HPC Allocations at the Oak Ridge Leadership Computing Facility

Rebecca Hartman-Baker
hartmanbakrj@ornl.gov

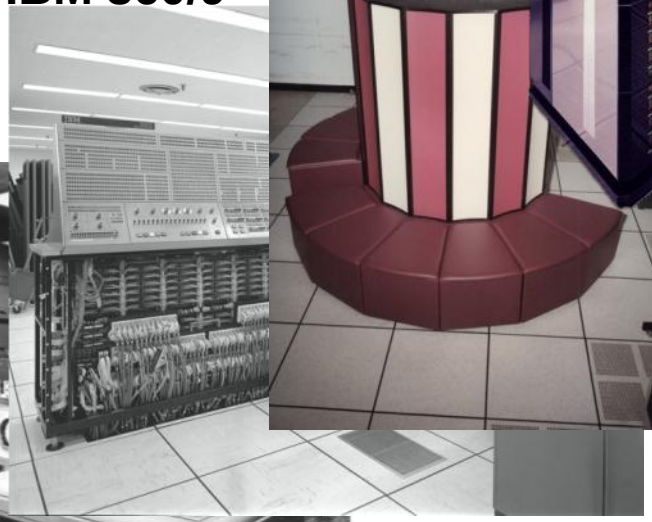
Judith Hill
hilljc@ornl.gov

ORNL has a long history in High Performance Computing

ORNL has had 20 systems on the TOP500® SUPERCOMPUTER SITES lists

1954
ORACLE

1969
IBM 360/9



1985
Cray X-MP

1992-1995
Intel Paragons

1996-2002
IBM Power 2/3/4

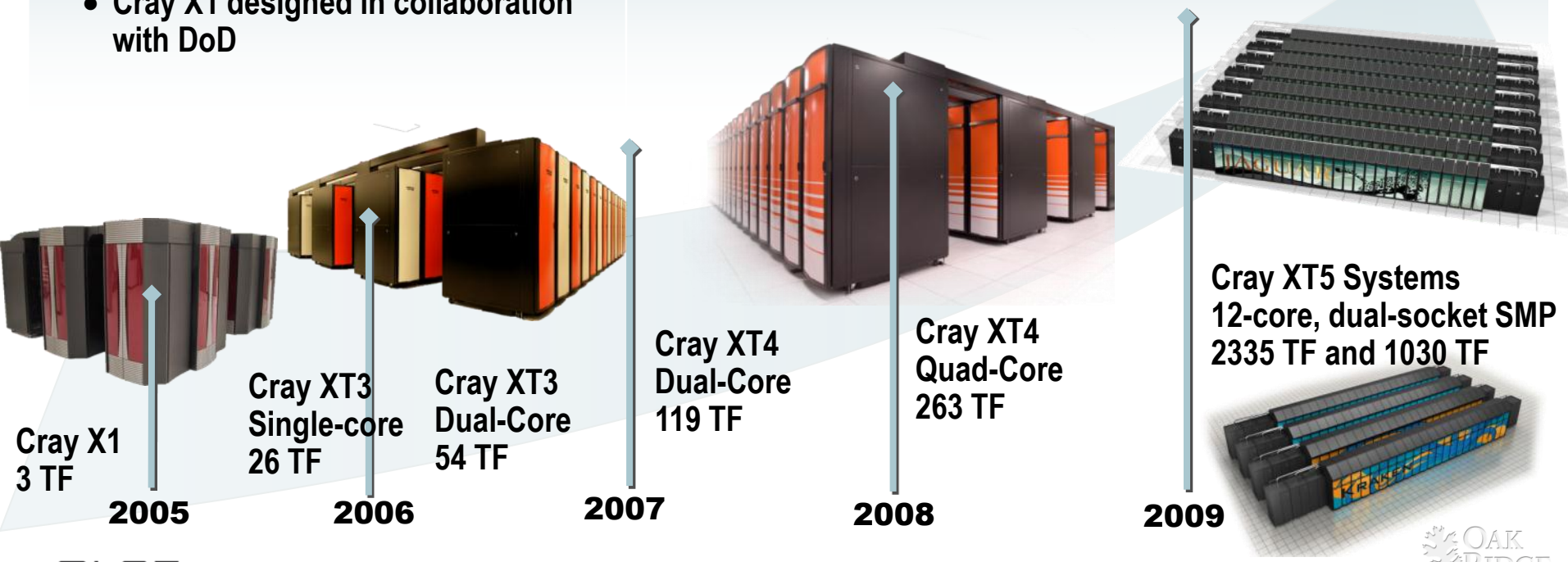
2007
IBM Blue Gene/P

2003-2005
Cray X1/X1E

We have increased system performance by 1,000 times since 2004

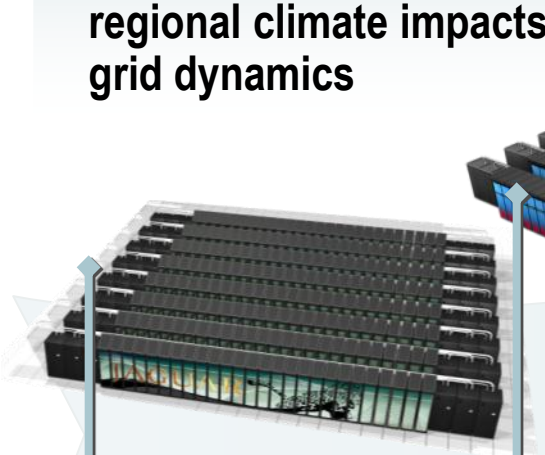
Hardware scaled from single-core through dual-core to quad-core and dual-socket , 12-core SMP nodes	Scaling applications and system software is the biggest challenge

• NNSA and DoD have funded much of the basic system architecture research
  • Cray XT based on Sandia Red Storm
  • IBM BG designed with Livermore
  • Cray X1 designed in collaboration with DoD
	• DOE SciDAC and NSF PetaApps programs are funding scalable application work, advancing many apps
	• DOE-SC and NSF have funded much of the library and applied math as well as tools
	• Computational Liaisons key to using deployed systems



