

Flexible Accelerator-based Many-Core Architecture

**Harnessing on-chip Resources to Optimize Performance for
Sequential and Data-Parallel Applications**

Hsien-Hsin Sean Lee

**Associate Professor
School of Electrical and Computer Engineering
Georgia Tech**

<http://arch.ece.gatech.edu>

Different Game Play for Many-Core Computing

- What is new?
 - Everything is On-Die



Performance Scalability

Area Scalability

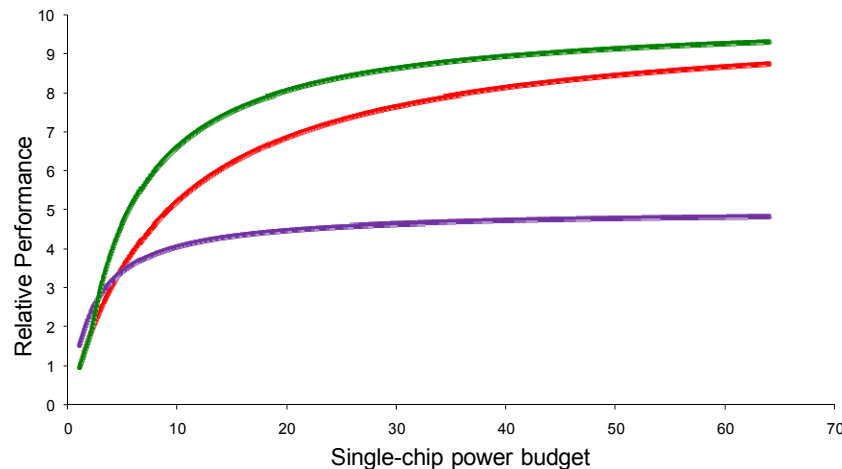
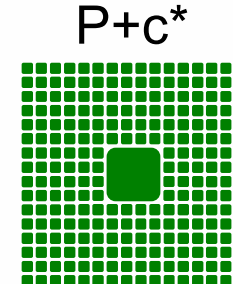
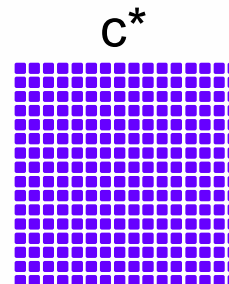
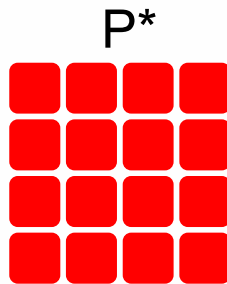
- Cores vs. Sustained Horsepower
- How many cores fit on a die?

Power Scalability

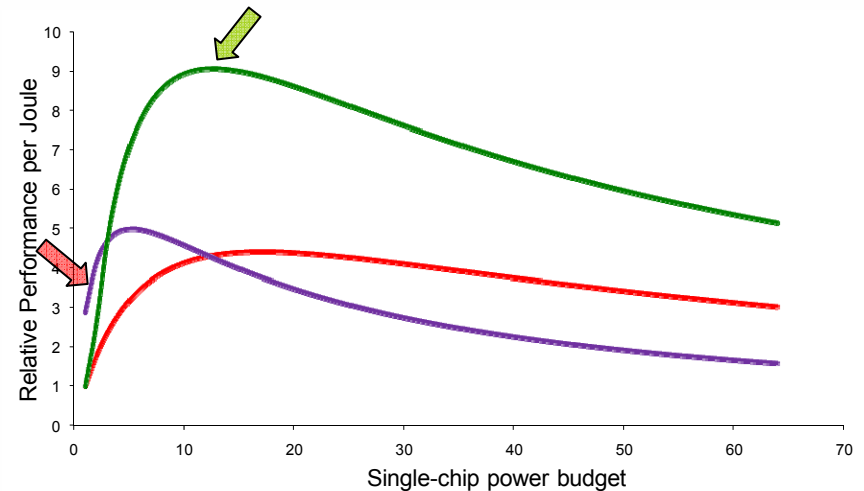
- Power envelope
- Power budget per core?

