

# The TIDB2 Meteo Experience

**Experience with the TIDB2 database  
interface in managing  
meteorological observation and forecast  
data**

João Simões – ECMWF, IM (Portugal)

Maria Monteiro - IM (Portugal)

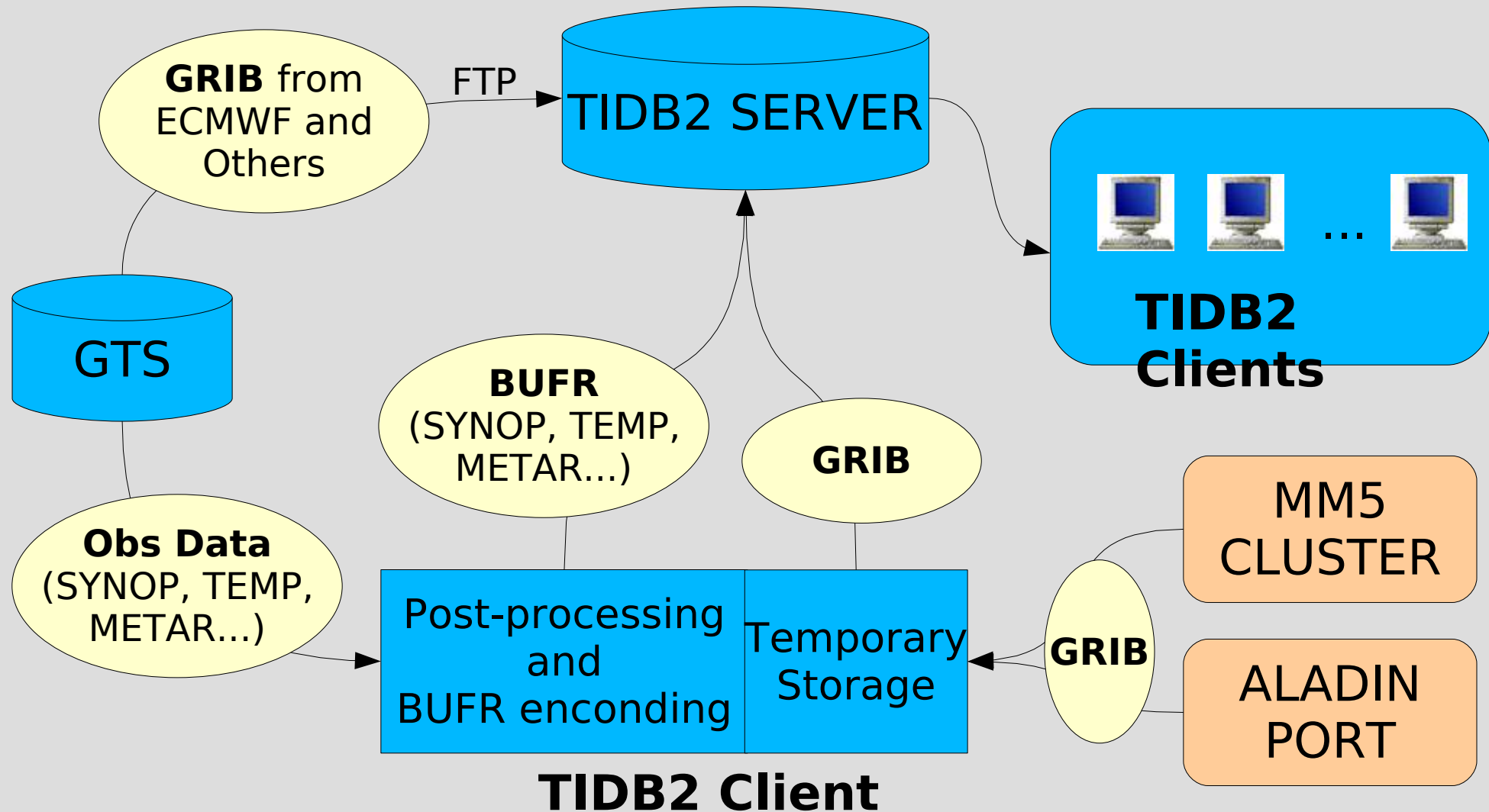
António Amorim - FCUL (Portugal)

