Panel Evaluation Report for the XSEDE Project
Year 5 and Final Report
June 14-16, 2016

Executive Summary

XSEDE is a NSF-funded virtual organization that supports the sharing of computing resources, data, and expertise with scientists around the nation and the world. In addition to supporting the integration of computing resources and services into a single platform, XSEDE invests significant resources in community engagement, education, and training. As XSEDE is finishing its final year of a five-year effort, it is preparing to move to XSEDE2, its next phase of effort.

The review panel was charged with evaluating the draft final project report of the XSEDE project, with an emphasis on PY4/PY5 activities, lessons learned, and the transition plan to XSEDE2. The goal was to provide feedback to the XSEDE team to maximize the value and potential impact of the final report for the community and NSF.

In addition, the review panel provided comments on Lessons Learned to the XSEDE team as they transition to XSEDE2. Areas of lessons learned include the overall question of how science, community engagement, and workforce development were best enabled by XSEDE in the respective project components of: 1) Governance, 2) Extended Collaborative Support (ECCS), 3) Resource Allocation (RAS), 4) Community Engagement & Enrichment (CEE), 5) Community Infrastructure (XCI), and 6) Operations (Ops).

Overall, the panel recognizes that XSEDE has developed and:

- ***Deployed a resilient cyberinfrastructure that is having a significant impact on multiple stakeholders including scientists, Service Providers (SPs), campuses, and research communities***
- ***Is executing impactful outreach and community engagement efforts***
- ***Is improving the operations, governance, and evaluation processes for a distributed CI, which can be of great value to the community and other virtual organizations***

Summaries of specific project components:

Governance: XSEDE is strong in this area. It is growing into a federation of high-end facilities and campus-based computing resources that support the diverse needs of many scientists and research communities. Increasingly, there is seamless access to data and analysis capabilities as XSEDE completes its final year in service before transitioning to XSEDE2. Under the present management structure, XSEDE's critical infrastructure for operations is **reaching a level of stability and maturity that will allow the project to focus on sustainability** and continued improvement. XSEDE is continuing to integrate new services into its infrastructure and actively working to deliver these services to a broader community.

ECSS: defines **world-class best practices and capabilities** in facilitating scientific discovery and broadening participation. Its **leadership and staff are to be greatly commended**. Better tracking of effort and ROI with continued pursuit of proactive and **innovative approaches are essential** as they transition into XSEDE2 to ensure that the project can continue to demonstrate relevancy and value.

RAS: The panel commends the achievements of RAS. Resource allocation plays an increasingly important role across XSEDE and its SPs. Adoption of a unified interface to identify resources, as well as create and monitor allocations, improves the user experience and provides cost savings across the organization. **The panel recognizes that elevating RAS to a Level 2 Area in XSEDE2 is appropriate and aligns with the increasing importance of RAS**.

CEE: The XSEDE team has developed a robust and responsive community engagement effort and begun to incorporate their lessons learned into changes in their practice. In particular, the team is to be **commended for their tremendous work on supporting underrepresented users**.

XCI: XSEDE is deploying the foundational cyberinfrastructure that binds the resources of the NSF supercomputing resources together into a coherent whole. **XCI is well positioned to do an excellent job in XSEDE2**. It defines and deploys the software components that make it possible for users and SPs to easily and securely manage, access and use resources. The **XCI component of XSEDE has done a very good job** of:

1) Clearing technical debt inherited from TeraGrid
2) Integrating and deploying critical services and software
3) Updating software engineering processes

Operations: The panel was impressed by the agility, quality and capabilities of the operations group that provide the critical underlying platforms to enable the overall success of XSEDE. Ops has significantly reduced outages, and optimized their systems and services to deliver a higher quality, more secure platform that underpins the entire XSEDE program. Technical debt has been paid down, resulting in significant cost savings and streamlining of activities. **The XSEDE Operations group is to be highly commended on excellent progress in their mission**.

Major Accomplishments of XSEDE in PY4 and PY5:

The panel highlights that XSEDE has:

1. Deployed and operates a robust and responsive infrastructure, delivers critical enterprise-level services, and provides an effective framework for coordination services across SPs and stakeholders
2. Collaborated with science teams over extended periods to increase scientific productivity and expand the benefits and effectiveness of NSF's advanced cyberinfrastructure
3. Implemented engagement efforts that have developed and executed a model approach towards broadening participation across the HPC community
4. Implemented and is effectively using Key Performance Indicators (KPIs) and other metrics to drive decisions around user engagement activities

5. Recognized that XCI's initial software development process did not meet the needs of the project, and has created new processes that are more responsive to evolving user needs and expectations
6. Enhanced communication and coordination across XSEDE and SPs that has resulted in creating a responsive and highly secure infrastructure

Transferable Lessons Learned

1. **Management of Virtual Organizations:** A key lesson learned is that the identification and implementation of an appropriate governance structure that integrates centralized leadership and coordination with distributed decision making and project execution. This can be leveraged by other large distributed projects and virtual organizations.
2. **Design for Sustainability**: A major challenge for a large distributed project is to design for sustainability beyond the lifetime of the award. In transitioning from TeraGrid to XSEDE, the project learned that the engineering practices that they adopt will have a major impact on the future sustainability of XSEDE2. Specifically, the outcomes of design and deployment must always yield systems that are well tested, documented, and easily extended to meet new requirements. Good software becomes a community resource.

Major Recommendations Moving Forward to XSEDE2

XSEDE should:

1. Improve the capturing, tracking, and highlighting of its **impacts for all major activities** in the context of project outcomes. These impacts go beyond the high-profile science use cases and include impacts on stakeholders, the SPs, campuses, and other CI-based projects, as well as end users. Efforts should include:
    - Documenting best practices
    - Developing consistent and more effective software delivery, documentation, and automated testing mechanisms
    - Identifying meaningful impact metrics, beyond KPIs
2. Better **understand the needs of its diverse user community** and develop outreach mechanisms to engage a broader set of users who currently may not feel they could benefit from XSEDE (*i.e.*, no computation or data are too small)
3. Build in **ROI measurements** and evaluate ROI as perceived by multiple constituencies (XSEDE, SPs, campuses, end-user scientists)

Table of Contents

1 Governance and Management

XSEDE continues to grow into a federation of high-end facilities and campus-based computing resources that support the diverse needs of many research communities. It provides seamless access to data and analysis capabilities as it completes its final year in service before transitioning to XSEDE2. Under present management, the critical infrastructure for XSEDE operations is reaching a level of stability and maturity that will allow the project to focus on sustainability and continuous improvement, integrating new services into the infrastructure, and actively working to deliver these services to a broader community. XSEDE is strong in this area.

