



## Geospatial workflows on the D-Grid infrastructure

Christopher Kunz  
Ralf Gröper, Tino Fleuren, Christian Grimm



OGF25  
March 4, 2009



SPONSORED BY THE

Federal Ministry  
of Education  
and Research

## Agenda

- GDI-Grid Spatial Data Infrastructure Grid
  - Project overview and goals
- GDI-Grid Architecture and Implementation
  - The GDI-Grid Architecture
  - The GDI-Grid Workflow Engine
  - Security Considerations

## Agenda

- GDI-Grid Spatial Data Infrastructure Grid
  - Project overview and goals
- GDI-Grid Architecture and Implementation
  - The GDI-Grid Architecture
  - The GDI-Grid Workflow Engine
  - Security Considerations

## GDI-Grid basics

- „Geodaten-Infrastruktur-Grid“ - „Spatial Data Infrastructure Grid“
- Project funded by Federal Ministry of Education and Research
  - Funding from 2007/07 until 2010/06
- Strong focus on sustainable development
  - 11 principal partners
    - 8 scientific institutions
    - 3 industry partners
  - Several associated partners
  - Public-private partnership

# Contact and project partners

Gabriele von Voigt, Christian Grimm

Regionales Rechenzentrum Niedersachsen (RRZN)

Leibniz Universität Hannover

Schloßwender Straße 5

30159 Hannover

<http://www.gdi-grid.de/>

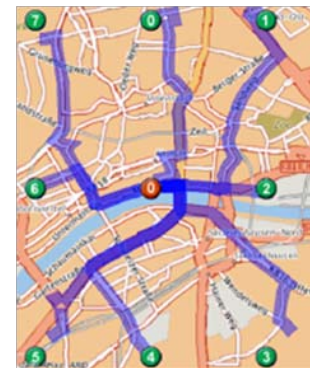


## Project goals in GDI-Grid

- Integration of existing SDI applications with the Grid will...
  - Satisfy increasing demand for computational power in SDIs
  - Facilitate access to distributed compute and storage resources
  - Connect base technologies for both worlds
    - In particular OGC web services and Globus Toolkit
  - Enable seamless processing of geospatial data on Grid infrastructures
- Development of generic services for...
  - Integration
  - Fusion
  - Management
  - Processing
- ...of geospatial data in the D-Grid infrastructure
- Performance review
  - Analysis of user requirements
  - survey for marketability and practicability of the developments

## Scenarios and proof of concept

- Three practical scenarios serve as the „proof of concept“ for the GDI-Grid project
  1. Flood simulation
  2. Noise propagation simulation
  3. Optimized routing for disaster management and emergency services
- These scenarios cover a wide area of SDI applications
- Demand for data and computing resources in all scenarios
- Large potential benefit by using the Grid



## GDI-Grid implementation

- Practical implementation under development
  - Grid connectors for OGC web services
  - Solutions for authorization and rights management
  - Example workflows for scenarios
- Implementation is web service based
  - WSRF and OGC web services
  - Workflow engine orchestrates services
- Challenges for implementation: Bridging web service standards
  - More detail see below

## Agenda

- GDI-Grid Spatial Data Infrastructure Grid
  - Project overview and goals
- **GDI-Grid Architecture and Implementation**
  - The GDI-Grid Architecture
  - The GDI-Grid Workflow Engine
  - Security Considerations

## Requirements in GDI-Grid

- Support OWS and Grid clients
  - OWS not compatible with WSRF (or other WS standards)
  - WSRF-Based GeoServices bridge between both worlds
- Challenge 1: Enable Workflow Engines
  - OpenBPEL enhanced by Uni Marburg to support WSRF
- Challenge 2: Security
  - Grid: Sophisticated Grid Security Infrastructure
  - OWS: No security mechanisms specified at all

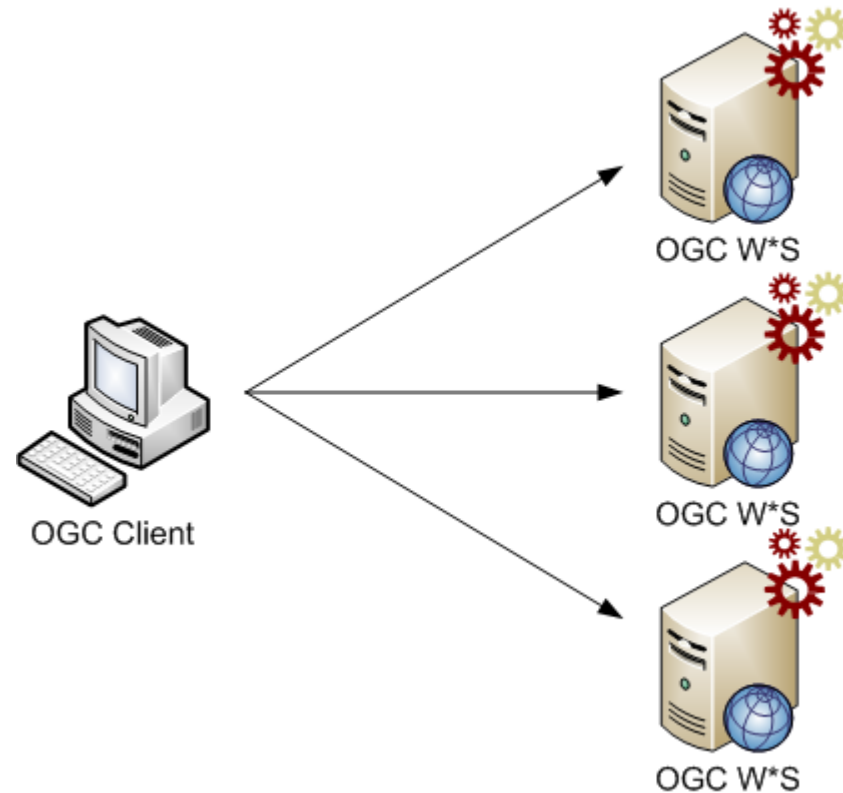
## Grid and geospatial web services

- Current Grid infrastructures rely on web services
  - Relevant standard: WSRF - Web Services Resource Framework
- Spatial data infrastructures also rely on web services
  - Standards body: Open Geospatial Consortium
  - Standards for different types of web service
    - nomenclature mostly W\*S (Web \* Service)
    - \* = {Feature, Coverage, Map, ...}
- Adaption between these web service standards
  - Connector between Grid and non-Grid web services needed
  - Security?
- Orchestration of web services to workflows
  - BPEL for workflow description
  - Workflow engine for workflow composition and management

## Agenda

- GDI-Grid Spatial Data Infrastructure Grid
  - Project overview and goals
- GDI-Grid Architecture and Implementation
  - **The GDI-Grid Architecture**
  - The GDI-Grid Workflow Engine
  - Security Considerations

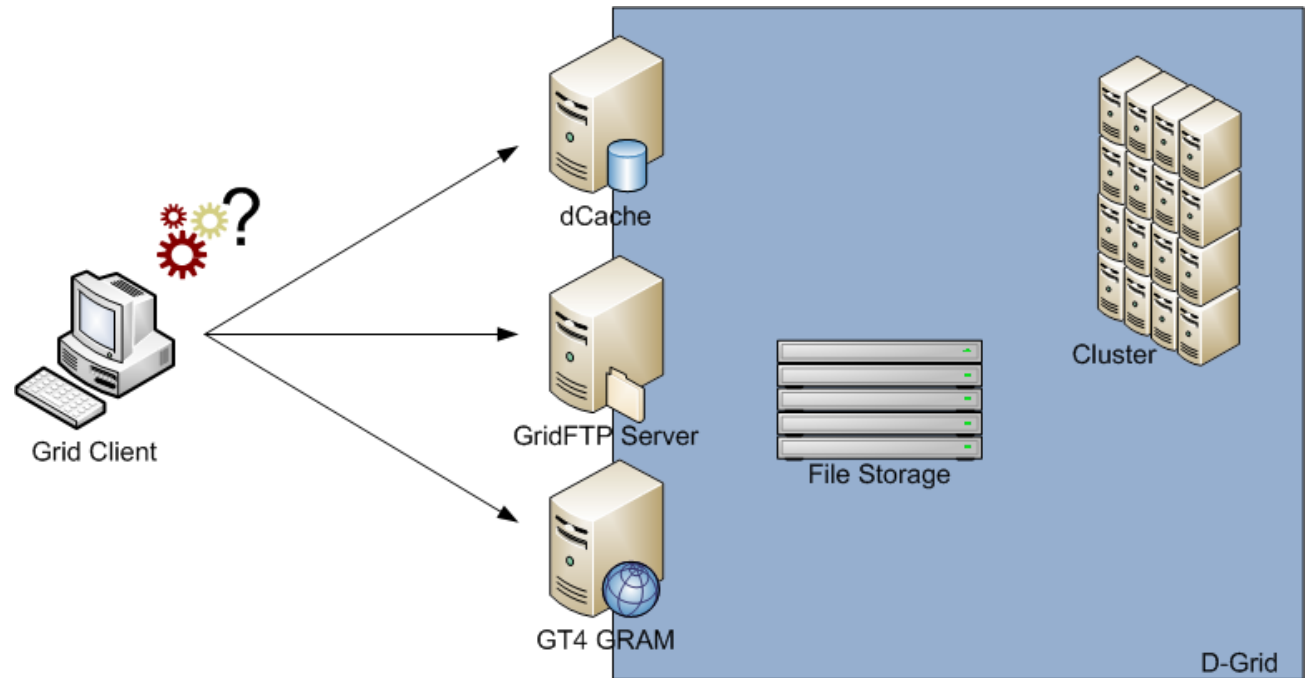
## OGC Web Services without the Grid




- No Grid features
- No security
- No workflows
- Domain-specific services
- Program Logic typically locally on service

