FUTURE OF EVENT-DRIVEN SIMULATION & SUPERCOMPUTER SYSTEMS

Chris Carothers

Center for Computational Innovations

Rensselaer Polytechnic Institute

Extreme Heterogeneous Supercomputer Design

2017 OLCF Leadership System

Hybrid CPU/GPU architecture

Vendor: IBM (Prime) / NVIDIA™ / Mellanox®

At least 5X Titan's Application Performance

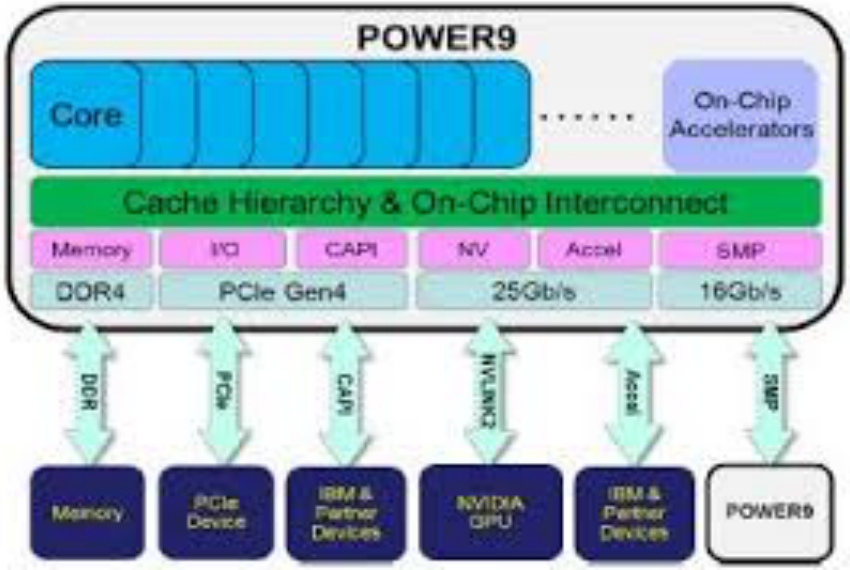
Approximately 3,400 nodes, each with:

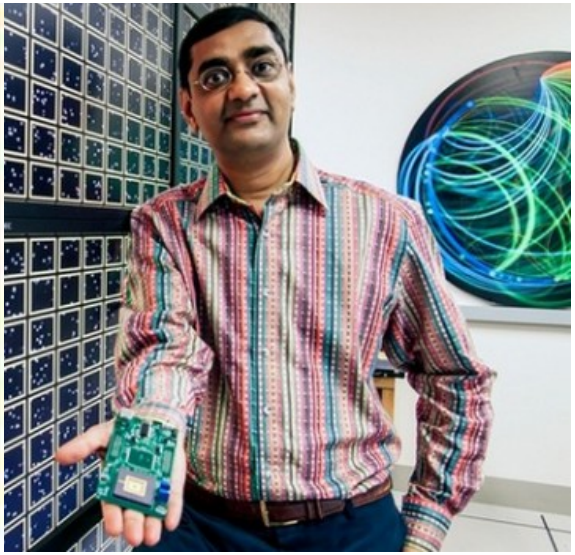
- Multiple IBM POWER9™ CPUs and multiple NVIDIA Volta® GPUs. CPUs and GPUs completely connected with high speed **NVLink**
- Large coherent memory: over 512 GB (**HBM** + DDR4)
  - all directly addressable from the CPUs and GPUs
- An additional 800 GB of NVRAM, which can be configured as either a **burst buffer** or as extended memory
- over 40 TF peak performance

Dual-rail Mellanox® EDR-IB full, non-blocking fat-tree interconnect

IBM Elastic Storage (GPFS™) - 1TB/s I/O and 120 PB disk capacity.

FPGA, NVMe & Neuromorphic Accelerators…



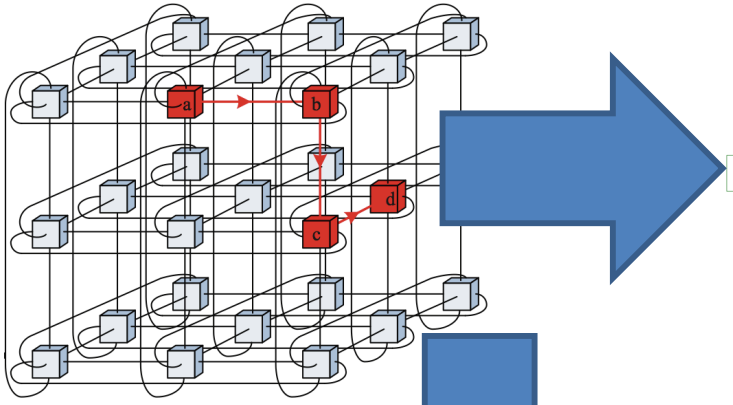


All devices exchange data via coherent shared memory!!

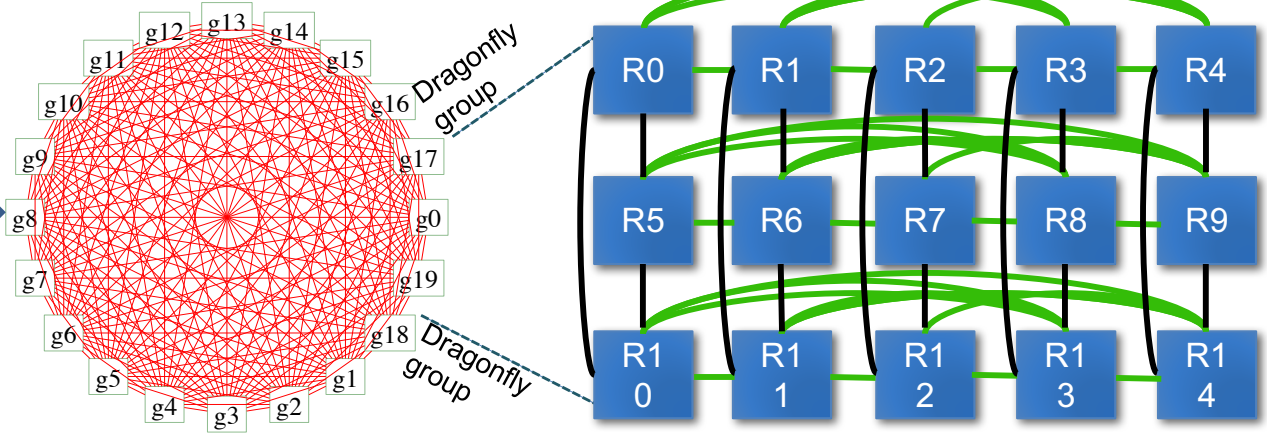
Structure of Modern Interconnect Topologies

