An Introduction to Basic Research on High-End Processor in China

Guo-jie Li

Institute of Computing Technology, CAS

June 14, 2010

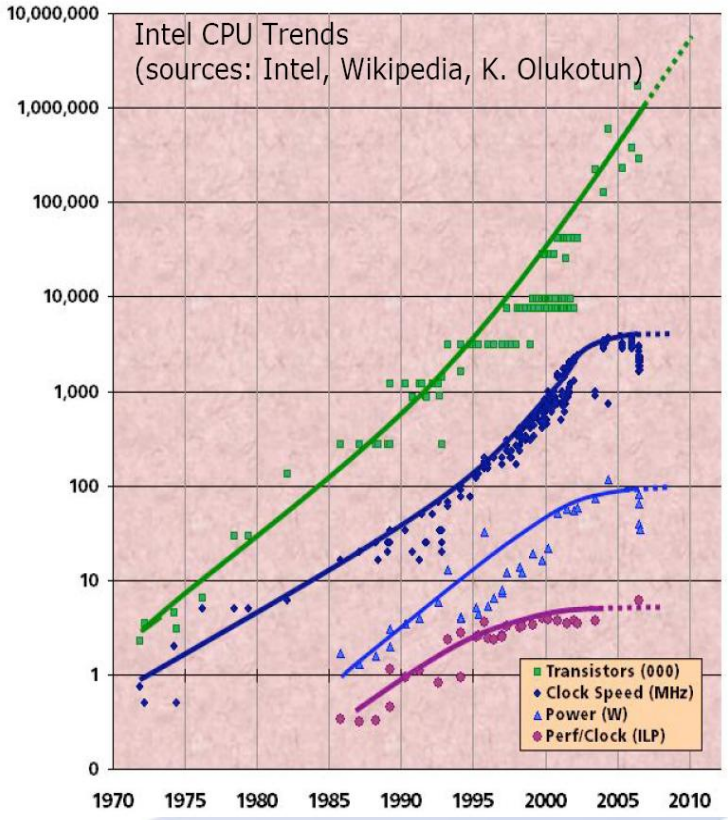
中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES

Agenda

- Background
- Our Research
  - Research Work Overview
  - Recent Papers
- Future Work

Processor Scaling Mismatches Moore's Law

- Moore's Law is still in effect
- But processor microarchitecture does not scale well
  - Frequency scaling ends
  - Aggressive ILP exploitation becomes inefficient
  - Power dissipation meets limitation
- → **Parallel Microprocessor**



New Theory, New Architecture and New Methodology to Scale Microprocessor Design

- China Basic Research Project

  No. 2005CB321600 (fund: ¥ 35 million)

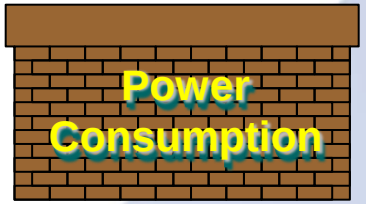
- Problems Formulation:

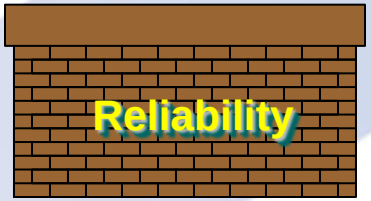
  3 Technology Walls for Future High-End Microprocessor

  – Information system **complexity**

  – Information processing **energy consumption**

  – **Reliable** information system



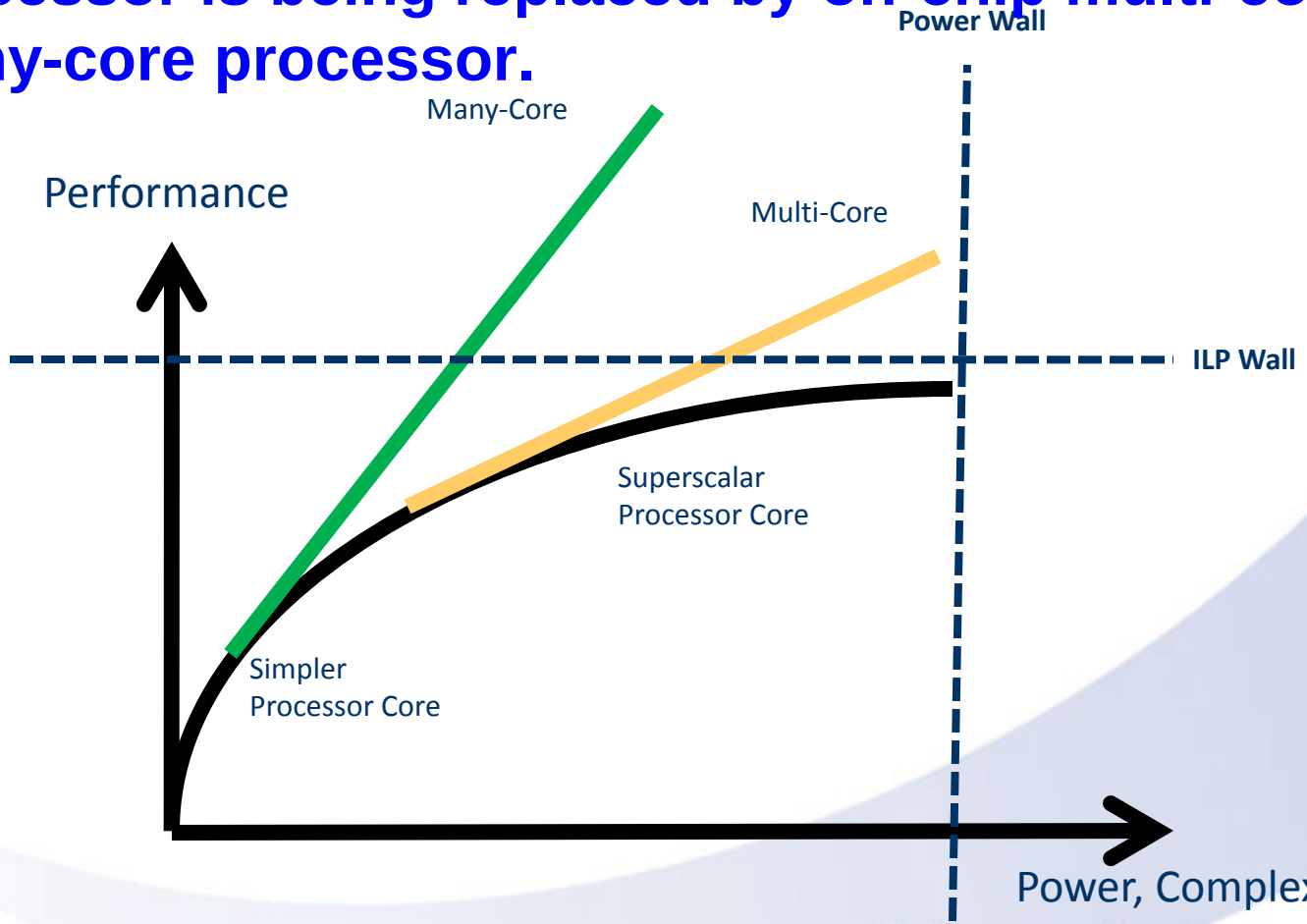


Agenda

- Background
- Our Research
  - **Research Work Overview**
  - Recent Papers
- Future Work

