Curated Power Grid Data Center

Mark Rice
Pacific Northwest National Lab
mark.rice@pnnl.gov
June 7-8, 2016
Washington, DC
Motivation

• No infrastructure to easily share utility power system information with research and education community to enable future power systems research.
• No automated process providing tracking of who has what types of data.
• Time wasted by researchers to repeatedly ask for data and to perform common tasks – such as removing erroneous readings from a sensor stream.
• Non-automated data access and data cleaning increases recurring cost and manual staff labor.
• Operational burdens on utility staff related to data access.
Project Objective

• Power grid data repository (PGDR) will provide researchers and developers with access to **fully-curated** data sets with **known data quality** handled in an automated process using well-tested and **well-documented** algorithms.

• The PGDR will provide a flexible web portal and workbench for users:
  – Explore the data
  – Create suitable subsets
  – Download data to local system
GOSS Summary

• GOSS is an open-source middleware architecture designed as a prototype future data analytics and integration platform

• Benefits
  – **Supports heterogeneity** – ease of integration with new/existing power grid applications developed in different languages
  – **Data source abstraction** – separates data sources from applications and provides a unified application programming interface (API) for access
  – **Rapid development** – Quickly make new data/events available to other applications integrated with GOSS
  – **Real-time** – subscription to streaming data and events
  – **Reliability** – provides redundant data access for improved reliability
  – **Security** – role and data based access control
  – **Scalability & Performance**
GOSS Architecture

**Application Layer**
- Web portal

**Communication Layer**
- Either synchronous or asynchronous

**Middleware**
- Security; Data routing, Data transformation; Data interface

**Data Source Layer**
- Heterogeneous data
Accomplishments

• Drafted a report for system design requirements and specification for full scale system
• Reviewed initial draft of requirement and specifications report with BPA
• Developed a bench-scale version of the system leveraging GridOPTICS Software System (GOSS) testing key features of the PGDR
• Established business practices for data access
• Completed the final version of report after incorporating comments from BPA
PGDR Implemented Features

- Data Management
- Data Upload
- Meta Data Capture
- Data Quality
- Application & Tools Integration
- HPC Integration
- Web Portal

- Data Format Transformation
- Data Download
- Multiple Format Support
- Data Publication
- User Accounts & Security
- Performance and Scalability
PGDR Implemented Features

- Data Management
- Data Upload
- Meta Data Capture
- Data Quality
- Application & Tools Integration
- HPC Integration
- Web Portal

- Data Format Transformation
- Data Download
- Multiple Format Support
- Data Publication
- User Accounts & Security
- Performance and Scalability

* Features added by CERTS funding
PGDR Prototype Implementation

• Developed bench scale prototype system.
• Implemented a subset of the requirements using GOSS middleware platform.
• PGDR report includes listing of requirements against GOSS capabilities explaining how GOSS fulfills those requirement.
Web Portal

- Developed a web based user interface for researchers to use data repository capabilities.
- Allows easy data upload, download and tracking.
- Built on top of GOSS using web services.

- User can upload a power grid model file and tag metadata.
- User can edit model – changes are tracked and displayed.
- User can view/download the data associated with a model.
## Deliverables

<table>
<thead>
<tr>
<th>Type</th>
<th>Title</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>Confirm Utility Partnership</td>
<td>11/20/2015</td>
</tr>
<tr>
<td>External</td>
<td>Bench scale system to prototype PGDR</td>
<td>4/15/2016</td>
</tr>
<tr>
<td>External</td>
<td>Final Report</td>
<td>7/31/2016</td>
</tr>
</tbody>
</table>
Planned Activities and Schedule

• Possible Future Tasks:
  – Move from a bench scale system to full scale data repository.
  – Implement new capabilities in PGDR:
    • Provenance
    • Converting aggregated data in to raw
    • Ensemble of power flow cases
• Follow on work in GRID DATA ARPA-E program
  • Synthetic scenario creation (SDET)
  • Model data repository (DR POWER)