

**Apache** *airavata*



# Apache Airavata: Enabling Science with Science Gateways

Marlon Pierce, Suresh Marru and Collaborators

SC15, Austin TX

[sgg@iu.edu](mailto:sgg@iu.edu)



# Science Gateways: User Interfaces to Supercomputers

Apache Airavata: Services that Enable  
Science Gateways to Work



sgg@iu.edu





# Analytical Ultracentrifugation Data Analysis

Borries Demeler

University of Texas Health Science  
Center, San Antonio

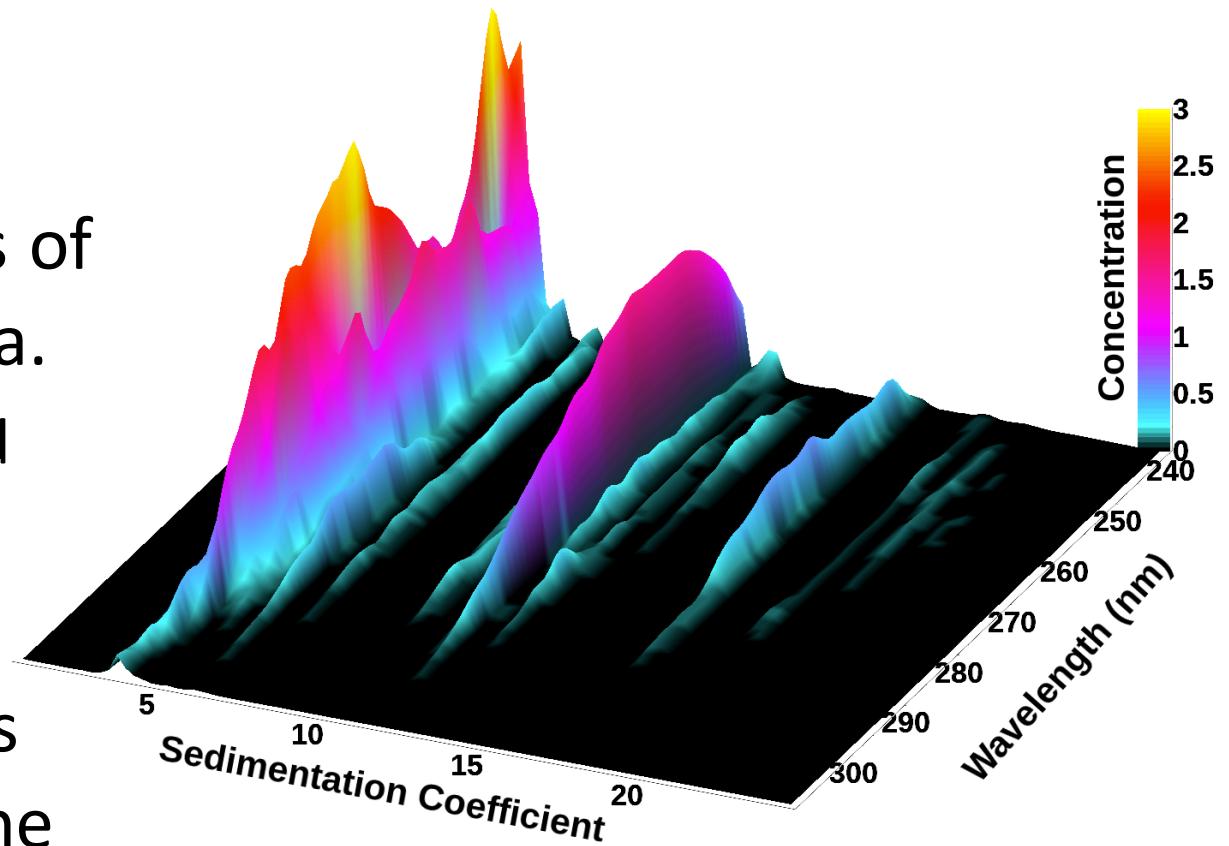


[sdg@ut.edu](mailto:sdg@ut.edu)



# Revolution in Ultracentrifugation

- New multi-wavelength instruments producing 100's of times more data.
  - Unprecedented accuracy
  - New science
- Supercomputers needed to do the analysis.



sgg@iu.edu

 SciGaP

## Cauma3d Test Database (uslims3\_cauma3d)

[Home](#) [UltraScan III](#) [UltraScan II](#) [LIMS](#) [SOMO](#) [Wiki](#)

## Welcome to the TeraGrid Science Gateway for UltraScan!

This website offers access to the UltraScan Laboratory Information Management System (USLIMS), a [TeraGrid Science Gateway](#) supported by an allocation through a TeraGrid community account. This system provides web and database support for users of the [UltraScan software](#). You can use this portal to access data associated with your sedimentation experiments, and share your data with collaborators. Authorized users can also use this site to model analytical ultracentrifugation experiments with UltraScan's high-performance analysis modules by submitting analysis jobs to computing clusters available at the [University of Texas Health Science Center](#) and TeraGrid sites at the [Texas Advanced Computing Center](#) and at [UT Health San Antonio](#).

Advanced Computing Center and at UT Health San Antonio are provided by the University of Texas at Austin, and are funded by a TeraGrid community account (see below for full details).

