Implementation Issues
for a Data Management Server
to support GSICS

Peter Miu

19th February, 2008

EUMETSAT

Purpose of the Presentation

Propose the requirements for the data management server to support GSICS.

Use this as the basis for discussion for the implementation of a data management server.

Finalise on a set of agreements between the GSICS partners on this implementation.

EUMETSAT

The data management server shall allow GSICS partners to upload source data and products into a 'rolling archive' for use by other GSICS partners.

The source data and products shall be presented in a webpage and is downloadable via a web-browser.

The operator of the server can optionally archive the products received.
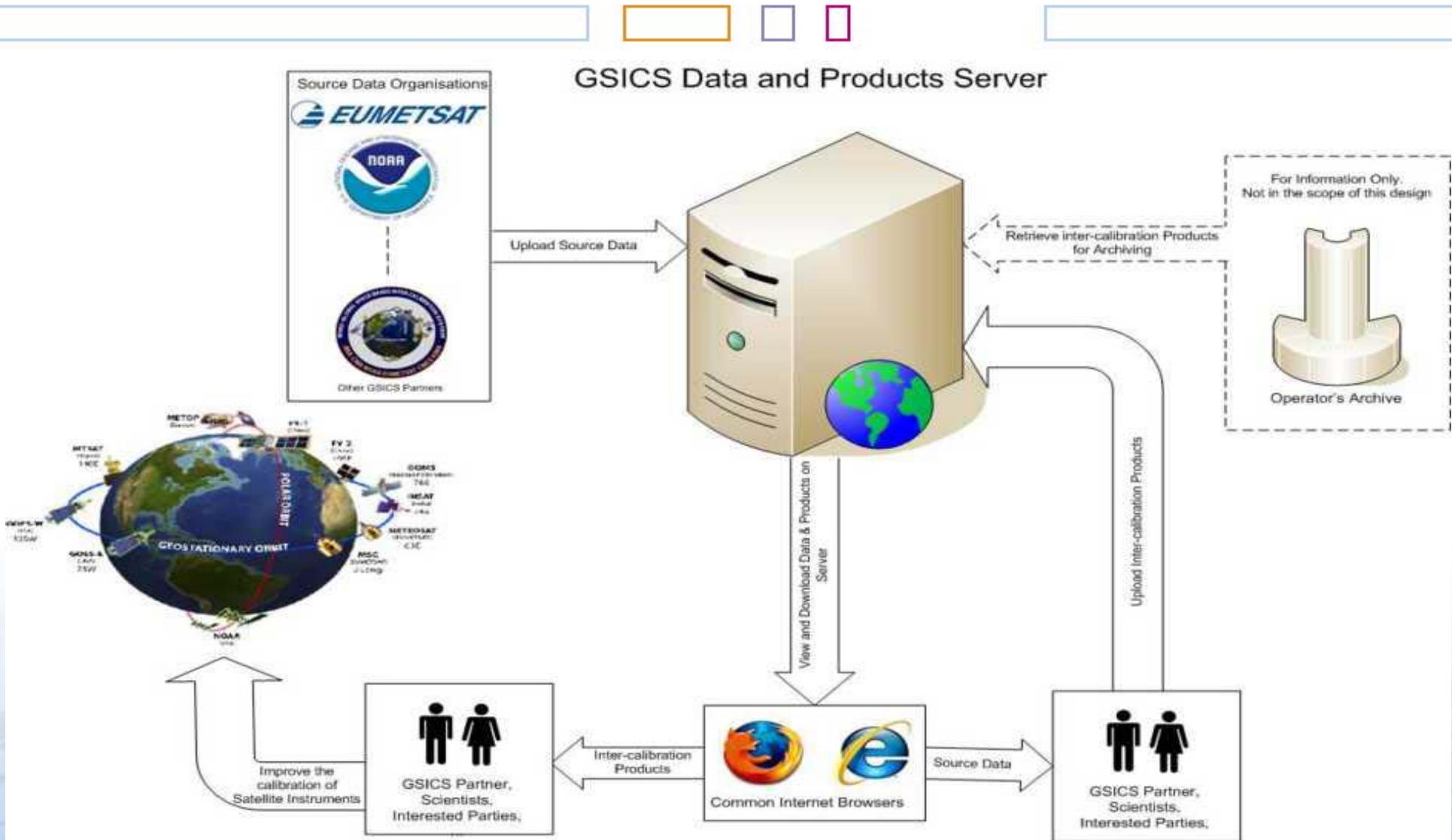
This is recommended for data preservation and it will benefit a wider user community.

Whether archiving requirements falls under the implementation of the data management server is up for debate as this is not specified in the primary goals of GSICS (see

http://www.orbit.nesdis.noaa.gov/smcd/spb/calibration/icvs/GSICS/index.html)

EUMETSAT

Generic view of the Data Management Server

GSICS Data and Products Server

Source Data Organisations

EUMETSAT

NOAA

Other GSICS Partners

Upload Source Data

For Information Only.
Not in the scope of this design

Retrieve inter-calibration Products
for Archiving

Operator's Archive

View and Download Data & Products on Server

Upload Inter-calibration Products

Improve the calibration of Satellite Instruments

GSICS Partner,
Scientists,
Interested Parties,

Inter-calibration Products

Common Internet Browsers

Source Data

GSICS Partner,
Scientists,
Interested Parties,

EUMETSAT

Proposed Data Management Server's Website Content

The server shall present the user with a catalogue of data and products that can be downloaded.

A simple tree structure with data and product attributes 'branches' can be used for organising the catalogue.

Downloadable content is represented as a node at the end of a branch.

Selecting a node shall provide its metadata information, as well as how to download it.

Optionally, additional links can be provided to invoke web applications that can visualise or interrogate its contents.

EUMETSAT

Inter-Operability is defined to be :

The ability of a system to work with, or use the parts or equipment of another system.

For the GSICS Data Management Server, the following components shall be inter-operable.

Source data formats.

Product formats.

EUMETSAT

Source Data Format.

The filename and its format shall be pre-determined by the GSICS partners.

Ideally a common format shall be defined, each partner shall supply their data in this format.

It shall be concise i.e. small file sizes to minimise transfer problems over the Internet.

All source data format guides shall be downloadable from the GSICS website and/or the data management server.

EUMETSAT

Product Format.

Scientific products created shall have a pre-determined set of meta-data values associated with it. These values shall be stored in the format itself.

Other Products that can be uploaded to the data management server are:

Software used for product generation.

Documentation such as algorithms, product format guides, etc.

How these products are represented on the data management server or the GSICS website is open for debate.

Server design; hardware, software, dataflow specification i.e.

What data.

Frequency the data is expected.

Shelf life of the data; how long the data shall exist on the server.

Source data format specifications including filename formats.

Products specifications including filename formats.

Proposed format of the product, software, documents up-loadable back to the server.

Shelf life of these products.

Archiving recommendations (?).

EUMETSAT

End of Presentation

Questions

EUMETSAT