Green Multicore Computing

WOI'21
WASEDA OPEN INNOVATION FORUM 2021

Prof. Hironori Kasahara, IEEE Fellow, IPSJ Fellow
Senior Executive Vice President, Waseda University
IEEE Computer Society President 2018
URL: http://www.kasahara.cs.waseda.ac.jp/

1980 BS, 82 MS, 85 Ph.D. , Dept. EE, Waseda Univ.
1985 Visiting Scholar: U. of California, Berkeley, 1986 Assistant Prof, 1988 Associate Prof, Waseda
1989-90 Research Scholar: U. of Illinois, Urbana-Champaign, Center for Supercomputing R&D,
1997 Prof., 2004 Director, Advanced Multicore Research Institute, Waseda
2017member: the Engineering Academy of Japan (2020 Board), Science Council of Japan

<Committees in Societies and Government: 263>
IEEE Computer Society: President 2018, Executive Committee(2017-2019), BoG(2009-14), Strategic Planning Committee Chair 2018, Multicore STC Chair (2012-), Japan Chair(2005-07),
IPSJ Chair: HG for Magazine. & J. Edit, Sig. on ARC.
【METI/NEDO】 Project Leaders: Multicore for Consumer Electronics, Advanced Parallelizing Compiler, and Green Computing, Chair: Computer Strategy Committee
【Cabinet Office】 CSTP Supercomputer Strategic ICT PT, Japan Prize Selection Committees, etc.
【MEXT】Info. Sci. & Tech. Committee, Supercomputers (Earth Simulator, HPCI Promo., Next Gen. Supercomputer K) Committees, JST Moonshot Project G3 Robot & AI Vice Chair,
【COCN】Board Member in Council of Competitiveness Nippon, etc.

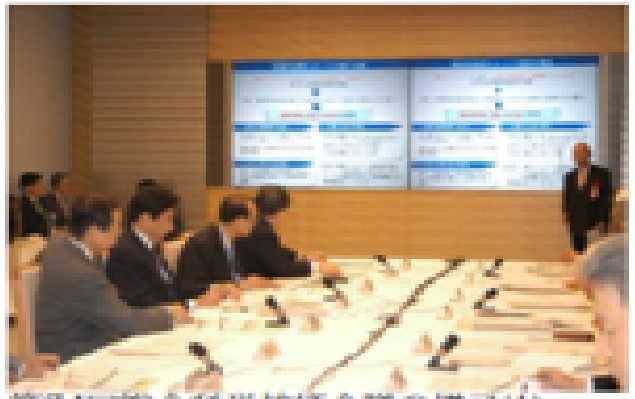
<Research Accomplishment>
Reviewed Papers: 221, Invited Talks: 190,
Granted Patents: 54 (Japan, US, GB, China),
Articles in News Papers, Web News, TV etc.: 625

<AWARD>
1987 IFAC World Congress Young Author Prize
1997 IPSJ Sakai Special Research Award,
2005 STARC Academia-Industry Research Award,
2008 LSI of the Year Second Prize,
2008 Intel Asia Academic Forum Best Research Award,
2010 IEEE CS Golden Core Member Award
2014 Minister of Edu., Sci. & Tech. Research Prize
2015 IPSJ Fellow, 2017 IEEE Fellow, Eta Kappa Nu
2019 Spirit of IEEE Computer Society Award,
2020 IPSJ Contribution Award

Demo of NEDO Green Multicore Processor for Real Time Consumer Electronics at Council of Science and Engineering Policy on April 10, 2008

http://www8.cao.go.jp/cstp/gaiyo/honkaigi/74index.html

第74回総合科学技術会議【平成20年4月10日】



第74回総合科学技術会議の様子(1)

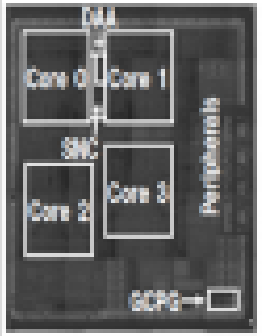
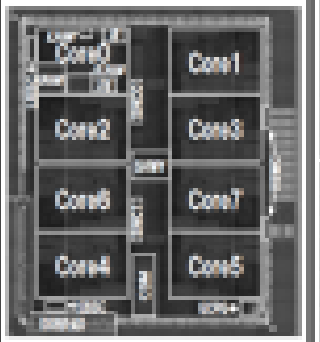
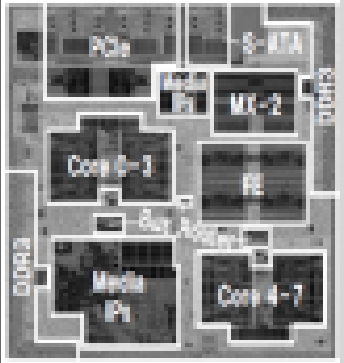
第74回総合科学技術会議の様子(2)

第74回総合科学技術会議の様子(3)

第74回総合科学技術会議の様子(4)

Codesign of Compiler and Multiprocessor Architecture since 1985

4 core multicore RP1 (2007), 8 core multicore RP2 (2008) and 15 core Heterogeneous multicore RPX (2010) developed in NEDO Projects with Hitachi and Renesas

RP-1 (ISSCC2007 #5.3)	*RP-2* (ISSCC2008 #4.5)	*RP-X* (ISSCC2010 #5.3)
		
