



Argonne Training Program on Extreme-Scale Computing

Introduction to ATPESC

Ray Loy
ATPESC Program Director



Outline

 **Overview of the ATPESC program**

 **Application Guidance**

 **Questions/Discussion**

Motivation for ATPESC

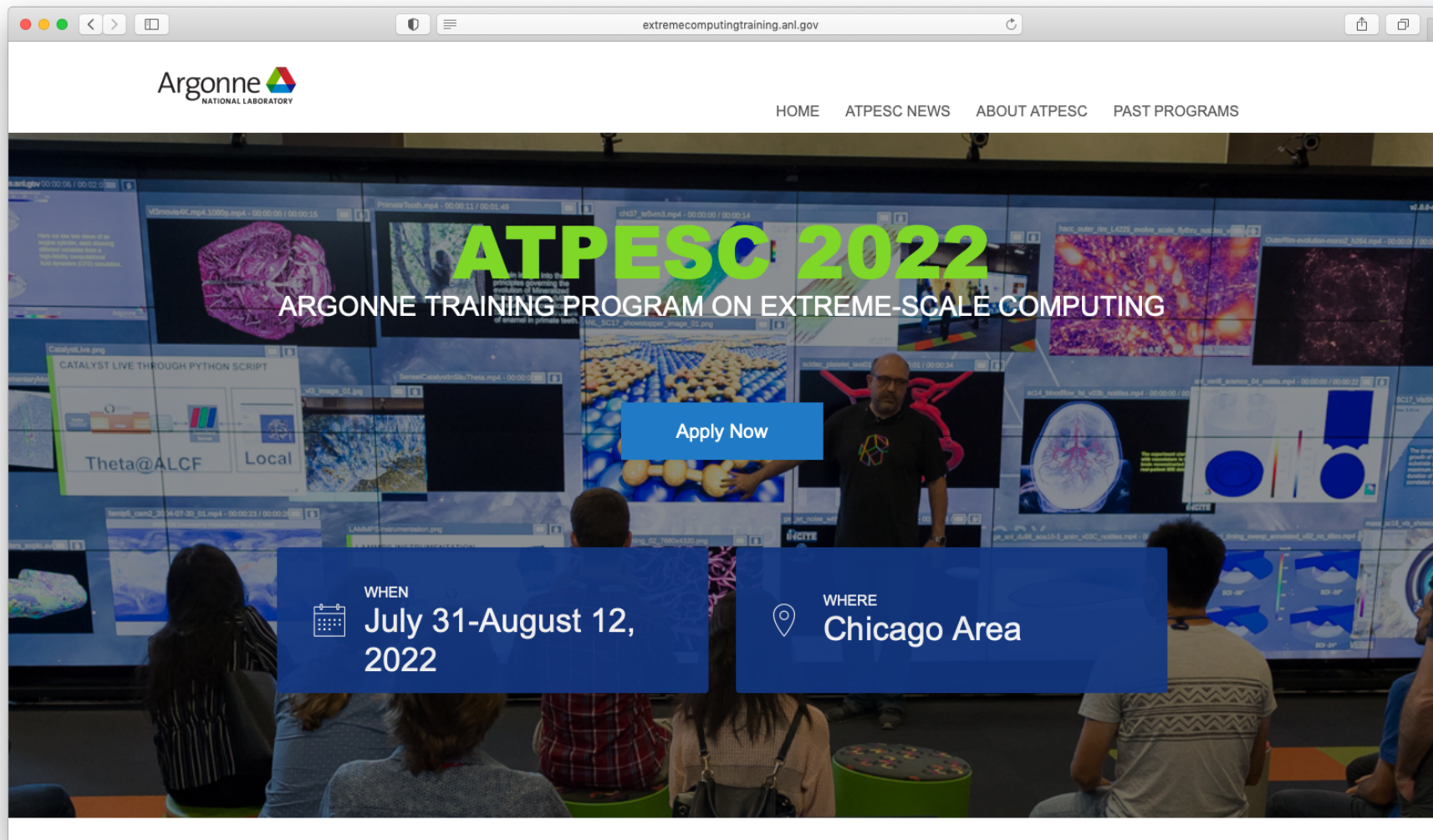
- Today's most **powerful supercomputers** have **complex hardware architectures** and **software environments**
 - and even greater complexity is on the horizon on next-generation and exascale systems
- The **scientific and engineering applications** developed for these systems are themselves **complex**
- There is a **critical need for specialized, in-depth training for the computational scientists** poised to facilitate breakthrough science and engineering using these systems