## DISCLAIMER:

We do not take any responsibility for your use of this facility. It is your responsibility to always make arrangements for such a service before you use it.

Funding for this facility is provided through grants from the following institutions:

- [Department of Biochemistry, UTHSCSA](#)
- User fees collected from collaborative projects
- San Antonio Life Science Institute
- [The National Science Foundation](#) (Demeler)
- [The National Institutes of Health](#) (Demeler)

When publishing, please credit our facility:

**Calculations were performed on the UltraScan Laboratory Information Management System (USLIMS) at the Bioinformatics Core Facility at the University of Texas Health Science Center at San Antonio and the Texas Advanced Computing Center supercomputer (TACC) (Award #MCB070038 (to Borries Demeler)).**

Please enter the link to each manuscript:

Before logging in, if you have not done so already, please log in. This will make it easier to use the secure password feature.

**Borries Demeler, Ph.D.**  
Associate Professor  
UltraScan Project Director

## Navigation

[Welcome!](#)[Admin Info](#)

## Project

[Projects](#)[Images](#)[Reports](#)[Sharing](#)

## Analysis

[Queue Setup](#)[2DSA Analysis](#)[2DSA Custom Grid](#)[GA Analysis](#)[RunID Info](#)

## Status Monitor

[Queue Status](#)[Cluster Status](#)

## General

[Change My Info](#)[Database Login Info](#)

## Partners

[Contacts](#)[Webmaster](#)[Data Security](#)[Logout](#)

Launch analysis and monitor through a browser

## 2DSA Analysis

Initialize 2DSA Parameters - demo1\_veloc\_rs.RA.2.A.260.auc; Edit profile: 1308301540; Dataset 1 of 1

## S-Value Resolution

1	S-Value Minimum
10	S-Value Maximum
60	S-Value Resolution (total grid points)

## f/f0 Resolution

1	f/f0 Minimum
4	f/f0 Maximum
60	f/f0 Resolution (total grid points)

## Uniform Grid Repetitions Setup

6	Uniform Grid Repetitions
---	--------------------------

## Monte Carlo Iterations

1	Monte Carlo Iterations
---	------------------------

Value: 1	Minimum: 1	Maximum: 100
----------	------------	--------------

## Fit Time Invariant Noise

On  
 Off

[Show Advanced Options](#)[Edit Profiles](#) [Change Experiment](#)

## Dataset control:

- Current dataset number: 1
- Run Name: demo1\_veloc\_rs.RA.2.A.260.auc
- Number of datasets: 1

## Select Cluster

Cluster	Status	Queue Name	running/queued
<input checked="" type="radio"/> stampede	*	normal	* / *
<input type="radio"/> lonestar	unknown	normal	0 / 0
<input type="radio"/> trestles	*	normal	* / *
<input type="radio"/> juropa	*	default	* / *
<input type="radio"/> alamo	unknown	default	0 / 0
<input type="radio"/> bcf	up	default	0 / 0

[Submit](#)

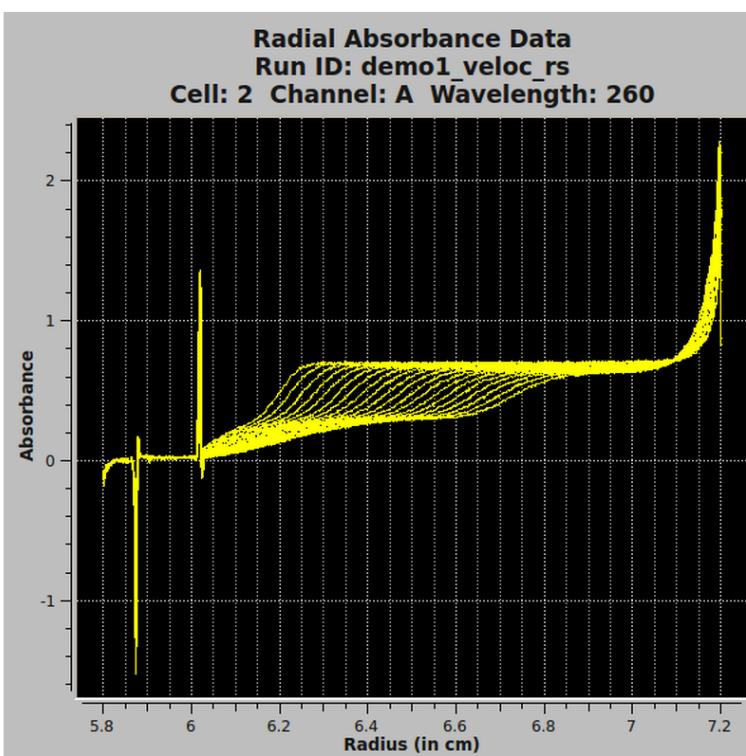
Desktop analysis tools are integrated with the Web portal.

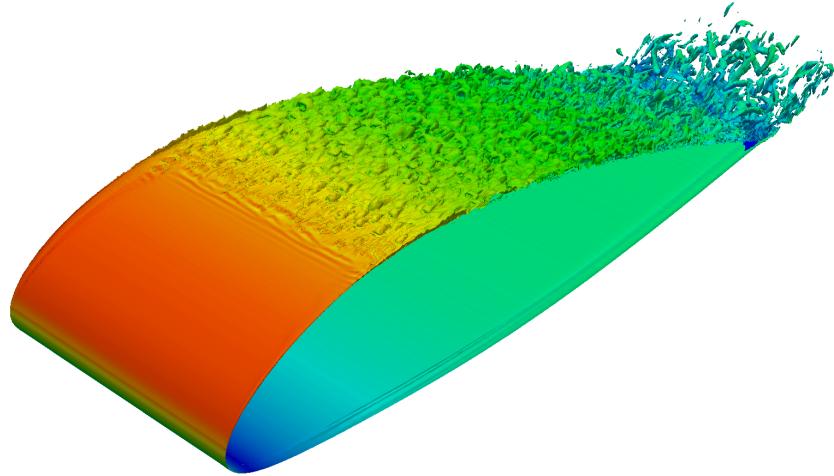
The UltraScan science gateway enables experimental scientists to analyze data on supercomputers.

UltraScan uses Apache Airavata for managing analyses on HPCs across the world.

**Run ID: demo1\_veloc\_rs**  
**Cell 2, Channel A, Wavelength 260**  
**Edited Dataset:**

**Raw Scan Data (PNG Plot)**  
Filename:cnvt.2A260.raw.png





# Large Eddy Simulations of Turbulent Flow

Cameron Smith, Steve Tran, and  
Onkar Sahni

Rensselaer Polytechnic Institute



[sgg@iu.edu](mailto:sgg@iu.edu)



# Science Use Case: PHASTA

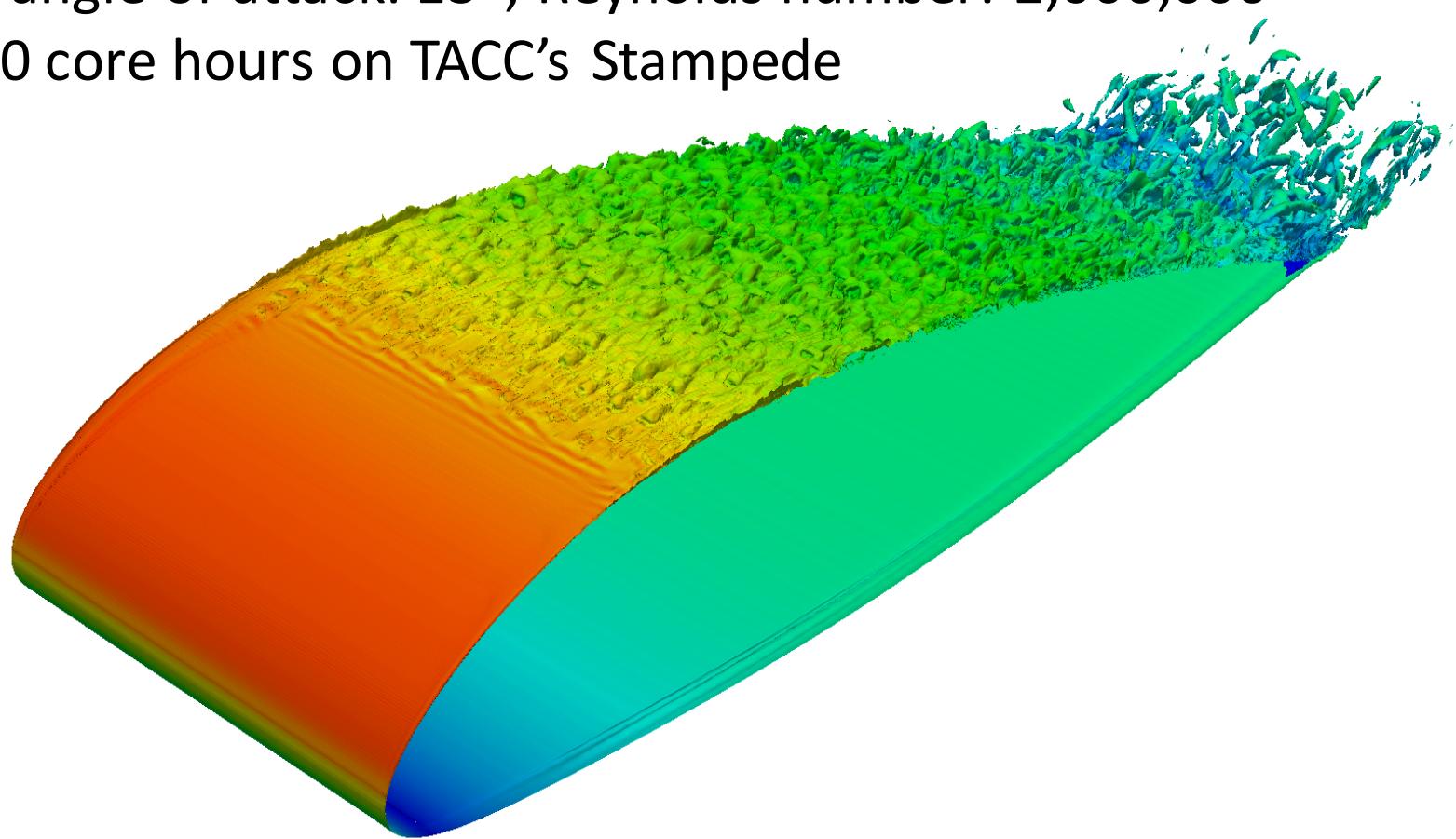
- One of the challenges in studying turbulent flows is the multiscale nature of turbulence.
  - Direct Numerical Simulation (DNS) resolves turbulent flow structures up to dissipation scale, but is not practical with current HPC resources on engineering problems of interest.
- Large Eddy Simulation (LES) is a powerful tool which is capable of accurately resolving and modeling turbulence at a reasonable cost.
  - LES resolves those flow features that can be captured by the numerical discretization/grid while modeling their interactions with the unresolved (subgrid scale) flow features.
  - To account for inhomogeneous flows, Lagrangian-averaged dynamic Smagorinsky model is employed in stabilized finite element simulation.
- Example case: Flow over Aerospatiale-A airfoil
  - Static angle of attack  $13^\circ$ , Reynolds number 2,000,000
  - Flow transitions to a turbulent flow and exhibits marginal flow separation near the trailing edge

