TeraGrid Usage Modalities

Slides prepared and presented by
Daniel S. Katz
d.katz@ieee.org
Director of Science, TeraGrid GIG
Senior Fellow, Computation Institute, University of Chicago &
Argonne National Laboratory
Affiliate Faculty, Center for Computation & Technology, LSU
Adjunct Associate Professor, ECE, LSU

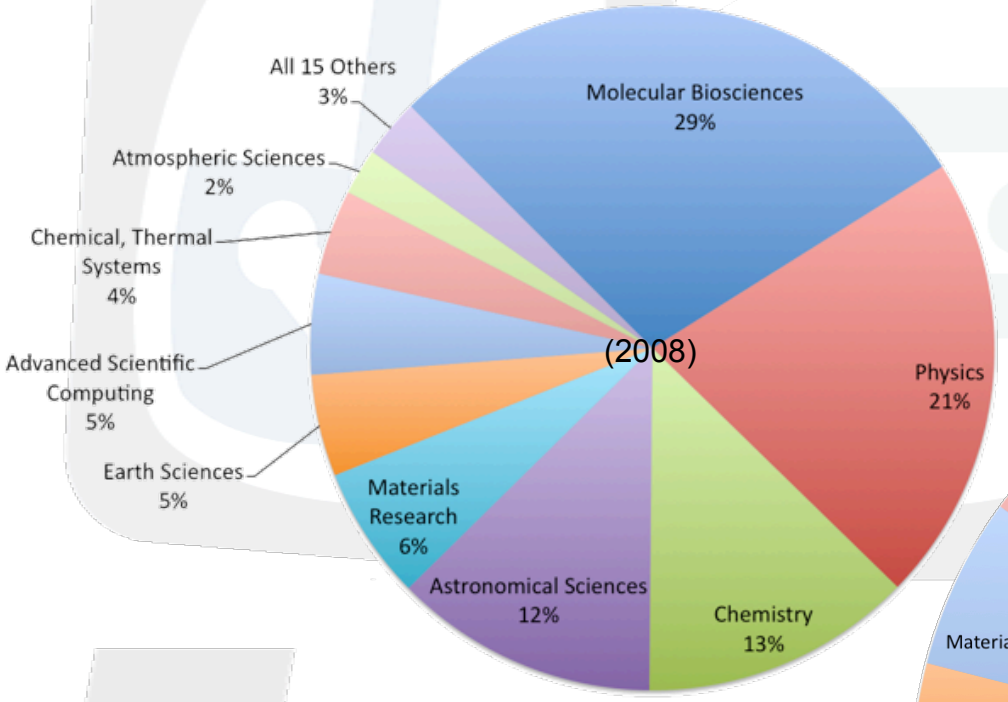
Collaborative with TeraGrid, specifically including:
David Hart (SDSC -> NCAR), Chris Jordan (TACC),
Amit Majumder (SDSC), J.P. Navarro (UC/ANL),
Warren Smith (TACC), John Towns (NCSA),
Von Welch (NCSA -> Indiana), Nancy Wilkins-Diehr (SDSC)

