SLAP: Making a Case for the Low-Powered Cluster by leveraging Mobile Processors

Dukyun Nam, Jik-Soo Kim, Hoon Ryu, Gibeom Gu, and Chan Yeol Park
Supercomputing Center at the Korea Institute of Science and Technology Information (KISTI)

From Petascale to Exascale

- Power consumption in supercomputers is becoming a challenging problem
  => Cost and Space !
- Exascale computers that are expected delivered in 2019-2020 should fulfill 20MW of power consumption limit
- Conventional multi-core processors are *not* sufficient
  => GPU or Xeon Phi ?

Utilizing Low-Powered Processors

- An alternative approach can be leveraging *low-powered and lightweight mobile processors* to build supercomputing clusters
- Potentially minimize the overall system size along with cooling facilities
- European project called Mont-Blanc has been conducted to design a new type of system built from energy efficient solutions used in embedded and mobile devices.

Our Approach

- Empirical Study of building our own SLAP (*Scalable*, *Low-powered*, *Autonomous & Robust*, *Pluggable*) cluster based on energy efficient mobile processors
- Especially focusing on the *usability* and *reliability* of our prototype system
  - HPL Benchmark
  - A real scientific application
  - GlusterFS for storage cluster

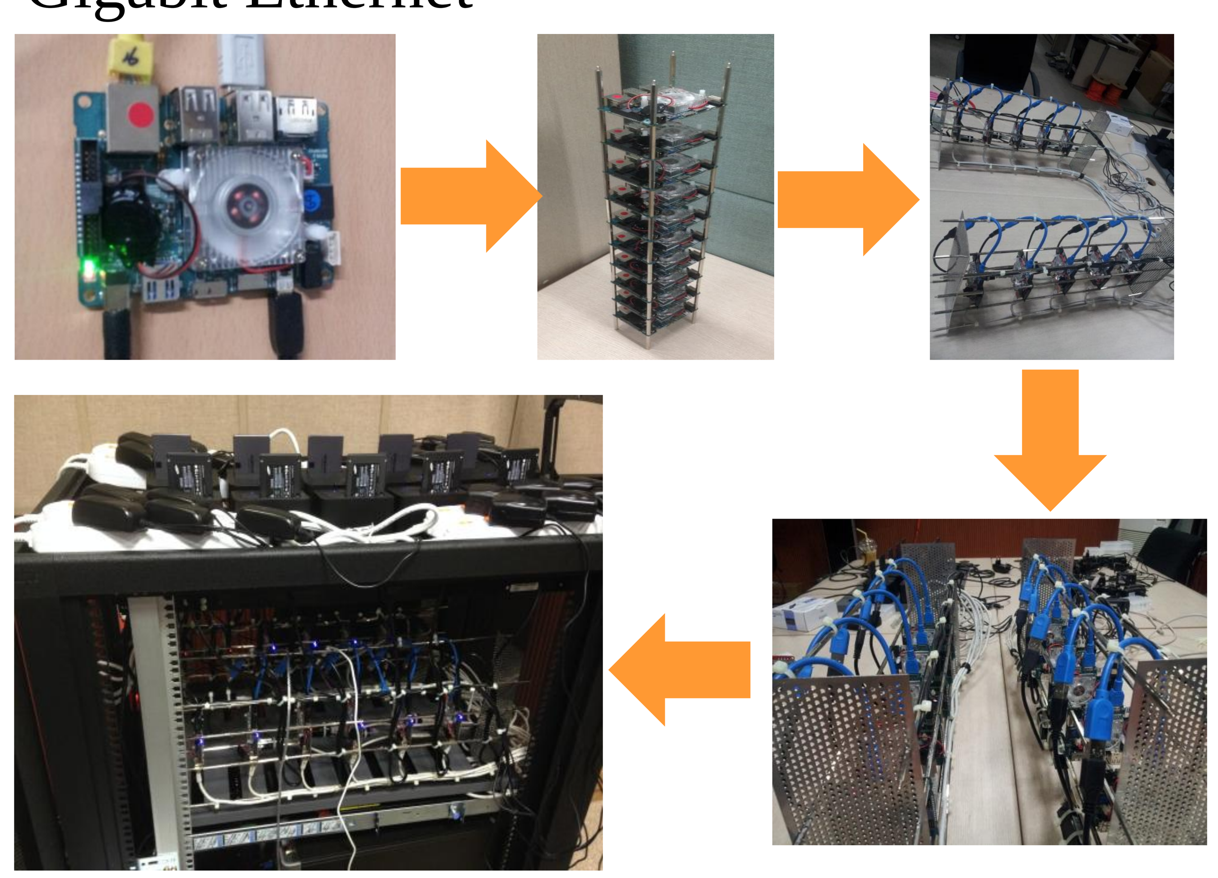
Building the SLAP Cluster

- A small scale of 10 cluster nodes based on mobile processors (Samsung Exynos 5)

