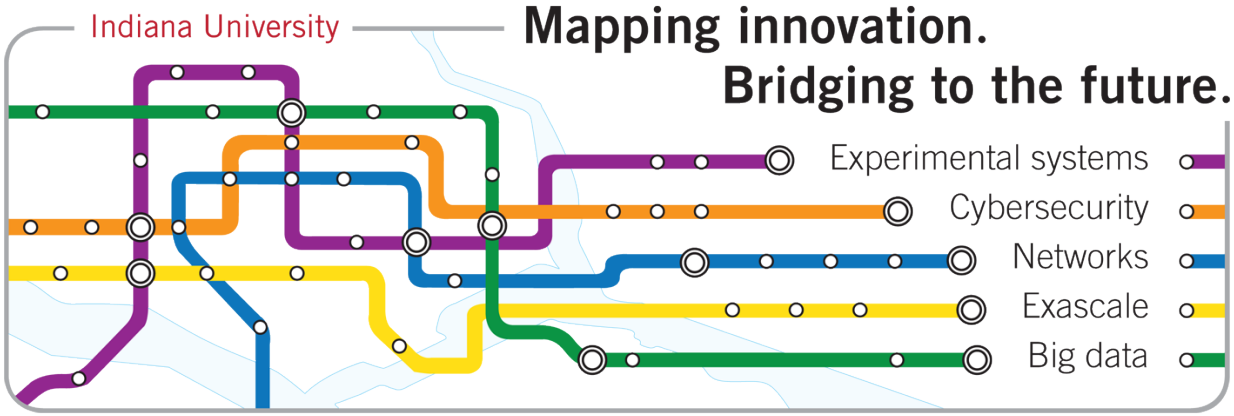
Open community software: Building science gateways and workflows

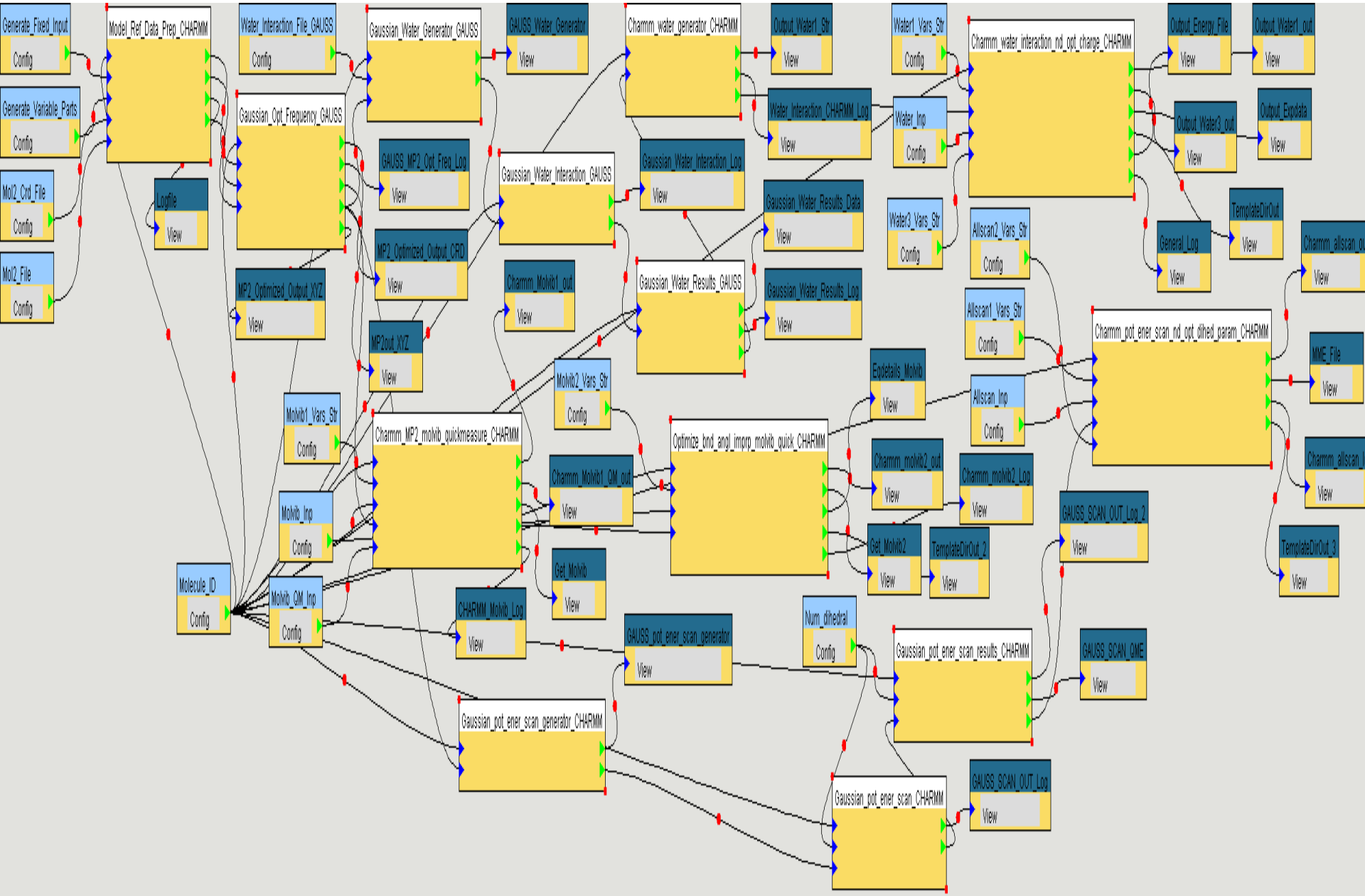
Marlon Pierce, Suresh Marru
Science Gateway Group
Research Technologies, UITS
November 16, 2012



