Performance & Energy Optimization @ OpenMP

Md Abdullah Shahneous Bari
Abid M. Malik
Millad Ghane
Ahmad Qawasmeh
Barbara M. Chapman

11/28/15

Layout of the talk

- Overview
- Motivation
- Factors that affect the performance & Energy Optimization
- Experimental Results
- Conclusion & Future Work

OpenMP

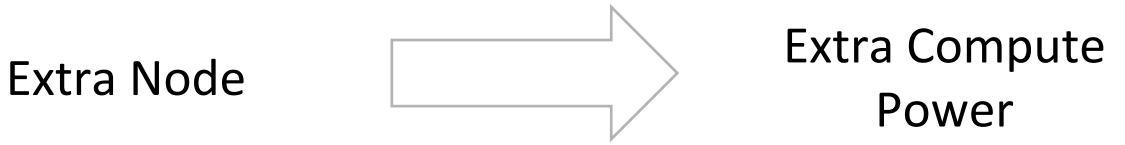
- De-facto standard for shared memory parallel programming
- Thread based parallelism
- Mainly two kinds of parallelism
  - Regular parallelism (work sharing constructs)
  - Irregular parallelism (task based constructs)

Main Barrier Towards Exascale Computing…

- Power, power and power
- 20MW power limit for *exascale* machines (DOE)
- Usually processor vendors concern
- But to reach the *exascale* limit software stack have to chip in
- Any solution????

Power Constrained Computing (Overprovisioning)

- Usually not all application use maximum node power all the time
- Capping the power at lower limit
- Allows extra node to be added at the similar power budget



Power Constrained Computing(Contd.)

- More focus on overall system level performance
- Some related work,
  - Sarood et al. [1]
  - Patki et al. [2]
  - Rountree et al. [3]

1. Sarood, Osman, et al. "Optimizing power allocation to CPU and memory subsystems in overprovisioned HPC systems." *Cluster Computing (CLUSTER), 2013 IEEE International Conference on*. IEEE, 2013.
2. Patki, Tapasya, et al. "Exploring hardware overprovisioning in power-constrained, high performance computing." *Proceedings of the 27th international ACM conference on International conference on supercomputing*. ACM, 2013.
3. Rountree, Barry, et al. "Beyond DVFS: A first look at performance under a hardware-enforced power bound." *Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW), 2012 IEEE 26th International*. IEEE, 2012.

Why *OpenMP*???

- Current Issue: Less focus on per-node performance
- Challenge: To reach the peak throughput, per-node performance must be improved
- *OpenMP* is the most popular language of choice for intra node parallelism

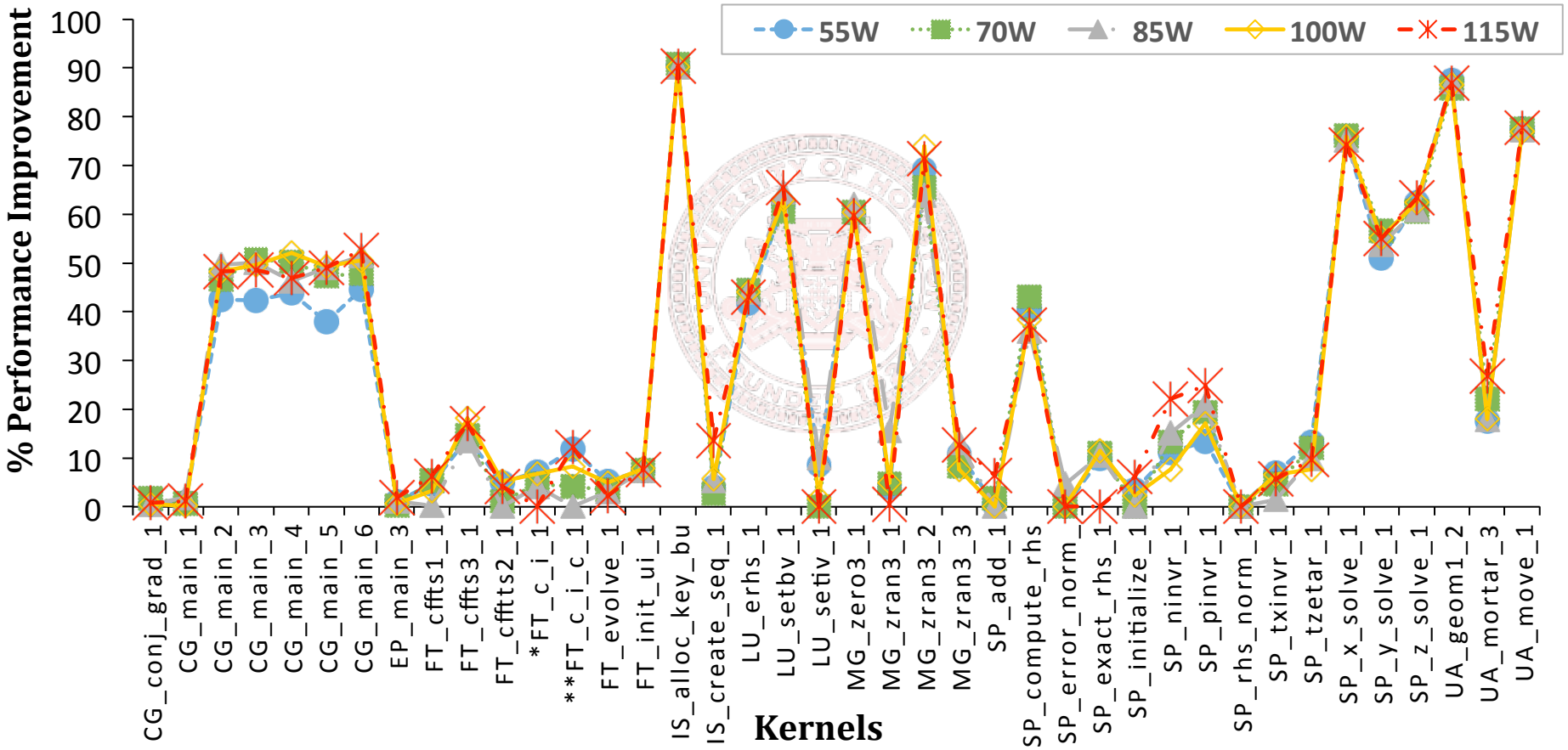
Factors That Impact Work Sharing Parallelism…

- How many workers are working? ~ Thread
- How the work is scheduled? ~ Scheduling Policy
- How much work they are given at one time? ~ Chunk Size
- How the data is laid out for the workers? ~ Thread Affinity
- What do the workers do during their break? ~ Wait Policy

Experimental Details

- Selected parameters
  - No. Of threads (2, 4, 8, 16, 24, 32)
  - Scheduling policy (STATIC, DYNAMIC, GUIDED)
  - Chunk size(1, 8, 32, 64, 128, 256, 512)
  - Wait policy (active, passive)
  - Thread affinity (OMP_PLACES + OMP_PROC_BIND)
- Power cap levels
  - $(55, 70, 85, 100, 115)$w
- Used technology:
  - Intel RAPL (for power capping & energy measurement)
  - OMPT for kernel level measurement
- Benchmark ~ NPB

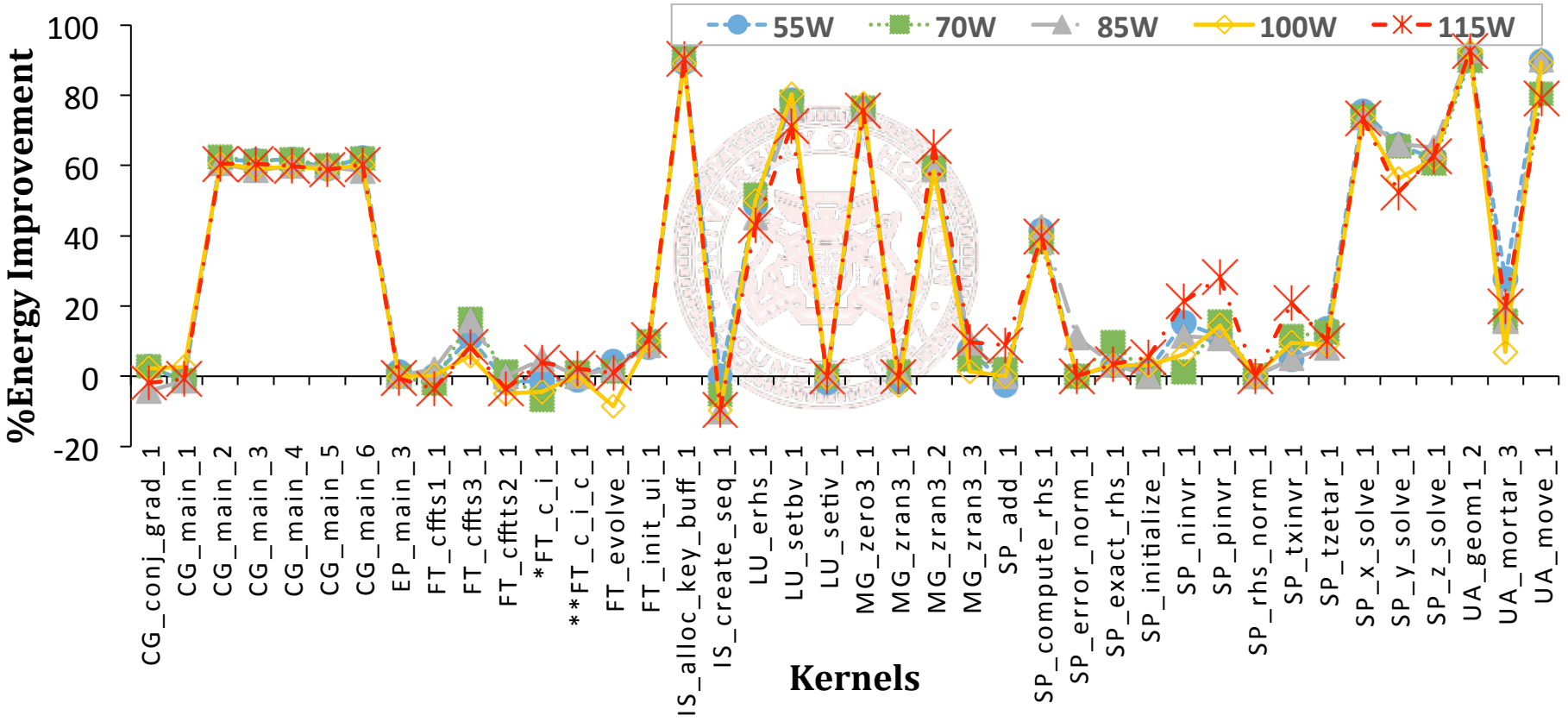
Performance improvement using the best configuration compared to default across all kernels

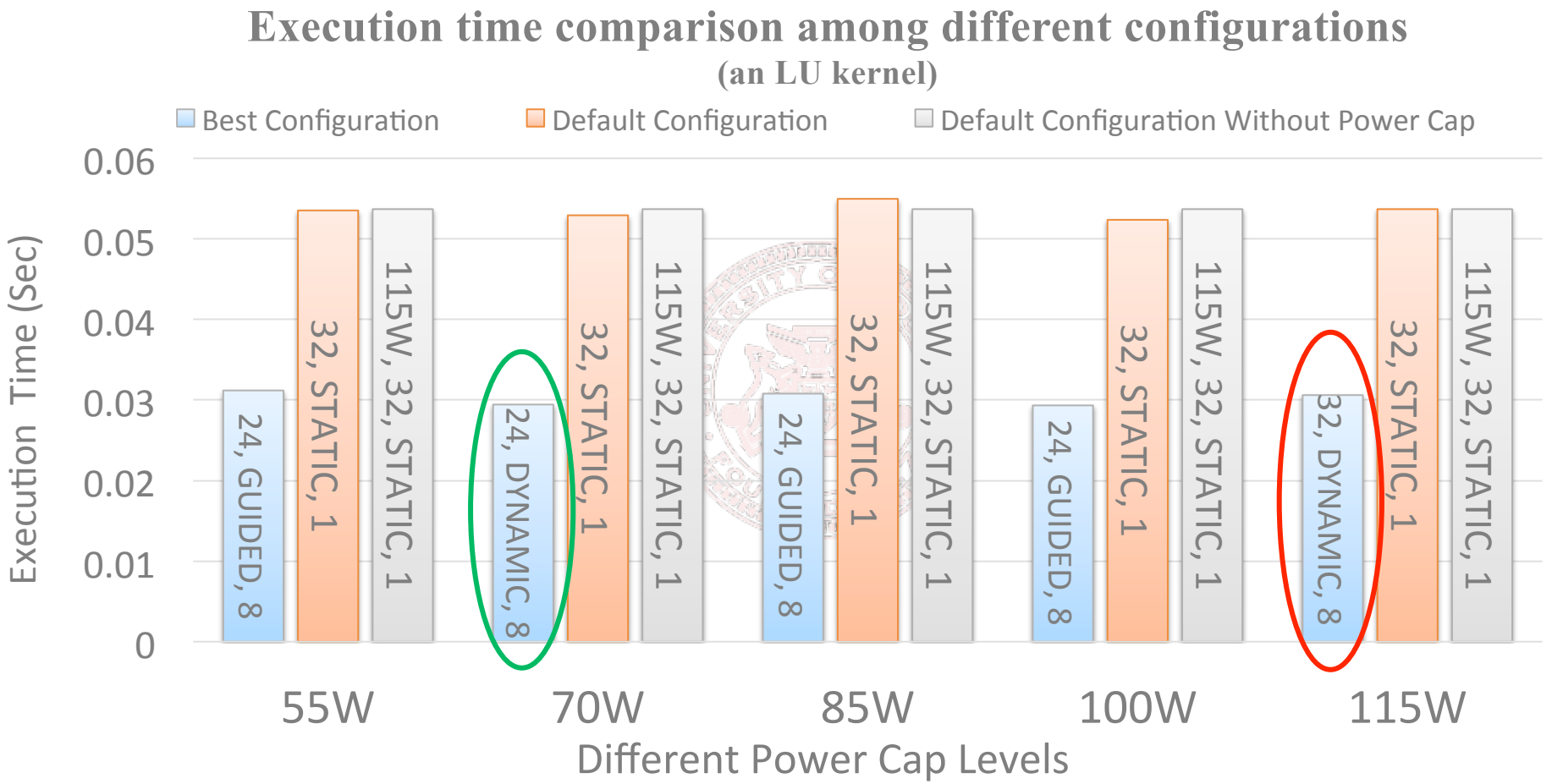


HPC TOOLS

OpenMP

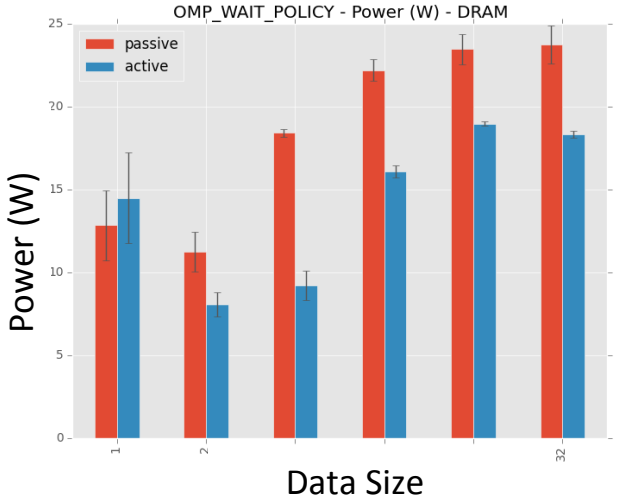
Energy consumption improvement using the best configuration compared to default across all kernels

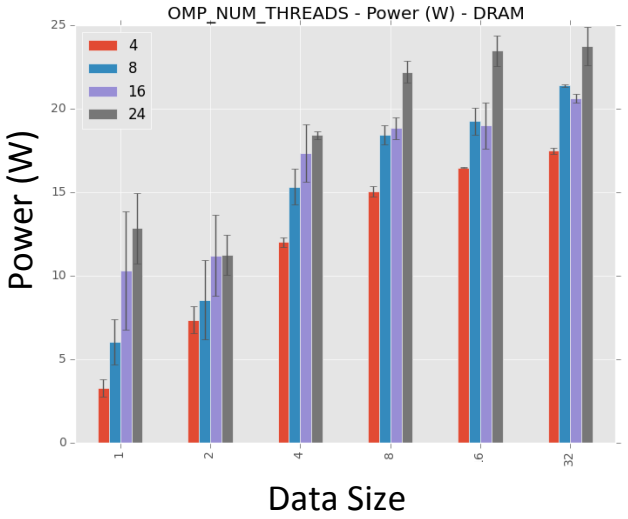


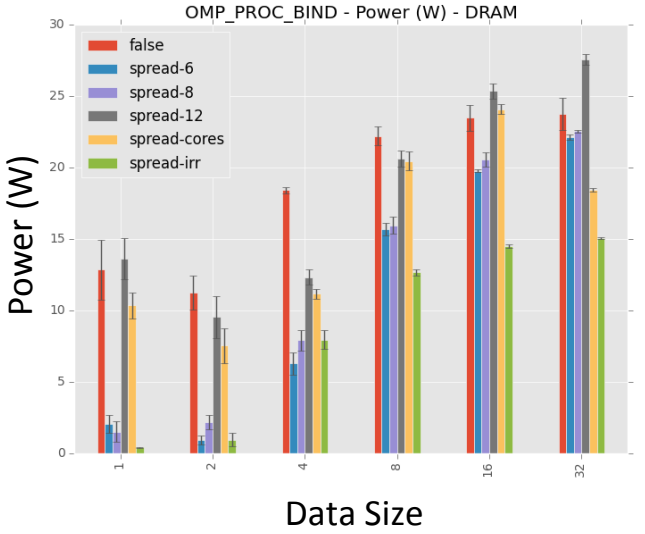

Execution time comparison among different configurations
(an LU kernel)
Best Configuration
Default Configuration
Default Configuration Without Power Cap
Execution Time (Sec)
0.06
0.05
0.04
0.03
0.02
0.01
0
24, GUIDED, 8
32, STATIC, 1
115W, 32, STATIC, 1
24, DYNAMIC, 8
32, STATIC, 1
115W, 32, STATIC, 1
24, GUIDED, 8
32, STATIC, 1
115W, 32, STATIC, 1
24, GUIDED, 8
32, STATIC, 1
115W, 32, STATIC, 1
32, DYNAMIC, 8
32, STATIC, 1
115W, 32, STATIC, 1
55W
70W
85W
100W
115W
Different Power Cap Levels

OpenMP ICVs on DRAM Power

- Developing a model for power consumption of openmp applications







These results are based on STREAM benchmark.

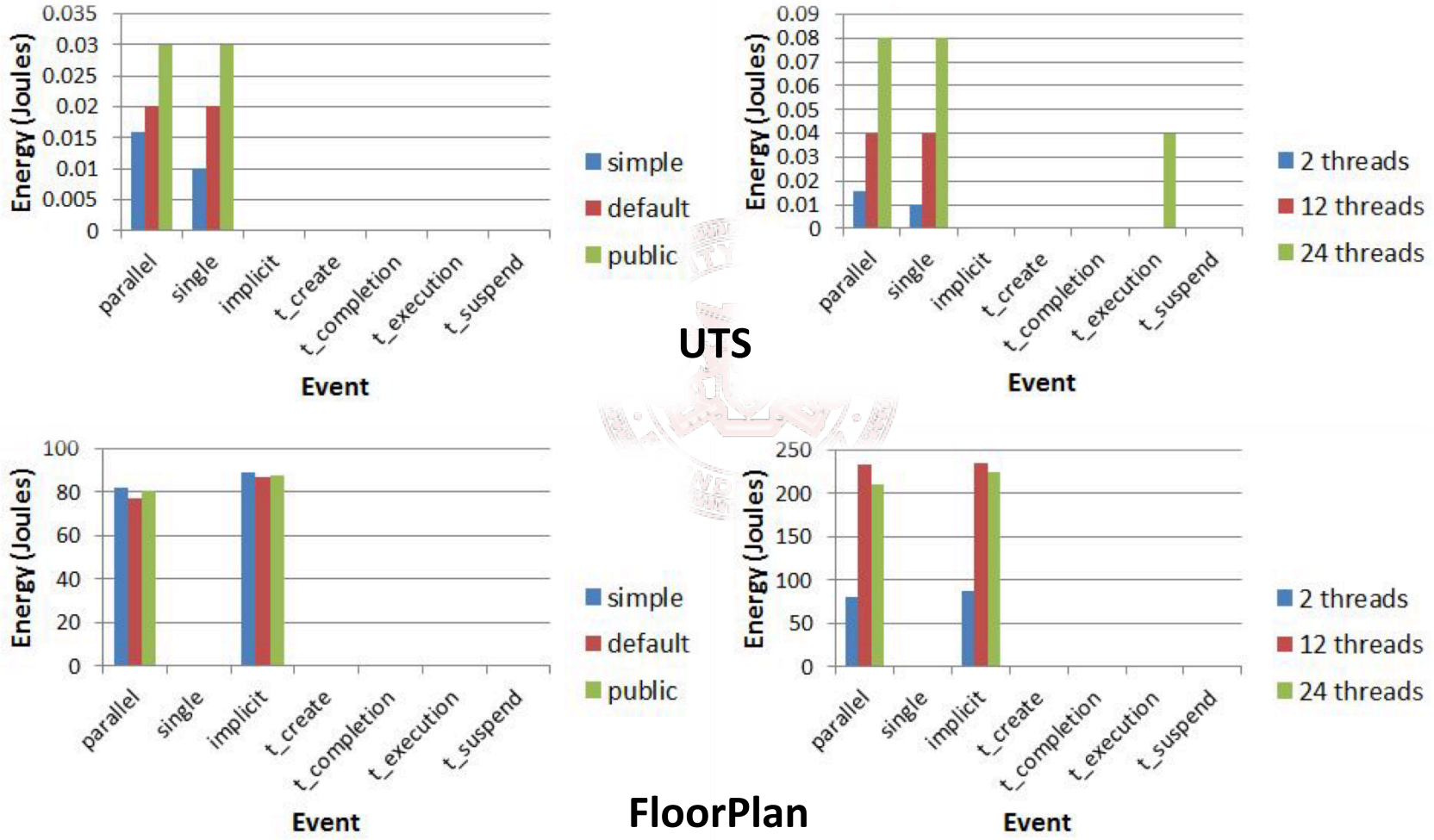
Data Size X means the array size for STREAM benchmark is 19,200,000*X.

Courtesy: **Millad Ghane**

HPC TOOLS

OpenMP

Impact of threads & scheduling policy in task based parallelism



Courtesy:
Ahmad Qawasmeh

Ongoing Work

- Dynamic adaptation (APEX),
  - Active harmony
  - Modeling
- Across different software stack (OpenMP runtime),
  - Openuh
  - GCC
  - Intel
- Across different hardware architecture
  - Intel sandybridge
  - IBM power8

Future Work

- More concrete configuration selection
- DRAM capping
- Fine grain (core level) control
- Other energy efficient techniques,
  - DVFS, frequency modulation etc.
- Combining it with a inter-node (MPI) programming models for hybrid applications

Summary

- Overview
- Motivation
- Factors that affect the performance & Energy Optimization
- Experimental Results
- Conclusion & Future Work