

# GISolve

*A Grid-based problem solving environment for computationally intensive geographic information analysis*

Shaowen Wang

Department of Geography &  
Grid Research & educatiOn group @ IoWa (GROW)  
ITS Academic Technologies – Research Services  
The University of Iowa  
Iowa City, IA 52242, USA

*July 24, 2005*

# Collaborators

- Marc P. Armstrong, Professor, Department of Geography and Program in Applied Mathematical and Computational Sciences
- Jun Ni, Ph.D., Associate Research Scientist, ITS Academic Technologies – Research Services
- Yan Liu, Graduate Research Assistant, ITS Academic Technologies – Research Services

# Purpose

Purpose

Background

Design

Implementation

Case Study

Evaluation

Conclusions

- To address why the Grid is important to geographic information analysis research
- To demonstrate the **design** and **implementation** of GISolve – a Grid-based problem solving environment for computationally intensive geographic information analysis

# Background

Purpose

Background

Design

Implementation

Case Study

Evaluation

Conclusions

- The amount of geographic information
  - Ever increasing
    - Application driven
  - GPS, LBS, RS
- Computationally intensive geographic analysis
  - Heuristic search
  - Simulation
  - Optimization
  - Statistical methods
- Grid computing
  - Cyberinfrastructure
  - Open Grid Service Architecture (OGSA)
    - Web services
- Problem solving environments
  - Grid portals

# GIScience Grid Portal

Purpose

Background

Design

Implementation

Case Study

Evaluation

Conclusions

The screenshot shows the 'GIScience Grid Portal' interface. At the top, it says 'Grid-Based Geo-Middleware (GGM)' with 'Help' and 'Logout' buttons. The main content area is titled 'Select a GIScience Computational Service'. It contains two options: 'Inverse Distance Weighted Interpolation (IDW)' with an unchecked checkbox, and 'Spatial Statistics (SS)' with a checked checkbox. A 'NEXT >>>' button is located at the bottom right of the selection area.

The screenshot shows the 'GIScience Grid Portal' configuration screen. At the top, it says 'Grid-Based Geo-Middleware (GGM)' with 'Help' and 'Logout' buttons. The main content area is titled 'DSA' and 'Decomposition and Scheduling'. Under 'Decomposition', there are two options: 'Spatially Adaptive' with a checked checkbox and 'Non-Adaptive' with an unchecked checkbox. Under 'Scheduling', there are two options: 'Static' with a checked checkbox and 'Dynamic' with an unchecked checkbox. A 'SUBMIT' button is located at the bottom right of the configuration area.

# Grid Resources – A User View

Purpose

Background

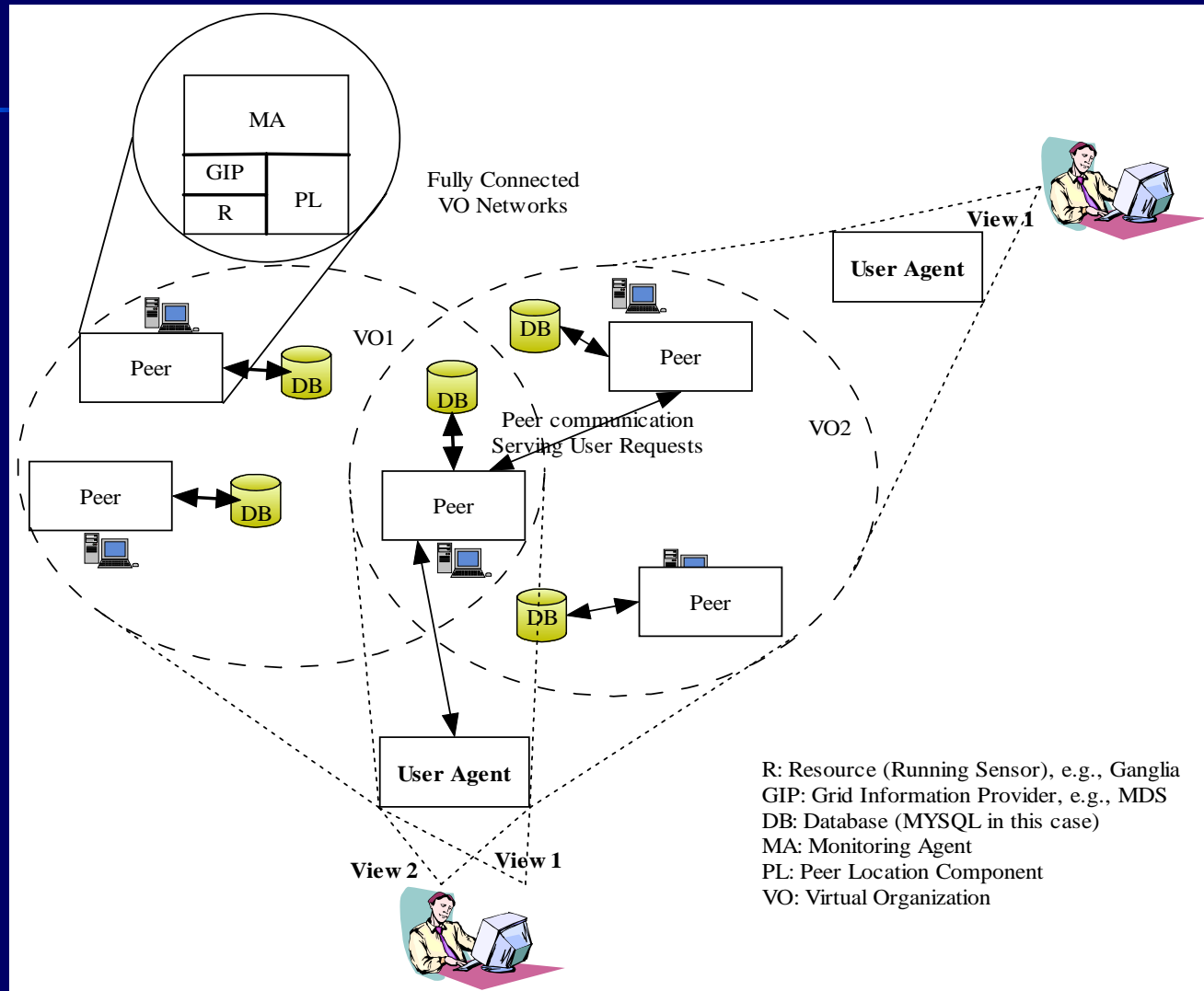
Design

Implementation

Case Study

Evaluation

Conclusions



# Grid Complexity

Purpose

- Grid middleware limitation

- Provides a selected set of protocols and services
  - Not directly user- or application- oriented

Background

Design

Implementation

- Heterogeneous

- Dynamic

- Administrated in different security domains

- Faults

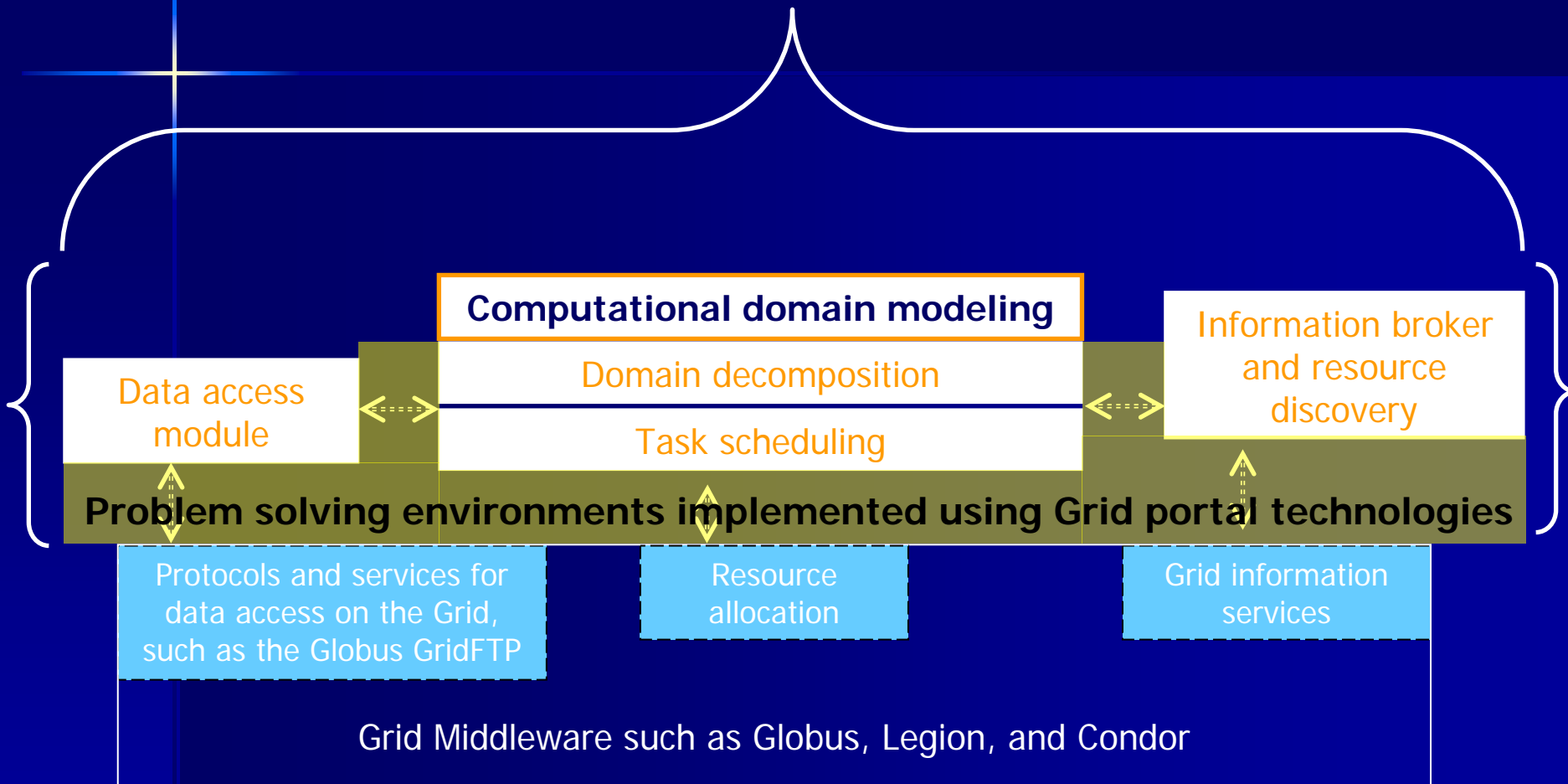
- Usually unpredictable

Case Study

Evaluation

Conclusions

# GISolve Architecture





# GISolve Services

Purpose

Background

Design

Implementation

Case Study

Evaluation

Conclusions

- Security
- Decomposition and task scheduling
- Geographic data access
- Resource information brokering

# Classification of GISolve Services

Purpose

Background

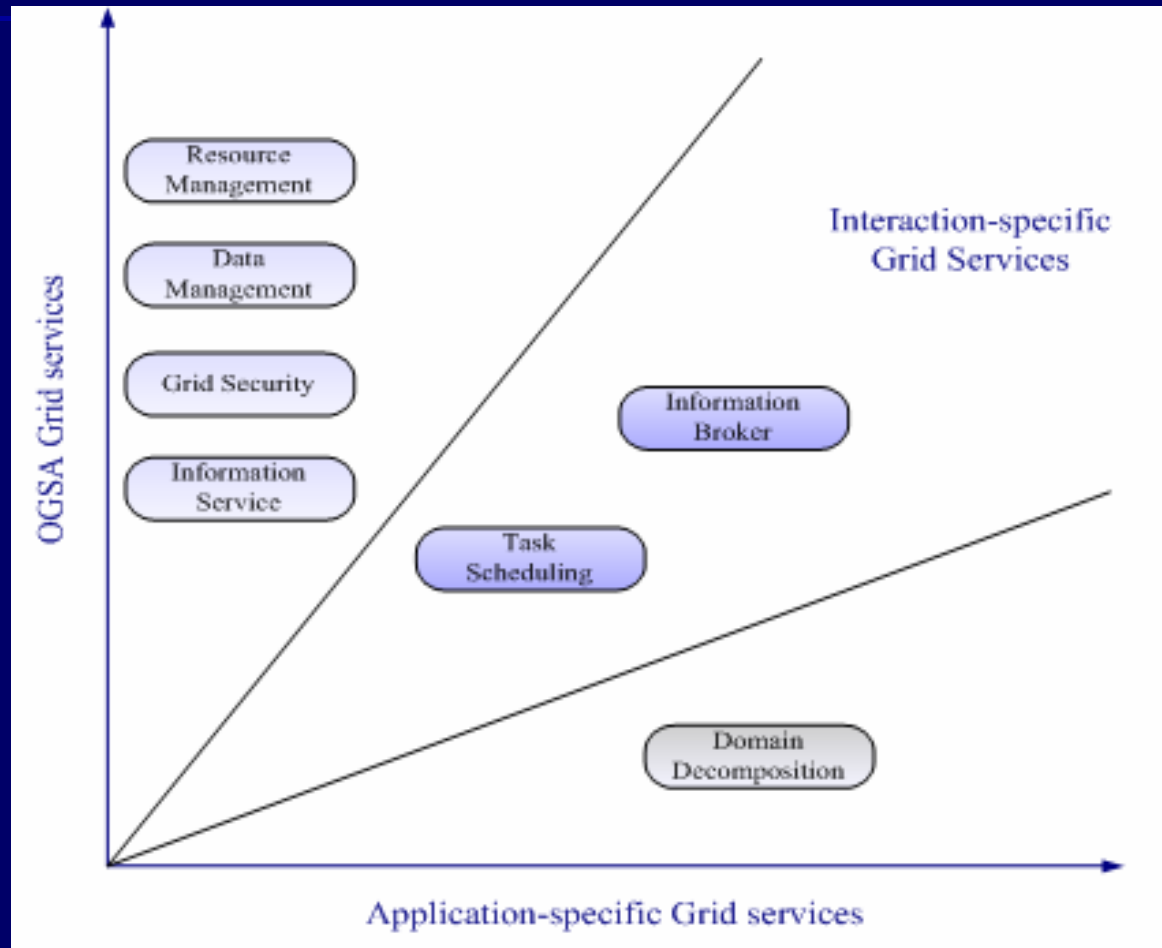
Design

Implementation

Case Study

Evaluation

Conclusions



# GISolve Workflow

Purpose

Background

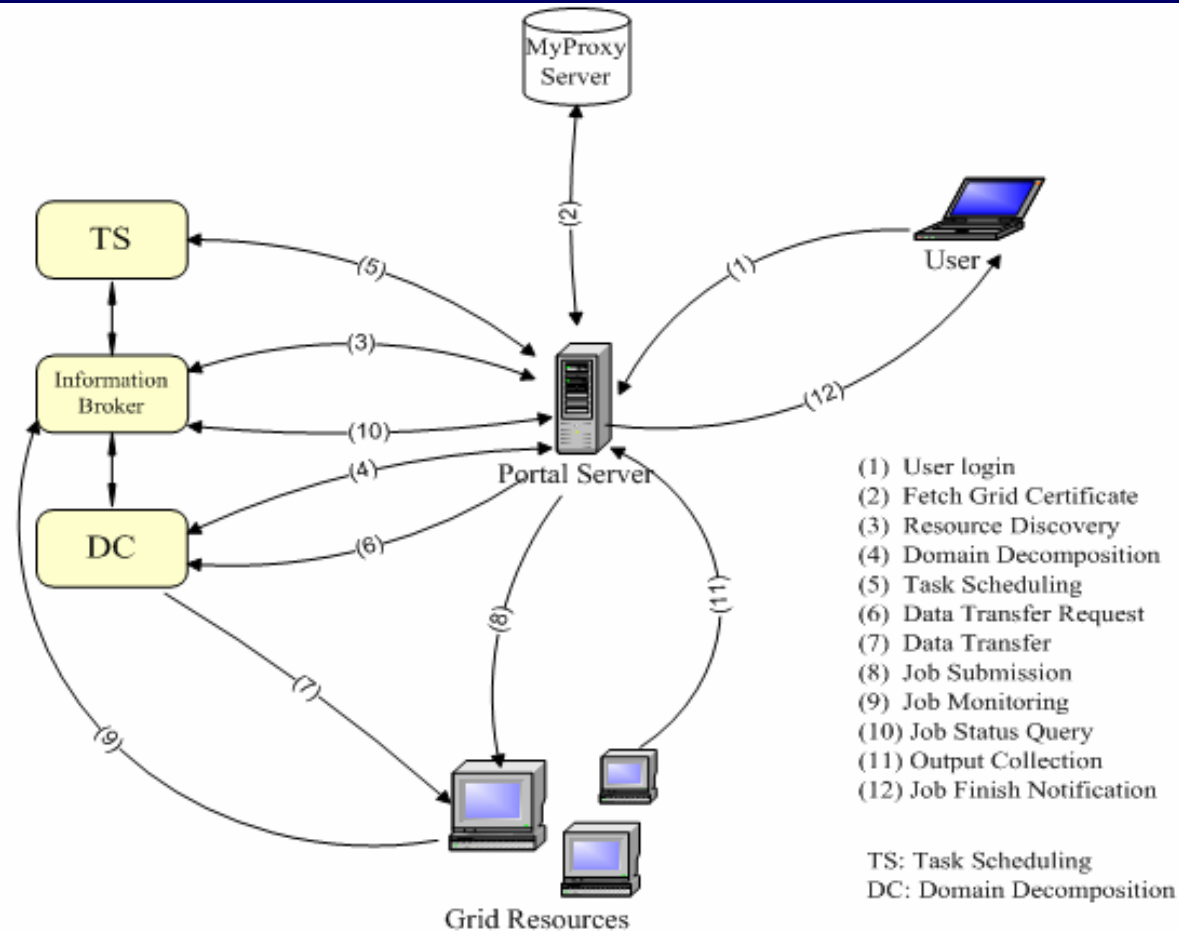
Design

Implementation

Case Study

Evaluation

Conclusions



# A Three-Layer Model of Grid Portal Technologies

Purpose

Background

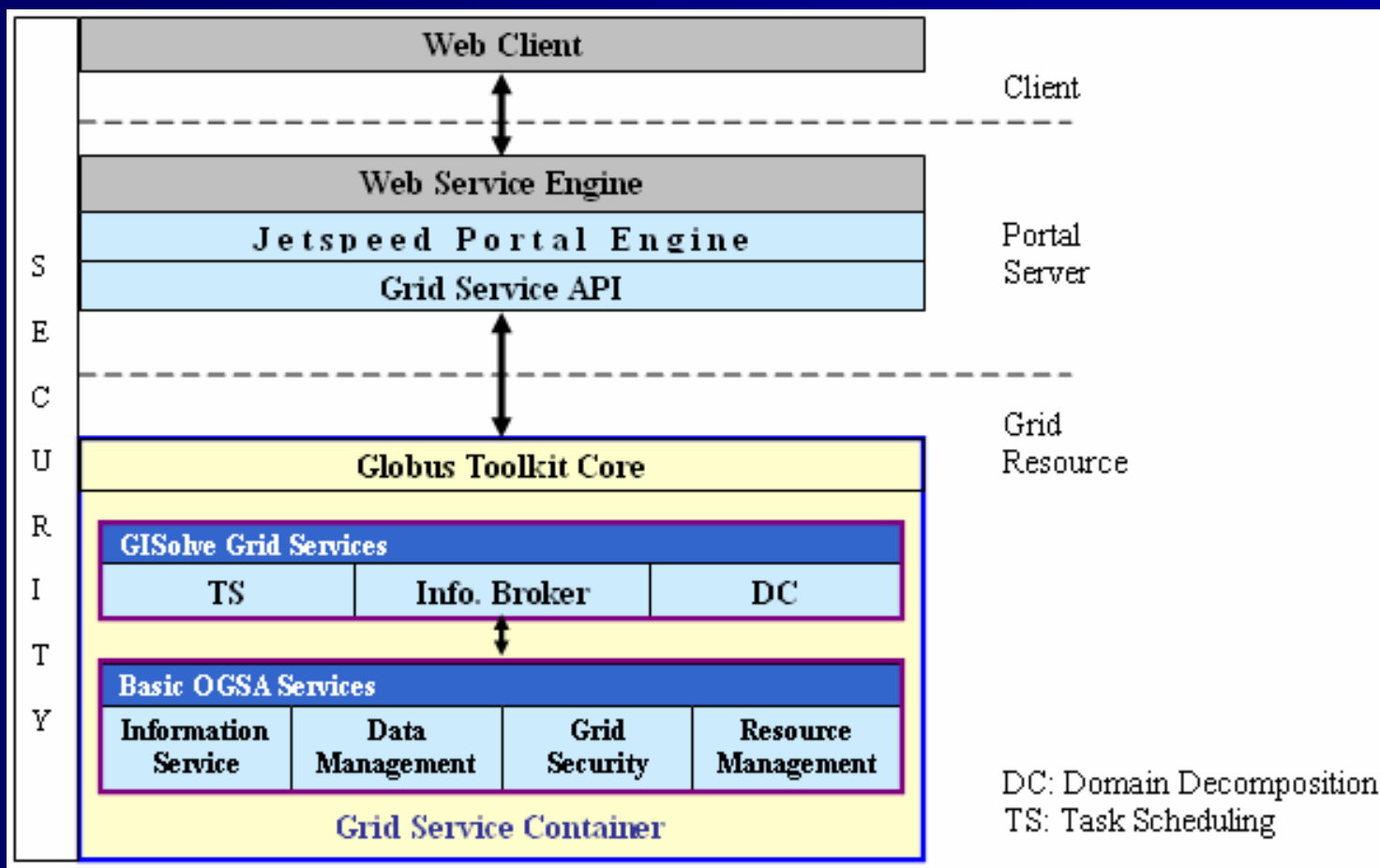
Design

Implementation

Case Study

Evaluation

Conclusions



# Technology Specification

Purpose

- Jetspeed portal server

- Jetspeed: portal engine
- Turbine: MVC (Model-View-Control) framework
- Velocity: dynamic html generation
- Tomcat: portal container

Background

Design

Implementation

- Multi-user support

- User registration and management
- User-based state management

Case Study

Evaluation

Conclusions

- Configurable portal interface

- Customized layout
- Display control of each portlet interface

# APIs Implemented

Purpose

## ■ APIs developed

- [org.gisolve.globus]: Globus Toolkit enhancement
- [edu.uiowa.gisolve.ogsa]: Grid service (interface definition, user-level Grid service implementation)
- [org.apache.jetspeed.modules.actions.gisolveportlets]: Jetspeed portlet action implementation
- [gisolveportlets]: portal support for Grid service APIs (state management, Grid service client APIs)

Background

Design

Implementation

Case Study

Evaluation

Conclusions

## ■ APIs integrated

- Java COG (GT2), GT3/GT4 Core, OGSA (WS GRAM, RFT), MyProxy, LDAP, JNI, JPL

# Defining GISolve Services

Purpose

## ■ A task scheduling service gWSDL:

Background

Design

Implementation

Case Study

Evaluation

Conclusions

```
<definitions name="TSService targetNamespace=http://uiowa.edu/grow/namespaces/2004/04/gisolve/TS ... />
  <!-- import external types -->
  <import location="../ogsi/ogsi.gwsdl" .../>
  <!-- self-defined types -->
  <import location="gidTypes.xsd" namespace="http://uiowa.edu/grow/namespaces/2004/04/gisolve/gidtypes"/>
  <!-- local types schema -->
  <types><xsd:schema .../></types>
  <!-- message definition -->
  <message name="SchedInputMessage">...</message>
  <message name="SchedOutputMessage">...</message>
  <!-- service operation definition -->
  <gwsdl:portType name="TSPortType" extends="ogsi:GridService">
