Overview of XSEDE Systems Engineering

John Towns
PI and Project Director, XSEDE
Executive Director, Science & Technology, NCSA
jtowns@ncsa.illinois.edu

XSEDE
Extreme Science and Engineering Discovery Environment

License terms

- Please cite as: Towns, John, *Overview of XSEDE Systems Engineering,* June 2015, [http://www.slideshare.net/jtownsil/overview-of-xsede-systems-engineering]
- ORCID ID: http://orcid.org/0000-0001-7961-2277
-
- This can be done under the following conditions: attribution – You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

XSEDE in Context

- XSEDE is an award made under the eXtreme Digital solicitation
  - TeraGrid Phase III: eXtreme Digital Resources for Science and Engineering (XD), NSF 08-571
- Consistent with NSF's vision and strategy statements

XSEDE – accelerating scientific discovery

XSEDE's Vision:

a world of digitally-enabled scholars, researchers, and engineers participating in multidisciplinary collaborations while seamlessly accessing computing resources and sharing data to tackle society's grand challenges.

XSEDE's Mission:

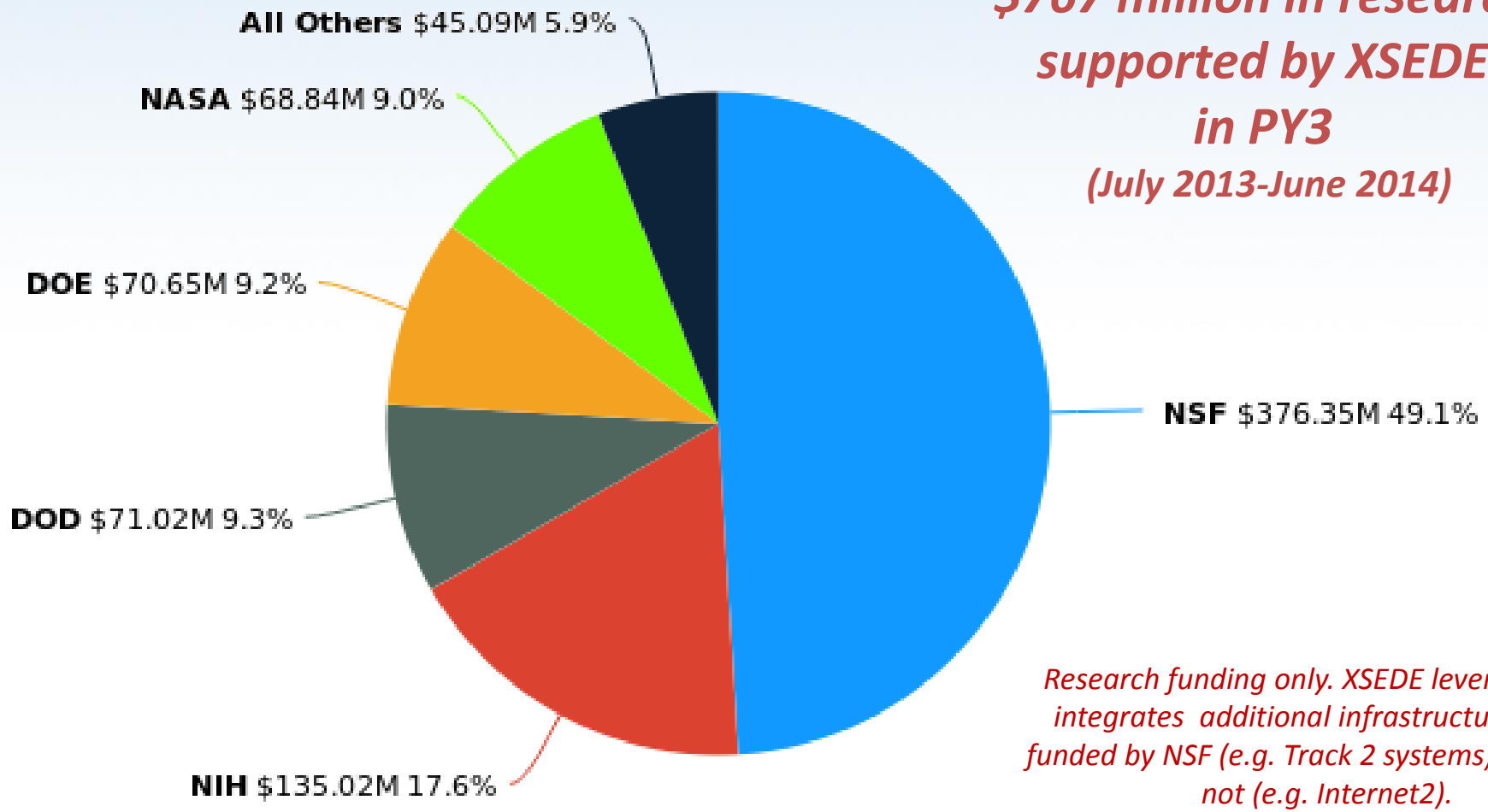
to substantially enhance the productivity of a growing community of scholars, researchers, and engineers through access to advanced digital services that support open research;

