The Low Power Multicore and Its Software for Embedded to High Performance Computing

Hironori Kasahara

IEEE Computer Society

President Elect 2017, President 2018

Professor, Dept. of Computer Science & Engineering

Director, Advanced Multicore Processor Research Institute

Waseda University, Tokyo, Japan

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

Waseda Univ. Green Computing R&D Center

Hironori Kasahara Voted 2017 IEEE Computer Society President-Elect

LOS ALAMITOS, Calif., 30 September 2016 – Hironori Kasahara, a Professor of Computer Science at Waseda University in Tokyo, and Director of the Advanced Multicore Research Institute, has been voted **IEEE Computer Society 2017 President-Elect.**

Kasahara is a former member of the IEEE-CS Board of Governors, has served as chair of the IEEE-CS Multicore STC and CS Japan Chapter, and board member of the IEEE Tokyo Section. Kasahara will serve as the 2018 IEEE CS President for a one-year term beginning 1 January 2018. Kasahara garnered 3,278 votes, compared with 2,804 votes cast for Hausi A. Müller, a Professor of Computer Science and Associate Dean of Research, Faculty of Engineering at University of Victoria, Canada, and a member of IEEE-CS Board of Governors.

The President oversees IEEE-CS programs and operations and is a nonvoting member of most IEEE-CS program boards and committees. The 2016 election had a 12.69% turnout, with 6,357 ballots cast. The turnout was higher than the 2015 election with and 12.68% turnout (6,239 ballots cast) and the 2014 election with a 12.66% turnout (6,728 ballots cast).

2016 IEEE Computer Society Election Results

Press Release | Ballot counts

Posted 29 September 2016

Hironori Kasahara selected 2017 President-Elect (2018 President)

Hironori Kasahara has served as a chair or member of 225 society and government committees, including a member of the CS Board of Governors; chair of CS Multicore STC and CS Japan chapter; associate editor of IEEE Transactions on Computers; vice PC chair of the 1996 ENIAC 50th Anniversary International Conference on Supercomputing; general chair of LCPC; PC member of SC, PACT, PPoPP, and ASPLOS; board member of IEEE Tokyo section; and member of the Earth Simulator committee.

He received a PhD in 1985 from Waseda University, Tokyo, joined its faculty in 1986, and has been a professor of computer science since 1997 and a director of the Advanced Multicore Research Institute since 2004. He was a visiting scholar at University of California, Berkeley, and the University of Illinois at Urbana–Champaign's Center for Supercomputing R&D.

Kasahara received the CS Golden Core Member Award, IFAC World Congress Young Author Prize, IPSJ Fellow and Sakai Special Research Award, and the Japanese Minister's Science and Technology Prize. He led Japanese national projects on parallelizing compilers and embedded multicores, and has presented 210 papers, 132 invited talks, and 27 patents. His research has appeared in 520 newspaper and Web articles.

Past IEEE Computer Society Presidents

Chairs of the IRE Professional Group on Electronic Computers

1951-53 Morton M. Astrahan
1953-54 John H. Howard
1954-55 Harry Larson
1955-56 Jean H. Felker
1956-57 Jerre D. Noe
1957-58 Werner Buchholz
1958-59 Willis H. Ware
1959-60 Richard O. Endres
1960-62 Arnold A. Cohen
1962-64 Walter L. Anderson

Chairs of the AIEE Committee on Large-Scale Computing Devices

1946-49 Charles Concordia
1949-51 John Grist Brainerd
1951-53 Walter H. MacWilliams
1953-55 Frank J. Maginniss
1955-57 Edwin L. Harder
1957-59 Morris Rubinoff
1959-61 Ruben A. Imm
1961-63 Claude A. Kagan
1963-64 Gerhard L. Hollander

Chairs & Presidents of the IEEE Computer Society

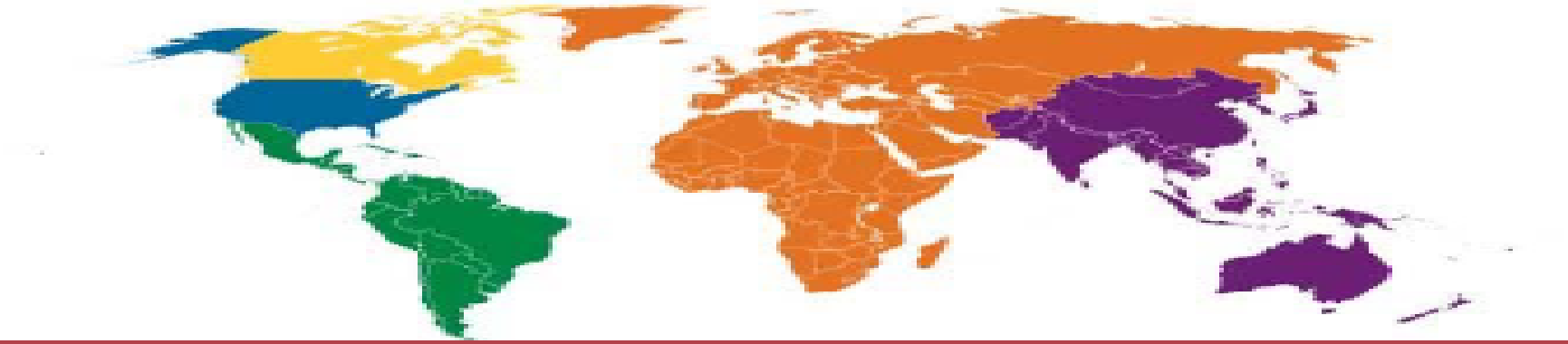
1964-65 Keith Uncapher
1965-66 Richard I. Tanaka
1966-67 Samuel Levine
1968-69 Charles L. Hobbs
1970-71 Edward J. McCluskey
1972-73 Albert S. Hoagland
1974-75 Stephen S. Yau
1976 Dick B. Simmons
1977-78 Merlin G. Smith
1979-80 Tse-Yun Feng
1981 Richard E. Merwin
1982-83 Oscar N. Garcia
1984-85 Martha Sloan
1986-87 Roy L. Russo
1988 Edward A. Parrish
1989 Kenneth A. Anderson
1990 Helen M. Wood
1991 Duncan H. Lawrie
1992 Bruce D. Shriver
1993 James H. Aylor
1994 Laurel V. Kaleda
1995 Ronald G. Hoelzeman
1996 Mario R. Barbacci
1997 Barry W. Johnson
1998 Doris L. Carver
1999 Leonard L. Tripp
2000 Guylaine M. Pollock
2001 Benjamin W. Wah
2002 Willis K. King
2003 Stephen Diamond
2004 Carl K. Chang
2005 Gerald L. Engel
2006 Deborah M. Cooper
2007 Michael R. Williams
2008 Rangachar Kasturi
2009 Susan K. (Kathy) Land,
2010 James D. Isaak
2011 Sorel Reisman
2012 John W. Walz
2013 David Alan Grier
2014 Dejan S. Milojicic
2015 Thomas M. Conte
2016 Roger U. Fujii
2017 Jean-Luc Gaudiot
2018 Hironori Kasahara

IEEE Computer Society

60,000+ members, volunteer-led organization,
200 technical conferences, industry-oriented "Rock Stars" ,
17 scholarly journals and 13 magazines, awards program,
Digital Library with more than 550,000 articles and papers ,
400 local and regional chapters, 40 technical committees,

- Region 1 (Northeastern US)
- Region 2 (Eastern US)
- Region 3 (Southern US)
- Region 4 (Central US)
- Region 5 (Southwestern US)
- Region 6 (Western US)
- Region 7 (Canada)
- Region 8 (Africa, Europe, Middle East)
- Region 9 (Latin America)
- Region 10 (Asia and Pacific)

- IEEE-USA (Regions 1-6)



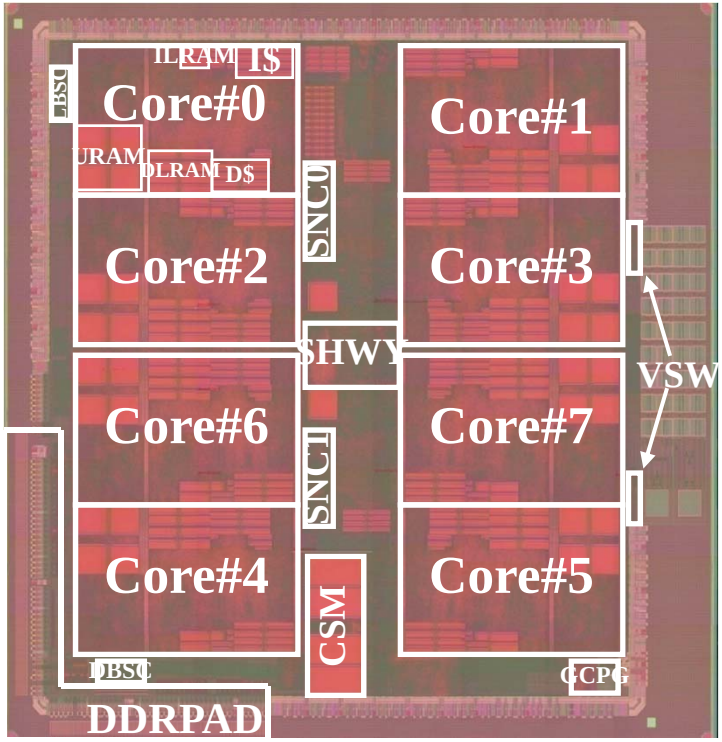
IEEE Computer Society BoG (Board of Governors) Feb.1, 2017

Toward 2018

1. Refining content and services to further improve the satisfaction of CS members;
2. Considering an incentive for volunteers to further accelerate CS activities and promptly provide technical benefits for people around the globe;
3. Offering more attractive services for practitioners in industry;
4. Providing the world's best educational content and historical treasures for future generations, which only the CS can create with our pioneering researchers (for example, the Multicore Compiler Video Series found at www.computer.org/web/education/multicore-video-series);
5. Thinking about sustainable membership fees while considering the diversity of economic situations within the 10 regions;
6. Cooperating with other IEEE societies and sister societies in a timely and efficient manner;
7. Intelligibly introducing the latest computer-related technologies to younger generations, including children, so that they can realize their technological dreams.

Multicores for Performance and Low Power

Power consumption is one of the biggest problems for performance scaling from smartphones to cloud servers and supercomputers (“K” more than 10MW) .



