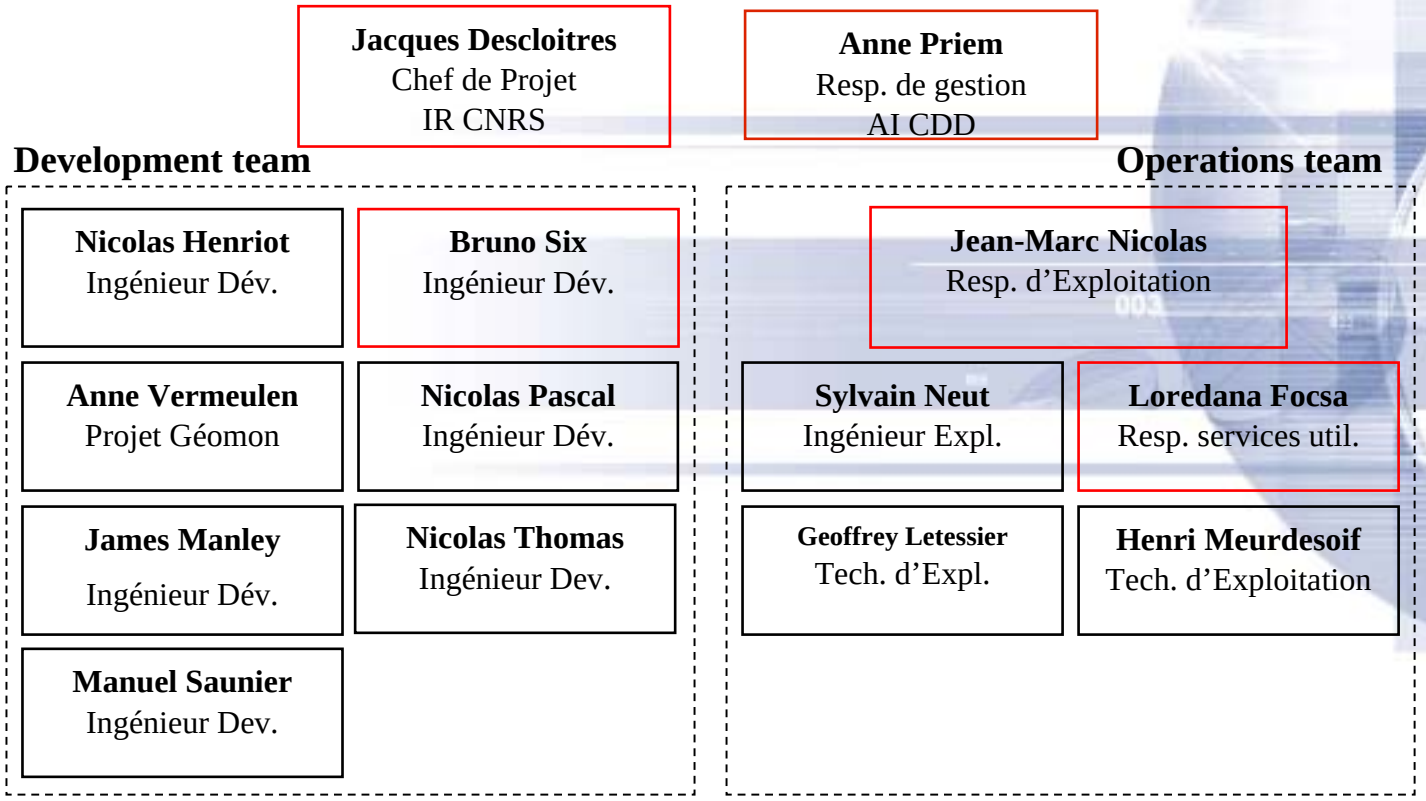
ICARE governance is : CNES, CNRS/INSU,University of Lille and Nord-Pas-de-Calais Regional Counci

Icare Thematic Center (2/3)

- **ICARE activities :**
  - Provide development services and tools
  - Develop and maintain computing facilities (hardware, software, distribution services)
  - Level 2 and 3 ground segment for current (PARASOL) or future missions (Megha-Tropique)
- **Some ICARE Services :**
  - operational processing
  - data collection from other providers
  - archive and distribution
  - science code development
  - visualization and analysis tools
  - documentation and user support
- **ICARE's key concept** is a symbiotic relationship between research labs and the Data and Services Center

Icare Thematic Center (3/3)

A relatively small unit based on a development team and an operations team :



Hardware Architecture (1/2)

