CISL

Computational & Information Systems Laboratory

Yellowstone allocations and writing successful requests

November 27, 2012

David L. Hart, CISL
dhart@ucar.edu

Welcome to the Petascale

- **Yellowstone environment**
- **Allocations opportunities at NWSC**
  - University, CSL, NCAR, and Wyoming-NCAR alliance
- **Tips for writing successful allocation requests**

Yellowstone

NWSC high-performance computing resource

- **Batch Computation**
  - 72,288 cores total – 1.504 PFLOPs peak
  - 4,518 IBM dx360 M4 nodes – 16 cores, 32 GB memory per node
  - Intel Xeon® E5-2670 (Sandy Bridge EP with AVX) – 2.6 GHz clock
  - 144.6 TB total DDR3-1600 memory
  - 28.9 Bluefire equivalents
- **High-Performance Interconnect**
  - Mellanox FDR InfiniBand full fat-tree
  - 13.6 GB/s bidirectional bw/node
  - <2.5 µs latency (worst case)
  - 31.7 TB/s bisection bandwidth
- **Login**
  - 6 IBM x3650 M4 Nodes; Intel Xeon® E5-2670 @ 2.6 GHz
  - 16 cores & 128 GB memory per node

GLADE

- **10.94 PB usable capacity → 16.42 PB usable (1Q2014)**

  Approximate initial file system sizes

scratch	≈ 5 PB	shared, temporary space
work	≈ 1 PB	individual work spaces
projects	≈ 3 PB	long-term, allocated space
collections	≈ 2 PB	RDA, CMIP5 data

- **Disk storage subsystem**
  - 4,560 3-TB NL-SAS drives
    - 76 IBM DCS3700 controllers
  - Add 2,280 3-TB NL-SAS drives (1Q2014)
- **GPFS NSD servers**
  - **90 GB/s** aggregate I/O bandwidth
    - 20 IBM x3650 M4 nodes
- **Data mover nodes (GPFS, GLADE-HPSS connectivity)**
  - 10-GbE & FDR interfaces; 4 IBM x3650 M4 nodes
- **High-performance I/O interconnect**
  - Mellanox FDR InfiniBand full fat-tree
  - 13.6 GB/s bidirectional bandwidth/node

Geyser and Caldera

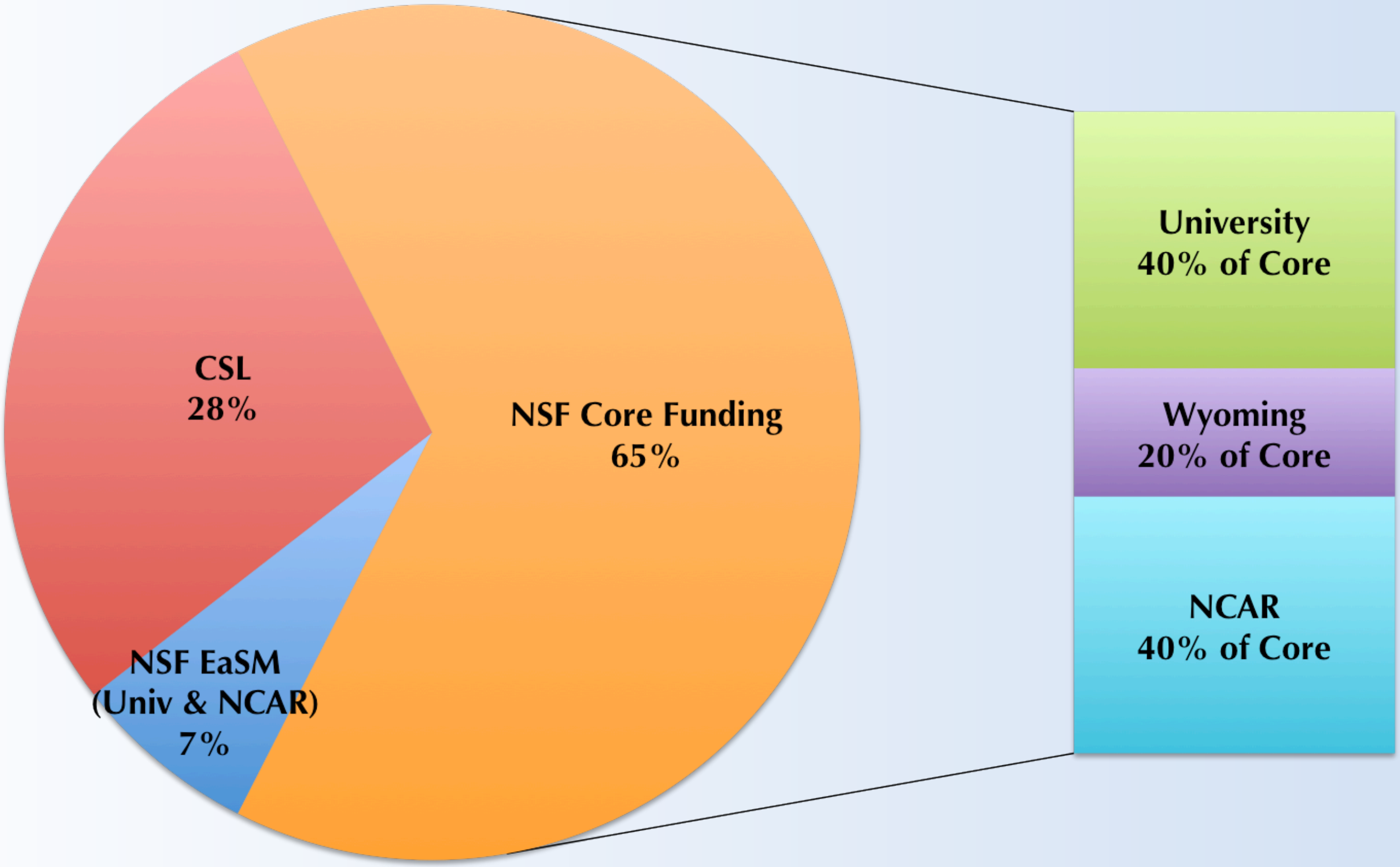
NWSC Data Analysis & Visualization Resource

