Center for Computational Physics University of Tsukuba

SCIMA: Software Controlled Integrated Memory Architecture for HPC

Background

- Memory wall problem
- Conventional Cache is not good in HPC
  - unwilling line conflict
  - fixed size of Off-Chip Memory access

Solution: SCIMA (Software Controlled Integrated Memory Architecture)

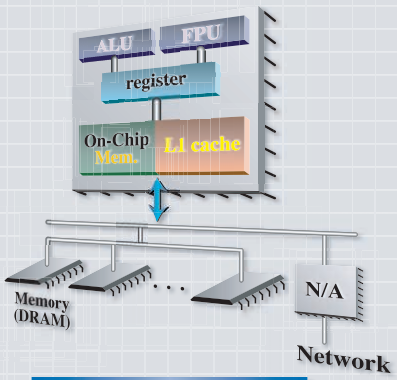
- Strategy: software controllability
- Addressable On-Chip Memory in addition to conventional cache
  - On-Chip Memory and cache are reconfigurable
- Explicit data transfer between On-Chip Memory and Off-Chip Memory by page-load/page-store instruction
  - burst transfer and stride transfer are supported

advantages of SCIMA

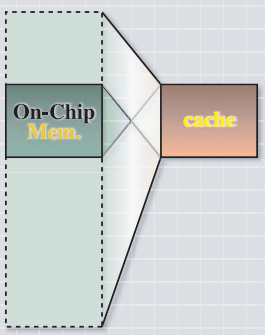
T_b: CPU busy time
T_l: Latency stall time
T_t: Throughput stall time

On-Chip Memory Features	T_b	T_l	T_t
software controllability	–	–	↓
page-load/page-store(burst)	↑	↓	–
page-load/page-store(stride)	↑	↓	↓
scheduling for page-load/page-store	–	↓	–
Latency Tolerating Techniques of Cache	T_b	T_l	T_t
larger cache line	–	↓	↑
lock-up free cache	–	↓	↑
cache prefetching	↑	↓	↑

Schematic View



Overview of SCIMA



Address Space