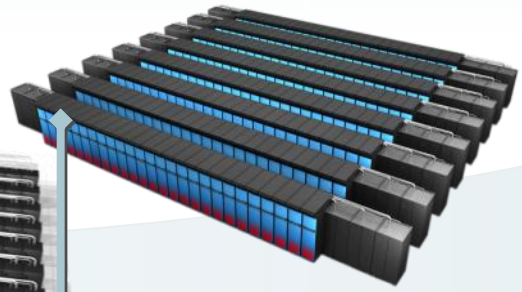
Our science requires that we advance computational capability 1000x over the next decade

Mission: Deploy and operate the computational resources required to tackle global challenges	Vision: Maximize scientific productivity and progress on the largest scale computational problems
• Deliver transforming discoveries in climate, materials, biology, energy technologies, etc. • Ability to investigate otherwise inaccessible systems, from regional climate impacts to energy grid dynamics	• Providing world-class computational resources and specialized services for the most computationally intensive problems • Providing stable hardware/software path of increasing scale to maximize productive applications development



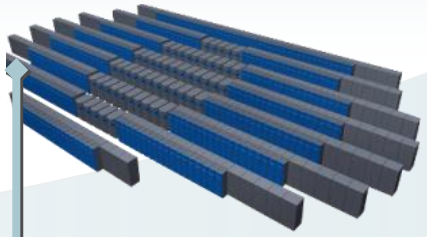
Cray XT5 2+ PF
Leadership system for science

2009



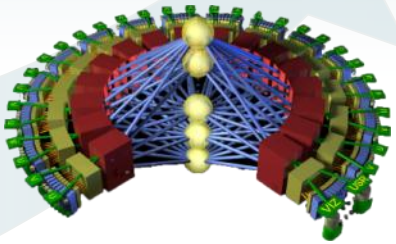
OLCF-3: 10-20 PF
Leadership system with some HPCS technology

2012



OLCF-4: 100-250 PF
based on DARPA HPCS technology

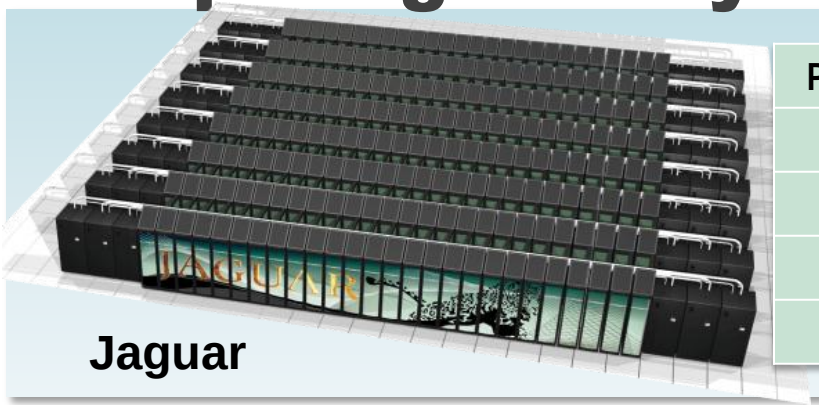
2015



OLCF-5: 1 EF

2018

Today, we have the world's most powerful computing facility



Jaguar

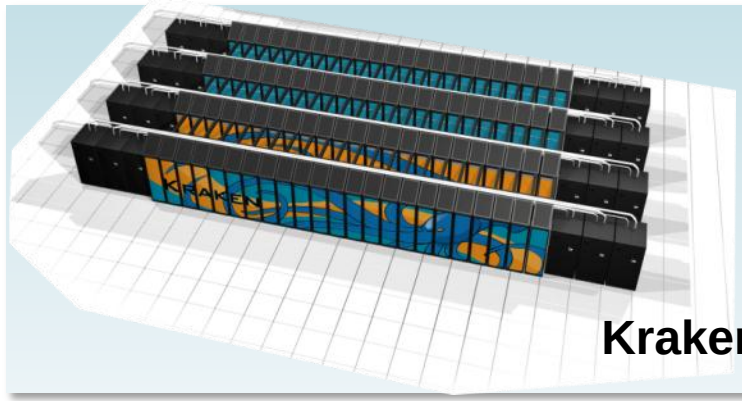
Peak performance	2.33 PF/s
Memory	300 TB
Disk bandwidth	> 240 GB/s
Square feet	5,000
Power	7 MW

Nov. 2010

TOP500 SUPERCOMPUTER SITES

#2

Dept. of Energy's most powerful computer



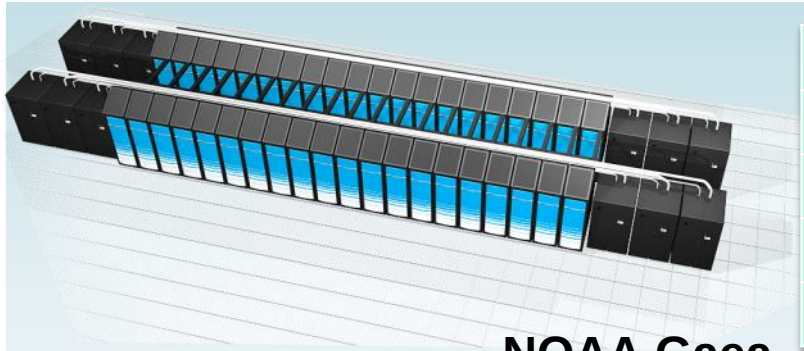
Kraken

Peak performance	1.2 PF/s
Memory	147 TB
Disk bandwidth	> 50 GB/s
Square feet	2,300
Power	3 MW

TOP500 SUPERCOMPUTER SITES

#8

National Science Foundation's most powerful computer



NOAA Gaea

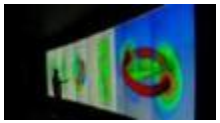
Peak Performance	1.1 PF/s
Memory	248 TB
Disk Bandwidth	104 GB/s
Square feet	1,600
Power	2.2 MW

