

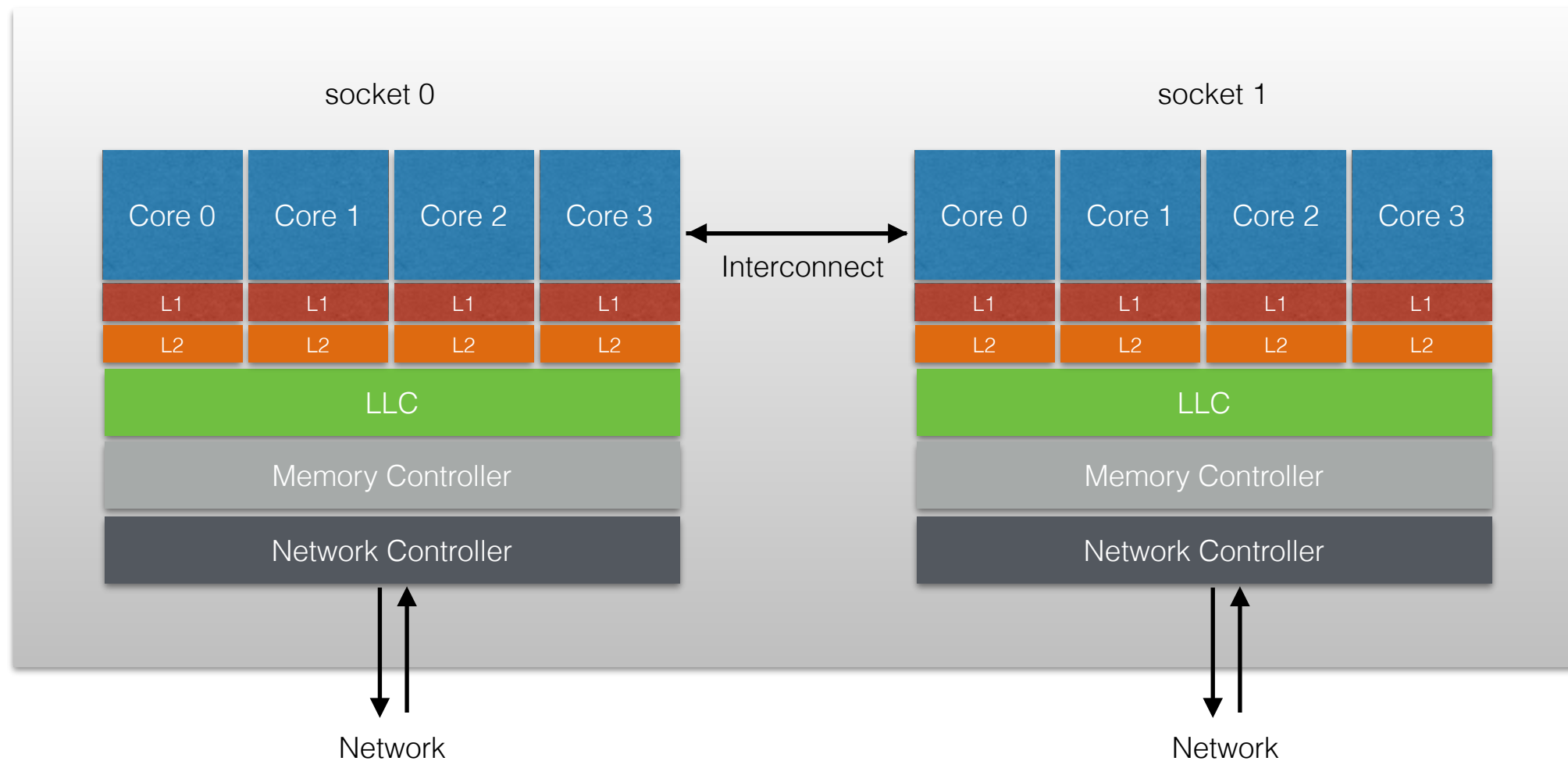
# Hardware Architecture

Presented by Alex Kosenkov

# Agenda

- CPU and NUMA
- Interconnects
- GPGPU

# CPU and NUMA



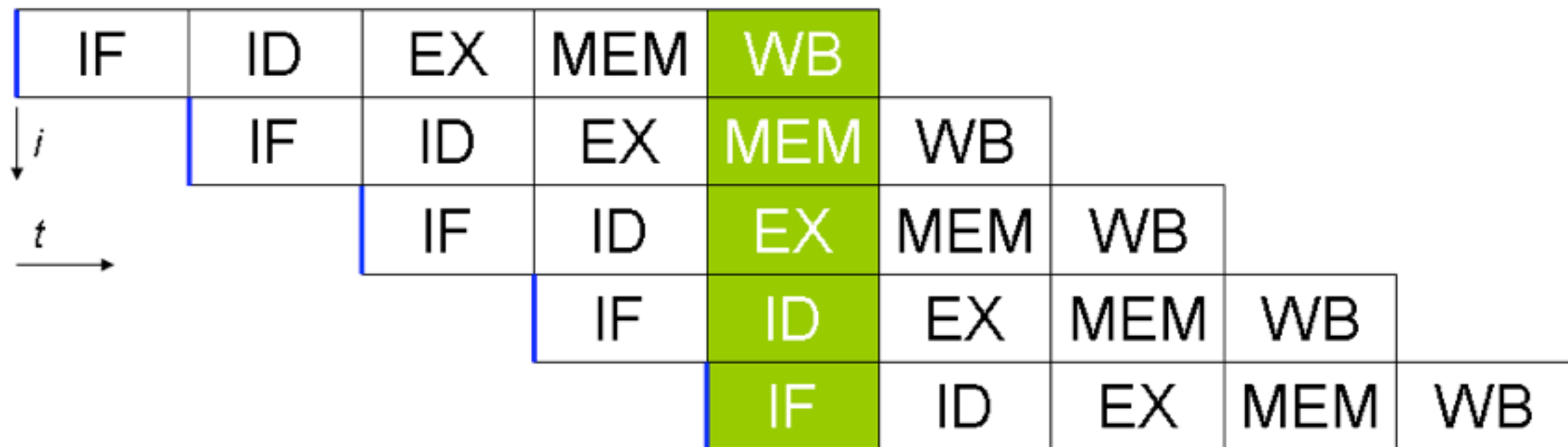
# CPU and NUMA: latency

- Processor registers: x1 (~1 cycle)
- L1 cache: x4 (64 KB)
- L2 cache: x10 (256 KB)
- L3 cache: x40-300 (12 MB)
- Memory: x250-400 (25 GB/s)
- Infiniband: x250 (37 GB/s)

# CPU and NUMA: RISC vs CISC

- CISC : complex instruction set computer
- RISC : reduced instruction set computer

# CPU and NUMA: Pipeline



# CPU and NUMA: Cache

L1 (instructions, data, TLB), L2, L3.

Structure unit is cache line: 128 bytes (payload + tag + flags)

Key mechanisms:

- Fetch & Prefetch (associativity types)
- Evictions (LRU)
- Write policy (write-through / write-back)
- Cache coherence (i.e. snooping)