[Figure 1: System Architecture of SLAP Cluster]

H/W Specification

- A single node spec
  - ODROID-XU board consisting of Samsung Exynos 5410 SoC CPU, 2GB LPDDR3 RAM, 10/100 Ethernet Controller, USB 3.0 and 2.0 ports
  - 64GB eMMC, 500G SSD as local storages
- Built a cluster of 10 nodes connected via 1 Gigabit Ethernet



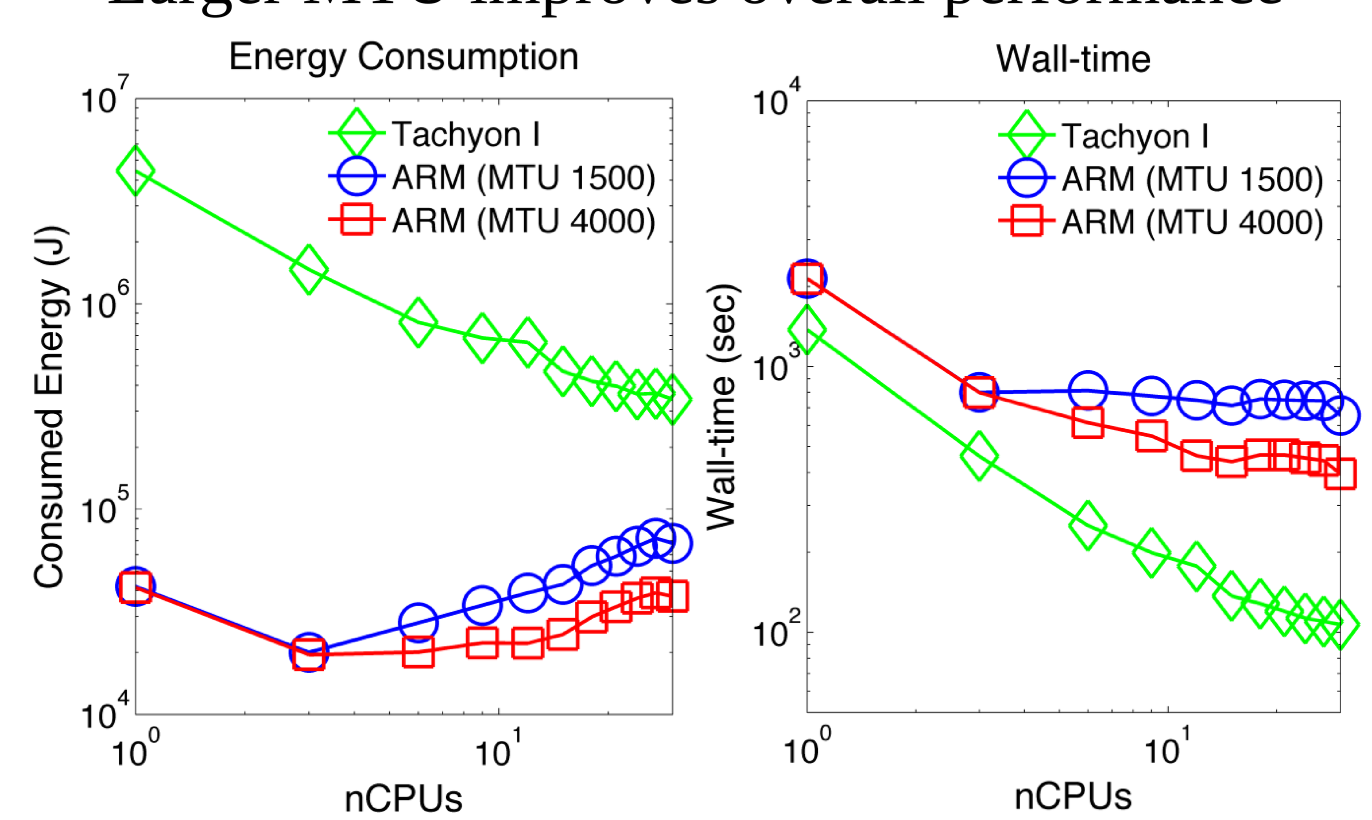
[Figure 2: From a single node to the final prototype]

HPL Benchmark Test

- A single node performance
  - Maximum 7.598GFlops => 2.739 GFlops/Watt (similar to *Green500 18th*)
  - Power consumption: 1.5Watts (idle time), Max 9.44Watts (CPU+Memory+GPU consume 4.49 Watts [47.6%])
- 10-nodes cluster performance
  - Maximum 47.44GFlops (83.25% of linear scaling) => 1.35 GFlops/Watt
  - Performance drops mainly due to limited scalability of interconnect and some waste of energy other than CPU, Memory, or GPU

