Toward Footprint-Aware Power Shifting for Hybrid Memory Based Systems

THE UNIVERSITY OF TOKYO

Eishi Arima
The University of Tokyo

Toshihiro Hanawa
The University of Tokyo

Martin Schulz
Technical University of Munich

Technical University of Munich

Introduction

- ***Power consumption*** is one of the major concerns in HPC
  - Desired features: ***power capping*** and ***power shifting*** [1]
- ***Memory bandwidth/capacity limitations*** are also severe
  - Promising solution: ***hybrid main memory architecture***
    - DRAM + NVRAM (Intel Optane SSD[2])
    - MCDRAM + DDR4 (in Intel Knights Landing[3])

Goal of this research

A sophisticated ***power management scheme*** for hybrid memory based systems

Motivation

Key observation: the ***effective bandwidth decreases*** as the footprint size scales (**Fig. 1**) on a hybrid memory system

As a result, the performance bottleneck changes ***depending on the footprint size*** (**Fig. 3**)

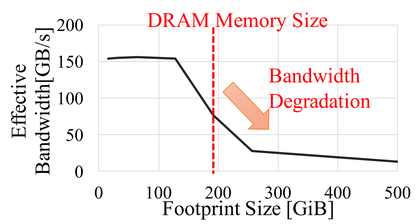

Fig. 1. Footprint size vs bandwidth

```
# pragma omp parallel for simd
for (i = 0; i < N; i ++) {
   A[i] = A[i] * B[i] ... * B[i];
}
```

- Footprint size ∝ N
- Arithmetic Intensity ∝ # of *B[i]

Fig. 2. Tested Code

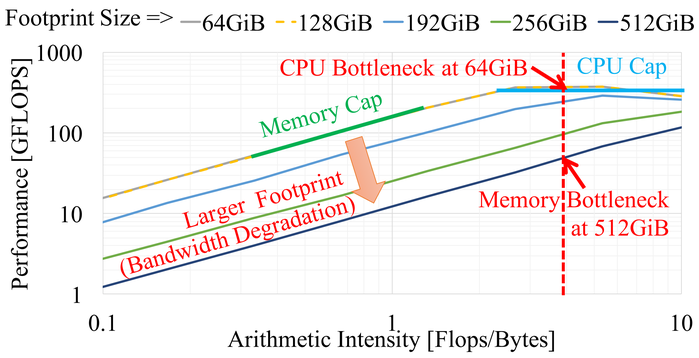

Fig. 3. Measured rooflines [4] for various footprint size

Concept

Objective: Maximizing performance (***Perf***) under a total power constraint (P_{total}) by controlling ***CPU/memory power*** (P_{cpu}/P_{mem}) for a given ***footprint (or problem) size*** (***S***)

- **Formulation:**

  max $Perf(P_{cpu}, P_{mem}, S)$
  s.t. $P_{cpu} + P_{mem} \leqq P_{total}$

* NVRAM power control is not considered – future work

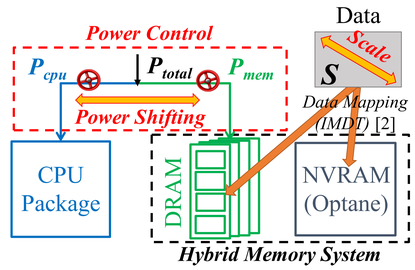

Fig. 4. Our environment

Solution

Footprint Aware Power Shifting: Shifting power between ***Pcpu*** and ***Pmem*** in accordance with the ***footprint size*** (**S**)

- We should allocate ***more power*** on the ***bottleneck component, which highly depends on the footprint size*** (**Fig. 3**)

Evaluation Setting

Experiment

We test various combinations of {P_{cpu}, P_{mem}} and choose the best one for each footprint size (***S***) under a given P_{total} (=P_{cpu}+P_{mem})

Environment

- **System Configuration:** Summarized in **Table 1**
- **Power Management:** Running Average Power Limit (RAPL)[5]
- **Workloads:** FFT, Lulesh, and the synthetic streaming code shown in **Fig.2** (Streaming)

Table 1. System Configuration

CPU Package	Xeon Gold 6154 Processor (Skylake), 18 cores, 3.0GHz, TDP 200W x2 sockets
Memory System	**DRAM:** DDR4-2666 x12 modules, 12ch, 192GiB **NVRAM:** Intel Optane SSD P4800X, 375GB, 2.4GB/s(read), 2.0GB/s(write) x2 cards **Data management:** IMDT [2]
OS	Cent OS 7.4
Compiler	Intel C++/Fortran Compiler 17.0.4 **Options:** -O3, -qopenmp

Power Shifting Result

Synthetic streaming code

- **CPU intensive case (Fig. 5):** Shifting power from P_{cpu} to P_{mem} improves performance because the CPU becomes less critical
- **Memory intensive case (Fig. 6):** Shifting power from P_{mem} to P_{cpu} improves performance – the DRAM needs less power due to frequent NVRAM accesses

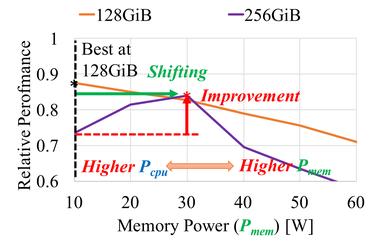

Fig. 5. Streaming (F/B: 10.7)
P_{total} (= P_{cpu} + P_{mem}): 220 [W]

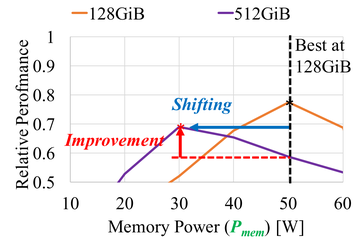

Fig. 6. Streaming (F/B: 0.17)

Benchmarks

- We observed performance improvement also for benchmarks

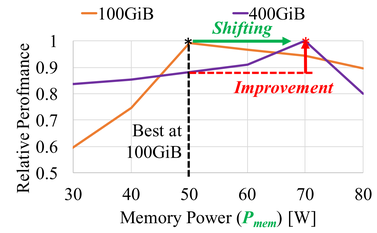

Fig. 7. FFT
P_{total} (= P_{cpu} + P_{mem}): 260 [W]

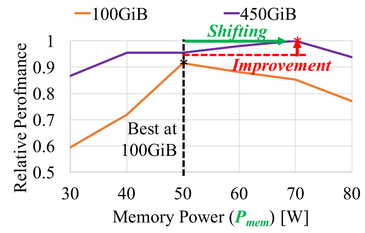

Fig. 8. Lulesh

* Performance is normalized to that without power capping for each data size

Conclusions and Future Work

Conclusions

Footprint-aware power shifting is promising to improve the performance of power constrained hybrid memory based systems

Future Work

- Developing a software framework, a performance model, and a power allocation algorithm to realize our proposal
- Including NVRAM power management
- Evaluating with other hybrid memory based systems such as Intel Knights Landing

Contact

[name] Eishi Arima
[organization] The University of Tokyo
[email] arima@cc.u-tokyo.ac.jp
[web] http://www.cspp.cc.u-tokyo.ac.jp/arima/index-e.html

Acknowledgement

This work is partly supported by JSPS Grant-in-Aid for Research Activity Start-up (JP16H06677), JSPS Grant-in-Aid for Early-Career Scientists (JP18K18021), and Research on Processor Architecture, Power Management, System Software and Numerical Libraries for the Post K Computer System of RIKEN.

References

[1] Lefurgy, C., et al. "Power Capping: A Prelude to Power Shifting". Cluster Computing 11, 2 (2008), 183–195.
[2] Intel. "Intel® Memory Drive Technology, Set Up and Configuration Guide", 2017.
[3] Jeffers, J., et al. "Intel Xeon Phi Processor High Performance Programming: Knights Landing Edition". Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2016.
[4] Williams, S., et al. "Roofline: An Insightful Visual Performance Model for Multicore Architectures". Communications of the ACM 52, 4 (2009), 65–76.
[5] Intel. "Intel® 64 and ia-32 Architectures Software Developer's Manual", System Programming Guide, 2017.