TOP500 SUPERCOMPUTER SITES

#32

National Oceanic and Atmospheric Administration's most powerful computer

OLCF Center-wide file system: Spider



Everest Powerwall

Remote Visualization Cluster

End-to-End Cluster

Application Development Cluster

Data Archive 25 PB

HPSS

Scalable I/O Network (SION)
4x DDR Infiniband Backplane Network

XT5

Other Resources

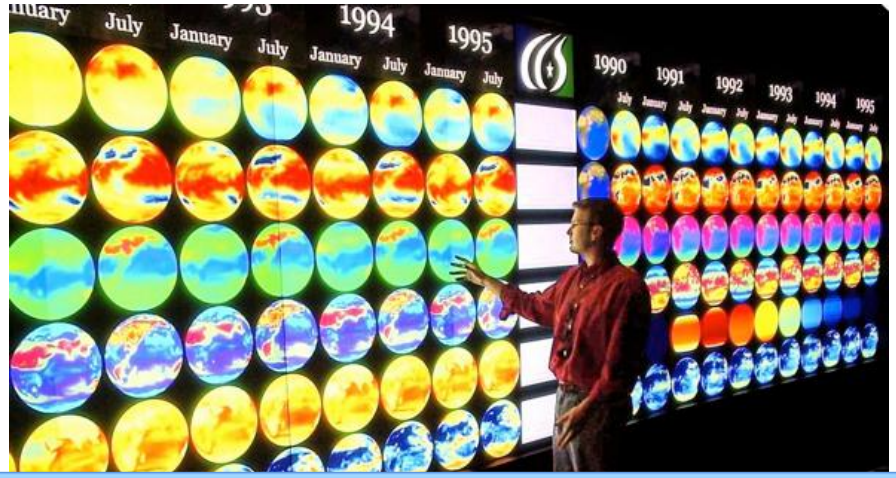
Login

Spider

ORNL's External login nodes and shared storage provide a single entry for users into a cluster of supercomputers

OLCF Visualization Resources

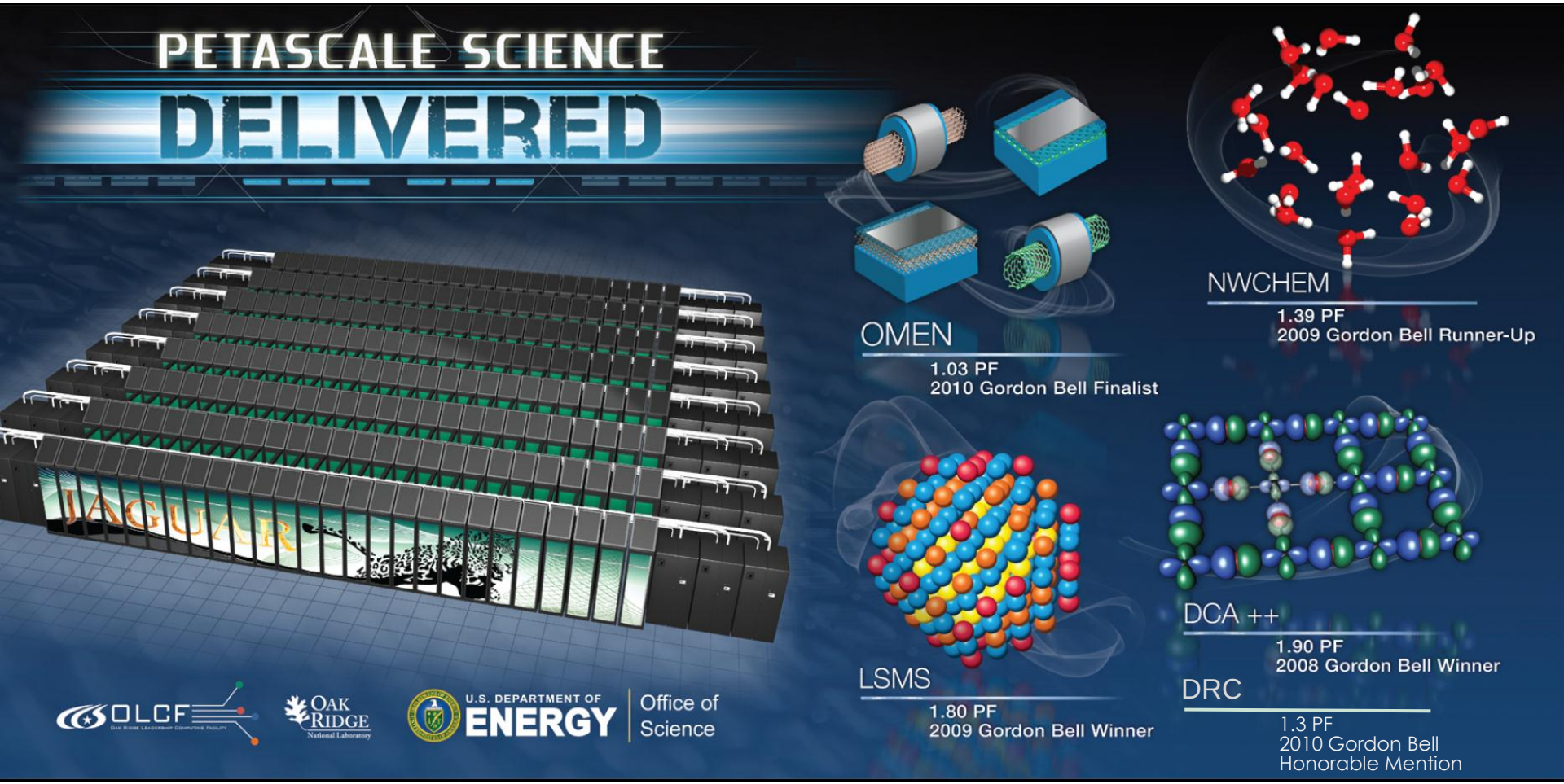
- Hardware
  - Lens cluster
    - 32-node, 512-core, 2TB aggregate memory
    - 1 NVIDIA GTX8800 (768MB) & 1 Tesla C1060 (4GB) per node
  - EVEREST powerwall
    - 30’ x 8’ display with 11,520 x 3,072 resolution
- Software
    - Including VisIt, EnSight Gold (DR), ParaView, POV-Ray, AVS/Express, R MPI, IDL, VirtualGL, NX



Every INCITE project is assigned a single “visualization liaison” as a single point of contact for all post-analysis data processing

