Ever-Increasing Variability

Reduced Guardband Causes Timing Errors

Online Meta-data Characterization

Distributed Variability- and Load-Aware Scheduler (DVLS)

* DVLS saves energy up to 44% (30% on average) compared to RRS

* DVLS saves time up to 47% (33% on average) compared to RRS

Key Idea

Time and Energy Saving

Application:

#pragma omp parallel {
#pragma omp task
run_add (t, TASK_SIZE);
#pragma omp taskwait
for (int t = 0; t < T;j; t++)
#pragma omp task
run_shift (t, TASK_SIZE);
#pragma omp taskwait
for (int t = 0; t < T;j; t++)
#pragma omp task
run_mul (t, TASK_SIZE);
#pragma omp taskwait
for (int t = 0; t < T;j; t++)
#pragma omp task
run_div (t, TASK_SIZE);
#pragma omp taskwait
}

Applicationx

Shared memory cluster

Core

Low-latency Interconnect

L1 I$ mem (TCDM)

EDA? C

Core 0 ... Core 15

... Task type 4 Task type 4

Task Execution Cost (taski, corej) = #I (taski, corej) + #R (taski, corej)

Meta-data Lookup Table for Applicationx

CORES

Decoupled Master Queue

Load-balanced Distributed Queues

DVLS saves the computational load of centralized scheduler from one master thread to multiple slave threads

* Master thread simply pushes tasks into a decoupled queue

* Slave threads pick up a task from decoupled queue and push to one of the well-matched distributed queues

* For a taski, DVLS assigns a corej where TEC (taski, corej) + load, is minimum across the cluster

Variability in Embedded Shared Memory Clusters

Abbas Rahimi†, Daniele Cesarini‡, Andrea Marongiu‡*, Rajesh K. Gupta†, Luca Benini†*

†UC San Diego, ‡University of Bologna, *ETH Zurich

NSF Expedition in Computing, Variability-Aware Software for Efficient Computing with Nanoscale Devices

http://variability.org