

# The GPU Clusterware Project

Andreas Diavastos

Constantia Alexandrou, Stelios Erotokritou, Christos Kallidonis, Giannis Koutsou, Andreas Panteli, Giannos Stylianou, George Tsouloupas

CaSToRC

Computation-based Science & Technology Research Center

<http://www.cyi.ac.cy/castorc>

OVERVIEW

## Project Overview

An inherent issue with co-processor accelerators is that their utilization requires arbitration from a CPU, either for resource management or performance monitoring. The project investigates the state-of-art in co-processor software to develop and implement management tools for efficient utilization of those devices as distinct compute resources

## Project Objectives

- Deliver application benchmarks for novel, co-processor accelerated hybrid architectures
- Include co-processor awareness:
  - Resource allocation
  - Computer time accounting
  - Resource monitoring

OBJECTIVES

INTEL XEON PHI

- Processor Name: Xeon Phi 7120P
- No. Cores : 61
- No. Threads : 4 / core (total 244)
- Frequency : 1.238 GHz
- Peak Perf. (d) : 1.208 TFlop/s
- Memory Band. : 352 GB/s
- Consumption : 300 W



- GPU Type : GK110
- No. CUDA Cores : 2496
- No. Threads : 2496
- Frequency : 706 MHz
- Peak Perf. (d) : 1.17 TFlop/s
- Memory Band. : 208 GB/s
- Consumption : 225 W



NVIDIA TESLA K20

MONITORING SYSTEM

## Ganglia Monitoring System

- Scalable distributed monitoring system
- Handles up to 2000-node clusters
- Cluster Monitoring
- Per node/accelerator monitoring
- Support for NVidia GPUs
- Support for Xeon Phis



## Co-Processor Aware Resource Manager

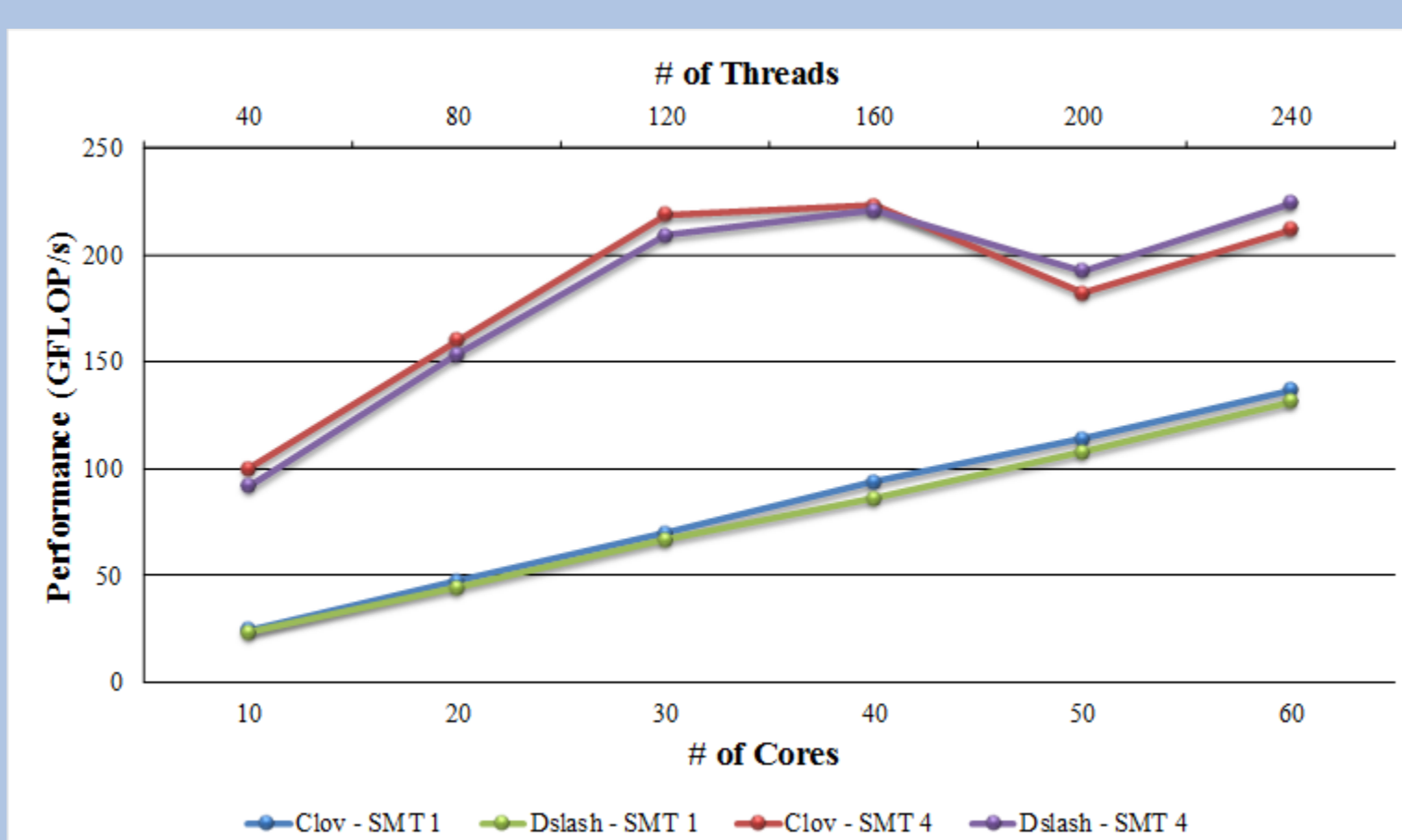
- Open-source workload manager for Linux Clusters
- Pinning of accelerators to users to allow sharing
- Aware of multiple resources:
  - Xeon CPUs
  - Xeon Phi Accelerators
  - K20m GPUs



RESOURCE MANAGER

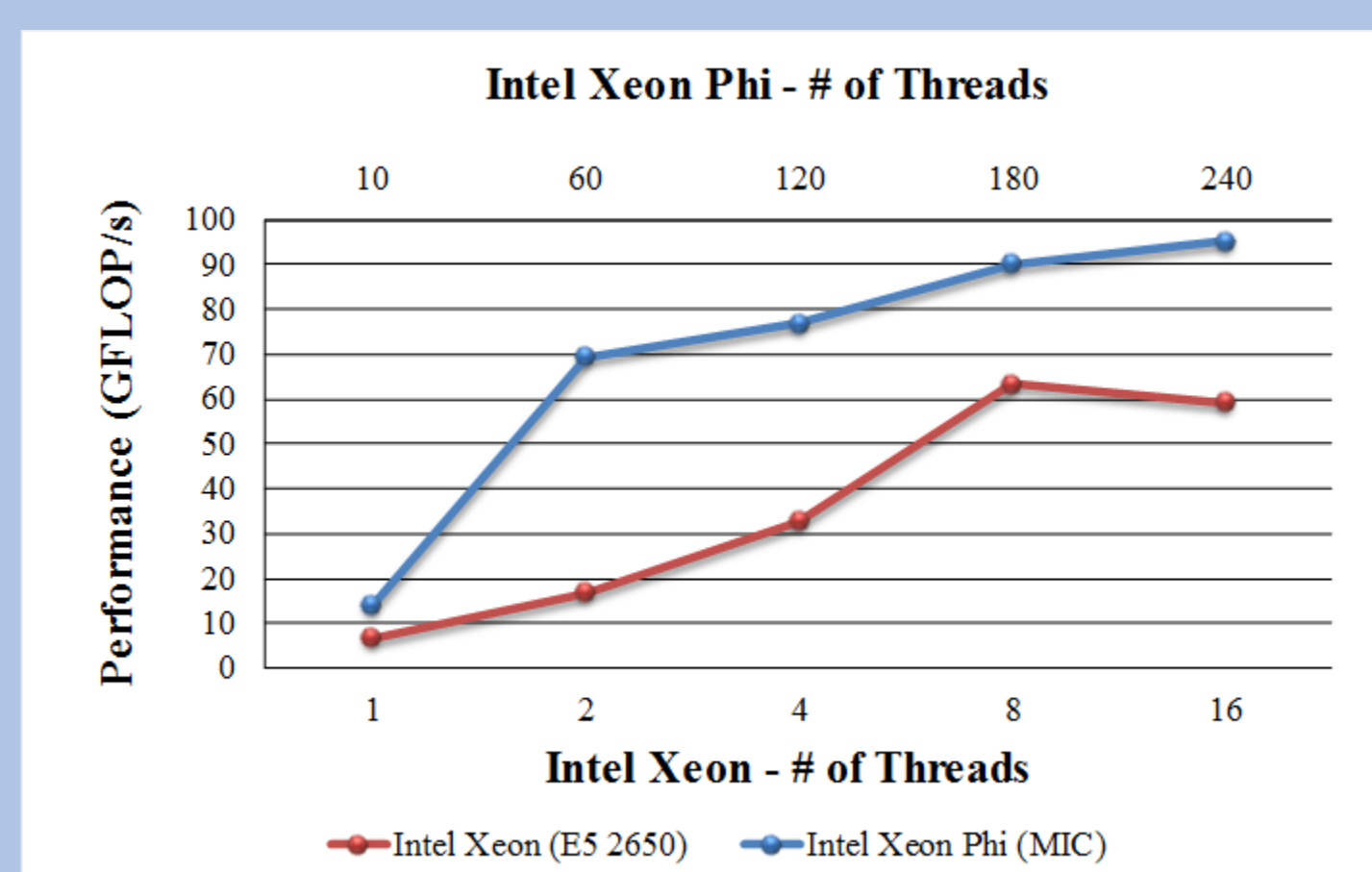
XEON PHI PERFORMANCE

## Qphix<sup>1</sup>



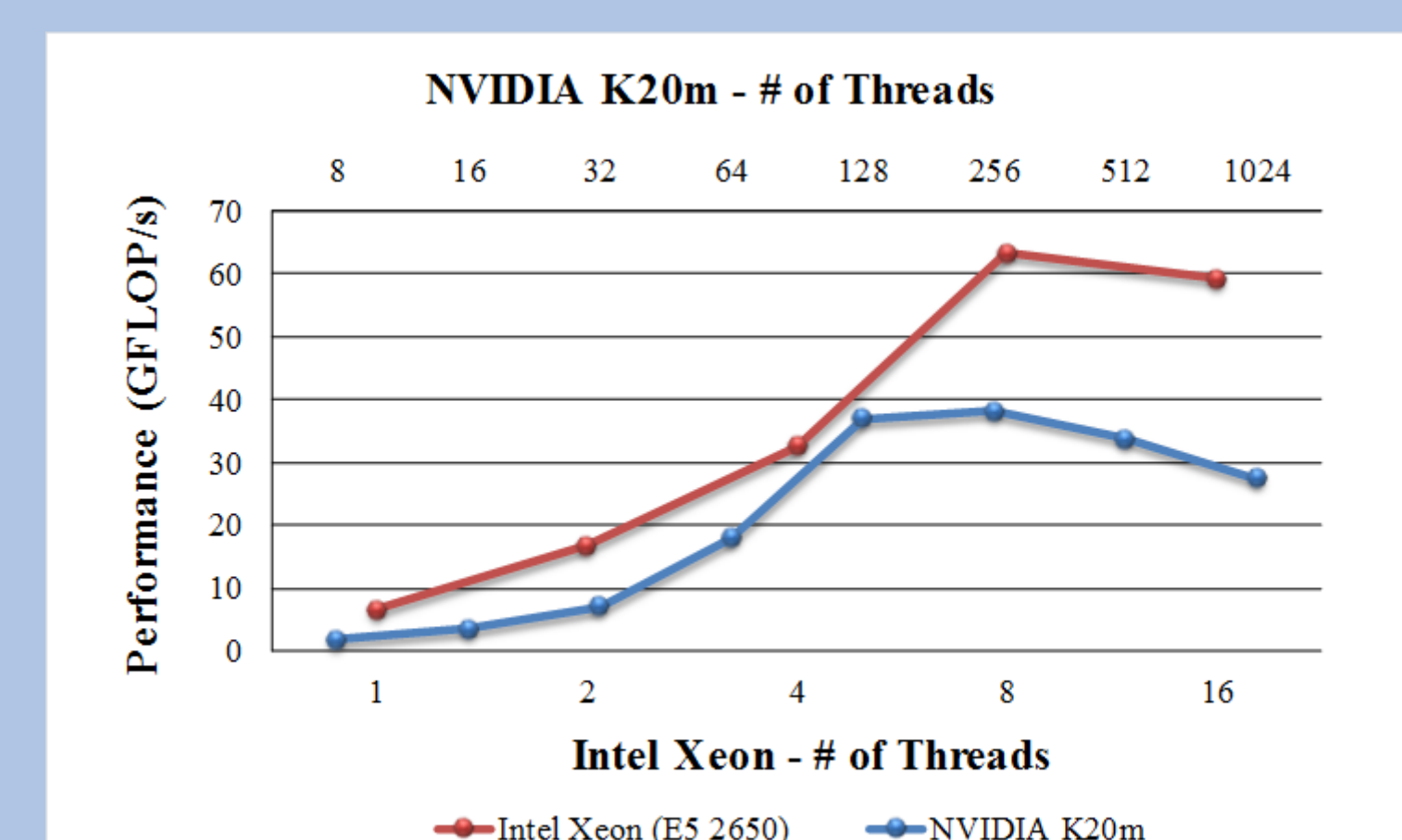
- Input Size: 48x48x24x64
- SMT provides data locality
- SMT offers better performance

## Laplacian<sup>2</sup>



- Input Size: 2048x2048
- Laplacian scales well on Xeon Phi and outperforms Intel Xeon CPUs

## Laplacian<sup>2</sup>

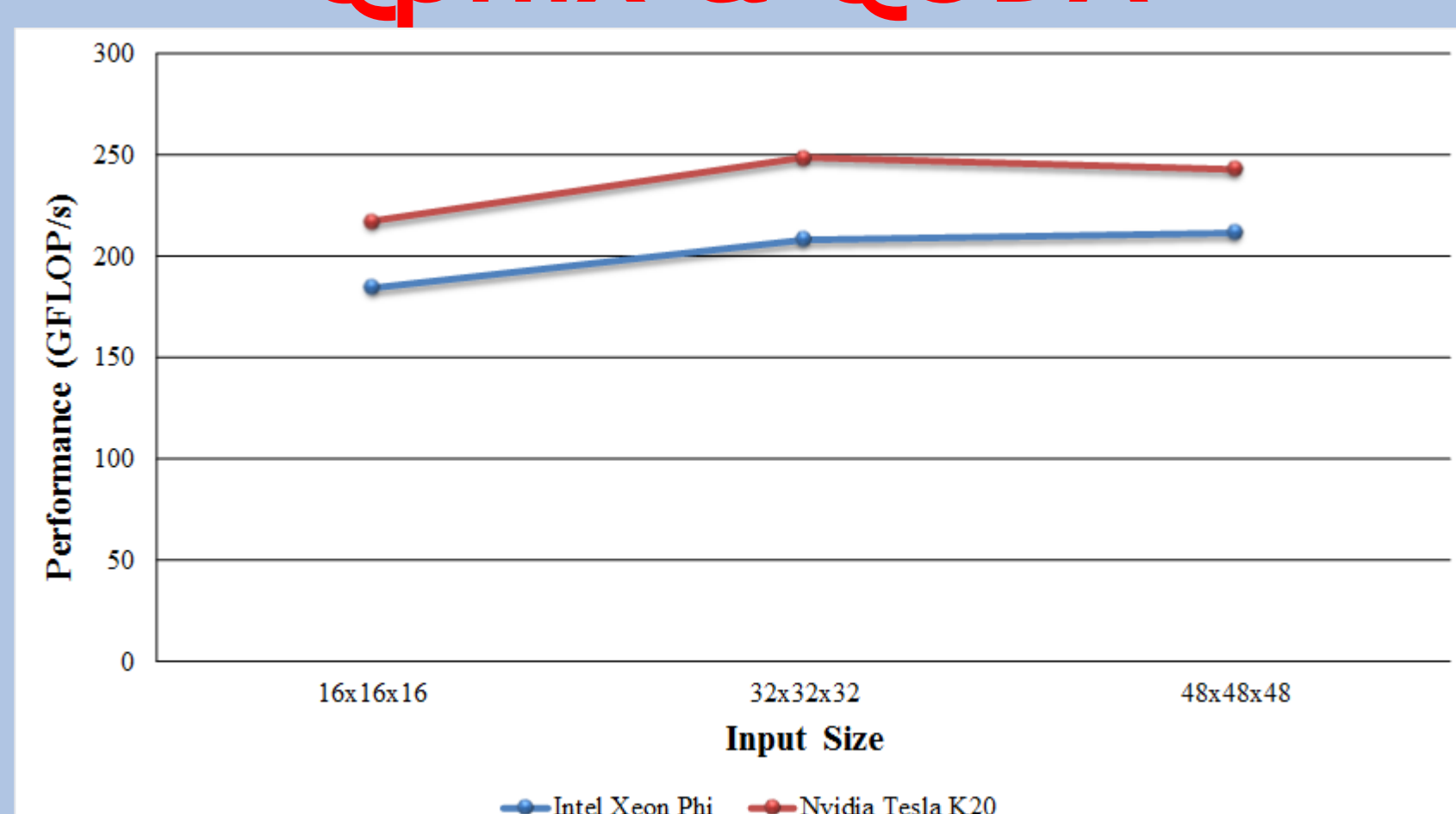


- Input Size: 2048x2048
- Moderately optimized for GPU execution

TESLA K20 PERFORMANCE

GPU vs. XEON PHI PERFORMANCE

## Qphix & QUDA<sup>3</sup>



- Performance scalability for different input sizes
- Both accelerators scale with increasing size
- GPU: 7% of peak (3.52TFlop/s)
- MIC: 9% of peak (2.416TFlop/s)

## Future Work

- More detailed provisioning tools
- Multi-node & Multi-accelerator performance results
- Application Benchmarks overhead analysis
- Detailed resource accounting through Slurm

## References

<sup>1</sup> <https://github.com/JeffersonLab/qphix>  
<sup>2</sup> <https://github.com/gpucw>  
<sup>3</sup> Babich, Ronald, Michael A. Clark, and Bálint Joó. "Parallelizing the QUDA library for multi-GPU calculations in lattice quantum chromodynamics." High Performance Computing, Networking, Storage and Analysis (SC), 2010 International Conference for. IEEE, 2010.

FUTURE DIRECTIONS



THE CYPRUS INSTITUTE



The GPU Clusterware Project, Grant No. TIE/ΠΛΗΠΟ/0311(BIE)/09, is co-financed by the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation