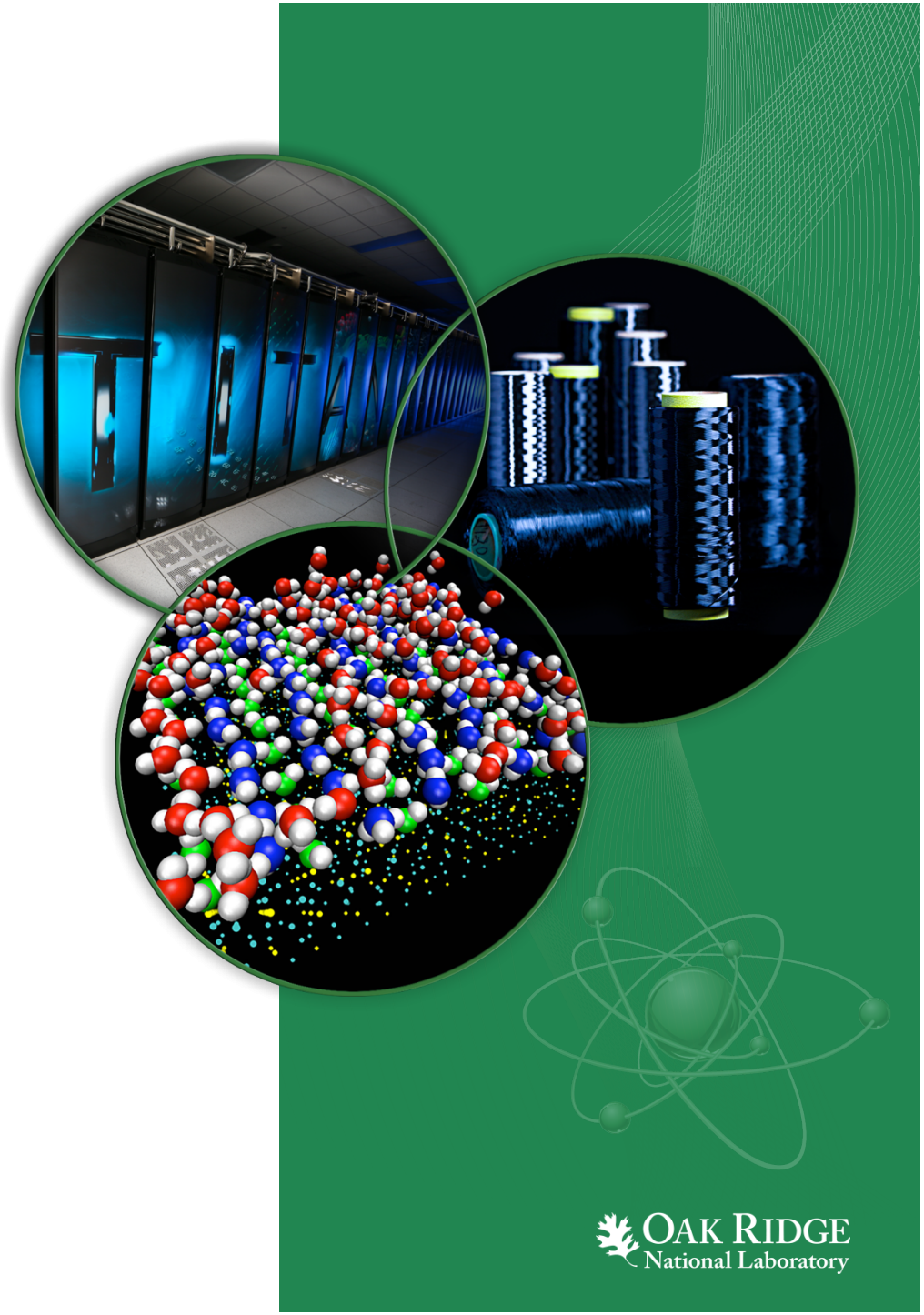
Thread Layout on Eos & Rhea

Robert D. French
NCCS User Assistance



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

OAK RIDGE National Laboratory

Thread Affinity

- Defines which CPUs a thread prefers
  - Each thread has an “Affinity Mask”
- Kernel migrates threads based on mask
- aprun/mpirun set Rank affinity
- OpenMP runtime suggests thread affinity to OS

