The Instrumented Oil Field
Towards Dynamic Data-Driven Management of the Ruby Gulch Waste Repository

Manish Parashar
The Applied Software Systems Laboratory
ECE/CAIP, Rutgers University
http://www.caip.rutgers.edu/TASSL
(NSF ITR-DDDAS 04 (UT-CSM, UT-IG, RU, OSU))
The Instrumented Oil Field: The Team

Mary Wheeler
Hector Klie
Clint Dawson
Mrinal Sen
Paul Stoffa

CSM and UTIG
University of Texas at Austin

Manish Parashar
Viraj Bhat
Vincent Matossian

Electrical and Computer Engineering Dept.

Joel Saltz
Umit Catalyurek
Tahsin Kurc

Biomedical Informatics Department

Roelof Versteeg

Supported by:
Dynamic Data/Information-driven Scientific Investigation

- Seamless Interaction and Aggregation
  - Computation, Data, Resources, Experts

- Oil Reservoir Design
  - Applications/Services:
    - Reservoir models, economic models, data analysis services, visualization services, …
  - Data Sources:
    - History data archives, real-time market data, real-time oil field data, …
  - Resources:
    - Instrumented oilfield (sensors/actuators), thin clients, compute servers, data servers, …
  - Experts:
    - Field engineer, Petroleum engineer, Scientist, Mathematician, Economist, etc.
Knowledge-based Data-driven Management of Subsurface Geosystems: The Instrumented Oil Field (ITR/DDDAS)

Detect and track changes in data during production.
Invert data for reservoir properties.
Detect and track reservoir changes.

Assimilate data & reservoir properties into the evolving reservoir model.
Use simulation and optimization to guide future production.
A New Approach: Dynamic, Data Driven Reservoir Management

Dynamic Decision System
- Optimize
  - Economic revenue
  - Environmental hazard
  - …
Based on the present subsurface knowledge and numerical model

- Update knowledge of model
- Improve numerical model

Dynamic Data-Driven Assimilation
- Acquire remote sensing data
- Plan optimal data acquisition
- Improve knowledge of subsurface to reduce uncertainty

Subsurface characterization

Data assimilation

Management decision

Start

Processing Middleware

Grid Data Management

Autonomic Grid Middleware
Vision: Diverse Geosystems – Similar Solutions

Landfills

Models

Simulation

Control

Data

Oilfields

Undersea Reservoirs

Underground Pollution
Dynamic Data Driven Simulation Framework: Models, Methods

• Integrated Parallel Accurate Reservoir Simulation: IPARS
  – *Multiple individual physical models and algorithms for multiphase flow and transport.*
    • Provides linear solvers with state of the art preconditioners.
    • Couplings with geomechanics and chemistry
  – *Multiblock approach (subdomain can treat unstructured grids)*
  – *Multi-physics, multi-numeric, multi-scale capabilities*

• Seismic Data Simulation: FDPSV
  – *Simulation of seismic data gathering*
  – *Simulates sound traces shot from sound sources and captured by receivers*
    • *Can scale up to thousands of sources and receivers*

• Optimization Tools
  – *Very Fast Simulated Annealing (VFSA)*
  – *Finite Difference Stochastic Optimization (FDSA)*
  – *Simultaneous Perturbation Stochastic Optimization (SPSA)*
Dynamic Data Driven Simulation Framework: Data Management

• Data Virtualization: STORM
  – Large data querying capabilities
  – Distributed data virtualization
  – Indexing, data cluster/decluster, parallel data transfer

• Metadata Service: Mobius
  – XML metadata definition
  – XML database creation and federation

• Data Analysis/Processing Workflows: DataCutter
  – Filter-stream based framework for combined task/data parallelism
  – On demand data product generation
Dynamic Data Driven Simulation Framework: Autonomic Middleware Substrate (AutoMate)

- **Grid Computational Engine: Seine/MACE/Armada**
  - Enable scalable, dynamically adaptive parallel applications
  - Enable complex (dynamic) application/multiblock coupling and parallel data redistribution
  - Adaptive, application and system sensitive runtime management

- **Programming system for self-management: Accord**
  - Specify application components/services that can adapt their behavior and interactions/compositions at runtime using high-level rules
  - Runtime engine for efficient, scalable, correct and consistent rule enforcement

- **Content-based middleware service: Meteor/Squid**
  - Content based service discovery and composition
  - Scalable associative messaging and coordination

- **Grid Computational Collaboratory: Discover**
  - Seamless and secure (collaborative) access to and interactions between users, applications, and services
A new generation of integrated and seamless simulations

Optimizing Oil Production on the Grid

Objective function

Autonomic Monitoring Management Control

Data mgmt./assimilation

Visualization

Static data

Dynamic data

Clients

Collaboration
Effective Oil Reservoir Management: Well Placement/Configuration

- Why is it important
  - Better utilization/cost-effectiveness of existing reservoirs
  - Minimizing adverse effects to the environment
Effective Oil Reservoir Management: Well Placement/Configuration

• **What needs to be done**
  – Exploration of possible well placements and configurations for optimized production strategies
  – Understanding field properties and interactions between and across subdomains
  – Tracking and understanding long term changes in field characteristics

• **Challenges**
  – Geologic uncertainty: Key engineering properties unattainable
  – Large search space: Infinitely many production strategies possible
  – Complex physical properties and interactions.
  – Complex numerical models
An Autonomic Well Placement/Configuration Workflow

AutoMate Programming System/Grid Middleware

Generate Guesses
SPSA
VFSA
Exhaustive Search

Send Guesses

Optimization Service

Send guesses

IPARS Factory

Start Parallel IPARS Instances
Instance connects to DISCOVER

MySQL Database
If guess in DB: send response to Clients and get new guess from Optimizer
If guess not in DB instantiate IPARS with guess as parameter

History/ Archived Data

Sensor Data

AutoMate Programming System/Grid Middleware

DISCOVER
Notifies Clients
Clients interact with IPARS

client
client
Pervasive Portals for Collaborative Monitoring and Steering

Production Run for Monitoring and Steering

Autonomic Oil Well Placement/Configuration

Permeability

Pressure contours
3 wells, 2D profile

Contours of $NEval(y,z,500)(10)$

Requires $NYxNZ$ (450) evaluations. Minimum appears here.

