FIll in sections below, then cut and paste into the report on research.gov. Due date: September 28th, 2018. Plan to submit this on September 21st. I use Heading 1, 2, etc for the report sections. I use [TEXT GOES HERE] to indicate where you should put text. Text in red is instructions from research.gov.

Milestones: please see what we submitted to Dan K and address these in the report: https://docs.google.com/document/d/1Z7tuKSUjZCCQpwQOSIq3oURTfnHnWGFWx8-uIP139Aw/edit#heading=h.jch3hpwyme5

For outreach, see https://docs.google.com/document/d/1KsgYZqF3aLU6Vs0Xf1i2Gu9_k5wlW81hxBNgjCt3Fg8/edit#heading=h.v0qfwhj3h1bw

Accomplishments

For NSF purposes, the PI should provide accomplishments in the context of the NSF merit review criteria of intellectual merit and broader impacts, and program specific review criteria specified in the solicitation. Please include any transformative outcomes or unanticipated discoveries as part of the Accomplishment section.

What are the major goals of the project? 8000 char limit.

List the major goals of the project as stated in the approved application or as approved by the agency. If the application lists milestones/target dates for important activities or phases of the project, identify these dates and show actual completion dates or the percentage of completion.

The overall goal of the SciGaP collaboration is to develop a Science Gateway Platform as a service that can be used to power both the funded partner gateways (UltraScan, CIPRES, NSG) as well as the gateway community at large. To meet this goal, we have implemented SciGaP as a multi-tenanted service (currently hosted on IU's Intelligent Infrastructure) that provides a well-defined Application Programming Interface (API) for client gateways. Apache Airavata software powers persistent SciGaP services, along with third party technologies such as RabbitMQ, Apache Zookeeper, Apache Helix, Apache Thrift and Keycloak. Gateways can use SciGaP services through this API to outsource the generic capabilities such as managing users, task execution on supercomputers, workflow management, experiment archiving, and digital archive sharing.

Project Year 5 saw major advances in this platform. This is reflected in the continued increased number of supported gateways, size of XSEDE allocations, and number of supported grants; details are in the attached project metrics document. We highlight our outreach efforts to start new science gateways in Year 5. This is built on our scalable, multi-tenanted approach to gateways, which enables us to quickly spin up gateways that use SciGaP services. Most of these make use of the PHP Gateway for Airavata (PGA), a reference implementation of the SciGaP API that we initially developed to support tutorials but which has taken on a life of its own,

The following are the project's recurring and Year 5 milestones. See also supplemental information in our metrics document.

Recurring Deliverables:

- *Engage in planned outreach activities:* We provide greater detail in the appropriate report sections below. We highlight our PEARC18 tutorial.
- *Conduct yearly gateway community user surveys*: This report was conducted for the

CIPRES gateway community user base. 8,681 survey invitations were distributed, and there were 936 respondents. 83% of respondents reported that CIPRES helped them do something that would be difficult or impossible otherwise, 90% said their work would be slowed or halted without access to these resources. 95% reported the resource benefits their research in a tangible way. Users with very modest levels of use reported CIPRES provides value; many users who consumed less than 100 core hours still planned to publish, or reported publications that were supported by CIPRES.

- *Follow 6-week release pattern for SciGaP software patches:* as in Years 2-4, we have moved to a frequent update model for SciGaP platform services and a less frequent software release model for Apache Airavata. Service releases are organized by branches and tags, with the master branch supporting production services. Repository statistics (commits, contributors, etc) are available from the Airavata GitHub mirrors https://github.com/apache/airavata/, https://github.com/apache/airavata-php-gateway, and https://github.com/apache/airavata-django-portal. The Django portal effort is our replacement for the PHP-based PGA.

Year 5: SciGaP Metrics Summary. See supplemental metrics information for a more thorough discussion of metrics.

- *Recruit more gateways, focus on simplifying integration with documentation, APIs, etc. .* During Year 5, we served 30 client gateways, as detailed in the supplemental metrics document.
- *Support new capabilities needed by evolution of UltraScan, CIPRES, NSG.* During the last year, we improved our job management components to better handle UltraScan's usage requirements. We also completed the integration of our revised user management and authentication system, based on Keycloak.org. The NSG had a requirement for a scalable mechanism for vetting new users. We incorporated a Web of Trust mechanism into the Workbench Framework, as described under Major Activities below. In collaboration with Indiana University's Research Technologies, we developed an integrated LDAP+gateway approach that allows us to do end-to-end, per-user authentication from the browser to the backend resource through SciGaP middleware.
- *Complete Net+ General Availability Phase.* As described in our previous reports, Net+ is better at negotiating academic price reductions with established vendors such as Box than at serving as a business incubator; we need the latter service. We are leveraging the Science Gateways Community Institute's sustainability efforts to meet this milestone. We have also been awarded an NSF Partnerships for Innovation-Technology Translation award that will build on SciGaP.org funded work to work with commercial partners on science gateways. Nancy Maron from the Science Gateways Community Institute will be engaged as a consultant on SciGaP sustainability as part of this award.

What was accomplished under these goals (you must provide information for at least one of the 4 categories below)? 8000 char limit for each of the 4 categories.

For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results, including major findings,

developments, or conclusions (both positive and negative); and 4) key outcomes or other achievements. Include a discussion of stated goals not met.