90nm, 8-layer, triple-Vth, CMOS	90nm, 8-layer, triple-Vth, CMOS	45nm, 8-layer, triple-Vth, CMOS
97.8 mm² (9.88 x 9.88 mm)	104.8 mm² (10.51 x 9.88 mm)	153.8 mm² (12.4 x 12.4 mm)
1.0V (internal), 1.8/3.3V (IO)	1.0-1.4V (internal), 1.8/2.5V (IO)	1.0-1.2V (internal), 1.2-3.3V (IO)
600MHz, 4.32 GIPS, 15.9 GFLOPS	600MHz, 8.64 GIPS, 33.9 GFLOPS	648MHz, 13.7GIPS, 115GOPS, 36.2GFLOPS
11.4 GOPS/W (32b換算)	18.3 GOPS/W (32b換算)	37.3 GOPS/W (32b換算)

Prime Minister FUKUDA is touching our multicore chip during execution.

OSCAR Parallelizing Compiler

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

Multigrain Parallelization (LCPC1991,2001,04)

coarse-grain parallelism among loops and subroutines (2000 on SMP), near fine grain parallelism among statements (1992) in addition to loop parallelism

Data Localization

Automatic data management for distributed shared memory, cache and local memory
(Local Memory 1995, 2016 on RP2,Cache2001,03)
Software Coherent Control (2017)

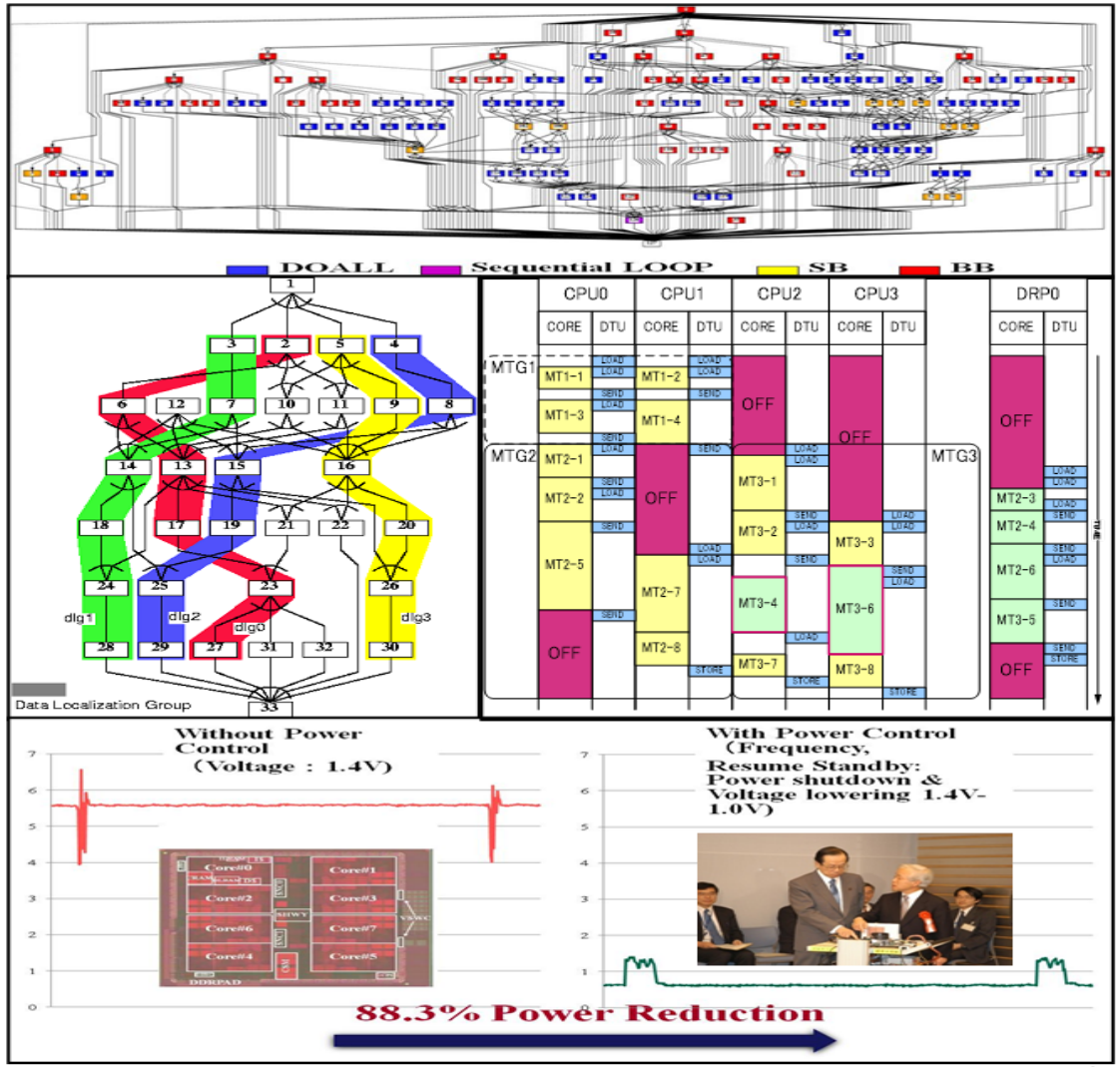
Data Transfer Overlapping (2016 partially)

Data transfer overlapping using Data Transfer Controllers (DMAs)

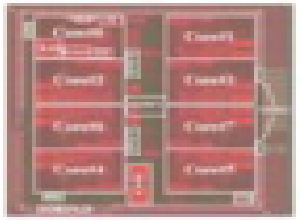
Power Reduction

(2005 for Multicore, 2011 Multi-processes, 2013 on ARM)

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

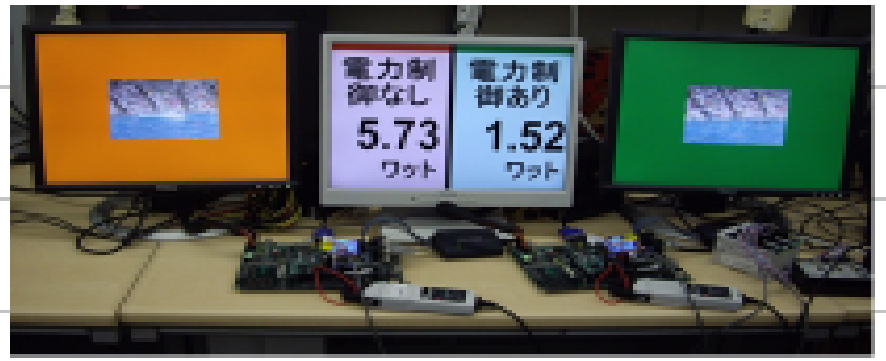


Power Reduction of MPEG2 Decoding to 1/4 on 8 Core Homogeneous Multicore RP-2 by OSCAR Parallelizing Compiler

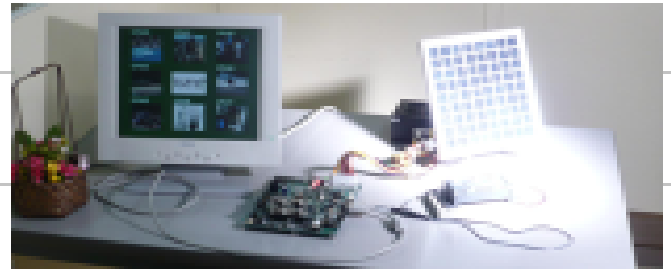


MPEG2 Decoding with 8 CPU cores

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]

IEEE Computer Society

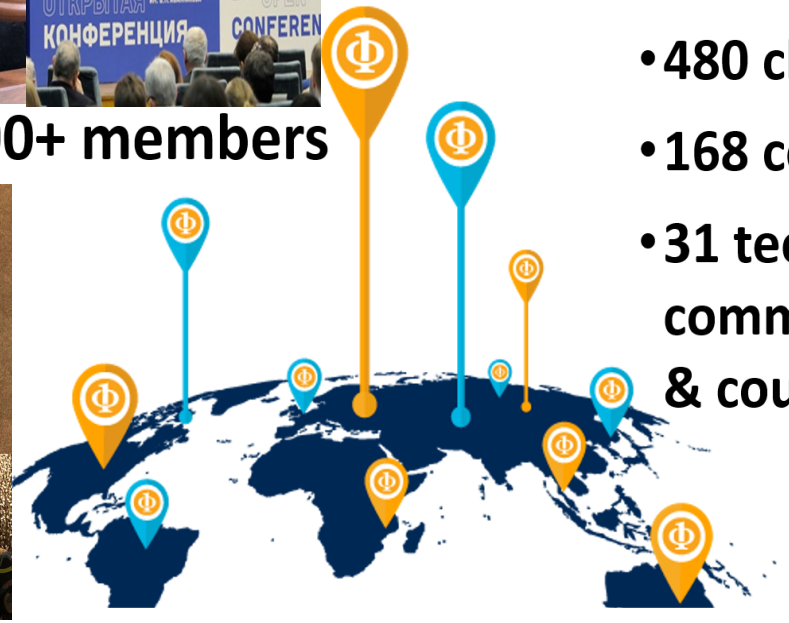
The first President from the outside of USA and Canada in 72 years history of IEEE CS

Bjarne Stroustrup: Morgan Stanley & Columbia Univ.
2018 IEEE Computer Society Computer Pioneer Award
IEEE COMPSAC2018 Keynote & Award Ceremony

