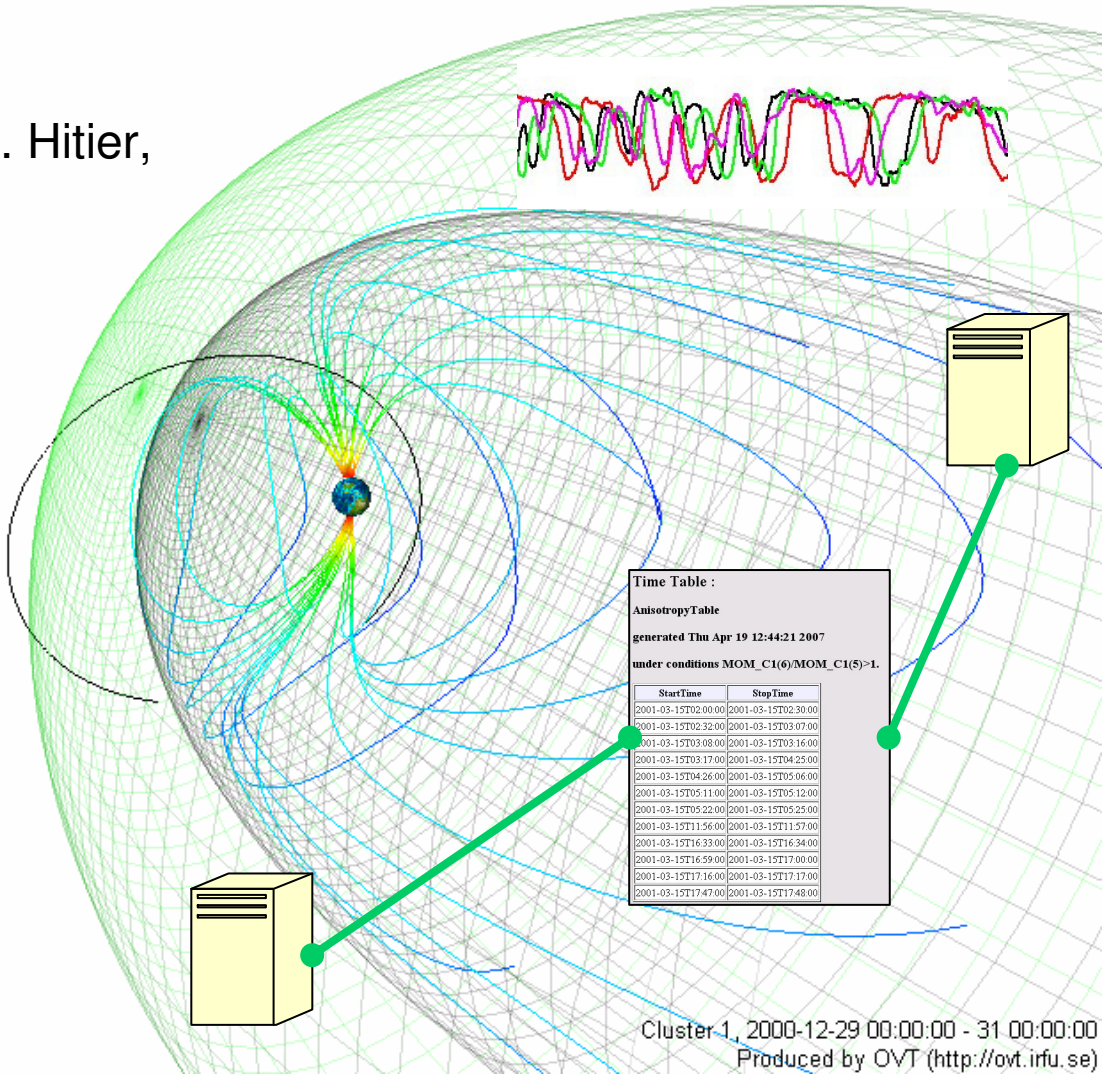
Science archives need to communicate more than data : the example of AMDA at CDPP

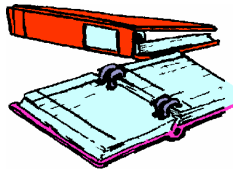
V. Génot, C. Jacquey, E. Budnik, R. Hitier, M. Bouchemit, M. Gangloff
CDPP@CESR, Toulouse, France

