JÜLICH
FORSCHUNGSZENTRUM

On Enabling Hydrodynamics Data Analysis of Analytical Ultracentrifugation Experiments

18. June 2013 | Morris Reidel, **Shahbaz Memon**, et al.

JÜLICH
FORSCHUNGSZENTRUM

Outline

- Background
- Ultrascan Application
- Ultrascan Software Components
- Airavata: Science Gateway Framework
- Current Architecture - Identified Limitations
- Integrated Architecture
- Conclusions

Background

- Science gateways have emerged as a lightweight layer for accessing distributed and heterogeneous computational and data resources
- They offer users with easy access to perform science without knowing underlying details of their remote application runs
- Many scientific communities are motivated for providing user facing scientific gateways to their researchers

EGI and XSEDE Impressions

"The most popular and most rapidly developing tools used by EGI communities to interact with the European Grid Infrastructure" [egi.eu]

"Gateways enable entire communities of users associated with a common discipline to use national resources through a common interface that is configured for optimal use." [xsede.org]

XSEDE
Extreme Science and Engineering Discovery Environment

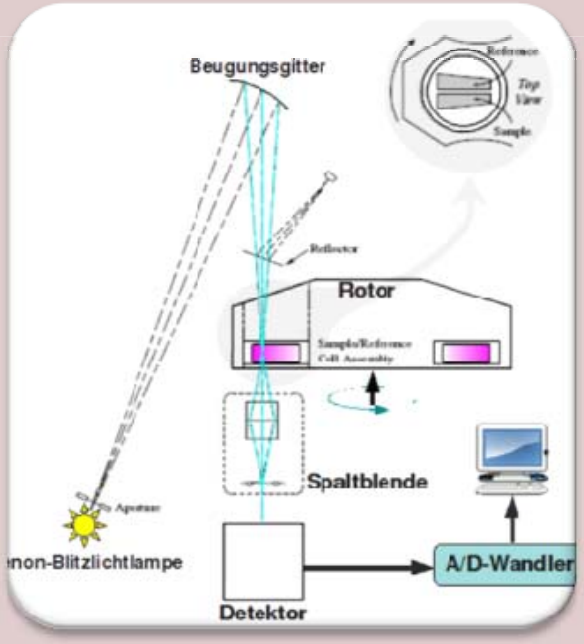
The Scientific Case: Ultrascan 3 (US3)

- US3 interprets complex data coming from high resolution analytical ultracentrifugation (AUC) experiments using sedimentation velocity centrifugation methods

- It provides better understanding of biological macromolecules and synthetic polymers that allows in the investigation of structural relationships of biological systems in many diseases such as different kind of cancers

- Users can fit their experiments using a 2-dimensional spectrum analysis, genetic algorithm optimization and Monte Carlo analysis to obtain fine grained details and analyze noise components in the data

Beckman Optima XL-A



Absorption optics schematic

View in rotor chamber

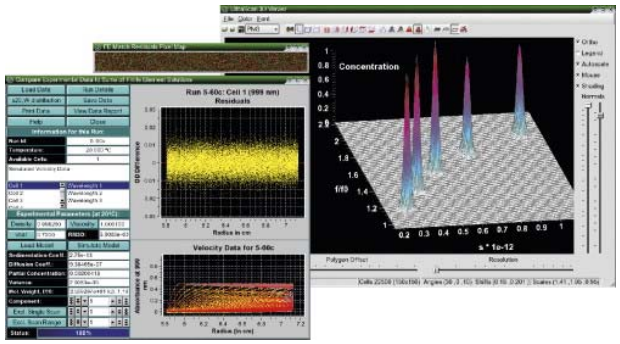
Image Courtesy: Institute of Physical Biology, Universität Düsseldorf

