Intel Xeon Phi Computing

Aiichiro Nakano

Collaboratory for Advanced Computing & Simulations
Department of Computer Science
Department of Physics & Astronomy
Department of Chemical Engineering & Materials Science
Department of Biological Sciences
University of Southern California

Email: anakano@usc.edu

Goal: Multithreading on Intel Xeon Phi

Two Supercomputing Parties in the US

GPU

Titan: Oak Ridge Nat'l Lab

17.6 Petaflop/s

AMD Opteron + NVIDIA K20x

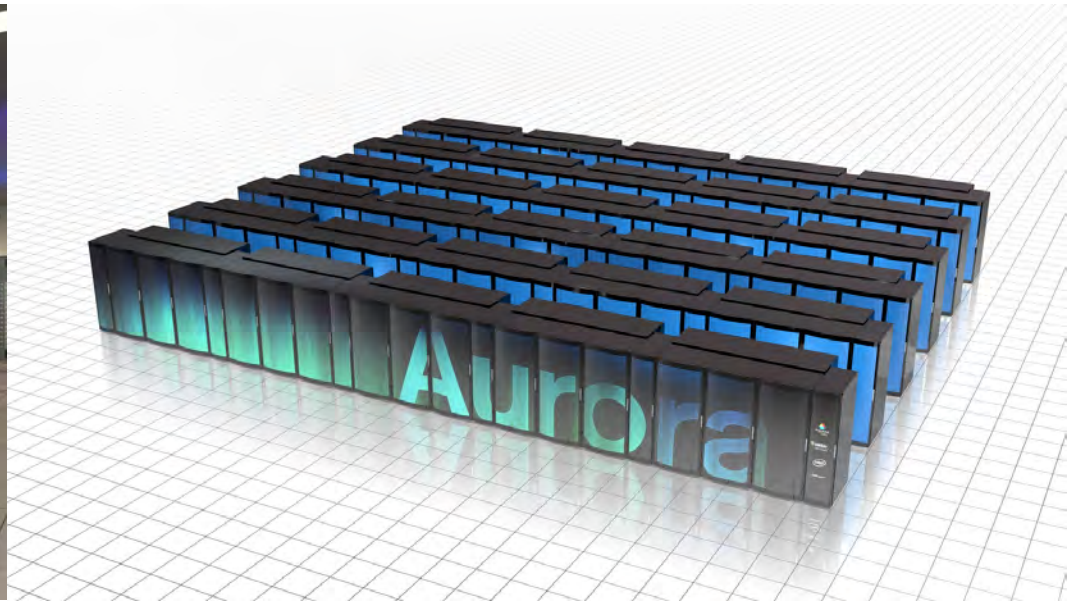
Summit: 5-10x performance (2018)

CRAY THE SUPERCOMPUTER COMPANY

AMD

NVIDIA

Phi



Aurora: Argonne Nat'l Lab (2021)

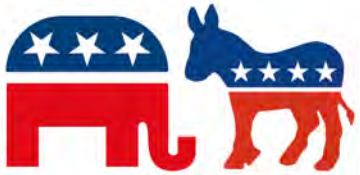
Exaflop/s

Intel Xeon Phi

CRAY THE SUPERCOMPUTER COMPANY

intel

GPU *vs.* Phi



Current & Future Computing Platforms

- **Two DOE supercomputing awards to develop & deploy metascalable ("design once, scale on future platforms") simulation algorithms (2017-2020)**

U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

- **NAQMD & RMD simulations on full 800K cores**

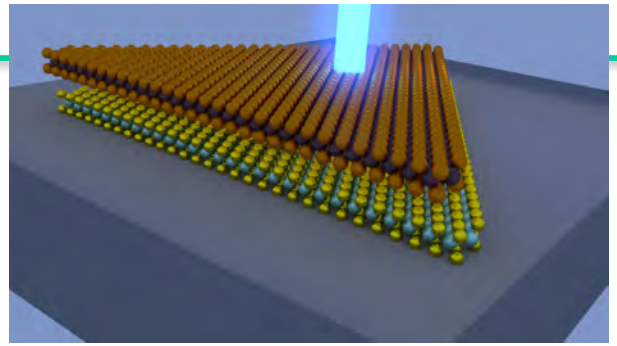
Innovative & Novel Computational Impact on Theory & Experiment

