New Corona System & CTS-2 Update

March 2019 LC User Meeting

Matt Leininger
CTS-2 POC

March 28, 2019

LLNL-PRES-770721

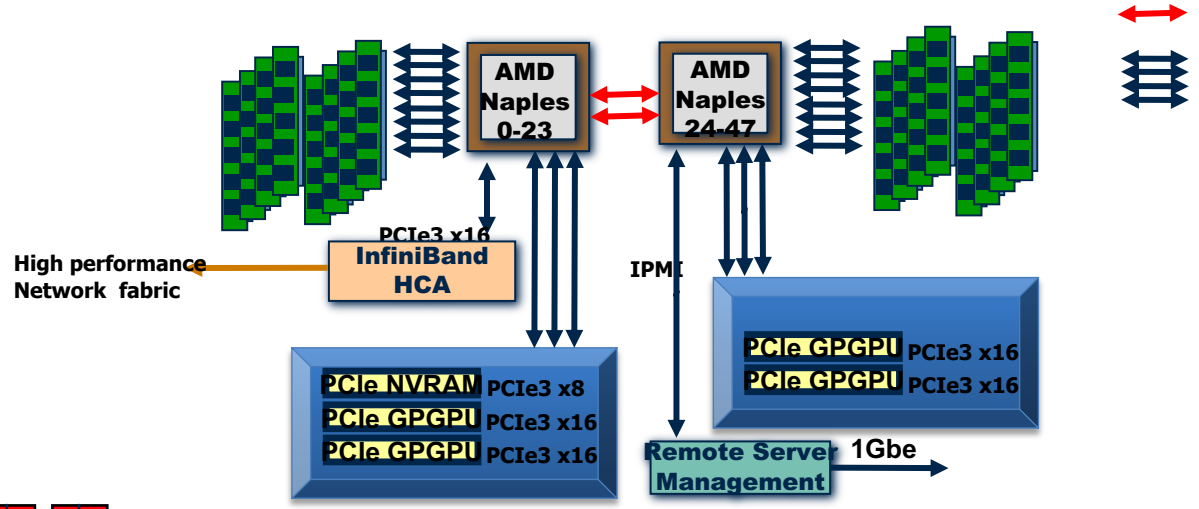
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