IEEE ISSCC08: Paper No. 4.5, M.ITO, … and H. Kasahara, “An 8640 MIPS SoC with Independent Power-off Control of 8 CPUs and 8 RAMs by an Automatic Parallelizing Compiler”

Power ∝ Frequency * Voltage^2

(Voltage ∝ Frequency)

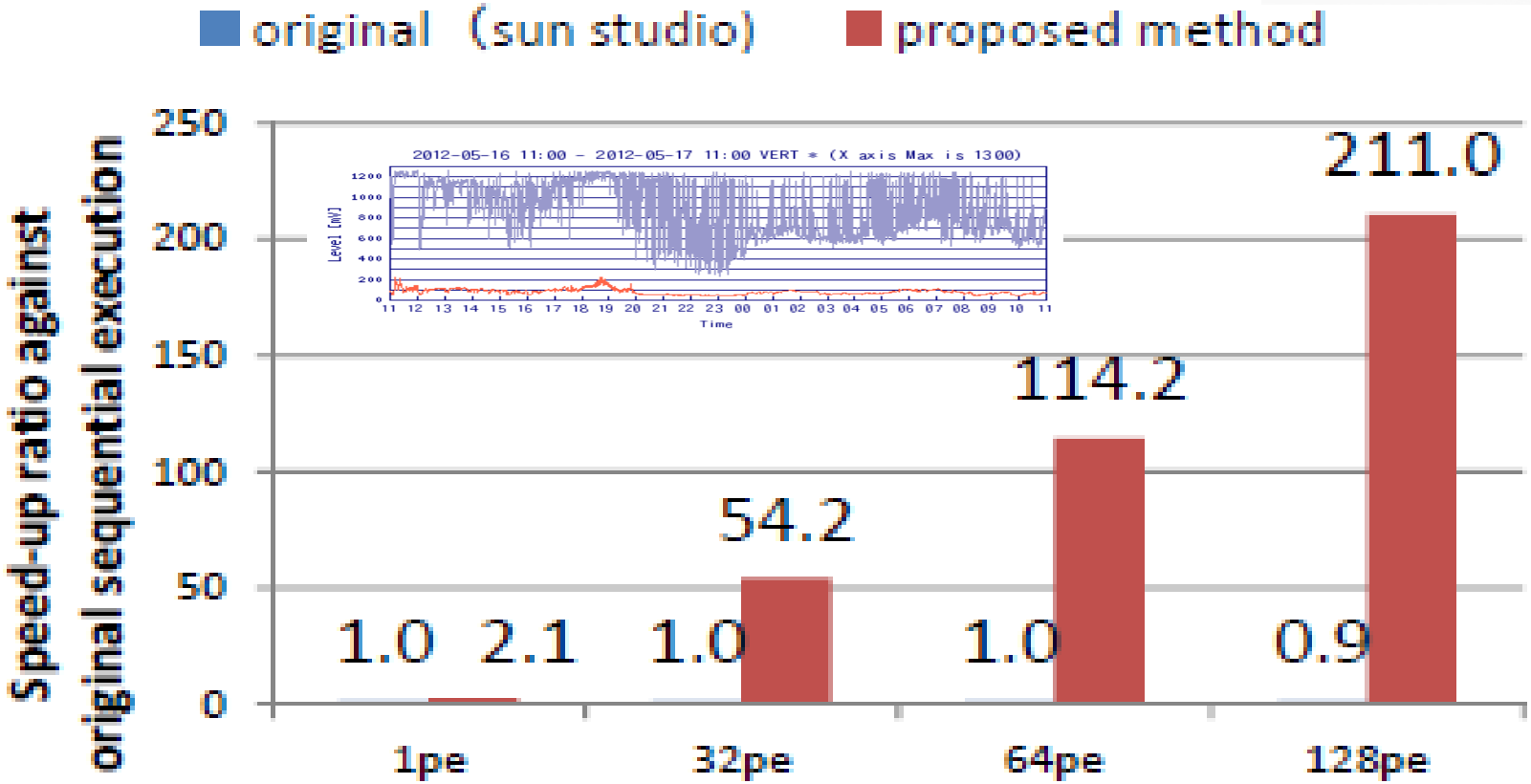
⇒ Power ∝ Frequency^3

If Frequency is reduced to 1/4 (Ex. 4GHz→1GHz), Power is reduced to 1/64 and Performance falls down to 1/4 .

<Multicores>

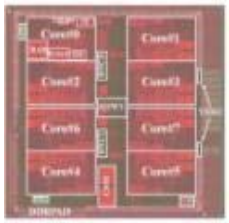
If 8cores are integrated on a chip, Power is still 1/8 and Performance becomes 2 times.

Earthquake Simulation "GMS" on Fujitsu M9000 Sparc CC-NUMA Server



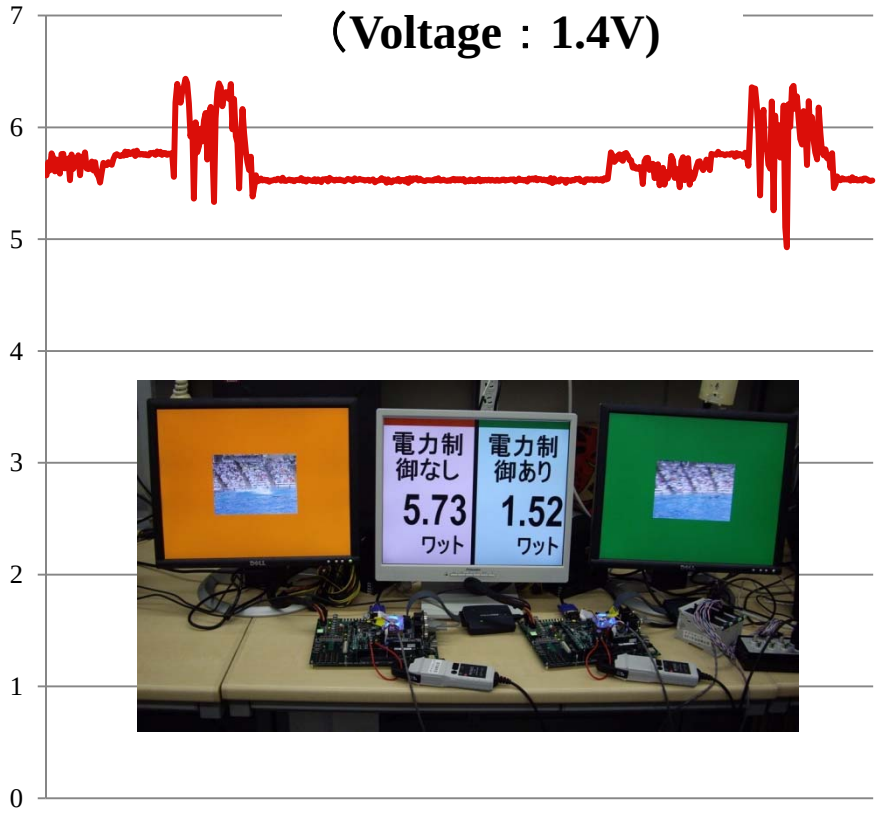
With 128 cores, OSCAR compiler gave us 100 times speedup against 1 core execution and 211 times speedup against 1 core using Sun (Oracle) Studio compiler.

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

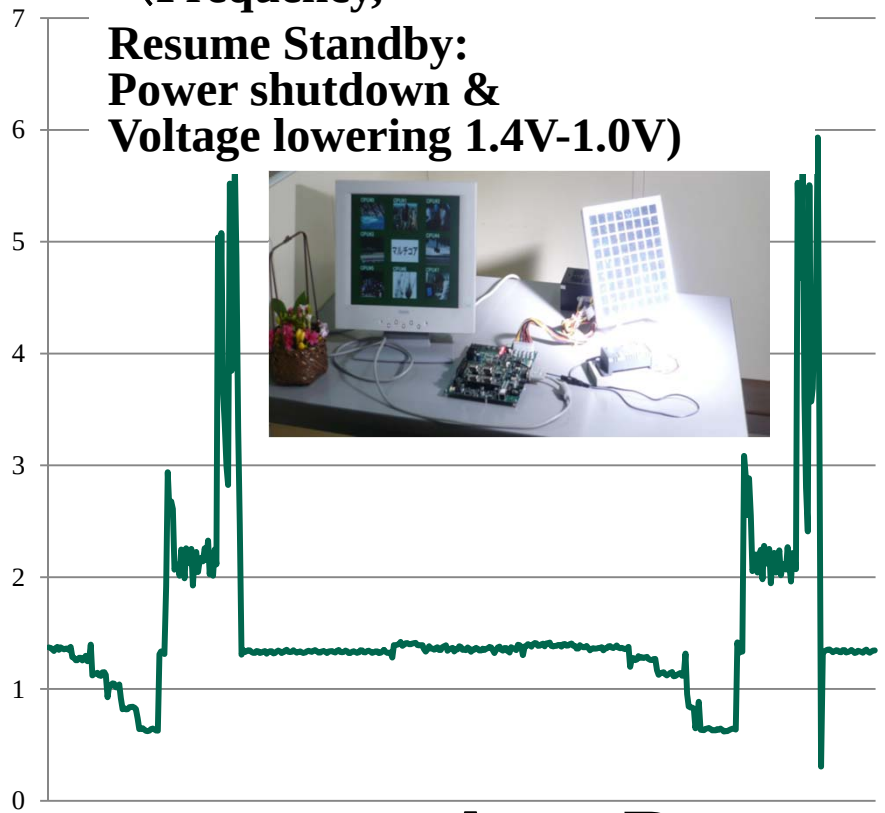


MPEG2 Decoding with 8 CPU cores

Without Power Control （Voltage：1.4V)



With Power Control （Frequency, Resume Standby: Power shutdown & Voltage lowering 1.4V-1.0V)



Avg. Power 5.73 [W]

73.5% Power Reduction

Avg. Power 1.52 [W]

