ODI
Open Data Interface
ESTEC/RFQ 3-12487/08/NL/AT

One year duration
October 2008 – October 2009
Contractor: Swedish Institute of Space Physics
(Peter Wintoft, Lars Eliassson)
Sub-Contractor: DH Consultancy
(Daniel Heynderickx)
ESTEC Contact: Hugh Evans

Contact: peter@lund.irf.se
Web: www.lund.irf.se/odi

Main items

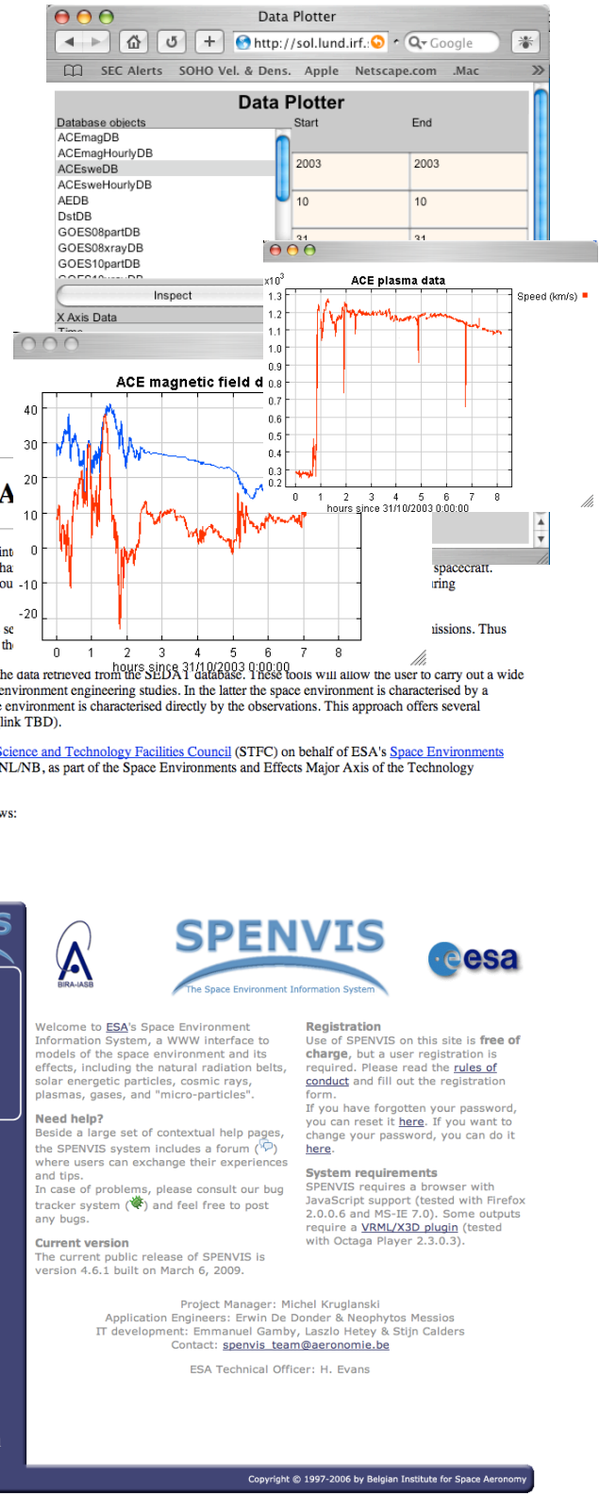
- MySQL database.
- CDF/ISTP/PRBEM compliant.
- SAAPS/SEDAT/SPENVIS data & metadata.
- SAAPS/SEDAT/SPENVIS will be updated to access data from the ODI database.

SAAPS/SEDAT/SPENVIS

- Spacecraft Anomaly Analysis and Prediction System (SAAPS) www.lund.irf.se/saaps

