Complexity Computational Environment: Data Assimilation SERVO Grid

SERVO — Solid Earth Research Virtual Observatory

PI: Andrea Donnellan

May 2005 Annual Review

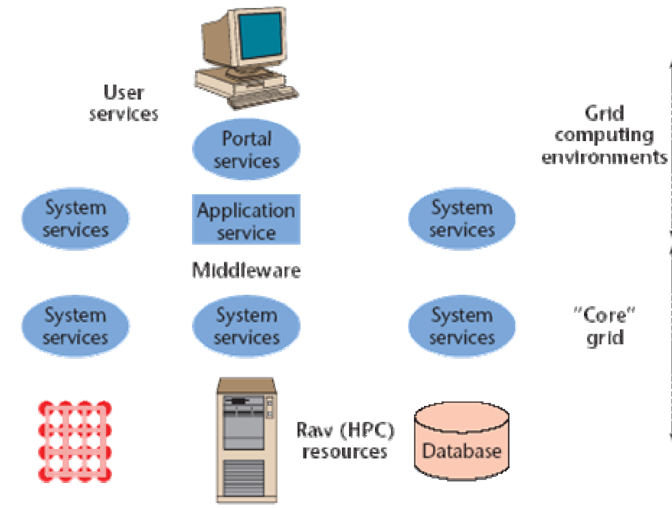
Complexity Computational Environments: Data Assimilation SERVO Grid

PI: Andrea Donnellan, Jet Propulsion Laboratory

Objective

Develop the first real-time, large-scale, data assimilation grid implementation for the study of earthquakes that will:

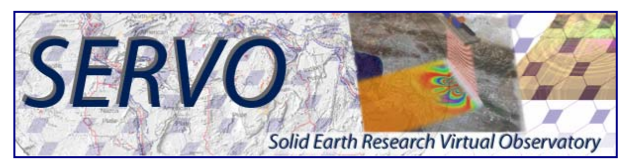
- Assimilate distributed data sources and complex models into a parallel high-performance earthquake simulation and forecasting system
- Simplify data discovery, access, and usage from the scientific user point of view
- Provide capabilities for efficient data mining
- Act as the first step in the Solid Earth Research Virtual Observatory (SERVO)



Three-tiered approach isolates the user from the computational resources

Year 2 (2004-2005) Accomplishments

- Completed detailed design of CCE architecture.
- Completed prototype data assimilation on SERVOGrid with one application an one coarse graining approach.
- Data assimilation and coarse graining interfaces defined in Web service framework.
- Developed meta-query mediator and translator.
- XML metadata services available as Grid services.
- Analyzed pattern space to reduce dimensionality on parallel computer.
- Completed analysis of system performance.



- We have already published a total of 21 abstracts/papers
- One undergraduate and eight graduate students and are involved in this project

CoI: Geoffrey Fox & Marlon Pierce (Indiana University)
Dennis McLeod (USC), John Rundle (UCD), Jay Parker (JPL)

TRL_{in} = 3; $TRL_{current}$ = 5

Earth Science Technology Office

SERVO Objectives

Develop the first real-time, large-scale, data assimilation grid implementation for the study of earthquakes that will:

- Assimilate distributed data sources and complex models into a parallel high-performance earthquake simulation and forecasting system
- Simplify data discovery, access, and usage from the scientific user point of view
- Provide capabilities for efficient data mining
- Act as the first step in the Solid Earth Research Virtual Observatory (SERVO)

SERVO Deliverables to NASA

- A demonstration of the assimilation of multiple distributed data sources (a typical data grid problem) into a major parallel high-performance computing earthquake forecasting code, using Web (Grid) service technology.
- A Complexity Computational Environment (CCE) to manage and integrate data and simulations and also provide data understanding and mining tools that integrate XML metadata and large-scale federated database repositories.
  - The architecture of the CCE consists of distributed, federated data systems, data filtering and coarse graining applications, and high performance applications that require coupling.

Scientific Importance

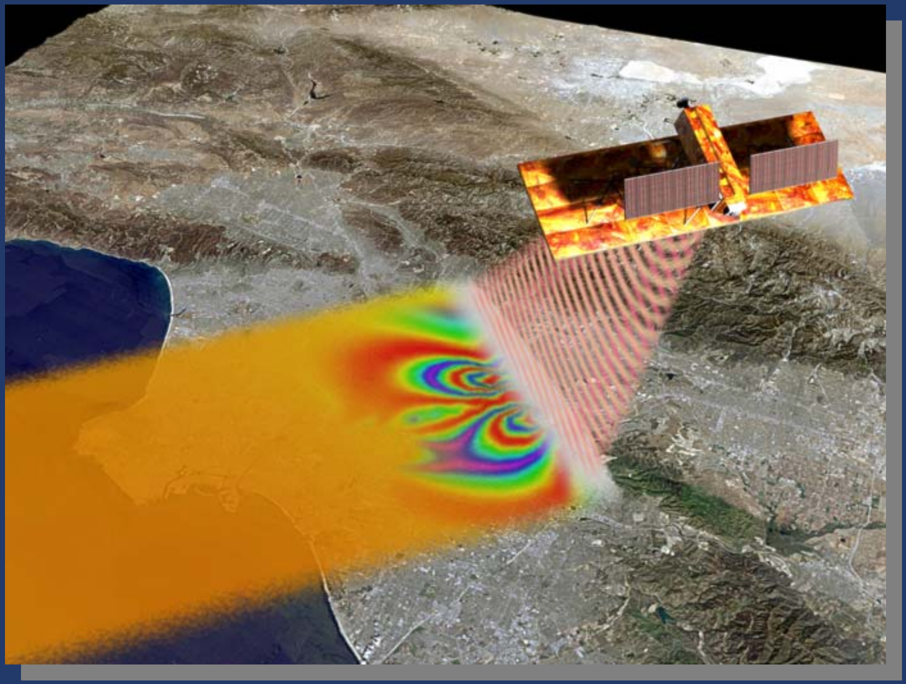
A solid Earth research environment is required to better understand earthquake processes, which cause an annualized U.S. loss of $4.4 billion / year.   This creates the necessary research infrastructure to efficiently model complex earthquake systems.

Surface deformation measurements fill a critical gap for understanding earthquake processes, and such diverse data sets must be incorporated into high performance models and analytical tools.

The models and tools elucidate the unobservable or subtle underlying physics of earthquake processes.

Earth Science Community

This work will develop the necessary infrastructure for future gravity and InSAR missions



We are Developing Five Technologies in this CCE

Architecture of the Complexity Computational Environment (CCE)

- Indiana University (Geoffrey Fox, Marlon Pierce), and JPL (Andrea Donnellan, Jay Parker)

1. Web Services
   - Indiana University (Geoffrey Fox, Marlon Pierce and students)
2. Metadata Services
   - USC (Dennis McLeod and students)
3. Federated Database System
   - USC (Dennis McLeod and students)
4. Data Assimilation Infrastructure
   - JPL (Jay Parker, Greg Lyzenga and student), UC Davis (John Rundle and students)
5. Datamining Infrastructure
   - JPL (Robert Granat, Maggi Glasscoe), UC Davis (John Rundle and students)

Team Interaction

- Team holds weekly telecons and meeting minutes are posted on a “docushare library” accessible via the web
- Team holds biannual SERVO all-hands meetings (e.g. in 2004, met for two days in Sept 2004 at USC, and half a day at the Dec 2004 AGU meeting)
- Sub-teams meet at other technical conferences, and informally as needed
- Frequent e-mail interaction between team members
- Progress shown at this annual review illustrates how the project work is fully integrated

Milestones met to date

- Developed XML schemas to support services and data structures of complexity computational environment (CCE).
- Developed geophysics meta-ontology.
- Completed stand-alone modules supporting database extraction and allowing direct integration into modeling simulation codes.
- Assessed techniques using limited datasets and determined appropriate resolution scaling.
- Completed exploring coupling methodologies to guide development of cross-scale tools.
- Completed detailed design of CCE architecture.
- Completed prototype data assimilation on SERVO Grid with one application and one coarse graining approach.
- Data assimilation and coarse graining interfaces defined in Web service framework.
- Developed meta-query mediator and translator.
- XML metadata services available as Grid service.
- Analyzed pattern space to reduce dimensionality on parallel computer.
- Completed analysis of system performance.

Community Grids Lab SERVOGrid CCE Review

Geoffrey Fox and
Marlon Pierce
Indiana University

SERVO Apps and Their Data

As summarized below, many SERVO codes use observational data measurements as input and create geo-located results.

- **Disloc**:
  handles multiple arbitrarily dipping dislocations (faults) in an elastic half-space.
  - Relies upon geometric fault models.

- **GeoFEST**:
  Three-dimensional viscoelastic finite element model for calculating nodal displacements and tractions. Allows for realistic fault geometry and characteristics, material properties, and body forces.
  - Relies upon fault models with geometric and material properties.

SERVO Apps and Their Data (cont.)

– **Virtual California:**
  Program to simulate interactions between vertical strike-slip faults using an elastic layer over a viscoelastic half-space.
  - Relies upon fault and fault friction models.

– **Pattern Informatics:**
  Calculates regions of enhanced probability for future seismic activity based on the seismic record of the region
  - Uses seismic data archives

– **RDAHMM:**
  Time series analysis program based on Hidden Markov Modeling. Produces feature vectors and probabilities for transitioning from one class to another.
  - Used to analyze GPS and seismic catalogs.

Our Approach to Building Grid Services

- There are several competing visions for Grid Web Services
  - WSRF (US) and WS-I+ (UK) are most prominent
  - An overview of these issues is available here http://grids.ucs.indiana.edu/ptliupages/presentations/ogsaukjan05.ppt
- We follow the WS-I+ approach
  - Build services on proven basic standards (WSDL, SOAP, UDDI)
  - Expand this core as necessary
    - GIS standards implemented as Web Services
    - Service orchestration, lightweight metadata management

Our Approach to Building Grid Services (cont.)

- We stress innovative implementations
  - Web Services are essentially message-based.
  - SERVO applications require non-trivial data management (both archives and real-time streams).
  - We can support both streams and events through NaradaBrokering messaging middleware.
  - HPSearch uses and manages NaradaBrokering events and data streams for service orchestration.
  - Upcoming improvements to the Web Feature Service will be based on streaming to improve performance.
  - Sensor Grid work is being based on NaradaBrokering.

Our Approach to Building Grid Services (cont.)

- Core NaradaBrokering development stresses the support for Web Service standards
  - WS-Reliability, WS-Eventing, WS-Security
  - AIST SERVO leverages this work

Geographical Information System Services as a Data Grid

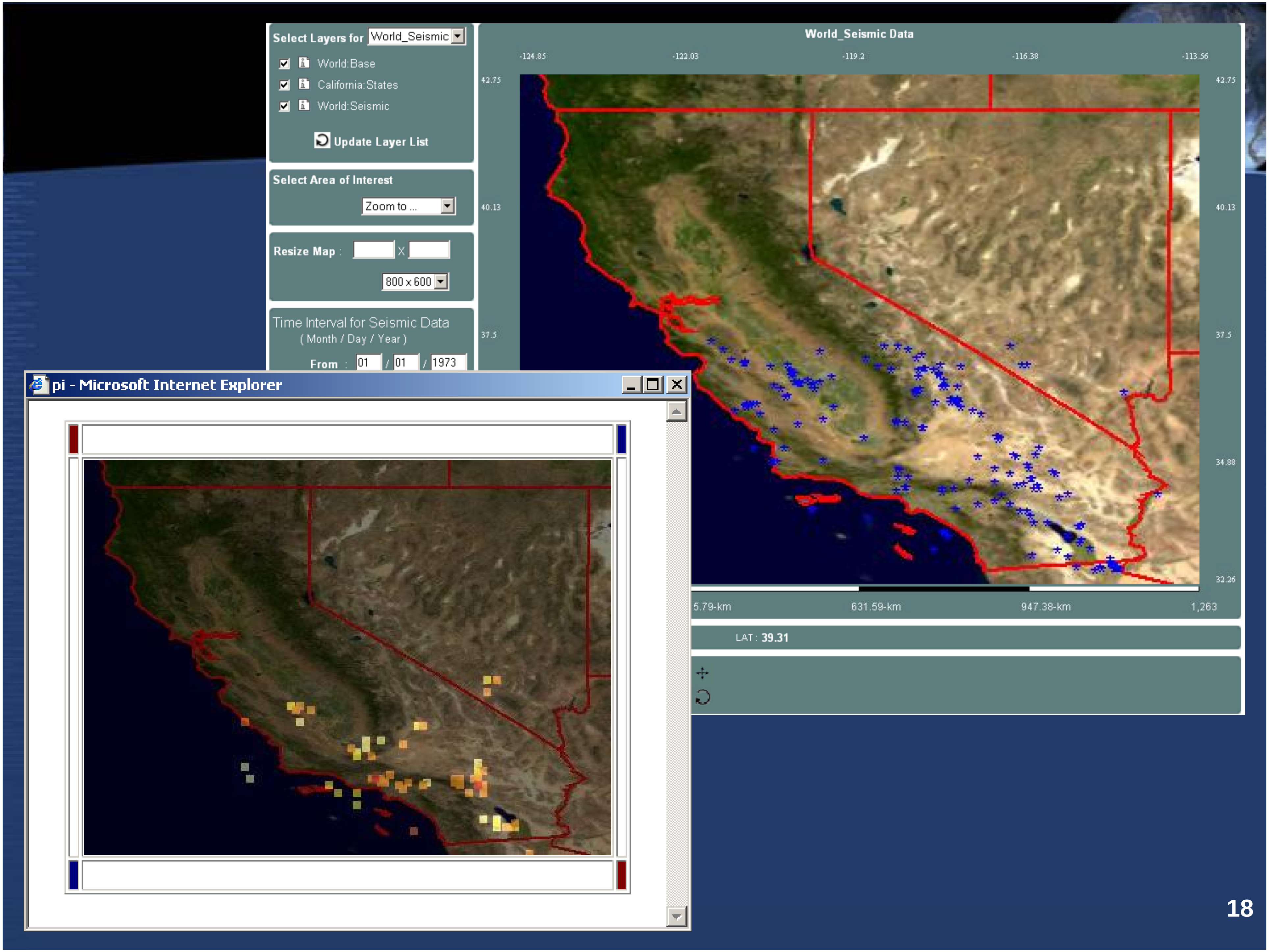
- The Data Grid components of SERVO are best implemented using standard GIS services.
  - Use Open Geospatial Consortium standards
  - Maximize reusability in future SERVO projects
  - Provide downloadable GIS software to the community as a side effect of SERVO research.

- We implemented two cornerstone standards
  - Web Feature Service (WFS): data service for storing abstract map features
    - Supports queries
    - Faults, GPS, seismic records
  - Web Map Service (WMS): generate interactive maps from WFSs and other WMSs.
    - Maps are overlays
    - Can also extract features (faults, seismic events, etc) from user GUIs to drive problems such as the PI code and (in near future) GeoFEST, VC.

Geographical Information System Services as a Data Grid (cont.)

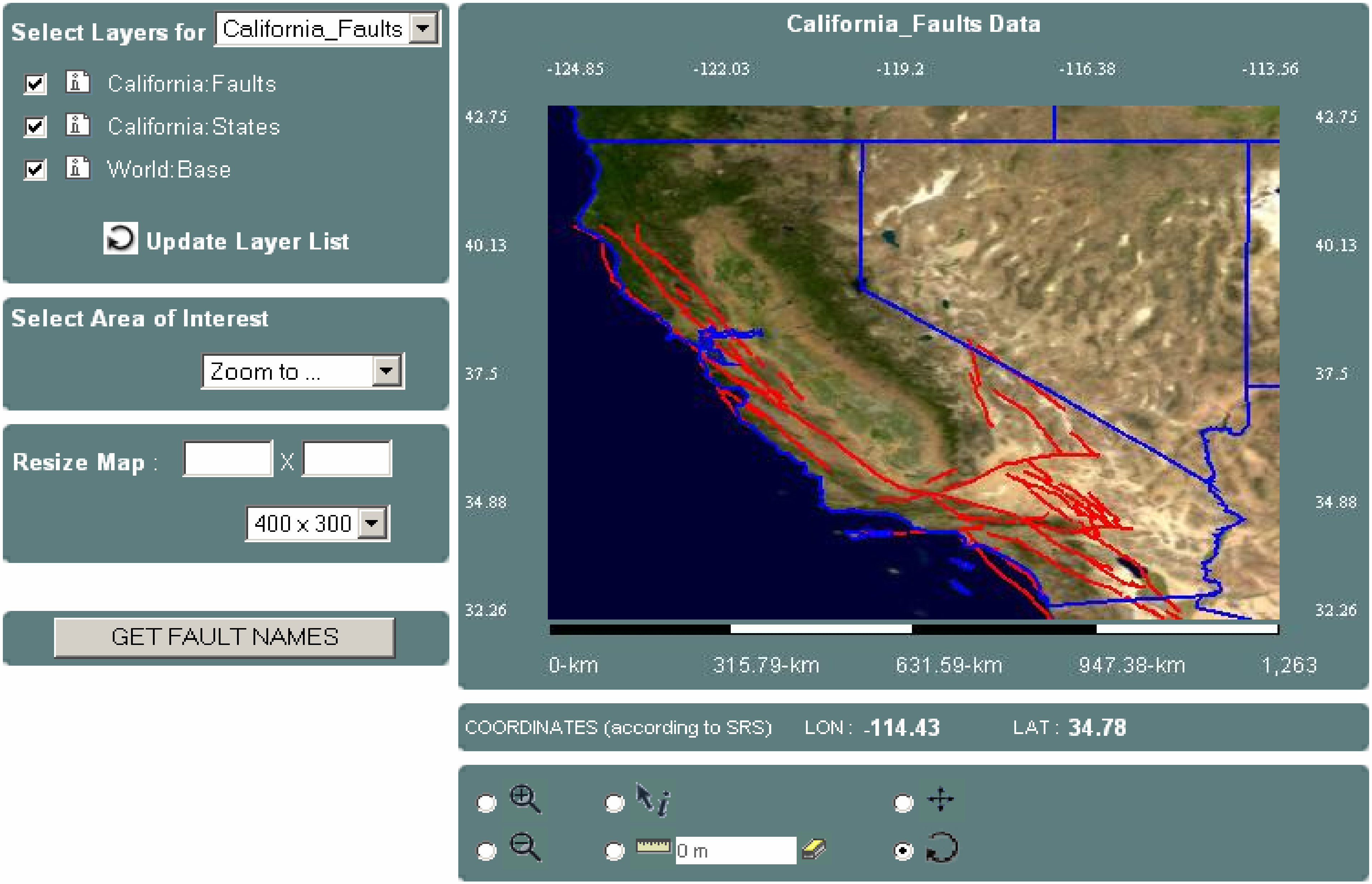
- We built these as Web Services
  - WSDL and SOAP: programming interfaces and messaging formats
  - Users can work with the data and map services through programming APIs as well as browser interfaces.
  - Running demos and downloadable code are available from www.crisisgrid.org.

- We are currently working on these steps
  - Improving WFS performance
  - Integrating WMS clients with more applications
  - Making WMS clients publicly available and downloadable (as portlets).
  - Implementing SensorML for streaming, real-time data.

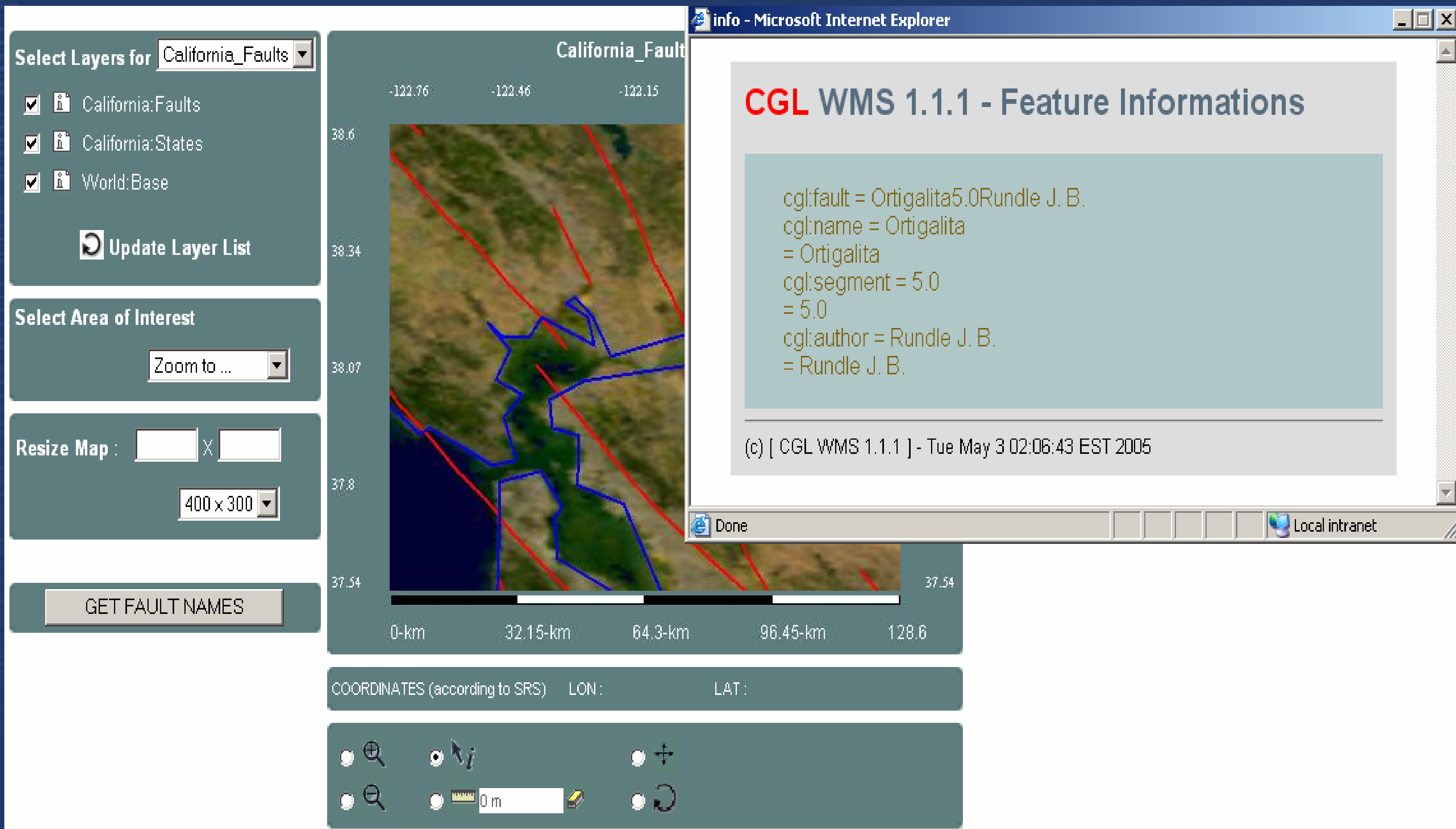


Select Layers for World_Seismic
World:Base
California:States
World:Seismic
Update Layer List
Select Area of Interest
Zoom to ...
Resize Map :
800 x 600
Time Interval for Seismic Data
(Month / Day / Year)
From : 01 / 01 / 1973
World_Seismic Data
-124.85
-122.03
-119.2
-116.38
-113.56
42.75
40.13
37.5
34.88
32.26
5.79-km
631.59-km
947.38-km
1,263
LAT : 39.31
pi - Microsoft Internet Explorer

A Screen Shot From the WMS Client



When you select (i) and click on a feature in the map



Metadata and Information Services Work

- We like the OGC but their metadata and information services are too specialized for GIS data.
  - Web Service standards should be used instead
- For basic information services, we developed an enhanced UDDI (Universal Description, Discovery and Integration)
  - UDDI provides registry for service URLs and queryable metadata.
  - We extended its data model to include GIS capabilities.xml files.
    - You can query capabilities of services.
  - We added leasing to services
    - Clean up obsolete entries when the lease expires.
- We are also implementing WS-Context
  - Store and manage short-lived metadata and state information
  - Store "personalized" metadata for specific users and groups
  - Used to manage shared state information in distributed applications
    - See "Performance Analysis" slides for more information
- See http://grids.ucs.indiana.edu/~maktas/fthpis/

Service Orchestration with HPSearch

- GIS data services, code execution services, and information services need to be connected into specific aggregate application services.
- HPSearch: CGL's project to implement service management
  - Uses NaradaBrokering to manage events and stream-based data flow
- HPSearch and SERVO applications
  - We have integrated this with RDAHMM and Pattern Informatics
    - These are "classic" workflow chains
  - UC-Davis has re-designed the Manna code to use HPSearch for distributed worker management as a prototype.
  - More interesting work will be to integrate HPSearch with VC.
- This is described in greater detail in the performance analysis presentation and related documents.
  - See also supplemental slides.

Upcoming Work

- Integrate HPSearch with Virtual California for loosely coupled grid application parameter space study.
  - HPSearch is designed to handle, manage multiple loosely coupled processes communicating with millisecond or longer latencies.
- Improve performance of data services
  - This is the current bottleneck
  - GIS data services have problems when you do non-trivial data transfers
  - But streaming approaches and data/control channel separation can dramatically improve this.

Upcoming Work (cont.)

- Provide support for higher level data products and federated data storage
  - CGL does not try to resolve format issues in different data providers
    - See backup slides for a list for GPS and seismic events
    - GML is not enough
  - USC's Ontronic system researches these issues
- Provide real time data access to GPS and other sources
  - Implement SensorML over NaradaBrokering messaging
  - Do preliminary integration with RDAHMM
- Improve WMS clients to support sophisticated visualization

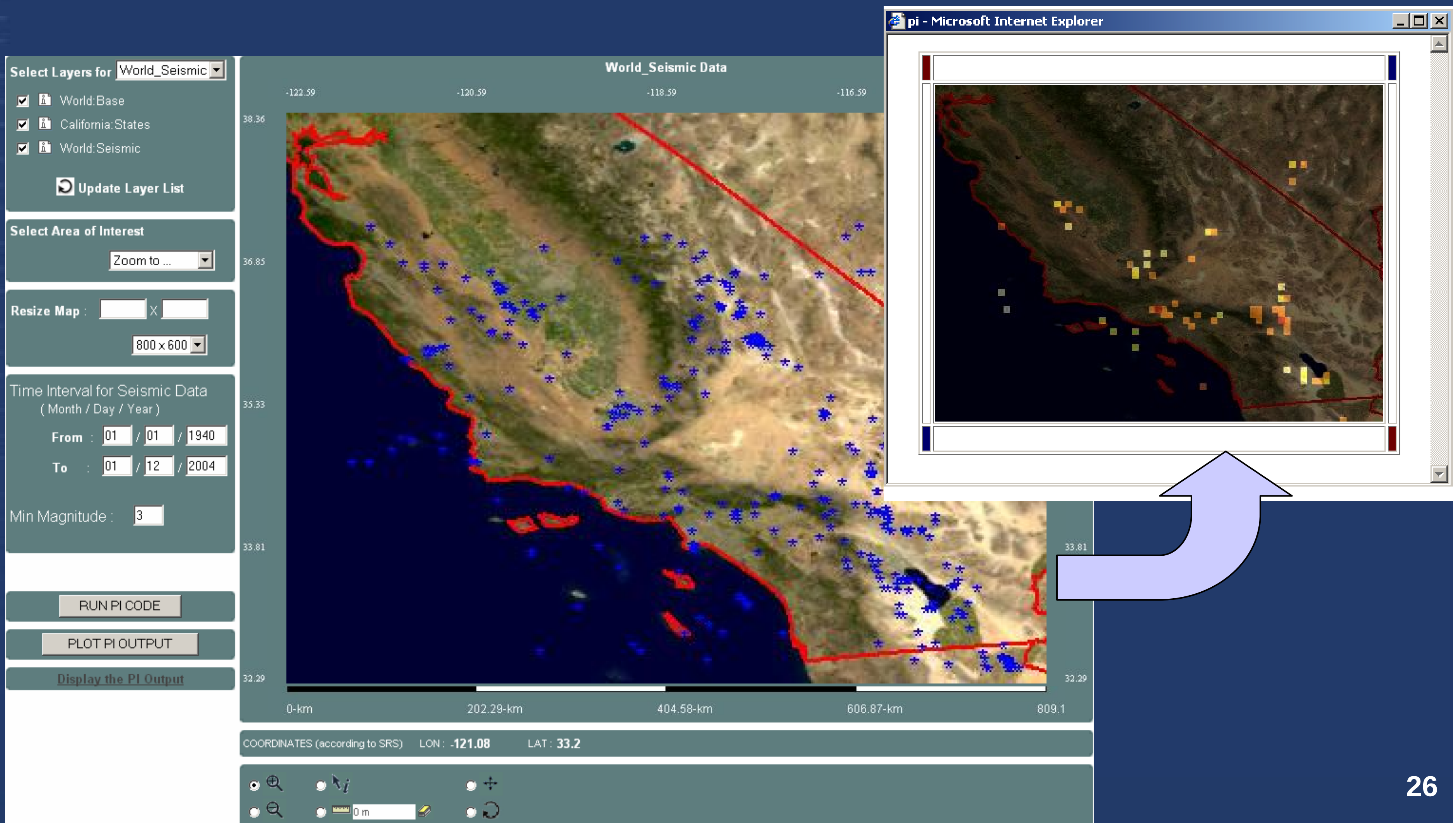
SERVOGrid System Performance

Jay W. Parker
NASA Jet Propulsion Laboratory

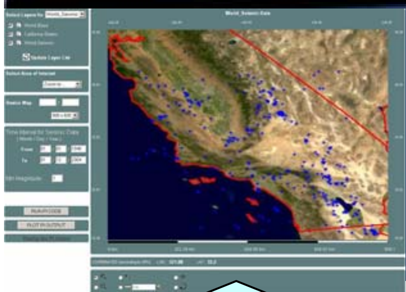
Mehmet Aktas,
Galip Aydin,
Harshawardhan Gadgil,
Marlon Pierce,
and Ahmet Sayar
Indiana University

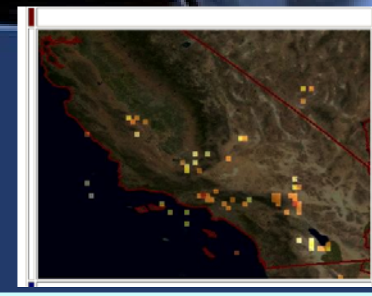
A Typical SERVOGrid Application: Pattern Informatics

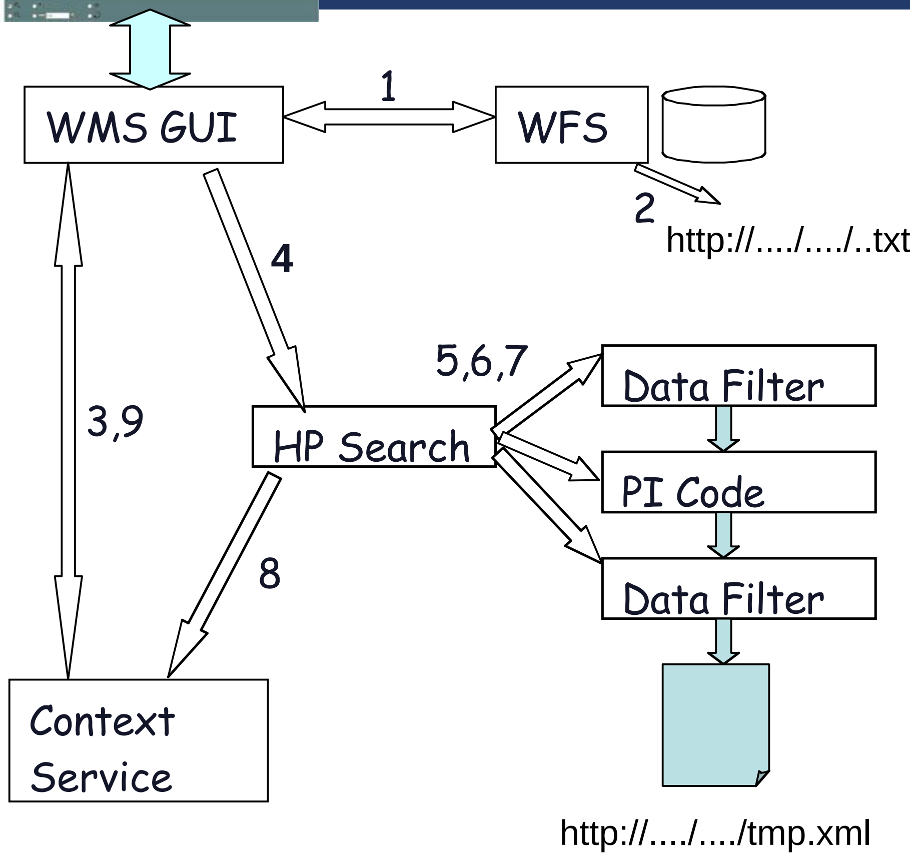
Past Earthquakes May Help Forecast Today’s Hazard



Making Pattern Informatics a SERVOGrid Service







1: Web Map Service (WMS) queries Web Feature Service (WFS) for user's request (region, time span,. . .)

2: WFS dumps results into a .txt file

3, 4: WMS starts a session, invokes HPSearch with a workflow script for PI run (w/ unique session id)

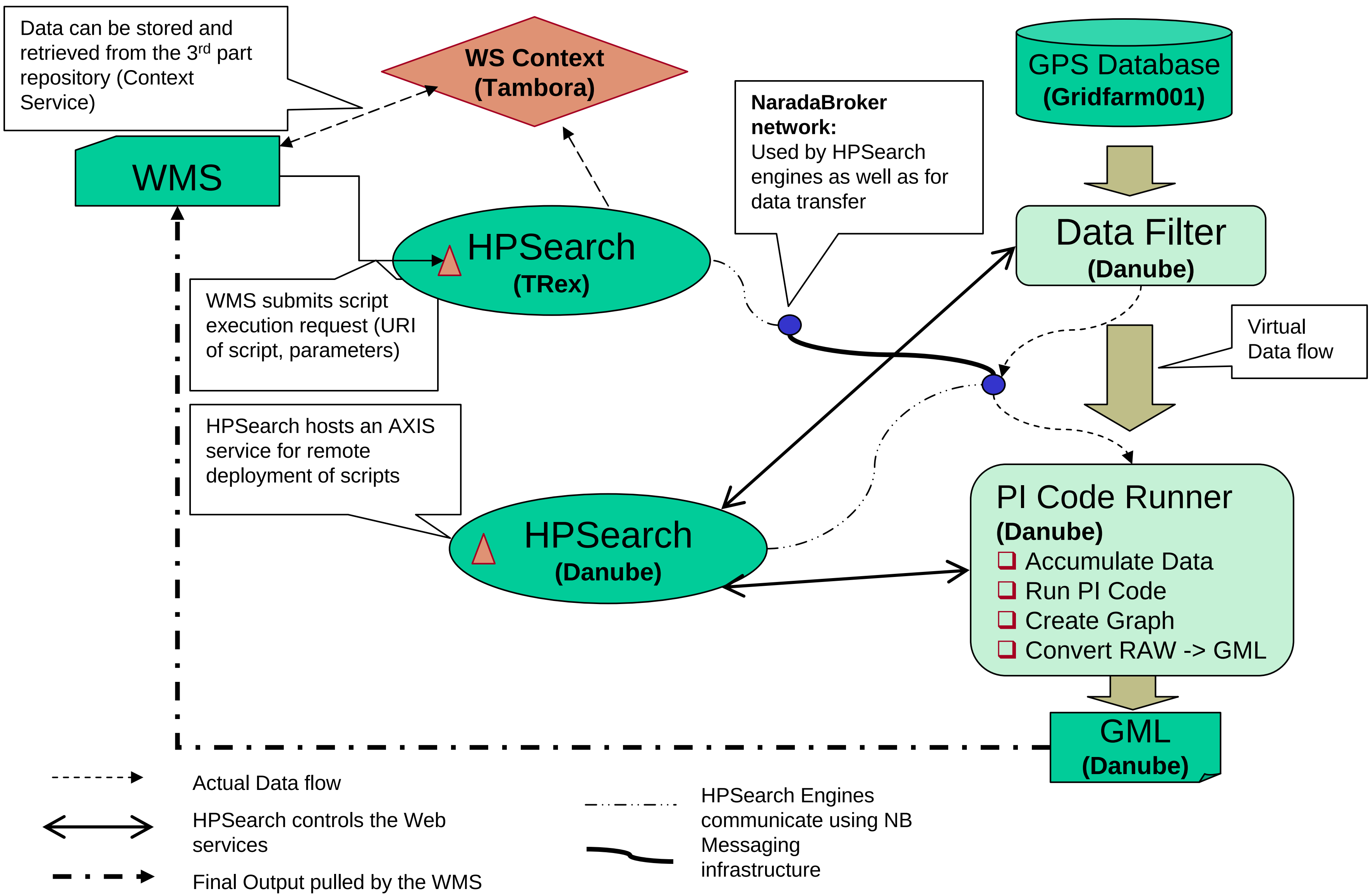
5,6,7: HPSearch runs it, generates GML output file.

8: HPSearch writes output file URI into Context Service (CS)

9: WMS polls CS for ready, URI

10: WMS uses URI to get result, make map.

SERVOGrid: Behind the Scenes



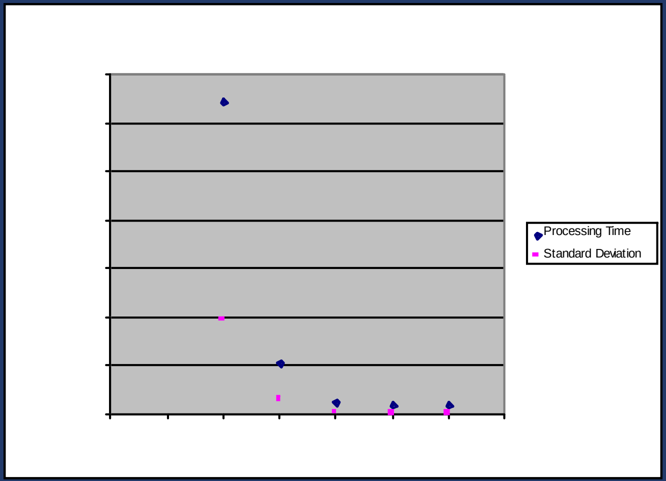
SERVOGrid Performance

Feature Service limit found, bypassed system is sound.

- *As threshold moves down, load increases as expected:*

Magnitude Bound	Count	GML File Size (KB)
5.0	19	11
4.5	67	36
4.0	209	106
3.5	587	287
3.0	1790	880

- *Too slow at high load:*



- *Culprit found – file-based transfer bottleneck in buildGMLstring:*

M=3.0	Mean (nsec)
T[init]	2.39687E+07
T[decodeRequest]	3.28368E+06
T[executeQuery]	3.67557E+09
T[buildQueryResultObj]	2.79935E+06
T[buildGMLobject]	2.42187E+09
T[buildGMLstring]	1.22319E+11
T[formatGML]	4.54205E+08
T[totalProcessTime]	1.29094E+11

- *Fix: developing streaming version*
- *Tests of HPSearch, Context Service, Map Service show SERVOGrid ready for expansion.*

Initial Customers Who Will Test SERVOGrid

- Already getting some basic testing in classes
  - Prof. Gerry Simila, Cal State University, Northridge
- Geologic data
  - Prof. Lisa Grant, UC Irvine
- Crustal modeling
  - Dave Manaker, UC Davis (not affiliated with Co-I John Rundle)
  - Mark Simons, Caltech
- Will continue to expand to other users
  - Prof. Brad Hager, MIT

Ontology-based Federated Information Management for Seismology and Geoscience

Dennis McLeod
Anne Chen,
Sang-Soo Sung,
Dongwoo Won

University of Southern California

Motivation

- Understand the meaning and format of heterogeneous data sources and requirements of simulation and analysis codes
- Desire to interoperate various codes with various information sources (subject to security)
- Problem of semantic and naming conflicts between various federated datasets
  - Discovery, management, integration and use of data difficult
  - Presence of many large federated datasets in seismology
  - Different interpretations and analysis of the same datasets by different experts

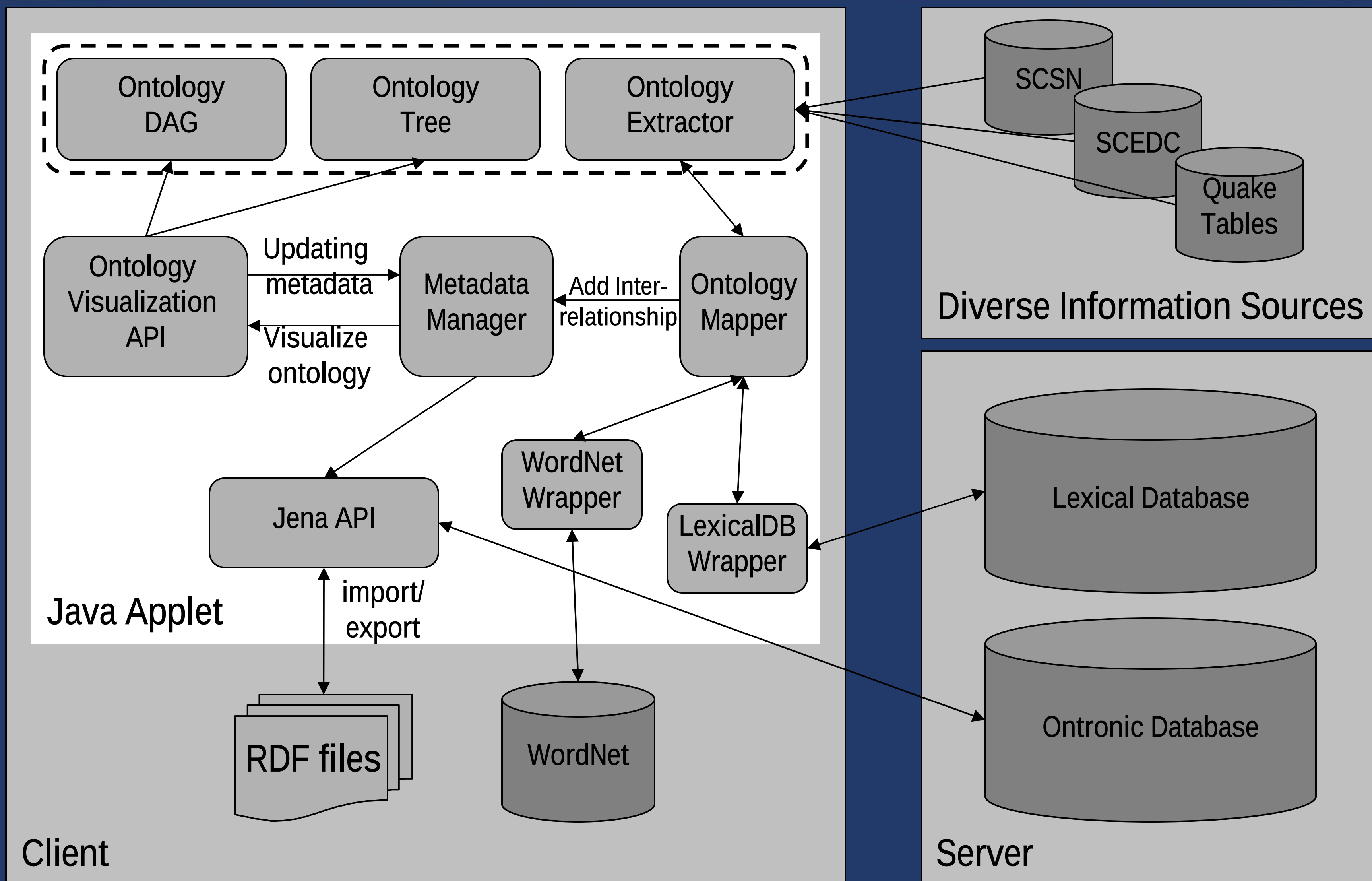
Research Goal

- Design and develop a method and a tool that can:
  - Support interoperation of data and software
  - Support data discovery
  - Semi-automatically extract ontologies from federated datasets
    - Ontology: concepts and inter-relationships
  - Mine for patterns in data to discover new concepts in these federated ontologies
  - Employ generalized geo-science ontology
    - Sources: Geon, GSL, Intellisophic, …

Our Approach

- A semi-automatic ontology extraction methodology from the federated relational database schemas
- Devising a semi-automated lexical database system to obtain inter-relationships with users' feedback
- Providing tools to mine for new concepts and inter-relationships
- Ontronic: a tool for federated ontology-based management, sharing, discovery
- Interface to Scientist Portal

Ontronic Architecture



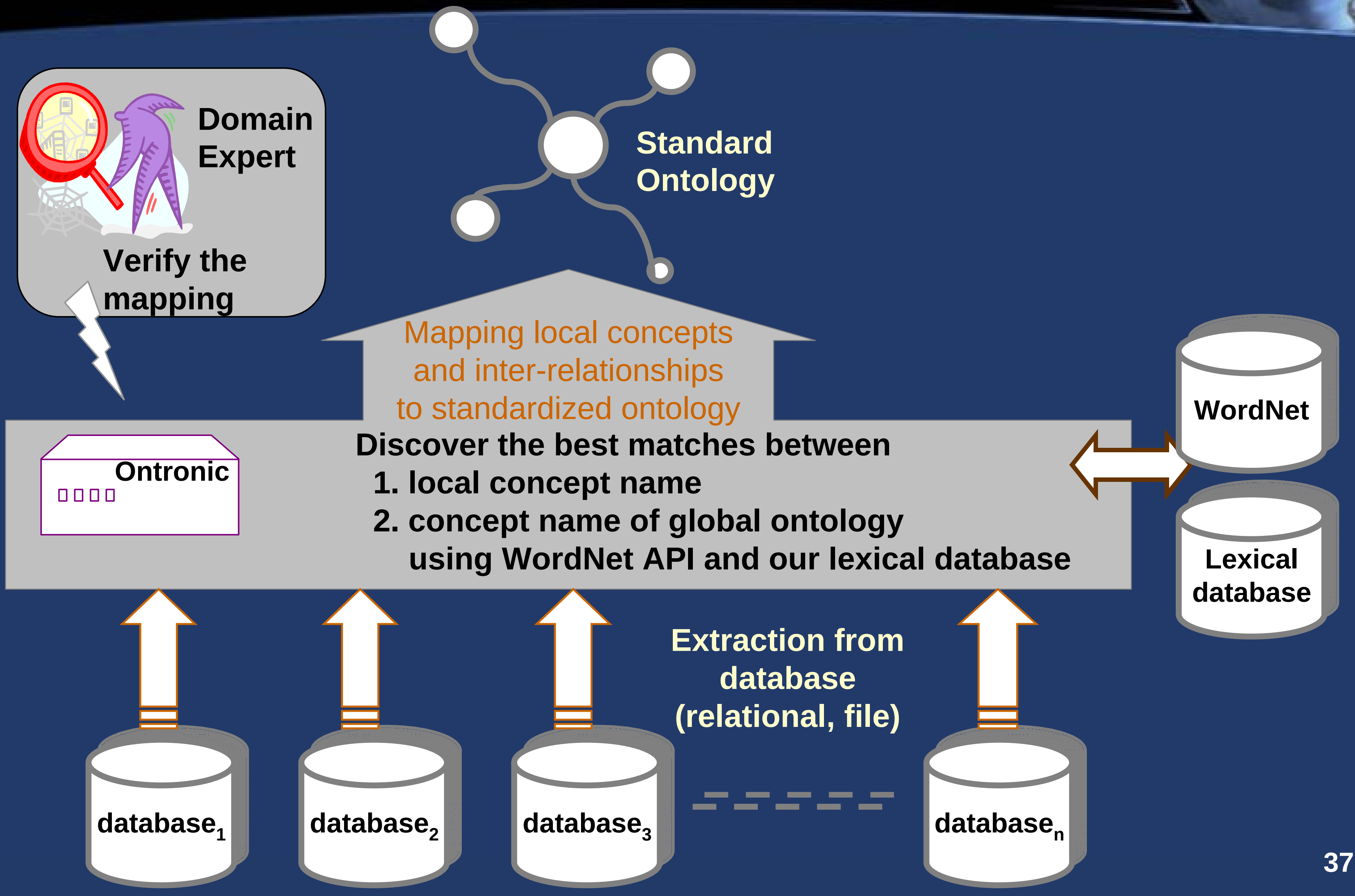
Mapping Algorithm

- A "Standard" ontology for the domain
- Extracting ontologies from the federated datasets
  - e.g., using relational metadata to extract the table and column names

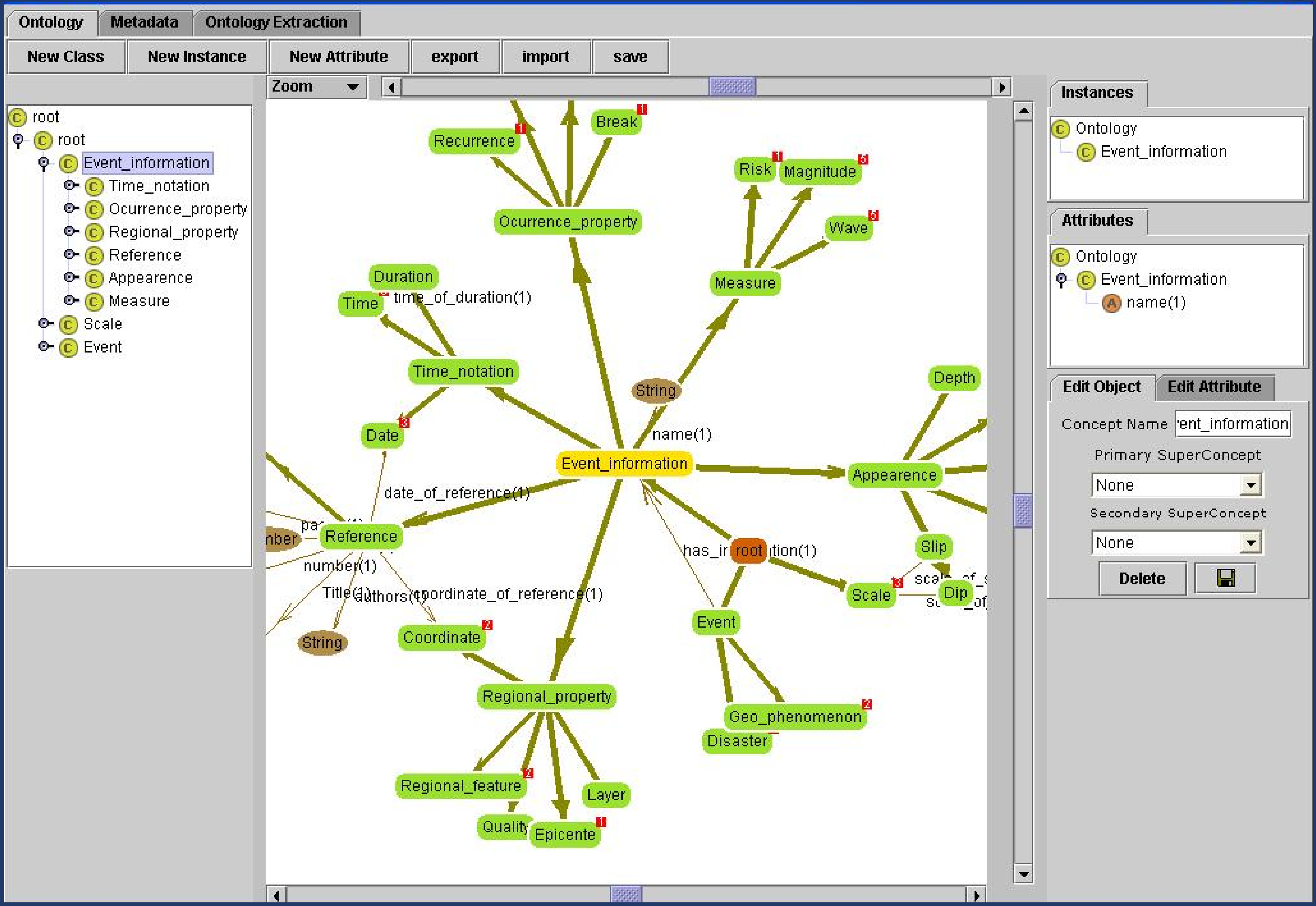
    (or file structures)
- Mapping and storing relationships
  - Mapping algorithm

```
Ontologies extracted from the Federated datasets are denoted by Fi
The Global ontology is denoted by Gi
For each Fi
For each Concept Ci in Fi
Begin
  Try an exact string match to each concept in Gi
  If no matches were found then
    Lookup the Lexical database by Fi
      If no results are found in this lookup then
        Lookup WordNet for synonyms Si of Fi
        Find the closest synonym to Fi in Si by string matching
        If no synonyms were found then
          Ask for user input on this mapping
          Store this mapping in the Lexical database
        Else
          Store the mapping in the Lexical database
        Else
          Store the mapping in the Lexical database
  Else
  Store the mapping in the Lexical database
End
```

Mapping Process



Visual Ontology Manager in Ontronic



Evaluation Plan

- Initially employ three datasets:
  - QuakeTables Fault Database (QuakeSim)
  - The Southern California Earthquake Data Center (SCEDC)
  - The Southern California Seismology Network (SCSN)
- From these large scale and inherently heterogeneous and federated databases, we will evaluate
  - Semi-automatic extraction
  - Checking correctness
  - Evaluating mapping algorithm

Information Management Summary

- Developed a new technique allowing federated datasets of to be queried through a well defined standard ontology using a lexical database
  - Developing a heuristic-based adaptive algorithm for mapping
  - Applied and tested on key geo-science datasets

- Key Contributions:
  - allowing users to transparently query federated databases in terms of a well defined standard ontology
  - Key component of web-services-based scientist portal for interoperation

Data Assimilation Infrastruture

John Rundle
UC Davis

Data Assimilation Using Evolutionary Grid Search Optimization: Towards Model Steering

Ensemble forecasting methods are only accurate to the degree that observed data has been assimilated into them, allowing for model updating and model steering.

- Basic Method: Use parallel grid search, both for model steering and data assimilation.
- System (and model) follows a dynamical path through state space, so observations can be used to periodically adjust model parameters, to bring the model path as close as possible to the path represented by the observed system.
- Cost function, or fitness function defines the misfit between the model path through time and the path represented by the observed data. Must account both for variation in model parameters, as well as variations in initial conditions for process.
- Evolutionary programming approach. Model tuning occurs by parallel grid search, combined with stochastic variation.
- Prototype the process using simpler, but still nonlinear, models.

Steps for Data Assimilation in Virtual California

Basic model dynamics

$$\frac{\partial s}{\partial t} = \frac{\Delta\sigma}{K_T}\left\{\alpha + \delta(t - t_F) + \beta\,\delta\left(\frac{\partial\sigma}{\partial t} - \eta\right)\right\} + noise$$

Define model parameters to be tuned (for each of N fault segments)

$$K_T, \alpha, \beta, \eta$$

Define cost function

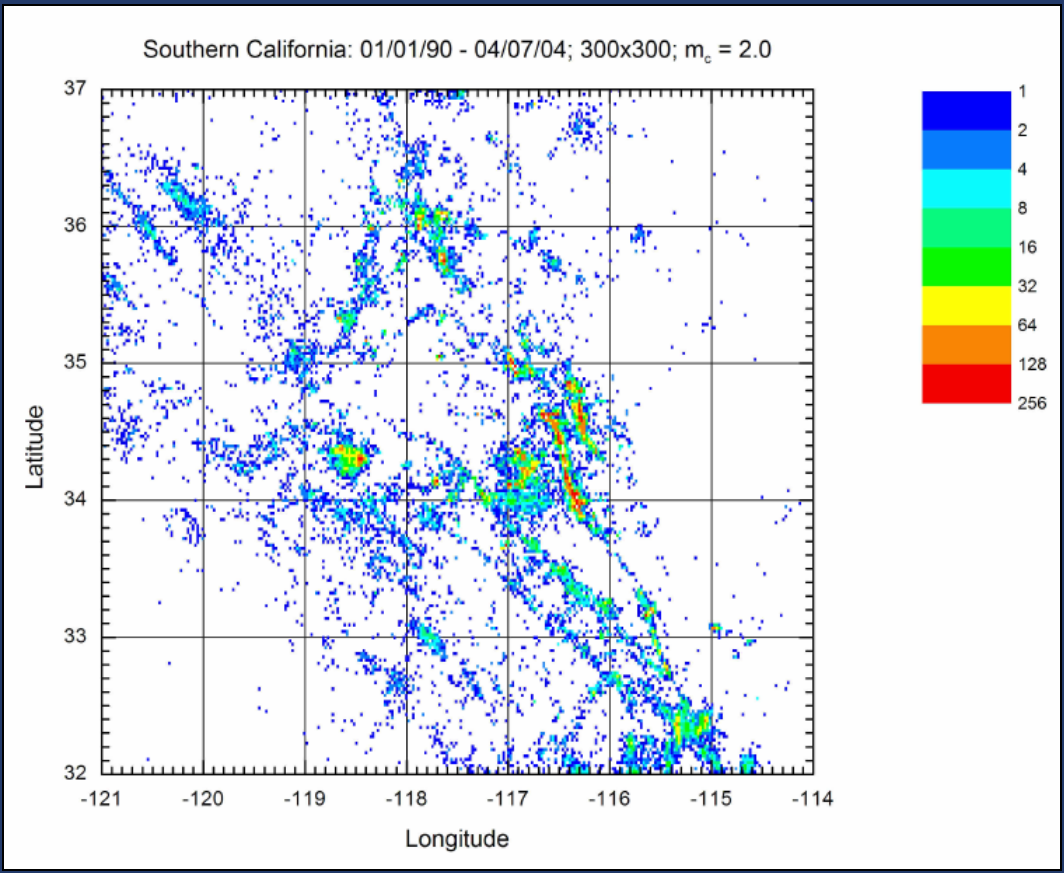
$$C_i(\tau) \equiv \frac{1}{T}\int_0^T dt\,\left\{z_i^O(t)\; z_i^S(t+\tau)\right\}$$

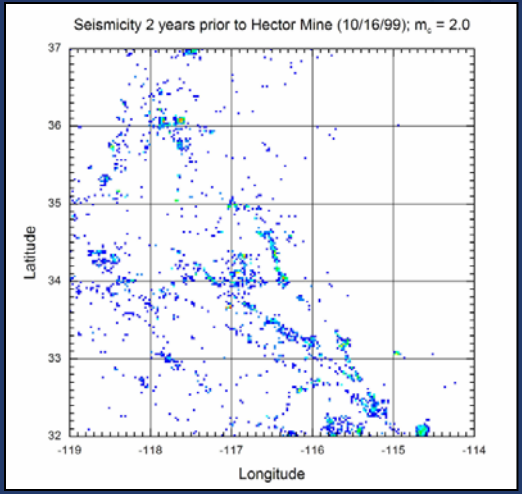
$z_i^O(t) \equiv$ Observed time series

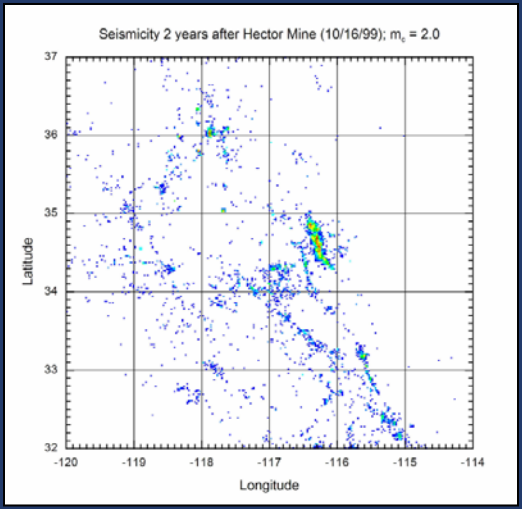
$z_i^S(t+\tau) \equiv$ Simulation time series (offset τ)

Assimilating Coarse-Grained Seismicity Data into Simple Model

Prototype for Complexity Computational Environment SERVOGrid







Steps:

- Coarse grain a spatial region with a spatial grid
- Using models, analyze the earthquake activity time series in each grid box with the idea of using changing space-time seismicity patterns to forecast future activity of large earthquakes

Activity associated with October 19, 1999 Hector Mine earthquake

Potts Model for Coarse-Grained Seismicity Forecasting

$$H \equiv -\frac{1}{2}\sum_{k,k'} J_{k,k'}\left(S\ \delta(s_k s_{k'}) - 1\right) - \sum_{k} h_k \left(S\,\delta(s_1,1) - 1\right)$$

Maximum Entropy - Minimum Free Energy Dynamics → $\dfrac{\partial s_k}{\partial t} = -\dfrac{\partial H}{\partial s_k}$

Here $s_k(t)$ can be in any of the states $s_k(t) = 1,\ldots,S$ at time t, $\delta(s_k, s_k{}')$ is the Dirac delta, and the field h_k favors box k to be in the low energy state $s_k(t) = 1$. This conceptually simple model is a more general case of the *Ising model* of magnetic systems, in which case S = 2. In our case for example, the state variable $s_k(t)$ could be similarly chosen to represent earthquake seismicity, GPS displacements of velocities, or InSAR fringe values.

Prototyping Our Approach

1. **Begin with Manna model** first, then q = 2 state Potts model {same as Ising model H(J,h) }. **Manna model** is an integer version of q = 2 state Potts model, with site variables mod(1).

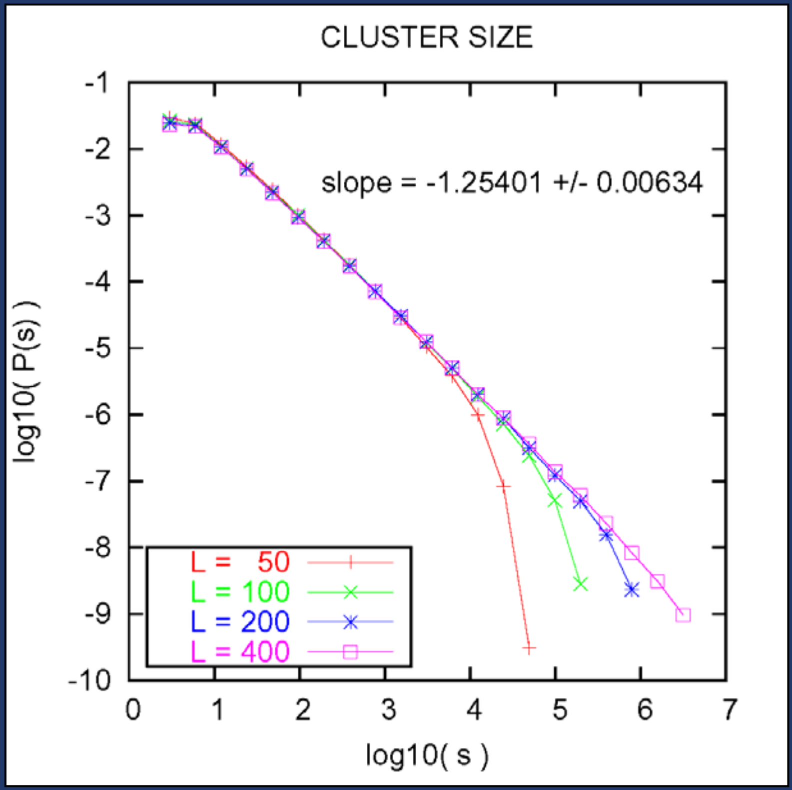
2. **Dynamics**. Erect a simple square lattice Z in the plane. Adopt a simple dynamics based on maximum entropy (minimum free energy). Pick a set of parameters (J,h) that give reasonable time series on each grid point, and produce a synthetic data set, both with and without noise

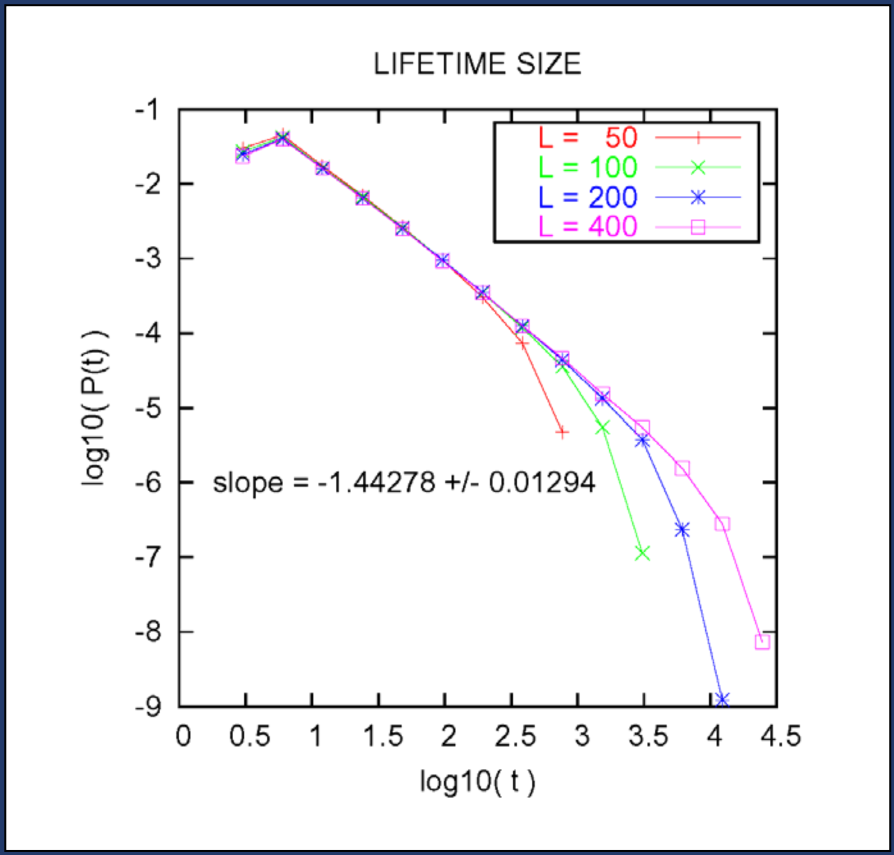
3. **Model Steering Procedure**: Define cost function. Run parallel system of models with stochastic variations in model parameters. Use cost function to determine which model optimally fits test data. Re-compute a new family of stochastic models to iteratively improve model parameters at each time step. Use this process to determine how closely the known values of J,h can be recovered.

4. **Scaling**. Repeat with larger lattices in the presence of noise to determine robustness of algorithm, and how it scales with lattice size.

Example of Manna Model: Scaling of Cluster Size and Cluster Lifetimes

The scaling of these models mimics many features of real data, as well as of Virtual California simulations. Many scales of length and time complicate the process of ensemble forecasting, prediction, and model steering.





Next Steps

1. **Finalize grid search procedure with Manna Model and Potts Model. Formulate procedural runs on Columbia machine as appropriate. Refine PI forecast methods and integrate with Grid services approach.**

2. **Migrate Procedures to Virtual California. Finish development of parallel implementation of Virtual California. Test on Columbia platform.**

3. **Model Steering Procedure**: Refine cost functions, particularly as they relate to Virtual California. Run parallel system of models with stochastic variations in model parameters. Ingest historic and recent data as appropriate.

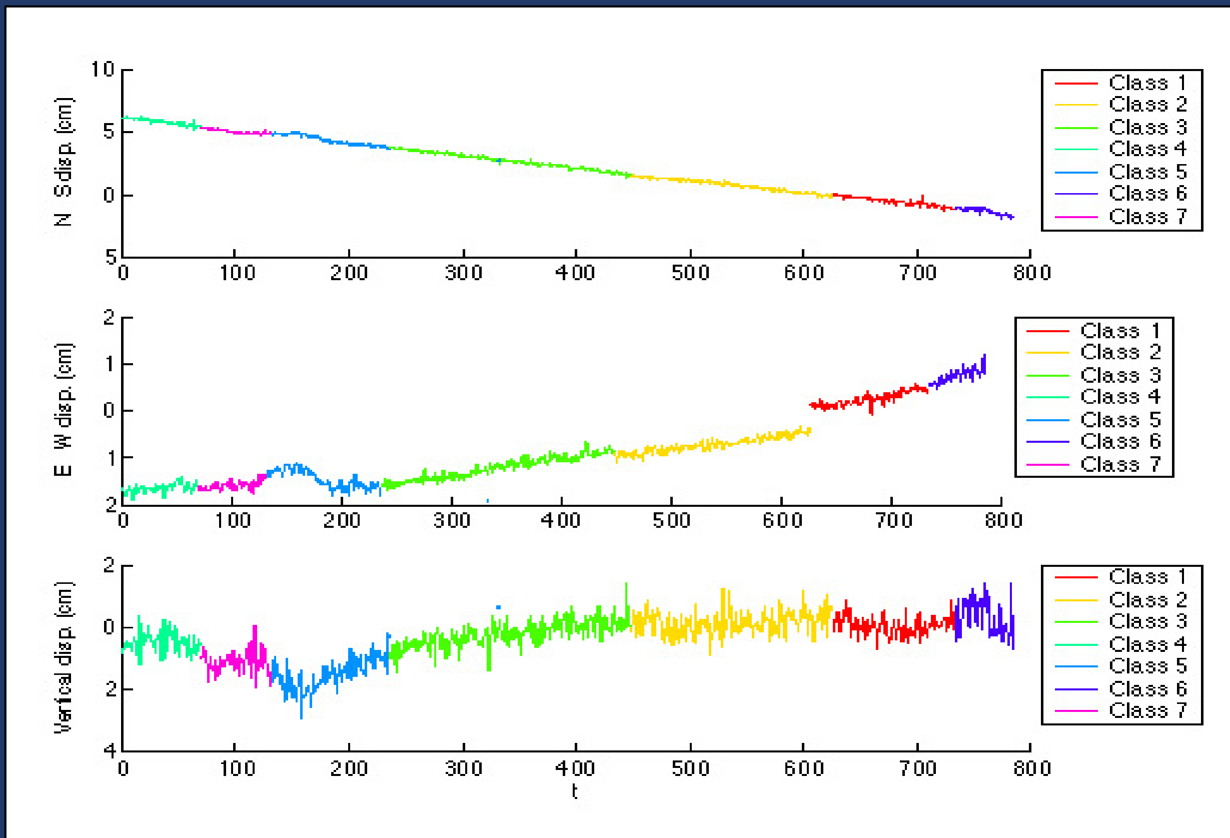
4. **Scaling**. Repeat with larger models in the presence of noise to determine robustness of algorithm, and how it scales with lattice size.

Datamining Infrastructure

Robert Granat
NASA JPL

John Rundle, et. al.
UC Davis

RDAHMM: GPS Time Series Segmentation



GPS displacement (3D) length two years.

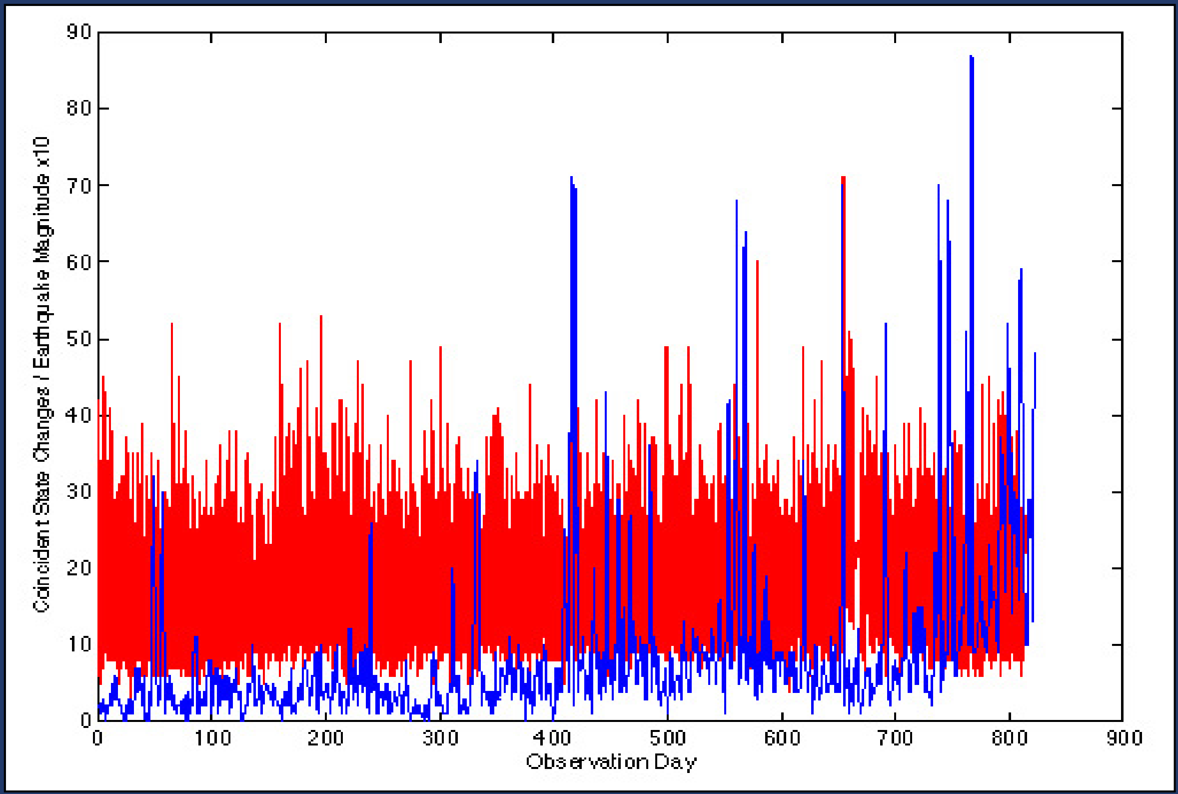
Divided automatically by HMM into 7 classes.

Features:

- Dip due to aquifer drainage (days 120-250)
- Hector Mine earthquake (day 626)
- Noisy period at end of time series

- Complex data with subtle signals is difficult for humans to analyze, leading to gaps in analysis
- HMM segmentation provides an automatic way to focus attention on the most interesting parts of the time series

RDAHMM: SCIGN GPS Network Analysis



Now segment all 127 GPS stations

In blue: Number of stations that change state on a given day

In red: Seismic activity

Days with many state changes often *do not* correlate with large earthquakes.

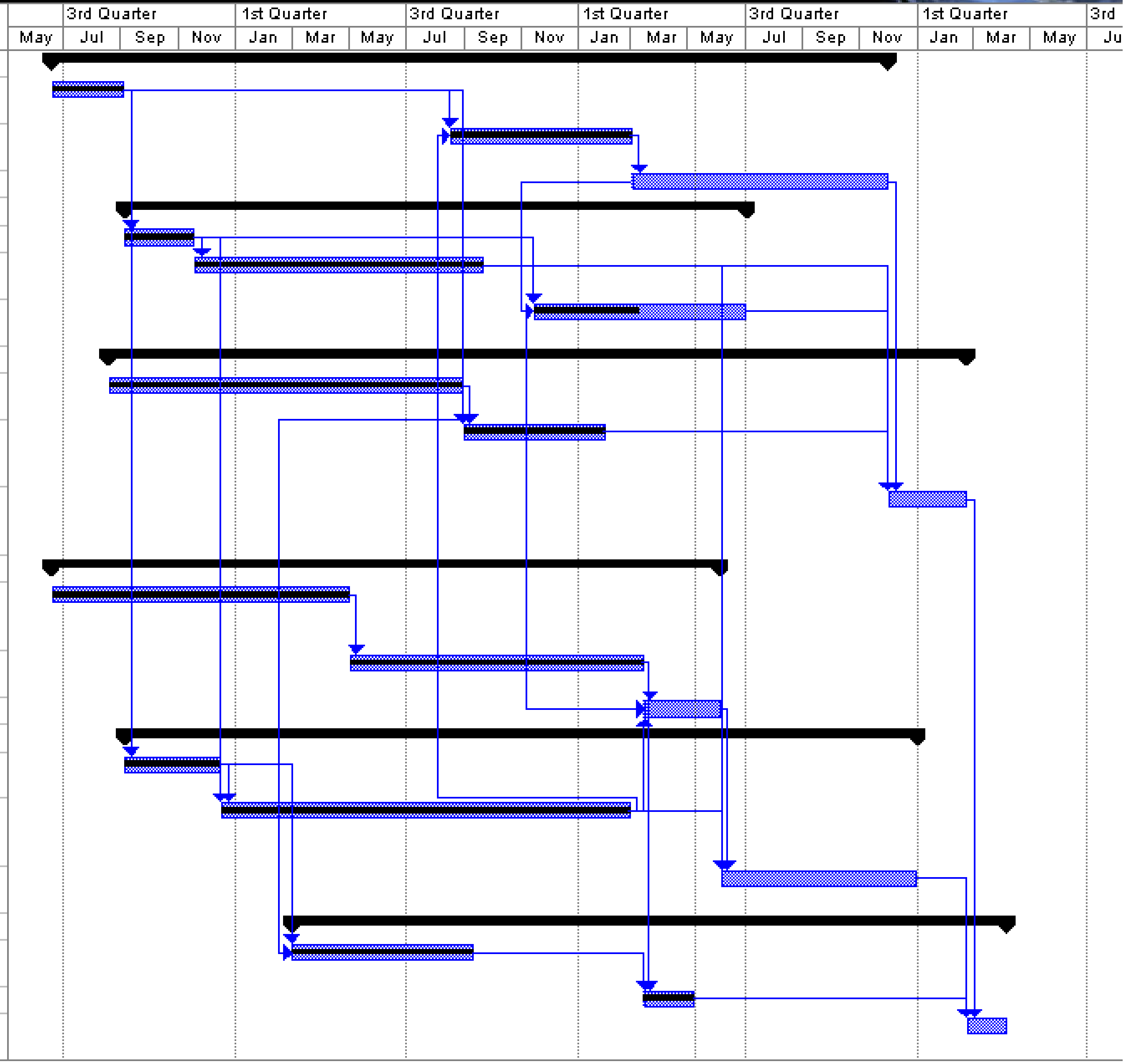
- Have found a way to detect regional aseismic signals
- This software is being integrated with the Quakesim web portal
- Scenarios for use with real time streaming data through the web portal are currently being investigated

Project Management

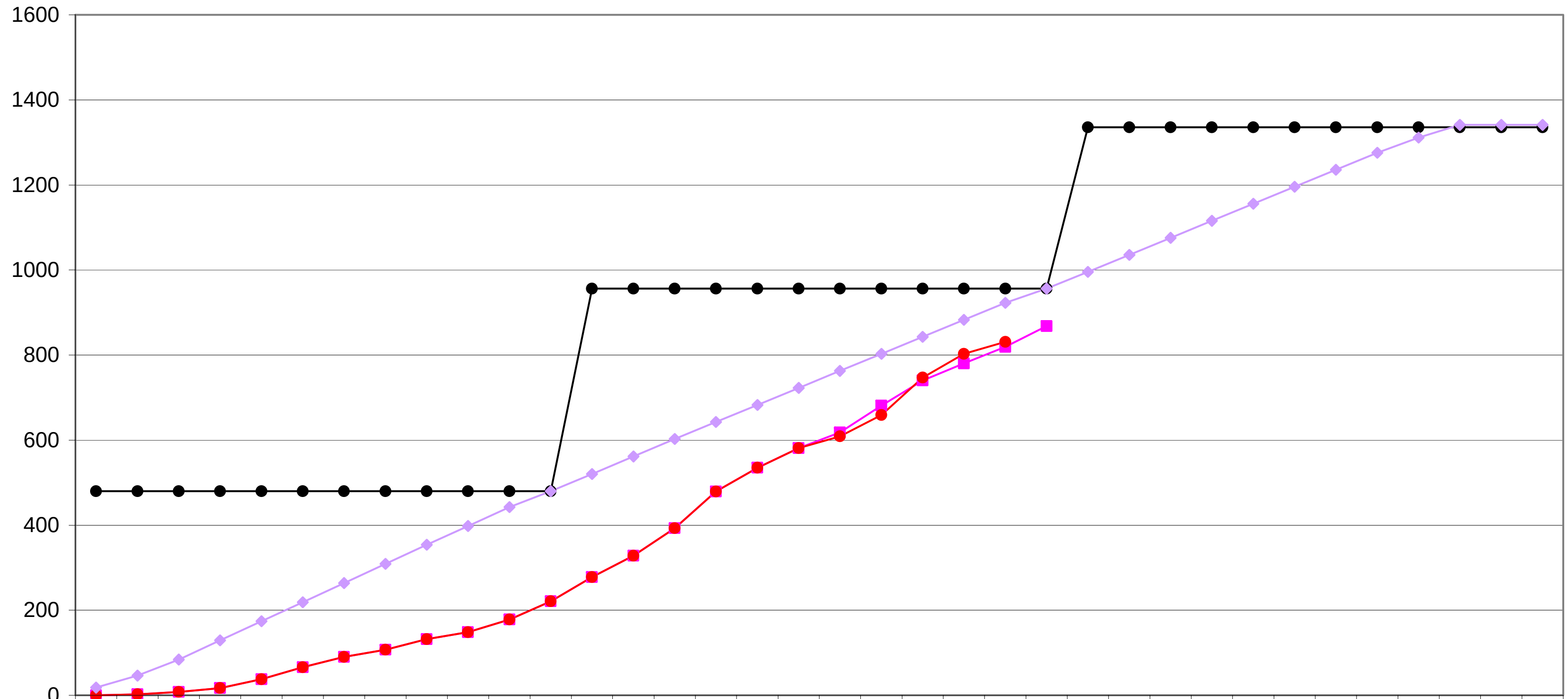
Andrea Donnellan
Michele Judd
NASA JPL

Project Plan

ID		Task Name
1		Metadata Services
2	✓	Develop XML Schemas to support services and data structures of CCE.
3	✓	XML metadata services available as grid services.
4		Services linked to federated database.
5		Federated Database System
6	✓	Develop geophysics meta-ontology.
7	✓	Develop meta-query mediator and translator.
8		Federated database available as Grid service.
9		Data Assimilation Infrastructure
10	✓	Explore coupling methodologies to guide development of cross-scale tools.
11	✓	Prototype data assimilation on SERVO Grid with one application and one coarse graining approach.
12		Full implementation of data assimilation SERVO Grid for multiple applications and multiple coarse
13		Datamining Infrastructure
14	✓	Assess techniques using limited datasets; Determine appropriate resolution scaling.
15	✓	Analyze pattern space to reduce dimensionality on parallel computer.
16		Automate with operational data.
17		Web Services
18	✓	Standalone modules supporting simulation and datagrid components of
19	✓	Data assimilation and coarse graining interfaces defined in Web service framework.
20		Integrate multiple models and data across several platforms.
21		Complexity Computational Environment
22	✓	Detailed design of CCE architecture.
23	✓	Analysis of system performance.
24		Technical report and published papers identifying system characteristics.



Financial Status: Cost



	Jun '03	Jul '03	Aug '03	Sep '03	Oct '03	Nov '03	Dec '03	Jan '04	Feb '04	Mar '04	Apr '04	May '04
Cum Cost Plan	0	2	8	17	38	66	90	107	132	148	178	221
Cum Actuals	0	2	8	17	38	66	90	107	132	148	178	221
Funding Plan	479.	479.	479.	479.	479.	479.	479.	479.	479.	479.	479.	479.
Original Cost Plan Cum	18.0	46.0	84.0	129.	174.	219.	264.	309.	354.	398.	442.	479.

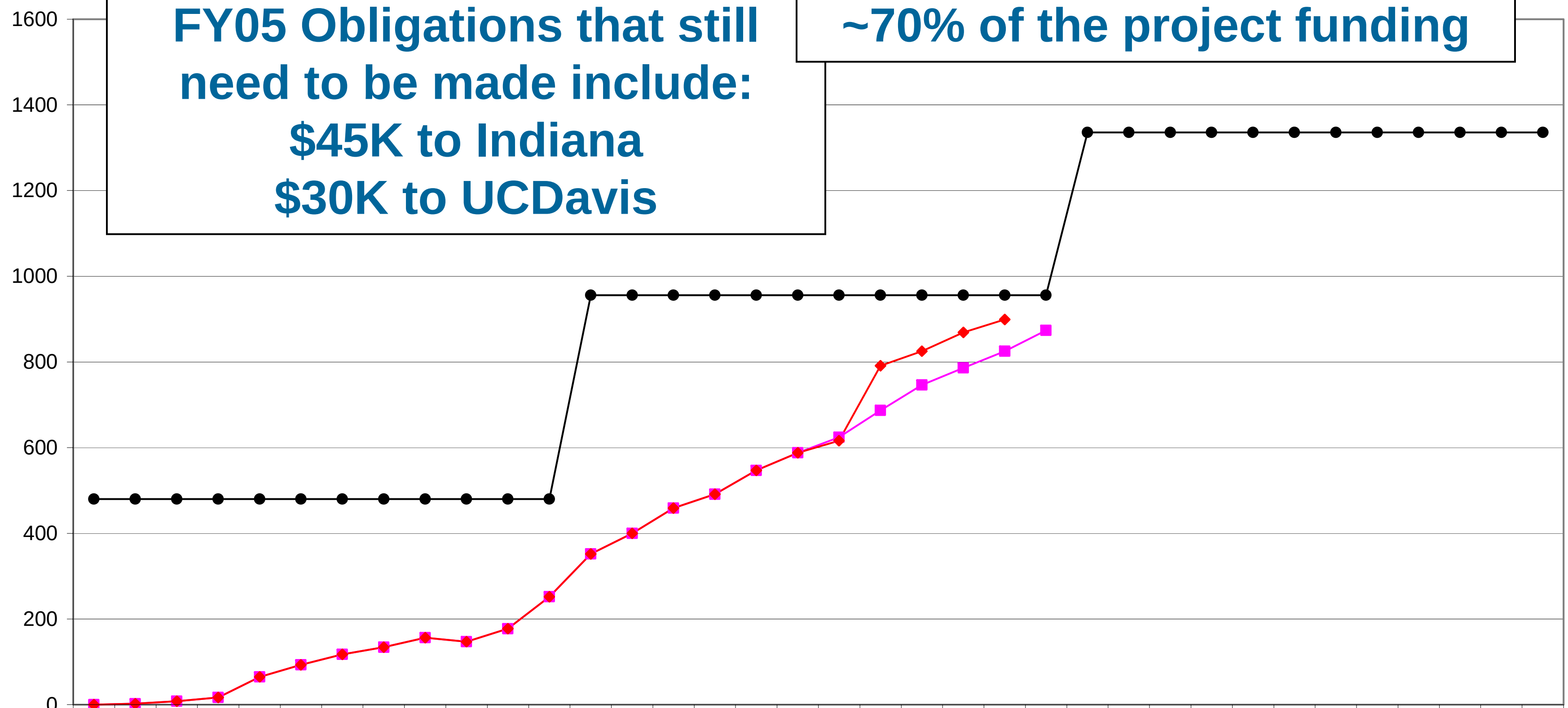
	Jun '04	Jul '04	Aug '04	Sep '04	Oct '04	Nov '04	Dec '04	Jan '05	Feb '05	Mar '05	Apr '05	May '05
Cum Cost Plan	278	328	393	479	535	581	618	681	740	780	819	868
Cum Actuals	278	328	393	479	535	581	609	659	747	803	831	
Funding Plan	955.	955.	955.	955.	955.	955.	955.	955.	955.	955.	955.	955.
Original Cost Plan Cum	520.	561.	602.	642.	682.	722.	762.	802.	842.	882.	922.	955.

	Jun '05	Jul '05	Aug '05	Sep '05	Oct '05	Nov '05	Dec '05	Jan '06	Feb '06	Mar '06	Apr '06	May '06
Cum Cost Plan												
Cum Actuals												
Funding Plan	133	133	133	133	133	133	133	133	133	133	133	133
Original Cost Plan Cum	995.	103	107	111	115	119	123	127	131	134	134	134

Financial Status: Obligations

Subcontracts account for ~70% of the project funding

FY05 Obligations that still need to be made include:
$45K to Indiana
$30K to UCDavis



	Jun '03	Jul '03	Aug '03	Sep '03	Oct '03	Nov '03	Dec '03	Jan '04	Feb '04	Mar '04	Apr '04	May '04
Cum Oblig Plan	0	2	8	17	65	93	117	134	156	147	177	252
Cum Actuals	0	2	8	17	65	93	117	134	156	147	177	252
Funding Plan	479.	479.	479.	479.	479.	479.	479.	479.	479.	479.	479.	479.

	Jun '04	Jul '04	Aug '04	Sep '04	Oct '04	Nov '04	Dec '04	Jan '05	Feb '05	Mar '05	Apr '05	May '05
Cum Oblig Plan	352	400	459	491	547	588	624	687	746	786	825	874
Cum Actuals	352	400	459	491	547	588	616	791	825	869	899	
Funding Plan	955.	955.	955.	955.	955.	955.	955.	955.	955.	955.	955.	955.

	Jun '05	Jul '05	Aug '05	Sep '05	Oct '05	Nov '05	Dec '05	Jan '06	Feb '06	Mar '06	Apr '06	May '06
Cum Oblig Plan												
Cum Actuals												
Funding Plan	1335	1335	1335	1335	1335	1335	1335	1335	1335	1335	1335	1335

TRL Level Moved from 3 to 5

- Since our last annual review, we have moved to TRL 5
- We expect to move from TRL 5 to TRL 6 by the close of the project in 2006

TRL 5 DEFINITION

System/subsystem/component validation in relevant environment

Thorough testing of prototyping in representative environment. Basic technology elements integrated with reasonably realistic supporting elements. Prototyping implementations conform to target environment and interfaces.

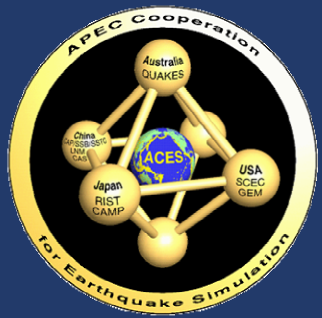
Issues / Concerns

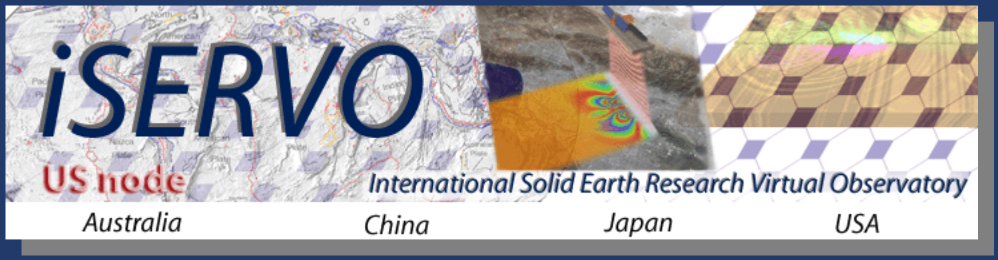
The flow of funding from NASA/JPL has been a continual source of concern for the University Co-Investigators. On top of the continuing resolution delay in funding, the process of transferring the funds out of JPL itself has been difficult.

JPL costs continue to be a concern as it is clear that we did not accurately allot a fair level of support for that portion of the work.

Cooperation with Others

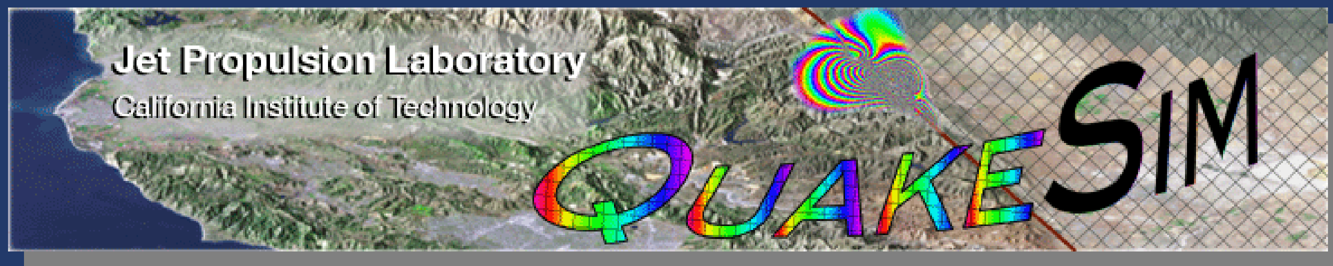
- Collaboration with international partners from Australia, Japan, and China has led to development of iSERVO (international Solid Earth Research Virtual Observatory).
- Organized an iSERVO Colloquium at the ACES workshop in July 2004.





- Made initial contact with Rob Raskin to understand the current stage of his ontology work with the OCEAN ESIP/POET project.

Cooperation with Others



- Working with "active tectonics" (QuakeSim) CT project team members
- NSF Center for Computational Infrastructure for Geodynamics
  - Actively collaborating to develop standards and compatible approaches for planned center
  - Investigating possible co-location of teams in Pasadena, CA

Cooperation with Others

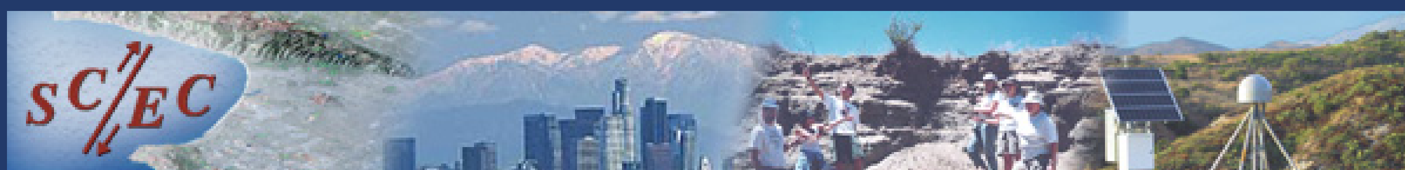
Our Indiana University Co-Is are heavily involved with the Global Grid Forum and the Open Middleware Infrastructure Institute

Geoffrey Fox ➡

- Global Grid Forum
  - Acting leader of the GGF Community Council
  - Member of the Grid Forum Steering Group
  - Co-leader of the Grid Computing Environments and Semantic Grid Research Groups
- Open Middleware Infrastructure Institute (OMII) Advisory Member

Cooperation with Others

- Working with the Open GIS Consortium to develop consistent standards under GML.

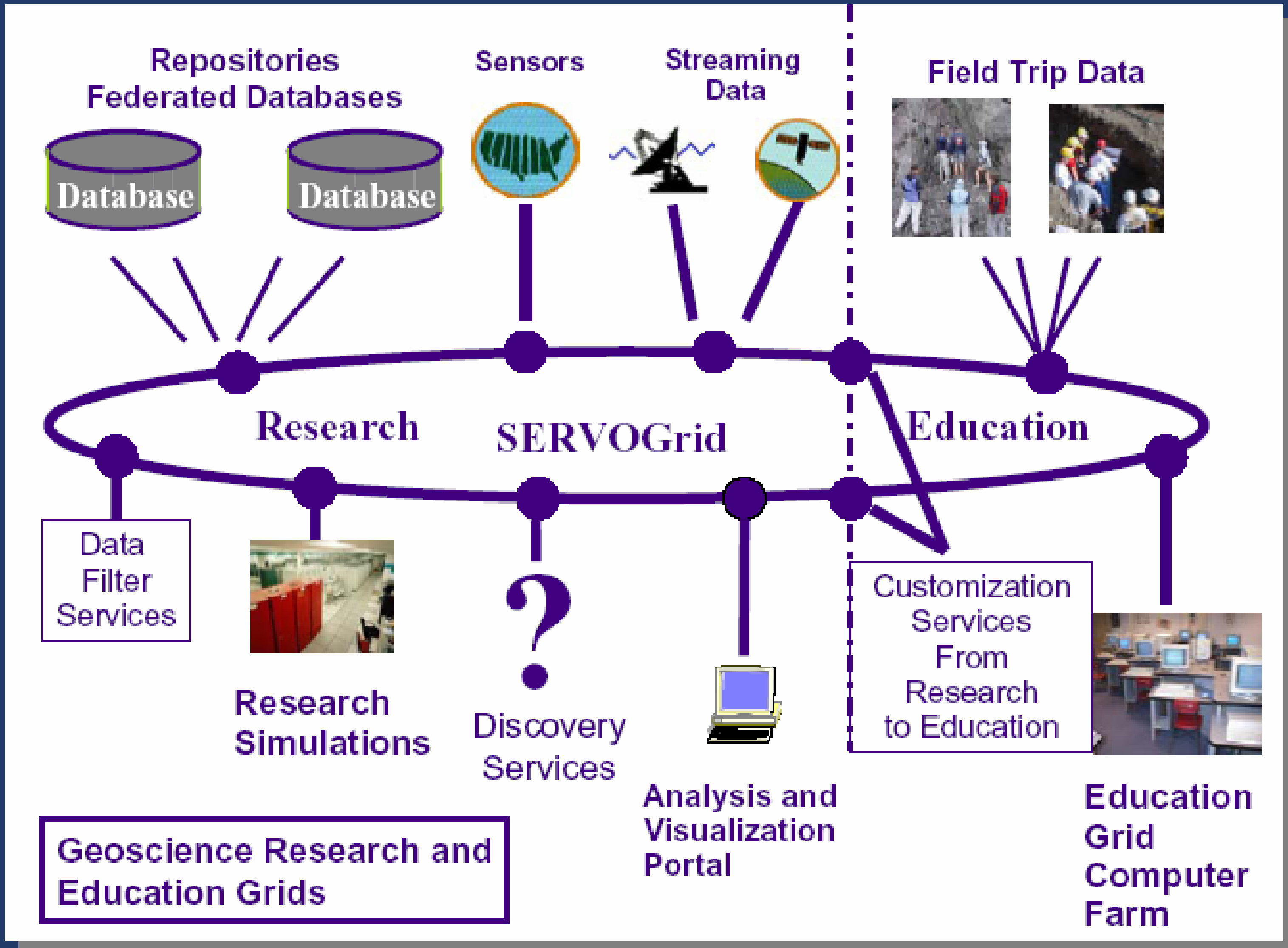


- Collaboration with the Southern California Earthquake Center Information Technology Research Project on computational infrastructure. They are primarily focused on earthquake wave propagation (which is complementary to the AIST work we are doing).

Cooperation with Others

- We are collaborating with SCIGN to build a Sensor Grid
    - Based on OGC's SensorML family of specifications
    - SensorML specification funded by NASA AIST
  - This will allow us to build real-time support for SERVO
    - RDAHMM is an excellent candidate
  - We also have interesting approaches for handling the streams and implementing SensorML.
- Full report describing the collaboration work to date is available at
  - http://complexity.ucs.indiana.edu/~gaydin/sopac/summary.doc
  - Sample code is in the same directory

SERVOGrid links to Education



Students working on SERVOGrid

- **Mehmet Aktas**, graduate student at Indiana University, has been working on Resource Description Framework (RDF) -based ontologies for describing AIST infrastructure (computers, data sources, codes, etc). Mehmet has also been investigating distributed RDF queries, which will allow researchers to do smart searches over distributed RDF fragments (the RDF can be fragmented over several hosts).

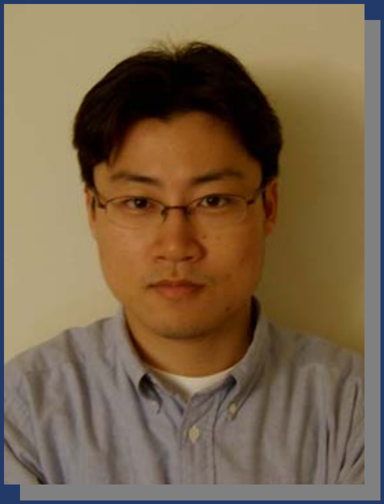
- **Ahmet Sayar**, graduate student at Indiana University, has been working on file-based workflow, and in particular, on a scripting service based on Apache Ant. This will allow us to generate workflow scripts that can be used to couple sequences of applications that are linked by files: the input of one application is the output of another.

- **Harshawardhan Gadgil**, graduate student at Indiana University, has been working on data stream-based workflow. Although his work was begun independently of AIST, the synergies are such that we will be infusing this research into the project.

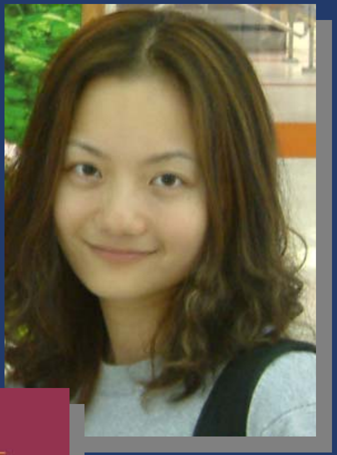
Students Working on SERVOGrid

- **Galip Aydin** is a graduate student at Indiana University who is implementing a Web Service compliant version of the WFS. He is using this to serve standard GML geographic features, including faults from the QuakeTables database.

- **Dongwoo Won** is a PhD student in the Computer Science Department at the University of Southern California. He has worked on the federated database, on ontology federations, and the definition of Earth science ontologies, using various sources. He has also been working with his USC colleague Sang-Soo on Ontronic and on the issues of information discovery and mining.



Students working on SERVOGrid



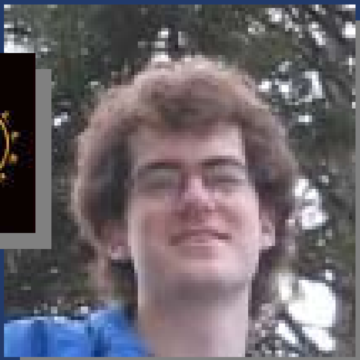
- **Anne Chen**, Ph.D. student at USC, has been focusing on our approach and architecture for access to heterogeneous datasets, and on wedding state-of-the-art database technology with web services.

- **Sangsoo Sung**, Ph.D. student at USC, joined the project January 2004; he is focusing on representing the meaning/semantics of data in our various data repositories/sources, and on (partially) integrating these data.

Students working on SERVOGrid

- **Jordan van Aalsburg**, graduate student at U.C. Davis is working on this project with a postdoctoral fellow, Dr. Robert Shcherbakov. They are researching a variety of the standard data assimilation techniques based on applying Kalman filter and neural network back-propagation techniques to deformation and seismicity data.

- **Eli Bogart** is an undergraduate student at Harvey Mudd college who is using the ESTO-developed GeoFEST code to numerically generate stress Green's functions for Virtual California. We hope to use this approach to incorporate viscoelastic behavior and inhomogenous material properties into the Virtual California simulations.

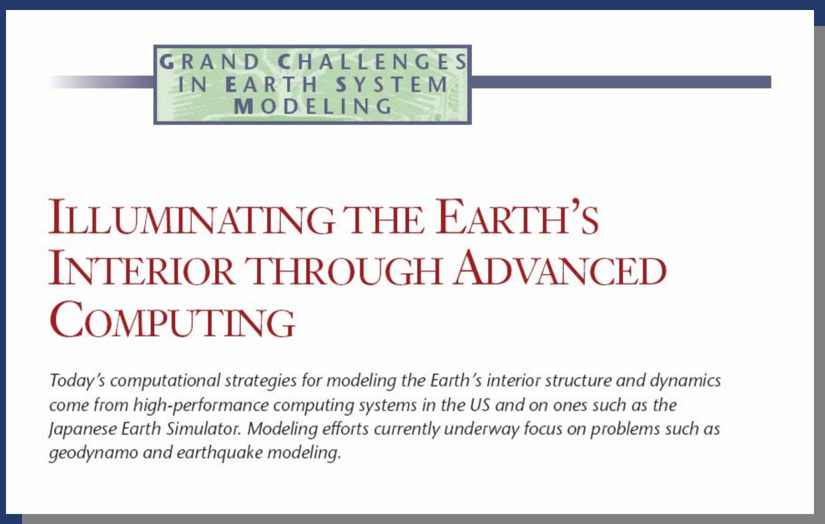


Papers/Abstracts/Posters

- Chen, A.Y., A. Donnellan, D. McLeod, G. Fox, J. Parker, J. Rundle, L. Grant, M., Pierce, M. Gould, S. Chung, S. Gao, "Interoperability and Semantics for Heterogeneous Earthquake Science Data", International Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data, Sanibel Island, FL, October 2003.
- Donnellan, A., J.B. Rundle, G. Fox, M. Pierce, D. McLeod, J.W. Parker, T. Tullis, L. Grant, "The Solid Earth Research Virtual Observatory", Eos Trans. AGU, 84(46), Fall Meet. Suppl., Abstract U21B-03, 2003.
- Granat, R., R. Clayton, S. Kedar, Y. Kaneko "Regularized Deterministic Annealing Hidden Markov Models for Identification and Analysis of Seismic and Aseismic Events", Eos Trans. AGU, 84(46), Fall Meet. Suppl., Abstract NG41C-0076, 2003.

Papers/Abstracts/Posters

– Mora, P., A. Donnellan, G. Fox, M. Pierce, M. Matsu'ura, D. McLeod, X. Yin, "The international Solid Earth Virtual Research Observatory (iSERVO) Institute Seed Project", Eos Trans. AGU, 84(46), Fall Meet. Suppl., Abstract NG12B-04, 2003.


GRAND CHALLENGES IN EARTH SYSTEM MODELING

ILLUMINATING THE EARTH'S INTERIOR THROUGH ADVANCED COMPUTING

Today's computational strategies for modeling the Earth's interior structure and dynamics come from high-performance computing systems in the US and on ones such as the Japanese Earth Simulator. Modeling efforts currently underway focus on problems such as geodynamo and earthquake modeling.

– Donnellan, A., Rundle, J., Ries, J., Fox, G., Pierce, Parker, J., Crippen, R., M., Dejong, E., Chao, Kuang, W., McLeod, D., Matu'ura, M., Bloxham, J., "Illuminating the Earth's Interior through Advanced Computing", Computing in Science and Engineering, Volume 6, Number 1, Pages 336-44, January/February 2004.

– Aktas, M.S., M.E. Pierce, G.C. Fox, " Designing Ontologies and Distributed Resource Discovery Services for an Earthquake Simulation Grid", GGF11, Semantic Grid Applications Workshop, Honolulu, June 2004.

Papers/Abstracts/Posters

- Granat, R., “Regularized Deterministic Annealing EM for Hidden Markov Models,” **Doctoral Dissertation, University of California, Los Angeles**, June, 2004.
- Donnellan, A., G. Fox, J. Rundle, T. Tullis, D. McLeod, L. Grant, J. Parker, M. Pierce, G. Lyzenga and R. Granat, The US QuakeSim and SERVO Projects, Third **International Conference on Continental Earthquakes** invited keynote lecture, Beijing, China, July, 2004.
- Granat, R., Statistical Modeling of Geodetic Networks for Detecting Regional Events, Fourth **ACES Workshop**, Beijing, China, July, 2004.
- Pierce, M., G. Fox, G. Aydin, and M. Aktas, Data Modeling and Information Management for Earthquake Applications and Grids, Fourth **ACES Workshop**, Beijing, China, July, 2004.
- Pierce, M., and G. Fox, Grids and Web Services for Earthquake Simulation, Fourth **ACES Workshop**, Beijing, China, July, 2004.

Papers/Abstracts/Posters

- Donnellan, A., J. Rundle, G. Fox, D. McLeod, L. Grant, T. Tullis, M. Pierce, J. Parker, G. Lyzenga, R. Granat, M. Glasscoe, QuakeSim and the Solid Earth Research Virtual Observatory, *PAGEOPH*, submitted.
- Grant L.B., **A. Donnellan**, D. McLeod, M. Pierce, G.C. Fox, A.Y. Chen, M.M. Gould[1], S.S. Sung, P.B. Rundle, A Web-Service Based Universal Approach to Heterogeneous Fault Databases, *Computing in Science and Engineering Special Issue on Multi-Physics Modeling,* in press.
- Donnellan, A., G. Fox, J. Rundle, T. Tullis, D. McLeod, L. Grant, J. Parker, M. Pierce, G. Lyzenga, and R. Granat, Enabling Earthquake Science with the Solid Earth Research Virtual Observatory (SERVO), **Southern California Earthquake Center Fall Meeting**, Palm Springs, 2004.

Papers/Abstracts/Posters

- Fox, G., M. Pierce, J. Rundle, A. Donnellan, J. Parker, R. Granat, G. Lyzenga, D. McLeod, and L. Grant, The International Solid Earth Research Virtual Observatory, Eos Trans. **AGU**, 85 (47), 2004 Fall Meet. Suppl., Abstract SF31B-07, 2004.
- Granat, R., Hidden Markov Models for Detecting Aseismic Events in Southern California, Eos Trans. **AGU**, 85 (47), 2004 Fall Meet. Suppl., Abstract NG43A-0436, 2004.
- Pierce, M., G. Fox, and S. Pallickara, Implementing Geographic Information System Grid Services Using Distributed Messaging Systems, Eos Trans. **AGU**, 85 (47), 2004 Fall Meet. Suppl., Abstract NG34A-07, 2004.
- Aktas, M., M. Pierce, G. Fox, and D. Leake, A Web based Conversational Case-Based Recommender System for Ontology Aided Metadata Discovery, November 8, Proceedings of the 5th IEEE/ACM International Workshop on Grid Computing (**GRID 2004**), 2004.

Papers/Abstracts/Posters

- Fox, G., S. Pallickara, M. Pierce, and H. Gadgil, Messaging and the Web Service Distributed Operating Environment, Submitted to Proceedings of the **Royal Society (UK) Special Issue on Scientific Grid Applications**.
- Fox, G., and M. Pierce, Web Service Grids for iSERVO, Proceedings of the **International Workshop on Geodynamics: Observation, Modeling and Computer Simulation, University of Tokyo**, Japan, October 14 2004.
- Aydin, G., M. Pierce, G. Fox, M. Aktas and A. Sayar, Implementing GIS Grid Services for the International Solid Earth Research Virtual Observatory. Extended abstract submitted to Proceedings of the **4th ACES Workshop, Beijing, China, June 2004.**
- Aktas, M., G. Aydin, A. Donnellan, G. Fox, R. Granat, L. Grant, G. Lyzenga, D. McLeod, S. Pallickara, J. Parker, M. Pierce, J. Rundle, A. Sayar, and T. Tullis, iSERVO: Implementing the International Solid Earth Research Virtual Observatory by Integrating Computational Grid and Geographical Information Web Services. Submitted to **Pure and Applied Geophysics for ACES** special Issue.
- Gadgil, H., G. Fox, S. Pallickara, M. Pierce, and R. Granat, A Scripting based Architecture for Management of Streams and Services in Real-time Grid Applications. Conference paper submission to **Cluster Computing and Grid 2005 (CCGrid05).** 2005.

Presentations

Andrea Donnellan gave an invited keynote address on *The US QuakeSim and SERVO Projects* at the **3rd International Conference on Continental Earthquakes** in Beijing, China, in July 2004.

Co-Investigator **John Rundle was honored as the invited Lorenz Lecturer** where he spoke on, “Process, Pattern, Prediction: Understanding Complexity in Driven Earth Systems” and explained his SERVOGrid work to 400+ geophysicists on December 14th.

Presentations

- Donnellan, A., "Living on a Restless Planet," IEEE Aerospace Conference, Big Sky, Montana, March 9, 2005.
- Pierce, "SERVOGrid: Grid Services and Portals to Support Earthquake Science", Geoinformatics Science Group (Indiana University, Purdue University and IUPUI), Indiana University, October 2003
- Fox, G., CEE and Grid Architectures. **Community Grids Lab** Seminar on Collaborative Technologies September 15 2004.
- Fox, G., Web Service Grids for iSERVO International Workshop on Geodynamics: Observation, Modeling and Computer Simulation. **University of Tokyo**, Japan, October 14 2004.
- Fox, G., Architecture of Web Service Grids. **IIT Chicago Computer Science Colloquium** October 25 2004.

Presentations

- Pierce, M., Developing SERVOGrid: e-Science for Earthquake Simulation. **DOD Programming Environment and Training Internet Seminar with Jackson State University**, Oct 26, 2004.
- Fox, G., Pallickara, S., and Parastatidis, S., Towards Flexible Messaging for SOAP Based Services. **SC04** Technical Program, Pittsburgh November 9 2004. (Described fundamental work to support SOAP messaging directly with the NaradaBrokering framework.)
- Fox, G., and Pierce, M., Data Grids for HPC: Geographical Information Systems Grids. **DOD Programming Environment and Training Internet Seminar with Jackson State University**, Dec 7, 2004. (Described work building GIS Web Service Grids.)
- Fox, G., Possible Architectural Principles for OGSA-UK and other Grids <ogsaukjan05.ppt> **UK e-Science Core Programme Town Meeting** London Monday 31st January 2005 “Defining the next Level of Services for e-Science” http://grids.ucs.indiana.edu/ptliupages/presentations/ogsaukjan05.ppt

Future Outlook: Final year

- Services linked to federated database.
- Federated database available as Grid service.
- Full implementation of data assimilation SERVO Grid for multiple applications and multiple coarse graining methods.
- Automate with operational data.
- Integrate multiple models and data across several platforms.
- Technical report and published papers identifying system characteristics.

Questions for the SERVOGrid Team?

Backup and Supplemental Material

Acronyms

ACES	APEC (the Asia Pacific Economic Cooperation) Cooperation for Earthquake Simulation
CCE	Complexity Computational Environment
CGL	Community Grids Laboratory
DAML	DARPA Agent Markup Language
ESG	Earth System Grid
GeoFEST	Geophysical Finite Element Simulation Tool
GEON	Geosciences Network
GIS	Geographic Information Systems
GML	Geographic Markup Language
GMT	Generic Mapping Tool
HMM	Hidden Markov Model
InSAR	Interferometric Synthetic Aperture Radar

Acronyms continued

NV	Name Value
OGC	Open GIS Consortium
OGSA-DAI	Open Grid Services Architecture Data Access and Integration
OWL	No acronym, it is a web ontology language
PDPC	Phase Dynamical Probability Change
RDAHMM	Regularized Deterministic Annealing Hidden Markov Model
RDF	Resource Description Framework
TSE	Task Scheduling Engine
UDDI	Universal Description, Discovery, and Integration
WFS	Web Feature Service
WMS	Web Map Service
WS-I+	Web Service Interoperability
XML	Extensible Markup Language

NaradaBrokering

NaradaBrokering

Managing Streams

- NaradaBrokering
  - Messaging infrastructure for collaboration, peer-to-peer and Grid applications
  - Implements high-performance protocols (message transit time of 1 to 2 ms per hop)
  - Order-preserving, optimized message transport with QoS and security profiles for sent and received messages
  - Support for different underlying protocols such as TCP, UDP, Multicast, RTP
  - *Discovery Service to locate nearest brokers*

HPSearch and NaradaBrokering

- HPSearch uses NaradaBrokering to route data streams
  - Each stream is represented by a topic name
  - Components subscribe / publish to specified topic
    - The WSProxy component automatically maps topics to Input / Output streams
    - Each `write (byte[] buffer)` and `byte[] read()` call is mapped to a NaradaBrokering event

WFS and Data Sources

Where Is the Data?

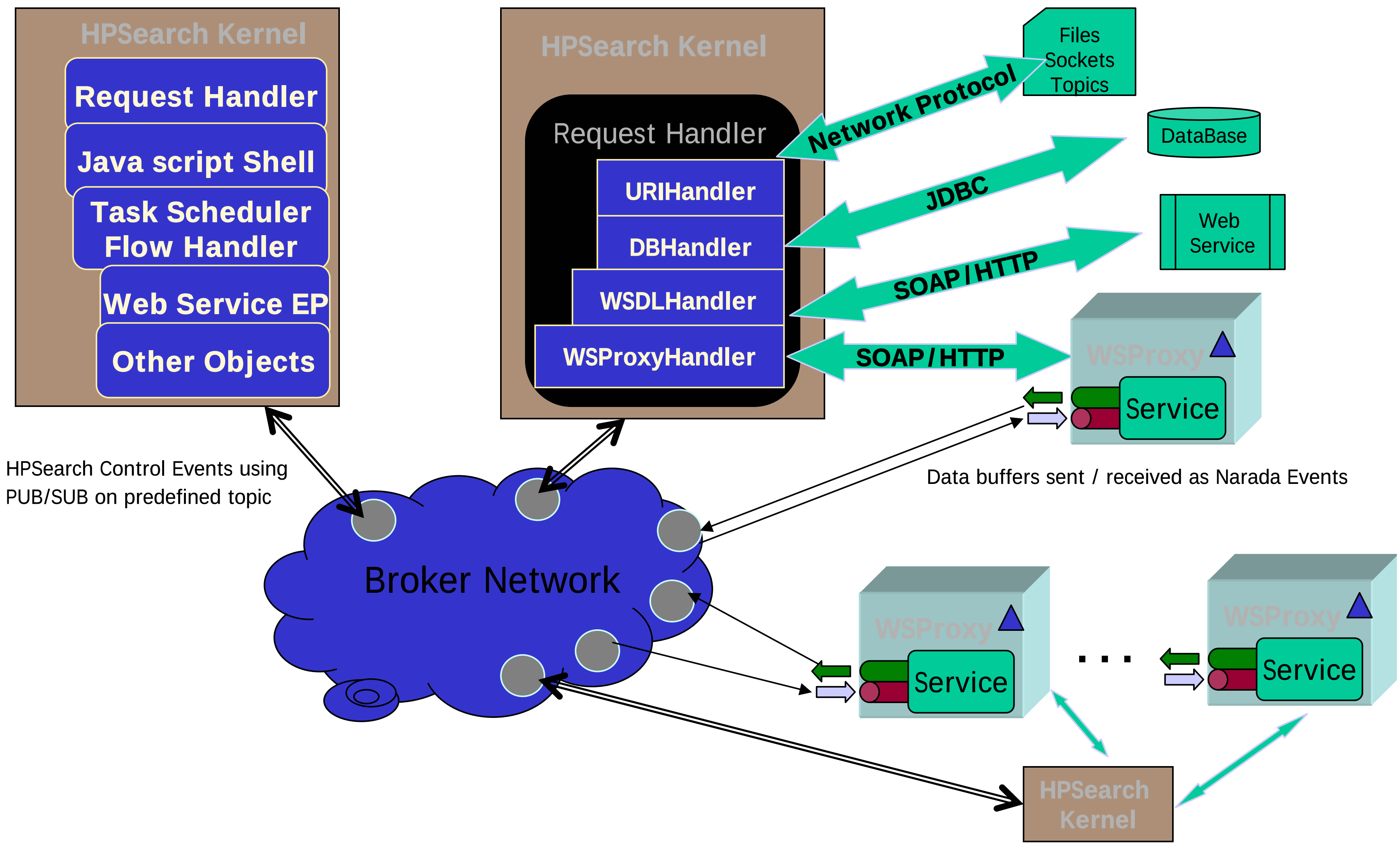
- QuakeTables Fault Database
  - SERVO's fault repository for California.
  - Compatible with GeoFEST, Disloc, and VirtualCalifornia
  - http://infogroup.usc.edu:8080/public.html
- GPS Data sources and formats (RDAHMM and others).
  - JPL: ftp://sideshow.jpl.nasa.gov/pub/mbh
  - SOPAC: ftp://garner.ucsd.edu/pub/timeseries
  - USGS: http://pasadena.wr.usgs.gov/scign/Analysis/plotdata/
- Seismic Event Data (RDAHMM and others)
  - SCSN: http://www.scec.org/ftp/catalogs/SCSN
  - SCEDC: http://www.scecd.scec.org/ftp/catalogs/SCEC_DC
  - Dinger-Shearer: http://www.scecdc.org/ftp/catalogs/dinger-shearer/dinger-shearer.catalog
  - Haukkson: http://www.scecdc.scec.org/ftp/catalogs/hauksson/Socal
- This is the raw material for our data services in SERVO

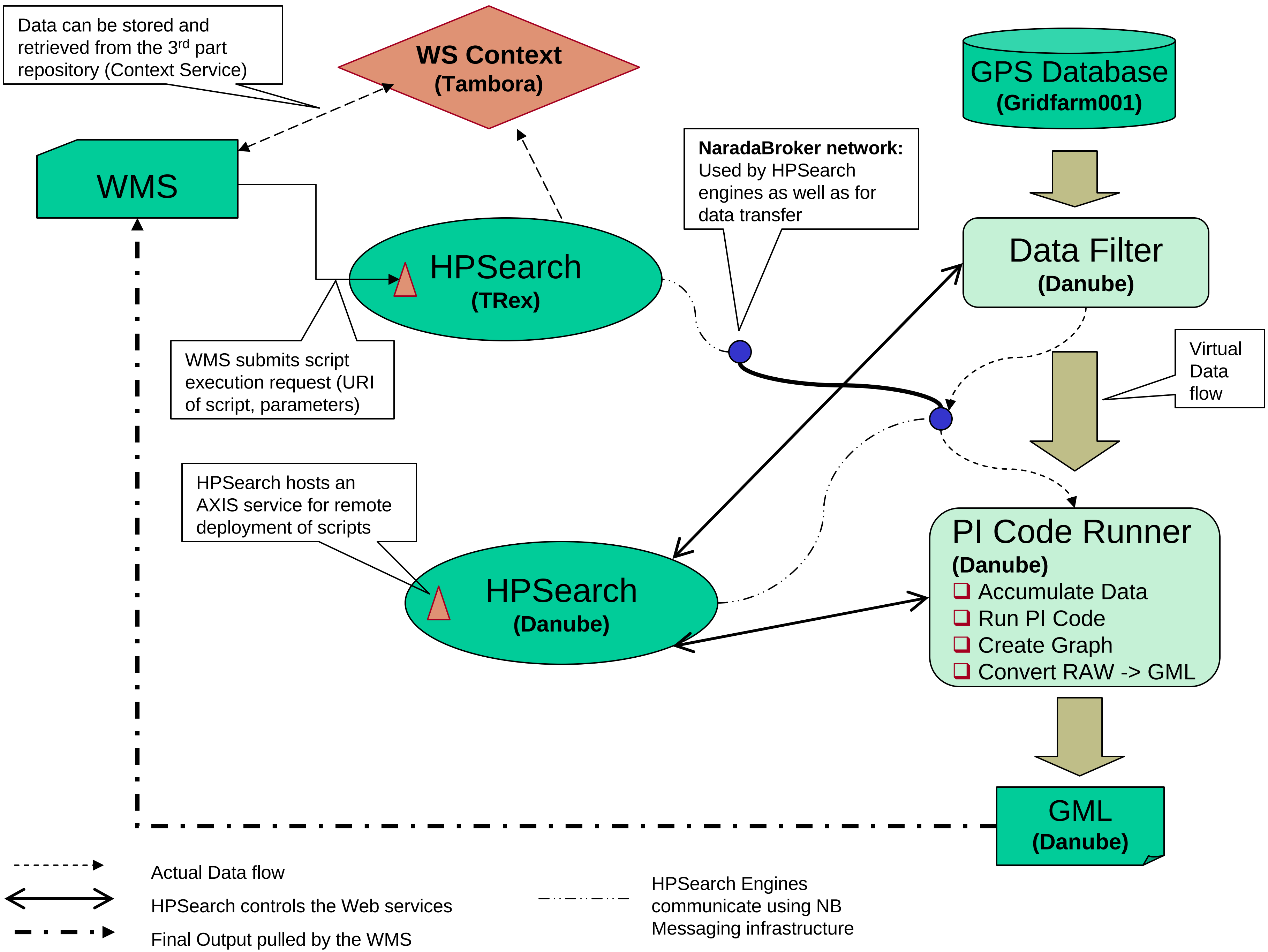
WFS by the Numbers

- The following data is available in the SERVO Web Feature Services
  - These were collected from public sites
  - We have reformatted to GML
- Data
  - Filtered GPS archive (297 stations) from : 48.02MB
  - Point GPS archive (766 stations): 42.94MB
  - SCEDC Seismic archive: 34.83MB
  - SCSN Seismic archive: 26.34MB
  - California Faults (from QuakeTables Fault DB): 62KB
  - CA Fault Segments (from QuakeTables Fault DB): 41KB
  - Boundaries of major European Cities: 12.7KB
  - European map data: 636KB
  - Global Seismic Events:14.8MB
  - US Rivers: 11KB
  - US Map-State Borders: 1.13MB
  - US State Capitals:5.75KB
- WFS URLs
  - http://gf1.ucs.indiana.edu:7474/axis/services/wfs?wsdl
  - http://gf1.ucs.indiana.edu:7474/wfs/testwfs.jsp

HPSearch Diagrams

HPSearch Architecture Diagram





Data can be stored and retrieved from the 3rd part repository (Context Service)
WS Context
(Tambora)
GPS Database
(Gridfarm001)
WMS
NaradaBroker network:
Used by HPSearch engines as well as for data transfer
HPSearch
(TRex)
Data Filter
(Danube)
WMS submits script execution request (URI of script, parameters)
Virtual Data flow
HPSearch hosts an AXIS service for remote deployment of scripts
PI Code Runner
(Danube)
Accumulate Data
Run PI Code
Create Graph
Convert RAW -> GML
HPSearch
(Danube)
GML
(Danube)
Actual Data flow
HPSearch controls the Web services
Final Output pulled by the WMS
HPSearch Engines communicate using NB Messaging infrastructure

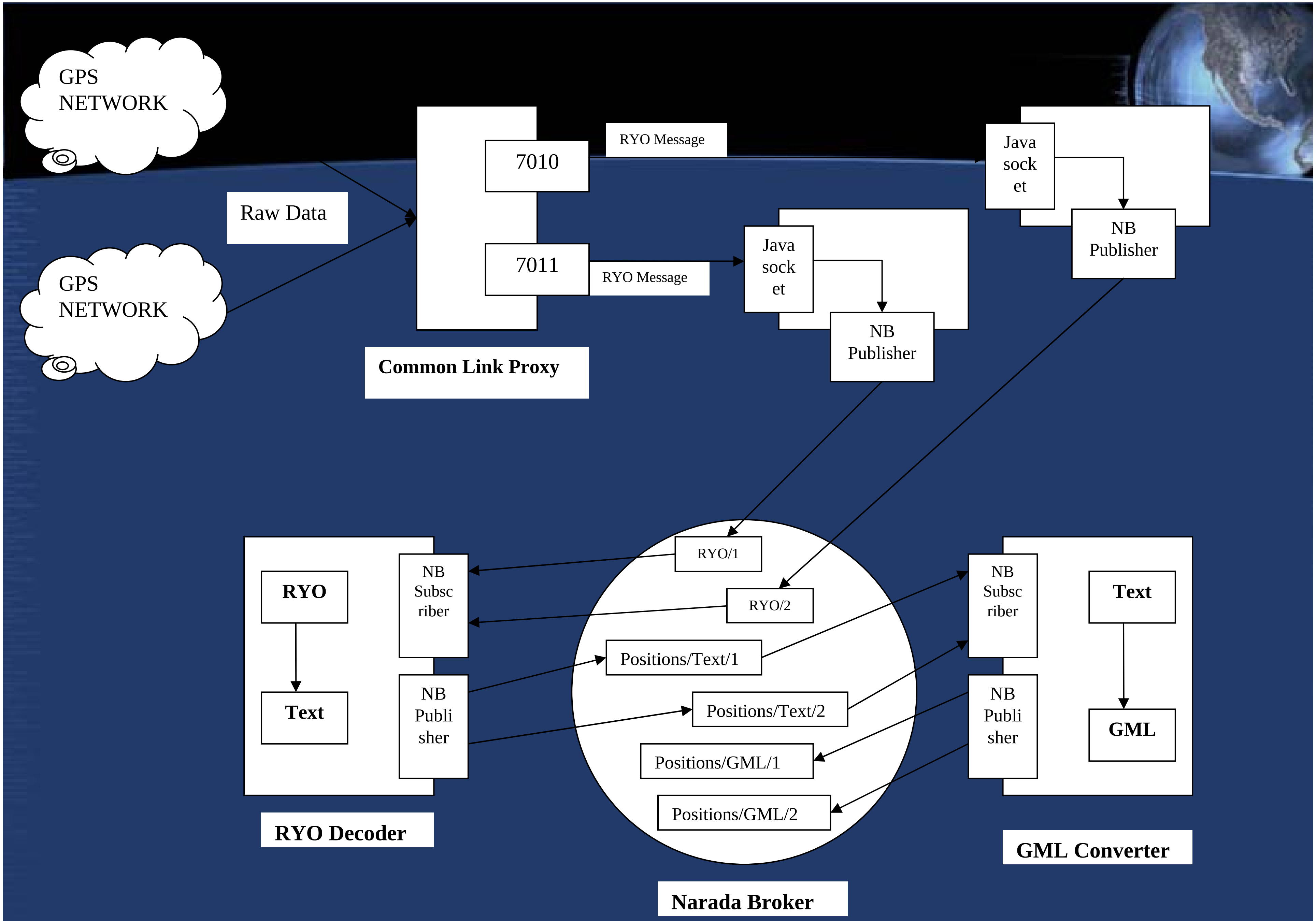
SensorML Work

Collaborations with SCIGN (Aydin)

- Installed NaradaBrokering node
  - Can be used to manage multiple streams using topic-based publish/subscribe techniques
- Uses chains of publishers and subscribers to
  - Deliver binary RYO data directly to subscribing clients, OR
  - Deliver RYO to Text translators that publish to interested clients, OR
  - Deliver RYO text to GML translators, for delivery to interested clients
- Used topic-based stream organization
  - RYO binary format on SOPAC/GPS/Positions/XXX/RYO
  - Text processed RYO messages available on SOPAC/GPS/Positions/XXX/Text
  - XXX is the location of the station (San Diego, Riverside)
- Full report available from here
  - http://complexity.ucs.indiana.edu/~gaydin/sopac/summary.doc
  - Sample code is in the same directory

In Development: SensorML

- CGL has great deal of experience with building software to support audio/video streams.
  - Code, lessons learned can be applied to real time data GPS data streams.
- We are collaborating with Scripps to build a Sensor Grid
    - Based on OGC's SensorML family of specifications
    - SensorML specification funded by NASA AIST
  - This will allow us to build real-time support for SERVO
    - RDAHMM is an excellent candidate
  - We also have interesting approaches for handling the streams and implementing SensorML.
- This work is just underway
  - We only preview it here.


GPS NETWORK
GPS NETWORK
Raw Data
7010
7011
RYO Message
RYO Message
Common Link Proxy
Java socket
NB Publisher
Java socket
NB Publisher
RYO
Text
NB Subscriber
NB Publisher
RYO Decoder
RYO/1
RYO/2
Positions/Text/1
Positions/Text/2
Positions/GML/1
Positions/GML/2
Narada Broker
NB Subscriber
NB Publisher
Text
GML
GML Converter

SensorML and NaradaBrokering

- Common Link Proxy: this is provided by SCIGN.
  - We connect to these sockets, grab the data, and publish it.
- NaradaBroker: this is the publish/subscribe manager node.
  - The enclosed squares (RYO/1, Positions/Text/1, etc) are topics.
- RYO Decoder: this translates RYO binary to text.
  - Subscribes to binary stream
  - Publishes back to the Text topics.
- GML Converter: translates text to GML
  - Subscribes to Text stream topics
  - Publishes on GML stream topics

People

Community Grids and the Grid Community

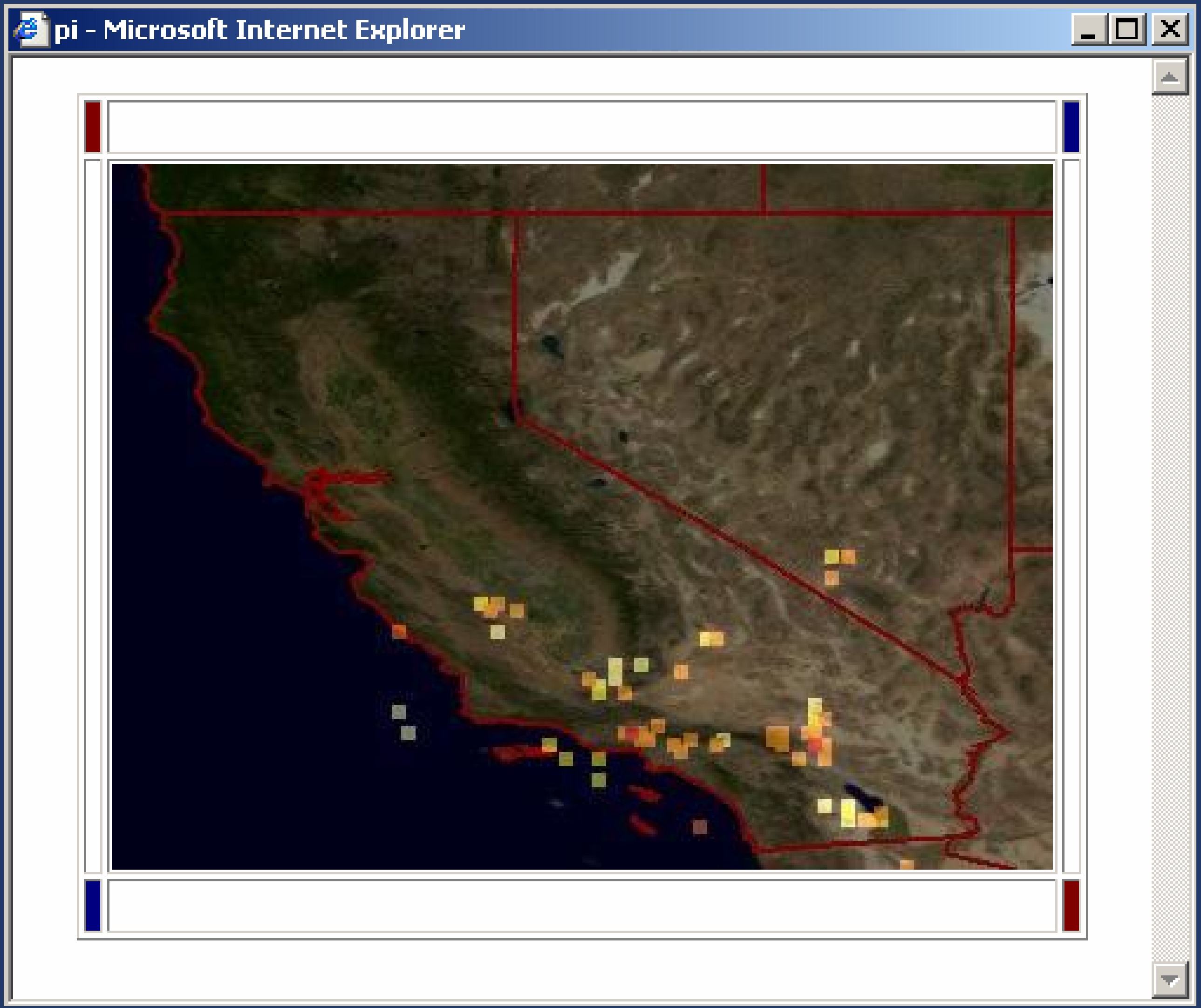
- **Geoffrey Fox**
  - Global Grid Forum
    - Acting leader of the GGF Community Council
    - Member of the Grid Forum Steering Group
    - Co-leader of the Grid Computing Environments and Semantic Grid Research Groups
  - UK e-Science Advisory Group Member
  - Open Middleware Infrastructure Institute (OMII) Advisory Member
  - Los Alamos D Division Advisory Board
    - Sensors, decision support systems, GIS Grids
- **Marlon Pierce**
  - Secretary of the GGF Semantic Grid Working Group
  - Party whip for the Grid Computing Environments Working Group

Who Does What?

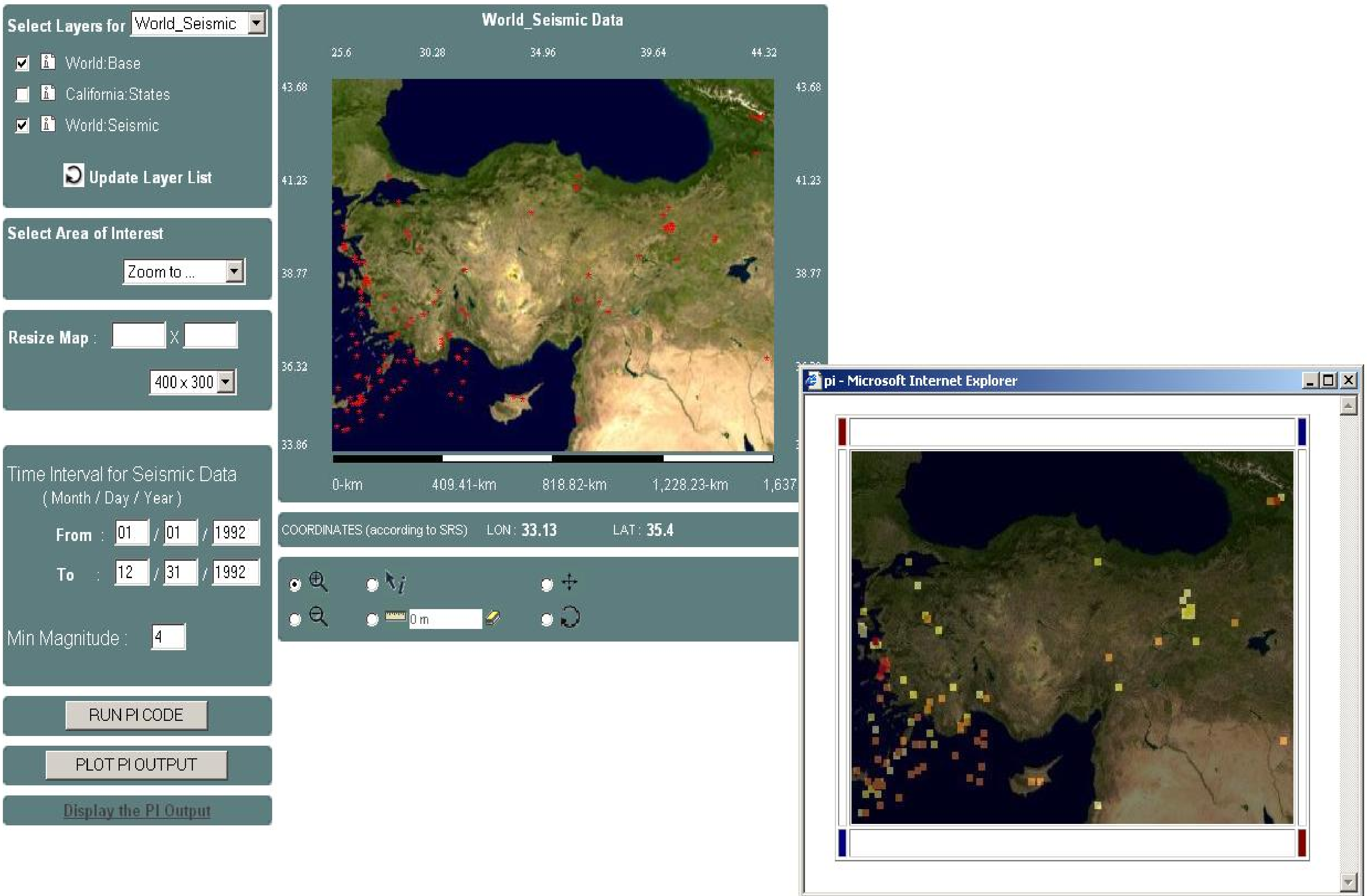
- Galip Aydin: WFS and SensorML
- Harshawardhan Gadgil: HPSearch
- Ahmet Sayar: WMS
- Mehmet Aktas: WS-Context and UDDI

More Screenshots

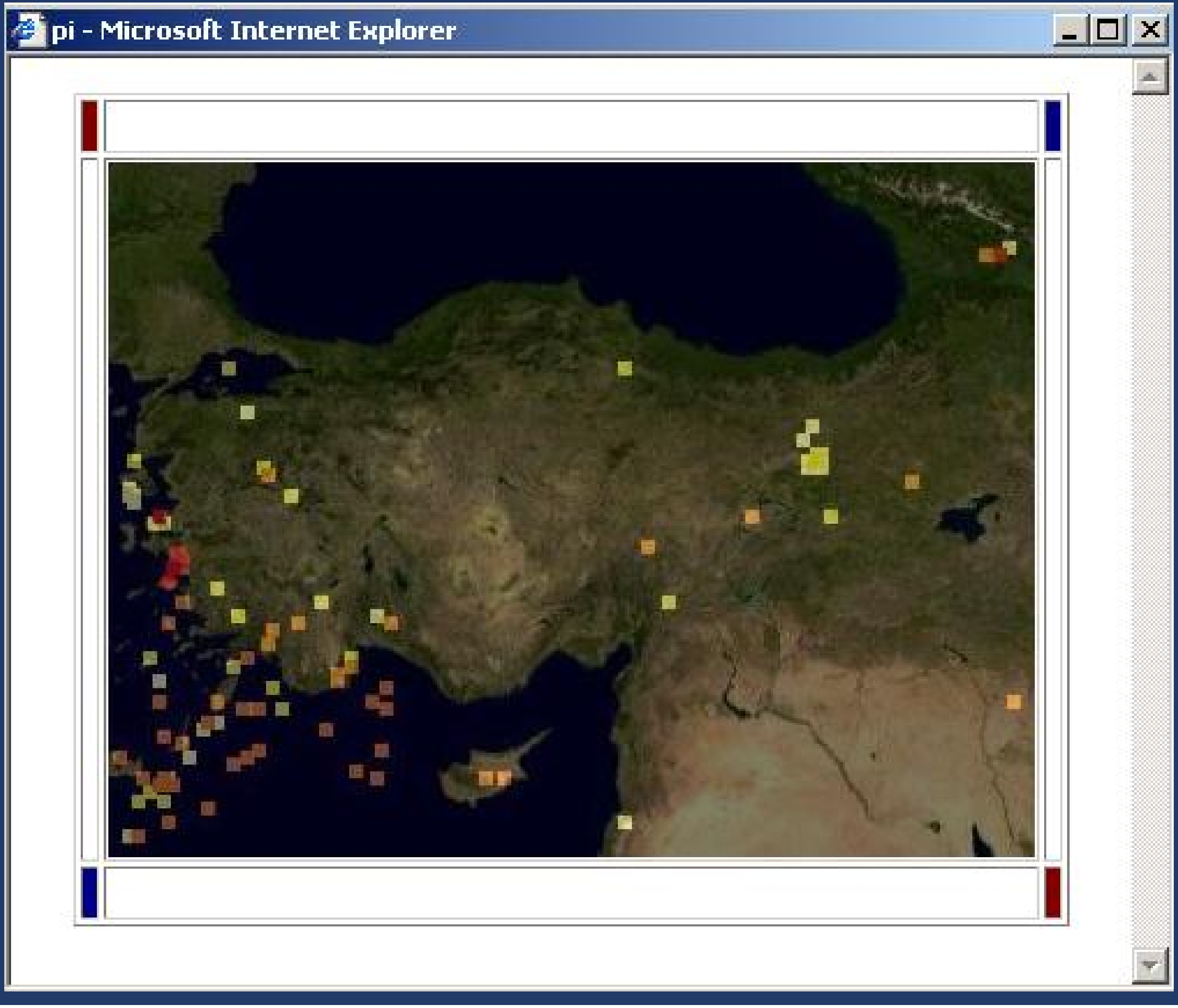
Sample-2 (PI Output Plotting)



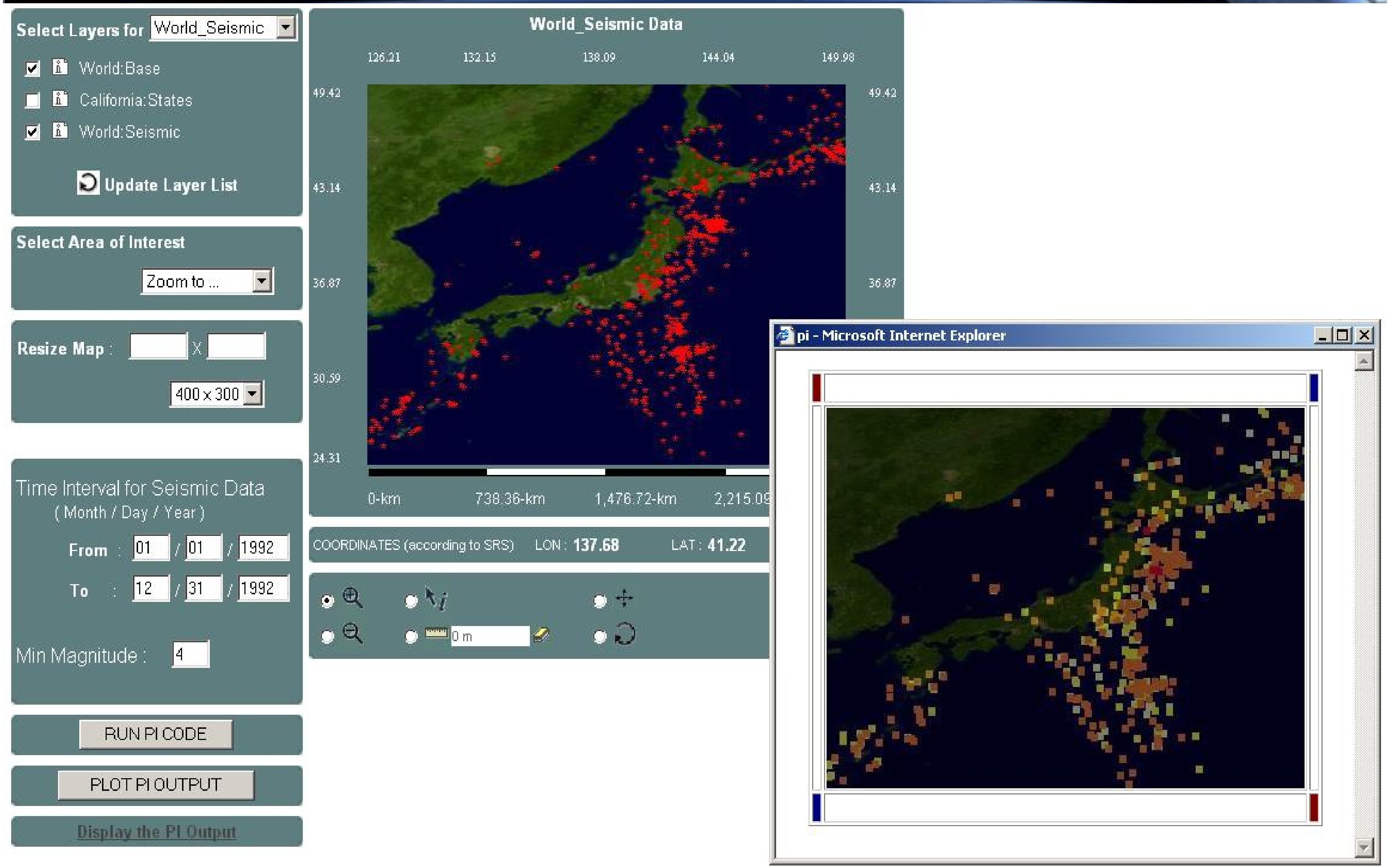
Turkey


Select Layers for
World_Seismic
World:Base
California:States
World:Seismic
Update Layer List
Select Area of Interest
Zoom to ...
Resize Map :
X
400 x 300
Time Interval for Seismic Data
(Month / Day / Year)
From : 01 / 01 / 1992
To : 12 / 31 / 1992
Min Magnitude : 4
RUN PI CODE
PLOT PI OUTPUT
Display the PI Output
World_Seismic Data
25.6
30.28
34.96
39.64
44.32
43.68
41.23
38.77
36.32
33.86
0-km
409.41-km
818.82-km
1,228.23-km
COORDINATES (according to SRS)
LON : 33.13
LAT : 35.4
0 m
pi - Microsoft Internet Explorer

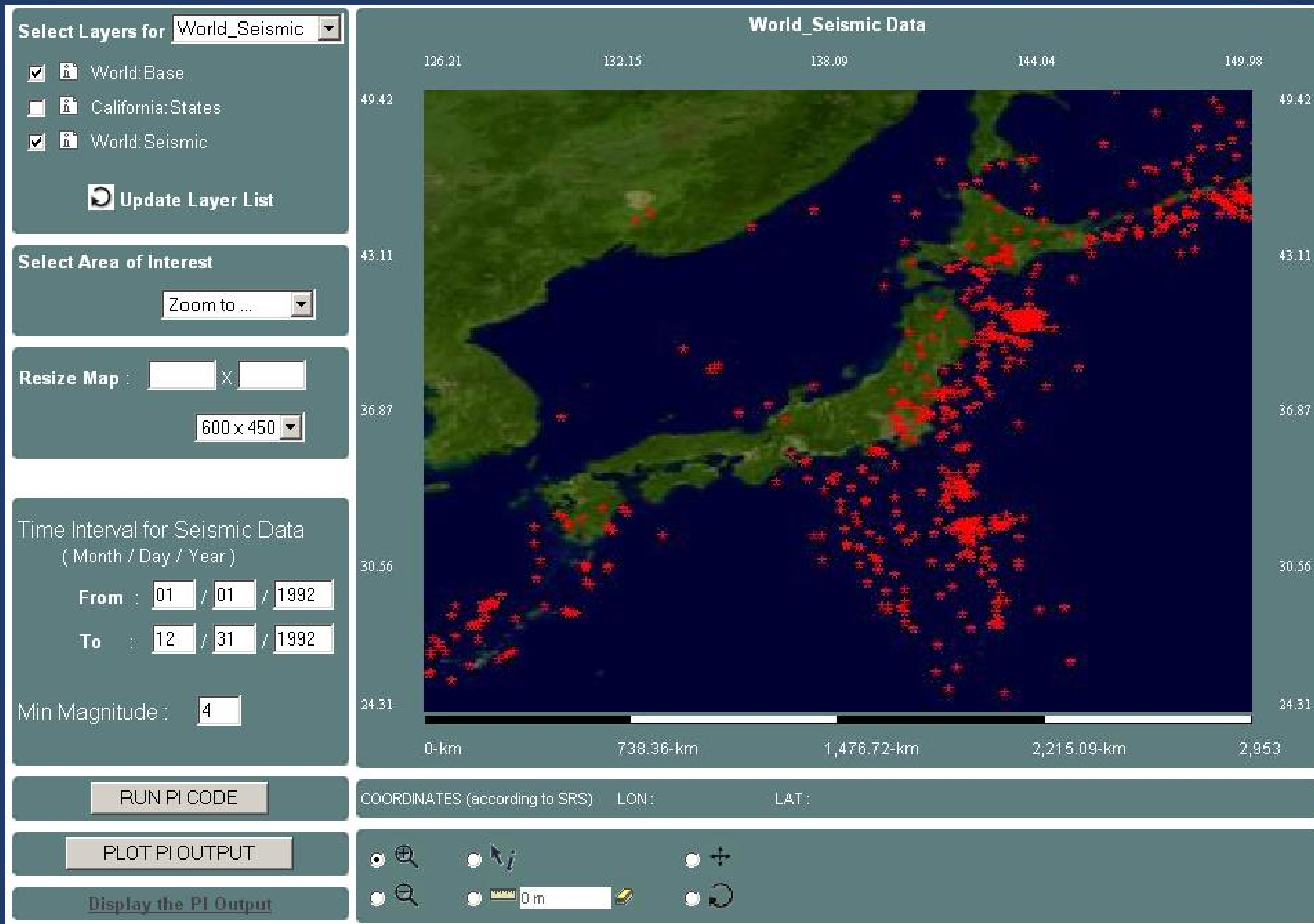
Turkey-2



Japan


Select Layers for
World_Seismic
World:Base
California:States
World:Seismic
Update Layer List
Select Area of Interest
Zoom to ...
Resize Map :
x
400 x 300
Time Interval for Seismic Data
(Month / Day / Year)
From : 01 / 01 / 1992
To : 12 / 31 / 1992
Min Magnitude : 4
RUN PI CODE
PLOT PI OUTPUT
Display the PI Output
World_Seismic Data
126.21
132.15
138.09
144.04
149.98
49.42
43.14
36.87
30.59
24.31
0-km
738.36-km
1,476.72-km
COORDINATES (according to SRS)
LON : 137.68
LAT : 41.22
0 m
pi - Microsoft Internet Explorer

Japan-1


Select Layers for
World_Seismic
World:Base
California:States
World:Seismic
Update Layer List
Select Area of Interest
Zoom to ...
Resize Map :
X
600 x 450
Time Interval for Seismic Data
(Month / Day / Year)
From : 01 / 01 / 1992
To : 12 / 31 / 1992
Min Magnitude : 4
RUN PI CODE
PLOT PI OUTPUT
Display the PI Output
World_Seismic Data
126.21
132.15
138.09
144.04
149.98
49.42
43.11
36.87
30.56
24.31
0-km
738.36-km
1,476.72-km
2,215.09-km
2,953
COORDINATES (according to SRS)
LON :
LAT :
0 m

2005 (Sub)Milestones

Tasks		1Q			2Q			3Q		
		1	2	3	4	5	6	7	8	9
Extracting Unknown Useful Knowledge from Large-scale Datasets	Develop an earthquake-domain standard ontology using Ontronic									
	"Understand" structured dataset semantics									
	Find key taxonomies from the structured datasets									
Semantic Mapping between Ontologies and Relational Databases	Discover the constraints that hold between a concept extracted from the structured datasets and a concept from the standard ontology									
	Match the recognized concepts with ones from a more general specification of related concepts by looking up: the lexical database, WordNet									
	Design an algorithm that discovers the best mapping using: ✓ The top *k* most frequent instances ✓ The advanced concet pattern matching									
	Learn to define the inter-relationship between two concepts via the feedback from domain-expert(s)									
	Merge the resulting structure with other similar knowledge representations									
Ontology-based Federated Database Service on the Web Service Context	Develop the Web Service (WSDL) on Ontronic to receive a query from authorized users									
	Query deformation based on mapping between the global ontology and databases									
	Distributed query execution and information filtering and Generate WSDL result-set									