Microprocessor Architecture Revolution

- **Limited by power and complexity, superscalar processor is being replaced by on-chip multi-core / many-core processor.**

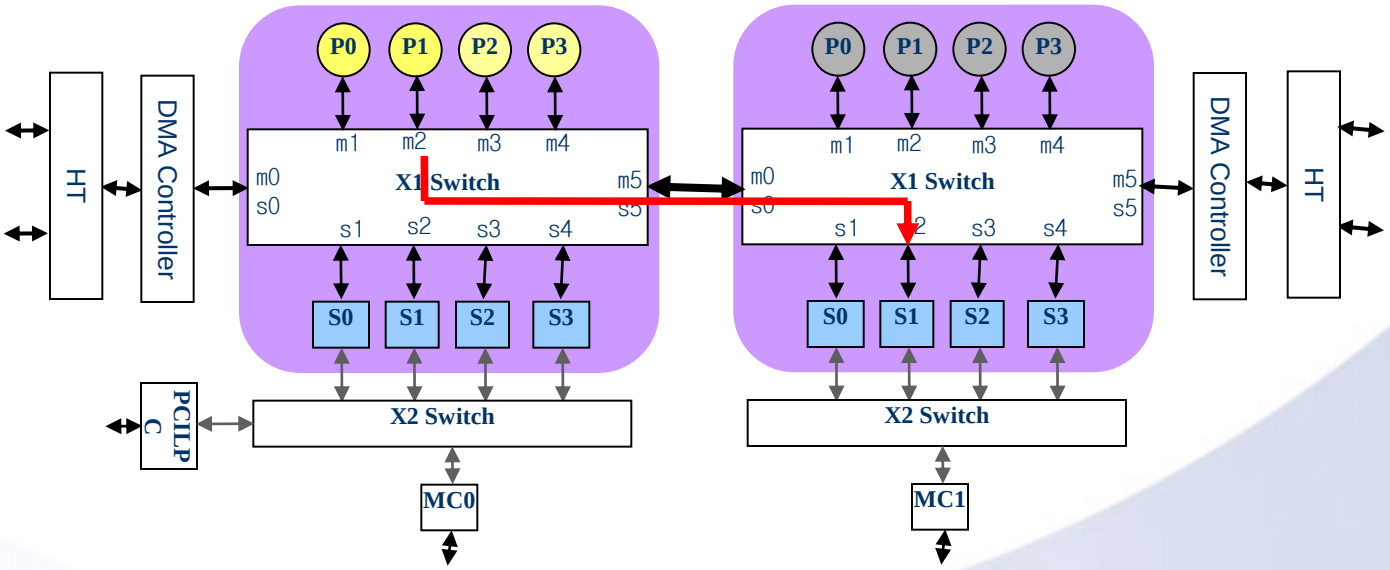


Overview of Our Research Work

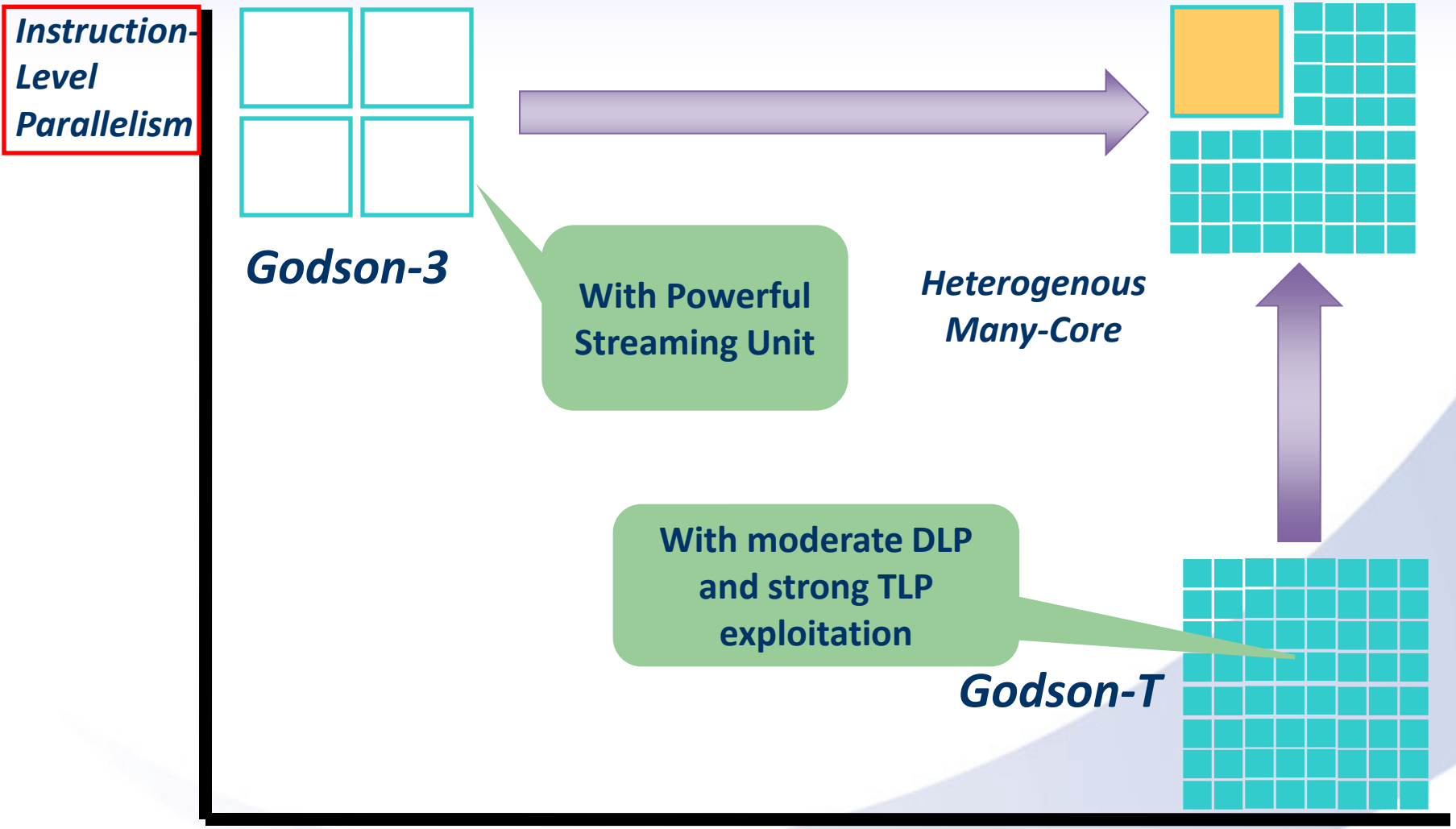
- Concluded as “1+2+N”
  - ***1***：scalable, configurable parallel microarchitecture
  - ***2***：multi-heavy-core (Loongson-3B)
    many-light-core (Godson-T)
  - ***N***：a number of new theory, new architecture and new methodology to address the technology walls

Scalable and Configurable TGAP Architecture (Tera-op Godson Architecture Prototype)

- Scalable parallel microarchitecture
- Configurable on-chip network and memory
- Flexible resource isolation



Exploiting Polymorphic Parallelism

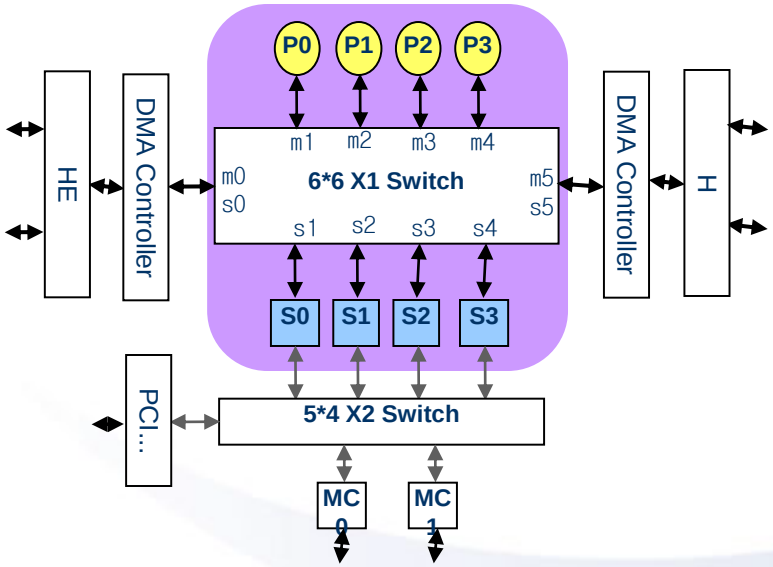


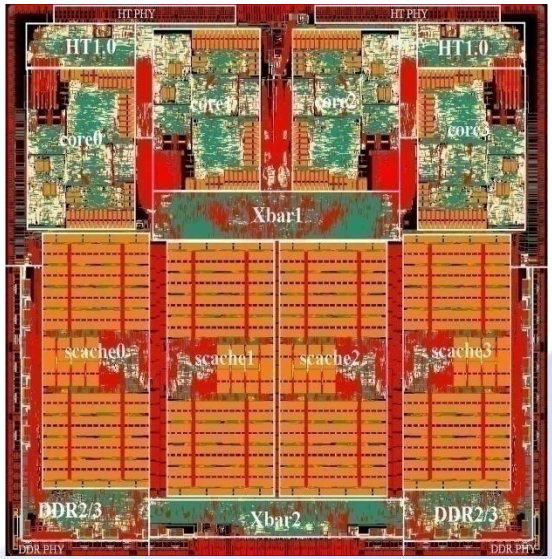
Two High-Performance Microprocessor Prototypes

- loongson3: research on efficient heavy-core technology and nanometer IC design, test and verification
  - Loongson 3A：4-core，65nm，1GHz，16GFLOPS
  - Loongson 3B：8-core，65nm，1GHz，128GFLOPS
  - Loongson 3C：16-core，28nm，1.5GHz，384GFLOPS，tape out in 2011
  - Godson-T: research on massively-parallel computing technology
  - 64 tiles，4 memory controllers，sample(16 tiles) taped out now.
  - New technologies to handle memory latency and bandwidth, reliability, testability, etc.

Loongson-3A

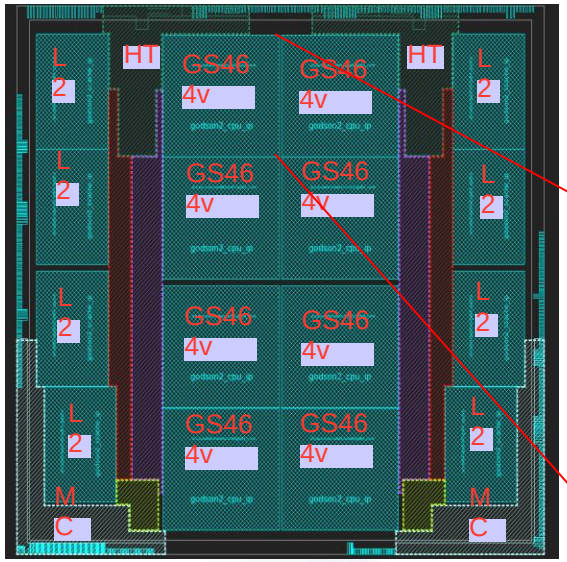
- 4 four-issue 64-bit heavy-core in a node;
- High throughput I/O and memory controller are integrated;
- 65nm, 425 million transistors, 174mm², 1GHz, 10-watt

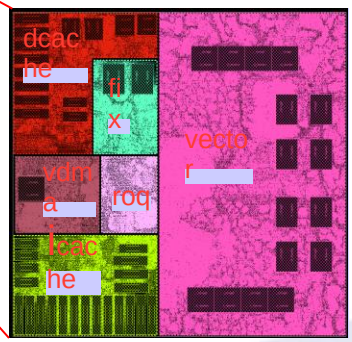


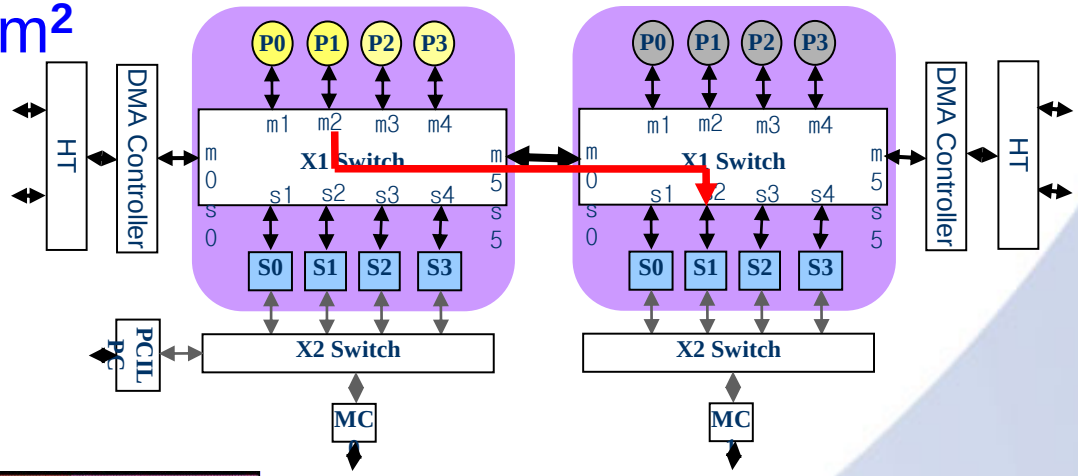


Loongson-3B

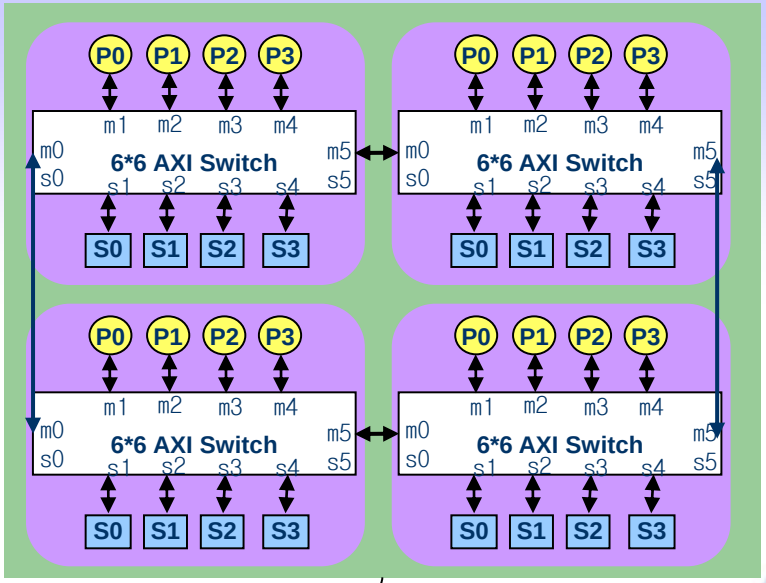
- 2 nodes, total 8 cores
- 128 GFLOPS provided by powerful streaming unit
- 65nm, 1GHz, 300mm²

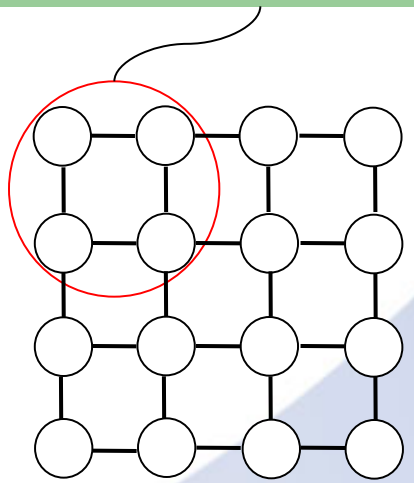






Loongson-3C

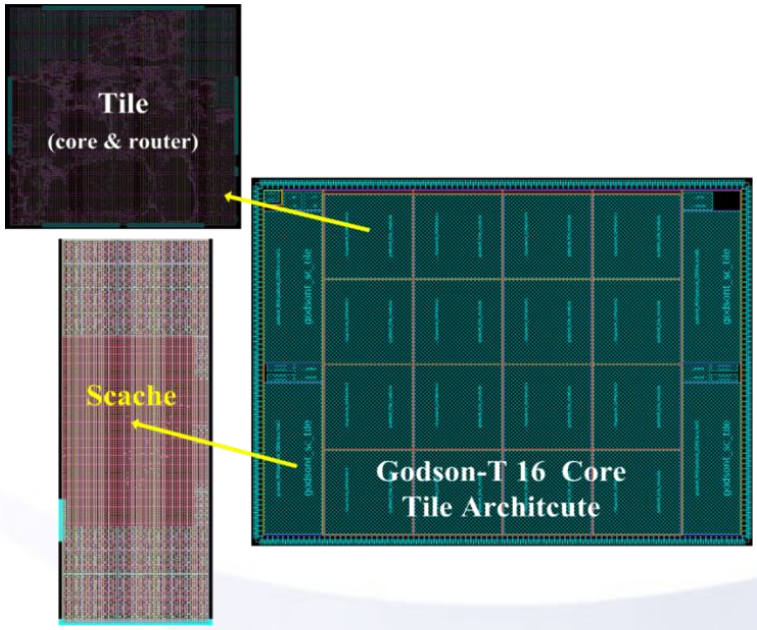


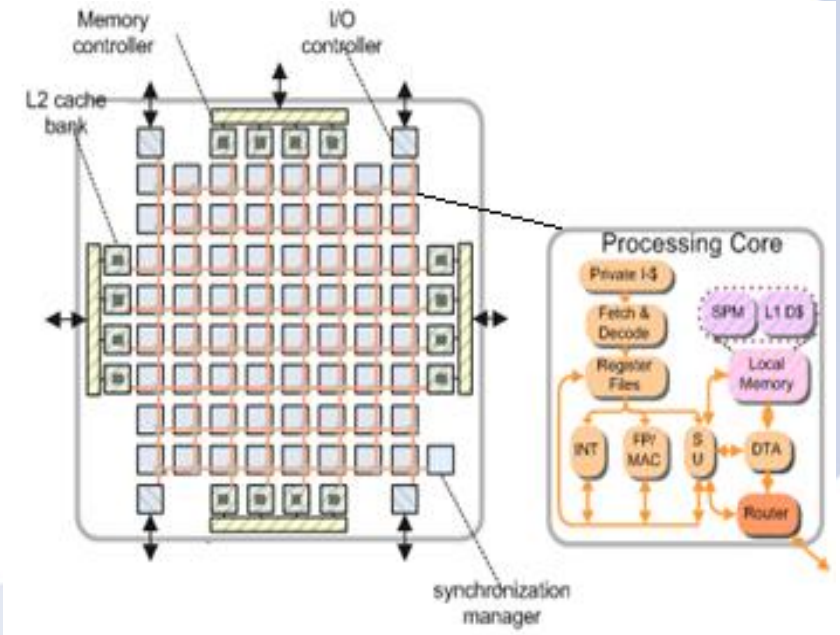


- 4 nodes, 16 cores
- 128 64-bit FMADD units
- 1.5 GHz@32nm, 200mm^2
- Peak Performance
  - **DP: 384 GFLOPS**
  - **SP: 768 GFLOPS**
  - **16-bit: 1.5TOPS**
  - **8-bit: 3TOPS**
- Target for computation-intensive applications, such as scientific computation, visual computation, ...

