Comparison of Single-ISA Heterogeneous versus Wide Dynamic Range Processors for Mobile Applications

Hamid Reza Ghasemi[1], **Ulya R. Karpuzcu**[2], Nam Sung Kim[3]

[1]University of Wisconsin [2]University of Minnesota [3]University of Illinois

33rd IEEE International Conference on Computer Design
ICCD 2015

Motivation

- Mobile applications demand a wide dynamic operating range
- How to provide a wide dynamic operating range?
  - Wide Dynamic Range (WDR) processors
    - Limited Dynamic Voltage Frequency (DVFS) scaling
      - Diminishing dynamic V (and thus F) range
      - Challenging minimum safe V scaling for memory
  - Single-ISA Heterogeneous Processors featuring two types of cores
    - Complex out-of-order (OoO) cores optimized for performance
    - Simple in-order cores optimized for power
  - Dynamic Resource Scaling (DRS)
    - Scaling processor resources to match workload characteristics
    - Selective activation of resources

Executive Summary

- We compare and contrast single-ISA heterogeneous and WDR architectures in terms of
  - Design challenges
  - Energy efficiency

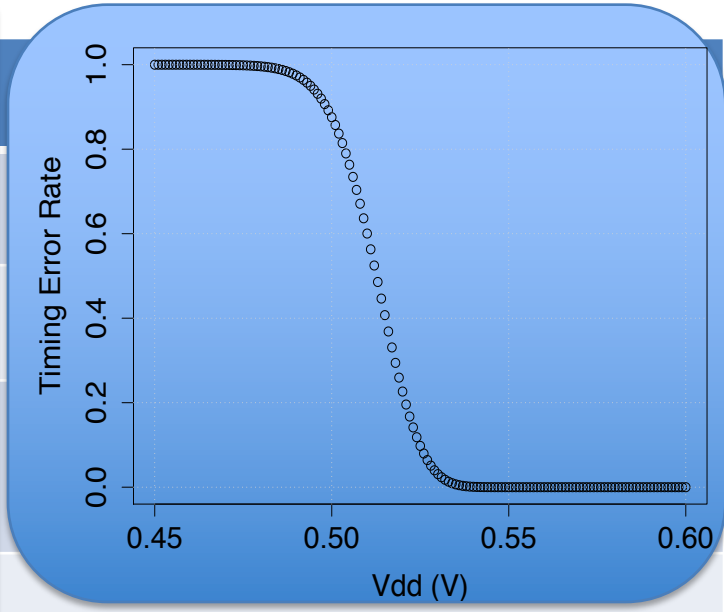
  considering implementations augmented w/ DRS.
- We observe that WDR processors augmented w/ DRS and single-ISA heterogeneous architectures can deliver similar energy efficiency.

Design Challenges

	WDR	Single-ISA Heterogeneous
Software support		Task migration Switching control between cores
Microarchitecture	Big OoO core	(big) OoO core + (little) in-order core
Cache design	Separate V domain Larger SRAM cells (possibly of higher transistor count)	Private L2 caches (for clusters of big and little cores)
Fully associative resources	Larger cells Separate V domains	
Combinational logic	Larger devices Separate V domains	
Hardware overhead	Multi-level shifters Power gating support Dedicated power management unit (PMU)	Coherent interconnection network

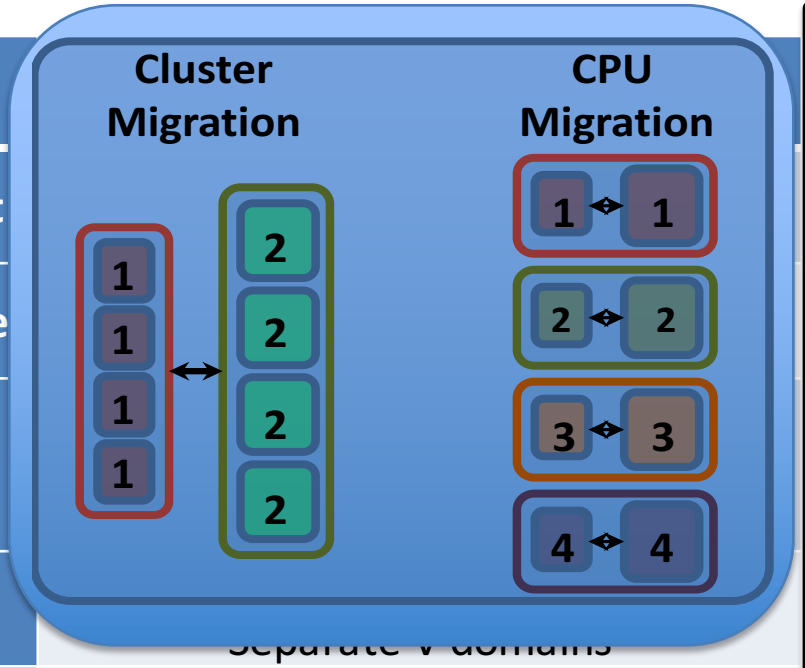
Design Challenges

	WDR
Software support	
Microarchitecture	**Big OoO core**
Cache design	Separate V domain Larger SRAM cells (possibly of higher transistor count)
Fully associative resources	**Larger cells** **Separate V domains**
Combinational logic	**Larger devices** Separate V domains
Hardware overhead	**Multi-level shifters** **Power gating support** **Dedicated power management unit (PMU)**



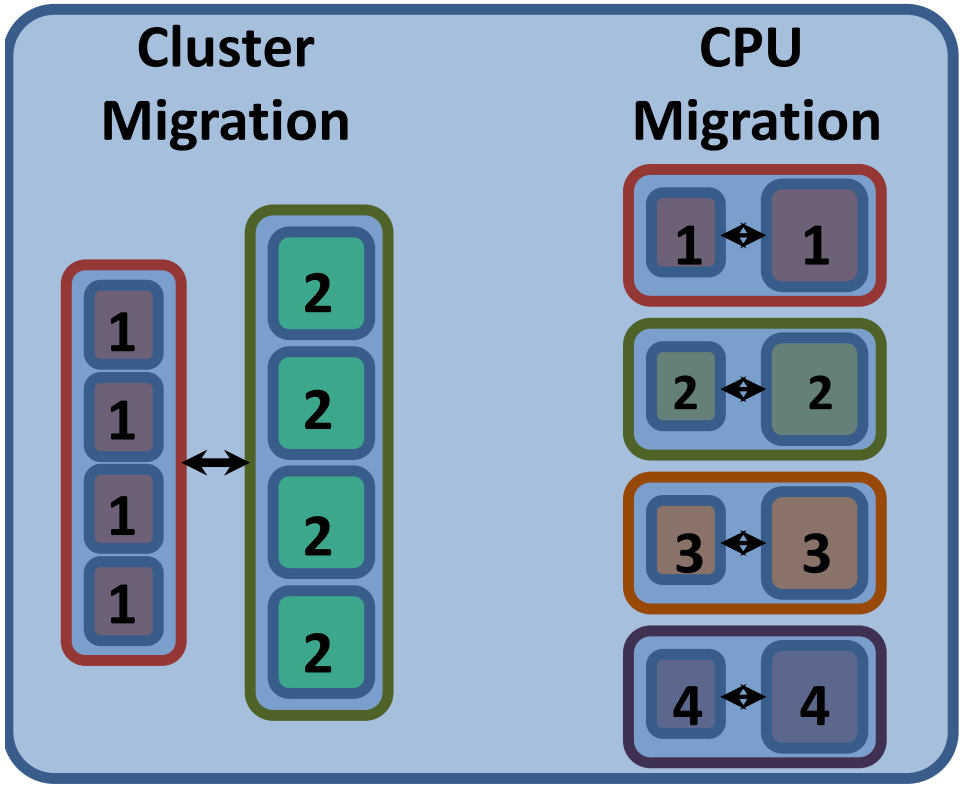
Coherent interconnection network

Design Challenges



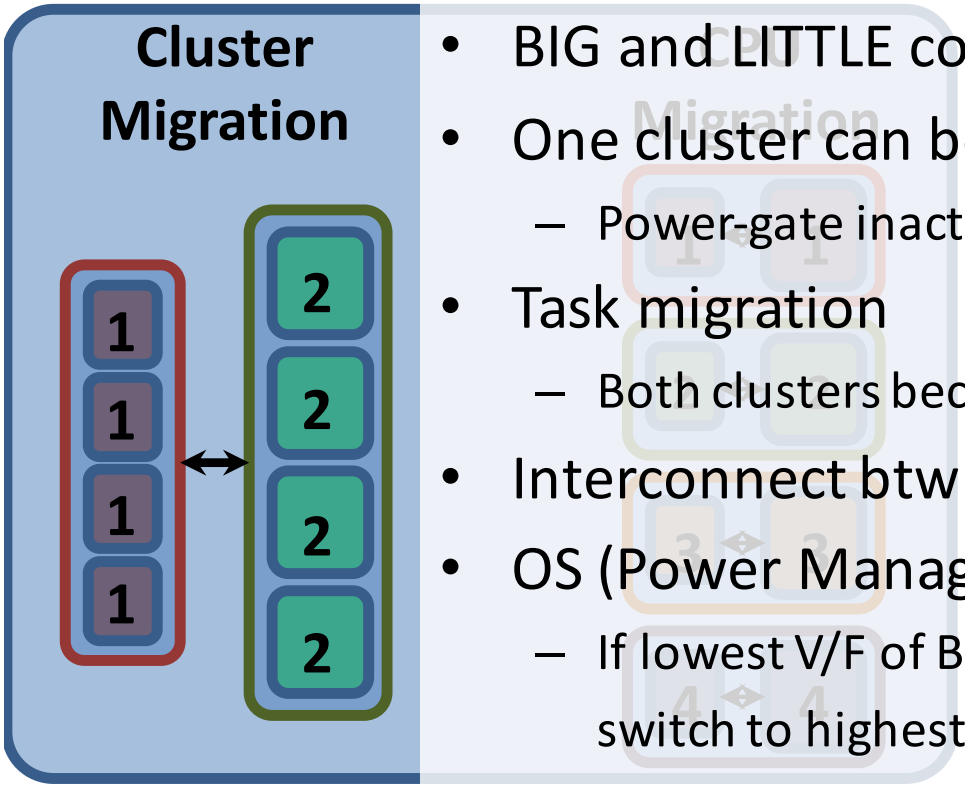
	Single-ISA Heterogeneous
Software support	**Task migration** **Switching control between cores**
Microarchitecture	**(big) OoO core** **+ (little) in-order core**
Cache design	Private L2 caches (for clusters of big and little cores)
Fully associative resources	
Combinational logic	Larger devices Separate V domains
Hardware overhead	Multi-level shifters Power gating support Dedicated power management unit (PMU)
	Coherent interconnection network

big.LITTLE: DynamicTask Scheduling



How to ensure data coherency upon migration?

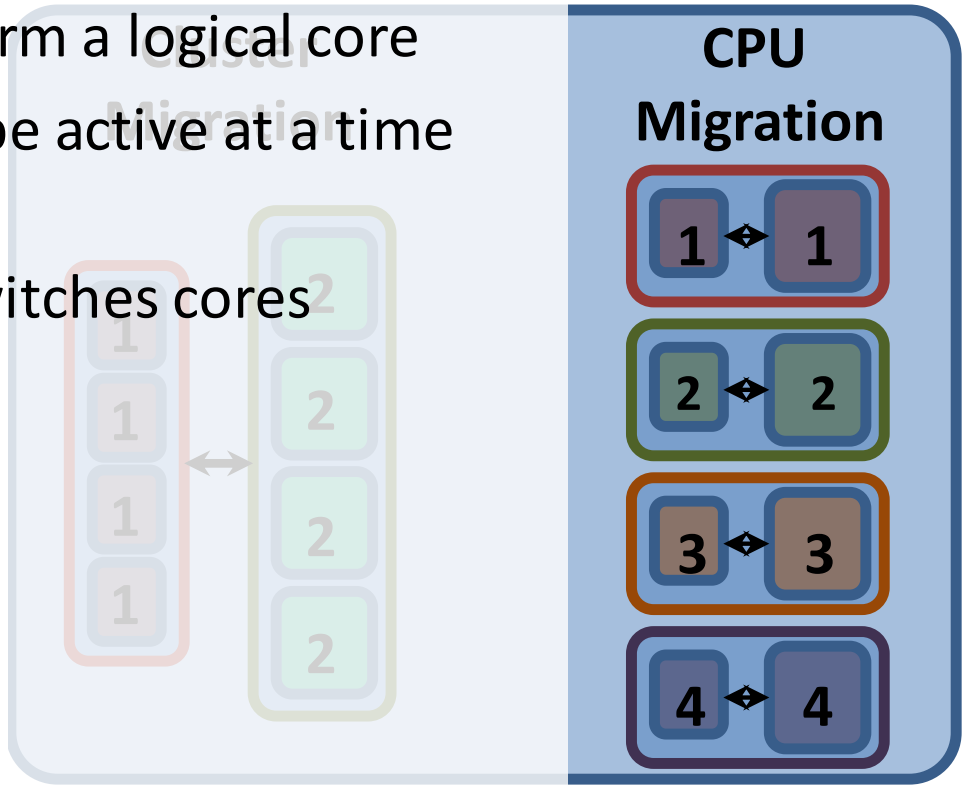
big.LITTLE: DynamicTask Scheduling



- BIG and LITTLE cores form separate clusters
- One cluster can be active at a time
  - Power-gate inactive cluster
- Task migration
  - Both clusters become active
- Interconnect btw clusters sets migration overhead
- OS (Power Management Unit) switches clusters
  - If lowest V/F of BIG delivers lower energy efficiency, switch to highest V/f of LITTLE

big.LITTLE: DynamicTask Scheduling

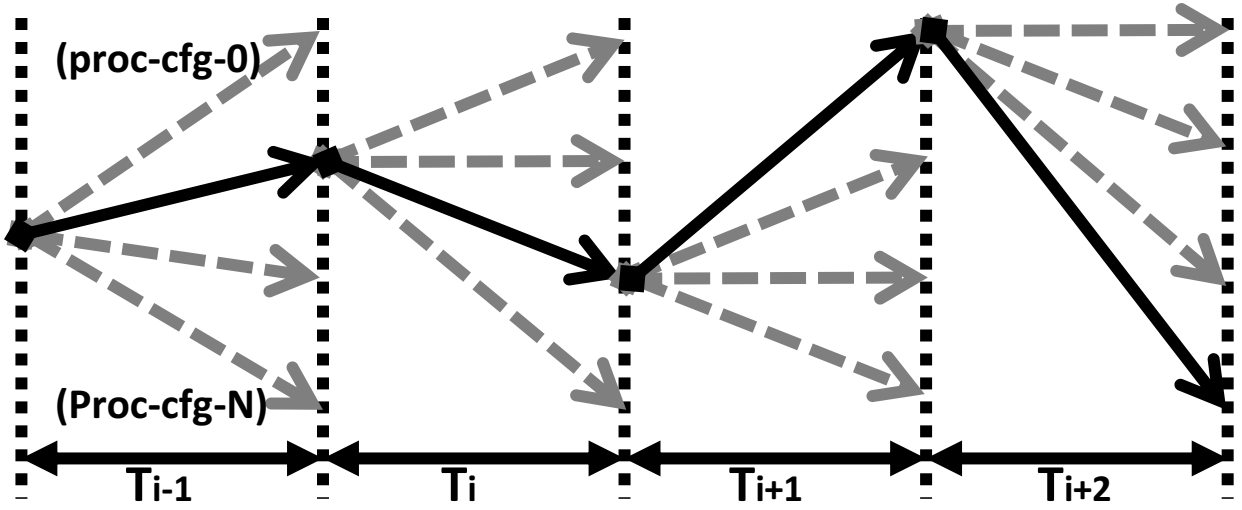
- BIG and LITTLE cores paired to form a logical core
- One physical core (per pair) can be active at a time
  - Power-gate inactive core
- OS (Power Management Unit) switches cores
  - Track load on a per core basis



Evaluation Setup

- Platforms
  - WDR as ARM Cortex-A15
  - ARM's big.LITTLE as single-ISA heterogeneous processor Cortex-A15 + Cortex-A17
- Gem5 cycle accurate full-system simulator
  - Augmented with oracular power managament policy
- McPAT to model power consumption at 22nm
- Benchmarks
  - Moby suite
    - Includes Popular Android applications from Google Play Store
  - Single-threaded applications from SPEC2006

