

DISTANCE-AWARE ROUND-ROBIN MAPPING FOR LARGE NUCA CACHES

Alberto Ros¹ Marcelo Cintra² Manuel E. Acacio¹
José M. García¹

¹Departamento de Ingeniería y Tecnología de Computadores
Universidad de Murcia
{a.ros,meacacio,jmgarcia}@ditec.um.es



²School of Informatics
The University of Edinburgh
mc@inf.ed.ac.uk

December 17, 2009



OUTLINE

- 1 MOTIVATION
- 2 DARR: DISTANCE-AWARE ROUND-ROBIN MAPPING
- 3 FIRST-TOUCH MAPPING AND PRIVATE CACHE INDEXING
- 4 EVALUATION RESULTS
- 5 CONCLUSIONS

OUTLINE

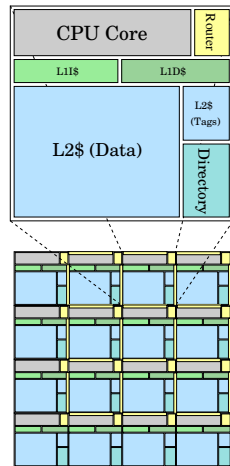
- 1 MOTIVATION
- 2 DARR: DISTANCE-AWARE ROUND-ROBIN MAPPING
- 3 FIRST-TOUCH MAPPING AND PRIVATE CACHE INDEXING
- 4 EVALUATION RESULTS
- 5 CONCLUSIONS

TRENDS

- The number of transistors currently offered in a single die allows us to design small-scale CMPs (e.g. IBM Power6, Sun T2)
 - Soon, the increasing integration scale will lead to medium or large-scale (many-cores) CMPs

TRENDS

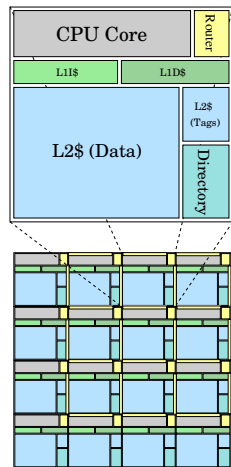
- The number of transistors currently offered in a single die allows us to design small-scale CMPs (e.g. IBM Power6, Sun T2)
 - Soon, the increasing integration scale will lead to medium or large-scale (many-cores) CMPs
- **Tiled CMPs** are a scalable alternative for building CMPs.
 - Designed as arrays of **replicated tiles**.
 - Cores connected through a **direct network**.



TILED-CMPS

ARCHITECTURE ASSUMED IN THIS WORK

- Each tile contains:
 - A processing core.
 - A private L1 cache (both instruction and data caches).
 - A **shared or private** L2 cache bank, and a directory.
 - A network interface (router).
- All tiles are connected through a scalable point-to-point interconnection network.

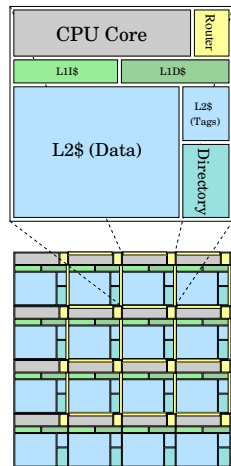


TILED-CMPS

SHARED VS. PRIVATE L2 CACHE ORGANIZATION

PRIVATE ORGANIZATION

- 😊 L2 hits have short latencies (local accesses).
- ☹️ Blocks potentially replicated in multiple L2 banks.
- ☹️ Load balancing problems.



TILED-CMPS

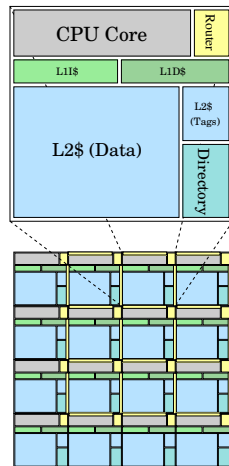
SHARED VS. PRIVATE L2 CACHE ORGANIZATION

PRIVATE ORGANIZATION

- 😊 L2 hits have short latencies (local accesses).
- ☹️ Blocks potentially replicated in multiple L2 banks.
- ☹️ Load balancing problems.

SHARED ORGANIZATION (NUCA ARCHITECTURE)

- 😊 Better use of the aggregate L2 cache capacity.
- ☹️ Long latencies when compared to a private L2 design.
 - The access latency to the L2 depends on where the requested block is mapped.



