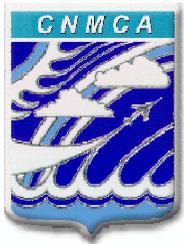

DISTRIC.O.

An operational system for post-processing and distribution of numerical model outputs



Enrico Fucile

Italian Meteorological Service

Centro **N**azionale di **M**eteorologia e
Climatologia **A**eronautica

Overview

- System requirements
- Overall System description
- Relational Database
- PHP module
- Activities
 - Loading
 - Distribution
 - Elaboration
- Web Interface
- Future development
- Conclusions



DISTRIC.O.

DISTRibuzione **C**atena **O**perativa

**OPERATIONAL SUITE
DISTRIBUTION**

Operational at CNMCA since 2001



DISTRICO – Early Requirements

- **Distribution** of model outputs in GRIB format
- **Transformation** of GRIB fields from one grid to another (e.g. from a rotated lat-lon grid to a geographical lat-lon grid)
- Processing a **large number** of GRIB fields coming from **several models**
- Graphical **real time monitoring** of activities (loading, processing and distribution)

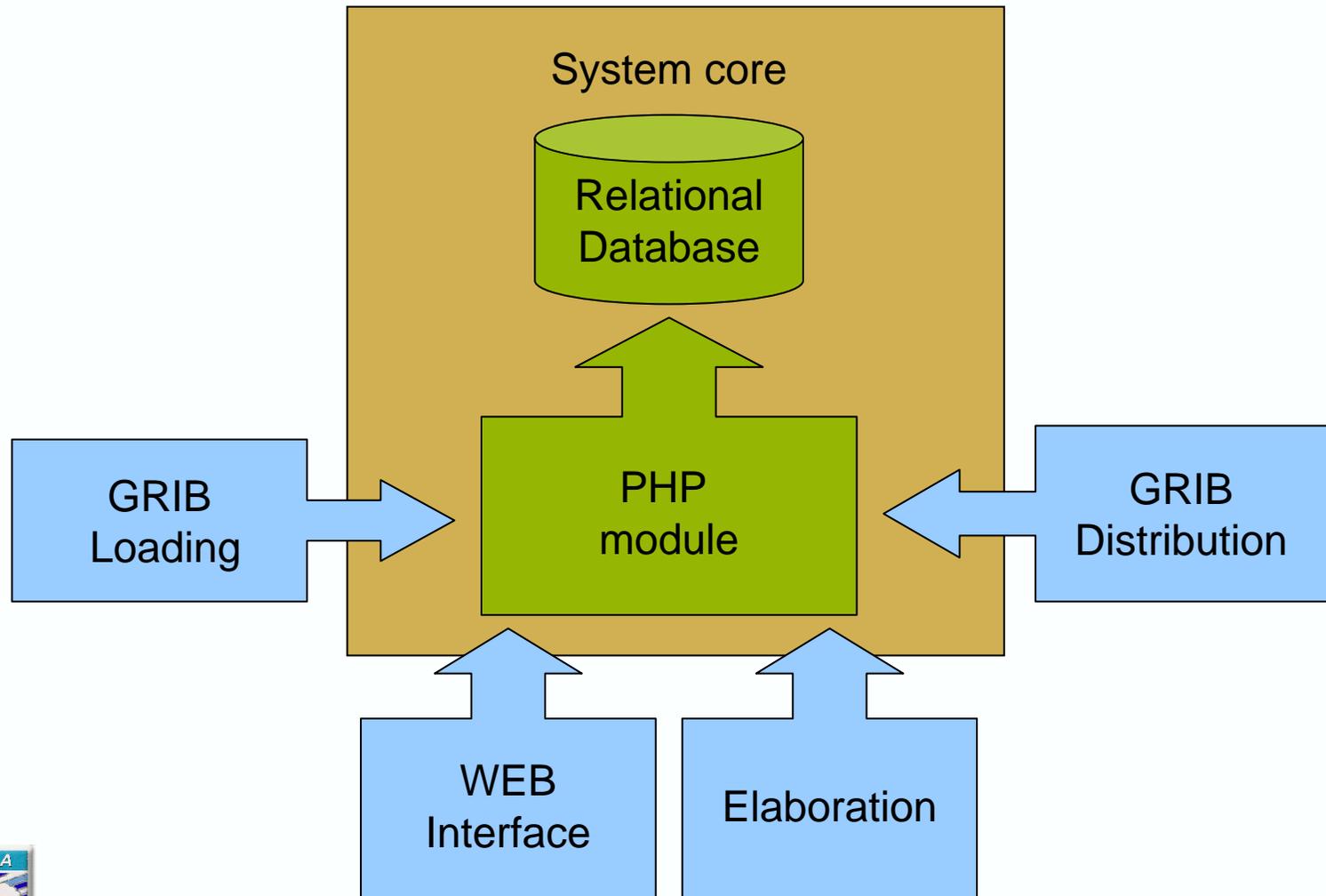


DISTRICO – Added Requirements

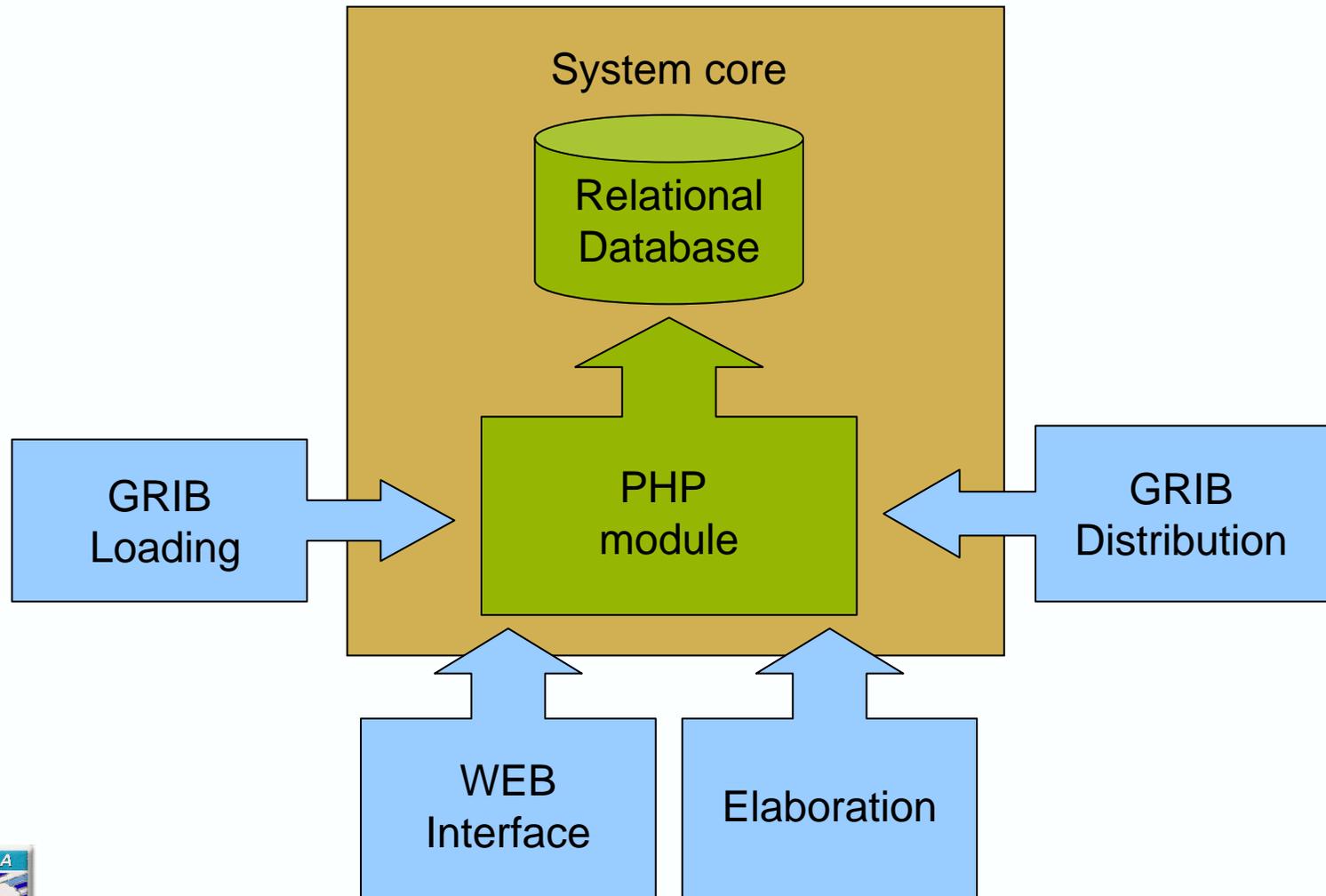
- Easy access to system facilities through a **scripting language (PHP)**
- A data model designed for post-processing not just for grid interpolation (e.g. plotting)
- **Friendly graphical interface providing management tools**
- **Extraction of single point data from a GRIB (meteograms)**



DISTRICO Schema



DISTRICO – Relational Database



DISTRICO – Why a Relational Database?

