

Toward for Exa-scale and Beyond from Parallelizing Compiler Aspect

Hironori Kasahara

Professor, Dept. of Computer Science & Engineering

Director, Advanced Multicore Processor Research Institute

Waseda University (早稲田大学), Tokyo, Japan

IEEE Computer Society

President Elect 2017, President 2018

URL: <http://www.kasahara.cs.waseda.ac.jp/>

Waseda Univ. GCSC

Key Technologies

➤ **Performance:**

➤ **Multigrain Parallelization**

- **Hierarchical Coarse Grain Task Parallelization, Loop Parallelization and Vectorization**
- **Data Localization and Overlapping Data Transfer Using DMA**

➤ **Architecture:**

- **Many cores with Accelerators, DMAC (DTU) and Distributed Shared memory**
- **Global Address Space**
- **Hierarchical processor grouping**
- **3 Dimensional Integration of memory and TSV (Through Silicon Vias)**

➤ **Power:**

- **Compiler controlled DVFS including Clock Gating and Power Gating**
- **Non-volatile Memory is helpful for Power Gating**

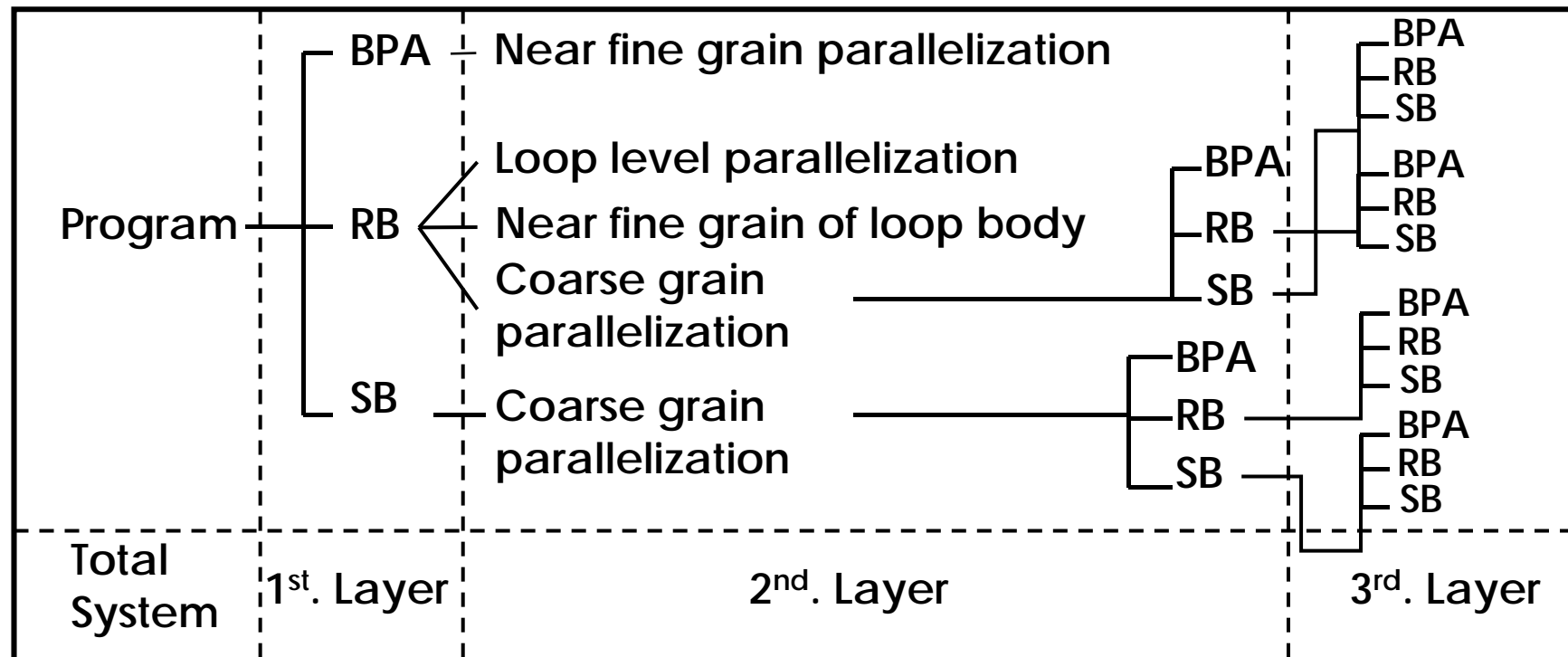
➤ **Programmability**

- **Automatic Parallelization by Compiler**
- **User's advices if compiler could not parallelize sufficiently**

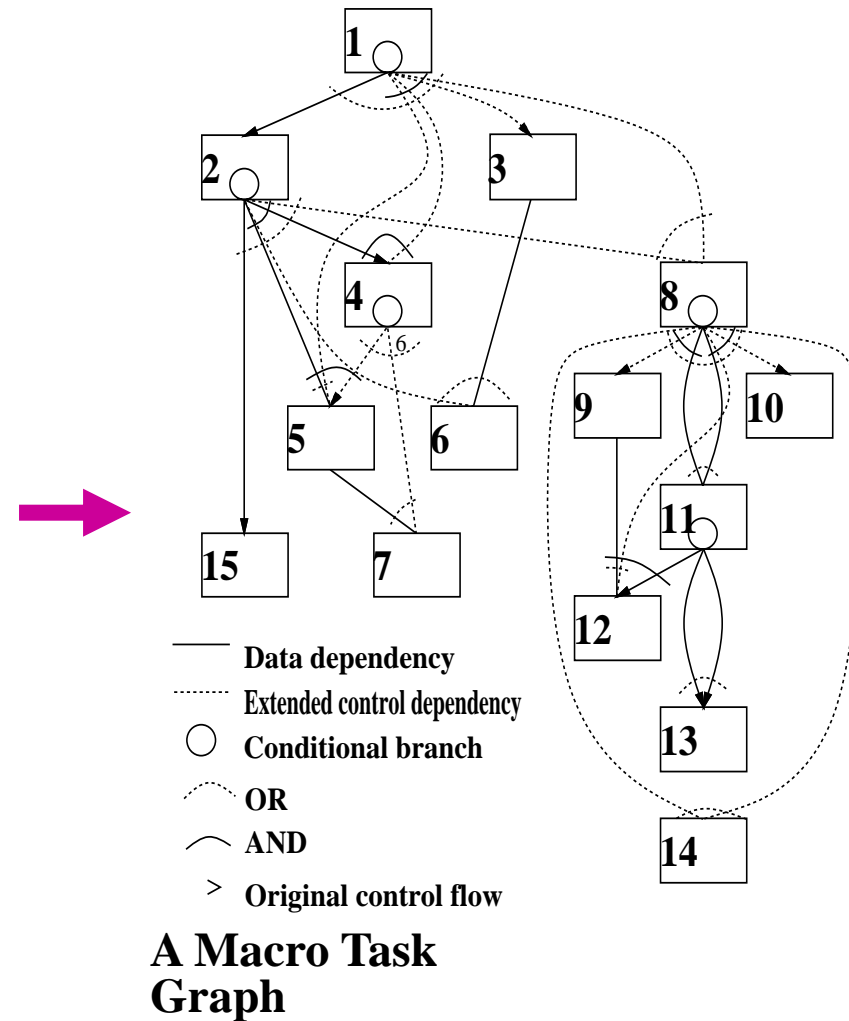
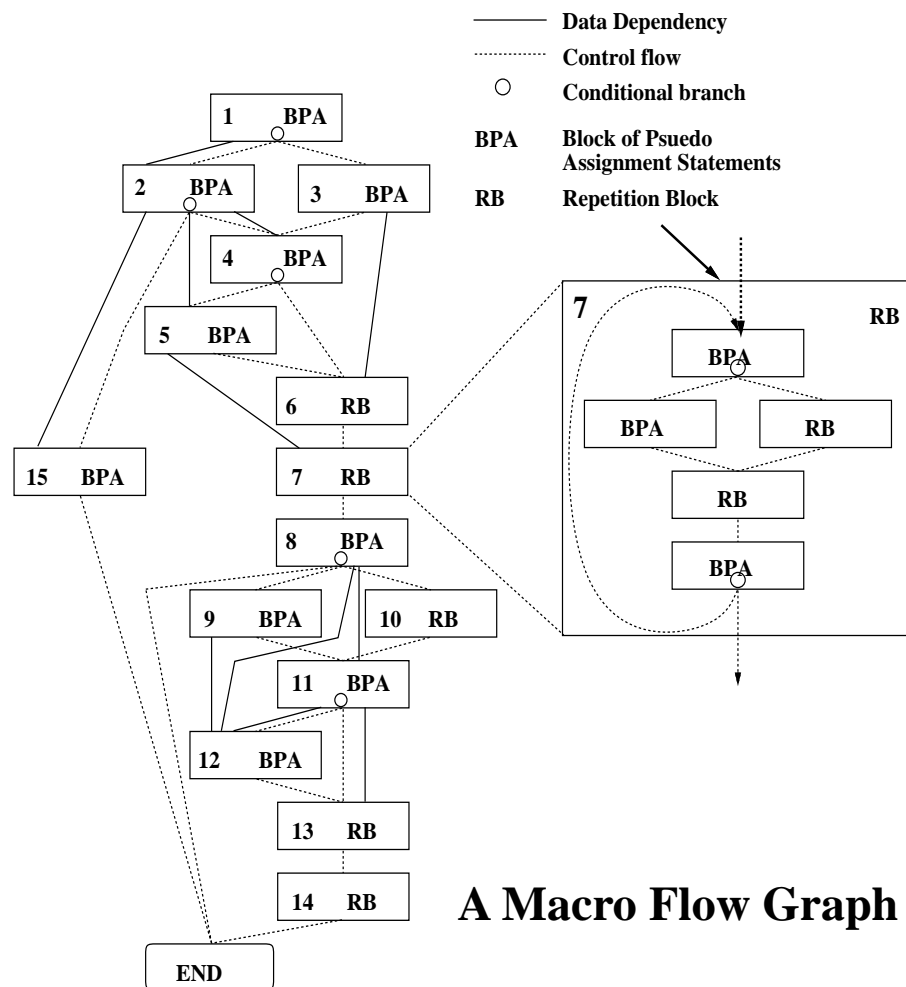
Generation of Coarse Grain Tasks

■ Macro-tasks (MTs)

- **Block of Pseudo Assignments (BPA): Basic Block (BB)**
- **Repetition Block (RB) : natural loop**
- **Subroutine Block (SB): subroutine**

