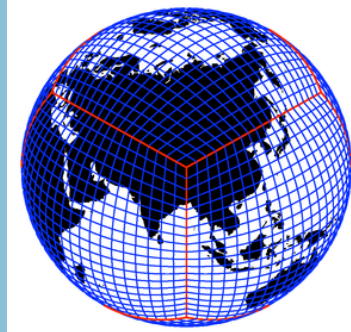


A Comparison of a spectral element dynamical core on several large platforms

John Dennis, Brian Dobbins, Chris Kerr,
Youngsung Kim

Application Scalability and Performance
Group, CISL

Motivation of HOMME optimization effort



- Atmosphere dynamical core (HOMME)
 - CAM: 35% of time (vert levels=32, # of tracers=25)
- Much easier to optimize than physics 😊
- Benchmark code
 - CORAL (CAM-SE)
 - NSF625
- Useful for evaluating full system performance

Group/Team

- Rich Loft, Division Director (NCAR)
- John Dennis, Scientist (NCAR)
- Chris Kerr, Software Engineer, contractor
- Youngsung Kim, Software Engineer (NCAR) / Graduate Student (CU)
- Brian Dobbins, Software Engineer (NCAR)
- Raghu Raj Prasanna Kumar, Associate Scientist (NCAR)
- Sheri Mickelson, Software Engineer (NCAR) / Graduate Student (CSU)
- Ravi Nanjundiah, Professor (IISc)

Related/Collaborative Activities

- Funding from Intel Parallel Computing Center (IPCC-WACS)
- NESAP (NERSC Exascale Science Application Program)
 - Bi-weekly: NERSC-Cray-NCAR telecon on CESM & HOMME performance (Feb 2015)
- Weekly Intel-TACC-NREL-NERSC-NCAR telecon
 - Concall focused on CESM/HOMME KNC performance

Current optimization focus

Xeon and Xeon Phi based platforms

- Broadwell (BDW) [i.e. Cheyenne]
- Knights Landing (KNL) [i.e. Cori Phase II]

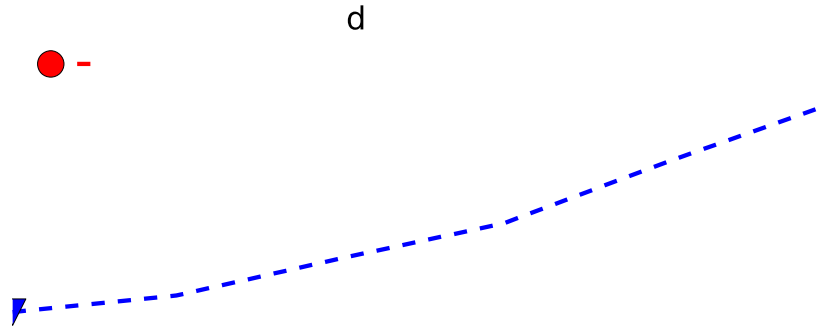
Impact of optimization at scale

- Edison [Ivybridge + Cray Aires]
- Cheyenne [Broadwell + EDR Infiniband]
- Cori [KNL + Cray Aires]