R. Conseil, D. Heulet, and C. Huc
CNES



Example: event search

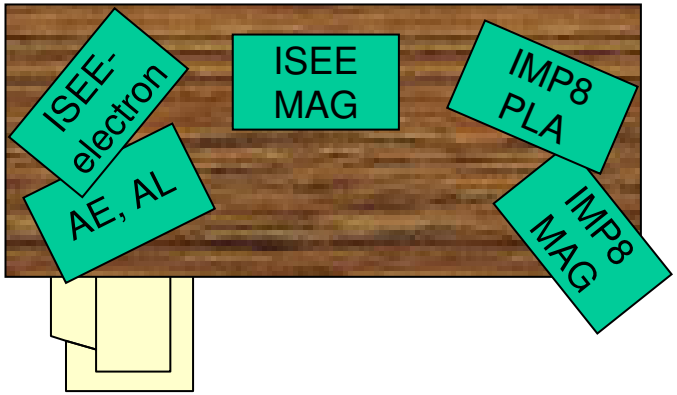
Old fashion: “paper” search



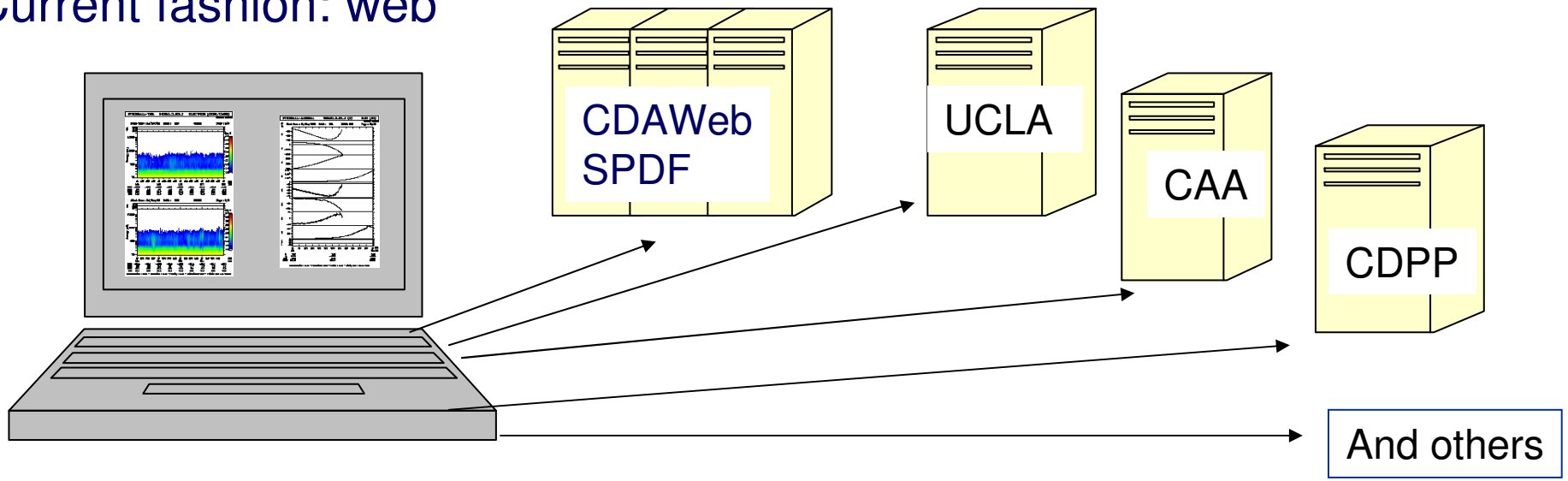
AE, AL

ISEE-data

IMP-8



Current fashion: web



Event search takes time and energy

Meeting user needs ...

1/ An archive should be active

-with value added services which motivate the user to come back regularly and provide inputs on potential improvements
-to encourage science feedback
→ promote a **user community** (***idea 1***)

2/ An archive should be friendly

-offer data directly in the natural format ready to be plotted or more generally manipulated
-be close to users' science
→ do not provide only data files but also direct science **parameters** (***idea 2***)

3/ An archive is only one link of a chain

-while VOs are in construction or just started, users still juggle from databases to archives;
-they need a *common media to hold, transmit and compare their events of interest*
→ event lists/**time tables** are portable, rich and compact objects suitable for this purpose (***idea 3***)

Combining these 3 ideas, AMDA system is one answer to meeting user needs in the Space Physics community

AMDA = Automated Mutiple Dataset Analysis

Multiple Dataset

the study of the (multi-scale) dynamics of plasma objects requires to perform the integrated analysis of multi-point and multi-instrument data

Automated (or semi-automated) Analysis

- Search, characterisation, classification of events
- Extensive mining of large database (statistics)
- “Historical” studies (example: through solar cycles)
- Building catalogs
- Building virtual constellations

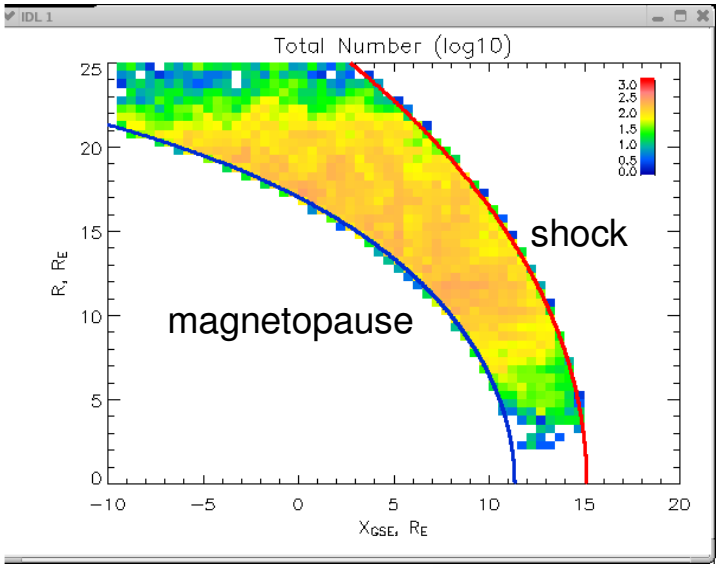
AMDA fact sheet

- First prototype : construction started in June 2006 after a ‘phase-A’
- Operational service opened recently (mid-april)
- Built on an existing data system (DD_System, A. Fedorov)
- Built in IDL, C, FORTRAN, Javascript (interface)
- Works on standardised and “simple” data :
  Cluster, Ace, Geotail, Isee, Imp-8, interball-Tail, Polar, indices, ...
- Includes models (Tsyganenko, Shock, MP, NS, …)
- Preliminary use of CDAWeb WebServices
- Forum for user feedback
- Already in use for scientific studies