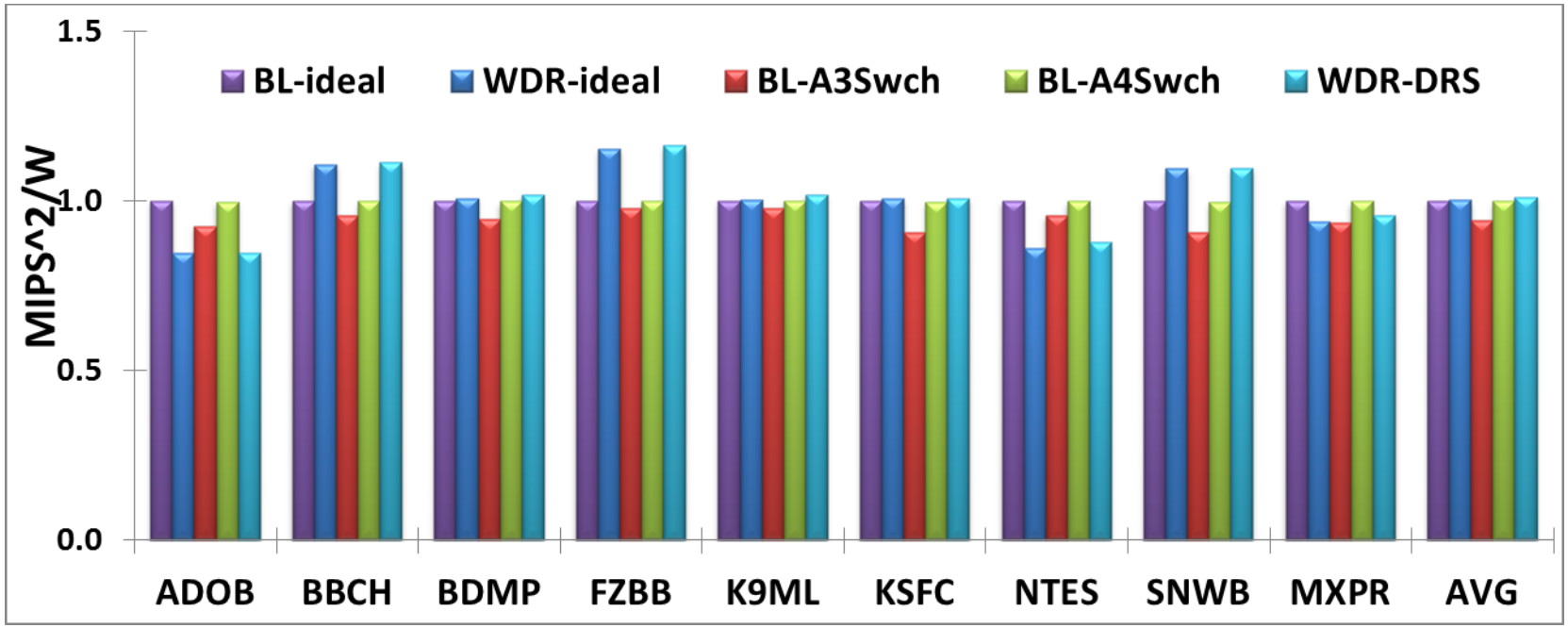
Oracular Greedy Power Management



- Covers big.LITTLE, WDR, WDR+DRS platforms
- Find the most energy-efficient processor configuration per interval
  - Simulate all processor configurations per quantum
  - Select the configuration of maximum energy efficiency
  - Next interval starts from the checkpoint induced by selected configuration

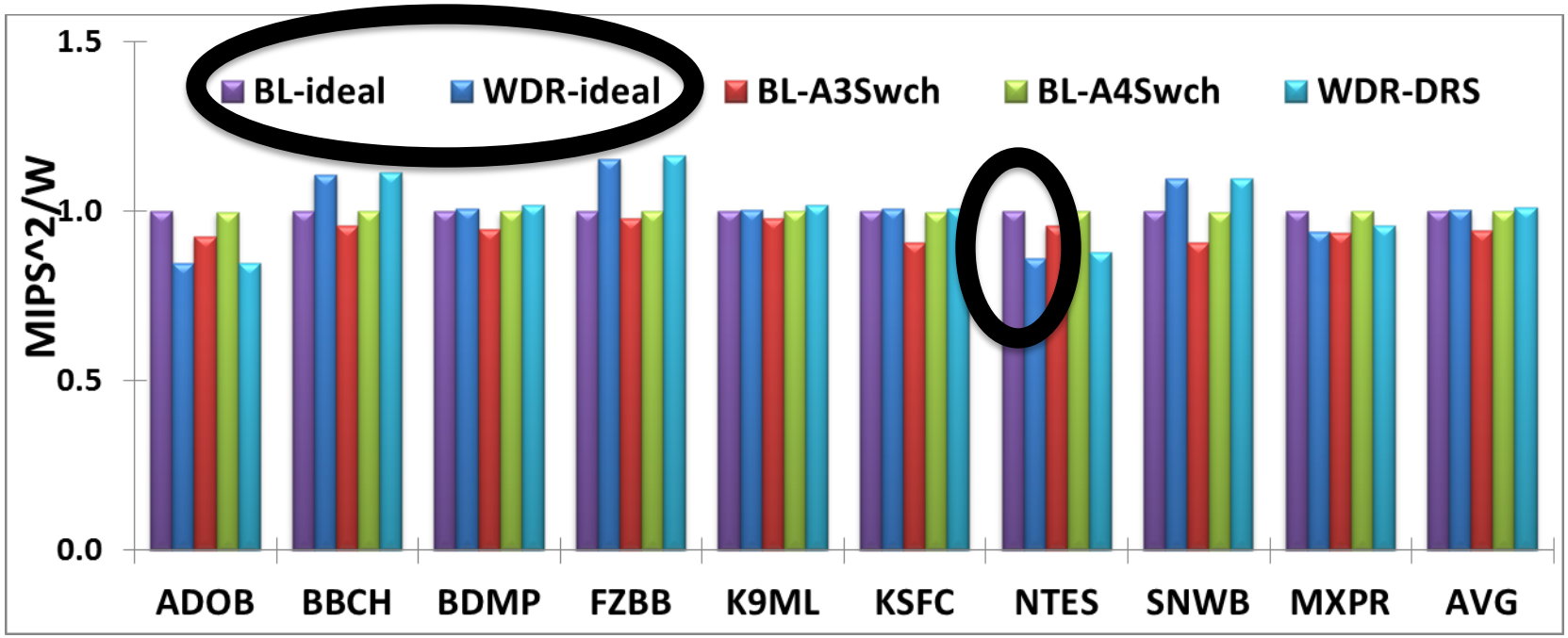
33rd IEEE International Conference on Computer Design
ICCD 2015

