

V-GISC/SIMDAT

Alfred Hofstadler, ECMWF

Data exchanges are the foundation of the core meteorological activities. Observation data is continually collected by earth observation systems such as automatic or manned station, aircrafts, ships, satellites, etc. These data are fed into real-time databases and used to monitor the weather conditions and as input to numerical weather prediction systems running on high performance computer facilities. The outputs of such models are then redistributed to weather forecast centres to be used by forecasters for the creation of value added products targeted at end-users. The observation data and the model outputs are also archived for research use and climate studies.

Over the years, the meteorological community has developed a multitude of information systems, each one designed to meet the specific requirements of a particular meteorological programme. This multiplicity of systems has resulted in incompatibilities, inefficiencies and often in the duplication of efforts. In 2002 the World Meteorological Organization (WMO) approved the concept of the Future WMO Information System ([F]WIS). The WIS will provide a single coordinated global infrastructure for the collection and sharing of information in support of all WMO and related international programmes. The core components of the WIS will be the GISCs (Global Information System Centres) and the DCPCs (Data Collection and Production Centres).

The national weather services of France, Germany and the UK in cooperation with the European Centre for Medium-Range Weather Forecasts (ECMWF) and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) plan to develop a common infrastructure for the collection and sharing of distributed meteorological data: the Virtual Global Information Systems Centre (V-GISC). The V-GISC concept will be developed within the meteorology activity of the SIMDAT project. The V-GISC/SIMDAT project will help to develop a virtual and consistent view to this vast amount of distributed data and provide a secure, reliable and efficient access to them, in order to support research and operational activities of the meteorological community and other related environmental communities.

Some key elements of the project are:

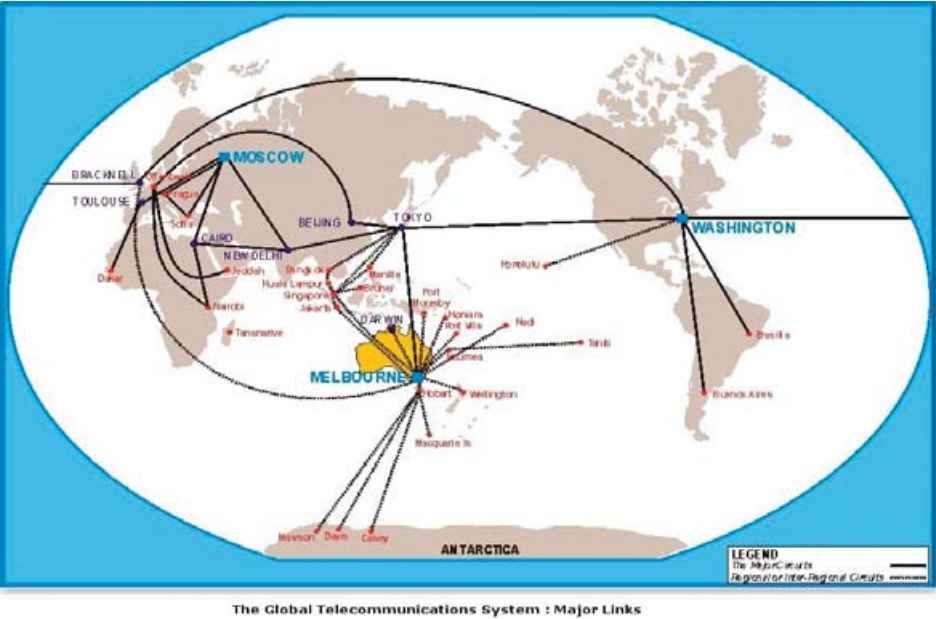
- Improve visibility and access to data through a comprehensive discovery service based on metadata development,
- Add value to existing datasets by enabling diverse databases to be used as a unique virtual resource,
- Offer a variety of reliable delivery services,
- Provide a global access control policy managed by the partners and integrated into their existing security infrastructure.

Within the V-GISC/SIMDAT project the technology and meteorology partners will together build, operate and support a prototype of the data GRID using the output of the SIMDAT technology activity. It is planned to evolve this test bed to a comprehensive infrastructure that should form the basis for the European V-GISC by the end of the project.

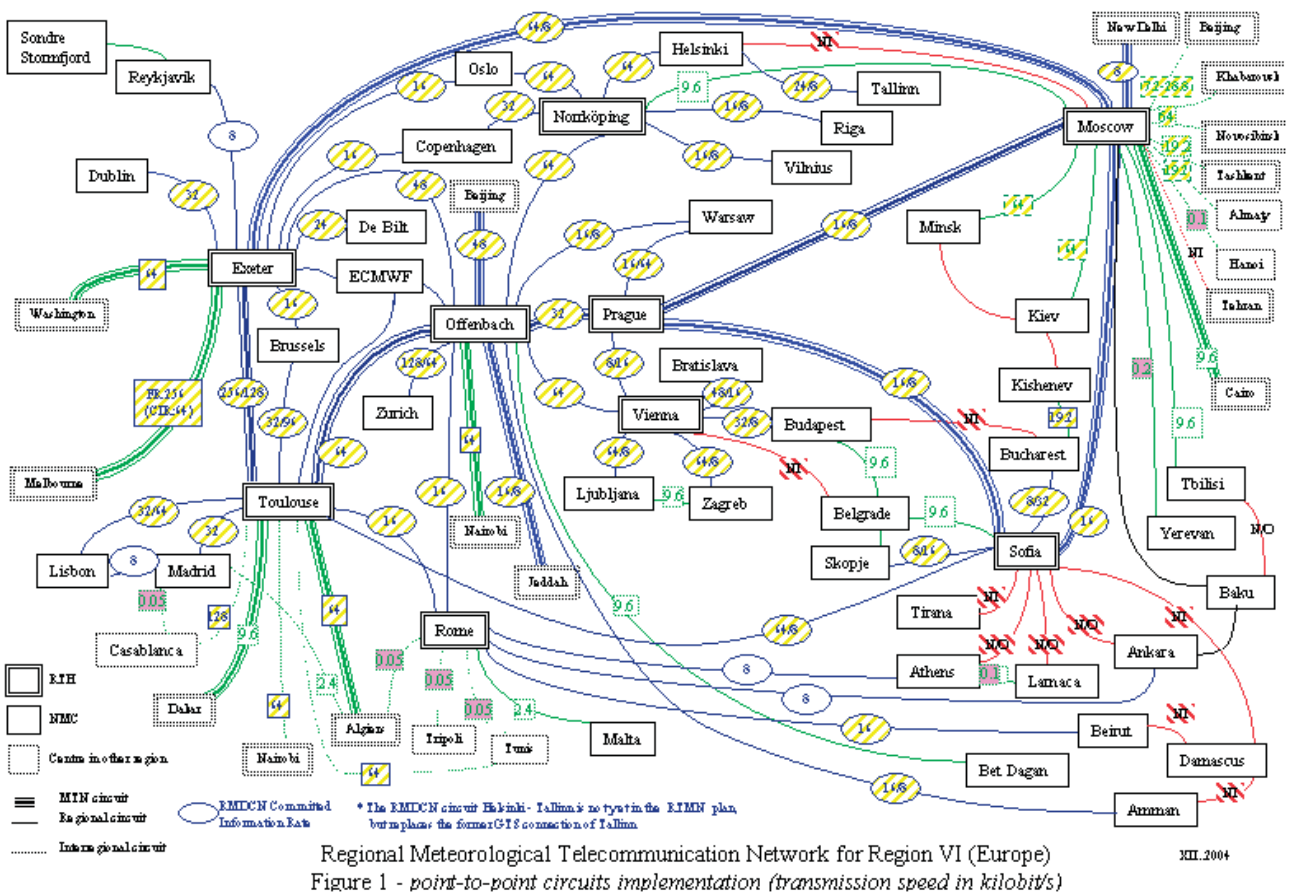
Background

- May 2002: Thirteenth session WMO Regional Association VI "...agreed that the concept of a Virtual GISC had merit..."
- June 2002: V-GISC in RA-VI Kick-off Meeting
 - Partners: DWD, Meteo France, UK Met-Office, EUMETSAT, ECMWF
- 2003: SIMDAT project proposal submitted to EU
- 1 September 2004: contract with EU signed for 4 year project
- October 2004: V-GISC steering group agrees to move V-GISC development into the SIMDAT project
- November 2004: SIMDAT Kick-off meeting
- November 2005: First V-GISC demonstrator

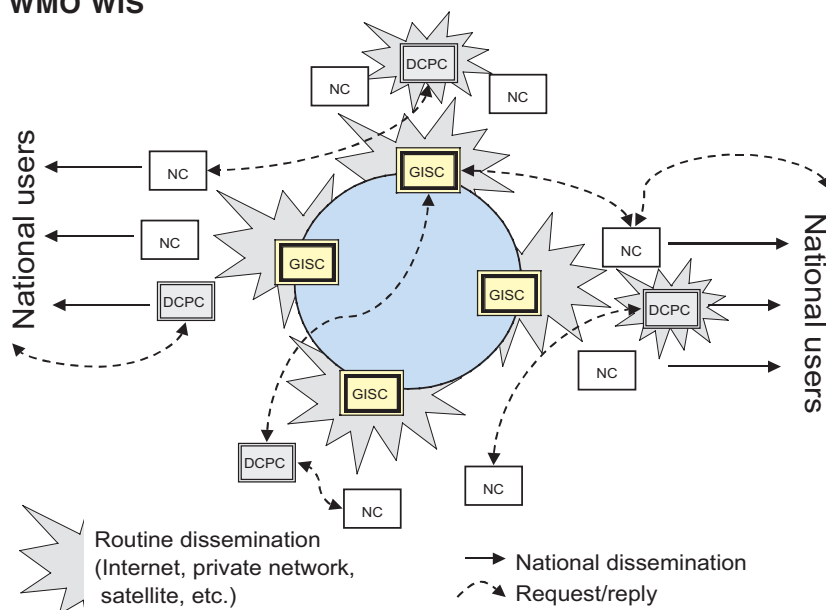
WMO GTS - Major links



WMO GTS - Region VI



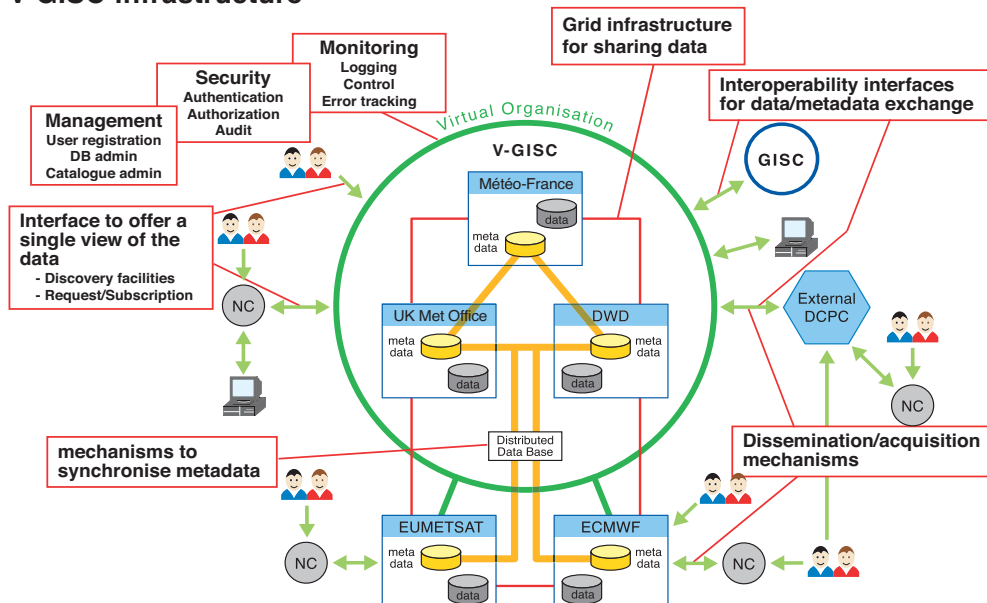
WMO WIS



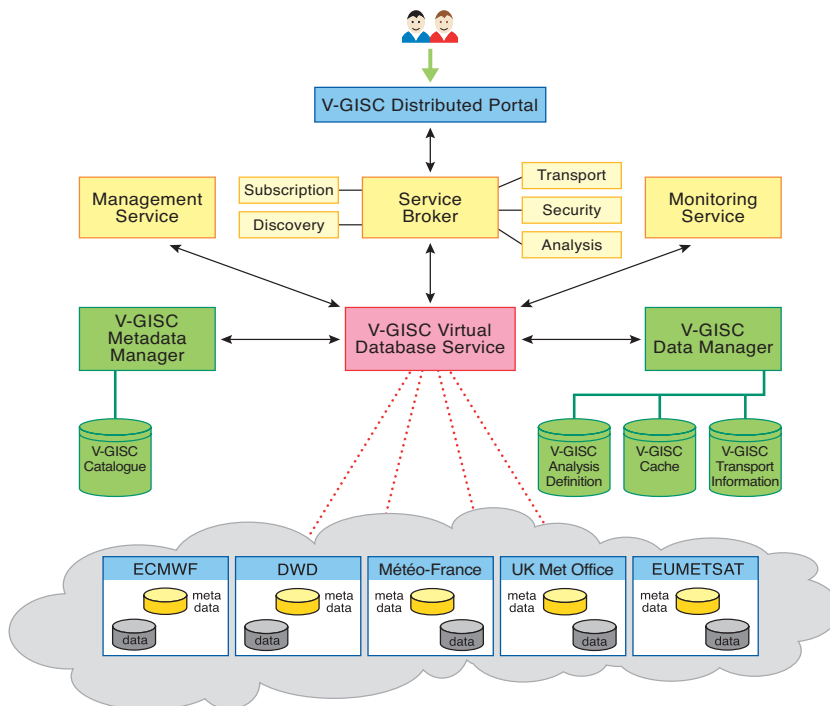
V-GISC Project Aims

- Instead of three GISCs in Region VI have one V-GISC
- The V-GISC will be seen as a normal GISC and will fulfil the WMO Information System technical requirements
- A complex problem: To build a Virtual GISC, an integrated and scalable framework for the collection and sharing of distributed data that will offer:
 - A single view of meteorological information which is distributed amongst the 5 partners
 - Improve visibility and access to meteorological data through a comprehensive discovery service based on metadata development
 - Offer a variety of reliable delivery services (routine dissemination and collection of data)
 - Provide a global access control policy managed by the partners and integrated into their existing security infrastructure
 - Quality of services, reliability and security
 - Processing services and shared data manipulation facilities
- The software developed within the project will be made available to WMO

V-GISC infrastructure



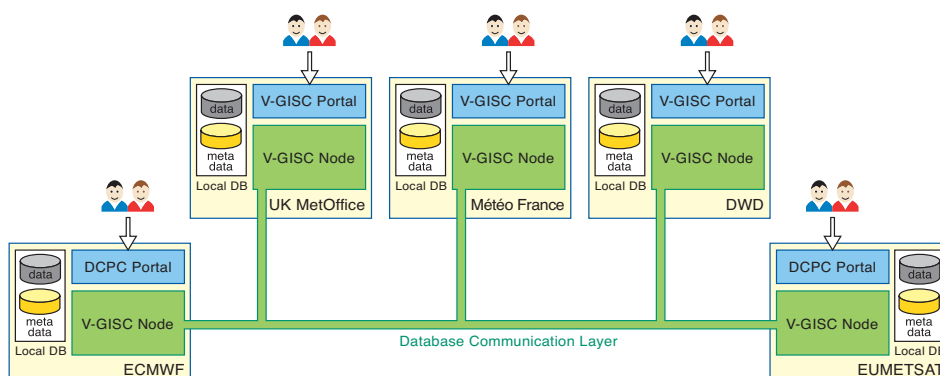
V-GISC Conceptual view



- Through the Distributed Portal users search for and retrieve data, subscribe to services subject to authentication and authorization
- The Virtual Database Service provides a single view of partner databases

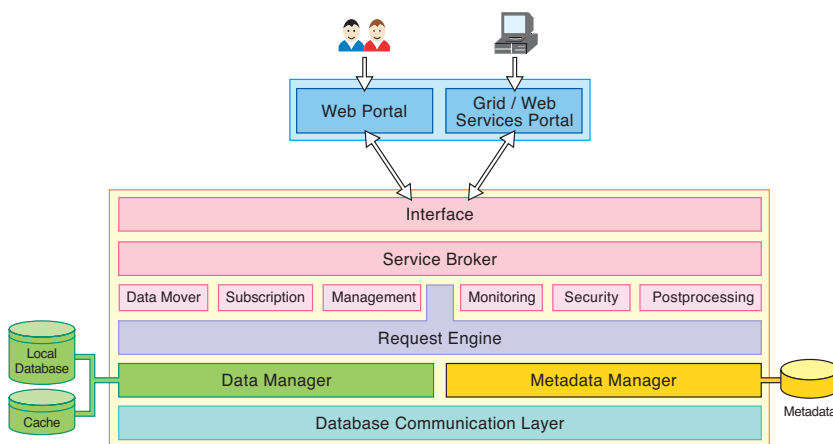
- Virtual Database
 - Provide a unified view of all the shared datasets through a distributed catalogue
 - Maintain the distributed catalogue amongst the partners using synchronization mechanisms
 - Provide interfaces to legacy databases
 - Implement data replication mechanisms
 - Preserve the integrity of the data
- Access Facilities
 - Collection & Dissemination services that support secure, efficient and reliable transport mechanisms
 - Quality of Service (QoS): Traffic Prioritization, Queuing mechanisms, Scheduling
 - Discovery service by browsing the catalogue or using a keyword search engine
 - Interactive and batch interfaces
- VO
 - Security Services
 - Users management
 - Data policy management
 - Monitoring and control

V-GISC Distributed Architecture



- Distributed components
 - V-GISC node is installed on each partner site
 - All nodes are interconnected through a dedicated secure communication channel: Database Communication Layer (DCL)
 - All the nodes exchange messages through the DCL
- Decentralized architecture
 - No central point where all the nodes are declared
 - No single point of failure
- Self-organized network
 - The network dynamically accepts new nodes and is aware of node disconnections
 - The network organizes its topology and indicates to the entering new nodes their position within the network
 - No manual intervention on the nodes to accept new peers

V-GISC Node – Functional Design



V-GISC Node

- Each node maintains a copy of the global catalogue describing data available through the V-GISC
 - The catalogue synchronization is done using the DCL
- Each node maintains a cache used to replicate data and to efficiently serve the users
- A node is interfaced with the local legacy databases
- A node has a Web Portal for interactive access
- A node has a Grid / Web Service Portal for batch access and integration of the V-GISC in a bigger Grid
- A node implements all services offered by the V-GISC

SIMDAT Outlook

- Finalise the Connectivity phase
- Coordination with other GISC and DCPC developer
- Prototype presentation at CBS-Ext., Seoul, November 2006
- Introduce acquisition of real-time data _ link with GTS
- Develop subscription service _ Push mode
- Develop Virtual Organisation
 - Monitoring and management of the system
 - User management and data access control
- Develop discovery mechanism
- SIMDAT project will run until August 2008
- Basic WIS infrastructure should exist by 2008