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

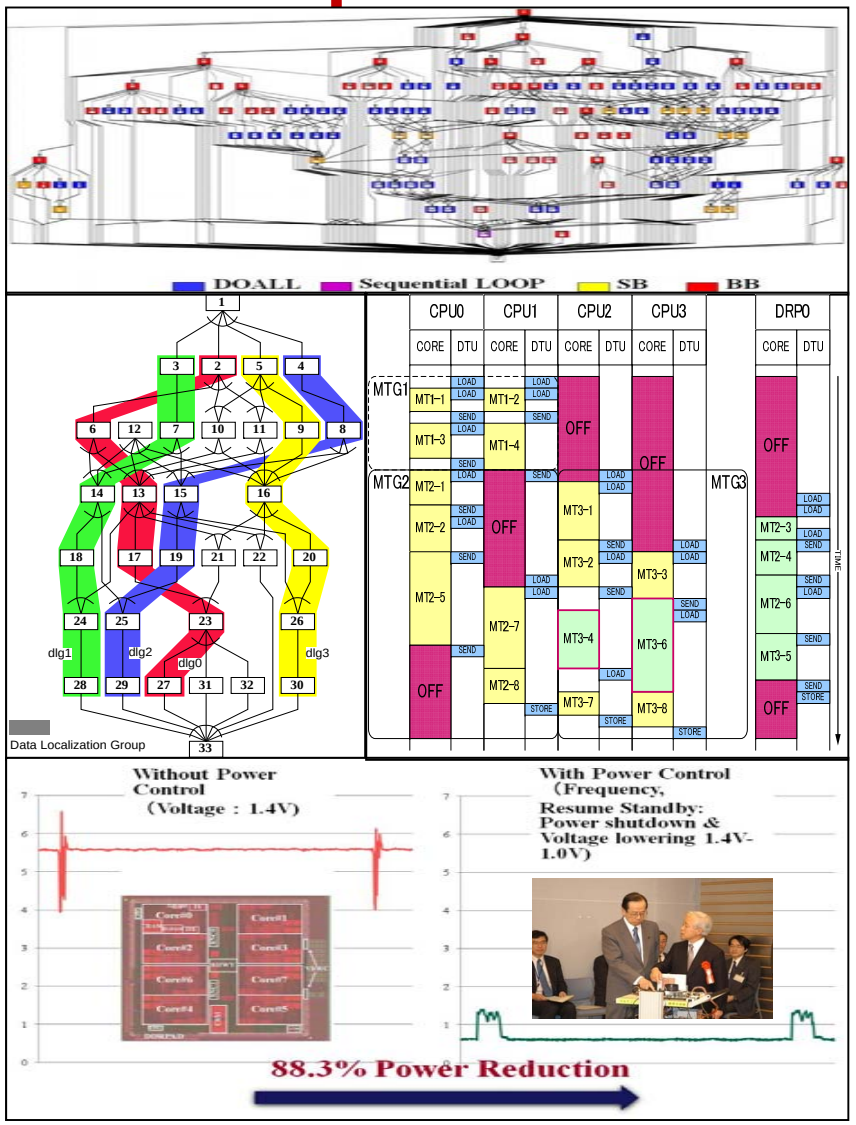
Automatic data management for distributed shared memory, cache and local memory

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.



Multicore Program Development Using OSCAR API V2.0

Sequential Application Program in Fortran or C
(Consumer Electronics, Automobiles, Medical, Scientific computation, etc.)

Homogeneous

Hetero

Manual parallelization / power reduction

Accelerator Compiler/ User
Add "hint" directives before a loop or a function to specify it is executable by the accelerator with how many clocks

Waseda OSCAR Parallelizing Compiler
- Coarse grain task parallelization
- Data Localization
- DMAC data transfer
- Power reduction using DVFS, Clock/ Power gating

OSCAR API for Homogeneous and/or Heterogeneous Multicores and manycores
Directives for thread generation, memory, data transfer using DMA, power managements

Parallelized API F or C program

Proc0
Code with directives
Thread 0

Proc1
Code with directives
Thread 1

Accelerator 1
Code

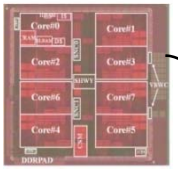
Accelerator 2
Code

Generation of parallel machine codes using sequential compilers

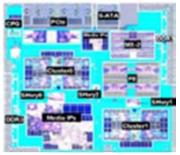
Low Power Homogeneous Multicore Code Generation
API Analyzer | Existing sequential compiler

Low Power Heterogeneous Multicore Code Generation
API Analyzer (Available from Waseda) | Existing sequential compiler

Server Code Generation
OpenMP Compiler



Homegeneous Multicore s from Vendor A (SMP servers)



Heterogeneous Multicores from Vendor B

Shred memory servers

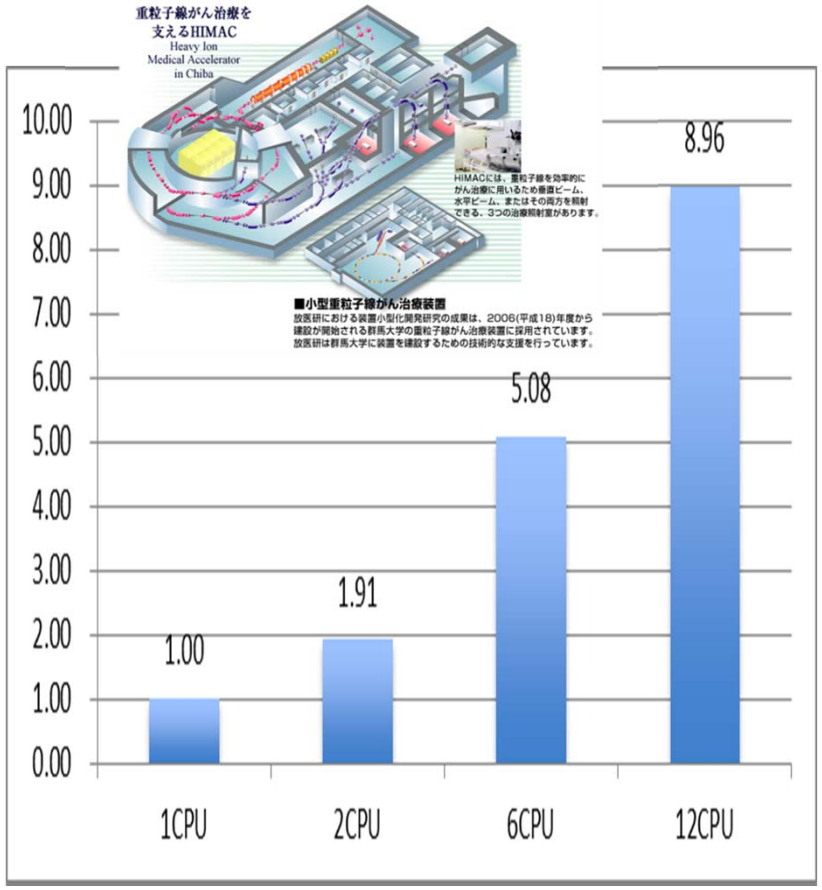
Executable on various multicores

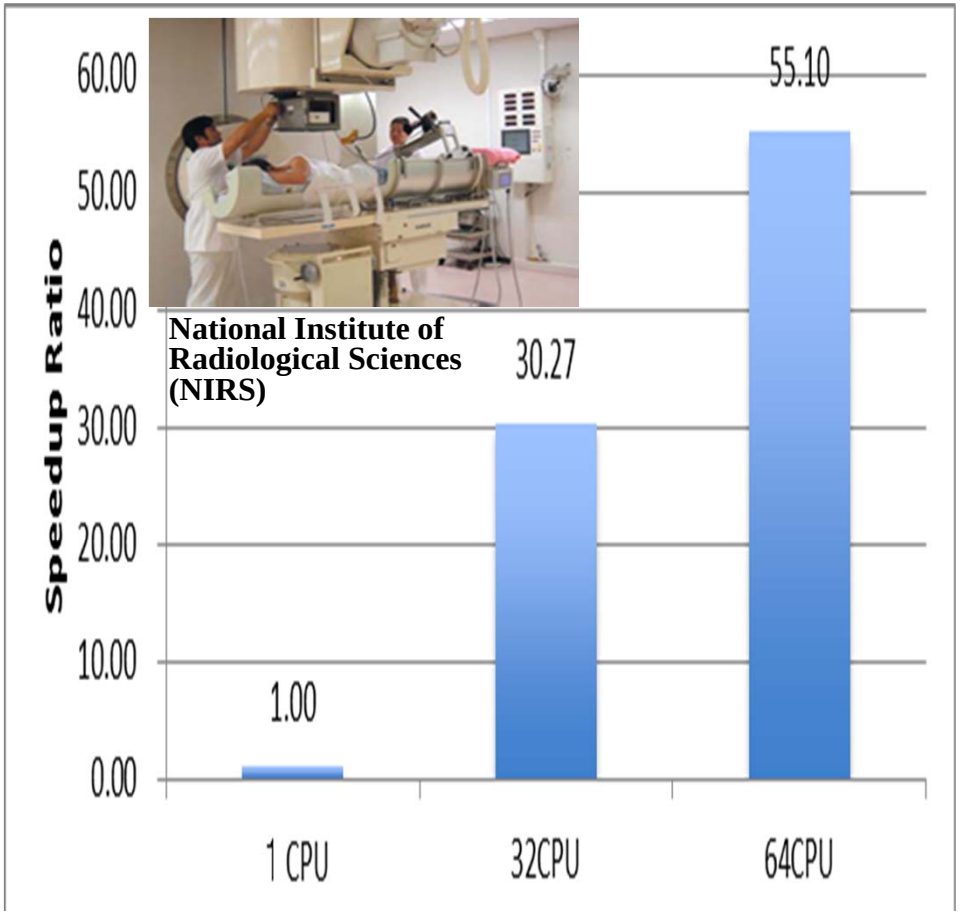
Hitachi, Renesas, NEC, Fujitsu, Toshiba, Denso, Olympus, Mitsubishi, Esol, Cats, Gaio, 3 univ.

OSCAR: Optimally Scheduled Advanced Multiprocessor
API : Application Program Interface

Cancer Treatment
Carbon Ion Radiotherapy

(Previous best was 2.5 times speedup on 16 processors with hand optimization)





8.9times speedup by 12 processors

Intel Xeon X5670 2.93GHz 12 core SMP (Hitachi HA8000)

