SciGaP Annual Report Metrics

Summary

Science Metric	Year 1	Year 2	Year 3	Year 4	Year 5
Enabled	>23	>22			

Publications					
Supported Allocations (#, SU)	5 awards, 1,440,000 SU	7 awards, 1,841,129 SU			
Collaborating Grants (#, $)	1 Grant, $263,000	4 Grants, > $2,000,000			

Operational Metric	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Gateways Using SciGaP	2	6 (see XSEDE allocations table below)			
Number of Users	127 (2013) 96 (2014 to Date)	130 (2014), 90 (2015, to date)			
Number of Experiments	19,595 (2013) 16,520 (2014 through September 20th)	1,274,232 (2014) 154,511 (2015 to date)			
Allocation Usage	784,064 (2013) 291,859 (2014)	1,615,746 (2014) 1,327,605 (2015 to date)			
Resources Available to Users	6 clusters made available	13 XSEDE, Campus, and International Resources Gordon (SDSC), Stampede			

		(TACC), Comet (SDSC), Alamo2 (UTHSCSA), Jureca (JSC), Big Red II (IU), Karst (IU), Mason (IU), Lonestar (TACC) Trestles (SDSC), Juropa (JSC), Alamo1 (UTHSCSA), Blacklight (PSC) all decommissioned.			
SciGaP Service Uptime	99.982% (pre-SciGaP Airavata Server) 97.774% (SciGaP Server)	99.99% (Nagios Monitoring)			
Job Failure Rates	10% (~80% of failed jobs resulted from application failures on the target resource)	8.68% (~75% of the job failures were failures on the target resource or incorrect application inputs)			
Jira Tickets Created, Closed	515, 288	370, 148			
New Apache	8 committers,	2 committers,			

Airavata Committers and PMC Members	0 PMC members	2 PMC members			
Apache Airavata Downloads (Unique IPs of all releases)	127	480			

Science Metrics

Number of Publications Enabled

Totals

In Year 2, we completed integration with the Workbench Framework (which is used to build NSG and CIPRES) as planned and are completing the transition of NSG to operations, so we do not report these numbers. In Year 3, we will support CIPRES, NSG, and SEAGrid in SciGaP production infrastructure. We include these numbers as a baseline for future reports.

- SciGaP Team Publications: 14
- SciGaP Supported Science Publications: 22
- NSG, CIPRES, SEAGrid Enabled Publications: Approximately 200, including papers in press.

Details are below.

Team Publications

Cyberinfrastructure and scientific publications from previous years directly supported by this grant have been sited at least 63 times (Google Scholar, 2013-2015). 6 papers supported by this award were cited at least 6 times.

1. Gesing, Sandra, Rion Dooley, Marlon Pierce, Jens Kruger, Richard Grunzke, Sonja Herres-Pawlis, and Alexander Hoffmann. "Science gateways-leveraging modeling and simulations in HPC infrastructures via increased usability." In High Performance Computing & Simulation (HPCS), 2015 International Conference on, pp. 19-26. IEEE, 2015.
2. Lawrence, Katherine A., Michael Zentner, Nancy Wilkins-Diehr, Julie A. Wernert, Marlon Pierce, Suresh Marru, and Scott Michael. "Science gateways today and tomorrow: positive perspectives of nearly 5000 members of the research community." Concurrency and Computation: Practice and Experience(2015).