What is the TeraGrid

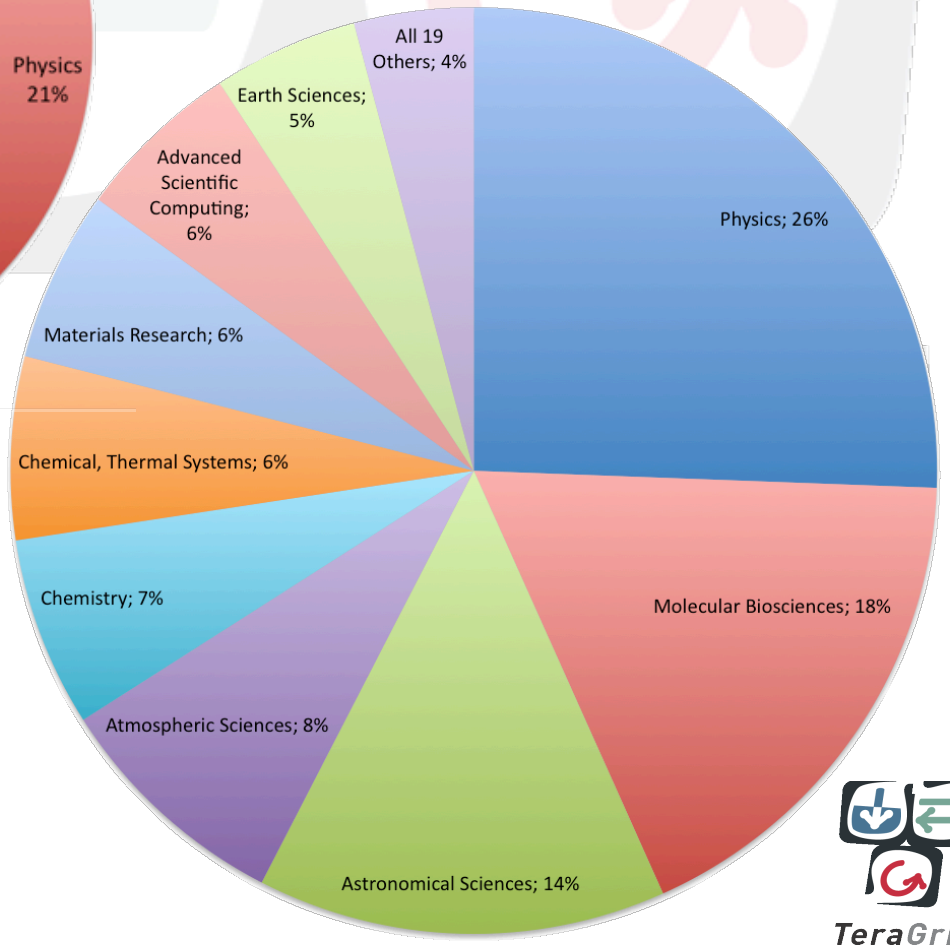
- World's largest (arguably) distributed cyberinfrastructure for open scientific research, supported by US NSF
- Integrated high performance computers (>2 PF HPC & >27000 HTC CPUs), data resources (>3 PB disk, >60 PB tape, data collections), visualization, experimental facilities (VMs, GPUs, FPGAs), network at 11 Resource Provider sites
- Freely allocated to US researchers and their collaborators through allocations
  - Researchers request time, peers review and determine merit, TG staff fit requests to resources
- Mission:
  - DEEP: provide powerful computational resources to enable research that can't otherwise be accomplished
  - WIDE: grow the community of computational science and make the resources easily accessible
  - OPEN: connect with new resources and institutions
- Integration: Single: portal; sign-on; help desk; allocations process; advanced user support; EOT; campus champions; accounts, allocations, usage database; coordinated middleware and other software

TeraGrid™

Who Uses TeraGrid (2009)



(2008)



4800 Active users
(from 361 institutions)
2300 new users
27B NUs delivered

TeraGrid™

How One Uses TeraGrid

POPS (for now)

User Portal

Science Gateways

Middleware

Command Line

RP 1

RP 2

RP 3

TeraGrid Infrastructure (Accounting, Network, Authorization,…)

TGCDB

Compute Service

Viz Service

Data Service

Slide modified from Dane Skow and Craig Stewart

Science Gateways

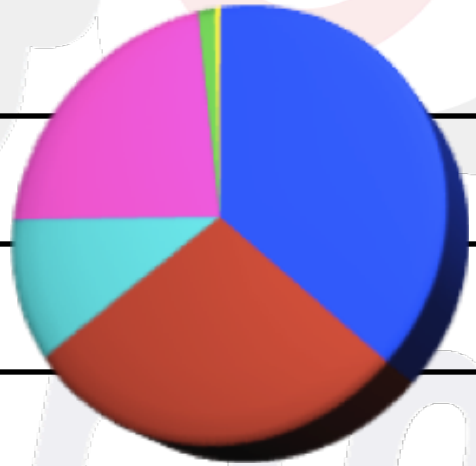
- An intermediate layer between users and providers
  - Provides scientific capabilities provided through familiar interface
    - Mostly web portal or web or client-server program
- Designed by communities; provides interfaces understood by those communities
  - Also provide access to greater capabilities (back end)
  - Without user understand details of capabilities
  - Scientists know they can undertake more complex analyses and that's all they want to focus on
  - TeraGrid provides tools to help developer
- Requires developer who knows community and resources and IT

Slide courtesy of Nancy Wilkins-Diehr

TeraGrid™

How TeraGrid Was Used (2006)

Usage Modality	Community Size (rough est. - number of users)
Batch Computing on Individual Resources	850
Exploratory and Application Porting	650
Workflow, Ensemble, and Parameter Sweep	250
Science Gateway Access	500
Remote Interactive Steering and Visualization	35
Tightly-Coupled Distributed Computation	10



TeraGrid

Usage Modalities

- Why do we want this data?
  - We want to know what our users are trying to do
    - Perhaps what they ideally want to do
    - At least how they are using our infrastructure now
  - We can use this information to develop new tools, improve existing tools, change operational practices and policies
- How was this data obtained in 2006?
  - Piecemeal at best
    - Including some data gathering, some guessing, some word-of-mouth, etc.
  - For gateways, TeraGrid asked each gateway how many users they had, so data was second-hand

Usage Modality Requirements

- Define a space of modalities, with each orthogonal modality as a dimension, with a set of possibly values
- Must be able to measure each value, ideally directly using the TeraGrid infrastructure, at least by inference from user intent
  - Use units that are common or translatable (core-hour, NU, TB/yr)
- Must be able to tie measurement of an activity (user, job, file, etc.) in one dimension to measurement of same activity in other dimension(s)
- Caveat
  - Fairly limited in use of CI currently – focused on running compute jobs, small amount on storing and moving data, small amount on human expertise, little-to-nothing about sensors, experimental facilities, etc.