Energy Efficiency



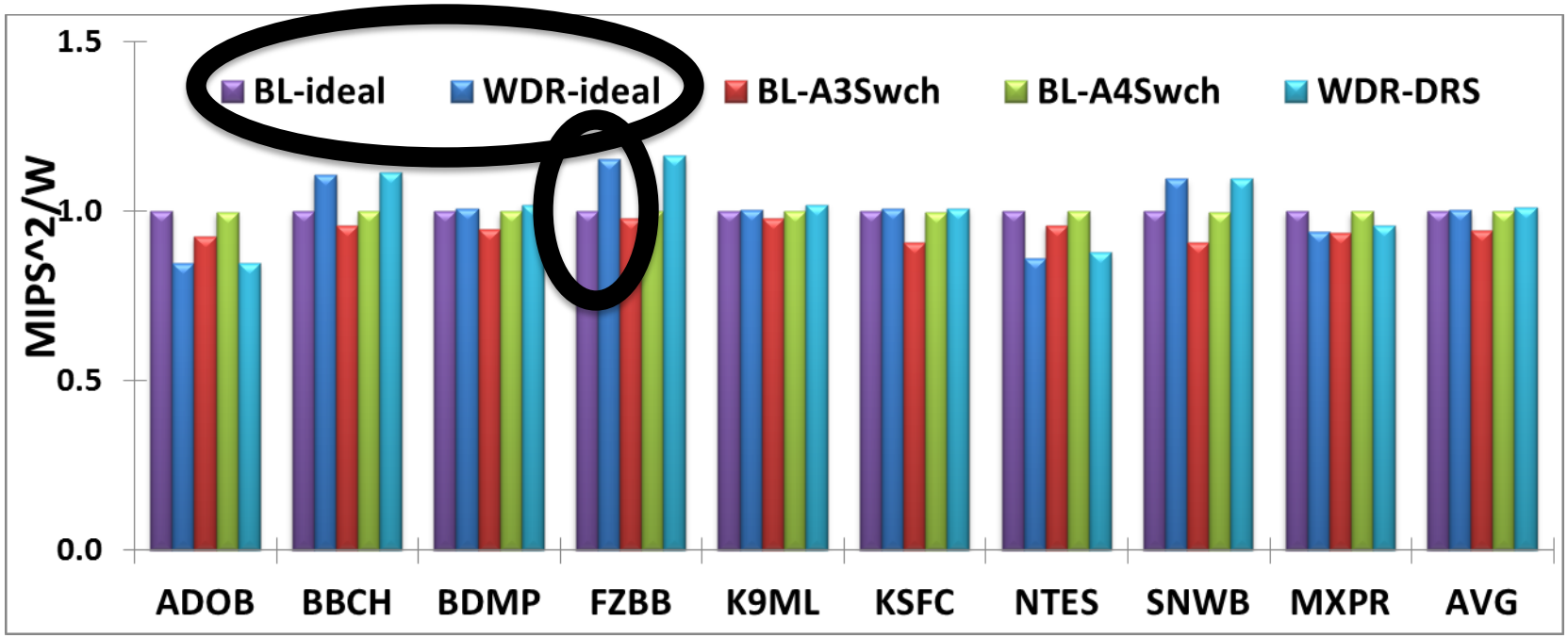


Energy Efficiency



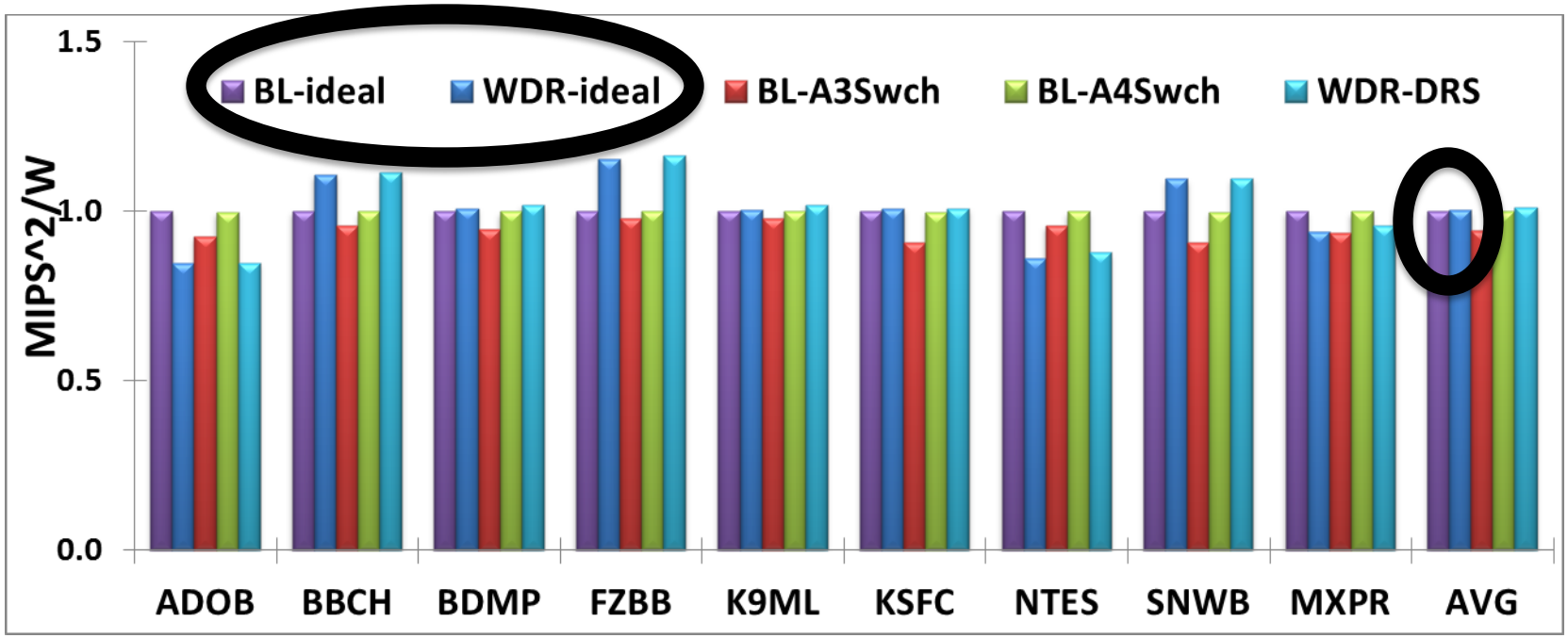
NTES: BL-ideal is more energy-efficient than WDR-ideal by 16%

Energy Efficiency



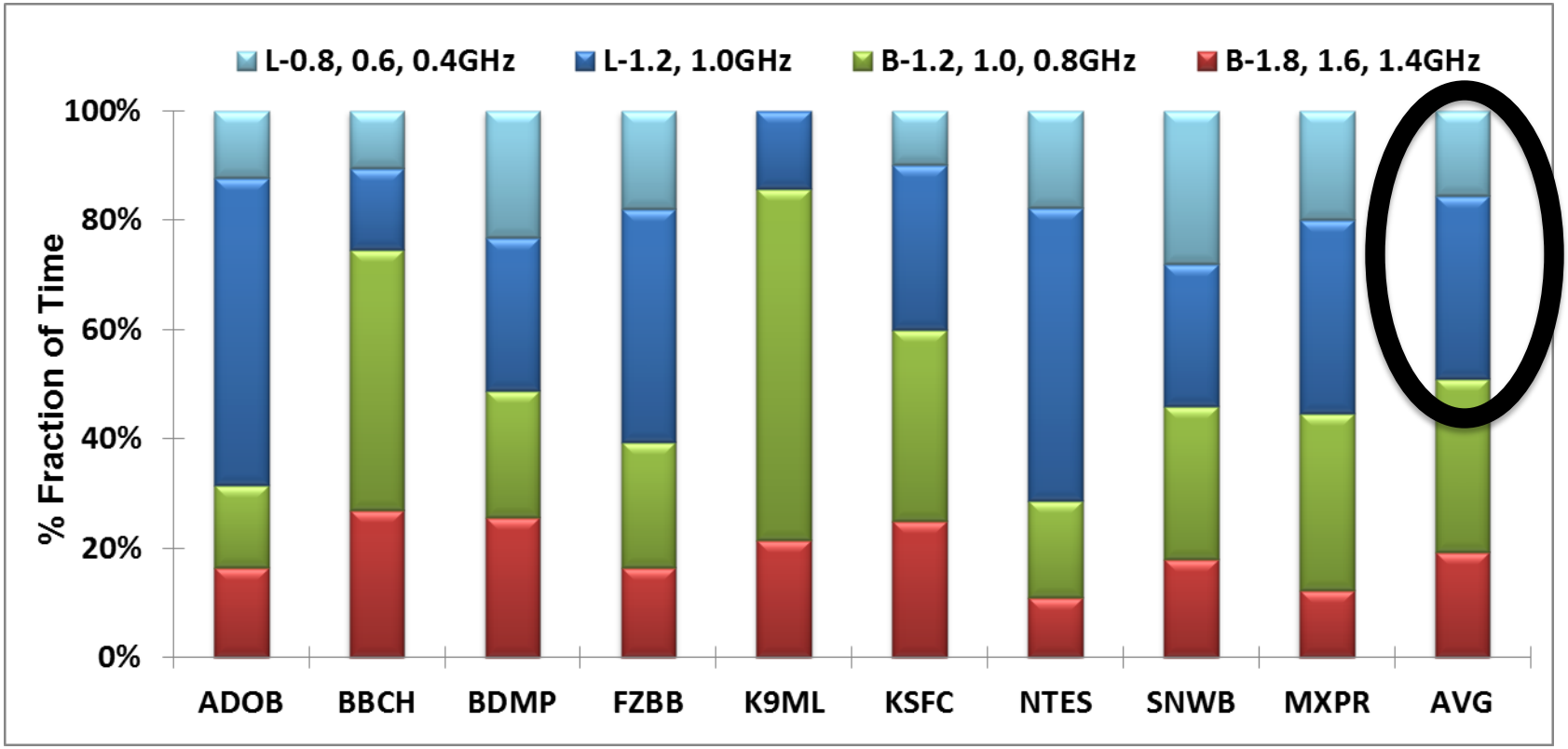
FZBB: WDR-ideal is more energy-efficient than BL-ideal by 15%

Energy Efficiency



On average, WDR-ideal is as energy-efficient as BL-ideal

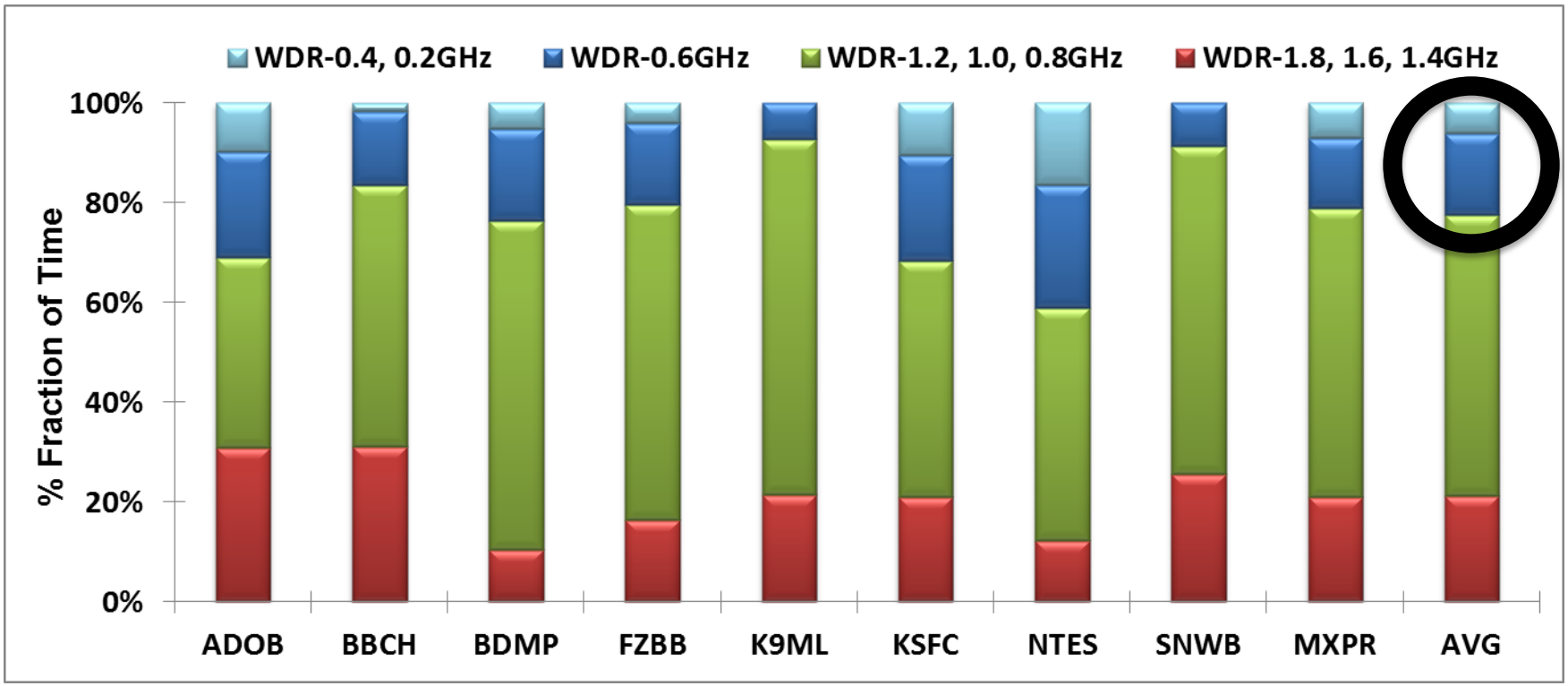
% time spent in different V/F states: big.LITTLE



