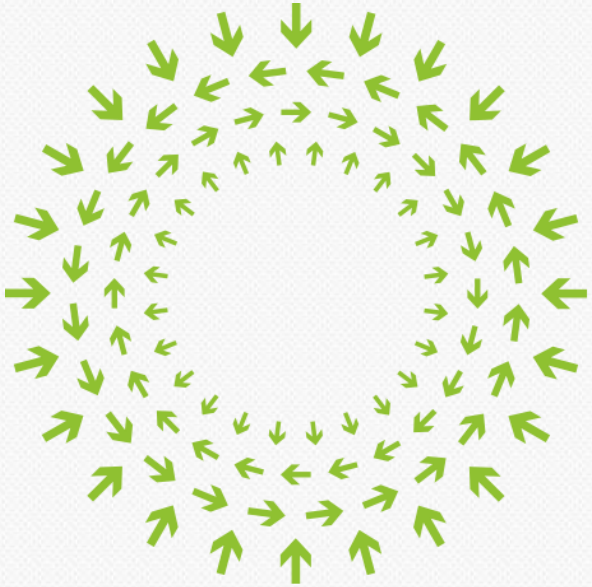


OPEN

Compute Engineering Workshop

March 9, 2015

San Jose

OPEN
Compute Project

OCP: High Performance Computing Project

Overview of Project/Status

Mar 2015

Devashish Paul – Co Lead, HPC Project Director
Marketing, Systems Solutions, IDT

Overview of OCP HPC Project

Mission

The HPC project has been established in the Open Compute Project to service the needs of the High Performance Computing, Supercomputing, Financial Trading and Low Latency Analytics segments

Service it with open hardware platforms delivering solutions in the market from system level down to silicon.

Deliver specs and market solutions

Unique to performance needs of

low latency multi processor computing

Overview of OCP HPC Project

Key Values

- Deliver Open Hardware platforms for HPC industry which often must do custom case by case deployments
- Serve as a central point of collaboration for HPC industry in terms of performance and cost optimization of HPC compute and networking platforms
- Group will act as an industry leader to help drive future innovation in HPC market in a collaborative open industry context
- Path for industry standard open interconnect silicon
- Path to Open processor architectures

Become the Industry “Node” for Open Collaboration in HPC

Project Charter Items

Fully open **heterogeneous computing** , networking and fabric platform

Optimized for multi-node **processor agnostic** any to any computing using x86, ARM, PowerPC, FPGA, ASICs, DSP, and GPU silicon on hardware platform

Enables rapid innovation in low latency high Performance Computing and Big Data analytics through open non-lock-in computing, interconnect, and software stack.

Energy efficient compute density

Distributed and central **storage** for large data manipulation (non spinning disk) with low latency

Operating System – Linux based operating systems and developer tools and **open APIs**

Path to **Open Silicon** and Open APIs, initially leveraging existing industry standards, later developing its own silicon

Re use developments from OCP Server group and Open Rack where appropriate

Leverage **industry standard interconnects**, no proprietary interconnects for main fabric and networking

Vendor Agnostic Low Latency Multi Processor Computing

Relationship to Other OCP Groups

Key Values

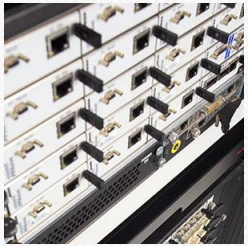
- Where possible use designs from Server/Storage for compute
- Comply with OCP Rack/Hardware/Electrical
- Improve networking designs for latency and scalability, investigate using alternate technologies already in production for near term goals
- Come up with new OCP Open interconnect fabric/clustering silicon spec focused on scalability, low latency and energy efficiency
- Open APIs etc
- Innovations from HPC group may flow back to other OCP groups where appropriate

Leverage existing OCP server-storage-networking etc, drive silicon industry too

Scope

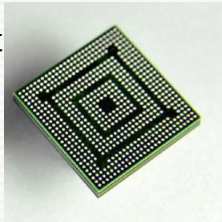
In Scope Technology Categories

- Low latency top of rack switching
- Combined Switch and micro servers
- Combined compute and switching
- Low latency scalable storage
- Connectivity from HPC Fabrics/clustering technology to out of cluster networking via ethernet
- OCP mechanicals (19 and 21 inch)
- APIs and software interfaces
- Open Hardware compute and switching Development of OCP HPC Interconnect silicon spec (can build on existing mainstream technologies such as PCIe, RapidIO, Infiniband, Ethernet)
- Investigation into Open industry processor specs for HPC Market



Out of Scope Technology Categories

- Items already covered in server, storage, networking and other groups that are not optimized



Focus Market Segments

- High Performance Computing
- Supercomputing
- Data Center Low Latency Analytics
- Mobile Network Edge Computing
- Low Latency Financial Trading

Mainly focus where systems span multiple processors

Need Low Latency East West Traffic

HPC Project Market Segments: Trends/Customer Requests

High Performance Computing

- Processor agnostic heterogenous computing with accelerators
- no ASIC budget
- ARM based compute
- Off the shelf interoperable solutions

Supercomputing

- Energy Efficient Multi Processor Compute
- Fundamental physics and biology govt research
- Integrate SoC
- Scalable Fabric
- Petascale to Exascale
- Top 500 computing

Financial Trading

- real time decision making on trends
- FPGA and GPU based computing
- Reliable fault tolerant interconnect
- Mission critical
- Off the shelf solutions
- Firmware

Low Latency Analytics

- intelligent public cloud computing
- cloud based deep machine learning
- pattern recognition
- scalable compute at data center scale

Mobile Edge Computing

- local caching
- offload backhaul
- Layer 7 analysis in real time
- ultra low latency with Baseband Unit
- Interop with base stations

Low Latency East West Traffic with Energy Efficiency

OCP HPC Fabric Interconnect Silicon

- HPC needs huge scale of any to any processing nodes
- Latency is a primary concern for this market as well as those that need analytics
- Energy footprint is an issue
- Low hanging fruit is to eliminate latency and power from NIC's and other interconnect devices
- Need native protocol termination on processing endpoints, like processors, DSP, GPU, FPGA
- Diverse industry initiatives to create proprietary clustering fabrics at many startups and large processor vendors
- Prefer to start with some industry standard options that scale, have low latency, multi vendor collaboration etc
- Take best attributes of PCIe, Infiniband, RapidIO, Ethernet technologies to reach exascale computing

Interconnect is key path to low latency exascale systems

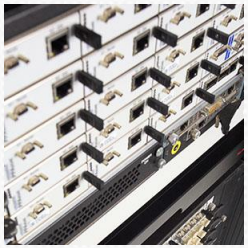
4 Phase Commercialization

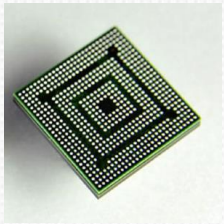
Phase 1: 6-12 months, board designs, leverage as much as possible, server group compute, look into low latency networking

Phase 2: 12-24 months HPC optimized heterogeneous computing-ARM, GPU, x86, DSP etc

Phase 3: 12-24 months: Deliver open silicon specs for interconnect as well as processor architectures for low latency optimized computing

Phase 4: Investigate Silicon Photonics





All phases commence now with delivery dates over time

Upcoming 2015 Potential HPC Submissions

Phase 1: board level/NIC/switching appliances

- X86 based low latency computing
- DSP + ARM computing
- FPGA based low latency Server
- Scalable GPU with low latency interconnect
- Top of Rack Switching
- Low latency NIC
- 4U server with low latency interconnect

Phase 3

- Open Interconnect
- Open Processor instruction set

Submissions

For low latency

Computing segments

Get involved

Mailing List

- http://lists.opencompute.org/mailman/listinfo

Conference Calls

- 1-2x per month
- Europe + NA East
- NA West + Asia/South Asia

Engineering Workshops

- Regional OCP Engineering workshops (host one and work with foundation to schedule)

www.opencompute.org/get-involved

Thanks

www.opencompute.org/get-involved