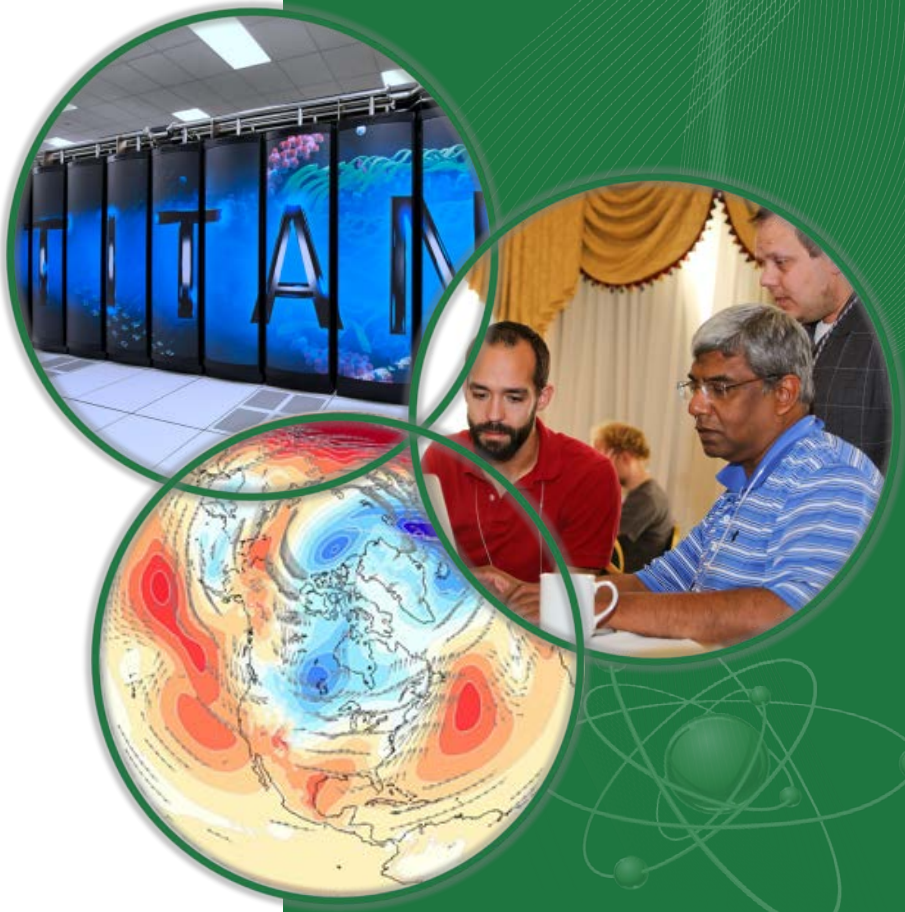
OLCF
Center for Accelerated Application Readiness



Tjerk P. Straatsma
Group Leader, Scientific Computing

Workshop on Performance Portability
Phoenix, AZ, April 19, 2016

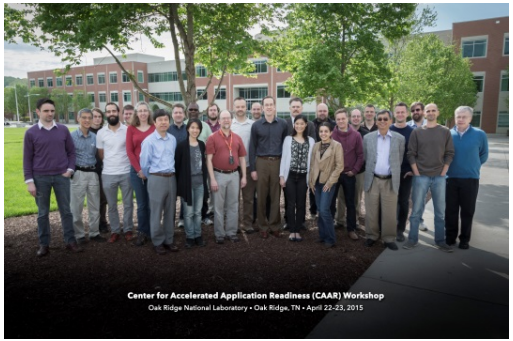
ORNL is managed by UT-Battelle
for the US Department of Energy

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

Center for Accelerated Application Readiness (CAAR)

Goals

- Partnerships between developers, OLCF and CoE
- Applications ready to go when Summit in production
- New applications to attract new PIs into OLCF user programs
- Science at scale to demonstrate HW, SW and Apps
- Experience benefiting users and developers
- Performance portability where possible/appropriate
- Coordination with NERSC/ALCF to avoid duplication of work
- Applications will meet INCITE Computational Readiness
- CSEEN distinguished postdoctoral associate program



Application Developer Team involvement

- Knowledge of the application
- Work on application in development “moving target”
- Optimizations included in application release

Early Science Project

- Demonstration of application on real problems at scale
- Shake-down on the new system hardware and software
- Large-scale science project is strong incentive to participate

Technical support from IBM/NVIDIA Center of Excellence is crucial

- Programming environment often not mature
- Best source of information on new hardware features
- On-site staff from IBM (TBD) and NVIDIA (Jeff Larkin)

Access to multiple resources, including early hardware

Joint training activities

Portability is a critical concern

Summit compared to Titan

Feature	Summit	Titan
Application Performance	5-10x Titan	Baseline
Number of Nodes	~3,400	18,688
Node performance	> 40 TF	1.4 TF
Memory per Node	>512 GB (HBM + DDR4)	38GB (GDDR5+DDR3)
NVRAM per Node	800 GB	0
Node Interconnect	NVLink (5-12x PCIe 3)	PCIe 2
System Interconnect (node injection bandwidth)	Dual Rail EDR-IB (23 GB/s)	Gemini (6.4 GB/s)
Interconnect Topology	Non-blocking Fat Tree	3D Torus
Processors	IBM POWER9 NVIDIA Volta™	AMD Opteron™ NVIDIA Kepler™
File System	120 PB, 1 TB/s, GPFS™	32 PB, 1 TB/s, Lustre®
Peak power consumption	10 MW	9 MW

Current CAAR Projects

ACME: Climate Research: Advancing Earth System Models
PI: Dr. David Bader, Lawrence Livermore National Laboratory
Science Domain: Climate Science
CAAR Liaison: Dr. Matt Norman
CSEEN Postdoc: TBD
NESAP

DIRAC: CAAR Oak Ridge Proposal for getting the Relativistic Quantum Chemistry Program Package DIRAC ready for SUMMIT
PI: Prof. Dr. Lucas Visscher, Free University Amsterdam, the Netherlands
Science Domain: Relativistic Quantum Chemistry
CAAR Liaisons: Dr. Dmitry Liakh, Dr. Tjerk Straatsma
CSEEN Postdoc: Dr. Amelia Fitzsimmons

FLASH: Using FLASH for Astrophysics Simulations at an Unprecedented Scale
PI: Dr. Bronson Messer, Oak Ridge National Laboratory
Science Domain: Astrophysics
CAAR Liaisons: Dr. Bronson Messer
CSEEN Postdoc: Dr. Thom Papatheodore

GTC: Particle Turbulence Simulations for Sustainable Fusion Reactions in ITER
PI: Prof. Dr. Zhihong Lin, University of California - Irvine
Science Domain: Plasma Physics
CAAR Liaison: Dr. Wayne Joubert
NESAP

HACC: Cosmological Simulations for Large-scale Sky Surveys
PI: Dr. Salman Habib, Argonne National Laboratory
Science Domain: Cosmology
CAAR Liaison: Dr. Bronson Messer
NESAP, ESP

LS-DALTON: Large-scale Coupled-cluster Calculations of Supramolecular Wires
PI: Prof. Dr. Poul Jørgensen, Aarhus University
Science Domain: Quantum Chemistry
CAAR Liaison: Dr. Dmitry Liakh
INCITE

NAMD: Molecular Machinery of the Brain
PI: Dr. Klaus Schulten, University of Illinois at Urbana-Champaign
Science Domain: Biophysics
CAAR Liaison: Dr. Antti-Pekka Hynninen
NESAP

NUCCOR: Nuclear Structure and Nuclear Reactions
PI: Dr. Gaute Hagen, Oak Ridge National Laboratory
Science Domain: Nuclear Physics
CAAR Liaison: Dr. Gustav Jansen
CSEEN Postdoc: Dr. Micah Schuster

NWCHEM: Developing Coupled Cluster Methodologies for GPUs
PI: Dr. Karol Kowalski, Pacific Northwest National Laboratory
Science Domain: Computational Chemistry
CAAR Liaison: Dr. Dmitry Liakh
IBM Postdoc: Dr. David Appelhans
NESAP

QMCPACK: Materials Science Research for High-Temperature Superconductors
PI: Dr. Paul R. C. Kent, Oak Ridge National Laboratory
Science Domain: Materials Science
CAAR Liaison: Dr. Ying Wai Li
CSEEN Postdoc: TBD

RAPTOR: Fluid Dynamics Research to Accelerate Combustion Science
PI: Dr. Joseph Oefelein, Sandia National Laboratory, Livermore
Science Domain: Engineering/Combustion
CAAR Liaison: Dr. Ramanan Sankaran
CSEEN Postdoc: Dr. Kalyana Gottiparthi

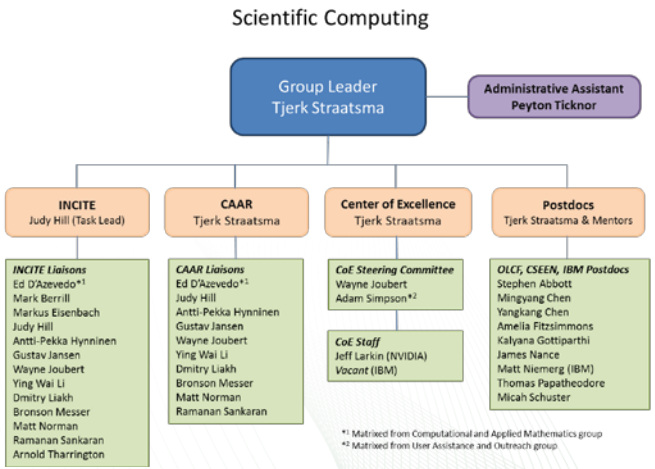
SPECFEM: Mapping the Earth's Interior Using Big Data
PI: Dr. Jeroen Tromp, Princeton University, Princeton University
Science Domain: Seismology
CAAR Liaison: Dr. Judy Hill
CSEEN Postdoc: Dr. Yangkang Chen

XGC: Multiphysics Magnetic Fusion Reactor Simulator, from Hot Core to Cold Wall
PI: Dr. CS Chang, Princeton Plasma Physics Laboratory, Princeton University
Science Domain: Plasma Physics
CAAR Liaison: Dr. Ed D'Azevedo
CSEEN Postdoc: Dr. Stephen Abbott
NESAP

Scientific Computing Group

Mission

The Scientific Computing group is a leader in enabling and conducting a broad array of computational science on state-of-the-art supercomputing architectures through direct collaboration with users and colleagues, active engagement with scientific communities, and advocacy and intercession with facility and vendor partners.



OLCF Center of Excellence (CoE)

CoE Steering Committee

Jim Sexton, IBM, Secretary
John Gunnels, IBM
Leopold Grinberg, IBM
Don Grice, IBM

Cyril Zeller, NVIDIA
Jeff Larkin,NVIDIA
Steven Rennich, NIVIDIA

Tjerk Straatsma, ORNL, CoE Chair
Wayne Joubert, ORNL
Adam Simpson, ORNL
Judy Hill, ORNL
David Bernholdt, ORNL
Oscar Hernandez, ORNL

Bronis de Supinsky, LLNL, CoE Chair
Rob Neely, LLNL
Bert Still, LLNL

CoE Technical Team

John Gunnels, IBM (ACME, DIRAC, FLASH, GTC, HACC, LSDALTON, NAMD, NUCCOR, NWCHEM, QMCPACK, RAPTOR, SPECFEM, XGC)
Leopold Grinberg, IBM (ACME, DIRAC, FLASH, GTC, HACC, LSDALTON, NAMD, NWCHEM, QMCPACK, SPECFEM, XGC)
Bob Walkup, IBM (ACME, DIRAC, FLASH, GTC, HACC, LSDALTON, NAMD, QMCPACK, SPECFEM, XGC)
Constantinos Evangelinos, IBM (LSDALTON)
Sanzong Zhang, IBM (SPECFEM)
Xing Liu, IBM (NUCCOR)
Brian Rogers, IBM (RAPTOR)
David Appelhans, IBM (FLASH, NWCHEM)

Jeff Larkin,NVIDIA (ACME, DIRAC, FLASH, GTC, HACC, LSDALTON, NAMD, NUCCOR, QMCPACK, RAPTOR, SPECFEM, XGC)
Peng Wang, NVIDIA (GTC, XGC)
Nsakharnykh, NVIDIA (NAMD)
Peter Messmer, NVIDIA (SPECFEM)
Jakob Progsch, NVIDIA (SPECFEM)
Levi Barnes, NVIDIA (NWCHEM, RAPTOR)
Justin Luitjens, NVIDIA (HACC)

CoE Technical Team on-site at ORNL

Vacant, IBM
Matt Niemerg, IBM Postdoc (starts end of March)

Jeff Larkin, NIVIDIA

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

Two Tracks for Future Large Systems

Tianhe-2 (NUDT): TH-IVB-FEP
Intel Xeon E5-2692 12 C 2.2 GHz
TH Express-2
Intel Xeon Phi 31S1P

Titan (Cray): Cray XK7
AMD Opteron 6274 16C 2.2 GHz
Cray Gemini
NVIDIA K20x

Sequoia (IBM): BlueGene/Q
Power BQC 16C 1.6 GHz

K computer (Fujitsu)
SPARC64 VIIIfx 2.0 GHz
Tofu

Mira (IBM): BlueGene/Q
PowerPC A2 16C 1.6 GHz

Piz Daint (Cray): Cray XC30
Intel Xeon E5-2670 8C 2.6 GHz
Cray Aries
NVIDIA K20x

Edison (Cray): Cray XC30
Intel Xeon E%-2695v2 12C 2.4 GHz
Aries

Many Core

- 10's of thousands of nodes with millions of cores
- Homogeneous cores
- Multiple levels of memory – on package, DDR, and non-volatile
- Unlike prior generations, future products are likely to be self hosted

Cori at NERSC

- Self-hosted many-core system
- Intel/Cray
- 9300 single-socket nodes
- Intel® Xeon Phi™ Knights Landing (KNL)
- 16GB HBM, 64-128 GB DDR4
- Cray Aries Interconnect
- 28 PB Lustre file system @ 430 GB/s
- Target delivery date: 2016

Aurora at ALCF

- Self-hosted many-core system
- Intel/Cray
- Intel® Xeon Phi™ Knights Hill (KNH)
- Target delivery date: 2018

Hybrid Multi-Core

- CPU / GPU Hybrid systems
- Likely to have multiple CPUs and GPUs per node
- Small number of very fat nodes
- Expect data movement issues to be much easier than previous systems – coherent shared memory within a node
- Multiple levels of memory – on package, DDR, and non-volatile

Summit at OLCF

- Hybrid CPU/GPU system
- IBM/NVIDIA
- 3400 multi-socket nodes
- POWER9/Volta
- More than 512 GB coherent memory per node
- Mellanox EDR Interconnect
- Target delivery date: 2017

ASCR Computer Upgrades at NERSC, OLCF and OLCF

Attributes	NERSC Now	OLCF Now	ALCF Now	NERSC Upgrade	OLCF Upgrade	ALCF Upgrade	
Name	Edison	Titan	Mira	Cori 2016	Summit 2017-2018	Theta 2016	Aurora 2018-2019
Peak (PF)	2.6	27	10	>30	150	>8.5	180
Peak Power (MW)	2	9	4.8	<3.7	10	1.7	13
Memory	357 TB	710 TB	768 TB	~1 PB DDR4+HBM 1.5 PB NVRAM	>1.74 PB DDR4+HBM 2.8 PB NVRAM	> 480 TB DDR4+HBM	>7 PB HBM+NVRAM
Node Peak (TF)	0.460	1.452	0.204	>3	>40	>3	> 3.4
Node Processors	Intel Ivy Bridge	AMD Opteron Nvidia Kepler	64-bit PowerPC A2	1) Intel KNL 2) Intel Haswell	IBM Power9 Nvidia Volta	Intel KNL	Intel KNH
System size (nodes)	5,600	18,688	49,152	9,300 + 1,900	~3,500	>2,500	>50,000
Interconnect	Aries	Gemini	5D Torus	Aries	Dual Rail EDR-IB	Aries	Intel Omni-Path
File System	7.6 PB Lustre 168 GB/s	32 PB Lustre 1 TB/s	26 PB GPFS 300 GB/s	28 PB Lustre 744 GB/s	120 PB GPFS 1 TB/s	10 PB Lustre 210 GB/s	150 TB Lustre 1 TB/s

CAAR: Architecture and Performance Portability

ALCF, NERSC and OLCF Joint Activities and Resources

- ALCF, NERSC and OLCF participated in each other's proposal reviews
- ALCC Award to support NESAP, CAAR and ESP
- Common applications teams in NESAP, CAAR and ESP will collaborate
- Leveraging training activities at NERSC, OLCF and ALCF
- SC15 workshop "Portability Among HPC Architectures for Scientific Applications" on Sunday, November 15 was chaired by Tim Williams (ALCF), Katie Antypas (NERSC) and Tjerk Straatsma (OLCF)
- All three ASCR facilities have representation on the standards bodies for programming models that facilitate portability (OpenACC and OpenMP)
- We are working with vendors to provide programming environments and tools that enable portability, as part of CORAL and Trinity procurements
- Organized portability workshops
- Portability Research Project shared between OLCF, ALCF, NERSC and their CoE's

Oakland, September 24-25, 2014

Oak Ridge, January 27-29, 2015

Synergy between Application Readiness Programs

NESAP at NERSC - *NERSC Exascale Science Application Program*

- Call for Proposals – June 2014
- 20+26 Projects selected
- Partner with Application Readiness Team and Intel/Cray
- 8 Postdoctoral Fellows

CAAR at OLCF - *Center for Accelerated Application Readiness*

- Call for Proposals – November 2014
- 13 Projects selected
- Partner with Scientific Computing group and IBM/NVIDIA Center of Excellence
- 8 Postdoctoral Associates

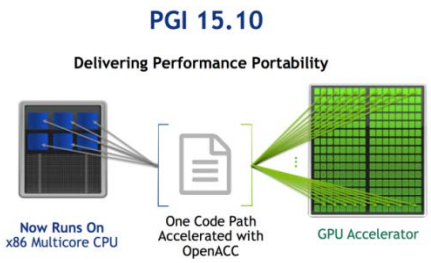
ESP at ALCF - *Early Science Program*

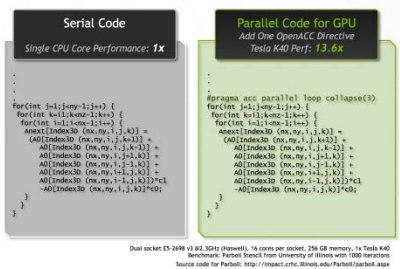
- Call for Proposals – May 2015
- 6 Projects selected in first round
- Partner with Catalyst group and Intel/Cray Center of Excellence
- Postdoctoral Appointee per project

Architecture and Performance Portability

Application portability among NERSC, ALCF and OLCF architectures is critical concern of ASCR

- Application developers target wide range of architectures
- Maintaining multiple code version is difficult
- Porting to different architectures is time-consuming
- Many Principal Investigators have allocations on multiple resources
- Applications far outlive any computer system





Improve data locality and thread parallelism

- GPU or many-core optimizations improve performance on all architectures
- Exposed fine grain parallelism transitions more easily between architectures
- Data locality optimized code design also improves portability

Use portable libraries

- Library developers deal with portability challenges
- Many libraries are DOE supported

MPI + Compiler Directives

- Current approaches include OpenMP and OpenACC
- OpenMP 4.x could emerge as common programming model
- All ASCR centers are on the OpenMP standards committee

Encourage portable and flexible software development

- Use open and portable programming models
- Avoid architecture specific models such as Intel TBB, NVIDIA CUDA
- Use good coding practices: parameterized threading, flexible data structure allocation, task load balancing, etc.

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

Questions
str@ornl.gov