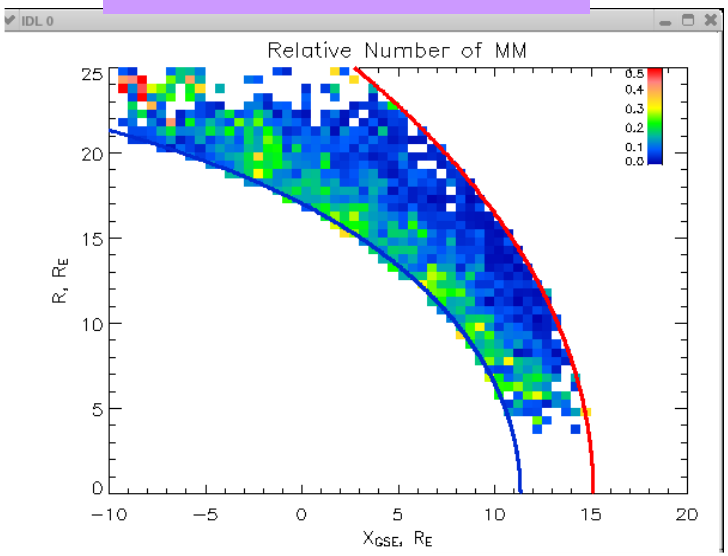
Visit
cdpp_amda.cesr.fr

Science with AMDA

Total number of 5 min magnetosheath crossings



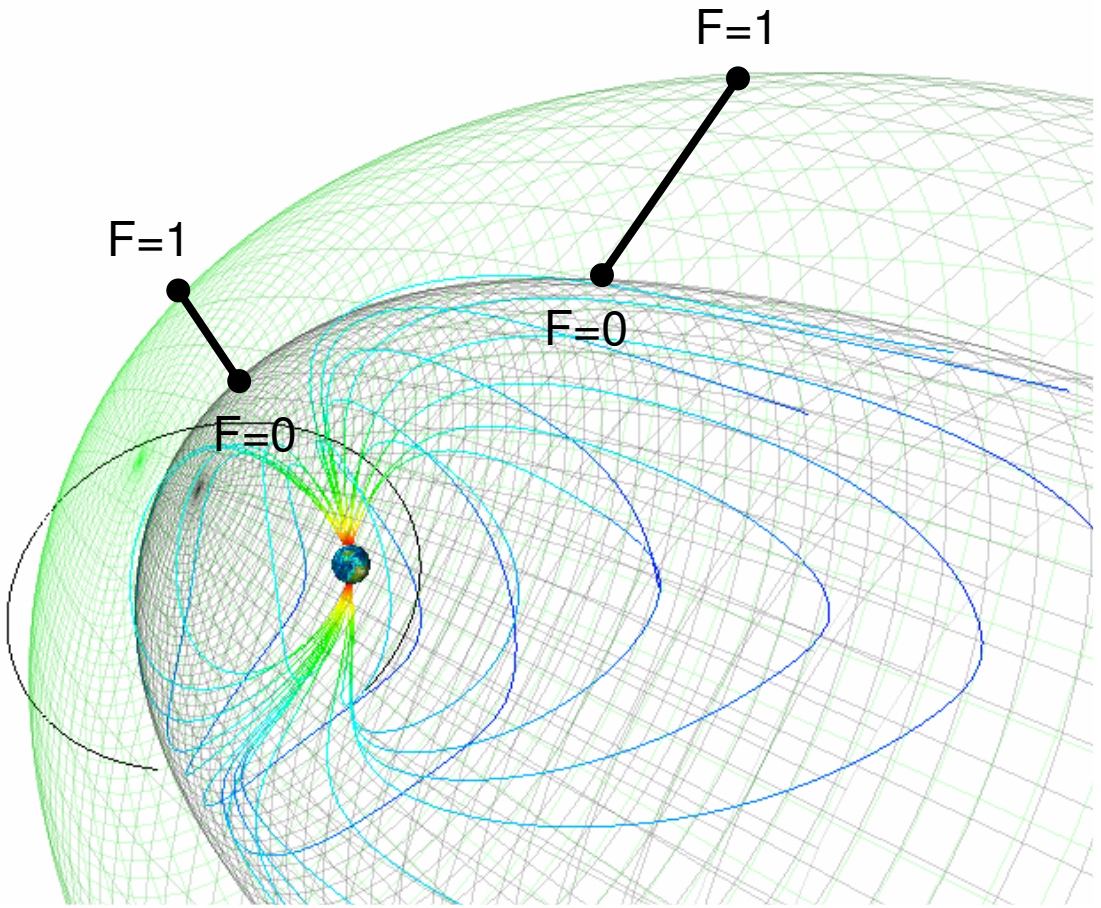
Relative number of mirror mode events



$$F_{\mathrm{mipm}} = \frac{r - r_{\mathrm{MP}}(\vartheta_{\mathrm{mipm}} | \rho V^2, B_z)}{r_{\mathrm{BS}}(\vartheta_{\mathrm{mipm}}, \varphi_{\mathrm{mipm}} | M_{\mathrm{a}}, M_{\mathrm{s}}, \vartheta_{\mathrm{bv}}) - r_{\mathrm{MP}}(\vartheta_{\mathrm{mipm}} \mid \rho V^2, B_z)}.$$

Verigin et al., 2001, 2003, 2006

For a position ***r*** inside the magnetosheath, the fractional distance is between 0 (MP) and 1 (BS)

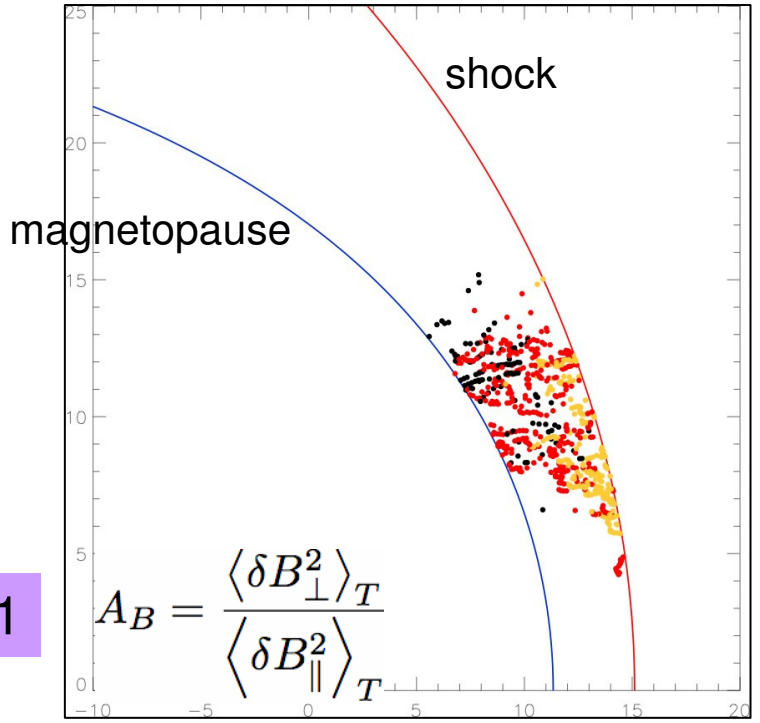


EGU General Assembly
Vienna, Austria, 15-20 April 2007

Statistical study of Alfvénic fluctuations in the Earth magnetosheath

O. Alexandrova[(1)], E. Budnik[(2)], V. Génot[(2)], C. Lacombe[(1)], C. Jacquey[(2)], I. Dandouras[(2)], E. Lucek[(3)]

[(1)]LESIA/CNRS, Observatoire de Paris, France
[(2)]CESR/CNRS, Toulouse, France
[(3)]Imperial College, London, UK