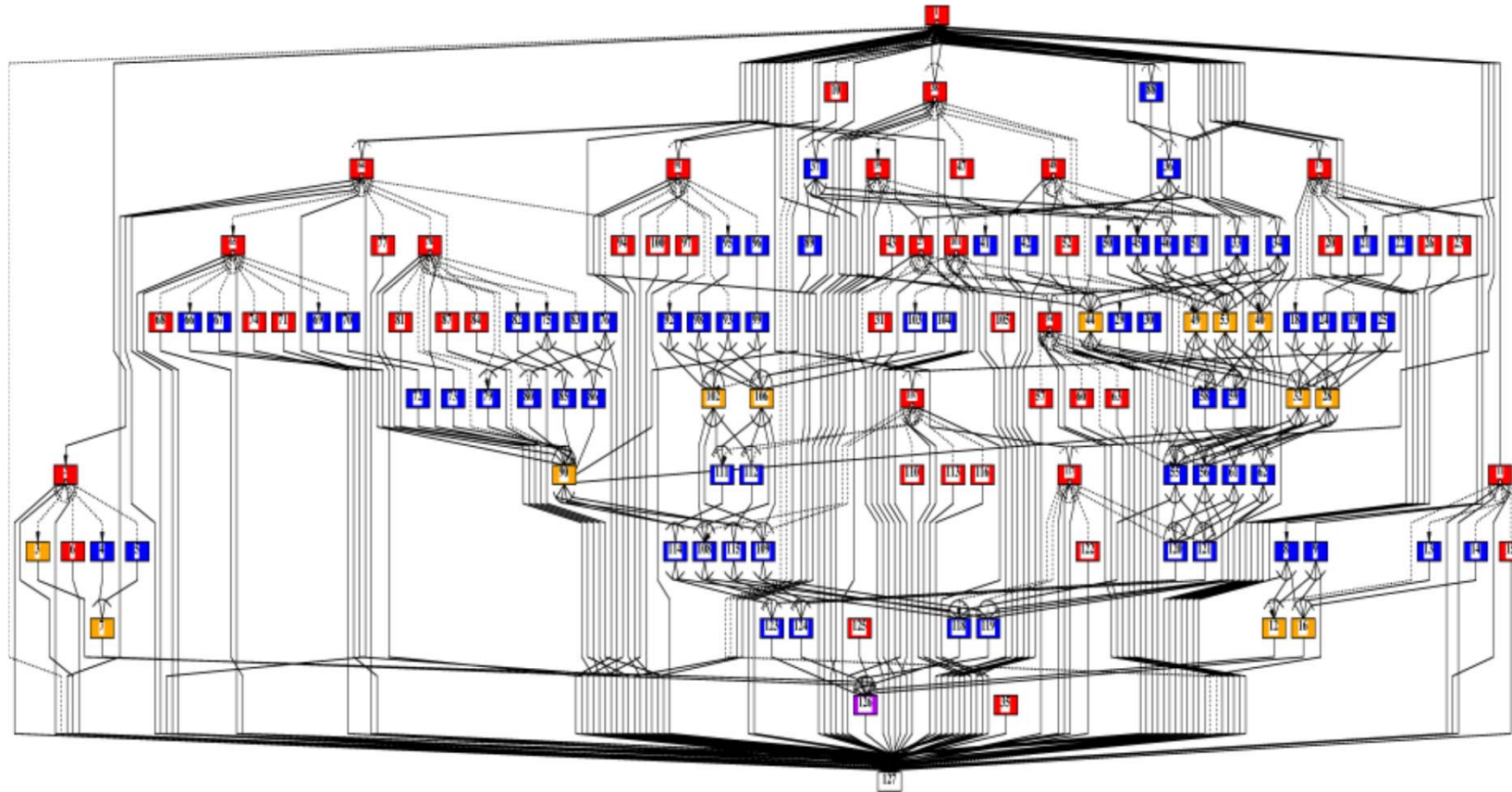


Earliest Executable Condition Analysis for Coarse Grain Tasks (Macro-tasks)



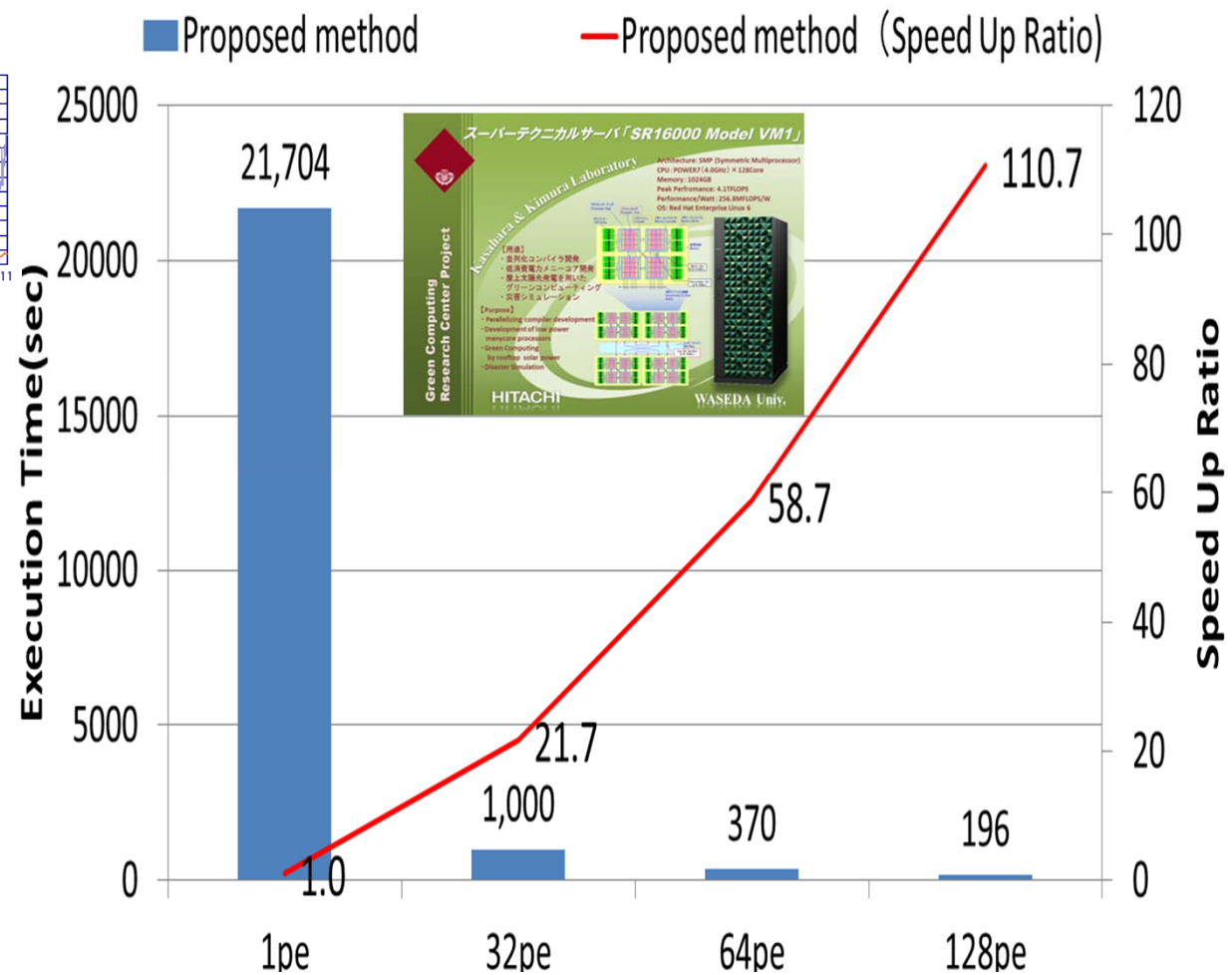
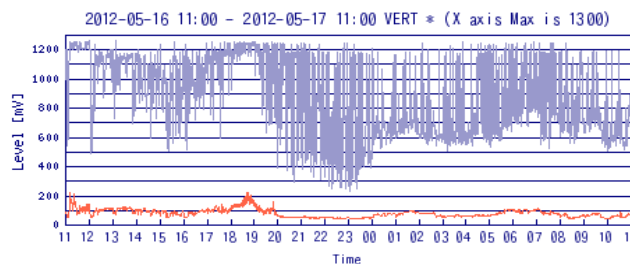
MTG of Su2cor-LOOPS-DO400

- Coarse grain parallelism $\text{PARA_ALD} = 4.3$

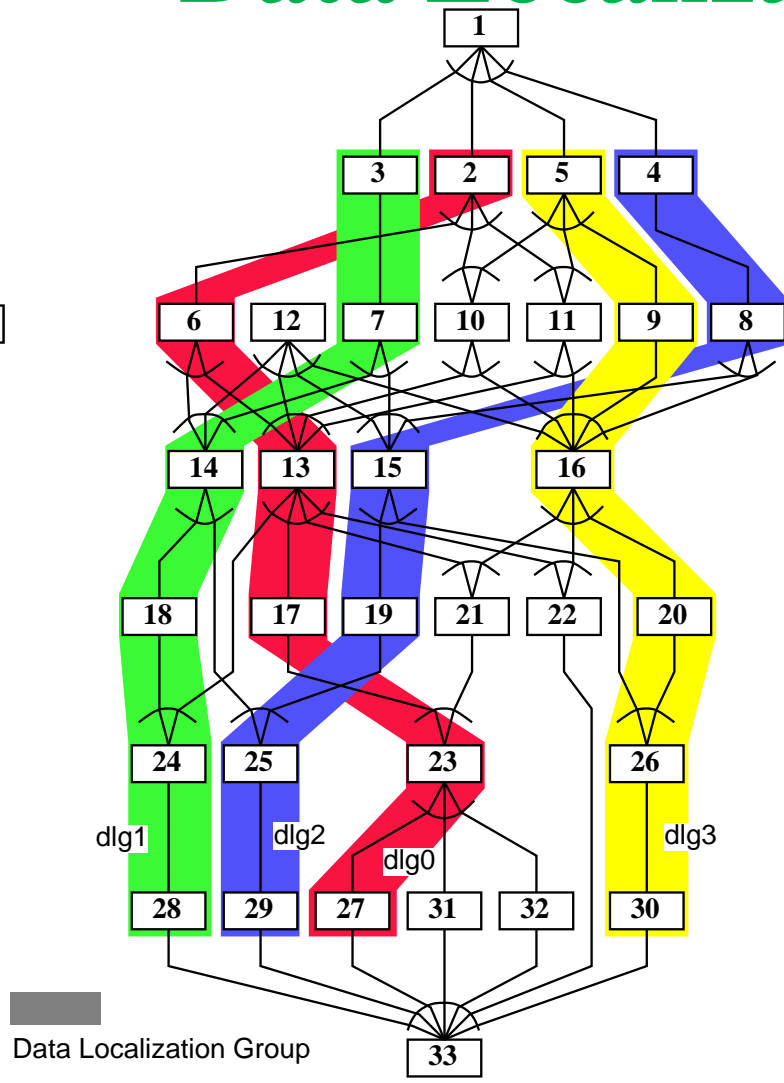
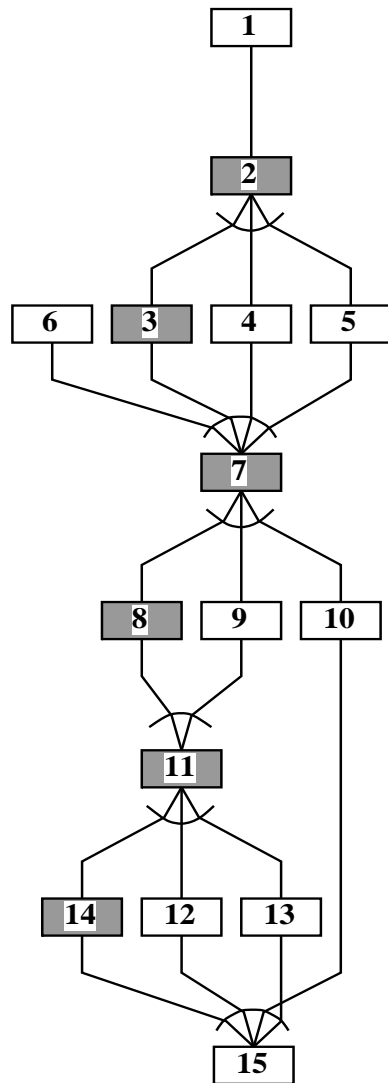


■ DOALL ■ Sequential LOOP ■ SB ■ BB

110 Times Speedup against the Sequential Processing for GMS Earthquake Wave Propagation Simulation on Hitachi SR16000 (Power7 Based 128 Core Linux SMP)



Data Localization

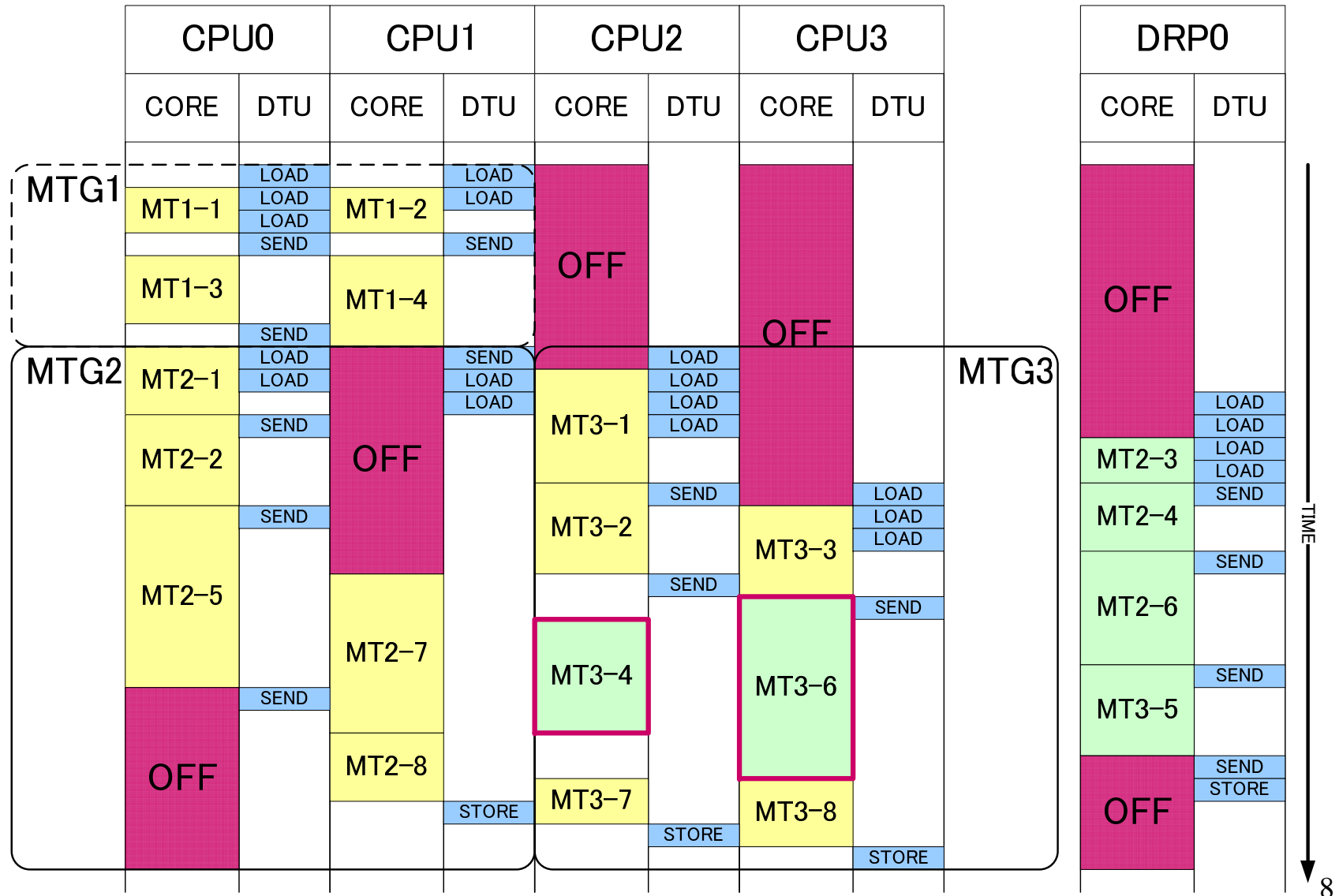


■ Data Localization Group

PE0	PE1
12	1
2	3
6	7
4	14
8	18
15	5
19	9
25	11
29	10
13	16
17	20
22	26
21	30
23	24
27	28
	32
	31

A schedule for two processors

An Image of Static Schedule for Heterogeneous Multi-core with Data Transfer Overlapping and Power Control



Key Technologies

➤ **Performance:**

➤ **Multigrain Parallelization**

- **Hierarchical Coarse Grain Task Parallelization, Loop Parallelization and Vectorization**
- **Data Localization and Overlapping Data Transfer Using DMA**

➤ **Architecture:**

- **Many cores with Accelerators, DMAC (DTU) and Distributed Shared memory**
- **Global Address Space**
- **Hierarchical processor grouping**
- **3 Dimensional Integration of memory and TSV (Through Silicon Vias)**