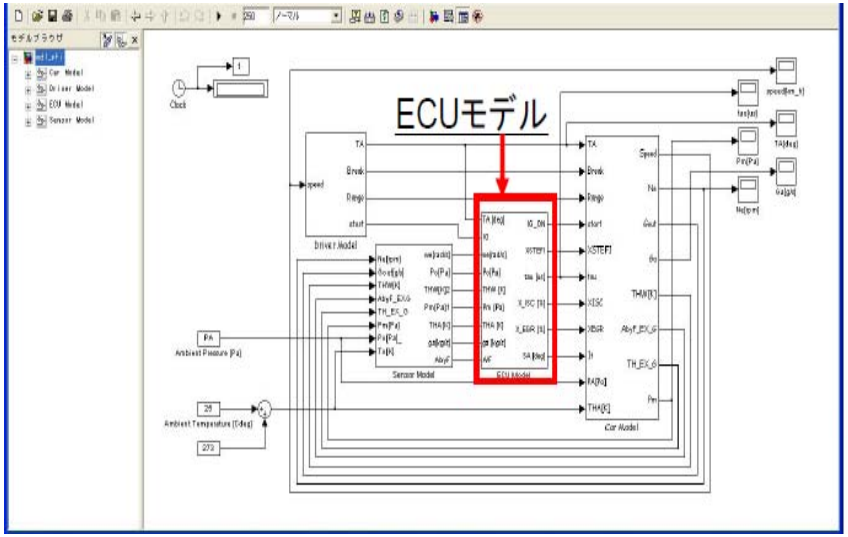
55 times speedup by 64 processors

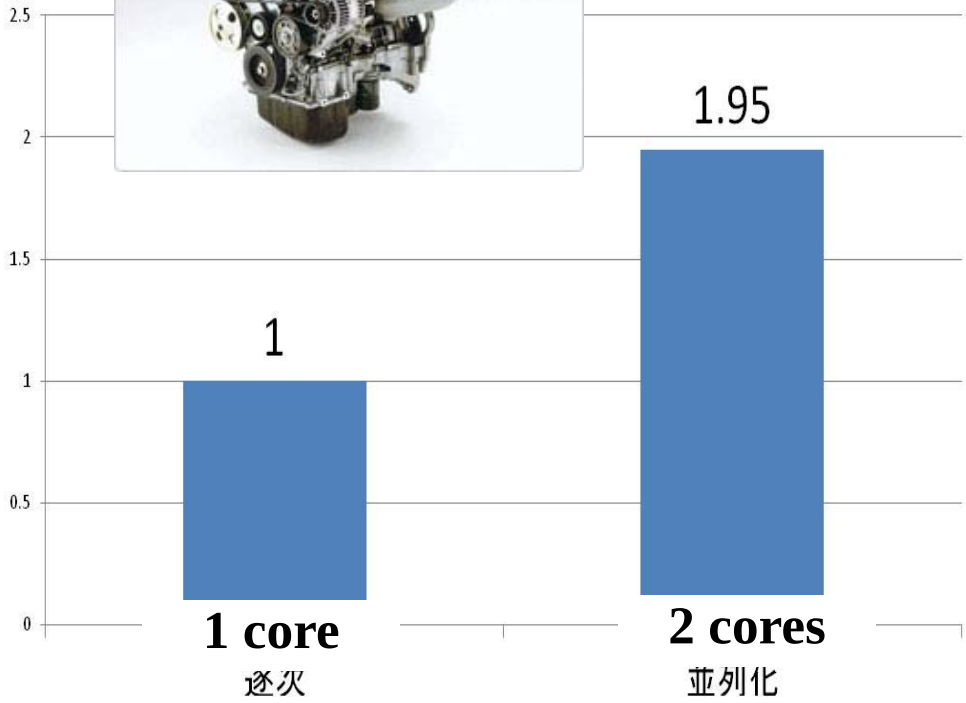
IBM Power 7 64 core SMP (Hitachi SR16000)

Engine Control by multicore with Denso

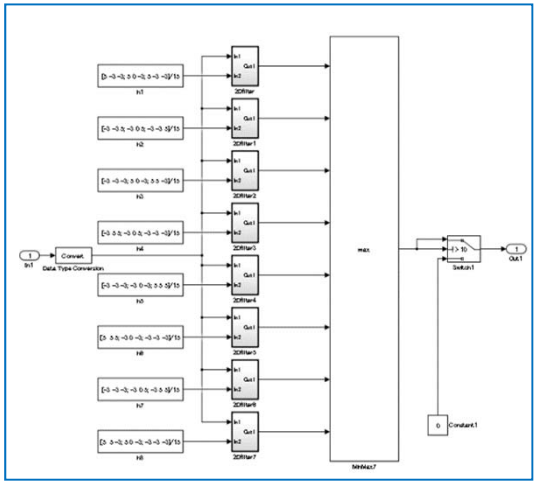
Though so far parallel processing of the engine control on multicore has been very difficult, Denso and Waseda succeeded 1.95 times speedup on 2core V850 multicore processor.

Hard real-time automobile engine control by multicore





OSCAR Compile Flow for Simulink Applications



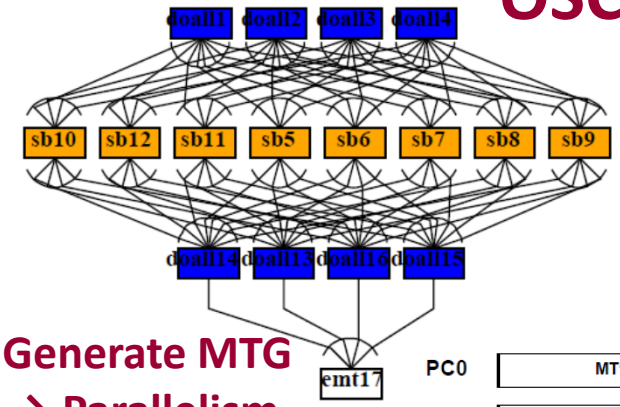
Simulink model

Generate C code
using Embedded Coder

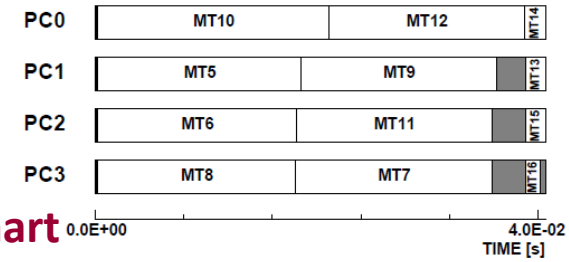
```
/* Model step function */
void VesselExtraction_step(void)
{
  int32_T i;
  real_T u0;

  /* DataTypeConversion: '<S1>/Data Type Conversion' incorporates:
   *  Inport: '<Root>/In1'
   */
  for (i = 0; i < 16384; i++) {
    VesselExtraction_B.DataTypeConversion[i] = VesselExtraction_U.In1[i];
  }

  /* End of DataTypeConversion: '<S1>/Data Type Conversion' */

  /* Outputs for Atomic SubSystem: '<S1>/2Dfilter' */

  /* Constant: '<S1>/h1' */
  VesselExtraction_Dfilter(VesselExtraction_B.DataTypeConversion,
    VesselExtraction_P.h1_Value, &VesselExtraction_B.Dfilter,
    (P_Dfilter_VesselExtraction_T *)&VesselExtraction_P.Dfilter);

  /* End of Outputs for SubSystem: '<S1>/2Dfilter' */

  /* Outputs for Atomic SubSystem: '<S1>/2Dfilter1' */

  /* Constant: '<S1>/h2' */
  VesselExtraction_Dfilter(VesselExtraction_B.DataTypeConversion,
    VesselExtraction_P.h2_Value, &VesselExtraction_B.Dfilter1,
    (P_Dfilter_VesselExtraction_T *)&VesselExtraction_P.Dfilter1);
```

C code

OSCAR Compiler



(1) Generate MTG
→ Parallelism



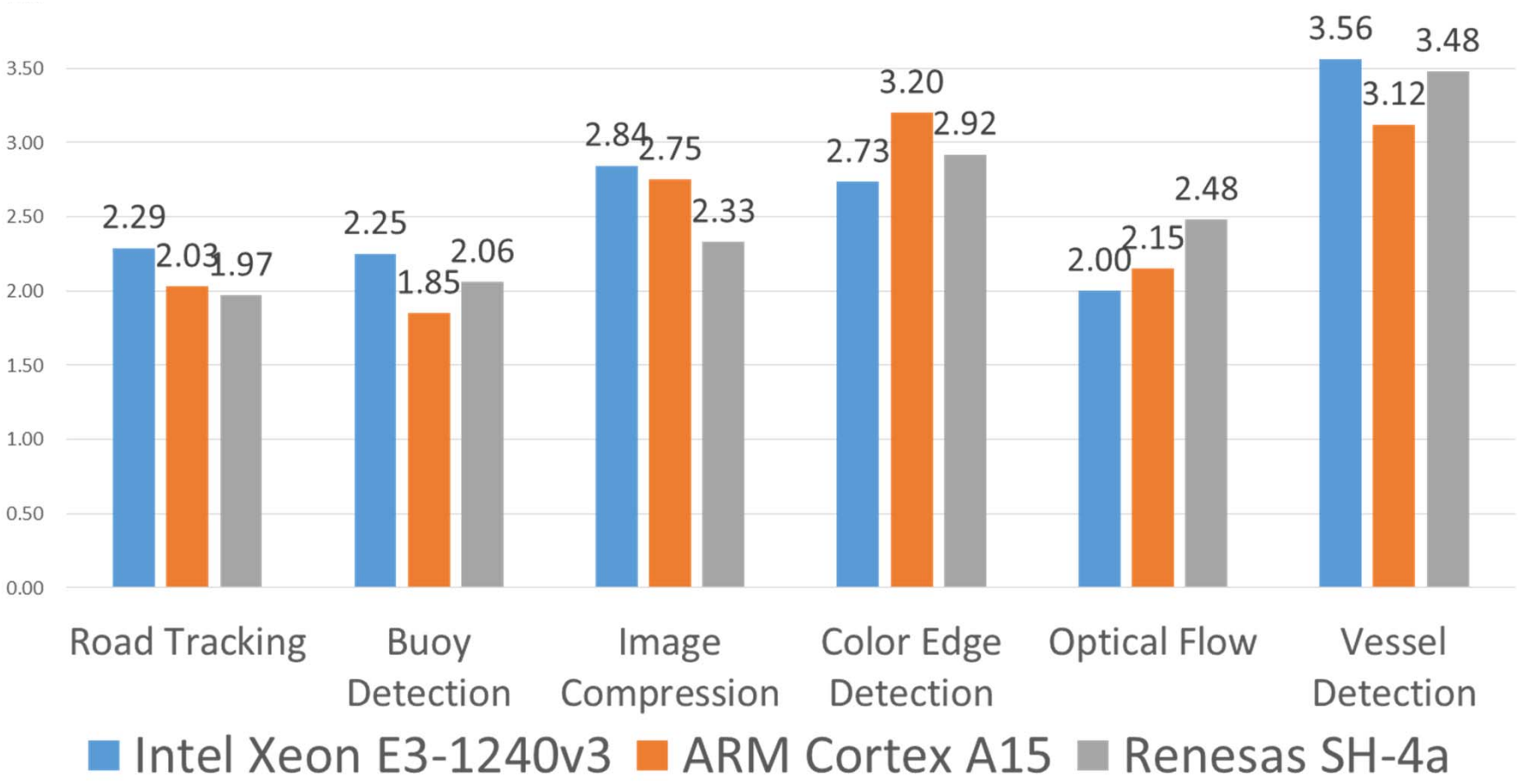
(2) Generate gantt chart
→ Scheduling in a multicore

```
void VesselExtraction_step ( )
{
    int thr1 ;
    int thr2 ;
    int thr3 ;
    {
        oscar_thread_create ( & thr1 ,
            thread_function_001 , (void*)1 ) ;
        oscar_thread_create ( & thr2 ,
            thread_function_002 , (void*)2 ) ;
        oscar_thread_create ( & thr3 ,
            thread_function_003 , (void*)3 ) ;

            VesselExtraction_step_PE0 ( ) ;

        oscar_thread_join ( thr1 ) ;
        oscar_thread_join ( thr2 ) ;
        oscar_thread_join ( thr3 ) ;
    }
}

void thread_function_001 ( void )
{
    VesselExtraction_step_PE1 ( ) ;
}
```