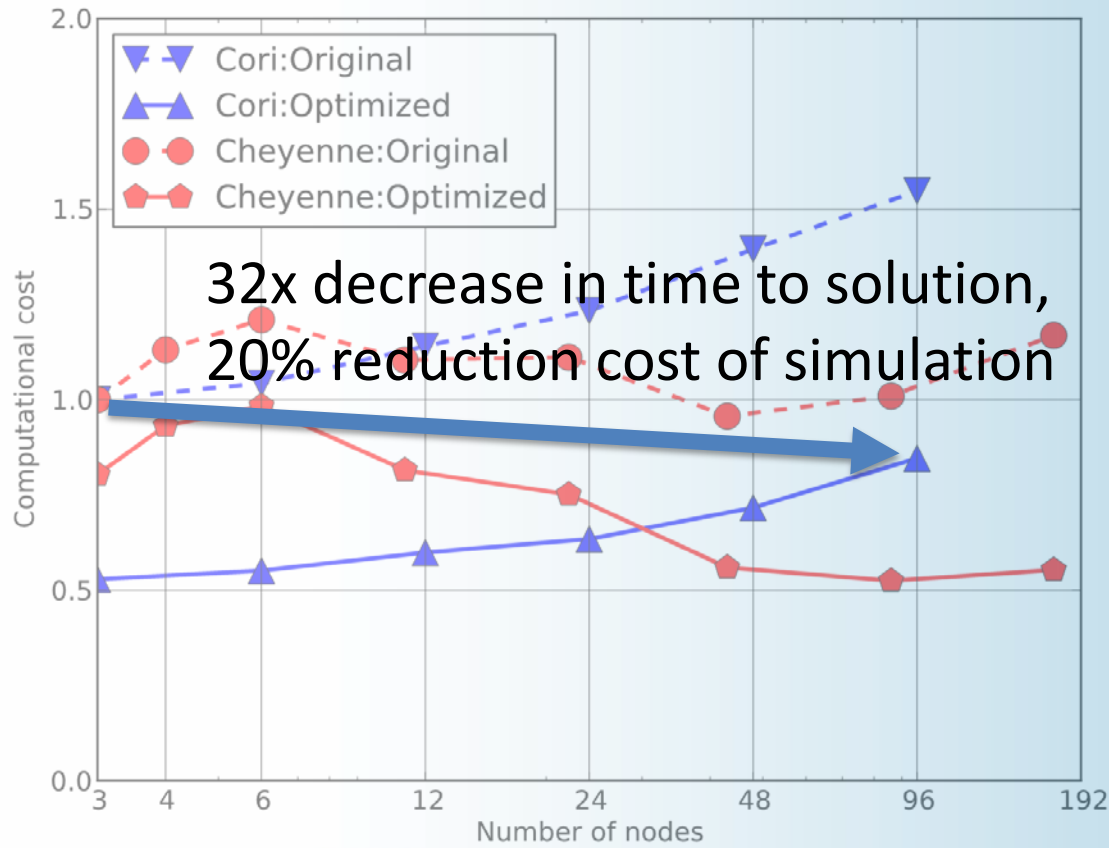
Optimization phases

1. Threading memory copy in boundary exchange [Jamroz]
2. Restructure data-structures for vectorization [Vadlamani & Dennis]
3. Rewrite message passing library/ specialized comm ops [Dennis]
4. Rearrange calculations in euler_step for cache reuse [Dennis]
5. Reduced # of divides [Dennis]
6. Restructured/alignment for better vectorization [Kerr]
7. Rewrote and optimized limiter [Demeshko & Kerr]
8. Redesign of OpenMP threading [Kerr & Dennis]
9. Flexible MPI message passing back-ends [Dennis]
 1. MPI_Put/Get (MPI3)
 2. MPI neighborhood collectives (MPI3)
10. Replaced all functions with subroutines [Kerr & Dennis]
11. Custom OpenMP barrier [Dobbins]

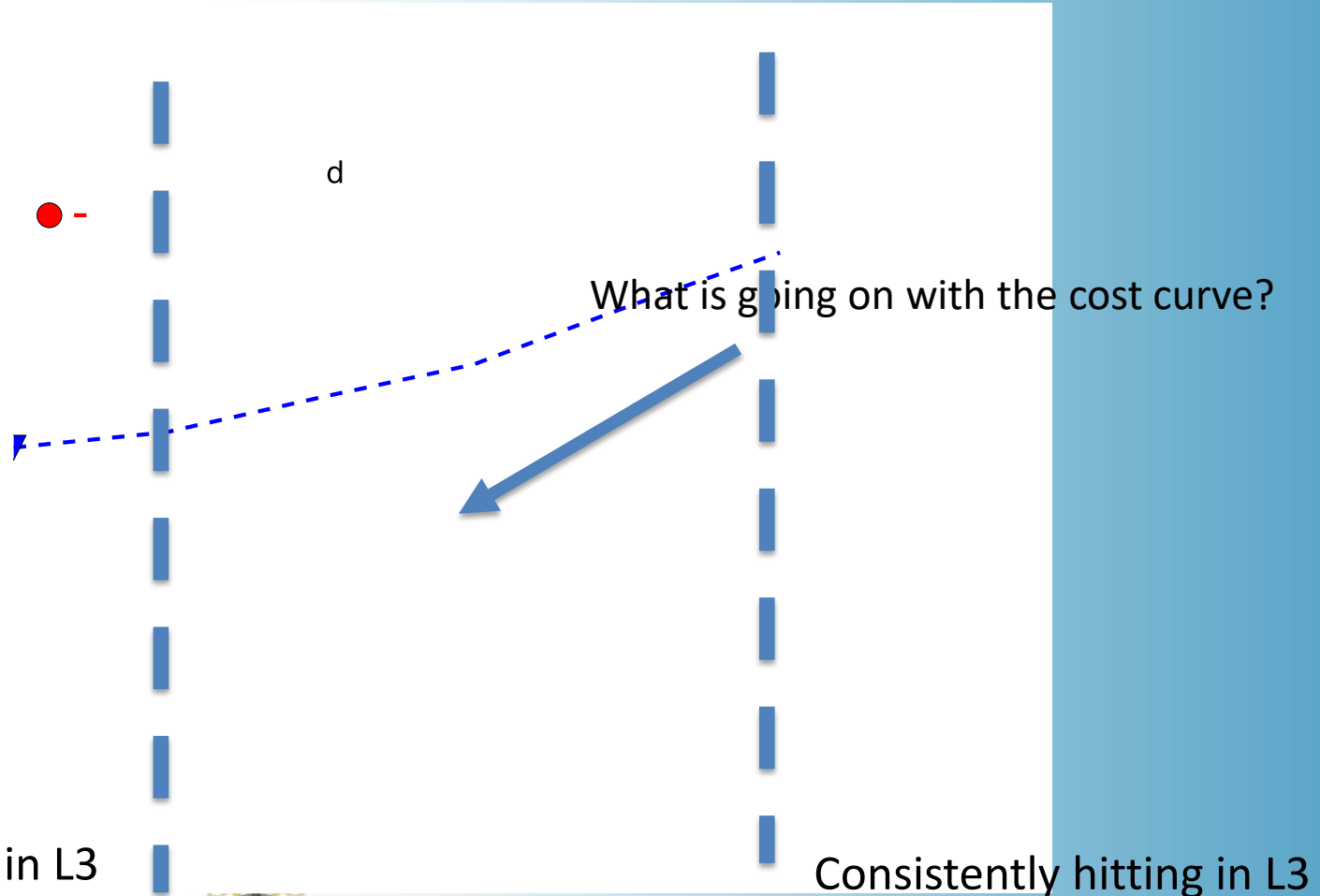
Simulation cost for HOMME on Xeon and Xeon Phi @ 100 km



Simulation cost for HOMME on Xeon and Xeon Phi @ 100 km



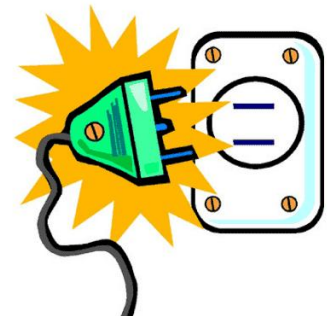
Simulation cost for HOMME on Xeon and Xeon Phi @ 100 km



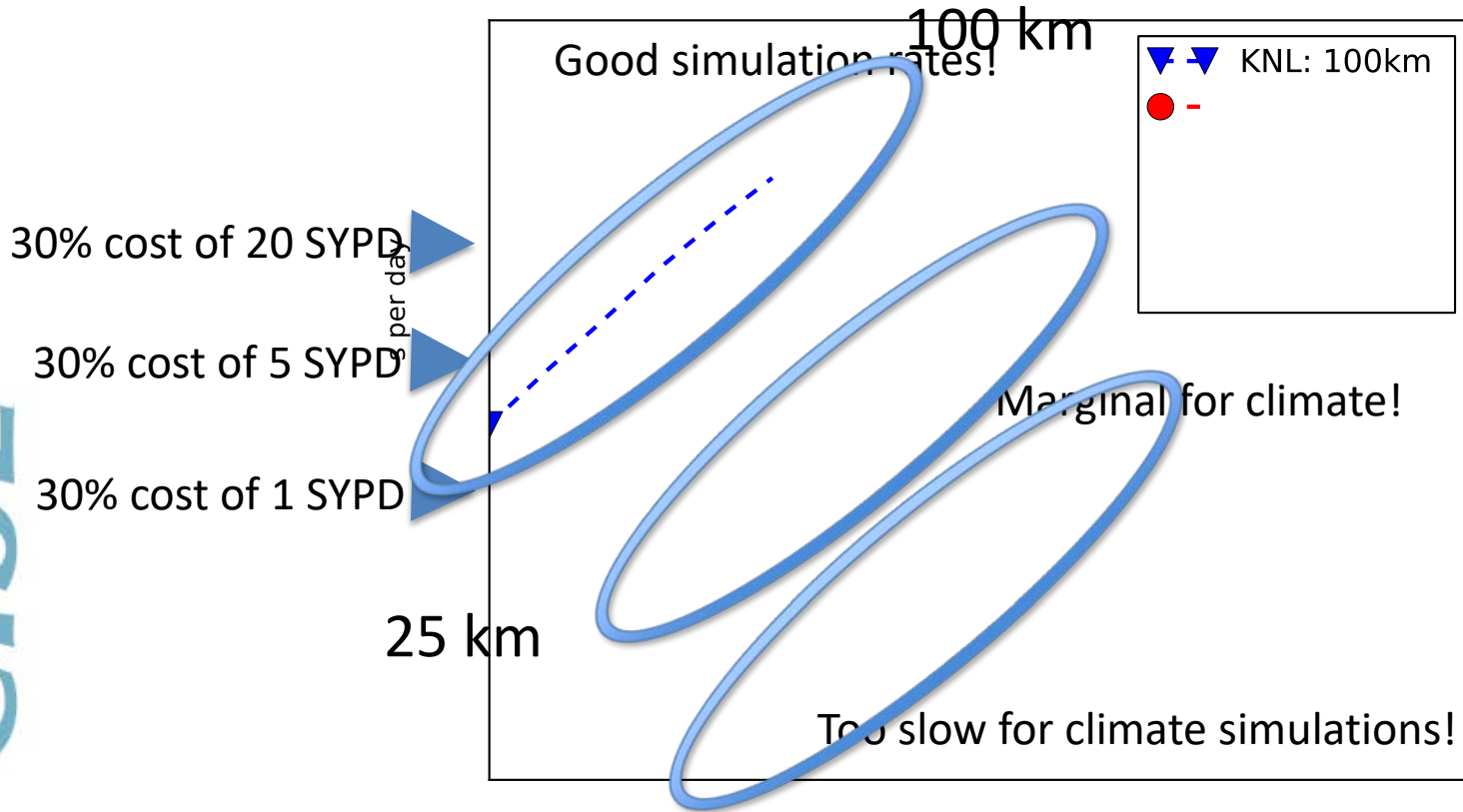
Energy usage for HOMME (NGGPS-like) on Xeon and Xeon Phi @ 12 km



ELECTRICITY



Simulation rate for HOMME on Xeon and KNL



Additional activities

- HOMME optimization work has completed
- Optimized code will be released in CESM2
- Backport currently underway into CESM1.3
- Evaluate Cheyenne, Cori, Sunway TaihuLight
- Case study:

Optimizing the HOMME Dynamical-Core for Multi-Core Platforms, J.M. Dennis, B. Dobbins, C. Kerr, Y. Kim, submitted to IJHPCA on Oct 15, 2017

Conclusions/Future work

- Concerted/sustained effort reduced cost of HOMME on Xeon and Xeon Phi
 - 2x speedup on KNL and BDW
 - Optimizations reduces time to solution by 32x for similar simulation cost
- The accumulation of incremental improvements will enable transformative capabilities!
- Future activities
 - Evaluate new Intel compiler
 - Engage MPI library community
 - Evaluate single node of upcoming systems
 - AMD EYPC
 - Intel Skylake
 - Cavium Thunder X2
 - IBM Power 9

Questions?

dennis@ucar.edu