MAPPING POLICIES FOR THE SHARED ORGANIZATION

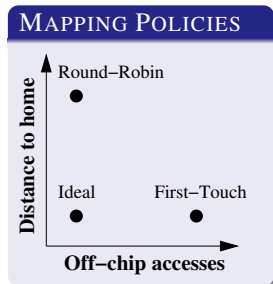
- Commonly performed by taking some bits from the block address leading to a **Round-Robin** mapping.
 - Optimal utilization of cache storage.
 - Does not care about the distance between requesting cores and home banks \Rightarrow long access latency.

MAPPING POLICIES FOR THE SHARED ORGANIZATION

- Commonly performed by taking some bits from the block address leading to a **Round-Robin** mapping.
 - Optimal utilization of cache storage.
 - Does not care about the distance between requesting cores and home banks \Rightarrow long access latency.
- A **First-Touch** mapping policy can lessen this latency.
 - But can cause imbalance among cache banks \Rightarrow high cache miss rate.

MAPPING POLICIES FOR THE SHARED ORGANIZATION

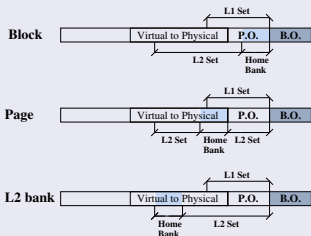
- Commonly performed by taking some bits from the block address leading to a **Round-Robin** mapping.
 - Optimal utilization of cache storage.
 - Does not care about the distance between requesting cores and home banks \Rightarrow long access latency.
- A **First-Touch** mapping policy can lessen this latency.
 - But can cause imbalance among cache banks \Rightarrow high cache miss rate.



HOME INTERLEAVING AND MAPPING POLICIES

- Home banks for memory blocks are defined by taking some bits from the block address
 - This defines the granularity of the home interleaving

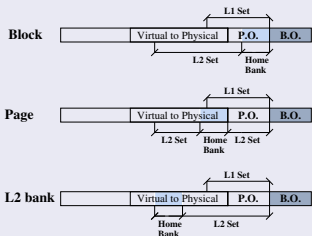
HOME INTERLEAVING



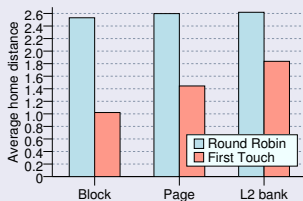
HOME INTERLEAVING AND MAPPING POLICIES

- Home banks for memory blocks are defined by taking some bits from the block address
 - This defines the granularity of the home interleaving
- The **performance** and **feasibility** of the mapping policies depend on this granularity

HOME INTERLEAVING



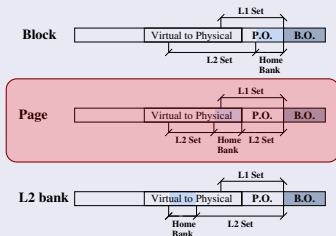
HOME DISTANCE



HOME INTERLEAVING AND MAPPING POLICIES

- Home banks for memory blocks are defined by taking some bits from the block address
 - This defines the granularity of the home interleaving
- The **performance** and **feasibility** of the mapping policies depend on this granularity

HOME INTERLEAVING



HOME DISTANCE



- We choose **OS** management and **page-grained interleaving** for our proposal

OUTLINE

- 1 MOTIVATION
- 2 DARR: DISTANCE-AWARE ROUND-ROBIN MAPPING**
- 3 FIRST-TOUCH MAPPING AND PRIVATE CACHE INDEXING
- 4 EVALUATION RESULTS
- 5 CONCLUSIONS

DARR MAPPING POLICY

- We propose a **Distance-Aware Round-Robin** mapping policy which is managed by the OS.
 - The unit of mapping is a memory page.
 - Does not add extra hardware.
- Our policy tries to map the pages to the local bank of the first requesting core.
 - ⇒ Reduces access latency.
- But also introduces a threshold to bound the deviation of the distribution of memory pages among cache banks.
 - ⇒ Lessens off-chip accesses.