Ultrascan Application – US3

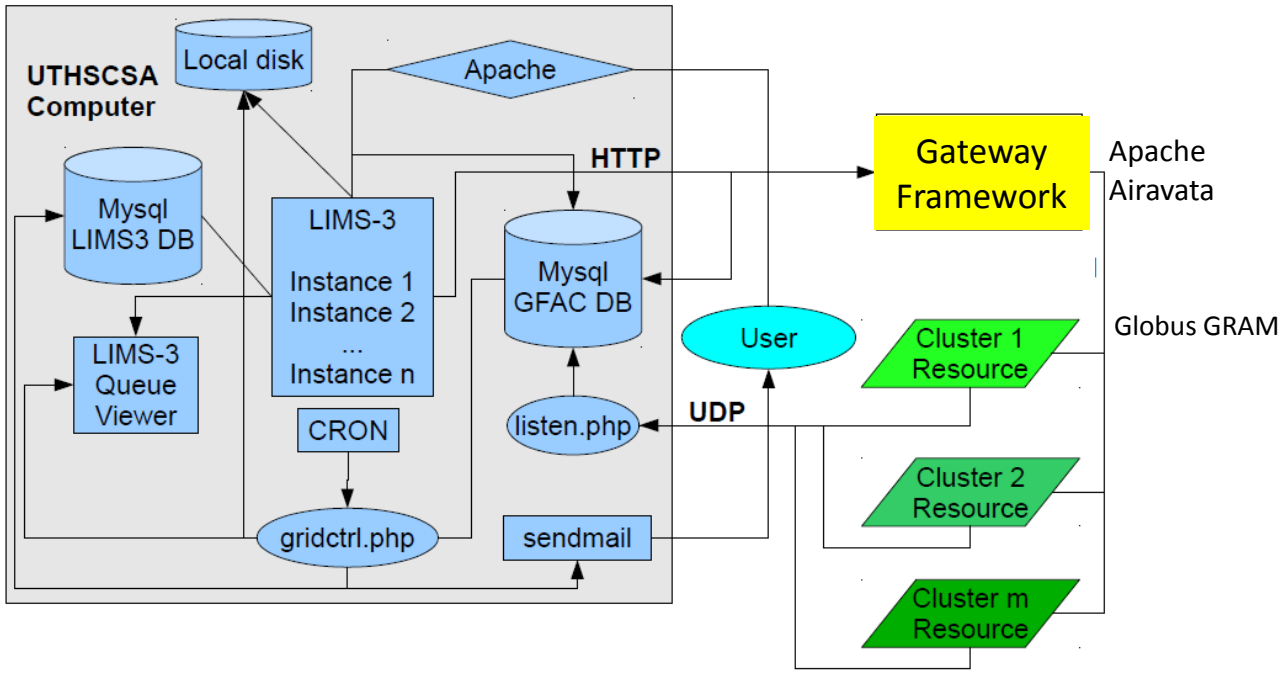
- An open source analysis and modeling software suite for biophysicists, biochemists, and material scientists
- US3 is an OpenAUC based analysis environment for analytical ultracentrifugation (AUC) experiments
- US3 provides an interface to manage data before and after AUC analysis

US3 Software Components

- MPI Data analysis package
  - Installed on HPC resources
- US-LIMS (Laboratory Information and Management System)
  - Scientific Gateway to the Ultrascan3 deployments
  - A web-based interface for collaboration, analysis job submission to remote HPC resources
  - Manage all data related to AUC analysis
  - Manage users and roles
- Desktop visualization and analysis package
  - Client side visualization suite
  - Data editing (preprocessing/cleaning)
  - Package Optimization
  - Finite Element Modeling



US-LIMS Gateway Architecture



What Is Apache Airavata?

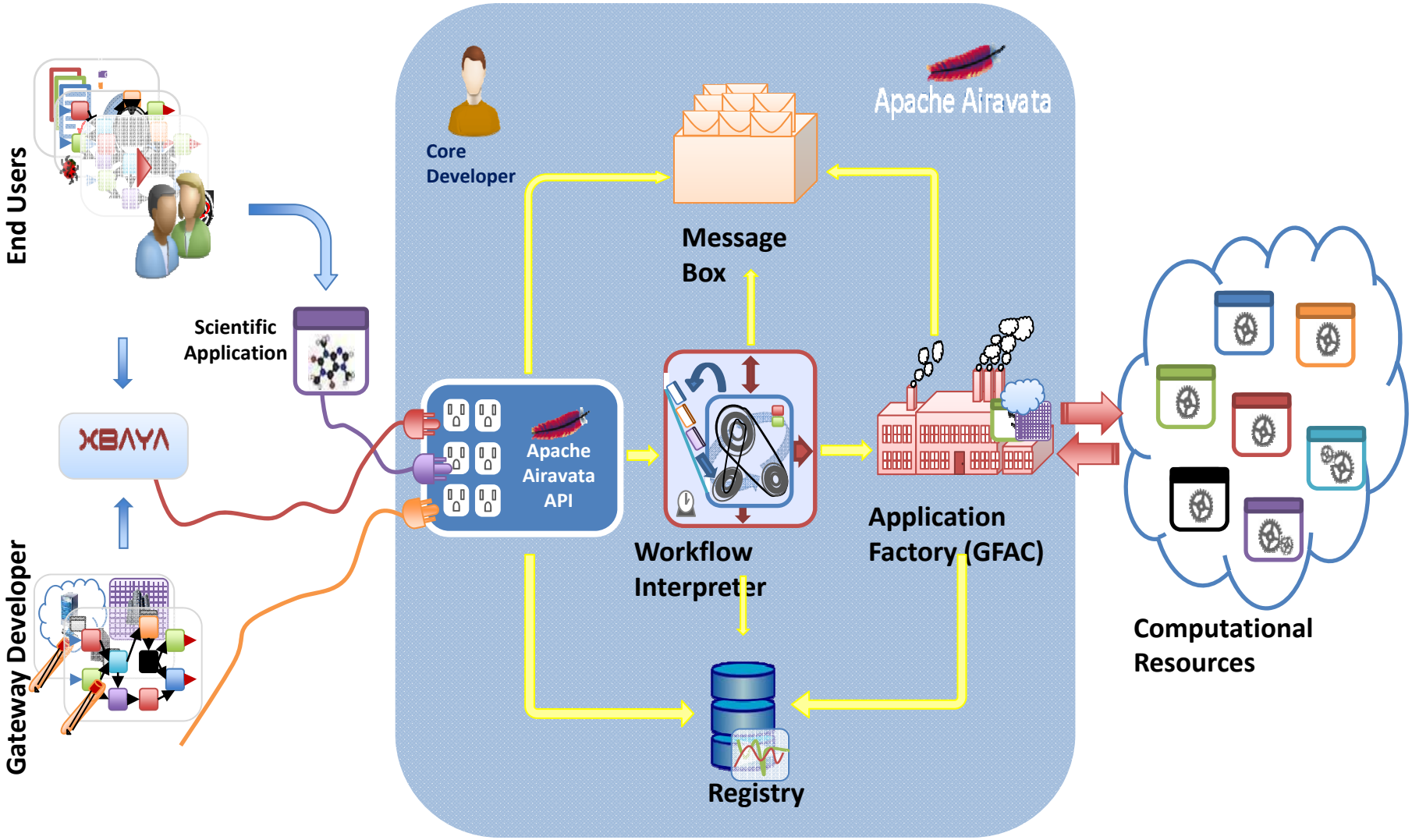
- Science Gateway software system to
  - Compose, manage, execute, and monitor distributed, computational workflows
  - Wrap legacy command line scientific applications with Web services.
  - Run jobs on computational resources ranging from local resources to computational grids and clouds

Courtesy: Marlon Pierce, OSG All Hands 2013

Apache Airavata

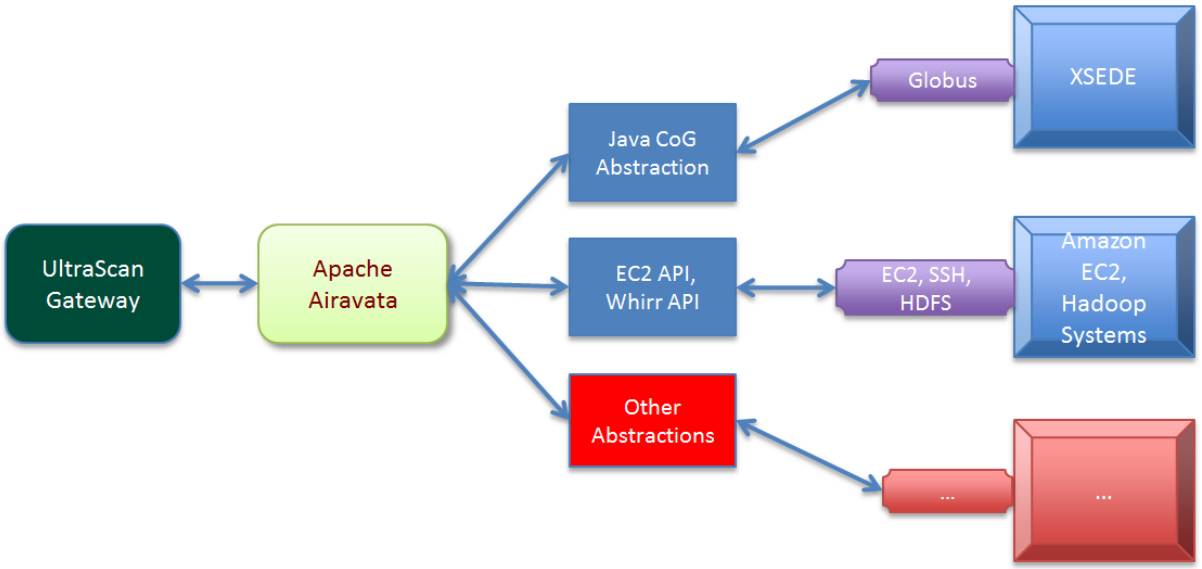
Domain	Description
Astronomy	Image processing pipeline for One Degree Imager instrument on XSEDE
Astrophysics	Supporting workflow of Dark Energy Survey simulations working group on XSEDE
Bioinformatics	Supported workflow executions on Amazon EC2 for BioVLAB project
Biophysics	Manage large scale data analysis of analytical ultracentrifugation experiments on XSEDE and campus resources
Computational Chemistry	Manage workflows to support computational chemistry parameter studies for ParamChem.org on XSEDE
Nuclear Physics	Workflows for nuclear structure calculations using Leadership Class Configuration Interaction (LCCI) computations on DOE resources

Apache Airavata Architecture

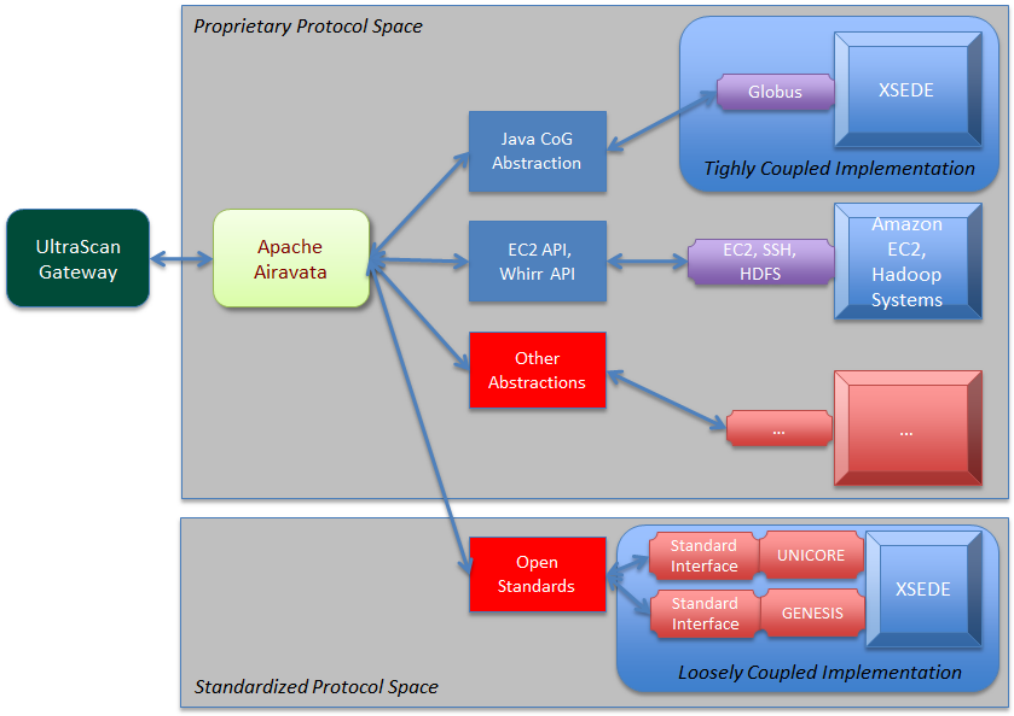


Courtesy: Marlon Pierce, OSG All Hands 2013

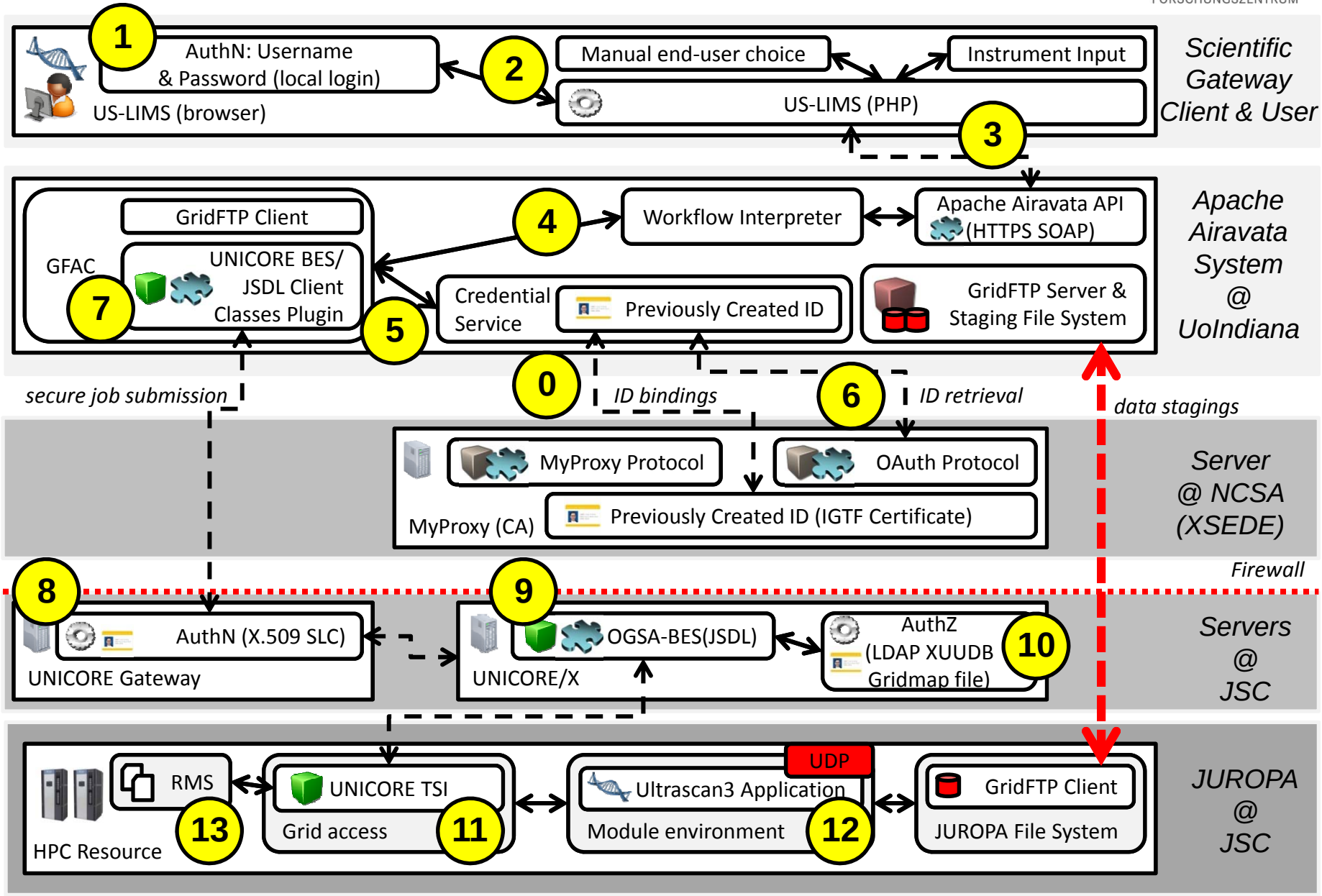
Current Architecture



Evolved Architecture



Integrated Architecture



JUROPA

- JUROPA (Jülich Research on Petaflop Architectures)
- Computing Power: 207 teraflops
- Processor Nodes: 2208 with two processors each
- Processor Type: Intel Xeon X5570 quad core, 2.93 GHz
- Batch Sytem: Moab / Torque
- Middleware: UNICORE

Conclusions and Future Work

- Implementation is under production testing by Airavata Team
- Besides JUROPA deployment this implementation has also been proposed as a pilot activity in XSEDE
- Push based job status notifications to Gateway users
- Better integration with OAuth4MP to support individuals
- Improper handling of community (pooled) accounts
- Introduce UNICORE file transfer mechanisms (ByteIO, UFTP) - complimentary to GridFTP
- Enable other science gateways to use UNICORE through Apache Airavata framework (ParamChem, GridChem, CIPRES, NeroScience, etc.)

References

- Ultrascan 3: http://www.ultrascan3.uthscsa.edu/
- UNICORE: http://www.unicore.eu/
- Borries Demeler et al., *Ultrascan Gateway Enhancements,* Terarid ’11, Salt Lake City, Utah, USA
- Apache Airavata: http://airavata.apache.org

JÜLICH
FORSCHUNGSZENTRUM

Questions?