Governance and Management includes all strategic planning and coordination activities, project management, change and risk management, reporting, sustainability planning, and policy development and implementation, and comprises the following components:

1. Senior Management Team, which includes the Chairs of the User Advisory Committee (UAC) and Service Providers Forum (SPF)
2. XSEDE Advisory Board (XAB)
3. Membership of the User Advisory Committee (UAC)
4. Membership of the Service Providers Forum (SPF)

1.1 Achievements in PY4 and PY5

The leadership team has injected considerable forward momentum into the project, with the refined management and reporting structures better able to support the project as demonstrated in the draft project report, and augmented by presentations and discussion at the reverse site visit. The Governance model has successfully aligned the PI institution with the needs and interests of the SPs, the user community, and other stakeholders through formal committees at the executive level and informal processes for communications at other levels. The panel recognized the progress made by the project in the following areas:

- The XSEDE senior management team has provided an infrastructure that enhances scientific productivity as intended in a climate where best practices and software are shared.
- XSEDE is becoming data driven in their business decision-making, and is proactively involving researchers from underserved minorities and underserved-minority serving institutions.
- Clear policies for decision making and escalation of issues have been developed and documented to ease the administrative burden.
- The evaluation program and process, which incorporates the Annual User Survey, the Enhanced Longitudinal Study, the Climate Study, and the KPI/continuous improvement reports, present a comprehensive and well-structured approach to measure progress and ensure that XSEDE is achieving its strategic goals.
- KPIs are aligned with the strategic goals and the WBS, and KPIs have clearly defined owners. This enhances the likelihood for success of the measurement and continuous improvement program.

- As in previous years, the KPIs need additional development, although it is clear that the XSEDE staff is familiar with and increasingly productive in a regime of data-driven decision making. It is important, however, to understand the rationale for how a particular indicator influences decision making related to the strategic goal with which it is aligned.
- The panel commends the project leadership team for generating a major increase in trust and confidence within all components of the project, a significant achievement in marked contrast to its predecessor, TeraGrid.
- A key goal of XSEDE during the upcoming project period is to transition to a more modular organizational structure so that it could be recast as multiple independently managed projects in the future. There is commendable evidence of this approach during the latter stages of XSEDE.

1.2 Lessons Learned

- The management team is highly qualified and brings a considerable depth of experience from the first five years of XSEDE. They present a strong collective of expertise in research computing and large-scale cyberinfrastructure. The distributed nature of the project and the distributed governance, which has been designed to preserve appropriate levels of SP autonomy, contribute to the resilience of XSEDE. The proposed increase in the proportion of time each manager must dedicate to XSEDE takes into account local needs of as well as the complex coordination of the project.
- XSEDE has strong leadership, including the PI, co-PIs, and L2 and L3 leaders, affording continuity of leadership. They have introduced industry best practices to ensure that appropriate policies and oversight are in place throughout the project, including critical areas such as cybersecurity.
- The management team is moving to better use a framework for data-driven decision-making that is maturing with the development of KPIs and the new Metrics Dashboard that includes Project Change Requests (PCR). This dashboard allows greater transparency in the statistics and demographics presented, and/or enhancing communications capabilities between its users.

1.3 Transition to XSEDE2

The panel identified the following opportunities for enhancements in XSEDE's approach to PY6:

- Effort must be made to formally establish how XSEDE will assess the scientific impact of projects and publications that result from using the resources of the national cyberinfrastructure. The decision framework should continue to be developed and enhanced to help management prioritize tasks and allocate operations efforts to continually improve upon increasing science impact.
- The panel noted the increased effort commitment for Level 2 and Level 3 management, providing them the time necessary for essential management tasks, including organizational improvements. This enables the Level 2 areas to be independently reviewed, and is intended to promote improved business operations through the provision of appropriate support to financial and sub-award management--an effort that was underestimated for XSEDE. As a result, Level 2, 3 and business operations

increased by 15% in XSEDE2 (PY6). Although a substantial increase, the panel was sympathetic to the variety of unforeseen activities that were not funded in XSEDE1.

1.4 Recommendations

- The management team continues to struggle to articulate the difference between measuring *outcomes* (e.g. the number of people affected, or "reach") and measuring *impacts* (e.g. time saved by researchers as a result of engaging ECSS) in developing KPIs. Ideally, the collection of KPIs for each WBS area should contain a balance of both. The team is collecting extensive data, but this is very much a bottom-up approach, with their ties to 'impact' limited at best. Qualitative and quantitative assessments may both be essential to characterizing impact. The panel recommends some brainstorming with the project evaluator about how to extract one or two measures of impact --and then doing a "deep dive". The panel feels that the XSEDE teams needs to understand and develop effective impact metrics in order to effectively demonstrate and convey the long-term value and legacy of the effort to users, NSF, and the public.
- XSEDE has a responsibility to keep abreast of innovative developments in cloud-hosted services, VMs, containers, etc., and potentially integrate more of them into its technical infrastructure. Similarly, the panel feels that XSEDE should be more proactive in providing technology leadership, and were surprised that this was not more evident within the draft final report and reverse site visit.
- The management team must continue to respond to the changing nature of the applications and community codes that utilize the national cyberinfrastructure. This requires an allocation of appropriate time and effort to ensure that staff support the proper integration activities. Management must also continue to adapt to a changing landscape of technological innovation, and ensure that the appropriate relationships are maintained to identify and respond to externally driven change.

2 Extended Collaborative Support (ECSS)

ECSS defines world-class best practices and capabilities in facilitating scientific discovery, and broadening participation and benefit. Its leadership and staff are to be greatly commended. Better tracking of effort and ROI with continued pursuit of proactive and innovative approaches are essential as they transition into XSEDE2.

