Measuring and Modeling Data Movement Power on Real Hardware

Vignesh Adhinarayanan

Introduction

- Power and energy have become a first-class design constraint in all areas of computing

Modified from Image by Vernon Chan CC BY 2.0

Image by Glenn Batuyong CC BY 2.0

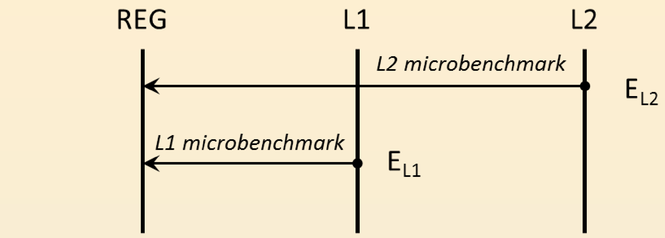
Image by Patrick Strandberg CC BY-SA 2.0

- Data movement thought to be a major source of power consumption

Objective

Measure and understand the nature of data-movement power on real hardware

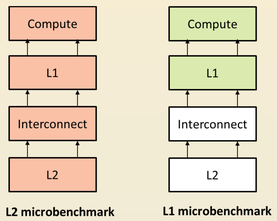
Related Work



G. Kestor et al., "Quantifying the energy cost of data movement in scientific applications," IISWC 2013

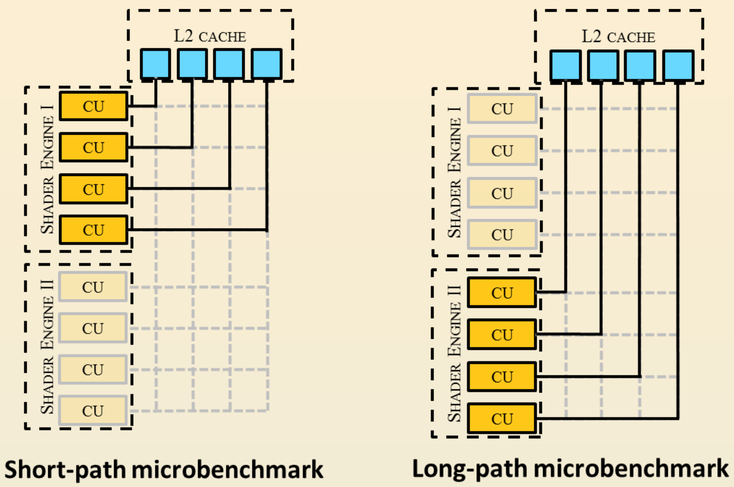
Energy cost of moving data from L2 to L1 = $\mathbf{E_{L2} - E_{L1}}$

Limitation



- L2 access power cannot be isolated

Our Approach



Design microbenchmarks based on data-movement distance as wire distance affects power

Challenges and Solutions

- OpenCL™ lacks native support to pin threads to programmer-specified compute units

```
__kernel void l2_read( __global float *data,
                       __global float *output) {
        int gid = get_global_id(0);
        int wid = get_group_id(0);
        if (wid >= 0 && wid <= 10) {
                // Read data from L2
        }
}
```

(a) Initial OpenCL code snippet

```
s_min_u32     s0, s0, 0x0000ffff  // 000000000014: 8380FF00 0000FFFF
s_mul_i32     s0, s16, s0         // 000000000018: 93000010
s_add_i32     s0, s0, s1          // 00000000001C: 81000100
v_add_i32     v0, vcc, s0, v0     // 000000000020: 4A000000
s_add_i32     s0, s16, s2         // 000000000028: 81000210
s_cmp_gt_i32  s0, -1              // 00000000002C: BF02C100
s_cbranch_scc0  label_0011        // 000000000030: BF840004
```

(b) Equivalent assembly code

```
00 FF 80 83 FF FF 00 00
10 00 00 93 00 01 00 81
00 00 00 4A 10 02 00 81
00 C1 02 BF 04 00 84 BF
00 FF 04 BF 81 00 00 00
C1 80 01 85 01 00 82 BF
```

(c) Equivalent binary (in hex)

```
00 FF 80 83 FF FF 00 00
10 00 00 93 00 01 00 81
00 00 00 4A 04 32 00 B9
00 C1 02 BF 04 00 84 BF
00 FF 04 BF 81 00 00 00
C1 80 01 85 01 00 82 BF
```

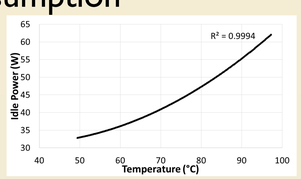
(d) Modified binary (in hex)

```
__kernel void l2_read( __global float *data,
                       __global float *output) {
        int gid = get_global_id(0);
        int cu_id = get_cu_id(0);
        if (cu_id >= 0 && cu_id <= 10) {
                // Read data from L2
        }
}
```

(e) Equivalent OpenCL code

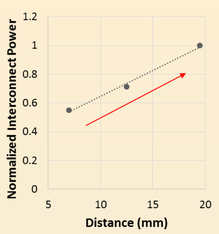
Solution: binary hacking

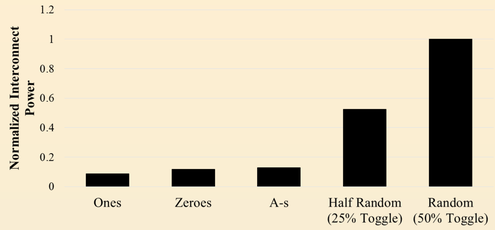
- Temperature of device during tests will affect the power consumption



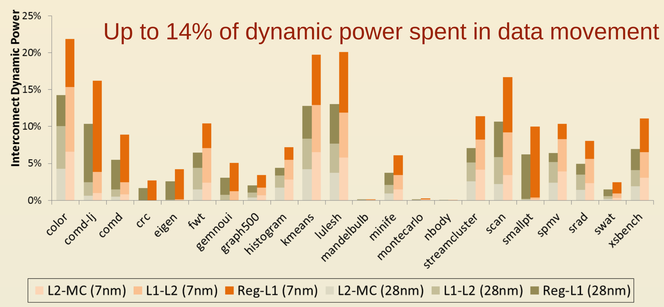
Solution: Model idle power separately and subtract from measured power

Summary of Results

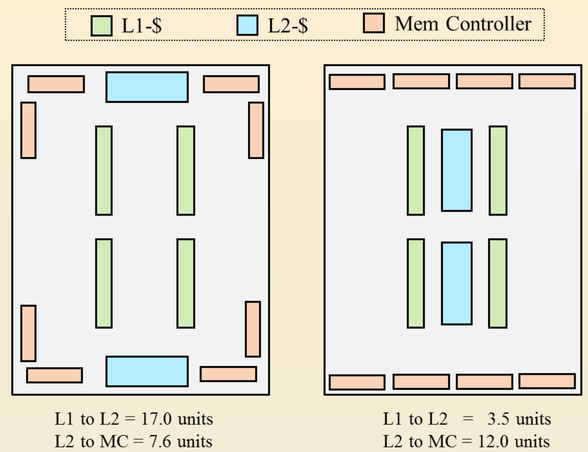




- Linear relationship between (i) distance and power (ii) toggle rate and power
- Asymmetrical power consumption for 0s and 1s



Optimization(s)



- On-chip data movement power reduces by 48% for an L1-L2 distance optimized layout
- More results in the paper

V. Adhinarayanan, I. Paul, J. Greathouse, W. Huang, A. Pattnaik, W. Feng, **"Measuring and Modeling On-Chip Interconnect Power on Real Hardware,"** IEEE International Symposium on Workload Characterization **(IISWC)**, September 2016. **Best Paper Award**