ResGrid: A Grid-Aware Toolkit for Reservoir Uncertainty Analysis


Presented by: Gabrielle Allen
Center for Computation & Technology
& Dept. Computer Science
Louisiana State University

CCGrid 2006

“UCOMS”

- “Ubiquitous Computing & Monitoring System for Discovery & Management of Energy Resources”
- DOE/Louisiana BOR funded
  - Petroleum engineering
  - Wireless sensor networks
  - Grid technologies
- Applications
  - Reservoir simulation
    - Uncertainty analysis, sensitivity studies, history matching
  - Real-time well surveillance
  - Drilling performance analysis with high-rate data
Oil Industry in Louisiana

- Major oil producing state in US:
  - 5th in production
  - 8th in reserves
  - Home to 2 of 4 strategic petroleum reserves
  - 17 petroleum refineries (capacity 2.8M barrels/day)
  - Ports receive ultra large oil tankers
  - 20,000 oil producing wells, around 4K offshore.

Katrina’s energy blow

Hurricane Katrina’s effect will be felt long after her winds have subsided, in the oil and gas fields of the Gulf of Mexico. Nearly a third of the oil and 10% of the natural gas produced in the U.S. originate in the Gulf, and a lengthy disruption would increase prices. Locations of drilling rigs and pipelines show the Gulf’s importance.

Reservoir Studies

- Assessments and predictions of oil/gas reservoir performance, depending on
  - Geological heterogeneity
  - Engineering choices
- Used for development and operational decisions
  … models assess different production scenarios.
- Applications:
  - Well placement & performance prediction
  - Sensitivity analysis & uncertainty assessment
  - History matching (model verification and tuning)
Reservoir Simulation

- Mathematical model for fluid flow in a reservoir involves density, permeability ($K$), mobility, pressure ($P$), production rate ($q$), porosity & saturation, where $m$ denotes either oil, water or gas.

$$\nabla \cdot (\rho_m K \lambda_m \nabla P_m) - q_m = \frac{\partial (\phi \rho_m S_m)}{\partial t}$$

- Many geological parameters cannot be measured or modeled and are unknowns.

- We are using UTChem (3D, multiphase, multicomponent, compositional, variable temperature, FD simulator)

Core Computational Needs

Compute:
- Large-scale computation: Seismic inversion, flow numerical simulations, Ggeostatistics, analysis, …

Data:
- Large data sets (TB-PB): Seismic, Geological & Geophysical (G&G), Well logging, Simulation results, …

Security:
- Commercial benefits lead to high security for all data, proprietary codes, etc.

Workflow:
- Parameter selection, model construction, data movement, model deployment, results analysis etc.

Move towards dynamic, data driven scenarios, including direct input from sensor data.
Reservoir Uncertainty Analysis

- Understand the effect of uncertainty in reservoir studies to guide development and operational decisions
  - Uncertainty in different (geological) parameters (factors)
  - Response surfaces are built from observational and simulation data to model parameter effect
  - Experimental design techniques help reduce the parameter space which needs to be explored.

Factors (parameters) in reservoir studies are classified into
- Controllable: Can be varied by process implementers, e.g. Well Location, injection rate, …
- Observable: Can be relatively accurately measured but not controlled, e.g. Depth to a structure, …
- Uncertain: Cannot be accurately measured or controlled, e.g. Permeability far from wells, …

Factors:
- Geological (pressure, permeability, water saturation, critical gas saturation, gas/water end points, gas/water Corey exponents, non-Darcy coefficient, aquifer size, anisotropy ratio, …)
- Engineering (completion length ration, tubing head pressure, tubing diameter, …).
Reservoir Uncertainty Analysis

- **Responses** to factors are obtained by models or measurement. Reservoir studies concerned with responses affecting value, e.g.
  - Peak oil rate
  - Cumulative oil recovery

- A response surface model is an empirical fit to reservoir simulation results:

\[
\hat{y}_j(x) = \hat{\beta}_{j,1} + \sum_{i=1}^{k} \hat{\beta}_{j,i+1} x_i + \hat{\beta}_{j,i+k+1} x_1 x_2 + \hat{\beta}_{j,k+2} x_1 x_3 \\
+ \ldots + \hat{\beta}_{j,1+k(k+1)/2} x_{k-1} x_k + \sum_{i=1}^{k} \hat{\beta}_{j,k+1+k(k+1)/2} x_i^2
\]

---

Experimental Design

- Techniques for varying factors to minimize number of simulations for calculating response models.
- Factors are scaled to [-1, +1] in such a way to best give a linear response (e.g. log scaling)
- A design is a set of factor combinations used to construct system response over a range of factors.

A: 2-level factorial design
   \(2^k\) experiments
B: 3-level factorial design
   \(3^k\) experiments
C: Box-Behnken design
   optimized

---

CENTER FOR COMPUTATION & TECHNOLOGY AT LOUISIANA STATE UNIVERSITY
General Workflow

- Start
  - Seismic Inversion (Sparse spike, correlation wavelet, probabilistic inversion)
  - “Massage” the results (models) into reservoir simulation model
  - Large-scale Reservoir Simulations (Flow Numerical Simulation)
  - Post processing: Response models, Model discrimination, Sensitivity computing, Posterior factor distributions
  - End

Current Workflow

- Example case: Eleven geological factors e.g. initial pressure + three engineering factors e.g. tubing diameter with either 3 or 4 levels.
- Factorial design:
  - $4^6 \times 3^8 = 26,873,856$ reservoir simulations
  - 100 days on 1024 proc cluster (at 6 mins per run)
- Even with experimental design many runs needed
Motivation for this Work

• Reservoir simulation is one of the largest users of computing power
  – Large, complex, uncertain models
  – High risks and rewards
• Can be performed more efficiently if
  – Moderate-sized jobs can be farmed out onto a heterogeneous, underutilized grid
  – Large jobs can be run in parallel on a grid
• Efficiency gains can be used to assess risks, estimate parameters, run larger and more complex models, and optimize developments

Methodology

Design and build a grid enabled toolkit for
  – Experimental Design Framework:
    • Select relevant models, records factor settings, controls execution, creates response models.
  – Post processing, analysis and visualization
    • Including RSMCB (Response surface models, Monte Carlo Simulation, and Bayesian techniques)

This toolkit is called **ResGrid**
Components

- Build on integrating existing middleware
- Frameworks
  - Grid Application Toolkit (GAT)
  - GridSphere Portal Framework
  - Grid Portlets, GridPort, GAT adaptors
- Core underlying middleware
  - Globus GRAM, GridFTP, GSI, MDS
  - Condor-G
- Application level
  - UTChem, HDF5, Amira/OpenDX
- Synergies with other application projects at CCT

ResGrid Usage Scenario

- Five key components:
  - Portal
  - Reservoir modeling
  - Large scale distributed reservoir simulation
  - Analysis
  - Visualization
Programming Grid Applications

- Question: Why are there so few grid applications used routinely?
- Answer:
  - Lack of a simple, stable and uniform high-level programming interface that integrates the most common grid programming abstractions
  - Need to hide underlying complexities, heterogeneities, and changes from applications

Grid Application Toolkit (GAT)

- Abstract programming interface between applications and Grid services
- Designed for applications (move files, run remote task, migrate, write to remote file)
- Led to GGF Simple API for Grid Applications (SAGA)

<table>
<thead>
<tr>
<th>Default Adaptors</th>
<th>Basic functionality, will work on single isolated machine (e.g. cp, fork/exec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Globus Adaptors</td>
<td>Core Globus functionality: GRAM, MDS, GT-RLS, GridFTP</td>
</tr>
<tr>
<td>GridLab Adaptors</td>
<td>GRMS, Mercury, Delphi, iGrid</td>
</tr>
<tr>
<td>Under Develop</td>
<td>Scp, DRMMA, Condor, SGE, SRB, Curl, RFT.</td>
</tr>
</tbody>
</table>
Grid Application Toolkit

• Standard API and Toolkit for developing portable Grid applications independently of the underlying Grid infrastructure and available services
• Implements the GAT-API
  – Used by applications (different languages)
• GAT Adaptors
  – Connect to capabilities/services
  – Implement well defined CPI (mirrors GAT-API)
  – Interchangeable adaptors can be loaded/switched at runtime
• GAT Engine
  – Provides the function bindings for the GAT-API

GAT API Subsystems

<table>
<thead>
<tr>
<th>File Subsystem</th>
<th>Monitoring and Event Subsystem</th>
<th>Information Exchange Subsystem</th>
<th>Resource Management Subsystem</th>
<th>Utility Subsystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATFile</td>
<td>GATRequestListener</td>
<td>GATAdvertService</td>
<td>GATSoftwareDescription</td>
<td>GATSelf</td>
</tr>
<tr>
<td>GATEndpoint</td>
<td>GATRequestNotifier</td>
<td></td>
<td>GATJobDescription</td>
<td>GATContext</td>
</tr>
<tr>
<td></td>
<td>GATPipeListener</td>
<td></td>
<td>GATJobDescription</td>
<td>GATStatus</td>
</tr>
<tr>
<td></td>
<td>GATPipe</td>
<td></td>
<td>GATResource</td>
<td>GATPreferences</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GATResourceBroker</td>
<td>URL, Time, ...</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GATReservation</td>
<td></td>
</tr>
</tbody>
</table>
GAT++ = SAGA

- Global Grid Forum Research and Working Groups to develop a “Simple API for Grid Applications” (growing community support)
  - Focus on scientific and engineering applications
  - Focus on simplicity
- Similar to GAT
  - Better thought out API (based on use cases)
  - Asynchronous calls, bulk operations, QoS
  - More functional areas (e.g. streaming)
- Implementations under development now.

SAGA (GAT) & OGSA?

- Complementary and not competitive approaches
  - OGSA defines interfaces at the service and middleware level, SAGA at the application level
  - OGSA is primarily an architecture, SAGA is an API
  - SAGA implementations interface OGSA-compliant services (but not limited to)
  - OGSA is for middleware developers, SAGA is for application developers
Example: Data Replication Tool: getdata

- Light weight client to locate and replicate data files

---

Case Study

- The objective is to use experimental design to examine flow behavior of geostatistical models.
- Task farming is ideally suited because the independent runs from the design are easy to distribute.
- A suite of simulation models investigates differences between flow responses of alternative geostatistical formulations:
  - sequential Gaussian
  - LU-decomposition
  - a hybrid of the two.
- A range of geostatistical parameters was evaluated.
A single bed of sandstone, which is 20,000 x 20,000 feet.

More than 585 permeability measurements were made.

The semivariogram has a very-defined range and sill. The semivariogram is used to compute the conditional 2D simulations.

\[ \gamma = 0.09 + 0.91 \exp(-h/a) \]

2D Simulation Summary

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Simulation Runs</th>
<th>Data Size (Gigabytes)</th>
<th>Execution Time (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Grid</td>
</tr>
<tr>
<td>LU Decomposition Simulation</td>
<td>1280</td>
<td>25.6</td>
<td>120</td>
</tr>
<tr>
<td>Sequential Gaussian Simulation</td>
<td>1280</td>
<td>25.6</td>
<td>120</td>
</tr>
<tr>
<td>Spectrum Simulation</td>
<td>1280</td>
<td>25.6</td>
<td>120</td>
</tr>
<tr>
<td>Hybrid Simulation</td>
<td>1280</td>
<td>25.6</td>
<td>120</td>
</tr>
<tr>
<td>Total</td>
<td>5120</td>
<td>102.4</td>
<td>480</td>
</tr>
</tbody>
</table>
The Response Surface Model (RSM) of LUSIM is flatter than the other two models. Furthermore, according to F-test and t-test for the RSM, the SGSIM may overestimate stochastic fluctuations and is a “biased” algorithm to forecast the petrophysical attribution of reservoir.
Case Study: Conclusions

- The Sequential Gaussian Simulation is fast and widely used by reservoir engineers but may significantly overestimate stochastic fluctuations, compared with the more rigorous LU-decomposition methods.
- Mis-evaluation of variance may have heretofore unappreciated impacts on uncertainty assessment methods.

SURAggrid

- SuraGrid is a beyond-regional initiative in support of regional strategy.
  - Grid Infrastructure connecting states in South-Eastern US.
  - Conceived to enable seamless sharing of resources between various institutions to further scientific research and Grid development.
- Partnering with IBM and other vendors to procure HPC Resources across various SURA sites to build a strong computational resource backbone.
- Many sites connected via high speed optical networks like NLR and LONI.
- http://www.sura.org/suragrid
Future Plans

- Complete the ResGrid toolkit
  - Portal interface (GridSphere)
  - Visualization
- Larger computation facilities and Grid environment, including SURAGrid, LONI, and Tera-Grid
  - Assess use of GAT/SAGA
  - More advanced scheduling decisions (compute/data)
- Extend the functionality of the uncertainty analysis scenario
  - More uncertainty factors
  - Automatic adaption of factor range and values
  - Automated experimental design
- Extend toolkit to more scenarios for petroleum engineers and geoscientists to run large suites of models in a stable environment, e.g. history matching.

More Information

- UCOMS Project
  - [http://www.ucoms.org](http://www.ucoms.org)
  - [http://www.cct.lsu.edu/projects/ucoms](http://www.cct.lsu.edu/projects/ucoms)
- Grid Application Toolkit
  - [http://www.gridlab.org/GAT](http://www.gridlab.org/GAT)
- GGF SAGA
  - [https://forge.gridforum.org/projects/saga-rg/](https://forge.gridforum.org/projects/saga-rg/)