CVRG: Making Cardiovascular Collaborations Easier in the Cloud

Stephen J. Granite, MS, MBA
CVRG Program Manager
NHLBI R24-HL085343
http://www.cvrgrid.org
The CVRG Project

Mission:
Serve the needs of basic and clinical cardiovascular researchers by providing seamless, customized, secure access to study datasets and analysis tools

Team:
Institute for Computational Medicine, Johns Hopkins University (Winslow, Miller)
Image Lab, Wake Forest University (Carr, Ge)
Center for Comprehensive Informatics, Emory University (Saltz, Post)
Computation Institute, University of Chicago (Foster, Madduri)
Relevant Biomedical Projects

- Multi-Ethnic Study of Atherosclerosis (MESA)
- Coronary Artery Disease Risk In Young Adults (CARDIA)
- Cardiac Translational Implementation Project (CTRIP)
- Hypertrophic CardioMyopathy Consortium (HCMC)
- Jackson Heart Study (JHS)
- Pediatric Heart Network (PHN)
Why Software as a Service (SaaS)?

What we have learned from our users…

• NIH study design often leads to data silos
• Accessing data is hard, not interactive
• Sites themselves have minimal to no IT Support
• Solution? – Access tools/data via the browser

Commercial SaaS Examples

Research SaaS Examples
The CVRG Approach to Technology Development

• Don’t re-invent the wheel
  – When possible, re-use or extend software components from National Center for Biomedical Computing (NCBC) & other open-source projects (e.g., i2b2, Globus Online, CILogon, etc.)
  – Form collaborations to develop new software components (e.g., contributions back to CILogon)

• Address the special needs of the CV community
  – Time-evolving imagery, ECG and EP data
  – Specific data analysis workflows
  – Multi-scale data integration

• Let the CV community drive technology development
CVRG Web Interfaces: Initial Approach

• Alpha CVRG Portal leveraged caGrid Portal and Web Single Sign-On (WebSSO)
  – caGrid Portal built upon Liferay 4.1 open-source Portal Server
  – caGrid WebSSO used Java Architectures Special Interest Group Central Authentication Service (JA-SIG CAS)

• Portlets implemented:
  – Beta Signup
  – ECG Storage/Analysis Workflow
  – Imaging Storage & Analysis
  – Heart shape and motion analysis
Initial Approach Lessons Learned

• Portal (e.g., Liferay 4.1)
  – All the work is done on the server
  – Changing the tabs requires a web browser refresh
    • Resets the portlet displayed so data entry starts all over again

• New requirements appeared during development
  – Need to dynamically display and annotate an ECG
  – Dynamic refresh of portlets themselves difficult
    • Ability to display an embedded scrolling graph next to impossible
  – Call to tools outside of a portlet required extensions to Liferay itself
    • Portal environment customization should be minimized
Web 2.0 Investigation/Revised Approach

- **Web 2.0 interfaces/tools**
  - Work can be divided between the server & web browser
  - Components can be placed in web pages or executed standalone, making changes dynamic

- **Yahoo User Interface (YUI)**
  - Open-source JavaScript library
  - Used by i2b2 in Web Client

- **Google Web Toolkit & API Library (GWT)**
  - Open-source Java project with plugin for Eclipse
    - Allows Java developers to produce rich JavaScript tools without knowledge of syntax
  - Used by LabKey & WebProtege
Revised Approach Lessons Learned

• Web 2.0 applications are more dynamic
  – Flexibility in deployment require development of authorization/authentication components as well

• Portal/Web 2.0 combination maximizes strengths/minimizes weaknesses
  – Portal serves as an interface access point
    • Allows addition/removal of interfaces without a portal reboot
  – Interfaces can be run on servers separate from the portal
    • Appear centralized when they actually are distributed
Tools for Cardiovascular Research

XNAT-CVI
ECGrid Toolkit
OpenClinica
Galaxy (AEC2)

Web Data Services (JHU & Amazon EC2)

CVRG Portal
Web-Browser Software as a Service

Web Analysis Services (AEC2 & JHU)

Ontology & Annotation (RESTful Web Service)

CILogon

National Heart, Lung and Blood Institute
The CVRG Portal

Customized Tool Set

Tool Interfaces Appear in Window
Liferay Authentication Extension/Hooks

- Liferay 6 has improved greatly from Liferay 4
  - JA-SIG CAS and Siteminder come standard with Liferay 6
    - Easier to integrate Portal into intranet environments
  - Liferay 6 extensions can be added, w/o changing portal code
    - However only one extension plug-in can be used at a time
    - All extensions must be methods in one plug-in
- CVRG has added Shibboleth and CILogon extensions
  - Shibboleth works for JHU authentication
  - CILogon has been integrated into the Portal
    - Portal accounts with matching e-mail addresses must exist for full login
    - Portal auto-account creation (new users enter additional information, for portal auditing/usage reporting)
Research Image Management

- Possible savings of buy-in costs for local institutional PACS
- Quick to deploy, ~20 minutes to configure CTP client at site
- No need for local expertise with XNAT

UCLA → JHU XNAT-CVI → Wake Forest

~5-10TB / project

~100 GB echo / week
Automated Quality Control

Auto QC check for Gad scans in No Gad cohort
2 subjects with Gad protocol (but No Gad)
CVRG Workflows & Galaxy

- Broad range of genetic analysis tools
- Graphical workflow interface, for creating/sharing analyses
- Added Globus Online, Bioconductor packages & instantiated on Amazon EC2
**CVRG ECGrid Toolkit**

- ECG-based risk screening
- Batch processing, in-browser visualization & annotation
- Data analysis done via web services, allowing asynchronous on a remote computational resource

- Algorithm screens for risk
- Clinicians annotate arrhythmias (Bioportal)
- Follow up MRI identifies structural lesion
We Are Growing

New Projects & Proposals

• NIRS Blood Flow Regulation (NHLBI)
• NIRS Blood Flow Regulation Comparative Effectiveness Study (NHLBI)
• Hypertrophic Cardiomyopathy Specimen and Clinical Data Research Resource (NHLBI)
• Cloud Computing Cardiac Radiomics (NHLBI) Biomarkers, Blood Flow Auto-Regulation and Brain Injury in the Preterm Neonate (NICHD*)
We Need to Scale Up

*Scale up - commercial vendors, private clouds, but...*

- Data security is paramount
- Potential complexity of negotiating a "Business Associate Contract" satisfying legal concerns of each project institution
- Need to satisfy concerns of project steering committees
- No “standard analyses”, difficult to estimate project computational and data access costs for our users
## Acknowledgements

### Johns Hopkins University
- Michael Miller
- Raimond Winslow
- Laurent Younes
- Siamak Ardekani
- Stephen Granite
- David Hopkins
- Kyle Reynolds
- Michael Shipway
- Christian Jurado
- Will Hayes
- Brandon Benitez

### Emory University
- Joel Saltz
- Andrew Post
- Tahsin Kurc
- Richard Willard
- Himanshu Rathod

### Wake Forest University
- J. Jeffrey Carr
- Yaorong Ge
- Ricardo Peral

### NHLBI
- Jennifer Larkin

NHLBI R24 HL085343

### University of Chicago
- Ian Foster
- Ravi Madduri
- Liu Bo
Questions?