On average: 52% of time in big core, 48% in LITTLE core V/F states

33rd IEEE International Conference on Computer Design
ICCD 2015

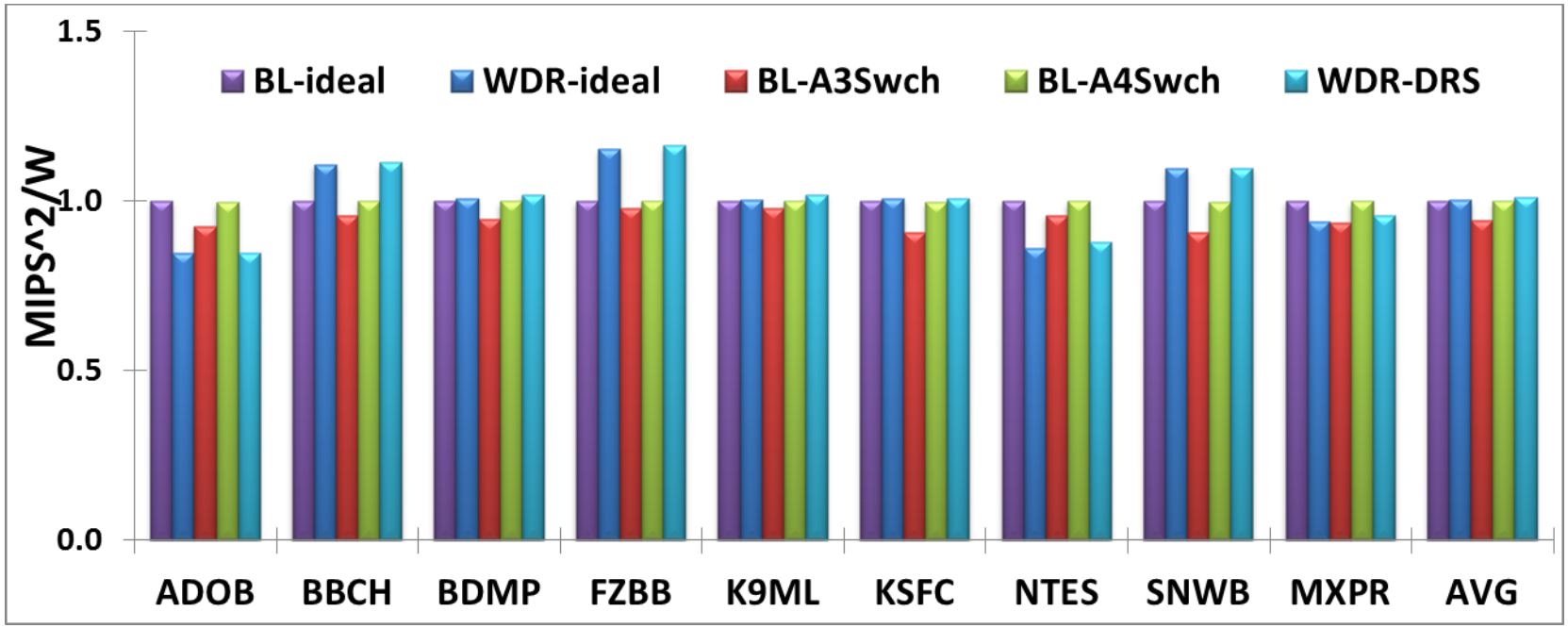
% time spent in different V/F states: WDR