OAK RIDGE National Laboratory

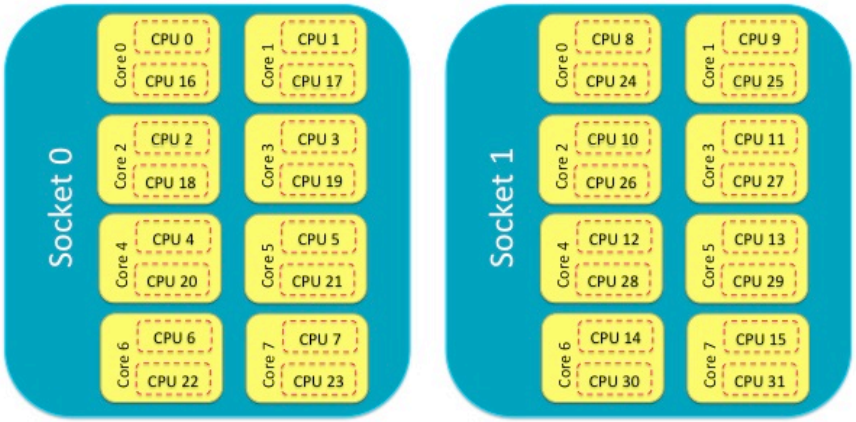
Thread Affinity

- Each compiler provides an OpenMP runtime
- Default affinity varies between vendors
  - Can be adjusted via environment vars or aprun options