Semiconductor Engineering Application Evaluation

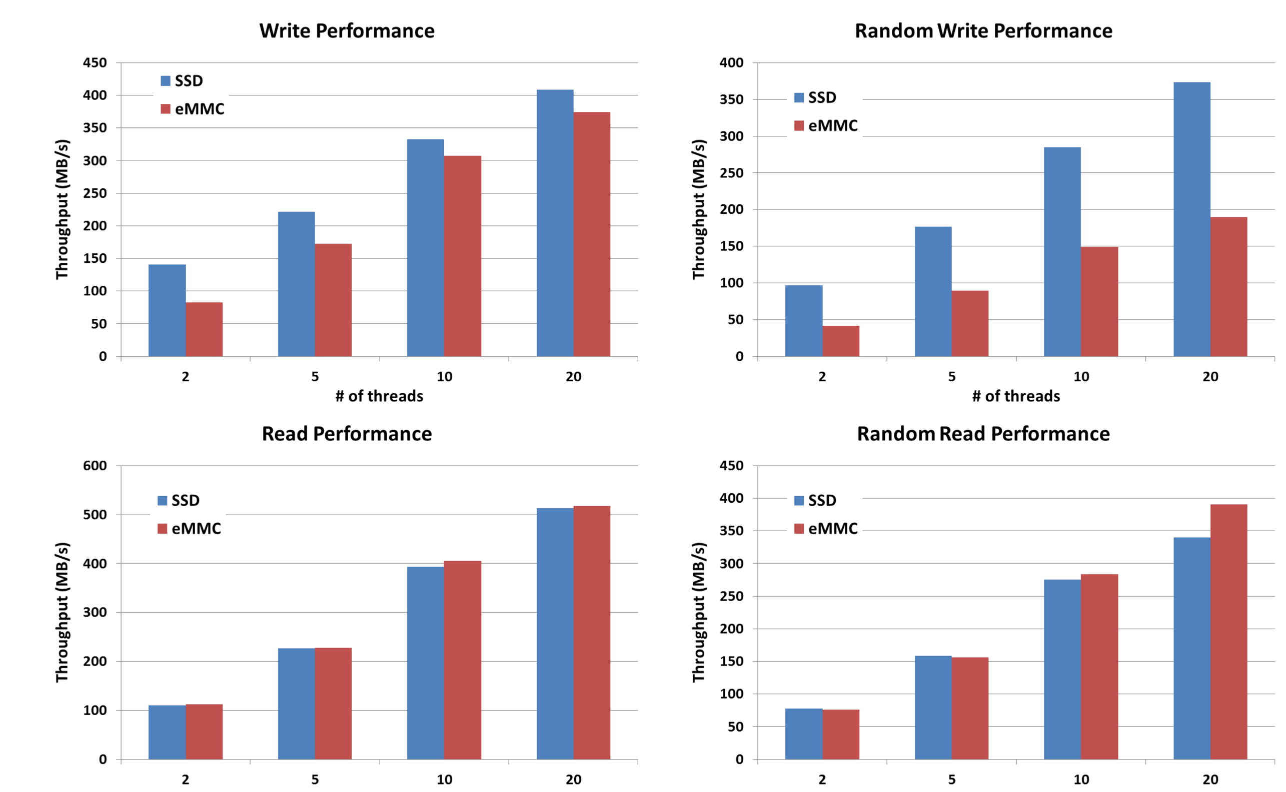
- Finding optical gap of a dome-shaped InAs/GaAs Quantum
- Dot with the LANCZOS algorithm; Compute 5 eigenvalues of a symmetric 400k x 400k complex matrix
- Benchmarking against KISTI Tachyon I
  - Reduction of energy consumption: *5~70 times*
  - Increase of wall-time: 4~6 times
- Larger MTU improves overall performance



[Figure 3: LANCZOS algorithm evaluation results]

GlusterFS(v3.2.7) Evaluation

- Constructing a storage cluster consisting of two logical volumes (eMMC & SSD)
- Evaluation Results
  - Aggregated throughput (10 SLAP cluster nodes) = 5 GByte/s (40% compared to the 20 storage nodes in CHEP'13)
  - Read/Re-Read: eMMC ≒ SSD
  - Random Read: eMMC > SSD (max 15%)
  - Write/ReWrite: eMMC < SSD (max 6~9%)
  - Random Write: eMMC < SSD (max 97%)
  - For Write-once, Read-many data analysis applications, eMMC can be better (suited for "*cache*")
  - For Write-intensive applications, SSD can be better (suited for "*buffer*")



[Figure 4: GlusterFS evaluation results]

Current Status & Future Work

- We could find potential benefits, possibility along with limits of applying embedded technology on supercomputing area
- As the mobile processor and ARM server technologies evolve, we will continue to investigate applying state-of-the-art techniques to save energy
- We plan to build a 8-node server appliance with 64-bit ARM mobile processor (Exynos 7)