2.1 Achievements in PY4 and PY5

The scale and distributed nature of ECSS cost-effectively enables actions, speed of response, and provides breadth and depth of expertise not possible by single SPs. In XSEDE, ECSS is comprised of five components (ESRT, NIP, ESCC, ESSG, ESTEO) with individual staff being called on to contribute project wide. This results in substantial cross-pollination and, for example, brings experts capable of implementing and optimizing complex science algorithms and software on modern computer systems into the classrooms of students across the nation. The expert staff of ECSS (~35 FTE in XSEDE according to the final report, though the slides identify 28 FTE across 70 individuals) collaborates with science teams over extended periods to

increase scientific productivity, and expand the benefit and effectiveness of the NSF's advanced cyber infrastructure.

In PY4 and PY5, ECSS continued its excellent and high-impact work, and has delivered some truly transformational impacts. It is not clear if these accomplishments (of individuals and teams) are adequately recognized within the project and promoted to the larger community via professional organizations, nominations for external awards, etc. In the last full project year (PY4), ECSS completed 65 projects in 26 different fields of science with a 4.58 (out of 5) average satisfaction rating and 3.93 (out of 5) average impact rating. The rudimentary ROI analysis suggests an average saving of 1.5 "researcher years" as a result of an investment of an average of 0.25 ECSS staff years, but some elements of this analysis are arbitrary and undermine the analysis (*e.g.*, selecting 3 years to represent the time saved by those who "could not have done it without XSEDE's assistance"). A more rigorous approach needs to be developed.

The success of the Novel and Innovative Projects (NIP) program is noteworthy. It employs multiple mechanisms and channels for the proactive identification and computational maturation of projects and, crucially, associates people from disciplines and communities that have historically underutilized ACI. In PY4, NIP-related allocations included 18 startup projects, 12 XRAC requests, 53 first-time users, and 27 first-time XRAC users.

Better tracking and root-cause analysis could help fully realize the potential of science gateways for democratizing access to ACI and enhancing scientific productivity. For the first time, science gateway users exceeded traditional (command line/batch) users and now exceed them by 1.4x, with this trend expected to continue. A small number of large gateways seem to dominate the statistics, and there was an incomplete discussion about the apparently short life of most gateways. Perhaps this is appropriate given the low-cost of instantiating a gateway. XSEDE must be commended for leading this transition, with multiple elements of its CI in addition to ESSGW being crucial enablers.

ECSS experienced its first oversubscription of services in Spring 2016. This is surprising because of its value to science projects and the sheer size of its budget and staff. Lack of awareness of ECSS services was identified in PI feedback as a root cause, and resulted in XSEDE executing a major push to increase awareness in Fall 2015, which resulted in the 4% oversubscription. The review panel concurs with renewal projects being given lower priority during project selection. It is worth noting that of the ECSS projects that support more than one University-based research activity, 30% are from the University of Illinois. This may be indicative of the limited awareness of ECSS within the broader scientific community.

2.2 Lessons Learned

- The ECSS Lessons Learned have clearly enabled the team to make significant (but not complete) progress towards effectively executing projects with academic PIs, who may not always have the expected level of engagement. The scale and distributed nature of ECSS have been crucial to its success and broad impact. Awareness within the

community has been identified as an issue, and additional work is necessary and metrics/KPIs should be developed to address this issue.

- The project took too long to identify awareness of ECSS as an issue. Over-reliance on metrics/KPIs to track project performance may prevent the team from identifying opportunities for improvement, particularly at the strategic level.
- The lessons identified by ECSS overall seem to be at low, detailed levels rather than at higher levels. This may be because the project has not yet been heavily resource constrained. In XSEDE2, the team will need to be more attentive to the big picture.
- The present inability to more fully identify the impacts and ROI of ECSS, and the need to make hard choices, should map into lessons learned concerning tracking effort per project and capturing more data concerning impact.

2.3 Transition to XSEDE2

The five components of ECSS in XSEDE continue into XSEDE2. The change in level of effort (28.8 FTE reported in Towns' slides) is unclear since there is conflicting data for XSEDE (was it 35 or 28 FTE?). A major reduction would seem incommensurate with the impact of ECSS.

In XSEDE, ECSS has built a winning approach that has both deep and broad benefits, and cuts across nearly all XSEDE activities. The project is indeed sufficiently nimble to respond to changing needs, but, as demonstrated by NIP and articulated by the project PIs, a proactive approach is necessary. ECSS must continually refresh itself in response to changing technology landscapes and users' needs or it will be stale if it looks the same 5 years from now.

The project team has only just started to think through the implications and opportunities of the scientific software innovation institutes. Two institutes are anticipated to be funded this year, gateways and chemistry/biophysics/materials, as part of the NSF SI^2 program. As new major contributors to CI, it is essential for XSEDE2 to establish synergistic relationships with these and other institutes that will be launched in the future. It is also important for XSEDE to leverage other SSE and SSI projects funded by the SI^2 program. The team should consider conceptualizing a software institute as a companion project for XSEDE focused on development.

In many areas of the project, the transition to XSEDE2 marks the completion of a transition from a reactive to proactive mode of execution. The panel recognized that it is important that aspects of the project, especially ECSS, go beyond being proactive and transition to anticipatory and leadership roles, for example, in better assisting the community to navigate technology transitions. It is unclear what of this is in project scope, or if anything should change in how NSF operates/positions XSEDE in the context of its broader ACI. As noted by the PI, the community and ecosystem are too large for XSEDE to take sole responsibility. By forming partnerships with other national/international peers and with the SPs from NSF and other federal agencies, it can exert significant leadership.

The impact of ECSS and its staff resource deployment/allocation do not seem to be adequately tracked. More effective use of the recently established publication database is one component of this, but other metrics should be developed for tracking the impact of ECSS engagement over time. For instance, grants and awards received by associated PIs could be correlated. It is also essential to better understand the impact of ECSS on the overall productivity of the CI. The collected data will inform crucial investment decisions by the Foundation, such as "is it better to expand the infrastructure or ECSS?" Many of the same techniques used to evaluate education/training programs can perhaps be applied here.

Careful consideration will need to be given to establishing priorities for ECSS projects and navigating the sometimes conflicting goals of maximum science impact and broadening participation. While the team is clearly very aware of the changes underway that touch all elements of the CI and its allocation, the team does not seem to have yet reflected in sufficient depth on the implications.

In the case of ECSS, the rapid growth of cloud and/or data-intensive computing will mean hiring/retraining multiple new staff to handle these domains, as well as reducing the level of effort directed towards current application areas. In XSEDE, data-intensive computing was consciously underemphasized due to initial budget constraints and a historical emphasis on computing. Contract staff is only a short-term solution given the identified need to retain successful and impactful individuals for the long term.

2.4 Recommendations

1. ECSS should continually seek to expand its success by being proactive and adopting new approaches including prioritizing activities with high-level objectives and by better tracking effort and ROI.
2. ECSS should seek additional innovations to scale impact beyond its fixed staff resources.
3. In the coming year, XSEDE should actively seek opportunities to assist in the successful launching of the S2I2 institutes and establish synergistic relationships with them and other SI^2 projects. In particular, XSEDE should identify elements of management and operations that are transferable, leverage its awareness and relationships of/with the broader academic community, and identify mechanisms for engagement and collaboration.
4. XSEDE should form partnerships with other national/international peers and with the SPs from NSF and other federal agencies, and by doing so can exert significant leadership in multiple spaces (technical, strategic, awareness, etc.) despite the community and ecosystem being too large for XSEDE to take sole responsibility.

3 Resource Allocation (RAS)

Overall, the panel commends the achievements of RAS. Resource allocation plays an increasingly important role across XSEDE and the SPs. Adoption of a unified interface to identify resources, and create and monitor allocations improves the user experience as well as

provides cost savings across the organization. The panel feels that the elevation of RAS to a Level 2 Area in XSEDE2 is appropriate and aligns with the increasing importance of RAS.

3.1 Achievements in PY4 and PY5

Resource allocations are central to XSEDE's primary goal of enabling the advancement of science and engineering discovery. The XSEDE team has provided quantitative KPIs and has started incorporating trend analyses that indicate healthy overall growth in usage and likely imply growing scientific impact over the past 5 years (PY1-PY5). Some of the critical KPIs include user satisfaction, system availability, percentage of recommended Service Units (SUs) that are actually allocated, and allocation proposal success rates. Data analysis trends also provide a persuasive case that there is substantive penetration into underrepresented disciplines and user communities. One of the key achievements noted in the Final Report includes the sheer volume of successfully processing 3,600 quarterly research requests. Also, the success of the XSEDE Resource Allocations Service (XRAS), a new resource allocation and usage management system, which is being adopted by non-XSEDE organizations including NCAR, CADENS, and soon Blue Waters, as Software as a Service (SaaS), is a first for XSEDE.

3.2 Lessons Learned

Perhaps the most important lesson learned during PY1-PY5 is that the resource allocation process must be guided more carefully due to the growing request volume and diversity of resources provided by SPs. This is becoming critical as new types of computing resources are becoming available to users (e.g., cloud/VM). A better resource selection system than what was employed during PY1-PY5 will be required. A manual process, though acceptable at first, is not sustainable as usage and demand continues to increase. A semi-automated "guide" will be needed to assist users in identifying the most appropriate resource(s) for their needs. While KPIs have been developed to address both availability and user satisfaction, it is essential that XSEDE continue to assess technical productivity improvements, scientific productivity improvements, educational impact, and overall scientific impact systematically, for example, by identifying exemplar success stories. Other important associated lessons learned include raising visibility of alternate user resources and access points (e.g., gateways and non-XSEDE systems), and more effectively educating users about XSEDE resources and services. An XSEDE-wide group has arrived at policy recommendations to address some of these issues, for example, handling large requests and associated size caps; however, a number of other critical issues such as smoothing the variability in the number of allocation requests is now under discussion.

3.3 Transition to XSEDE2

An important change in the transition to PY6 is the elevation of RAS to L2 to meet the spectrum of challenges of resource allocation, especially in view of the expanding portfolio of resource types (HPC, cloud, etc.) and because demand is increasing faster than resource capacity. This is accomplished through (1) managing resources in collaboration with SPs, (2) enhancing the allocation service infrastructure, and (3) planning for future needs. It was instantiated in part by repositioning activities of Operations (allocations and policies) and User Support (allocations, accounting, and account management). Important past lessons learned will be directly carried

over to PY6, including assisting users to identify appropriate resources matching their needs and use cases. The allocations process is being updated by utilizing a PY5 allocation policy review and improving the XSEDE Central Database (XDCDB). The adoption of XRAS at other organizations has led to the development of a cost/sustainability model during PY6 in anticipation of longer-term use at more organizations. In general, the panel sees a reasonably well thought out transition into PY6 from lessons learned in PY1-PY5.

3.4 Recommendations

RAS must:

1. Help users identify appropriate resources (HPC, HTC, Cloud, etc.)
2. Consider how they can help users monitor resource utilization and code efficiency (e.g. user requests a 32-core 256MB VM, but during occupancy, only uses one core)
3. Identify how RAS could be leveraged for data management allocation that will become more important as users are requesting VMs, block storage, and cloud storage devices.
4. Determine how a RAS KPI could describe the effectiveness of resource allocation in terms of *impact*, rather than just tracking physical parameters that monitor resource utilization, code efficiency, and data management allocation

4 Community Engagement & Enrichment (CEE)

The XSEDE team has developed a robust and responsive community engagement effort and has already begun to incorporate the lessons learned into proposed changes in their practice. In particular, the team is to be commended for its tremendous work on supporting underrepresented users.

4.1 Achievements in PY4 and PY5

Broadening Participation: The CEE team has taken a model approach towards diversity, both in how and what it delivers, and ensuring that broadening participation is hard baked into the organization. The team is strongly encouraged to publish this work in the higher-education literature, as many institutions are struggling with similar approaches to inclusion. Doing so will make it easier to approach NSF/EHR and foundations for additional resources.

Training: The XSEDE team recognized that SPOCs are a better model for student retention and content delivery than MOOCs that deliver only minimal education and training. This is in line with educational research that suggests that for technical content, in particular, MOOC-formats struggle to make significant impacts on learning.

User Engagement: The XSEDE team has expanded the use of the ticket system to drive team accountability. An increased emphasis on transparency demonstrates XSEDE's responsiveness to the community and will build trust, leading to more effective communication of user needs. The team should continue to communicate to stakeholders how they are using this process to be more responsive to user needs.

Metrics and KPIs: The panel commends the team for using KPIs and other metrics to drive decisions around user engagement activities. By leveraging internal KPIs to identify which types of user engagement activities should be dropped (e.g., focus discussions and birds of a feather meetings), and which should be continued (personal, one--to-one contact), the team is well positioned to be flexible in a changing organizational and funding environment. That said, KPIs, as defined by the XSEDE team, often do not translate into *authentic impact measurements* (see below), a critical step for XSEDE going forward.

4.2 Lessons Learned

The panel is concerned that for education and training the XSEDE team sometimes conflates KPIs (which largely measure the effectiveness of organizational progress) and authentic impact measurements. KPIs do play an important role in setting staff expectations and monitoring effort, and they represent significant steps forward in the evolution of XSEDE assessment. However, at this stage of development, with the support of external evaluators and a strong educational research literature in this area, there is an expectation that education impact, which will likely involve more than KPIs, is going to be one of the key program markers moving forward. The Panel suggests that the team look at the SOARS program (Significant Opportunities in Atmospheric Research) at UCAR, one of the nation's longest and most impactful research programs for underrepresented students. They have an evaluation model that directly addresses impact. See, e.g., doi:10.1038/ngeo2501 (and perhaps contact Rebecca Haacker for further discussions).

While the panel recognizes that supporting underrepresented students often requires 'high touch', it is worth noting that the one--on--one contact model that XSEDE currently employs is not a scalable solution for user engagement. The panel commends XSEDE's effort to partner with groups providing specialized and successful training (e.g., Software and Data Carpentry).

4.3 Transition to XSEDE2

The reorganization of Education & Outreach and components of User Engagement into CEE reflects a maturation of the organization, and more accurately reflects community needs and team tasking. Driven by the decrease in support to this unit resulting in a FTE drop of ~2.5 staff, the team has done an admirable job identifying poor performing programs to cut Specifically, the team has decided to drop the XSEDE Scholars program due to poor outcomes (~85% of Scholars did not use XSEDE; many did not have an XSEDE account).

4.4 Recommendations

- Campus Champions:
    - Campus Champions need articulation from XSEDE leadership to their institutional leadership as to the Champion's role and importance. A letter from XSEDE to cognizant campus leaders will help support Campus Champions and allow them to grow in their roles of supporting XSEDE locally and regionally.
    - The panel notes it is worth working with Campus Champions to engage upper level administration at their home institutions to make them aware of how their

researchers are using XSEDE and how institutions can provide feedback to NSF in support of XSEDE.
    - Campus Champions have grown in their number and scope. They should be encouraged to consider attaching to a broader organizational structure/organization rather than trying to self-organize. Suggestions may include organizing within ACI-REF, a subgroup of Internet2, an ACM SIG (such as SIGUCCS), etc.
- The engagement of campus--level diversity specialists as part of a Diversity Forum is solid. There is a concern that since the HPC community is relatively small, the Diversity Forum may not rise to a level of importance to allow those senior campus representatives to participate in the Forum. The panel suggests that the team think about how to incentivize this effort or at least outline why senior leadership should want to participate. In addition, XSEDE may need to look within specific colleges to find the right level of diversity professional. The panel notes that the XSEDE team suggests that a model for the Diversity Forum be the SP Forum; however, the SP Forum has not necessarily been the greatest success and there may be other models that are better suited for this important effort.
- The panel suggests that XSEDE would benefit from a strategy to engage disciplines that currently do not use the XSEDE platform (e.g., computer science, polar science), in a similar fashion as they have engaged traditionally underrepresented users.
- XSEDE should more explicitly focus on 'workforce development' as a gateway to industry. Throughout the project, the reference to workforce is implicit, but would benefit from being explicitly articulated through the XSEDE CEE efforts.
- The panel encourages increased exploration of educational impact, particularly longitudinally tracking students, rather than their current metrics that focus on tracking users and their activities in the portal. The panel feels that the XSEDE team has the expertise in place for measuring impact, and this should be leveraged, especially with regards to measuring the impact of training on professional development and career choices of students. These metrics are considered essential to demonstrate XSEDE's value to the educational community and workforce development.
- Given the high cost of 'high touch' student engagement, the team is encouraged to look at new partners, models and activities that allow them to --reach larger audiences and new methods that can be explored, tested, and refined in order to scale the XSEDE efforts. One promising area to consider is the utilization of alumni as mentors.
- To grow student participation activities, the panel suggests that XSEDE engage local tech companies at their national meeting sites as much as a year ahead of time, to sponsor student activities. In addition, the XSEDE team might approach NSF/EHR for funds that allow the XSEDE team to look at how students from diverse backgrounds are engaging XSEDE, and how XSEDE supports them professionally. This will articulate how these students are transitioning from community colleges to 4--year programs to professional degrees. This would permit the team to explore and test one or two innovative ideas for increasing broad participation.
- To enable campus engagement to grow and flourish beyond the XSEDE funding stream, the panel suggests working with one showcase example to explore the launch of a

regional, self--sufficient campus engagement group, including developing a roadmap with metrics. In addition, the panel suggests engaging university consortia (e.g., SURA, AUI, UCAR, USRA) to support a sustainable model of campus engagement.

5 XSEDE Community Infrastructure (XCI)

5.1 Achievements in PY4 and PY5

The project team delivered eight critical enterprise-level services by the end of the project that are essential to the operation and use of XSEDE. These include services for the management and operation of XSEDE, such as the resource allocation service (XRAS), the user portal and website (XUP), the community software repository (XCSR), and the central database (XDCDB). In addition, they rolled out hardened versions of services that are critical to the users, including the Globus identity management, authorization services, and resources such as single sign-on, remote login, and resource discovery.

A major accomplishment has been the deployment of the highly-reliable Globus Data Transfer and Data Sharing Services, including the services to support data movement from campuses through Globus Connect. As part of the campus bridging services, they delivered an updated version of the rocks-based Cluster management software and the national integration toolkits. Fifty-four campuses have used the cluster tools and one hundred have used parts of both software tools. They have clearly been lifesavers for a number of campus IT organizations: “We wouldn’t be able to manage this cluster without having these XSEDE cluster management tools.”