July 26, 2018, Keynote, Hitotsubashi Hall

July 25, 2018 Award Ceremony
Rihga Royal Hotel Tokyo

- 84,000+ members

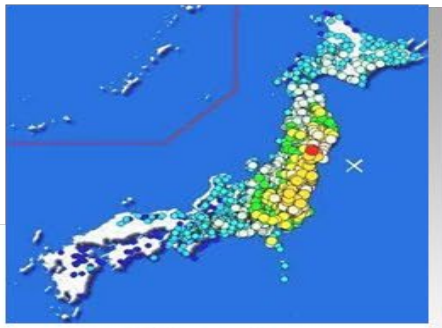


- 480 chapters
- 168 countries
- 31 technical committees & councils

IEEE

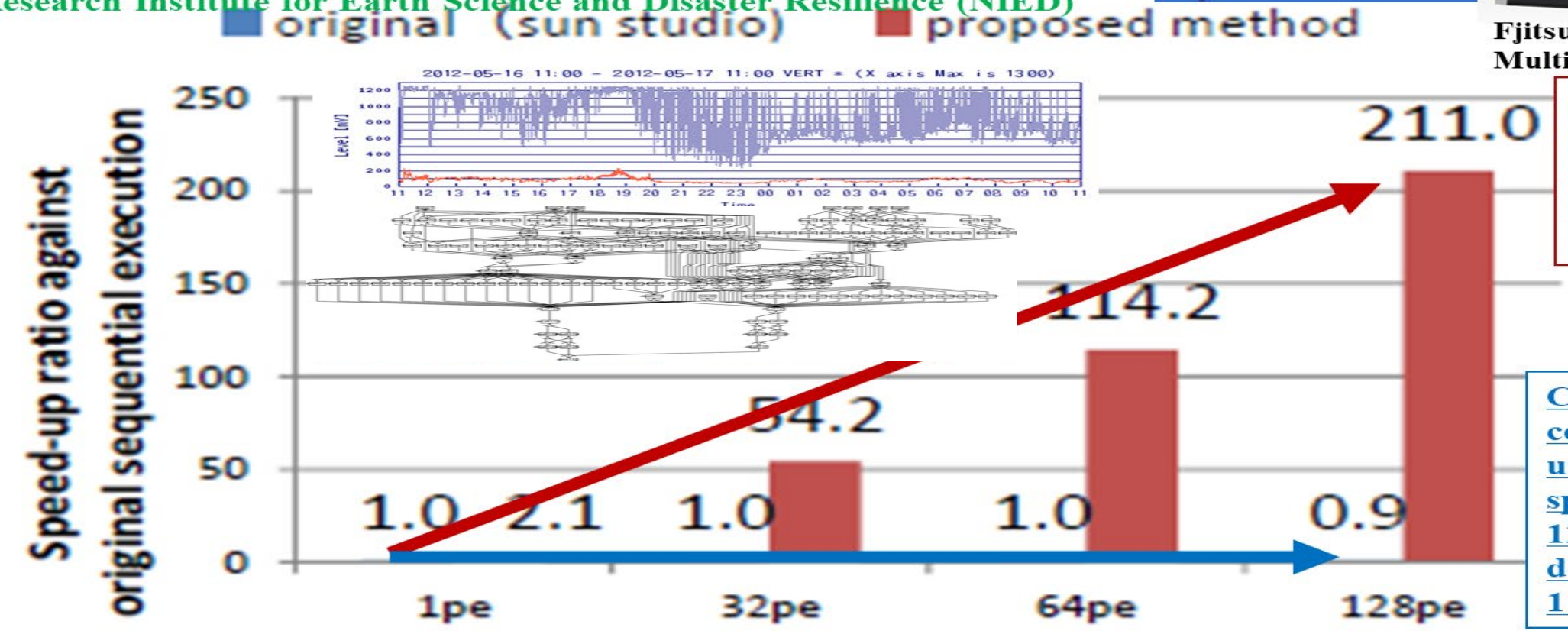
Parallel Soft is important for scalable performance of multicore (LCPC2015)

- Just more cores don't give us speedup
- Development cost and period of parallel software are getting a bottleneck of development of embedded systems, eg. IoT, Automobile



Fjitsu M9000 SPARC Multicore Server

Earthquake wave propagation simulation GMS developed by National Research Institute for Earth Science and Disaster Resilience (NIED)



OSCAR Compiler gives us 211 times speedup with 128 cores

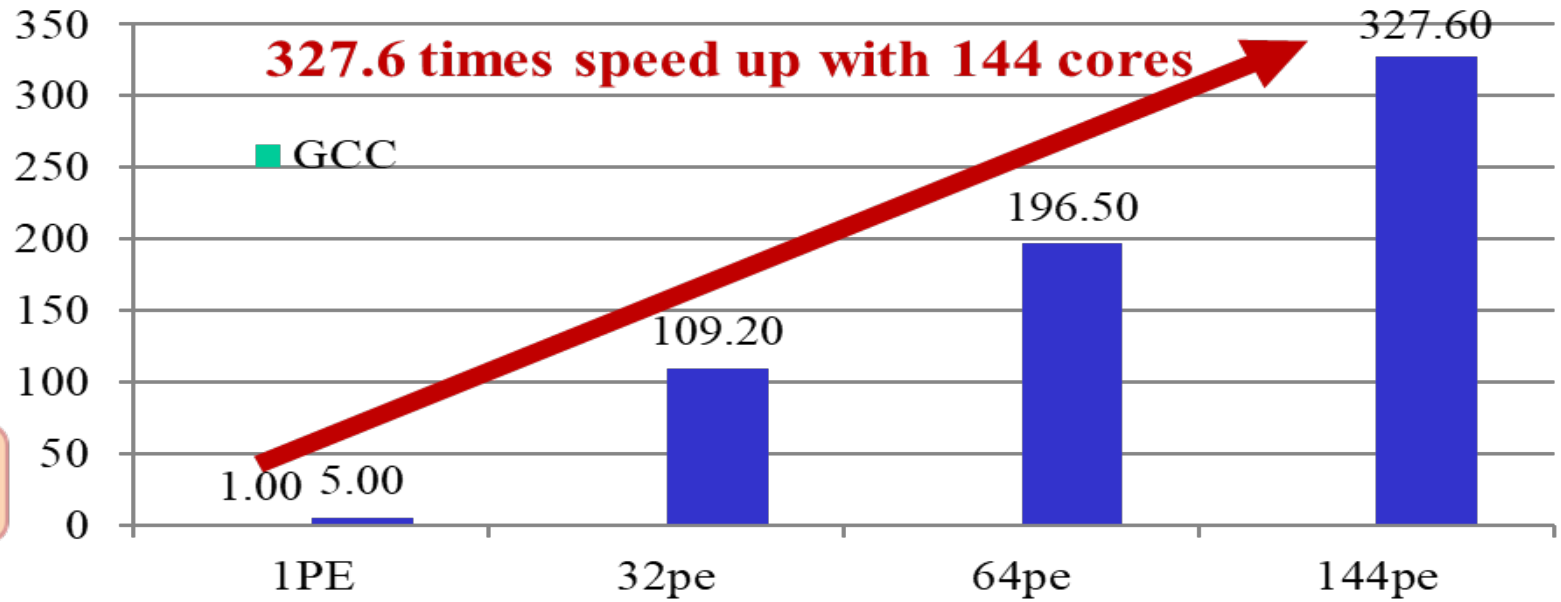
Commercial compiler gives us 0.9 times speedup with 128 cores (slow-downed against 1 core)

- Automatic parallelizing compiler available on the market gave us no speedup against execution time on 1 core on 64 cores
  - Execution time with 128 cores was slower than 1 core (0.9 times speedup)
- Advanced OSCAR parallelizing compiler gave us 211 times speedup with 128cores against execution time with 1 core using commercial compiler
  - OSCAR compiler gave us 2.1 times speedup on 1 core against commercial compiler by global cache optimization

Performance on Multicore Server for Latest Cancer Treatment Using Heavy Particle (Proton, Carbon Ion)

327 times speedup on 144 cores

Hitachi 144cores SMP Blade Server BS500:
Xeon E7-8890 V3(2.5GHz　18core/chip) x8 chip



- **Original sequential execution time 2948 sec（50 minutes）using GCC was reduced to 9 sec with 144 cores（327.6 times speedup）**
  - **Reduction of treatment cost and reservation waiting period is expected**

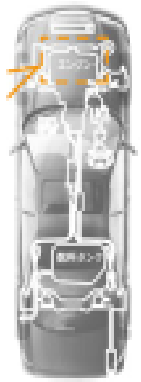
Hironori Kasahara has started collaboration with robot Group and research of hard-realtime Control and Simulation since 1982.

H. Kasahara, S. Narita, "Parallel Processing of Robot Arm Control Computation on a Multimicroprocessor System", IEEE Journal of Robotics and Automation, Vol. RA-1, No. 2, Jun. 1985.

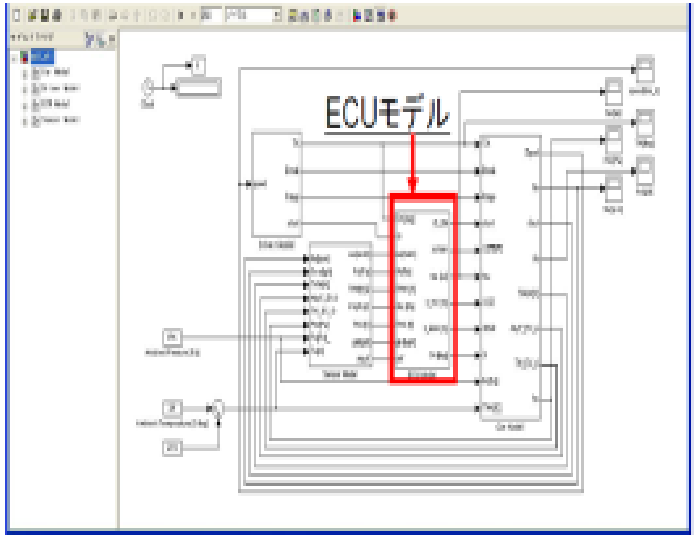
H. Kasahara, H. Fujii, M. Iwata, "Parallel Processing of Robot Motion Simulation", Proc. IFAC 10th World Congress, pp.329-336, Jul. 1987. "IFAC World Congress Young Author Prize", 1987.

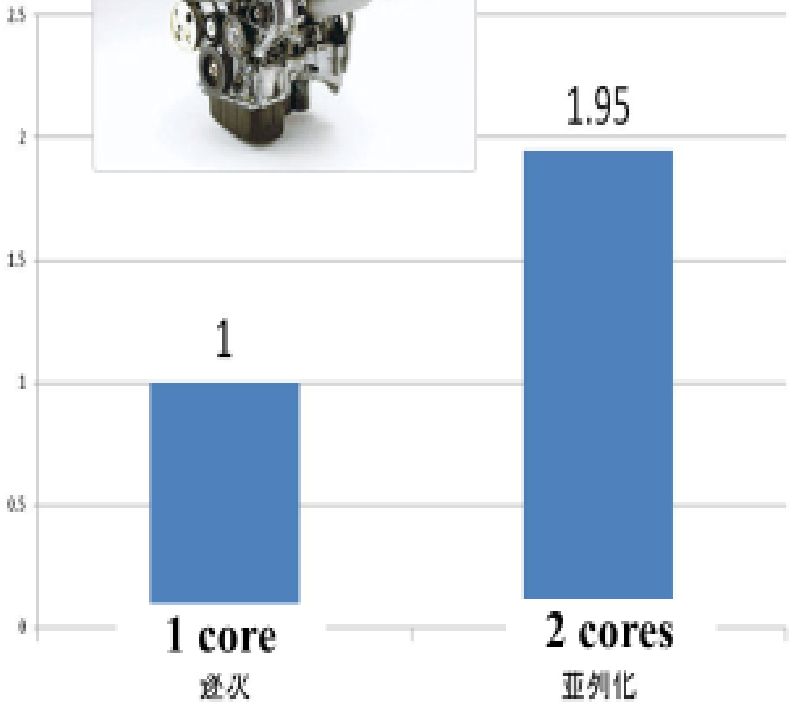
Engine Control by multicore with Denso

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.