- Central repository accessed through a client-server application
- Large number of inhomogeneous fields coming from different models, runs, areas, at different levels... **A good structure of data needed.**
- Access provided by an high level language as **Structured Query Language**



DISTRICO – Relational Database

- GRIB stored into **B**inary **L**arge **O**Bject (BLOB) fields
- GRIB identified with a **serial ID** and the **reference date**
- Elaboration tables refer to grib through ID
- IBM Informix - **R**elational **D**ata**B**ase **M**anagement **S**ystem



DISTRICO – Relational Database

Almost static. It changes only when new unregistered fields enter into the system

REGISTERED GRIB							
id_reg_grib	id_centre	id_process	id_grid	id_table	id_param	run	...
1001	80	103	23	2	11	0	...
1002	80	107	12	2	2	0	...
1003	98	203	34	2	1	12	...
...

GRIB		
id_reg_grib	ref_date	grib_object
1002	2005-11-02	BLOB
1002	2005-11-03	BLOB
1003	2005-11-02	BLOB
1003	2005-11-03	BLOB
...

It grows daily if older data are not deleted.

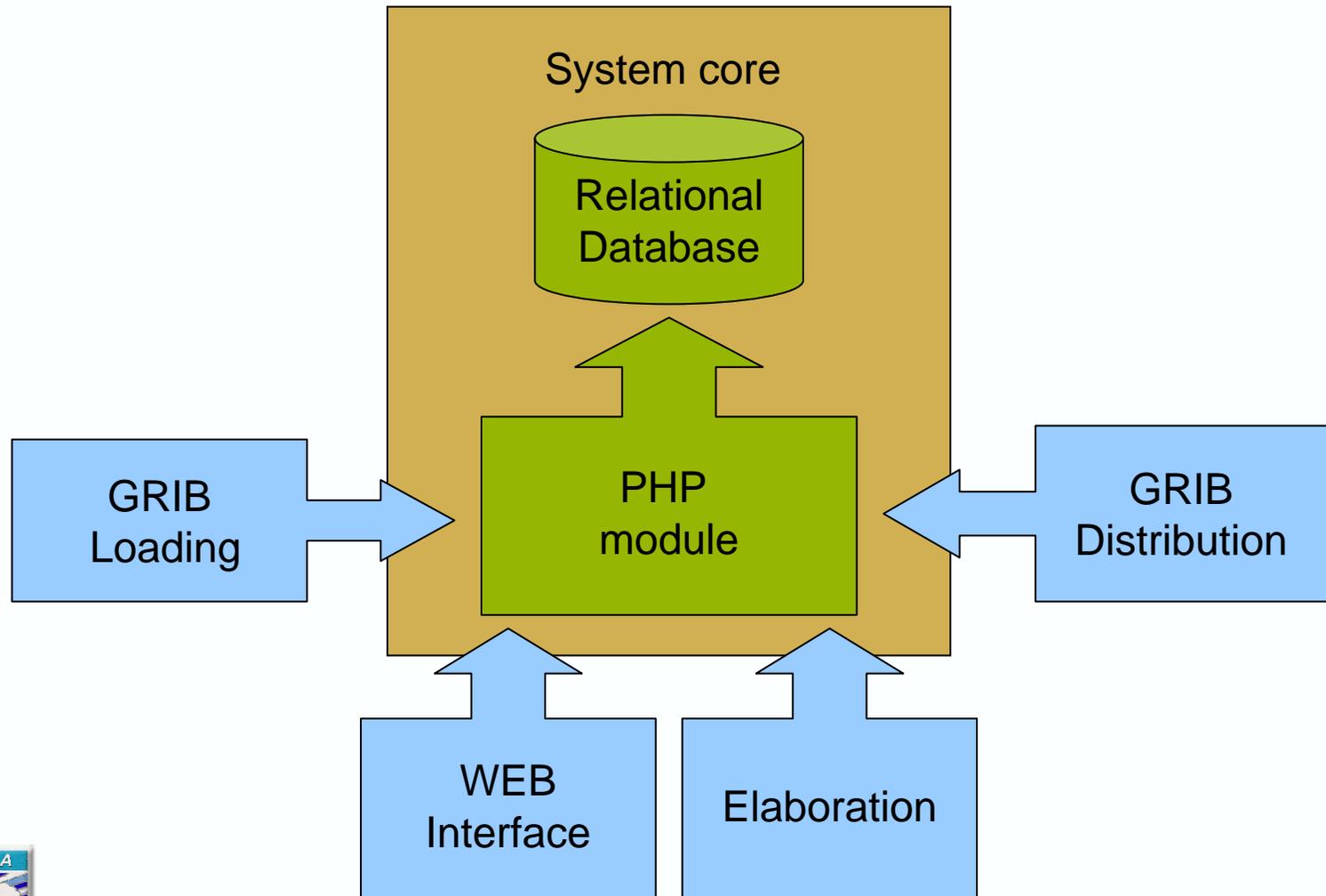


DISTRICO – Relational Database

- **90000** records (**7 GB**) in the **GRIB** table each day
- Data **older than 10 days** are **deleted**
- **Deterministic Models:**
 - ECMWF Global and WAVE (run 0 and 12)
 - EUROHRM (CNMCA) (run 0 and 12)
 - EUROLM (ECMWF) (run 0)
 - HRM (run 0 and 12)
 - HRM on operations areas
 - LAMI (CINECA) (run 0 and 12)
- 3D-VAR analysis (CNMCA)
- **EPS** fields are **not yet inserted** into the DB (not enough bandwidth to download fields)