1) Major activities

Project Wide: Our major activities include the centralization of SciGaP's online services, which in turn supports our ability to scale up to support many gateways simultaneously; 30 gateways are currently using SciGaP services. In Year 5, we collectively supported over 15,000 users, made 35 different computing resources (including both campus and XSEDE resources) available to gateways, supported 24 XSEDE allocations totaling over 42 million SUs, and supported 12 new grants totaling over $10M. We collectively authored 19 journal, conference, and peer-reviewed poster publications, and collected over 280 citations about our cyberinfrastructure. Eight of our publications produced under this grant have at least 8 citations. Our team supported over 1100 science publications.

We co-authored 8 peer-reviewed posters and 2 peer-reviewed papers with SciGaP client gateway providers at PEARC18. The SciGaP team also organized a PEARC18 tutorial, "Science Gateways for New Users" that featured the PI and Co-PIs.

Core development activities included the complete overhaul of the Apache Airavata job management framework, based on experiences gained through the previous four years of operations. Apache Airavata's registry component was also refactored to support better integration with the Sharing Service, enabling us to use the general purpose sharing service capabilities with any internal data object.

We devoted significant effort to upgrading the PGA, which has emerged as an important asset to the project, providing a face to SciGaP services and a starting point for many new gateways. During Project Year 5, we began a complete rewrite of the PGA using the Django framework.

Education efforts play an important role in SciGaP. We continued our graduate independent study course in Fall 2017(14 students) and Fall 2018 (22 students), with an "advanced" version of the course in Spring 2018 (4 students); see http://courses.airavata.org/ for details. We hired two Research Assistants from the Spring 2017 course (Sneha Tilak and Eldho Mathulla) and two hourly employees (Sachin Kariyattin and Stephen Paul Adithela) from the Fall 2017 course. Students in the Spring 2018 course prototyped several advanced features for Apache Airavata. The Fall 2017 course placed increased emphasis on contributions to the Apache Airavata framework. The more stringent requirements resulted in fewer students. We co-authored 1 full paper submission and 2 posters at PEARC18 with students from the course.

UltraScan Gateway: The UltraScan team primarily focused on completing the implementation of a robust multi-wavelength submission method. We identified several performance bottlenecks in the submission process when very large numbers of multi-wavelength datasets for AUC experiments were involved, and addressed these shortcomings, mainly by hardware upgrades and reconfigurations of the local job submission mechanism. Older servers in use at UTHSCSA were not fast enough to handle multiple jobs submitted in rapid succession for multi-wavelength experiments and were ultimately replaced with an XSEDE virtual machine server running on Jetstream set up in collaboration with Eric Coulter of the IU SciGaP team. XSEDE Jetstream allocations are now used for this service. The

service can be scaled to any required performance level, and is thus conveniently scalable for the UltraScan Science Gateway. With the arrival of the Beckman Optima AUC instruments, multi-wavelength experiments are now more common, and our gateway implementation required significant work to adapt to this new environment. A major overhaul of our finite element codes was also undertaken to address the submission of multi-speed experiments. This type of analysis is highly beneficial for systems with broad particle distributions. This also required significant changes in the fitting logic and the submission of experiments. A major effort was made to disseminate the results of this work by holding multiple workshops in the US and abroad to teach rigorous biophysics analysis with our software and Science Gateway. This included courses at the University of Delaware, at Mayo Clinic, at the Hormel Institute in Minnesota, two week-long courses in San Antonio, a 1-week workshop at Christchurch, NZ at the University of Canterbury, and a 2 week workshop at Aalto University in Espoo/Helsinki in Finland. We are currently working on implementing a new Science Gateway for UltraScan at the University of Helsinki and at the University of Lethbridge in Alberta, Canada. Our goal is to expand the work accomplished under this grant from XSEDE also to Compute Canada.

CIPRES Science Gateway: Major activities of the CIPRES Science Gateway include providing advice, driving requirements, and coding support for Airavata development. CIPRES developers worked to develop a scalable method for vetting new users. This is a driving requirement of the NSG. The design provides a capability that allows members of a science community to perform distributed approval of new gateway account creation. A software package was created for managing a web-of-trust. The software uses the GNU Privacy Guard (GnuPG) to manage a network of approvals in the GnuPG trust database. Existing gateway users can use a web interface to pre-approve new users that they deem to be trustworthy. The gateway can check the database for pre-approved users to allow or deny account registration requests.

This software can be installed as static html and CGI scripts under a web server. Integration with a gateway can be achieved by providing access, through either a flat text file or read-only mysql database access, to user email addresses and hashed passwords. Integration with the CIPRES Workbench code was accomplished by modifying the registration page to check the trust database. The trust check will be integrated into the NSG registration form by having the backend code check the trust database, then on success, notifying the gateway administrators of approval status and sponsoring member. This service can be implemented by other subscribing gateways as required.