and to coordinate and add significant value to the leading cyberinfrastructure resources funded by the NSF and other agencies.

XSEDE Factoids: high-order bits

- 5 year, US$121M project
  - plus US$9M, 5 year Technology Investigation Service
    - separate award from NSF
  - option for additional 5 years of funding upon major review after PY3
- No funding for major hardware
  - coordinate, support and create a national/international cyberinfrastructure
  - coordinate allocations, support, training and documentation for >$100M of concurrent project awards from NSF
- ~112 FTE /~250 individuals funded across 20 partner institutions
  - this requires solid partnering!

Total Research Funding Supported by XSEDE in Program Year 3



$767 million in research supported by XSEDE in PY3

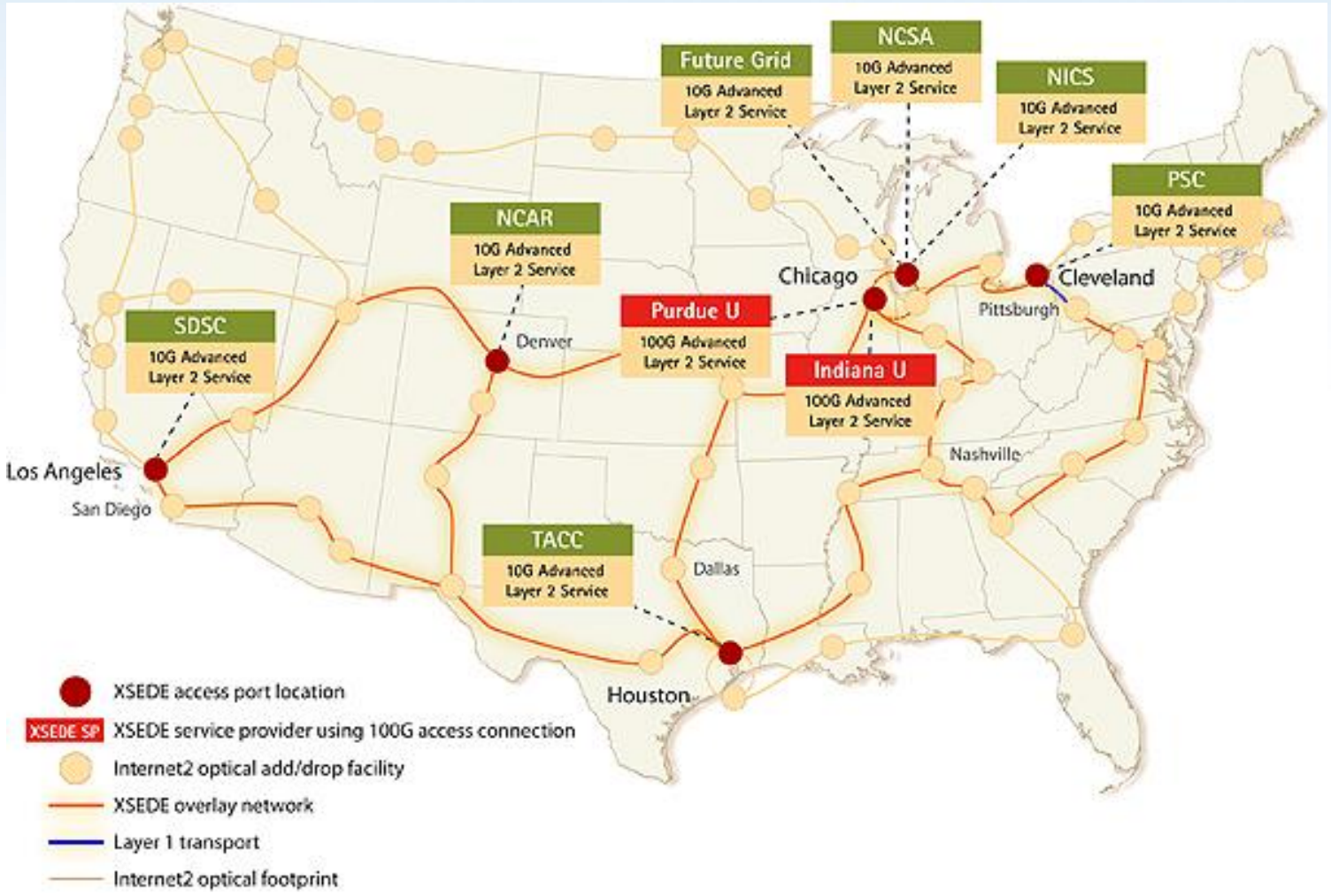
(July 2013-June 2014)

Research funding only. XSEDE leverages and integrates additional infrastructure, some funded by NSF (e.g. Track 2 systems) and some not (e.g. Internet2).

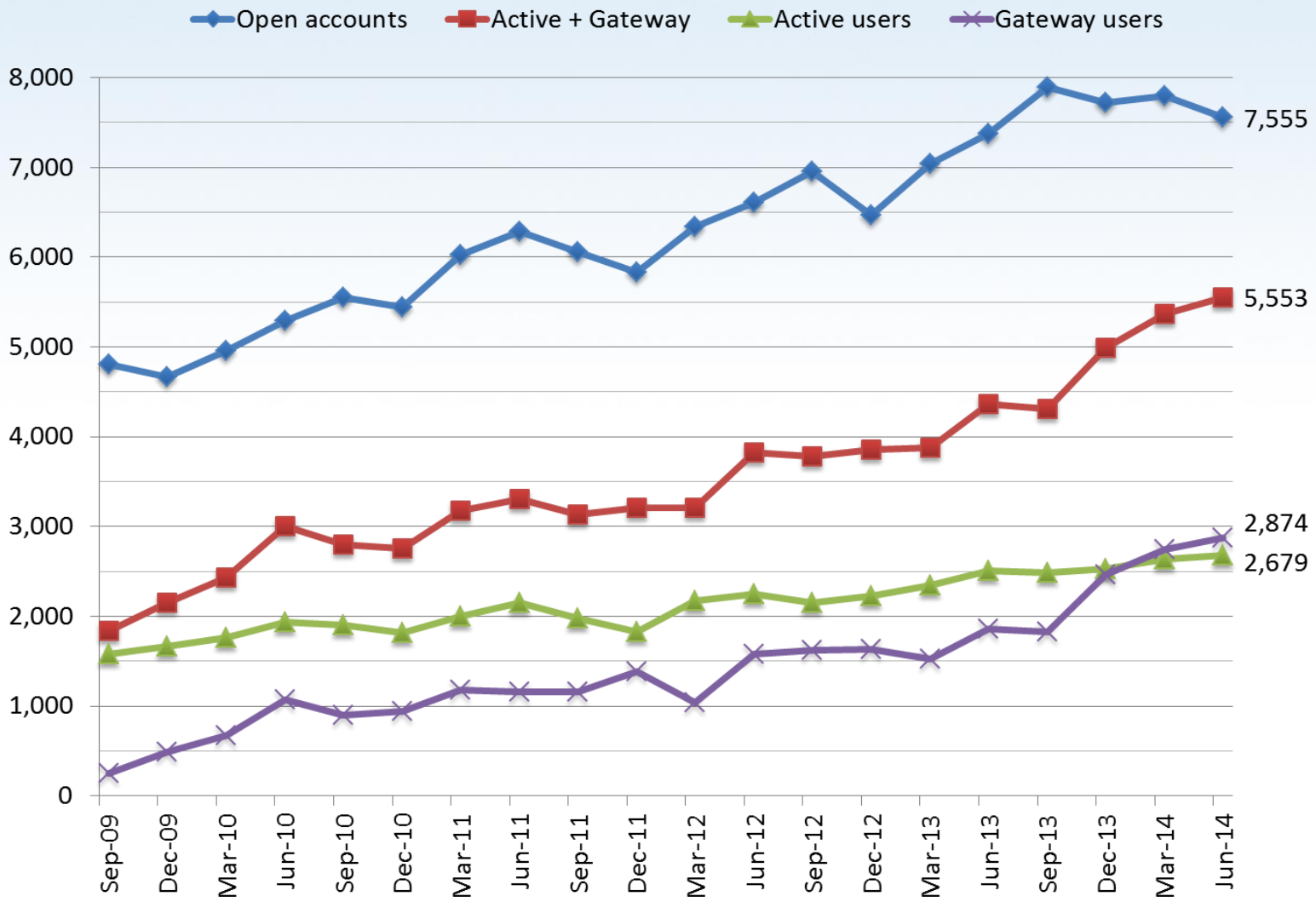
What is XSEDE?

- An ecosystem of advanced digital services accelerating scientific discovery
  - support a growing portfolio of resources and services
    - advanced computing, high-end visualization, data analysis, and other resources and services
    - interoperability with other infrastructures
- A virtual organization (partnership!) providing
  - dynamic distributed infrastructure
  - support services and technical expertise to enable researchers engineers and scholars
    - addressing the most important and challenging problems facing the nation and world
- More than just a project funded by the National Science Foundation
  - XSEDE is a path-finding experiment in how to develop, deploy and support e-science infrastructure

XSEDEnet – Using Internet2's AL2S Service



XSEDE Computational User Census



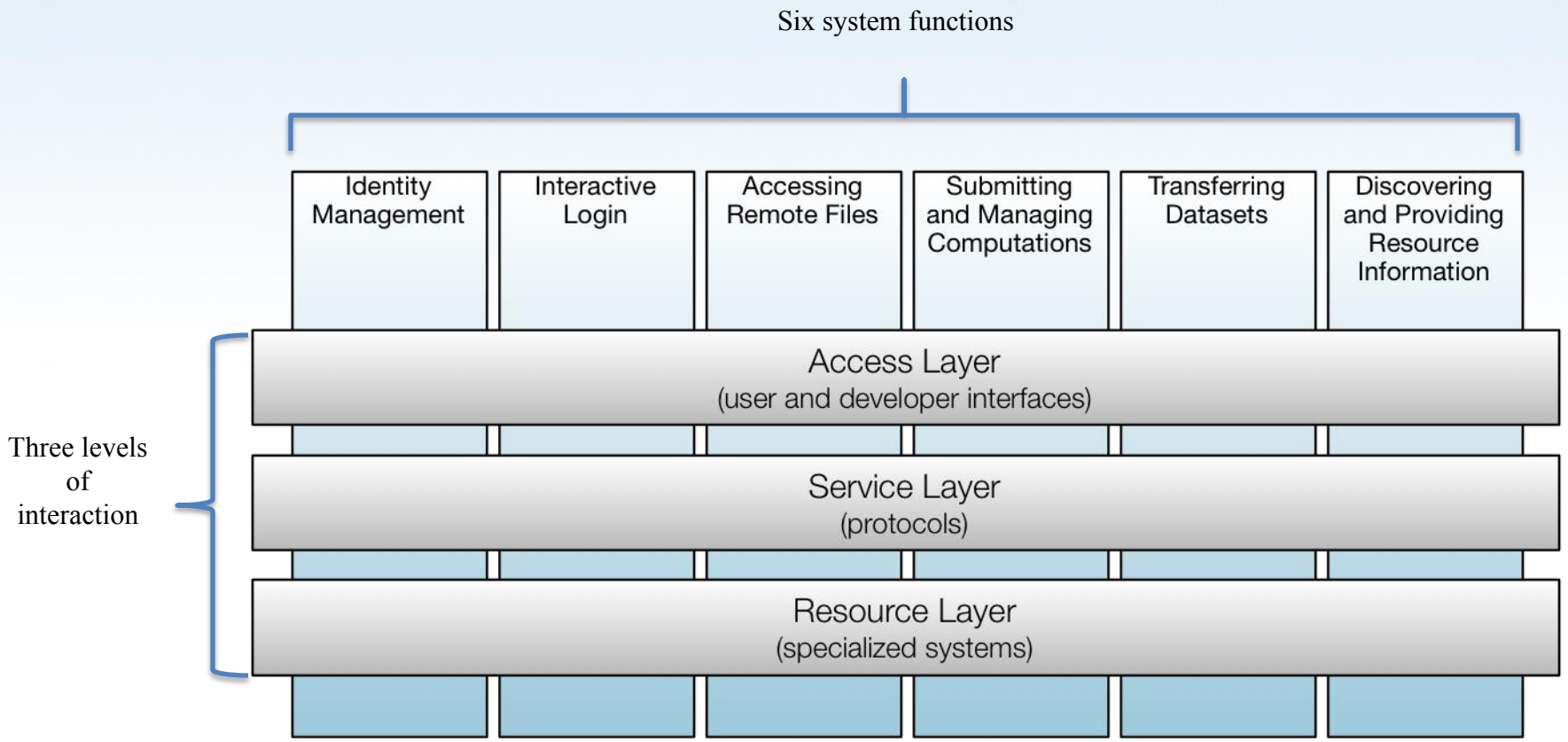
What do you mean by “Advanced Digital Services?”

- Often use the terms “resources” and “services”
  - these should be interpreted very broadly
  - most are likely not operated by XSEDE
- Examples of resources
  - compute engines: HPC, HTC (high throughput computing), campus, departmental, research group, project, …
  - data: simulation output, input files, instrument data, repositories, public databases, private databases, …
  - instruments: telescopes, beam lines, sensor nets, shake tables, microscopes, …
  - infrastructure: local networks, wide-area networks, …
- Examples of services
  - collaboration: wikis, forums, telepresence, …
  - data: data transport, data management, sharing, curation, provenance, …
  - access/used: authentication, authorization, accounting, …
  - coordination: meta-queuing, …
  - support: helpdesk, consulting, ECSS, training, …
  - And many more: education, outreach, community building, …

XSEDE – What does it do?

- Each individual XSEDE resource is designed to satisfy its user needs as fully as possible.
- XSEDE's system-wide functions support a more basic user need: *the need to have a relationship with the collection of resources as a whole and to maintain that relationship over time as their specific data and computing needs change and as system components change.*
- In addition to staff and programs, XSEDE provides and operates a system that supports this mission.

The XSEDE system



How did we get this system?

1. We operated the legacy system: Legacy TeraGrid system + early XSEDE additions/replacements
2. We talked to users and solicited specific use cases: Engaged a broad set of users and user representatives to document (~70) stakeholder use cases detailing what users need to do using XSEDE digital ecosystem.
3. We looked for the "80%" commonality: Identified ten "canonical" use cases that provide most of what XSEDE users and communities need and offer clear strategies for the rest.
4. We defined a foundational architecture: Organized the architecture around canonicals (six basic system functions) and applied our expertise to define three levels of interaction: user access interfaces (Web, CLI, API), service interfaces (protocols), and resource-specific interfaces.
5. We redesigned the *allocations* and *identity management* functions: Agreement between all impacted XSEDE working groups on updates to the federated allocations system and the federated security mechanisms.

From requirements to use cases

- **PY0/Proposal:** User needs documented as formal *Requirement Statements*.
- **PY1 Q1-Q3:** Pivoted to a stakeholder use case-driven process developed by Felix Bachmann from CMU's Software Engineering Institute. Now have much clearer way to describe user needs and to trace the architecture to those user needs.
- **PY1 Q4–PY2 Q3:** Engaged stakeholders to produce ~*70 use cases* in their project areas:
  - Data: Data Analytics, Data Management, and Visualization
  - Campus Bridging
  - Computing: HPC, HTC, and Scientific Workflows
  - Connecting Instrumentation
  - Federation and Interoperability
  - Science Gateways
- **PY2 Q3–PY4:** Distilled a new set of *10 canonical use cases* containing the common user needs that define the foundational architecture.

What is the problem?

- Campus Bridging use case 5

UCCB 5.0	Support for distributed workflows spanning XSEDE and campus-based data, computational, and/or visualization resources
Description	Enable distributed workflows – interactively or in batch mode - possibly spanning XSEDE and campus CyberInfrastructure resources, without user intervention after workflow is initiated

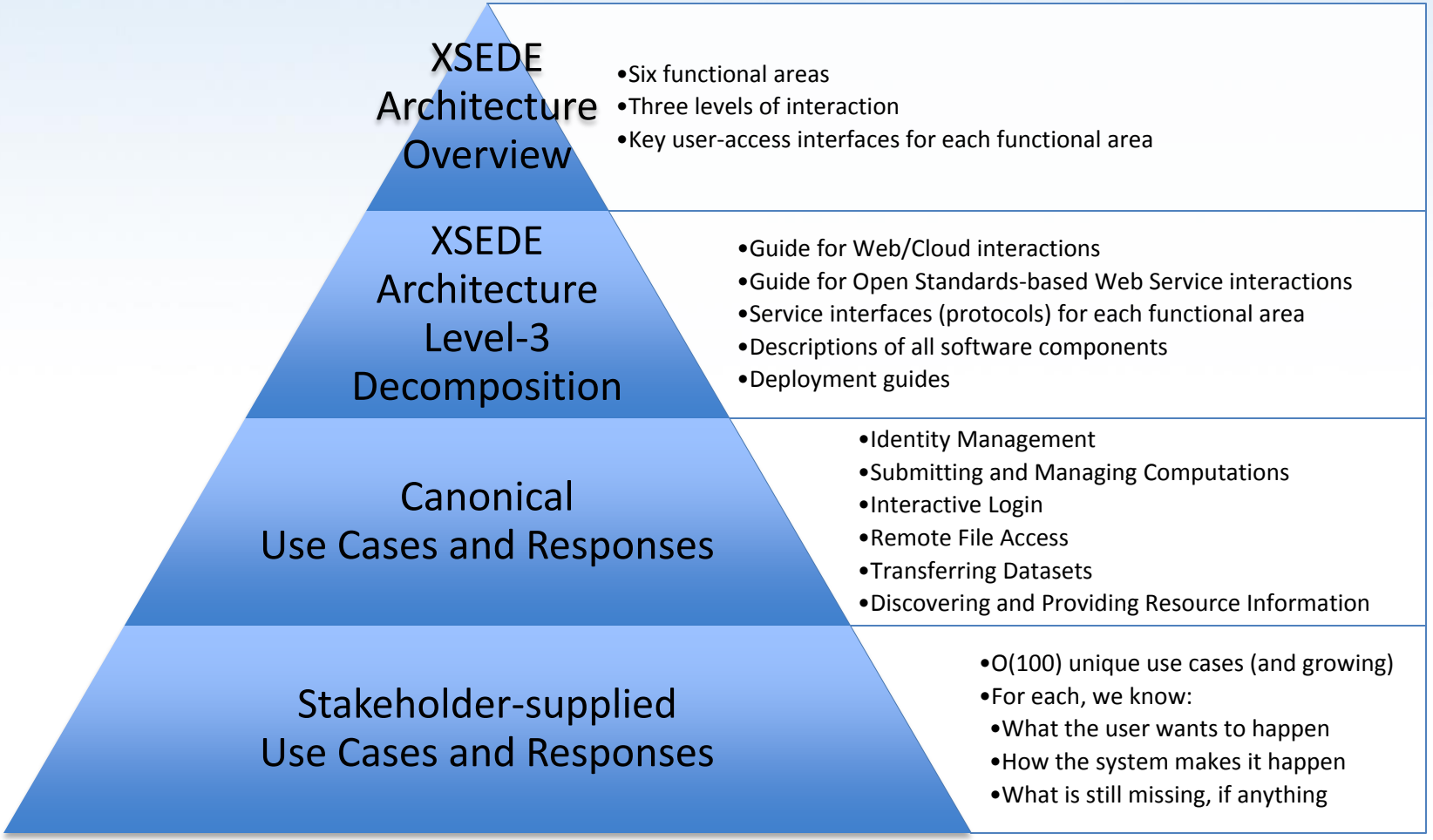
Scenario <QAS-CB5.1> Execute an automated workflow, possibly spanning XSEDE and campus cyberinfrastructure resources, without user intervention after workflow is initiated		
Scenario	A user wants to execute a workflow specified with an automated workflow system, using resources that may include XSEDE and campus (non-XSEDE) resources.	
Attribute	Reliability	
Attribute Concern	Probability that the workflow will complete without user intervention.	
Scenario Refinement	Stimulus	A host running workflow task fails causing the task to fail.
	Stimulus Source	The host fails.
	Environment	The workflow is specified in a language such as DAGMAN, and each DAGMAN task is correctly specified. There are sufficient resources are available to run the tasks even if one host fails. The user is properly authenticated before executing the workflow and the credentials are valid over the duration of the workflow. The user has sufficient allocations permission on resources to execute the workflow. There is sufficient disk space for jobs to execute.
	Artifact	A DAGMAN workflow engine
	Response	User expects the workflow to complete without their intervention. User might expect to be informed about the failure. Systems administrators may need to be notified about the failure.
	Response Measure	98% of workflows complete despite a host failing.

Scenario <QAS-CB5.2> Expired credentials while executing DAGMAN workflow		
Attribute Concern	User or host credential expired	
Scenario Refinement	Stimulus	A credential times out during the execution of the workflow.
	Stimulus Source	Time passes.
	Environment	The workflow is specified in a language such as DAGMAN, and each DAGMAN task is correctly specified. There are sufficient resources are available to run the tasks even if one host fails. The user has sufficient allocations permission on resources to execute the workflow. There is sufficient disk space for jobs to execute.
	Artifact	A DAGMAN workflow engine
	Response	The user is given the opportunity to re-authenticate and continue executing the workflow. Tasks that were executing will be restarted: checkpoint/restart of the task is the programmer's responsibility.
	Response Measure	99.9% No completed work done so far is lost, i.e., data generated by completed tasks, the graph state (tasked completed so far).

From use cases to architecture

Six system functions	Ten "canonical" use cases
Identity Management	Canonical 6.0 – authenticate to one or more SP resources, SP services, and XSEDE central services Canonical 9.0 – user & group management
Submitting and Managing Computations	Canonical 1.0
Transferring Datasets	Canonical 2.0
Accessing Remote Files	Canonical 3.0
Interactive Login	Canonical 4.0
Discovering and Providing Resource Information	Canonical 7.0 – subscribe for resource Information Canonical 8.0 – search for resource Information Canonical 11.0 – publish resource Information Canonical 12.0 – update resource Information

XSEDE architecture documents



From architecture to components

1. Use case-driven process: let the users drive the requirements
2. Prioritization: UREP-prioritized use cases
3. Traceability: traceability from use cases to operational capabilities
4. Documentation: SWOT analysis identified deficient documentation and led to a new L1-2 architecture description
5. Metrics: established KPIs to measure and monitor performance
6. Tuned reviews: streamlined Active Design Reviews to focus on sufficient implementation details, leaving detailed design for later engineering steps
7. Tuned testing: joint SD&I and Operations testing where appropriate to accelerate delivery
8. Improved user documentation: formalized user documentation preparation in the engineering process

Redesigned identity management

Goals:

- Support modern standards
- Sustainability
- Federate identities
- Science gateways
- Campus bridging

Work-in-progress through early 2016…

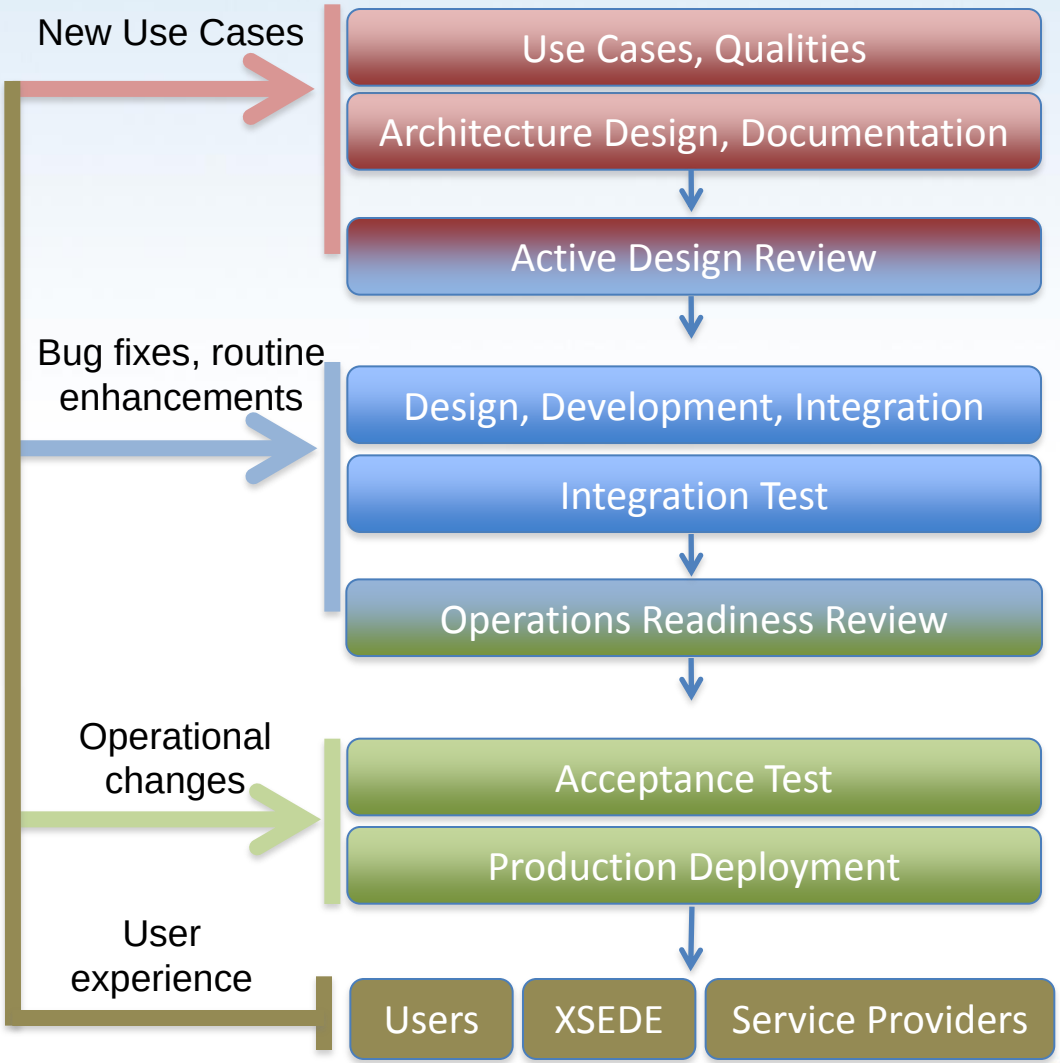
Globus Auth:

