OSCAB Automatic Parallelizing Compiler
Automatic Speedup and Power Reduction
Kasahara & Kimura Lab, Waseda University, TOKYO
http://www.kasahara.cs.waseda.ac.jp

Parallel Processing of MATLAB/Simulink by OSCAR Compiler
on Intel, ARM & Renesas multi cores
Kasahara & Kimura Lab, Waseda University, TOKYO

Automatic Parallelization of MATLAB/Simulink by OSCAR Compiler

Automatic Localization
LR: Localizable Region (data can be passed through LDM)
CAR: Commonly Accessed Region (possibly data transfers are required among processors)
DLG: Data Localization Group (group of loops accesses the same iteration ranges)

Speedups by OSCAR Compiler’s Local Memory Management

(1) Generate MTG → Parallelism
(2) Generate gantt chart → Scheduling in a multicore
(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)

SC16 members
A. Maruoka
Y. Mushu
S. Karino
K. Miyamoto
T. Kawata
K. Yamamoto
T. Shirakawa
Y. Oki
T. Kitamura
M. Takamura
K. Kimura
H. Kasahara

Task graph of OSCAR API Program
Execution environment: Hitachi SR16000 Model VM1 (IBM POWER7 Processor: 128core)
**Platinum Multicore Architecture**

- CPU
- Vector ACC (Vector Reg, LD/ST Unit, Scalar unit)
- Core 1
- Core 2
- Core 3
- LDM (DSM, Interconnect)
- Network Interface
- On-chip Interconnection Network
- On-chip CSM

**OSCAR Compiler Flow w/ OSCAR API**

- OSCAR Compiler
  - CPU Program
  - OSCAR Compiler
    - Multi-Grain Parallelization
    - Local Memory Management
    - Data Transfer Optimization
    - Power Management
    - Auto Vectorization
  - Vector ACC Program
  - Clang/LVM
  - CPU Object Code
  - Executable File
    - For quick implementation of vector back end
    - Vector + Inside Scalar Object Code

**Vector Accelerator**

- Features
  - Attachable for any CPUs (Intel, ARM, IBM)
  - Data driven initiation by sync flags

- Function Units [tentative]
  - Vector Function Unit
    - 8 double precision ops/cycle
    - 64 characters ops/clock
    - Variable vector register length
    - Chaining LD/ST & Vector pipes
  - Scalar Function Unit
  - Registers [tentative]
    - Vector Register 256Bytes/entry, 32entry
    - Scalar Register 8Bytes/entry
    - Floating Point Register 8Bytes/entry
    - Mask Register 32Bytes/entry

**Performance w/ OSCAR Simulator**

- Evaluate performance of Vector Accelerator with cycle accurate simulator
- Simulation Settings
  - 16 single precision ops/cycle
  - Local Data Memory Bandwidth 32 byte/clock
  - All data located on Local Data Memory in initial state

Applications
- Matmul:
  - Size: 256x256, Data type: float
  - 2D Convolution:
  - Size: 256x256, Data type: float, kernel size: 3x3

**Power Reduction of Face Recognition on Intel Haswell 3 cores by OSCAR Compiler - Reduced Power to 2/5 on Intel-**

- Kasahara & Kimura Lab, Waseda University, TOKYO
- http://www.kasahara.cs.waseda.ac.jp

**Saving Energy by Compiler**

- PE0: MT1 Frequency: 100%
- PE1: MT2 Frequency: 100%
- PE2: MT3 Frequency: 100%

- Data Dependence:
  - MT1: time
  - MT2: Dead Line
  - MT3: Critical Path

**Measuring Environment**

- CPU: Intel Core i7 4770K
- No. of Cores: 4
- Frequency: 3.5GHz~0.8GHz
- Motherboard: ASUS H81M-A

**Speedup and Power reduction on Intel Haswell 3 Cores**

- Speedup at Fastest Execution Mode
  - Speedup
    - Matmul: x20.1
    - 2D Convolution: x22.2

- Average Power Consumption at Power Reduction Mode
  - Without power control: 41.72
  - With power control: 16.14
  - Reduced to 2/5 (-61.31%)

**Face Recognition program**

- Detecting Human Face in Input Image from Camera and Marking

**OSCAR Compiler**

- Intel Haswell
- Power Reduction