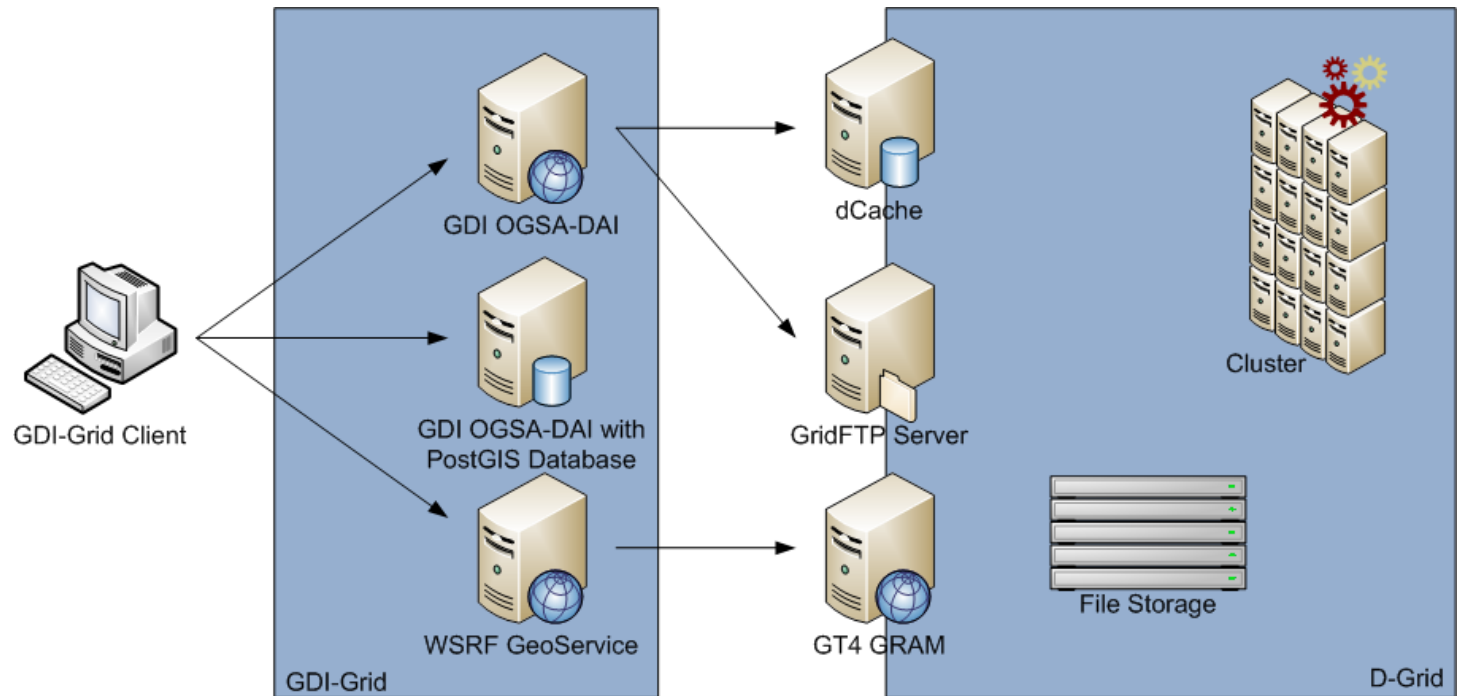



# The Grid without Geo-Services



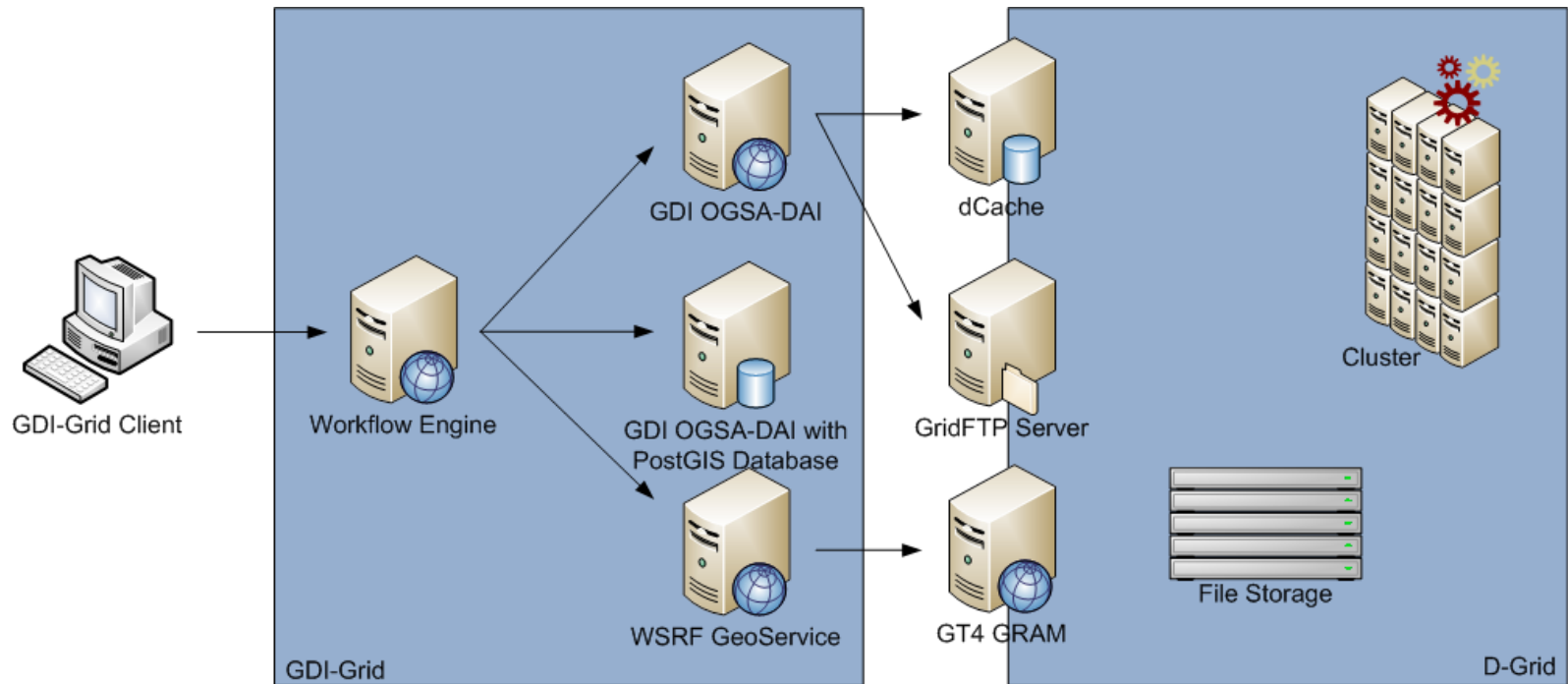
- Grid features
- GSI security
- No workflows
- No Domain-specific services
- Where to put the program logic? 


## WSRF GeoServices and the Grid



- Grid features
- GSI security
- No workflows
- Domain-specific services
- Program logic pre-installed on D-Grid compute resources 

## WSRF GeoServices with Workflow Support



- Grid features
- GSI security
- WSRF-Enabled Workflow Engine orchestrates GeoServices and Data Access
- Domain-specific services
- Program logic pre-installed on D-Grid compute resources 



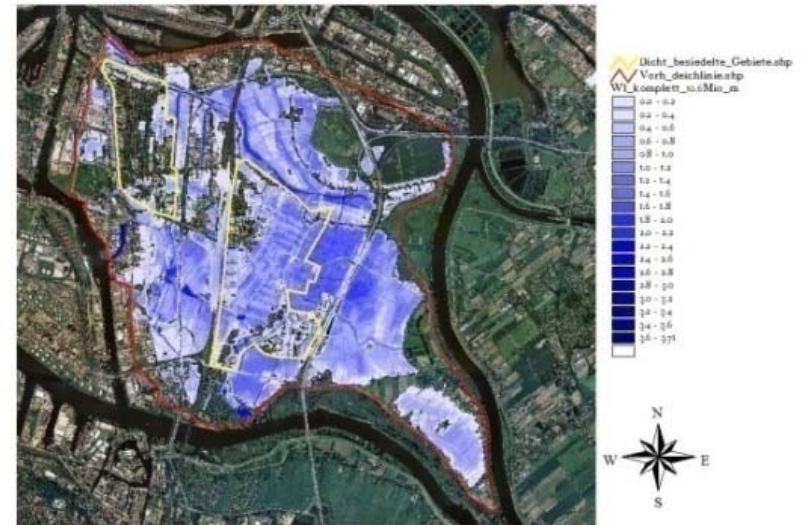
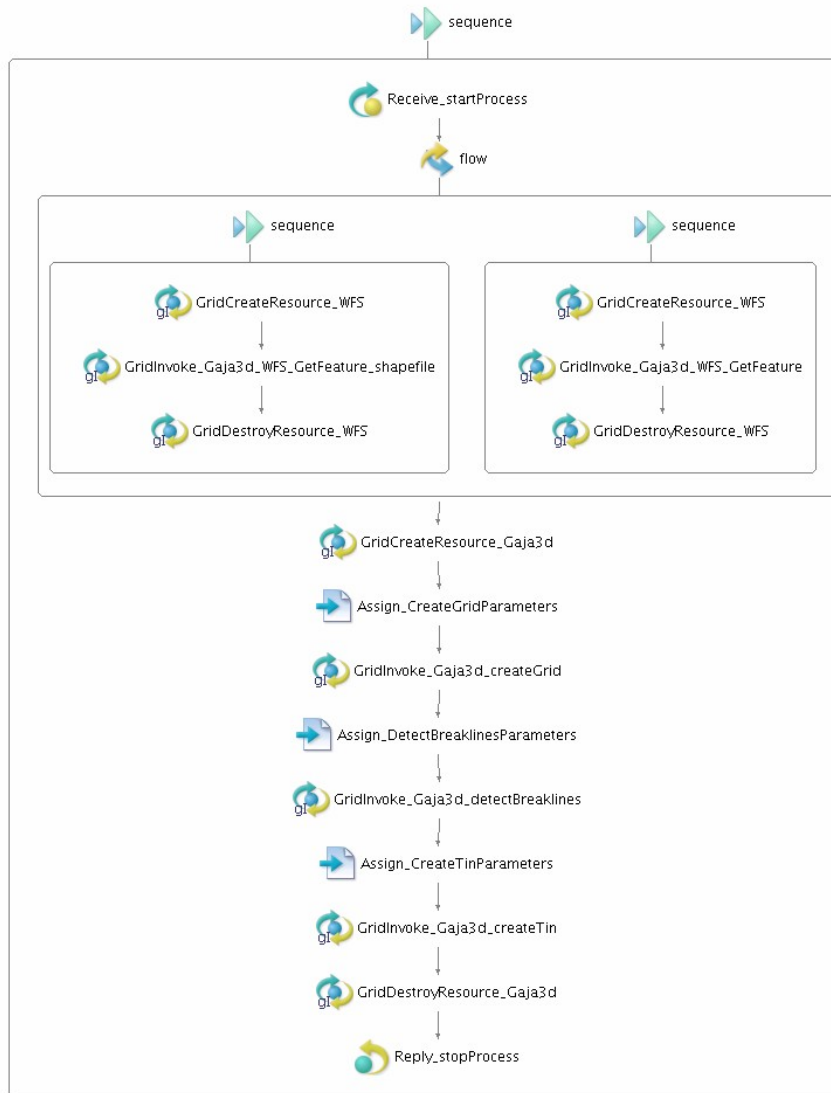
## Agenda

- GDI-Grid Spatial Data Infrastructure Grid
  - Project overview and goals
- GDI-Grid Architecture and Implementation
  - The GDI-Grid Architecture
  - **The GDI-Grid Workflow Engine**
  - Security Considerations

# Geospatial Workflows

- Workflow design should be easy
  - No knowledge of underlying service technologies needed
  - Supported by graphical tools
  - Quick adaption to new requirements
  - Using validated functionalities to construct new applications
  - Reuse workflows by other users
  - No new middleware should be implemented
    - ~> Workflow design by orchestrating services
- ~> **Workflow Description Languages**
- Business Process Execution Language (BPEL)
  - De-facto industry standard
  - supported by several vendors like Microsoft, IBM, Sun Microsystems
  - Open source implementations of BPEL engines

# Workflow Scenario Flood simulation



## Different types of services

### 1. External Standard Web services

- WSDL, SOAP
- BPEL is designed for these Web services
- Challenge: Not implemented by OGC

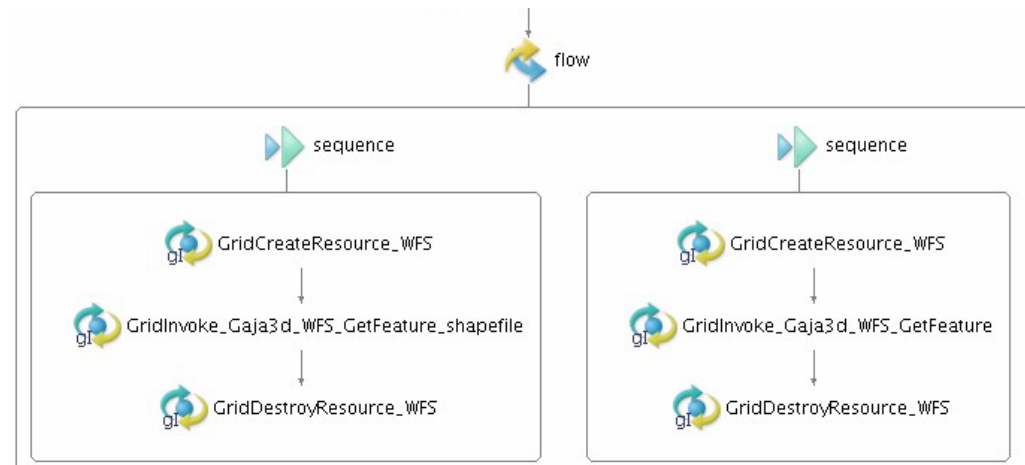
### 2. WSRF-based Grid services

- Web Services Resource Framework (WSRF)
  - Stateful service
  - Dynamic service creation
  - dynamic endpoint references (EPR)
  - Grid Security Infrastructure
- Challenge: Grid-enabling of existing services

## Different types of services

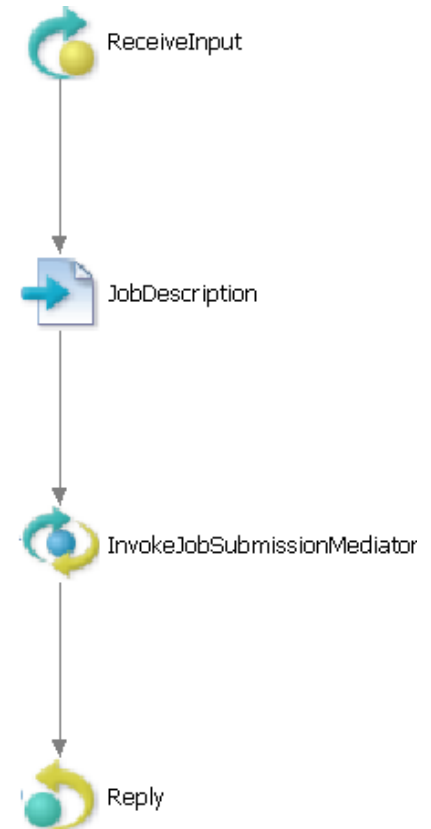
### 3. External OGC Web services

- Services offered by third parties
  - Third parties charge fees
  - Third parties do not offer Grid-enabled versions
  - Third parties hide their source code
- Numerous supporting services
  - Grid-enabling can be time-consuming
  - Not all services need the compute power of the Grid
- Challenges
  - Leaving the Grid boundaries
  - Different security concepts
  - OGC services can not be executed by BPEL



## Different types of services

- Command-line tools and shell scripts
  - Legacy (geospatial) applications that run on a cluster
  - Jobs executed by WS-GRAM
  - WSGRAM
    - part of the Globus Toolkit
    - submits, monitors and cancels jobs
    - provides a uniform interface to batch scheduling systems, such as PBS, SGE, or Condor
- Challenges
  - BPEL cannot call WS-GRAM services
  - BPEL cannot handle notifications



# Challenge 1

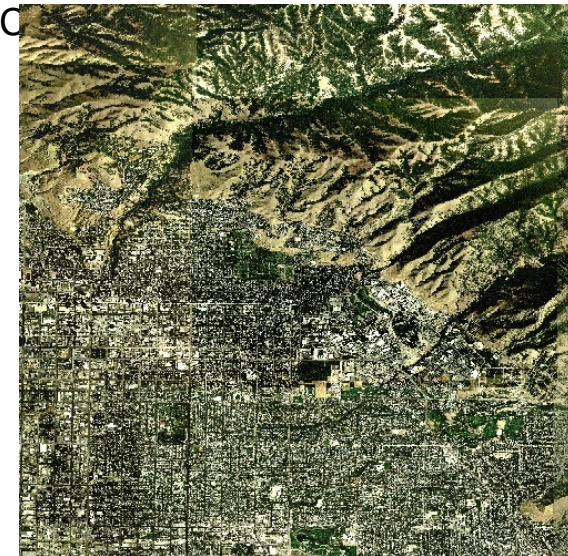
## BPEL and WSRF-based Services

- BPEL Enhancement: Mage BPEL4Grid
  - University of Marburg, Tim Dörnemann et al. [2007]
  - (D-Grid, InGrid)
- Features of Mage BPEL4Grid
  - Support for WSRF and Grid Security Infrastructure (GSI)
  - Ability to invoke both Grid and Web Services,
  - Encrypt and sign messages and delegate credentials.
  - Monitoring of the lifetime of proxy certificates
  - Automatically renew proxy certificates when using MyProxy

## Challenge 2

# Accessing (Grid external) OGC services

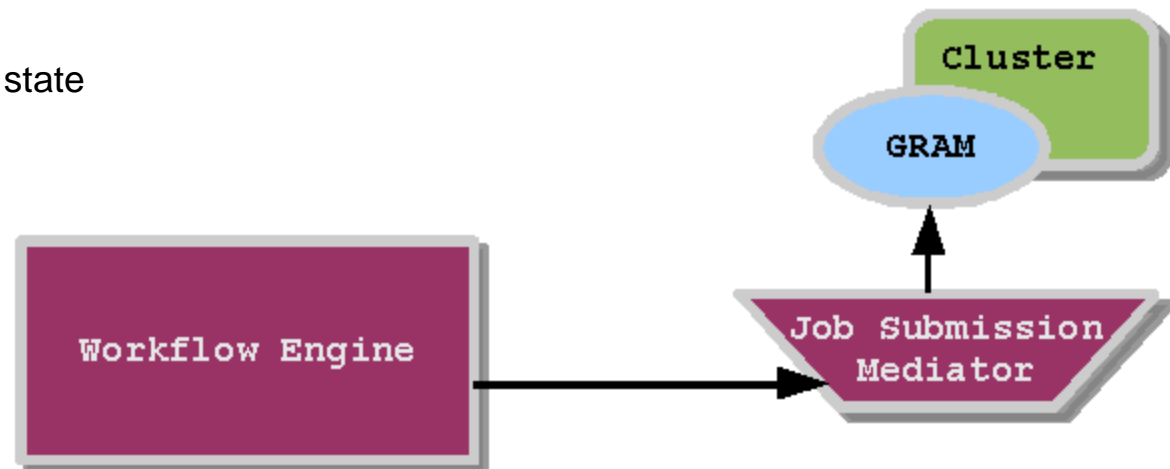
- Example: Web Coverage Service (WCS)
  - OGSA-DAI W\*S adapter developed by SEE-GEO (Edinburgh University)
- Input parameters
  - OGSA-DAI's base URI
    - e.g. <https://ogsadai.uni-hannover.de:8443/wsrf/services/dai/>
  - HTTP-GET request
    - e.g. [http://localhost:8080/deegree-wcs/services?service=WCS](http://localhost:8080/deegree-wcs/services?service=WCS&request=GetCoverage&coverage=saltlakecity&format=jpeg)
      - ...  
&request=GetCoverage
      - ...  
&coverage=saltlakecity
      - ...  
&format=jpeg
  - Destination of the result
    - e.g. <gsiftp://gridftp.uni-hannover.de:2811/tmp/saltlake.jpg>



## Challenge 3

# BPEL and WS-GRAM: Job Submission Mediator (JSM)

- The JSM is a Grid service that
  - Can be integrated as an activity in a BPEL workflow
  - Receives the job description
  - Triggers job execution
  - Submits jobs to WS-GRAM or scheduling service
  - Subscribes to WS-GRAM notification service
  - Translates Messages between BPEL process and WS-GRAM
    - Error handling
    - Monitoring of job state

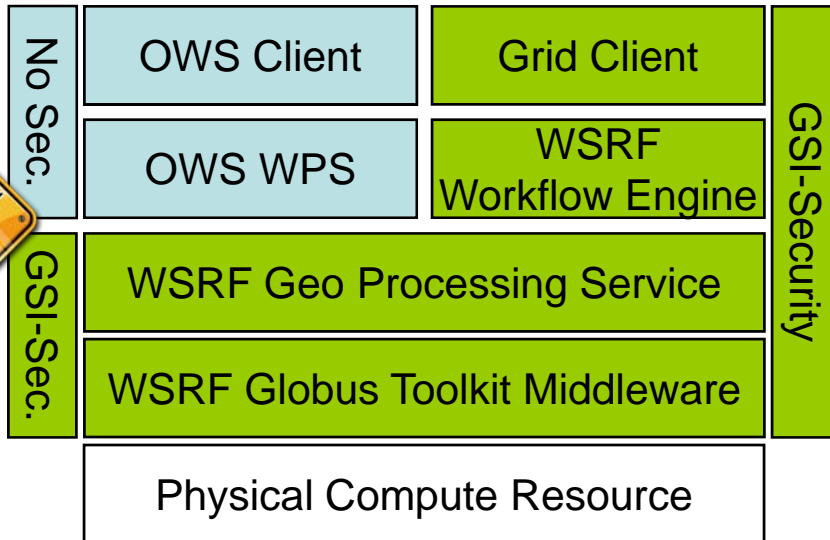


## Agenda

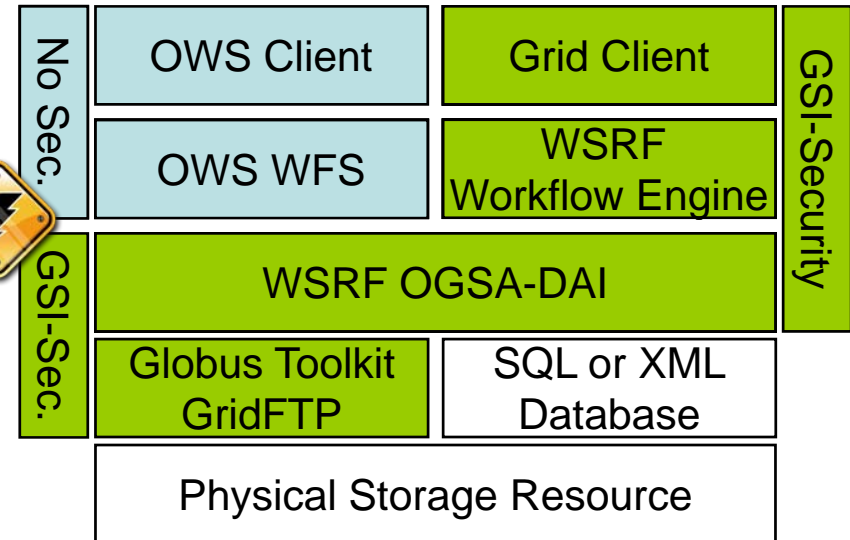
- GDI-Grid Spatial Data Infrastructure Grid
  - Project overview and goals
- GDI-Grid Architecture and Implementation
  - The GDI-Grid Architecture
  - The GDI-Grid Workflow Engine
  - **Security Considerations**

## A Look on Security

### Compute



### Data Access



OGC: Open Geospatial Consortium

OWS: OGC Web Service

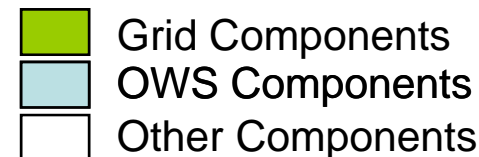
WPS: Web Processing Service

WFS: Web Feature Service

GSI: Grid Security Infrastructure

WSRF: Web Services Resource Framework

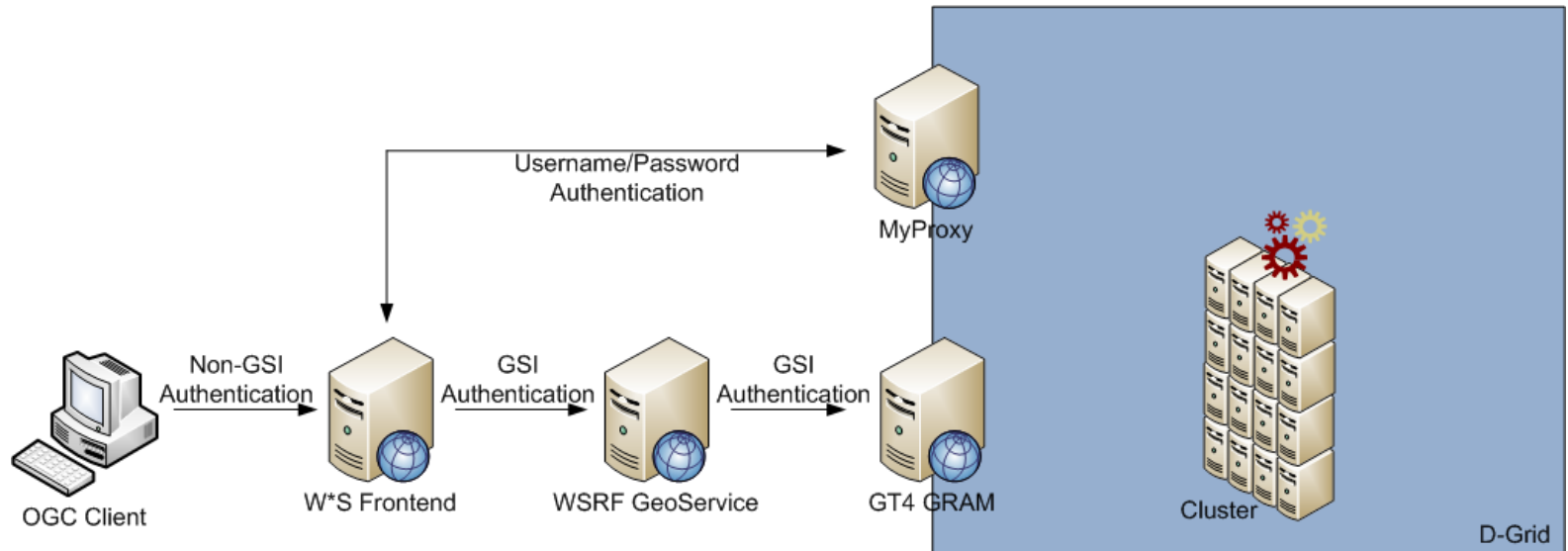
OGSA-DAI: Open Grid Service Architecture – Data Access and Integration



## Implementing the GSI in OGC Web Services

- Situation for GDI-Grid Components
  - Acceptance of proxy credentials as an authentication token
    - Or: Authentication by other means and acquire proxy credentials from MyProxy
  - Authenticate to D-Grid infrastructure using delegated or acquired Proxy Credential
- For other legacy OGC components
  - connectors / adaptors to D-Grid resources needed
- Approach in GDI-Grid
  - GDI-Grid GeoServices are WSRF Services deployed in Globus Container
    - support GSI out of the box
  - MAGE GDT Service Generator offers easy integration of GSI

## Using MyProxy to Acquire Proxy Credential



- User has to delegate Proxy Credential to MyProxy service first
  - Graphical client tool developed at RRZN
- Username/Password for MyProxy are given to W\*S as URL parameters
  - SSL encrypted data channel is obligatory
  - Not standardized by OGC
- W\*S Frontend acquires Proxy Credential from MyProxy and delegates it to WSRF GeoService

## Summary

- Geospatial Applications in the Grid pose several challenges
  - Orchestration
    - Neither OGC nor Globus offer support for workflows
  - Security
    - GSI vs. no Security
  - Bridging different incompatible standards
    - OGC Web Services: Standard for geospatial Web services
    - WSRF Web Services
- Orchestration of WSRF Web Services
  - OpenBPEL + WSRF enhancements
- Integration with D-Grid infrastructure
  - Integrate existing Grid Services
  - Make domain specific program logic available on D-Grid's resources