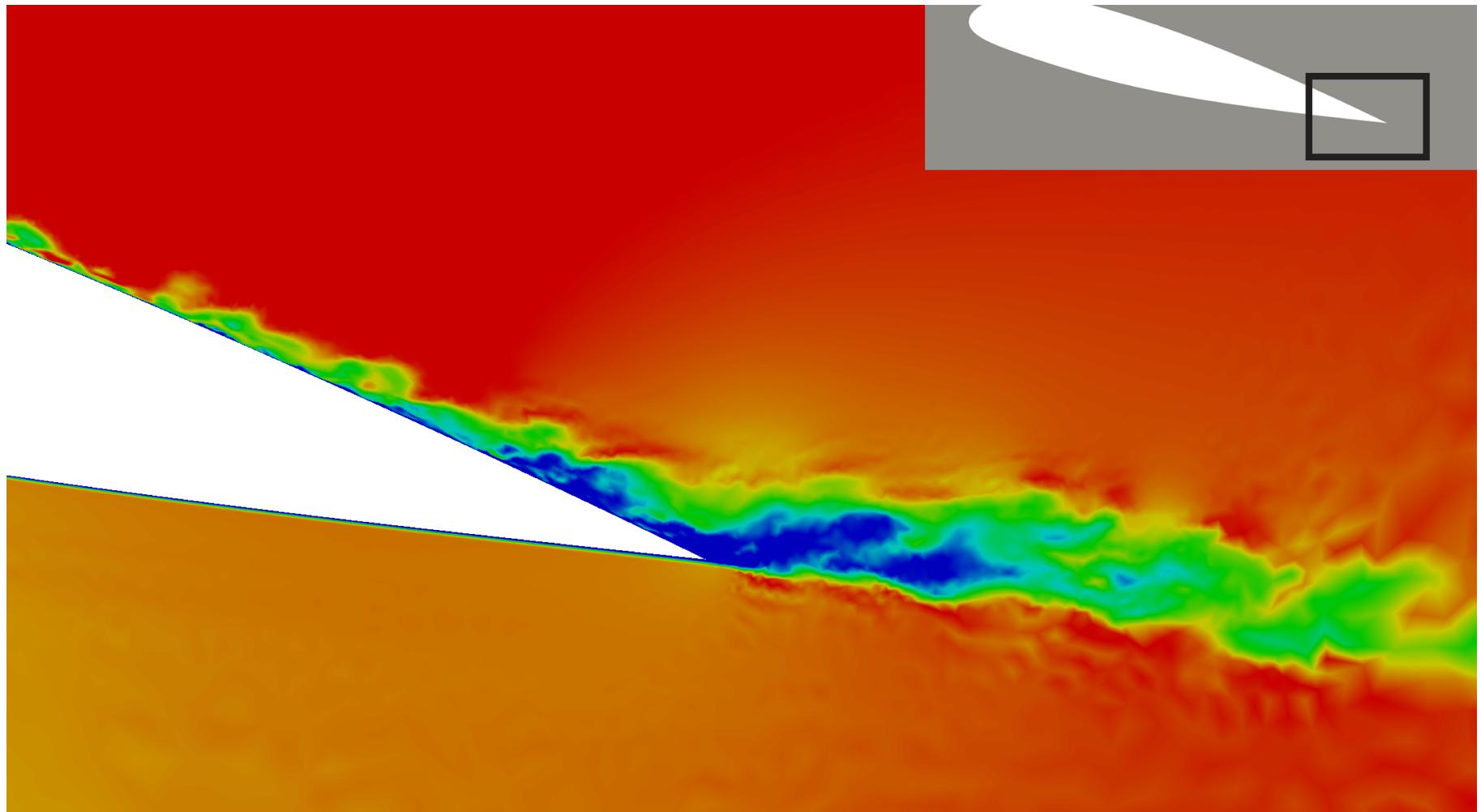


sgg@iu.edu



- Direct Numerical Simulation (DNS) not practical for many engineering problems.
- Large Eddy Simulation (LES) models turbulence at a reasonable cost.
  - Lagrangian-averaged dynamic Smagorinsky model
- Example: flow over Aerospatiale-A airfoil
  - Static angle of attack:  $13^\circ$ ; Reynolds number: 2,000,000
  - 75,000 core hours on TACC's Stampede





sgg@iu.edu

 SciGaP

## PHP Gateway with Airavata

PGA is a science gateway built with the Airavata API. You can reference PGA as you integrate Airavata into your own gateway, or you can create your gateway on top of PGA by cloning it at the link below. PGA is known to work well in the Chrome, Firefox, and Internet Explorer browsers.

[See the code ↗](#)

[XSEDE 2015 tutorial documentation ↗](#)



SciGaP is a hosted service with a public API that science gateways can use to manage applications and workflows running on remote supercomputers, as well as other services. Gateway developers can thus concentrate their efforts on building their scientific communities and not worry about operations.

Science Gateway Platform as a Service (SciGaP) provides application programmer interfaces (APIs) to hosted generic infrastructure services that can be used by domain science communities to create Science Gateways.

[Learn more ↗](#)



powered by

## Apache Airavata

Apache Airavata is a software framework which is dominantly used to build Web-based science gateways and assist to compose, manage, execute and monitor large scale applications and workflows on distributed computing resources such as local clusters, supercomputers, national grids, academic and commercial clouds. Airavata mainly supports long running applications and workflows on distributed computational resources.

[Learn more ↗](#)

Apache Airavata manages multi-stepped workflow, multiple code versions.

PHASTA team uses Apache Airavata and the PGA to run simulations on TACC's Stampede.

### Experiment Summary ↗

Experiment Id	Phast-Exp-Stampede_4e07a9e1-31a2-4210-b3b6-d32bd48cf498
Name	Phast-Exp-Stampede
Description	Phasta exp running on Stampede
Project	November/12/2015
Application	Phasta_P
Compute resource	stampede.tacc.xsede.org
Experiment Status	COMPLETED
Job Status	COMPLETE
Creation time	2015-11-12, 12:43 PM - GMT-0500 (EST)
Last Modified Time	2015-11-12, 12:44 PM - GMT-0500 (EST)
Enable Auto Schedule	false
Wall time	30
CPU count	16
Node count	1
Queue	normal
Inputs	<a href="#">geom.xmt_txt ↗</a> <a href="#">geom.smd ↗</a> <a href="#">geom.sms ↗</a> <a href="#">solver.inp ↗</a>
Outputs	Phasta-Output-TAR : <a href="#">PHASTA_Output.tar.gz ↗</a> Phasta-Standard-Error : <a href="#">Phasta_P.stderr ↗</a> Phasta-Standard-Out : <a href="#">Phasta_P.stdout ↗</a>

[Clone](#)



# Science Gateways for Small Angle Scattering

Prof. Emre Brookes, University of  
Texas Health Science Center, San  
Antonio



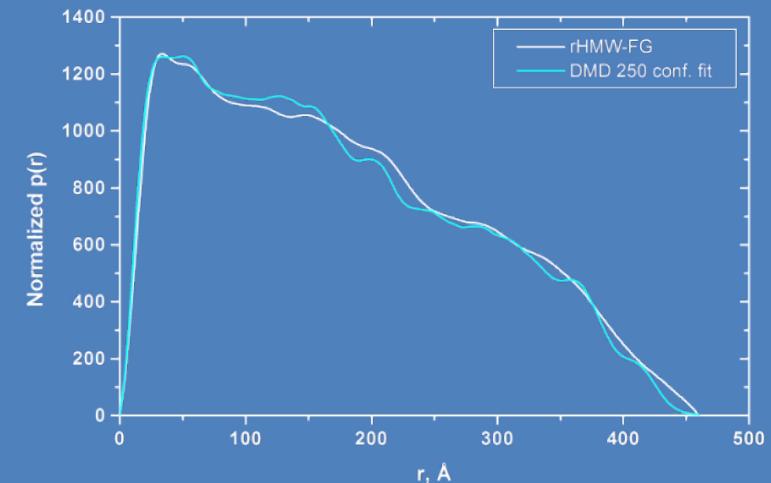
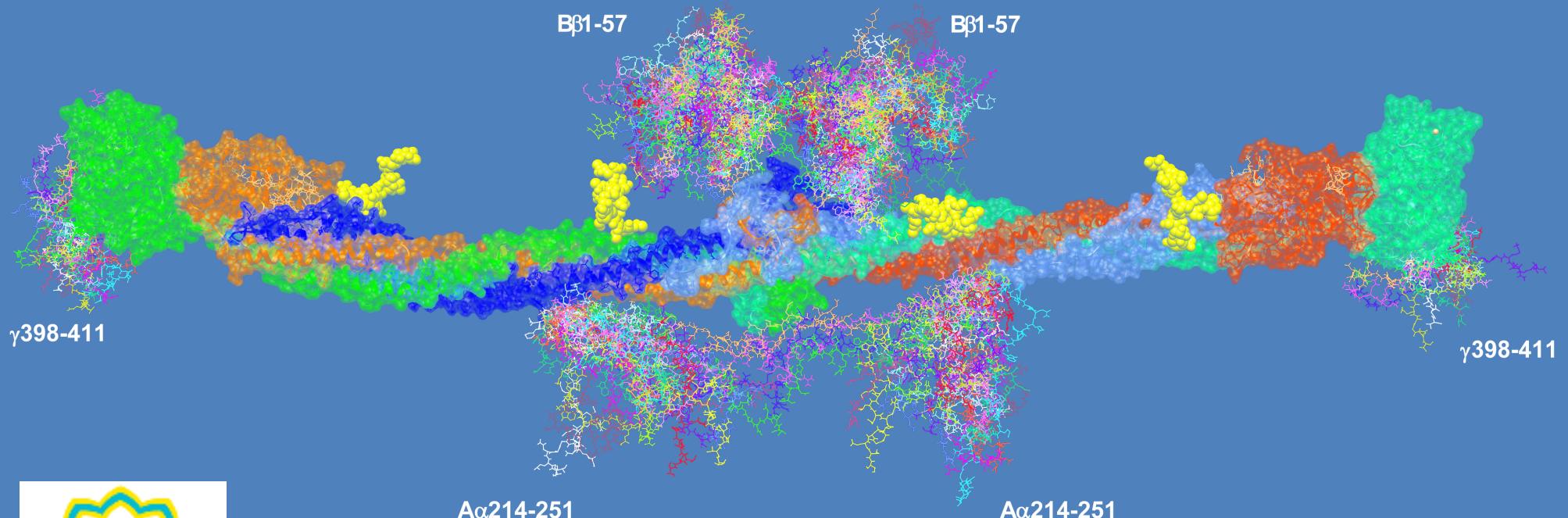
sgg@iu.edu