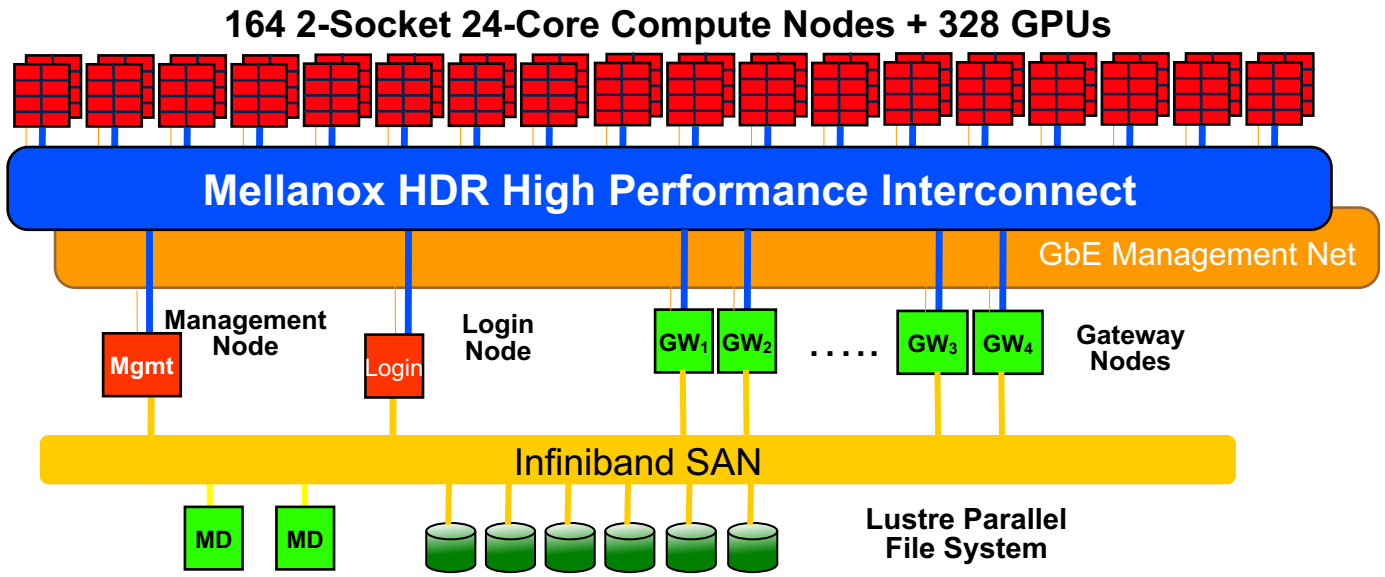
Lawrence Livermore National Laboratory

Corona is a Follow-on to Catalyst: First AMD GPU Cluster for HPC, ML, and Data Science

Node

- AMD Naples 24-core 2.0 GHz
- Memory: 256 GB; 5.3 GB/core
- Memory BW: > 300 GB/s DDR
- 1.6 TB NVMe
- Mellanox HDR100
- 4 GPU per compute node





System Nodes

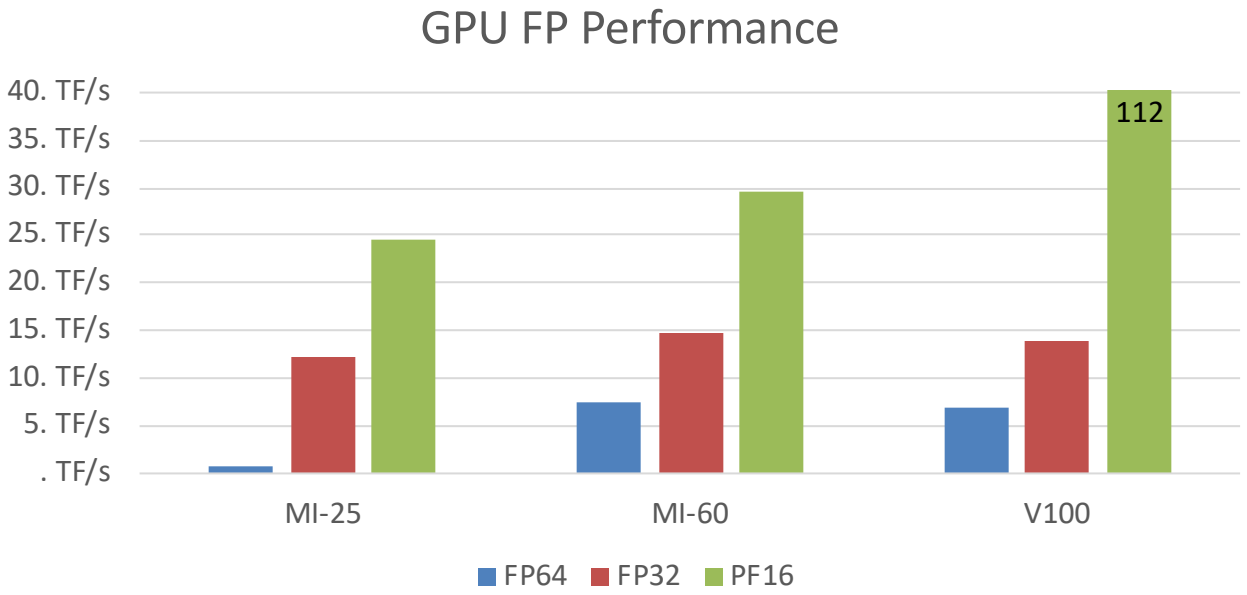
- 82 CPU-only nodes
- 82 CPU+GPU
- 4 Gateways
- 1 Login
- 1 Management

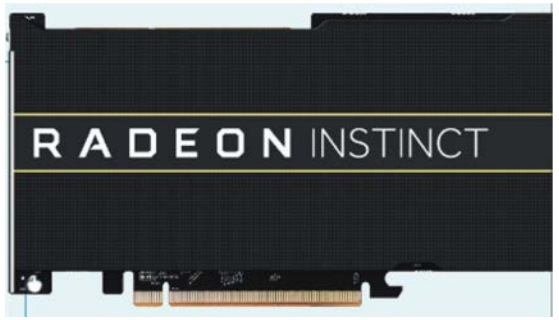
AMD

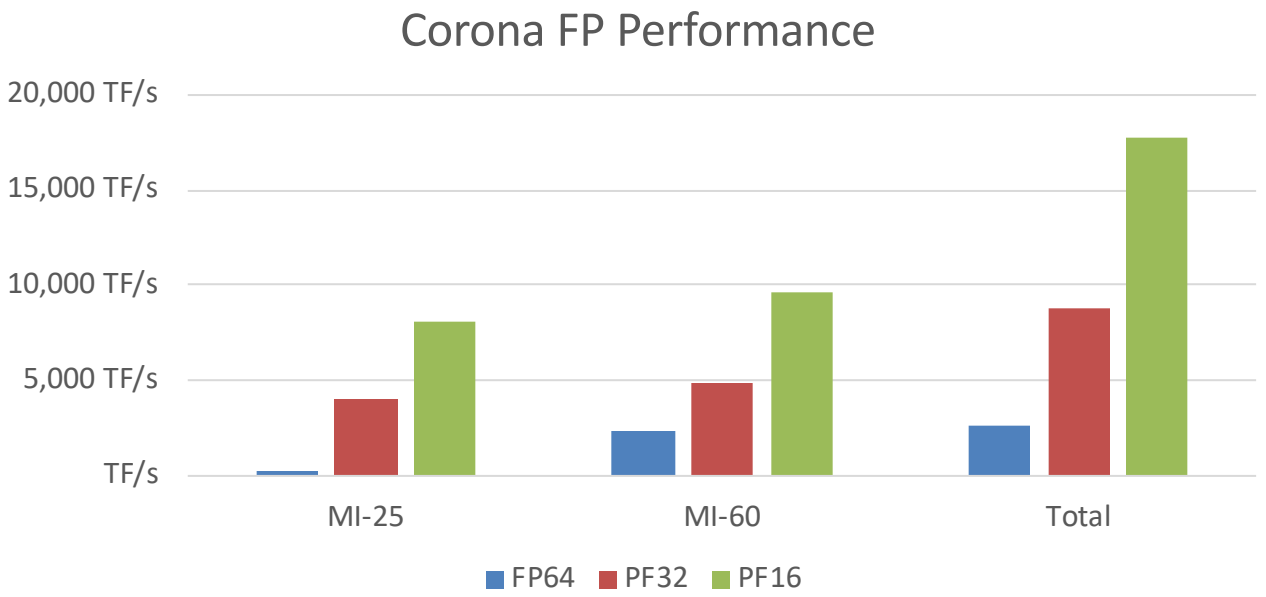
Mellanox TECHNOLOGIES

PENGUIN COMPUTING

Corona Highlights







Considering adding 328 AMD MI-60 GPUs to Corona

Corona NVMe

NVMe HGST 2N200 3 DWPD	Read	Write
Sequential @ 128 KiB	3.35 GB/s	2.1 GB/s
Random @ 4 KiB	835K IOPs	200K IOPs
Total	549 TB/s; 137M IOPs	344 TB/s; 32.8M IOPs

Corona Software Environment

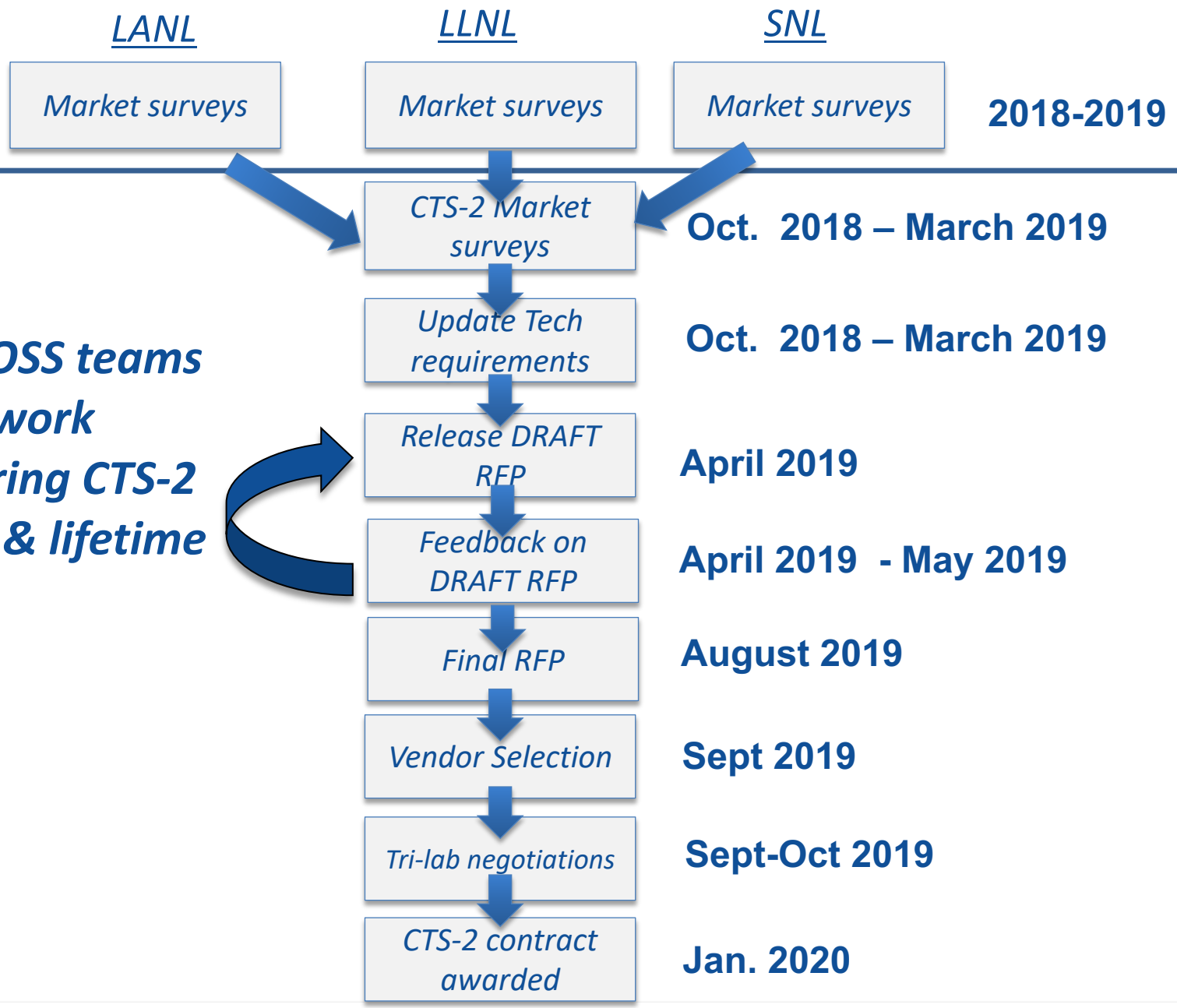
- Tri-Lab Operating System Software HPC environment as base foundation
  - TOSS 3.x based on RHEL 7.x
  - Provides smooth transition for TOSS team and LLNL HPC users
  - Includes AMD drivers, compilers, etc.
  - Slurm + Flux scheduler and resource manager
- Additional software for Data science & Machine Learning
  - Containers supported
  - Working with early users to explore other software

Corona is onsite and undergoing burn-in. Early User access in April.

Commodity Technology Systems

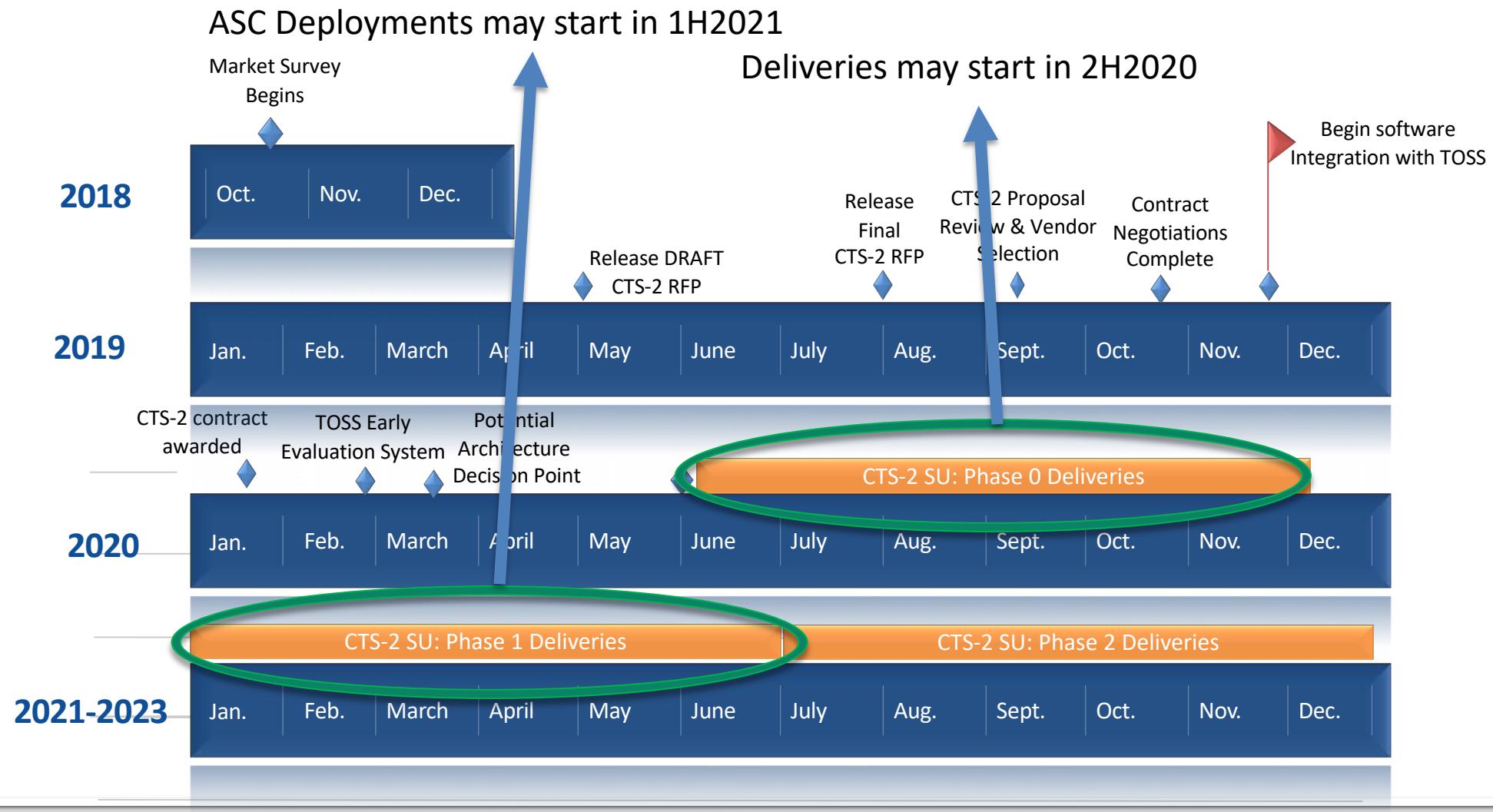
- Status of CTS-2 procurement
- Approximate Timeline
- Potential Architectures

