

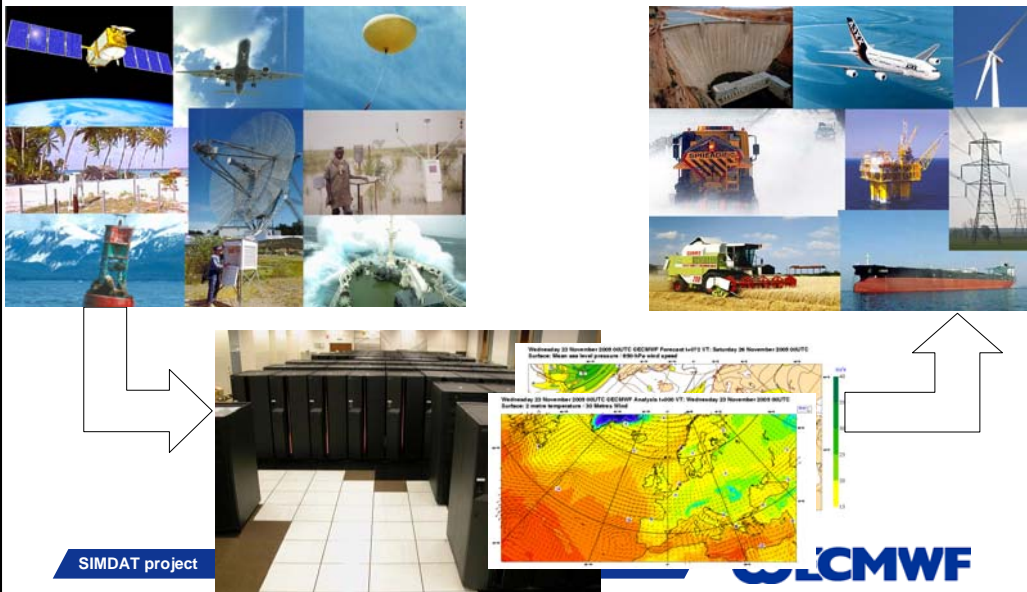
# SIMDAT Project Meteorology Activity Elements for building the WIS

Guillaume Aubert – Baudouin Raoult, ECMWF

SIMDAT project



## Real-time data flow: From Observation to Numerical Weather Prediction to Decision Making



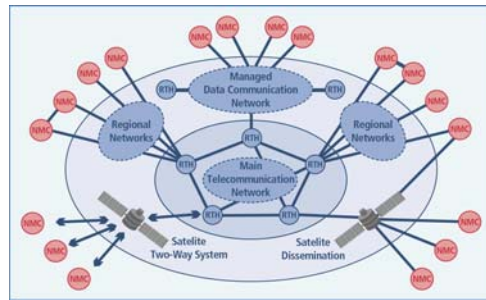
## Current Situation: GTS

- **Global Telecommunications System (GTS)**

- Observations
- Forecasts
- Warnings

- **Private Network**

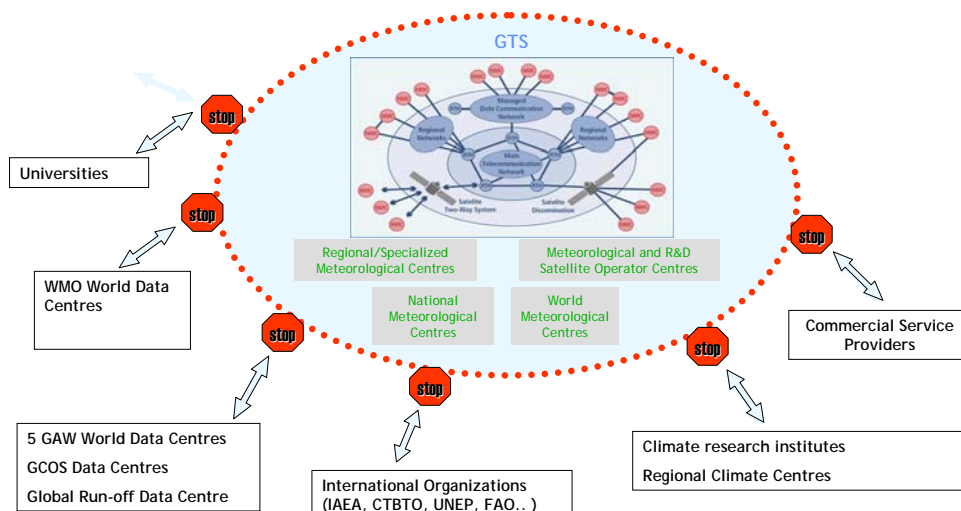
- Node to Node network
- NOT an Internet



SIMDAT project



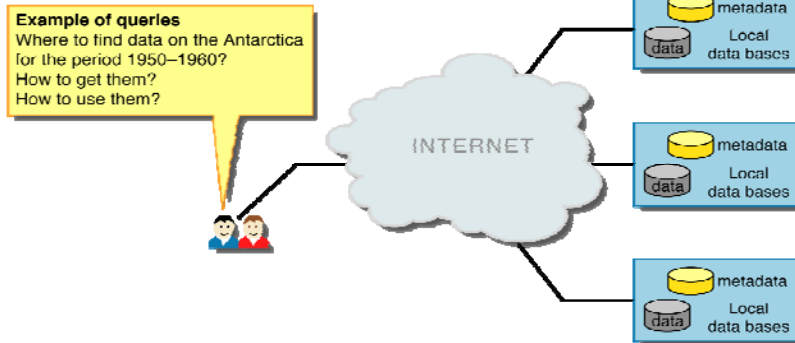
## Current Situation: GTS



SIMDAT project



## Current Situation (cont.)



- At present, WMO Programmes do not offer appropriate response to such queries

SIMDAT project



## WMO Information System - Vision

- **Real-time collection and dissemination:**
  - Real-time “push” through dedicated telecommunication for operation-critical data
- **Timely delivery of data and products:**
  - Delayed mode “push” through dedicated telecommunication means and public data networks, especially the Internet
- **Data discovery and retrieval service:**
  - “Pull” through the Internet (HTTP, FTP,...)
- **Unified procedures**
  - More efficient data exchange
- **Coordinated and standardized metadata**
  - Data interoperability between programmes
  - Improved data management
  - ISO 191xxx series

SIMDAT project



## Structure of WIS

- **Functional centres interconnected by data communication networks**
- **National Centres (NC)**
  - Links national data providers and users to regional and global data exchange nodes
- **Data Collection and Production Centres (DCPC)**
  - Provides for regional and mission-oriented exchange of data and products
  - Supports information push and pull
- **Global Information System Centres (GISC)**
  - Provides for 24/7 reliable global exchange services
  - Collects and provides metadata
  - Supports data and information discovery and retrieval

SIMDAT project



## SIMDAT Project

- **EU funded project**
  - 4 Activities: Aero, Pharma, Automotive and Meteo
- **Meteorology activity of SIMDAT:**
  - 5 Partners: DWD, Meteo-France, UKMO, EUMETSAT and ECMWF
  - European effort for developing the WIS
- **Coordination and main developments: ECMWF**
- **Distribution of the SIMDAT software**
  - Contract with the EU, already specifies that the software will be made available to the entire WMO community
  - Wider distribution, looking at an Open Source licence
- **An evaluation copy can be requested from ECMWF**

SIMDAT project



## Meteorology Activity: Project Aims

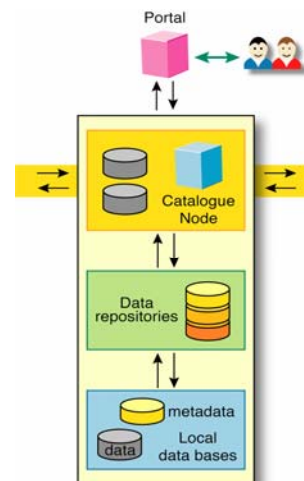
- **To build an integrated and scalable framework for the collection and sharing of distributed data**
  - Targeting meteorology, hydrology, climate and other environmental data
- **To provide a unified view of all available data**
- **To provide a transparent access to distributed resources**
  - Discovery service, Cataloguing service, Subscription service,...
- **To implement a non-intrusive system**
  - Provide access to existing local databases
  - Provide a global access control policy managed by the partners and integrated into their existing security infrastructure

SIMDAT project



## Prototype: Architecture

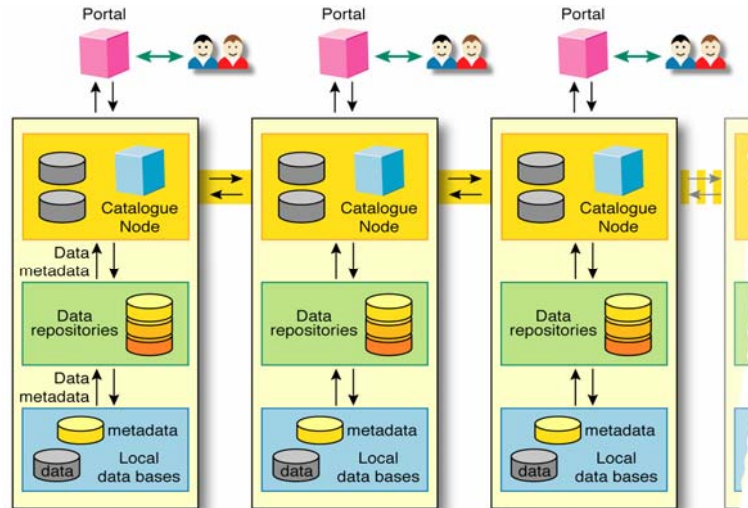
- **Based on three components (*plus one*):**
  - Portal
  - Catalogue nodes
  - Data repositories
  - (*Local database*)



SIMDAT project



## Prototype: Connectivity

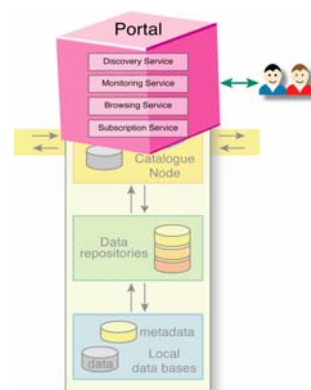


SIMDAT project



## Components: Portal

- **Web based user Interface of the system**
- **Offers discovery facilities**
  - Catalogue Browsing
  - Searching (keyword, time and space)
- **Allows data selection and request submission**
- **Offers per user request management**
  - Progress monitoring, ...
- **Offers data download facility**



SIMDAT project



## Focus on usability

- **Usability from an end-user point of view**
  - Portal based on concepts used by major search engines: no learning necessary
- **Usability from a data publisher point of view**
  - Effort to minimise prerequisite technological knowledge (Metadata, XML, ...)

## Finding data

- **Google like search**
  - **Fast:** people are now used get results in a few milliseconds
  - **Scoring:** people are used to find the most relevant hits in the first result page
- **Yahoo like directory**
  - Datasets can be categorized
  - A dataset can be in several categories at once


## Portal: Searching...

[Home](#) [Search](#) [Directory](#) [Requests](#)

**Search >**

**Keywords:**

**Location:**



Map showing search location. Coordinates: N 27.1, W -105.5, E -68.9, S -8.4. The map highlights a region in South America. Navigation controls are visible on the left. Text on the map includes 'POWERED BY Google' and 'terms of Use'.

Search by Location (e.g. Paris, France)

**Period:**

From:  To:  (yyyy-mm-dd)

SIMDAT project



## Portal: Search results

[Home](#) [Search](#) [Directory](#) [Requests](#)

**Home > Search Results >**

Search results for **ozone** from 2000-01-01 to 2006-12-31  
1-10 of 73 First Previous [Next](#) [Last](#) (0.9 seconds)

[ucar.ncar.acd.mozart.mozart2.mozart v2 1 maccm3](#)

- [EARTH SCIENCE > Atmosphere > Atmospheric Chemistry/Oxygen Compounds > Ozone >](#)

MOZART-2 includes detailed NO<sub>x</sub>-CH<sub>4</sub>-NMHC chemistry for the troposphere, with stratospheric chemistry constrained to climatologies. [EARTH SCIENCE > Atmosphere > Air Quality > Tropospheric Ozone](#) MOZART is a global 3-D chemical transport model driven by offline meteorological fields, developed at NCAR, MPI-Meteorology, and NOAA/GFDL.

[ucar.acd.model.eval.data.polaris](#)

- [EARTH SCIENCE > Atmosphere > Aerosols >](#)
- [EARTH SCIENCE > Atmosphere > Atmospheric Chemistry >](#)
- ...

POLARIS was a series of high-altitude airborne investigations to understand the behavior of polar stratospheric **ozone** as it changes from very high concentrations in the spring down to very low concentrations in autumn. The data will help with our understanding on the distribution, chemistry, and physics of stratospheric **ozone** after the vortex breakup, during the continuous daylight conditions of summer. NetCDF files of 1-minute averaged merged data sets for each flight of the ER-2 aircraft during the NASA POLARIS field experiment. These data sets include measurements that were part of aircraft campaigns, for the most part sponsored by NASA/GTE, NASA/UARP and NSF/NCAR. The NetCDF files contain 1-minute merged datasets, containing all available measurements, for each flight. These files have been created primarily for use in chemical model

SIMDAT project



## Portal: Browsing by Category

**SIMDAT** [Home](#) [Search](#) [Directory](#) [Requests](#)

[Home](#) > [Directory](#) > **EARTH SCIENCE** >

- [Atmosphere](#) (2 products, 19 categories)
- [Biogeochemistry](#) (2 categories)
- [Biosphere](#) (1 product, 5 categories)
- [Climate Indicators](#) (2 categories)
- [Climate change](#) (3 categories)
- [Cryosphere](#) (2 categories)
- [Human Dimensions](#) (3 categories)
- [Hydrosphere](#) (2 categories)
- [Land Surface](#) (7 categories)
- [Oceans](#) (2 products, 17 categories)
- [Paleoclimate](#) (1 category)
- [Radiance Or Imagery](#) (1 category)
- [Solid Earth](#) (1 product, 2 categories)
- [Spectral/Engineering](#) (4 categories)

SIMDAT project



## Portal: Legible Metadata

**Metadata** [Retrieve Data](#) [Show XML](#)


[Home](#) > **Metadata** >

**Title:** *Meso-scale NWP product from MESO\_LAPS*

**Abstract:** Meso\_LAPS (Meso-scale Limited Area Prediction System) is a part of the Bureau's regional NWP system. The prediction component of LAPS is a grid point primitive equation model with a grid spacing of 0.25deg in the horizontal, and 29 levels in the vertical, from the surface to 50 hPa. The lateral boundary conditions are obtained from the GASP system. The analysis component involves a 6 hourly data assimilation cycle. The system operates with a short data cut-off time to ensure timely delivery of forecast guidance to forecast offices. Consequently, the data analysis stage places special emphasis on the use of locally derived satellite observations. The data provided here is 3 hourly forecasts of the selected surface fields for a 48 hours period starting from the given base date/time.

**Period:** 2006-10-31 00Z to 2006-11-02 12Z

**Bounding Box:** 5°N 95°E 55°S 170°E

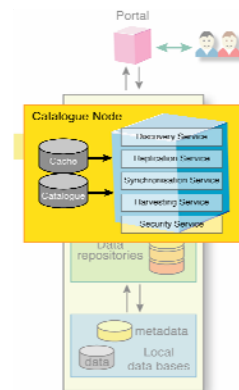


**Categories:** [Centre](#) > [BoM](#) > [NWP Outputs](#) > [Meso-scale Model product](#) > [0.25x0.25 degree](#) > [NWP Outputs](#) > [BoM](#) > [Meso-scale Model product](#) > [0.25x0.25 degree](#) >

**Keywords:** Operational Meso-scale NWP Product in netCDF format

## Components: Catalogue Node

- **Provides connectivity between Partners**
  - Can reach any other Catalogue Node
- **Holds the metadata catalogue**
  - Provide discovery services to the Portal
- **Implement peer-to-peer synchronization of metadata with other Catalogue Nodes**
- **Forward data requests**
  - To its Data Repositories
  - To peer Catalogue Nodes
- **Stream retrieved data between Data Repository and End User**



SIMDAT project



## Peer-to-peer synchronization

- **Each Catalogue Node synchronizes its catalogue with its neighbours in the Meshed Network**
- **Based on a journaling mechanism**
  - Create, update, delete records
- **Transactional access to the journal and metadata catalogue**
  - Maintains catalogue consistency
- **Clean separation between the metadata management and the synchronization**
- **Demo ...**

SIMDAT project



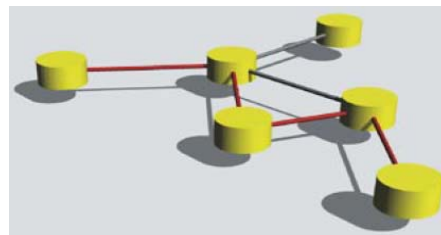
## Mesh Network: Peer-to-peer routing

- **Year I Prototype:**

- All the nodes had to be fully interconnected to get a complete catalogue (N x N interconnections)
- It is not scalable, not manageable as the number of nodes is growing

- **Year II Prototype:**

- Implements a Mesh Network
- Each node is connected to 2-3 peers at most
- Each node can reach any nodes which is part of the network



SIMDAT project



## Components: Data Repository

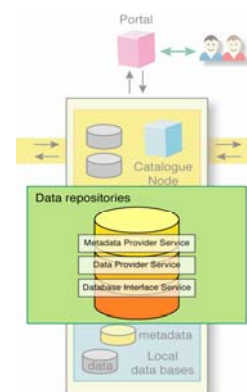
- **Provides a unified interface between the Catalogue Node and local databases**

- **Acts as a metadata provider**

- Publish metadata for harvesting by the Catalogue Node

- **Acts as a data provider**

- Accept data requests from the Catalogue Node
- Translate data requests into request for local databases
- Implement asynchronous handling of data requests
  - Support for access to off-line data



SIMDAT project

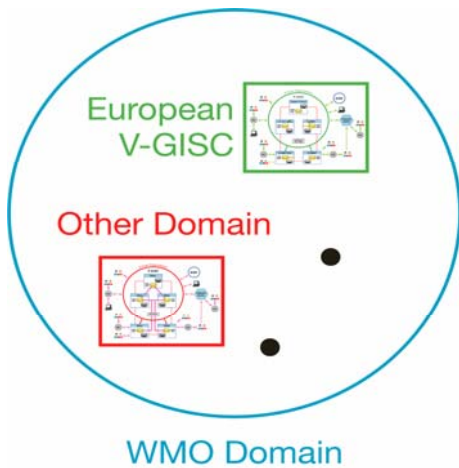


The image displays two screenshots of the METEO FRANCE metadata interface. The left screenshot shows metadata for 'Satellite imageries of MTSAT' with fields for Title, Abstract, Period, and Bounding Box, along with an XML snippet and a globe showing satellite coverage. The right screenshot shows metadata for 'GTS stations alphanumeric products' with similar fields, an XML snippet, and a globe showing station locations. A central diagram illustrates the data flow: 'Local data bases' feed into 'Catalogues', which are accessed by 'Users'. Red arrows indicate the flow from local databases to catalogues and then to users. Blue arrows point from the XML metadata to the globe, indicating the geographic context of the data.

## Virtual Organisation

- **We want security**
  - Users needs to be authenticated (who are they?)
- **We want to enforce data policies**
  - Can this user access this data?
  - Users have “roles”, e.g. this user is a “researcher”
  - Data have “policies”, e.g. this data is accessible by “researchers”.
- **These issues have to be solved in a distributed environment**
  - Non centralised solution: user may login at one site (authenticated)...
  - ... to get data from another (authorized)
- **Problems**
  - How to we make sure that all sites have the same understanding of roles and policies (e.g. what is a “researcher”)
  - How do we solve technically the fact that authentication and authorisation do not take place at the site.

## Virtual Organisations: Domains & Trust

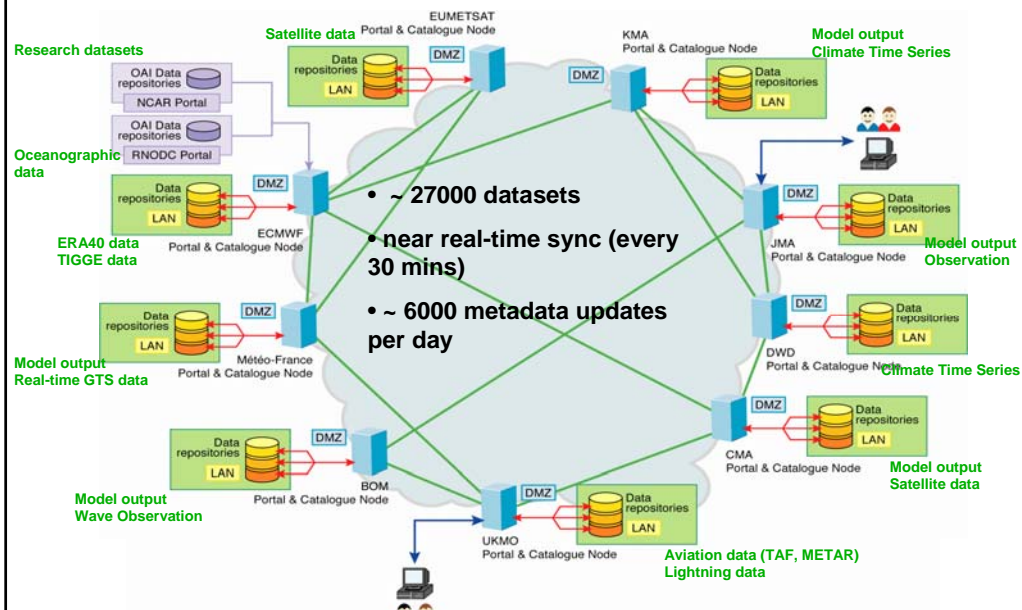


- A domain represents a group of organisations with a common data access policy.
- Organisations within a domain trust each other to authenticate users
- Authorization is performed by the Site hosting the data

SIMDAT project



## Status: 11 sites connected



## Data available in demonstrator

BoM	NWP outputs, Observations	GRIB, ASCII
CMA	NWP output, satellite data	GRIB, BUFR
DWD	Climate Time Series, NWP outputs	XML, HTML, GRIB
ECMWF	NWP outputs (ERA15, ERA40, TIGGE), Sample GTS observations (1 day)	GRIB1, GRIB2, BUFR, PDF
EUMETSAT	METEOSAT 8 images	JPEG
JMA	NWP outputs, MTSAT images, Observations	GRIB, BUFR, animated GIFs
KMA	Climate Time Series, NWP outputs	XML, GRIB
Météo France	NWP outputs Real-time GTS data	GRIB, ASCII
NCAR	~ 6000 datasets	NetCDF, other
RNODC	Oceanographic data (BATHY, SHIP, TESAC)	ASCII
UKMO	Aviation data (METAR, TAF)	XML

SIMDAT project



## Support of WIS requirements

Routine collection of observation data	Work in progress.
Automated dissemination of all observed data and products, both real-time and non real-time : Push model	Work in progress.
Ad-hoc requests for data and products : Pull model	Done.
Support of different user profiles and data policies	Work in progress.
Support of data and network security	Done. Based on SSL
Support the integration of diverse datasets	Done. Support for Any data type, from any data repository.
Reliable infrastructure	Done.
Technologically sustainable and appropriate to local expertise	Done. Based on Standards. Run on a PC. Open Source components.
Modular, flexible and scalable	Done.

SIMDAT project



## Plans

- **Security [Q3 2007]**
  - Tools for user and policy administration [Q2 2007]
  - Data access policy [Q3 2007]
- **Routine Collection of observations [Q2 2008]**
  - Push mechanism for critical data [Q2 2007]
  - Data Replication/Caching [Q2 2008]
- **Subscription service [Q1 2008]**
- **Monitoring Tools [Q4 2007]**
- **Data delivery of very large datasets [Q3 2008]**

SIMDAT project



## Conclusion

- **Virtual distributed database**
- **Flexible, non intrusive architecture**
- **Peer-to-peer synchronisation mechanisms implemented**
- **Mesh Network communication implemented**
- **Generates interest in meteorology and other environmental communities**
- **Software running continuously, catalogues are constantly updated**
- **Data available from 11 sites worldwide**

SIMDAT project



**Thank you !**

SIMDAT project