Neuroscience Gateway: Major activities of the Neuroscience Gateway (NSG) include adding new features, installing and implementing tools and pipelines as requested by the user community, providing feedback regarding specific NSG requirements for the SciGaP project, outreach to the user community, and involving undergraduate and high school students for internship. Examples of new activities include 1) installing and making available the primarily Matlab based EEGLAB tool (PI Scott Makeig, UCSD)- via NSG for processing EEG data; 2) installing and making available the DynaSim tool via the NSG (PI Jason Sherfey , Boston University) - an open-source MATLAB/GNU Octave toolbox for rapid prototyping of neural models, 3) ongoing work on “web of trust” feature where “trusted NSG users” can validated new users to get NSG account 4) enabling workshops and tutorials to utilize NSG. Multiple

workshops, and poster presentations were done as a part of the outreach, and two female high school students participated as student interns during the summer of 2018. T NSG team also presented papers and posters, and did workshops and research talks at the Society for Neuroscience Annual meeting in Washington D.C. in November 2017,at the Computational Neuroscience Annual meeting in Seattle in July, 2018, and at the PEARC18 conference in Pittsburgh in July, 2018.

2) Specific Objectives

Our project's specific objective this year was to broaden the adoption of SciGaP services. We pursued this objective through outreach efforts that targeted gateways that a) support local university clusters,and b) support specific scientific communities, typically backed by specific scientific applications. Based on our experience, we see greater uptake in scientific collaborations and will make these the focus in Year 6. This takes advantage of previous years' work to bring a multi-tenanted gateway hosting capability in Apache Airavata and the ability to clone PGA instances to quickly stand up new gateways. We have also realized that the PGA reference implementation gateway is too limited to provide a basis for many of the new clients we added and too rigid to support the broader use of the Sharing Service, so we began a redesign of the PGA based on the Django framework.

3) Significant results, including major findings, developments or conclusions

Project Wide: During this period we were able to significantly increased our support for additional gateways beyond the original three gateways, thus demonstrating the technical viability of our approach as well as the market for SciGaP-style science gateways.

Apache Airavata and the PGA: The central SciGaP concept of hosted services was amply demonstrated in Year 5, with over 30 gateways simultaneously using SciGaP services. Core improvements to Airavata included the overhaul of the job management infrastructure and the refactoring of the registry metadata management system, in part to support better integration with the Sharing Service. As described above, we began the replacement of the PHP-based PGA with the next generation, Django-based PGA. This will, among other improvements, give us better integration with the Sharing Service for fine-grained sharing of resources and better user-interface support for groups. The increased demand placed on our job management infrastructure by the increasing number of client gateways led us to redesign and reimplement the job management infrastructure using Apache Helix; this work was begun in the previous year as an outgrowth of the Spring 2017 Science Gateway Architectures advanced class. Finally, in a collaboration enabled by supplemental funds from the XSEDE project, we worked with Eric Coulter to develop and deploy a flexible mechanism for deploying virtual private clusters for numerous gateway clients (including UltraScan and SEAGrid) on the Jetstream cloud.

UltraScan Gateway: The switch to newer hardware solved all of our performance bottlenecks and provided a robust solution for multi-wavelength job submissions. The software and hardware improvements together now assure a smooth workflow for such experiments, and

have proven to be a solid production environment. Upgrades in the finite element codes and fitting software for our software to support modeling of multi-speed experiments revealed an unanticipated discovery: While simulated data incorporating speed-dependent rotor stretch could be perfectly fitted, fits of actual experiments failed to reproduce the meniscus positions accurately, causing fits to produce poor results. This was traced to a speed dependent compression of the solution column. In retrospect, this physical effect is not unexpected, but we were not aware that the new instrumentation had actually sufficient radial precision to measure these changes accurately. As a consequence of this discovery we are therefore now pursuing this effect to develop methods for measuring compressibility of solvents with this instrument. The dissemination and training efforts have paid off by increasing usage of our Gateway and have increased interest in analytical ultracentrifugation and enhanced knowledge about solution biophysical methods.

NeuroScience Gateway: NSG continues to add new tools such as the EEGLAB and DynaSim tools both of which are Matlab based tools. Given that Matlab was made available in the year 4, we now see multiple Matlab based tools becoming part of NSG's tools based on request from the user community. Due to Matlab and EEGLAB we now see more data processing types jobs are run via NSG in addition to large scale neuronal models and parameter sweep type simulations. As a result more high throughput and cloud computing mode of resources are being used by NSG users. In this past year work was done on "web of trust" implementation. Currently a new NSG user is manually vetted by the three NSG team members before a new account is approved. The new "web of trust" work will allow trusted NSG users to approve new NSG users with whom they are professionally connected and know them as neuroscientists who will use NSG. NSG is being used more and more for teaching and education both in the US and in EU - example of these include NEURON summer course, the NIH funded computational neuroscience training at University of Missouri,the NSF funded CyberTraining project for Data-Intensive Neuroscience Learning and Research, and the EU Human Brain Project (HBP) hosted Computational Neuroscience course and the HBP School on the Brain Simulation Platform.

CIPRES Science Gateway: The CIPRES project continues to support web applications that are significant REST consumers. We have implemented two new codes - IQ Tree, and Exabayes. as RESTful services, the former for the Mesquite/Zephyr project. We collaborated with the COSMIC2 gateway to create tools to manage storage and transfer of large data files through Globus On-line. The distributable software underlying the CIPRES Gateway has been implemented by the COSMIC2 gateway, which was recently awarded a grant from NSF ABI, and the CyNeuro gateway, a neuroscience-oriented gateway at the University of Missouri, which is currently supported by NSF OAC.