CTS-2 activities leading to RFP and Contract

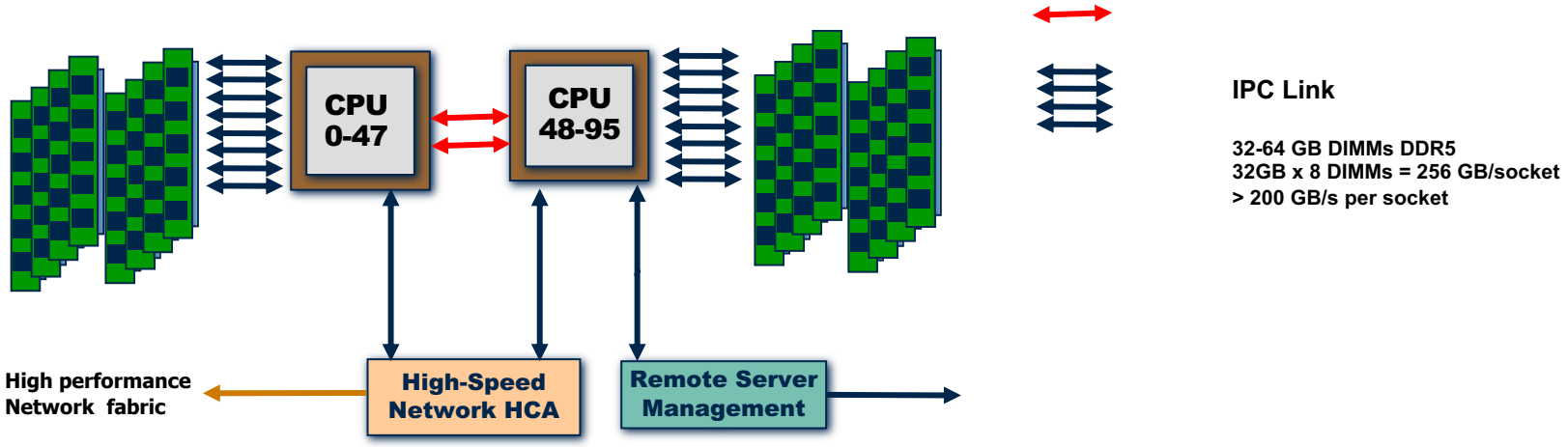


CTS-2 and TOSS teams continue to work together during CTS-2 deployment & lifetime support

DRAFT
CTS-2 Procurement Timeline



Potential CTS-2 Node Design



CPU Architecture & Software Readiness are key aspect of CTS-2 Selection

- Intel Xeon, AMD Epyc, Marvell ThunderX, IBM Power all viable processors
- Maturity of platform?
- TOSS support
- Maturity of system software and overall software ecosystem?
- Cost/performance of platform?

What about GPU systems and HBM memory?

Bringing ATS features to CTS-2

- GPU are becoming more widely adopted
- Past commodity procurements were dominated by CPU-only SU's
- GPU system will be available under CTS-2
  - Programs responsible for determining the mix of CPU-only + GPU nodes/clusters best address workloads
  - How much GPU memory do you need?
  - What is the ratio of CPU's to GPU's?
  - Is hardware support for unified memory required?
  - Can all codes utilize GPU's?
  - Can all workloads utilize GPU's – 3D vs 2D?

Bringing ATS features to CTS-2

- Give me the fast GPU memory but on CPU's!!!
- Today's GPU utilize High Bandwidth Memory (HBM v2 or HBM2)
- CPU + HBM may be a nice architecture for CTS
- Time to market is likely 2022+
- High Bandwidth Memory provides
  - ~3X more bandwidth per socket
  - ~4X less memory capacity per socket
  - 1-1.5 GB/core – adapt applications accordingly
- CTS-2 will include options for CPU+HBM if/when available

Questions?

- Matt Leininger
  matt@llnl.gov