3. Brookes, Emre H., Nadeem Anjum, Joseph E. Curtis, Suresh Marru, Raminder Singh, and Marlon Pierce. "The GenApp framework integrated with Airavata for managed compute resource submissions." Concurrency and Computation: Practice and Experience (2015).
4. Pierce, Marlon E., Suresh Marru, Lahiru Gunathilake, Don Kushan Wijeratne, Raminder Singh, Chathuri Wimalasena, Shameera Ratnayaka, and Sudhakar Pamidighantam. "Apache Airavata: design and directions of a science gateway framework." Concurrency and Computation: Practice and Experience (2015).
5. Pierce, Marlon, Suresh Marru, Borries Demeler, Raminderjeet Singh, and Gary Gorbet. "The apache airavata application programming interface: overview and evaluation with the UltraScan science gateway." In Proceedings of the 9th Gateway Computing Environments Workshop, pp. 25-29. IEEE Press, 2014.
6. Marru, Suresh, Marlon Pierce, Sudhakar Pamidighantam, and Chathuri Wimalasena. "Apache Airavata as a Laboratory: Architecture and Case Study for Component-Based Gateway Middleware." In Proceedings of the 1st Workshop on The Science of Cyberinfrastructure: Research, Experience, Applications and Models, pp. 19-26. ACM, 2015.
7. Heiland, Randy, Scott Koranda, Suresh Marru, Marlon Pierce, and Von Welch. "Authentication and Authorization Considerations for a Multi-tenant Service." InProceedings of the 1st Workshop on The Science of Cyberinfrastructure: Research, Experience, Applications and Models, pp. 29-35. ACM, 2015.
8. Lawrence, Katherine A., Nancy Wilkins-Diehr, Julie A. Wernert, Marlon Pierce, Michael Zentner, and Suresh Marru. "Who cares about science gateways?: a large-scale survey of community use and needs." In Proceedings of the 9th Gateway Computing Environments Workshop, pp. 1-4. IEEE Press, 2014.
9. Brookes, Emre H., Nadeem Anjum, Joseph E. Curtis, Suresh Marru, Raminder Singh, and Marlon Pierce. "GenApp module execution and airavata integration." In Proceedings of the 9th Gateway Computing Environments Workshop, pp. 9-12. IEEE Press, 2014.
10. Nakandala, Supun, Sachith Dhanushka Withana, Dinu Kumarasiri, Hirantha Jayawardena, H. M. N. Dilum Bandara, Srinath Perera, Suresh Marru, and Sudhakar Pamidighantam. "Schema-independent scientific data cataloging framework." In Moratuwa Engineering Research Conference (MERCon), 2015, pp. 289-294. IEEE, 2015.
11. Carnevale, Ted, Amit Majumdar, Subha Sivagnanam, Kenneth Yoshimoto, Vadim Astakhov, Anita Bandrowski, and Maryann Martone. "The neuroscience gateway portal: high performance computing made easy." BMC Neuroscience 15, no. Suppl 1 (2014): P101.
12. Sivagnanam, Subhashini, Amit Majumdar, Kenneth Yoshimoto, Vadim Astakhov, Anita Bandrowski, MaryAnn Martone, and Nicholas Carnevale. "Early experiences in developing and managing the neuroscience gateway." Concurrency and Computation: Practice and Experience 27, no. 2 (2015): 473-488.
13. Miller, Mark A., Terri Schwartz, Paul Hoover, Kenneth Yoshimoto, Subhashini Sivagnanam, and Amit Majumdar. "The CIPRES workbench: a flexible framework for

creating science gateways." In Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure, p. 39. ACM, 2015.
14. Miller, M. A., Schwartz, T., Pickett, B. E., He, S., Klem, E. B., Scheuermann, R. H., Passarotti, M., Kaufman, S., and O'Leary, M. A. (2015) A RESTful API for Access to Phylogenetic Tools via the CIPRES Science Gateway. *Evolutionary Bioinformatics* **11**, 43-48.

SciGaP-Supported Publications

These are scientific (rather than Cyberinfrastructure) publications and posters that were enabled by the SciGaP project.

1. Cockburn, Darrell W., Nicole I. Orlovsky, Matthew H. Foley, Kurt J. Kwiatkowski, Constance M. Bahr, Mallory Maynard, Borries Demeler, and Nicole M. Koropatkin. "Molecular details of a starch utilization pathway in the human gut symbiont Eubacterium rectale." *Molecular microbiology* 95, no. 2 (2015): 209-230.
2. Tsutsui, Yuko, Jennifer M. Johnson, Borries Demeler, Michael Kinter, and Franklin Alan Hays. "Conformation-Dependent Human p52Shc Phosphorylation by Human c-Src." *Biochemistry* (2015).
3. Ranaivoson, Fanomezana M., Qun Liu, Francesca Martini, Francesco Bergami, Sventja von Daake, Sheng Li, David Lee, Borries Demeler, Wayne A. Hendrickson, and Davide Comoletti. "Structural and Mechanistic Insights into the Latrophilin3-FLRT3 Complex that Mediates Glutamatergic Synapse Development." *Structure* 23, no. 9 (2015): 1665-1677.
4. Gorbet, Gary E., Joseph Z. Pearson, Aysha K. Demeler, Helmut Cölfen, and Borries Demeler. "Next-Generation AUC: Analysis of Multiwavelength Analytical Ultracentrifugation Data." *Methods in Enzymology* (2015).
5. Ivanov, Dmitri, Zinaida Yudina, Rory Johnson, Nikolaos Biris, Alexander Taylor, P. Hart, and Borries Demeler. "The Role of RING Dimerization in the E3 Ubiquitin Ligase Activity of the TRIM5α Retroviral Restriction Factor." *The FASEB Journal* 29, no. 1 Supplement (2015): 888-9.
6. Kim, Haram, Emre H. Brookes, and Borries Demeler. "A performance predictor for UltraScan supercomputer calculations." In *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure*, p. 42. ACM, 2015.
7. Pearson, Joseph, Frank Krause, Dirk Haffke, Borries Demeler, Kristian Schilling, and Helmut Cölfen. "Next-Generation AUC Adds a Spectral Dimension: Development of Multiwavelength Detectors for the Analytical Ultracentrifuge." *Methods in Enzymology* (2015).
8. Yudina, Zinaida, Amanda Roa, Rory Johnson, Nikolaos Biris, Daniel A. de Souza Aranha Vieira, Vladislav Tsiperson, Natalia Reszka et al. "RING Dimerization Links Higher-Order Assembly of TRIM5α to Synthesis of K63-Linked Polyubiquitin." *Cell reports* 12, no. 5 (2015): 788-797.
9. Bhattacharya, Akash, Steven L. Alam, Thomas Fricke, Kaneil Zadrozny, Jaroslaw Sedzicki, Alexander B. Taylor, Borries Demeler et al. "Structural basis of HIV-1 capsid recognition by PF74 and CPSF6." *Proceedings of the National Academy of Sciences*

111, no. 52 (2014): 18625-18630.
10. Swygert, Sarah G., Benjamin J. Manning, Subhadip Senapati, Parminder Kaur, Stuart Lindsay, Borries Demeler, and Craig L. Peterson. "Solution-state conformation and stoichiometry of yeast Sir3 heterochromatin fibres." *Nature communications* 5 (2014).
11. Swygert, Sarah G., Benjamin J. Manning, Subhadip Senapati, Parminder Kaur, Stuart Lindsay, Borries Demeler, and Craig L. Peterson. "Solution-state conformation and stoichiometry of yeast Sir3 heterochromatin fibres." *Nature communications* 5 (2014).
12. Demeler, Borries, Tich-Lam Nguyen, Gary E. Gorbet, Virgil Schirf, Emre H. Brookes, Paul Mulvaney, Ala'A. O. El-Ballouli et al. "Characterization of size, anisotropy, and density heterogeneity of nanoparticles by sedimentation velocity." *Analytical chemistry* 86, no. 15 (2014): 7688-7695.
13. Gorbet, Gary, Taylor Devlin, Blanca I. Hernandez Uribe, Aysha K. Demeler, Zachary L. Lindsey, Suma Ganji, Sabrah Breton et al. "A parametrically constrained optimization method for fitting sedimentation velocity experiments." *Biophysical journal* 106, no. 8 (2014): 1741-1750.
14. Pham, Johnny D., Borries Demeler, and James S. Nowick. "Polymorphism of Oligomers of a Peptide from β-Amyloid." *Journal of the American Chemical Society* 136, no. 14 (2014): 5432-5442.
15. Zhang, Shengnan, Kenneth M. Roberts, and Paul F. Fitzpatrick. "Phenylalanine binding is linked to dimerization of the regulatory domain of phenylalanine hydroxylase." *Biochemistry* 53, no. 42 (2014): 6625-6627.
16. Ramos, Isbaal, Noelia Fernández-Rivero, Rocío Arranz, Kerman Aloria, Ron Finn, Jesús M. Arizmendi, Juan Ausió, José María Valpuesta, Arturo Muga, and Adelina Prado. "The intrinsically disordered distal face of nucleoplasmin recognizes distinct oligomerization states of histones." *Nucleic acids research* 42, no. 2 (2014): 1311-1325.
17. Wright, David W., Ruodan Nan, Gar-Kay Hui, Joseph E. Curtis, Emre H. Brookes, and Stephen J. Perkins. "CCP-SAS-Novel Approaches for the Atomistic Modelling of Small Angle Scattering Data in Biology." *Biophysical Journal* 108, no. 2 (2015): 191a.
18. Rocco, Mattia, and Emre Brookes. "Dynamical Aspects of Biomacromolecular Multi-resolution Modelling Using the UltraScan Solution Modeler (US-SOMO) Suite." In *The Future of Dynamic Structural Science*, pp. 189-199. Springer Netherlands, 2014.
19. Brookes, Emre H. "An open extensible multi-target application generation tool for simple rapid deployment of multi-scale scientific codes." In *Proceedings of the 2014 Annual Conference on Extreme Science and Engineering Discovery Environment*, p. 53. ACM, 2014.
20. Chae, Heejoon, Sungmin Rhee, Kenneth P. Nephew, and Sun Kim. "BioVLAB-MMIA-NGS: microRNA–mRNA integrated analysis using high-throughput sequencing data." *Bioinformatics* (2014): btu614.
21. Smith, Cameron W., Steven Tran, Onkar Sahni, Farhad Behafarid, Mark S. Shephard, and Raminderjeet Singh. "Enabling HPC simulation workflows for complex industrial flow problems." In *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure*, p. 41. ACM, 2015.
22. Shephard, Mark S., and Cameron W. Smith. "HPC Simulation Workflows for Engineering

Innovation." In *Proceedings of the 2014 Annual Conference on Extreme Science and Engineering Discovery Environment*, p. 56. ACM, 2014.

CIPRES-Enabled Science Publications

A complete list can be found here: http://www.phylo.org/index.php/portal/enabled_publications. In 2015, CIPRES enabled 136 scientific publications with 44 additional papers in press for 180 total.

NSG-Enabled Science Publications

The NSG gateway, which went into production in early 2013, has enable 7 peer-reviewed publications so far, and 4 posters in the field of computational neuroscience by NSG users. In addition the NSG team presented two posters (at Society for Neuroscience 2014 annual meeting and XSEDE15) during the year and two talks (at Computational Neuroscience 2015 and NEURON 2015) at workshops. A complete lists is availalbe from http://www.nsgportal.org/citation.html. NSG-enabled publications will be included in our publication count in our 2015-2016 annual report.

SEAGrid-Enabled Publications

A complete list is available from https://seagrid.org/papers/index.shtml. SEAGrid has supported 16 chemistry and material science publications during the reporting period. SEAGrid-enabled publications will be included in our 2015-2016 annual report.

SciGaP-Enabled XSEDE Allocations

Allocation	**Allocation Size (SU)**	**Comments**
SciGaP	200,000	Startup award
UltraScan	208,295 Comet 41,659 Gordon 208,295 Stampede 250,000 (TACC internal allocation)	XRAC award (2015)
NSG	1,251,163 Comet 394,916 Stampede 799,870 Trestles	XRAC award (2015)
CIPRES	4,287,135 Gordon 10,560,677 Comet	XRAC award
Parallel Unstructured Mesh	50,000	XSEDE Consultation led by

Infrastructure: Mark Shephard (RPI), Cameron Smith (RPI)		David O'Neal; Pierce is building a gateway as part of this effort.
Computational support for small angle scattering for advanced analyses of structural data in chemical biology and soft condensed matter: Emre Brookes (UTHSCSA)	75,000	GenApp and Airavata based gateways to support the NSF-funded CCP-SAS gateway.
Continuation Request for XSEDE Research Allocation on Blacklight at the Pittsburgh Supercomputing Center (PSC): Noushin Ghaffari	76,062	ECSS support effort using Apache Airavata
Direct numerical simulations and analysis of compressible and incompressible turbulence: Diego Donzis	681,818	ECSS support effort using Apache Airavata
Gateway-based flow simulations for remote users: Arne Pearlstein	50,000	ECSS support effort using Apache Airavata, SEAGrid, and GenApp
The Science and Engineering Grid (SEAGrid) Gateway: Sudhakar Pamidighantam	6,000,000 (requested)	Request to be made in October 2015

Grants Awarded to Gateway Users and Collaborators

Awardee	**Award #**	**Amount**	**Comment**
Emre Brookes, UTHSCSA	NSF 1265817	$263,926.00	Emre's award actually dates from June 1, 2013, predating SciGaP but not the OGCE collaboration.
Ari Berman, BioTeam	USDA	> $1,000,000	Working with BioTeam's Bhanu Rekepalli to prototype a gateway using

			Airavata and PGA for the USDA.
Amitava Majumdar, SDSC	NSF 1458840	$774,000.00	See http://www.nsf.gov/awardsearch/showAward?AWD_ID=1458840. Awarded August 2015.
Google Summer of Code Students	NA	$44,000	Google pays $5,500 directly to mentored students. Project information available from https://www.google-melange.com/gsoc/projects/list/google/gsoc2015. Use "Airavata" search title.

Other Meritorious Activities

These include promotions, awards, recognition, etc given to team members and collaborators.

1. High school student Haram Kim co-authored an accepted publication to XSEDE15 with Borries Demeler and Emre Brookes.

Operational Metrics

Gateways Using SciGaP Services

SciGaP services are hosted on IU's Quarry Gateway Hosting Service (https://portal.xsede.org/iu-quarry). This list includes only the Gateways using SciGaP and does not include gateways using Airavata software elsewhere. We only report usage for UltraScan.

1. UltraScan: http://www.uslims3.uthscsa.edu/index.php
2. PGA: http://test-drive.airavata.org. This hosts the PHASTA code gateway as well as several demonstration applications.
3. PHASTA: this is an XSEDE ECSS project to support finite element simulations and uses the PGA. Cameron Smith at RPI is the contact.
4. TURBSIM: this is an XSEDE ECSS project to support atmospheric simulations and uses the PGA. Diego Donzis at Texas A&M is the contact
5. UltraScan SOMO: gateway for small angle scattering codes usn GenApp front end developed by Emre Brookes, UTHSCSA
6. BioTeam gateway, using PGA and Airavata to support USDA researchers. Bhanu

Rekepalli is the point of contact.

The PGA is a reference implementation in PHP of the Airavata API that is used for tutorials and interested gateways are encouraged to reuse. We run a persistent version at https://testdrive.airavata.org.

NSG integration is wrapping up and will be reported next year. We also have several gateways supported through XSEDE ECSS activities that are in development.

Number of Users

Gateway	**# of Users**	**Comments**
UltraScan	132 (2014), 90 (2015 to date)	Includes LIMS usage for US, German, and Indian-based researchers. 2014 was a 50% increase over 2013.
CIPRES	7033	Not using SciGaP centralized infrastructure but a useful baseline for future years.
NSG	263	Initial testing for use of SciGaP in being completed.
SEAGrid	40	Not using SciGaP centralized infrastructure but will use SciGaP in future.

Number of Computational Experiments

Gateway	**# of Experiments**	**Comments**
UltraScan	581,278 (2014), 83,679 (2015 to date). Resources used: 2014: 620,459 SUs 2015: 649,120	New multiwavelength and two new international LIMS instances (Germany, India) greatly increased usage. New PMG composite job submission accounting reduces reported number of experiments by up to 50 fold in 2015, so a direct comparison with 2014 is not possible. Actual usage is

		actually higher in 2015 compared to 2014
CIPRES	194,683 computational jobs were submitted to XSEDE.. Approximately 20% failed because of user/input errors.	Not using SciGaP centralized infrastructure but a useful baseline for future years.
NSG	3263 computational jobs were submitted to XSEDE HPC resources from Oct 1, 2014 till Sept 17, 2015	Not using SciGaP centralized infrastructure but a useful baseline for future years.

Application	**# of Users**	**Allocation Used**	**Comments**
UltraScan	132 (2014) 90 (2015 to date)	620,459(2014) 649,120(2015)	Substantial usage of non-XSEDE resources.
BEAST DPPDIV GARLI jModeltest2 MAFFT MigrateN MrBayes PhyloBayes RAxML	1,425 35 507 1,045 381 63 3,157 102 4,190	6,006,266 105,433 758,470 89,290 20,877 981,458 4,147,212 1,068,285 4,707,747	46,579 submissions 239 submissions 6,390 submissions 461 submissions 2,255 submissions 1,083 submissions 47,403 submissions 544 submissions 71,061 submission
NEURON Brian NEST PyNN PGENESIS TVB FreeSurfer	# of active users were about 70 out of the total 263 registered users	1,497,559 145 4935 834 6863 260532 5203	3244 submissions 22 submissions 21 submissions 36 submissions 23 submissions 161 submissions 22 submissions

Resources Used

Gateway	**Resources**	**Comments**
UltraScan	Alamo (UTHSCSA), BCF (UTHSCSA), Comet (SDSC),	Alamo hardware was 100% updated. Jureca, Comet and

	Gordon (SDSC), Jureca (Juelich), Juropa (Juelich), Lonestar (TACC), Stampede (XSEDE), Trestles (XSEDE)	Gordon were added. Juropa and Trestles were deprecated.
SEAGrid	Raven (U Cardiff), Big Red 2 (IU), Comet, Stampede, Trestles, Blacklight and Gordon	Planned for integration with SciGaP in Year 3.
NSG	Comet, Stampede, Trestles	Continuing integration
CIPRES	Comet, Gordon, Trestles (deprecated) TSCC	Scheduled for integration in Year 3.

SciGaP Service Uptime

Airavata Jira Tickets

JIRA Tickets between 10/01/2014 - 09/10/2015

Type	**Created**	**Closed**
All	370	148
Bugs	173	99
Improvements	63	25
New Features	16	4
Epic/Story/Task/Subtask	113	20
Other	5	0

SciGaP Job Matrix

All jobs were run by UltraScan between 10/01/2014 - 09/10/2015

Total Jobs	Success	Failed	Canceled	Other	% Success	% Failure
14987	12822	777	945	443	91.31%	8.68%

* To calculate % Success/Failure, canceled jobs were reduced from total count.

Failure Matrix: All errors reported are for UltraScan operations.

Error Type	**Total Error**
Application Failure	7
GridFTP Connection issue	0
Myproxy Service issue	0
Gram service failure	0
Other failure	32
UNDETERMINED/AIRAVATA_INTERNAL_ERROR	183
FILE_SYSTEM_FAILURE	10
Output is not generated properly : Empty Output	497
Unable to create file	1
The job contact string does not match any which the job manager is handling	16
Path <output/juropa.fz-juelich.de_application.stdout> not found on storage	27
Path <output/juropa.fz-juelich.de_application.stderr> not found on storage	3
Could not invoke service.. Nested exception is org.codehaus.xfire.fault.XFireFault: Couldn't send message	1
TOTAL	777

New Committers and Project Management Committee Members

Committers are contributors who have been granted full access to the Airavata code base. PMC members are committers of merit who have been granted full (binding) voting writes in the project. The following additions were made between September 2013 to August 2014

1. Committers: Nipurn Doshi, Hasini Gunasinghe

2. PMC Members: Supun Nakandala, Shameera Rathnayaka

Number of Downloads of Airavata Software

This information is available from https://repository.apache.org/index.html (login required, Apache ID required, more information available on request). During the period of September 2014-September 2015, Airavata had over 480 IP-unique downloads.

Additional available statistics:

- All Airavata mailing lists: http://airavata.markmail.org/
- Airavata Developer mailing list: http://airavata.markmail.org/search/?q=#query:%20list%3Aorg.apache.incubator.airavata-dev+page:1+state:facets
- Airavata Architecture mailing list: http://airavata.markmail.org/search/?q=#query:list%3Aorg.apache.airavata.architecture+page:1+state:facets
- Apache Airavata GitHub mirrors:
    - https://github.com/apache/airavata
    - https://github.com/apache/airavata-php-gateway
- OpenHub Analysis:
    - https://www.openhub.net/p/airavata
    - https://www.openhub.net/p/airavata-php-gateway