Performance, Energy Efficiency of Many-core

Multi-core / Many-core
Scaling Options



Power-Equivalent "Performance"



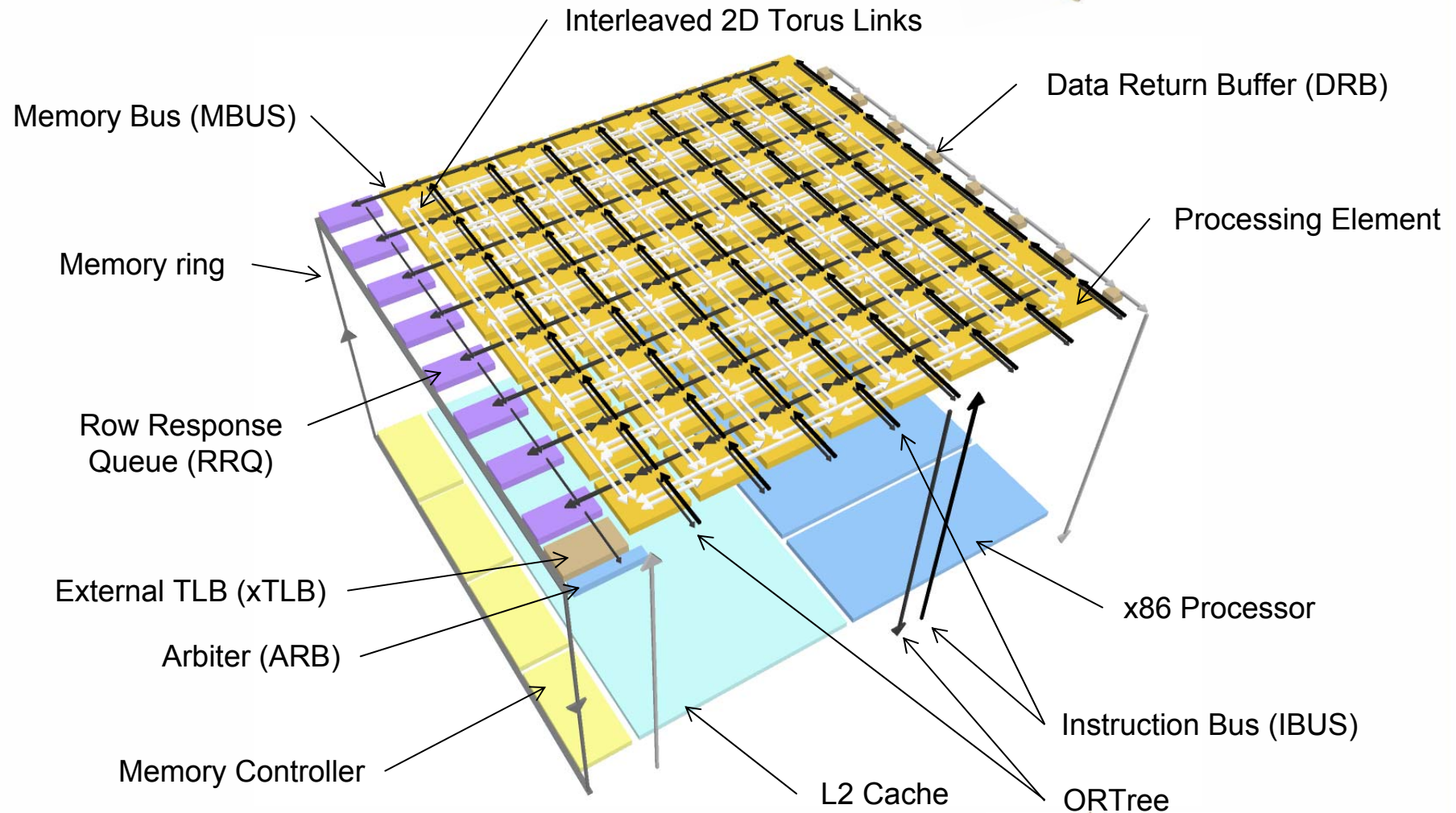
Power-Equivalent "Performance Per Joule"

$f = 0.9$ (parallelizable fraction)

$\text{Perf}(P) = 2 \times \text{Perf}(c)$; $\text{Power}(P) = 4 \times \text{Power}(c)$

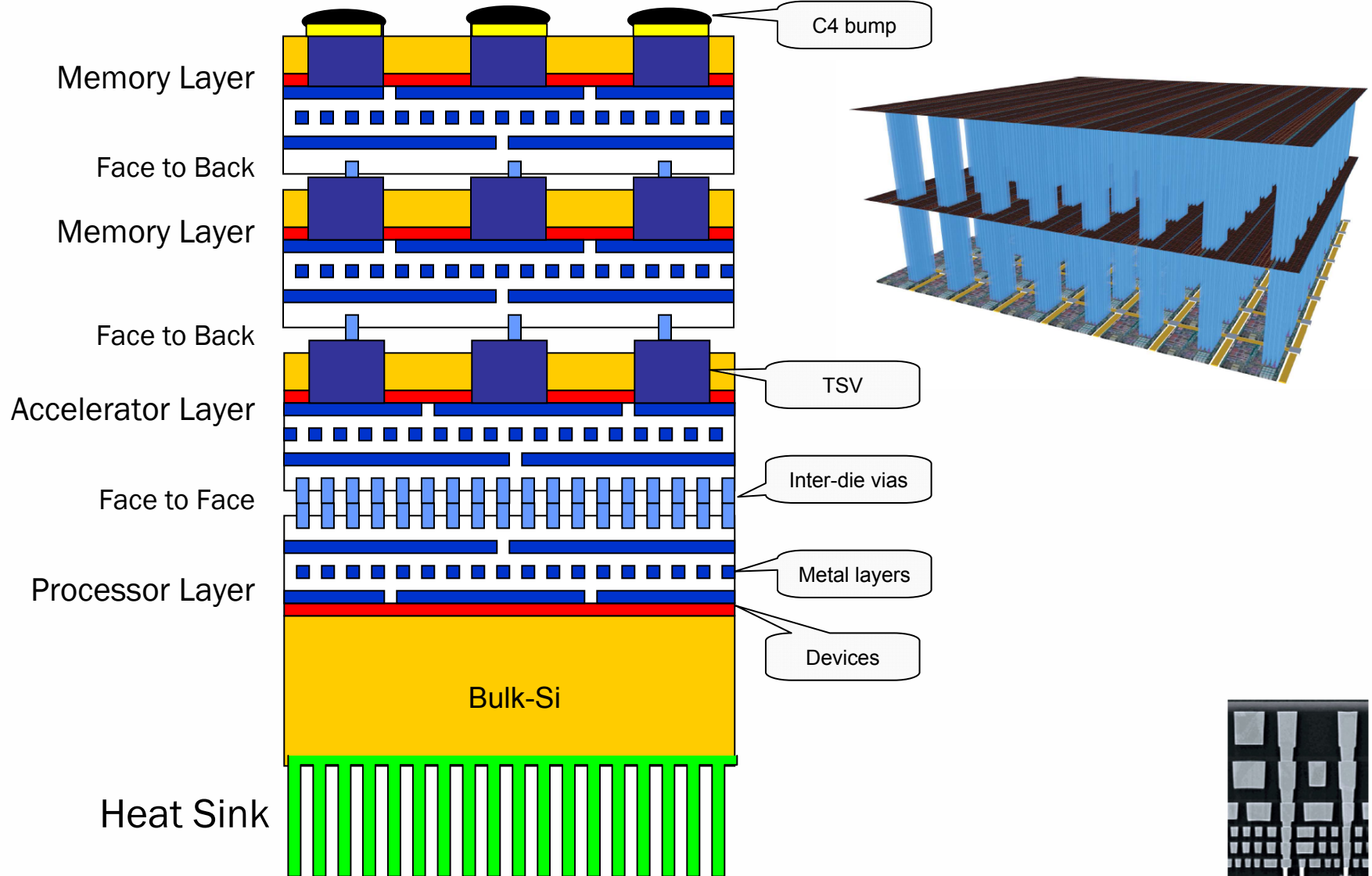
POD: 3D-Integrated Broad-Purpose Accelerator

POD: Parallel-On-Demand

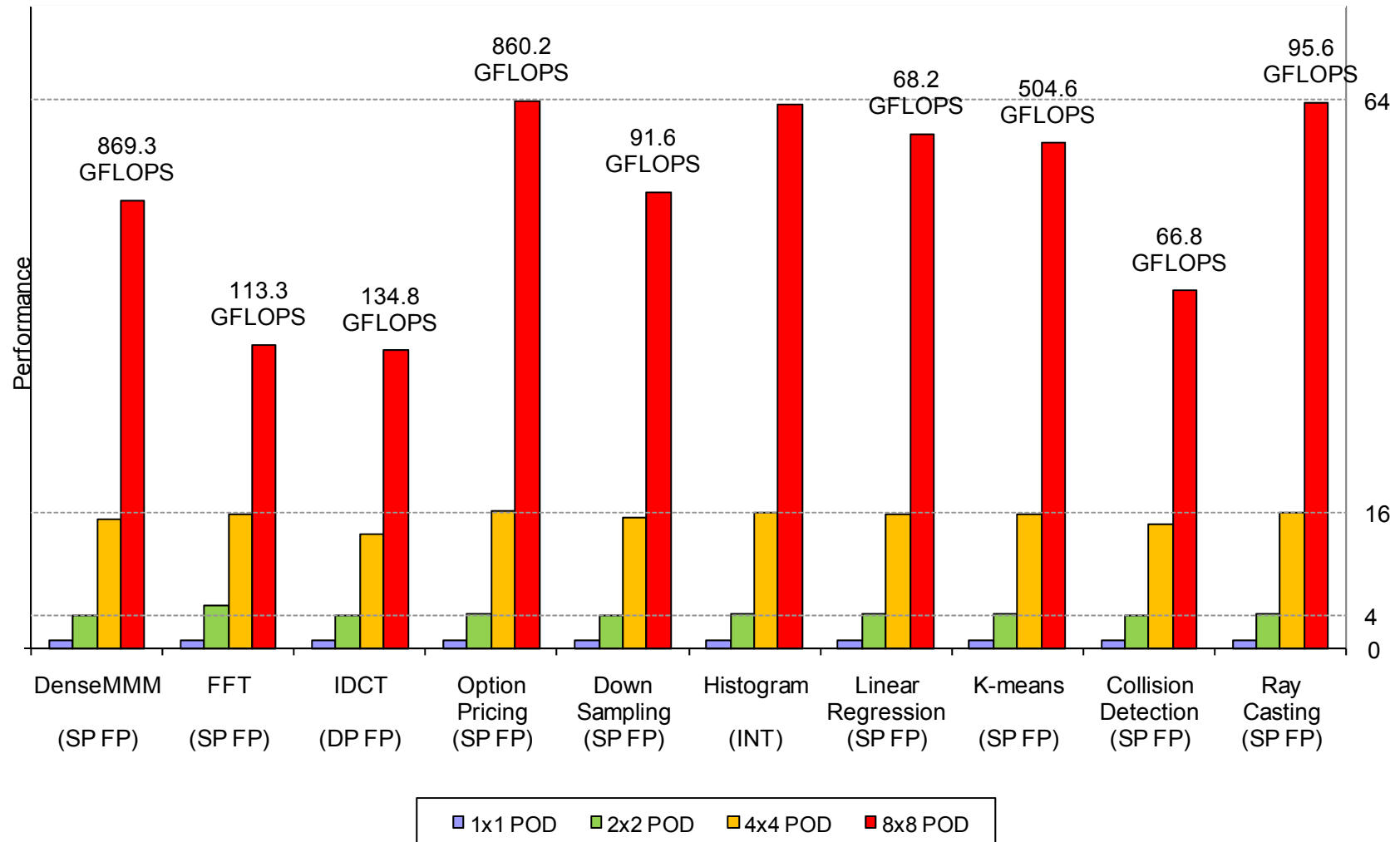
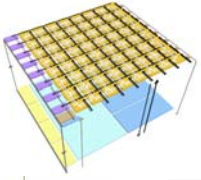


[Woo, Lee, et al., IEEE MICRO 08]