TeraGrid™

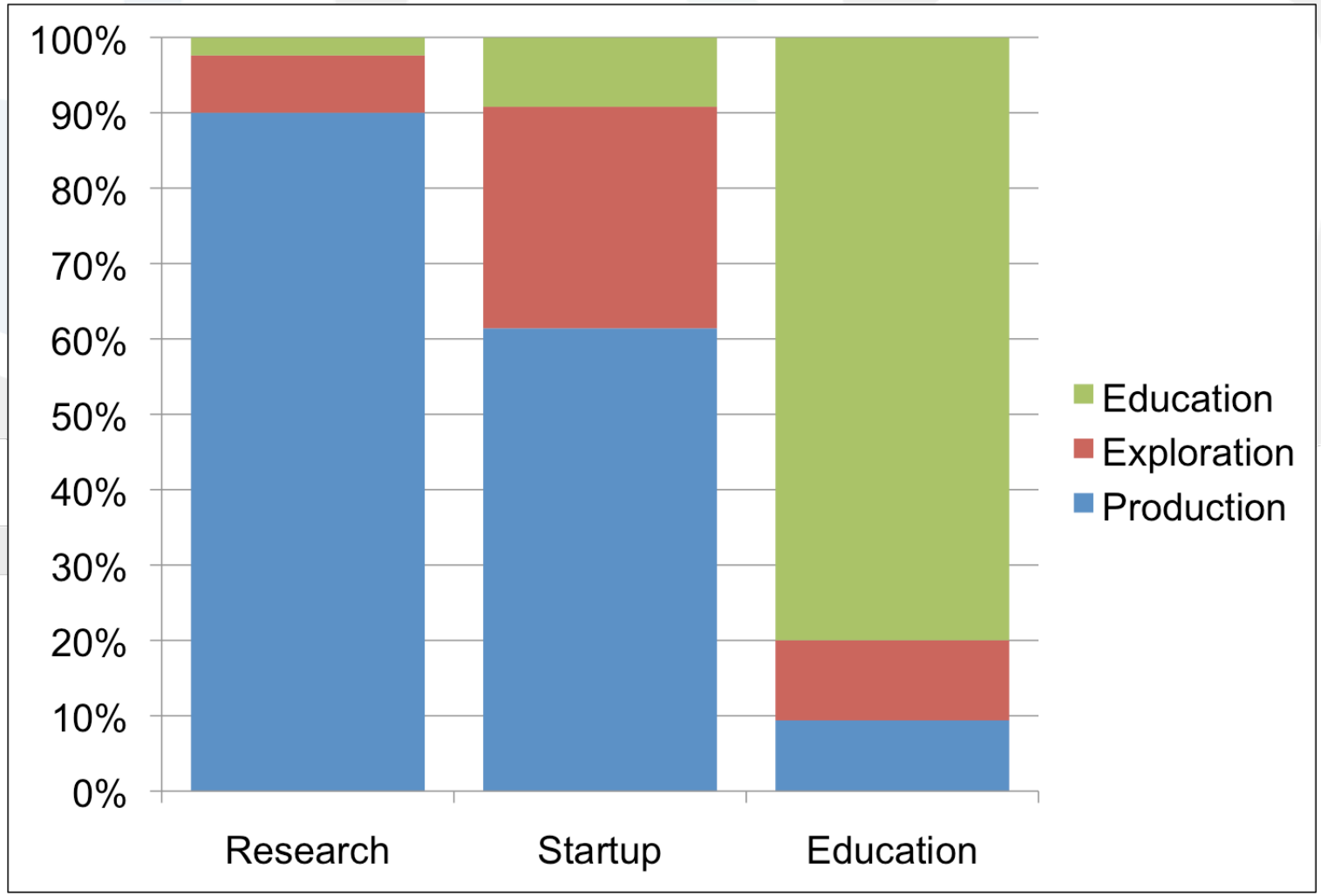
Current Draft Set of Modalities

- User intent
- When-to-run
- Submission mechanism
- Resources
- Job coupling
- Support
- Level of software development