VFSA solution: “walk”:
found after 20 (81) evaluations
Solution for 7 different initial guesses

Convergence history
Optimal Well Placement

Comparison of optimization approaches

<table>
<thead>
<tr>
<th>Method</th>
<th>Nelder-Mead</th>
<th>GA</th>
<th>VFSA</th>
<th>FDSA</th>
<th>SPSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best solution</td>
<td>-1.018e8</td>
<td>-1.073e8</td>
<td>-1.083e8</td>
<td>-1.062e8</td>
<td>-1.075e8</td>
</tr>
<tr>
<td>Average number of function evaluations</td>
<td>99.95</td>
<td>104.02</td>
<td>75.5</td>
<td>57.0</td>
<td>37.8</td>
</tr>
</tbody>
</table>

Optimal solution: $F^* = -1.098E8$

Learned lessons:

- Robust stochastic algorithms increases the chances to find (near) optimal solutions (VFSA)
- Several trials of a fast algorithm pay off against sophisticated algorithms (SPSA)
- Need to develop hybrid strategies
Grid-based Optimization

A 3-level parallelism on the Grid:

- Stochastic realization 1
  - Initial guess 1
  - Initial guess 2
  - Initial guess m
  - IPARS 1
  - IPARS 2
  - IPARS m
  - IPARS m+1
  - IPARS m+2

- Stochastic realization 2
  - Initial guess 1
  - Initial guess 2
  - Initial guess m
  - IPARS 1
  - IPARS 2
  - IPARS m
  - IPARS m+1
  - IPARS m+2

- Stochastic realization n
  - Initial guess 1
  - Initial guess 2
  - Initial guess m
  - IPARS 1
  - IPARS 2
  - IPARS m
  - IPARS (n-1)m+1
  - IPARS (n-1)m+2

n*m parallel independent runs of IPARS

• Divided in 3 Operable Units. OU3 is the Ruby Gulch Waste Rock Repository: a valley with about 20 million cubic yard of waste rock. The waste rock generated AMD (acid mine drainage) which impacted drinking water supplies

• Water captured at toe of repository for treatment in water treatment plant. Treatment costs are substantial over repository lifetime based on observed outflows in 1997-1999

• Cost driven solution: cap 70 acre waste rock repository to reduce AMD production
Monitoring system hardware

- Multi electrode resistivity system (523)
  - One data point every 2.4 seconds from any four electrodes
- Temperature & Moisture sensors in four wells
- Flowmeter at bottom of dump
- Weather-station
- Manually sampled chemical/air ports in wells

- Current state: data is automatically collected and transmitted from data acquisition systems to web accessible relational database – data is accessible to user within hours of being collected
  - Approx 40K measurements/day
- Design lifetime: 30 years
Gilt Edge Mine Superfund Site, South Dakota
September 25, 2003

OU3: Ruby Gulch Waste Rock Repository
Gilt Edge Site
Dynamic Data-Driven Waste Management

**Observations**
- Ruby Gulch Waste Repository
- Sensors
- AutoMate
- STORM/Datacutter

**Data Assimilation**
- Surrogate/Reduced model
- IPARS

**Optimization**
- Control algorithms

**Controllable input**
Many Challenges and Opportunities

• Applications and algorithms
  – *model development and calibration*
  – *uncertainty estimation*
  – *parameter selection and optimization*

• Measurement and actuation systems
  – “real-time” *data collection and transport*
    • *in-network aggregation, assimilation*
  – *data selection and application integration, data quality management*
  – *actuation*

• Systems software
  – *programming systems/models for data integration and runtime self-management*
  – *data management mechanisms for real time, space and data quality constraints,*
  – *runtime execution services that guarantee reliable execution with predictable and controllable response time*
First steps …

• Coupled air-water models
  – *Model diurnal/seasonal variations in the outflow measurements observed*

• Wide-area model/simulations coupling
  – *Abstractions, parallel data redistribution, node-to-node data transport*

• In-network data aggregation mechanisms
  – *Evaluated using the Orbit 400 node sensor testbed*

• Runtime data steaming middleware using model-based control/optimization strategies
  – *Minimize impact on simulations, eliminate data loss*

• Reservoir and seismic data archives
  – *30TB of seismic dataset, relatively small volume of oil reservoir data*
Conclusion

• **DDDAS:** Enabling the next generation knowledge-based, data-driven, dynamically adaptive applications on the Grid
  - can enable accurate solutions to complex applications; provide dramatic insights into complex phenomena

• **The Instrumented Oil Field: DDDAS for the management and control of subsurface geosystems**
  - Models, algorithms, numerics – IPARS/Mace/Seine
  - Programming system, middleware – Accord/Meteor/Rudder/Squid
  - Data management, assimilation – Storm/Mobius/DataCutter
  - Collaborative monitoring, interaction, control - Discover

• **Dynamic data-driven waste management**
  - Many challenges and opportunities

• **More Information, publications, software**
  - [www.caip.rutgers.edu/~parashar/](http://www.caip.rutgers.edu/~parashar/)
  - parashar@caip.rutgers.edu