DISTRICO – PHP module

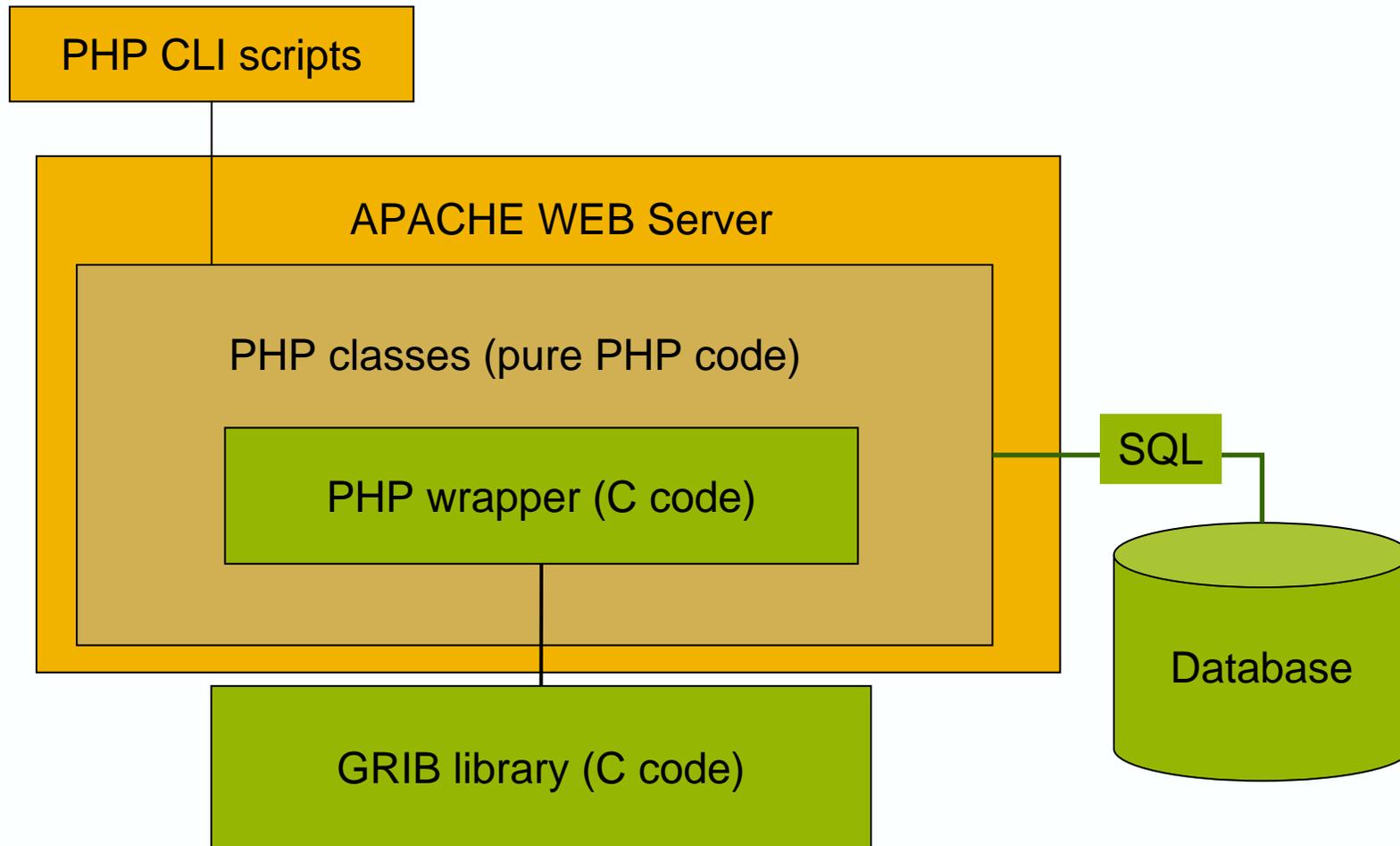


DISTRICO – Why PHP?

- A lot of **extensions** available (gd, xml, network, DB connectivity modules...)
- **Easy to extend** with compiled C and Fortran code (easy **reuse of existing code** and easy extension with high performance routines for intensive calculations)
- Used both as **command line interpreter (CLI)** and embedded in a **WEB server**
- **Object based syntax** (good code maintenance)



DISTRICO – PHP module



DISTRICO – PHP module

GET GRIB AND DECODE

```
$grib = new GRIB;
$grib->With("id_center",80);
$grib->With("id_model",2);
$grib->With("ref-date","2005-11-16");
$grib->With("run",12);
$grib->With("id_parameter",11);
...
$fh=fopen("filename","r"); // optional file handler if not
                             // given default database is used
while ($grib->GetNext( optional file handler $fh )) {
    $grib->Decode();
    $grib->PrintPDS(); //PDS also available as $grib->pds[parm]
    $grib->PrintGDS(); //GDS also available as $grib->gds[parm]
    $grib->Unpack();
    for ($i=0; $i < ( $grib->gds[ni] * $grib->gds[nj] ); $i++)
        print $grib->field[$i];
}
```



DISTRICO – PHP module

INTERPOLATION

```
...  
// NEW GRID Object  
$grid = new GRID;  
$grid->Set("ni",161);  
$grid->Set("nj",241);  
$grid->Set("lat1",60000);  
$grid->Set("lon1",-30000);  
$grid->Set("lat2",20000);  
$grid->Set("lon2",30000);  
...  
while ($grib->GetNext()) {  
    $grib->SetNewGrid($grid);  
    $grib->DBPut();  
}
```

