



Cluster, THEMIS and the Virtual Magnetospheric Observatory

Raymond J. Walker, Jan Merka, Todd A. King,
Thomas Narock, Steven P. Joy, Lee F. Bargatze,
Peter Chi and James Weygand

THEMIS Cluster Workshop
University of New Hampshire
September 23, 2008

The Virtual Magnetospheric Observatory

- A web-accessible environment that provides unified discovery and retrieval for distributed magnetospheric data.
- A system that remains current by actively harvesting information about available resources.
- A source of domain experts to assist providers in making resources available through the VMO.
- A resource for providers in producing archive quality data collections.
- A provider of services to the science community for finding, retrieving and using data.
- A long term repository for data collections.

The VMO

- The VMO is a federated (GSFC and UCLA) data system focused on magnetospheric physics.
- The VMO offers a comprehensive view of available resources.
- The VMO provides value added services tailored to the magnetospheric researcher.
- The VMO enables data to be maintained by the experts - a distributed architecture.
- The VMO is built on a standards based framework.

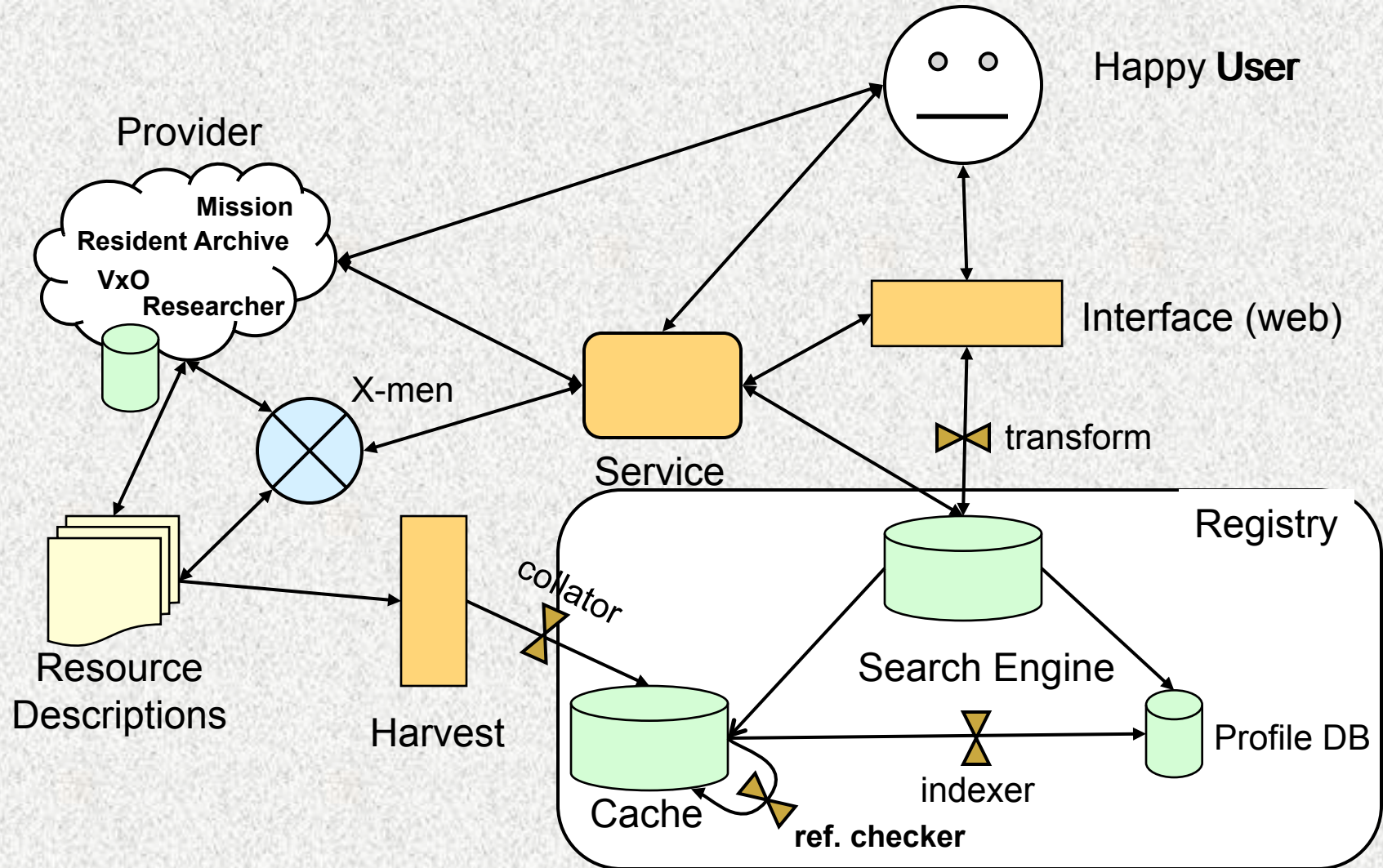
Our Approach

- Adopt a system design that recognizes that the data are best kept with experts.
- Contact providers and assisting with generation of metadata.
- Train domain experts (research scientists) also to be information engineers (X-Men).
- Involve the X-Men in the generation and review of resource descriptions.
- Encourage international partners and open data sharing.

SPASE and the VMO

- We use the Space Physics Access Search and Extract (SPASE) metadata standard.
 - All inventoried resources have a SPASE resource description.
 - The registry server is SPASE-based.
 - Services use SPASE resource IDs to locate data.
 - SPASE resource descriptions are delivered to the user (may be transformed by services)

Information Flow in the VMO



X-Men

- Identify data and/or service providers (project, resident archive, individual).
- Assign a lead.
- Identify each product set
 - i.e. Catalog, Numerical Data, Display Data, resolution, cadence.
- Work with data/service providers to generate SPASE descriptions.
- Schedule and track.

Inventory by Resource Type

(Goddard, UCLA and SMWG)
(Sept 2008)

Type	Current	Change From July 2008
Display Data	36	0
Catalog	1	+1
Instrument	1,123	+656
Numerical Data	1,323	+413
Observatory	819	+444
Person	4,666	+4,468
Registry	9	0
Repository	73	+67
Granule	184,170	+53,529

Projects In The System

- ACE
- AMPTE CCE, IRM, UKS
- Cluster
- CRRES
- DE1,2
- FAST
- Geotail
- GOES
- ICESTAR
- IGPPLANL
- IMAGE
- IMP8
- Interball
- ISEE-1,2,3
- McMAC
- MEASURE
- Polar
- SAMPEX
- THEMIS
- THEMIS-EPO
- WDC
- Wind

Cluster Update

- Phil Richards, James Weygand
- SPASE Metadata has been prepared for all four spacecraft and for Cluster personnel
- The CIS and FGM instruments have been described.
- Numerical data files for prime parameters have been described for FGM, PEACE and RAPID
- Granules are not yet available.

THEMIS Update

- Jan Merka, Peter Chi
- THEMIS ground magnetometer data
 - Observatory, Instrument, Numerical Data – current
 - Granules – up to July 30, 2007
- THEMIS spacecraft data
 - Observatory, Instrument, Numerical Data - FGM, EFI, ESA – current

Tools

- Data Providers
 - Tools for creating SPASE metadata
 - Tools for creating registries
 - Tools for harvesting metadata
- Data Users
 - Services to aid in finding and accessing data.

The Repository of Last Resort

- Some providers have asked the VMO to host their data.
 - Typically smaller research groups that lack institutional support.
 - We always agree
 - We have established a 6TB RAID array as the initial repository.
 - We send data bricks to transport the data (network connections are too slow).

Moving Towards an Archive

- Resident Archives have been established for SAMPEX and Polar.
- Work with the Resident Archives to produce archive quality collections.
 - Data Formats
 - Documentation
 - Organization
 - Integrity
 - Tracking

Problems

- Metadata creation has been much slower than anticipated. Describing data granules has been very slow. Only now are data bases being added at a rapid pace.
- Services to make the data more useful are coming slowly. Most are for data providers not users.
- More work is needed on the user interface and display of the results.

Summary

- The VMO is operating – <http://vmo.nasa.gov> or <http://vmo.igpp.ucla.edu>
- Our infrastructure is becoming stable – tools are available to help data providers become part of the virtual observatory system.
- With the help of the X-Men data bases are being added to the system.
- We are beginning to offer user services.