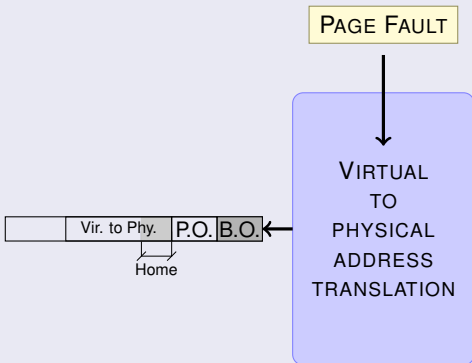
DARR MAPPING POLICY

VIRTUAL TO PHYSICAL ADDRESS TRANSLATION

PAGE FAULT

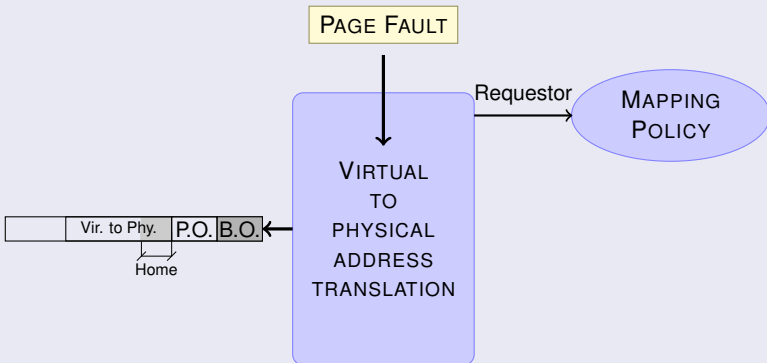
DARR MAPPING POLICY

VIRTUAL TO PHYSICAL ADDRESS TRANSLATION



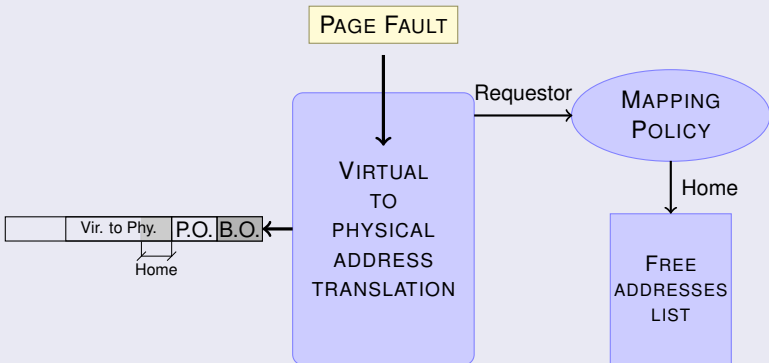
DARR MAPPING POLICY

VIRTUAL TO PHYSICAL ADDRESS TRANSLATION



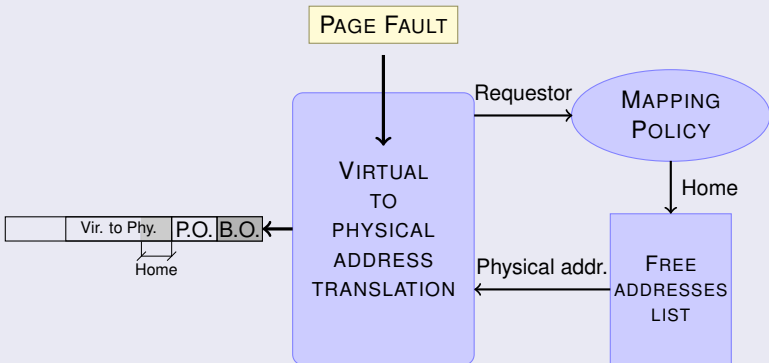
DARR MAPPING POLICY

VIRTUAL TO PHYSICAL ADDRESS TRANSLATION



DARR MAPPING POLICY

VIRTUAL TO PHYSICAL ADDRESS TRANSLATION



DARR MAPPING POLICY

EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2

2x2 tiled CMP

P_0	P_1
P_2	P_3

DARR MAPPING POLICY

EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2

2x2 tiled CMP 1. $P_0 \rightarrow 0x00$

P_0	P_1	1	0
P_2	P_3	0	0

- 1 P_0 accesses a block within a page 0x00 which faults in memory.

DARR MAPPING POLICY

EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2

2x2 tiled CMP 1. $P_0 \rightarrow 0x00$ 2. $P_1 \rightarrow 0x01$

P_0	P_1	1	0	1	1
P_2	P_3	0	0	0	0

- 1 P_0 accesses a block within a page 0x00 which faults in memory.
- 2 P_1 accesses a new page (0x01) \Rightarrow OK.

DARR MAPPING POLICY

EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2

2x2 tiled CMP 1. $P_0 \rightarrow 0x00$ 2. $P_1 \rightarrow 0x01$ 3. $P_1 \rightarrow 0x00$

P_0	P_1	1	0	1	1	1	1
P_2	P_3	0	0	0	0	0	0

- ① P_0 accesses a block within a page 0x00 which faults in memory.
- ② P_1 accesses a new page (0x01) \Rightarrow OK.
- ③ P_1 accesses page 0x00 \Rightarrow Nothing happens.

DARR MAPPING POLICY

EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2

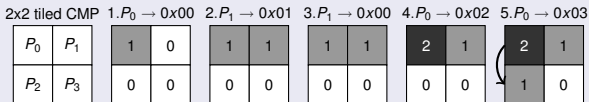
2x2 tiled CMP 1. $P_0 \rightarrow 0x00$ 2. $P_1 \rightarrow 0x01$ 3. $P_1 \rightarrow 0x00$ 4. $P_0 \rightarrow 0x02$

P_0	P_1	1	0	1	1	1	1	2	1
P_2	P_3	0	0	0	0	0	0	0	0

- ① P_0 accesses a block within a page 0x00 which faults in memory.
- ② P_1 accesses a new page (0x01) \Rightarrow OK.
- ③ P_1 accesses page 0x00 \Rightarrow Nothing happens.
- ④ P_0 accesses a new page (0x02) \Rightarrow threshold reached.

DARR MAPPING POLICY

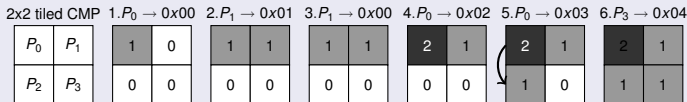
EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2



- 1 P_0 accesses a block within a page 0x00 which faults in memory.
- 2 P_1 accesses a new page (0x01) \Rightarrow OK.
- 3 P_1 accesses page 0x00 \Rightarrow Nothing happens.
- 4 P_0 accesses a new page (0x02) \Rightarrow threshold reached.
- 5 P_0 accesses a new page (0x03) \Rightarrow to another bank.

DARR MAPPING POLICY

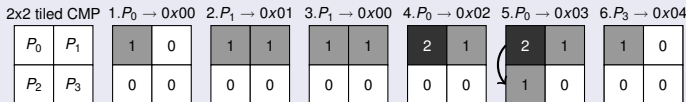
EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2



- P_0 accesses a block within a page 0x00 which faults in memory.
- P_1 accesses a new page (0x01) \Rightarrow OK.
- P_1 accesses page 0x00 \Rightarrow Nothing happens.
- P_0 accesses a new page (0x02) \Rightarrow threshold reached.
- P_0 accesses a new page (0x03) \Rightarrow to another bank.
- P_3 accesses a new page (0x04) \Rightarrow counters decreased.

DARR MAPPING POLICY

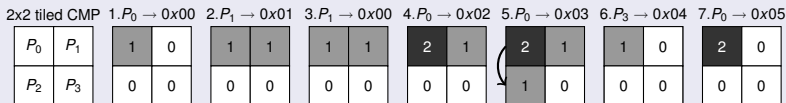
EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2



- P_0 accesses a block within a page $0x00$ which faults in memory.
- P_1 accesses a new page ($0x01$) \Rightarrow OK.
- P_1 accesses page $0x00$ \Rightarrow Nothing happens.
- P_0 accesses a new page ($0x02$) \Rightarrow threshold reached.
- P_0 accesses a new page ($0x03$) \Rightarrow to another bank.
- P_3 accesses a new page ($0x04$) \Rightarrow counters decreased.

DARR MAPPING POLICY

EXAMPLE FOR A 2X2 TILED CMP AND A THRESHOLD VALUE OF 2



- 1 P_0 accesses a block within a page 0x00 which faults in memory.
- 2 P_1 accesses a new page (0x01) \Rightarrow OK.
- 3 P_1 accesses page 0x00 \Rightarrow Nothing happens.
- 4 P_0 accesses a new page (0x02) \Rightarrow threshold reached.
- 5 P_0 accesses a new page (0x03) \Rightarrow to another bank.
- 6 P_3 accesses a new page (0x04) \Rightarrow counters decreased.
- 7 P_0 accesses a new page (0x05) \Rightarrow OK.