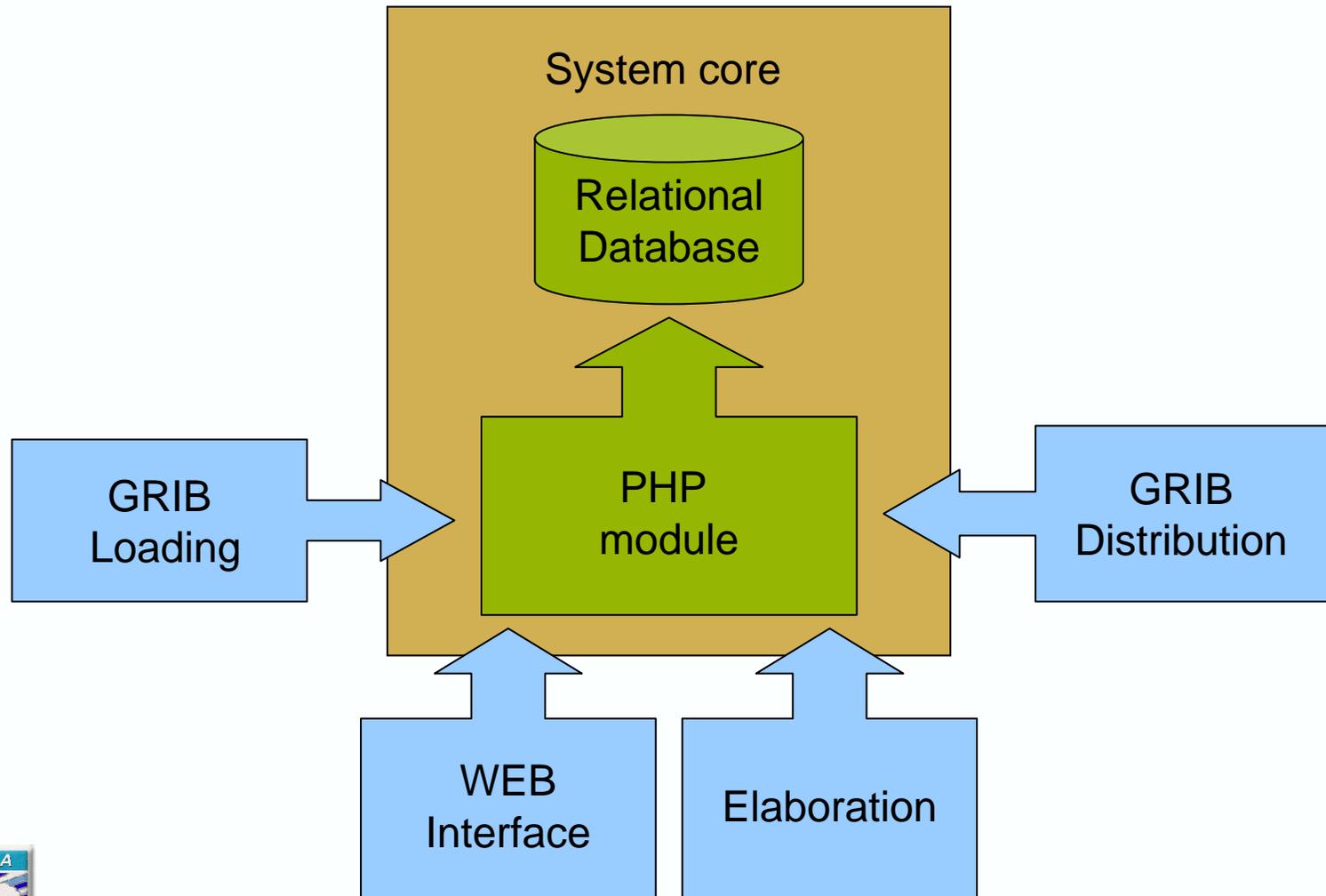


DISTRICO – PHP module

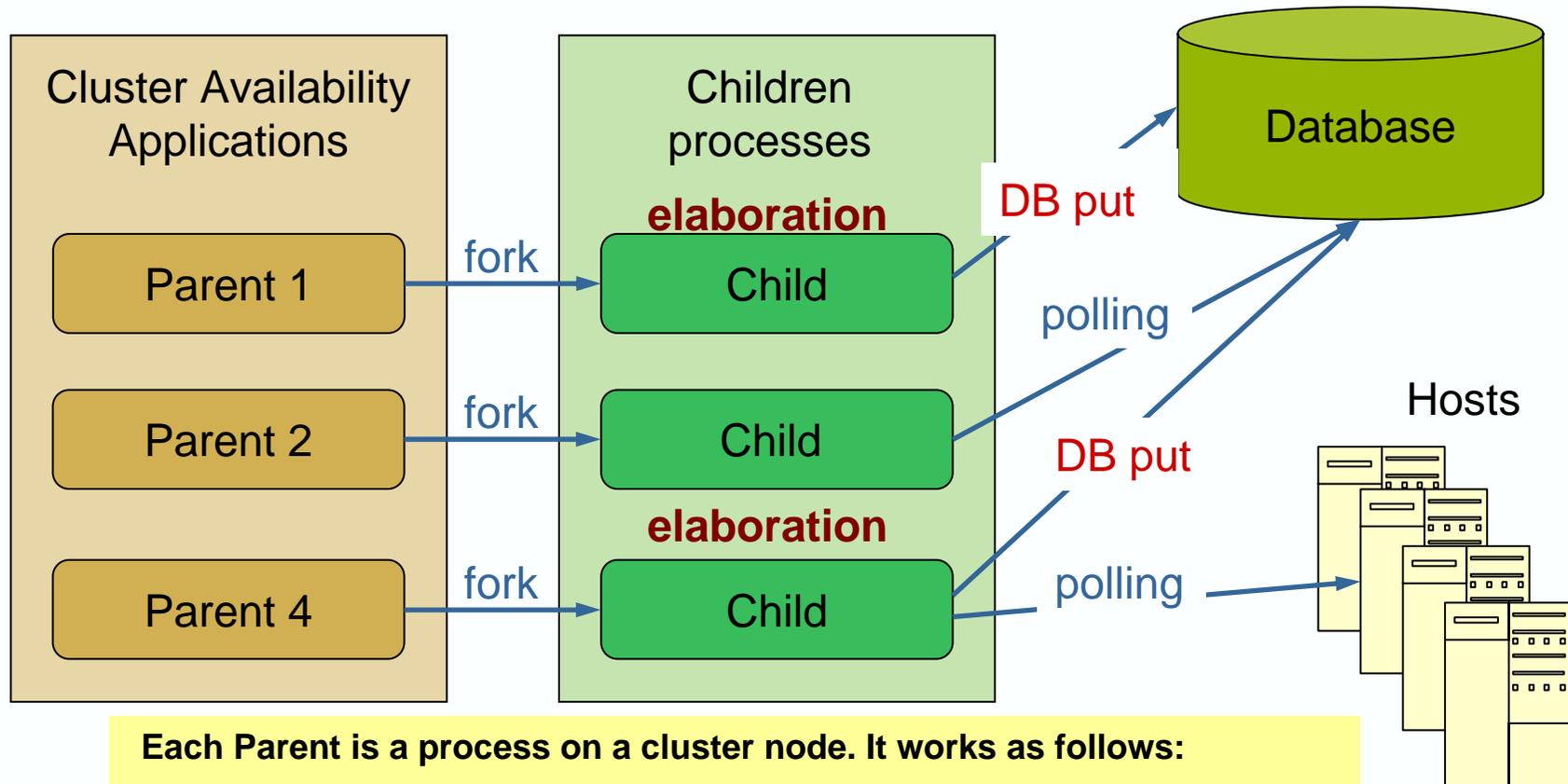
```
...
while ($grib->GetNext()) {
    $lat=40000; // latitude in millidegrees
    $lon=11000; // longitude in millidegrees
    $NearestPoints=$grib->FindNearest($lat,$lon);
    foreach ($NearestPoints as $Point) {
        print $Point->latitude;
        print $Point->longitude;
        print $grib->field[$Point->index];
    }
}
}
```



DISTRICO – Activities model



DISTRICO – Activities model

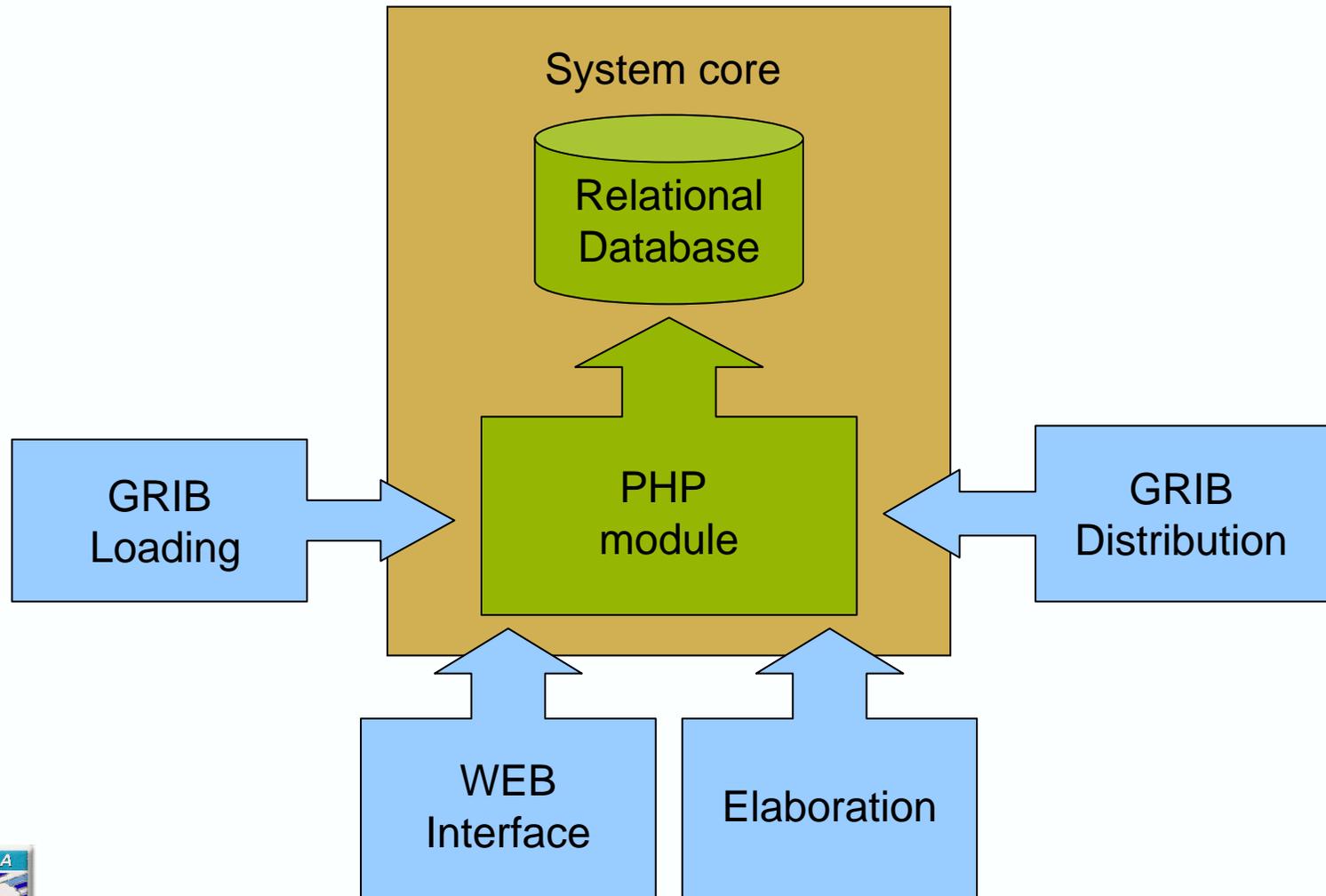


Each Parent is a process on a cluster node. It works as follows:

1. Fork a child
2. Wait for child to exit
3. Sleep for a configurable amount of seconds
4. Restart from 1.



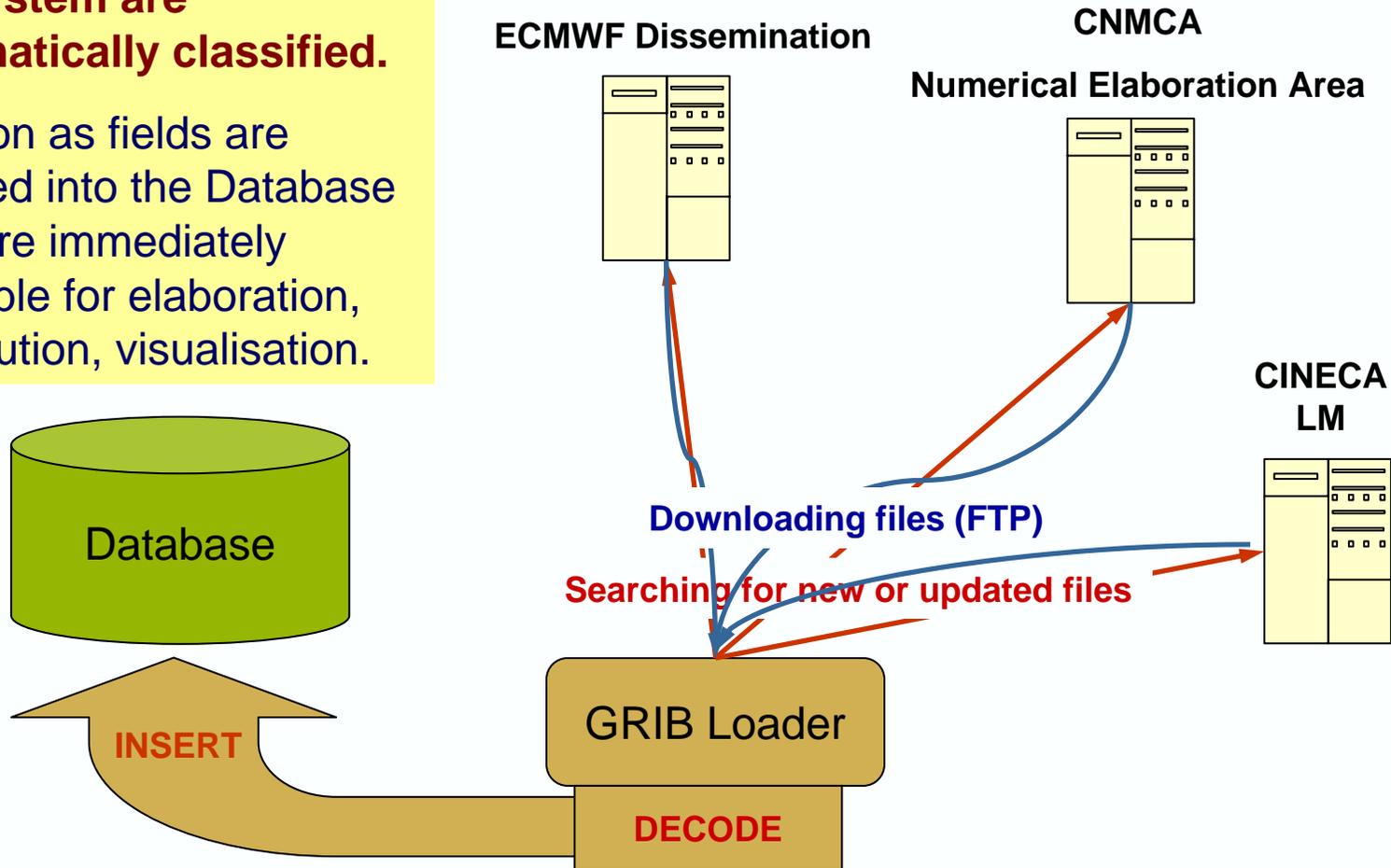
DISTRICO – GRIB Loading



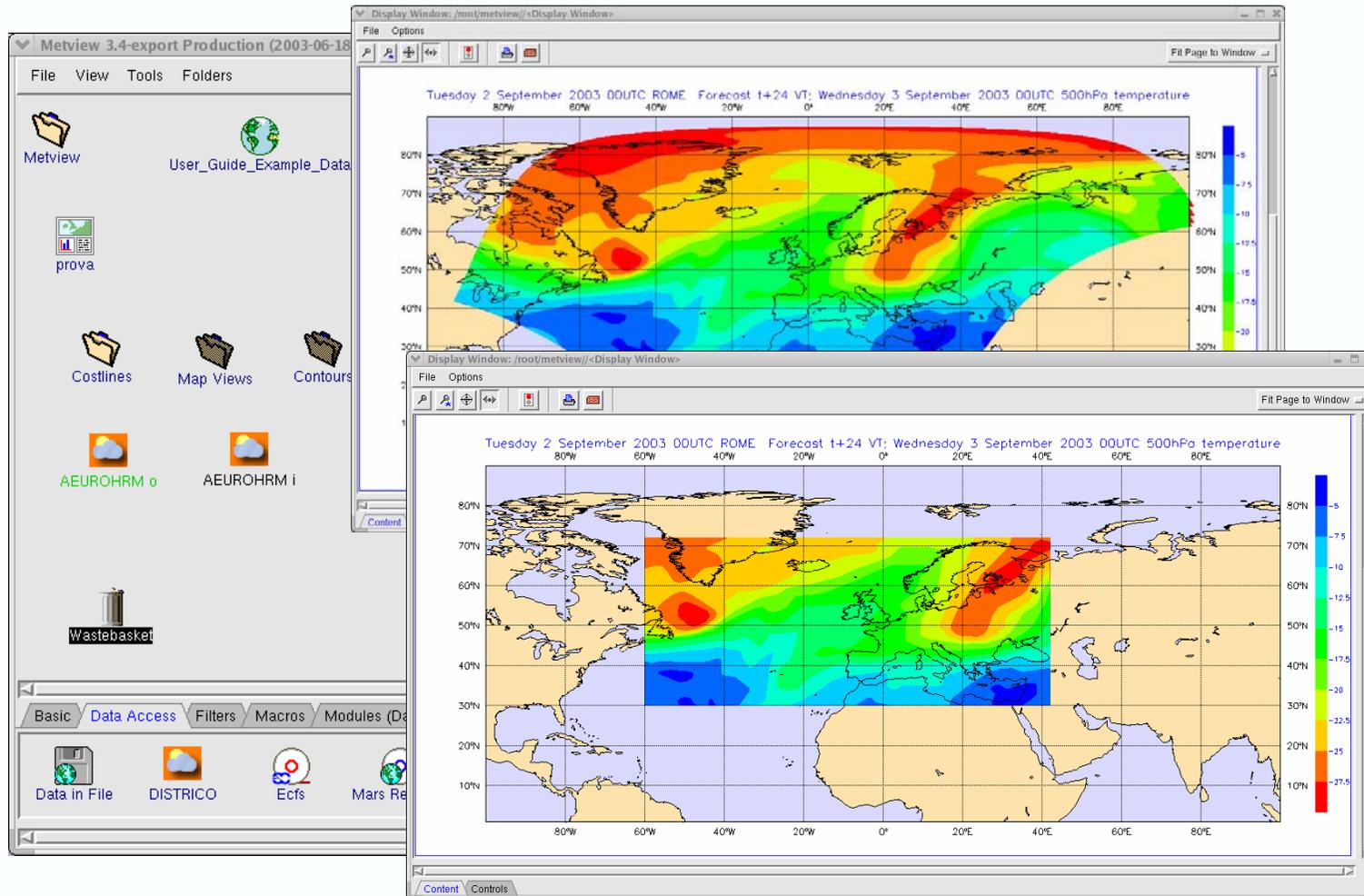
DISTRICO – GRIB Loading

GRIB fields unknown to the system are automatically classified.

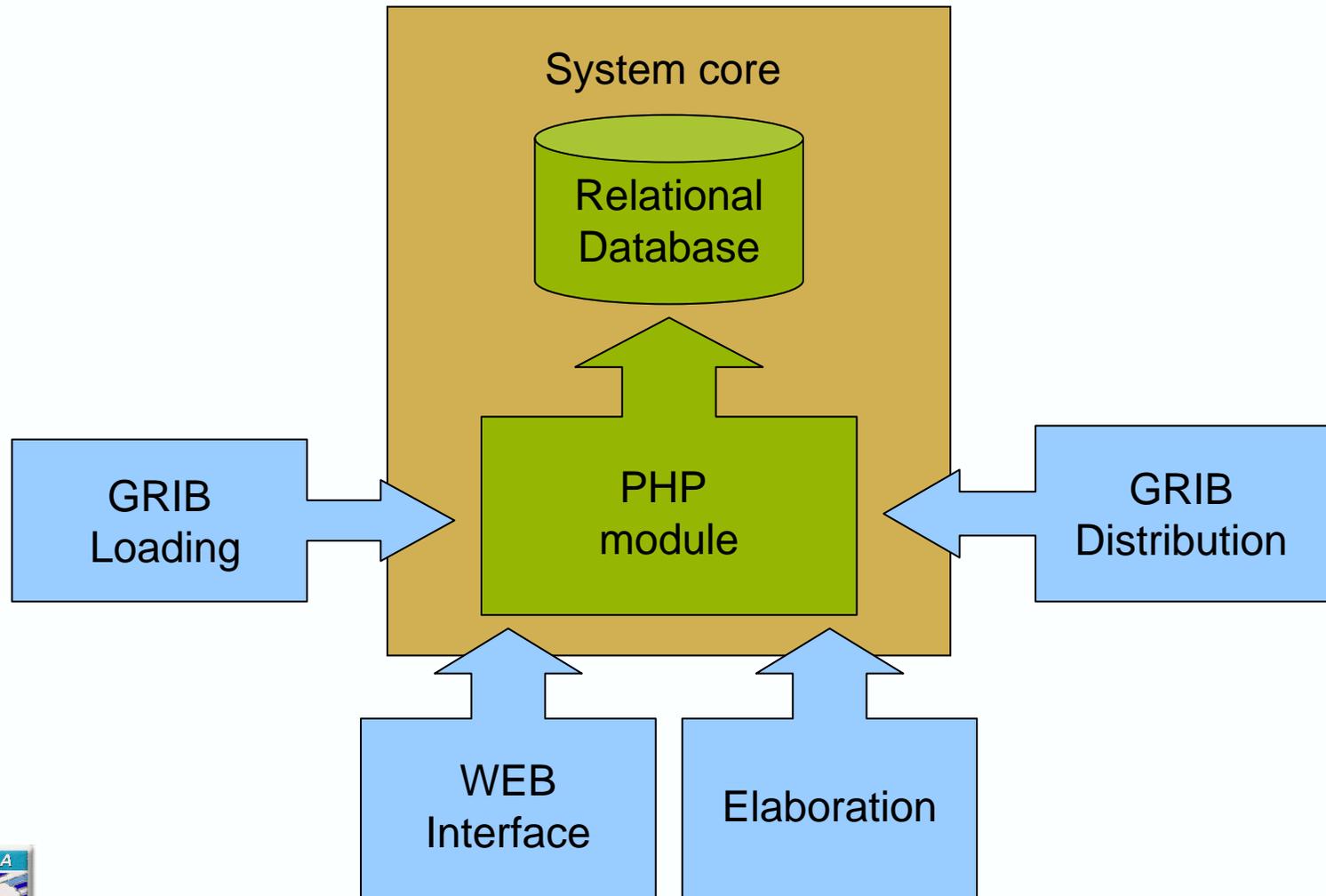
As soon as fields are inserted into the Database they are immediately available for elaboration, distribution, visualisation.



DISTRICO – Fields Visualisation



DISTRICO – GRIB Distribution



DISTRICO – GRIB Distribution

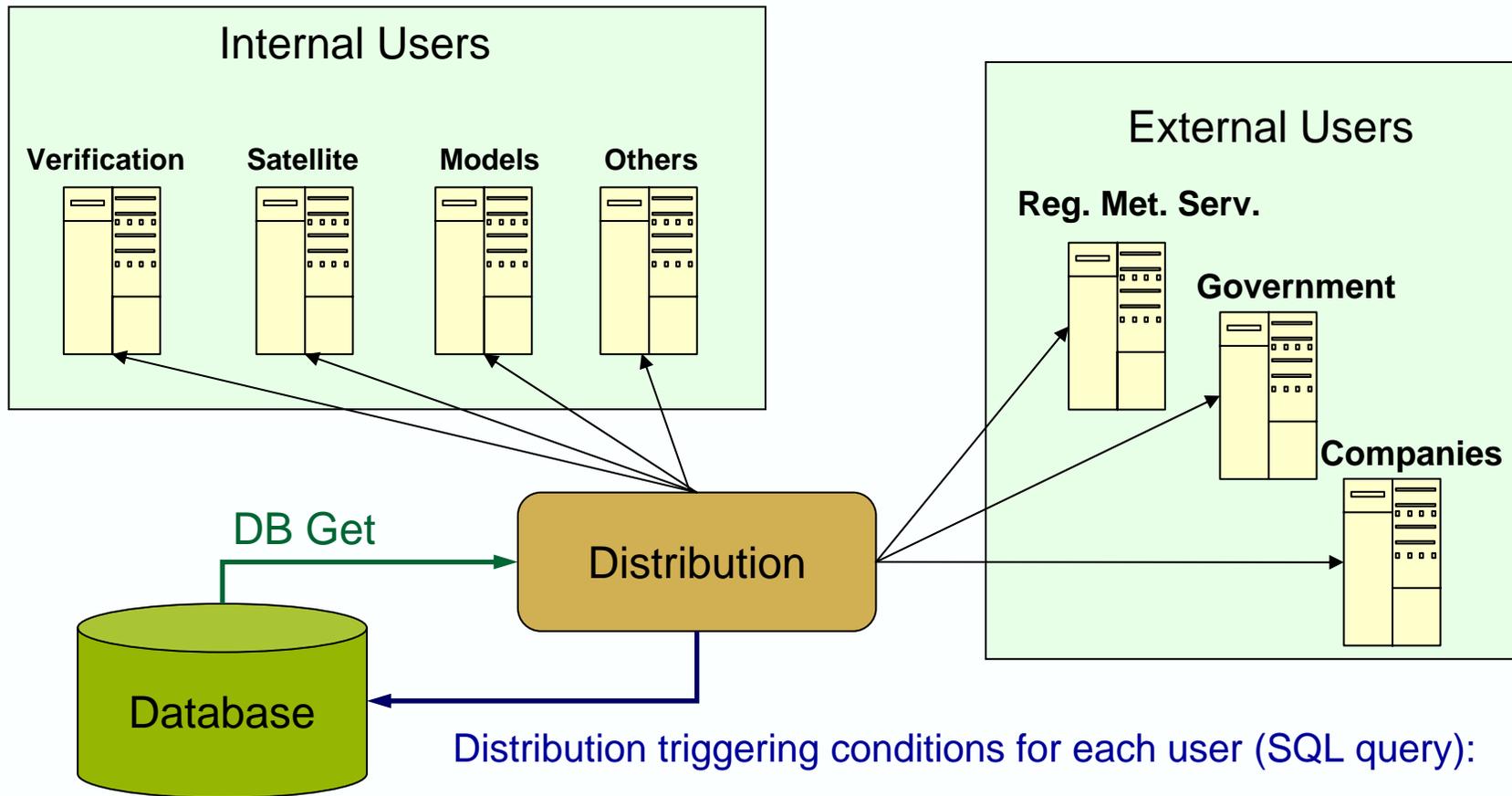
REGISTERED GRIB							
id_reg_grib	id_centre	id_process	id_grid	id_table	id_param	run	...
1001	80	103	23	2	11	0	...
1002	80	107	12	2	2	0	...
1003	98	203	34	2	1	12	...
...

Distribution	
id_reg_grib	id_user
1002	2
1003	2
1003	3
1003	4
...	...

USER		
id_user	name	...
1	“Sezione satelliti”	...
2	“Presidenza del Consiglio”	...
3	“Comune di Venezia”	...
4	“Regione Puglia”	...
...



DISTRICO – GRIB Distribution

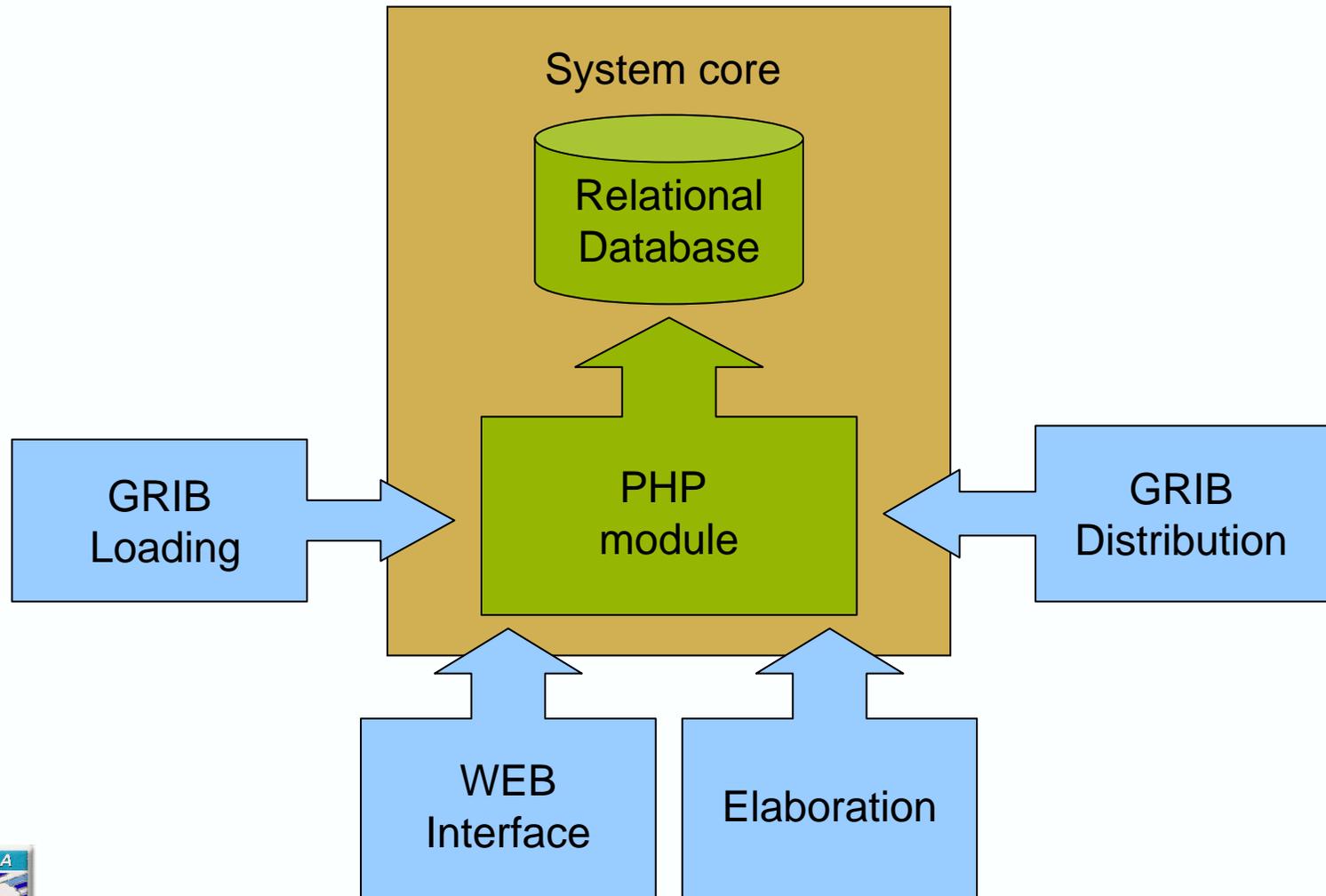


Distribution triggering conditions for each user (SQL query):

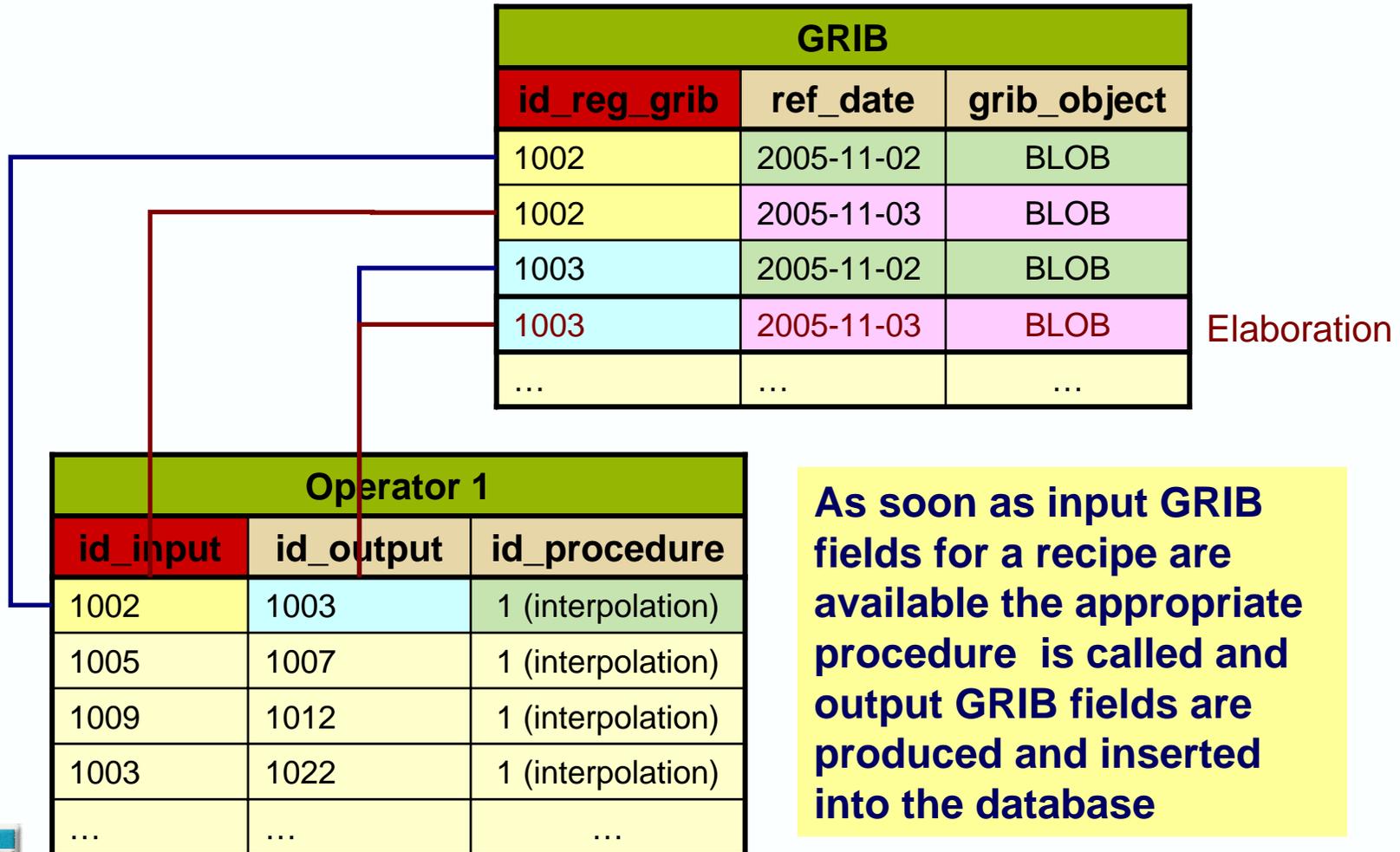
1. All fields requested for a specified model, run and time step are on DB
2. Fields were not already distributed



DISTRICO - Elaboration



DISTRICO - Elaboration



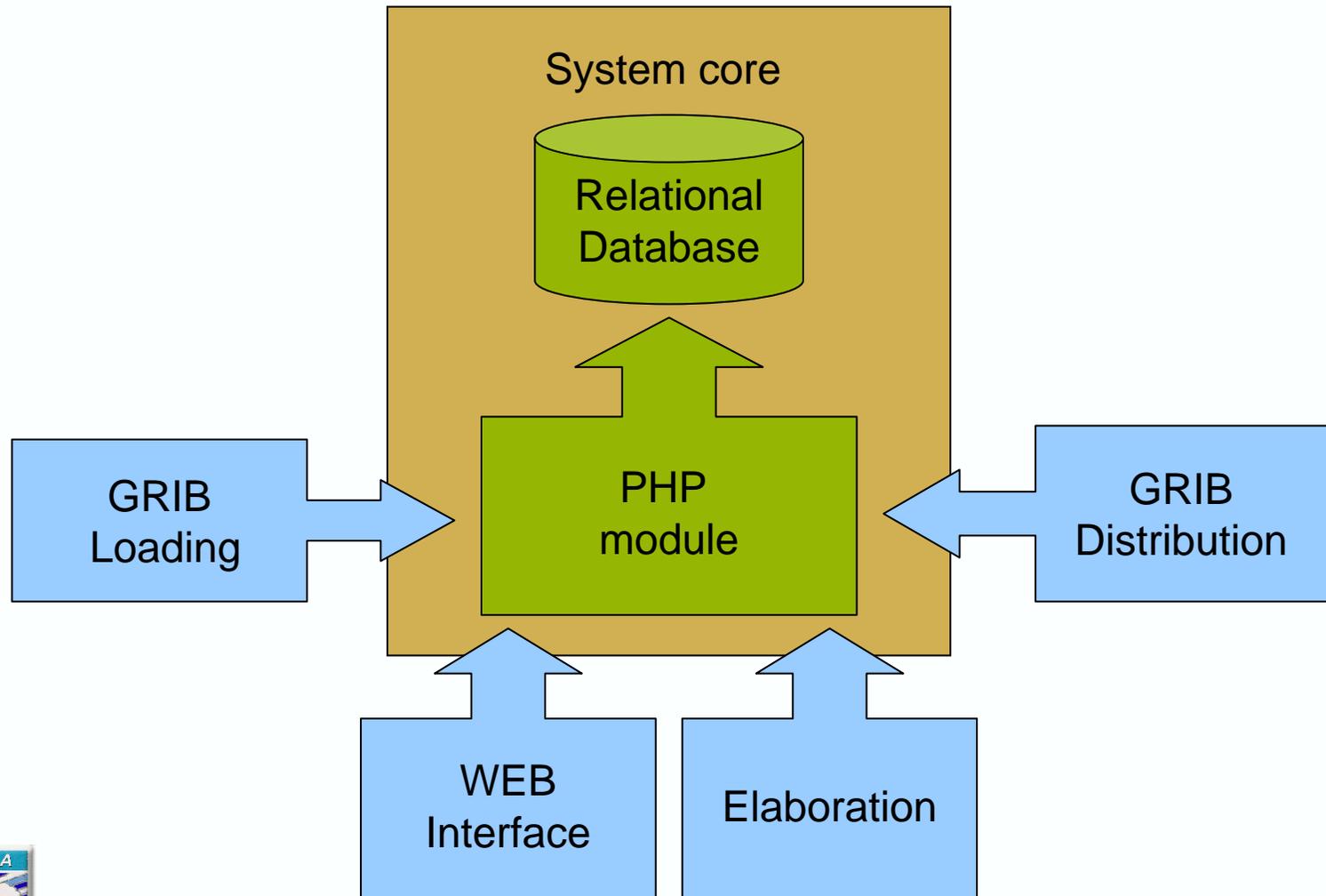
DISTRICO - Elaboration

All kind of elaboration can be carried on referring to GRIB data through their **ID**.

- **Interpolation** from one grid to another.
- Production of fields with **elementary operations** (cumulated fields, differences,...)
- Production of **all kind of charts** using MAGICS
- Production of **meteograms** in textual and graphical format



DISTRICO – WEB Interface



DISTRICO – WEB Interface

- The **WEB interface** was initially designed for **system administration** not for users access
- It provides access to the **elaboration and distribution tables** so that an operator can easily define a new distribution stream or add a new set of fields to the elaboration activities
- **Monitoring views** are available for loading, distribution and elaboration
- Views are available for **monitoring status of processes and machines**



DISTRICO – WEB Interface

- **Apache** Web Server
- All **pages** are **dynamically created** getting content from the database tables
- A **model, view, controller** application was developed
- **Administration** can be done **remotely** accessing through a VPN



DISTRICO – WEB Interface



DISTRICO – WEB Interface

Districo Amministrazione - Microsoft Internet Explorer

File Modifica Visualizza Preferiti Strumenti ?

DistriCO

Giovedì 3 Novembre 2005

Disseminazione dell'utente Regione Puglia

LAMI N corsa 0

PRESSURE REDUCED TO MSL Pa (prod=2 tab=2)
Mean sea-level liv1=0 liv2=0 (id iterv temp=0)
griglia=82
0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57
60 63 66 69 72

TEMPERATURE deg K (prod=11 tab=2)
Specified heigt level (above ground) liv1=0 liv2=2 (id iterv temp=0)
griglia=82
0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57
60 63 66 69 72

MAXIMUM TEMPERATURE deg K (prod=15 tab=2)
Specified heigt level (above ground) liv1=0 liv2=2 (id iterv temp=0)
griglia=82
0

MAXIMUM TEMPERATURE deg K (prod=15 tab=2)
Specified heigt level (above ground) liv1=0 liv2=2 (id iterv temp=2)
griglia=82
0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57
60 63 66 69

MINIMUM TEMPERATURE deg K (prod=16 tab=2)
Specified heigt level (above ground) liv1=0 liv2=2 (id iterv temp=0)
griglia=82
0

MINIMUM TEMPERATURE deg K (prod=16 tab=2)
Specified heigt level (above ground) liv1=0 liv2=2 (id iterv temp=2)
griglia=82
0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57

Utenti

- ANPAX
- ARSI2
- ARST1A
- ARSXX
- ASS00
- ASS06
- ASS12
- ASS1B
- GESIX
- CNTR2
- CNTRL
- CONVE
- CSIMX
- CTRL1
- CTRL2
- ERSAX
- ESRIN
- ETNAX
- EURO0
- EURO2
- HRM00
- HRM12
- ICI4S
- ICIXX
- INGVC
- LINAT
- LISIN
- LD120

Strumenti

- Aggiungi utente
- Elimina utente
- Modifica utente
- Aggiungi prodotto
- Elimina prodotto
- Attiva distribuzione
- Disattiva distribuzione
- Inverti attivazione
- Riesegui distribuzione

PUGLI

Regione Puglia
Priority=10
Host=meteoftp
Username=ecmwf
Directory=
/users/ecmwf/bin

199 campi disseminati

Griglia 82

latitudine primo = 48000
latitudine ultimo = 35000
longitudine primo = 6000
longitudine ultimo = 20500
num. punti lon. = 146
num. punti lat. = 131
delta lon. = 100
delta lat. = 100
lon. polo sud = 0
lat. polo sud = 0
rotazione = 0



DISTRICO – WEB Interface

Processo	Numero campi registrati
Analisi Non Inizializzata corsa 0	115
Analisi Non Inizializzata corsa 3	1
Analisi Non Inizializzata corsa 6	148
Analisi Non Inizializzata corsa 9	1
Analisi Non Inizializzata corsa 12	148
Analisi Non Inizializzata corsa 15	1
Analisi Non Inizializzata corsa 18	148
Analisi Non Inizializzata corsa 21	1
EUROLM (ECMWF) corsa 0	1871
Global Model corsa 0	3051
Global Model corsa 6	46
Global Model corsa 12	5454
Global Model corsa 18	46
LAMI corsa 0	10244
LAMI corsa 12	13387
LAMI N corsa 0	20312
LAMI N corsa 12	15054
Nuovo EUROHRM Asincrono corsa 0	13219
Nuovo EUROHRM Asincrono corsa 12	10422
Nuovo HRM corsa 0	6372
Nuovo HRM corsa 12	6370

Totale campi registrati 106411

Processo	Data	Corsa	Numero di grib
Analisi Non Inizializzata	2005-10-31	0	115
Analisi Non Inizializzata	2005-10-31	3	1
Analisi Non Inizializzata	2005-10-31	6	148
Analisi Non Inizializzata	2005-10-31	9	1
Analisi Non Inizializzata	2005-10-31	12	148
Analisi Non Inizializzata	2005-10-31	15	1
Analisi Non Inizializzata	2005-10-31	18	148



Future development - WEB interface

- New release should provide **tools to a user for plotting** a set of charts and submitting them to the production system
- Providing **aided access to MAGICS** functions through a friendly WEB interface
- **Embedding MAGICS** into PHP (using SWIG?)
- On the fly meteogram plotting



Future development – Data handling

- Implementation of **GRIB Edition 2** (database structure must be partly changed, system design remains unchanged)
- Use of **Smart Binary Large Objects**, user defined types and functions on **IBM Informix RDBMS** (**embedding C code into the database engine** so that single points or vertical profiles can be obtained by SQL queries)



Conclusions

- A **Relational Database** is **efficiently** used to store large number of fields in GRIB format
- Putting all fields into a database provide access through high level language as **SQL**
- Using a scripting language as **PHP** **speeds up the system development** and makes **easier** its maintenance
- **C and Fortran** routines can be easily **embedded** into the PHP scripting language providing **high performance** and reliable calls



Conclusions - Scalability

- **Database is scalable** up to PetaByte depending on the RDBMS. IBM Informix assures that this is possible using the fragmentation of tables and other techniques **without loss of performance**.
- **Distributed database can be easily managed** into the PHP module hiding as much as possible multiple data sources to end users.
- **Activities model is scalable** because each activity can be divided into smaller tasks and distributed to different machines.



Questions ?

