National Cancer Institute

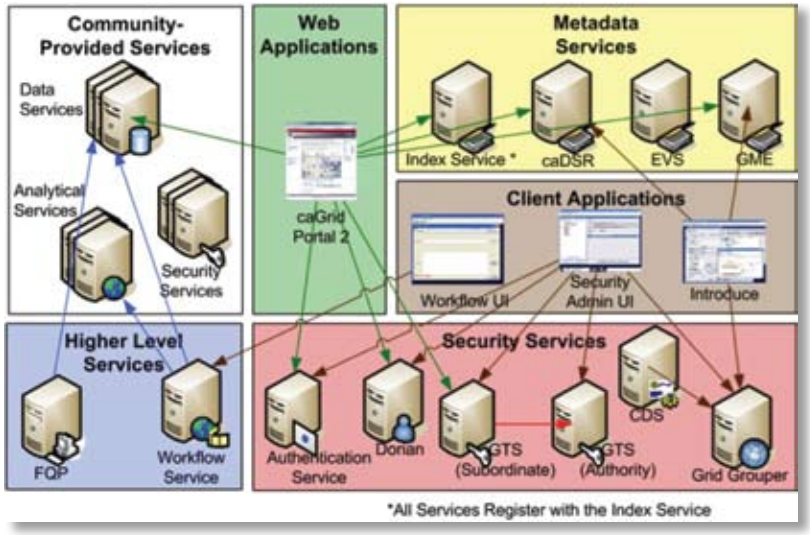
caGrid

caBIG®-interoperable systems architecture

Interoperability has become an essential driver of biomedical research: interoperable IT systems provide researchers with access to data, tools, and computational capacity in order to power information exchange and collaboration, and to speed the pace of research. caGrid is the underlying information technology grid infrastructure developed by the National Cancer Institute to enable connectivity among interested organizations, allowing researchers to tap into the growing collection of cancer research and outcomes data while supporting individual and collaborative investigations. caGrid helps researchers securely share information and analytic resources using locally managed access control policies and by using strongly typed data objects in XML format.

Features

- Nationally deployed, standards-compliant, data and analysis grid that any caBIG®-interoperable system can plug into or draw from
- Semantically discoverable and described services
- Rich runtime metadata infrastructure
- Uniform data query and retrieval across systems
- Federated security and identity management to support controlled access to systems
- Business Process Execution Language (BPEL) workflow engine for data analysis pipeline construction and execution
- Support for additional workflow tools, including the Taverna 2.0 workbench
- Federated, cross-domain data mining and integration
- Standardized programming interfaces for application developers
- Graphical service development toolkit



caGrid structure

Categories of Use:

- ☐ Biospecimens
- ☐ Clinical Trials Management
- ☐ Data Analysis & Statistical Tools
- ■ **Data Sharing**
- ☐ Genome Annotation
- ■ **Infrastructure**
- ☐ Imaging
- ☐ Microarrays
- ☐ Pathways
- ☐ Proteomics
- ☐ Translational Research
- ☐ Vocabularies

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

caGrid

caBIG®-interoperable systems architecture

caBIG cancer Biomedical Informatics Grid®

Architecture Overview

- **Application type:** Grid infrastructure
- **System requirements:** Java installation and network connectivity

Installation and Administration

- **caGrid Security Administration:** Experience in IT security, basic understanding of digital certificate concepts, and group-based authorization concepts
- **Infrastructure:** Desktop/Server, network
- **Long-term administration:** Network/system administration, if running services

Key Contributors:

- NCI Center for Biomedical Informatics and Information Technology (NCI CBIIT)
- The Ohio State University
- Emory University
- Argonne National Laboratory

Resources

Tool Overview Page	https://cabig.nci.nih.gov/sharable/cagrid_overview
Primary Workspace	https://cabig.nci.nih.gov/workspaces/Architecture/caGrid/
caGrid Knowledge Center	https://cabig-kc.nci.nih.gov/CaGrid/KC/
caBIG® Tool Inventory	https://cabig.nci.nih.gov/inventory
caBIG® Support Service Providers	https://cabig.nci.nih.gov/esn/service_providers
NCI Center for Bioinformatics Applications Support	ncicb@pop.nci.nih.gov

NIH Publication No. 10-7641
Printed August 2010