Science Gateway Challenges

- Science Gateways are user environments for interacting with computing resources.
- Gateways can be built in many ways
    - Can run on desktops.
    - Can run in Web browsers.
    - Can use every Web technology known to humanity.
- Science Gateways need to do many things.
    - Run jobs on supercomputers.
    - Add value to online data collections.
    - Support collaborations
- Many gateways are powered by scientific workflows.
- **Service oriented software allows us to work with many different gateways.**

ParamChem Workflow in Airavata



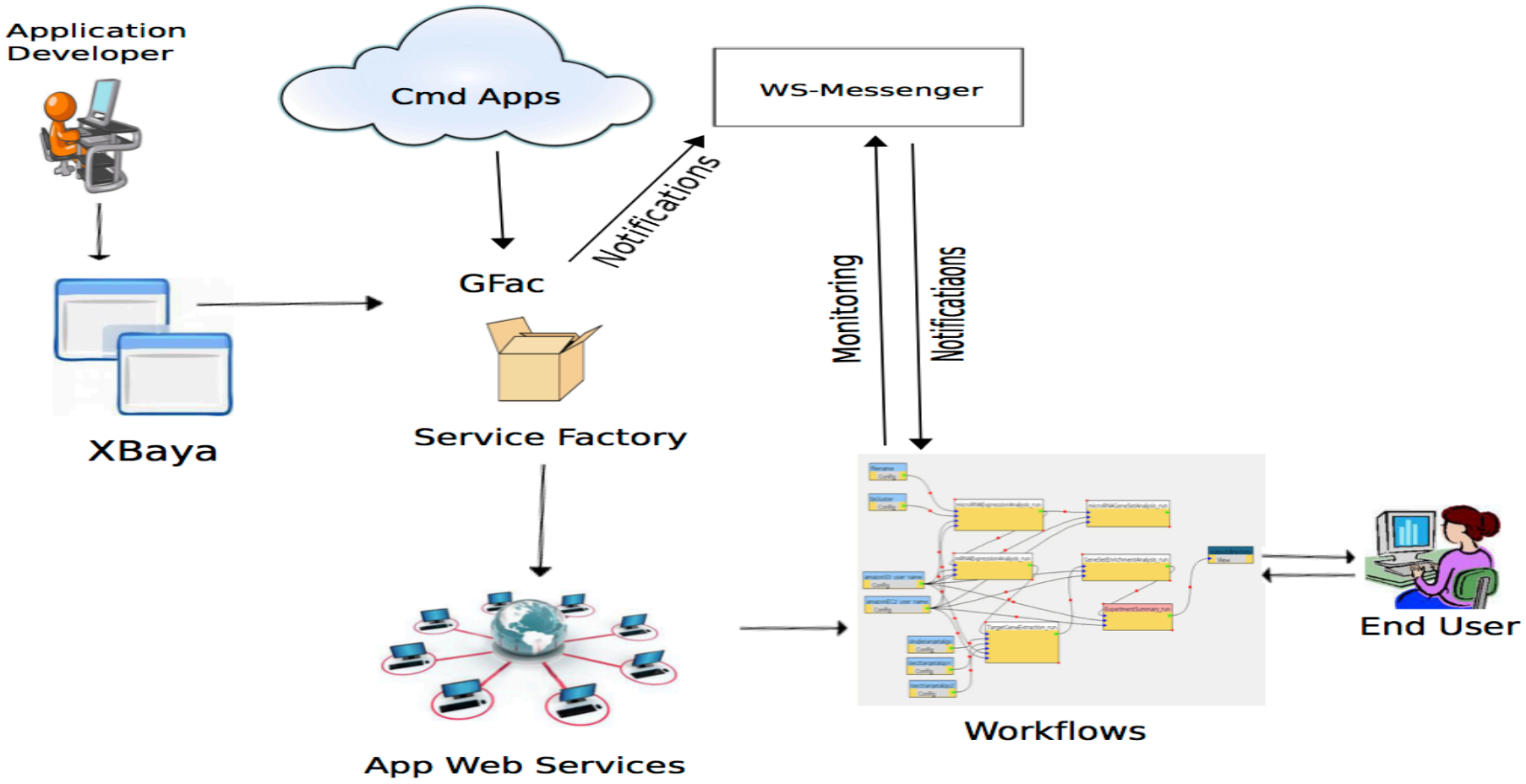
Apache Airavata

- Science Gateway software framework to:
  - Compose, manage, execute, and monitor computational workflows
  - Wrap legacy command line scientific applications with Web services.
  - Run jobs on computational resources ranging from local resources to computational grids and clouds

Apache Airavata Components

Component	Description
XBaya	Workflow graphical composition tool.
Registry Service	Insert and access application, host machine, workflow, and provenance data.
Workflow Interpreter Service	Execute the workflow on one or more resources.
Application Factory Service (GFAC)	Manages the execution and management of an application in a workflow
Airavata API	Single wrapping client to provide higher level programming interfaces.

Apache Airavata High Level Overview



Can I Contribute to Apache Airavata?

- Absolutely…
- Join the mailing list and find out what needs to be done: dev@airavata.apache.org.
- Check out the Airavata Jira
- Learn the Apache Way.
- Get voted into the project as a committer and PMC member.

Apache Rave Overview

Rave is an Apache Top Level Project for building a Web portal on the Open Social and W3C Widget specifications.

- Initially, joint effort of Mitre, Hippo Software, SURFnet, and the OGCE project
- Several new members added to PMC

Goal 1: Provide a useable, packaged, downloadable OpenSocial portal.

- Get started with minimal hassle.

Goal 2: Provide a platform for non-invasive developer extensions, customizations

- Science gateways, for example

Rave Building Blocks

Rave is implemented in JavaScript, Java with Spring MVC

- Bean initialization specified in XML configuration files.
- Inversion of Control makes it easy to swap out implementations.
- Disciplined MVC through Java annotations

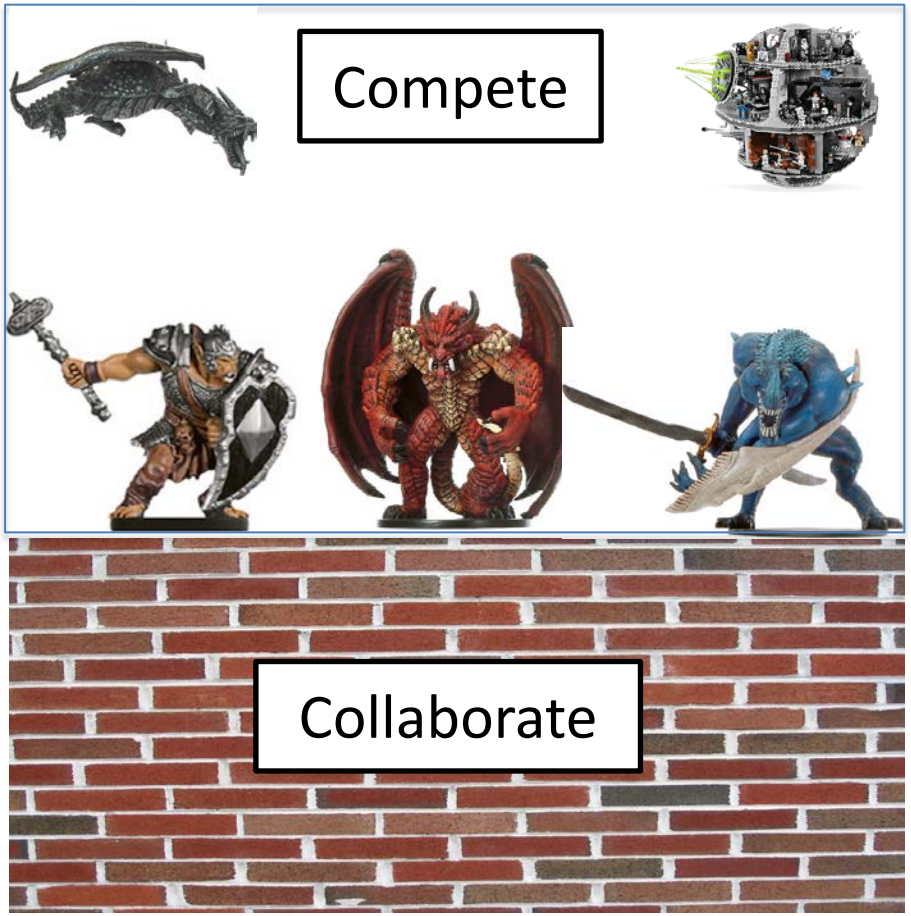
Builds on Apache Shindig and Wookie

- Provide layout management, user management, administration tools, production backend data systems, etc.

Open Source Software, Open Communities

Open Community Software and Governance

- Open source projects need governance.
- Incentives for projects to diversify their developer base.
- Govern how
  - Software is released
  - Contributions are handled.
  - Credit is shared.
- Our approach: Apache Software Foundation



Can I Get Some Help Building Gateways?

XSEDE Extended Collaborative Support Services

HPC SYSTEMS | VIS SYSTEMS | STORAGE SYSTEMS | HTC SYSTEMS | SERVICES

RESOURCE NAME	SITE	MANUFACTURER / PLATFORM	MACHINE TYPE	PEAK TERAFLOPS	DISK SIZE (TB)	LINKS	AVAILABILITY
Wispy	Purdue U	HP DL140g3	Cluster	0.0	0.0		Production through 2013-07-31
Gordon ION	SDSC	Appro	Cluster	0.0	4000.0	User Guide	Production through 2015-03-01
Ranger	TACC	Sun Constellation System	Cluster	579.3	1730.0	User Guide	Production through 2013-02-04
Kraken-XT5	NICS	Cray XT5	MPP	1174.0	2400.0	User Guide	Production through 2014-04-01
Lonestar4	TACC	Dell PowerEdge Westmere Linux Cluster	Cluster	302.0	1000.0	User Guide	Production through 2014-01-31
Keeneland-KIDS	Georgia Tech	HP and NVIDIA	Cluster	0.0	0.0	User Guide	Production through 2014-08-31
Steele	Purdue U	Dell 1950 Cluster	Cluster	66.59	130.0	User Guide	Production through 2013-07-31
Gordon Compute Cluster	SDSC	Appro	Cluster	341.0	4000.0	User Guide	Production through 2015-03-01
Trestles	SDSC	Appro	Cluster	100.0	140.0	User Guide	Production through 2014-06-30
Quarry	Indiana U	Dell AMD	SMP	0.0	335.0	User Guide	Production through 2016-06-30