The XSEDE team made major improvements to their software engineering process, Requirements Analysis & Capability Delivery (RACD), for defining which services are needed by users and a how tools can be adopted. Called the Capability Delivery Plan (CDP), the new process is more agile than their previous process, allowing them to shrink steps to the point that they have been able to complete the task of tool proposal to tool delivery in one quarter of the time it took with the old process. However, the team realizes delivering software of sufficient quality (passing testing and meeting defined requirements) is balanced against the risk of failing to meet user expectations.

Assessment and Metrics: The project team has made a substantial effort to assess the impact of the services and tools that have been deployed. One of the criteria for evaluating a potential service has been the ability of the service to log uses. Logs are the most important way to see if a component is actually used and if it is working properly. However, this only provides an indirect way to understand how satisfied and helpful the service is to the user. In the case of software packages, the metric of downloads is often used. Unfortunately, this is not always an accurate mechanism to understand if the software is deployed and used. For that, specific feedback to the team from the users is required. This feedback has been good for the cluster software.

5.2 Lessons Learned

The XSEDE Community Infrastructure development and deployment activity underwent a major evolution during the period of the award. They started with a large body of software tools left in place from the TeraGrid project that were not easy to deploy and maintain. Building and deploying services that can provide a consistent set of user interfaces and reliable capabilities across the extremely heterogeneous collection of resources provided by the SPs is an extremely difficult but essential task. They looked at 57 user-facing use cases and identified 11 canonical cases. To address these, the project developed and deployed over 20 distinct services to support computing, data, and user management on XSEDE managed resources. Unfortunately, many any of these failed the tests of usefulness or universality and had to be dropped. For example, the XSEDE-wide file system was dropped when it proved to be complex, expensive to maintain, and unreliable. Due to a lack of stakeholder engagement throughout its development and testing phases, it did not satisfy users who were unaware of the performance implications of such a distributed file system. However, a positive outcome has been the deployment of the eight critical enterprise-level services, and the user file transfer and security services described above.

The project initially put in place some extremely ambitious software engineering methods that proved to be not well suited to a project like XSEDE, and recognized they needed to change their approach. RACD (mentioned above) was a major step forward. They are conceptually segregating the users in the requirements stage (with CEE) from the developers in the design stage (with XCI) by having CEE serve as liaison.

5.3 Transition to XSEDE2

The plan for XSEDE2 contains a refined software assessment, engineering, integration, and delivery process that reduces the XSEDE process from 10 steps to a lean set of 5 steps. This shortens the time to deliver software and is an excellent step forward. However, they will need to make sure that the quality of the outcome remains high through frequent user engagement. This will require finding the right balance between flexible and firm engineering processes. In addition, the team recognizes that maintaining and supporting deployed services takes significant resources, and has budgeted accordingly.

Three areas in which XSEDE2 plans to deliver new service capabilities are:

1. The use of group/VO credentials for unified access across all resources
2. Support for distributed science workflows
3. Support for the use of distributed (academic, private and public) cloud resources

The expansion of services to the cloud platform and additional virtualization technologies (e.g., containers) is critical. The team needs to learn from commercial cloud providers and continually assess and address user needs and demands. Given the strong shift to easily deployed microservices, the team is looking at container-based virtualization technology.

There will be additional improvements in XSEDE's cluster tools, which have been rapidly adopted by campus HPC centers. Specifically, they must provide the option to integrate with

XSEDE's identity authentication and authorization tools so users of a campus system can transparently move computation and data to XSEDE resources.

XSEDE must partner with the community to specify features identified as critical to improve XSEDE2 since they are not permitted to spend resources on development, per se. These partners will need to obtain external funding through traditional requests for proposals, and it is critical that XSEDE communicate requirements to potential development groups as widely as possible. Their current "response" to a call for proposals web page is not as widely visible as it could be.

5.4 Recommendations

1. ROI measurement must be built in from start: evaluate ROI as perceived by multiple constituencies (XSEDE, SPs) and build in instrumentation that enables assessment of service delivery users.
2. Due to rapid changes in technology, XSEDE must maintain agility in order to be responsive to user needs. They will need continuous stakeholder engagement to assess their changing needs, and be able to respond appropriately.
3. XSEDE serves two user communities: SPs and end-user scientists. While requirements come from both, solutions are delivered and deployed through SPs with different services impacting one or both groups. It is critical that XSEDE understand which stakeholder(s) is (are) impacted by new services and features; having appropriate conceptual diagrams will help communicate this internally and externally.
4. It is essential that XSEDE engage with the broadest possible community, clearly communicate their future needs and feature requirements, and work as closely with possible with these groups since XSEDE does not develop new capabilities,
5. As XSEDE2 tackles the challenge of integrating cloud technology, it is critical that they examine the methods used by the public cloud providers around containers. Much of this is open source, well tested and understood.
6. As more software is distributed to SPs and campuses, XSEDE needs to have consistent software delivery, documentation, and automated testing mechanisms. The latter will be particularly important for the 100+ sites leveraging XSEDE's software to manage user identities, data movement, and local HPC clusters.
7. Documenting best practices for campus CI providers/SPs would be extremely useful. (A review of DataOne's achievements in this area is suggested.)

6 Operations Group

The panel was impressed overall by the agility, quality and capabilities of the Operations group that forms the critical underlying platform to enable the overall success of the entire XSEDE program. The Operations group significantly reduced outages, and optimized their systems and services to deliver a higher quality, more secure platform that underpins the entire XSEDE program to support, enhance and provide world-class leadership in computational research, discovery and science. Technical debt has been paid down, resulting in significant cost savings

and streamlining of their activities. The XSEDE Operations group are to be highly commended on excellent progress in their mission to support a number of complex and diverse SPs while also reducing complexity and enhancing the overall security and availability of their systems and services.

6.1 Achievements in PY4 and PY5

Service Delivery: The uptime goal of the 15 critical services was two nines (99%), which is 3.65 days of downtime per year. The Ops Team achieved close to three nines (99.9%), which is only 8.76 hours of downtime annually. They deployed RT for their ticketing system (replacing a legacy system) that handles ~12,000 tickets per year with an average of 7 hours turn around per ticket. Approximately half of these tickets were for SPs, the majority of which were handled by XSEDE frontline and represent a significant pass through cost savings to SPs. The account management system is now automated, permitting new account creation to happen nearly instantly, compared to 5 days previously. These improved deployment and management systems enable the onboarding of new SPs to happen more easily and rapidly.

Security: There were no known security events that crossed resources on the XSEDE platform. While individual SPs had incidents, these were contained and did not spread across centers (unlike TeraGrid). This was attributed to close coordination and communication across sites, week scans, and audit testing. The review panel appreciates the dedicated time and effort required to achieve this.

System Development: XSEDE successfully transitioned from NLR to Internet2 Advanced Layer Two Services (AL2S), and deployed a 10-12 node 10GbE perfSONAR mesh for network performance monitoring and tuning. Over 50 enterprise/core services, of which fifteen are classified as critical (including maintaining failover), were deployed and managed. The Ops Team is deploying additional services to track and log data movement across SPs. However, they realize that they need to set realistic goals about the type of data that can be collected due to the heterogeneous policies at different institutions. Logging on disparate systems is not trivial, and deploying XSEDE-wide Globus GridFTP monitoring is viewed as a success. These data are collected and stored in an XSEDE-wide database (XDCDB), which is leveraged for automated report generate for each SP's quarterly and annual reports and represents a significant cost savings to SPs.

6.2 Lessons Learned

The Ops Team realizes that coordination across SPs is a significant benefit to the program, allowing for increased security, a coordinated approach to common challenges, and enabling the opportunity for further increased interoperability. As noted above, through these efforts, there were no known security incidents that spread across SPs. Coordination and shared best practices (when used) across sites have greatly helped combat security problems, and they are able to provide a faster response to evolving threat models and reduce the frequency of successful attacks on individual SPs. This coordination across SPs is essential when dealing with new issues. As highlighted by the misuse of resources by a user (specifically, bitcoin mining), excellent coordination across SPs leads to new monitoring, best practices, and better

and more consistent acceptable use policies. Continued efforts to further improve coordination, transparency, and communication across XSEDE and SPs are required to handle emerging issues. In addition, system/failover testing is essential, and these improvements have significantly contributed to the 99.9% uptime for critical and core services.

6.3 Transition to XSEDE2

The Ops Team plans to deploy two factor identity management, for all users, more consistent logging of data movement across all services (GridFTP, scp, etc.), and XSEDEnet health and performance tuning SDN across the WAN - all of which are considered to be essential in the future. These deployments will afford the team the ability to identify network hot spots, such as where researchers are still using standard secure copy (scp) when they could use improved technologies such as GLOBUS to transfer data. Each L3 manager has turned over due to a variety of factors (e.g. reorganization, retention, etc.). The review panel is convinced that this was part of normal operations and not as a result of any specific underlying issue. *The panel commends the group and their activities as the project continues to streamline their overall processes and systems to support XSEDE leadership in research and science.*

Systems Operational Support: The XSEDE team is in the process of continuing to streamline their operations, while maintaining a clear focus on user needs. The XSEDE2 help desk will not be staffed 24x7, as weekend-overnight calls are rare, though critical monitoring will continue. Net result: $100,000 worth of savings per year. In addition, decreased outage time has led to fewer tickets, allowing the team to focus on more sophisticated systems to further improve their services. The team should continue to monitor the mean time to resolution of tickets closely to ensure that the quality of customer support remains high. A high degree of failover testing, periodic failover tests are now used for the XSEDE Central Database (XDCDB), and critical systems are now more available. When questioned about how the service desk are incentivized, the response was that these staff would be a tier 1.25 - 1.5 in a “regular IT” organization and love working with the most sophisticated systems in the world, they have the best incentives to deliver a high quality product.

XSEDE2 Operations Priorities: The future target areas for the Ops Team are to provide better management and coordination, with more secure operations as new cyberinfrastructure is deployed (e.g., VM, cloud, and SDN). In addition, the Ops Team will focus on data transfer support services amongst SPs and campuses with QoS for data transfer and SDN for novel data distribution and resource management. The Ops Team needs an improved training program so they can continue to remain current with evolving technologies. Working to consolidate the 50 central services with advanced monitoring and failover will improve system resilience and maintenance. To integrate new capabilities, some commercial outsourcing has been attempted (e.g. Duo for two factor authentication), but does not appear to be a general strategic item for XSEDE.

6.4 Recommendations

XSEDE should:

1) Highlight security, performance and capability successes delivered by the operations group within the final report. Efficiently handling tickets on behalf of SPs provides a clear value add for the SPs. Operations is a core and vital part of the XSEDE mission and needs to be highlighted accordingly near the front of the report.
2) Document the excellent balance of creating a shared governance and security model across XSEDE and SPs. This is a critical community activity and many lessons were clearly learned and need to be articulated.
3) Articulate how XSEDE employs or will employ DEVOPS. There was no mention of the industry term "development operations" (DEVOPS) or reference to "infrastructure as code" methodologies to enhance system wide configuration management (*e.g*. Chef, Puppet, Ansible) or continuous integration services (*e.g.* Jenkins) in the report or presentations. "Development" in this context is not traditional "software development", rather it is an integrated dev/test/deployment platform for systems infrastructure, improvements, and changes. Omitting DEVOPS from the operations strategy may put the complex XSEDE infrastructure in potential jeopardy, even more so with new advanced platforms such as JetStream and Comet. Modern, agile configuration management, and DEVOPS practices that leverage "infrastructure as code" techniques will position the XSEDE operations team for greater future success. Accordingly, an XSEDE wide DEVOPS strategy needs to be clearly articulated.
4) Develop a strong training program for the Ops Team so staff may continue to keep up to date with evolving technologies, especially when coupled with an agile DEVOPS strategy.
5) Highlight savings in Operations for from moving from a 24x7 support and helpdesk function to one that more accurately mirrors the actual demand. The significant reductions in costs are commended by the panel. XSEDE should continue monitoring system and support utilization for additional savings.

7 Advice to XSEDE for Structuring the Final Report

The review panel **commends the XSEDE PI and co-PIs on their leadership of this major infrastructure project**. XSEDE continues to enable science discovery in a secure environment that at a minimum could have taken much longer to develop. The panel also recognizes the **significant impact on multiple stakeholders including scientists, Service Providers, campuses, and research communities**, as well as its **outreach and community engagement efforts**. Infrastructure for the sake of infrastructure does not have much appeal, which is why XSEDE must continually put emphasis on the science. Furthermore, it is critical that XSEDE continue to engage with its stakeholders and the broader community.