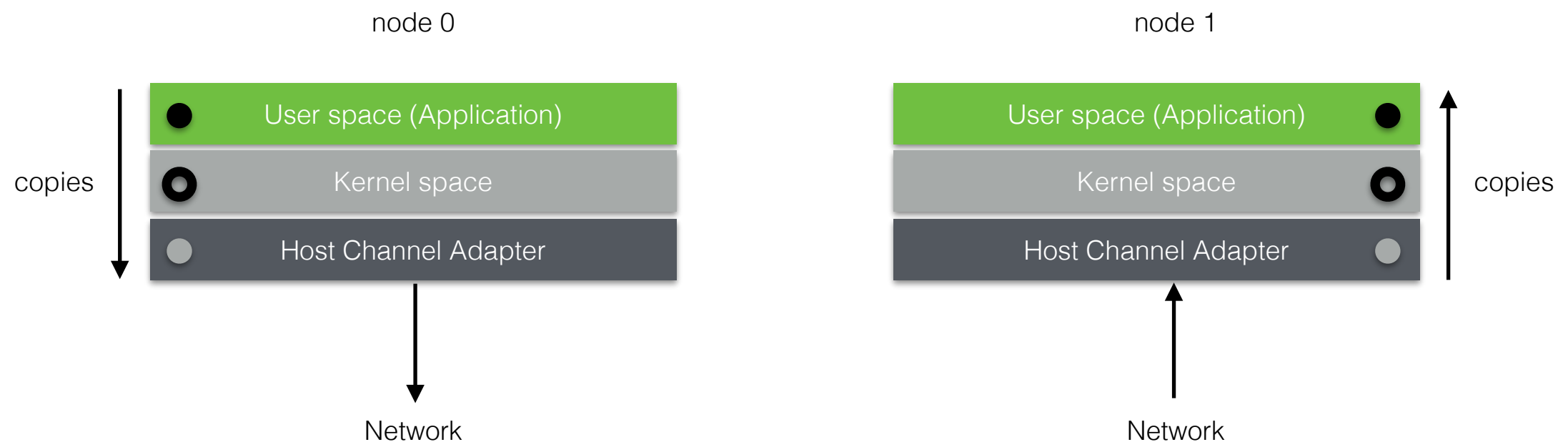
# Interconnects

- QPI / HyperTransport
- PCI Express
- Infiniband



# Interconnects: InfiniBand

## Without RDMA



# Interconnects: InfiniBand

With RDMA



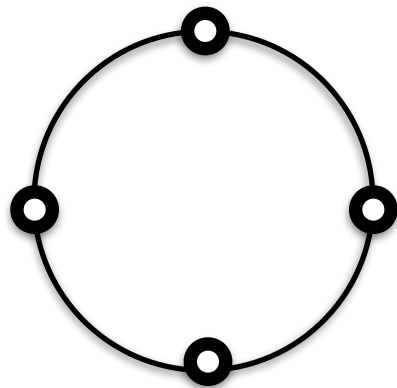
# Interconnects: InfiniBand

## Structure:

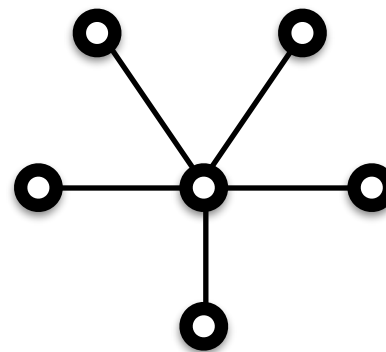
- HCA, Switch, SM, Router, Gateway (64 bit addressing, 4096 byte payload).
- Multicast-based (and point to point).
- Auto-tuneable (Subnet Manager builds routing table).
- Partitioning (QoS and Virtual Lanes).

# Interconnects: Topologies

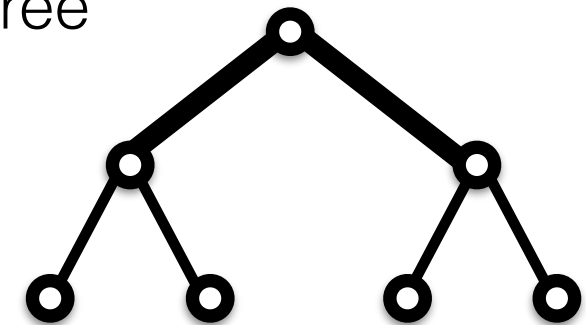
Ring



Star

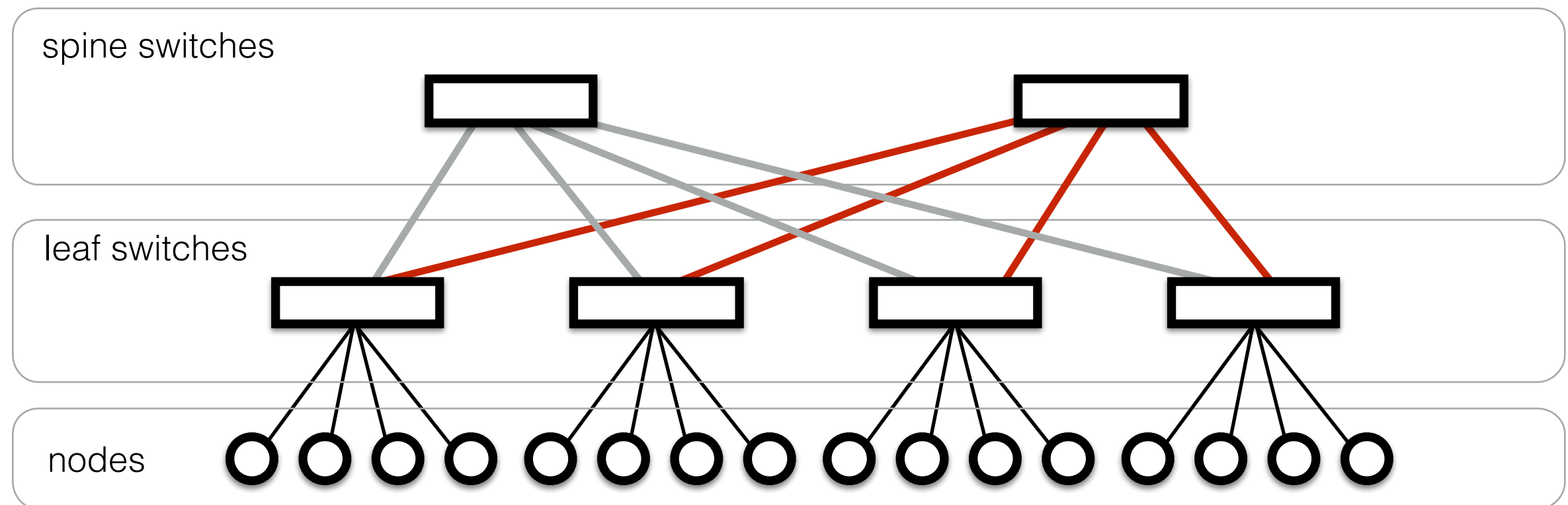


Fat tree



# Interconnects: Topologies

## Two-level fat-tree

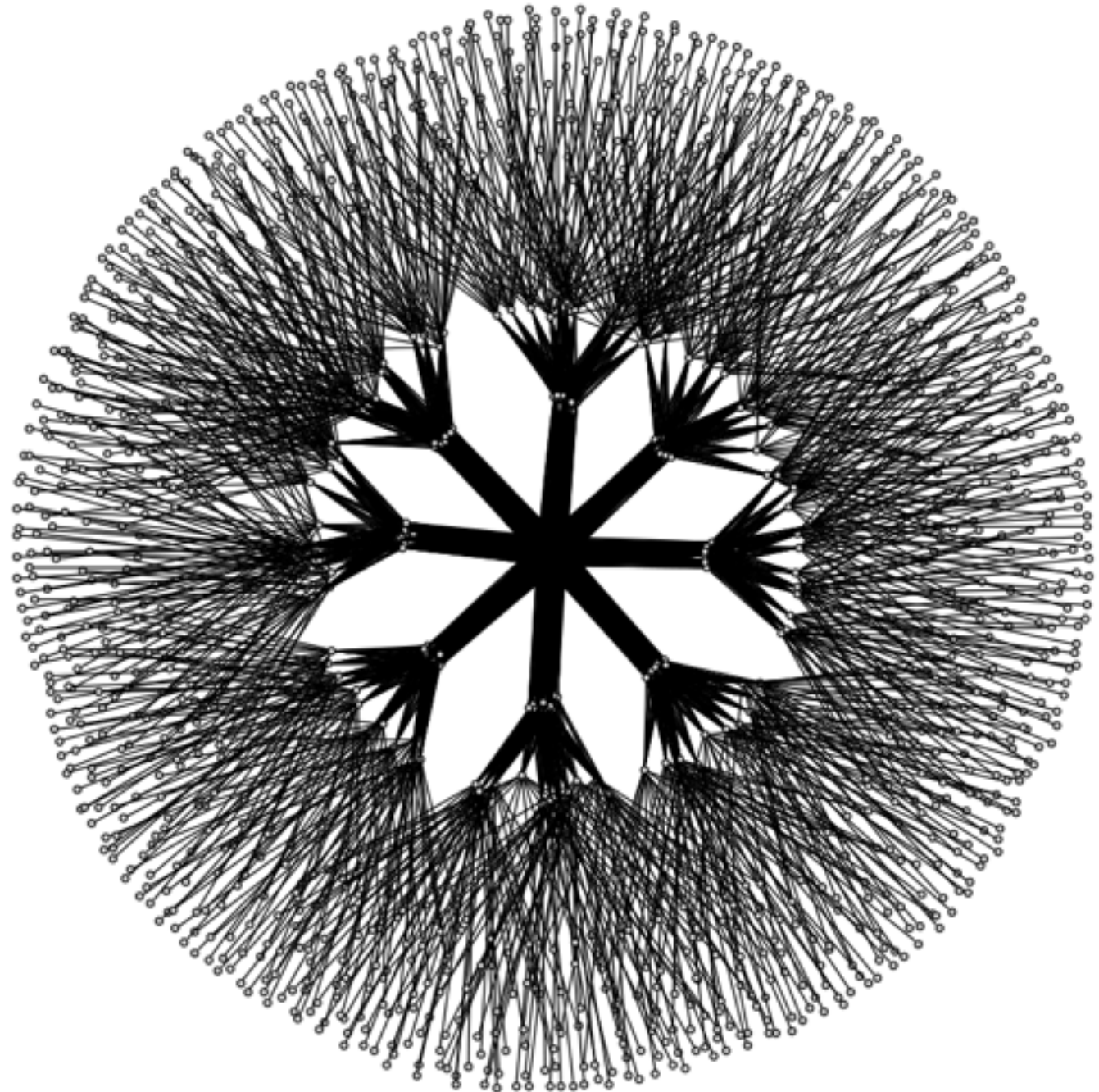


# Interconnects: Topologies

Atlas @ Lawrence Livermore National Laboratory