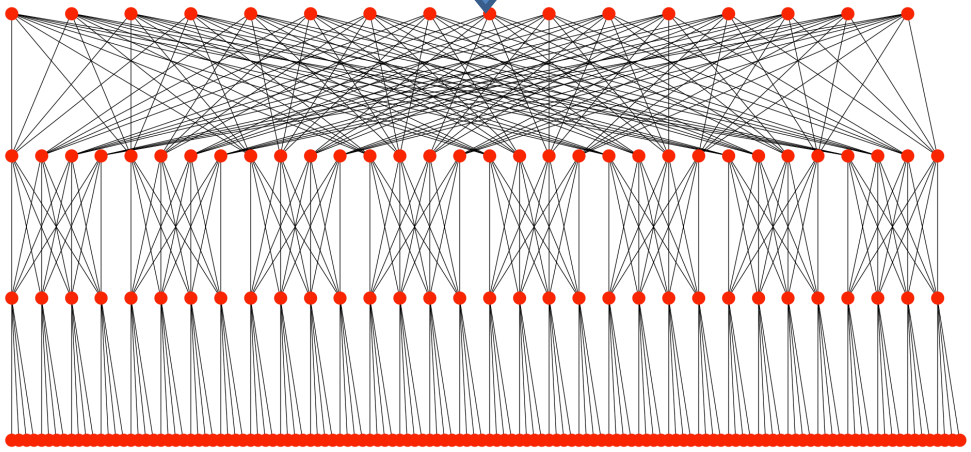
"Good Bye" Torus Network

- Homogeneous structure
- Dedicated routers per compute node

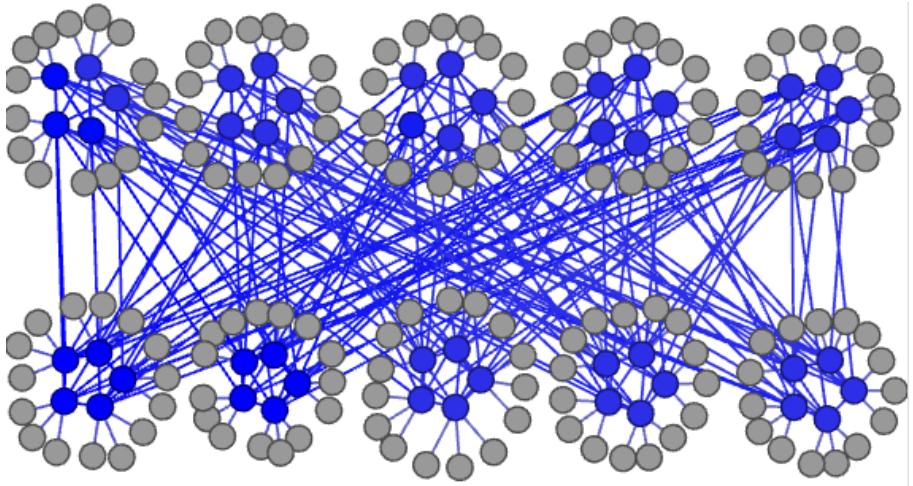


"Hello" Dragonfly Network for high node count systems





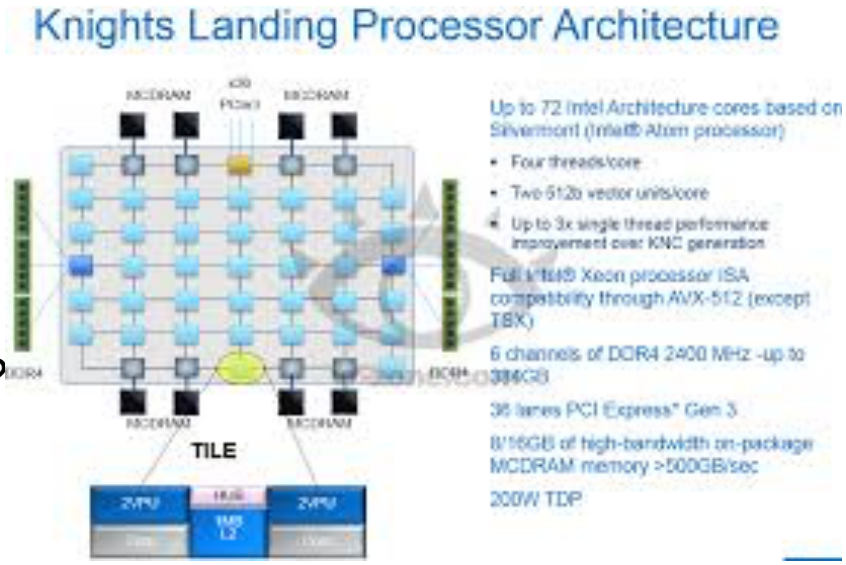
"Revised" Fat Tree Network for low node count systems



Extreme Low Diameter "Slimfly" Network Could be a great option for PDES like apps

Challenge for PDES

- Fortunately, “many-core” systems like Theta provide a path forward.
  - Need to re-worry about how to manage shared memory as we did for the KSR-2 and SGI Origin machines
- How does PDES take some advantage of GPU hardware ?
  - Offload approach ?
  - Complete system on GPU ?
- Can PDES be tolerant of external job network interference ?
- How does PDES leverage node-local & burst buffer storage?
  - Checkpoint/restart data?
  - Backing storage for live LP state for big models?
  - Intermediate results prior to running UQ tasks?
- How can PDES leverage ML/neuromorphic accelerators ?
  - Auto-performance tuning ?
  - System failure prediction ?



Thank you!

Summer of CODES Workshop – ANL
https://press3.mcs.anl.gov/summerofcodes2017/