User Intent

- Definition
  - Why the user is doing this
- Values
  - Production
  - Exploration/porting
  - Education
- How to measure
  - Ask the user at the time of their allocation request
    - Estimate fraction of each value you plan
  - Multiply actual runs by these fractions
- Issues
  - Not very accurate

User Intent



June 2010 – July 2011 User-reported planned intent vs. type of allocation

TeraGrid

When-to-run

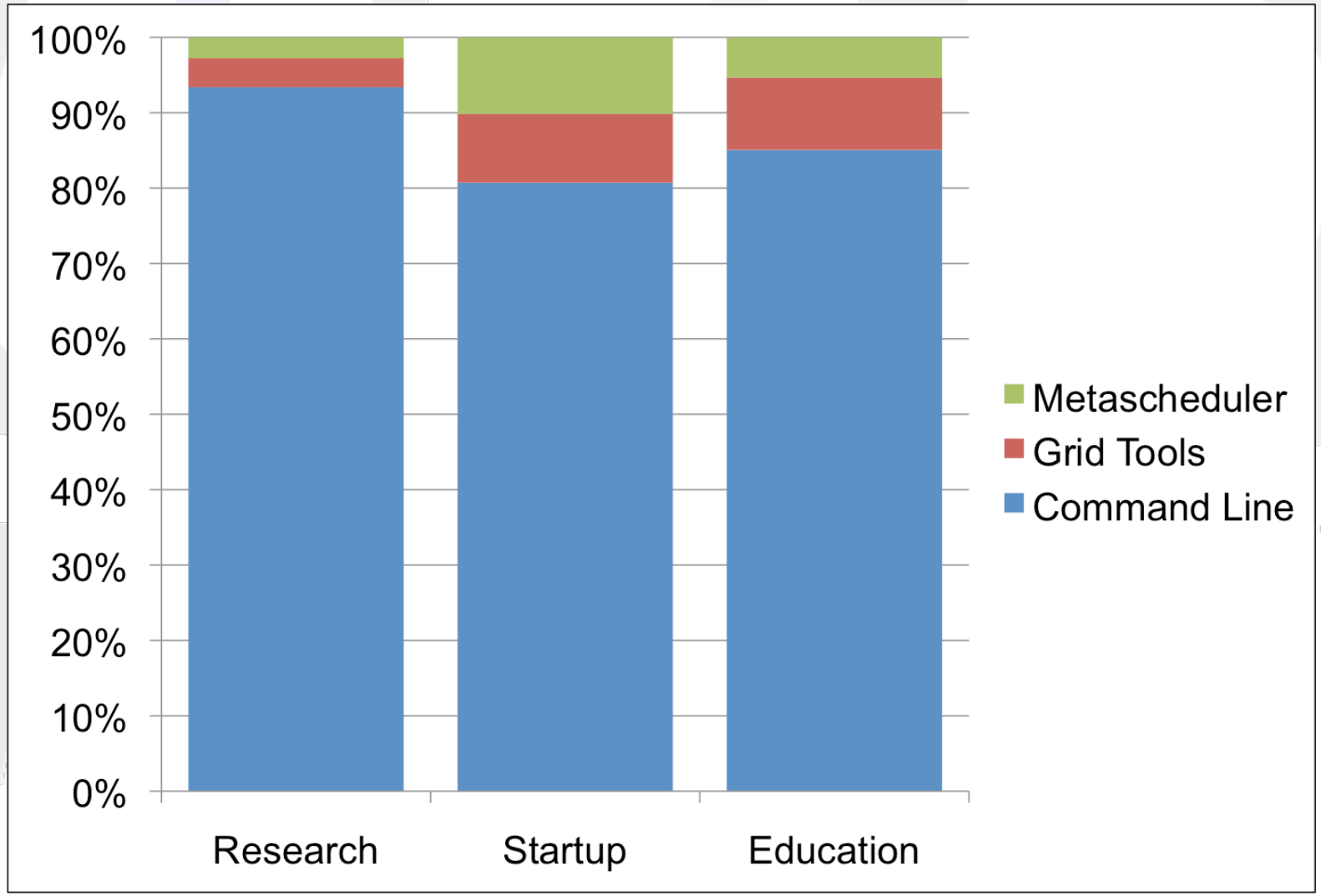
- Definition
  - When the user's activity should (needs to) start
- Values
  - Batch
  - Interactive
  - High Priority
  - Reservation
- How to measure
  - Local job scheduler measures all jobs – we assume by default all are batch
  - Subtract off everything else, as measured by tools/local job scheduler
- Issues
  - Potentially many tools that need to be modified to count activities

Submission Mechanism

- Definition
  - How the user's activity is started
- Values
  - Command line
  - Grid tools
  - Science gateways
  - Meta-scheduler
- How to measure
  - Tools report this to TGCDB
- Issues
  - Science gateways use grid tools underneath
  - Meta-scheduler is a grid tool?