OAK RIDGE National Laboratory

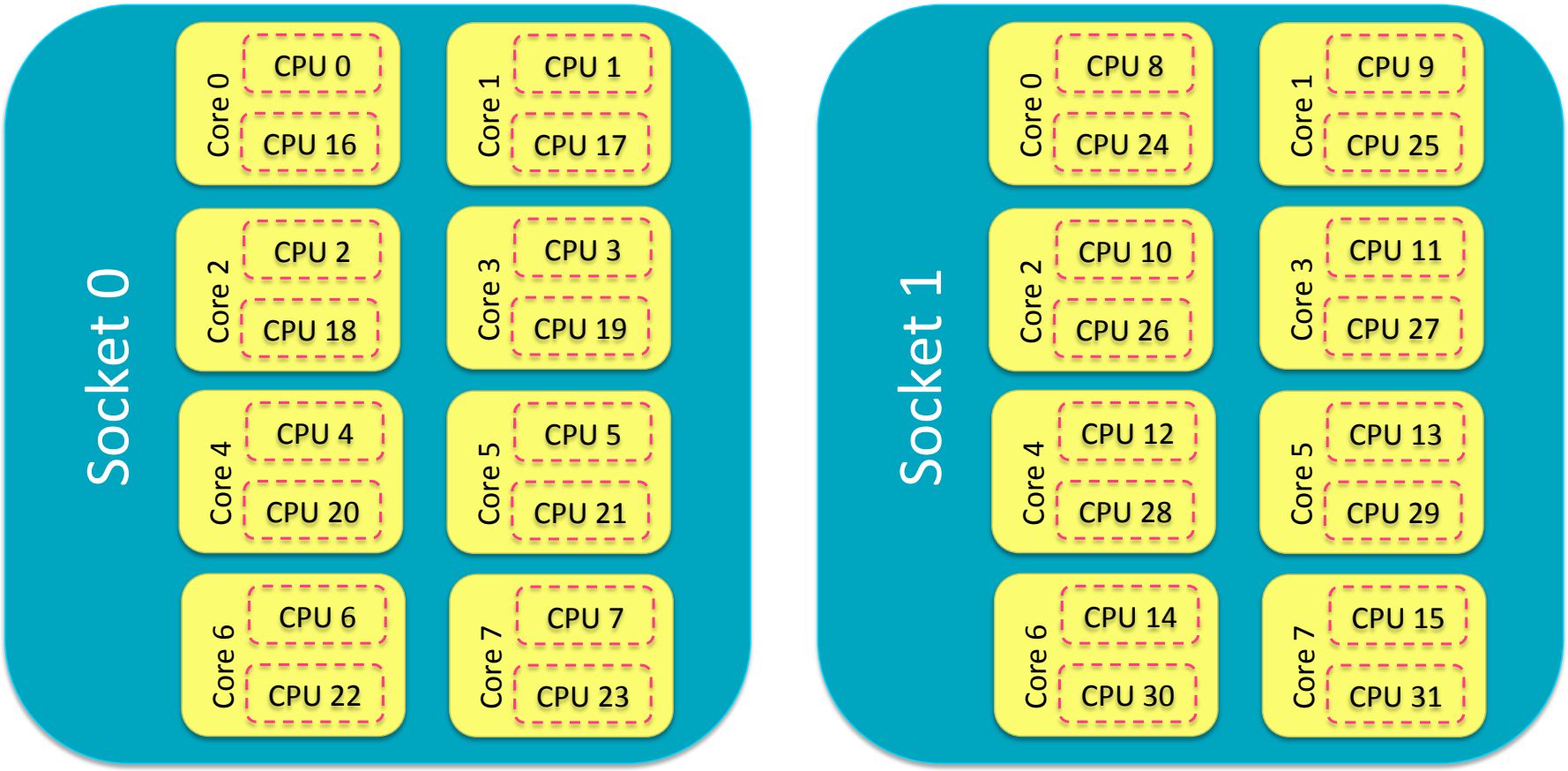
HyperThreading Overview

- Both machines feature two (2) Intel Xeon *packages*
  - Eos has Xeon E5-2670s
  - Rhea has Xeon ER-2650s
- Eight (8) physical *cores* each
- Each core provides two (2) virtual *CPUs* to the OS



OAK RIDGE National Laboratory

Virtual CPU Layout



OAK RIDGE National Laboratory

Thread Affinity Probe

- Spawns threads according to OMP_NUM_THREADS
- Each thread identifies its host CPU
- Threads collaborate on FP intensive task
  - Not rigorous, but more threads should mean less time
- Goal is 1 Thread per physical core
  - i.e. want to “avoid” hyperthreading