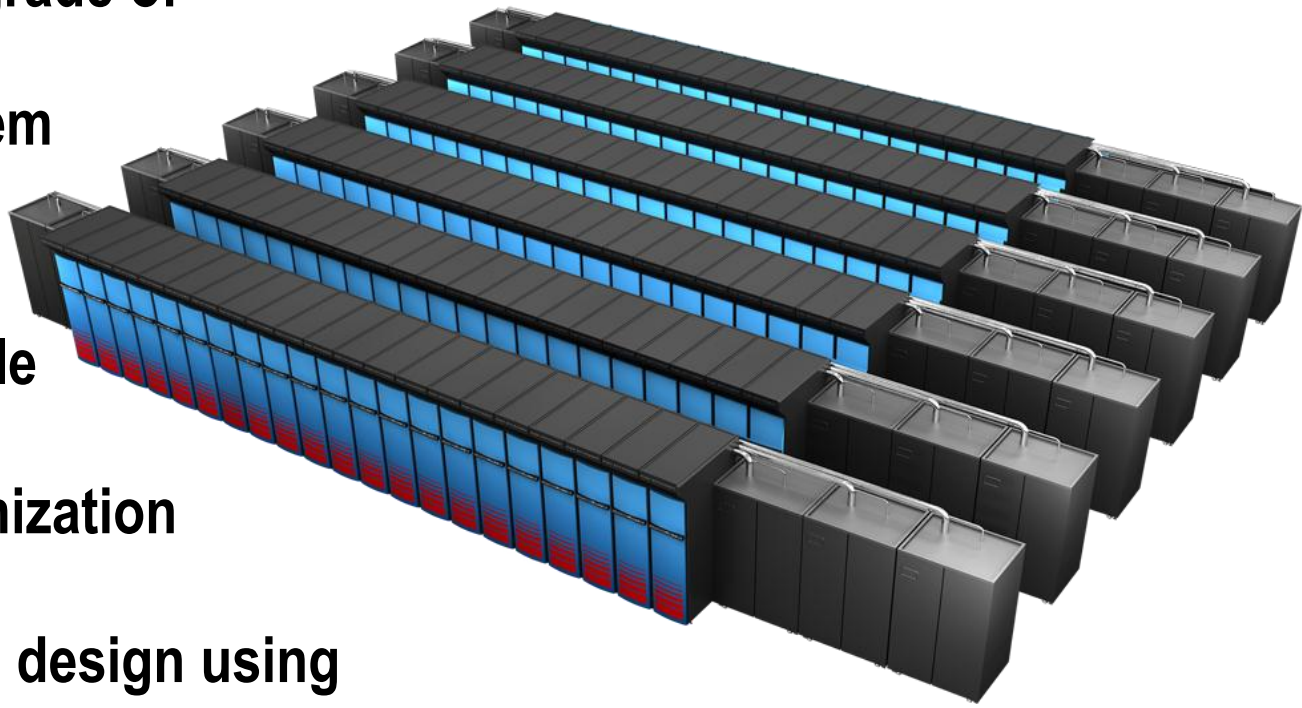
Support visualization tools	**Convert data**
Provide parallel data analysis support	**Produce publication-ready images**
Perform statistical analyses	**Produce movies and animations**
Research new data exploration techniques	**Highlight science successes to visitors**
Develop custom visualization tools and algorithms	**Large display support**

We are delivering Petascale Science Today! Five applications running over 1 Petaflops



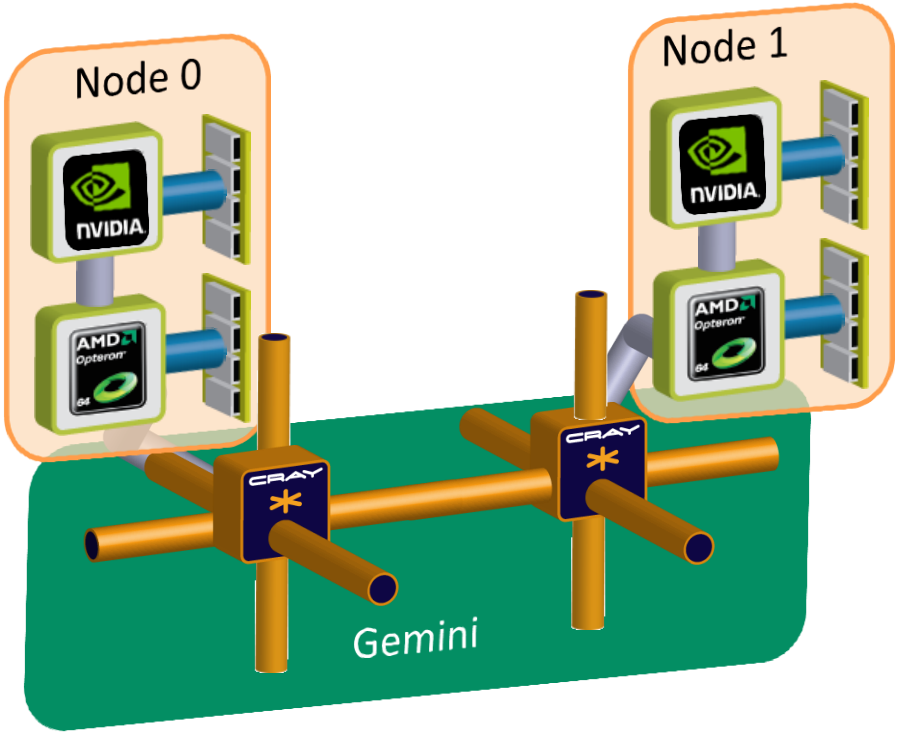
ORNL's "Titan" 20 PF System Goals

- Similar number of cabinets, cabinet design, and cooling as Jaguar
- Operating system upgrade of today's Linux Operating System
- Gemini interconnect
  - 3-D Torus
  - Globally addressable memory
  - Advanced synchronization features
- New accelerated node design using GPUs
- 20 PF peak performance
  - 9x performance of today's XT5
- Larger memory
- 3x larger and 4x faster file system