3D Die-Stacking Fabrication Technology

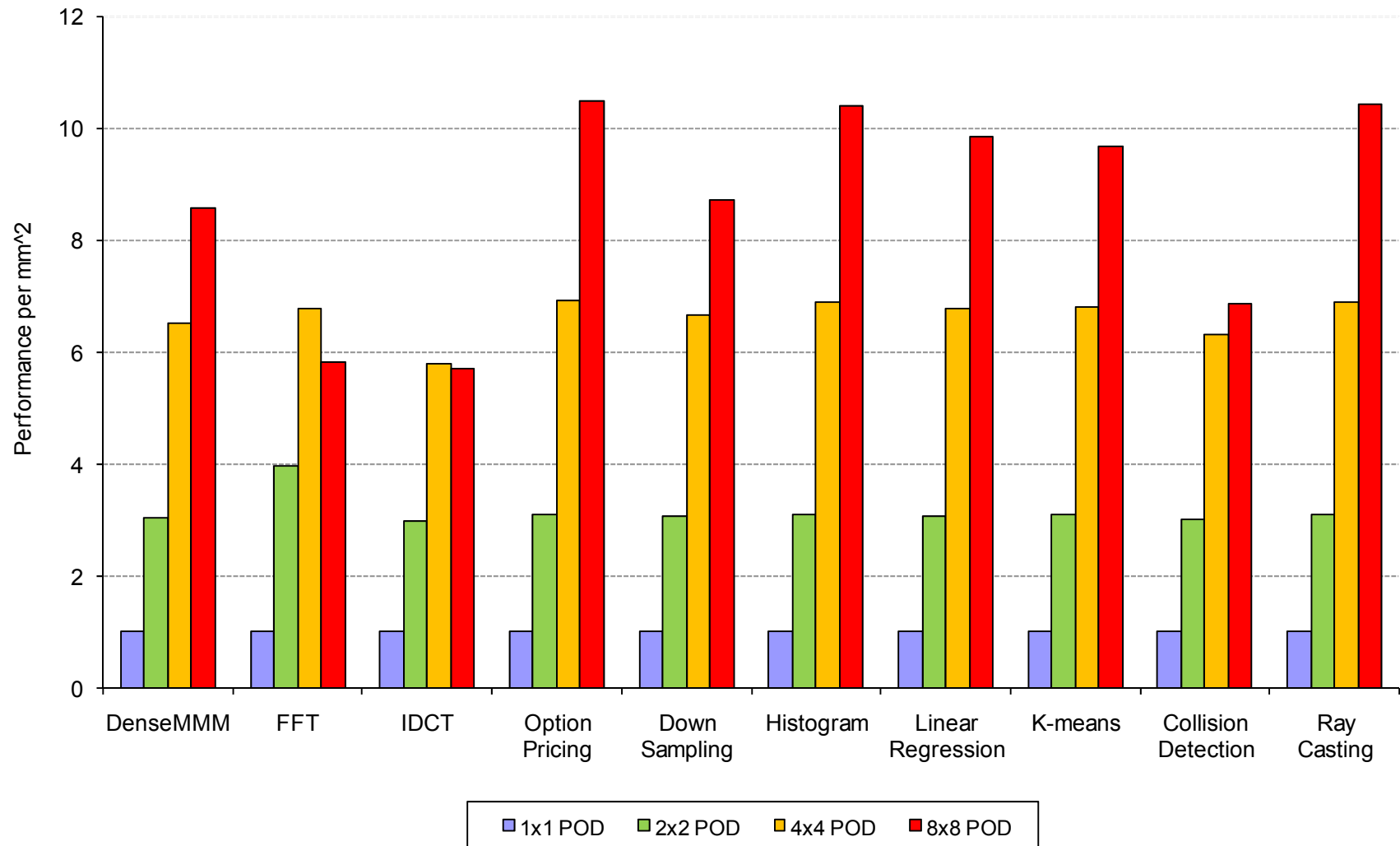
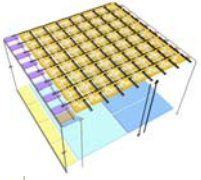


POD: Absolute Application Performance



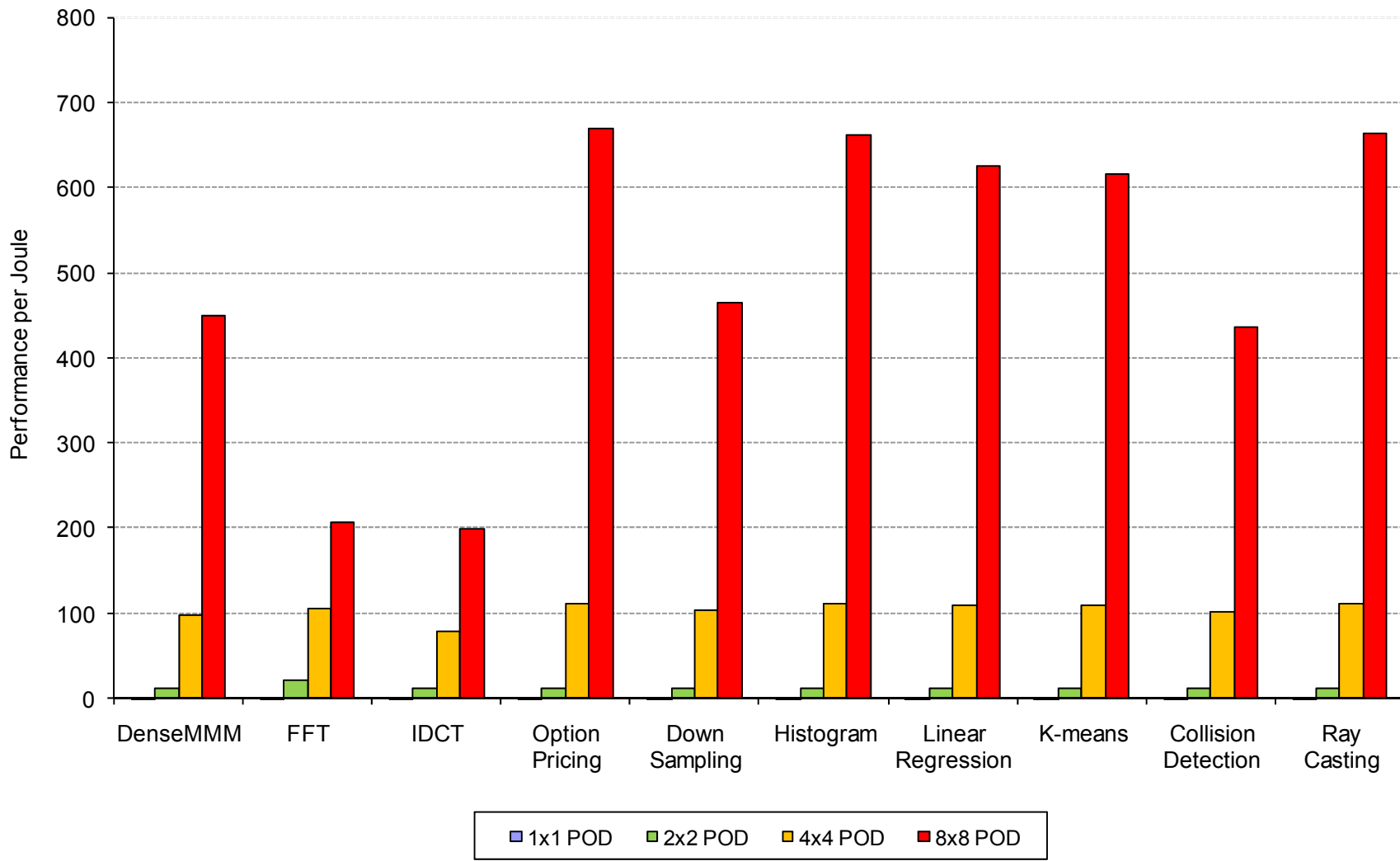
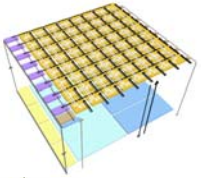
[Woo, Lee, et al., IEEE MICRO 08]

POD: Scalable “Performance per mm²”



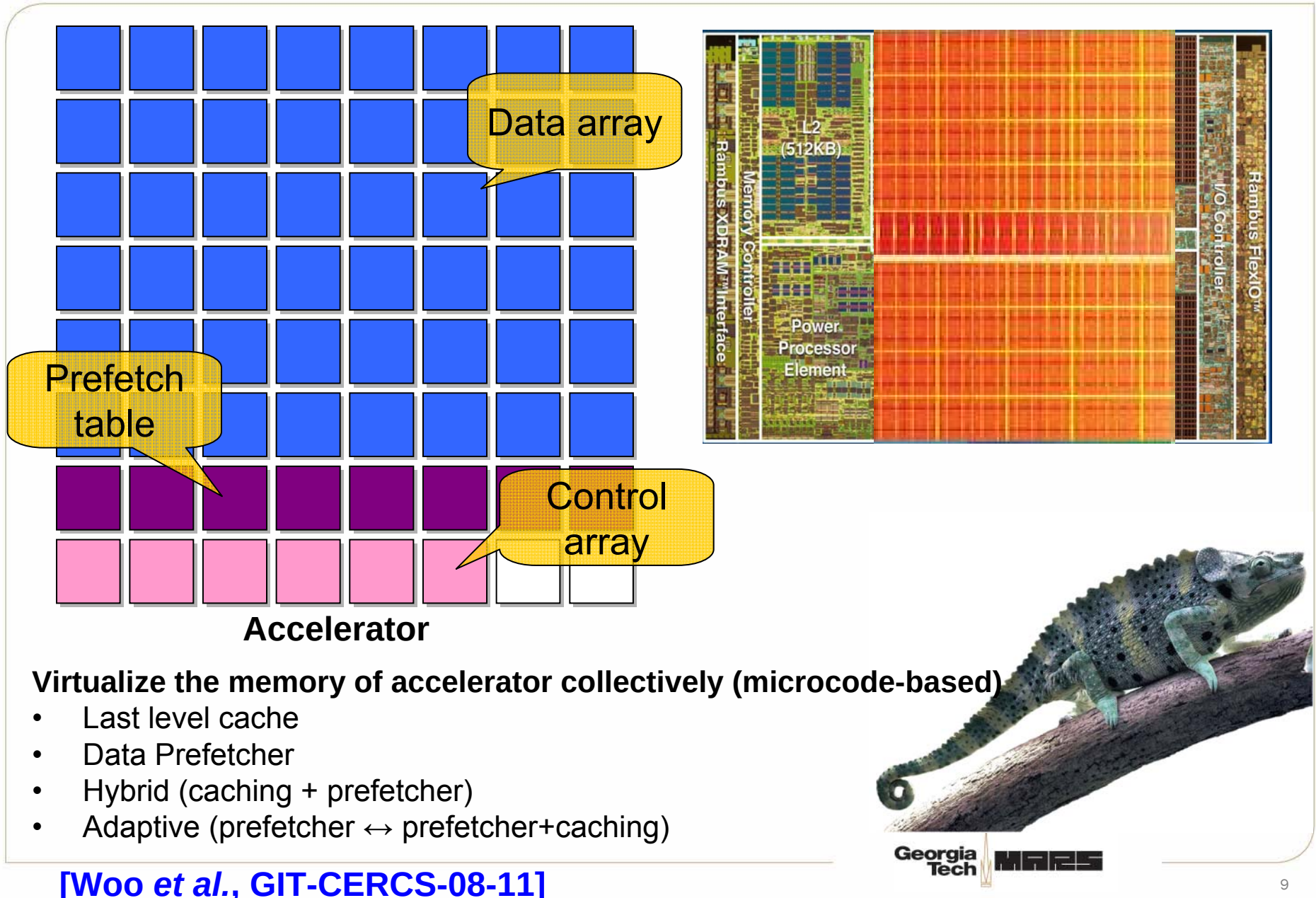
[Woo, Lee, et al., IEEE MICRO 08]

POD: Scalable “Performance per Joule”

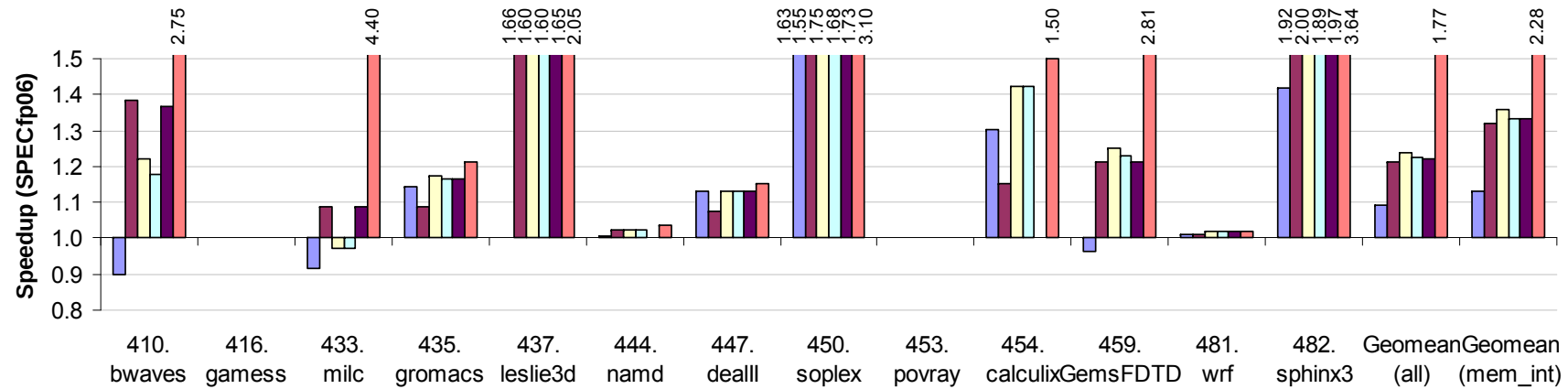
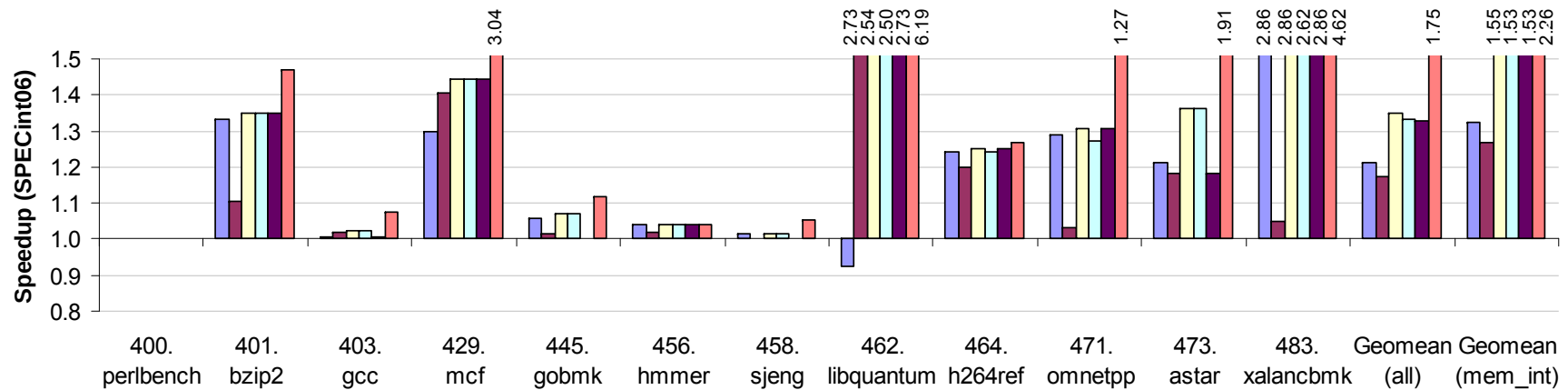


[Woo, Lee, et al., IEEE MICRO 08]

Chameleon: Virtualizing Accelerator's Idle Resources



Chameleon: Performance Evaluation



■ wblp_nuca
 ■ stride1
 ■ hybrid_7MB_128KB
 ■ hybrid_6MB_1MB
 ■ adaptive_7MB_128KB
 ■ perfect

7MB LLC

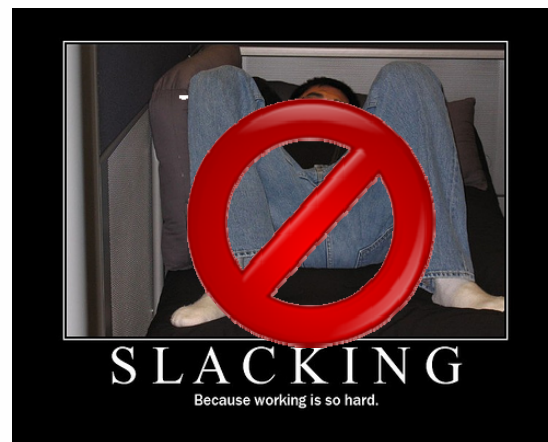
1MB Pref
Table

[Woo et al., GIT-CERCS-08-11]



A Flexible Polymorphic Many-Core Architecture

- Reduce NRE cost for various market products via optional 3D stacking
- Accelerate data-parallel algorithms
- Achieve high energy efficiency
- Accelerate sequential applications by harnessing idle resources
- Turn accelerators into extended memory subsystem
- Adaptively migrate among different operation modes



Georgia Tech
ECE MARS Lab
<http://arch.ece.gatech.edu>

