

HA-PACS/TCA: Tightly Coupled Accelerators for Low-Latency Communication

Overview of Tightly Coupled Accelerators (TCA) Architecture

GPGPU is now widely used for accelerating scientific and engineering computing to improve performance significantly with less power consumption.

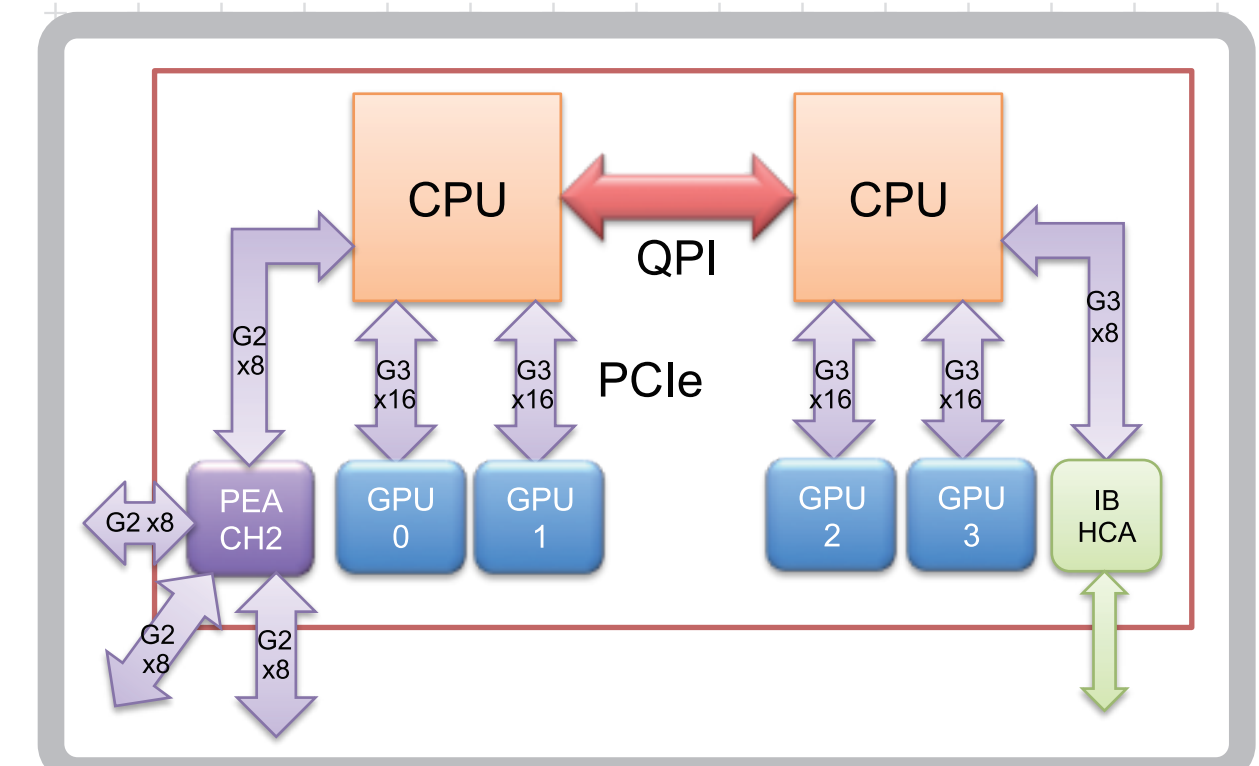
However, I/O bandwidth bottleneck causes serious performance degradation on GPGPU computing. To solve this problem, **TCA (Tightly Coupled Accelerators)** enables direct communication among multiple GPUs over computation nodes.

PEARL (PCI Express Adaptive and Reliable Link) is a concept for direct communication by using PCI Express technology. **PEACH2 (PCI Express Adaptive Communication Hub ver. 2)** chip and board have been developing. PEACH2 is implemented by FPGA (Field Programmable Gate Array) for flexible control and prototyping.