    <operation name="schedule">
      <input message="tns:SchedInputMessage"/>
      <output message="tns:SchedOutputMessage"/>
      <fault name="Fault" message="ogsi:FaultMessage"/>
    </operation>
    <sd:serviceData name="TSState" ... />
  </gwsdl:portType>
```

# OGSI to WSRF

## WSDL

```
<?xml version="1.0" encoding="UTF-8"?>
<definitions name="DCService" ... ..
  xmlns:wsrp="http://docs.oasis-open.org/wsrp/2004/06/wsrp-WS-ResourceProperties-1.2-draft-01.xsd"
  xmlns:wsrpw="http://docs.oasis-open.org/wsrp/2004/06/wsrp-WS-ResourceProperties-1.2-draft-01.wsdl"
  xmlns:wslpp="http://www.globus.org/namespaces/2004/10/WSDLPreprocessor" >
  <wsl:import namespace="http://docs.oasis-open.org/wsrp/2004/06/wsrp-WS-ResourceProperties....wsdl"
    location=" ../wsrf/properties/WS-ResourceProperties.wsdl"/>
  <types>... ..<xsd:element name="DCResourceProperties">... ..</xsd:element></types>
  <!-- WSRF Grid service PortType definition -->
  <portType name="DCPortType" wslpp:extends="wsrpw:GetResourceProperty"
    wsrp:ResourceProperties="tns:DCResourceProperties">
    <operation name="decompose">... ..</operation>
  </portType>
```

## gWSDL

```
<?xml version="1.0" encoding="UTF-8"?>
<definitions name="DCService" ... ..
  xmlns:ogsi="http://www.gridforum.org/namespaces/2003/03/OGSI"
  xmlns:gwsdl="http://www.gridforum.org/namespaces/2003/03/gridWSDLExtensions"
  xmlns:sd="http://www.gridforum.org/namespaces/2003/03/serviceData" >
  <!-- import ogsi Grid Service gwsdl. Location is under gt3/schema/ogsi/ -->
  <import location=" ../ogsi/ogsi.gwsdl" namespace="http://www.gridforum.org/namespaces/2003/03/OGSI"/>
  ... ..
  <!-- Grid service interface definition -->
  <gwsdl:portType name="DCPortType" extends="ogsi:GridService">
    <operation name="decompose">... ..</operation>
    <sd:serviceData name="DCState" type="xsd:int" ... .. />
  </gwsdl:portType>
```



# GISolve Services in Action

Purpose

## ■ Deployed GISolve services:

Background

Design

Implementation

```
[globus4@rsgridportal globus4]$ globus-start-container
Starting SOAP server at: https://128.255.162.167:8443/wsrf/services/
With the following services:
```

Case Study

... ..

```
[12]: https://128.255.162.167:8443/wsrf/services/gisolve/TSService
```

... ..

```
[47]: https://128.255.162.167:8443/wsrf/services/gisolve/DCService
```

Evaluation

Conclusions

# Case Study – $G_i^*(d)$ Statistic

Purpose

Background

Design

Implementation

Case Study

Evaluation

Conclusions

$G_i^*(d)$

=

$$\frac{\sum_j w_{ij}(d)x_j - W_i^* \bar{x}}{s\{[(nS_{1i}^*) - W_i^{*2}]/(n-1)\}^{1/2}}$$

## ■ References

- Getis, A., and Ord, J.K., 1992, The analysis of spatial association by use of distance statistics. *Geographical Analysis*, **24**(3): 189-206.
- Ord, J. K., and Getis, A., 1995, Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, **27**(4): 286-306.

# Synthetic Datasets

Purpose

Background

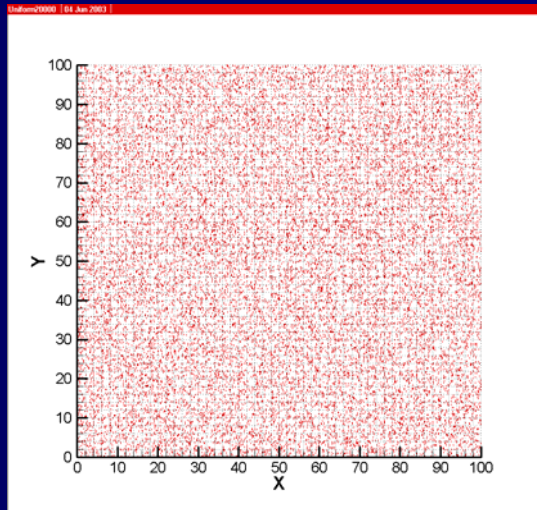
Design

Implementation

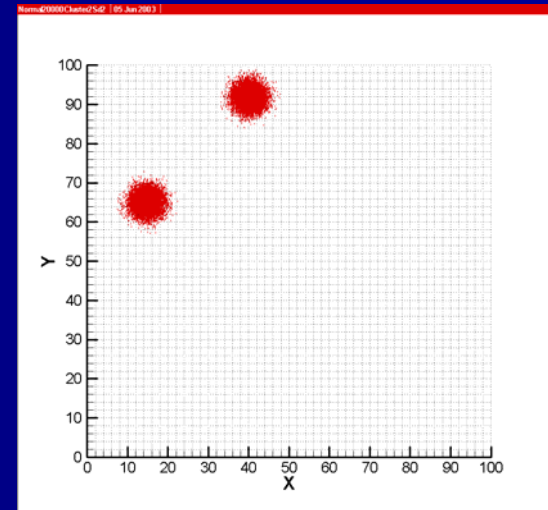
Case Study

Evaluation

Conclusions



20000-point a uniform random distribution



20000-point : two clusters, each of which has a normal distribution with a standard deviation of 2

# User Interface

Purpose

Background

Design

Implementation

Case Study

Evaluation

Conclusions

**User Profile**

User name: **Shaowen Wang**  
 VO: Academic Technologies - Research Services  
 From: The University of Iowa

**Spatial Domain Decomposition Result**

The dataset has **20000** points and has been decomposed to **11** jobs.

**GISolve Session Summary**

Dataset	Date	Status
normal20000cluster2sd2	09:08:34, 06/10/2004	Initiated
normal20000cluster2sd10	09:08:59, 06/10/2004	Finished
normal20000cluster6sd2	10:08:45, 06/10/2004	DC done
normal20000cluster6sd10	23:01:24, 06/13/2004	DC done
normal20000cluster2sd2	16:14:52, 06/17/2004	TS done
normal20000cluster2sd2	19:12:23, 06/17/2004	Running

**Computation Progress**

Job ID	Data Transfer	Job Execution	Result Transfer	Status
3-0-45	100%	60%	0%	Running
3-46-157	100%	100%	100%	Done
3-158-255	100%	34%	0%	Running
6-0-7	100%	0%	0%	Queued
6-8-17	100%	0%	0%	Queued
6-18-47	100%	78%	0%	Running
6-48-62	100%	0%	0%	Queued
6-63-63	100%	80%	0%	Running
8-0-2	100%	68%	0%	Running
8-3-9	100%	100%	100%	Done
8-10-15	100%	34%	0%	Running
<b>Overall:</b>	<b>100%</b>	<b>37%</b>	<b>18%</b>	<b>41%</b>

**Jetspeed POWERED**

GISolve. Last updated on June 22, 2004.

# An Example of Spatial Computational Domain Decomposition

Purpose

Background

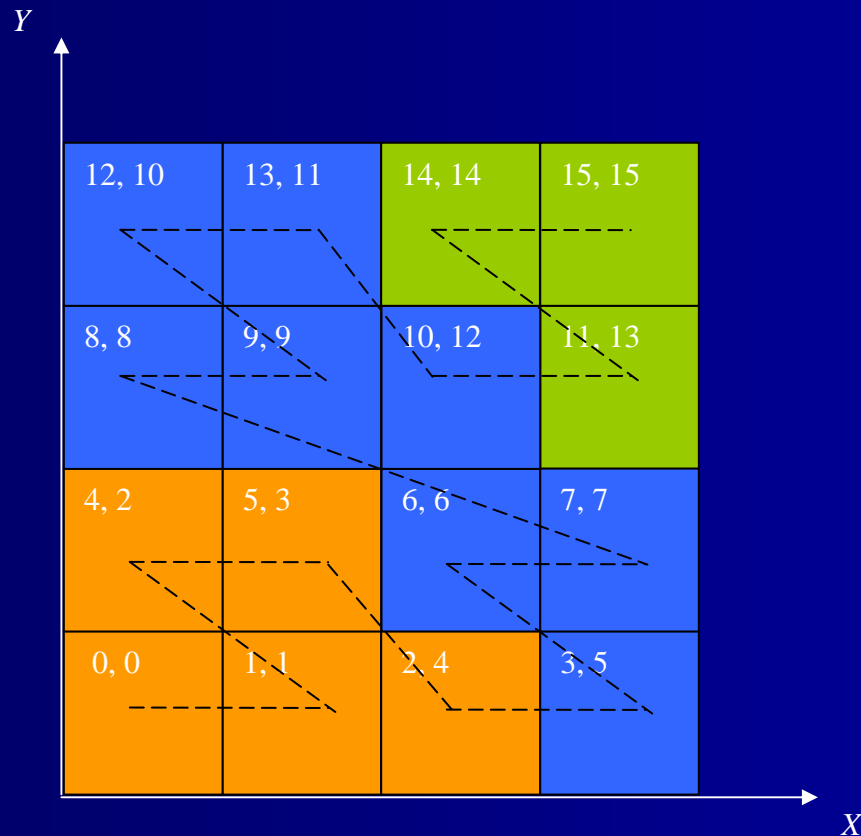
Design

Implementation

**Case Study**

Evaluation

Conclusions



# A Task Scheduling Service

Purpose

- Decoupled from domain decomposition services

Background

Design

- *NP*-complete problem

- Max-min algorithm implemented

Case Study

- Using the theoretical estimate to compute the Expected Time to Compute (ETC) matrix

Evaluation

Conclusions

# A Diagrammatic Example of Static Task Scheduling

Purpose

Background

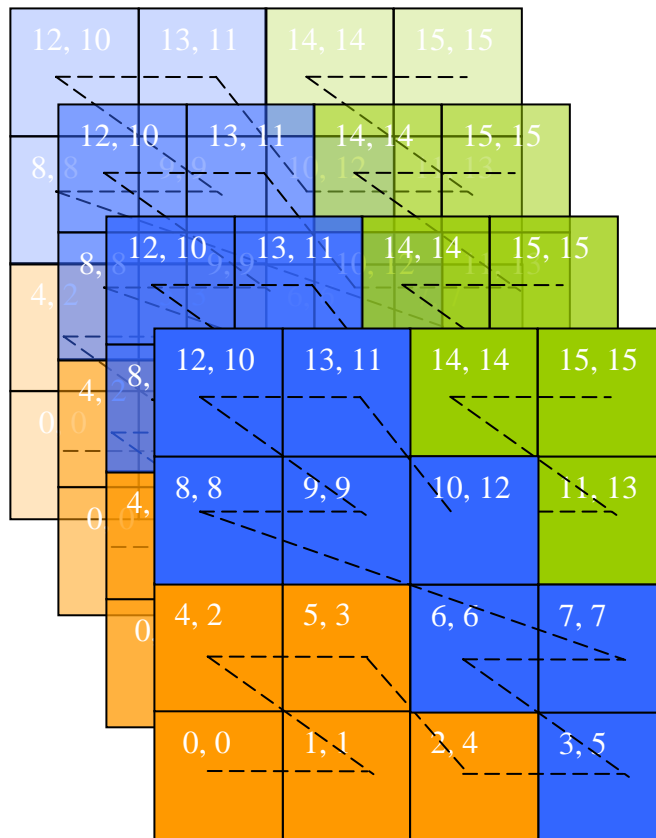
Design

Implementation

Case Study

Evaluation

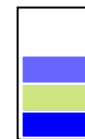
Conclusions



Resource1



Resource2



Resource3



# Grid Testbed Implementation - HawkGrid

Purpose

Background

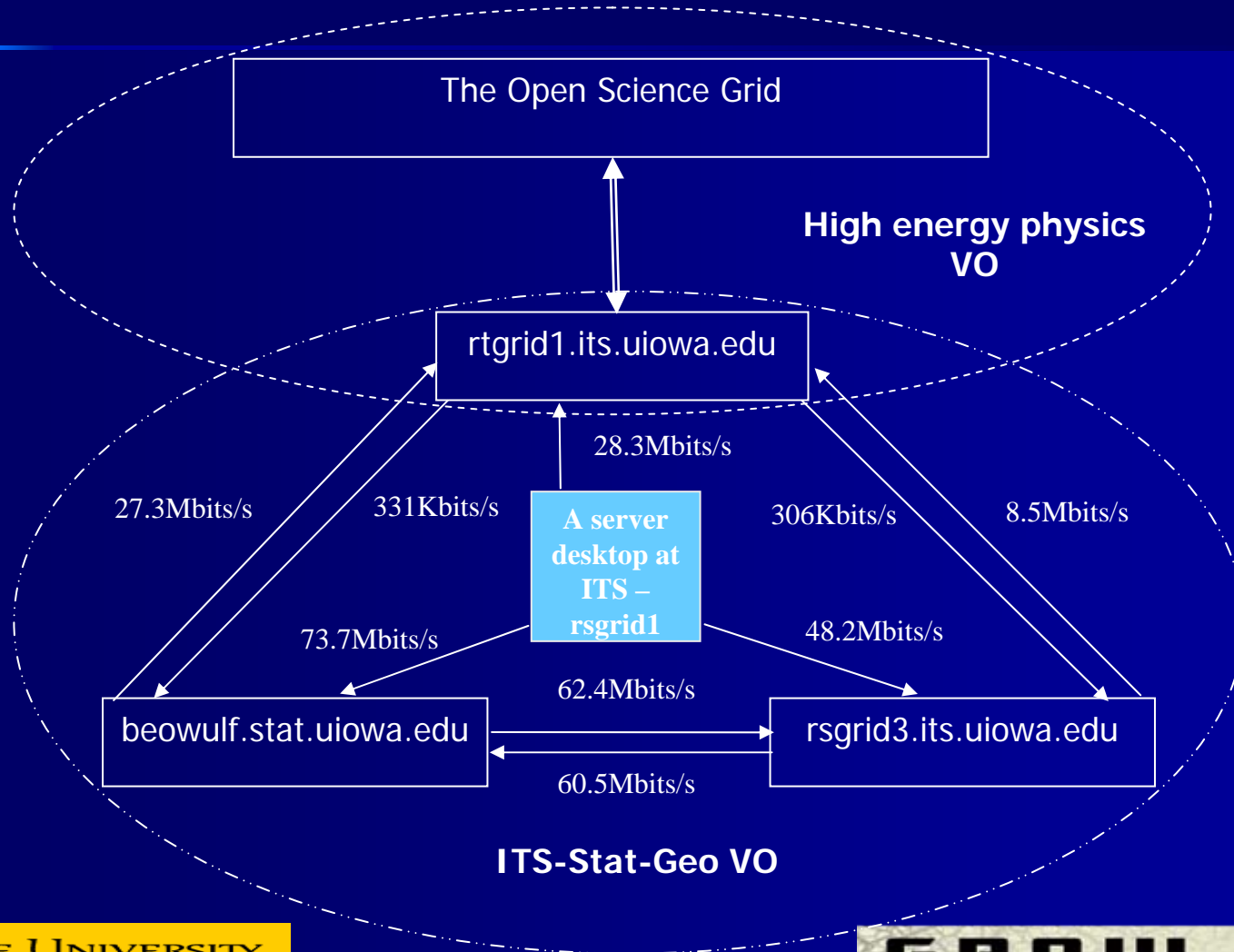
Design

Implementation

Case Study

Evaluation

Conclusions





Purpose

Background

Design

**Implementation**

Case Study

**Evaluation**

Conclusions

# DEMO

# Performance

Purpose

Background

Design

Implementation

Case Study

Evaluation

Conclusions

- The problem cannot be solved by any **single** computer in our Grid testbed
  - **3.2G memory**
- Solved within 10 minutes through GISolve

# Conclusions

Purpose

Background

Design

Implementation

Case Study

Evaluation

Conclusions

- GISolve demonstrates how the Grid can benefit research on computationally intensive geographic information analyses
- GISolve integrates OGSA-based Web services to support the computational aspects of GIServices

# Ongoing Research

- Interoperability of GISolve services
- Adaptive domain decomposition services
- Evaluation of GISolve performance
- Extension of the types of geographic information analyses GISolve supports

# Acknowledgement

- This research is partially supported by the HawkGrid project funded by the Office of Vice President for Research at The University of Iowa
- Computational resources used for experiments include a cluster supported by the U.S. NSF iVDGL (international Virtual Data Grid Laboratory) project