2 LTO librairies (backup)

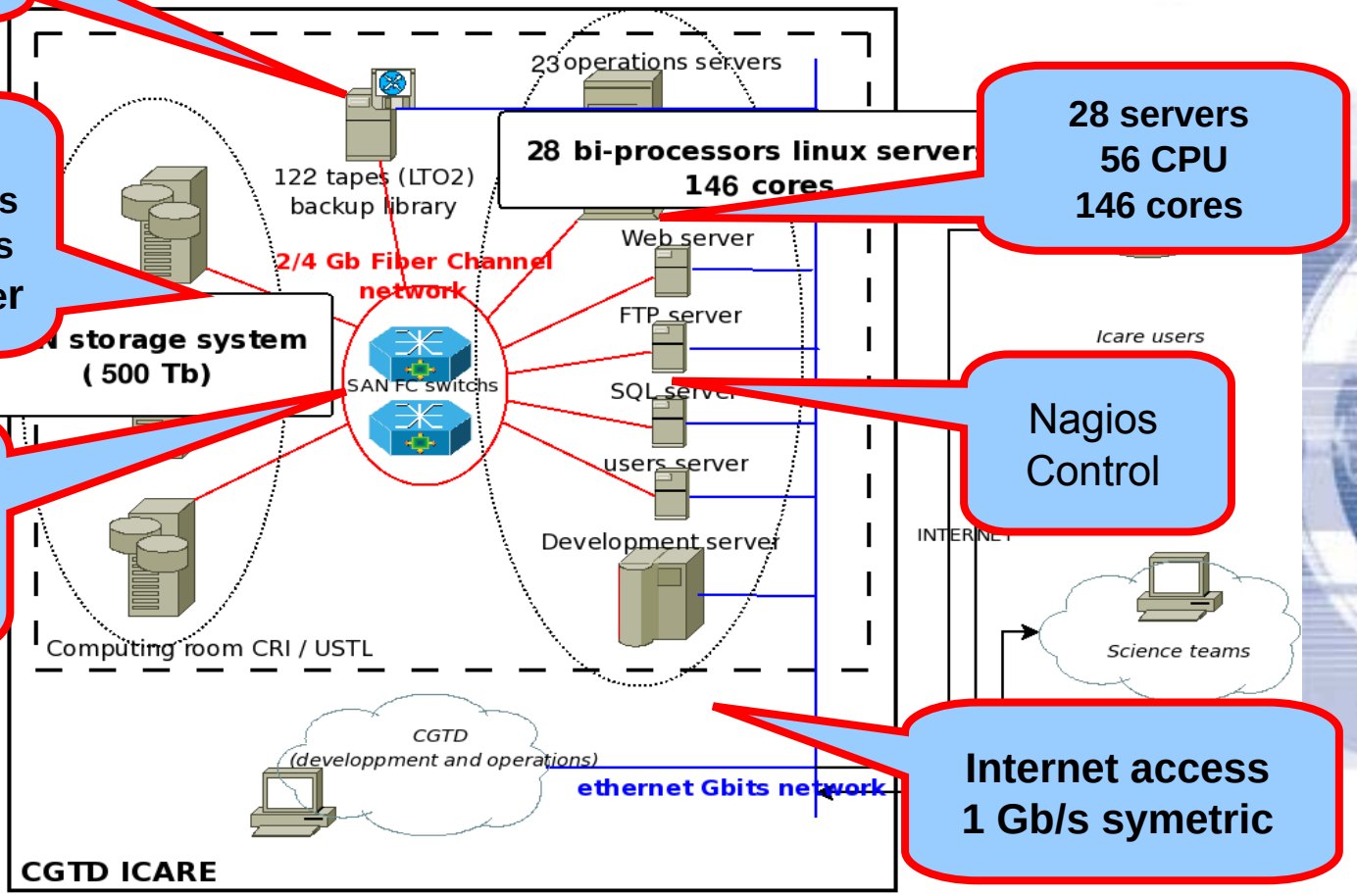
500 TB
15 storage arrays
1033 Hard Disks
1 km optical fiber

2 x 4 FC switch
2 SVC nodes

28 servers
56 CPU
146 cores

Nagios Control

Internet access
1 Gb/s symetric



Hardware Architecture (2/2)

SAN Storage System heterogeneous : IBM and Dell/EMC[2]

13 IBM System Storage Disk Arrays DS4000, DS4700, DS4800

2 Dell/EMC (CX3-20 and CX4-240)

2 IBM BladeCenter with 20 servers-based blade

5 IBM Xseries

Storage clustering using IBM solution GPFS (General Parallel File System

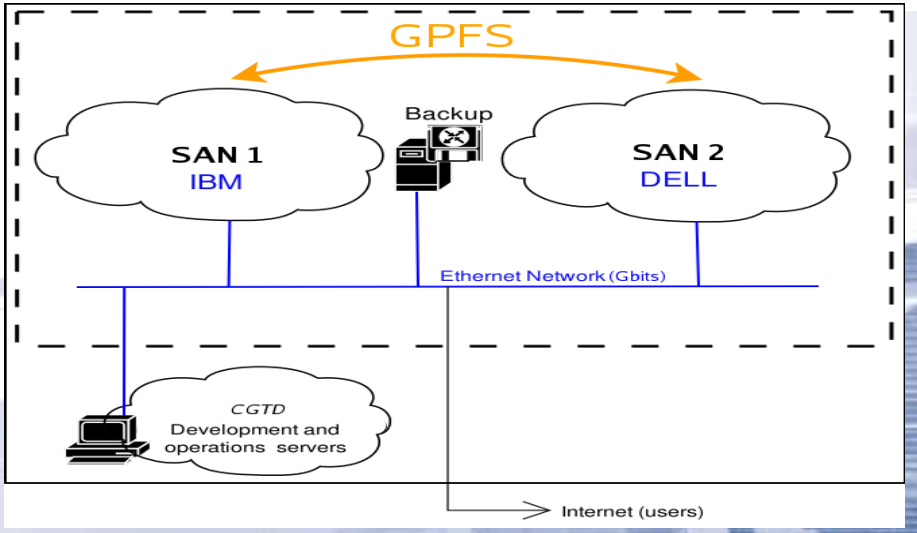
Two distinct high-performance storage clusters are connected using GPFS throught network.

GPFS allows large filesystems (many PetaB) and servers shared access

2 configurations :

-> **Direct access** (FiberChannel) : all disks are SAN-attached to all nodes

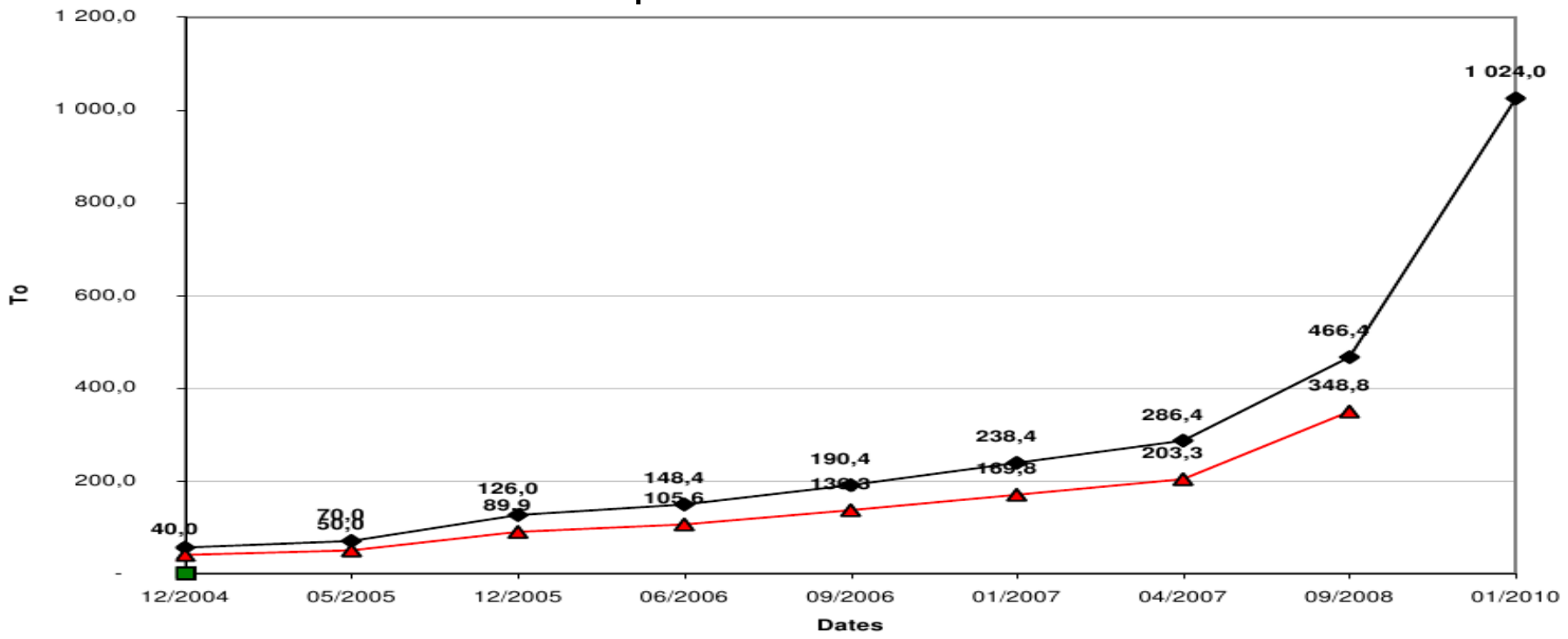
-> **Network Access** : nodes not directly attached to the disk can access FS using **NSD(Network Shared Disk)** server



430 TB available on 33 filesystems

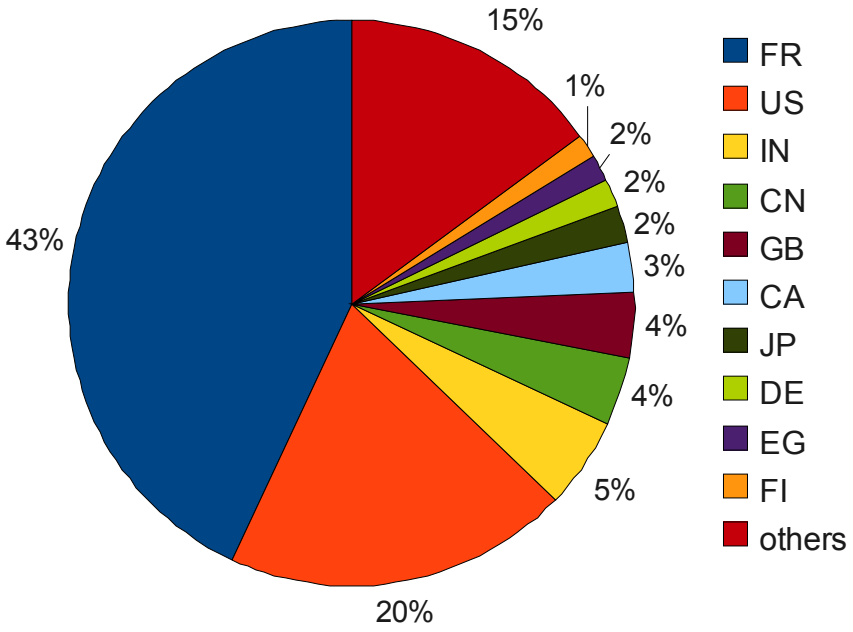
The tapes are used for backup only **(**IBM TS3310 LTO 3 with 3 drives, 122 storage slots) **~ 800 LTO3**

Exponential storage volume for last four years
Up to 1 PetaB in 2011



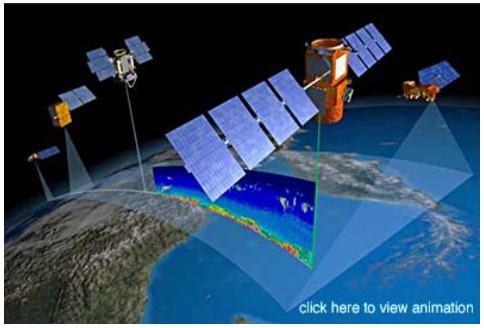
Evolution of storage capacity of ICARE archive

390 registered users at ICARE

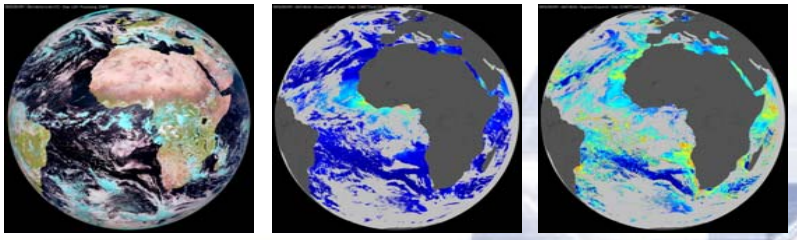


- Free registration
- But every user should be registered
- All users can access all open datasets (including MODIS, CALIPSO, PARASOL, AMSU, ...)
- Access data from ftp / web or user server.