# Conformational variability / Fibrinogen

- Fibrinogen is an important component of the coagulation cascade, as well as a major determinant of blood viscosity and blood flow
- A centrosymmetric dimer made by 3 pairs of chains
- US-SOMO/DMD simulations of the conformational variability for comparison to experimental data

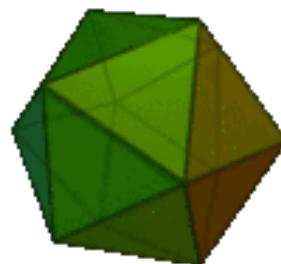
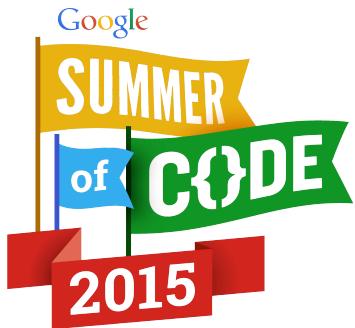
Images credit: Mattia Rocco



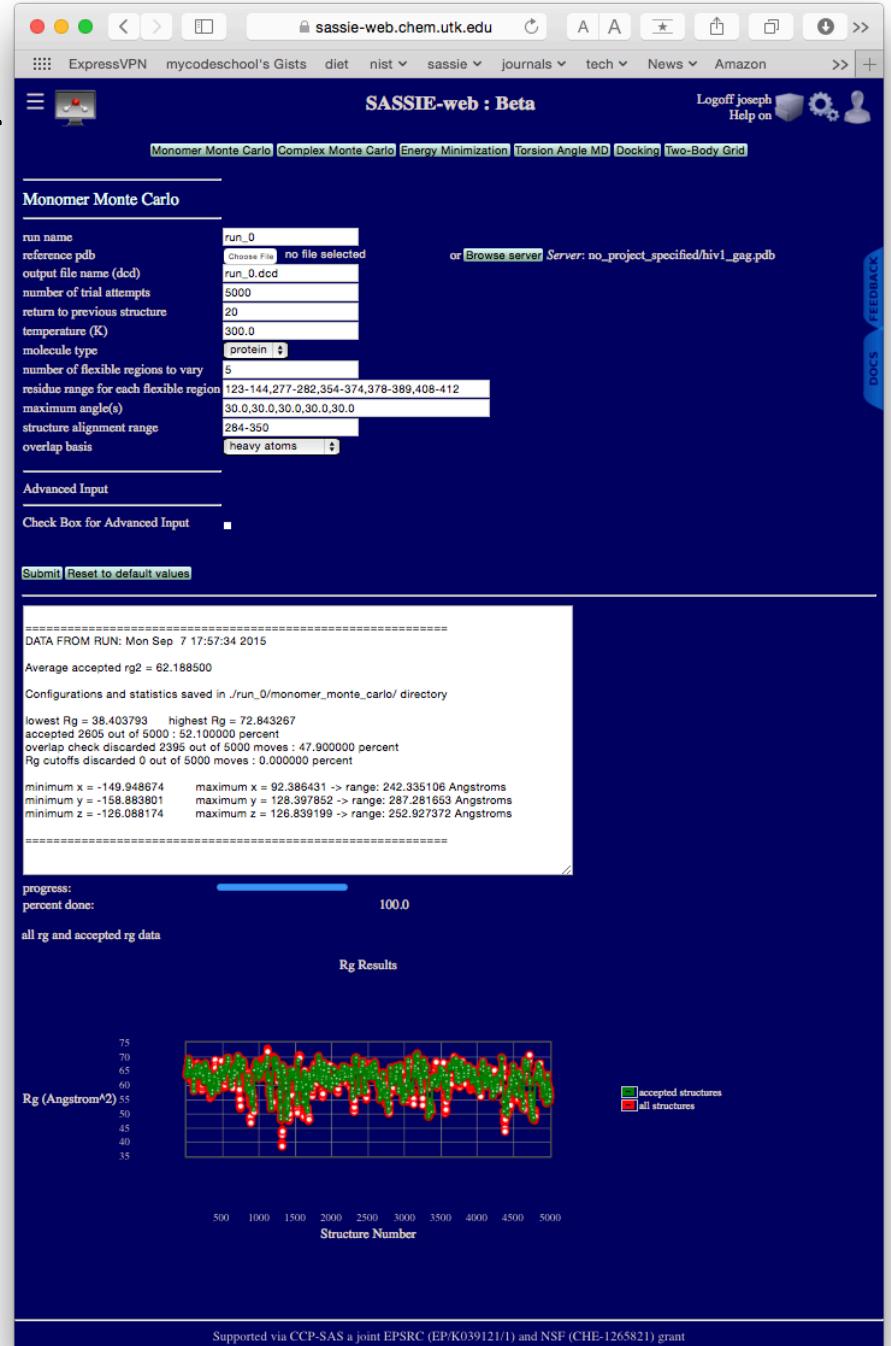
sgg@iu.edu

SciGaP

- GenApp generates user interfaces for the wide range of small angle scattering codes.
  - Python, Tcl/Tk, HTML5, Java
  - QT, Web Apps, iOS, Android
- GenApp uses Apache Airavata's Thrift APIs to manage workflows on HPC resources.
- Code: <http://genapp.rocks>
- Example:
- <https://sassie-web.chem.utk.edu/sassie2/>



sgg@iu.edu



Supported via CCP-SAS a joint EPSRC (EP/K039121/1) and NSF (CHE-1265821) grant

# Science Gateways and Cloud Research on Jetstream

Prof. Madhusudhan Govindaraju  
State University of New York  
Binghamton



[sgg@iu.edu](mailto:sgg@iu.edu)



Chrome File Edit View History Bookmarks People Window Help

Mesos scigap-1r.jetstream-cloud.org:5050/#/ Renan

**Mesos**

	CPU	Mem
Total	36	88.0 GB
Used	12	12.0 GB
Offered	0	0 B
Idle	24	76.0 GB

Active Tasks

ID	Name	State	Started	Host
Marathon	Marathon	Running	2015-11-10 12:00:00Z	scigap-1r.jetstream-cloud.org:8080

Suspend Scale Restart App Destroy App

Tasks Configuration Debug

Refresh

ID	Status	Version	Updated
idv-cloud-withpsw.3e38b7b1-873c-11e5-964a-0242711de25c	Started	2 days ago	11/9/2015, 6:47:33 PM
idv-cloud-withpsw.b45354a3-87d2-11e5-964a-0242711de25c	Started	2 days ago	11/10/2015, 12:45:23 PM

noVNC scigap-1r.jetstream-cloud.org:31116 Renan

Connected (unencrypted) to: 199dedb13890:1

Send CtrlAltDel Shutdown Reboot Reset

Dashboard

File Edit Displays Data Tools Help

Unidata IDV - Globe View - One Pane

File Edit Displays Data Tools Help

View Projections

15:00:53 GMT Latitude: 2.5 Longitude: 173.7 Altitude: 0.0 m

15:00:53 GMT

< Workspace 1 > 11 Nov, Wed 15:00:53 < > idv@199dedb13890: ~ Dash



sgg@iu.edu

SciGaP

# Science Gateway Platform as a Service

Mark Miller, Amit Majumdar  
San Diego Supercomputer Center



sgg@iu.edu



## Missing results?

Send us the [job handle](#),  
and we may be able to  
help.

## More Information

About Us

Requirements

Known Issues

Usage Statistics

User Locations

Enabled Publications

The CIPRES Science Gateway now offers PhyloBayes MPI and DPPDIV, along with fast RAxML and MrBayes codes.

First Time Users: Please review the [XSEDE Primer](#) and our [Fair Use Policy](#).

## CIPRES Login:


 
[Forgot Password?](#)

## iPlant users login here:


[What is this?](#)

**Status:** Routine Maintenance will occur on Mar 14, 2014; because this will kill running jobs, submissions that could enter that time period will be held in the queue until maintenance is complete.

## CIPRES Gateway News

**PhyloBayes MPI and MrBayes 3.2.2 now available**

10/7/2013

**Command Change Issues for MB 3.2.1**

8/23/2013

# Simplified access to phylogenetics codes on powerful XSEDE resources

## Folders

- +  [sarahm](#)
- +  [amy](#)
- +  [pamela](#)
- +  [andy](#)
- +  [James MrBayes](#)
- +  [peggy](#)
- +  [phylobayes](#)
- +  [Jana](#)
- +  [John Philips](#)
- +  [paul mrbayes 3.2.2](#)
  -  [Data \(1\)](#)
  -  [Tasks \(3\)](#)
- +  [phylobayes](#)
- +  [oldmrbayes](#)
- +  [joe](#)
- +  [susana](#)
- +  [aelys](#)
- +  [vanessa](#)
- +  [readseq](#)
- +  [truncated\\_treee](#)
- +  [selma](#)
- +  [diogo](#)
- +  [torsten](#)
- +  [mike\\_hawaii](#)
- +  [raxml\\_test](#)



Task Summary Select Data Select Tool Set Parameters

MrBayes 3.2.2 on XSEDE: Tree Inference Using Bayesian Analysis - run on XSEDE ([John P. Huelsenbeck and Fred Ronquist](#))

[Simple Parameters](#)

OPEN / CLOSE

My Data Contains a MrBayes Data Block (CHECK THIS OR MrBayes BLOCK ENTRIES WILL BE OVERWRITTEN!!!) \*

I confirm that there is not an "autoclose = no" statement in my MrBayes block \*

My MrBayes Block specifies runs=

My MrBayes Block specifies nchains=

Maximum Hours to Run (click here for help setting this correctly) \*

My Data Type Is (only one data type can be used through the web form, see help below) \*  nucleic acid

Set the Seed Number Set seed=

Set the Swapseed=

Use scientific notation

How many decimals should we print?