On average: 77% in V/F states with F>=0.8GHz

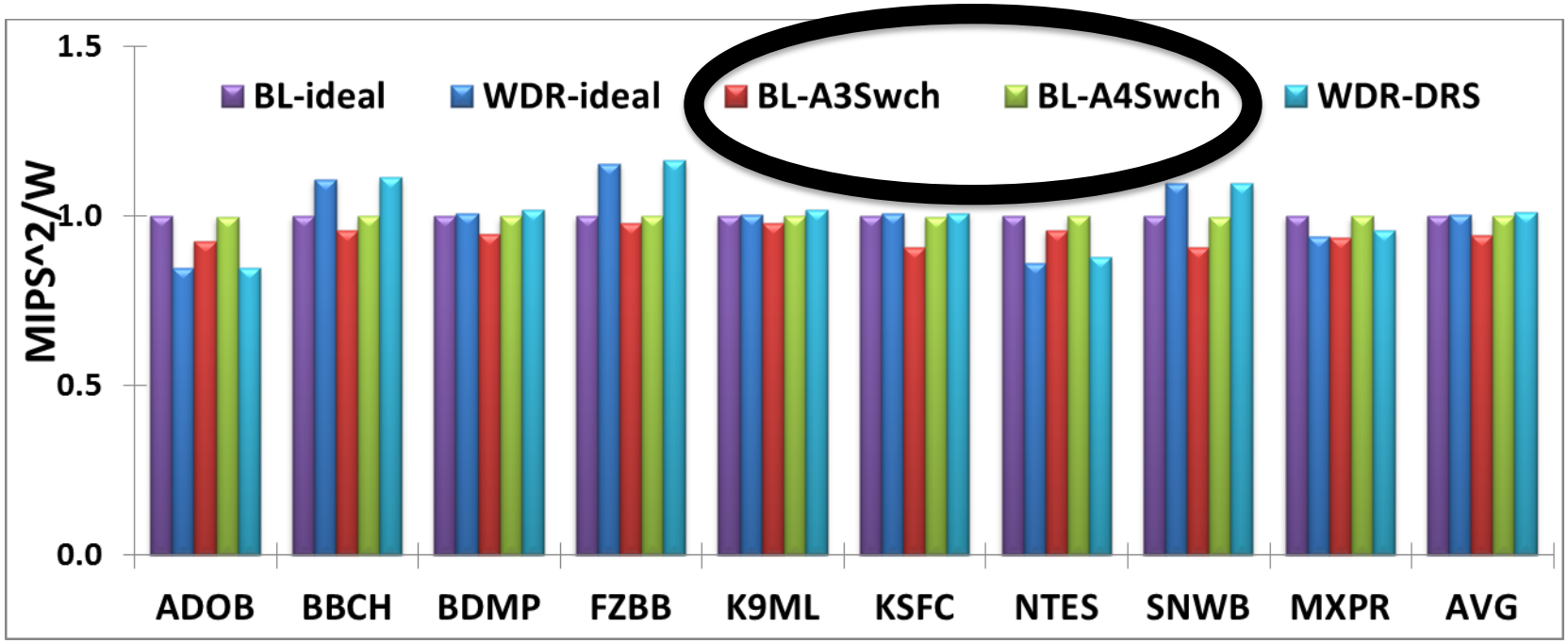
23% in NTV states with F<=0.6GHz

Switching Overhead: big.LITTLE



Both clusters stay on during switching

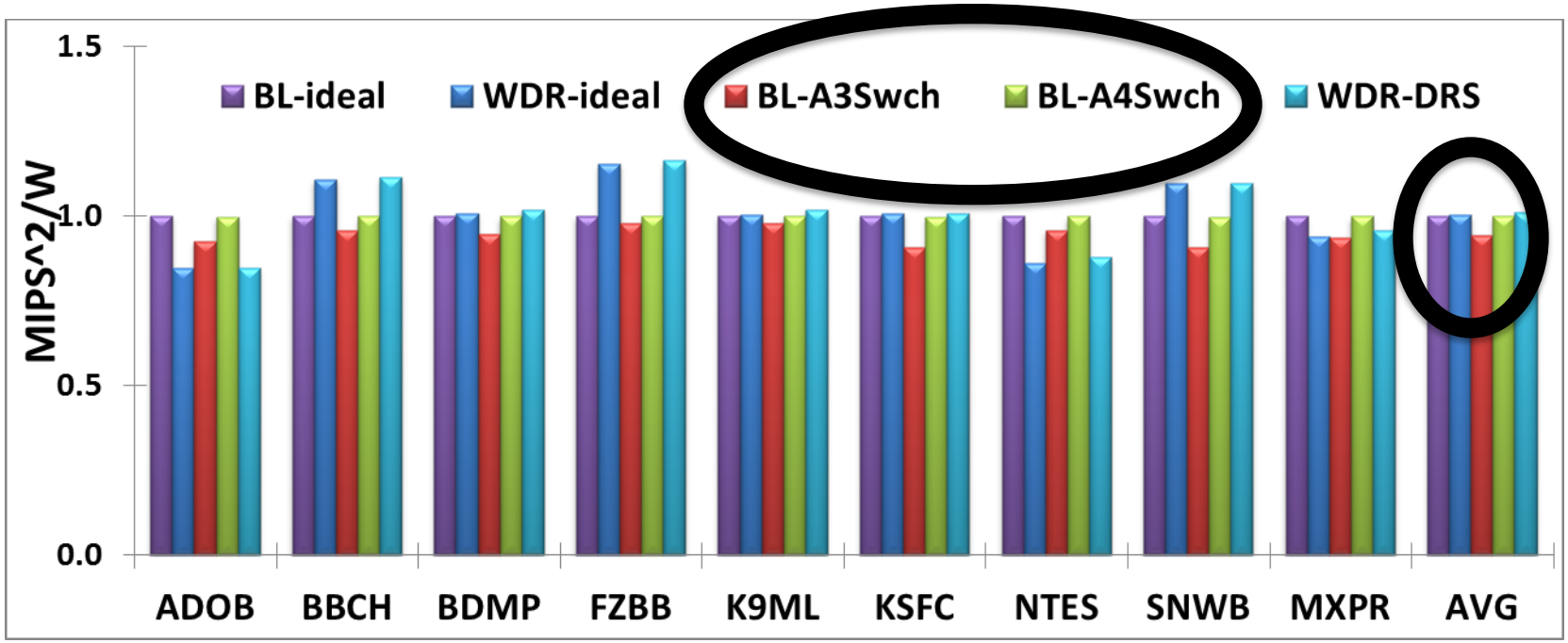
Switching Overhead: big.LITTLE



A(MBA)3: Software intervention needed to orchestrate write-back
1.4 (0.4)ms for 2MB (512KB) L2