TeraGrid™

Submission Mechanism



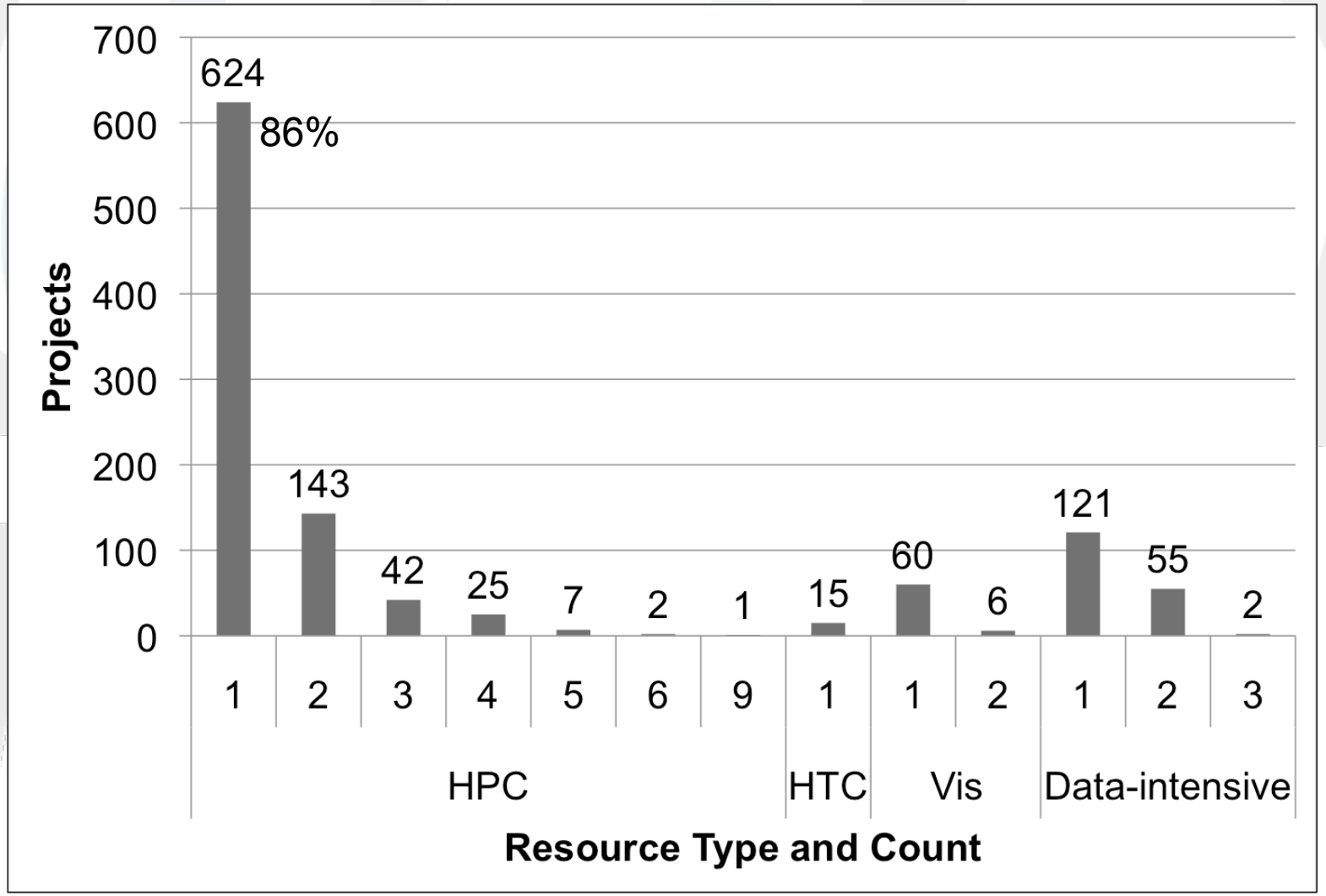
June 2010 – July 2011 User-reported planned submission mechanisms vs. type of allocation

Resources

- Definition
  - What resource types are needed to "run" the activity
- Values
  - 1 HPC resource, multiple HPC resources, 1 HTC resource, visualization resource, data-intensive resource, archival storage resource, multi-site storage resource, non-TG resource
- How to measure
  - Pull from TGCDB directly
- Issues
  - Incomplete?
  - Helps identify resource type needs, resource investment direction
  - Helps identify multi-resource interoperation needs

TeraGrid™

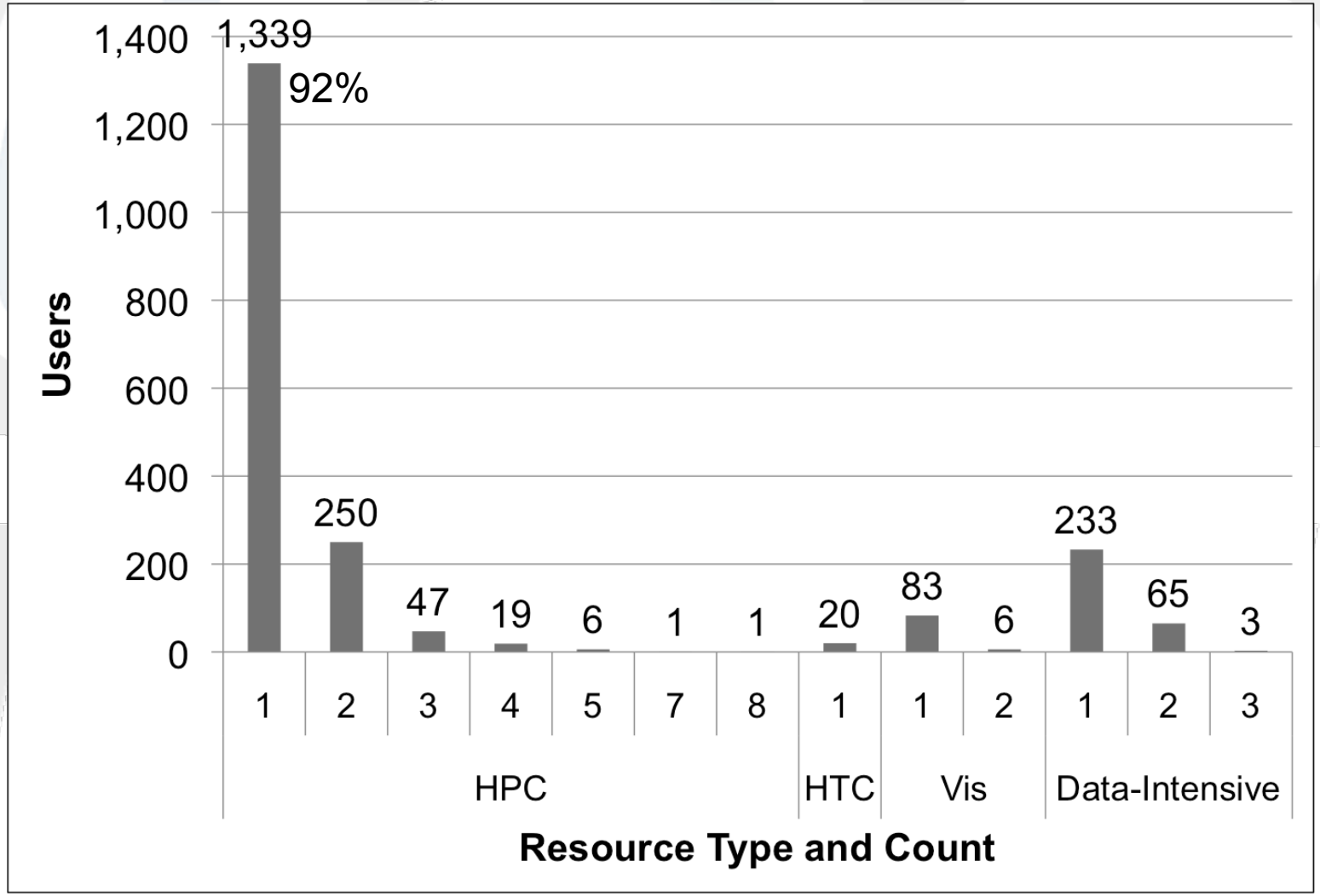
Resources - Projects



Sept. 2010 – Dec. 2010 Number of projects using various types and numbers of resources