OAK RIDGE National Laboratory

Thread Affinity on Eos

OAK RIDGE
National Laboratory

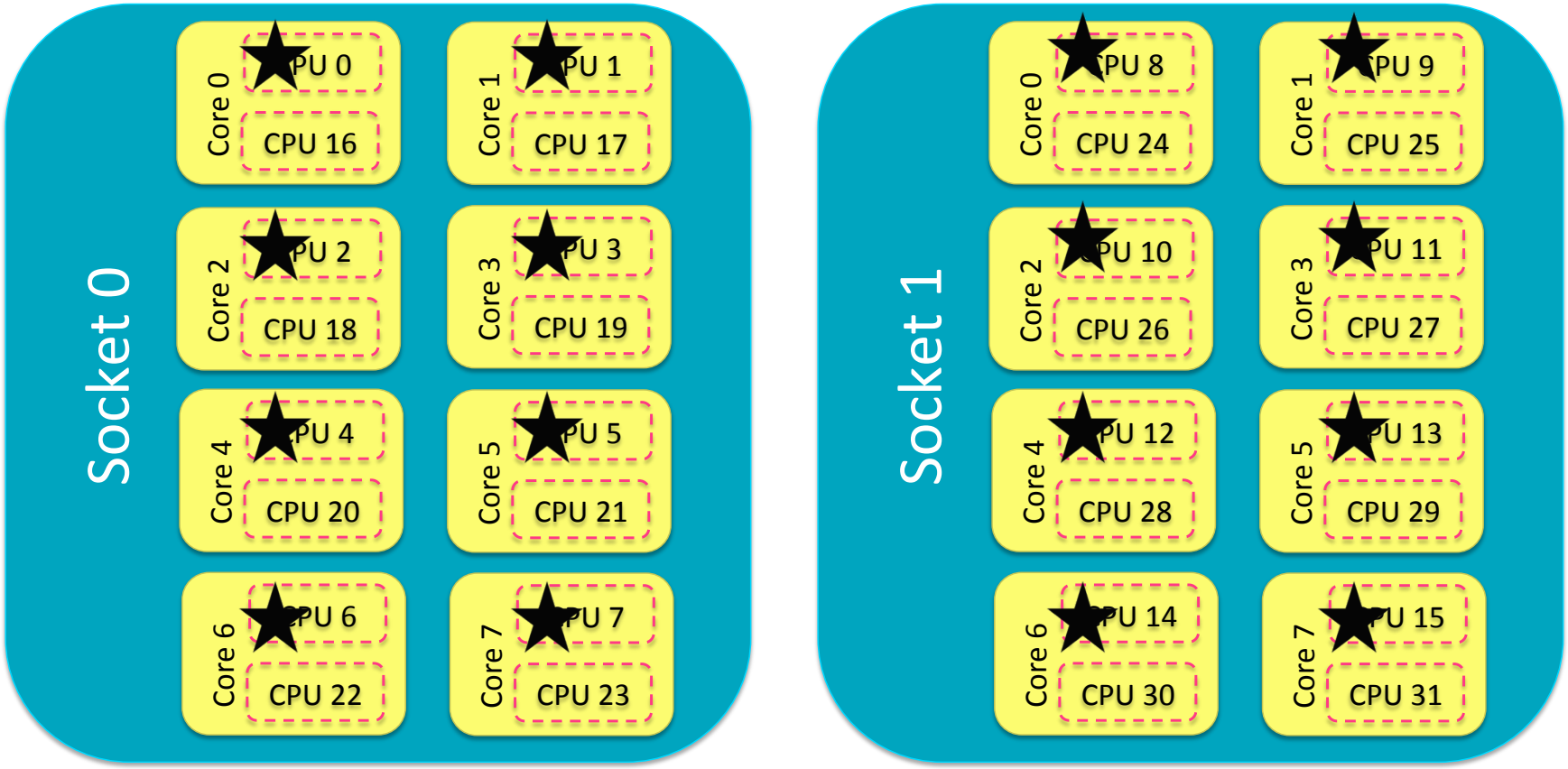
Eos - 2 Threads per Rank

- 8 MPI Ranks per node
- Declare 2 OpenMP threads
  - `$ export OMP_NUM_THREADS=2`
- Launch with thread depth of 2
  - `$ aprun -n8 -N8 -d2 ./eos.exe`

OAK RIDGE National Laboratory

Eos - 2 Threads per Rank (PGI)

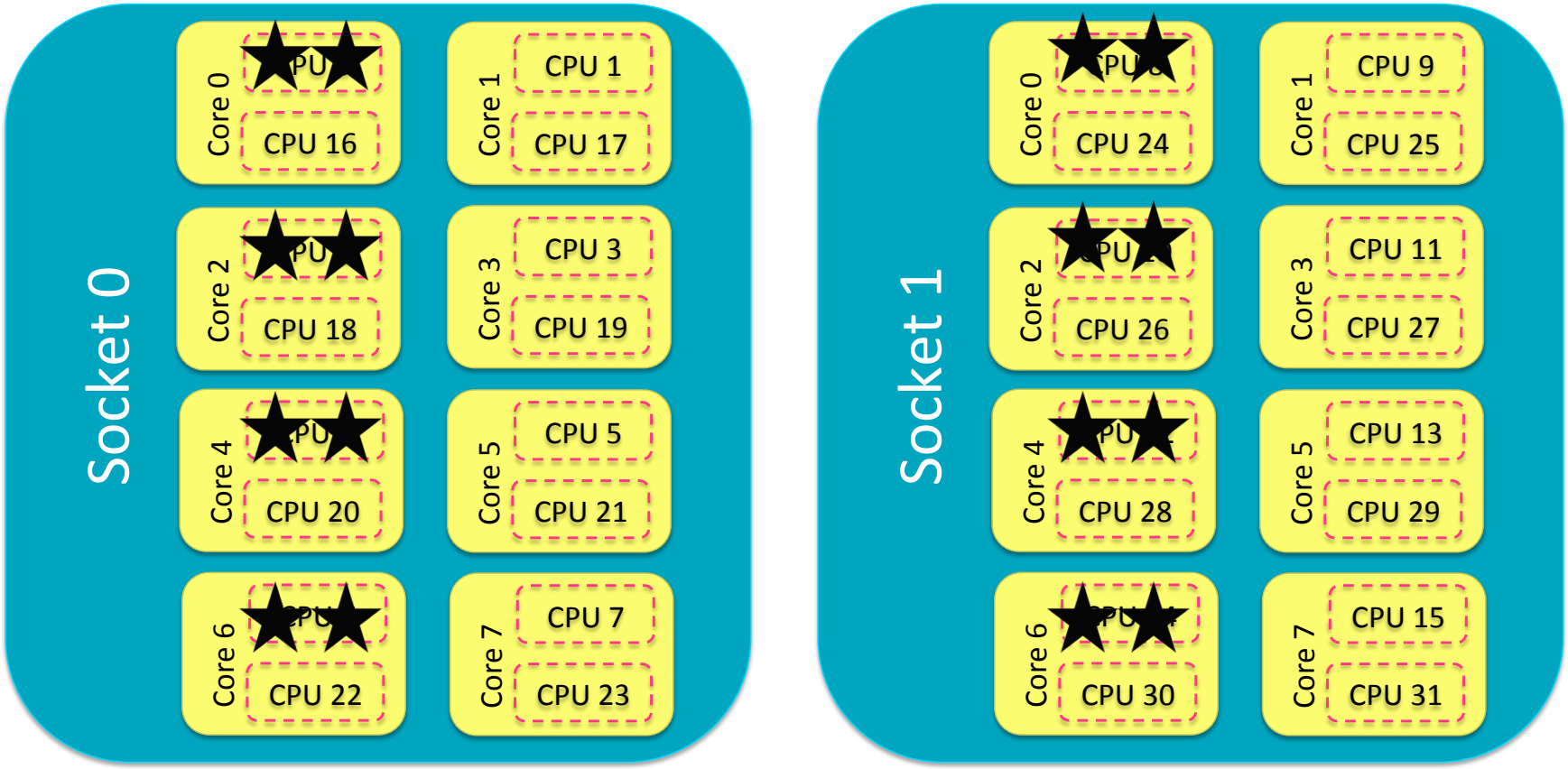
- Thread Affinity assigned by PGI runtime



OAK RIDGE National Laboratory

Eos - 2 Threads per Rank (Intel)

- Thread Affinity assigned by Intel runtime



OAK RIDGE National Laboratory

Eos - 1 Rank, 16 Threads

- PGI, Cray, GNU compilers:
  - `$ OMP_NUM_THREADS=16`
  - `$ aprun -n1 -N1 -d16 ./a.out`
  - 1 thread per physical core
- Intel 13 compiler:
  - `$ OMP_NUM_THREADS=16`
  - `$ aprun -n1 -N1 -d16 -cc none ./a.out`
  - Threads may alternate within virtual CPU pair
  - Does not appear to affect performance

OAK RIDGE National Laboratory

Eos - 2 Ranks, 8 Threads Each

- PGI, Cray, GNU compilers:
  - `$ OMP_NUM_THREADS=8`
  - `$ aprun -n2 -N2 -d8./a.out`
  - 1 thread per physical core
- Intel 13 compiler:
  - `$ OMP_NUM_THREADS=8`
  - `$ aprun -n2 -N2 -d8 -cc numa_node./a.out`
  - 1 Thread per physical core

OAK RIDGE National Laboratory

Eos - 4 Ranks, 4 Threads Each

- PGI, Cray, GNU compilers:
  - `$ OMP_NUM_THREADS=4`
  - `$ aprun -n4 -N4 -d4 ./a.out`
  - 1 thread per physical core
- Intel 13 compiler:
  - `$ OMP_NUM_THREADS=4`
  - `$ aprun -n4 -N4 -d4 -cc numa_node./a.out`
  - 1 thread per physical core

OAK RIDGE National Laboratory

Eos Summary

- PGI, GNU, Cray compilers
  - Set thread depth, run as normal
- Intel Compiler
  - Set thread depth and
    - -cc numa_node if using multiple ranks per node
    - -cc none if using 1 rank per node
    - Layout may fluctuate within virtual CPU pairs

OAK RIDGE National Laboratory

Thread Affinity on Rhea

OAK RIDGE
National Laboratory

Rhea - 1 Rank, 16 Threads

- All compilers:
  - `$ OMP_NUM_THREADS=16`
  - `$ mpirun –np 1 –bind-to-core –bycore –cpus-per-proc 16 ./a.out`
  - Use –report-bindings to display which cores are bound to each rank

OAK RIDGE National Laboratory

Rhea - 2 Ranks, 8 Threads Each

- All compilers:
  - `$ OMP_NUM_THREADS=8`
  - `$ mpirun –np 2 –bind-to-core –bycore –cpus-per-proc 8 ./a.out`
  - Use –report-bindings to display which cores are bound to each rank

OAK RIDGE National Laboratory

Rhea - 4 Ranks, 4 Threads Each

- All compilers:
  - `$ OMP_NUM_THREADS=4`
  - `$ mpirun –np 4 –bind-to-core –bycore –cpus-per-proc 4 ./a.out`
  - Use –report-bindings to display which cores are bound to each rank

OAK RIDGE National Laboratory

Rhea Summary

- Same launch behavior for all compilers
- Codes compiled with Intel may alternate between virtual CPU pairs
  - Does not appear to impact performance

OAK RIDGE National Laboratory

Questions?

OAK RIDGE
National Laboratory