(3) Generate parallelized C code
using the OSCAR API
→ Multiplatform execution
(Intel, ARM and SH etc)

Speedups of MATLAB/Simulink Image Processing on Various 4core Multicores

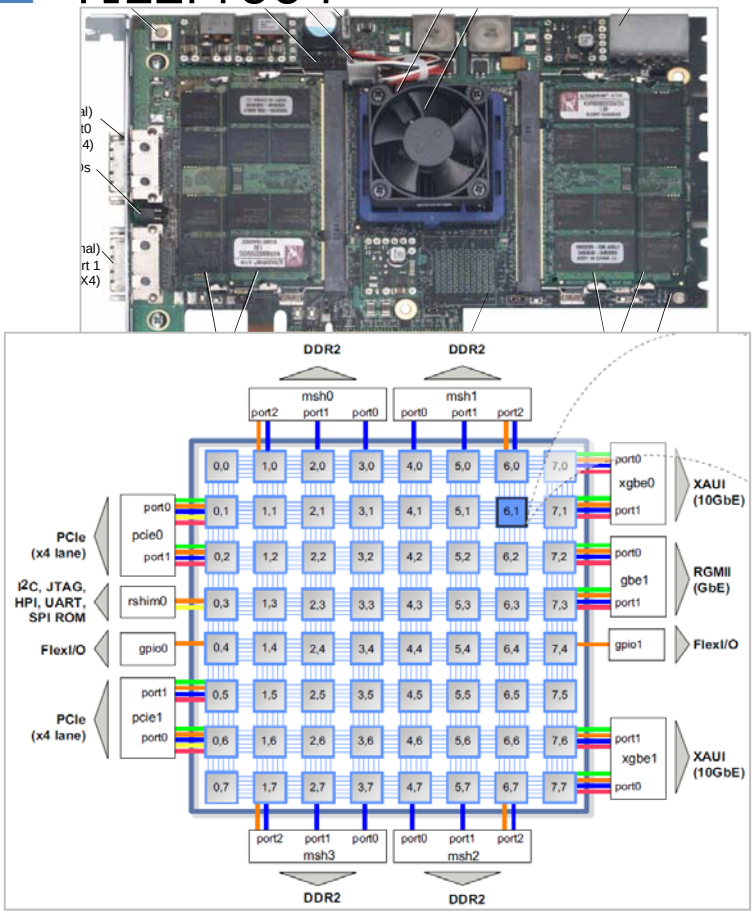
(Intel Xeon, ARM Cortex A15 and Renesas SH4A)

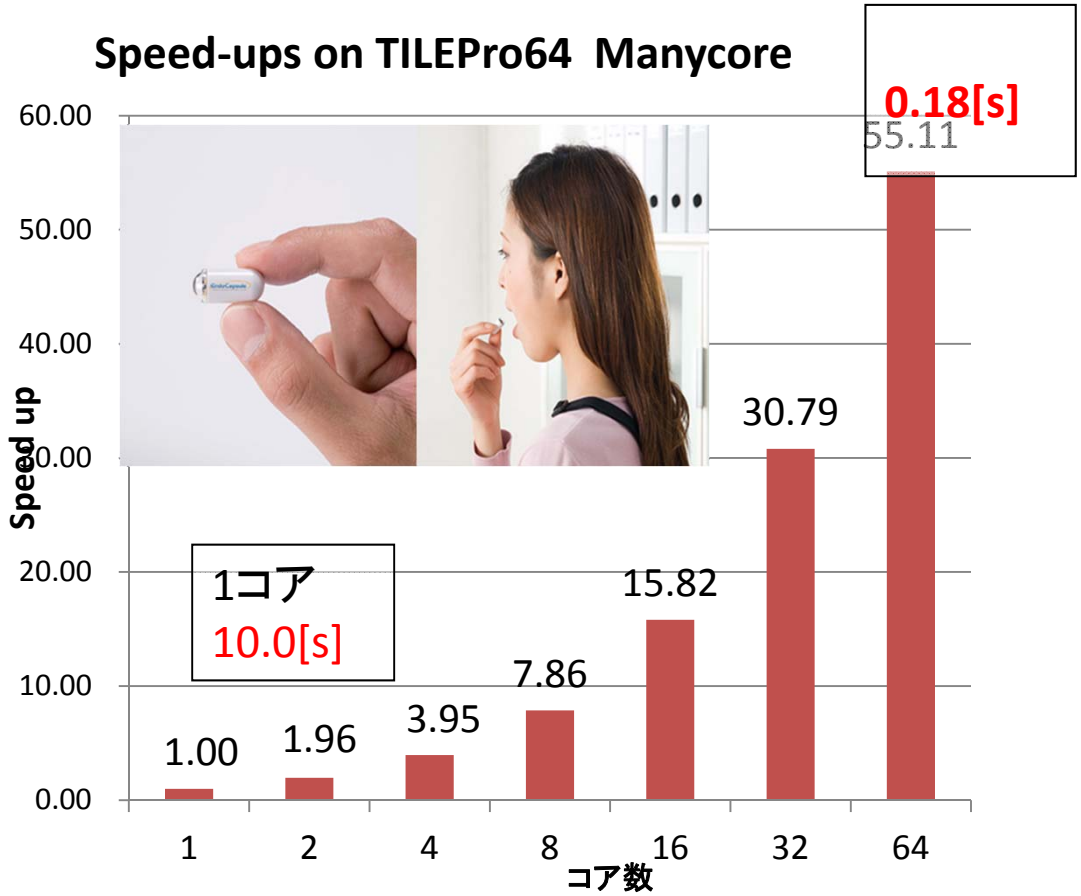


Road Tracking, Image Compression : http://www.mathworks.co.jp/jp/help/vision/examples
Buoy Detection : http://www.mathworks.co.jp/matlabcentral/fileexchange/44706-buoy-detection-using-simulink
Color Edge Detection : http://www.mathworks.co.jp/matlabcentral/fileexchange/28114-fast-edges-of-a-color-image--actual-color--not-converting-to-grayscale-/
Vessel Detection : http://www.mathworks.co.jp/matlabcentral/fileexchange/24990-retinal-blood-vessel-extraction/

Automatic Parallelization of Still Image Encoding Using JPEG-XR for the Next Generation Cameras and Drinkable Inner Camera

