Project Description

CC*IIE Networking Infrastructure: Accelerating science, translational research, and collaboration at the University of Pittsburgh through the implementation of network upgrades.

Introduction

The University of Pittsburgh is a nonsectarian, coeducational, state-related, public research university. It is internationally respected as a center for learning and research. With an enrollment of nearly 35,000 students, the University is one of the largest institutions of higher learning in Pennsylvania. More than 13,300 faculty, research associates, and staff support the needs and interests of the University. University-related spending is 1.74 billion dollars annually, providing an important economic impact on the area economy.

Information technology is an integral part of the academic and research mission of the University of Pittsburgh. The University's central IT organization, Computing Services and Systems Development (CSSD), provides the technology infrastructure, resources, and support for services utilized by nearly every student, faculty, and staff member.

The Center for Simulation and Modeling (SaM) at the University of Pittsburgh is dedicated to supporting and facilitating computational-based research. SaM supports Pitt's commitment to multi-disciplinary research by enabling the application of high-performance computing to the solution of problems through software, hardware resources, training, and workshops. Faculty and staff working within the center come from a wide range of disciplines including astronomy, biology, chemistry, economics, engineering, health and medicine. SaM consultants are the XSEDE campus champions and provide consultation and programming support for the use of the University high performance computing resources. Over 200 faculty and researchers participate in the shared HPC work environment representing a cross-section of the University spanning 36 different departments.

The University's high performance computing infrastructure is housed and operated at the CSSD-managed University data center. Structured to support a diverse set of computational needs across the research community, the HPC clusters provide over 8,000 cores in over 400 nodes to research projects requiring high-end computational processing power. The Center for Simulation and Modeling (SaM) coordinates activities and provides direct consultation on computational projects, assistance with code optimization and parallelization and technical expertise to participants. CSSD's dedicated HPC systems engineers provide systems and networking support for HPC infrastructure. CSSD also operates PittGrid, a Condor-based grid computing systems that harvests CPU cycles and provides support for grid-based research and code execution. *High performance computing systems, SaM, CSSD, and the University data center are key components of the Campus Cyberinfrastructure plan.*

Background

CSSD built and maintains PittNet as a multi-service, 10Gbps backbone network providing IP services to over 80 buildings in Pittsburgh's Oakland area. Standard building connections are 10Gbps with 1Gbps port service to over 43,000 ports. The campus core is comprised of 17 nodes distributed throughout campus providing MPLS/BGP services for building connections based on a VLAN/ZONE-centric architecture. Firewall services are integral to the architecture with 100% of PittNet ports protected, see Fig.1.

The University's Data Center in nearby Harmarville is operated by CSSD and houses enterprise, academic, web, and HPC clusters. The data center is connected by 2x10Gbps service supporting combined traffic to firewalled zones. The data center also houses CSSD's Network Operations Center, providing 24x7x365 support, monitoring, and troubleshooting.

CSSD operates numerous WAN-based services connecting the University's four regional campuses and off-campus facilities in the metro area. Campus edge services include 10GbE connections to 3Rox (Three Rivers Optical Exchange - regional GigaPoP) for routes to the major R&E networks and to KINBER's PennREN - the new statewide R&E network in Pennsylvania. Researchers use this connectivity for access to XSEDE, scientific gateways and other networks. Pitt is an active member of the research IT community with participation in Internet2 and memberships with InCommon and KINBER. *PittNet is a key component of the Campus Cyberinfrastructure plan.*

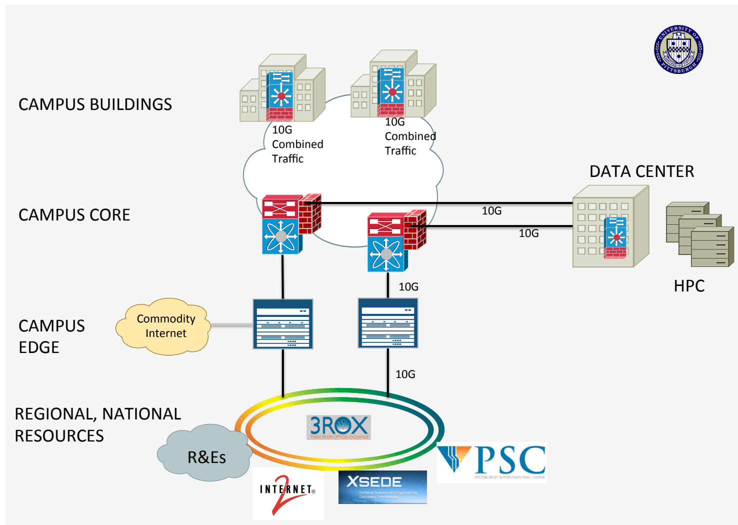

Fig 1 - PittNet Reference Architecture

<u>Challenges/Needs/Opportunities</u>

Campus clusters vs. data center. Pitt continues to invest in the HPC infrastructure at the University data center, with the goal of realizing economic and operational economies of shared resources over the traditional model of department/grant-owned hardware in campus buildings. CSSD has successfully targeted relocation of campus clusters to the data center, and the Center for Simulation and Modeling's (SaM) "condo-model" of shared investments in HPC resources has provided investigators options for contributing to the shared resources instead of purchasing their own. Despite this success, "campus clusters" will still play a role in University research. Scientific instruments, collectors, sensors, sequencers and other data-intensive technology on campus drive the demand for advanced networking capabilities. In order to realize a positive return on these investments, advanced networking solutions allowing researchers to exploit this capacity are necessary.

Collaboration. External agencies are driving an even greater expectation of collaboration and coordination with national cyberinfrastructure programs aimed at coherent, end-to-end approaches for cyberinfrastructure investments consistent with community best practices and emerging network backbone capacities. Today, agencies provide incentives for adoption of these advanced networking capabilities; over time, however, these types of advanced networking capabilities will be prerequisites for funding any type of grant that uses cyberinfrastructure.

Competing network demands. Over the years, PittNet has met the networking requirements of the academic, enterprise, residential, and research scientific communities on campus. Regular upgrades have kept the infrastructure current; a deployment strategy that balances security with performance and capacity serves diverse needs to the best of its intended design. Network upgrade projects are typically driven by operational metrics such as end-of-life hardware, line card/port capacities, and new OS features, resulting in multi-year projects that deliver an upgraded infrastructure consistent with the current architecture to all buildings on campus regardless of any unique cyberinfrastructure requirements at the floor or lab levels. While this "one size fits all" enterprise model has served well to date, "slices of campus cyberinfrastructure" are not served optimally, and the potential for data-intensive technologies generating bigger, faster flows of scientific data will have a disruptive effect on the general purpose network and lessen the utility of the new technologies. *Effectively managing the tradeoffs between security and performance is an important component of the Campus Cyberinfrastructure plan.*

High-end user needs. Dynamic allocation of network resources to serve unforecasted or unknown data flows is not currently supported. Research projects in several areas still use manual methods of data movement (physical hard drives) either because they have tried unsuccessfully or because they "know" the network is not "fast" enough. Some failed attempts are the result of insufficient bandwidth at key interconnection points. Others may be due to firewalls or ineffective use of tools for data transfer. Our current tools for troubleshooting these problems fall short in helping us satisfy this problem. Many research groups are unaware of solutions, developing "work arounds" that are inefficient and problematic along the chain of collaboration. Most concerning are groups of researchers that have simply not engaged in new efforts for research and collaboration due to preconceived notions or hearsay. *There is a compelling need for new approaches and solutions to meet the needs of science drivers and applications of research computing on our campus.*

Changing role for CSSD. CSSD is changing the way we do business in support of research computing. We intend to remove barriers, re-engineer with an end-to-end perspective, upgrade facilities and connections, and introduce new services to our campus research computing community that extend through regional and national networks. We will bring these solutions to our campus research computing communities instead of waiting for it. We will communicate these plans to our user community and share the experiences and results with our peers. We will partner with regional providers such as 3Rox and KINBER, to exploit new dynamic services as they become available. *Engaging researchers on new services and opportunities for collaboration is an important component of the Campus Cyberinfrastructure plan.*

Proposal

With support from the NSF's CC*IIE program, CSSD will deliver key network infrastructure and engineering improvements to benefit existing research projects at the University and position our institution to lead in the new era of innovation, community and transformational research envisioned in *Cyberinfrastructure Vision for 21st Century Discovery*. The existing projects that will benefit from this work immediately have a wide range of scientific data movement and computational requirements spanning campus, regional and national infrastructures. The projects will deliver new knowledge, data and tools to our campus community, collaborators, and K-12 STEM communities targeted by University outreach programs.

<u>Objectives of Grant Proposal *(Components of the Campus Cyberinfrastructure Plan)*</u>

- Utilize funding from the NSF's CC*IIE program to make key networking infrastructure upgrades to accelerate our campus cyberinfrastructure plan and deliver immediate benefits to sponsored research projects in need of enhanced networking capabilities.
- Improve research outcomes by delivering enhanced campus bridging capacity to the regional and national infrastructures, including compute resources, data stores, remote instruments and data transfer capabilities.
- Allow for envisioning of novel problems and solutions with new opportunities for exploiting distributed scientific grids and experimental networks such as GENI and NDDI/OS3E.
- Promote new collaborations with those in the university communities to participate in joint efforts and potential awards.
- Deliver improved campus-to-data center capacity to drive the use of shared computational resources and utility provided at the University data center.
- Promote the development of tools and mechanisms for common identity and access management systems to facilitate sharing and collaboration outside of Pitt.
- Establish a ScienceDMZ to become a well-known location for campus cyberinfrastructure connections and implement data transfer tools to facilitate the scientific data flows for on-campus and off-campus collaborators.
- Build a model for extending the ScienceDMZ into the campus network and test it for purpose and proofing. Based on the results, develop a new model for delivering high-performance, purpose-built connections to scientific data transfer nodes.

<u>Project Deliverables</u>

- Upgraded 100GbE campus connection to 3Rox (Three Rivers Optical Exchange - the regional network aggregation point operated and managed by the Pittsburgh Supercomputing Center) for increased bandwidth service to Internet2, PSC, XSEDE and other research and education resources, see Fig. 2.
- New 40Gbps (4x10G) service for connecting data transfer nodes and other campus cyberinfrastructure to the University's high performance computing clusters housed at the University's data center. Exclusive bandwidth service for research computing resources supporting current projects on campus and facilitating new inter/intra

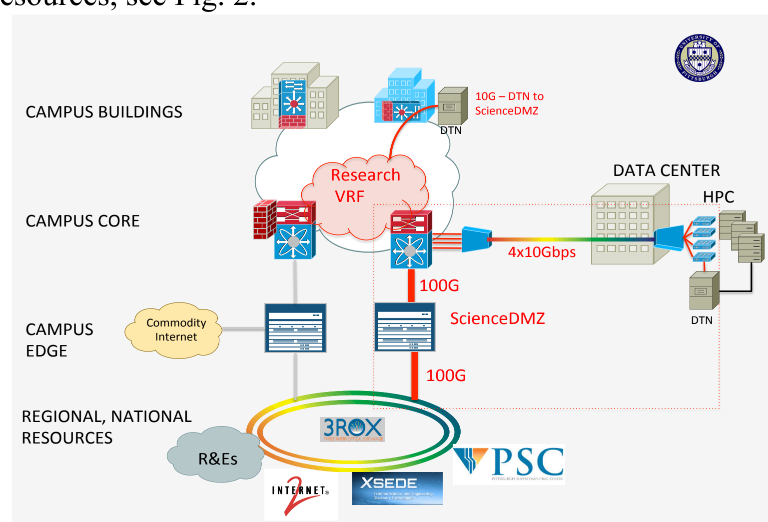

Fig 2: Proposed Upgrades

University collaborations with local, regional, and national collaborators.

- Establishment of a ScienceDMZ on the upgraded campus edge to facilitate data movement between on/off campus collaborators. Re-purpose an existing system to become a Globus endpoint with support for federated ID accessible to on/off-campus collaborators.
- New "research VRF service" for extension of the ScienceDMZ into the campus core network as an overlay to the campus backbone network from which connections to the ScienceDMZ and HPC resources can be extended to connect DTNs.
- Extension of the "research VRF" service into Old Engineering Hall (Pittsburgh campus) using existing unlit fiber and existing unused switch ports to connect an existing DTN to the ScienceDMZ. The new, dedicated connection will support an existing DTN connected to the ATLAS cluster and be used for data movement services to the ATLAS grid and HPC resources at the University's data center using the ScienceDMZ. No grant funds are being requested to implement this pilot.
- New perfSONAR unit to conduct active measurement and troubleshooting during the project and long term integration into the new campus infrastructure.

<u>Broader Impact</u>

- The upgrade of key network connections on campus will provide the foundation for continued network improvements and development of services tailored to meet future research computing needs. Bringing the dialog of purpose-built connections for data transfer utility to the researchers will help them consider new, more effective ways to transmit data and utilize University resources such as HPC, data storage, institutional repositories and collaborations including national cyberinfrastructure. The strengthened infrastructure will help transform the Pitt research culture as faculty and students adopt new strategies to transport, store, and preserve data. This capability will also serve as a proving ground for expanding use of other potentially beneficial services, such as the Pittsburgh Supercomputing Center's DataSupercell and others being investigated by CSSD.
- Numerous sponsored projects will benefit from the improved capacity to move data to and from 3Rox/PSC/XSEDE. New network capacity to University HPC resources will help current projects and spur new collaborations and educational activities for undergraduate research, teaching, training, and sharing of data with a larger community of researchers. The design and operations of the new network features will be included in SaM's HPC workshops and in meetings, presentations, and e-communications conducted by CSSD.
- The infusion of new technology targeted at bandwidth and purpose-built service for data movement will help CSSD maintain momentum with information security programs targeted at mitigating security risks and compliance in research labs and facilities. Further penetration of IPv6 service and development of tools such as perfSONAR and techniques to achieve "fasterdata" will benefit users and IT support teams.
- The enhanced networking capability will facilitate the University's ongoing efforts in expanding research and collaborations in regionally critical areas such as energy and shale research.
- The infrastructure will enhance undergraduate, graduate, and professional education and career development by improving access to advanced technologies within the University

HPC cluster and national XSEDE resources. Training and proficiency in HPC techniques and advanced programming using algorithms for advanced data analytics, informatics, and big-data management will help Pitt students contribute to their chosen fields and to society as a whole.

Project Management Approach/Staffing

The team of investigators represents the central IT organization (PI-Stengel, CoPI-Keslar), HPC consultants (CoPI-Pisciuneri), domain scientist for bioinformatics/genetics/medicine (CoPI-Barmada) and a key project group (CoPI-Pisciuneri, Givi group). Each investigator is well suited for communicating our project goals and activities to the widest set of users across the campus communities spanning health, medicine, engineering, and others. PI-Stengel, reporting to the University of Pittsburgh's Chief Information Officer (CIO) will sponsor the project in the PMO. Letters of commitment from other stakeholders and beneficiaries have been included in the supplemental documents section of the proposal.