Stampede	UT Austin	Dell Dell Power Edge C8220 Cluster with Intel Xeon Phi coprocessors	Cluster	9000.0	14336.0		Coming Soon
Blacklight	PSC	SGI UV 1000 cc-NUMA	SMP	36.0	150.0	User Guide	Production through 2013-06-30
Keeneland	Georgia Tech	HP and NVIDIA	Cluster	615.0	0.0	User Guide	Production through 2014-08-31

High Throughput Computing
Visualization
Storage
Networking
Software
User Guides
Metascheduling
SU Converter

XSEDE ECSS Science Gateways Program

Mission/purpose

- Science Gateways enable communities of users associated with a common discipline to use computational resources through a familiar and simpler interface.
- The missions of the Extended Support for Science Gateway (ESSGW) Group is to provide Extended Collaborative Support to existing and new Scientific Communities in developing, enhancing and maintaining Science Gateways in effectively using XSEDE Computational Resources.
- Outreach to potential communities and help fostering new gateways.
- Engage the gateway community through forums & discussions.

ECSS Gateway Examples

- Implementation of new workflows for automation of scientific processes
- Incorporation of new visualization methods
- Innovative scheduling implementation
- Integration of XSEDE resources into a portal or Science Gateway
- Move data from gateway to XSEDE resources
- Bridge Campus Resources with XSEDE through a gateway

Contact Information

- Marlon Pierce: marpierc@iu.edu
- Suresh Marru: smarru@iu.edu
- Science Gateway Group Website: http://pti.iu.edu/sgg
- Apache Airavata: http://airavata.apache.org
- Apache Rave: http://rave.apache.org

Indiana University Science Gateway Group