- **Geyser: Large-memory analysis system**
  - 16 IBM x3850 nodes – Intel Westmere-EX processors
  - 40 cores, **1 TB memory,** 1 NVIDIA GPU per node
  - Mellanox FDR full fat-tree interconnect
- **Caldera: GPU computation/visualization system**
  - 16 IBM x360 M4 nodes – Intel Sandy Bridge EP/AVX
  - 16 cores, 64 GB memory per node
  - 2 NVIDIA GPUs per node
  - Mellanox FDR full fat-tree interconnect
- **Knights Corner system (early 2013 delivery)**
  - Intel Many Integrated Core (MIC) architecture
  - 16 IBM nodes
  - 16 Sandy Bridge EP/AVX cores, 64 GB memory
  - 1 Knights Corner adapter per node
  - Mellanox FDR full fat-tree interconnect

Yellowstone Software

- Compilers, Libraries, Debugger & Performance Tools
  - **Intel** Cluster Studio (Fortran, C++, performance & MPI libraries, trace collector & analyzer) 50 concurrent users
  - **Intel** VTune Amplifier XE performance optimizer 2 concurrent users
  - **PGI** CDK (Fortran, C, C++, pgdbg debugger, pgprof) 50 conc. users
  - **PGI** CDK GPU Version (Fortran, C, C++, pgdbg debugger, pgprof) for DAV systems only, 2 concurrent users
  - **PathScale** EckoPath (Fortran C, C++, PathDB debugger) 20 concurrent users
  - Rogue Wave **TotalView** debugger 8,192 floating tokens
  - **IBM** Parallel Environment (POE), including IBM HPC Toolkit
- System Software
  - **LSF-HPC** Batch Subsystem / Resource Manager
    - IBM has purchased Platform Computing, Inc., developers of LSF-HPC
  - Red Hat Enterprise **Linux** (RHEL) Version 6
  - IBM General Parallel Filesystem (**GPFS**)
  - Mellanox Universal Fabric Manager
  - IBM xCAT cluster administration toolkit

Yellowstone allocation opportunities



Yellowstone funding sources

Yellowstone will be capable of *almost 600 million core-hours per year,* compared to 34 million for Bluefire, and each Yellowstone core-hour is equivalent to 1.53 Bluefire core-hours.

Wyoming-NCAR Alliance

- ***Deadline: May 15, 2013***
  - *For early feedback on requests prior to panel submission, contact* **bshader@uwyo.edu**
- **Yellowstone resources**
  - 75 million core-hours per year
  - U Wyoming managed process
- **Large requests reviewed by "WRAP" twice per year**
- **Small and classroom allocations are also be available**
- **Activities must have substantial U Wyoming involvement**
  - Allocated projects must have Wyoming lead
  - Extended list of eligible fields of science
  - Eligible funding sources not limited to NSF
  - Actively seeking to increase collaborations with NCAR and with other EPSCoR states.

www.uwyo.edu/nwsc

University allocations

- *Next deadline: April 1, 2013*
- **Large allocations will continue to be reviewed and awarded twice per year**
  - Deadlines (usually) in March and September
  - Approx. 85 million core-hours to be allocated at each opportunity
- **Small allocations are also available**
- **Limited to atmospheric, ocean, and related sciences**
- **Small allocation will be up to 200,000 core-hours**
  - For researchers with NSF award—appropriate for benchmarking, preparation for large request
  - One-time allocations for grad students, post-docs, new faculty without NSF award
  - Classroom allocations for instructional use
- ***www2.cisl.ucar.edu/docs/allocations/university***