TeraGrid

Resources - Users

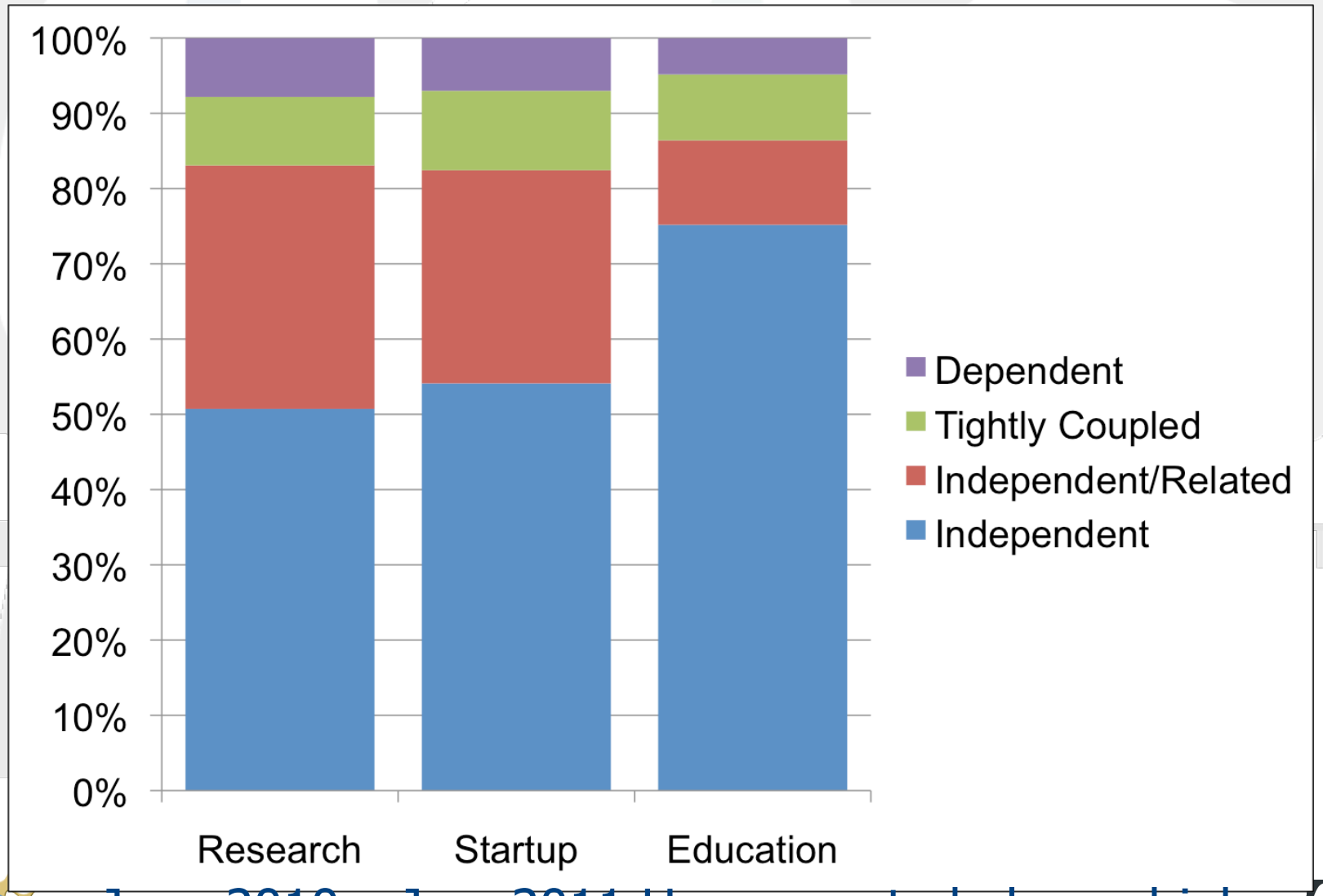


Sept. 2010 – Dec. 2010 Number of users using various types and numbers of resources

Job coupling

- Definition
  - How are multiple jobs (activity) coupled?
- Values
  - Independent/not coupled (e.g., single job)
  - Independent but related (e.g., an element of parameter sweep)
  - Tightly coupled (e.g., a part of distributed MPI or component application, jobs that must run simultaneously)
  - Dependent (e.g., an element of a workflow, jobs that depend on other jobs)
- How to measure
  - Ask users to estimate their fraction of the four types as part of their allocation request, multiply usage by these fractions
- Issues
  - Not very accurate

Job coupling



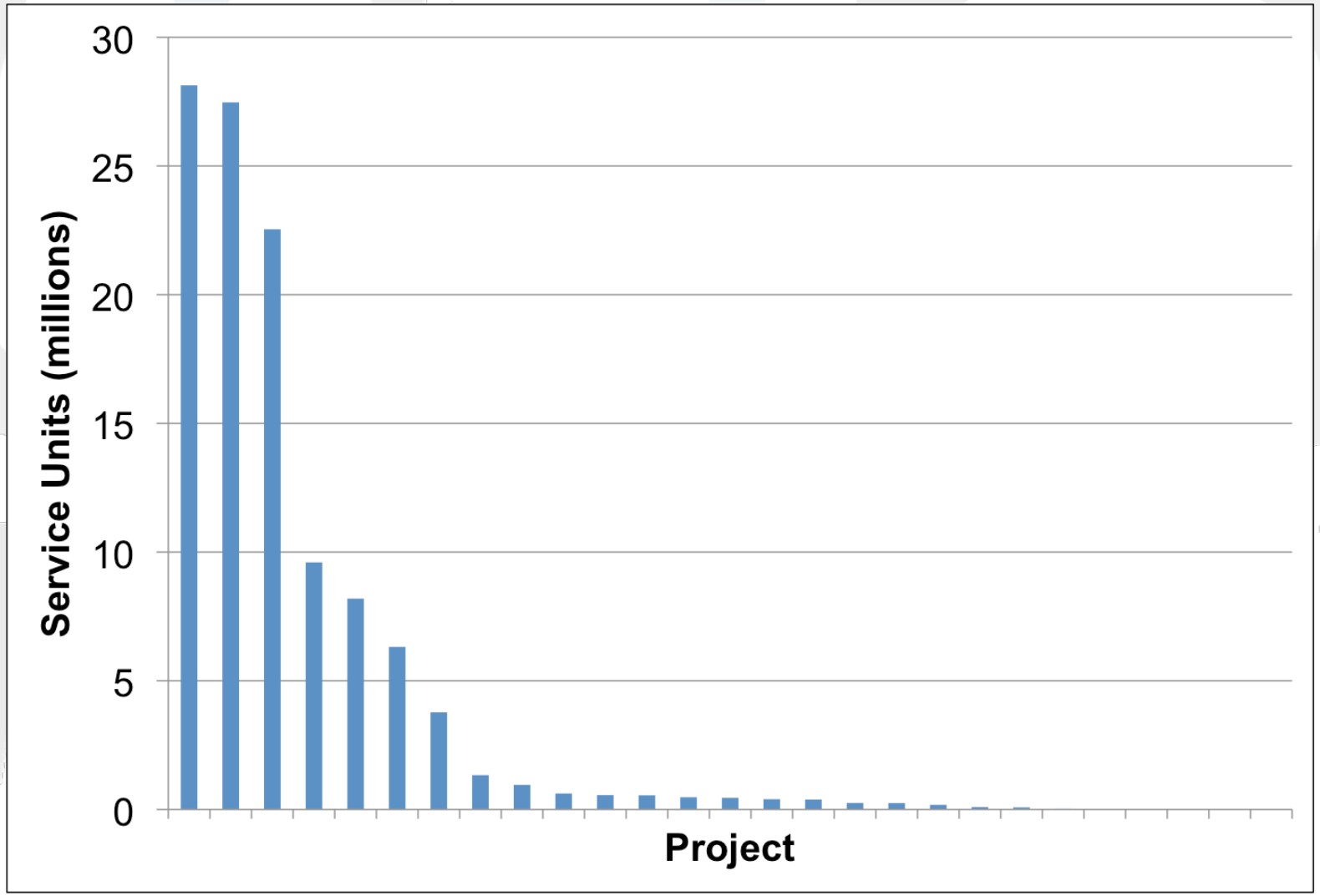
June 2010 – Jan. 2011 User-reported planned job coupling vs. type of allocation

Support

- Definition
  - Are we (TG) providing special support to the user? (Advanced user support is requested by the user as part of their allocation request, then peer reviewed, and possibly granted.)
- Values
  - Advanced support is being provided
  - Advanced support is not being provided
- How to measure
  - AUS is tracked in TGCDB – just pull this data
- Issues
  - Utility of this?
  - Other support we should measure?
    - Campus Champions?

TeraGrid™

Support



June 2010 – Jan. 2011 SUs consumed by projects supported by ASTAs during the ASTA

Level of software development

- Definition
  - How much software development the user has done
- Values
  - Custom (some work done by user)
  - Commodity (no work done by user)
- How to measure
  - Check path to binary executable?
  - Build table of exceptions by hand?
- Issues
  - Custom is not very specific
    - Probably, some additional values would be useful
  - Even though this is perhaps too simplistic, it's already hard to measure

Going further

- Do these modalities make sense in other venues?
  - Campuses? Other national/international infrastructures?
- Measuring them is much easier if supported by tools
  - Common needs help encourage tool developers to support us
- Does this lead to the need for new tools that need to be developed?
  - Or common policies across different infrastructures?

Final Thoughts

- Every HPC system is unique
  - Moving HPC applications without virtualization is extremely difficult
  - Virtualization hurts performance too much for many HPC applications
- Is this true of other resources – that they are unique?
- Do we want to hide this or expose it?
- Maybe worthwhile to think about:
  - What parts of CI can be made common with policies/tools?
  - What parts can't?
    - What do we do about this?

TeraGrid™