A(MBA)4: Hardware-based cache coherence
33µs

Switching Overhead: big.LITTLE

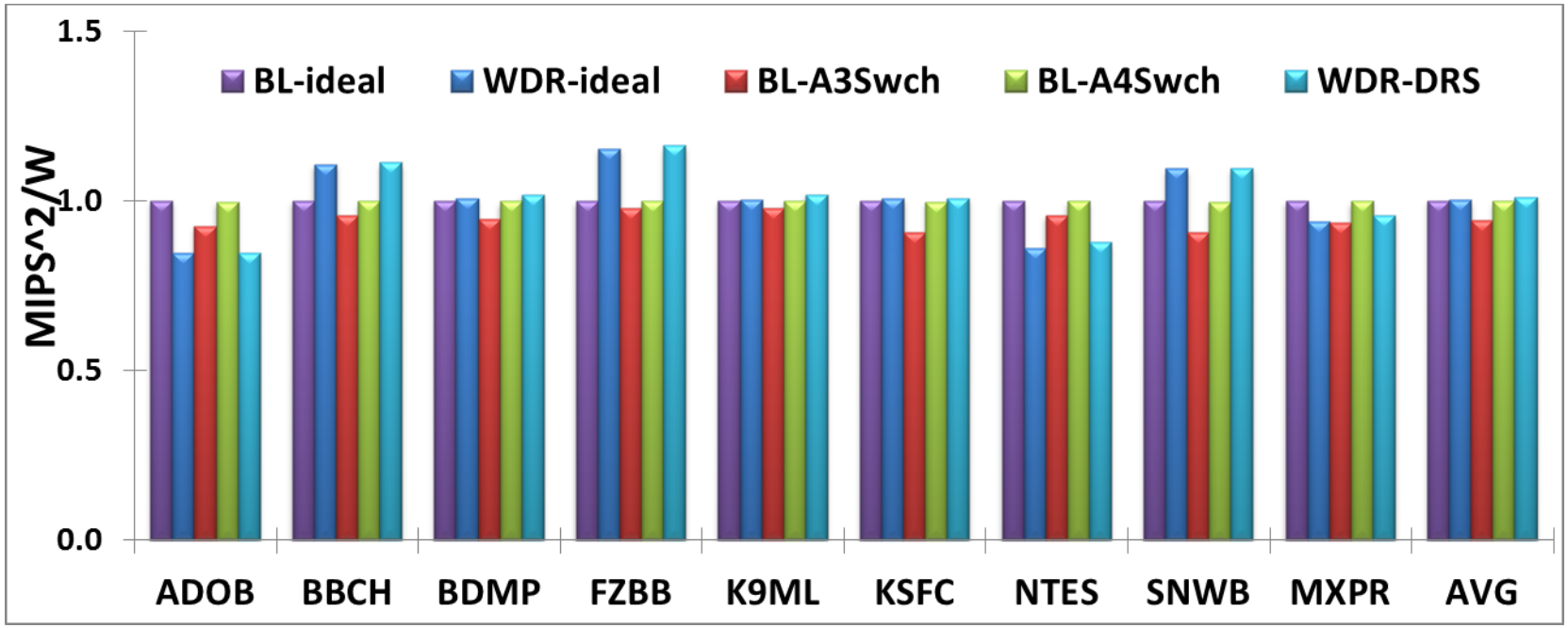


Energy efficiency (under ideal conditions) reduces on average by

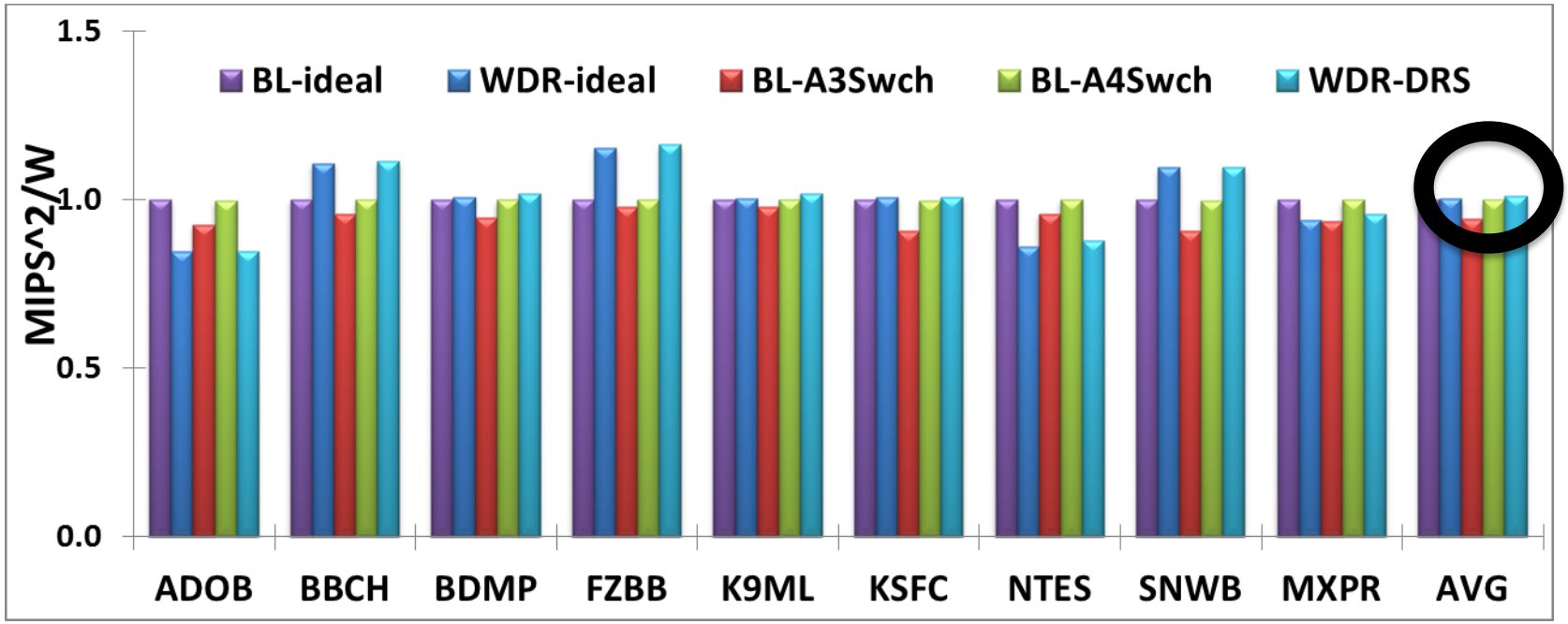
4% for A3

2%for A4

WDR + DRS

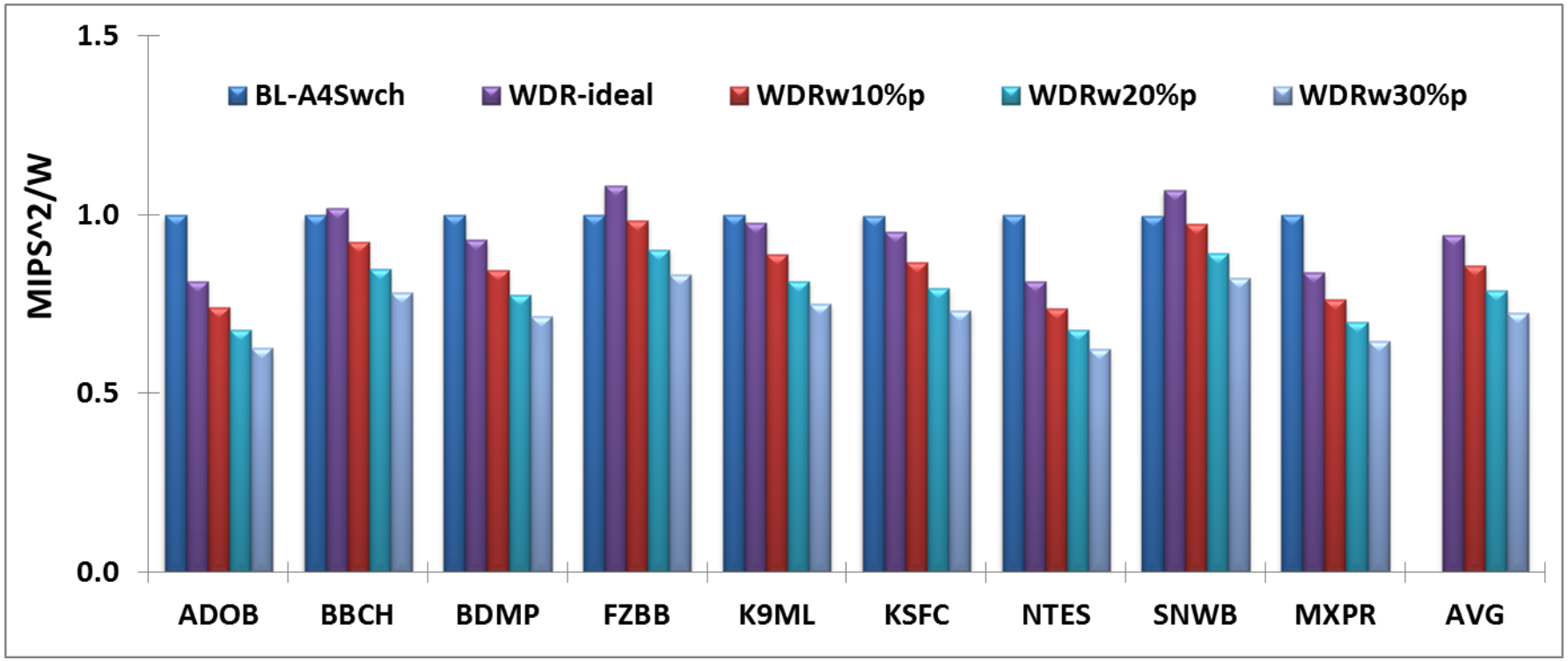

1.5
1.0
0.5
0.0
MIPS^2/W
BL-ideal
WDR-ideal
BL-A3Swch
BL-A4Swch
WDR-DRS
ADOB
BBCH
BDMP
FZBB
K9ML
KSFC
NTES
SNWB
MXPR
AVG

WDR + DRS

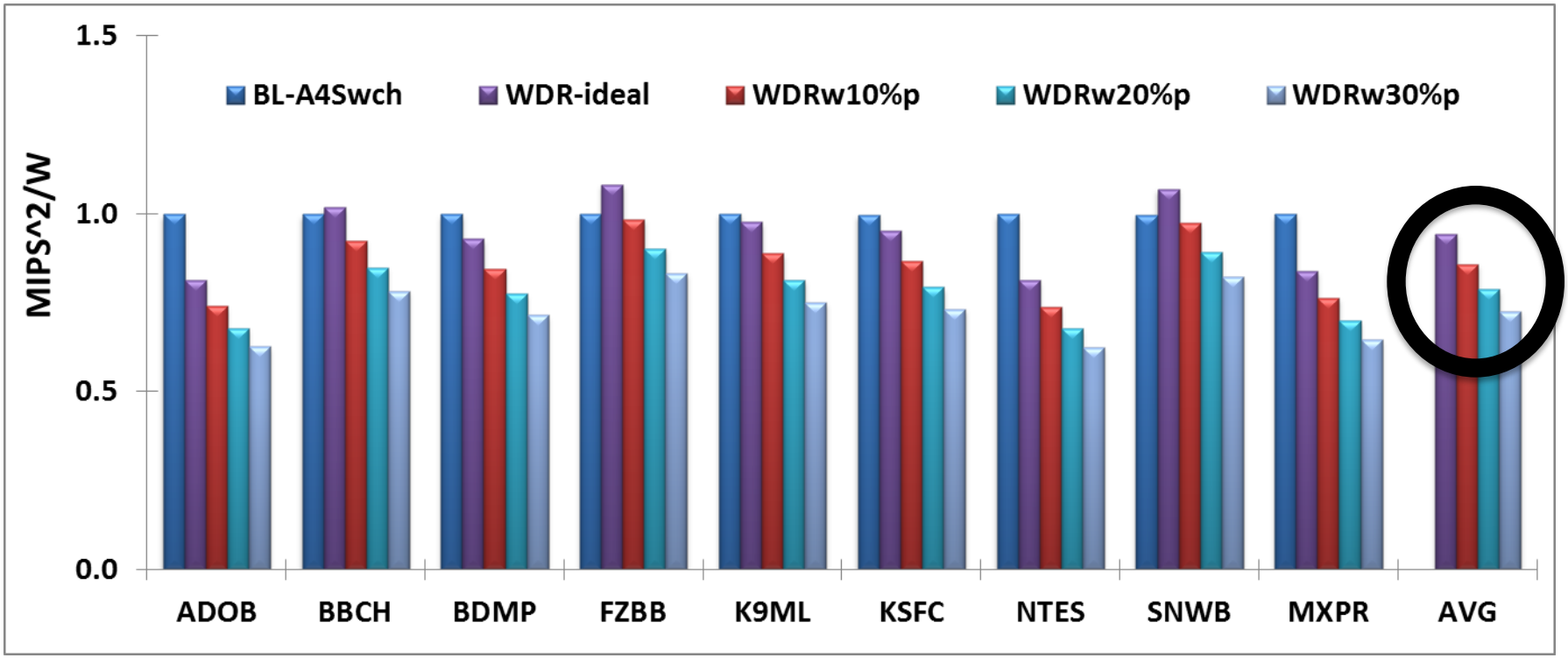


When augmented by DRS, WDR's energy efficiency increases by 2%

Power Overhead of Practical WDR



Power Overhead of Practical WDR



Energy efficiency can decrease by up to 25%

33rd IEEE International Conference on Computer Design
ICCD 2015

Conclusion

- We compare and contrast single-ISA heterogeneous and WDR architectures in terms of
  - Design challenges
  - Energy efficiency

  considering implementations augmented w/ DRS.
- We observe that WDR processors facilitating DRS can deliver energy efficiency very close to single-ISA heterogeneous processors.



Comparison of Single-ISA Heterogeneous versus Wide Dynamic Range Processors for Mobile Applications

Hamid Reza Ghasemi[1], **Ulya R. Karpuzcu**[2], Nam Sung Kim[3]

[1]University of Wisconsin [2]University of Minnesota [3]University of Illinois

33rd IEEE International Conference on Computer Design
ICCD 2015

Processor Configurations

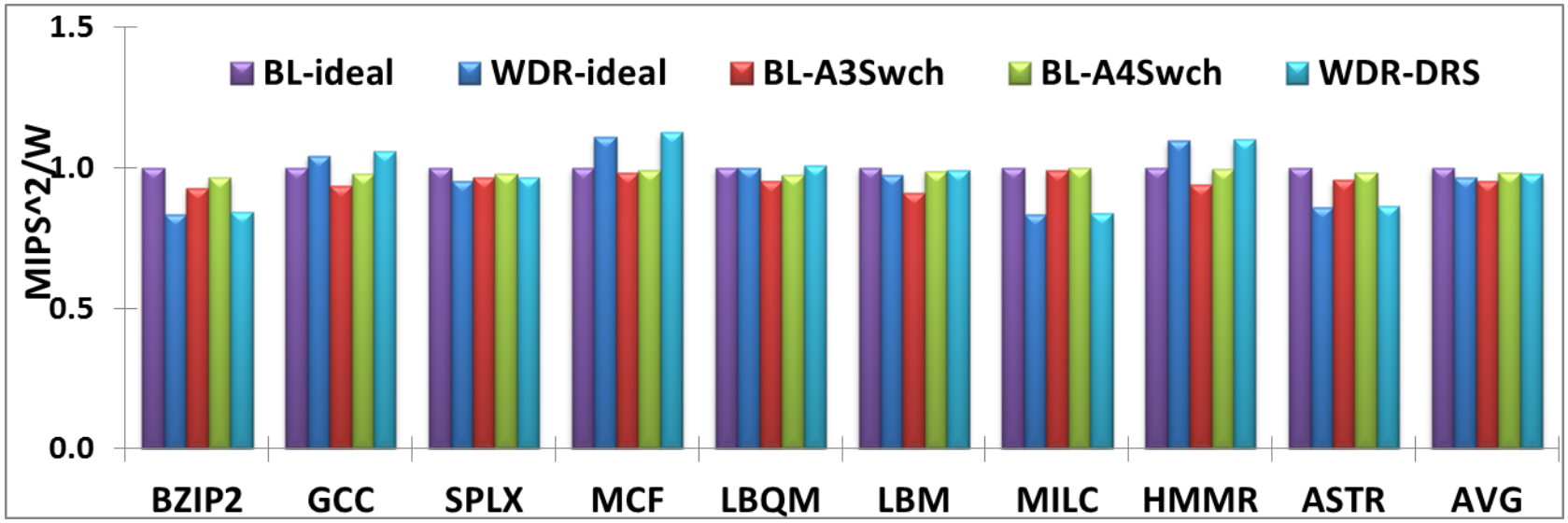
WDR	big.LITTLE		Memory Subsystem	
Core type: Cortex-A15	Big core type: Cortex-A15	Little core type: Cortex-A7	IL1/DL1	32KB/4-Way/64B 3cycles
3-wide issue	3-wide issue	2-wide issue	IL1/DL1	32KB/4-Way/64B 3cycles
160-entry register file	160-entry register file	32-entry register file	Coherency Protocol	Snoopy-based MESI
128 ROB	128-entry ROB	-	L2 little cluster	512KB/8-way/64B, 10 cycles
4 cores	4 cores in big cluster	4 cores in little cluster	L2 big cluster	2MB/8-way/64B, 12 cycles
Regular V/F states: (1.8GHz,0.9V)~ (0.8GHz,0.65V)	Regular V/F states: (1.8GHz,0.9V)~ (0.8GHz,0.65V)	Regular V/F states: (1.2GHz,0.9V)~ (0.4GHz,0.65V)	L2 WDR	2MB/8-way/64B, 12 cycles
Near-threshold V/F states: (0.6GHz,0.6V)~ (0.2GHz,0.53V)	-	-	-	-

Dynamic Resource Scaling Configurations

Configuration	Width	ALU	ROB	LSQ	L2 cache
Configuration 1	3	3	36	24	2 MB
Configuration 2	2	2	24	16	1 MB
Configuration 3	1	1	16	8	512 KB

To have a fair comparison against big.LITTLE, we consider the DRS configurations only at NTV

Single Threaded Applications



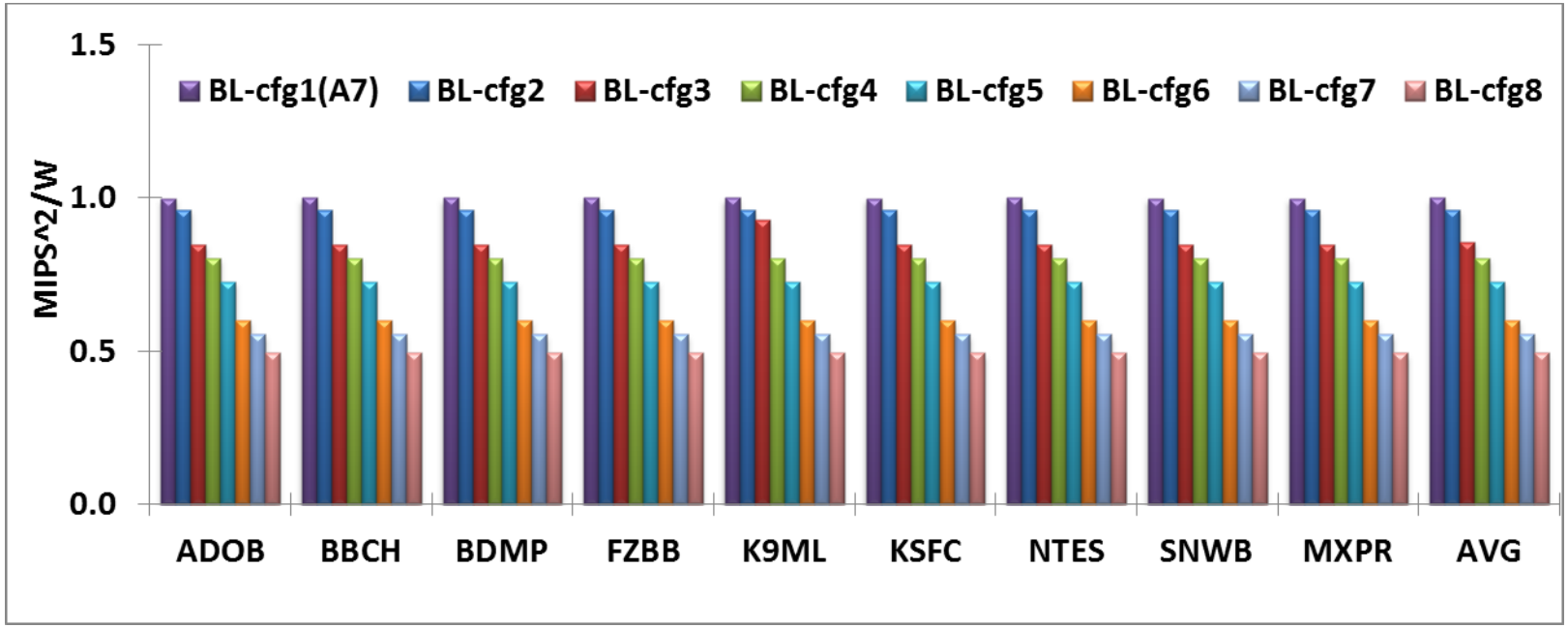
On average: WDR achieves 1-2% lower energy efficiency than BL-ideal

33rd IEEE International Conference on Computer Design
ICCD 2015

Different Configurations for LITTLE Core

Config.	issue width	ALU	L2 (KB)	L1D (KB)	L1I (KB)	INT Reg.	Float Reg.
Cfg1 (A7 core)	2	2	512	32	32	32	32
cfg2	2	1	512	32	32	32	32
cfg3	1	1	512	32	32	32	32
cfg4	1	1	256	32	32	32	32
cfg5	1	1	256	16	32	32	32
cfg6	1	1	256	16	16	32	32
cfg7	1	1	256	16	16	16	32
cfg8	1	1	256	16	16	16	16

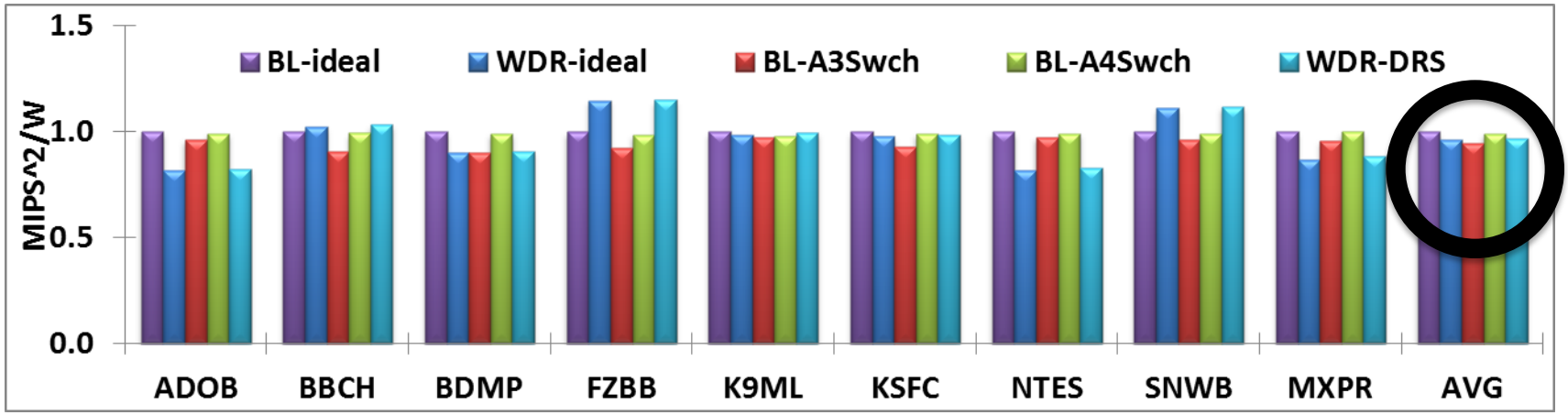
Sensitivity to Different Little Core Configurations



As the size of the LITTLE core decreases, the energy efficiency degrades by up to 47%

33rd IEEE International Conference on Computer Design
ICCD 2015

big.LITTLE w/ CPU migration vs. WDR w/ per-core V domains



Energy efficiency follows a similar trend to
big.LITTLE w/ cluster migration vs. WDR w/ single V domain

Comparison of Single-ISA Heterogeneous versus Wide Dynamic Range Processors for Mobile Applications

Hamid Reza Ghasemi[1], **Ulya R. Karpuzcu**[2], Nam Sung Kim[3]

[1]University of Wisconsin [2]University of Minnesota [3]University of Illinois

33rd IEEE International Conference on Computer Design
ICCD 2015