- Space Environment Database and Analysis Tool (SEDAT) www.ukssdc.ac.uk/sedat/

- Space Environment Information System (SPENVIS) www.spenvis.oma.be/



SEPEM

Welcome to the Solar Energetic Particle Environment Modelling (SEPEM) Project

Estec Contract No. 20162/06/NL/JD

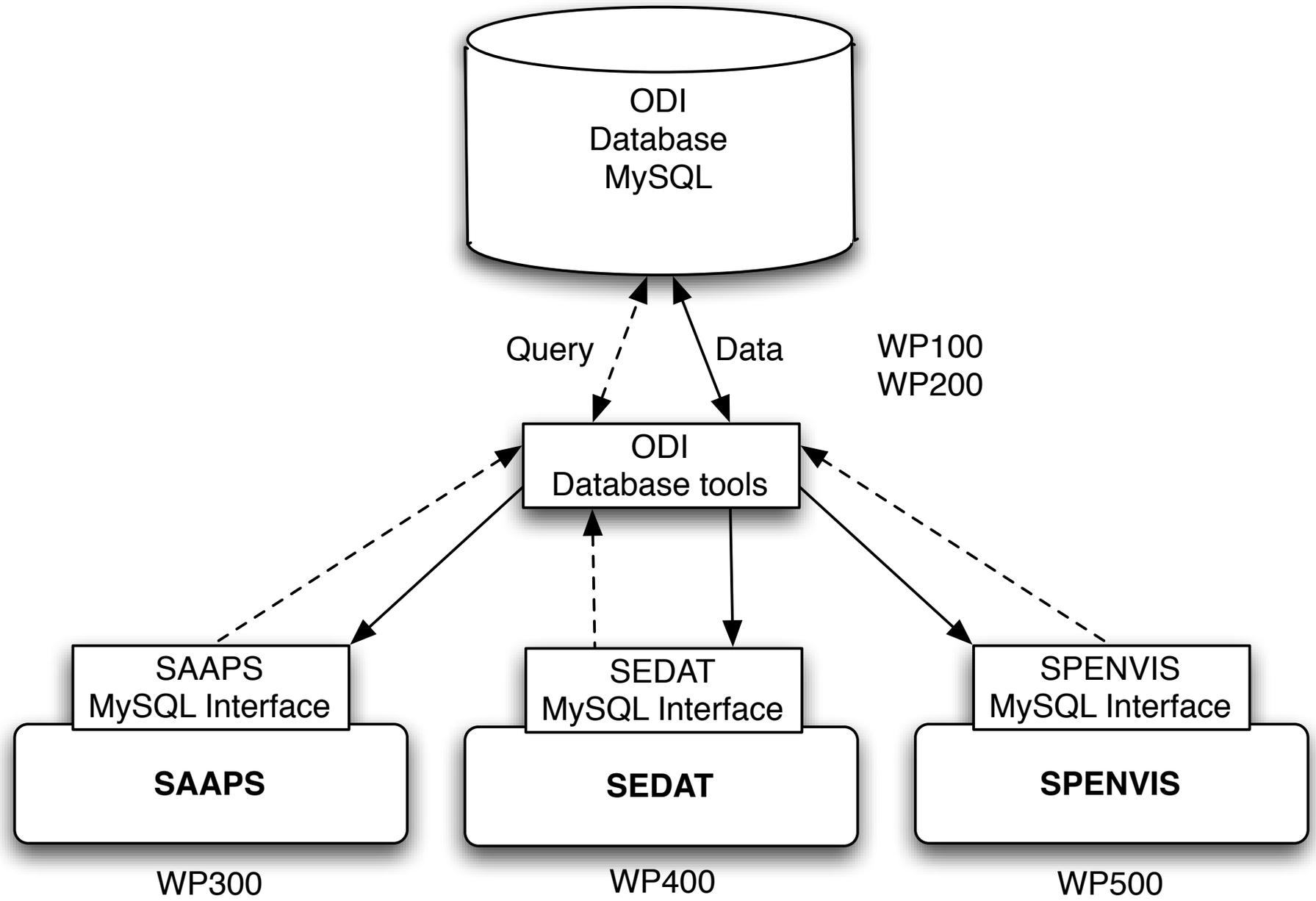
BACKGOUND: Various populations of energetic particles in the energy range from eV to more than 10^{21} eV fill our solar system. The most energetic among them are the constant flux of galactic cosmic rays and the sporadic solar energetic particle (SEP) events. The latter have energies ranging from a dozen of keVs to a few GeVs and are either linked to a solar flare or the shock wave driven by a coronal mass ejection. SEPs are mainly protons, electrons and alpha-particles, with small mixtures of ^{3}He nuclei and heavier ions up to iron. Space weather effects such as radiation damage to on-board electronics, solar arrays, living organisms, etc. are all due to energetic particles. SEPs are especially a significant hazard to Earth-orbiting missions, as well as for future interplanetary missions.