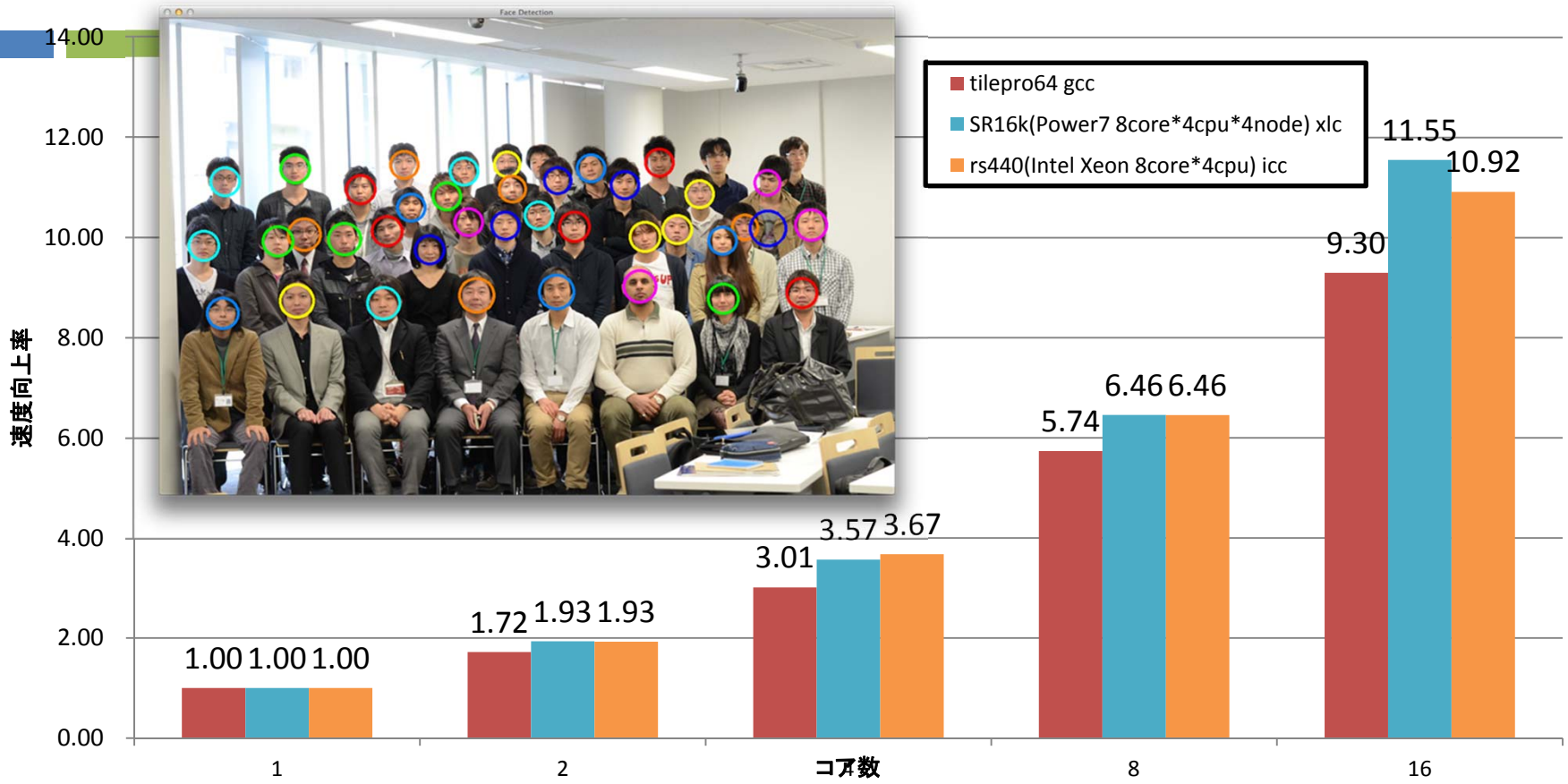
- TILEPro64





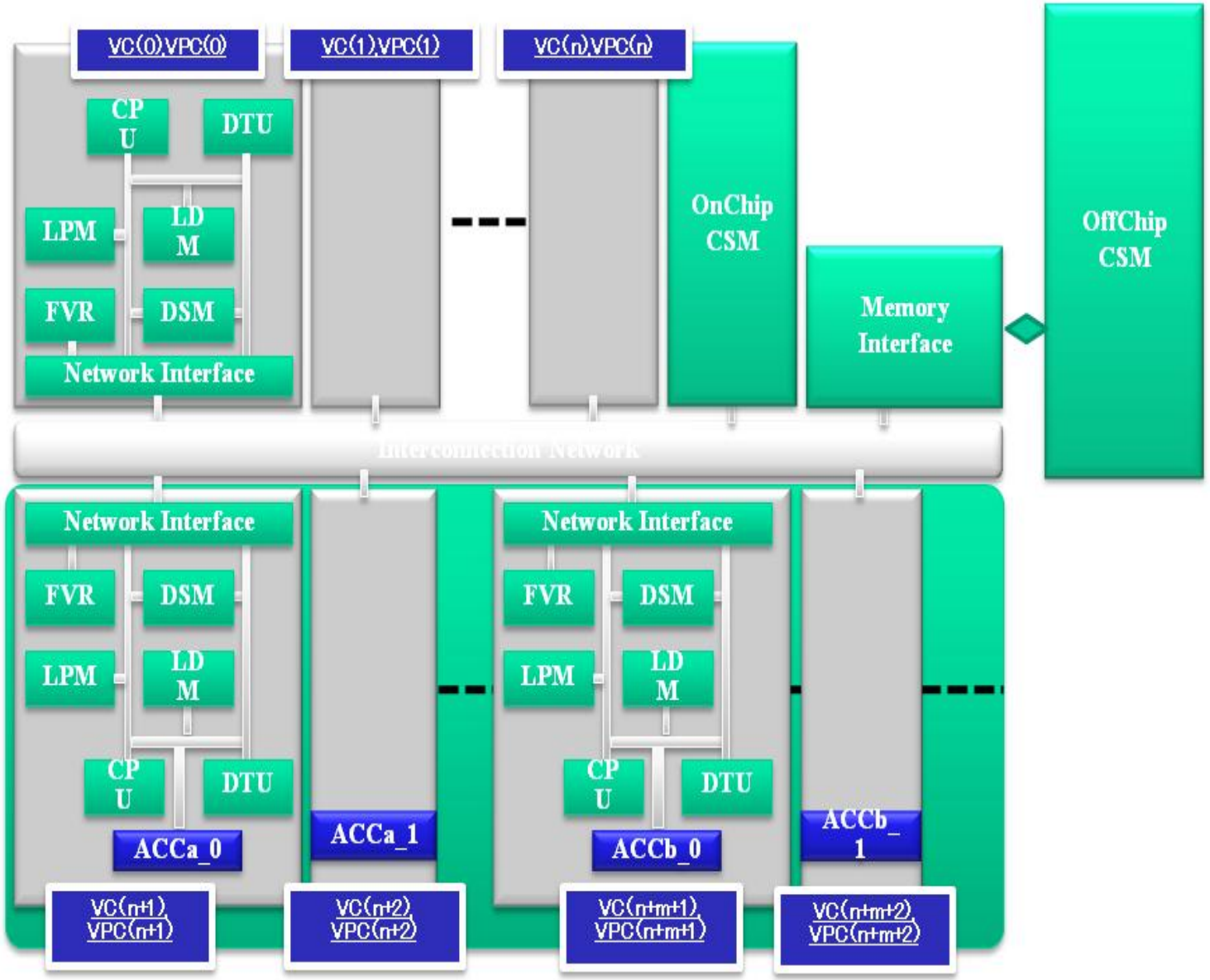
55 times speedup with 64 cores against 1 core

Parallel Processing of Face Detection on Manycore, Highend and PC Server



- OSCAR compiler gives us 11.55 times speedup for 16 cores against 1 core on SR16000 Power7 highend server.

OSCAR Heterogeneous Multicore



DTU
- Data Transfer Unit

LPM
- Local Program Memory

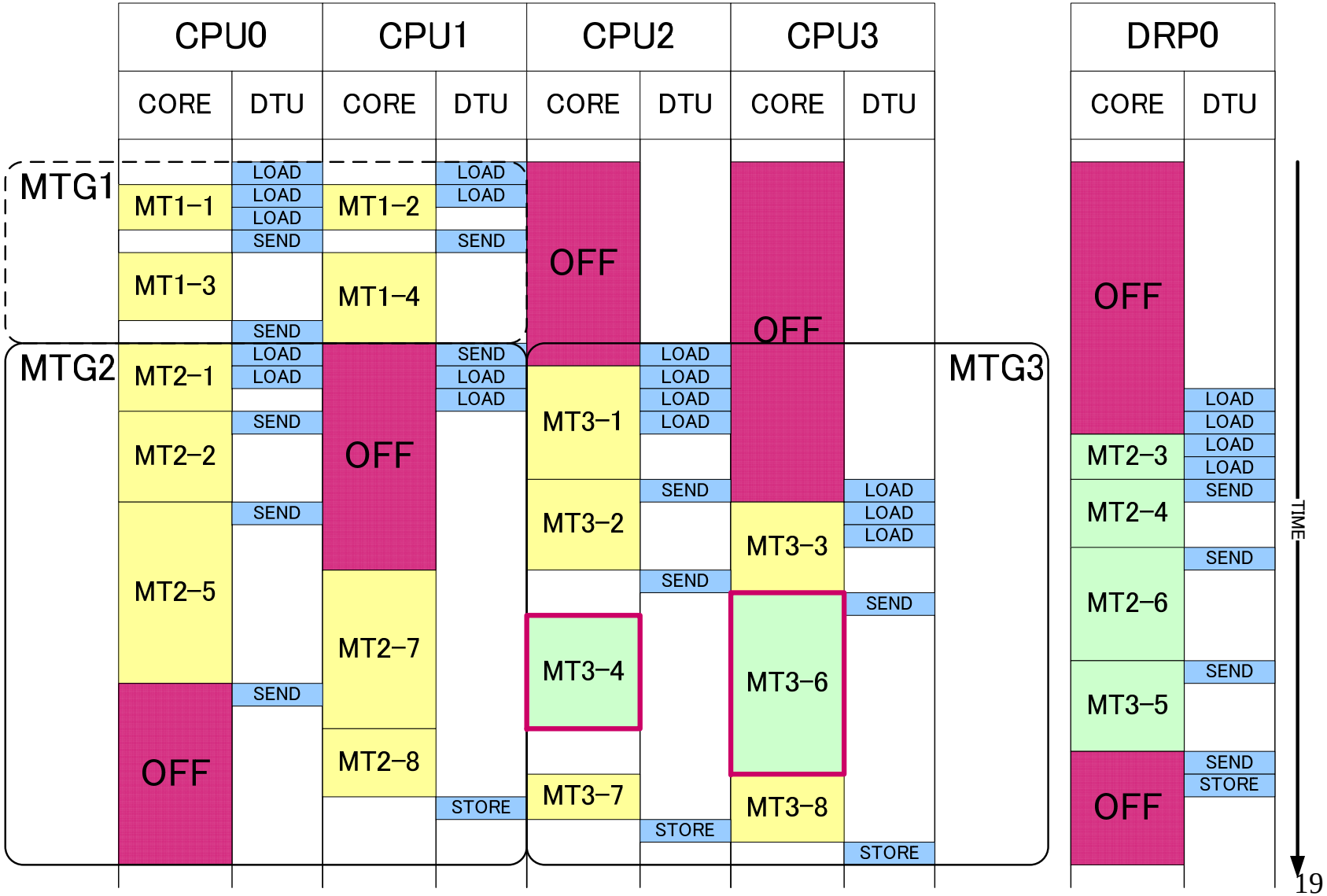
LDM
- Local Data Memory

DSM
- Distributed Shared Memory

CSM
- Centralized Shared Memory

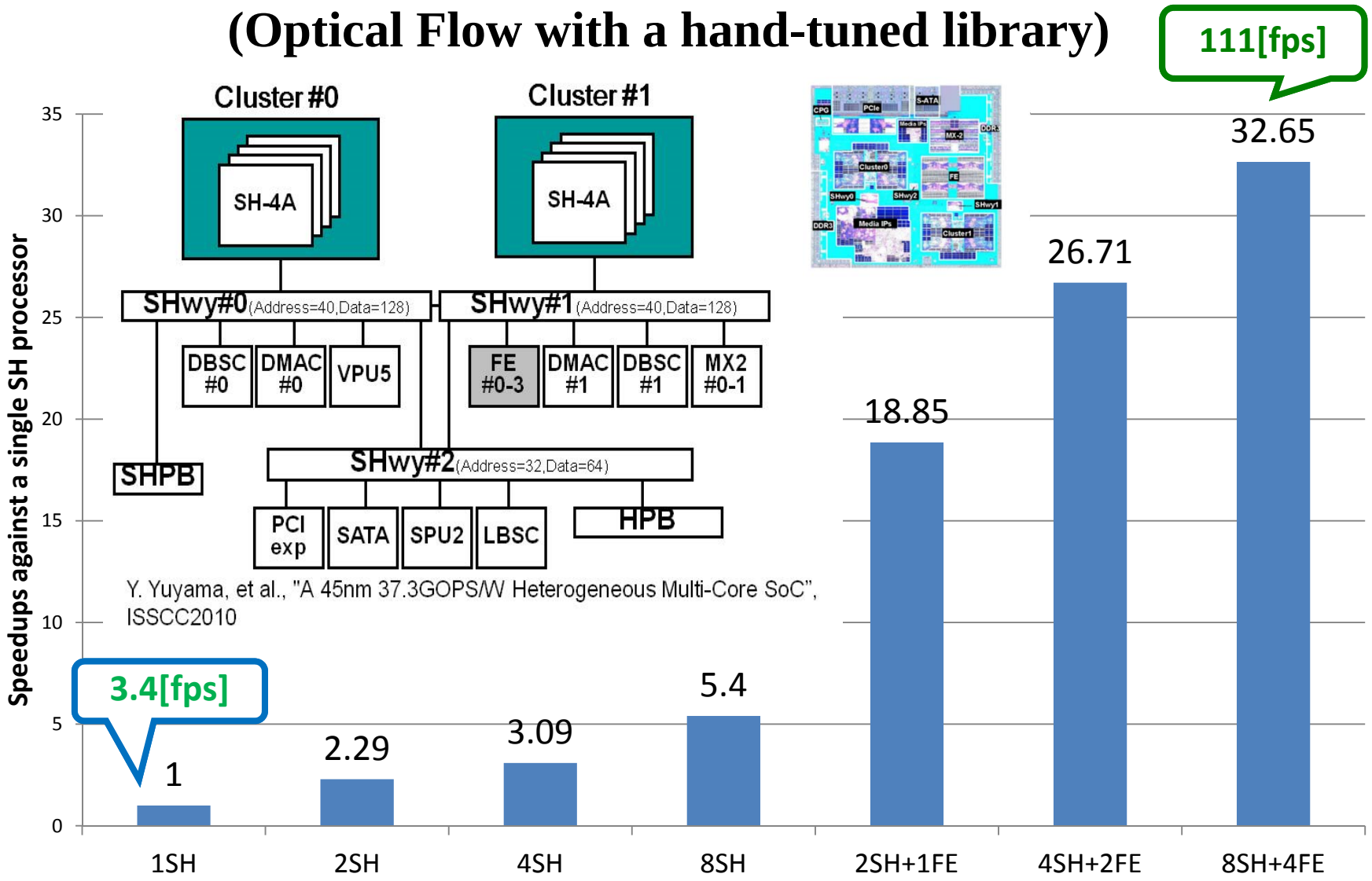
FVR
- Frequency/Voltage Control Register