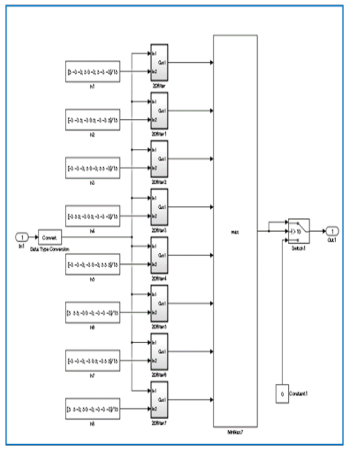


- Hard real-time automobile engine control by multicore using local memories
- Millions of lines C codes consisting conditional branches and basic blocks





OSCAR Compile Flow for Simulink Applications



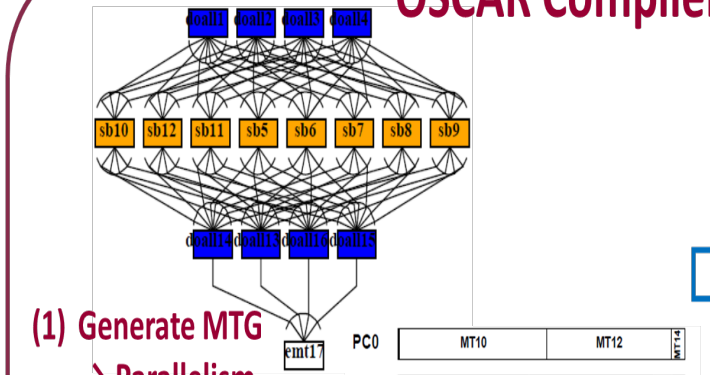
Simulink model

Generate C code
using Embedded Coder

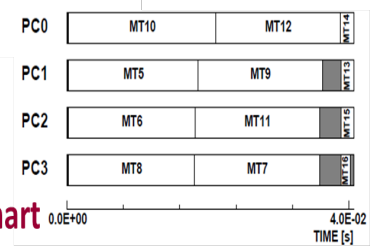
```
/* Model step function */
void VesselExtraction_step(void)
{
  int32_T i;
  real_T u0;

  /* 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



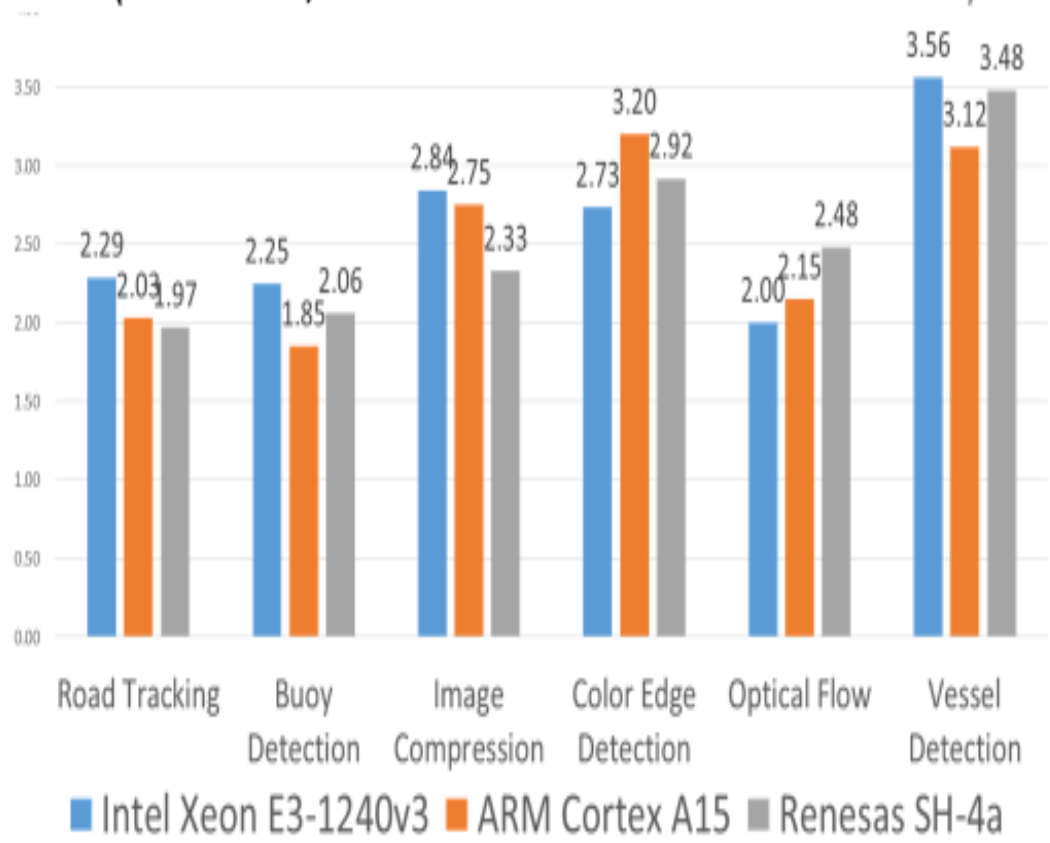
(2) Generate gantt chart
→ Scheduling in a multicore

```
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)

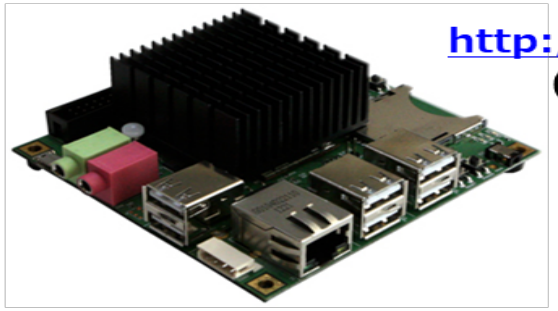
Speedups of MATLAB/Simulink Image Processing on Various 4core Multicores

(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/

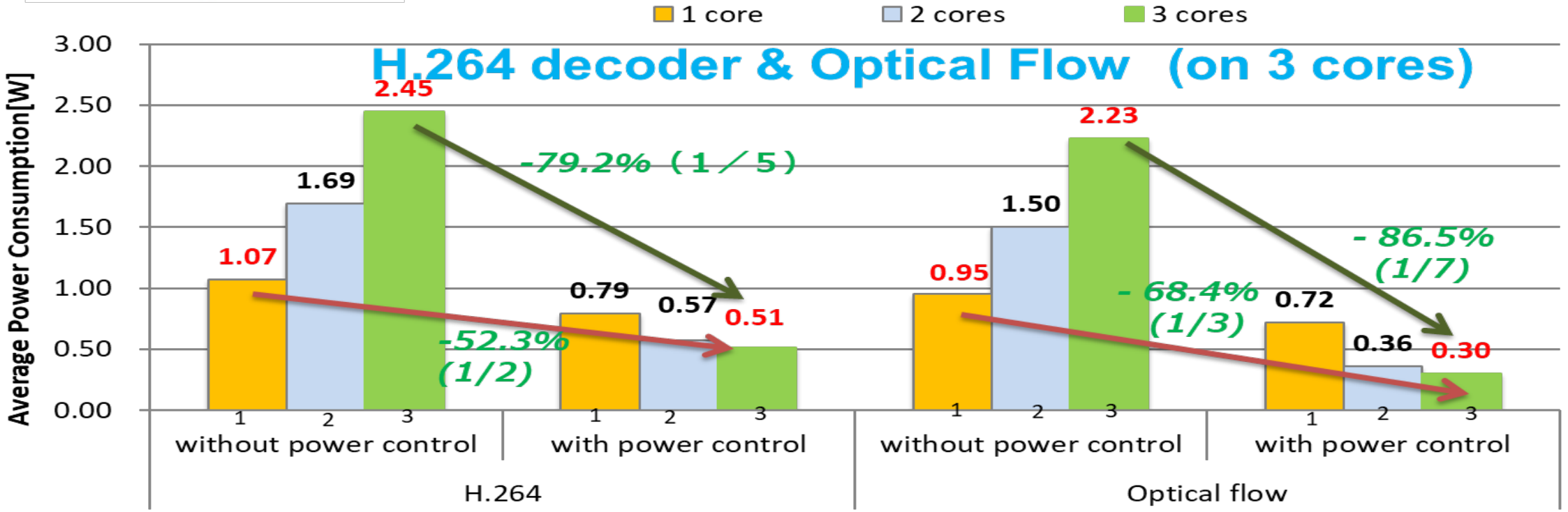
Automatic Power Reduction on ARM CortexA9 with Android