**11th ECMWF Workshop  
on Meteorological Operational Systems**  
ECMWF, November 2007

# General TIDB2 characteristics

- **Temporal based system:**
  - all data is stored with temporal information (associated with timestamp);
  - The history is kept for all objects (the objects are always added, never deleted).
- **Objects** (like GRIB and BUFR) **and database connections** (to Oracle, MySQL...) **handled via runtime plugins.**
- Binary data stored with customized “**auto-metadata**”.
- Support for many interfaces on multiple platforms.
- Integrated with PAIPIX Linux distribution.

# The (current) TIDB2 Meteo-Data Flow

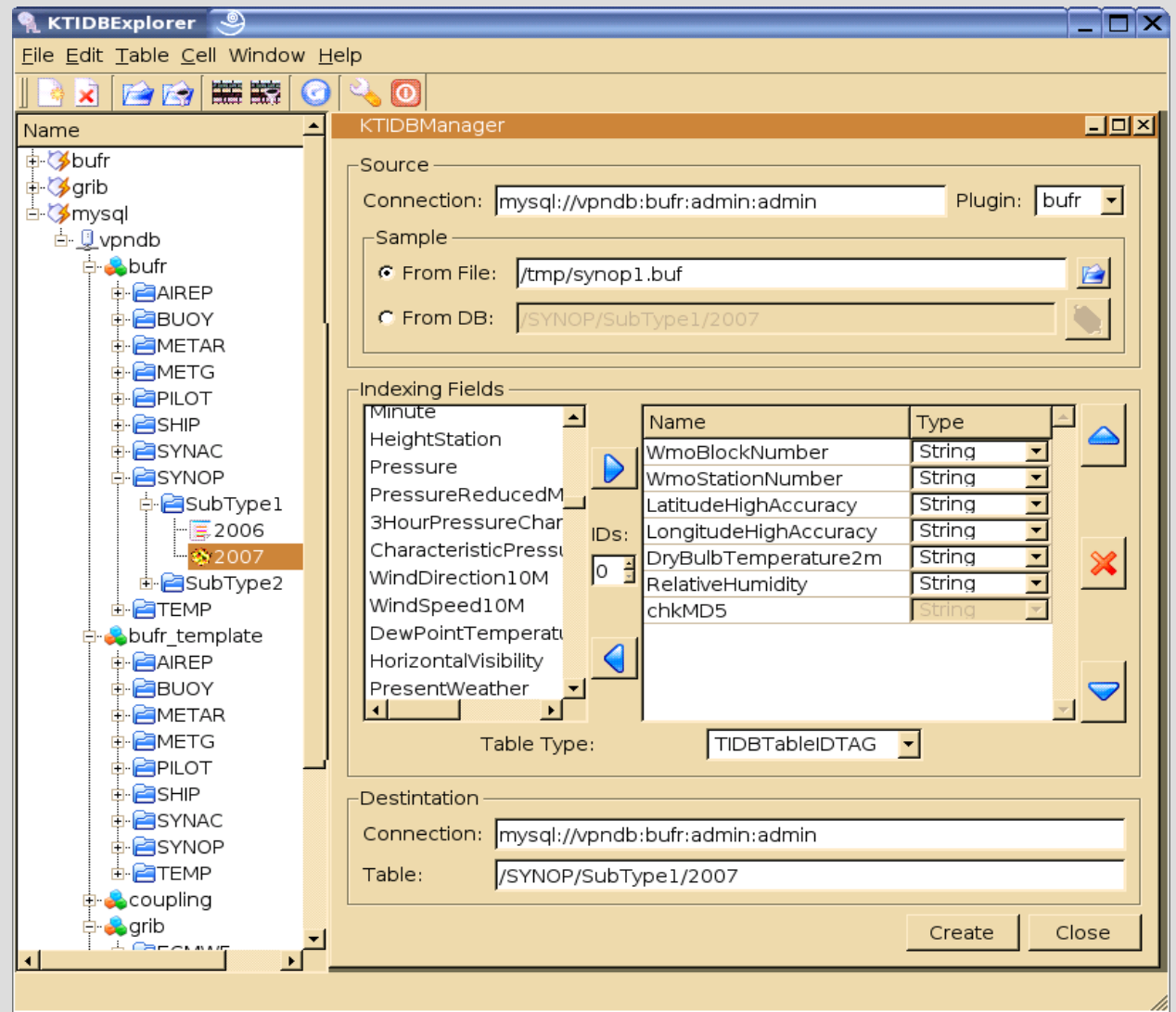


# The TIDB2 Server

- The TIDB2 server is:
  - an **AMD64x2** machine with 4GB of RAM;
  - running **PAIPIX** Linux;
  - using **MySQL** (TIDB2 allows to mix or change the RDBMS server at any time).
  - 1.2 TB of disk and 1TB of online data.
- It is being upgraded to two servers with redundancy:
  - Dual Quad-Core Xeon with 8GB of RAM;
  - 7.5TB RAID disk array each.

# Preparing to Store Meteo-Data

- Before we started to push data into the database we had to create some table infrastructure to store it!
- KTIDBExplorer provides a tool for such operation departing from the BUFR/GRIB metadata.



# Why Templates are Important...

- The structure of the tables has changed a lot since they have been created for the first time:
  - We make use of the schema evolution feature of TIDB2.
  - We store all schema versions as templates.
  - There is the need to recreate tables with a different schema and we want all metadata to be regenerated.
  - The tool provided for such task is “**tidbrefactor**”.