Tips for writing successful allocation requests

Allocation requests

- **Not just HPC, but DAV, HPSS, GLADE allocations**
  - Non-HPC resources ≈ 1/3 procurement cost
  - Ensure that use of scarce and costly resources are directed to the most meritorious projects
- **Balance between the time to prepare and review requests and the resources provided**
  - Minimize user hurdles and reviewer burden
  - Build on familiar process for requesting HPC allocations
- **Want to identify projects contributing to the NWSC Community Scientific Objectives**
  - *www2.cisl.ucar.edu/resources/yellowstone/science*

General submission format

- ***Please see specific opportunities for detailed guidelines!***
- **Five-page request typically**
  - A. Project information (title, lead, etc.)
  - B. Project overview and strategic linkages
  - C. Science objectives
  - D. Computational experiments and resource requirements (HPC, DAV, and storage)
- **Supporting information**
  - E. Multi-year plan (if applicable)
  - F. Data management plan
  - G. Accomplishment report
  - H. References and additional figures

Tips and advice (part 1)

- **Remember your audience: Computational geoscientists from national labs, universities and NCAR**
  - Don't assume they are experts in *your* specialty
- **Be sure to articulate relevance and linkages**
  - Between funding award, computing project, eligibility criteria, and NWSC science objectives (as appropriate)
- ***Don't submit a science proposal***
  - Describe the science in detail sufficient to justify the computational experiments proposed
  - Panel is not re-reviewing the science

Tips and advice (part 2)

- **Most of the request should focus on computational experiments and resource needs**
  - *Effective methodology:* Are you using the right computational tools for the job?
  - *Appropriateness of experiments:* Are the proposed experimental configurations necessary and sufficient to answer the scientific question? Are all key experimental parameters justified?
  - *Efficiency of resource use:* Are the resources being used efficiently for the proposed methodologies and experiments?
- **The amount of requested resources should be clearly and explicitly calculated based on the justification for the three preceding criteria**
  - E.g., University guidelines recommend a table with one row per experimental configuration

Justifying resource needs

- ***HPC*** **— familiar, if you've requested compute time elsewhere**
  - Cost of runs necessary to carry out experiment, supported by benchmark runs or published data
  - Yellowstone allocations will be made in "core-hours" (not GAUs).
  - Request a small allocation to conduct actual Yellowstone benchmark runs
    - Reasonably justified estimates based on runs from other systems will also be accepted.
    - Yellowstone core-hours can be calculated as: # of nodes x 16 x job duration (in hours).
- ***DAV*** **— will be allocated, similar to HPC practice**
  - Simple justification for standard interactive use: # users x 5,000 core-hours
  - Small allocations will get up to 5,000 core-hours upon request
  - Allocation review will focus on larger needs associated with batch use
    - E.g., for projects conducting GPGPU code development and testing

GLADE resource requests

- ***All disk is not provisioned equally***
- **Allocations only needed for *project space***
  - No need to detail use of scratch, work spaces
- **Describe why project space is essential**
  - That is, why scratch or work space insufficient
    - Show that you are aware of the differences
- **Relate the storage use to your workflow and computational plan**
  - Projects with data-intensive workflows should show they are using resources efficiently

HPSS resource requests

- **Goal: Demonstrate that HPSS (tape archive) use is efficient and appropriate**
  - Not store-everything-forever in a “data coffin”
  - Not treating it as a temporary file system
- **Note: Tape use up to 20 TB available upon request**
- **Explain new data to be generated**
  - Relate to computational experiments proposed
  - Describe scientific value/need for data stored
- **Justify existing stored data**
  - Reasons for keeping, timeline for deletion
- **Data management plan: Supplementary information**
  - Additional details on the plans and intents for sharing, managing, analyzing, holding the data

Summer Internships in Parallel Computational Science

Students work in NCAR's Supercomputing Lab with mentors on challenging R&D projects

SIParCS Class of 2011

- **11-week summer internship program**
  - May 20 – August 2, 2013
  - *Application deadline: Feb. 1, 2013*
- **Open to**
  - Upper division undergrads
  - Graduate students
- **In disciplines such as**
  - Computer Science and Software Engineering
  - Mechanical Engineering
  - Applied Math and Statistics
  - Earth System Science
- **Support**
  - Travel, housing
  - 11 weeks salary
  - Conference travel and publication costs
- **Number of interns selected**
  - Approx. 10-12

For more information:

http://www2.cisl.ucar.edu/siparcs

Computational & Information Systems Laboratory

http://www2.cisl.ucar.edu/resources/yellowstone

http://www2.cisl.ucar.edu/docs/allocations

cislhelp@ucar.edu *or* dhart@ucar.edu

QUESTIONS?