The XSEDE team needs to approach the PY1-PY5 report with an understanding of its full potential, importance and gravity. They should consider it not as "yet another project report" but rather as a **nearly unique opportunity to document and provide concrete evidence for**

the power of both high-performance and data-intensive computation in frontier scientific discovery. Additionally, the team can articulate the **amazing value to the nation of the NSF cyberinfrastructure**, the **sweeping transformations** being led by this particular NSF project in diverse research communities and workforces, and the **impact that the work has had on education and workforce development**. Most other reports can either only assert need or refer to distant and poorly documented case studies, whereas the **XSEDE report can make everything concrete and immediate**. With NSCI beginning to unfold and with multiple imperatives to **grow rather than merely sustain the NSF CI**, this report and its companions are **very timely and have potentially huge significance**.

The report needs to **understand its intended audience(s)** and must have a **clear message**. Indeed, with so many stakeholders it is likely necessary to take multiple perspectives. Having expended $120M, **the project report must communicate both *impact* and *value for money* in both presentation and content**.

The **structure and order** of the Final Report on XSEDE should be **revised and shortened**. Once the readership of the report is identified, it should be **targeted** to that audience, recognizing that the final report will be public. The report should be **ordered with the most important sections first, and be very readable**. **Missing elements** include an **Executive Summary** (which the panel understands will be added), sufficient material to be **appreciated by someone unfamiliar with XSEDE**, the **science cases**, **impact on science and the community**, an **organization chart**, the **high-level WBS with budget/FTEs**, etc. The project report should **stand on its own**, and contain enough information for any reader to **capture the character, activities, approach, outcomes, and strategic impact of XSEDE in its first five years**.

The panel feels that the **current draft dwells too much on the negative and undersells success**. Details on **lessons learned** and perhaps some of the **established technical best practices** might best be moved into a **supplementary document or appendix** to keep the main report **more concise and focused primarily on outcomes and success**. Writers are cautioned to ensure that **the report does not fall into the "too long, didn't read" (TLDR) category**.

Science Cases: The section after the Executive Summary/Introduction needs to include a sample of **representative compelling science cases and research resulting from using XSEDE resources**. The five or six examples should demonstrate impacts on science, business, phylogenetic trees, flood predictions, etc. [LIGO has conveniently provided a second timely example.] These can be brief with more information in appendices. There should be **inclusion of the multidisciplinary nature of some of the research and the cross fertilization of ideas that has occurred**. Including reference to a secure operating structure and improvements over TeraGrid (for example, zero uncontained security incidents to date) can be included at the PI's discretion.

Broader impacts and intellectual merit. The report should **address broader impacts and intellectual merit in separate sections**. **Examples to emphasize** include XSEDE's accomplishments in encouraging diversity, XSEDE support for the rise in social science and humanities computing, as well as traditional science and engineering uses. Areas that have emerged over the PY1-PY5 period and that have been cultivated thanks to XSEDE support should be noted, including **the rise of data analytics, digital humanities**, and so on.

Some general suggestions:

- **Topic sentences**: People tend to read topic sentences and not whole paragraphs. Make sure the first sentence in each paragraph or section contains the most important thought.
- **Lists and tables**: XSEDE management should use bulleted lists and tables to make the report more readable when possible, especially if there is an underlying order to what is being discussed.
- **Charts, graphs, infographics**: Pictures are indeed worth thousands of words. For example, the pie chart (from the briefing) showing researchers and their supporting agencies is an excellent example. The graphic showing users of XSEDE resources by discipline speaks volumes.

7.1 Final Report Advice for Specific Sections

- **XCI**
  1. The description of the **delivered services in the final report is vague and extremely difficult to understand** and complicates evaluating their impacts. The list of services and plans presented in the answers to the questions at the end of day 1 of panel review was clarifying.
  2. The **presentation of Campus Bridges** in the final report needs to be refined and reviewed, with an eye on consistency and concise writing while documenting current and future impact. As currently written, the description of **Campus Bridges** is spread over pages 8-36, making it **difficult to understand what was delivered and its impact on campuses and SPs**.
  3. A **high-level summary** would be very helpful in understanding the distributions of usage modes of resources.

- **Operations**
  1. At least one **overarching unified success story needs to be articulated** to show that the resilience and reliability of the systems supported by Operations are essential to XSEDE's continued delivery success.
  2. **The panel suggests that a few anecdotes about the importance of Operations** would strengthen the final report. For example:
     a. Coordination/communication across SPs
     b. Including a list of XSEDE services that have **led to SP cost savings** (e.g., ticketing system, network tuning, software deployment and testing)
     c. Security

3. Where possible, quantify savings as FTEs/time/dollars, as well as the overall return on investment (ROI).

- **RAS**

The final report needs to:

1. Focus on **highlighting impact** (as a complement to KPIs), cost savings through centralized RAS (highlighting savings to other groups leveraging the service, for example) and ROI.
2. Make a more quantitatively compelling case in the report **to demonstrate the increasing demand for resources**, using trends in historical allocation requests, recommendations, and grants.
3. Include a table for acronyms as in the PY6 report.

- **ECSS**

1. It is imperative for the report to **communicate the multiple central roles of this expert staff**, and the need to establish long-term career paths to retain and recruit talented individuals. They are not fungible. Multidisciplinary skills take time to acquire, must be kept current, and are deeply rooted in constantly evolving academic communities in which trusted relationships must be established and maintained.
2. In the team's response to the overnight question #4 ("what were the **top five most significant accomplishments** … scientific discovery ..."), all five accomplishments were due to ECSS activities, but **none of these activities were mentioned in the draft report**. Telling these stories and clearly communicating the crucial role(s) of all XSEDE components and people contributing to these successes is essential. Without the full team, and the underlying CI, several of these stories would have had very different outcomes or timelines.

- **CEE**

1. Resolve the fairly significant **disconnect between "lessons learned" as written in the report compared to how they were presented at the reverse site visit**. The presentation was **better, more complete and told a story**.
2. The stories of **community outreach and community building are missing from the report**, and lost among KPIs and numbers. Even the **thread on broadening participation efforts, which the panel found captivating, is completely absent from the final report**. These anecdotes and trends are compelling, and should be included in the final report.