4) Key outcomes or other achievements.

Our key outcomes and achievements are described above.

Other notable achievements: Co-PI Demeler was awarded the prestigious Canada 150 Research Chair for Biophysics and is in the process of relocating his laboratory to the University of Lethbridge and the University of Montana. He also completed a 2 month Erskine Fellowship as visiting professor at the University of Canterbury in Christchurch, NZ, and spent a two-week

Fulbright-sponsored visit at Aalto University in Espoo/Helsinki to teach analytical ultracentrifugation and initiate the expansion of the UltraScan Science Gateway to Finland. A new LIMS server and science gateway has been deployed at the University of Helsinki (Taito) and is expected to be in production mode at the end of September 2018. .

CIPRES received 319,000 job submissions (an increase of 16% over the previous year) and supported its 4,500th publication in mid- 2018. We expect to support 1,200+ publications in 2018 and total publications will exceed 5,200 by year's end.

Number of NSG users continue to grow and become more diverse, compared to the neuroscience researchers that were interested primarily in spiking neuronal simulation tools since its inception in 2013. In the recent years increasingly neuroscientists who are doing data processing are utilizing NSG. As a result NSG is contributing towards data driven brain research. NSG has also become a dissemination platform where neuroscientists are releasing their newly developed tools for the neuroscience community; example of these include DynaSim and the Human Neocortical Neurosolver (PI Stephanie Jones, Brown University).

What opportunities for training and professional development has the project provided? 8000 char limit.

Describe opportunities for training and professional development provided to anyone who worked on the project or anyone who was involved in the activities supported by the project. "Training" activities are those in which individuals with advanced professional skills and experience assist others in attaining greater proficiency.

If the research is not intended to provide training and professional development opportunities or there is nothing significant to report during this reporting period, please check "Nothing to Report" if applicable.

Project Wide: We organized a team-wide PEARC18 tutorial and conducted several additional local efforts described below.

Apache Airavata: The IU team taught Fall 2017 (14 students), Spring 2018 (4 students), and Fall 2018 (22 students) graduate-level course in science gateway technologies. We supervised four Google Summer of Code students. Two IU students who took the Spring 2017 course worked as research assistants for Spring 2018 semester before completing their Masters' degrees; they worked as hourly employees for Fall 2017. Dimuthu Upeksha Wannipurage joined the team as a core Apache Airavata developer and took the lead on the job management component redesign. Wannipurage originally worked with the IU team as a Google Summer of Code student in 2015.

UltraScan Gateway: As in years past, we provided workshops and training opportunities throughout the year, starting with a workshop for industry hosted by the University of Delaware, followed by two week-long workshops in San Antonio in December 2017 and May 2018, and at Mayo Clinic and Hormel Institute in Minnesota, a week-long hands-on workshop in Christchurch, and a two-week Fulbright sponsored workshop at Aalto University in Helsinki, Finland. A class (BIOC5085 "Hydrodynamic Methods") was taught at UTHSCSA in May of 2018, which is open to students from other institutions, and was attended by investigators from Albert Einstein College of Medicine and Elion Labs, Boulder, Colorado. Other students attended from the University of Texas Health Science Center and the Texas Biomedical Research Institute..

Another opportunity for training was realized through multiple Google Hangout sessions, where we helped new users in Europe, India, New Zealand and Australia to become familiarized with the Science Gateway by being able to provide customized training to multiple users not in the US. P.I. Demeler supervised the work of three postdocs and two graduate students who contributed to this work.

CIPRES Science Gateway: CIPRES was presented as part of a Gateway tutorial at the PEARC2018 conference.

Neuroscience Gateway: We have presented the NSG at multiple workshops to train users about how to use the gateway and also to collect feedback from users regarding features that users wanted for the future. NSG workshop was hosted at the Society for Neuroscience (SFN) annual meeting in November 2017 in Washington D.C.. This NSG workshop was a half day workshop and included presentation by NSG users as well as talks by some of the neuroscience tool developers from US and EU. About 30 attendees attended the workshop. We hosted a NSG workshop in July, 2018 at the Computational Neuroscience Annual meeting in Seattle, WA. This workshop was attended by about 25 people and talks were given by computational neuroscientists from USA and Canada. As a part of SDSC's Research Experience for High School (REHS) students program, two high school female students did internship at SDSC during the summer of 2018.

How have the results been disseminated to communities of interest? 8000 char limit.

Describe how the results have been disseminated to communities of interest. Include any outreach activities that have been undertaken to reach members of communities who are not usually aware of these research activities, for the purpose of enhancing public understanding and increasing interest in learning and careers in science, technology, and the humanities.

We engaged in the following outreach activities

- We presented a team-wide tutorial at PEARC18, which targeted new science gateway users.
- The IU team presented two booth presentations at SC17.
- PI Pierce, Co-PI Marru, and other IU team members presented five short papers on SciGaP efforts at Gateways 2017.
- Pierce, Marru, Co-PI Miller, and other community members co-authored and presented "Towards a Science Gateway Reference Architecture" at the International Workshop on Science Gateways 2018.
- Pierce, Marru, other IU team members, and collaborators presented ten posters at PEARC18.
- Co-PI Demeler visited Aalto University in Helsinki, Finland as a Fulbright specialist and he was also e sponsored by an Erskine visiting professorship from the University of Canterbury in Christchurch, New Zealand, to train students in analytical ultracentrifugation, and to train new users in hydrodynamics by using the UltraScan/Science Gateways from abroad. The purpose of this effort is to bring new users to the Science Gateway from foreign countries, and build up AUC expertise with our foreign user base.

- Co-PI Demeler gave the following talks and workshops: 6/18: From Experimental Design to Data Analysis: Problem solving with AUC. Aalto University, Department of Bioproducts and Biosystems, Helsinki, Finland; 6/18: Next Generation Analytical Ultracentrifugation: Measuring macromolecular interactions with multi-wavelength detection. Aalto University, Department of Bioproducts and Biosystems, Helsinki, Finland; 5/18: Analytical Ultracentrifugation Workshop and UltraScan training. The University of Texas Health Science Center, San Antonio, Texas; 4/18: Analytical Ultracentrifugation Workshop and UltraScan training. University of Minnesota, Hormel Institute, Austin, MN; 3/18: Next Generation AUC: Adding a spectral dimension to analytical ultracentrifugation. Invited Seminar, Dept. of Biological Sciences, University of Canterbury, Christchurch, New Zealand; 3/18: Solution Studies of Nanoparticles, Quantum Dots and Metal Nanoclusters, Invited Seminar, Dept. of Chemistry, University of Canterbury, Christchurch, New Zealand; 2/18: Analytical Ultracentrifugation Workshop and UltraScan training. University of Canterbury, Christchurch, New Zealand; 12/17: Analytical Ultracentrifugation Workshop and UltraScan training. The University of Texas Health Science Center, San Antonio, Texas; 11/17: Analytical Ultracentrifugation Workshop and UltraScan training. Mayo Clinic, Rochester, Minnesota; 11/17: Survey of Multi-Wavelength AUC Applications. University of Delaware, 11/17: Advanced Analytical Ultracentrifugation Workshop, Newark, Delaware; 11/17: Introduction to Experimental Design. University of Delaware, Advanced Analytical Ultracentrifugation Workshop, Newark, Delaware; 11/17: Hydrodynamic Calculations with the UltraScan SOlution MOdeler (US-SOMO) Software Suite. University of Delaware, Advanced Analytical Ultracentrifugation Workshop, Newark, Delaware; 09/17: Next Generation AUC: Multi-wavelength detection. University of Western Ontario, Canada, Invited Lecturer; 09/17: Next Generation AUC: Multi-wavelength detection. Atlanta State University, Georgia; 08/17: Measuring Interactions in the Solution Phase: Adding a Spectral Dimension to Analytical Ultracentrifugation. University of Lethbridge, Alberta, Canada; 8/17: Solution Studies at the Nanoscale with Analytical Ultracentrifugation. University of Montana, Missoula, Montana.
- The CIPRES team gave the following presentations: Mark Miller spoke as a panelist on “Sustaining Digital Resources” at the Annual Meeting of the Society for Systematic Biologists in June 3, 2018, and served as a presenter in the Tutorial “Introduction to Science Gateways” at the PEARC18 conference in Pittsburgh, July 22, 2018.
- The NSG team gave the following presentations: NSG was used by the students attending the Computational Neuroscience Training at the University of Missouri, June, 2018; NSG was used as a part of a EU HBP workshop titled, "Neuroscience for ICT: Applications to Computation and Robotics", in KALKSCHEUNE BERLIN, GERMANY, July 4-6, 2018; NSG was used by the students of the course " Neuroscience Simulation" taught by EPFL researchers, May, 2018; Neuroscience Information Framework (NIF) organized webinar on the Neuroscience Gateway Project, Subhashini Sivagnanam, September 29, 2017; K. Yoshimoto, "Neuroscience Gateway - Scalable infrastructure for computationally-intensive cognitive neuroscience", Tutorial at CogSci18, Madison, WI July 25-28, 2018; S. Sivagnanam, K. Yoshimoto, T. Carnevale, A. Majumdar, "The

Neuroscience Gateway - Enabling Large Scale Modeling and Data Processing in Neuroscience," Practice & Experience in Advanced Research Computing PEARC18, Pittsburgh, PA, July 22-26, 2018; A. Majumdar, S. Sivagnanam, K. Yoshimoto, T. Carnevale, "Neuroscience Gateway - Enabling Large Scale Simulations and Data Processing and Dissemination of Neuroscience Tools/Software", Poster, Organization of Computational Neuroscience (CNS) Annual Conference, Seattle, WA, July 13-18, 2018; S. Sivagnanam, "Enabling computational modeling and big data analysis through Neuroscience Gateway," Neural Interfaces Conference, Minneapolis June 2018.; S. Sivagnanam, A. Majumdar, K. Yoshimoto, T. Carnevale, "Neuroscience Gateway: Enabling Easy Path to Supercomputing for Neuroscience Research and Education," Poster, Neural Interfaces Conference 2018, Minneapolis, MN, June 25-27, 2018; A. Wagner, K Pezzoli, A. Majumdar, J. Bottum, N. Wilkins-Diehr, "Science Gateways and their impact on research and scholarship nationally and internationally," Internet2, Global Summit, May 6-9, 2018, San Diego, CA; A. Majumdar, S. Sivagnanam, K. Yoshimoto, T. Carnevale, "Neuroscience gateway - enabling easy path to supercomputing for neuroscience research and education," Poster BRAIN Initiative PI Meeting, Bethesda, MD, April 9-14, 2018; A. Majumdar, S. Sivagnanam, K. Yoshimoto (UCSD), N.T. Carnevale (Yale U.), A. Peyser (Jülich Supercomputer Center), "High Performance Computing (HPC) Resources for Parallel Simulations and Data Analysis: NSG and HPAC," Satellite Workshop, Society for Neuroscience Annual Meeting 2017, Washington D.C., November 2017; T. Carnevale, A. Majumdar, S. Sivagnanam, K. Yoshimoto, "The Neuroscience Gateway Portal: high performance computing for neuroscientists", Poster, Society for Neuroscience Annual Meeting, Nov 10-15, 2017, Washington D.C.

What do you plan to do during the next reporting period to accomplish the goals? 8000 char limit

Supporting files. You may upload pdf files with images, tables, charts, or other graphics in support of this section. You may upload up to 4 pdf files with a maximum file size of 5 MB each.

During the next reporting period, we expect to significantly increase the impact of SciGaP supported gateways, building on collaborations initiated through XSEDE Science Gateways and Science Gateways Community Institute support requests. We expect overall usage to increase, and we will also focus on increasing scientific and educational impact. We expect to move the revised job management component into full production, completing extensive testing with UltraScan. The new Django-based PGA is also planned to go into full production.

Financial sustainability is a crucial consideration for the project. Several new NSF and NIH projects have been awarded to team members as a direct result of the SciGaP project, as summarized in the supplemental metrics document. We are also planning to build on the scientific collaborations we have developed with numerous client gateways to pursue financial sustainability through co-funding efforts.

NSG and CIPRES team members, jointly with undergraduate students, will work to implement additional features such as allowing users to edit already uploaded files, allowing

bungling of jobs for tools such as EEGLAB, allowing users to open VNC connection to cloud resources etc.

Products

For NSF purposes, the PI should include and discuss in the Product section the goals associated with data management and access and note any significant changes in them, as well as specific plans for dissemination of data, software and other digital research products. When you report any of these items, please include any available identifiers and whether and how these products can be accessed or shared.

INSTRUCTIONS - List any products resulting from the project during the reporting period.

If there is nothing to report under a particular item, please check, "Nothing to Report" if applicable.

Your Output Summary for this Reporting Period:

Publications

For NSF purposes, each category of publication should identify any associated data, software, other supplementary material and their appropriate identifiers. "Other publications, conference papers and presentations" should include other "non-reviewed" publications, conference papers, and presentations.

Technologies or techniques

Websites

1. https://scigap.org/: general project information
2. http://airavata.apache.org/: Apache Airavata website
3. http://courses.airavata.org/

Other products, such as data or databases, physical collections, audio or video products, software or NetWare, models, educational aids or curricula, instruments, or equipment

1. https://github.com/SciGaP/: Project code repository.
2. https://github.com/airavata-courses: course website
3. https://github.com/apache/airavata: GitHub mirror for Apache Airavata source code.

Supporting Files

You may upload pdf files with images, tables, charts, or other graphics in support of this section. You may upload up to 4 pdf files with a maximum file size of 5 MB each.

[TEXT GOES HERE]: Probably nothing needed here.

Participants

For NSF purposes, for separately submitted and awarded collaborative proposals, the PI should report progress on his/her institution's portion of the collaborative effort only.

In each of the subsections below, note which collaborators or contacts are involved in data contribution and/or management.

If there is nothing significant to report during this reporting period, please check "Nothing to Report", if applicable.

* Required fields

What individuals have worked on the project?

Name	Most Senior Project Role	Nearest Person Month Worked
Marlon Pierce	PI	2
Borries Demeler	PI	3
Gary Gorbet	Programmer	12
Mark Miller	PI	2
Amit Majumdar	Co-PI	2
Paul Hoover	Programmer	2
Kenneth Yoshimoto	Programmer	5
Suresh Marru	Co-PI	6
Sudhakar Pamidighantam	Senior Staff	6
David Reagan	Programmer	2
Raminder Singh	Programmer	3
Lahiru Gunathilake	Graduate Student	6
Eroma Abeysinghe	Project Coordinator	3
Shameera Yodage	Graduate Student	6

What other organizations have been involved as partners?

All Participants: XSEDE, Science Gateways Community Institute

UltraScan Gateway: Juelich Supercomputing Center (Unicore integration in Airavata, and colocation services for LIMS server), Indian Institute of Science, Bangalore, India, LaTrobe University, Melbourne, Australia

NSG: Yale University; University College London, UK

CIPRES Science Gateway: NC State University, American Museum of Natural History, Oregon State University, J.Craig Venter Foundation, University of Wisconsin, University of Missouri.

Have other collaborators or contacts been involved?

Nancy Wilkins-Diehr's S2I2 Science Gateways Community Institute. Apache Software Foundation. Von Welch, Trusted CI

Some significant collaborators or contacts within the recipient's organization may not be covered by "What people have worked on the project?" Likewise, some significant collaborators or contacts outside the recipient's organization may not be covered under "What other organizations have been involved as partners?"