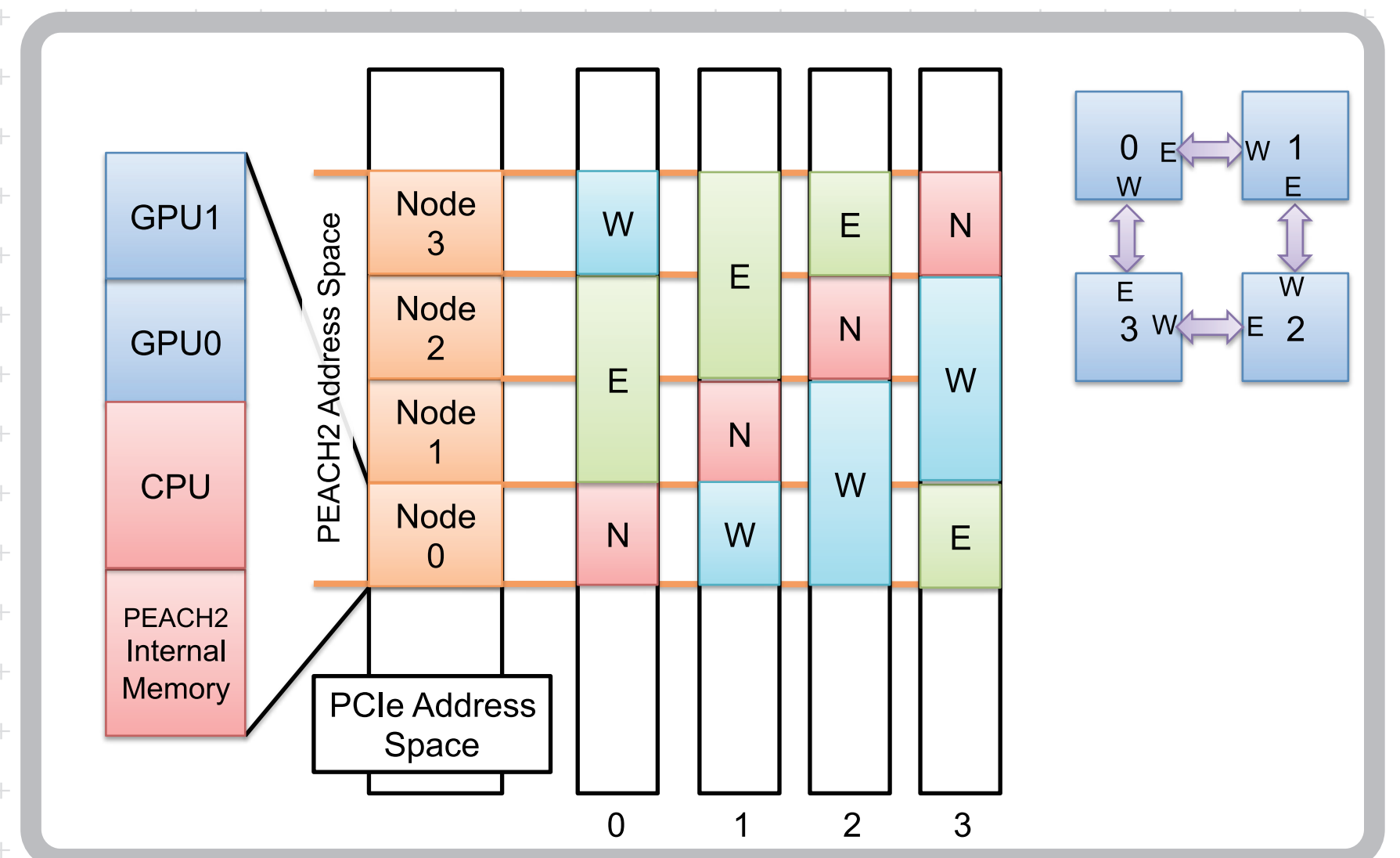
TCA provides the following benefits:

- Direct I/O among GPU memory over nodes
 - Reduce the overhead
- Shared PCI Express address space among multiple nodes
 - Ease to program

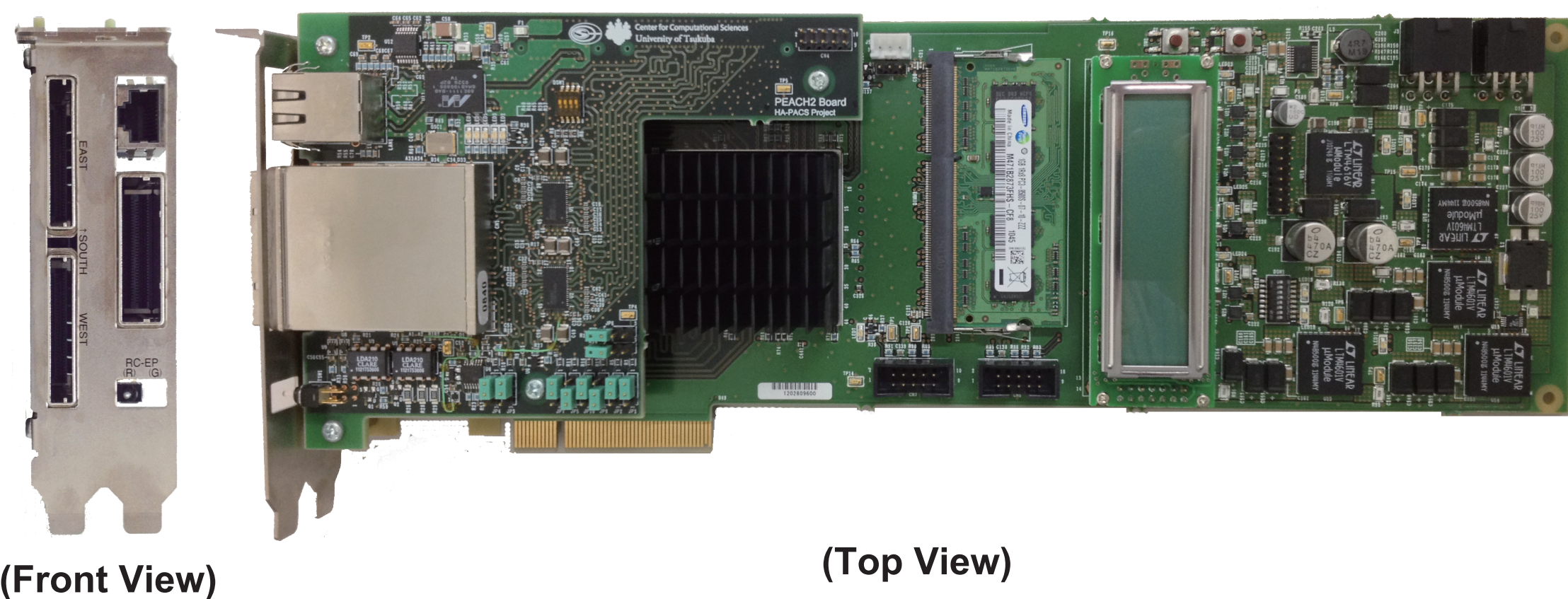
HA-PACS/TCA, which is an extended part of HA-PACS base cluster, will be constructed on Oct. 2013 using PEACH2 board in each node. HA-PACS/TCA will be managed with HA-PACS base cluster, and HA-PACS will be a Peta-scale class HPC cluster totally.



Block diagram of computation node of HA-PACS/TCA



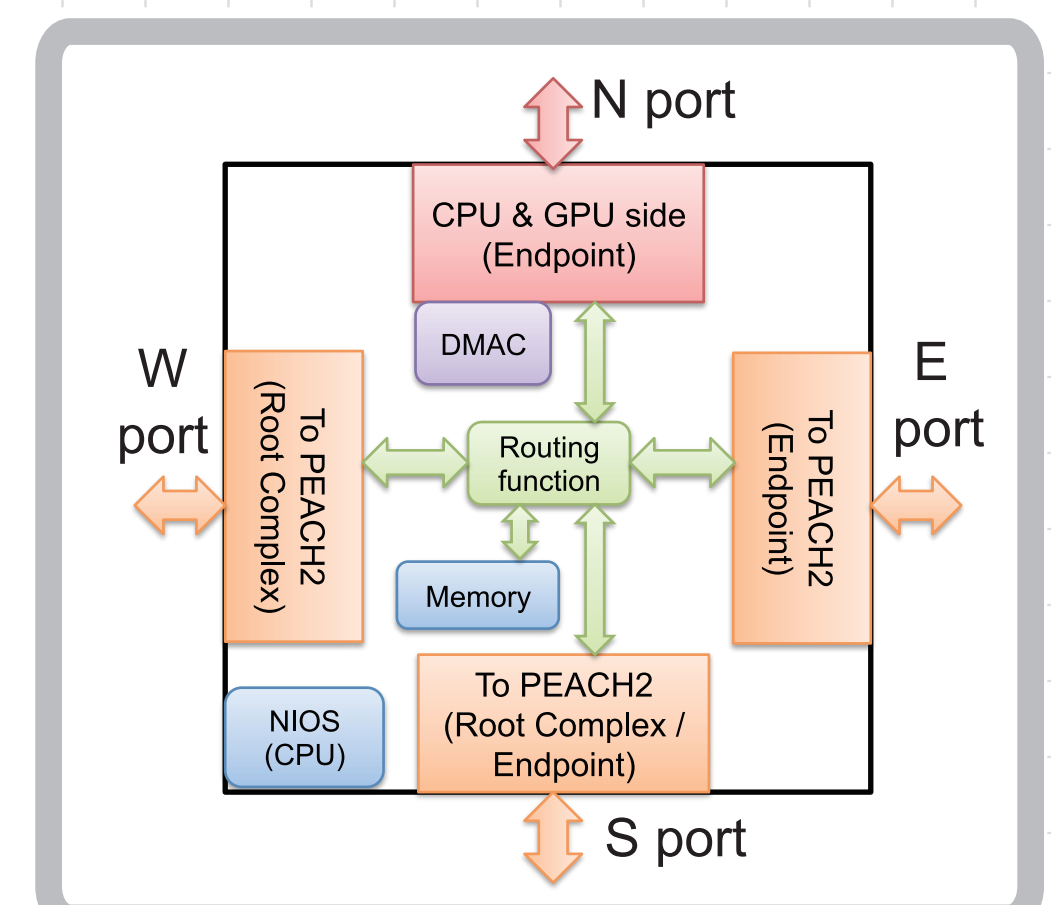
PEACH2 Address Space and Routing Mechanism



(Front View)

(Top View)

TCA Communication Board (PCIe CEM Spec., double height)

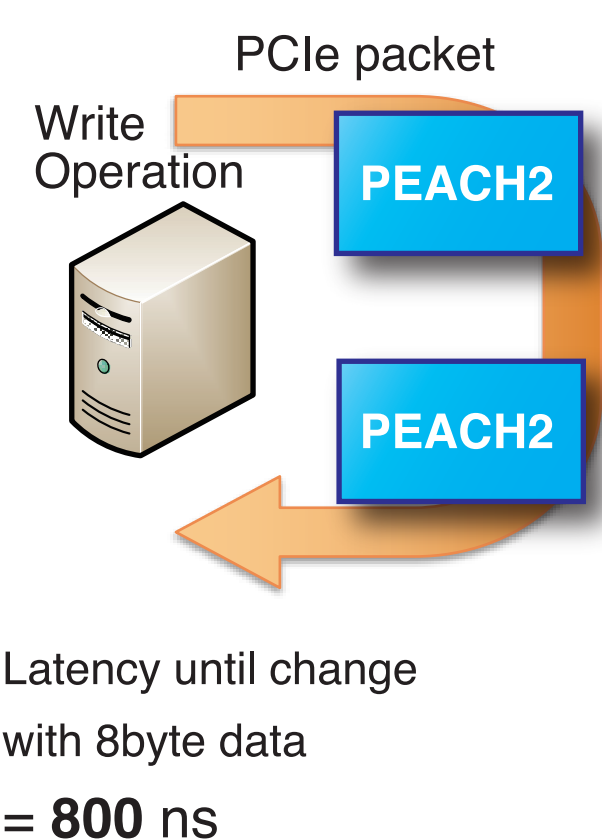


Block diagram of PEACH2 Chip

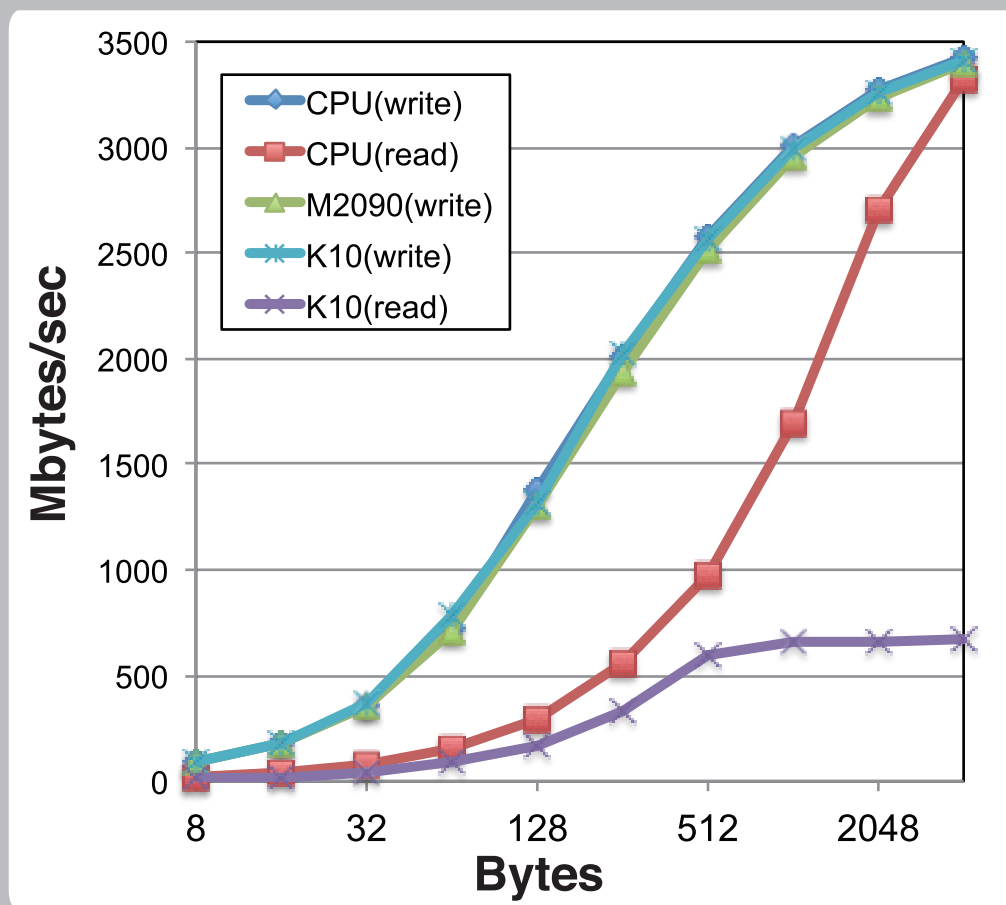
Preliminary Evaluation of TCA Communication