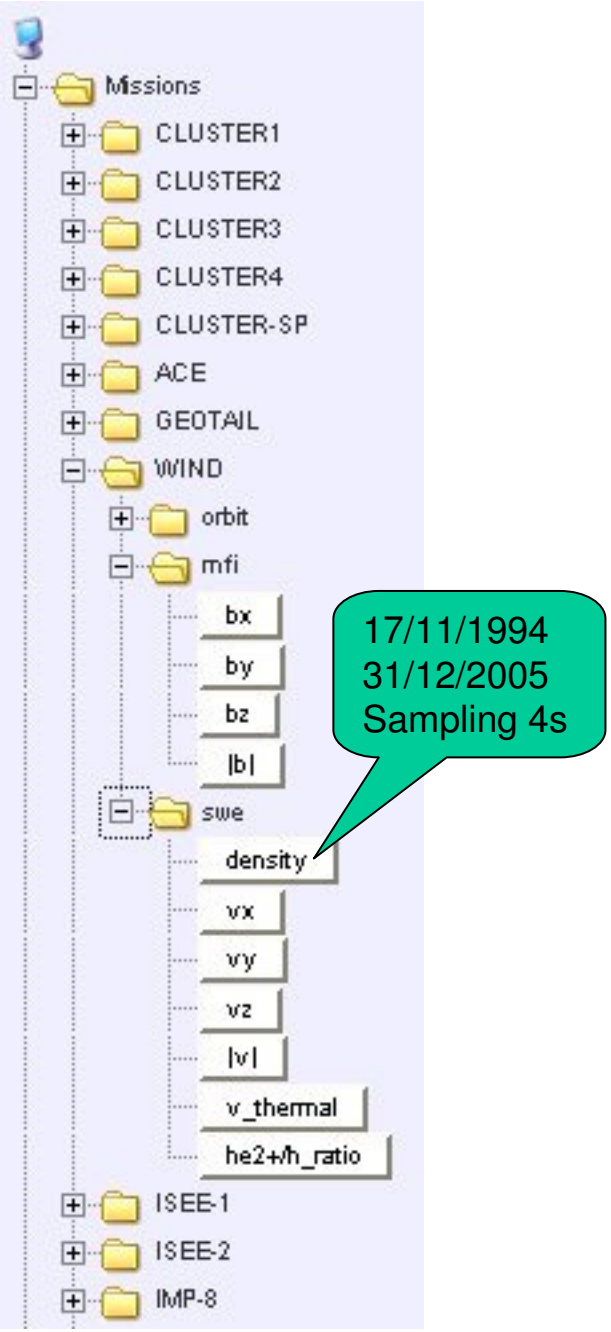
Condition = A_B>1

... other ways to make science with AMDA

1/ Multi-mission study of plasma sheet conditions minutes before substorm onsets (*Jacquey et al.*) :
- 25 years of data : ISEE, IMP, IRM, POLAR, GEOTAIL, WIND, ACE, CLUSTER
- combined approach:
    -automatic search = mathematical description of the temporal change in the AL index
    -semi automatic = visual inspection to classify cases in different catalogues

2/ The active plasma sheet: definition of 'events' and statistical analysis (*Louarn et al., 2006, ICS8*)
- define conditions : a lower threshold for the low frequency magnetic fluctuations and a minimal duration for each events
- produce time tables accordingly
- download a wide (resampled) dataset and work offline

So, what's inside AMDA ?



The parameter level

A parameter is any time series :
- magnetic or electric field vector
- particles moments
- indices
- ...
- *combinations of the previous*

