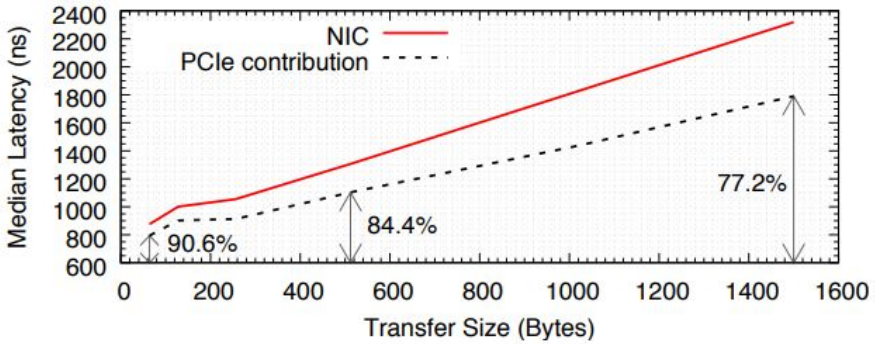
CCI-Bench:

Measuring Cache Coherent Interconnection

Jiacheng Ma & Gefei Zuo
EECS 589

Background

- ➔ Why CCI measurement is important?
  - ◆ "Understanding PCIe performance for end host networking", SIGCOMM 2018
  - ◆ The features of interconnection may influence future NIC design
  - ◆ Currently, PCIe dominates the latency
    - the PCIe subsystem contributes between 77% of the latency for 1500B packets, increasing to more than 90% for small packet sizes.



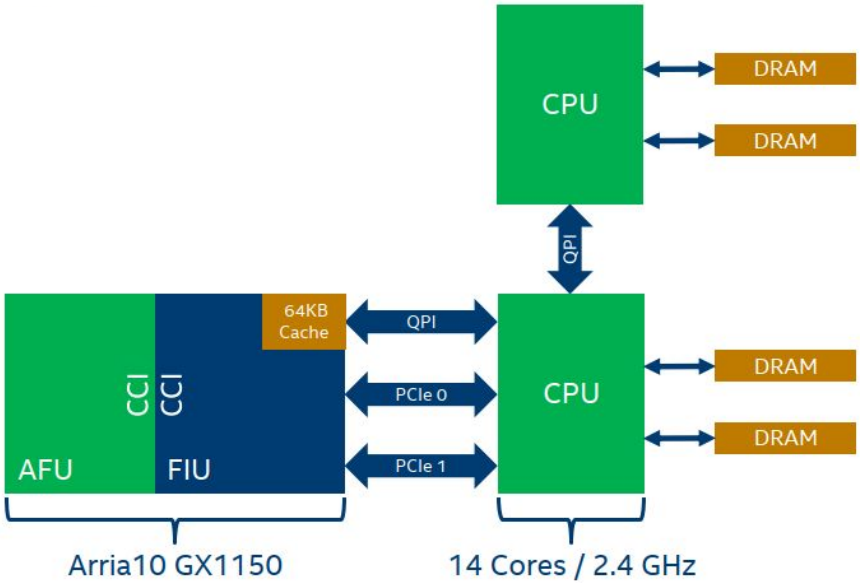
Background

- ➔ Cache Coherent Interconnection
  - Used to perform inter-CPU communication
  - Few available information
    - Low latency? (<100ns if cache hit, < 500 ns if cache miss ?)
    - Acceptable bandwidth? (<10 Gbs)
    - Hard to measure w/o FPGA: measure CCI between two CPUs?
  - Future NICs
    - NICs want ultra-low latency
    - Embrace the combination of PCIe and CCI interconnection

Background

- ➔ Target: CCI Protocols
  - ◆ QPI and UPI from Intel (the most widely used CCI protocol in the world)
  - ◆ Other CPU vendors have their own version of CCI
- ➔ Evaluation Platform
  - ◆ Intel Xeon+FPGA Platform
    - HARPv2 (with QPI, without IOMMU)
    - HARPv3 (with UPI, with IOMMU)
  - ◆ The only platform we can get close to CCI

Background



- ➔ Intel HARP
  - CPU and FPGA in the same chip
  - 1 UPI/QPI + 2 PCIe
- ➔ Use a packet based interface
  - Named CCI-P
  - RX/TX, similar to NIC
  - An unified abstraction of QPI/UPI/PCIe

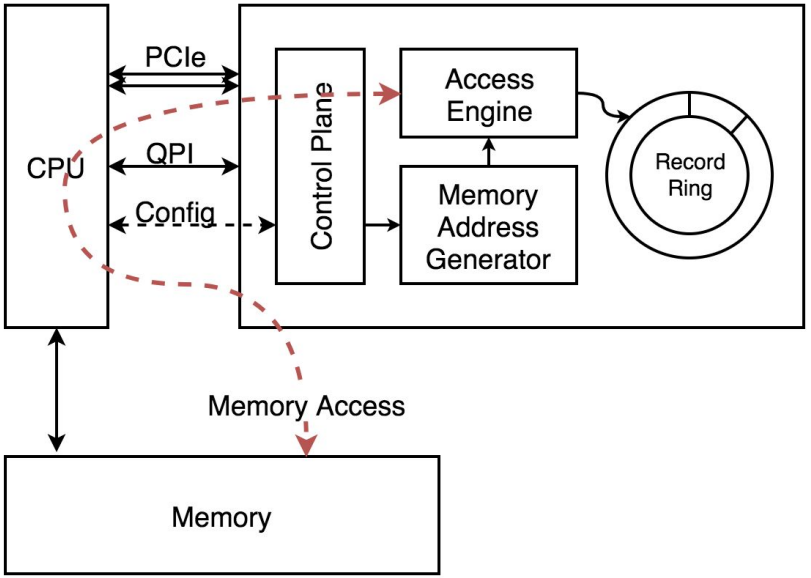
So how do we measure it?

A microbenchmark implemented in Verilog, running on FPGA, controlled by CPU.

- ➔ pcie-bench on NetFPGA can't be reused.
  - The reality of hardware programming we can't change.
- ➔ Metrics (for both PCIe and UPI/QPI)
  - Read / Write bandwidth / latency varied with data length
  - Cache Hints influence
    - Affect descriptor ring access and small packet receive
  - NUMA effect
  - IOMMU effect

Current Status

- ➔ A preliminary implementation of CCI-Bench
  - ◆ Architecture
    - Control Plane, Access Engine, Memory Address generator
  - ◆ Perform sequence or random memory access
  - ◆ Finish simulation
  - ◆ Can't run on hardware yet



Thanks.
Q&A