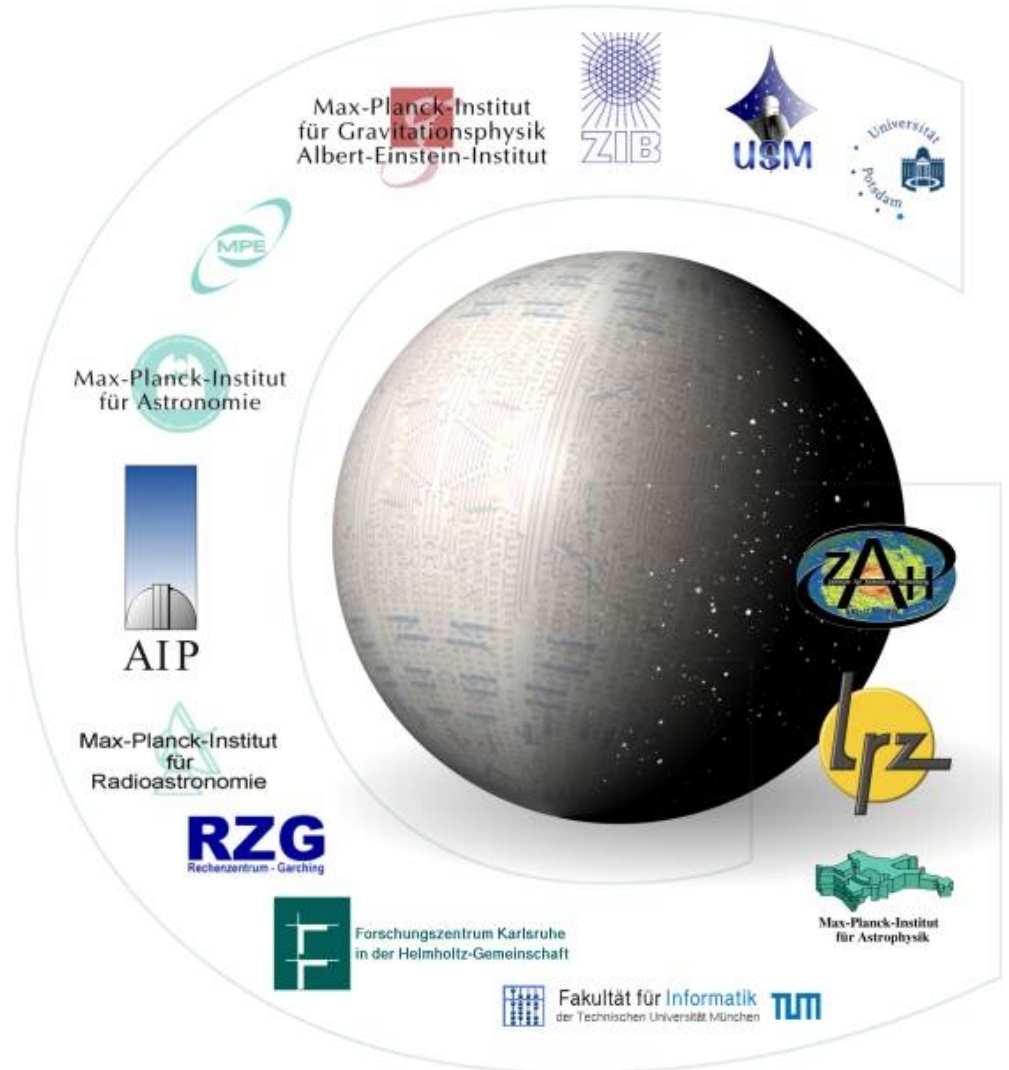
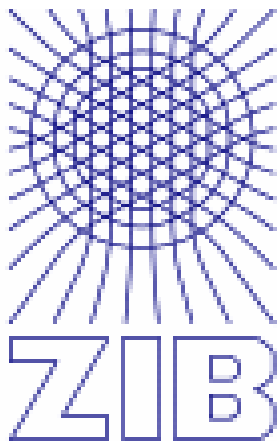




WG-2 Information Service

Mikael Höggqvist
hoeggqvist@zib.de





Motivation

- Metadata describes resources
 - ◆ discovery through search or browsing
 - ◆ metadata and raw data is mixed
 - ◆ generated by people and by programs
 - ◆ used by people and by programs
- AstroGrid-D metadata
 - ◆ resource information (CPU speed, Memory, ...)
 - ◆ grid state (jobs, files, ...)
 - ◆ application-specific (simulation iterations, ...)
 - ◆ scientific (data set annotations, provenance, image metadata, ...)



Requirements

- Extensible/flexible schemes
- Easy to extract and export metadata
- Protect from unauthorized access
- Support exact match and range queries
- Handle both static and dynamic metadata

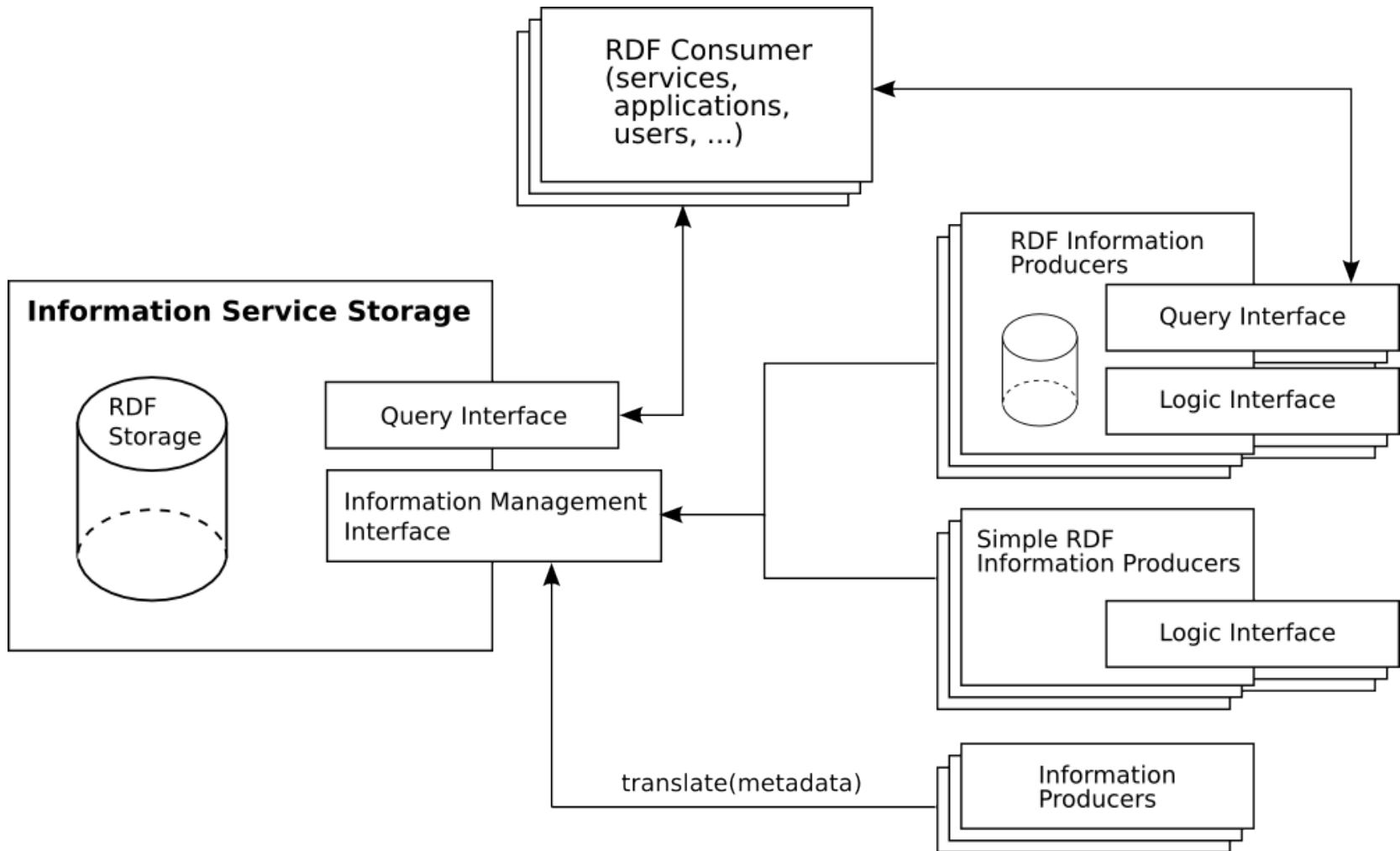


Approach

- Resource Description Framework (RDF) for representation of metadata
- SPARQL, a query language for RDF
- Service for query and persistent storage of metadata
- Metadata is naturally partitioned
 - ◆ Scientists at institutes are responsible for different applications
 - ◆ => query over multiple information service instances



Architecture overview





UseCase - Cactus

GridSphere Portal - Mozilla

File Edit View Go Bookmarks Tools Window Help

http://localhost:8080/gridsphere/gridsphere?cid=simviewportlet

gridsphere portal framework [Logout](#)
Welcome, Thomas Radke

Welcome Administration **Cactus Metadata**

Preferences **Simulations** Parameter Files Integration Tests

SimulationView

Query

- List all parameters (by name and value) of thom "IOUtil"
- List all simulations by parameter filename, job owner, run host/nprocs/date and - optionally - with their simulation title and PBS jobname
- List all simulations by parfilename name and run date, along with their update information sent periodically and before termination
- List all Wavetoy simulations (eg. simulations which have "wavetoy" in their parameter filenames), along with their run date and the value of IO::out_dir (if set)

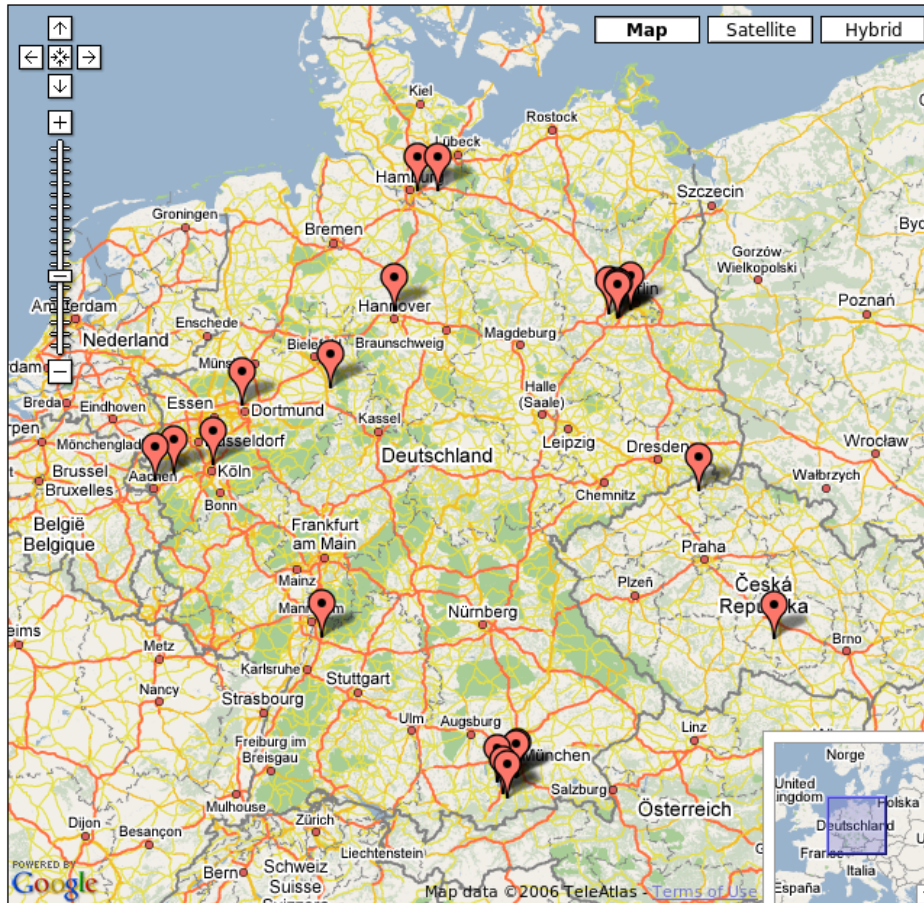
Choose Query

ParFilename	User	Host	nProcs	StartedAt	RunTitle	PBS_Jobname
CarpetIOJpeg.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:06:33 CET 2006	Carpet WaveToy Demo with HTTPD and live Visualization	
CarpetIOStreamedHDF5.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:06:39 CET 2006		
CarpetWaveToy-TwoComponents.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:06:40 CET 2006		
CarpetWaveToyCheckpointTest.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:06:52 CET 2006		
CarpetWaveToyRecoverTest.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:06:54 CET 2006		
HDF5socket_demo.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:06:59 CET 2006		
announce.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:07:26 CET 2006	Announce in the various ways	
byrange_full.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:07:27 CET 2006		
carpet-test.par	tradke	nidud.aei.mpg.de	1	Thu Nov 09 11:07:27 CET 2006		



UseCase - GridMap

GridMap (beta)



astrogrid.aei.mpg.de

Fork

default: 0/1 CPUs

Latitude, Longitude:



UseCase - Grid resources

- Information Providers for compute resources and robotic telescopes
- Integration with test-suite of compute resources
- Basis for meta-scheduling and brokering
- Data Stream Management (WG-4) has an Information Provider for Content Providers



Summary

- Metadata describes Grid Resources, Simulations, etc.
- Information Service provides Storage and query facilities
- AstroGrid-D metadata is naturally distributed
- Deliverables
 - D2.1 – Description of architecture and requirements
 - D2.2 – Production version in end of November
 - D2.3 – Implementation of Information Providers (March, 2007), WG-2 is working actively in assisting several groups (AIP, AEI, TUM, ...)