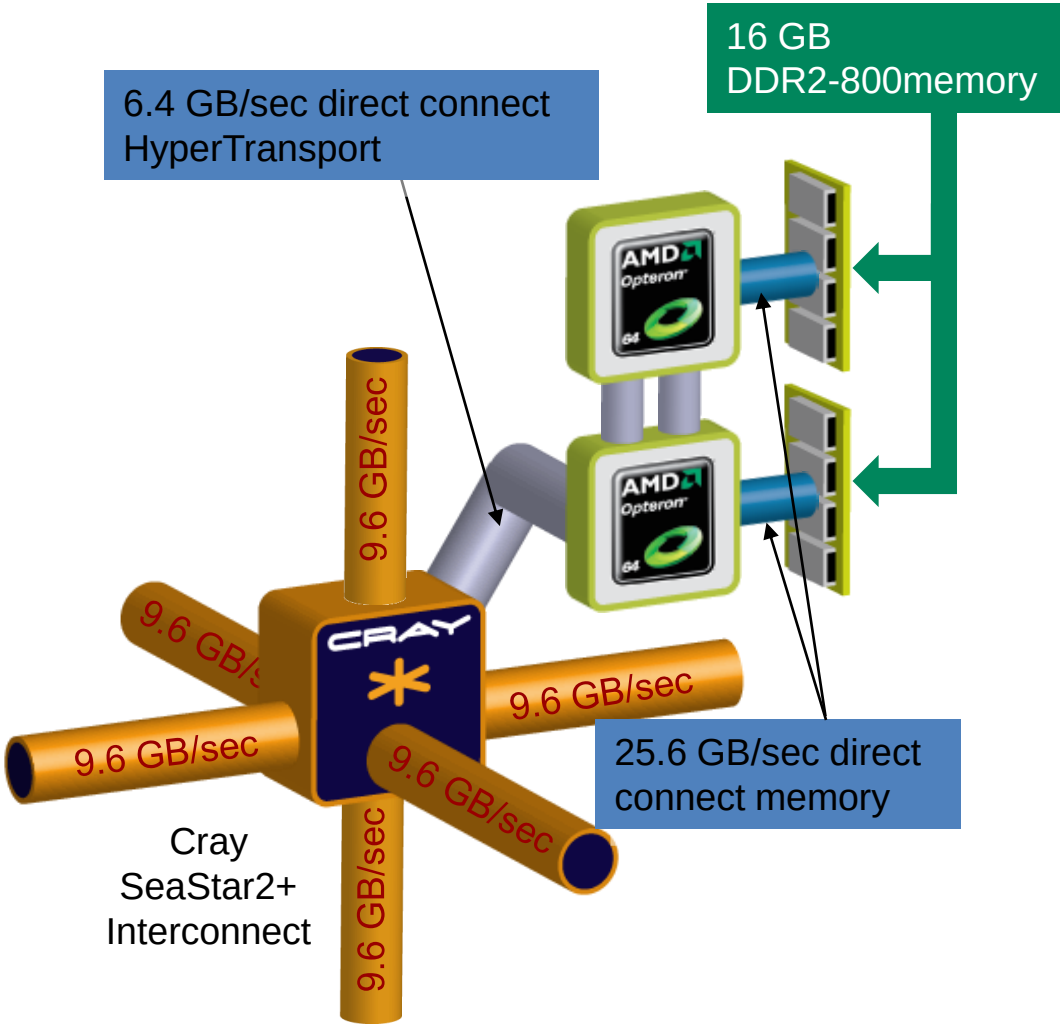
OLCF-3 node description

- New node for “Cray XE” infrastructure
  - Gemini interconnect
  - AMD Socket G34 processor
- 1 AMD socket G34 processor and 1 NVIDIA GPU per node
- Interlagos uses AMD socket G34 and new “Bulldozer” core
  - DDR3-1600 memory
  - HyperTransport version 3
- NVIDIA “Kepler” accelerator
  - Successor to Fermi



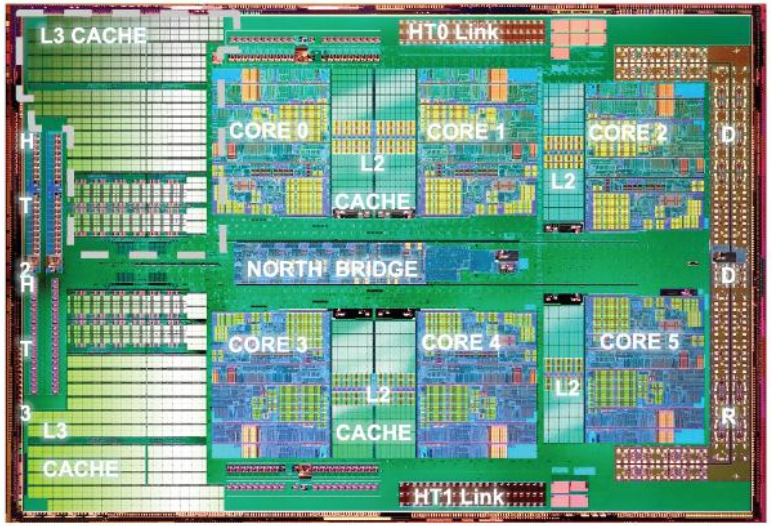
	Jaguar’s XT5 node	OLCF-3 node
Opteron sockets	2	1
Opteron memory (GB)	16	32
Interconnect	Seastar2	Gemini
Node peak GFLOPS	110	>1500

XT5 Node



Cray XT5 Node Characteristics	
Number of Cores	12
Peak Performance	125 Gflops/sec
Memory Size	16 GB per node
Memory Bandwidth	25.6 GB/sec

AMD Opteron 2435 (Istanbul) processors



I'm sold, how do I get time?

- INCITE – 10M+ CPU-hours
- ALCC – 1M+ CPU-hours
- OLCF Director's Discretion – <1M CPU-hours
- CSGF – <1M CPU-hours

What is INCITE?

U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

INCITE: Innovative and Novel Computational Impact on Theory and Experiment

Provides awards to academic, government, and industry organizations worldwide needing large allocations of computer time, supporting resources, and data storage to pursue transformational advances in science and industrial competitiveness.

INCITE is jointly run by the ALCF and OLCF, managed by Julia White

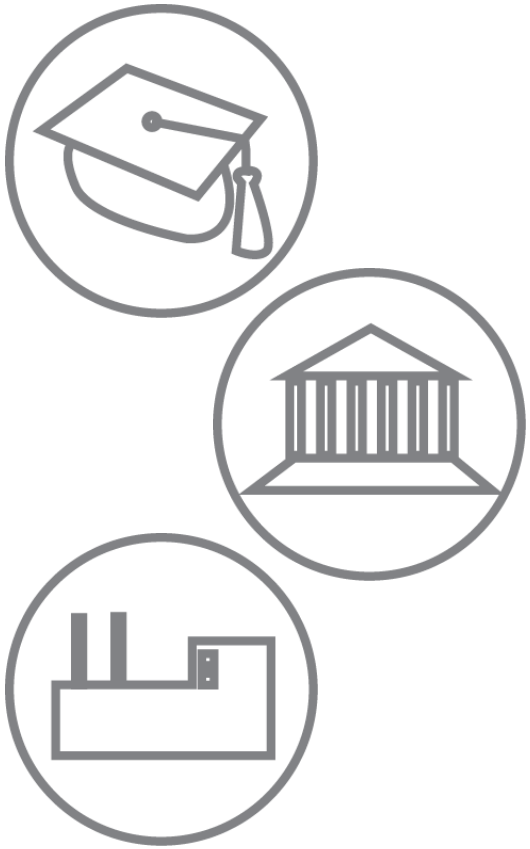
OAK RIDGE National Laboratory

Argonne NATIONAL LABORATORY

OAK RIDGE National Laboratory

INCITE
Innovative and Novel Computational Impact on Theory and Experiment

- Solicits large-scale, computationally intensive research projects
- Open to all scientific & engineering researchers and organizations worldwide
- Provides large computer time and data storage allocations on
  - ALCF IBM BlueGene/P “Intrepid”
  - OLCF Cray XT5 “Jaguar”



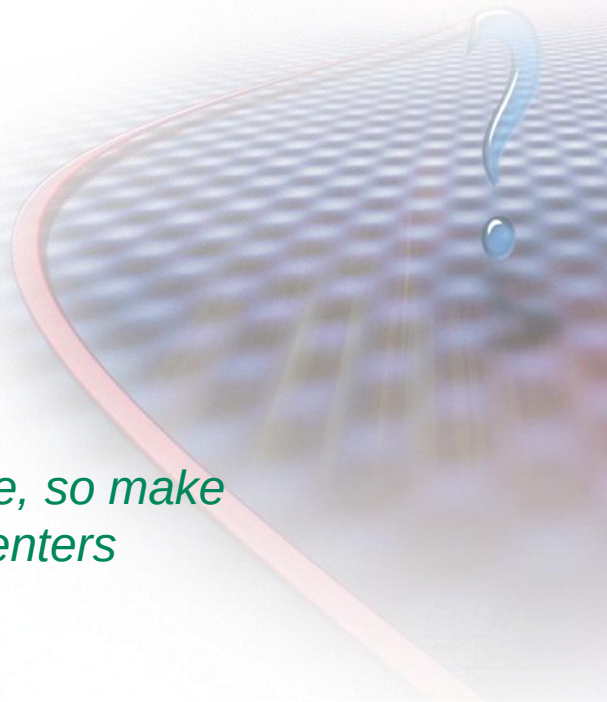
Projects suitable for INCITE

Does your project satisfy most of the following criteria?	
✓	High-impact science and engineering with specific objectives
	AND
✓	Computationally intensive runs that cannot be done anywhere else
✓	Jobs can use at least 20% of the system
✓	Campaign requires tens of millions of CPU hours
✓	Computations are efficient on INCITE's LCF systems

http://www.doeleadershipcomputing.org/

Key questions to ask

- Is another resource more appropriate?
- Is both the scale of the runs **and** the time demands of the problem of LCF scale?
- Do you need specific LCF hardware?
- Do you have the people ready to do this work?
- Do you have a post-processing strategy?
- Do you have a workflow?



Some of the above characteristics are negotiable, so make sure to discuss atypical requirements with the centers

ALCC

- ASCR Leadership Computing Challenge
- Awards ~30% of time at OLCF, ALCF, NERSC
- Emphasis on high-risk, high-payoff simulations in areas directly related to DOE's energy mission

http://science.energy.gov/ascr/facilities/alcc/

OLCF Director's Discretion

- Opportunity to enhance scalability and productivity of scientific codes
- Preparation for INCITE or ALCC

http://www.nccs.gov/user-support/access/

CSGF Allocation Program

- Subset of DD time specifically for CSGF fellows
- Only CSGF fellows can use the allocation for their own research
- Review process differs from usual DD

Proposal form: Outline

1	Personal Information
2	Citizenship Status
3	Mailing Address
4	Employment Information
5	Project Information
6	Project Requirements
7	Description of Research - Research Objectives - Project Milestones - Significance of Research - Computational Approach - Job Characterization - Application Parallel Performance - I/O Requirements
8	Proprietary and Sensitive Information
9	Export Control
10	Comments and Questions

http://www.nccs.gov/user-support/access/project-request/

Starting the proposal

- Personal Information
  - Identify as CSGF in HPC Experience biography
- Citizenship Status
- Mailing Address
  - So we can send you a Fob!
- Employment Information
- Project Information
  - The PI must be you and only you!
  - Funding source: select ASCR in addition to any other funding sources

Project Requirements

- **Up to 1 million CPU hours available - Mind the units!**
  - Processor (core) hours for system
  - Disk storage in gigabytes for both Home and Scratch space (see Sec 7)
  - Mass (tape) storage in gigabytes or terabytes (specify) (see Sec 7)

- Units are core-hours; you are charged for all cores on a node
- Jaguar: Hours are core-hours, but you must request cores in increments of nodes (i.e., 12 cores)

- Project role: choose "Other Project Role" and put in CSGF
- Project Duration: 12 months
- OLCF Support: Fundamental user assistance (probably)

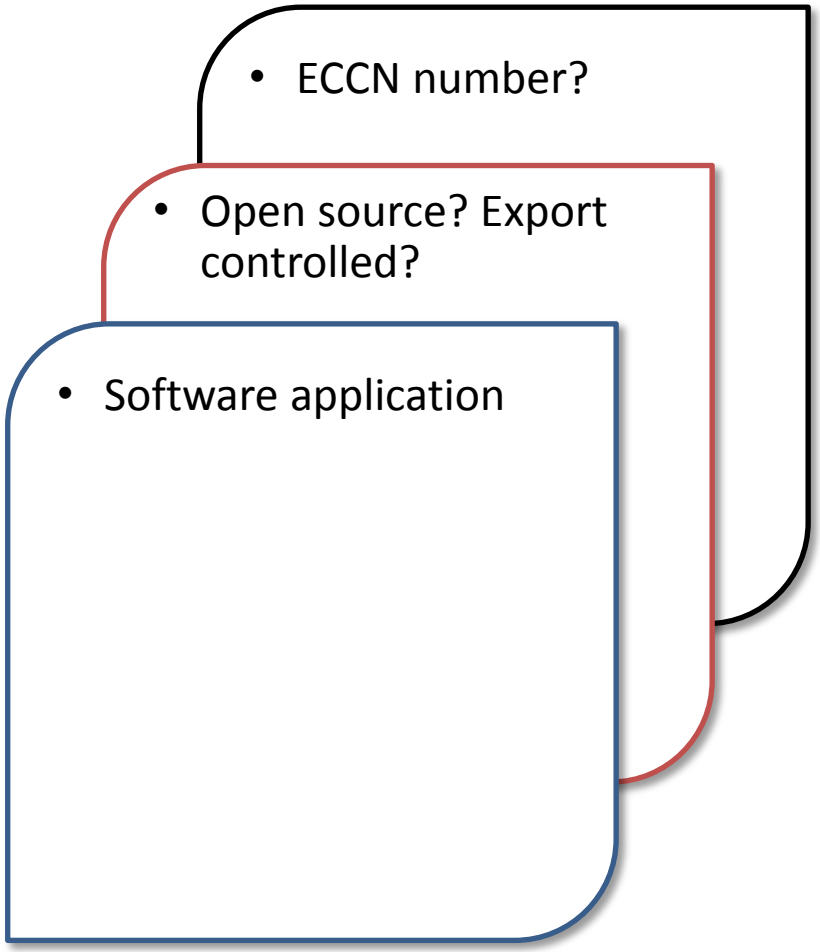
Description of Research: Research Objectives, Milestones, Significance

- **Audience**
  - Computational-science-savvy senior scientists/engineers, and faculty
  - Not everyone will be well versed in your approach
- **Story elements**
  - What the problem is, and its significance
  - Key objectives, key simulations/computations, milestones
  - Approach to solving the problem, its challenging aspects, preliminary results
  - Impact of a successful computational campaign — the big picture
  - Reasons why it is important to carry out this work now

Description of Research: Software Information

List all software application packages/suites to be used (Note: Long lists may reduce credibility)

- What will be used to set up computations?
- What are the codes for the main simulation/modeling?
- What will be used to analyze results?



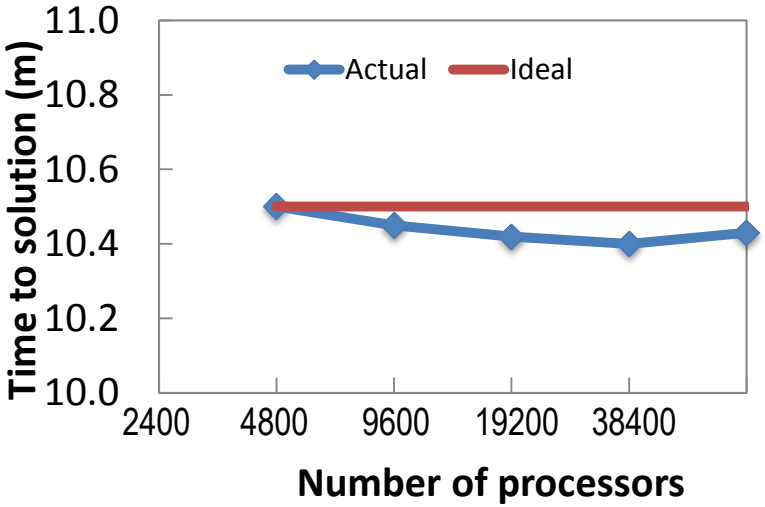
Description of Research: Computational approach

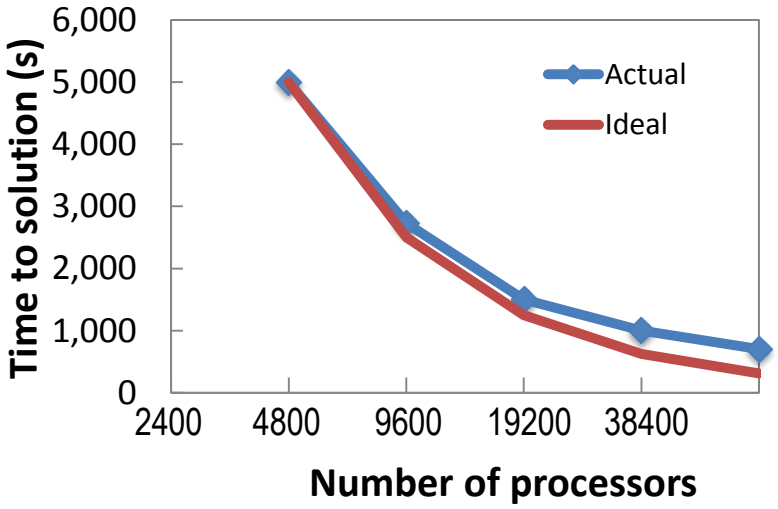
- **Programming languages, libraries and tools used**
  - Describe what you will need and what you will port
- **Description of underlying formulation**
  - Don't assume reviewers know all the math
  - Make it clear that the code you plan to employ is the correct tool for your research plan
  - If you plan to use a private version of a well-known code, delineate the differences

Description of Research: Job Characterization/Parallel Performance

- What jobs are you going to run?
- Why do you need this unique resource?
- Do you have evidence of scalability?

WEAK SCALING DATA	STRONG SCALING DATA
Increase problem size as resources are increased	Increase resources (nodes) while doing the same computation





Description of Research: I/O

- Restart I/O - Application initiated program restart data
  - I/O technique used, e.g., MPI I/O, HDF5, raw
  - Number of processors doing I/O, number of files
  - Sizes of files and overall dump
  - Periodicity of the checkpoint process
- Analysis I/O - Application written files for later analysis
  - I/O technique used, e.g., pNetCDF, pHDF5
  - Number of processors doing I/O, number of files
  - Sizes of files and overall dump
- Archival I/O - Data archived for later use/reference
  - Number and sizes of files
  - Retention length
  - If archived remotely, the transport tool used, e.g., GridFTP

Finishing the Application

- Proprietary and Sensitive Information
  - **Ideal** answers:
    - Does this project generate publishable data? **Yes**
    - Does this project use proprietary data? **No**
    - Does this project generate proprietary data? **No**
    - Does this project use or generate sensitive or restricted information? **No**
- Export Control
    - Does the proposed project involve any of the following areas? **No**
    - Does the proposed project use and/or create proprietary information, intellectual property, licensing, or will utilize controlled technology in any of these categories in pursuit of the project objectives? **No**
    - Are the proposed project and its intended subject matter deliverables considered Fundamental Research or Publicly Available Information as under National Security Decision Directive (NSDD) 189? **Yes**
- Comments and Questions
    - Do you agree to provide periodic updates of research accomplishments? **Yes**

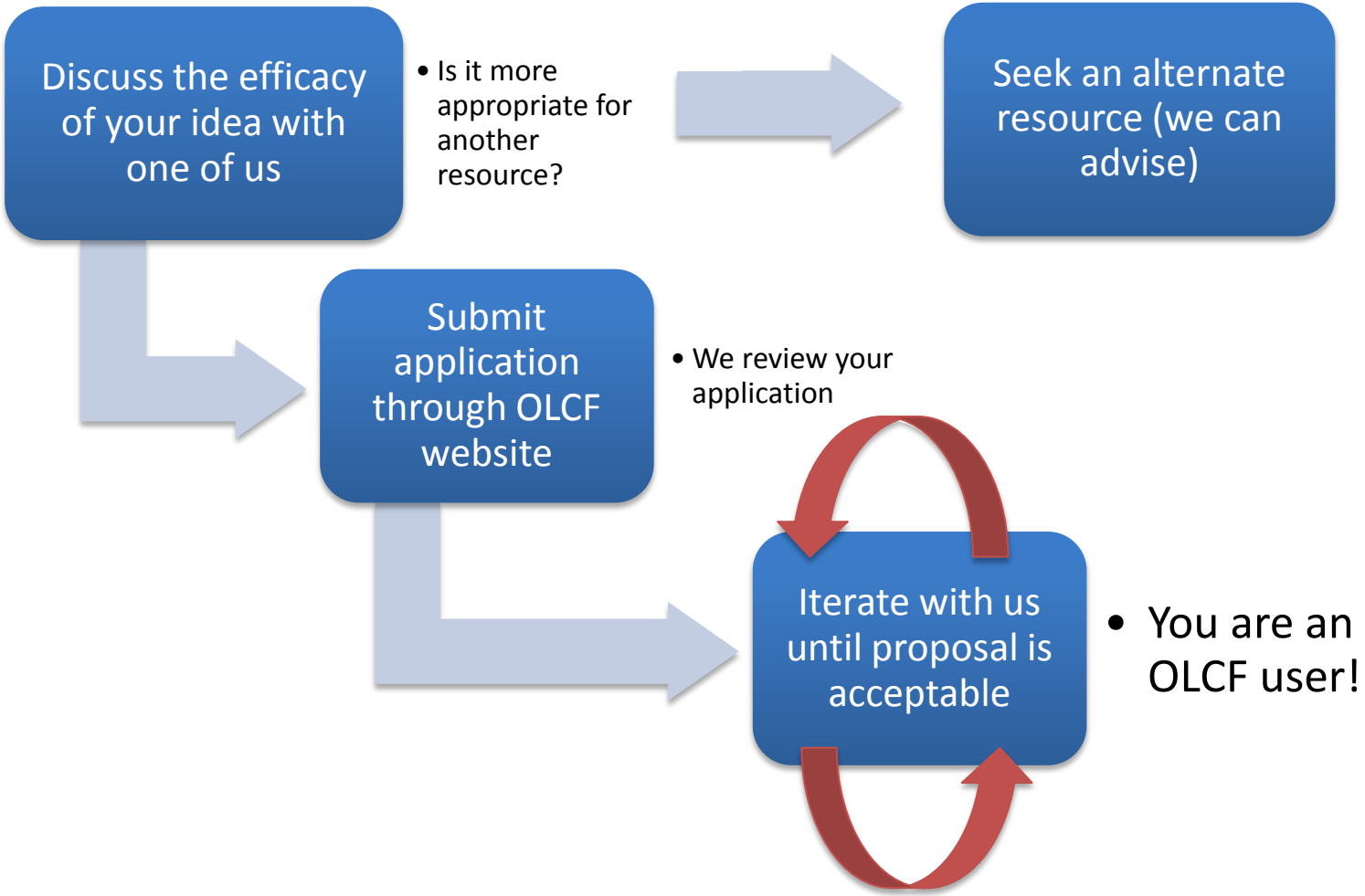
Final checks

- Write your research description in another editor and cut and paste into the form (you **can't** save the form and return later)
- Required fields must be completed for the form to be successfully submitted
- After submitting your proposal, you will not be able to edit it

http://www.nccs.gov/user-support/access/project-request/

Submit

Iterative Review Process



http://www.nccs.gov/user-support/access/project-request/

Freedom Isn't Free

- Reporting Requirements:
  - A final report highlighting the accomplishments with your allocated time is **required**
- Provide highlights on significant science/engineering accomplishments as they occur
- Submit annual renewal request
- Complete annual surveys
- Be good citizen on the computers
- Use the resources for the proposed work

It is a small world…

- OLCF resources will continue to grow as researchers around the world require larger systems for high-impact results
- Let the science agency that funds your work know how significant the OLCF will be to your work
- Contact us if you have questions: we want to hear from you

Relevant links

INCITE Program
http://www.doeleadershipcomputing.org/

ALCC
http://science.energy.gov/ascr/facilities/alcc/

Oak Ridge Discretionary Program
http://www.nccs.gov/user-support/access/project-request/

Contacts

For details about the INCITE program:

www.doeleadershipcomputing.org
INCITE@DOEleadershipcomputing.org

For details about the OLCF:

www.olcf.ornl.gov
help@nccs.gov, 865-241-6536

Contact Us:

Rebecca Hartman-Baker, hartmanbakrj@ornl.gov, 865-241-8989
Judith Hill, hilljc@ornl.gov, 865-241-1731