Title: "Petascale Simulations for Layered Materials Genome"

Principal Investigator: Aiichiro Nakano, University of Southern California
Co-Investigator: Priya Vashishta, University of Southern California

786,432-core IBM Blue Gene/Q

AURORA | EARLY SCIENCE PROGRAM



Early Science Projects for Aurora Supercomputer Announced

Metascalable layered materials genome

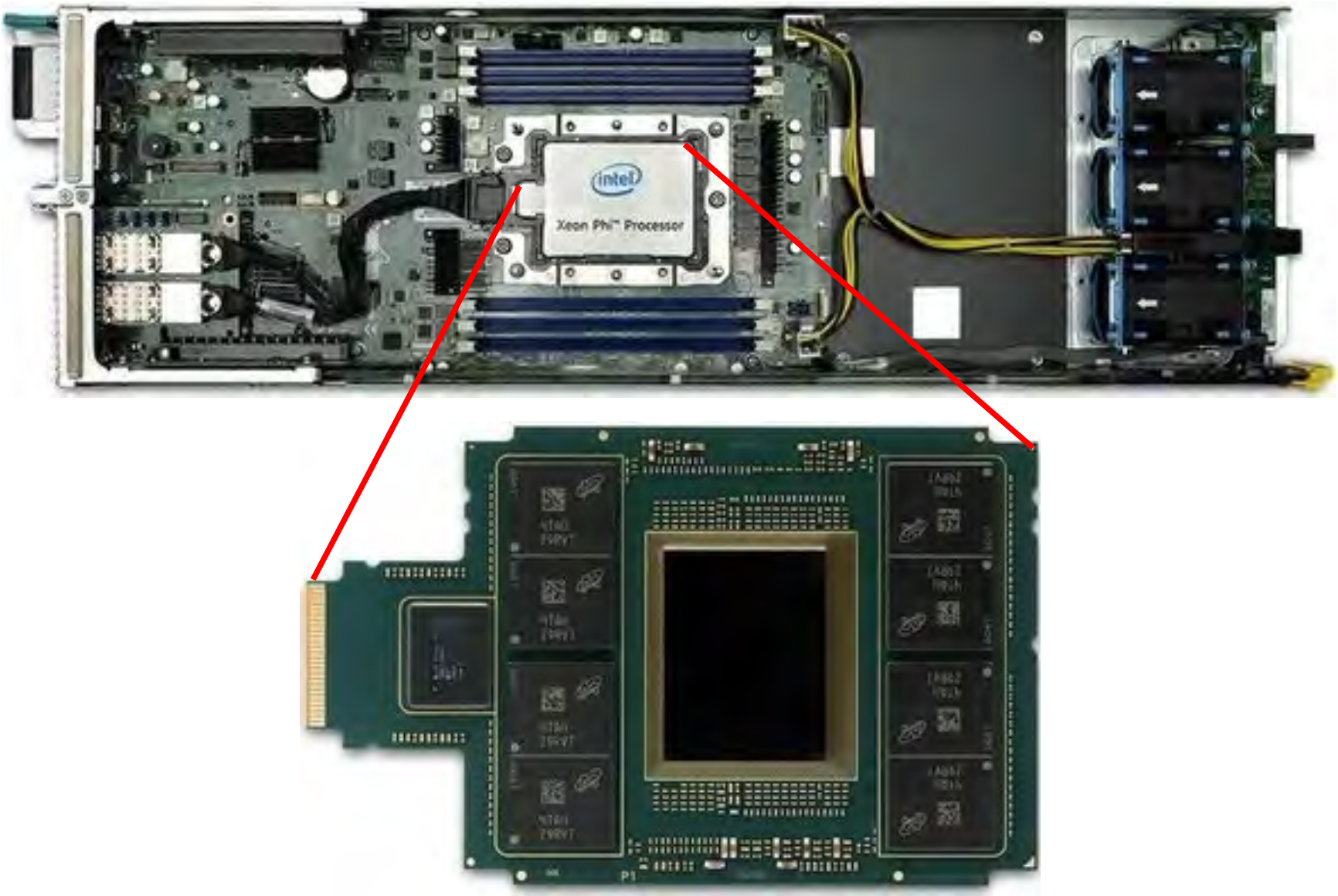
Investigator: Aiichiro Nakano, University of Southern California

Exaflop/s
Intel Xeon Phi Knights Hill (KNH) (2021)

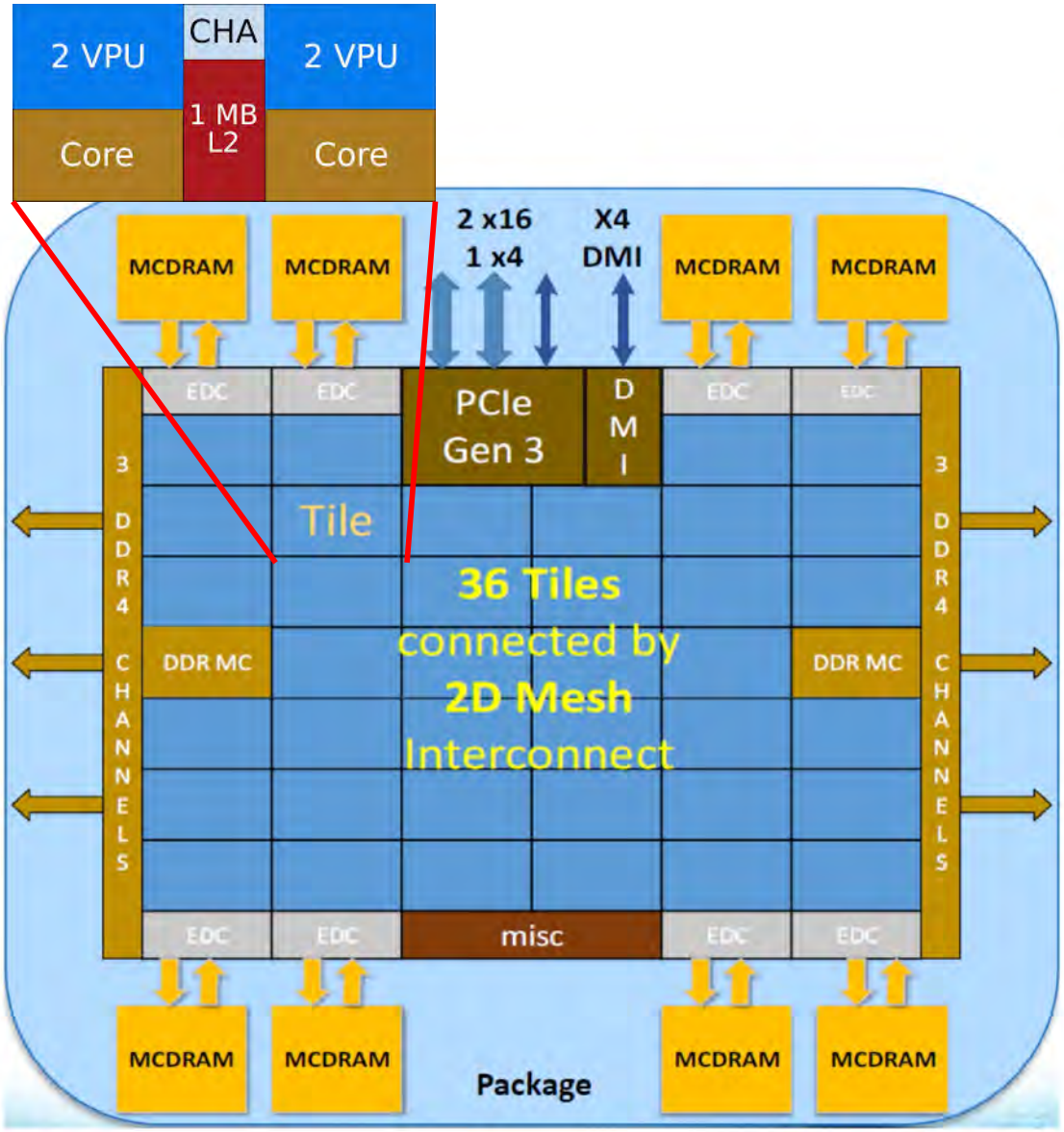
- **One of 10 exclusive users of the next-generation DOE supercomputer**

Intel Xeon Phi Processors

Current Knights Landing (KNL) is a predecessor of the Knights Hill (KNH) processor in Aurora



Knights Landing (KNL)



VPU: Vector processing unit (512 bits)
MCDRAM: Multi-channel dynamic random access memory (4× bandwidth of DRAM)

Chip
- 683 mm²
- 14 nm process
- 8 Billion transistors

Up to 72 Cores
- 36 tiles
- 2 cores per tile
- 2.4 TF per node

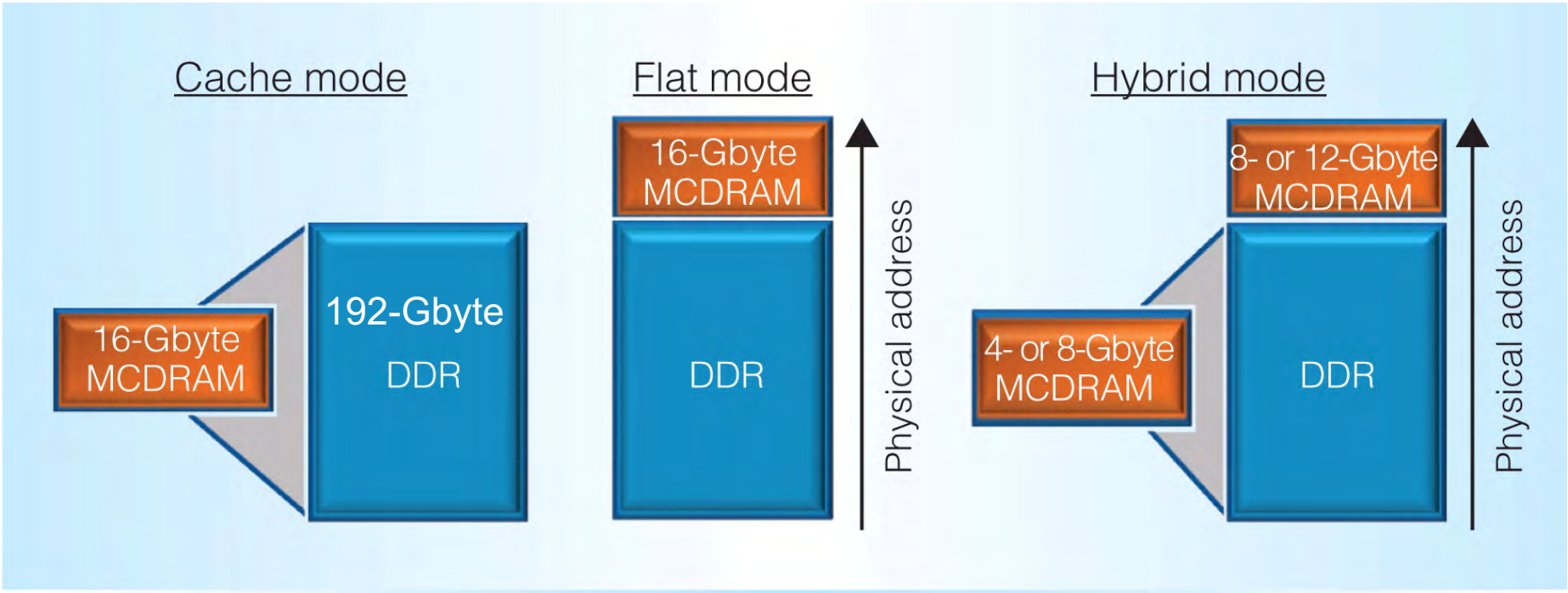
2D Mesh Interconnect
- Tiles connected by 2D mesh

On Package Memory
- 16 GB MCDRAM
- 8 Stacks
- 485 GB/s bandwidth

6 DDR4 memory channels
- 2 controllers
- up to 384 GB external DDR4
- 90 GB/s bandwidth

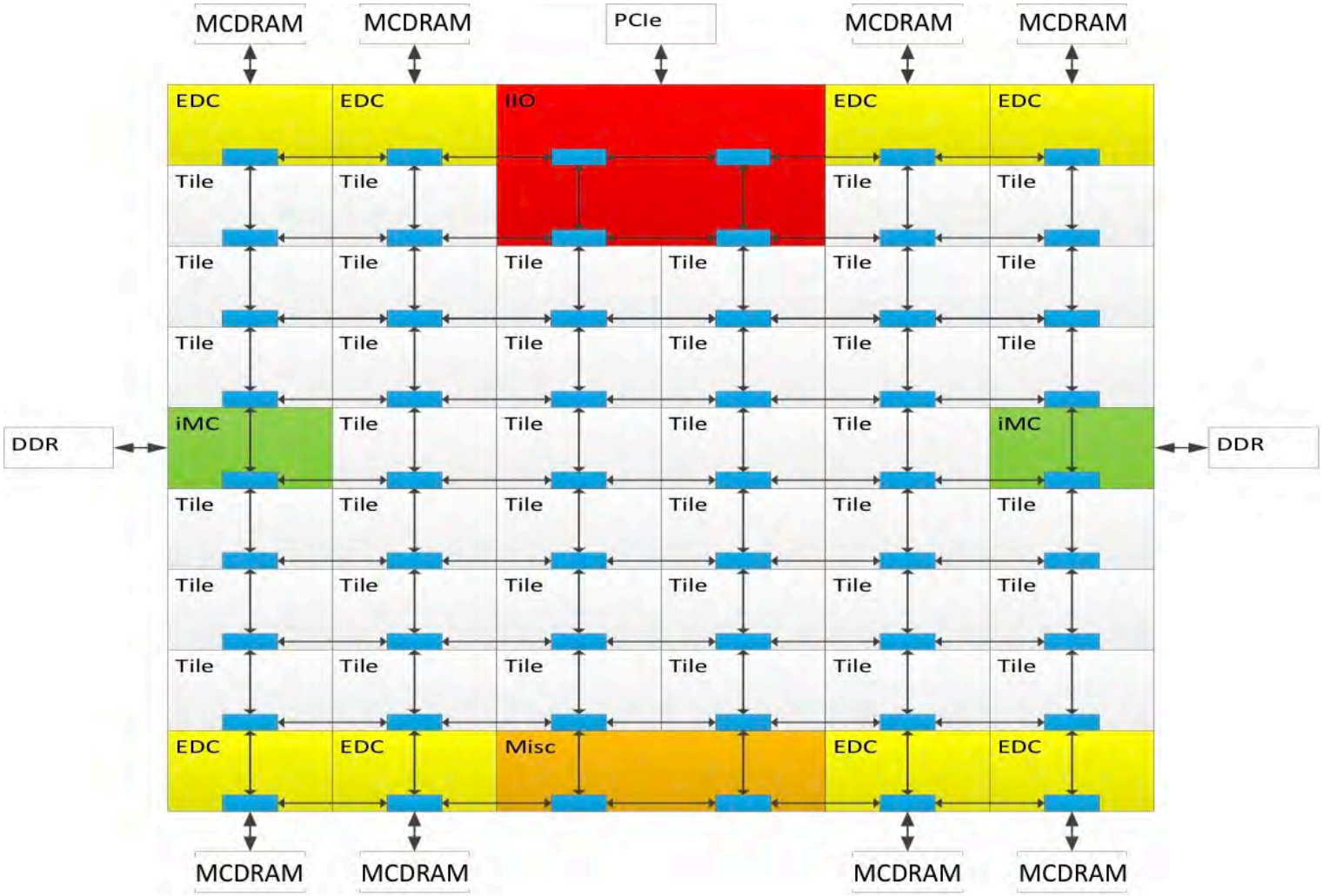
On Socket Networking
- Omni-Path NIC on package
- Connected by PCIe

Memory Modes



- **MCDRAM: Multi-channel dynamic random access memory (4× bandwidth of DRAM)**

On-Chip Mesh Interconnect



- **YX routing**
- **3 cluster modes: (1) all-to-all, (2) quadrant, (3) sub-NUMA (non-uniform memory access)**

Theta at Argonne National Laboratory

System:

- Cray XC40 system
- 3,624 compute nodes/ 231,936 cores
- 9.6 petaflop/s peak performance

Processor:

- Second generation Intel Xeon Phi, Knights Landing (KNL) 7230
- 64 cores (up to 72 cores)
- 1.3 GHz

Memory:

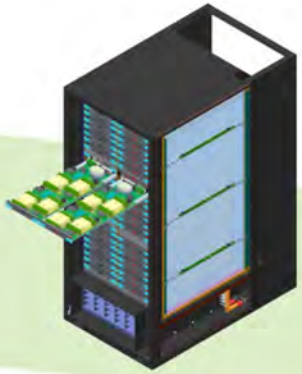
- 736 TB of total system memory
- 16 GB fast MCDRAM per node
- 192 GB DDR4-2400 per node

Network:

- Cray Aries interconnect
- Dragonfly network topology

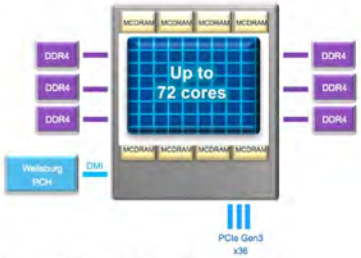
Theta Organization

System: 20 Cabinets
3264 Nodes, 960 Switches
10 groups, Dragonfly 7.2 TB/s Bi-Sec
9.65 PF Peak
56.6 TB MCDRAM, 679.5 TB DRAM



Cabinet: 3 Chassis
510.72 TF
3TB MCDRAM, 36TB DRAM

Chassis: 16 Blades
64 Nodes, 16 Switches
170.24 TF
1TB MCDRAM, 12TB DRAM



Node: KNL Socket
2.66 TF
16GB MCDRAM, 192 GB DDR4 (6 channels)

Compute Blade:
4 Nodes/Blade + Aries switch
10.64 TF
64GB MCDRAM, 768GB DRAM
128GB SSD

Sonexion Storage
4 Cabinets
Lustre file system
10 PB usable
210 GB/s

Argonne NATIONAL LABORATORY

KNL Parallel Programming

- **Standard MPI+OpenMP programming is supported***
- **Should take advantage of AVX-512 (512-bit or 8 double-precision) SIMD (single-instruction multiple-data) operations on VPUs (vector processing units)**
- **Should utilize fast on-chip MCDRAM (multi-channel dynamic random access memory) shared by 72 cores**

***Hyperthreading technology supports 4 simultaneous multithreads (SMTs) per core, with out-of-order execution of instructions**

Program with many threads on vector data!