# Creating Templates

- A **Template** is an empty table, similar to the one that will contain data, but contains only the table structure.

## SYNOP1

### Structure

Name	Type
Tag	String
Type	Int
Subtype	Int
WmoBlockNumber	Int
WmoStationNumber	Int
LatitudeHighAccuracy	Double
LongitudeHighAccuracy	Double
HeightStation	Double
DryBulbTemperature2m	Double
DewPointTemperature2m	Double
WindDirection10M	Double
WindSpeed10M	Double
HorizontalVisibility	Double
CloudCoverTotal	Int
PressureReducedMeanSeaL	Double
TotalPrecipitation	Double
PrecipitationPeriod	Int
chkMD5	String

## TEMP

### Structure

Name	Type
Tag	String
Type	Int
Subtype	Int
WmoStationNumber	Int
WmoBlockNumber	Int
Correction1	Int
LatitudeHighAccuracy	Float
LongitudeHighAccuracy	Float
HeightStation	Float
CloudCoverTotal	Float
Pressure	Float
TemperatureDryBulbTempe	Float
DewPointTemperature	Float
WindDirection	Float
WindSpeed	Float
chkMD5	String

## METAR

### Structure

Name	Type
Type	Int
Subtype	Int
IcaoLocationIndicator	String
Correction1	Int
LatitudeCoarseAccuracy	Float
LongitudeCoarseAccuracy	Float
HeightStation	Float
TemperatureDryBulbTempe	Float
DewPointTemperature	Float
WindDirection	Float
WindSpeed	Float
Pressure	Float
chkMD5	String

# of Indexes

# The “Refactor” tool

- TIDB2 provides the tool “**tidbrefactor**” to change the schema of an existing table according to a model.

Use:

```
tidbrefactor <[url]/source_table> <[url]/model_table>
```

Example:

```
tidbrefactor mysql://server:db:user:pass/table1  
mysql://server2:db2:user:pass/table2
```

- This tool works in 3 steps:
  - Modify the data table structure according to a template.
  - Reprocess all **BLOBs** stored with the metadata.
  - Regenerate all metadata automatically.