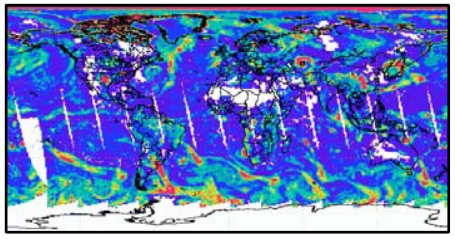
Data Holdings (1/4)



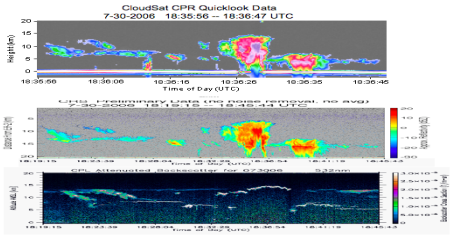
A-Train:
PARASOL/POLDER-3
Aqua/MODIS
CALIPSO/CALIOP
CALIPSO/IIR
CALIPSO/WFC
CloudSat/CPR
Aura/OMI



MSG/SEVIRI (aerosol products)



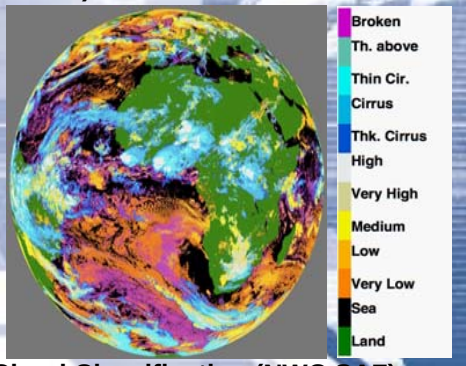
MODIS



CLOUDSAT

Geostationary:
MSG/SEVIRI
METEOSAT-7
GOES-E
GOES-W
GOES-SA
MTSAT
FY-2C

Ancillary:
ECMWF
GMAO



Cloud Classification (NWC SAF)

Microwave:
Aqua/AMSR-E
TRMM/TMI
DMSP/SSMI
DMSP/SSMI/S
METOP/AMSU-A
METOP/AMSU-B/MHS
NOAA/AMSU-A
NOAA/AMSU-B
Megha-Tropiques/SAPHIR
Megha-Tropiques/MADRAS
Megha-Tropiques/ScaRaB

No operator – a fully automated system developed at ICARE. **110 data sets** from over 20 satellite missions collected from **15 distant data providers**

Sources	*Access*	*Products*	*Nb. pr*	*Delay*	*Rolling period*	*Nb. files/d*	*Volume Go/d*
CNES	ftp	PARASOL L1/exo	4	D+1 to D	8 days	65	8.5
NASA/ASDC	ftp	CALIPSO	15	D+2 to D	15 days	400	28.0
NASA/LADS	ftp	MODIS	13	D+2 to D	15 days (L1) / 7	2400	90.0
Colorado St. U	ftp	CloudSat	10	D+3 to D	3 years	120	10.0
CMS / satmos	ftp	Geostationnaires	8	D+1 to D	20 days	2600	35.0
NOAA	ftp	AMSU (A/B), HIRS	29	D+2 to D	8 days	400	2.5
ESA	ftp	MERIS	2	D+1	8 days	75	18.0
Autres	ftp	OMI, TMI, AMSR_	12	D+2 to D	Mission lifetime	150	3.0
Total						***6210***	***195.0***

Data Holdings (3/4)

32 processing algorithms in operation

91 derived products and 54 associated “quicklook” produced by 40 processing code

Products	Nb. processing codes	Nb.files / day	Vol./day (Go)
PARASOL Aer	2 (L2/L3)	30	0.01
PARASOL Nuages	2 (L2/L3)	15	0.30
PARASOL Terre	2 (L2/L3)	15	0.30
PARASOL Browses	1 (L1/L2/L3)	450	0.03
CALIPSO browse	1 (L1,2)	105	0.20
IIR inter-etal.	2 (L1_c)	30	6.50
IIR L2	1 (L2)	90	3.50
CALIOP extraction	1 (L2)	30	0.80
CloudSat extraction	1 (L2)	105	1.60
Caliop-Cloudsat	1 (L2)	14	6.00

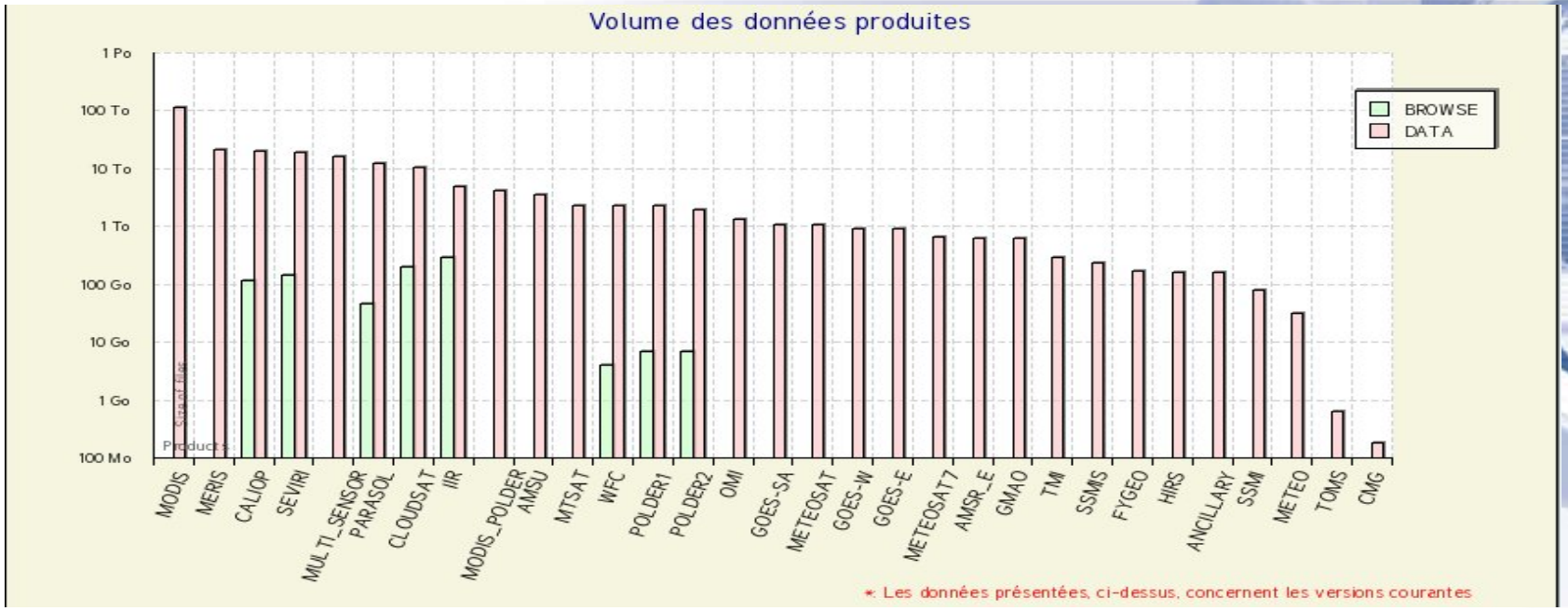
Products	Nb.processing codes	Nb.files / day	Vol./day (Go)
PARASOL-MODIS	1 (L2)	45	3.00
MSG aerosol (mer)	2 (L2/3)	96	0.10
MSG nuages	2 (L2)	96	4.00
SAF_NWC	1(L2)	2040	3.00
Brain	1 (L1)	140	15.50
Geo2hdf	1 (L1_c)	2000	12.00
AMSU	1 (L1_c)	150	1.00
EPSAT-SG	3 (L2/3)		
Total	***26***	***5376***	***54.83***

Data Holdings (4)

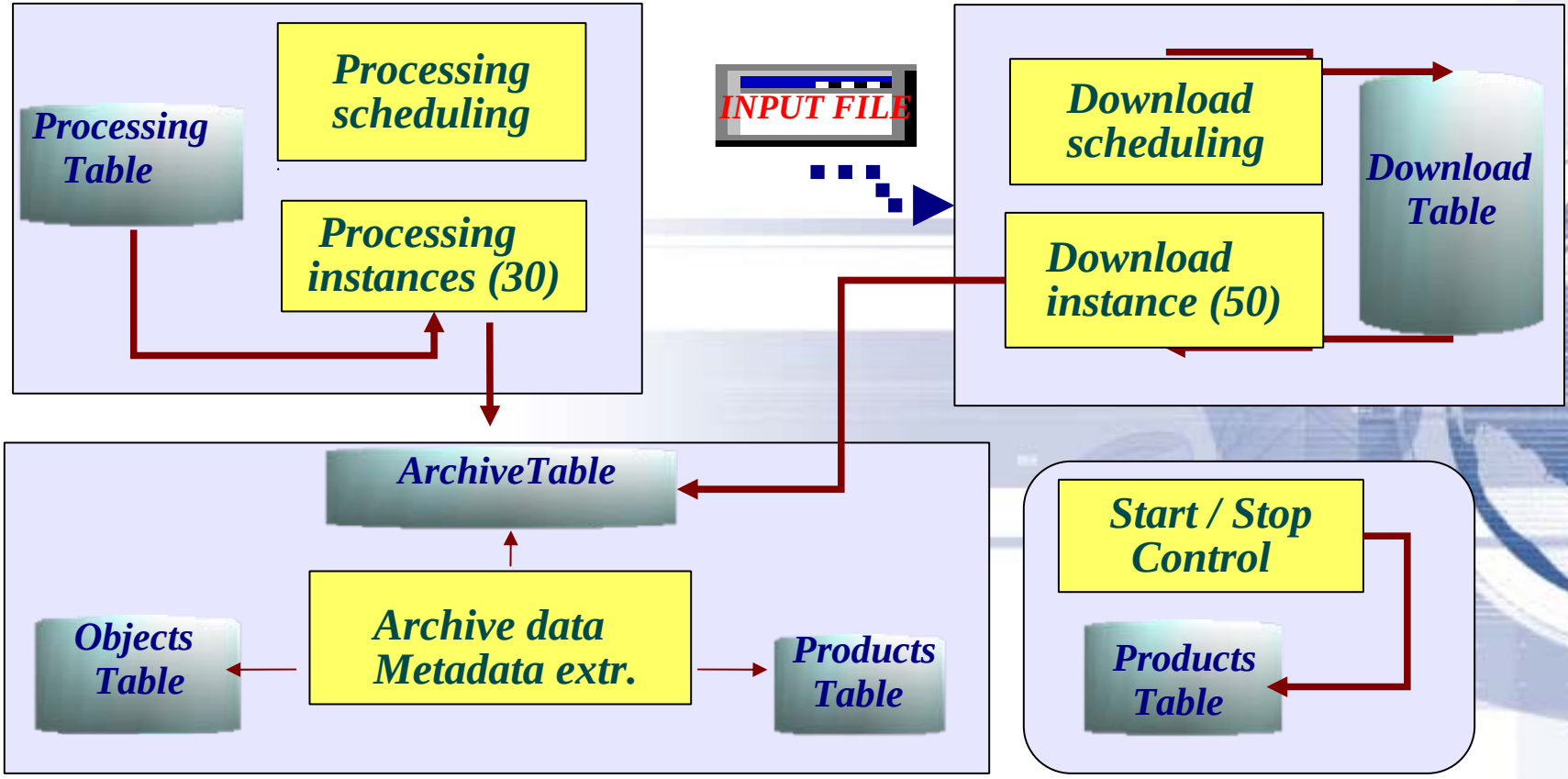
All data sets are available online by FTP, WEB and SSH

MODIS is the most voluminous of all data sets of Icare archive

Total volume per mission is variable, ~ 300 TB archived / distributed today



Production and Storage Mechanism



Distribution Services (1/3)

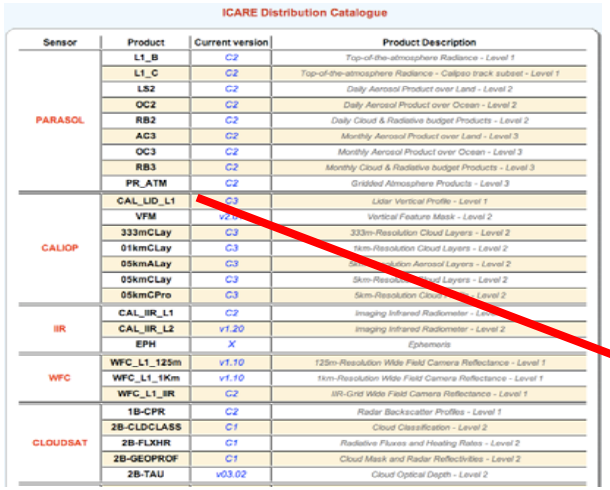
ICARE developed various distribution services with primary focus on:

- Data search and access
  - **Online archive**
  - **FTP** access for most products
  - Automatic product catalogue
  - Query interface
  - Distant computing resources for registered users (SSH access)
- Visualization and analysis tools
  - Browse imagery for all products
  - Browse-and-order approach
  - Visualization tools to display collocated products over the same grid
- 213 products archived and distributed – 53 associated “quicklook” products
- 390 registered users
- ~1TB distributed each month

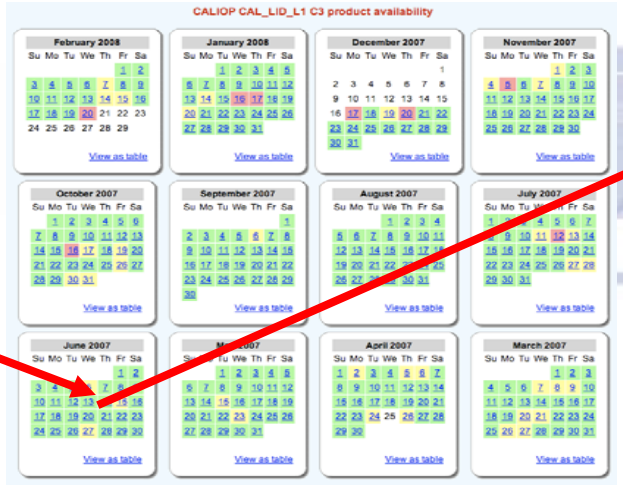
Distribution Services (2/3)

Online catalogue of all products distributed by ICARE

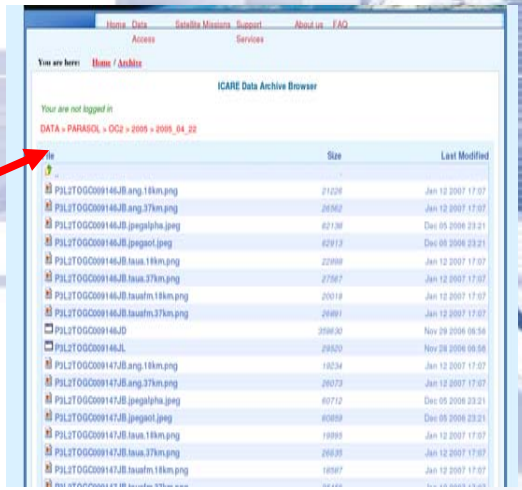
- Synoptic information about current product version, last update, and last data received
- Calendar layout for each product, with indication of daily data availability
- Direct access to online archive
- The catalogue is updated automaticaly



Product Version Catalogue



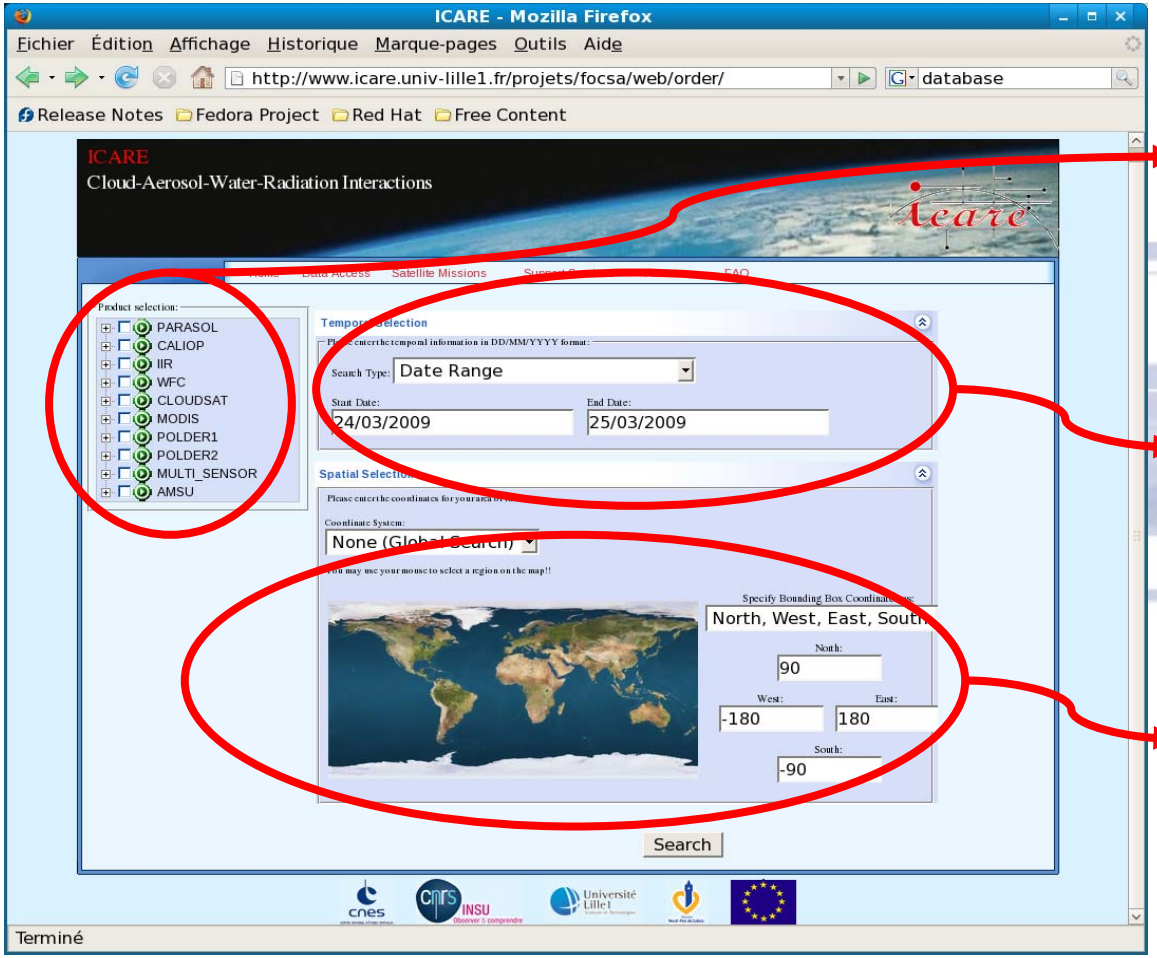
Product Daily Availibility



Product Catalogue and Web Access

Distribution Services (3/3)