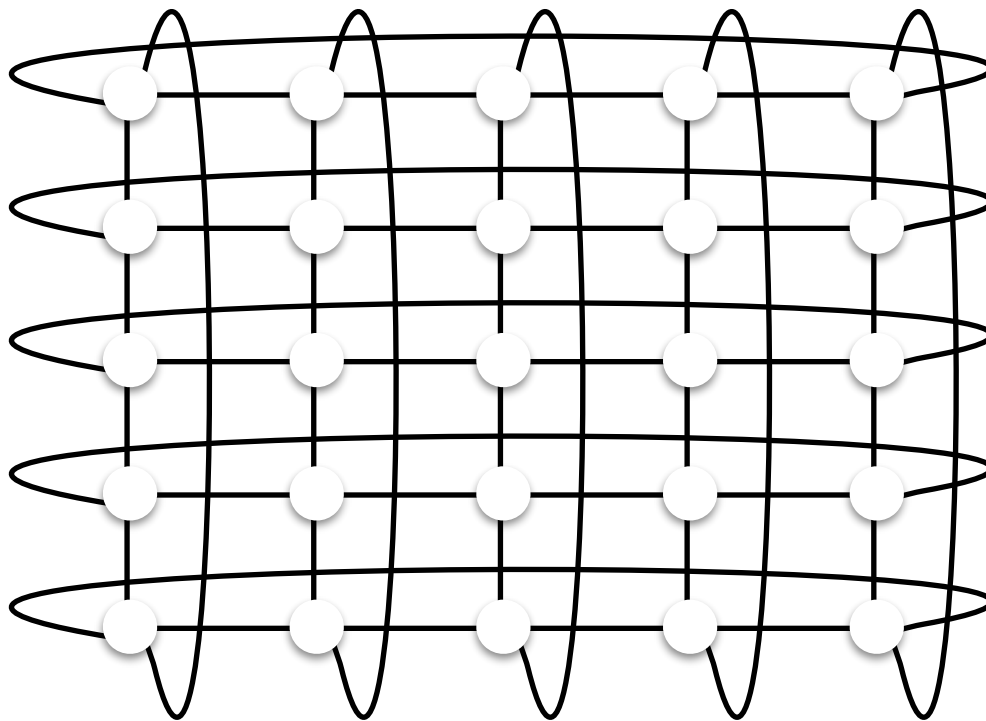
- 1142 nodes
- 192 24-port InfiniBand SDR crossbars
- full 3-stage folded Clos topology

From Torsten Hoefer's Network Topology:  
<http://htor.inf.ethz.ch/research/topologies/>

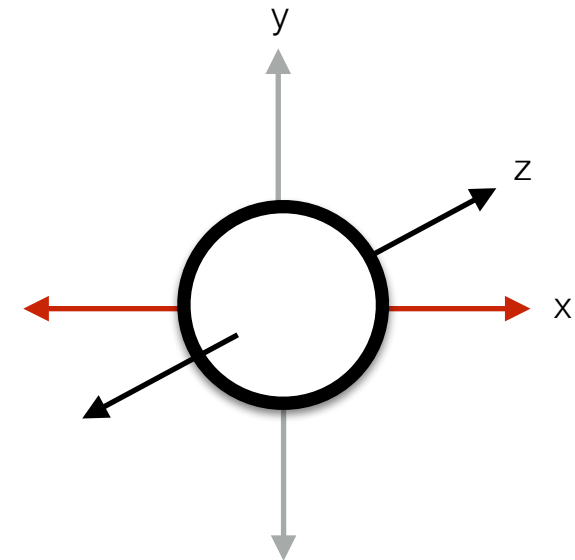


# Interconnects: Topologies

2D-Torus:



3D-Torus:



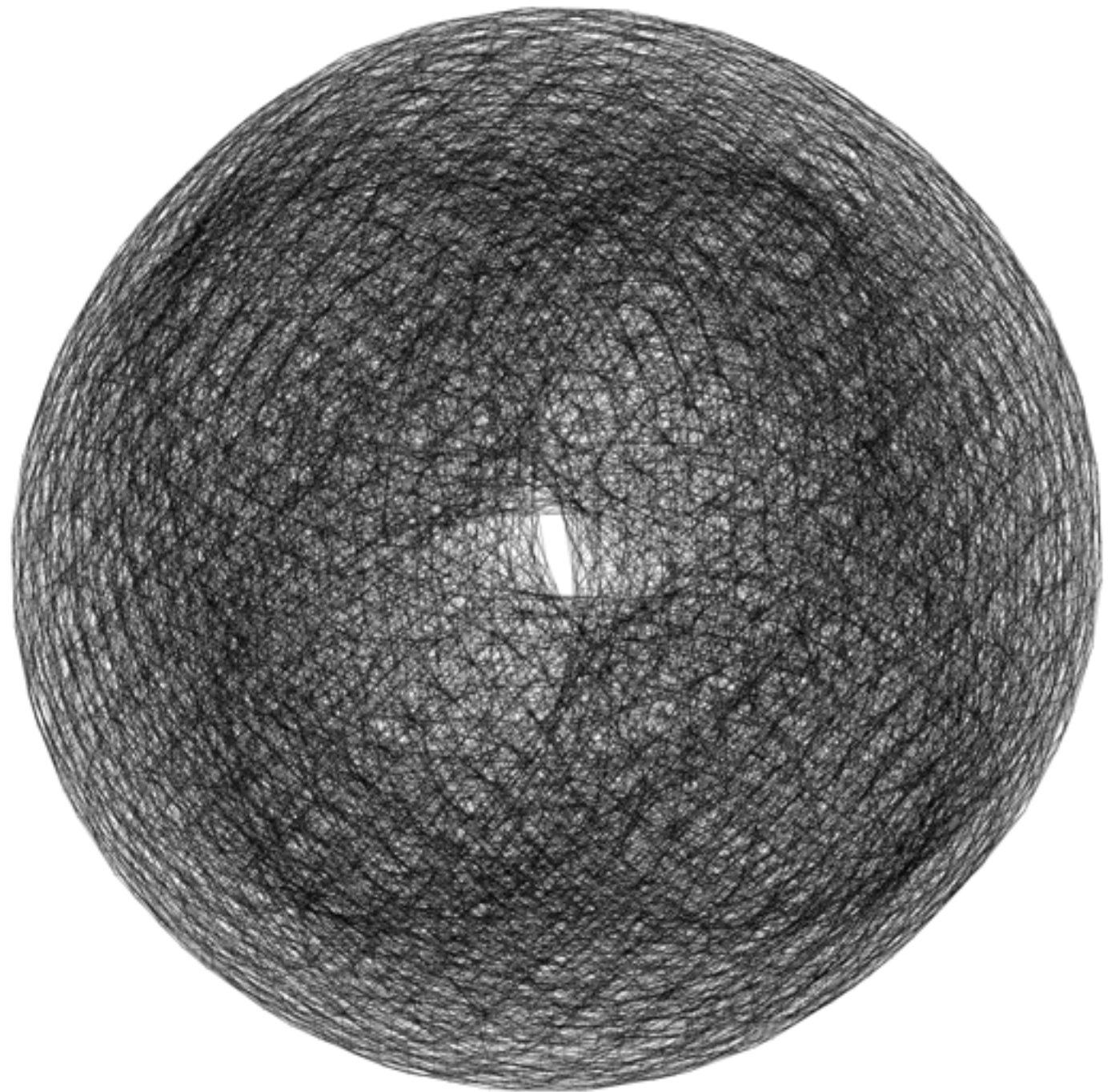


# Interconnects: Topologies

Jaguar XT-5 partition @ Oak Ridge National Lab

- 18851 nodes
- SeaStar 2 25x32x24 3D-Torus network

From Torsten Hoefler's Network Topology:  
<http://htor.inf.ethz.ch/research/topologies/>

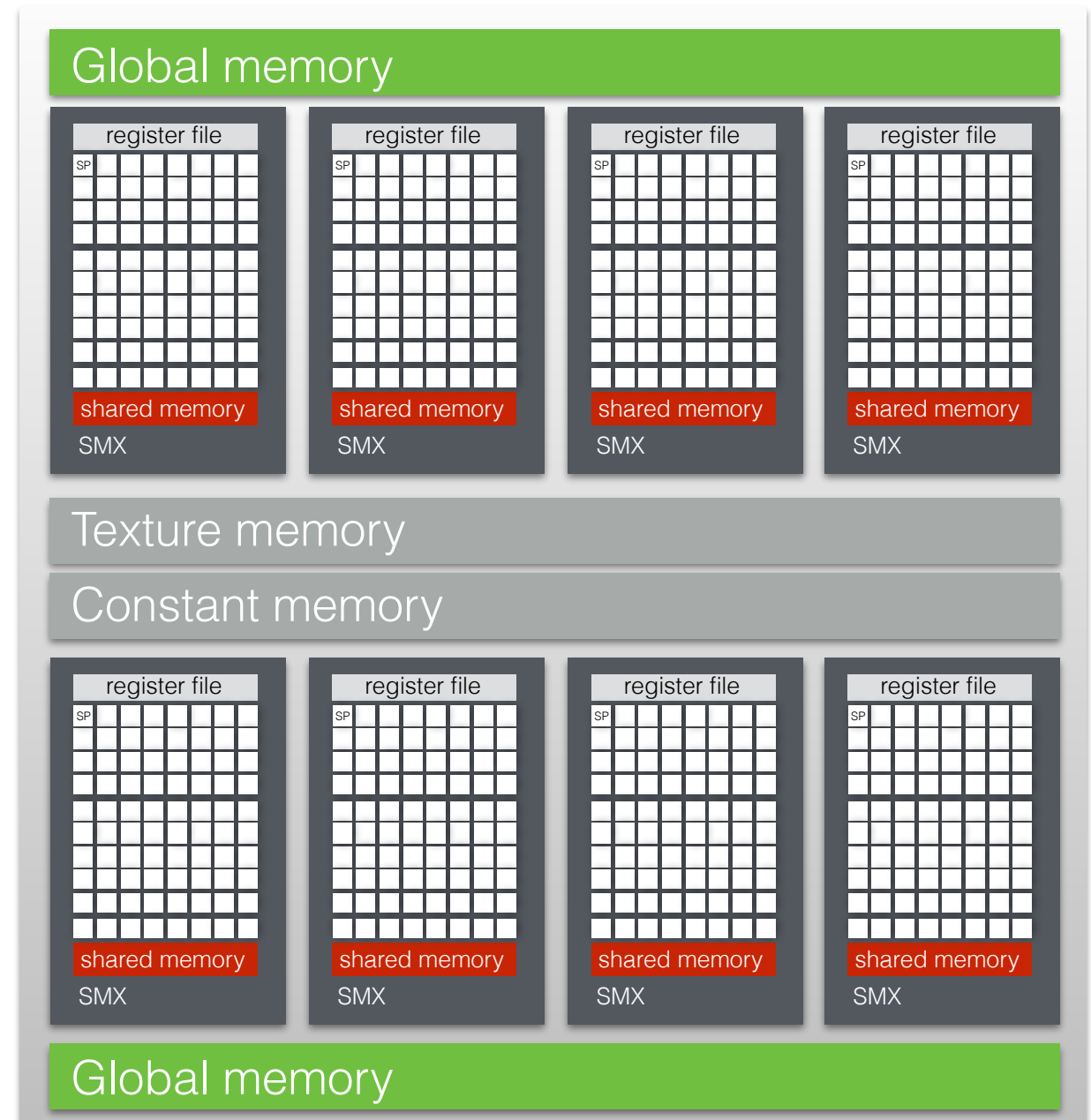




# GPGPU

Structure:

- Streaming multiprocessor (SMX)
- Streaming processor (SP)
- Registers
- Shared memory
- Texture and constant memory
- Global memory



# FPGA and ASIC

- FPGA: field-programmable gate array
- ASIC: application-specific integrated circuit

# Questions

End of the first part