# Some real “Refactor” use cases

- The “refactor” has been used as a maintenance task to:
  - Add/Remove indexing/metadata columns.
  - Change column names.
  - Change data types.
  - Change the table type.
  - Regenerate corrupted metadata due to a BUFR/GRIB decoding failure (ex.: missing tables, unsupported BUFR/GRIB format...).
- A slightly modified version of “**tidbrefactor**”, the “**tidbtrans**”, has been used to copy several GB of data between different databases and servers and RDBMS.

# Introducing “TAGs” in TIDB2 #1

- We faced the problem that data could be being inserted multiple times with slightly modifications/corrections.

	Since	Tag	Subtype	WmoBlock	WmoStation	Latitude	Longitude	Height	DryBulbTem	DewPointTem	WindDire	WindSpeed	Horizontal
1	2007-11-09/00:00	H0	1	8	524	33.07	-16.35	82.0	293.4	292.0	140.0	5.0	80
2	2007-11-09/01:00	H0	1	8	524	33.07	-16.35	82.0	293.1	291.8	140.0	5.0	80
3	2007-11-09/02:00	H0	1	8	524	33.07	-16.35	82.0	293.1	291.6	130.0	5.0	80
4	2007-11-09/03:00	0	1	8	524	33.07	-16.35	82.0	293.1	291.7	130.0	6.0	<missing
5	2007-11-09/03:00	H1	1	8	524	33.07	-16.35	82.0	293.1	291.7	130.0	6.0	80
6	2007-11-09/04:00	H0	1	8	524	33.07	-16.35	82.0	293.2	291.9	130.0	5.0	<missing
7	2007-11-09/05:00	H0	1	8	524	33.07	-16.35	82.0	293.2	292.0	120.0	6.0	<missing
8	2007-11-09/06:00	H0	1	8	524	33.07	-16.35	82.0	293.3	292.1	150.0	6.0	80
9	2007-11-09/08:00	H0	1	8	524	33.07	-16.35	82.0	293.2	291.9	120.0	6.0	90
10	2007-11-09/09:00	H0	1	8	524	33.07	-16.35	82.0	293.5	292.0	120.0	6.0	90
11	2007-11-09/10:00	H0	1	8	524	33.07	-16.35	82.0	293.7	292.0	130.0	6.0	90
12	2007-11-09/12:00	0	1	8	524	33.07	-16.35	82.0	294.0	292.4	130.0	7.0	90
13	2007-11-09/12:00	1	1	8	524	33.07	-16.35	82.0	294.0	292.4	130.0	7.0	90
14	2007-11-09/12:00	2	1	8	524	33.07	-16.35	82.0	294.0	292.4	130.0	7.0	90
15	2007-11-09/12:00	H3	1	8	524	33.07	-16.35	82.0	294.0	292.4	130.0	7.0	90
16	2007-11-09/13:00	H0	1	8	524	33.07	-16.35	82.0	294.0	292.5	160.0	6.0	80
17	2007-11-09/14:00	H0	1	8	524	33.07	-16.35	82.0	293.9	292.4	160.0	5.0	80
18	2007-11-09/15:00	H0	1	8	524	33.07	-16.35	82.0	294.6	291.6	140.0	4.0	90
19	2007-11-09/16:00	H0	1	8	524	33.07	-16.35	82.0	294.6	292.1	150.0	4.0	90
20	2007-11-09/17:00	H0	1	8	524	33.07	-16.35	82.0	294.2	291.5	150.0	3.0	90

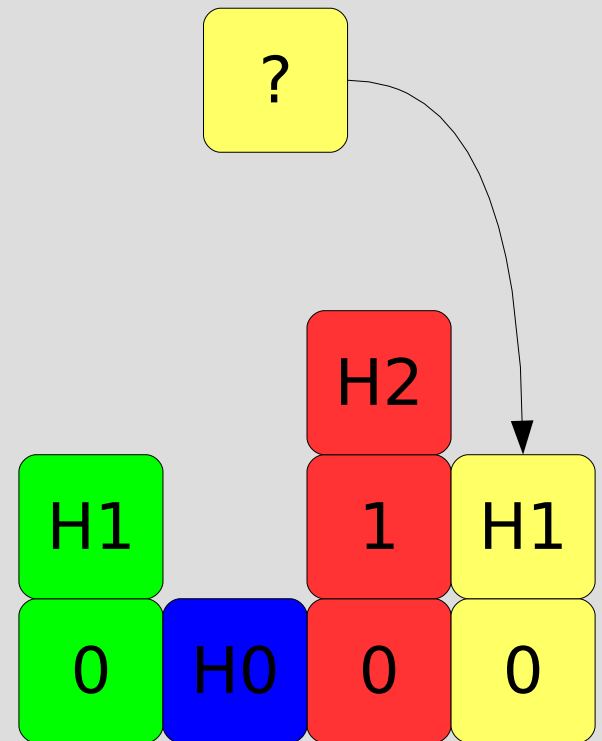
# Introducing “TAGs” in TIDB2

## #2

- **TIDB2** has a feature to not allow storing multiple times the very same data, **but...**
- It happens to have similar data stored multiple times on the database:
  - The data from the GTS is sent multiple times to the post processor, reprocessed and sent to the database.
  - There is a data correction and last data should be replaced.
  - The new data is more complete and should replace the last one.
  - Also makes easy to clean up the earlier versions of objects in a maintenance task

# How “TAGs” work?

- The just arrived objects are tagged as “H0”
- If there was already a similar object on the database it is tagged as “H1”, the existing object is retagged as “0” instead of “H0”.
- To get the last version of all objects we just need to grab the “H\*” objects.



- An object is called “similar” if it shares the same indexing information.

# Viewing BUFR data

- Looking at METAR BUFRs on the TIDB2 Server.

The screenshot shows the KTIDBExplorer application window. The main window displays a table titled "/METAR/2007::TABLE" with columns: Since, Subty, IcaoLoc, Corre, Latitud, Longitud, Height, Tempe, DewPo, WindD, WindS, Pressure, and bufrObj. The table contains 26 rows of data for the year 2007. A pop-up window titled "bufrObj(1)" is open, showing the structure of a BUFR object. It includes a "Times" section with "Since" and "Till" fields, a "Header" section with a table of Name and Value, and a "Data" section with a table of Name and Value.

	Since	Subty	IcaoLoc	Corre	Latitud	Longitud	Height	Tempe	DewPo	WindD	WindS	Pressure	bufrObj
1	2007-11-12/20:59.00	140	PTRO	0	7.33	134.47	33.0	300.2	298.2	210.0	3.1	100250.0	<BUFR C
2	2007-11-12/21:00.00	140	DTTA	0	36.83	10.23	3.0	289.2	286.2	240.0	3.1	101660.0	<BUFR C
3	2007-11-12/21:00.00	140	OSDI							0.0	1.0	94960.0	<BUFR C
4	2007-11-12/21:00.00	140	OSAP							0.0	3.1	97340.0	<BUFR C
5	2007-11-12/21:00.00	140	OSLK							0.0	2.1	101820.0	<BUFR C
6	2007-11-12/21:00.00	140	OSDI							0.0	1.0	94960.0	<BUFR C
7	2007-11-12/21:00.00	140	OSAP							0.0	3.1	97340.0	<BUFR C
8	2007-11-12/21:00.00	140	OSLK							0.0	2.1	101820.0	<BUFR C
9	2007-11-12/21:00.00	140	OSDI							0.0	1.0	94960.0	<BUFR C
10	2007-11-12/21:00.00	140	OSAP							0.0	3.1	97340.0	<BUFR C
11	2007-11-12/21:00.00	140	OSLK							0.0	2.1	101820.0	<BUFR C
12	2007-11-12/21:00.00	140	LPLA							0.0	3.1	100880.0	<BUFR C
13	2007-11-12/21:00.00	140	OYAA							0.0	2.6	101260.0	<BUFR C
14	2007-11-12/21:00.00	140	OYSQ							0.0	2.1	100460.0	<BUFR C
15	2007-11-12/21:00.00	140	OJAI							0.0	0.0	93490.0	<BUFR C
16	2007-11-12/21:00.00	140	OJAM							0.0	0.0	92980.0	<BUFR C
17	2007-11-12/21:00.00	140	OJAI							0.0	0.0	93490.0	<BUFR C
18	2007-11-12/21:00.00	140	OJAM							0.0	0.0	92980.0	<BUFR C
19	2007-11-12/21:00.00	140	OJAI							0.0	0.0	93490.0	<BUFR C
20	2007-11-12/21:00.00	140	OJAM							0.0	0.0	92980.0	<BUFR C
21	2007-11-12/21:00.00	140	OJAI							0.0	0.0	93490.0	<BUFR C
22	2007-11-12/21:00.00	140	OJAM							0.0	0.0	92980.0	<BUFR C
23	2007-11-12/21:00.00	140	SBBR							0.0	1.5	89280.0	<BUFR C
24	2007-11-12/21:00.00	140	SBCG							0.0	3.6	94480.0	<BUFR C
25	2007-11-12/21:00.00	140	SBPA							0.0	2.1	101660.0	<BUFR C
26	2007-11-12/21:00.00	140	SRBR							0.0	3.1	90310.0	<BUFR C

**bufrObj(1) Times**

Since	Till
2007-11-12/20:59.59	2007-11-12/20:59.59

**bufrObj(1) Header**

Name	Value
8	CentreID 98
9	Update 1
10	Flag 128

**bufrObj(1) Data**

Name	Value
1	IcaoLocationIndicator PTRO
2	TypeStation 1
3	Year 2007
4	Month 11
5	Day 12
6	Hour 20
7	Minute 59
8	LatitudeCoarseAccuracy 7.33
9	LongitudeCoarseAccuracy 134.47

# Viewing GRIB data

- Previewing a 2m Temperature **GRIB** on the TIDB2 Server.

The screenshot shows the KTIDBExplorer application interface. The main window displays a table with columns: Since, gribT, indica, indica, level, perio, latitud, iDirec, longiti, jDirec, latitud, longitu, and gribObj. The table contains 14 rows of data for the year 2007. A preview window titled 'gribObj(119)' is open, showing a 'Times' section with 'Since: 2007-11-11/00' and 'Till: 2007-11-12/22' circled in red. Below this is a 'Data' section with a table of values for columns 1, 2, and 3. A 'Header' section is also visible, listing various metadata fields like editionNumber, gribTablesVersio, etc. A red arrow points from the 'Since' and 'Till' fields to a text box on the right.

Since	gribT	indica	indica	level	perio	latitud	iDirec	longiti	jDirec	latitud	longitu	gribObj
2007-11-11/00:00	128	167	1	0	46	44.84	0.12	345.2	0.12	34.76	356.6	<GRIB DATA>
2007-11-11/00:00	128	170	11									
2007-11-11/00:00	128	171	11									
2007-11-11/00:00	128	172										
2007-11-11/00:00	128	184	10									
2007-11-11/00:00	128	185										
2007-11-11/00:00	128	186										
2007-11-11/00:00	128	187										
2007-11-11/00:00	128	188										
2007-11-11/00:00	128	201										
2007-11-11/00:00	128	202										
2007-11-11/00:00	128	239										
2007-11-11/00:00	128	240										
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									
2007-11-11/12:00	128	3	10									

1	2	3
44.84	345.2	289.37
44.84	345.32	289.3
44.84	345.44	289.32
44.84	345.56	289.31
44.84	345.68	289.31
44.84	345.8	289.30
44.84	345.92	289.29
44.84	346.04	289.27
44.84	346.16	289.26
44.84	346.28	289.23
44.84	346.4	289.20
44.84	346.52	289.17
44.84	346.64	289.14
44.84	346.76	289.11
44.84	346.88	289.08
44.84	347	289.05

- Time interval is set to: [TOR, TOR+Step]  
- TOR is the "time of run".

# The TIDB2 Interfaces

- **C++** - is the native TIDB2 interface, fast, fully featured and easy to use.
- **C/Fortran** – it was very useful to migrate the legacy applications.
- **Shell tools** – very suitable for integration with other general purpose systems, php web scripts, crontab like jobs, shell scripts...



# Example of a Migration of a Legacy Application

- This is the example of a very old application migrated from a VAX system, using the **fortran interface!**
- This application takes as input a fortran namelist and retrieves the correspondent observation from database.

```
$OBSOP
      lblock=.f.,
      lident=.t.,
      ident=07149,
      idate=20071002,
      larea=.f.,
      carea='global',
      ctime='0200/TO/1200',
      lctime=.t.,
      cobstype='s',
      lshow_bufnr=.t.

$END
```

```
pmos@cirrus 22:43:54 ~/suite_gribsEbufrs/bufrwork> ./lista_obs
[/SYNOP/SubType1/2007::TableIDTAG]
```

since_t	till_t	Tag Type Subtype WmoBlockNumber WmoStationNumber LatitudeHighAccuracy LongitudeHighAccuracy HeightStation DryBulbTemperature2m
KeyOld	KeyOld	String Int  Int  Int  Int  Double  Double  Double  Double
2007-10-02/03:00.00 2007-10-02/03:00.00 H0   0  1  7  149  48.72  2.38  90.0  289.9		
2007-10-02/06:00.00 2007-10-02/06:00.00 H0   0  3  7  149  48.72  2.38  90.0  289.1		
2007-10-02/09:00.00 2007-10-02/09:00.00 H1   0  1  7  149  48.72  2.38  90.0  290.1		
2007-10-02/12:00.00 2007-10-02/12:00.00 H0   0  1  7  149  48.72  2.38  90.0  293.3		

```
[4] rows
pmos@cirrus 22:43:54 ~/suite_gribsEbufrs/bufrwork>
```



# The TIDB2 GRIB-WebClient

- This client is a *php* web page, using the **shell tools**.

The screenshot displays the TIDB2 Web Client interface within a browser window. The browser's address bar shows the URL: `http://lismeteo.org:8003/tidb2/cliente/`. The page title is "TIDB2 Web Client".

The interface includes a form with the following fields and values:

- Year: 2007
- Model: ALADIN
- Run: 2007-11-12/12:00.00
- Parameter: [167] 2 metre temperature
- Level: 0
- H+: any
- Latitude: 44.84
- Step: 0.12

Below the form, there are additional settings:

- Area: Portugal
- Draw mode: Temperature
- Limit to: 40

A "Search database" button is located below these settings. The status bar at the bottom of the form area displays "Maps loaded, finished."


The main content area features a weather map of Portugal. The map is color-coded by temperature, with a color scale on the right ranging from -20 (dark blue) to 24.43 (yellow). The map shows a temperature gradient across the region, with higher temperatures (yellow) in the south and lower temperatures (green) in the north. The map is titled "2ª-Feira 12 Novembro 2007 12H VAR Previsão H+26 para 3ª-Feira 13 Novembro 2007 14h Temperatura a 2 metros".

At the bottom of the map area, there are navigation controls and a label "arib #27/40".

# Integration with SIMDAT VGISC #1

- SIMDAT Virtual Global Information Centre provides a shell scripting interface for data retrieval.
- We used TIDB2 *Shell Tools* interface for integration:
  - Standard unix command line tools to convert the *request* into time intervals and SQL query.
  - **tidbgetobject** to get the *bufr* data from the database;
  - **tidbviewobject** to view retrieved *bufrs* as *HTML*.
- It was a very simple task, took only a day to get the first working dataset!

# Integration with SIMDAT VGISC #2

**Simdat-VGISC** 

You are logged in as user [lpmg](#). [Logout](#) [Home](#) [Search](#) [Directory](#) [Requests](#)

**Welcome**  
... to the SIMDAT portal running at the European Centre for Medium-Range Weather Forecasts

**Home > Search Results >**  
Search results for **BUFR** 1 of 1 (0 seconds)

[BUFR SYNOP data stored in LPMG.](#)

- [LPMG > BUFR > SYNOP >](#)

This dataset contains SYNOP **BUFR** type1 data from portuguese and international observations.

**Metadata**

[Retrieve Data](#)

[Show XML](#)

This dataset is freely available.

**Home > Metadata >**

**Title:** *BUFR SYNOP data stored in LPMG.*

**Abstract:** This dataset contains SYNOP BUFR type1 data from portuguese and international observations.

**Period:** 2007-01-01 to 2007-12-31

**Bounding Box:** 83,6°N 166,9°W 73°S 139,4°E



**Home > Retrieve data >**

**Date**  
Please enter a date between 2007-01-01 and 2007-12-31  
2007-11-07 2007-11-08

**Area**

Output format  
bufr html

**Home > Your requests >**

**BUFR SYNOP data stored in LPMG. .**  
Date from 2007-11-07 to 2007-11-08. Area: (39.52;-10.98)(37.9;-7.28). Output format: html.

**Status:** COMPLETE **Started:** 2007-11-07 11:44:11 **Last update:** Right now

[Download results \(428.4 K\)](#)  
[details](#) [xml](#) [delete](#)

WmoBlockNumber	WmoStationNumber	TypeStation	Year	Month	Day	Hour	Minute	LatitudeHighAccuracy	LongitudeHighAccuracy	HeightStation	Pressure	PressureReducedMeanSeaLevel
8	532	1	2007	11	7	0	0	38.83	-9.33	130	100900	102560
8	558	1	2007	11	7	0	0	38.53	-7.9	246	99570	102530
8	579	1	2007	11	7	0	0	38.77	-9.13	105	101230	102470
8	562	1	2007	11	7	0	0	38.02	-7.87	247	99540	102470
8	535	1	2007	11	7	0	0	38.72	-9.15	95	101580	102500
8	571	1	2007	11	7	0	0	39.28	-7.42	590	95700	102670

# The Flexible TIDB2 Shell

## Tools #1

### Data Tools

- **tidbtableput** – store non BLOB data in the database.
- **putobject** – store an object into a specified table in the database.
- **tidbgetobject** - grab selected objects from specified table(s) in the database, store them as a collection of objects on a file at the local filesystem.
- **tidbviewobject** – use TIDB2 object plugins to view a local file (like a GRIB or BUFR collection) either in **txt** or *HTML* format.

# The Flexible TIDB2 Shell

## Tools #2

- **tidbtabledump** – dump the selected contents of specified table(s) in the database.
- **tidbdate2key** – convert a regular time expression into a TIDB2 key (used for indexing data).

## Management Tools

- **tidbrefactor** – alter the schema of tables.
- **tidbtrans** – copy a table to another database or server/RDBMS.
- **tidbtabledrop** – remove a table from the database.

# Getting help and downloading TIDB2

- A good documentation about **TIDB2** history, installation and API documentation could be obtain from:

[http://www.sim.fc.ul.pt/sim\\_en/Tidb2](http://www.sim.fc.ul.pt/sim_en/Tidb2)

You are always welcome to contact the developers!

- The last version of **TIDB2** can be downloaded from:

<http://isscvcs.cern.ch/cgi-bin/viewcvs-all.cgi/tidb2.tar.gz?root=atlastdaq&view=tar>

- Try the **PAIPIX** Linux distribution, with **TIDB2** and a lot of tools already configured and ready to run!

<http://www.paipix.org>

# The END

- **Thanks to...**
  - ECMWF
  - IM, Portugal
  - All of you!