OUTLINE

- 1 MOTIVATION
- 2 DARR: DISTANCE-AWARE ROUND-ROBIN MAPPING
- 3 FIRST-TOUCH MAPPING AND PRIVATE CACHE INDEXING**
- 4 EVALUATION RESULTS
- 5 CONCLUSIONS

SKIPPING HOME BITS

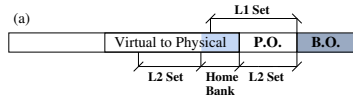
OS-managed mapping can hurt the hit rate of private L1 caches

- Mainly when a first-touch policy is implemented

SKIPPING HOME BITS

OS-managed mapping can hurt the hit rate of private L1 caches

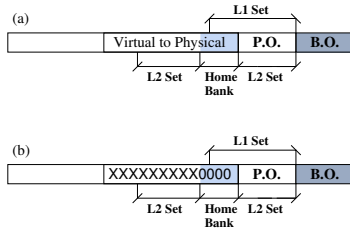
- Mainly when a first-touch policy is implemented
- The bits used to choose the **home bank** are the **less significant** bits of the virtual to physical field
- The bits used to index the **private caches** are commonly the **less significant** bits of the block address



SKIPING HOME BITS

OS-managed mapping can hurt the hit rate of private L1 caches

- Mainly when a first-touch policy is implemented
- The bits used to choose the **home bank** are the **less significant** bits of the virtual to physical field
- The bits used to index the **private caches** are commonly the **less significant** bits of the block address
- Most of the blocks that the L1 caches hold have these bits with the same value
 - They map to the same cache sets

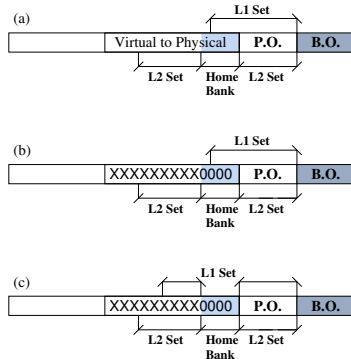


SKIPPING HOME BITS

OS-managed mapping can hurt the hit rate of private L1 caches

- Mainly when a first-touch policy is implemented

- The bits used to choose the **home bank** are the **less significant** bits of the virtual to physical field
- The bits used to index the **private caches** are commonly the **less significant** bits of the block address
- Most of the blocks that the L1 caches hold have these bits with the same value
 - They map to the same cache sets
- **Solution** ⇒ skip home bank bits for the L1 cache indexing



OUTLINE

- 1 MOTIVATION
- 2 DARR: DISTANCE-AWARE ROUND-ROBIN MAPPING
- 3 FIRST-TOUCH MAPPING AND PRIVATE CACHE INDEXING
- 4 EVALUATION RESULTS**
- 5 CONCLUSIONS

SIMULATION TOOLS

GEMS ⇒ MEMORY PARAMETERS

Processor	4GHz, in-order
Cache block size	64 bytes
Split L1 I & D caches	64KB, 4-way
L1 cache hit time	3 cycles
Shared L2 cache	512KB/tile, 16-way
L2 cache hit time	6 cycles
Memory access time	300 cycles
Page size	4KB

SiCoSys ⇒ NETWORK PARAMETERS

Network frequency	2GHz
Topology	2D mesh
Switching technique	Wormhole
Routing technique	Deterministic X-Y
Message size	4 or 1 flits
Routing and switch time	2 cycles
Link latency (one hop)	2 cycles
Link bandwidth	1 flit/cycle

SIMULATION TOOLS

GEMS ⇒ MEMORY PARAMETERS

Processor	4GHz, in-order
Cache block size	64 bytes
Split L1 I & D caches	64KB, 4-way
L1 cache hit time	3 cycles
Shared L2 cache	512KB/tile, 16-way
L2 cache hit time	6 cycles
Memory access time	300 cycles
Page size	4KB

SiCoSYS ⇒ NETWORK PARAMETERS

Network frequency	2GHz
Topology	2D mesh
Switching technique	Wormhole
Routing technique	Deterministic X-Y
Message size	4 or 1 flits
Routing and switch time	2 cycles
Link latency (one hop)	2 cycles
Link bandwidth	1 flit/cycle

- Policies evaluated:
 - **Block-RoundRobin**: Block granularity, HW round-robin mapping
 - **Page-RoundRobin**: Page granularity, OS round-robin mapping
 - **Page-FirstTouch**: Page granularity, OS first-touch mapping
 - **Page-DARR-#**: Page granularity, OS DARR mapping with threshold values (#) ranging from 2^0 to 2^{10}

BENCHMARKS

- Benchmarks (scientific applications):
 - SPLASH-2: FFT, Radix, Raytrace, Ocean, and Water-NSQ
 - Other: Unstructured

BENCHMARKS

- Benchmarks (scientific applications):
 - SPLASH-2: FFT, Radix, Raytrace, Ocean, and Water-NSQ
 - Other: Unstructured
- Workloads evaluated:
 - Parallel applications (16-tile CMPs):
 - FFT, Radix, Ocean, and Unstructured

BENCHMARKS

- Benchmarks (scientific applications):
 - SPLASH-2: FFT, Radix, Raytrace, Ocean, and Water-NSQ
 - Other: Unstructured
- Workloads evaluated:
 - Parallel applications (16-tile CMPs):
 - FFT, Radix, Ocean, and Unstructured
 - Multiprogrammed workloads (32-tile CMPs):

Ocean4

0	1	2	3	4	5	6	7
Ocean				Ocean			
8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23
Ocean				Ocean			
24	25	26	27	28	29	30	31

Radix4

0	1	2	3	4	5	6	7
Radix				Radix			
8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23
Radix				Radix			
24	25	26	27	28	29	30	31

Mix4

0	1	2	3	4	5	6	7
Ocean				Raytrace			
8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23
Water-NSQ				Unstructured			
24	25	26	27	28	29	30	31

Mix8

0	1	2	3	4	5	6	7	
Ocean		Raytrace		Ocean		Raytrace		
8	9	10	11	12	13	14	15	
16	Water-NSQ		Unstructured		Water-NSQ		Unstructured	
18	19	20	21	22	23	24	25	
26	27	28	29	30	31	32	33	

EVALUATION

MEMORY ACCESS FOOTPRINT FOR A FIRST-TOUCH POLICY

FFT

73	60	56	54
52	71	57	55
53	55	59	54
52	55	53	57

Ocean

307	239	255	243
238	240	240	239
239	239	244	240
239	267	239	246

Ocean4

576	468	468	469	475	468	468	469
472	470	468	500	478	471	469	490
480	472	470	469	473	469	478	470
468	489	469	491	468	470	468	490

Radix4

524	529	491	548	491	549	487	559
506	547	524	562	520	562	519	561
495	541	498	548	492	552	493	551
532	565	517	568	523	560	523	558

Radix

280	92	91	92
89	99	121	109
121	109	153	148
190	163	251	238

Unstructured

93	65	18	15
253	19	11	34
15	10	20	12
16	14	16	40

Mix4

658	469	468	470	130	113	96	66
474	469	468	506	108	39	62	81
0	17	17	29	54	43	49	45
29	26	29	87	54	45	45	83

Mix8

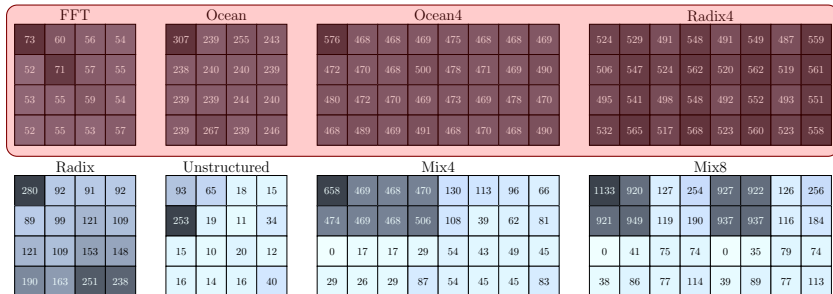
1133	920	127	254	927	922	126	256
921	949	119	190	937	937	116	184
0	41	75	74	0	35	79	74
38	86	77	114	39	89	77	113

#

NUMBER OF PAGES MAPPED
TO EACH CACHE BANK

EVALUATION

MEMORY ACCESS FOOTPRINT FOR A FIRST-TOUCH POLICY



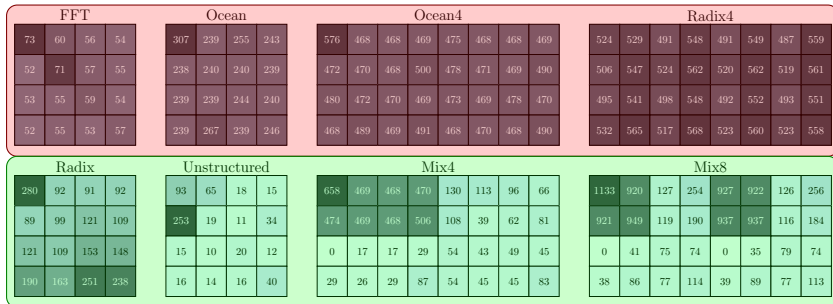
#

NUMBER OF PAGES MAPPED
TO EACH CACHE BANK

HOMOGENEOUS

EVALUATION

MEMORY ACCESS FOOTPRINT FOR A FIRST-TOUCH POLICY



#

NUMBER OF PAGES MAPPED
TO EACH CACHE BANK

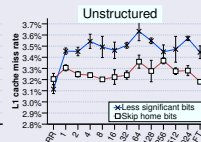
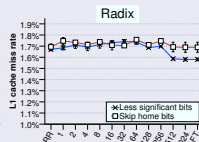
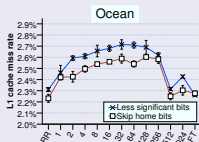
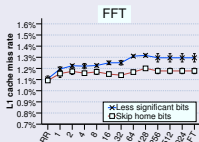
HOMOGENEOUS

HETEROGENEOUS

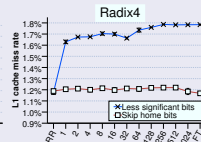
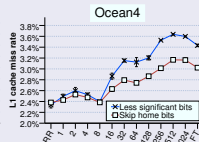
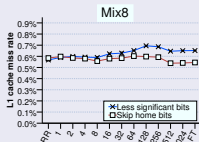
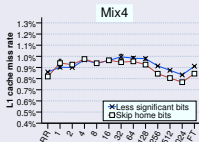
EVALUATION

PRIVATE CACHE INDEXING AND MISS RATE

PARALLEL APPLICATIONS



MULTIPROGRAMMED WORKLOADS

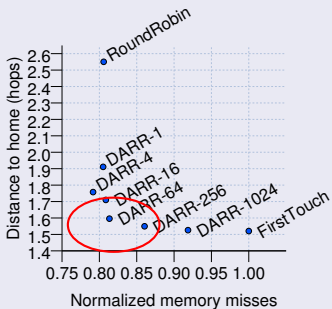


- Skip home bits has similar hit rate for Round-Robin, but better for First-Touch

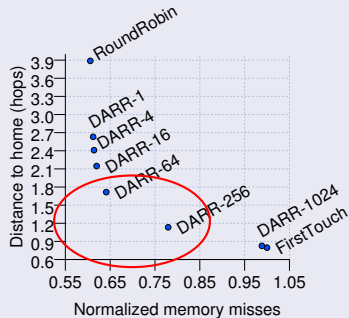
EVALUATION

TRADE-OFF BETWEEN LATENCY AND OFF-CHIP MISSES

PARALLEL APPLICATIONS



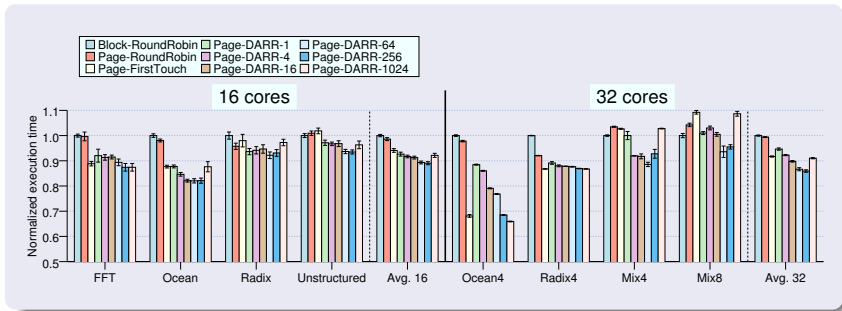
MULTIPROGRAMED WORKLOADS



- A threshold value between 64 and 256 obtains a good trade-off.

EVALUATION

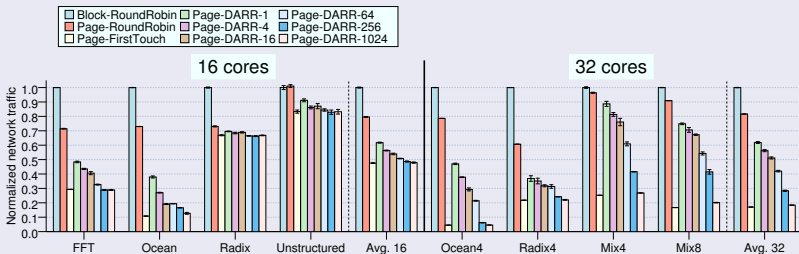
EXECUTION TIME



- Compared to **Round-Robin**: improvements of **11%** for parallel applications and of **14%** for multi-programmed workloads
- Compared to **First-Touch**: improvements of **5%** for parallel applications and of **6%** for multi-programmed workloads

EVALUATION

NETWORK TRAFFIC



- Compared to **Round-Robin**: traffic is reduced by **39%** for parallel applications and of **65%** for multi-programmed workloads
- Compared to **First-Touch**: traffic is increased by **3%** for parallel applications and of **31%** for multi-programmed workloads

OUTLINE

- 1 MOTIVATION
- 2 DARR: DISTANCE-AWARE ROUND-ROBIN MAPPING
- 3 FIRST-TOUCH MAPPING AND PRIVATE CACHE INDEXING
- 4 EVALUATION RESULTS
- 5 CONCLUSIONS

CONCLUSIONS

- The **Distance-Aware Round-Robin** mapping policy achieves a good trade-off between access latency and cache miss rate.
- It obtains the following improvements:
 - Compared to Round-Robin: **11%** for parallel applications and **14%** for multi-programmed workloads.
 - Compared to First-Touch: **5%** for parallel applications and **6%** for multi-programmed workloads.
- It also reduces network traffic compared to Round-Robin:
 - **39%** for parallel applications and **65%** for multi-programmed workloads.
- It does not require any extra hardware structure, which makes our proposal easy to implement in current tiled CMPs.

DISTANCE-AWARE ROUND-ROBIN MAPPING FOR LARGE NUCA CACHES

Alberto Ros¹ Marcelo Cintra² Manuel E. Acacio¹
José M. García¹

¹Departamento de Ingeniería y Tecnología de Computadores
Universidad de Murcia
{a.ros,meacacio,jmgarcia}@ditec.um.es



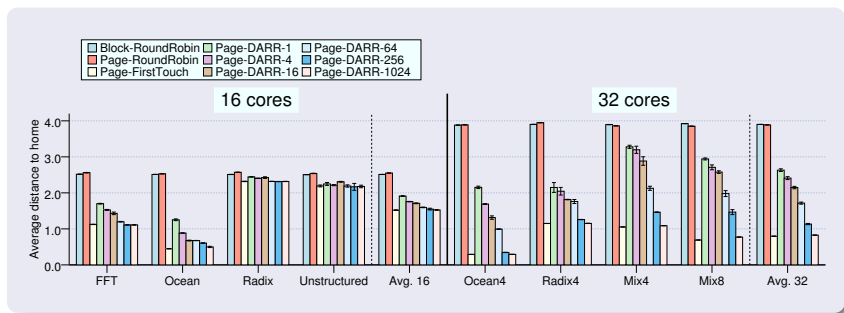
²School of Informatics
The University of Edinburgh
mc@inf.ed.ac.uk

December 17, 2009



EVALUATION

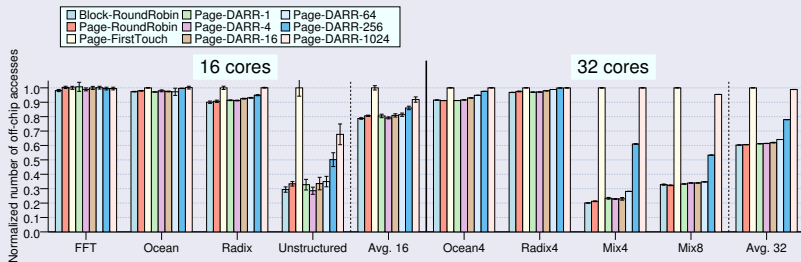
AVERAGE DISTANCE TO HOME BANKS



- **Round-Robin**: Average \Rightarrow 2.5 for a 4×4 mesh and 3.875 in a 4×8 mesh
- **First-Touch**: Less hops \Rightarrow 1.58 for parallel applications and 0.82 for multi-programmed workloads
- **DARR**: In between \Rightarrow Good results with a threshold value ≥ 64

EVALUATION

NUMBER OF OFF-CHIP ACCESSES



- **Round-Robin**: Lessens off-chip accesses
- **First-Touch**: Increases the number of off-chip accesses mainly for heterogeneous workloads
- **DARR**: In between \Rightarrow Good results with a threshold value ≤ 256