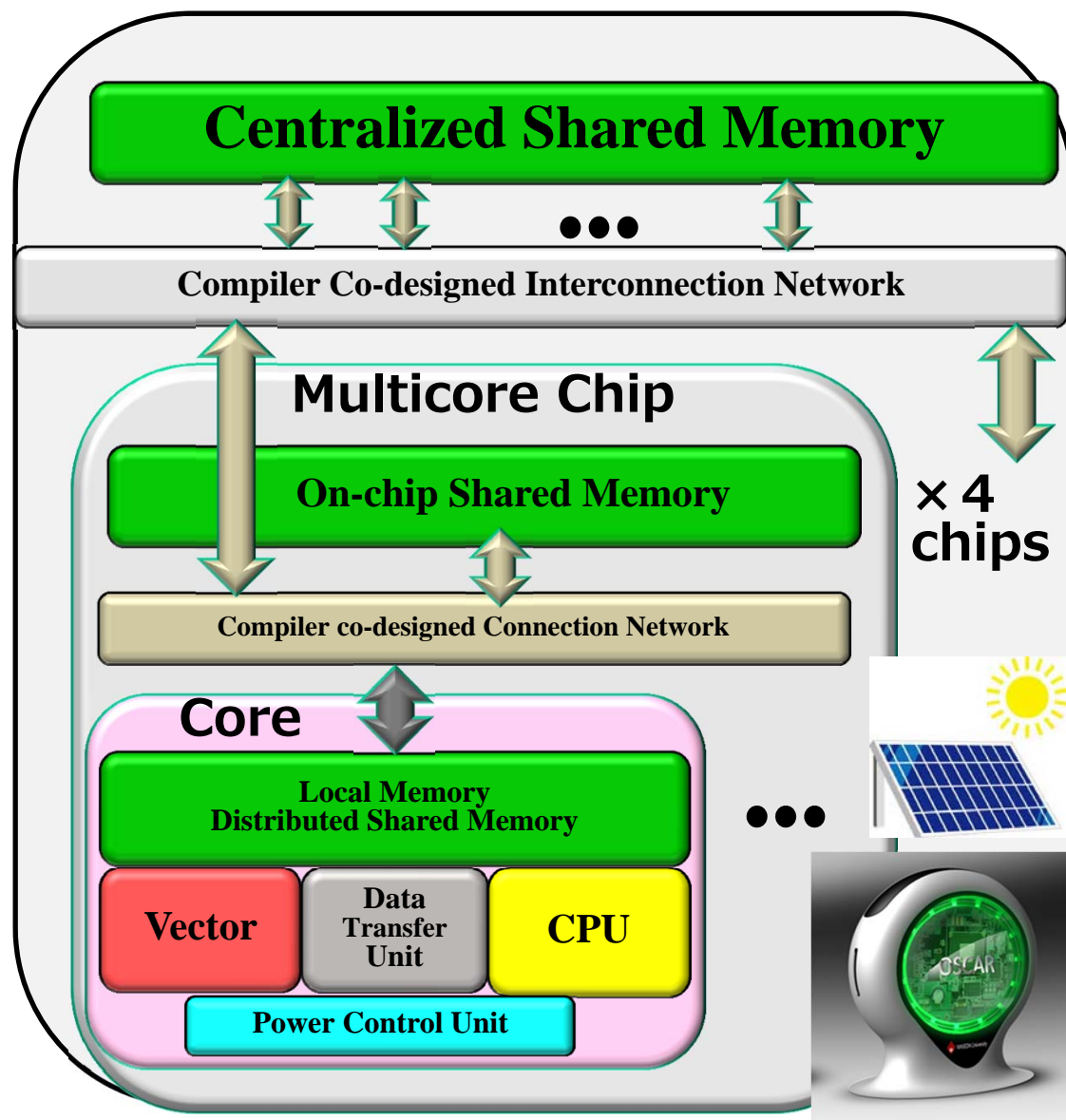
➤ **Power:**

- **Compiler controlled DVFS including Clock Gating and Power Gating**
- **Non-volatile Memory is helpful for Power Gating**

➤ **Programmability**

- **Automatic Parallelization by Compiler**
- **User's advices if compiler could not parallelize sufficiently**

OSCAR Vector Multicore to Support OSCAR Compiler's Parallelization and Power Reduction for Embedded to HPC.



- Compiler is designed to parallelize many applications.
- Next, hardware is designed to support compiler.

Architecture Supports:

- **Global Address Space:** Off-chip and on-chip centralized shared memories and local memories are mapped.
- **Flexible processor clustering** with multi-cast and group barrier sync,
- **Power reduction** (DVFS & power gating for each core.)

Key Technologies

➤ Performance:

➤ Multigrain Parallelization

- Hierarchical Coarse Grain Task Parallelization, Loop Parallelization and Vectorization
- Data Localization and Overlapping Data Transfer Using DMA

➤ Architecture:

- Many cores with Accelerators, DMAC (DTU) and Distributed Shared memory
- Global Address Space
- Hierarchical processor grouping
- 3 Dimensional Integration of memory and TSV (Through Silicon Vias)

➤ Power:

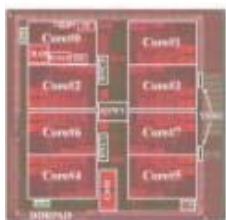
- Compiler controlled DVFS including Clock Gating and Power Gating
- Non-volatile Memory is helpful for Power Gating

➤ Programmability

- Automatic Parallelization by Compiler
- User's advices if compiler could not parallelize sufficiently

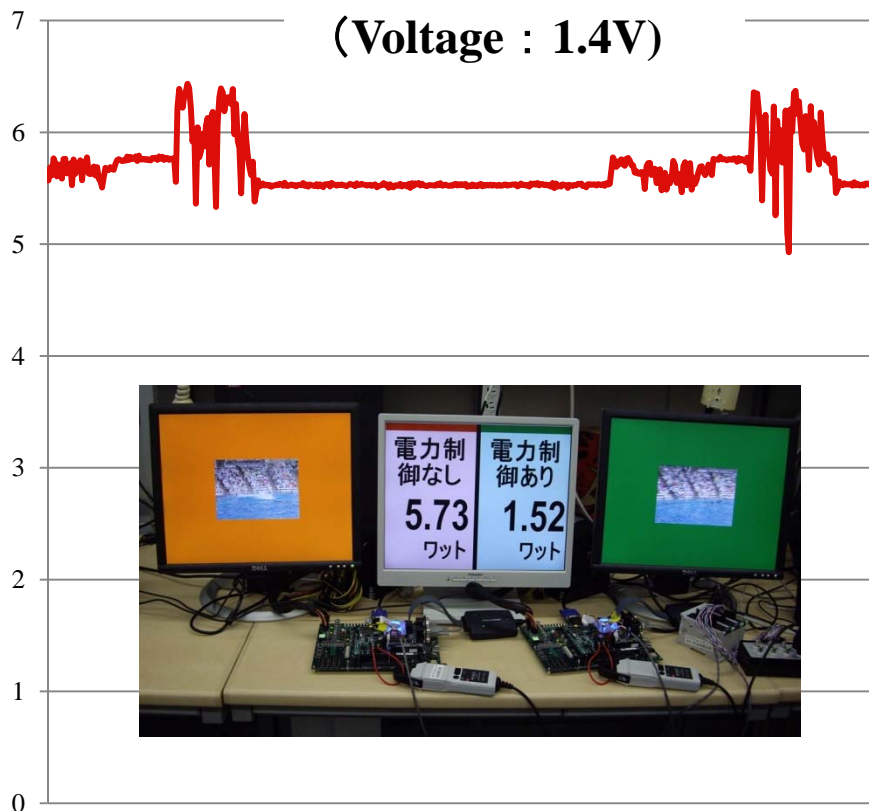
Power Reduction of MPEG2 Decoding to 1/4 on 8 Core Homogeneous Multicore RP-2 by OSCAR Parallelizing Compiler

MPEG2 Decoding with 8 CPU cores



Without Power
Control

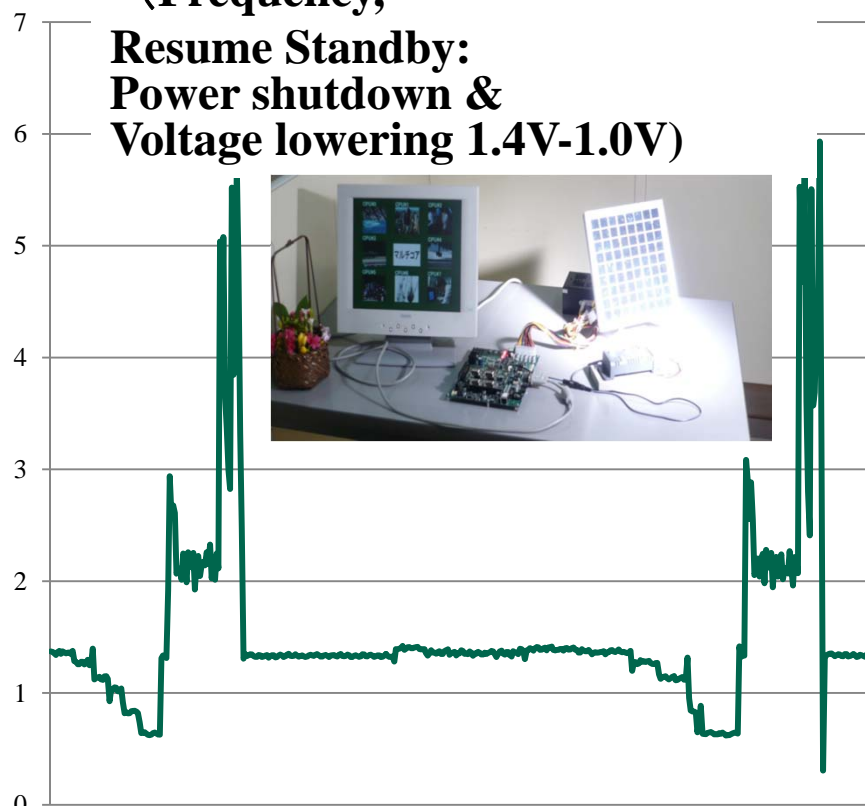
(Voltage : 1.4V)



Avg. Power
5.73 [W]

With Power Control
(Frequency,
Resume Standby:

Power shutdown &
Voltage lowering 1.4V-1.0V)



Avg. Power
1.52 [W]

73.5% Power Reduction



Key Technologies

➤ Performance:

➤ Multigrain Parallelization

- Hierarchical Coarse Grain Task Parallelization, Loop Parallelization and Vectorization
- Data Localization and Overlapping Data Transfer Using DMA

➤ Architecture:

- Many cores with Accelerators, DMAC (DTU) and Distributed Shared memory
- Global Address Space
- Hierarchical processor grouping
- 3 Dimensional Integration of memory and TSV (Through Silicon Vias)

➤ Power:

- Compiler controlled DVFS including Clock Gating and Power Gating
- Non-volatile Memory is helpful for Power Gating

➤ Programmability:

- Automatic Parallelization by Compiler
- User's advices if compiler could not parallelize sufficiently

OSCAR Parallelizing Compiler

To improve **effective performance**, **cost-performance** and **software productivity** and **reduce power**

Multigrain Parallelization

coarse-grain parallelism among loops and subroutines, near fine grain parallelism among statements in addition to loop parallelism

Data Localization

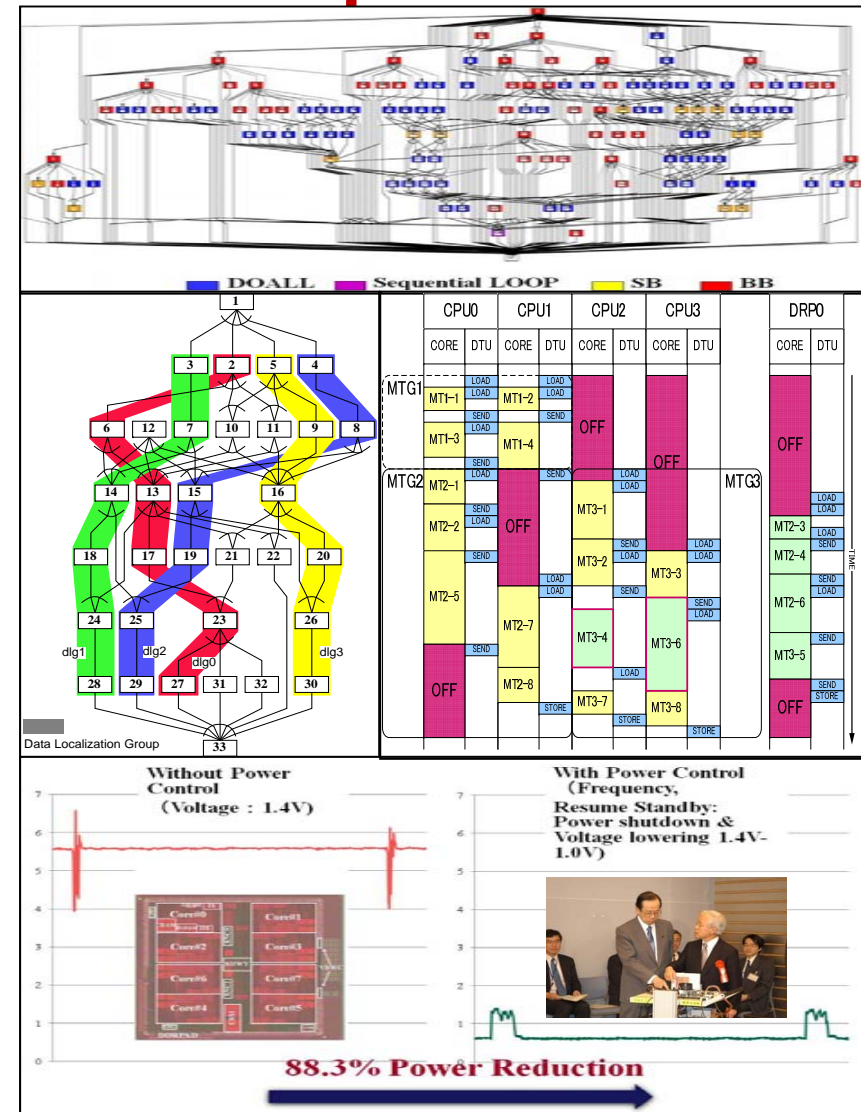
Automatic data management for distributed shared memory, cache and local memory

Data Transfer Overlapping

Data transfer overlapping using Data Transfer Controllers (DMAs)

Power Reduction

Reduction of consumed power by compiler control DVFS and Power gating with hardware supports.

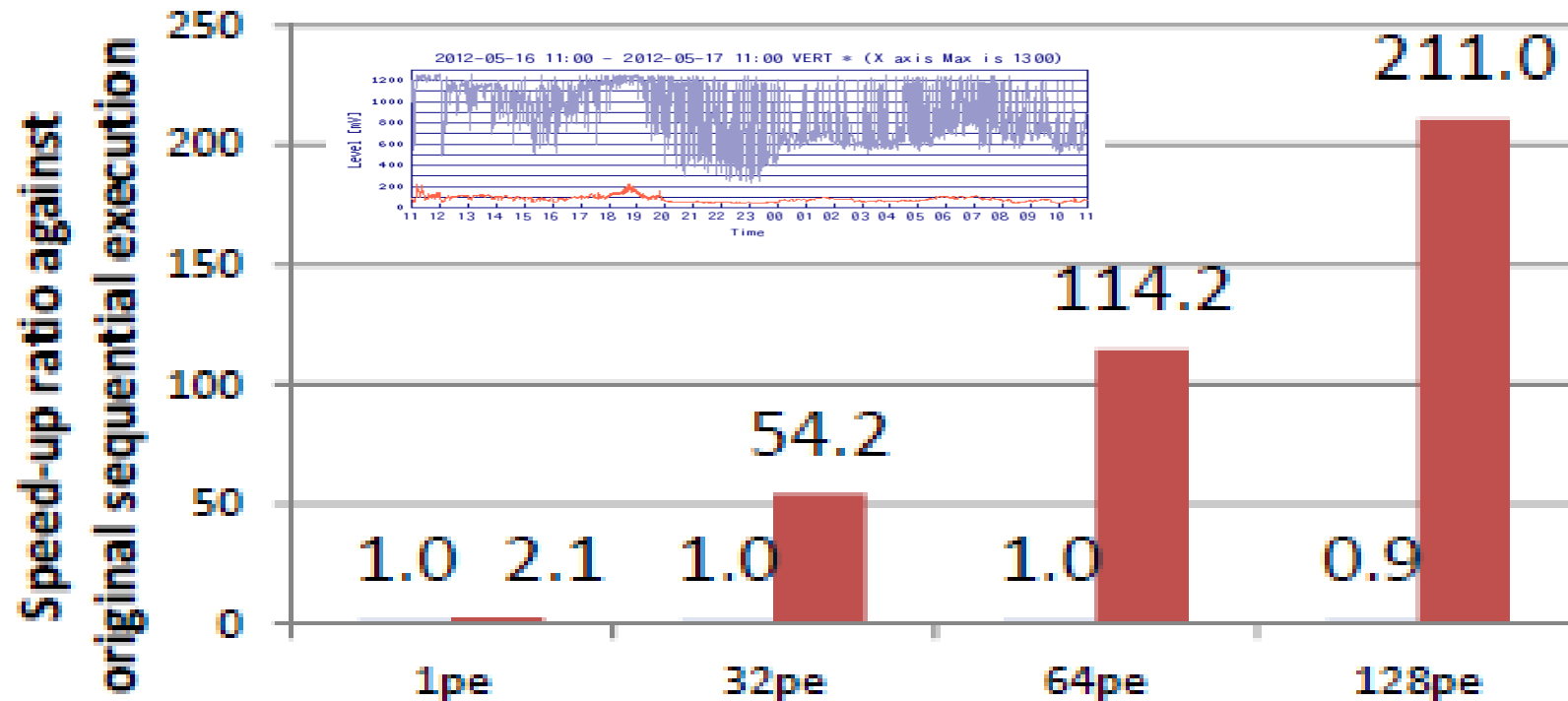




Earthquake Simulation “GMS” on Fujitsu M9000 Sparc CC-NUMA Server



■ original (sun studio) ■ proposed method



With 128 cores, OSCAR compiler gave us 100 times speedup against 1 core execution and 211 times speedup against 1 core using Sun (Oracle) Studio compiler.