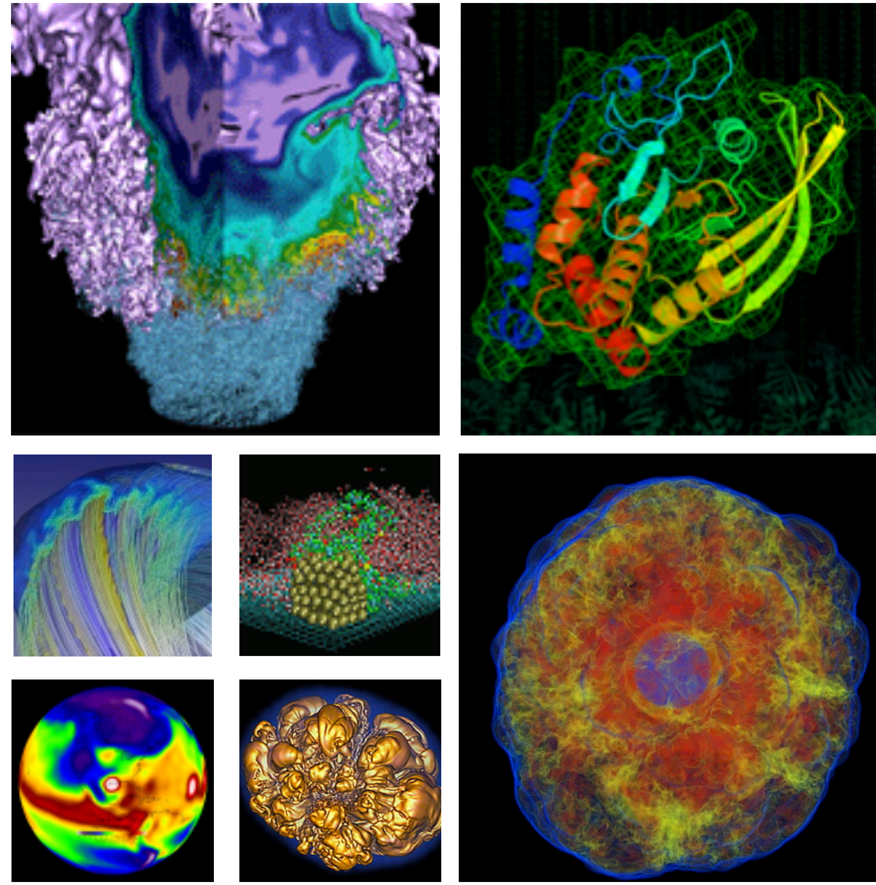
NERSC Initiative: Preparing Applications for Exascale



NERSC

Nicholas J. Wright
Advanced Technologies Group Lead

NERSC User Group Meeting
February 12. 2013

U.S. DEPARTMENT OF ENERGY | Office of Science

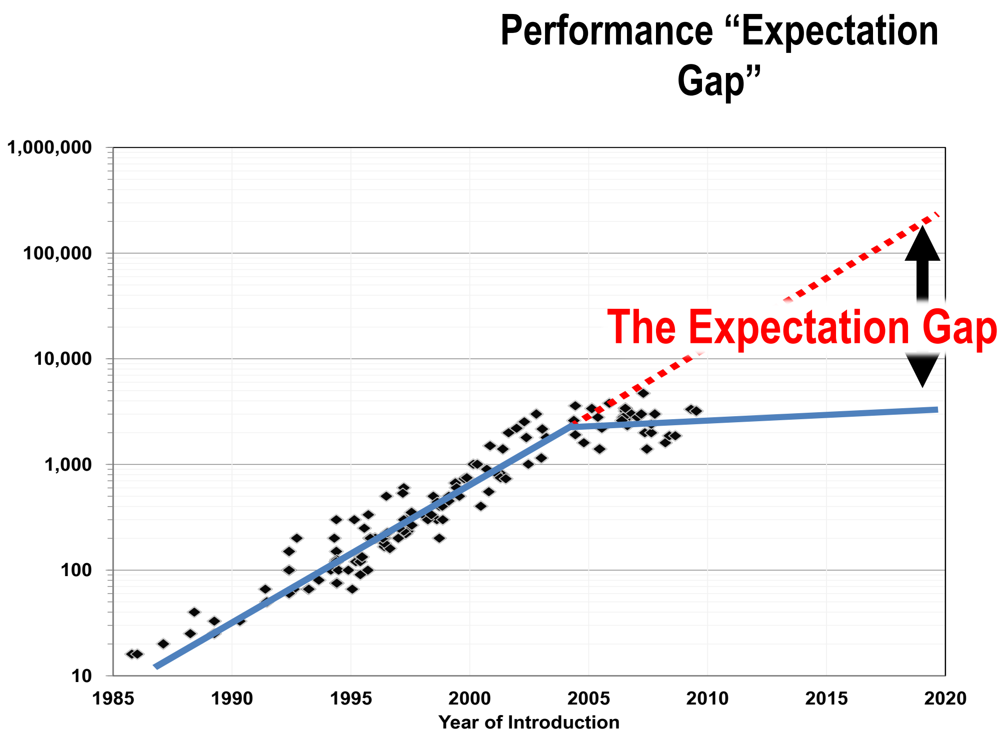
BERKELEY LAB

Computer Industry Roadmaps

- Technology disruption is underway at the processor and memory level. Computing challenges include:
  - Energy efficiency
  - Concurrency
  - Data movement
  - Programmability
  - Resilience

These will impact all scales of computing

- We can only meet these challenges through both hardware and software innovation
  - Rewrite application codes
  - Try to influence computer industry



Challenges to ~~Exascale~~ Performance Growth

1. System power is the primary constraint
2. Concurrency (1000x today)
3. Memory bandwidth and capacity are not keeping pace
4. Processor architecture is open, but maybe heterogeneous
5. Programming model - heroic compilers will not hide this
6. Algorithms need to minimize data movement, not flops
7. I/O bandwidth unlikely to keep pace with machine speed
8. Reliability and resiliency will be critical at this scale
9. Bisection bandwidth limited by cost and energy

Unlike the last 20 years most of these (1-7) are equally important across scales e.g., 1000 1-PF machines

Goal for Programmers at All Levels
(NNSA Exascale Roadmapping Workshop in SF, 2011)

- **Minimize the number of lines of code I have to change when we move to next version of a machine**
  - Evidence that current abstractions are broken are entirely related to effort required to move to each new machine
  - **Target is the FIRST DERIVATIVE of technology changes!!!**

- **What is changing the fastest? What do we want to make future programming models less sensitive to?**
  - Insensitive to # cores (but unclear if as worried about # of nodes)
  - Less sensitive to sources of of non-uniformity (execution rates and heterogeneous core types)
  - Memory capacity/compute ratio (strong'ish'-scaling)
  - Data Movement Constraints
    - Increasingly distance-dependent cost of data movement
    - Topological constraints (node-scale & system-wide)
    - Expressed as NUMA domains (within node)

Slide from John Shalf

Programming Models Strategy

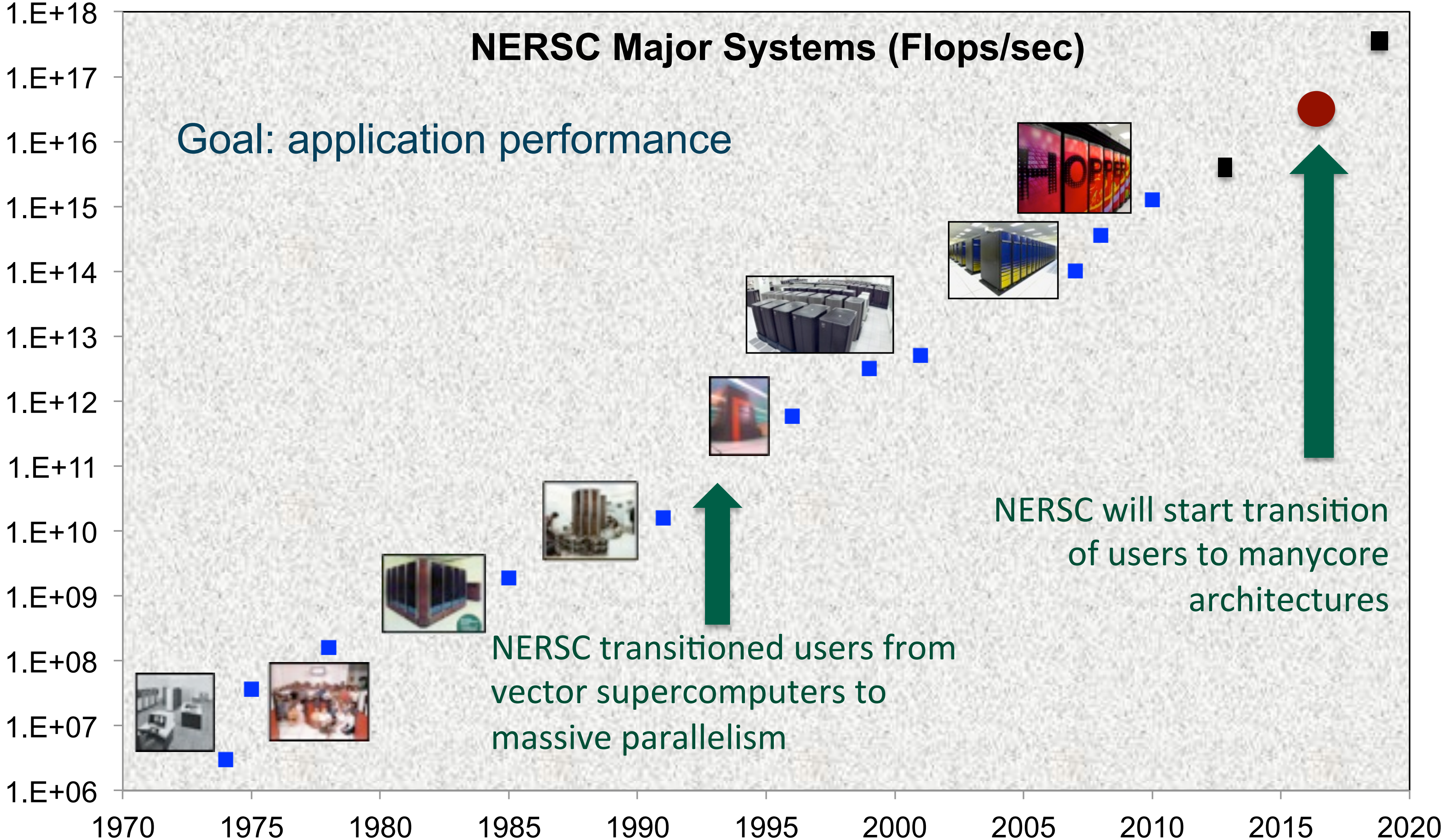
- **The necessary characteristics for broad adoption of a new pmodel is**
  - Performance: At least 10x-50x performance improvement
  - Portability: Code performs well on multiple platforms
  - Durability: Solution must be good for a decade or more
  - Availability/Ubiquity: Cannot be a proprietary solution
- **Our near-term strategy is**
  - Smooth progression to exascale from a user's point of view
  - Support for legacy code, albeit at less than optimal performance
  - Reasonable performance with MPI+OpenMP
  - Support for a variety of programming models
  - Support optimized libraries
- **Longer term, Berkeley Lab is willing to lead a multinational effort to converge on the next programming model**
  - Leverage research efforts (XStack, XTune, DEGAS) for advanced programming constructs
  - Assess existing models through participation in standards bodies (OMP, MPI, Fortran, C/C++) and assess emerging languages
  - Engage co-design community of vendors & HPC for cross-cutting solutions
  - Share results and build consensus

Strategy for transitioning the SC Workload to Energy Efficient Architectures

- **We will deploy testbeds to gain experience with new technologies and to better understand emerging programming models and potential tradeoffs.**
  - In particular, we will deploy a testbed representative of the NERSC-8 architecture as soon as it is determined
- **We will have in-depth collaborations with selected users and application teams to begin transitioning their codes to our testbeds and to NERSC-8**
  - We will choose partnerships based on level of interest, expected application usage, and algorithmic diversity
- **We will develop training and online resources to help the rest of our users based on our in-depth collaborations, as well as on results from co-design centers and ASCR research**
  - We will leverage our existing training vehicles, such as online webinars, as much as possible.
- **We will add consultants with an algorithms background who can help users when they have questions about improving the performance of key code kernels**

It is important to note that all users will be impacted by technology changes because the disruption is at the processor and memory level. Not making this change will stall scientific process because users will be stuck at current performance levels

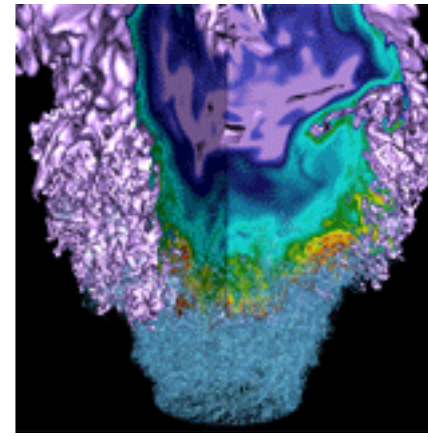
NERSC plan will take scientists through technology transition

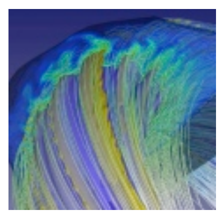


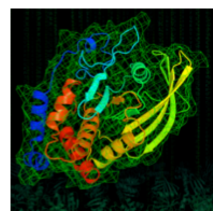
U.S. DEPARTMENT OF ENERGY | Office of Science

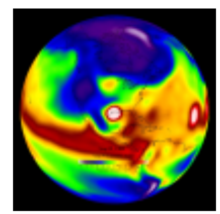
BERKELEY LAB

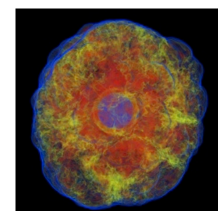
Application Readiness for NERSC-8

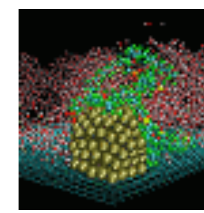












U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

Architecture Evaluation and Application Readiness team

Nick Wright: Lead
Amber (Molecular Dynamics)
(proxy: NAMD, DLPOLY, LAMMPS, Gromacs)

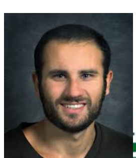
Katie Antypas: FLASH (explicit hydro)
(proxy: Castro, MAESTRO, S3D?, AMR)

Harvey Wasserman: POP
(proxy: CESM)

Jack Deslippe: Quantum Espresso, Berkeley GW (proxy: VASP, PARATEC, Abinit, PETot, Qbox, p2k)

Woo-Sun Yang: CAM (Spectral Element)
(proxy: CESM)

Matt Cordery: MPAS (Scalable Ocean Model)
(proxy: CESM)

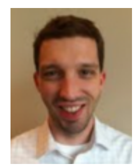


Lenny Oliker: GTC (PIC - Fusion)
(proxy: GTS, XGC, Osiris, g2s)

Brian Austin: Zori (QMC)
(proxy: qwalk)

Kirsten Fagnan: BLAST**,** Allpaths (Bioinformatics)

Burlen Loring: Vis, Iso-surface



Aaron Collier: MADAM-toast (CMB), Gyro (Fusion)

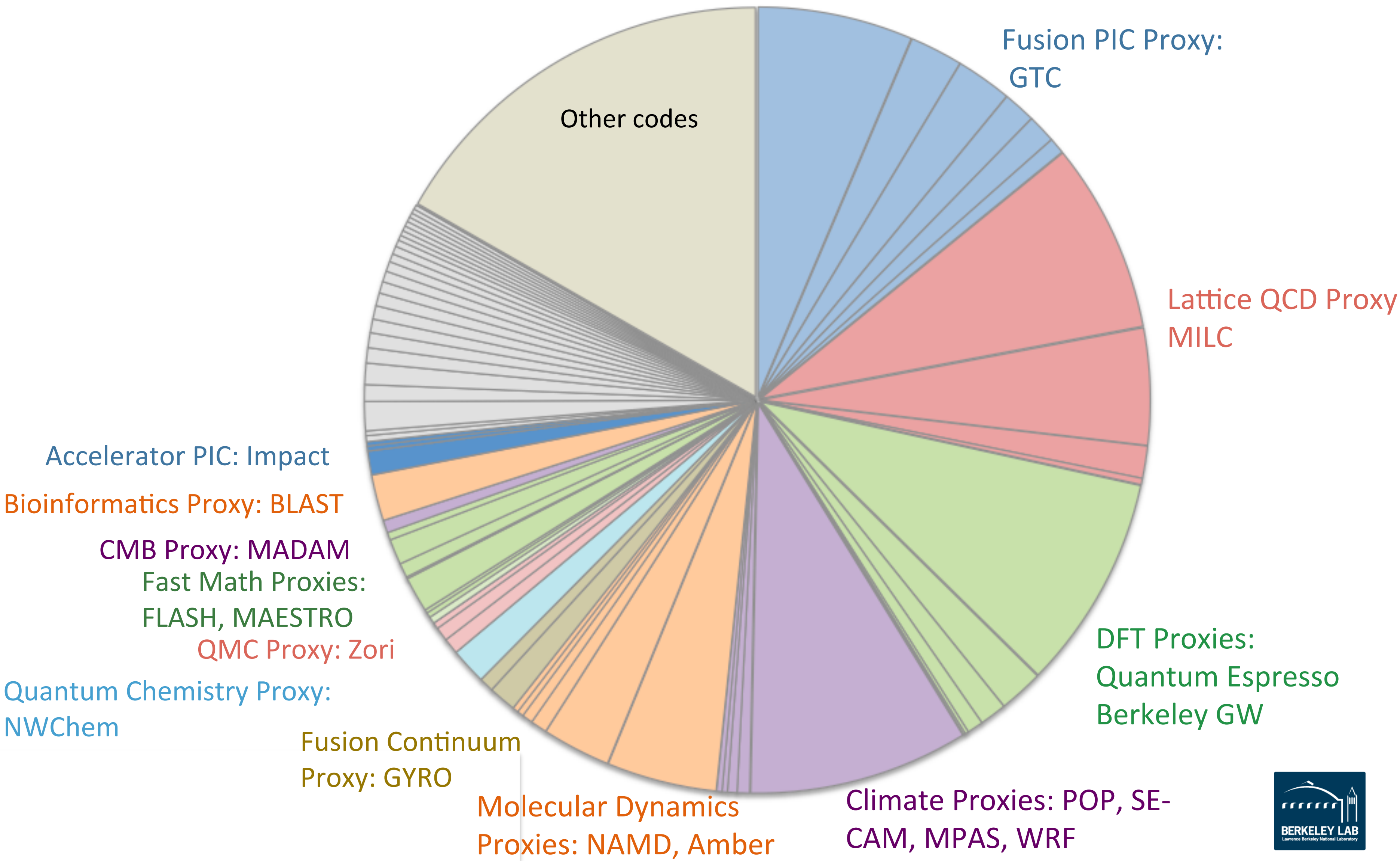
Hongzhang Shan: NWChem
(proxy: qchem, Gamess)

Helen He: WRF (Regional Climate)

Hari Krishnan: Vis, Iso-surface

The App Readiness team proxies cover almost 75% of the workload

Top Codes by Algorithm and Application Readiness Coverage



HPC motifs are well represented by App-readiness applications (or proxies)

Science Areas	Dense Linear Algebra	Sparse Linear Algebra	Spectral Methods (FFTs)	Particle Methods	Structured Grids	Unstructured or AMR Grids
Accelerator Science	Vorpal	Vorpal	**IMPACT**, Vorpal	**IMPACT**, Vorpal	**IMPACT**, Vorpal	Vorpal
Astrophysics		MAESTRO, CASTRO, **FLASH**			MAESTRO, CASTRO, **FLASH**	MAESTRO, CASTRO, **FLASH**
Chemistry	GAMESS, **NWChem**, qchem, QWalk, **Zori**	QWalk, **Zori**		Qwalk, **Zori**, NAMD, **AMBER**, Gromacs, LAMMPS		
Climate		lesmpi, global_fcst	**CAM**, MITgcm		**CAM, POP, WRF, MPAS**, MITgcm, lesmpi, global_fcst	
Combustion					MAESTRO, **S3D**	
Fusion	Xaorsa		Xaorsa, NIMROD	**GTC**, XGC, Osiris, gs2, GTS	**GTC**, Xaorsa, **gyro**, NIMROD, XGC, Osiris, gs2	
Lattice Gauge	MFD	MFD, **MILC**, chroma, hmc, qlua	**MILC**, chroma, hmc, qlua	**MILC**, chroma, hmc, qlua	**MILC**, chroma, hmc, qlua	
Material Science	PARATEC, cp2k, **QE**, VASP, **BerkeleyGW**, Abinit, qbox, PETot		PARATEC, cp2k, **QE**, VASP, **BerkeleyGW**, Abinit, qbox, PETot		PARATEC, cp2k, **QE**, VASP, **BerkeleyGW**, Abinit, qbox, PETot	

Work in progress

Application Readiness: Two-phased approach

1. **Explore characteristics of relevant architectures today**
    1. Collate community knowledge and evaluate current status
    2. Provide input into procurement decision
    3. Begin to determine best practices
2. **Once architecture is chosen**
    1. Each team member works on porting at least one application to advanced architecture
    2. Develop training materials to include more sophisticated examples, as well as the use of profiling tools and debuggers.
    3. Procure and deploy testbed systems for NERSC users and manage interactions with users

Conclusions

- **Disruptive technology changes are ~~coming~~ here !**
- **Your code will have to change !**
  - Hopefully in an evolutionary way
  - NERSC's strategy is designed smooth the path and maximize scientific productivity
- **NERSC-8 is expected to be an advanced architecture**
  - To achieve optimal performance code tuning and modification *will* be required
  - We are aiming to ensure that these changes will be robust
  - We hope that applications will only need to be re-complied to run, albeit with less than optimal performance

Bonus Topic: Energy to Solution

- **Will user be willing to optimize for energy used rather than time-to-solution?**
  - If you are given the appropriate tools
  -

NERSC

National Energy Research Scientific Computing Center

The path to Exascale: How and When to Move Users?

NERSC