ATPESC Overview

- Founded by Paul Messina in 2013
- Conceived as a 2-week retreat
- Renowned computer scientists and HPC experts from US national laboratories, universities, and industry serve as lecturers and guide hands-on sessions.
- Target audience: advanced doctoral students, postdocs, and early career computational scientists
- No fee to participate. Domestic travel, meals, and lodging provided.
- Competitive application process reviewed by committee
 - Must have experience in MPI and/or OpenMP and/or Data Science frameworks
 - Experience with at least one HPC system
 - Concrete plans to conduct CSE research on large-scale computers

ATPESC Website

extremecomputingtraining.anl.gov



Argonne – a part of DOE National Laboratory System

Office of Science Laboratories

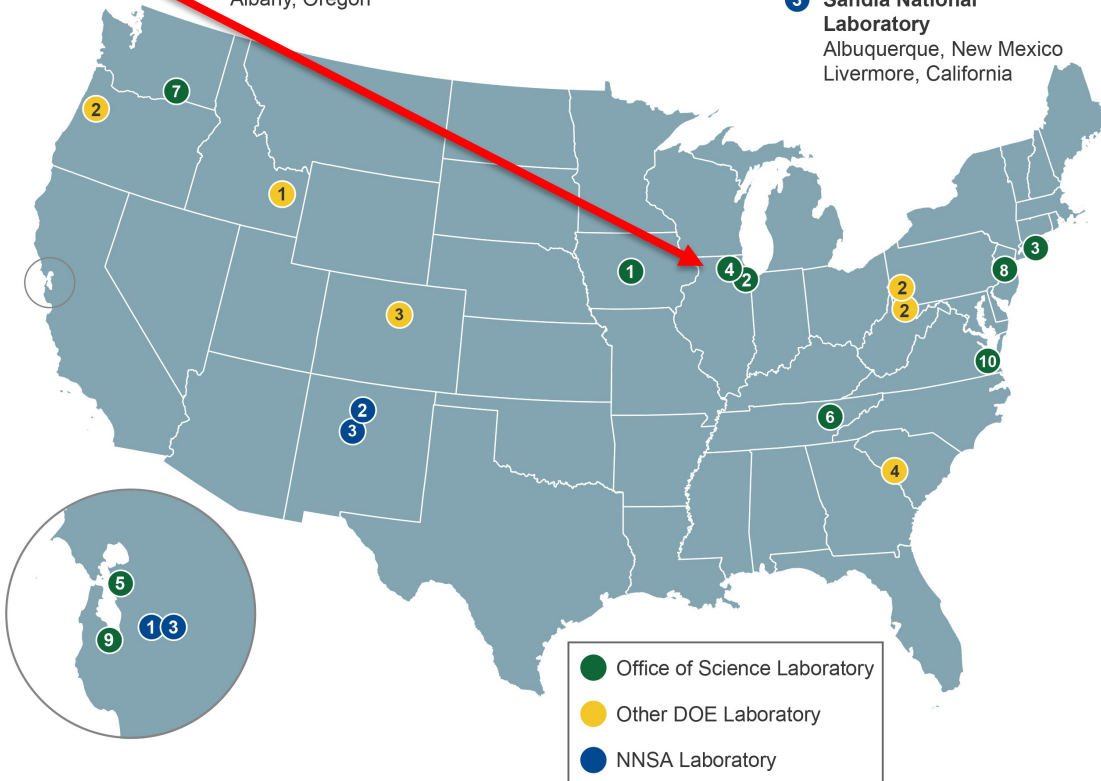
- 1 Ames Laboratory
Ames, Iowa
- 2 Argonne National Laboratory
Argonne, Illinois
- 3 Brookhaven National Laboratory
Upton, New York
- 4 Fermi National Accelerator Laboratory
Batavia, Illinois
- 5 Lawrence Berkeley National Laboratory
Berkeley, California
- 6 Oak Ridge National Laboratory
Oak Ridge, Tennessee
- 7 Pacific Northwest National Laboratory
Richland, Washington
- 8 Princeton Plasma Physics Laboratory
Princeton, New Jersey
- 9 SLAC National Accelerator Laboratory
Menlo Park, California
- 10 Thomas Jefferson National Accelerator Facility
Newport News, Virginia

Other DOE Laboratories

- 1 Idaho National Laboratory
Idaho Falls, Idaho
- 2 National Energy Technology Laboratory
Morgantown, West Virginia
Pittsburgh, Pennsylvania
Albany, Oregon
- 3 National Renewable Energy Laboratory
Golden, Colorado
- 4 Savannah River National Laboratory
Aiken, South Carolina

NNSA Laboratories

- 1 Lawrence Livermore National Laboratory
Livermore, California
- 2 Los Alamos National Laboratory
Los Alamos, New Mexico
- 3 Sandia National Laboratory
Albuquerque, New Mexico
Livermore, California



Together, the **17 DOE laboratories** comprise a preeminent federal research system, providing the Nation with strategic scientific and technological capabilities. The laboratories:

- Execute long-term government scientific and technological missions, often with complex security, safety, project management, or other operational challenges;
- Develop unique, often multidisciplinary, scientific capabilities beyond the scope of academic and industrial institutions, to benefit the Nation's researchers and national strategic priorities; and
- Develop and sustain critical scientific and technical capabilities to which the government requires assured access.

Source: https://science.energy.gov/~media/ /images/laboratories/DOE_Laboratories_Map_2014_Hi-res.jpg

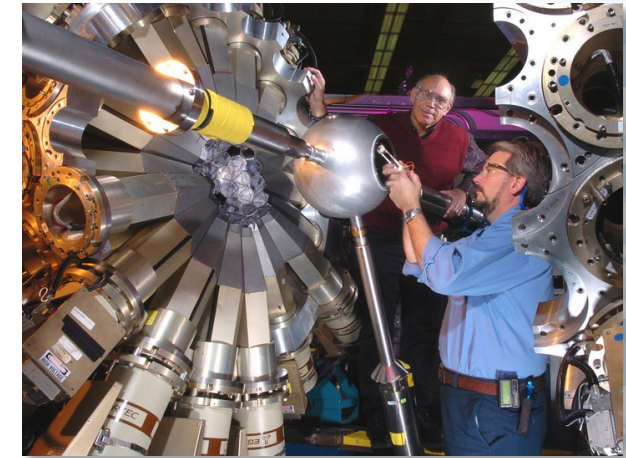


Major Scientific User Facilities at Argonne

Advanced
Photon
Source



Argonne Tandem Linear
Accelerator System



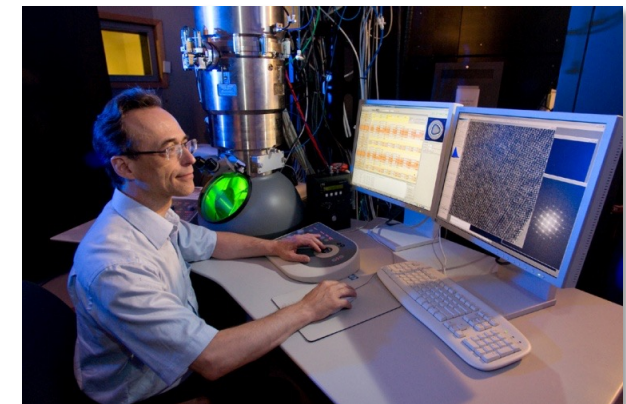
Center for
Nanoscale
Materials



Argonne
Leadership
Computing
Facility



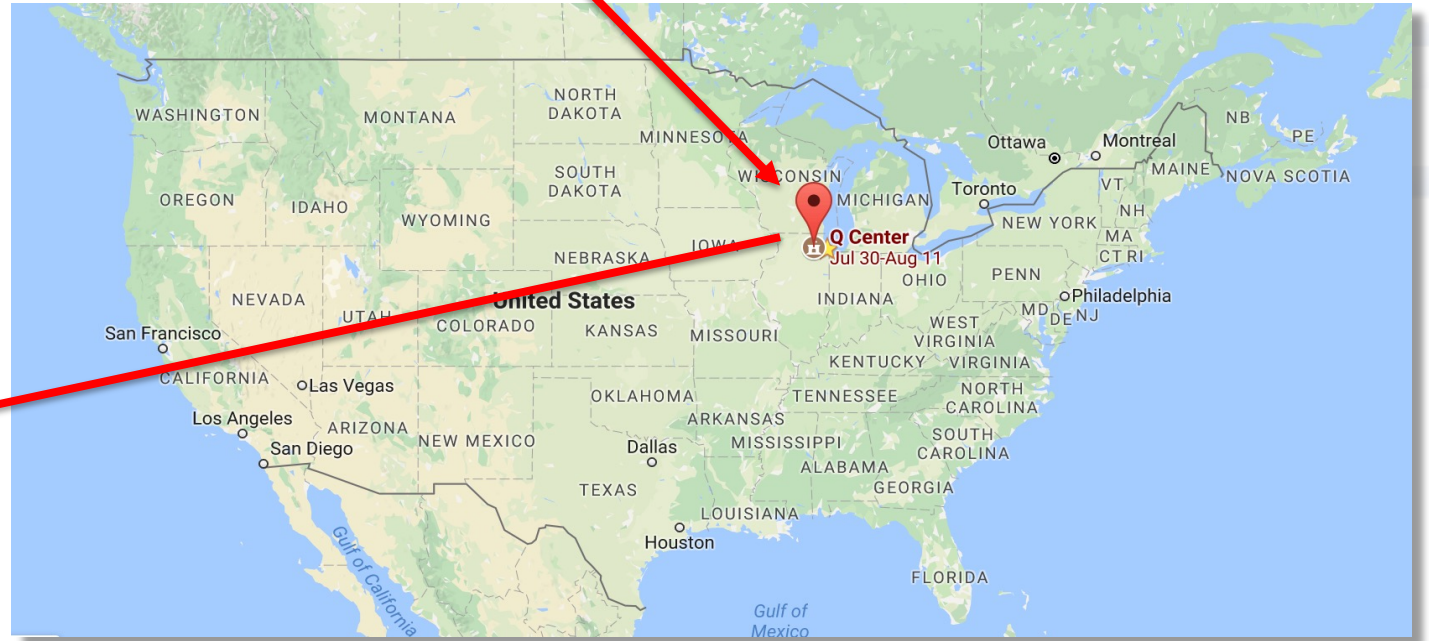
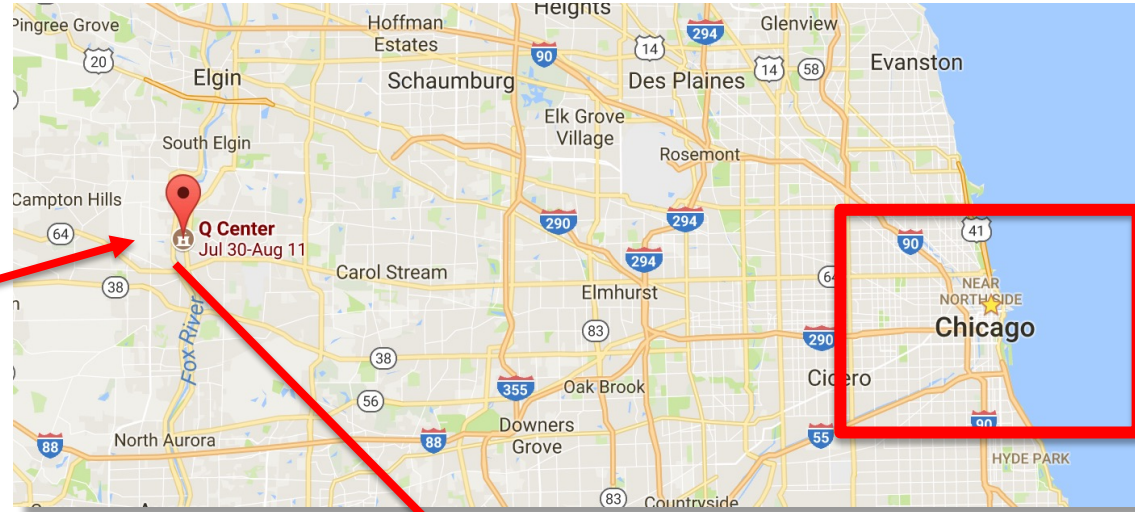
Electron
Microscopy
Center



The future... Aurora Exascale System



ATPESC 2019 Venue



Source: Google Maps, shutterstock, NASA



ATPESC 2020, 2021 – Virtual Events

Recording Total non-video participants: 67

Participants (71)

Argonne Training Program on Extreme-Scale Computing (ATPESC)
Data Analysis and Visualization

ENERGY Office of Science NIS

Participants: Ray Loy (Host), Joseph Insley, Kathy Gorgan (Co-host), Rose Lynch (Co-host), Yasaman Gh..., Albert Mollen, Alberto Roper Pol, Amanda Dufek

Chat: From Kathy Gorgan to Everyone: Good Morning, From Fazlay Rabb to Everyone: Good Morning :), From Soongil Kang to Everyone: Good morning

ALCF-Workshops

#announcements

Go to <https://accounts...Requests>

Appropriate allocation sizes (node hrs): Theta ~16k, ThetaGPU 1k, Cooley 8k

In the "detailed description" box make sure to include
I attended SDL 2020 or My team member (name) attended SDL 2020
I am preparing to submit an (ADSP, INCITE, ALCC) proposal

Ray Loy 2:05 PM
REMEMBER - At 5PM today, the workshop reservations will conclude. Any jobs left in the training queue after 5PM will never run, you should move them to a normal queue for the machine or delete them. Abandoned jobs in training may be deleted by the admins without further warning.

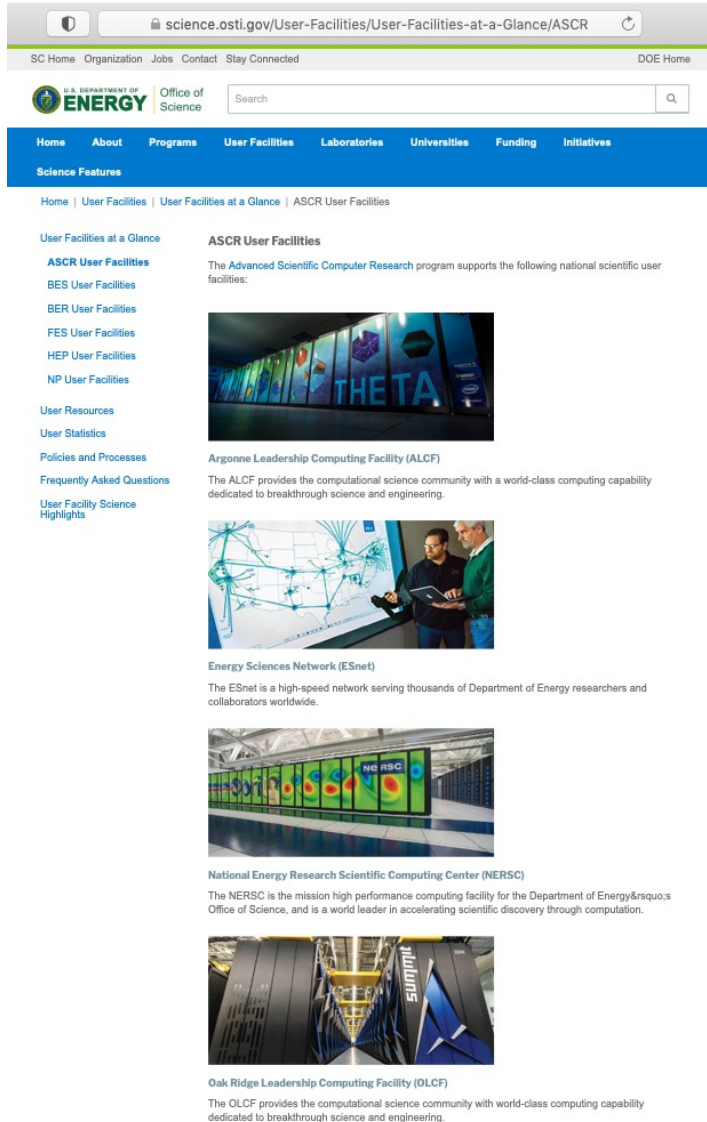
Ray Loy 4:59 PM
Which queues to use now that the reservations are over:
Theta - For small/short jobs, consider debug-cache-quad For complete info see <https://www.alcf.anl.gov/support-center/theta/job-scheduling-policy-theta#kn1-queues>
ThetaGPU - Assuming you are using GPUs, either debug-gpu or single-gpu See <https://www.alcf.anl.gov/support-center/theta/job-scheduling-policy-theta#gpu-queues>
Cooley - For small/short jobs, consider debug See <https://www.alcf.anl.gov/support-center/cooley/job-scheduling-policies-cooley>

Channels: #01_distributed_deep_learning, #02_data_pipelines, #03_distributed_hyper_opt, #04_profiling_deep_learning, #05_simulation_ml, #announcements, #general-discussion, #sdl-management, #sdl-speakers

Curriculum Tracks (*over 100 hours of content*)

- **Track 1:** Hardware Architectures
- **Track 2:** Programming Models and Languages
- **Track 3:** Data-intensive Computing and I/O
- **Track 4:** Visualization and Data Analysis
- **Track 5:** Numerical Algorithms and Software for Extreme-Scale Science
- **Track 6:** Performance Tools and Debuggers
- **Track 7:** Software Engineering
- **Track 8:** Machine Learning and Deep Learning for Science

ATPESC Computing Resources



science.osti.gov/User-Facilities/User-Facilities-at-a-Glance/ASCR

SC Home Organization Jobs Contact Stay Connected DOE Home

U.S. DEPARTMENT OF ENERGY Office of Science

Home About Programs User Facilities Laboratories Universities Funding Initiatives

Science Features

Home | User Facilities | User Facilities at a Glance | ASCR User Facilities

User Facilities at a Glance

- ASCR User Facilities
- BES User Facilities
- BER User Facilities
- FES User Facilities
- HEP User Facilities
- NP User Facilities

User Resources

User Statistics


Policies and Processes

Frequently Asked Questions

User Facility Science Highlights


ASCR User Facilities

The Advanced Scientific Computer Research program supports the following national scientific user facilities:



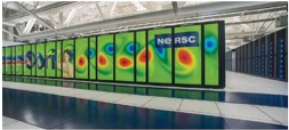
Argonne Leadership Computing Facility (ALCF)

The ALCF provides the computational science community with a world-class computing capability dedicated to breakthrough science and engineering.




Energy Sciences Network (ESnet)

The ESnet is a high-speed network serving thousands of Department of Energy researchers and collaborators worldwide.



National Energy Research Scientific Computing Center (NERSC)

The NERSC is the mission high performance computing facility for the Department of Energy's Office of Science, and is a world leader in accelerating scientific discovery through computation.



Oak Ridge Leadership Computing Facility (OLCF)

The OLCF provides the computational science community with world-class computing capability dedicated to breakthrough science and engineering.

ALCF – Theta/ThetaGPU and Cooley

OLCF - Ascent

NERSC – Cori et al.

Intel Devcloud

Other ATPESC Content

- Dinner Speakers
 - <https://extremecomputingtraining.anl.gov/dinner-talks/>
- Tour of selected Argonne National Lab Facilities
 - <https://extremecomputingtraining.anl.gov/atpesc-argonne-national-laboratory-tour/>


Goals for Attendees



Exposure to New ideas

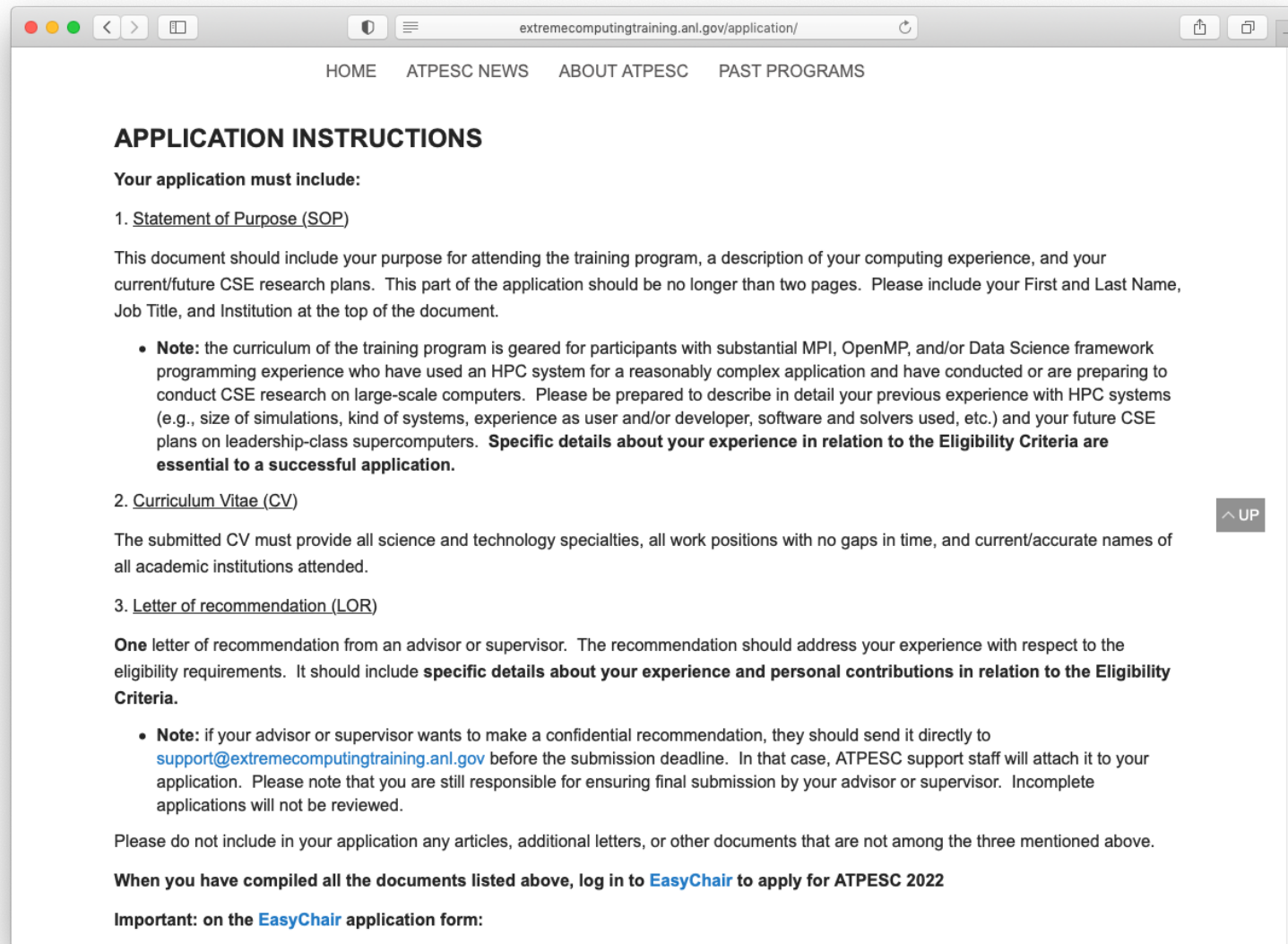


Take advantage of ATPESC Resources



Talk F2F with Lecturers, other Participants, Support

Application Instructions



The screenshot shows a web browser window with the URL extremecomputingtraining.anl.gov/application/. The page has a navigation menu with links for HOME, ATPESC NEWS, ABOUT ATPESC, and PAST PROGRAMS. The main heading is "APPLICATION INSTRUCTIONS".

Your application must include:

- 1. Statement of Purpose (SOP)**
This document should include your purpose for attending the training program, a description of your computing experience, and your current/future CSE research plans. This part of the application should be no longer than two pages. Please include your First and Last Name, Job Title, and Institution at the top of the document.
 - Note:** the curriculum of the training program is geared for participants with substantial MPI, OpenMP, and/or Data Science framework programming experience who have used an HPC system for a reasonably complex application and have conducted or are preparing to conduct CSE research on large-scale computers. Please be prepared to describe in detail your previous experience with HPC systems (e.g., size of simulations, kind of systems, experience as user and/or developer, software and solvers used, etc.) and your future CSE plans on leadership-class supercomputers. **Specific details about your experience in relation to the Eligibility Criteria are essential to a successful application.**
- 2. Curriculum Vitae (CV)**
The submitted CV must provide all science and technology specialties, all work positions with no gaps in time, and current/accurate names of all academic institutions attended.
- 3. Letter of recommendation (LOR)**
One letter of recommendation from an advisor or supervisor. The recommendation should address your experience with respect to the eligibility requirements. It should include **specific details about your experience and personal contributions in relation to the Eligibility Criteria.**
 - Note:** if your advisor or supervisor wants to make a confidential recommendation, they should send it directly to support@extremecomputingtraining.anl.gov before the submission deadline. In that case, ATPESC support staff will attach it to your application. Please note that you are still responsible for ensuring final submission by your advisor or supervisor. Incomplete applications will not be reviewed.

Please do not include in your application any articles, additional letters, or other documents that are not among the three mentioned above.

When you have compiled all the documents listed above, log in to [EasyChair](#) to apply for ATPESC 2022

Important: on the [EasyChair](#) application form:

Eligibility

- Substantial experience in MPI and/or OpenMP programming. Equivalent experience with Data Science frameworks will also be considered.
- Experience using at least one HPC system for a reasonably complex application.
- Plans to conduct computational science and engineering (CSE) research on large-scale computers.

Application Components

- Statement of Purpose
- Curriculum Vitae (CV)
- Letter of Recommendation (LOR)

The Key to success

Provide details

The #1 problem a reviewer faces is trying to guess your qualifications when the materials are not clear.

1. Substantial experience in MPI and/or OpenMP programming. Equivalent experience with Data Science frameworks will also be considered.

- Weak statements
 - "I took a class on MPI"
 - "I ran a code that uses MPI"
- Stronger statements
 - "I work on the ___ code. I implemented, tested, and benchmarked the message-passing routines for the boundary exchange of the ___ solver which has a communication pattern that communicates between ___ and the observed scaling was ___ ..."
 - "I set up the machine learning framework ___ to solve the ___ problem, using a training data set of size ___ and run on ___ nodes. This was then applied to ___ thousand data sets of size ___ ..."
- If you are part of a group-developed code or working on a community code, be specific about your individual contribution.

Experience using at least one HPC system for a reasonably complex application.

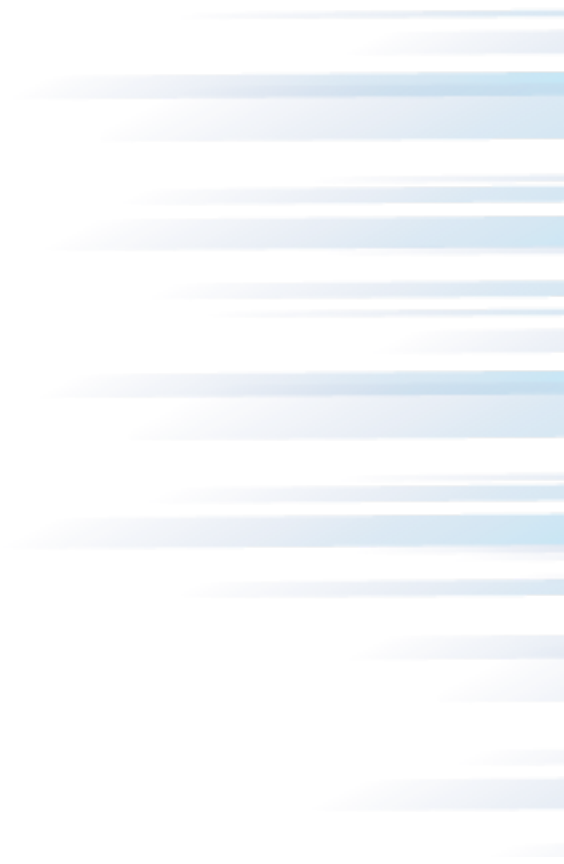
- We do not expect that you have access to a world-class system
 - University/corporate clusters are relevant
- Be clear if you are developing code and running it vs. running an existing code that you did not develop.
- What types of runs have you accomplished?
 - Core/node counts are informative
 - Did you perform scaling studies or other performance inquiries?
- What type of queuing system(s) were involved?
 - Did you construct job scripts?
 - Did you run a large campaign and employ any workflow techniques to manage jobs?

Plans to conduct computational science and engineering (CSE) research on large-scale computers.

- The training should be put to good use
- What is your current situation?
 - Are you currently part of a research effort?
 - What specific ways do you anticipate applying your ATPESC training?
- Where are you headed in your career?
 - If you are a student, what types of positions are you applying to?
 - In what ways do you expect your current work to carry over?

ATPESC 2022 Updates

- Subscribe to mailing list linked from the homepage
 - <https://extremecomputingtraining.anl.gov> (bottom of page)
- Direct inquiries to
 - support@extremecomputingtraining.anl.gov



Acknowledgments

Exascale Computing Project



EXASCALE COMPUTING PROJECT

Website: <https://exascaleproject.org>

This training and research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

Questions/Discussion