CSSD will use its Project Management Office (PMO) and professional project management resources to ensure the successful delivery of this project. A project manager will be assigned with responsibilities to lead all phases of the project.

The project team will be comprised of mostly CSSD resources including enterprise architects, network engineers, HPC systems engineers, systems engineers, security, and managers. Resources from SaM will also participate in the effort. The effort to implement this infrastructure, and coordinate and communicate to stakeholders falls within the normal duties of the PI, the Co-PIs, and the technical resources. No staffing costs are included in this proposal. The result will be a working system integrated into the existing network operations infrastructure.

Equipment Requested for Funding by the CC*IIE Program

New equipment requested in this proposal includes:

- New Juniper 100GbE line cards, modules, and optics for use in the new campus edge router (Juniper MX480) in the Cathedral of Learning to be used for the connection to 3Rox and the Cisco Nexus core switch in the Cathedral of Learning.
- New 100G line card, modules and optics for connecting the campus edge router to the campus core (Cisco Nexus 7009) in the Cathedral of Learning.
- New 24x10GbE line card in the (Cisco Nexus 7009) in the Cathedral of Learning to establish the "meet me point" for formation and connections to the ScienceDMZ and extension of the ResearchVRF into campus connections. This card will also be used to drive the new 4x10Gbps service to the University data center for research traffic.
- New DWDM MUX equipment in the Cathedral of Learning and the University data center for both "sides" of the new 4x10G channel service. Existing fiber will be used. This new service will be used exclusively for research computing
- New special purpose equipment to support implementation of perfSONAR toolkit.

The Cisco and Juniper switches in the Cathedral of Learning are listed as equipment in place and will accept the equipment being requested from the grant. CSSD is committed to supporting the grant equipment as part of normal operations and lifecycle planning moving forward.

Budget Summary

The proposed budget is summarized below. Line item details are included in the Budget Justification submittal and Fastlane. No indirect costs are included per the solicitation. No costs for staff are requested. Vendor quotes for each item are included in the supplemental documents section and reflect the best discounts available. The total requested award amount is $499,437.00

Category	Budget
Equipment in the Cathedral of Learning and University Data Center (RIDC)	$443,421
Other Direct Costs - 100GigE Port Fee on 3Rox Brocade Switch (Service Charge)	$55,200
Travel	$816
Total Request	$499,437

Project Boundaries

The boundaries of this project include the University of Pittsburgh's campus in Pittsburgh, Pennsylvania 3Rox/PSC (Mellon Institute), the University's data center (Harmarville), and the Cathedral of Learning and Old Engineering Hall (OEH) on the Pittsburgh campus. The University's regional campuses are not included in this scope. Campus interconnections to the University of Pittsburgh Medical Center (UPMC) are not in scope.

Project Schedule/Milestones

CSSD will implement this project using standard project management methodology for project delivery. Below is a high-level date and phases/milestones view of the proposed effort.

Date	Phases/Milestones
September 2014	Anticipated award date
September 2014	Project Management Office engaged
October 2014	Engage communications plan
October 2014	Execute hardware purchases
November 2014	Establish pre-upgrade performance baselines
December 2014 (winter recess)	Execute 3Rox upgrade
March 2015	Execute 4x10G service to data center
April 2015	Build Research VRF Service and Security Profiles for ScienceDMZ
April 2015	Build 10G extension to OEH and DTN on ATLAS
May 2015	Complete perfSONAR build and integration
June 2015	Performance testing, baseline reviews, tuning
July 2015	Adjustments, assessment, measurements
August 2015	Prepare NSF Report
September 2015	Completion

Measurements of Success

CSSD will include the following in the project plan as measures of success:

- Completion of the project plan on time and on budget.
- Completion of NSF required reporting.

- Expenditure of grant funds with appropriate supporting documentation.
- Successful integration of the technologies into service and promotion to production status.
- Increase in throughput or reduction in transfer time using post-upgrade data movement tests compared with pre-upgrade baseline tests showing for PGRR, turbulent flow simulation, ATLAS.
- Confirm Phys/Ast readiness for the Large Hadron Collider (LHC) energy run in 2015.
- Increase in the number of existing SaM projects using the DTN and ScienceDMZ for data movement instead of current methods.
- Documented number of discovery of "unknowns"- groups with data movement problems that we can engage and assist.
- Documented number of "conversions" to new DTN/ScienceDMZ model from current, inefficient data movement practices.
- Documented number of new collaborations (internal/external) spurred by our project and outreach.
- Documented number of new grant proposals that identify our new service(s) as a key differentiator or service to be exploited.
- Document any opportunities for researchers or projects interested in using GENI, NDDI/OS3E or Dynes in the future.
- Documented number of new "ResearchVRF" connections that CSSD can use for planning future network enhancements and considerations on plans for developing advanced networking solutions such as SDN.

Risks/Mitigations

- The system targeted for use as the DTN at the data center (app0) may be insufficient to realize the desired performance gains for data movement to the HPC clusters. An upgrade may be required. The Center of Simulation and Modeling will work with the project team in this assessment and determination of necessary mitigation.
- 3Rox/PSC is an external entity outside of our control. We have communicated our intent to 3Rox and received a formal quote for the 100GbE service project. 3Rox has provided a letter of commitment.
- The PSC-developed Slash2 nodes and distributed file system functionality are outside the control of CSSD. These nodes are being installed in early 2014 and should give PSC time to work out any problems and be ready for the bandwidth upgrades in early 2015.
- ATLAS Cluster - GridFTP DTN - The computer to be used as a GridFTP server for data transfers to the ATLAS grid over the new 10Gbps connection to the ScienceDMZ is not owned by CSSD. Personnel from Physics and Astronomy are in support of this effort and will work with us if problems arise.

Additional Factors

- CSSD will use our standard procedures for network upgrades including scheduling of work during normal maintenance windows and communicating these activities to the user communities. CSSD will target the 2014 winter recess when the University is closed to perform some of the work. Other work will be scheduled and delivered with minimal impact expected on production systems and research projects.

- CSSD operates a central directory infrastructure for user accounts as a central component to the IdM architecture. The University joined InCommon and is currently using EduRoam certificates as one component of the centralized scheme for authentication and authorization.
- All CSSD personnel will be using the standard two-factor authentication mechanisms for access control and identity management for network equipment administration.
- All new equipment and services will be integrated into the Network Operation Center's monitoring and systems. This includes NetCool, Cacti, and other tools used for monitoring and managing network infrastructure.
- Standard vendor hardware/software maintenance coverage for all new equipment is requested in the grant proposal. Upon completion of the term, CSSD will assume responsibility of funding hardware/software maintenance for the grant equipment.
- The special purpose equipment proposed in the grant request includes a computer to support the use of perfSONAR tools for active measurement and troubleshooting. The grant equipment will serve as a measurement point similar to the use case adopted by LHC Tier2 sites; the disk becomes part of a global registry of perfSONAR tools allowing affiliated sites and other interested parties to perform tests to the installation. BWCTL, NPAD, and OWAMP tools will be used. Other perfSONAR-PS services will be investigated. CSSD currently uses Cacti in our NOC.
- Currently, dynamic circuit network offerings are not available from the R&E providers 3Rox and KINBER. 3Rox will be receiving an ExoGENI rack in the future. As such, this grant does not include any dynamic circuit network services.
- The grant will leverage funding 3Rox received through the NSF ARI program (NSF grant 0963459) for upgraded MAN infrastructure to be 100GbE capable: adding a 100Gb/s wavelength on the metro DWDM infrastructure and installing a 100GbE-enabled router and two 100GbE Openflow enabled switches.

Research Projects That Will Benefit From the Proposed Upgrades

Pitt is a research intensive university, ranking fifth overall and third among public institutions in the U.S. National Science Foundation's ranking of federally funded research. Pitt's School of Medicine and its affiliates rank 5th among U.S. medical schools in National Institutes of Health (NIH) funding, and the University ranks No. 1 in funding from the NIH's National Institute of Mental Health.

Numerous sponsored research projects at the University will benefit from improved capacity to local, regional, and national cyberinfrastructure, including the Pittsburgh Genome Resource Repository, Data Management and Visualization in Petascale Turbulent Combustion Simulation, Pitt participation in the ATLAS Experiment at CERN, and Linearly-Scaling Biomolecular Simulations with Efficient Statistical Sampling.

The Pittsburgh Genome Resource Repository (PGRR)

Intellectual Merit: The PGRR project will deliver new methods and techniques for managing big-data to support use of TCGA data for personalized medicine.

Project Description

Broader Impacts: The creation of an approach that is scalable to other big-data national resources for personalized medicine research as they become available.
Led by Dr. M. Michael Barmada, Associate Professor of Human Genetics and Biomedical Informatics at the University of Pittsburgh, the Pittsburgh Genome Resource Repository will provide a data management and computing infrastructure for the development of scalable approaches to utilizing large-scale genomics data sets such as The Cancer Genome Atlas (TCGA) for personalized medicine research projects and the implementation of clinical genomics.

The Cancer Genome Atlas (TCGA) is one example of the rapidly emerging paradigm of large nationally sourced repositories of phenotype and genotype data in disease. Many large projects are creating datasets that promise to have a transformative effect on translational research and clinical practice, particularly in the area of personalized medicine. However, while development of these knowledge bases has proceeded, the large size of data has made many of them unwieldy and difficult to access and use. NIH has recently recognized these issue and produced white papers and RFA's on the need for cloud computing resources and computational infrastructure to integrate genotype and phenotype. Fortunately, researchers at the University of Pittsburgh realized the challenges in this area and set out to tackle these issues.

The Pittsburgh Genome Resource Repository (PGRR) will provide a data management and computing infrastructure to support use of large genomic data sets such as TCGA for personalized medicine, with the goal of creating an approach that is scalable to other Big Data national resources for personalized medicine research as they become available. PGRR includes collaboration of over 60 faculty and staff from the University of Pittsburgh's Institute for Personalized Medicine, the Department of Biomedical Informatics, the University of Pittsburgh's Center for Simulation and Modeling, the Pittsburgh Supercomputing Center, and the University of Pittsburgh Medical Center's Enterprise Analytics Program.

Publically available datasets such as TCGA represent extremely valuable resources for Personalized Medicine research projects permitting discovery of new biomarkers, validations of new methods, and education and training. In addition to research, TCGA and other large projects will become invaluable for the clinical program at UPMC, as they will be used to derive knowledge bases for clinical decision support. TCGA currently contains data from over 7,600 participant cases and is expected to grow to more than 10,000. The total size of this dataset is currently more than 850TB with growth projections to 2PB in 2014.

The TCGA data provides an excellent test-bed for Personalized Medicine/Big Data research efforts because:

- It requires partnership. Success with Big Data will require working collaboratively, and TCGA provides a first clear example to develop systems and processes necessary to meet the various needs of investigators and clinicians. Centralized infrastructure including HPC and Supercomputing environments are required. Large graph analytics machines like PSC's Sherlock, analysis infrastructure like Pitt's SaM - CLCBio's Genomics Workbench/Server and other UPMC data-mining infrastructures need to function coherently and non-redundantly.

- It changes the level of scale. With nearly a PB of raw and processed data, solutions to data storage problems are required. PGRR has selected PSC's DataSupercell storage service as the lowest cost, large-scale storage currently available.
- It teaches us how to deal with Big Data's Big Problems. Frequent file changes, reordering of data, and widely dispersed metadata limit the progress that most individual investigators will be able to make without centralized infrastructure.

PGRR teams continue to develop the necessary methods for managing the Big Data problems. However, the need to maintain the data sets as a coherent, centralized store is necessary. Data movement and staging of datasets to the other infrastructure tiers for analysis, mining, and simulation using traditional file transfer methods is problematic and burdensome given the different locations of the machines. PGRR is working with PSC and CSSD to implement the unique distributed files system solution Slash2 to provide a virtualized, storage-access layer across an IP network allowing for centralized access to the data sets without the need for manual data movement. Slash2 nodes at the University's data center and at PSC's machine room will connect using the existing IP network provided by 3Rox. This solution is expected to scale to meet the forecasted increase in the TCGA dataset sizes in the foreseeable future. However, this places additional demands on the IP network currently supporting both research traffic and enterprise traffic (University Data Center) and could prove to be a bottleneck in the near future.

The proposed upgrades for the University's connection to 3Rox (100GbE) and the University's Data Center (4x10Gbps) will be of distinct benefit to the PGRR project by maximizing the investment in the Data Supercell and Slash2 technologies. Exploiting new bandwidth and connection points within the ScienceDMZ edge will provide opportunities for enhanced collaboration and effectiveness of the PGRR program. These improvements can also be leveraged in the ongoing discussions with UPMC teams on more effective ways to access the TCGA data sets

Data Management and Visualization in Petascale Turbulent Combustion Simulation
Intellectual Merit: Massively parallel simulations of turbulent reacting flows.
Broader Impact: New algorithms facilitating filtered density function simulations on massively parallel, petascale platforms has the potential to redefine the state-of-the-art for high fidelity predictions of turbulent reacting flows. Open source code bases will be developed for public use.
The Data Management and Visualization in Petascale Turbulent Combustion Simulation project sponsored by the National Science Foundation (NSF Award Number: CBET12 50171) is led by Dr. Peyman Givi, James T. MacLeod Professor of Engineering, and supported by Dr. Patrick Pisciuneri, Research Assistant Professor on the Center for Simulation and Modeling team. The Laboratory for Computational Transport Phenomena at the University of Pittsburgh conducts massively parallel simulations of turbulent reacting flows on XSEDE resources.

Large scale turbulent reacting flows of interest in gas turbine combustion and propulsion will be simulated. Large eddy simulation (LES) will be the primary tool for turbulence predictions. The subgrid scale (SGS) modeling in LES will be based on the filtered density function (FDF) methodology. The FDF is superior to other SGS models in that the effects of SGS chemical reactions are accounted for in an exact manner. Because of this the FDF methodology has grown in popularity, and it has recently been implemented in many commercial and government codes, such as FLUENT/ANSYS, VULCAN, and US3D among others.

A typical LES/FDF solver consists of a mesh composed of Eulerian grid points for the LES solver overlaid by an ensemble of Monte Carlo Lagrangian particles for the FDF solver. For even moderately sized geometries, this usually involves tens of millions of grid points and hundreds of millions or even billions of Monte Carlo particles. At each point several flow quantities of interest are retained. For reacting flows the chemical composition at each point must also be retained. The amount of information to describe the chemical composition increases as the detail of the modeling of the chemical kinetics increases. The result is that any given *snapshot* of the simulation can range from hundreds of gigabytes to a few terabytes of data depending on the size of the flow geometry considered and the complexity of the chemical kinetics.
There are two major prerequisites to producing data at this scale:

- A highly scalable implementation of the LES/FDF methodology capable of efficiently utilizing supercomputers.
- Access to large scale supercomputers.

The first criterion has been met as outlined in several recent journal publications [1-3]. The second criterion traditionally is met through allocations for usage of the Extreme Science and Engineering Discovery Environment (XSEDE) resources.

Workflow and Data Movement.
The workflow begins with a set of initial and boundary conditions for a given reacting flow of interest. These conditions are used to set up and run a simulation. Snapshots of simulation data are output at fixed intervals. Thus, a complete simulation consists of several hundred snapshots of data.

The next step is the post-processing of the data generated by the simulation. Because transferring large amounts of data is slow, two approaches are generally taken:

- Only a subset of data is transferred to the client at a time, i.e. particular flow variables of interest.
- Visualization and post-processing are done remotely.

These two approaches are adequate for the purpose of producing simple plots and comparing with the available experimental data. Results reported in our previous work use some combination of these approaches [1-6].

However, these approaches are not adequate for curating the data for a given simulation. In this scenario, all snapshots must be transferred to a client for storage and subsequent access. The curation of the simulation data has many important benefits, including:

- The ability to make available entire data sets to other researchers for independent analysis.
- The ability to make available entire data sets for further collaboration with other researchers.
- The ability to explore entire data sets for interesting or previously uncaptured physics.

In order to add the curation process to our workflow adequate data transfer rates must be available to and from our client machines. Further, turbulent reacting flows are transformational, in the sense that as more computing power becomes available, larger and more detailed flow problems of interest will be accessible through simulation, thus exacerbating the need for adequate network transfer rates.

ATLAS Experiment

Intellectual Merit: Participation in the Large Hadron Collider experiment at CERN.
Broader Impact: Expanding knowledge and contributing to new discoveries in particle physics and high-energy research.
The University's Department of Physics and Astronomy operates a Tier-3 cluster consisting of approximately 200 cores in 11 rack mount machines and several terabytes of storage, integrated into a grid of computers which simulates, reconstructs, and analyzes data from the ATLAS experiment at the Large Hadron Collider at CERN. The activities include calibrating b-jet identification algorithms and analyzing samples of top quark decays recorded at the ATLAS detector. The present bandwidth limitations make it unattractive for the group to fully exploit the CPU and storage capacity of the cluster and to fully integrate it with the rest of a distributed infrastructure. While the workflow has been adapted to utilize more off-site computing and fewer in-house resources where possible, removing these barriers would have a significant impact. This is particularly important for the upcoming 2015 run, during which the volume of data is expected to increase by a factor of four. The proposed network upgrades will allow Pitt to remain competitive, assure the relevance of the department's infrastructure and continuation of a dynamic program at the energy frontier.

Linearly-Scaling Biomolecular Simulations with Efficient Statistical Sampling

Intellectual Merit: Large scale molecular dynamics simulations of biomolecular systems; grand challenge opportunities in the elucidation of the biophysics behind allostery and protein/protein binding.
Broader Impact: The Weighted Ensemble Simulation Toolkit with Parallelization and Analysis (WESTPA) open source software package provides a flexible and portable interface to any stochastic sampling engine. The beta release is in use by more than 30 universities.
It is widely appreciated that biomolecular simulations are hamstrung by a lack of configurational sampling, preventing access to biologically relevant timescales of interest [7]. Even record-setting molecular dynamics (MD) simulations using vast resources have studied extremely small systems and/or un-physiological conditions [8-10]. Alternatives to "straightforward" parallelization of brute-force MD are sorely needed.

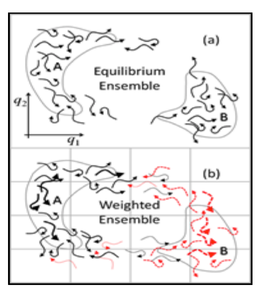

Fig 3: **Weighted ensemble parallelization.** Individual WE trajectories (arrows) are run on separate cores with intermittent communication. Additionally, WE is efficient because trajectories are better distributed in conformational space (b) compared to naïve parallelization (a).

The weighted ensemble (WE) simulation protocol [11, 12] is an ideal marriage of parallel computing and molecular simulation. Because WE employs quasi-independent MD trajectories run on separate cores with intermittent communication, the parallel scaling is essentially perfect (see below). Yet WE is better still: the communication among trajectories enables efficiency by improving the distribution of trajectories in configuration space, Fig 3. That is, in addition to the scaling performance, the overall amount of computing time in a WE simulation can be orders of magnitude less than a straightforward MD run to obtain the same statistical sampling [13-18]. Finally, as a "meta algorithm" controlling and organizing multiple simulations, WE can be employed with a broad class of dynamics types beyond MD, such as Markov chain Monte Carlo, Brownian dynamics, and hybrid quantum/molecular mechanics simulations [12].

WE simulations simultaneously can provide both equilibrium (state populations, potentials-of-mean-force) and non-equilibrium (rate constants) information [14]. The basis for the WE methodology is sketched in Fig. 3. Equilibrium dynamics can be represented as an ensemble of trajectories, but critical transition regions and transition events will be poorly represented, if at all, in straightforward MD due to their low probability. WE simulation, on the other hand, distributes trajectories more uniformly in configuration space permitting much better sampling of transitions while maintaining an exact statistical mechanics representation of the ensemble [12]. Less likely trajectories are over-represented but accordingly assigned lower weights (thinner arrows in Fig. 3). By decomposing equilibrium into two steady states [19] (black, red in Fig. 3), kinetic information among arbitrary states can be obtained even after a simulation is performed. The black trajectories are those currently in, or which were last in, state A, while red trajectories are currently or were last in B.

Investigators Chong and Zuckerman are leaders in developing WE methodology and applying it to biomolecular systems [12-18]. Led by the Chong group, the investigators have collaboratively developed a powerful and flexible WE implementation called "Weighted Ensemble Simulation Toolkit with Parallelization and Analysis" (WESTPA). WESTPA is ideally suited for large core-count computing clusters for the following reasons: a) the software runs with nearly perfect linear scaling when using up to thousands of cores in parallel on the XSEDE Stampede supercomputer, and b) thousands of cores with the same processor speed are required for maximum efficiency of the WE approach. The software, which is written in Python with time-critical portions in C, can be easily adapted for use with a variety of dynamics engine. It has already been used extensively with GROMACS [20], AMBER [21], and NAMD [22] MD engines. The "beta" version of WESTPA is already in use at more than 30 universities.

<u>Grand Challenges in Biomolecular Simulations Leads to New Insights and Large Data.</u>
Uncovering the biophysics of allostery and elucidating the atomistic details of protein-protein binding/unbinding events represents grand challenges in the field of biomolecular simulations. The WESTPA framework provides a unique opportunity to study these important molecular processes, which would otherwise be computationally inaccessible using conventional techniques.

The Zuckerman group are employing WE simulations to study hemoglobin, a key allosteric system [23-26]. Previous MD simulations have proven inadequate to reveal forward and reverse allosteric conformational changes [27-30] let alone ensembles of transition events or the physical process of binding. WE simulations, particularly given past successes with proteins [13, 15, 18], have substantial potential to open up the field of computational allostery, for hemoglobin and beyond.

The Chong group are employing the WE approach to simulate protein-protein binding (and unbinding) events in explicit solvent. In particular, her research focuses on one of the tightest known protein-protein complexes: the complex between the bacterial ribonuclease, barnase, and its intracellular inhibitor, barstar. In addition to the wild-type complex, the Chong group are simulating mutant versions of the complex to provide opportunities to validate both the relative and absolute computed association and dissociation rate constants with experimentally measured values [31]. While simulations of barnase-barstar binding have been performed previously with rigid proteins [32, 33] simulations with flexible proteins in explicit solvents have not yet been possible.

Currently, the HPC facility at the University's Center for Simulation and Modeling (SaM) is connected to campus with a shared 10G fiber and the University of Pittsburgh is connected to the Pittsburgh Supercomputing Center and Internet 2 via a 10G link at the slowest point. However, the pioneering projects described above, which address grand challenges in the field of biomolecular simulations, i.e. allostery and protein binding (Fig. 4), has the potential to generate tens to hundreds of terabytes of data using the local HPC resources and NSF-funded XSEDE supercomputing resources, i.e. Stampede at the Texas Advanced Computing Center (TACC). The data transfer of these large amounts of data between TACC and our campus cluster would benefit greatly from dedicated research network that includes an upgrade to 100G to Internet2. Increasing the rate of data transfer will enable richer utilization of XSEDE computational as well as archival resources. Furthermore, a more efficient network will enhance current and enable new collaborations with other research institutions that are using the WESTPA software to elucidate the underlying mechanisms behind key long-timescale processes. Simulation data can be shared among collaborators and data mined for auxiliary mechanisms that contribute to the overall biologically relevant rate-limiting step. Data mining of wild-type vs. mutant biological systems to uncover correlations between structure and function can generate new hypothesis in silico as potential targets for experimental mutation studies. Establishing the proposed ScienceDMZ, thus, will enable the WESTPA software tool to reach its full potential and demonstrate to the broader scientific community the ready portability of this approach with other advance computing algorithms and dynamics engines.

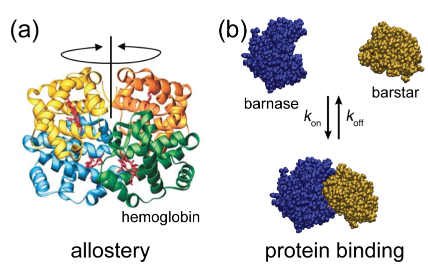

Fig. 4 **Biological systems to be explored with weighted ensemble.** We intend to study topics that are inaccessible using conventional simulation techniques such as allostery in hemoglobin (a), and fully flexible protein-protein binding and unbinding between barnase and barstar (b).