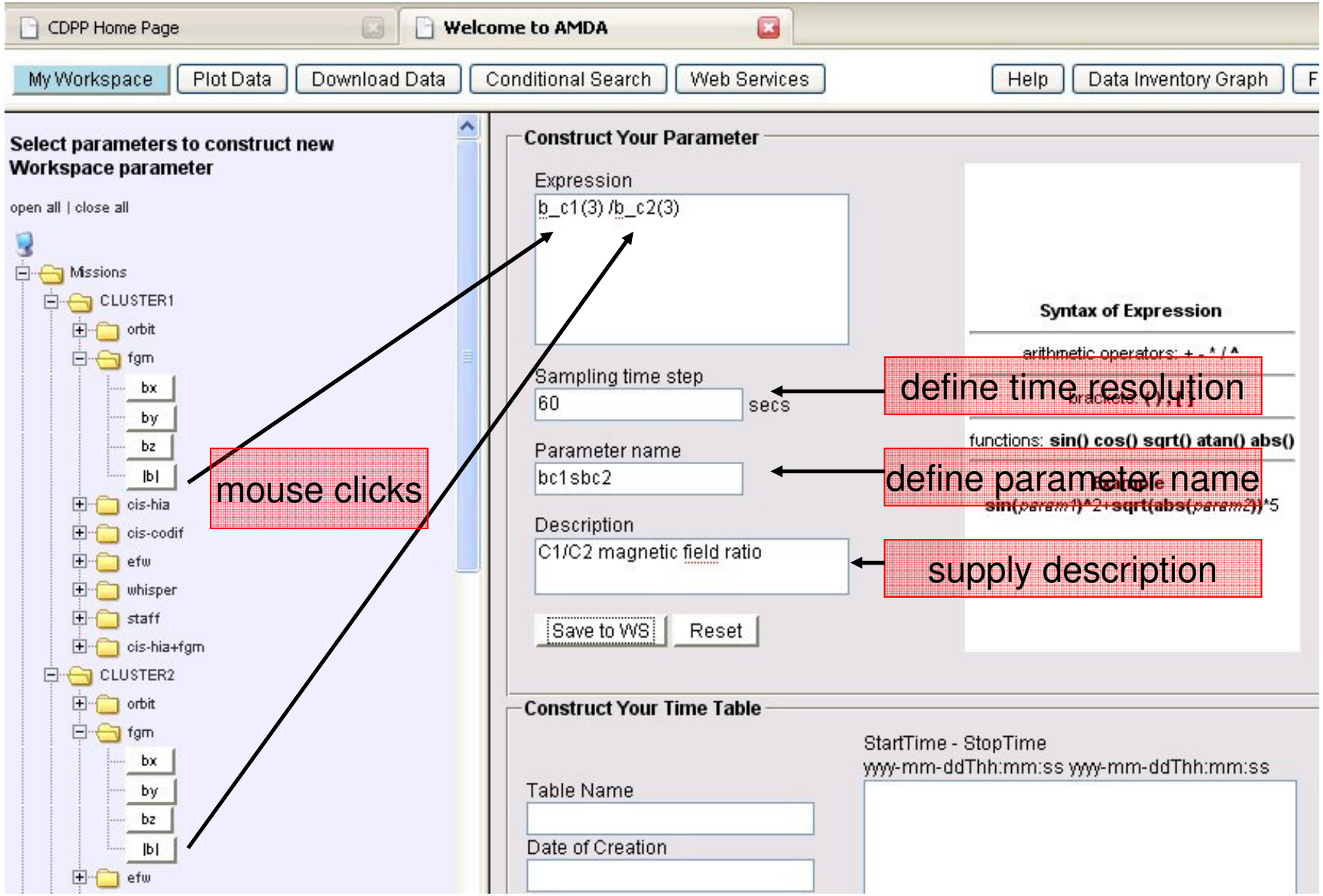
→ valid on a given time range

→ at a given time resolution

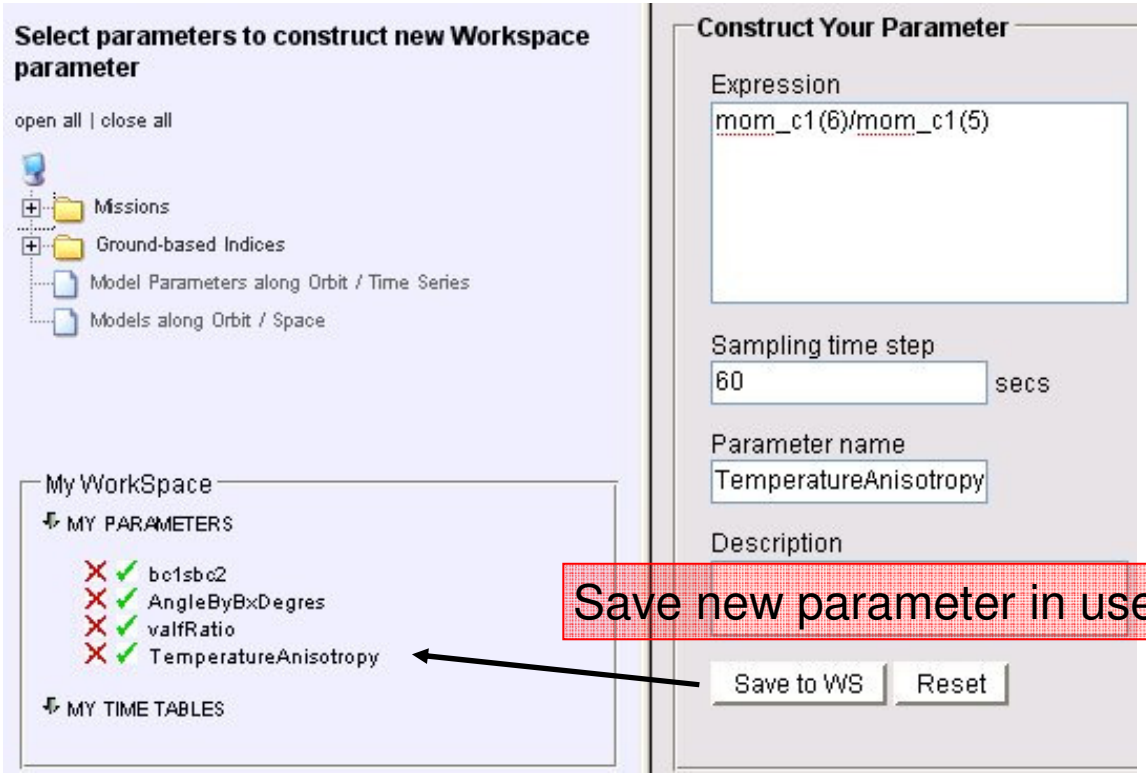
→ used for :
- plotting
- searching
- downloading, ...

The file organisation inside the archive is hidden to the user

New parameter contruction 1/2



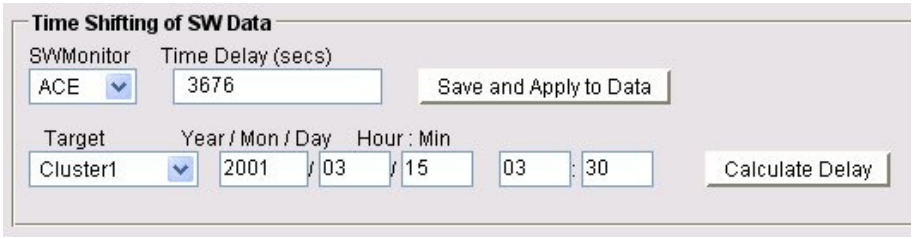
New parameter contruction 2/2



Save new parameter in user workspace

Parameters from
-different missions
-different instruments
-with different time resolution
may be conbined

The newly constructed parameter is now an element of the database.
It is computed on the fly only when it's needed (virtual object concept).



Tools are implemented :
-'time shifting' to account for time lag between instruments,
-'resampling' to account for resolution difference between instruments.

The Time Table level

an easy way to communicate information between archives

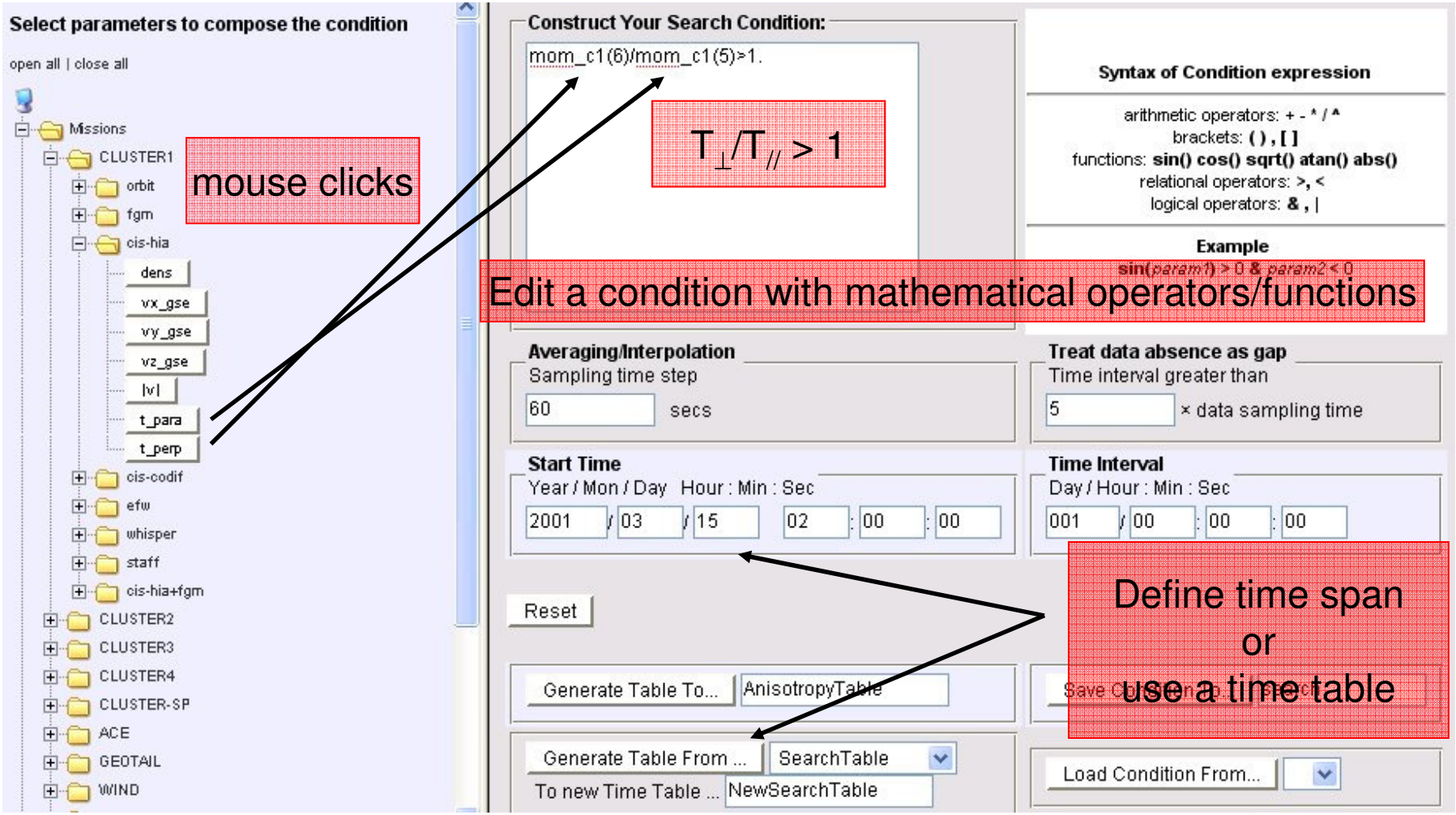
- Time tables, or event lists, are a collection of times when ‘something’ happens or time spans when some conditions are true.

- They are usually manually created and managed by scientists for their own need or the one of restricted community.

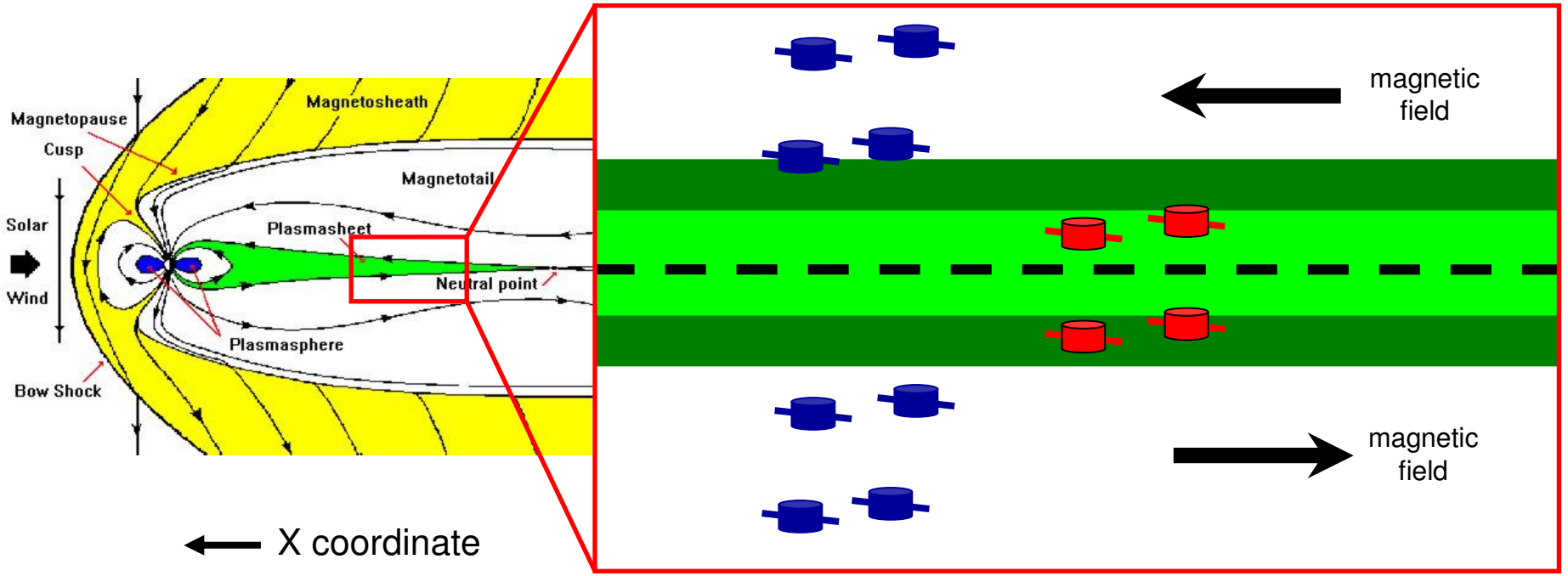
- AMDA enhances the potentialities of these objects by offering a generic tool to create, handle and communicate them between databases/archives and users.

These time-tables may be used for :
- extracting subdatabases,
- executing massive and/or interactive treatments,
- creating catalogs,
- serving as reference in the community

Create a time table : *1- Conditional Search*



An example: search of the events when the cross-tail current is “encircled” by CLUSTER

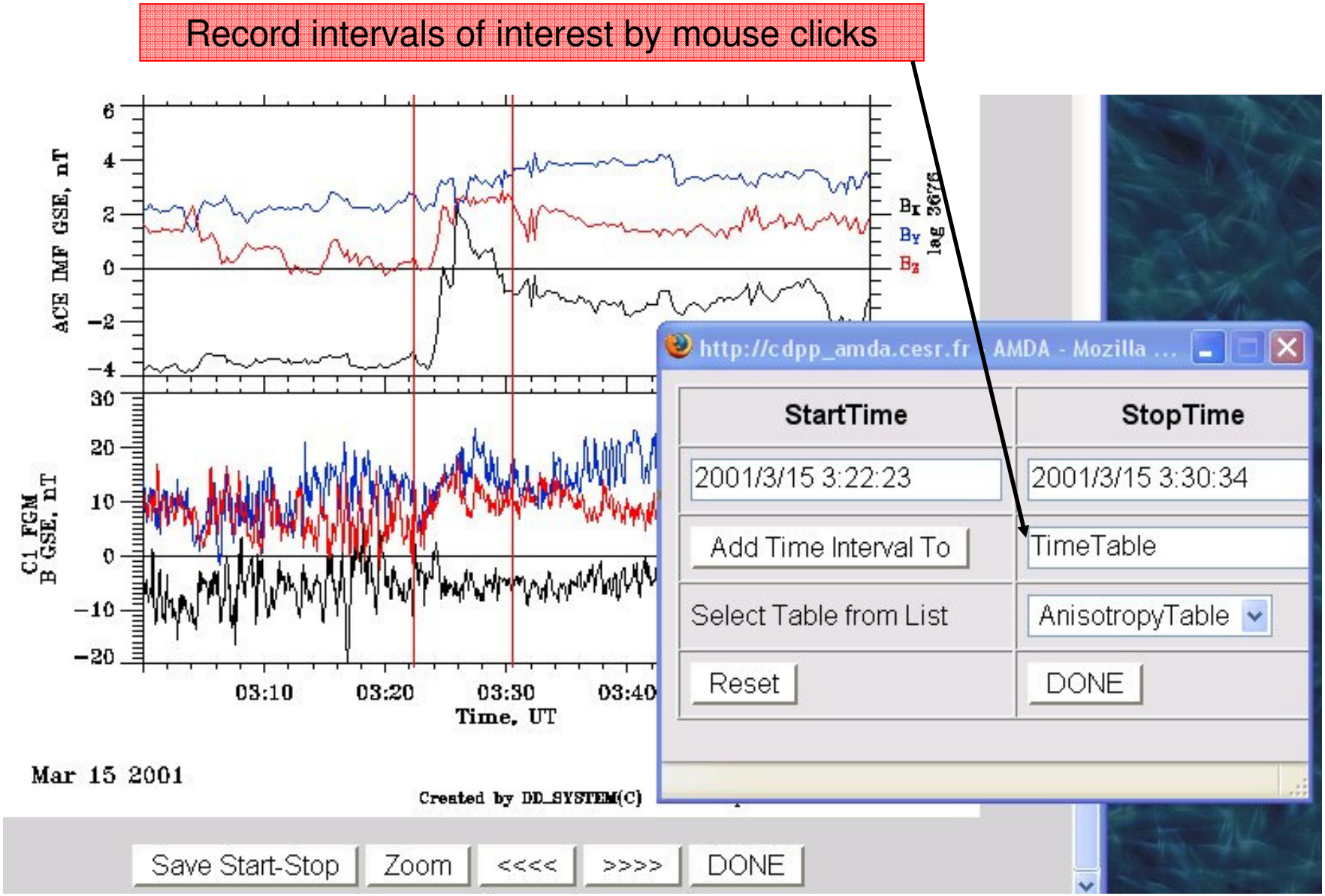


Looking for the time periods when 2 s/c are above the neutral sheet and the 2 other ones are below

Test:

- X1< -10 Re
- $BX_1 * BX_2 * BX_3 * BX_4 > 0$
- $\min([BX_1\ BX_2\ BX_3\ BX_4]) < 0$
- $\max([BX_1\ BX_2\ BX_3\ BX_4]) > 0$

Create a time table : *2- Visual Inspection*



AnisotropyTable

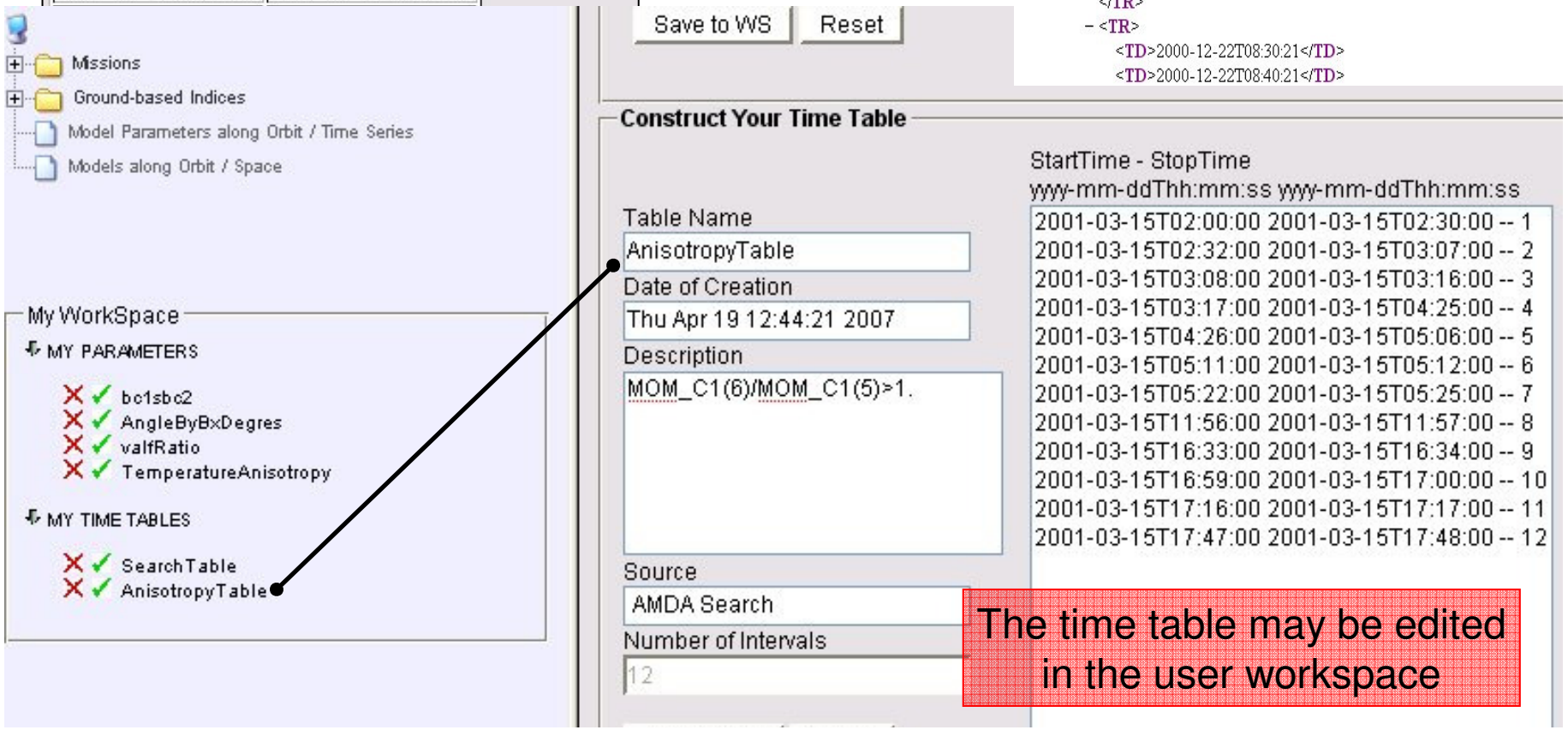
generated Thu Apr 19 12:44:21 2007

under conditions MOM_C1(6)/MOM_C1(5)>1.

StartTime	StopTime
2001-03-15T02:00:00	2001-03-15T02:30:00
2001-03-15T02:32:00	2001-03-15T03:07:00
2001-03-15T03:08:00	2001-03-15T03:16:00
2001-03-15T03:17:00	2001-03-15T04:25:00
2001-03-15T04:26:00	2001-03-15T05:06:00
2001-03-15T05:11:00	2001-03-15T05:12:00
2001-03-15T05:22:00	2001-03-15T05:25:00

Tables may be exported in VOTable format

```
<VOTABLE version="1.1" xsi:schemaLocation="http://www.ivoa.net/xml/VOTable/v1.1 C:\
- <DESCRIPTION>
    List of shocks from: http://www.sp.ph.ic.ac.uk/~eal/shocks-staffsa-031022.txt This list has
    from this time to get Start Time we add 5 minutes to this time to get Stop Time Creation Dat
  </DESCRIPTION>
- <RESOURCE>
    <DESCRIPTION> SPACECRAFT=CLUSTER </DESCRIPTION>
  - <TABLE>
    - <FIELD datatype="char" name="Start Time" ID="TimeIntervalStart" ucd="time.start">
        <DESCRIPTION>time tag for beginning of interval</DESCRIPTION>
      </FIELD>
    - <FIELD datatype="char" name="Stop Time" ID="TimeIntervalStop" ucd="time.stop">
        <DESCRIPTION>time tag for end of interval</DESCRIPTION>
      </FIELD>
    - <DATA>
      - <TABLEDATA>
        - <TR>
            <TD>2000-12-22T08:19:49</TD>
            <TD>2000-12-22T08:29:49</TD>
          </TR>
        - <TR>
            <TD>2000-12-22T08:30:21</TD>
            <TD>2000-12-22T08:40:21</TD>
```



```
<TABLE>
- <FIELD datatype="char" name="Start Time" ID="TimeIntervalStart" ucd="time.start">
    <DESCRIPTION>time tag for beginning of interval</DESCRIPTION>
  </FIELD>
- <FIELD datatype="char" name="Stop Time" ID="TimeIntervalStop" ucd="time.stop">
    <DESCRIPTION>time tag for end of interval</DESCRIPTION>
  </FIELD>
- <FIELD datatype="float" name="Geomagnetic Latitude" ID="GeomagLat">
    <DESCRIPTION>Geomagnetic Latitude</DESCRIPTION>
  </FIELD>
- <FIELD datatype="float" name="Geomagnetic Longitude" ID="GeomagLon">
    <DESCRIPTION>Geomagnetic Longitude</DESCRIPTION>
  </FIELD>
- <FIELD datatype="float" name="Local time of onset" ID="onset">
    <DESCRIPTION>Local time of onset</DESCRIPTION>
  </FIELD>
- <DATA>
  - <TABLEDATA>
    - <TR>
        <TD>2001-03-07T15:57:07</TD>
        <TD>2001-03-07T15:59:07</TD>
        <TD> 66.33</TD>
        <TD>181.57</TD>
        <TD>23.06</TD>
      </TR>
```

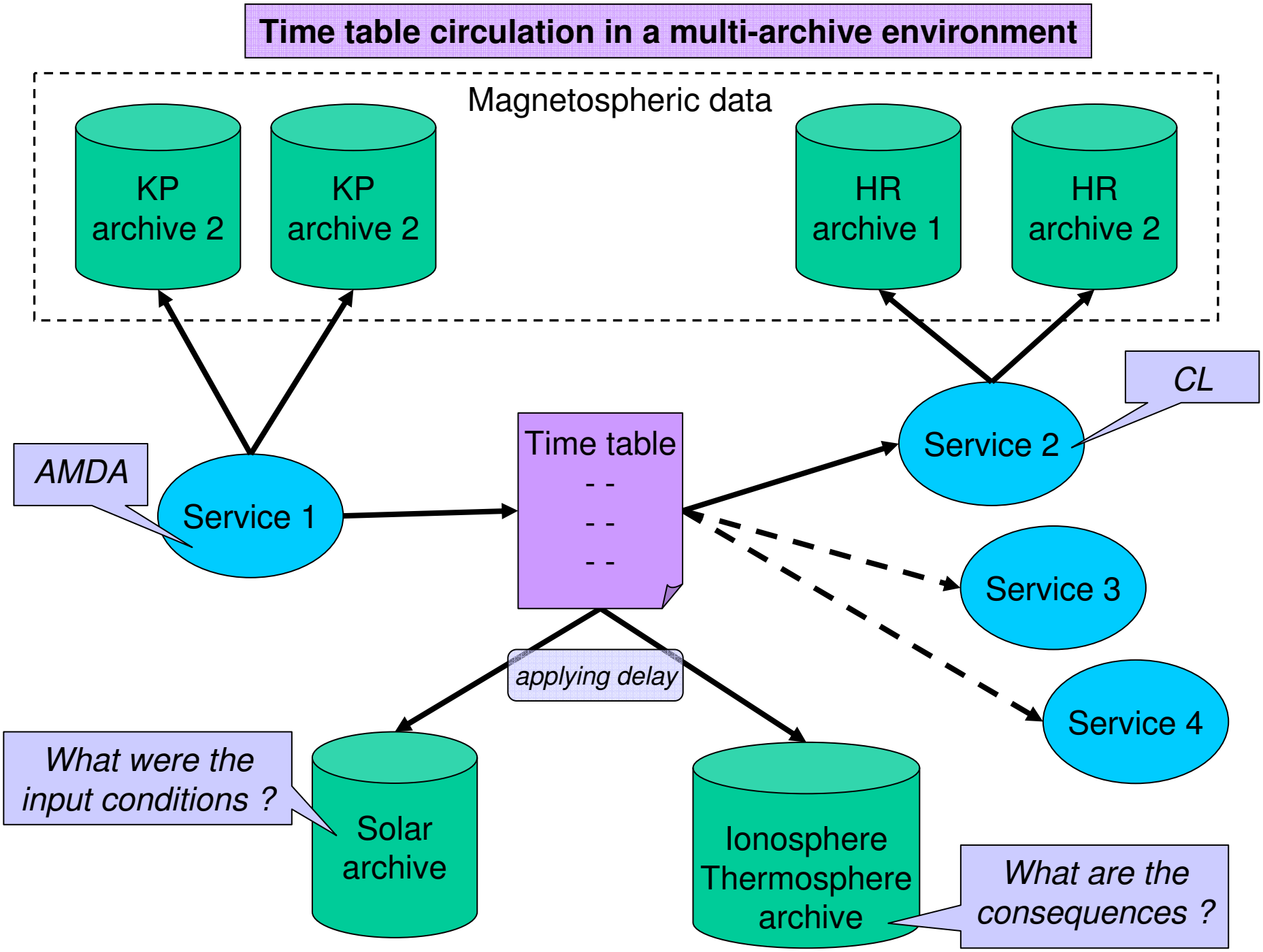
Complex tables may be produced :
- with multiple time tags
- associated parameters : average, max, min
- geographic information
- ...

Table description is a key issue for communication between archives

Example : List of Cluster magnetosheath crossings when ACE velocity is greater than 500 km/s during at least 5min with associated parameters such that maximum and average plasma beta and temperature anisotropy on these intervals.

→A careful description of this table would require :

- data description (instrument, resolution, ...)
- magnetosheath identification method (models, instrument modes, data)
- Cluster-ACE delay method (constant velocity, iteration)


Time table circulation in a multi-archive environment
Magnetospheric data
KP archive 2
KP archive 2
HR archive 1
HR archive 2
CL
AMDA
Service 1
Time table
- -
- -
- -
Service 2
Service 3
Service 4
applying delay
What were the input conditions ?
Solar archive
Ionosphere Thermosphere archive
What are the consequences ?

Communication with other archives

Parameter level

- protocol and communication design with close by archives :
  - -Data on Sitools system (Stereo, Themis KP @ CESR)
  - -MAPSKP (Cassini @ CESR)
- WebServices to CDAWeb : construct your own data tree
- AMDA interoperability with French solar databases (VHM) :
  - -MEDOC (SOHO)
  - -BASS2000 (Themis telescope)

Time table level

- tools for managing time tables
  - - ex.: union/intersection of tables coming from different services
- protocols for exchange