Evaluation Environment

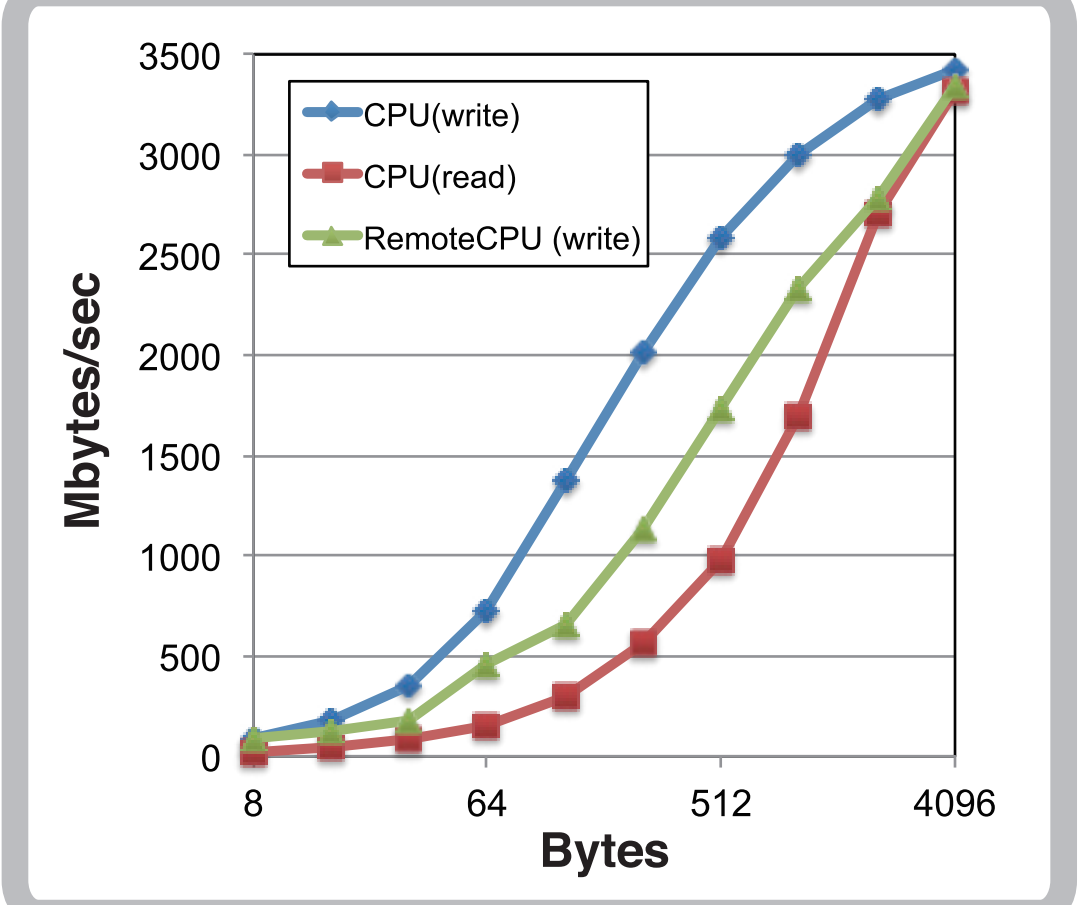
CPU	Intel Xeon E5-2670 2.6GHz (SandyBridge EP) 8 core x 2 socket
Memory	DDR3-1600 4ch. 128Gbytes
OS	CentOS 6.3
Kernel	2.6.32-279.9.1.el6.x86_64
GPU	NVIDIA M2090, K10
Driver	NVIDIA-Linux-x86_64-304.51
CUDA	5.0 preview
FPGA	Altera Stratix IV 530GX (EP4SGX530NF45C2)



Latency of TCA (driver level)



DMA Bandwidth between PEACH2 and CPU or GPU (255 times burst)



DMA Bandwidth between PEACH2 and Remote CPU (255 times burst)