Many of the currently used standard models of the solar energetic particle environment were developed based on results published more than 15 years ago. Modern user requirements, as well as recent observational data and scientific advances, mean that these standards are currently in need of review and updating. Accurate modeling of the SEP environment constitutes a priority requirement for astrophysics missions and human exploration. Incorporating recent scientific results and a complete set of cross-calibrated data the SEPEM project is working towards creating new engineering models and tools to address current and future needs.

OBJECTIVES: One of the important outputs of SEPEM is the creation of a standard solar energetic particle dataset. A further output of SEPEM for the user community will be a user-friendly webserver with access to the models being developed under this project, in particular:

- it will take advantage of new data and take into account recent advances in understanding the generation mechanism;
- it will enable automatic model and tools update and establish community consensus;
- rather than producing only mission-integrated fluence for a given confidence level (for dose, solar array degradation), models and tools will be designed and developed to produce new types of user products (suitable for SEU rate, system impact and radiation background), including, but not limited to, peak flux statistics, durations of high or arbitrary flux periods, and corresponding error and uncertainty estimates;

http://www.oma.be/SEPEM/ norma.crosby@oma.be


ODI
Database
MySQL
Query
Data
WP100
WP200
ODI
Database tools
SAAPS
MySQL Interface
SAAPS
WP300
SEDAT
MySQL Interface
SEDAT
WP400
SPENVIS
MySQL Interface
SPENVIS
WP500

ISTP & PRBEM

International Solar Terrestrial Physics
spdf.gsfc.nasa.gov/sp_use_of_cdf.html

Panel on Radiation Belt Environment Modeling
craterre.onecert.fr/prbem/home.html

- The ISTP and PRBEM guidelines are defined using CDF.
- ISTP defines a set of valid CDF attributes, but not variables (except "Epoch").
- PRBEM uses the ISTP attributes and defines a set of valid variables.

Datasets overview

ACE magnetic field
ACE plasma
AE index
Dst index
Kp index
Ap index
F10.7
Sunspot number
GOES-5 e and p
GOES-5 x-ray
GOES-6 e and p
GOES-6 x-ray
GOES-7 e and p
GOES-7 x-ray
GOES-8 e and p
GOES-8 x-ray
GOES-9 e and p
GOES-9 x-ray
GOES-10 e and p
GOES-10 x-ray
GOES-11 e and p
GOES-11 x-ray
GOES-12 e and p
GOES-12 x-ray
LANL el.
OMNI
SAMPEX/PET
MIR/REM
ACE/SIS
AMPTE/UKS
AZUR
CRRES/MEA
ISEE1-2
METEOSAT
STRV1B/REM
UARS/PEM
XMM/SREM
PROBA1/SREM
HELIOS1-2
SOHO/ERNE
IMP8/CPME
IMP8/GME
IMP8/EPE
INTEGRAL/SREM
SAC-C
Equator-S
Rosetta/SREM
NOAA POES/SEM
METOP-02/SEM
GIOVE-B/SREM

Mixed formats:
ASCII table, HDF, CDF/
ISTP, CDF/PRBEM.

MySQL

- MySQL is an open source Structured Query Language (SQL) relational database management system (www.mysql.com).
- Data are stored in tables with a given number of columns and one record per row.
- MySQL can not store vector or matrix data (exception: BLOB). It must be expanded over multiple columns or tables.
- Columns may be defined to hold only unique values thus avoiding storing of duplicate records (e.g. two records with the same epoch).
- Advanced search capabilities over multiple data sets by joining tables.
- "Fast" and "easy to use".
- Different interfaces (command line, web, ...).
- Different programming interfaces (C, Java, Perl, PHP, Python, IDL, Matlab, Mathematica, ...).

Common Data Format (CDF)

"... is a self-describing data format for the storage and manipulation of scalar and multidimensional data in a platform- and discipline-independent fashion."

cdf.gsfc.nasa.gov/

Global attributes

```
! Skeleton table for the "SREMRosetta_PACC_20090101.cdf" CDF.
! Generated: Wednesday, 18-Feb-2009 09:48:17
! CDF created/modified by CDF V2.7.2
! Skeleton table created by CDF V3.2.4

#header

                      CDF NAME: SREMRosetta_PACC_20090101.cdf
                 DATA ENCODING: IBMPC
                      MAJORITY: ROW
                        FORMAT: SINGLE

! Variables  G.Attributes  V.Attributes  Records  Dims  Sizes
! ---------  ------------  ------------  -------  ----  -----
    0/12           13             13         0/z      0


#GLOBALattributes

! Attribute          Entry       Data
! Name               Number      Type        Value
! ---------          ------      ----        -----

  "Project"             1:    CDF_CHAR     { "SREM aboard Rosetta" } .

  "Source_name"         1:    CDF_CHAR     { "Rosetta" } .

  "Discipline"          1:    CDF_CHAR     { "Space " -
                                             "Physics>Interplanetary " -
                                             "Science" } .

  "Data_type"           1:    CDF_CHAR     { "PACC>Processed " -
                                             "Accumulation Data" } .

  "Descriptor"          1:    CDF_CHAR     { "SREM>Standard Radiation " -
                                             "Environment Monitor" } .
```

CDF (cont.)

```
! Variable          Data      Number                    Record   Dimension
! Name              Type      Elements  Dims  Sizes     Variance Variances
! --------          ----      --------  ----  -----     -------- ---------

  "label_COUNTERS"
                    CDF_CHAR      3       1     15        F        T

  ! Attribute       Data
  ! Name            Type        Value
  ! --------        ----        -----

    "FIELDNAME"     CDF_CHAR    { "SREM counter label" }
    "VAR_TYPE"      CDF_CHAR    { "metadata" }
    "CATDESC"       CDF_CHAR    { "SREM counter label" } .

  ! NRV values follow...

    [1] = { "TC1" }
    [2] = { "S12" }
    [3] = { "S13" }
    [4] = { "S14" }
    [5] = { "S15" }
    [6] = { "TC2" }
    [7] = { "S25" }
    [8] = { "C1 " }
    [9] = { "C2 " }
    [10] = { "C3 " }
    [11] = { "C4 " }
    [12] = { "TC3" }
    [13] = { "S32" }
    [14] = { "S33" }
    [15] = { "S34" }
```

Metadata variable

```
! Variable          Data      Number                    Record   Dimension
! Name              Type      Elements  Dims  Sizes     Variance Variances
! --------          ----      --------  ----  -----     -------- ---------

  "EPOCH"           CDF_EPOCH     1       0               T

  ! Attribute       Data
  ! Name            Type        Value
  ! --------        ----        -----

    "FIELDNAME"     CDF_CHAR    { "Time since 0 AD of accumulation" }
    "VALIDMIN"      CDF_EPOCH   { 01-Oct-2001 00:00:00.000 }
    "VALIDMAX"      CDF_EPOCH   { 01-Jan-2010 00:00:00.000 }
    "LABLAXIS"      CDF_CHAR    { "Epoch" }
    "UNITS"         CDF_CHAR    { "ms" }
    "FILLVAL"       CDF_DOUBLE  { -1.0e+31 }
    "VAR_TYPE"      CDF_CHAR    { "support_data" }
    "DICT_KEY"      CDF_CHAR    { "time>epoch" }
    "SCALETYP"      CDF_CHAR    { "linear" }
    "MONOTON"       CDF_CHAR    { "INCREASE" }
    "CATDESC"       CDF_CHAR    { "Accumulation interval centered epoch" } .
```

Support data variable

Variable attributes

```
! Variable          Data      Number                    Record   Dimension
! Name              Type      Elements  Dims  Sizes     Variance Variances
! --------          ----      --------  ----  -----     -------- ---------

  "COUNTRATE"       CDF_DOUBLE    1       1     15        T        T

  ! Attribute       Data
  ! Name            Type        Value
  ! --------        ----        -----

    "FIELDNAME"     CDF_CHAR    { "SREM deadtime corrected count rates" }
    "VALIDMIN"      CDF_DOUBLE  { 0.0 }
    "VALIDMAX"      CDF_DOUBLE  { 4.29497e+09 }
    "LABLAXIS"      CDF_CHAR    { "Count rate" }
    "UNITS"         CDF_CHAR    { "1/sec" }
    "FILLVAL"       CDF_DOUBLE  { -1.0e+31 }
    "VAR_TYPE"      CDF_CHAR    { "data" }
    "DICT_KEY"      CDF_CHAR    { "particle_flux>number" }
    "SCALETYP"      CDF_CHAR    { "linear" }
    "CATDESC"       CDF_CHAR    { "SREM deadtime corrected count rates" }
    "DEPEND_0"      CDF_CHAR    { "EPOCH" }
    "DEPEND_1"      CDF_CHAR    { "label_COUNTERS" } .
```

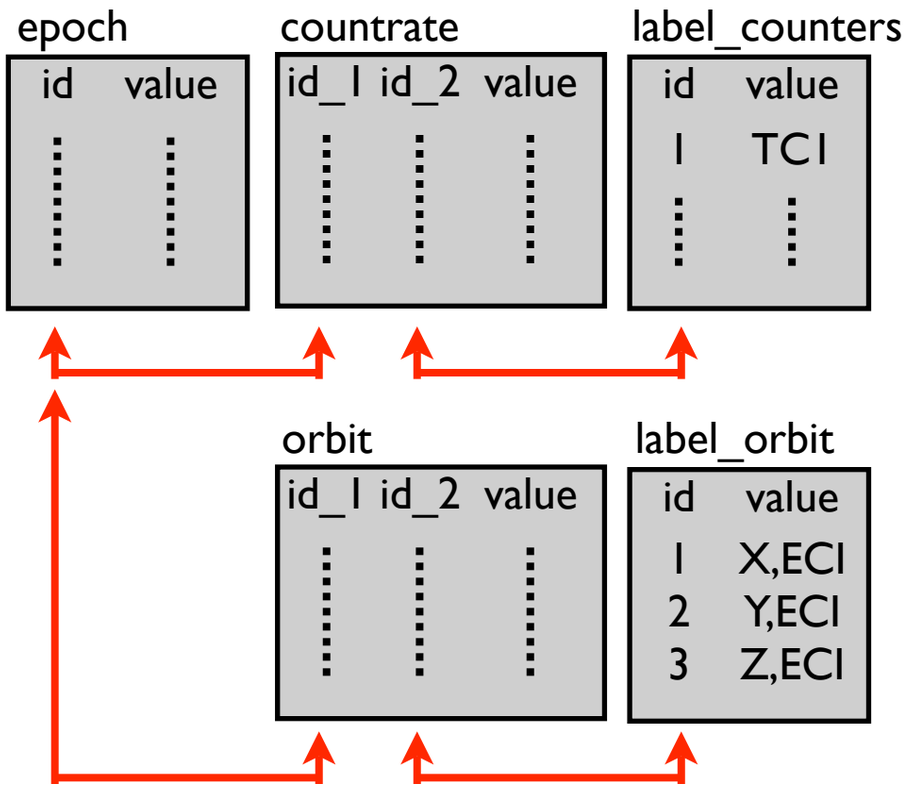
Data variable

Variable attributes

CDF multi-dim variables to SQL tables

Multi-table approach

- Each CDF variable gets one table.
- Relations between tables using SQL keys based on the CDF dependencies.

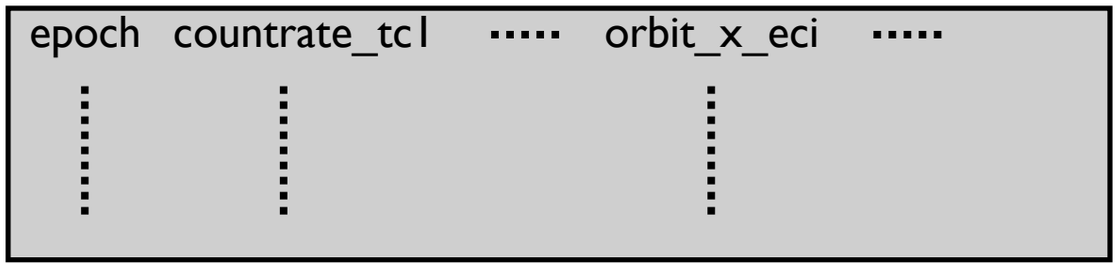


+Can store all types of CDF data sets.

- Performance?

Multi-column approach

- The "data" is stored in one SQL table per CDF data set.
- "Epoch" is a key variable.
- Each CDF "data" variable element gets one column.



+Simple table structure.

+Close resemblance to other data sets in table format.

- Metadata must be constant.

Status

- Code has been developed to automatically generate CDF/ISTP/PRBEM compliant data tables.
- A prototype multi-table approach database has been installed on the ESTEC server.
- The single-table approach is under development.
- A number of SEDAT perl scripts used for basic queries on the datasets are being adapted to use MySQL statements.
- Download scripts for GOES, IMP8 and ACE data have been developed; parser scripts to populate the database columns are under development.
- Tests have been performed with the IDL Dataminer to access the prototype database; IDL will be used for SEDAT and SPENVIS.
- Tests have been performed with Java JDBC to access data; Java is used in SAAPS.