A Portal for Computational Neuroscience

Amit Majumdar (PI), Maryann Martone (Co-PI), Subha Sivagnanam (Sr. Personnel)  
Kenneth Yoshimoto (Sr. Personnel), Anita Bandrowski (Sr. Personnel), Vadim Astakhov UCSD  
Ted Carnevale (PI), Yale School of Medicine  
NSF Awards: ABI #1146949; ABI #1146830

The NSG is a simple and secure online science portal that provides access to computational neuroscience codes on XSEDE HPC resources

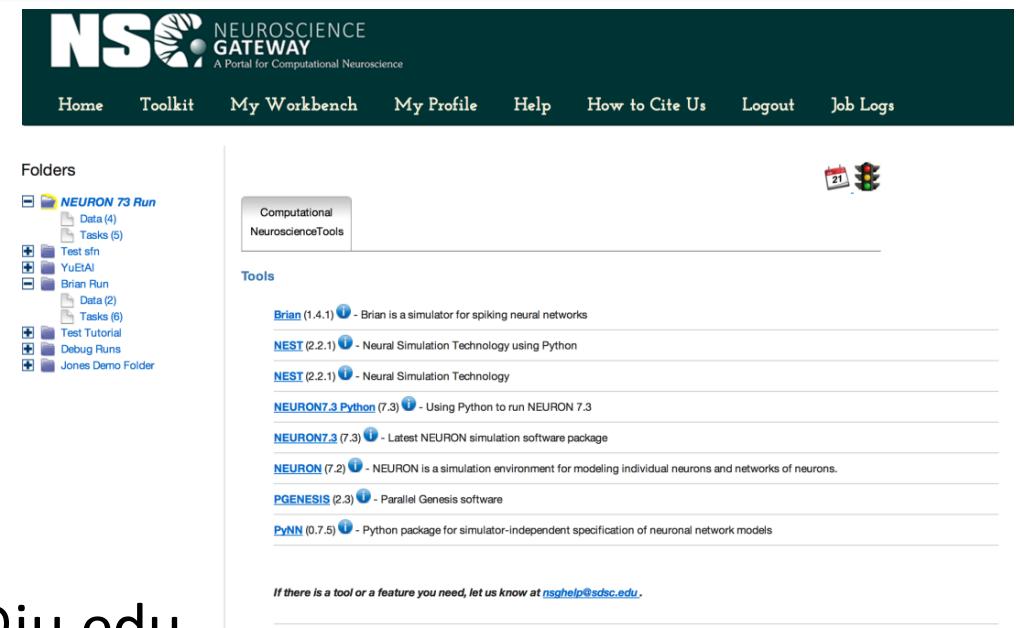
<http://www.nsgportal.org>

Easy user interface – providing easy model upload, running of codes

Complete set of neuronal simulation tools – **NEURON, GENESIS, Brian, NEST, PyNN** – widely used by computational neuroscientists

Ability to easily get to the results, download results

*Democratize computational neuroscience*



The screenshot shows the NSG portal homepage. At the top, there is a dark header bar with the NSG logo and the text "NEUROSCIENCE GATEWAY A Portal for Computational Neuroscience". Below the header, there is a navigation menu with links for Home, Toolkit, My Workbench, My Profile, Help, How to Cite Us, Logout, and Job Logs. The main content area has a purple background. On the left, there is a sidebar titled "Folders" containing a list of simulation runs and tools. On the right, there is a section titled "Tools" listing various simulation software packages. A small calendar icon and a traffic light icon are also visible in the top right corner of the main content area.

Folders

- NEURON 7.3 Run
  - Data (4)
  - Tasks (5)
- Test smn
- YuBAI
- Brian Run
  - Data (2)
  - Tasks (6)
- Brian Tutorial
- Debug Runs
- Jones Demo Folder

Computational Neuroscience Tools

Tools

- Brian** (1.4.1) - Brian is a simulator for spiking neural networks
- NEST** (2.2.1) - Neural Simulation Technology using Python
- NEST** (2.2.1) - Neural Simulation Technology
- NEURON7.3 Python** (7.3) - Using Python to run NEURON 7.3
- NEURON7.3** (7.3) - Latest NEURON simulation software package
- NEURON** (7.2) - NEURON is a simulation environment for modeling individual neurons and networks of neurons.
- GENESIS** (2.3) - Parallel Genesis software
- PyNN** (0.7.5) - Python package for simulator-independent specification of neuronal network models

If there is a tool or a feature you need, let us know at [nsghelp@sdsc.edu](mailto:nsghelp@sdsc.edu).

sgg@iu.edu