- Links identities from many sources
- Supports existing OAuth2, OpenID Connect services
- Robust funding model
- Large existing user base
- Already part of XSEDE's data transfer services

From components to a system

1. Operator coordination: Established Campus Bridging and Service Provider coordination. Conducting annual operations checklist reviews.
2. Engineering process: Established an end-to-end engineering process with S&SE, A&D, SD&I, and Operations collaboratively advancing XSEDE's digital ecosystem. This process includes planning and prioritizing, architecture and design, development/integration, significant testing phases, deployment, and retirement; and engage stakeholders in reviewing engineering products as they transition through phases.
3. Engineering process adoption: Selected elements of the engineering process used developing the XSEDE Resource Allocation Service (XRAS).
4. Engineering tools: Deployed a set of engineering software and services to manage use cases and engineering activities, distribute software, and publish (engineering) documents to the community.

Overview: An iterative engineering process



Area(s): A&D, S&SE

Implementation-ready architecture

Area(s): SD&I

Deployment-ready components

Area(s): Operations

Operational capabilities

User benefits from the engineering process

- ✓ Traceability from users (use cases) to operational capabilities
- ✓ Enhanced usability derived from architectural interoperability
- ✓ Better user documentation and developer documentation
- ✓ Service quality, scalability, and reliability

Sustainability benefits

- ✓ Written docs on what the system is supposed to do
- ✓ Written docs on how the system is supposed to work
- ✓ Record of key decisions (security choices, component evaluations, priorities)
- ✓ A solid basis for future work (by us and by others)

Ongoing challenges

- Managing engineering documents
- Stakeholder visibility into engineering process
- Pace of getting components through the engineering process
- Timely stakeholder participation of engineering reviews
- Resources to address big data, clouds, advanced SaaS, and other emerging capabilities

Looking ahead (years 6-10)

- Build on the foundational architecture
  - Emphasize key interface specs: Web (HTML, HTTP, CSS, JavaScript), TLS &X.509, OAuth2, OpenID Connect, REST
  - Continue leveraging Globus Auth and Transfer
  - Broaden use of sharing and group features
  - Keep executing (while refining) the engineering process
- Encourage and enable others to build as well
  - Documentation, training, examples

Priorities for PY6-PY10: Community Infrastructure

- Continue to evolve the XSEDE infrastructure
  - must provide support for and integration of "Track 2" resource
- Expose this architecture to the broader community
  - facilitate integration of broad range of services
    - provide discoverability
  - become the "connector of services" to support the research enterprise

Priorities for PY6-PY10: Toward "Sustainability"

- Developing services on offer to others: providing basic cyberinfrastructure services
  - expose services developed and put in place to operate the XSEDE environment
  - where necessary customize/extend for needs of other projects
  - charge incremental costs for operating/supporting services for other projects
- Objective is not to make money!
  - provides mechanism for other NSF project investments to leverage the XSEDE investment
  - can lead to significant cost saving across NSF CI investments
  - others can leverage this too: projects, institutions, regional consortia, …
- Pilot under way: NCAR with XRAS
  - will use XRAS to support allocation of NCAR resources
  - expressions of interest from some campuses

Meet me in St. Louis!

Save the date!

XSEDE15 WILL TAKE PLACE IN ST. LOUIS, JULY 26-30, 2015.

Thanks to you, XSEDE has become an established ecosystem of HPC resources. Help us take the next step.

The annual XSEDE conference brings together the extended community of individuals interested in advancing research cyberinfrastructure and integrated digital services for the benefit of science and society.

XSEDE

Extreme Science and Engineering Discovery Environment

xsede.org/xsede15 | #XSEDE15

Questions?


Our reach will forever
exceed our grasp, but,
in stretching our horizon,
we forever improve our world.
XSEDE
Extreme Science and Engineering
Discovery Environment