New paradigm in access to information

Marie-Françoise Voidrot-Martinez
Head of Internet and Mobiles app - Météo-France
Co–chair of the OGC Met Ocean Domain Working Group – OGC

With Contribution from
Arnaud Cauchy – Airbus DS Geo Intelligence
Jérôme Gasperi - CNES

WMO
CONTEXT
Specificities of met data

- Huge volumes
- Quickly obsolete
- Multi dimensional
- Heterogeneous
- With Limitations
- Lots of Metadata
- Global
- Interoperable
- ...

METEO FRANCE
Toujours un temps d'avance
Specificities of Use Cases

- Real time or Archive
- Decision aids
- Availability
- Reliability
- Quality (Metadata)

Impacts on

- Emergency and disasters
- Agriculture
- Transport
- Defence
- Business
- Aeronautics
- Global policies
- Education...
Some milestones

- Looking for data: Observation, Production...
- Distributing the data (Networks, trades off) and Internet
- Visualisation
- Merge, extraction of information
- Open data
- Cloud
- Interoperability
DATA ACCESS
Global Telecommunication System

- Private Network
  - Node to Node network
  - Based on a “push” type technology
  - Not flexible
  - Standardized Message switching system
Push or Pull?

- Push

- Pull
From GTS to WIS a big step

- **GTS**:  
  - Push  
  - between Met Services  
  - Formats optimised for telecoms

- **WIS**:  
  - Pull  
  - open to all  
  - Adding WKN formats
New Implementation of Push and Pull

1. **Producer**: Push Notification + Id
2. **Consumer**: Request needed data
3. **Producer**: Send the requested data
Criterias of selection

- **Push**
  - High SLA
    - High speed
    - High reliability
    - Real time update
  - Need of autonomy
  - Small data
  - Very used data
  - Confidentiality of use
  - Controlled access

- **Pull**
  - Regular SLA
    - Regular speed *
    - Backuped providers
    - No real time update
  - Backup architecture
  - Big data
  - Save storage costs
  - Save bandwidth
  - Open access
  - Not always used
  - Public Cloud

- **Push/Pull**
  - Regular SLA
    - Regular speed *
    - Backuped providers
    - Real time update
  - Backup architecture
  - Big data
  - Save storage costs
  - Save bandwidth
  - Controlled access
  - Not always used
  - Private Cloud

* Only for 1st access, then can be stored locally.
FROM DATA TO INFORMATION:
WHAT ABOUT PROCESSES?
From data to information: today

- **Getting raw Data**
  - Data
  - Server
  - Data
  - Info

- **Getting Information**
  - Data
  - Server
  - Data
  - Info
  - Info
  - Server
New Implementation of Information on demand

1. Push Notification that a product could be calculated +Id
2. Request for product
3. Build product
4. Send the requested information
From data to information

- Send the process and execute it near the data (**web Processing Service**)

![Diagram showing data processing near the source](image-url)
From data to information

- Send the process and execute it near the data (web Processing Service)

**Use case**

- a big satellite center develops a new process to track a signal of climate change
- they want to build a set of homogeneous products covering the globe based on a unique process (or a family of processes fitted to the input data from the different satellites) on a long historical period
- getting all archived raw data is too big
In fact a chain of processes

Typically to:

- extract (WCS or WCPS)
- filter (WPS)
- transform (WPS)
- process (WPS)

Cloud:

- SAAS Software As A Service and
- PAAS Platform As A Service
Data and Computing centers

- Data centres (real time and/or archives)
- Basic services to extract, transform the data Cloud SAAS
- Collaborative packages (processing, data mining, …)
- Computing resources available for tiers processing Cloud PAAS (or IAAS)
- Accounting and scheduling tools
- Analytics
- Co working tools ...
New trades off to relieve pressure on the networks, and lower the costs

- New trades off on data access between
  - push,
  - pull, 
  - notification and pull access

- New trades off on process between
  - run in real time as soon as the data is here,
  - run when needed
  - near the user
  - near the data to distribute information rather than raw data

- Requires the definition of SLA = Service Level Agreement
  - On our data, products and services, (rights and services commitments)
  - On our providers (requirements of quality level for data, information and services, …)

Avoid transferring raw data if not used
Avoid calculating products if not used before being obsolete
Save storage and bandwidth costs
As a result

Push

Data Never Used – Transferred for nothing

Pull

Data or info used

Info Never used - Calculated for nothing
Issues raised by pull architectures

- **Discovery**
  - Discovery of the data,
  - Discovery of the processes,

- **Selection of the products**
  - How to describe quality information, quality of the processes

- **Metadata**
  - catalogue metadata
  - products metadata
  - algorithm metadata

- **SLA**
  - reliability of the producer
  - reliability of the services
  - reliability of the network,

- **Dynamic Scalability**
  - Analytics
  - Cloud computing

- **Security**
  - Identity, Authentication
  - Data
  - Infrastructure
An explosion of servers
Validation and quality issues

- Fat Client
  - HMI layer
  - Graphic layer

- Data Server
  - Data layer

- Client
  - IHM Layer

- Internet
  - OGC Webservices

- Server
  - Graphic layer
  - Data layer

- RMDCN

RTM ECMWF
More services issued from integration as well as from development
Wrap up

Avoid transferring raw data if not used
Avoid calculating products if not used before being obsolete
Optimise storage and bandwidth costs

- Build catalogues of data and services
- Define relevant metadata (search and product)
- Redefine trade off
  - push,
  - pull,
  - notifications to get or build on demand
- Improve standards
  - For data
  - For services
  - For SLA
Build interoperability!

Past:
- Interoperability between meteorological services
- Compatibility with institutional customers

Future:
- More interoperability with heterogeneous domains
  - Hydrology
  - Agriculture
  - Roads
  - Politics
  - ...

2040
Tomorrow for Satellite Industry but Science Fiction for Computer Information systems!
Some challenges nevermind…

- Always very constrained Time response

- A service continuity anytime
  - With a huge legacy system
  - Critical safety missions

- Standards yes … but with flexibility to support permanent evolutions:
  - of technology (captors, computers, telecommunications,…)
  - of science
  - Of the users expectations
Crisis in chinese

Wei (danger) Ji (opportunity).

Just don’t ignore it  Get prepared