http://www.youtube.com/channel/UCS43INYEIkC8i_KIgFZYQBQ

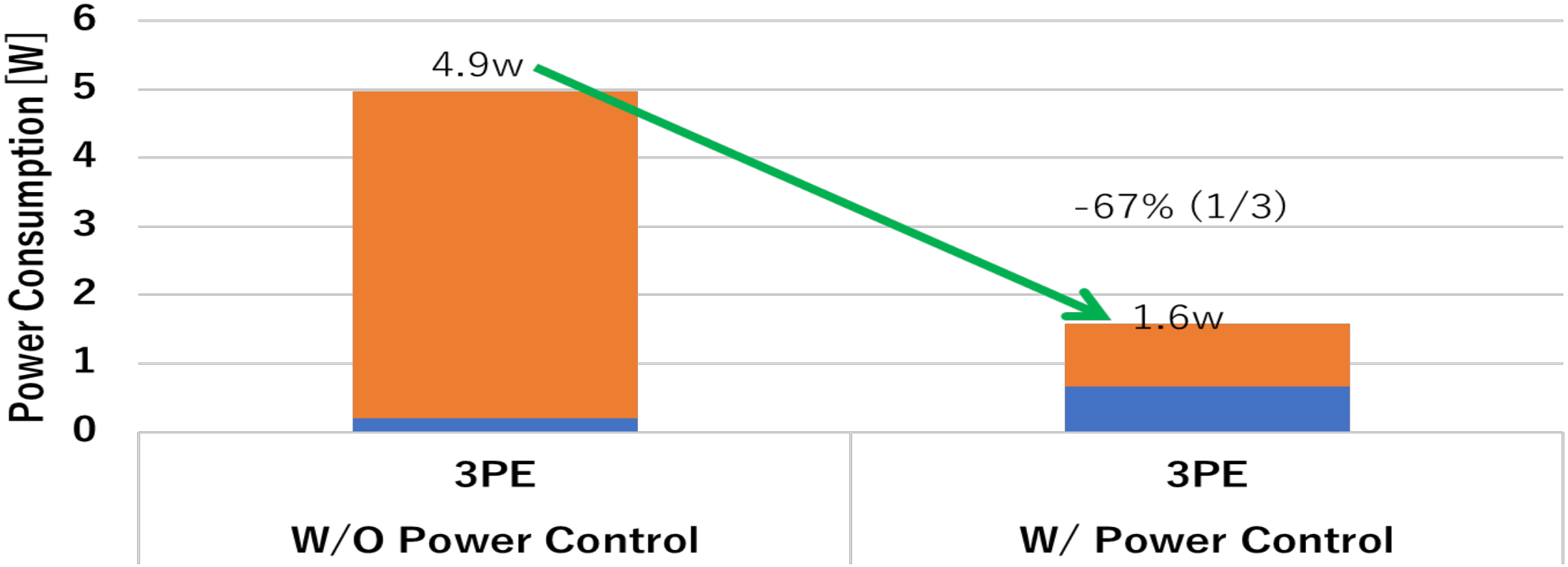
ODROID X2

Samsung Exynos4412 Prime, ARM Cortex-A9 Quad core
1.7GHz～0.2GHz, used by Samsung's Galaxy S3



Power for 3cores was reduced to 1/5～1/7 against without software power control
Power for 3cores was reduced to 1/2～1/3 against ordinary 1core execution

Automatic Power Reduction of OpenCV Face Detection on big.LITTLE ARM Processor



- ODROID-XU3
  - Samsung Exynos 5422 Processor
    - 4x Cortex-A15 2.0GHz, 4x Cortex-A7 1.4GHz big.LITTLE Architecture
    - 2GB LPDDR3 RAM

Frequency can be changed by each cluster unit

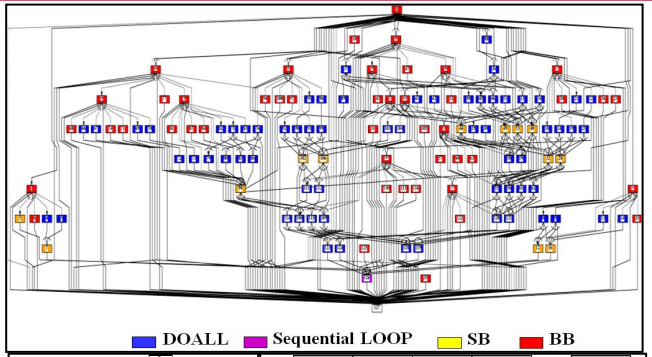
Patents related with OSCAR Parallelizing Compiler & Hardware

58 international patents in USA, UK, China, Japan to improve effective performance, cost-performance and software productivity and reduce power

High Performance & Low Power

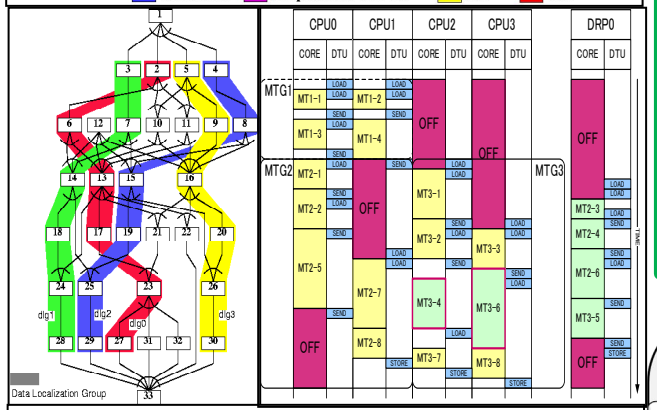
1) Multigrain Parallelization for Embedded and Heterogeneous Multicores & Hardware Supports (Synchronization)

coarse-grain parallelism among loops, subroutines & basic blocks among statements in addition to loop parallelism



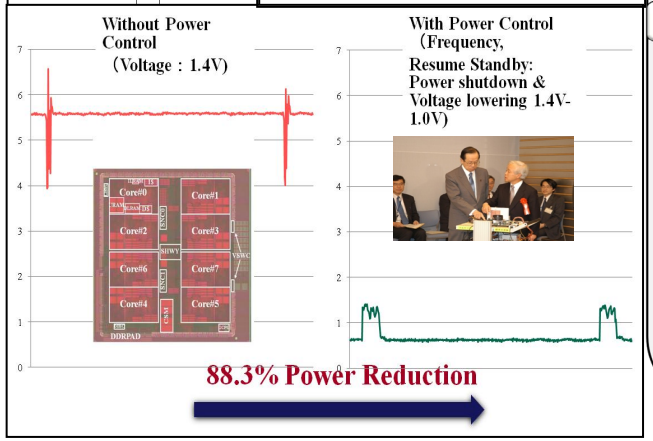
2) Data Localization: Cache & Local Memory Optimization

- Automatic data management for distributed shared memory, cache and local memory
- Software Cache Coherent Control
- Data Transfer Controller
- Data transfer overlapping using DMA hardware & its compiler



3) Automatic Power Reduction HW&SW

- Reduction of consumed power by compiler control DVFS and Power gating with hardware supports for Carbon Neutral.
- Accelerator See right figure



Green Accelerator

can be attached to any processor cores, RISCV, arm, Infineon Renesas, AMD, Intel, etc, without instruction extensions.

- It works with automatic local memory management and power reduction control by OSCAR Automatic vectorizing & parallelizing compiler