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



33 Times Speedup Using OSCAR Compiler and OSCAR API on RP-X

(Optical Flow with a hand-tuned library)



Power Reduction in a real-time execution controlled by OSCAR Compiler and OSCAR API on RP-X (Optical Flow with a hand-tuned library)

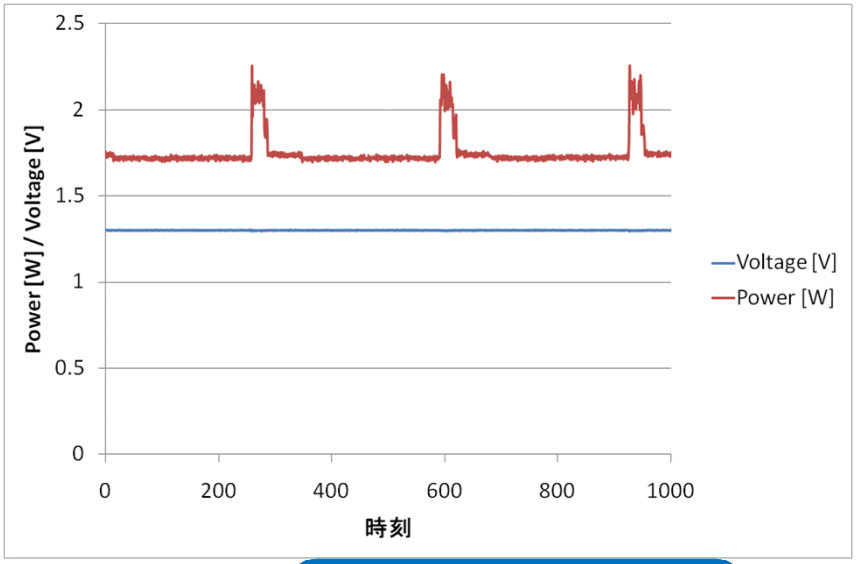
Without Power Reduction

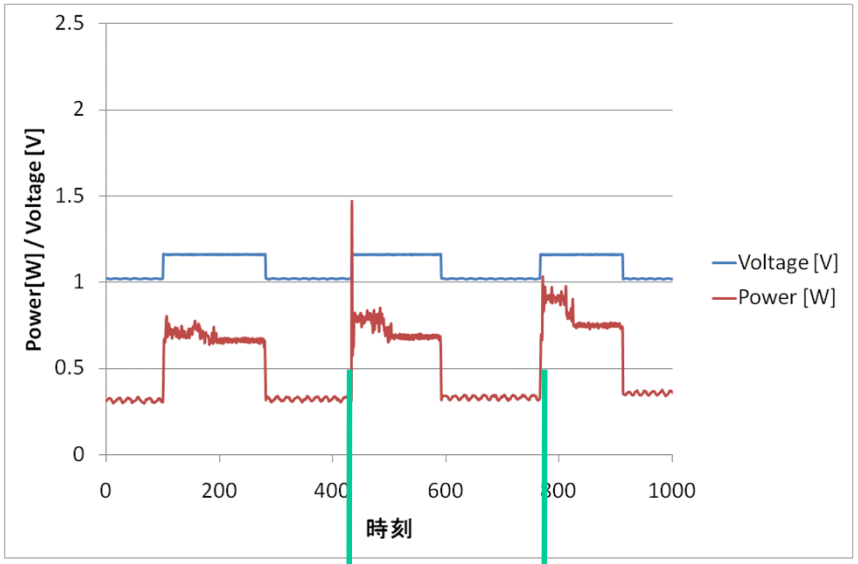
With Power Reduction by OSCAR Compiler

70% of power reduction

Average:1.76[W]

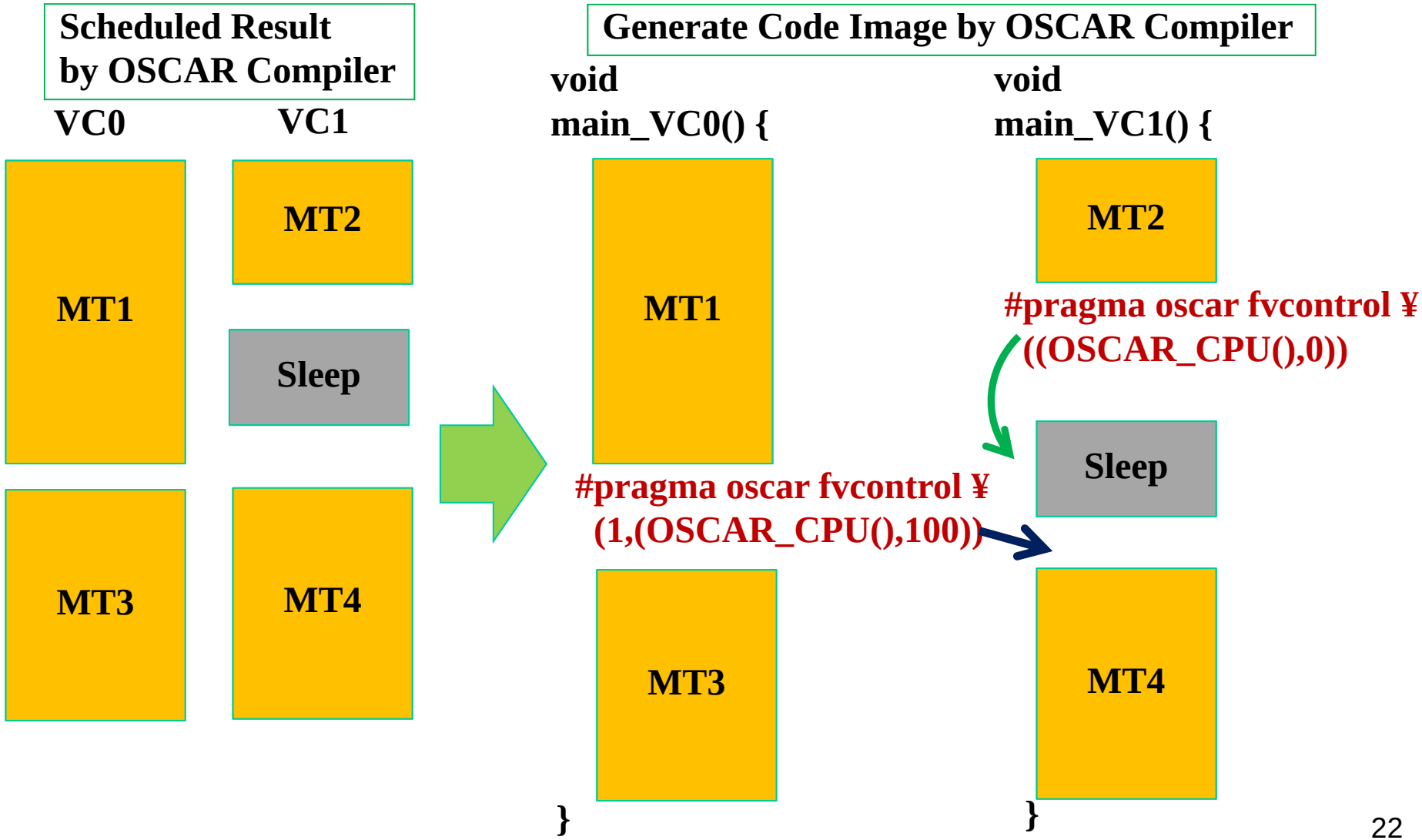
Average:0.54[W]





1cycle : 33[ms]
→30[fps]

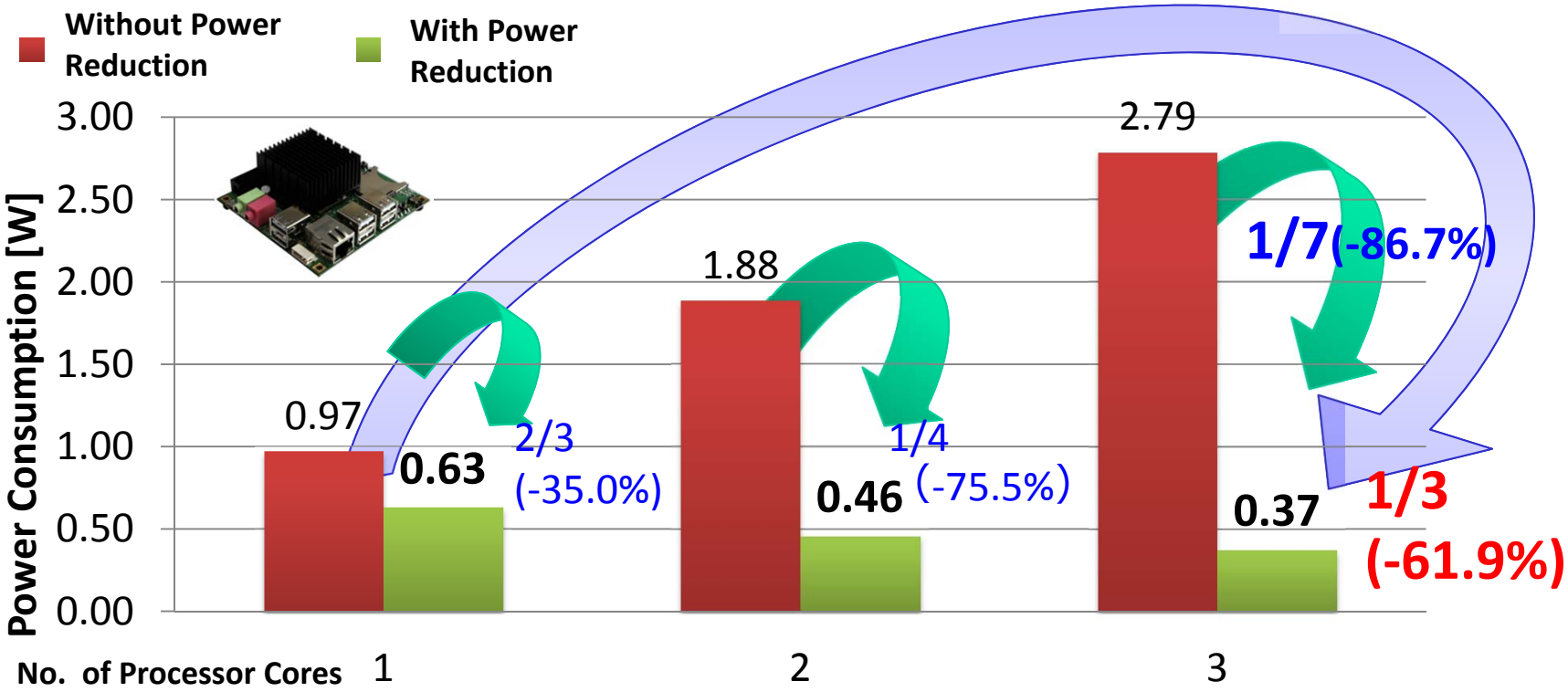
Low-Power Optimization with OSCAR API



Automatic Power Reduction for MPEG2 Decode on Android Multicore

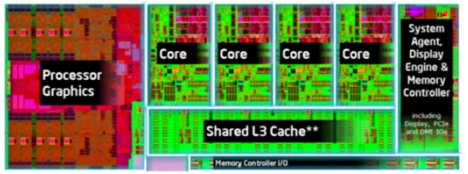
ODROID X2 ARM Cortex-A9 4 cores

http://www.youtube.com/channel/UCS43lNYEIkC8i_KIgFZYQBQ



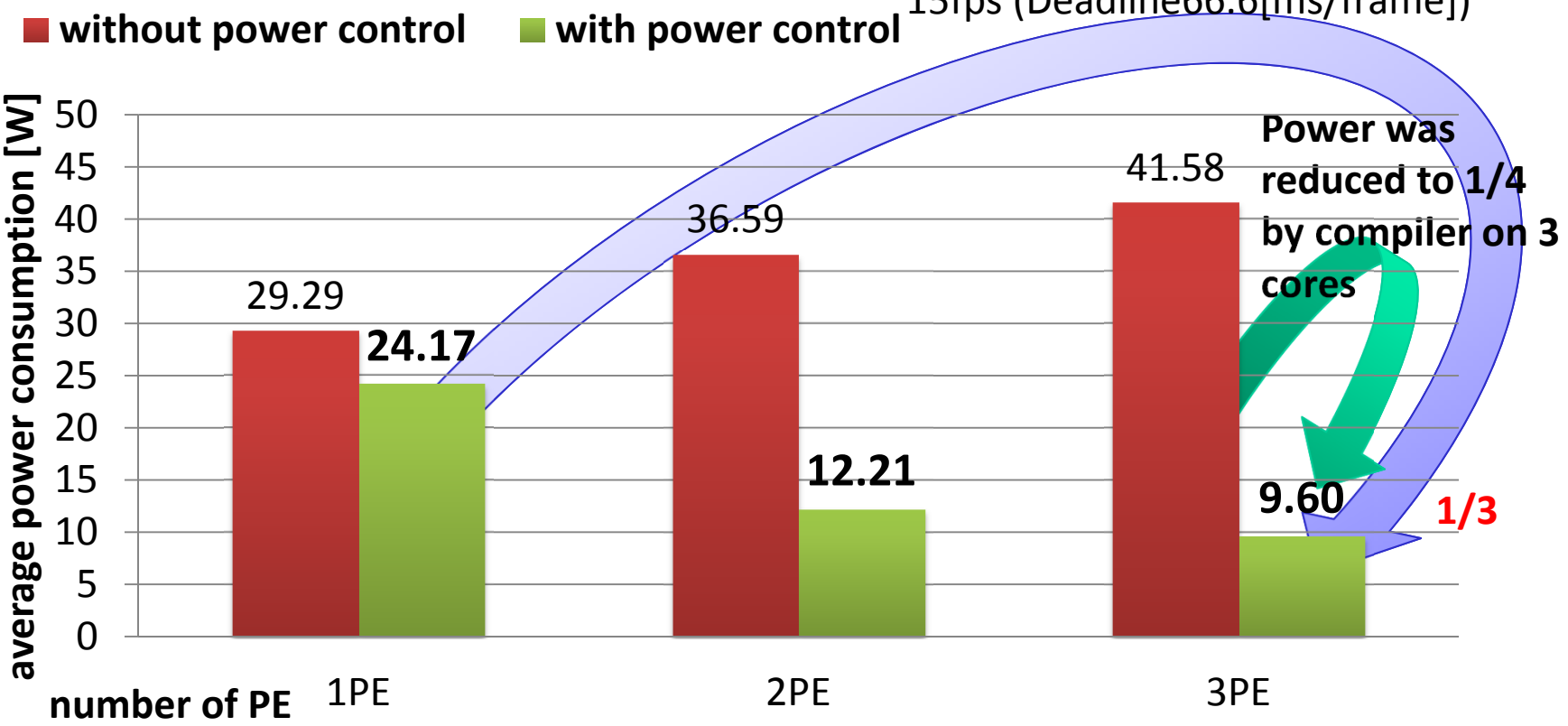
- On 3 cores, Automatic Power Reduction control successfully reduced power to 1/7 against without Power Reduction control.
- 3 cores with the compiler power reduction control reduced power to 1/3 against ordinary 1 core execution.

Power Reduction on Intel Haswell for Real-time Optical Flow



Intel CPU Core i7 4770K

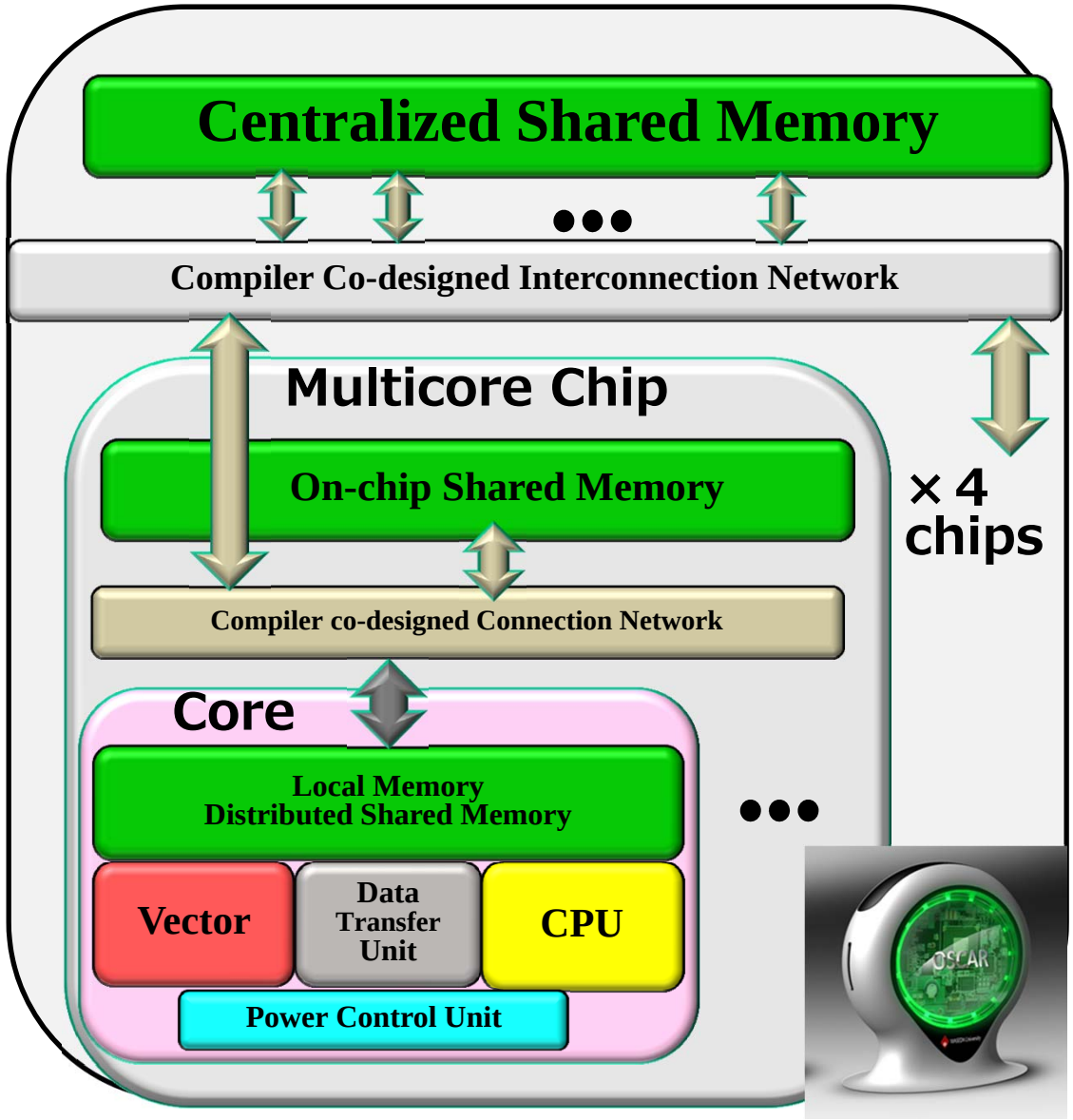
For HD 720p(1280x720) moving pictures
15fps (Deadline66.6[ms/frame])



Power was reduced to 1/4 (9.6W) by the compiler power optimization on the same 3 cores (41.6W).

Power with 3 core was reduced to 1/3 (9.6W) against 1 core (29.3W) .

OSCAR Vector Multicore and Compiler for Embedded to Severs with OSCAR Technology



Target:

- **Solar Powered with compiler power reduction.**
- **Fully automatic parallelization and vectorization including local memory management and data transfer.**

Summary

- Waseda University Green Computing Systems R&D Center supported by METI has been researching on low-power high performance Green Multicore hardware, software and application with government and industry including Hitachi, Fujitsu, NEC, Renesas, Denso, Toyota, Olympus and OSCAR Technology.
- OSCAR Automatic Parallelizing and Power Reducing Compiler has succeeded speedup and/or power reduction of scientific applications including "Earthquake Wave Propagation", medical applications including "Cancer Treatment Using Carbon Ion", and "Drinkable Inner Camera", industry application including "Automobile Engine Control", "Smartphone", and "Wireless communication Base Band Processing" on various multicores from different vendors including Intel, ARM, IBM, AMD, Qualcomm, Freescale, Renesas and Fujitsu.
- In automatic parallelization, 110 times speedup for "Earthquake Wave Propagation Simulation" on 128 cores of IBM Power 7 against 1 core, 55 times speedup for "Carbon Ion Radiotherapy Cancer Treatment" on 64cores IBM Power7, 1.95 times for "Automobile Engine Control" on Renesas 2 cores using SH4A or V850, 55 times for "JPEG-XR Encoding for Capsule Inner Cameras" on Tilera 64 cores Tile64 manycore.
  - The compiler will be available on market from OSCAR Technology.
- In automatic power reduction, consumed powers for real-time multi-media applications like Human face detection, H.264, mpeg2 and optical flow were reduced to 1/2 or 1/3 using 3 cores of ARM Cortex A9 and Intel Haswell and 1/4 using Renesas SH4A 8 cores against ordinary single core execution.