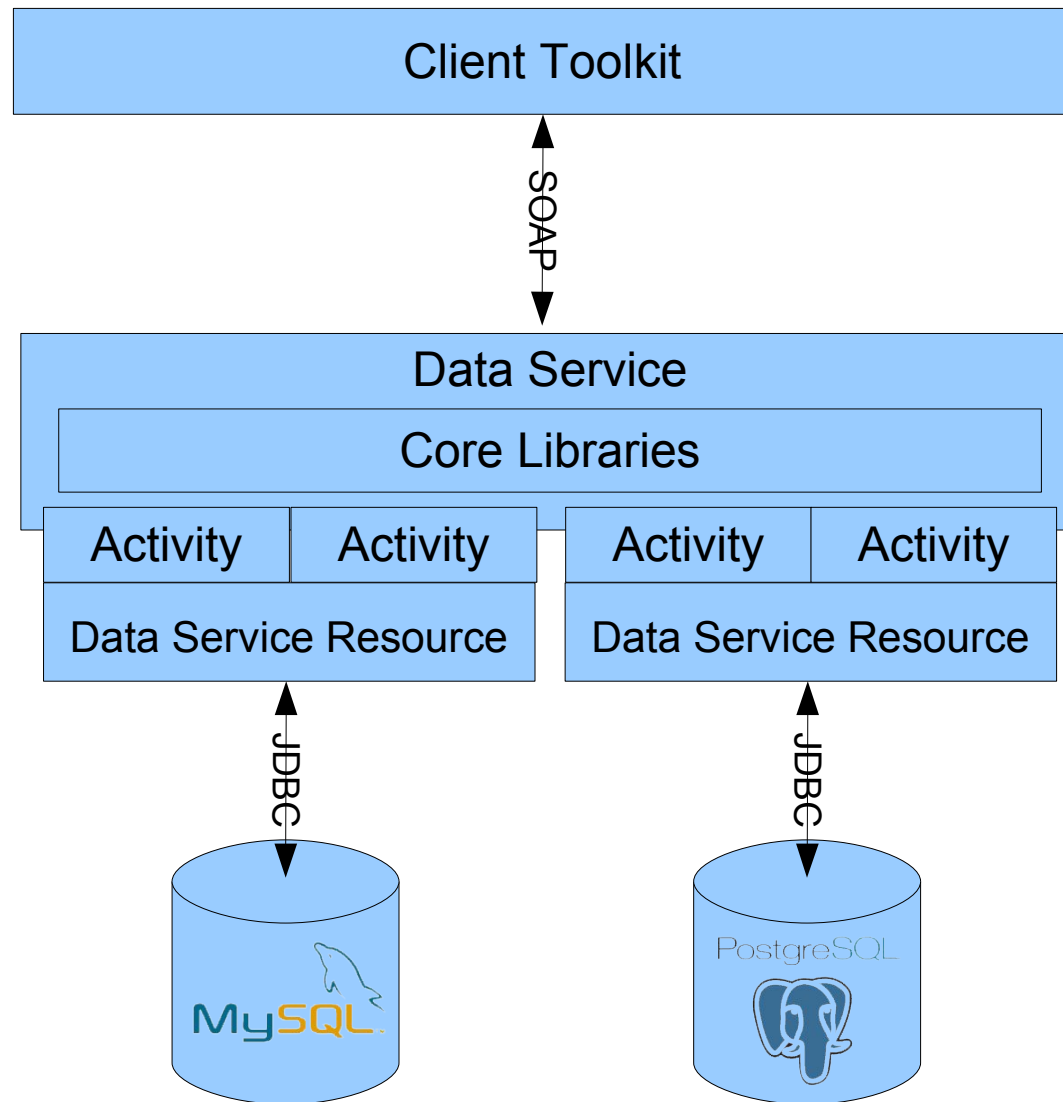


# OGSA-DAI

Schnittstelle zu Datenbanken im  
Grid

Thomas Jejkal

- Open Grid Services Architecture - Data Access and Integration
- 
- Entwickelt am EPCC: Edinburgh Parallel Computing Centre
- 
- Standardisierter Zugriff auf Daten(-banken) im Grid
- 
- Integriert in Globus Middleware als Web Service
- 
- Verbergen der Lokation und der Art von Daten(-banken)
- 
- OGSA-DAI WSRF 2.2 im D-Grid verfügbar



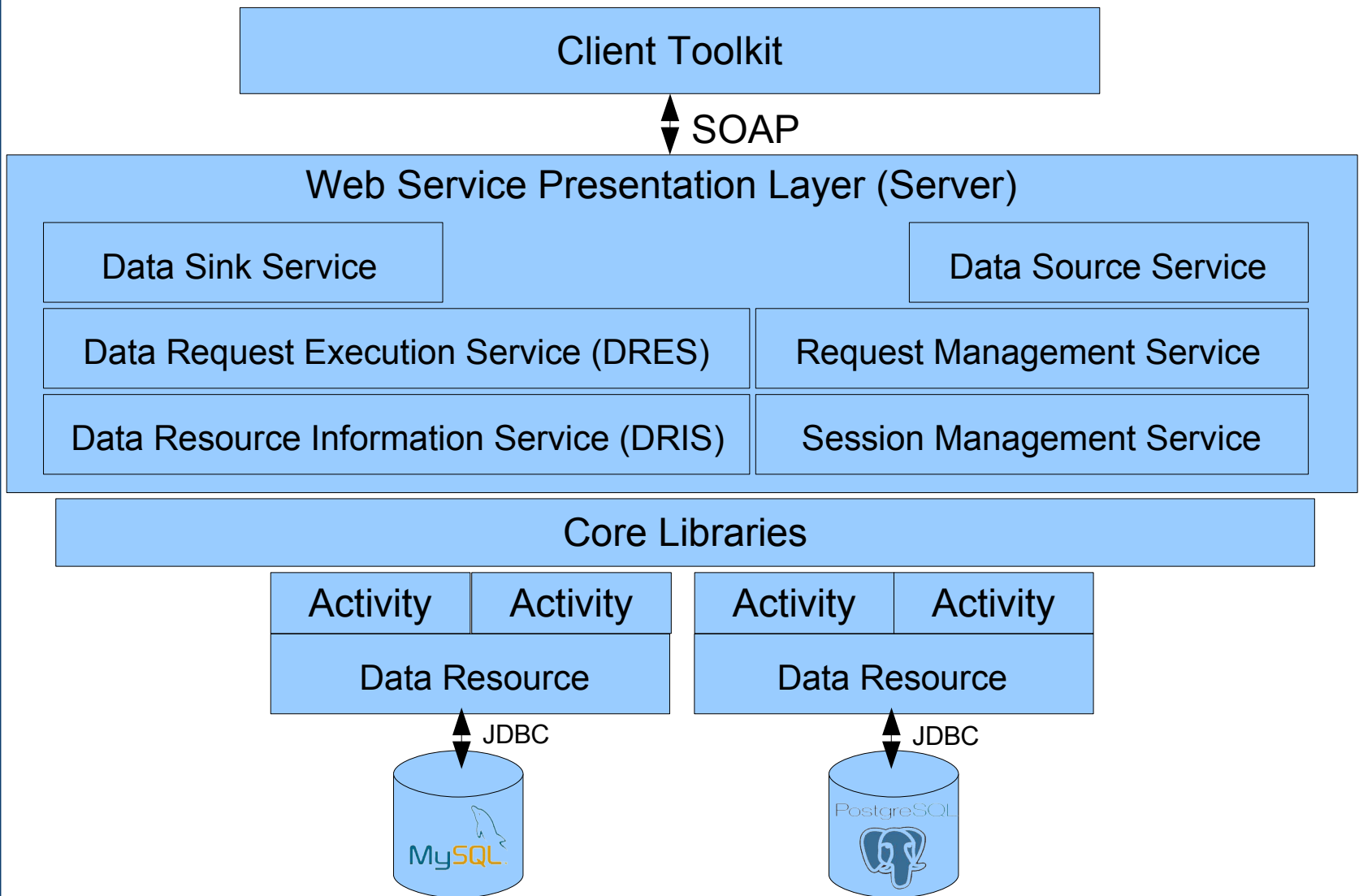
- Einheitliche Schnittstelle für alle Datenressourcen
- Abstrahierung von Typ und Lokation von Datenressourcen
- Erweiterung des Zugriffs um Grid Security Features
- Umfangreiches Client API für intuitiven Zugriff
- Nahezu beliebig und einfach erweiterbar
- Einfach in bestehenden Anwendungen integrierbar

- Beschreibung einer Referenzinstallation
  - Allgemeine Ressourcen für alle Communities
  - Einbindung von Sicherheitsmechanismen des D-Grid
  - Einbindung in D-Grid Monitoring
- Referenzinstallation auf 7 Sites verfügbar
  - GWD Göttingen, Uni Dortmund, FZ Jülich, Uni Heidelberg, HLR Stuttgart, FZ Karlsruhe, TU Dresden

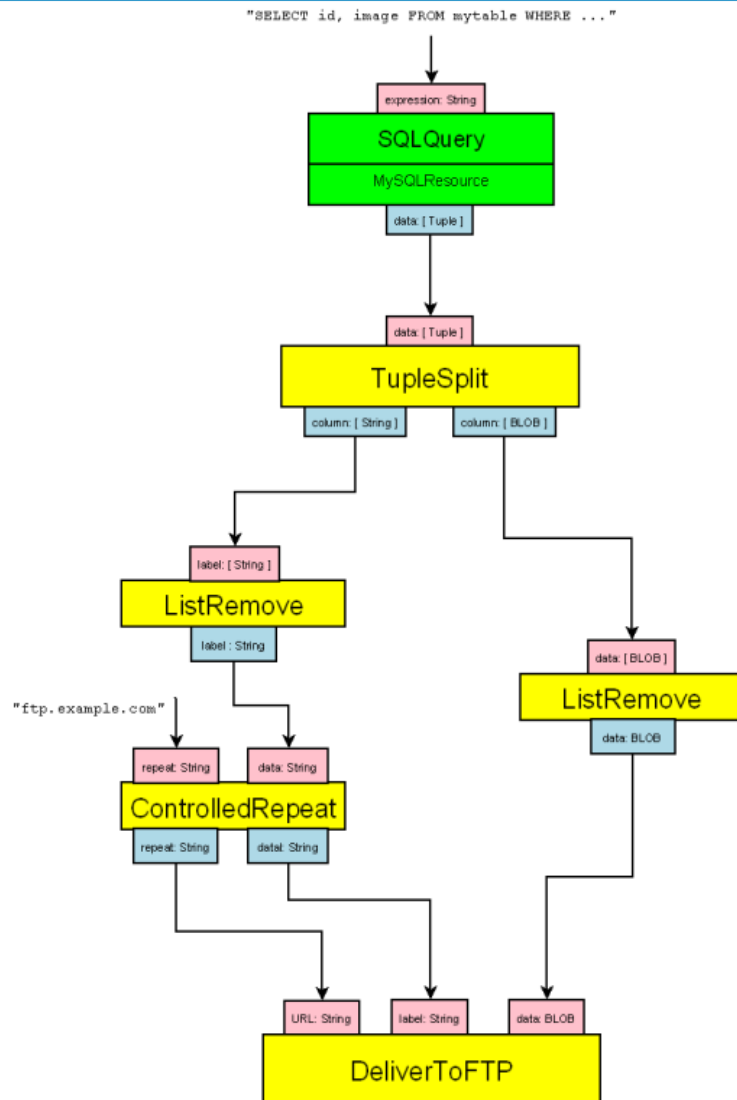
|          |                           |                                                     |
|----------|---------------------------|-----------------------------------------------------|
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |
| OGSA-DAI | juggle-ogsa.fz-juelich.de | DataResource with productName MySQL (38 activities) |

- Komplettes Top-To-Bottom Redesign
  - Anpassungen bei bisherigen Applikationen notwendig!
- Standardisierung von Interfaces und Extension Points
- Umstellung der Service und Ressourcenstruktur
  - Funktionsgebunden (Request, Information, Sink, Source, Management)
- Besserer Workflow-Support
- Management von Client-Seite aus möglich
- Volle Kompatibilität zu WSRF Standard
- Dynamische Konfiguration

# OGSA-DAI 3.0 Architektur



# Workflows mit OGSA-DAI



- Komplexe Workflows umsetzbar
- Bei Bedarf Site übergreifend
- Synchron und Asynchron
- Zwischen unterschiedlichen Datenressourcen möglich

- Anbindung an DGUS (<http://dgus.d-grid.de>) (April '08)
- Umstieg auf OGSA-DAI 3.0 (September '08)
- Bereitstellung von allgemeinen Clients (September '08)
- Behandlung der Problematik unterschiedlicher SQL Syntax (September '10)



<http://www.ogsadai.org.uk>



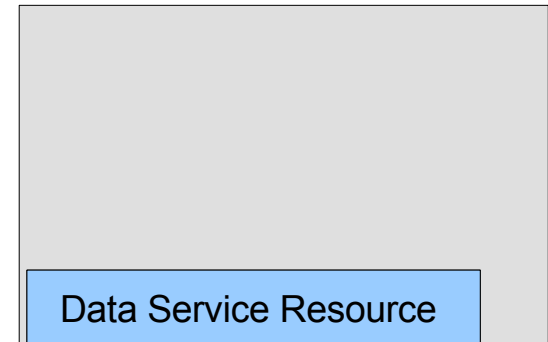
<http://dgi.d-grid.de/index.php?id=77>

<http://dgus.d-grid.de>



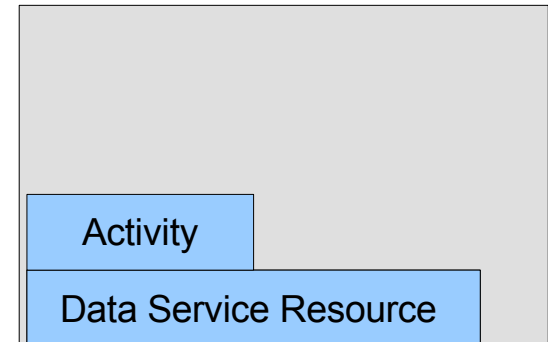
## Data Service Ressource

- Datenquellen/-senken auf die über OGSA-DAI zugegriffen wird
- Fokus auf relationale Datenbanken (MySQL, PGSQL, Oracle)
- Beliebig um neue Datenressourcen erweiterbar
- Zugriffsmöglichkeiten definiert durch Activities



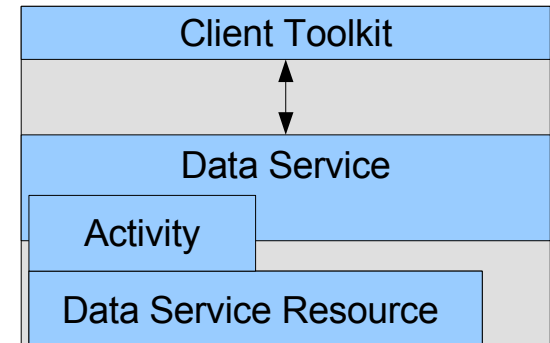
## Activities

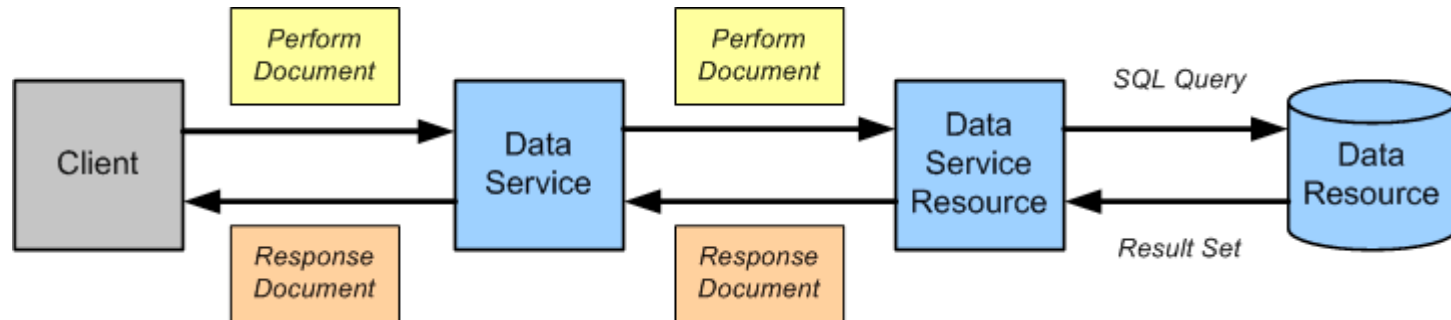
- Mögliche Interaktionen mit einer Datenressource
- Teilweise spezifisch für Ressourcen-Typen
- Unterteilt in Relational, XML, File, Delivery, Transformation, Resource Management, Indexed File Activities und Multi-Resource
- Kombination zu einfachem Workflow möglich



## Data Service

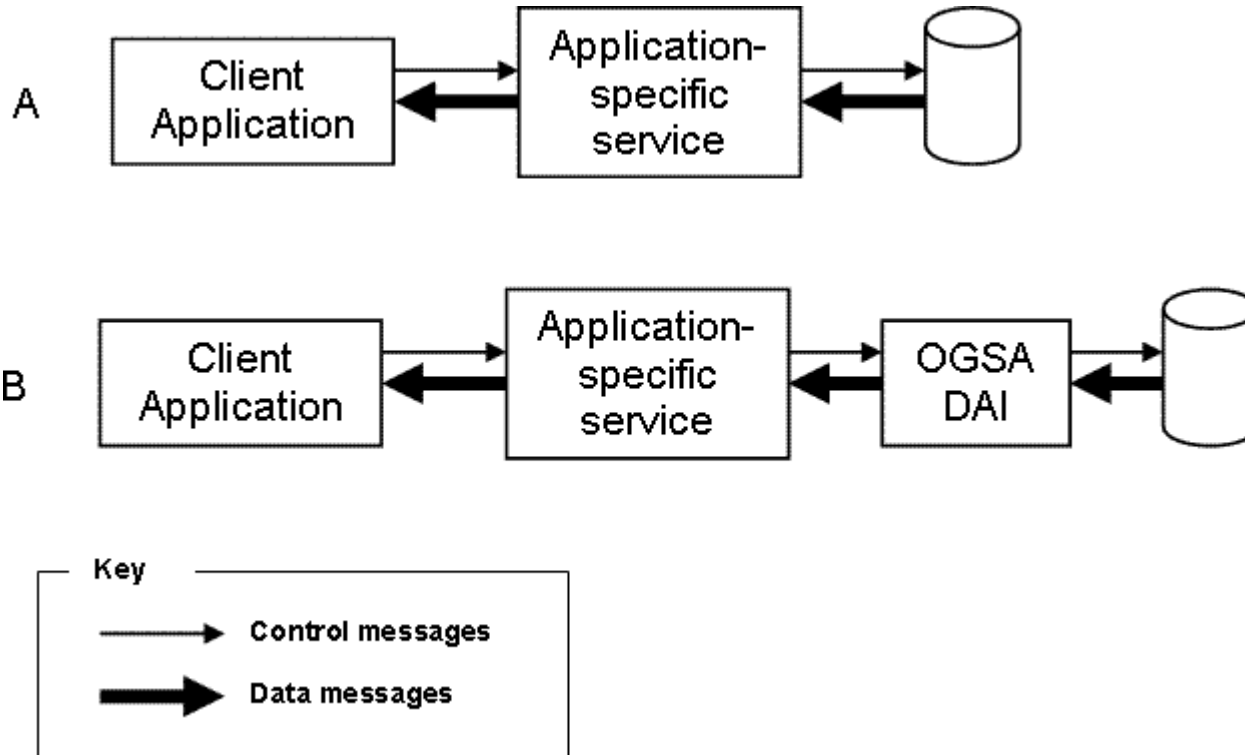
- WRSF-konformer Web Service
- Integriert in den Globus Toolkit 4 Web Service Container
- Schnittstelle zu Datenressourcen
- Umfangreiche Möglichkeiten zur Autorisierung und Authentifizierung basierend auf Grid Security Infrastructure (GSI)
- Zugriff über Client-Toolkit





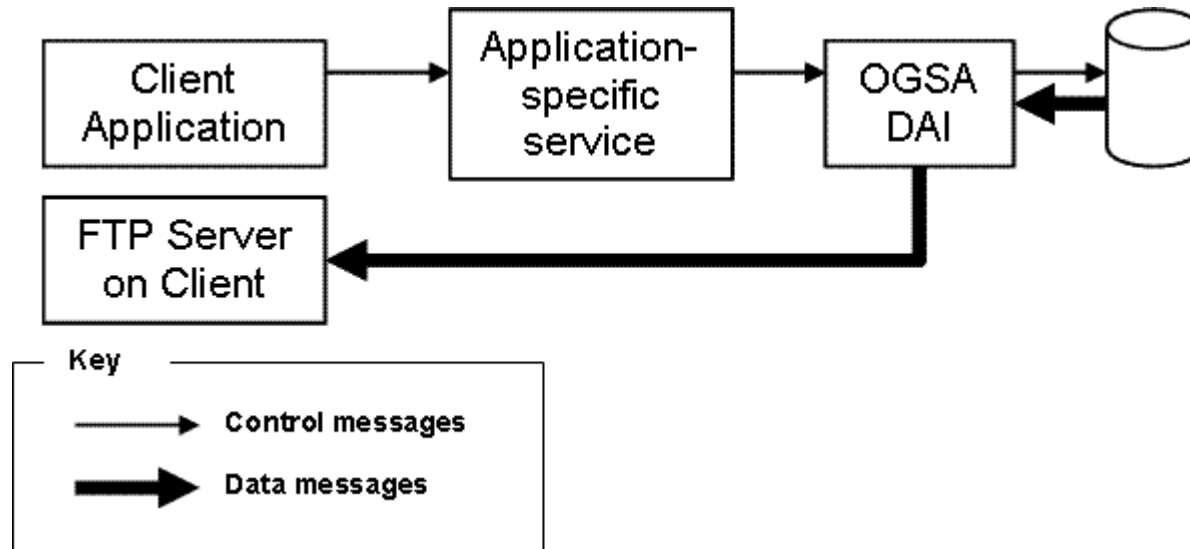
- Dokumentbasierter Zugriff
- Perform Document enthält linearen Workflow aus Activities
- Response Document kann Ergebnisdaten enthalten
- Delivery Activities dienen zur Übertragung von Ergebnisdaten zu z.B. GridFTP, FTP oder SMTP Servern

# Anwendungsszenarien (1)

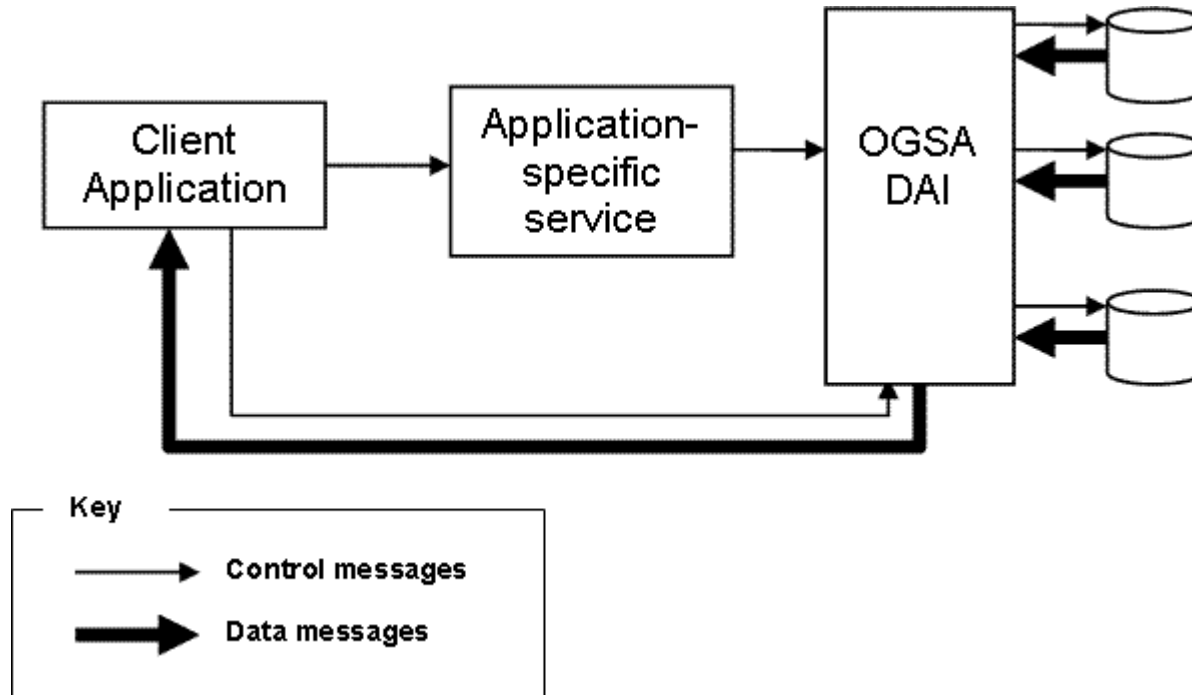


- Integration in vorhandene Services für den Datenzugriff
- Vorhandene Datenressourcen bleiben unverändert
- Vollkommen transparent für Client Anwendungen

# Anwendungsszenarien (2)



- Ergebnisdaten als File an beliebigen (Grid-)FTP Server geliefert
- Zur Weiterverwendung durch andere Applikationen (z.B. Grid Jobs)
- Sowohl Datenzugriff als auch File Transfer transparent für Client



- Gleichzeitiger Zugriff auf mehrere Ressourcen über einen OGSA-DAI Service
- Komplette Abstrahierung von Lokation und Typ der Datenressourcen