Impacts

INSTRUCTIONS - This component will be used to describe ways in which the work, findings, and specific products of the project have had an impact during this reporting period.

For NSF purposes, include, where appropriate, discussion of data resources and the acquisition of data skills. Include the emergence of new career paths, such as data scientists, or new disciplines.

If there is nothing significant to report during this reporting period, please check "Nothing to Report" if applicable.

Please make sure to read all instructions including NSF specific instructions, which can be found in the following link:

Required fields

What is the impact on the development of the principal discipline(s) of the project? *8000 char limit.*

Describe how findings, results, techniques that were developed or extended, or other products from the project made an impact or are likely to make an impact on the base of knowledge, theory, and research and/or pedagogical methods in the principal disciplinary field(s) of the project.

Project Wide: By unifying the common features of three different gateways into a core set of persistent, multi-tenanted services, SciGaP will greatly improve the ability of all its client gateways to offer reliable, sustainable services. As described in our metrics, our publications on SciGaP infrastructure and gateways have been cited over 280 times, and our project has h-index of 8, according to tabulations from Google Scholar. The paper, "Towards a Science Gateway Reference Architecture", was co-authored by SciGaP team members, members of the Science Gateways Community Institute, and other members of the community. This paper presents considerations for the future of the science gateway field.

Apache Airavata: We continued work to position Apache Airavata as a tool for undertaking core cyberinfrastructure and distributed computing research as well as an operational system; this is reflected in the IU courses offered through the School of Informatics and Computing as

well as our publications, listed elsewhere in the report. We continue to champion the Apache Software Foundation's open source, open community ideals as a way for federally funded software systems to be transparent, accountable, and sustainable.

UltraScan Gateway: In addition to the field of computer science, our contributions benefit the field of hydrodynamics, especially investigators in biomedical applications, material science, nanotechnology and structural, functional, and molecular biology. The new methods available now through the UltraScan gateway provide unsurpassed detail in the analysis by leveraging high-performance computing through the gateway infrastructure. These improvements are particularly important to the data analysis of data from the new multiwavelength detector, because of the sheer data density that cannot be handled without parallel processing. Additional analysis methods investigate particle size and shape distributions, as well as density and hydrodynamic radius.

CIPRES Science Gateway: In addition to advancing the science and practice of science gateway development, CIPRES accelerates progress in virtually every field of biology, from Virology to Phylogeography.The methods made available through the CIPRES Gateway provide users around the world with access to parallel codes for computationally intensive sequence alignment and tree inference problems. The access CIPRES provides speeds the analysis of results, and changes the landscape of what of which problems are computationally tractable for any given laboratory. The CIPRES Gateway has made resources available to more the 1100 institutions (Universities, Museums, Botanical Gardens, Governmental Organizations, Research Institutes) in 86 countries on 6 continents, including all 31 EPSCOR states/territories.

Neuroscience Gateway: The project has allowed computational and cognitive neuroscience researchers access to HPC/HTC/cloud resources via an administratively and technologically streamlined environment for uploading models, specifying job parameters, querying running job status, receiving job completion notices, and storing and retrieving output data. This has allowed researchers to model high-dimensional parameter space exploration, study models that involve stochasticity, process data using software such as Freesurfer, EELAB, Matlab etc.,simulate models with large neuronal networks, run parameter sweep studies, run BluePyOpt - the Blue Brain Python Optimization Library, etc. NSG went into production in early 2013 and since then it has about 750 users from about 20 different countries.

What is the impact on other disciplines? 8000 char limit

Describe how the findings, results, or techniques that were developed or improved, or other products from the project made an impact or are likely to make an impact on other disciplines.

As described previously, a science gateway is cyberinfrastructure that directly impacts science. Gateways supported by the SciGaP award have contributed to over 1100 scientific publications in the last year. This is in addition to the core cyberinfrastructure publications by our team funded by this grant, which have been cited over 280 times.

Science gateways also support multidisciplinary work, enabling scientists in one domain to use tools from another domain that the find that they need for their research. The general purpose software and infrastructure we are developing in SciGaP, while a subject of cyberinfrastructure research itself, is also being built to provide an operational system that will

be the platform for new science gateways in other scientific domains in later project years.

What is the impact on the development of human resources? 8000 char limit

Describe how the project made an impact or is likely to make an impact on human resource development in science, engineering, and technology.

Our approach to open source, open community development has created a pipeline of new developers for science gateways through the Science Gateway Architectures course and the Google Summer of Code program as well as enabling and acknowledging the contributions of peer cyberinfrastructure developers. SciGaP developers, by working in the open via Apache Airavata and GitHub, can readily demonstrate their skills to future employers or advisors. We added four Apache Airavata committers and two Project Management Committee (PMC) members in this project year. The committers were all students from the IU Science Gateway Architectures course, and one of the new PMC members is an alumnus of the Google Summer of Code.

All team members actively work with both graduate and undergraduate students. The IU team has taught basic science gateway concepts to over 70 students at IU during this reporting period. The undergraduate student supported as part of CIPRES during the past year is now a graduate student at the Hopkins Marine Laboratory of Stanford University. The UltraScan group has reached out to high schools through institutional liaisons with local high schools and consistently recruited high school students to engage them in computational science, physics and chemistry disciplines. High School students did summer internships with the NSG team and the project work can influence their field of study for undergraduate education and college selection.

What is the impact on physical resources that form infrastructure? 8000 char limit

Describe ways, if any, in which the project made an impact, or is likely to make an impact, on physical resources that form infrastructure, Including physical resources such as facilities, laboratories, or instruments.

Science gateways are the infrastructure that provide science-centric views of scientific computing infrastructure. At present, more users access the XSEDE compute resources through Science Gateways than through conventional command line logins. The CIPRES Science Gateway and the Neuroscience Gateway are important parts of this trend toward simplified access through gateways, and the role of these gateways in bringing new users to XSEDE has been documented by the XSEDE project.The UltraScan gateway and its spinoffs have had a significant impact on the ability manage, analyze and understand data produced by scientific instruments. Our challenge in the upcoming years of SciGaP is to extend these successes to enable new gateways providing access to new resources.

What is the impact on institutional resources that form infrastructure?

Describe ways, if any, in which the project made an impact, or is likely to make an impact, on institutional resources that form

infrastructure.

UltraScan Gateway: The developments for the UltraScan gateway significantly benefit core facilities and individual laboratories because they allow parallelized data analysis of multiple experiments simultaneously. This dramatically reduces analysis time. The LIMS backend significantly improves accuracy and the high-resolution detail obtained from our parallel methods allow core facilities like CAUMA to provide better service and higher information content to their investigators. Additional gateway instances are easily created thanks to the open source software stack of Airavata and new LIMS servers can be rapidly deployed to facilitate service to geographically remote sites (example: University of Lethbridge, Canada) and those who wish to integrate their own supercomputers (example: Aalto University/University of Helsinki, Finland).

CIPRES Science Gateway: The CIPRES Gateway benefits researchers at many institutions by supporting experiments that could not easily be conducted with local institutional resources. Often the need for resources at a given institution is confined to a single laboratory.he aggregate consumption of CIPRES resources at many institutions is less than 50,000 core hours per year. In such cases it is not cost-effective to maintain a cluster for on-site usage. The computational resources accessed through CIPRES make individual analyses faster, and allow multiple analyses to be run simultaneously, and provide “on-demand” clusters available to institutional users when needed. There is a tremendous benefit from economy of scale when the compute resources are constantly in full use, due to the aggregated demand across 1000 institutions. The result is a significant savings in local infrastructure and system administration costs.

Neuroscience Gateway: The NSG provides benefits to many institutions where computational neuroscience research is being pursued by graduate students and researchers dealing with complex and large neuronal models. These simulations require access to HPC/HTC/cloud resources and proper installation of optimized neuronal simulation tools on these resources and easy way to use the tools on the resources and retrieve output results. The NSG, by facilitating these, is allowing researchers from many institutions to carry out computational and cognitive neuroscience research and data processing without having to make HPC/HTC/cloud resources and associated expertise available within their institutions. NSG has also become a dissemination platform for neuroscientists. Neuroscientists are making their tools, pipelines and libraries available via NSG for the broader neuroscience community. Examples of these include DynaSim, Human Neocortical Neurosolver etc.

What is the impact on information resources that form infrastructure?

(Blank)

What is the impact on technology transfer? *8000 char limit*

Describe ways in which the project made an impact, or is likely to make an impact, on commercial technology or public use.

SciGaP's Apache Airavata software is open source software that, like all Apache-licensed software, can be adapted to commercial use. Moreover, as an Apache Software Foundation project, Airavata is owned by the foundation rather by the SciGaP team members' home institutions. This clarifies governance issues such as ownership of contributions and control of the project. Our objectives in Year 5 will be to explore supplemental funding streams by partnering with universities and commercial companies.

What is the impact on society beyond science and technology? *8000 char limit*

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world.

By bringing open source, open governance methodologies to cyberinfrastructure development, we hope to provide better conduits between academic communities and commercial software development. By working in an open, public forums, we give students and junior programmers a chance to demonstrate their programming skills, develop their problem solving abilities, as well as their abilities to interact productively as part of a team. But we also are engaging in the other direction, the SciGaP infrastructure is intended to simplify the process of Gateway development. As a result, the project will lower the barrier, and encourage the rendering of thoughts and expertise of non-academic developers and architects to SciGaP problems through Apache Airavata's architecture mailing list.

Changes

INSTRUCTIONS -

The PI is reminded that the grantee is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction. See agency specific instructions for submission of these requests.

If not previously reported in writing to the agency through other mechanisms, provide the following additional information or state, **"Nothing to Report",** if applicable:

* Required fields

Notifications and Request

For more information on Grantee Notifications to and Requests for approval from the National Science Foundation, please visit the Notifications and Requests section in FastLane or refer to Exhibit II-1 of the Award and Administration Guide (AAG).

Changes in approach and reasons for change 8000 char limit

Actual or Anticipated problems or delays and actions or plans to resolve them 8000 char limit

[TEXT GOES HERE]:

Changes that have significant impact on expenditures *8000 char limit*

[TEXT GOES HERE]:

Significant changes in use or care of human subjects 8000 char limit

[TEXT GOES HERE]:

Significant changes in use or care of vertebrate animals 8000 char limit

[TEXT GOES HERE]:

Significant changes in use or care of biohazards 8000 char limit

[TEXT GOES HERE]: