ODB API

Eoin Whelan
Irish Meteorological Service

Outline

- Background, Data & Software
- Practical

Background

- All you need is available here:
  - https://confluence.ecmwf.int/display/ODBAPI
- Data referred to as ODB or ODB-2
- Encoding and processing of obs data
  - SQL filtering and statistics engine
  - Command line tools
  - APIs for C/C++, Fortran and Python

ODB API Training

https://software.ecmwf.int/wiki/display/ODBAPI/

Piotr Kuchta

Peter.Kuchta@ecmwf.int

ECMWF

ODB API data structure / file format

- Tabular data structure, like a table in a relational database
- Line / row oriented
- Each file consists of one or more blocks, each block containing:
  - Header: column names, types, codec info, values of constant columns
  - Data section: packed sequence of rows (customizable, by default 10,000)
  - When reading files only one block needs to be unpacked into RAM at a time
- The above properties allow for
  - Processing files of arbitrary length in constant, small amount of memory
  - Data can be appended to existing files / files can be concatenated

Data format: blocks

- ODB API file consists of one or more blocks
- Each block consists of a header (metadata) and data section

ODB Gov
Repor
16001 - Automat
var
110 - surfac
58 - 2m relat

andate :integer	antime :integer	reportype :integer	date :integer	time :integer	varno :integer	obsvalue :real	....
20170108	120000	16001	20170108	150000	110	101220.00	
20170108	120000	16001	20170108	150000	58	0.992837	
20170108	120000	16001	20170108	150000	110	100110.00	
20170108	120000	16001	20170108	150000	110	100870.00	
20170108	120000	16001	20170108	150000	110	100980.00	
andate :integer	antime :integer	reportype :integer	date :integer	time :integer	varno :integer	obsvalue :real	....
20170108	120000	16001	20170108	150000	110	101230.00	
20170108	120000	16001	20170108	150000	58	0.992852	
20170108	120000	16001	20170108	150000	110	100110.00	
20170108	120000	16001	20170108	150000	110	100860.00	
20170108	120000	16001	20170108	150000	110	100550.00	

Files with heterogeneous metadata

- Over time new columns are added to ODB. When retrieving from MARS data spanning multiple cycles it may happen that part of the retrieved data has more columns than the rest.

andate :integer	antime :integer	reportype :integer	date :integer	time :integer	varno :integer	obsvalue :real
20170108	120000	16001	20170108	150000	110	101230.00
20170108	120000	16001	20170108	150000	58	0.992852
20170108	120000	16001	20170108	150000	110	100110.00
20170108	120000	16001	20170108	150000	110	100860.00
20170108	120000	16001	20170108	150000	110	100550.00

andate :integer	antime :integer	reportype :integer	date :integer	time :integer	varno :integer	obsvalue :real	new_f :bitfie
20170109	120000	16001	20170109	150000	110	101230.00	1
20170109	120000	16001	20170109	150000	58	0.992852	1
20170108	120000	16001	20170108	150000	110	100110.00	0
20170108	120000	16001	20170108	150000	110	100860.00	0
20170108	120000	16001	20170108	150000	110	100550.00	1

Local use

- Observation monitoring
  - Convert ECMA/CCMA using odb_migrator tool
- COPE, b2o, Metview

- ECMWF OFB/MFB feedback data available in MARS

Local use - installation

```
~$ VERSION=0.18.1
~$ wget https://software.ecmwf.int/wiki/download/attachments/61117379/odb_api_bundle-${VERSION}-Source.tar.gz
~$ gunzip odb_api_bundle-${VERSION}-Source.tar.gz
~$ tar -xvf odb_api_bundle-${VERSION}-Source.tar
~$ cd odb_api_bundle-${VERSION}-Source
~$ mkdir build
~$ cd build
~$ cmake .. -DCMAKE_INSTALL_PREFIX=/opt/metapp/odb_api/${VERSION}/gnu \
  -DENABLE_ODB_API_SERVER_SIDE=ON -DENABLE_FORTRAN=ON \
  -DENABLE_GRIB=OFF -DENABLE_ODB_SERVER_TIME_FORMAT_FOUR_DIGITS=ON \
  -DENABLE_PYTHON=ON -DENABLE_ODB=ON -DODB_SCHEMAS="ECMA;CCMA"
~$ make -j 2
~$ ctest
~$ make install
```

Practical

- Let’s have some fun with the odb command line tool …

… and some Irish observation (surface) feedback data from 2002

Preparation (ecgate)

```
#
# on ecgate: add DA Training PATH, environment variables and modules
#
. /home/ms/spsehlam/hlam/daTraining/user_env.sh
```

Get the data

```
#
# on ecgate … you may have a copy of Day_1 already
#
~$ cd $PERM
~$ cp –r /hpc/perm/ms/spsehlam/hlam/daTraining/Day_1 .
~$ cd Day_1
~$ ls
 data    filters
~$ cd data/odb2
```

odb: show headers & metadata

```
#
# odb header
#
~$ odb help
~$ odb help header
~$ odb header canECMAconv200209.odb
```

odb: data digging

https://confluence.ecmwf.int/display/ODBAPI/Examples

```
#
# odb sql
#
~$ odb help
~$ odb help sql
```

odb: data digging

https://confluence.ecmwf.int/display/ODBAPI/Examples

```
#
# odb sql
# List the following:
#   1. unique obstype
#   2. unique codetype
#   3. unique station identifier (statid)
```

Hint 1: use the **SELECT DISTINCT** SQL statement

odb: data digging

https://confluence.ecmwf.int/display/ODBAPI/Examples

```
#
# odb sql
# What and when was the “worst” 2 m temperature observation
```

Hint 1: **SELECT** date, time, statid, fg_depar

Hint 2: use **WHERE**

Hint 3: use **ORDER BY**