Search and order for A-train products



Multi-sensor, multi-product selection

Date and time selection

Spatial selection

Visualisation Services (1/4)

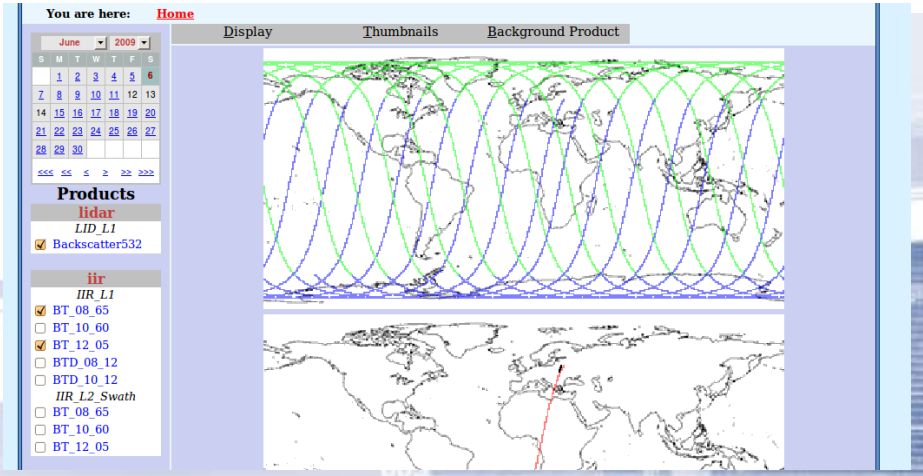
MSG/SEVIRI aerosols browser :

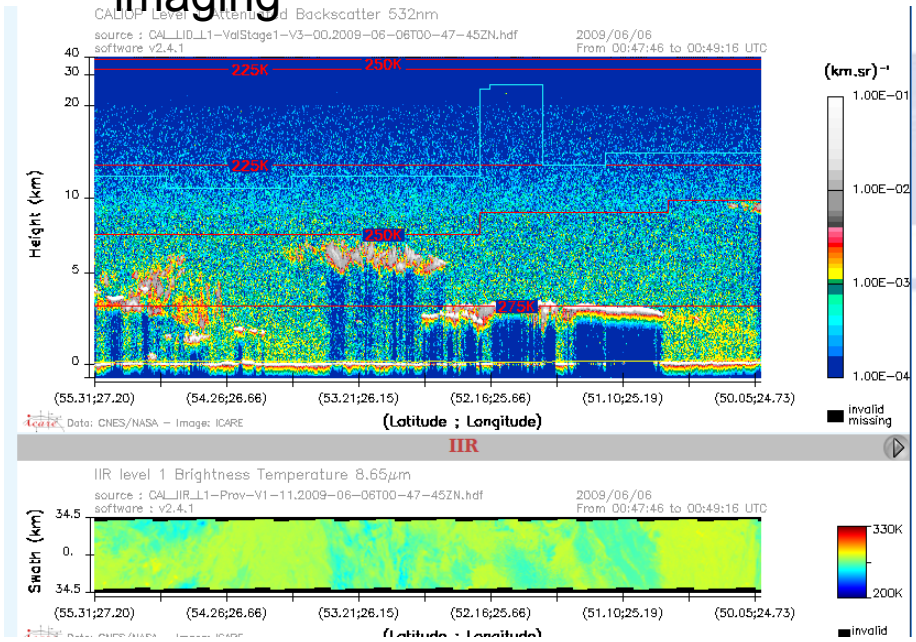
- True-color L1 composite (0.6, 0.8, 1.6 microns)
- Aerosol optical thickness (charge) and angstrom coefficient (~ particle size)
- For each slot and daily synthesis
- Zooming capabilities

Visualisation Services (2/4)

CALIPSO (Lidar + VIS and TIR camera) :

- Calendar, orbit and scene selection
- Lidar profile + VIS/TIR imaging





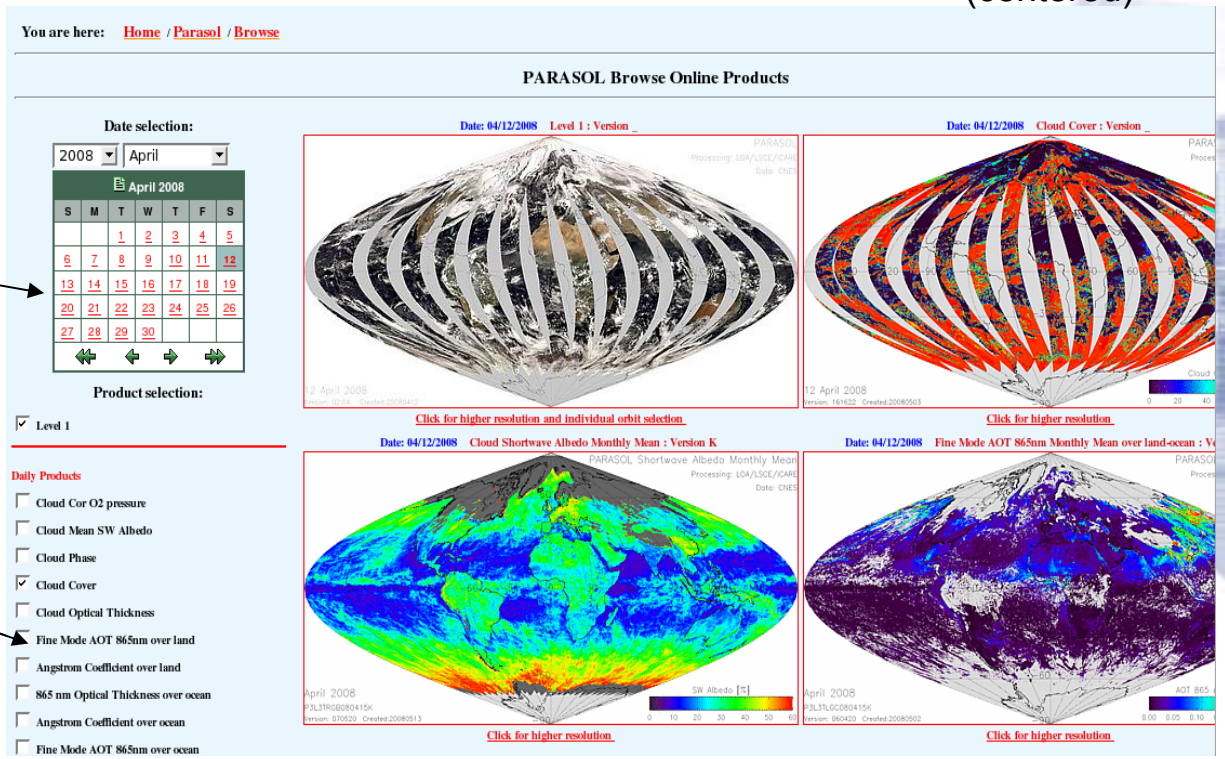
Visualisation Services (3/4)

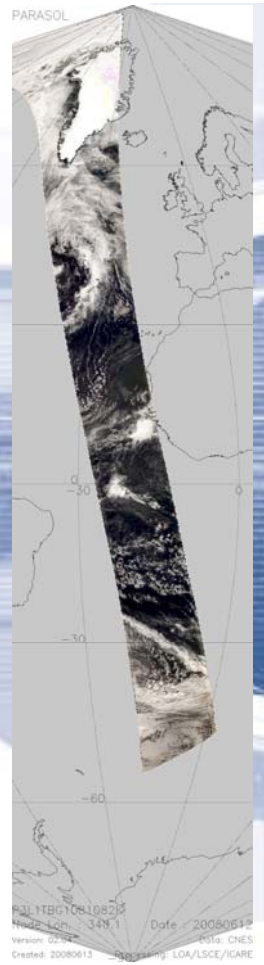
POLDER/PARASOL browser

Zooming on each orbit (centered)

Calendar selection for the three POLDER sensors

Physical parameter selection





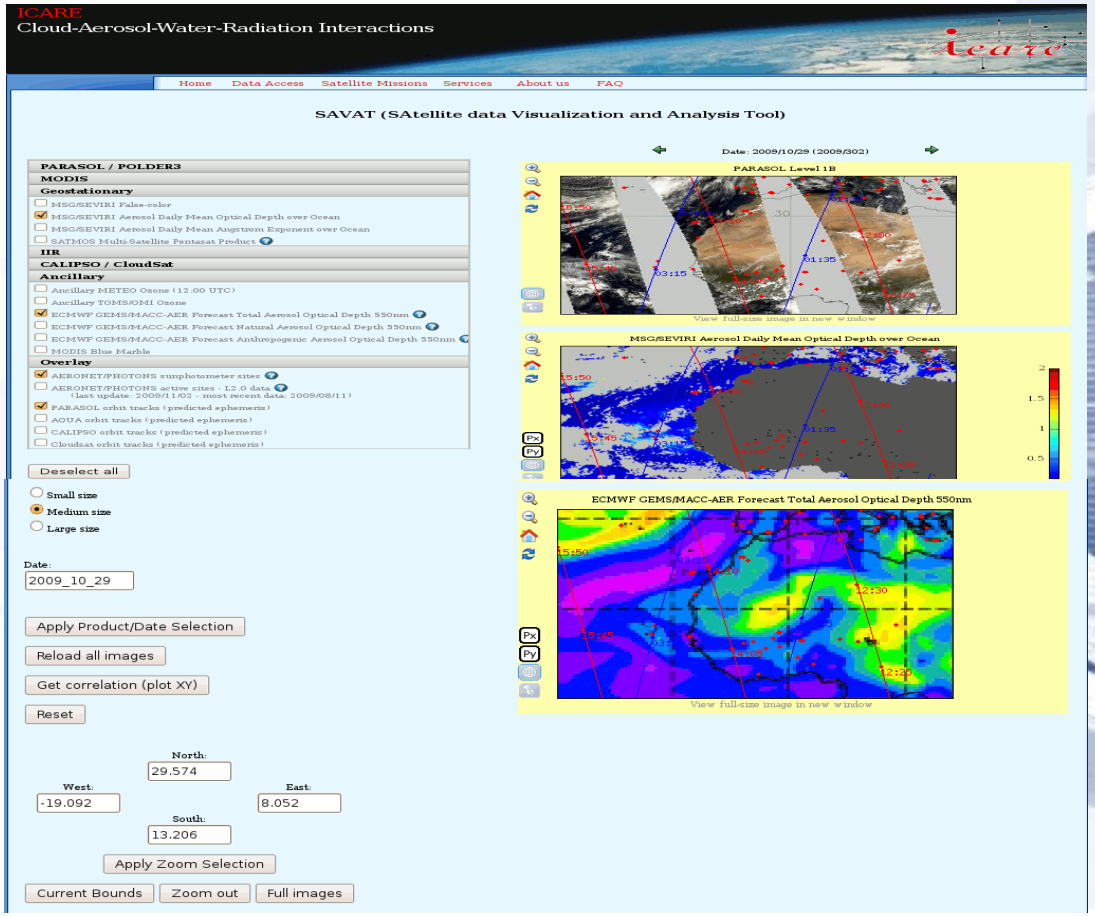
Visualisation Services (4/4)

Multi-sensor visualisation interfaces :

- L1/L2/L3 satellite data
- Models / analysis data
- Satellite overpath
- In-situ data (aerosols measurements from AERONET)
- Zooming capabilities
- X vs Y plotting availability

Under development :

- Access to real data in user-friendly formats (HDF, NetCDF)
- Projection issues



For more information and
all ICARE data and services access:

http://www.icare.univ-lille1.fr

Loredana.focsa@univ-lille1.fr
Jean-Marc.Nicolas@univ-lille1.fr
Jacques.Descloitres@univ-lille1.fr