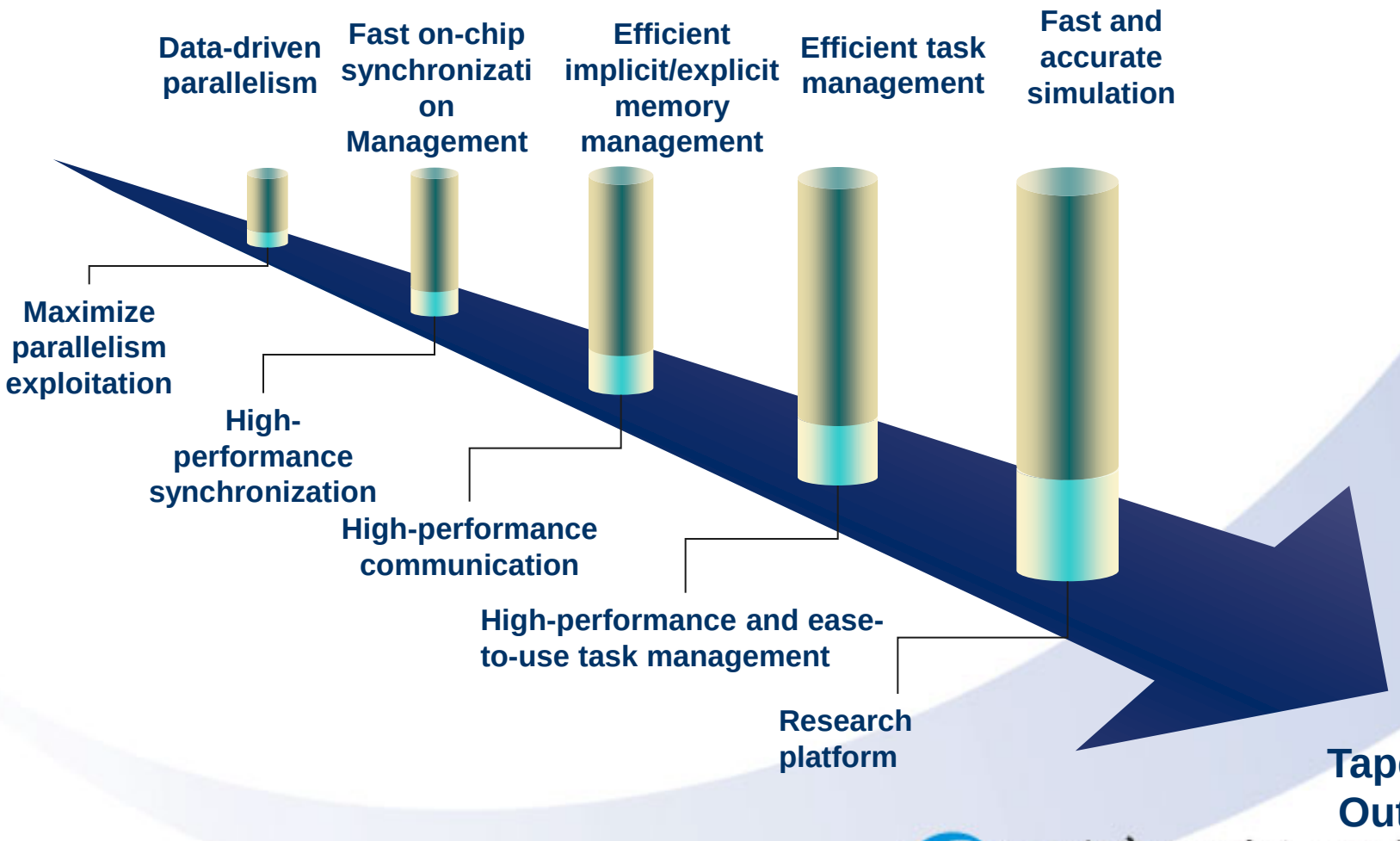
Godson-T Many-Core Prototype

- 64 light-weight processing tiles currently,
  256 tiles in 2015
- 16-tile sample:130nm, 230mm^2 , SMIC
- Target for domain-specific parallel acceleration



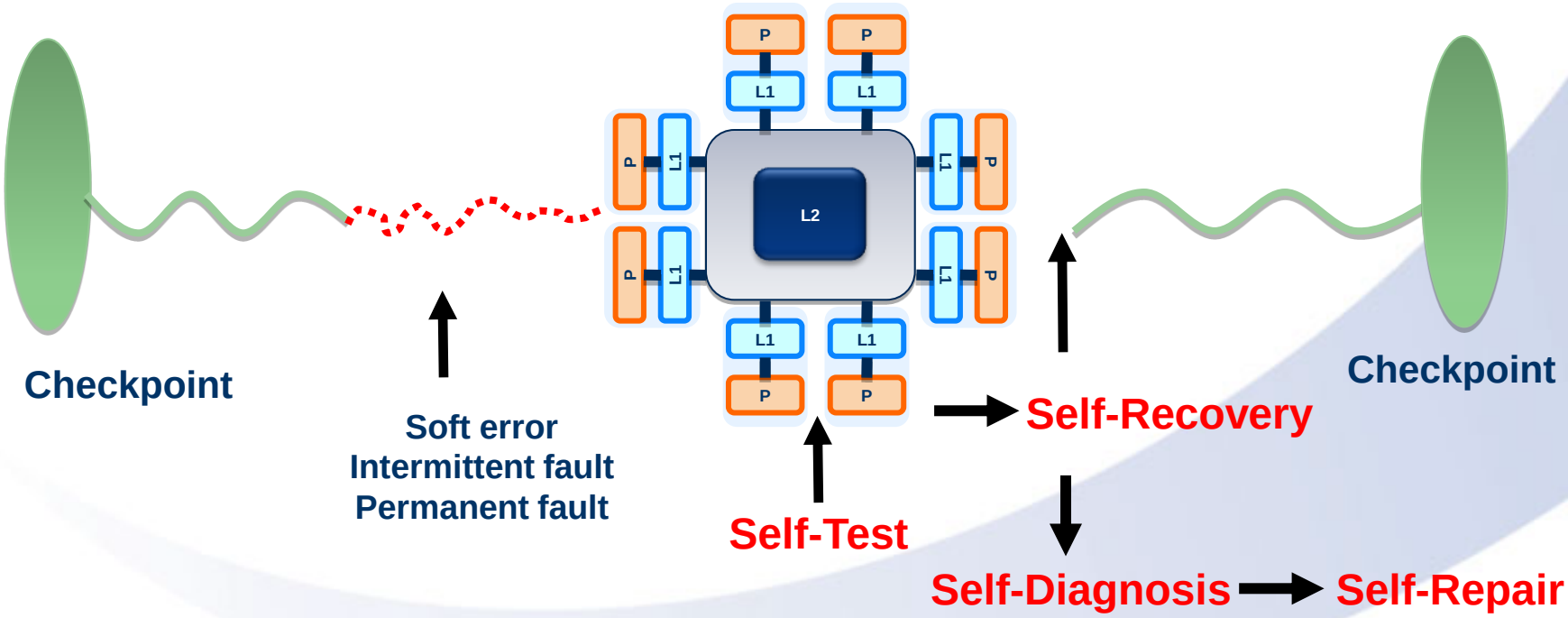


Godson-T Design and Implementation

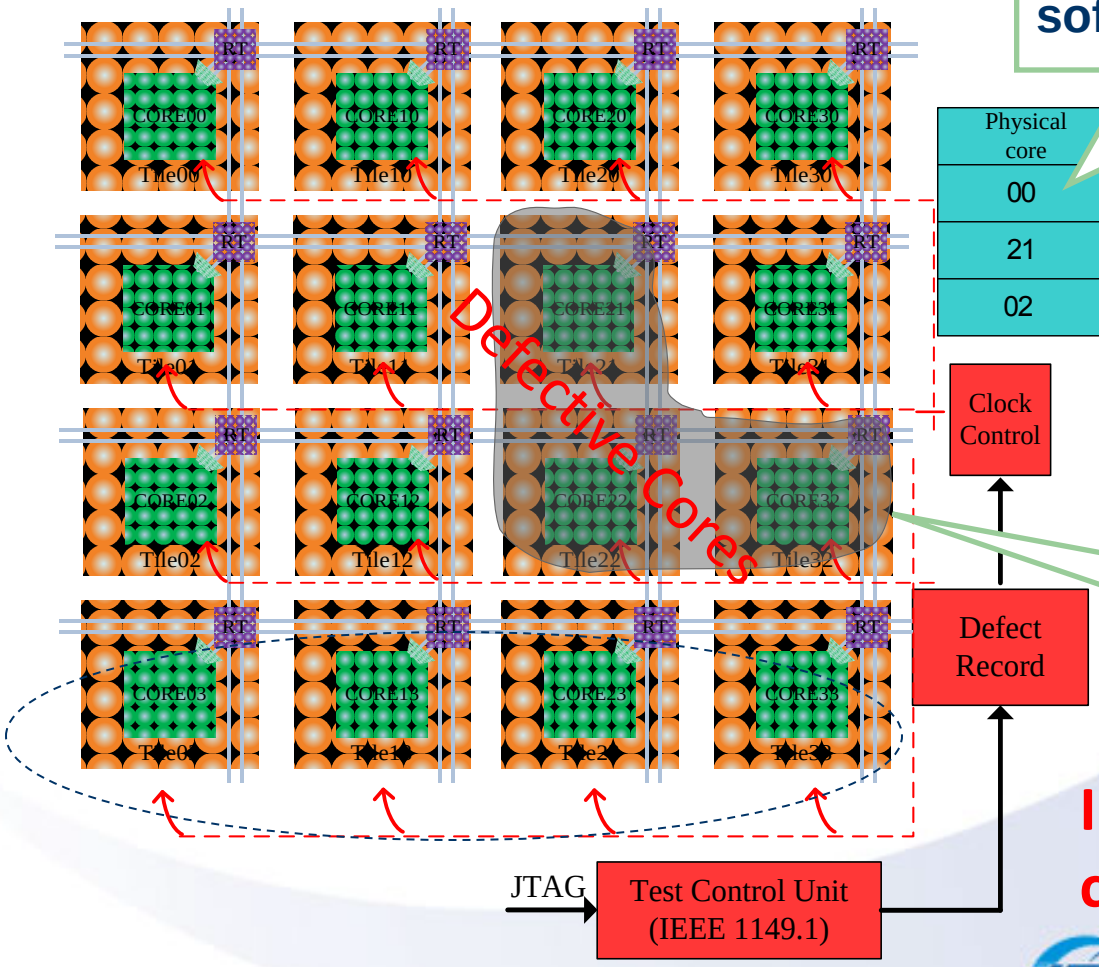


3S: Self-Test, Self-Diagnosis, Self-Repair

- Reliable Design for Future Parallel Architecture
- Component failure will be normal in many-core processor, we use 3S to tolerate variety of abnormality including soft errors, intermittent faults and permanent faults. Keep service with 3S!



3S in Godson-T 16-tile sample



Efficient mapping from virtual cores (seen from software) to physical cores

Physical core	Virtual Core
00	00
21	NULL
02	01

Clock Control

Defect Record

Repair defective cores with redundant cores

Increasing Godson-T chip yield through 3S

Agenda

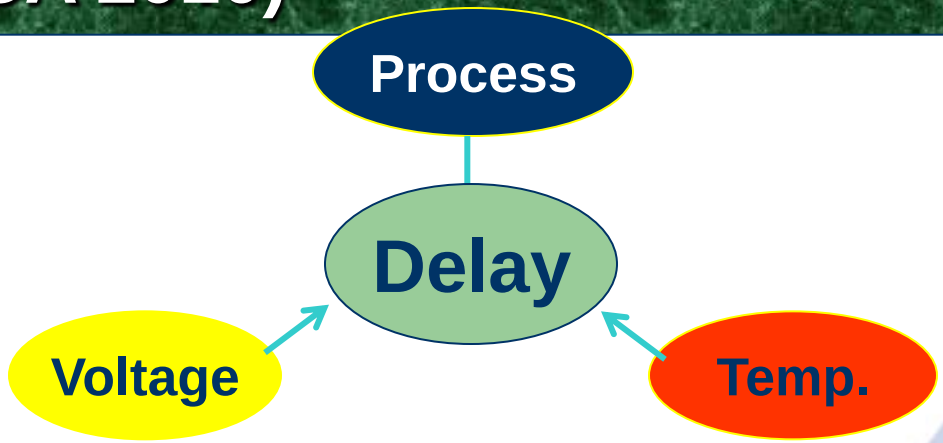
- Background
- Our Research
  - Research Work Overview
  - Recent Papers
- Future Work

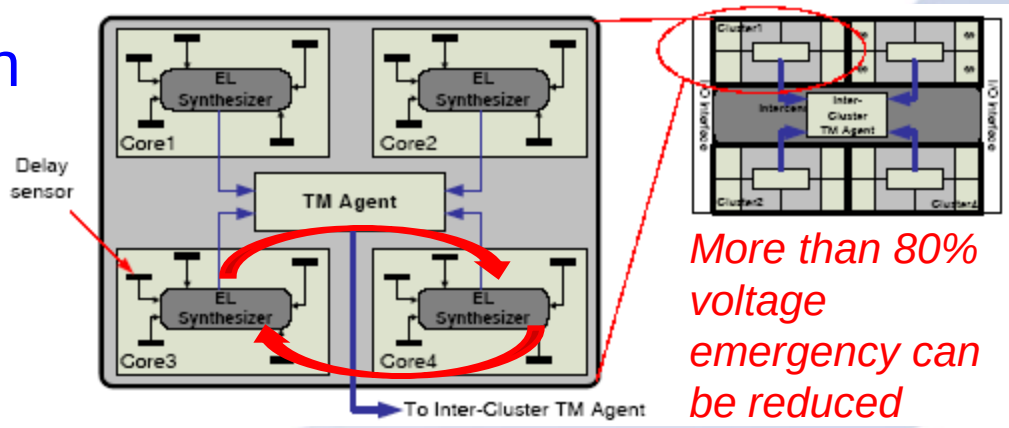
Recent Papers – Partial Representative List

- Co-optimizing Process, Voltage, and Temperature (PVT) variations in Multi-core Processor . ISCA 2010.
- LReplay: A Pending Period Based Deterministic Replay Scheme. ISCA 2010.
- DMA Cache: Using On-Chip Storage to Architecturally Separate I/O Data from CPU Data for Improving I/O Performance, HPCA 2010.
- Evaluating Iterative Optimization Across 1000 Data Sets. PLDI 2010.
- High Performance Comparison-Based Sorting Algorithm on Many-Core GPUs . IPDPS 2010.
- Fast Complete Memory Consistency Verfication, HPCA 2009.
- On Topology Reconfiguration for Defect-Tolerant NoC-Based Homogeneous Manycore Systems. IEEE Trans. On VLSI 2009.
- A unified online Fault Detection scheme via checking of Stability Violation. DATE 2009.
- Single-particle 3D Reconstruction from Cryo-Electron Microscopy Images on GPU. ICS 2009.
- Deterministic Diagnostic Pattern Generation (DDPG) for Compound Defects. ITC 2008.
- HMTT: A Platform Independent Full-System Memory Trace Monitoring System. SIGMETRICS 2008.
- A Parallel Dynamic Programming Algorithm on a Multi-core Architecture. SPAA 2007.
- ……

Co-optimizing Process, Voltage, and Temperature (PVT) variations in Multi-core Processor (ISCA 2010)

- First work to model process, voltage and temperature variations as a uniform delay variation.
- Efficiently address the fault prediction, detection and diagnosis problems.
- Employ thread migration to prevent 50% voltage emergencies.

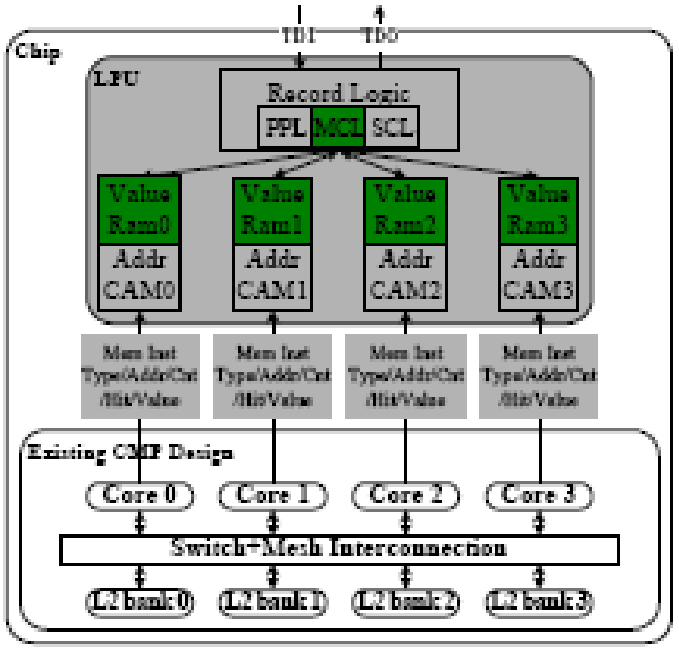




More than 80% voltage emergency can be reduced

LReplay: A Pending Period Based Deterministic Replay Scheme (ISCA 2010)

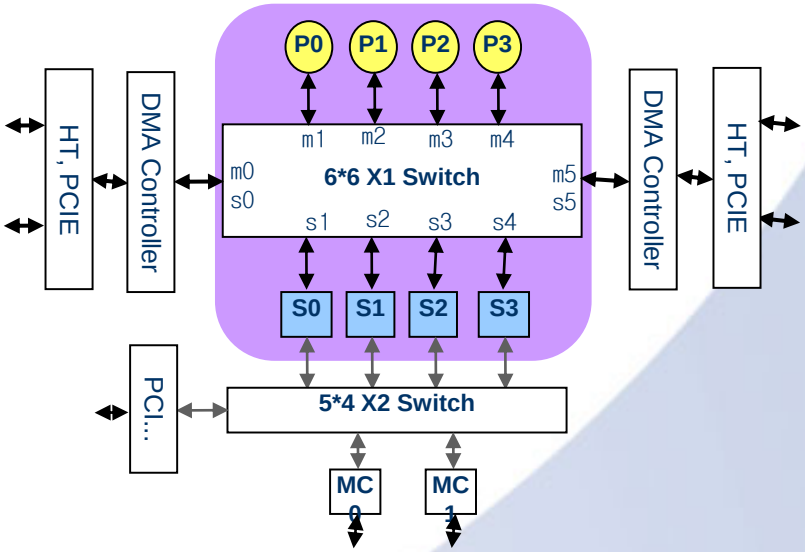
- Hardware-based deterministic replay scheme facilitating global clock in a chip
- Result
  - Overall log size of LReplay Only generate 0.55B per kilo-instructions,
  - Very low hardware cost and easy to implement



DMA Cache: Using On-Chip Storage to Architecturally Separate I/O Data from CPU Data for Improving I/O Performance (HPCA 2010)

- High-throughput computing becomes more popular
  - Computation-centric → Memory-centric →

    I/O-centric
- I/O-centric microarchitecture
  - Data is transferred between CPU and I/O via on-chip cache rather than memory
  - Applied to Looongson-3 chip, performance of SSD disk is improved by 40%



Fast Complete Memory Consistency Verification (HPCA 2009)

- Propose a novel method that largely reduces the time complexity of processor memory consistency verification.
  - Verification of 16-core Loongson-3 processor becomes feasible.

Fast Complete Memory Consistency Verification

Yunji Chen[1], Yi Lv[2], Weiwu Hu[1]*, Tianshi Chen[3], Haihua Shen[1], Pengyu Wang[1], Hong Pan[2]

[1]Key Laboratory of Computer System and Architecture
Institute of Computing Technology, Chinese Academy of Sciences
Beijing 100190, P. R. China

[2]State Key Laboratory of Computer Science
Institute of Software, Chinese Academy of Sciences
Beijing 100190, P. R. China

[3]Department of Computer Science and Technology
University of Science and Technology of China
Hefei, Anhui 230027, P. R. China

Abstract

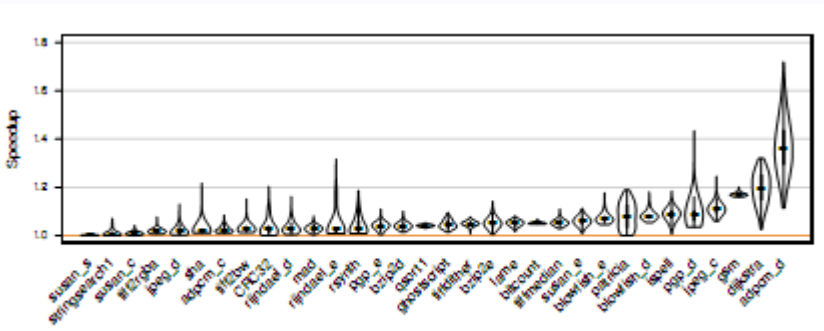
The verification of an execution against memory consistency is known to be NP-hard. This paper proposes a novel fast memory consistency verification method by identifying a kind of natural partial order: time order. In multi-processor system with store atomicity, time order restriction exists between two operations whose pending periods are disjoint: the former operation in time order must be observed by the later operation. Based on time order restriction, memory consistency verification is localized: for any operation, both inferring related orders and checking related cycle only need to take bounded operations into account.

ory in multi-processor system, memory subsystem must spend many resources on supporting memory consistency and cache coherence. Therefore, Memory consistency verification is an indispensable part of verifying memory subsystem.

Researchers have found that the verification of an execution against memory consistency is NP-hard with respect to the number of memory operations [3, 8]. To cope with the problem in practice, there are two kinds of solutions: micro-architecture dependent methods which exploit the help of extra observability in design to bring down the complexity [22, 25, 26], and micro-architecture independent methods which devise polynomial time algorithms that are sound but not necessarily complete [12, 23, 24, 30]. However, micro-

Evaluating Iterative Optimization Across 1000 Data Sets (PLDI 2010)

- Most iterative optimization studies find the best optimizations on the same data set, which prevents its usage in practice.
- We evaluated the effectiveness of iterative optimization across a large number (1000) of data sets.
- **Conclusion**
  - **Compiler optimizations achieves 86% speedup than ICC (83% for GCC)**
  - **optimizing programs across data sets much easier than previous anticipation**



(a) sorted by average data set-optimal speedup

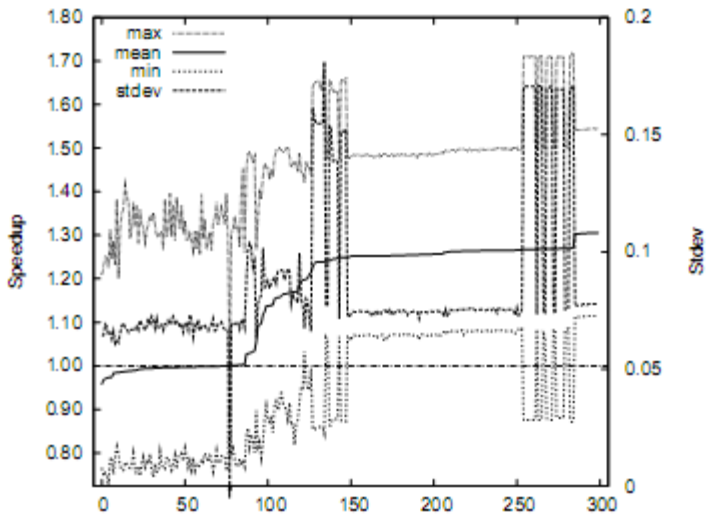


Figure 4. *Reactions to compiler optimizations (adpcm_d).*

High Performance Algorithm based on Novel Architecture (SPAA 2007)

- **Optimized Dynamic Programming Algorithm is cited by Professor Vigaya Ramachandran (SPAA 2008 and Phd. Thesis directed by him)**

Cache-efficient Dynamic Programming Algorithms for Multicores *

Rezaul Alam Chowdhury
The University of Texas at Austin
Department of Computer Sciences
Austin, TX 78712
shaikat@cs.utexas.edu

Vijaya Ramachandran
The University of Texas at Austin
Department of Computer Sciences
Austin, TX 78712
vlr@cs.utexas.edu

…A parallel algorithm for solving the parenthesis problem which…, but the algorithm is not cache-efficient. A cache-efficient multi-core algorithm for the IBM Cyclops64 processor was given in [25]

oblivious algorithm based on Valiant's context-free language recognition algorithm [27] was given for solving the recurrence. A parallel algorithm for solving the parenthesis problem which runs in $\mathcal{O}\left(n^{\frac{3}{4}} \log n\right)$ time and performs optimal $\mathcal{O}\left(n^3\right)$ work was given in [15], but the algorithm is not cache-efficient. A cache-efficient multicore algorithm for the IBM Cyclops64 processor was given in [25].

Algorithms and Data Structures for Cache-efficient Computation: Theory and Experimental Evaluation

by

Rezaul Alam Chowdhury, B.Sc.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

…A similar algorithm for simple-DP is also given in [117], and in [118] the algorithm is extended for cache-efficient execution on a multicore programming model based on IBM Cyclops64

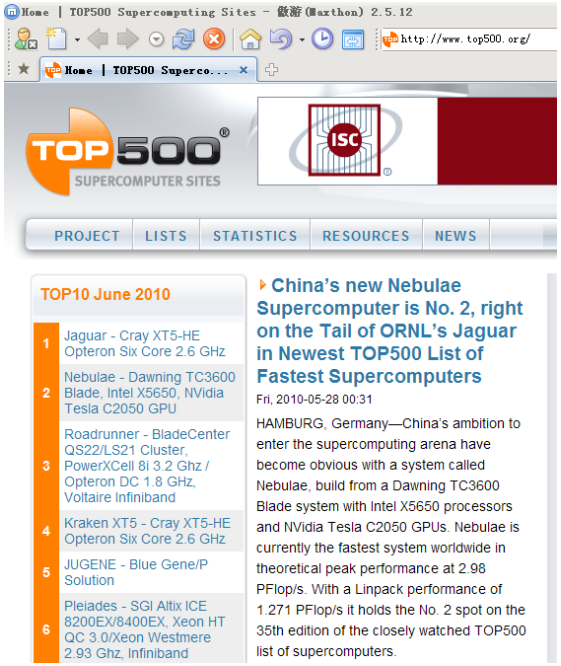
cludes algorithms for secondary structure prediction [78], matrix chain multiplication, optimal polygon triangulation and optimal binary search tree construction. A similar algorithm for simple-DP is also given in [117], and in [118] the algorithm is extended for cache-efficient execution on a multicore programming model based on IBM Cyclops64.

The cache-oblivious stencil computation technique presented in [54] can be used as a dynamic programming algorithm for computing the length of a longest common subsequence of two sequences of length n each in $\mathcal{O}\left(n^2\right)$ time, $\mathcal{O}\left(n\right)$ space and $\mathcal{O}\left(\frac{n^2}{BM}\right)$ I/Os. This method, however, does not compute the subsequence.

Research Impact on Top-Tier Conferences

Top-Tier Conference	Acceptance Rate	# of Mainland Publishing	# of ICT Publishing	Proportion	
ISCA	18%	9	5	56%	
HPCA	18%	3	2	67%	First in Mainland
SC	25%	4	3	75%	First in Mainland
SPAA	30%	1	1	100%	First in Mainland
ICS	26%	7	5	71%	First in Mainland
SIGMETRICS	18%	4	3	75%	First in Mainland
PLDI	20%	3	1	33%	

International Impact of ICT



Dawning Nebulae Supercomputer holds No. 2 on Top 500
1.271 Petaflops

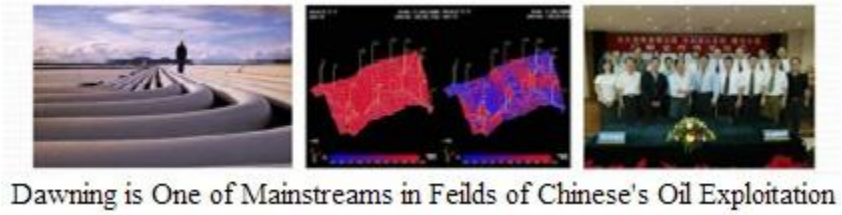
Godson-3 introduction on "Microprocessor"

Godson-3 on MIT "Technology Review"

中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES

Technology Transfer

- **Dawning supercomputers**
  - Occupied 27% in TOP100 Supercomputer in China (IBM 26%)
  - Contributions on oil exploration, national security, …
- **Loongson microprocessors**
  - Set up Loongson Corp. (initial capital: about $ 30M)
  - Low-cost PC in China sell 150K units in Jiangsu



Dawning is One of Mainstreams in Feilds of Chinese's Oil Exploitation

Dawning 4000 helps to fly Shenzhou Spaceship

Agenda

- Background
- Our Research
  - Research Work Overview
  - Recent Papers
- **Future Work**

HPC versus HTC

Conventional

FLOPS

Now

Throughput

Conventionally, HTC systems were implemented by HPC infrastructures. As requirements of throughput, energy-efficiency, scalability and reliability are increasing in emerging HTC systems, conventional wisdom is no longer suitable for next-generation data-center computing.

Expected Contribution: HTC Microprocessor

- **Kilo-Thread running on a chip**
  - To apply a new 973 project

Processor	# of Threads	Microarchitecture
AMD Opteron	12	12 cores
Intel Xeon	8	2~8 cores
Intel SCC	48	48 cores
IBM Power7	32	8cores, 4-way multithreading
IBM Wire-Speed Processor	64	16 cores. 4-way multithreading, special-purpose hardware acceleration
Sun UltraSparc T2	64	8 cores, 8-way multithreading
Tilera TILEPro	64	64 cores
HTC Processor of ICT	1024	1024 threads on a chip

Thanks

中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES