Supercomputers at CCS: COMA and HA-PACS/TCA

COMA

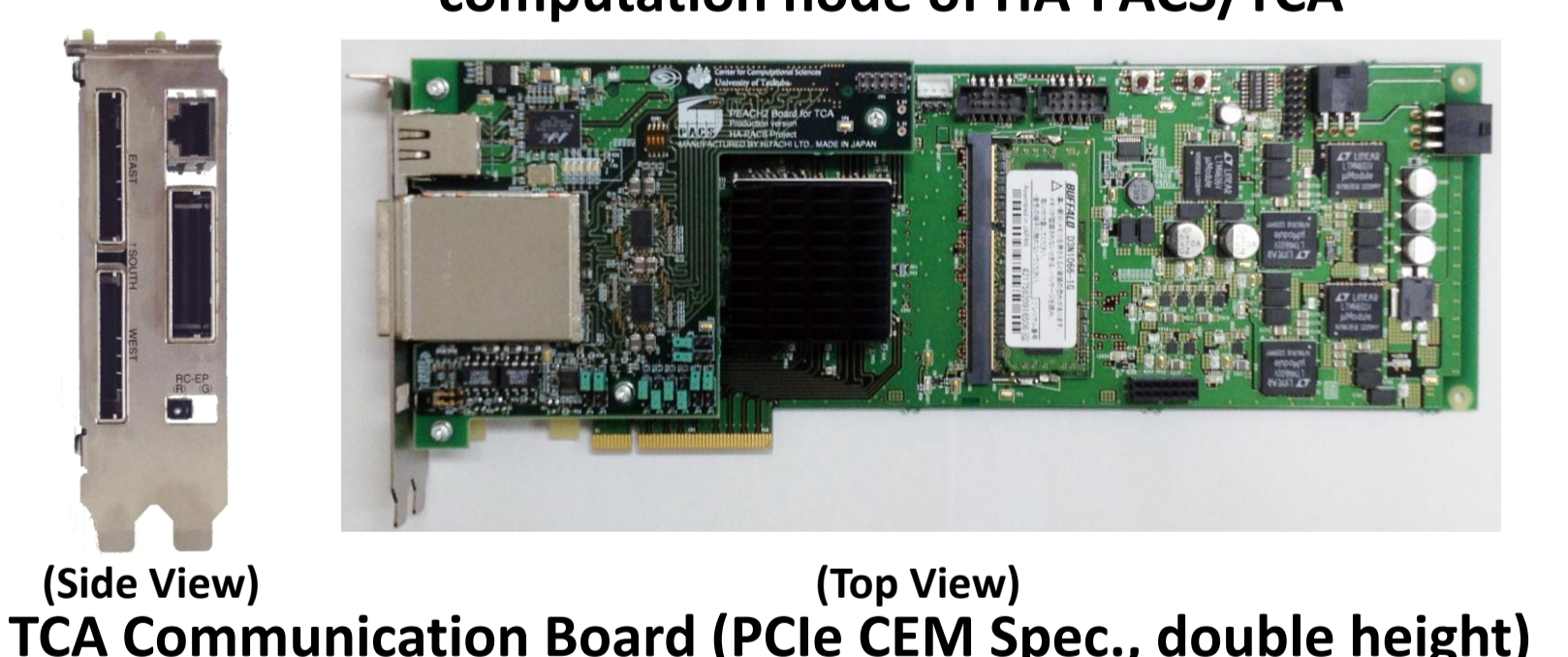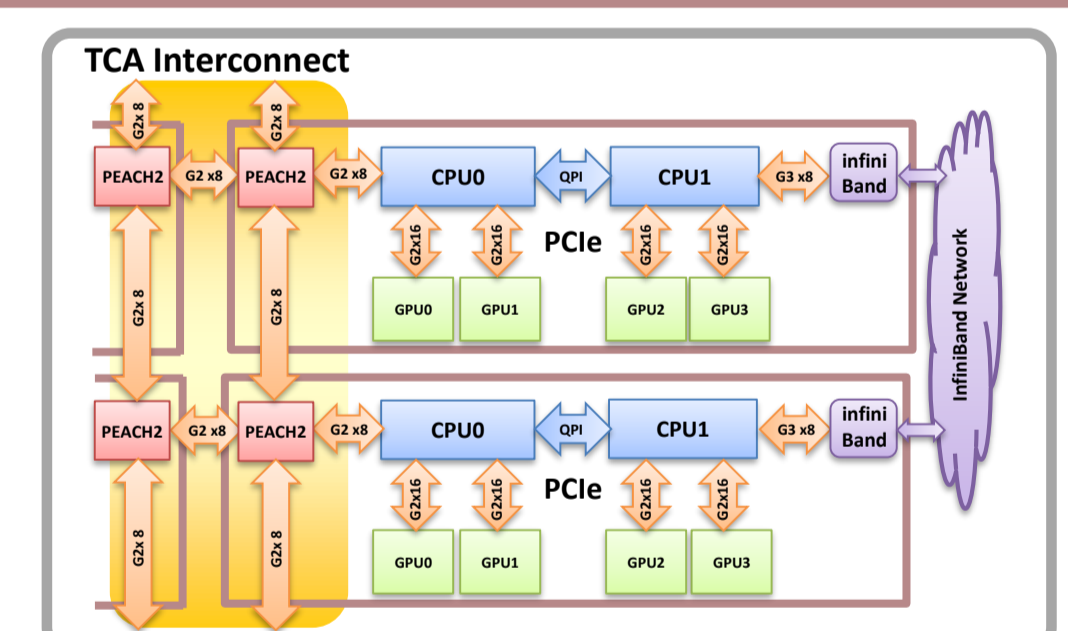| Item	Specification
Peak performance	1.001 PFLOPS (MIC: 843.9TF, CPU: 157.2TF)
# of nodes	393
File system	Lustre, 1.5 PB user area (DDN SFA12000)
Infiniband network switch	393 port FDR (Mellanox SX6536)
Total network bandwidth	2.75 TB/s
Language	Fortran90, C, C++
MPI	MVAPICH2, Intel MPI
System Management	Cray Advanced Cluster Engine, SLURM

Block diagram of computation node of COMA
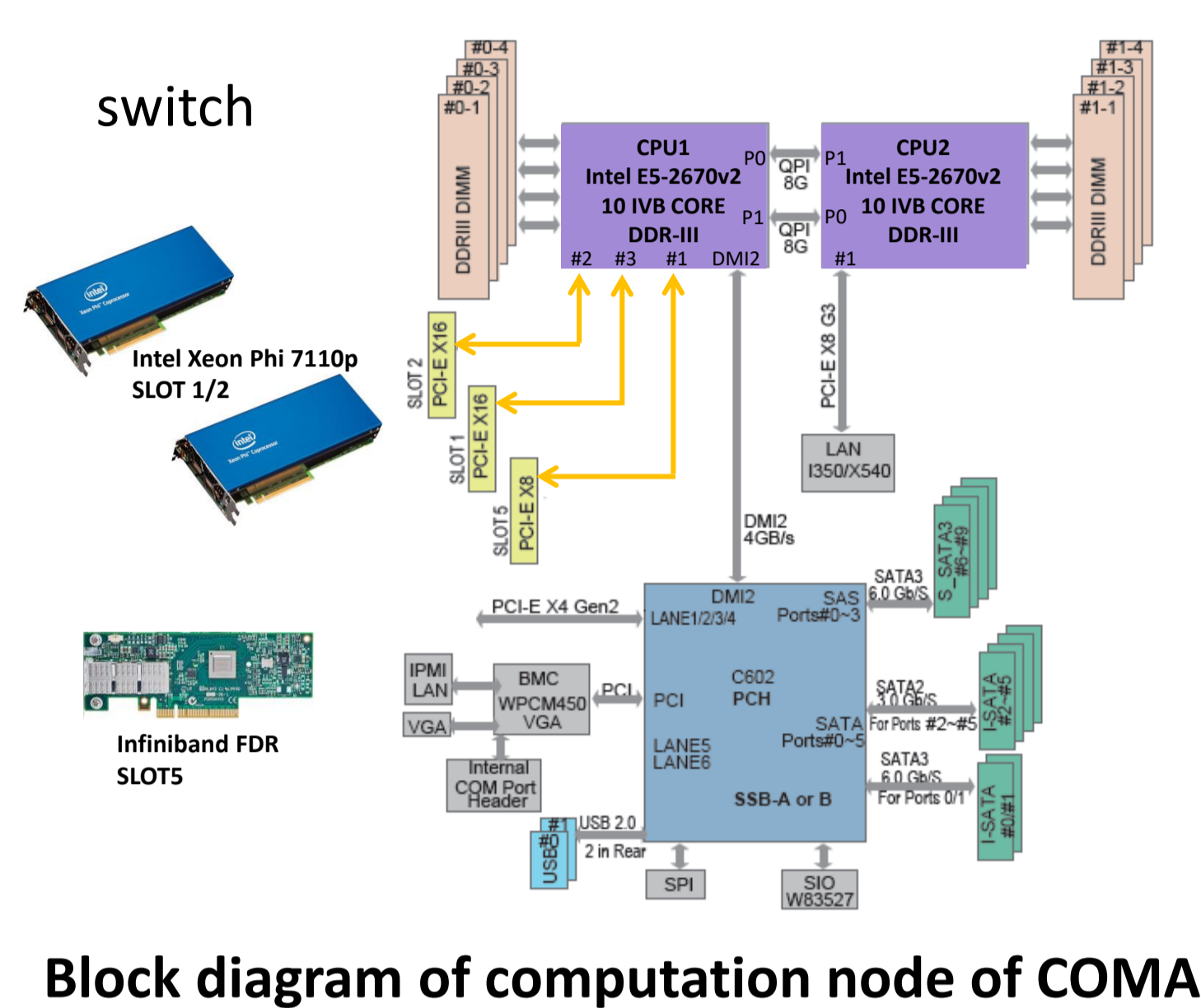
COMA (Cluster Of Manycore Architecture) is the 9th generation of PACS series supercomputer at CCS, of which computation node is equipped with two set of Intel Xeon Phi co-processor (Knights Corner) as well as two sockets of Intel Xeon CPU (Ivy Bridge). The interconnection network topology is a Fat-Tree with full bisection bandwidth supported by Mellanox InfiniBand/FDR. To avoid communication bottleneck between multiple PCIe devices over Intel QPI, all these devices are connected to single socket of Xeon CPU.

Either of "offloading", "native" and "symmetric" modes of Xeon Phi co-processor is available for users. Especially, the code development in the native mode is recommended to users to shift them to our new platform, Oakforest-PACS under management of JCAHPC as the largest Xeon Phi (Knights Landing) cluster in Japan with 25 PFLOPS of peak performance. Several computational science codes developed and tuned on COMA are ported to Oakforest-PACS with several times faster performance on each node.

HA-PACS/TCA: testbed for Tightly Coupled Accelerators concept

HA-PACS/TCA Specification

Computation node	CRAY 3623G4-SM
Motherboard	SuperMicro X9DRG-QF
CPU	Intel Xeon E5 2680 v2 (Ivy Bridge 2.8 GHz, 10core) x 2 socket
Memory	DDR3-1866MHz 4ch. 128 GB (119.4 GB/s)
Peak Performance	448 GFLOPS/node
GPU	NVIDIA Tesla K20X x 4 GPU
Memory	GDDR5 2600MHz, 6 GB/GPU (250 GB/s/GPU)
Peak Performance	5.24 TFLOPS/node
Interconnect	IB QDR x 2 rails (Mellanox Connect X-3)
TCA Interconnect	PEACH2 (FPGA: Altera Stratix IV 530GX)
# of Nodes	64
Peak Performance	364 TFLOPS (CPU: 28.7 TF, GPU: 335.3 TF)
LINPACK Benchmark	277 TFLOPS (Efficiency: 76%) 3.52 GFLOPS/W (3rd Nov. 2013 Green 500)

Block diagram of computation node of HA-PACS/TCA

(Side View) **(Top View)**
TCA Communication Board (PCIe CEM Spec., double height)

HA-PACS (Highly Accelerated Parallel Advanced system for Computational Sciences) is the 8th generation of PACS series supercomputer at CCS. It consists of Base Cluster and TCA parts. TCA (Tightly Coupled Accelerators) is a concept of advanced accelerated computing by coupling of GPU computing and low latency communication based on FPGA. As a proof-of-concept system of TCA, we implemented PEACH2 (PCI Express Adaptive Communication Hub ver.2) with Altera Stratix-IV FPGA. All the 64 nodes of HA-PACS/TCA part consists of two Intel Xeon CPU (Ivy Bridge), four NVIDIA GPU (K20X) and a PEACH2 board. As well as the testbed for TCA concept, HA-PACS/TCA is ranked at #3 in Green500 list of November 2013 with high performance/power ratio.

We have developed programming APIs for CUDA programmer and evaluated several benchmarks and applications, such as QUDA (LatticeQCDlibrary), Himeno benchmark (3D Poisson Solver), etc. Recently, we implemented GASNet for GPU, which is under development by Lawrence Berkeley National Laboratory, on PEACH2 to exploit high performance GPU-to-GPU communication especially for halo communication on 3D stencil computation.