

2차원 구조와 3차원 구조에 따른 멀티코어 프로세서의 온도 분석

최 홍 준*, 안 진 우*, 장 형 범**, 김 종 면***, 김 철 흥*

Thermal Pattern Comparison between 2D Multicore Processors and 3D Multicore Processors

Hongjun Choi*, Jinwoo Ahn*, Hyungbeom Jang**, Jong-myon Kim***, Cheolhong Kim*

요 약

동작 주파수의 증가는 싱글코어 프로세서의 성능을 크게 향상시키는 반면 전력 소모 증가와 높은 온도로 인한 신뢰성 저하 문제를 유발하고 있다. 최근에는 싱글코어 프로세서의 한계점을 극복하기 위한 대안으로 멀티코어 프로세서가 주로 사용되고 있다. 하지만, 멀티코어 프로세서를 2차원 구조로 설계하는 경우에는 내부 연결망에서의 전송 지연 현상으로 인해 프로세서의 성능 향상이 제약받고 있다. 내부 연결망에서의 전송 지연을 줄이기 위한 방안으로 멀티코어 프로세서를 3차원 구조로 설계하는 연구가 최근 큰 주목을 받고 있다. 2차원 구조 멀티코어 프로세서와 비교하여 3차원 구조 멀티코어 프로세서는 성능 향상과 전력 소모 감소의 장점을 지닌 반면, 높은 전력 밀도로 인해 발생된 발열 문제가 프로세서의 신뢰성을 위협하는 문제가 되고 있다. 3차원 멀티코어 프로세서에서 발생하는 발열 문제에 대한 상세한 분석이 제공된다면, 프로세서의 신뢰성을 확보하기 위한 연구 진행에 큰 도움이 될 것으로 기대된다. 그러므로 본 논문에서는 3차원 멀티코어 프로세서의 온도에 밀접하게 연관된 요소인 작업량, 방열판과의 거리, 그리고 적층되는 다이의 개수와 온도 사이의 관계를 자세히 살펴보고 높은 온도가 프로세서의 성능에 미치는 영향 또한 분석하고자 한다. 특히, 2차원 구조 멀티코어 프로세서와 3차원 구조 멀티코어 프로세서에서의 온도 문제를 함께 분석함으로써, 온도 측면에서 효율적인 프로세서 설계를 위한 가이드라인을 제시하고자 한다.

▶ Keyword : 멀티코어 프로세서, 3차원 집적회로, 온도, 열섬현상, 신뢰성

Abstract

Unfortunately, in current microprocessors, increasing the frequency causes increased power consumption and reduced reliability whereas it improves the performance. To overcome the power

• 제1저자 : 최홍준 • 교신저자 : 김철흥

• 투고일 : 2010. 10. 13, 심사일 : 2011. 04. 04, 게재확정일 : 2011. 05. 04.

* 전남대학교 전자컴퓨터공학부(School of Electronics and Computer Engineering, Chonnam National University)

** 고려대학교 컴퓨터전파통신공학부(Division of Computer and Communication Engineering, Korea University)

*** 울산대학교 컴퓨터정보통신공학부(School of Electrical Engineering, University of Ulsan)

※ 본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음
(NIPA-2011-C1090-1111-0008)

and thermal problems in the singlecore processors, multicore processors has been widely used. For 2D multicore processors, interconnection is regarded as one of the major constraints in performance and power efficiency. To reduce the performance degradation and the power consumption in 2D multicore processors, 3D integrated design technique has been studied by many researchers. Compared to 2D multicore processors, 3D multicore processors get the benefits of performance improvement and reduced power consumption by reducing the wire length significantly. However, 3D multicore processors have serious thermal problems due to high power density, resulting in reliability degradation. Detailed thermal analysis for multicore processors can be useful in designing thermal-aware processors. In this paper, we analyze the impact of workload distribution, distance to the heat sink, and number of stacked dies on the processor temperature. We also analyze the effects of the temperature on overall system performance. Especially, this paper presents the guideline for thermal-aware multicore processor design by analyzing the thermal problems in 2D multicore processors and 3D multicore processors.

▶ Keyword : multi-core processor, 3D integrated circuits, temperature, hotspot, reliability

1. 서론

싱글코어 프로세서는 동작 주파수가 빨라짐에 따라 성능은 크게 향상된 반면 전력 소모 증가와 고온으로 인한 칩의 신뢰성 저하 등의 문제점을 드러내며 한계에 도달하고 있다[1]. 특히, 싱글코어 프로세서의 하드웨어 복잡도 또한 물리적 한계에 도달하여 성능을 향상시키는데 많은 어려움을 겪고 있기 때문에 프로세서 설계자들은 멀티코어 프로세서라는 새로운 패러다임을 제시하였다. 멀티코어 프로세서 구조는 보다 강력한 성능과 소비 전력 절감, 그리고 여러 개의 작업을 보다 효율적으로 처리하기 위해 두 개 이상의 코어들을 하나의 패키지로 통합한 구조를 말한다[2]. 멀티코어 프로세서는 한 컴퓨터 내에 두 개 이상의 독립된 프로세서가 설치된 경우나 동시 멀티스레딩(SMT: Simultaneous Multi Threading)과 종종 비교된다. 하지만 두 개 이상의 코어가 하나의 단일 회로에 구성된 멀티코어 프로세서 구조는 각기 다른 회로에 독립적으로 구성된 프로세서를 연결하는 구조와는 엄연히 다르다고 할 수 있다. 또한, 동시 멀티스레딩은 외부에서 보았을 경우 코어가 2개 이상으로 보이는 하지만 실제로 존재하는 코어는 하나라는 점에서 멀티 코어 기법과는 근본적으로 다르다[3].

싱글코어 프로세서의 대안으로 제시된 멀티코어 프로세서는 다수의 코어를 탑재하기 때문에, 싱글코어 프로세서 보다 낮은 주파수에서 높은 성능을 발휘할 수 있으므로 소모 전력을 줄일 수 있고, 싱글코어 프로세서에서는 하나의 코어에서 소모되던 전력을 여러 코어에 분배시키기 때문에 높은 전력

밀도로 인해 발생하는 온도 문제 또한 완화시킬 수 있다. 그리고 멀티코어 프로세서에서는 코어의 개수가 증가할수록 데이터를 병렬 처리 할 수 있는 능력이 향상되어 처리 속도가 빨라지므로 성능 향상의 이점도 동시에 얻을 수 있다. 일반 데이터 처리에서 대용량의 멀티미디어 데이터 처리 위주로 변화하고 있는 최근의 컴퓨팅 시스템 환경에서 멀티코어 프로세서를 사용하면 계산 능력을 크게 향상시킬 수 있을 뿐만 아니라, 전력 소모 문제 또한 해결할 수 있으리라 기대된다[4].

하지만, 기존의 2차원 구조(2D Planar) 설계에서는 내부 연결망(interconnection)이 멀티코어 프로세서의 성능과 전력소모 측면에서 부정적인 영향을 미치고 있다[5]. 멀티코어 프로세서를 3차원으로 설계하는 기술은 2차원 구조에서의 내부 연결망 문제를 해결하는 데 매우 적합하기 때문에 대부분의 프로세서 회사들에서 활발하게 연구되고 있다[6-7]. 멀티코어를 3차원으로 설계하는 기술은 각 코어들을 여러 층에 수직으로 쌓으며, 각 층은 TSV(Through-Silicon Vias)를 통해 서로 연결하는 기술이다[8]. 3차원 설계 기술은 선로 길이(wire length)를 상당히 줄일 수 있으므로 칩 내부의 통신 속도를 향상시키고 선로에서 소모되는 전력 또한 상당히 줄일 수 있다. 하지만, 3차원 적층(stacked) 구조 멀티코어 프로세서는 각 코어들이 여러 층에 수직으로 쌓이므로 전력 밀도가 급격히 증가하여 열섬현상을 유발시킨다는 단점이 존재한다. 열섬현상은 프로세서의 작동 오류를 유발하고, 변형을 발생시켜서 칩의 신뢰성(reliability)에 부정적인 영향을 미친다[9].

3차원 멀티코어 프로세서 구조에서 발생하는 발열 문제에 관련된 요소들이 체계적으로 분석되어 있다면 온도를 제어하

여 프로세서의 신뢰성을 확보하기 위한 연구에 커다란 도움이 될 것이다. 3차원 멀티코어 프로세서에서 발생하는 온도는 작업량, 방열판과의 거리, 그리고 적층되는 다이의 개수와 밀접한 관계가 있다고 알려져 있다. 그러므로 본 논문에서는 위의 3가지 요소가 2차원 멀티코어 프로세서와 3차원 멀티코어 프로세서에 미치는 영향을 상세히 분석하고자 한다. 또한, 높은 온도가 시스템 성능에 미치는 영향을 분석하여 온도를 제어함으로써 성능 측면에서 얻을 수 있는 효과에 대해서도 알아보 고자 한다. 본 논문의 구성은 다음과 같다. 2장에서는 본 논문의 분석 대상에 대해 기술하고, 3장과 4장에서는 멀티코어 프로세서의 온도를 측정하기 위해 수행한 모의실험 환경과 자세한 실험 결과를 보여준다. 마지막으로, 5장에서는 본 논문의 결론과 향후 연구 방향을 기술한다.

II. 관련 연구

프로세서를 설계하는 데 있어서 이미 온도는 매우 중요한 제약 조건 중 하나로 인식되고 있으며, 이를 해결하기 위해 많은 연구자들이 온도를 고려하는 저온도 프로세서 개발을 진행하고 있다[10-11]. 특히, 온도 문제가 더욱 심각하게 발생하는 3차원 구조에서의 온도 문제를 해결하기 위해 다양한 연구들이 이루어지고 있다[9,12]. 다이의 수와 전력 밀도에 따른 온도를 분석하고자 국제 반도체 기술 로드맵(ITRS : International Technology Roadmap for Semiconductors)[13]에서 제공하는 데이터를 사용하여 3차원 구조에서의 온도를 분석하는 연구가 진행된 바 있다[12]. 또한, 2-다이 3차원 구조에서의 데이터 처리에 따른 온도 패턴 분석에 대한 연구 또한 수행되었다[9].

기존에 진행된 온도 분석 연구들과는 달리, 본 논문에서는 열섬현상이 발생하는 근본적인 원인을 검증된 툴들을 이용하여 다양한 관점에서 정보를 수집하고 분석하고자 한다. 본 논문에서 분석하고자 하는 측면은 아래와 같다.

1. 코어의 수
2. 작업량
3. 방열판과의 거리
4. 다이의 수

뿐만 아니라, 온도에 많은 영향을 미치는 다양한 요소들을 최대한 정확하게 반영하여 온도 모델링을 설계하였다는 점에서 신뢰할 수 있는 결과를 도출할 수 있을 것으로 판단된다. 특히, 온도에 영향을 미치는 요소들에 대한 분석과 더불어 온도가 성능에 미치는 영향에 대한 분석 또한 수행한다.

III. 분석 대상

2차원 멀티코어 프로세서 구조는 모든 코어를 하나의 다이로 이루어진 하나의 패키지에 통합한 것을 의미하고, 3차원 멀티코어 프로세서 구조는 코어들을 여러 개의 다이에 분산하여 배치하고 수직으로 쌓은 다이들을 하나의 패키지에 통합한 것을 의미한다. 동작 주파수에 의해 주로 성능이 결정되는 싱글 코어 프로세서와 달리 멀티코어 프로세서의 성능은 동작 주파수와 함께 메모리나 아키텍처에 대한 의존도가 매우 크기 때문에 본 장에서는 이에 대한 정의를 명시하고자 한다. 멀티코어 프로세서는 보통 메모리를 개별적으로 갖거나 공유, 또는 두 가지를 동시에 사용하기도 하는데 메모리를 공유하여 사용하는 경우에는 데이터 독립성을 유지하기 매우 어렵다[14].

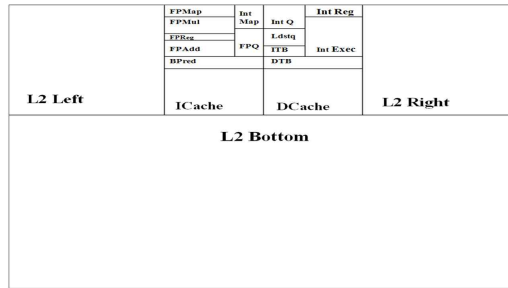


그림 1. Alpha 21364의 구조
Fig 1. Alpha 21364 architecture

본 연구에서의 초점은 멀티코어 프로세서에서의 온도에 있으므로 온도와 관련이 있는 실행 요소들 외에는 서로간의 연관성을 최소화시키기 위해, 분석대상인 멀티코어 프로세서의 각 코어는 개별적인 메모리 구조와 동일한 아키텍처를 갖고 있으며 슈퍼스칼라 실행, 파이프라인 처리, 멀티스레딩과 같은 최적화 조건에서 동작하는 것으로 가정한다. 또한, 멀티코어 프로세서의 효율성은 병렬 연산에 맞는 프로그래밍의 지원 여부에 따라 큰 차이가 있으므로 분석 대상들은 동일한 운영체제와 미들웨어를 사용하는 것으로 가정하고자 한다.

그림1은 본 논문에서 코어로 사용하는 Alpha 21364[15] 프로세서 구조를 보여주고 있다. 하나의 Alpha 21364 프로세서는 가로, 세로 모두 0.016미터 크기인 다이에 패키징 되어있는 것으로 가정한다. 멀티코어 프로세서 구조에서 코어는 캐쉬 메모리를 제외한 다이 위에 구성된 핵심 회로만을 의미하는 경우가 대부분이나 본 논문에서 분석대상으로 하는 멀티코어 프로세서는 각 코어가 개별적인 메모리만을 가지고 있다고 가

정하기 때문에 본 논문에서의 코어는 캐쉬 메모리를 포함하는 그림 1에서 보이는 Alpha 21364 전체를 의미한다.

1. 듀얼 코어 프로세서

그림 2와 그림 3은 서로 다른 형태의 듀얼코어 프로세서를 보여주고 있다.

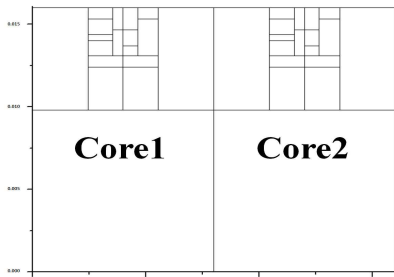


그림 2. 수평적 배치의 2차원 듀얼코어 프로세서
Fig 2. 2D dual-core processor of horizontal layout

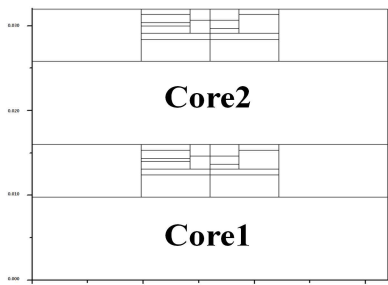


그림 3. 수직적 배치의 2차원 듀얼코어 프로세서
Fig 3. 2D dual-core processor of vertical layout

그림 2는 두 개의 코어를 수평적으로 배치한 것으로 좌측의 코어를 core1로 우측의 코어를 core2로 나타내고 있다. 수평적 배치의 2차원 듀얼코어 프로세서의 경우 다이는 가로축 0.032미터, 세로축 0.016미터의 크기를 가지고 있다. 그림 3에서 보이는 듀얼코어 프로세서는 그림 2와 마찬가지로 두 개의 코어를 가지고 있으나 두 개 코어의 배치가 수평이 아닌 수직적임을 알 수 있다. 수직적 배치의 듀얼코어 프로세서는 하단의 코어를 core1, 상단의 코어를 core2로 나타내고 있으며 가로축 0.016미터, 세로축 0.032미터의 길이를 가진 다이에 패키징 되어 있다.

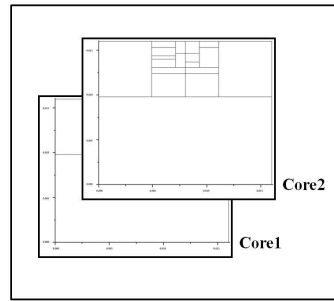


그림 4. 3차원 듀얼코어 프로세서
Fig 4. 3D dual-core processor

그림 4는 2층으로 적층된 2다이 3차원 듀얼코어 프로세서의 배치를 보여 주고 있다. 그림에서 각 코어는 가로, 세로 모두 0.016미터 길이의 다이에 패키징 되어 1,2층에 위치하고 있으며, 1층의 코어를 core1, 2층의 코어를 core2로 나타낸다.

2. 쿼드 코어 프로세서

그림 5는 4개의 코어를 가지는 2차원 쿼드코어 프로세서가 배치된 다이를 보여주고 있다. 그림에서 볼 수 있듯이, 2차원 쿼드코어 프로세서 다이에서 하단 좌측의 코어를 core1, 하단 우측의 코어를 core2, 상단 좌측의 코어를 core3, 그리고 상단 우측의 코어를 core4로 나타내며 다이의 길이는 가로, 세로 모두 0.032미터이다.

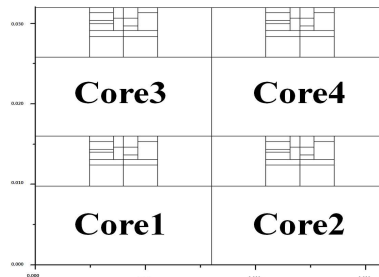


그림 5. 2차원 쿼드코어 프로세서
Fig 5. 2D quad-core processor

그림 6은 4개의 코어를 4개의 다이에 하나씩 패키징하여 4층으로 쌓은 4다이 3차원 쿼드코어 프로세서를 보여 주고 있다. 프로세서가 패키징된 다이의 크기는 가로, 세로 모두 0.016미터의 길이를 가지고 있으며 가장 하위 층의 코어부터 순차적으로 core1, core2, core3, core4로 나타낸다.

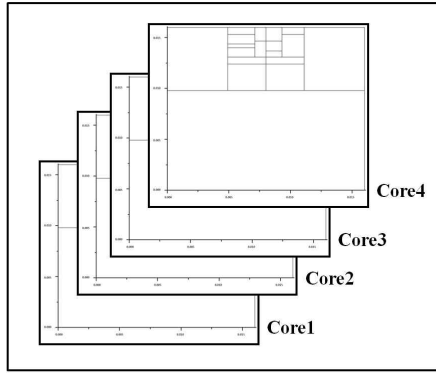


그림 6. 3차원 쿼드코어 프로세서
Fig 6. 3D quad-core processor

IV. 실험 환경

멀티코어 프로세서의 온도를 측정하는데 있어서 냉각 장치는 매우 중요한 요소이다. 실험 대상 프로세서의 기계적 냉각 장치로는 기존의 방열판을 그대로 사용한다. 본 논문에서 사용한 멀티코어 프로세서 구조의 성능과 온도를 정확하게 평가하기 위한 모의실험 환경과 온도 모델링 기법은 다음과 같다.

1. 모의실험 환경

본 논문에서는 멀티코어 프로세서의 성능과 온도를 평가하기 위한 도구로 Watch[16]와 HotSpot[17]을 사용하였다.

표 4. 모의실험 프로세서 구성 변수
Table 1. System parameter

실험인자	값
Functional Units	4 integer ALUs, 1 FP ALUs, 1 integer multiplier/divider, 1 FP multiplier/divider
L1 I-Cache	32KB, 4-way, 32byte lines, 1 cycle latency
L1 D-Cache	32KB, 4-way, 32byte lines, 1 cycle latency
L2 Cache	256KB, 8-way, 64byte lines, 12 cycle latency

HotSpot은 프로세서에서 발생하는 온도 값을 예측하는 시뮬레이터로 입력 값으로 각 코어의 전력 트레이스를 필요로 한다. 전력 트레이스 값은 SPEC CPU2000[18] 벤치마크 프로그램을 Watch에서 2억 개의 명령어를 수행하여 추출하였다. 모의실험 시 사용한 4-Way 슈퍼스칼라 프로세서의 구성 변

수는 아래의 표 1에 자세히 나타나있다. 명시하지 않은 값들은 기본 설정 값을 사용하였다.

본 실험에서는 SPEC CPU2000 벤치마크 프로그램에서 gcc, mcf 두 개의 애플리케이션을 선택하여 실험을 수행하였다. 선택된 두 개의 애플리케이션은 높은 명령어 레벨 병행성(ILP: Instruction Level Parallelism)과 메모리 요구성, 그리고 프로세서의 온도가 높게 발생한다는 특징을 가지고 있다. L2 캐시의 실패율(miss rates)은 프로세서의 성능에 미치는 영향이 크기 때문에 실험 시 중요하게 고려되어야 한다. 본 실험에서는 각기 다른 특징을 가진 애플리케이션을 수행하기 위하여, L2 캐시의 실패율이 높아 L2 캐시의 크기에 큰 영향을 받은 메모리 연계(Memory-bound) 애플리케이션(mcf)과 메모리에 큰 영향을 받지 않은 중앙처리장치 연계(CPU-bound) 애플리케이션(gcc)을 각각 선택하였다.

2. 온도 모델링

본 논문에서는 멀티코어 프로세서에서 발생하는 온도를 예측하기 위해서는 HotSpot 5.0을 사용한다. HotSpot을 이용하여 온도 값을 예측하기 위해 Wattch에서 측정된 각 구성요소의 전력 데이터를 입력으로 이용한다.

표 5. 온도 모델링 구성 변수
Table 2. Thermal modeling parameter

Parameters	Specific heat capacity (J/ m ³ K)	Resistivity (mK/W)	Thickness (m)	
Value	die-1	1.75e6	0.01	15e-5
	TIM-1	4e6	0.25	2e-5
	die-2	1.75e6	0.01	15e-5
	TIM-2	4e6	0.25	2e-5
	die-3	1.75e6	0.01	15e-5
	TIM-3	4e6	0.25	2e-5
	die-4	1.75e6	0.01	15e-5
	TIM-4	4e6	0.25	2e-5

플로어플랜 배치는 Alpha 21364 코어를 참조하여 구성요소의 영역과 순서를 먼저 설정한 후, hotfloorplan[19]을 이용하여 설계한다. 특히, 3차원 멀티코어 프로세서의 온도를 측정하기 위해서는 HotSpot 5.0에서 제공된 격자 형식(grid model)의 온도 평가 방법을 이용한다. 그리고 방열판은 가장 상위 계층에 위치하고 있으며 각 다이의 크기와 동일한 크기를 가지고 있는 것으로 가정하고 실험을 수행한다.

실험에서 사용한 3차원 멀티코어 프로세서의 온도 구성 변수 값은 표 2에서 자세히 보여준다. 표에서 die-1은 가장 하단

의 다이를 die-4는 가장 상단에 위치하는 다이를 각각 나타낸다. 방열판은 가장 위에 존재하기 때문에 die-1이 방열판과 가장 멀리 떨어져 있으며 die-4가 방열판과 가장 가까이에 위치하고 있다.

V. 실험결과 분석

프로세서 내에서 발생하는 열섬현상은 프로세서의 평균 온도가 아니라 어느 한 점에서 발생한 고온의 영향을 받기 때문에 실제 열섬현상을 해결하기 위해서는 가장 높은 온도를 줄여야 한다. 그러므로 본 논문에서는 프로세서의 평균 온도가 아닌 최고 온도를 대상으로 하여 분석을 수행한다. 뿐만 아니라, 본 논문에서는 3차원 멀티코어 프로세서에서의 온도에 대해 직접적으로 분석하기 위해서 기계적 냉각 기법과 기존의 동적 온도 제어기법을 그대로 사용한다. 본 논문에서 적용된 동적 온도 제어 기법에 대해서는 5.4절에서 설명하도록 한다.

1. 작업량

실험 결과 그래프의 가로축은 수행된 애플리케이션을 나타내는 것으로써 gcc는 g로 mcf는 m으로 표시하고 있으며 앞글자의 애플리케이션부터 core1에 할당된다. 예를 들면, gm의 경우 core1은 gcc를 수행하고, core2는 mcf를 수행한 경우를 말한다. 세로축은 코어에서 발생한 온도 중 최고 온도 값(°C)을 나타낸다.

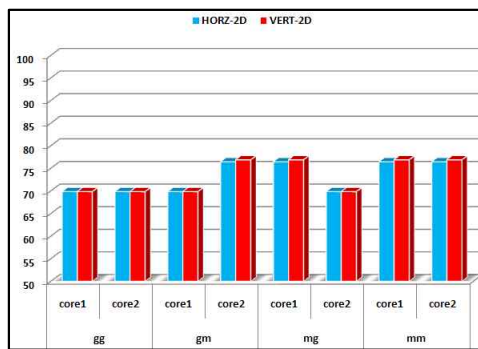


그림 7. 2차원 듀얼코어 프로세서에서의 각 코어 최고온도
Fig 7. Peak temperature on each core of 2D dual-core processor

그림 7은 2차원 수평적 배치의 듀얼코어 프로세서와 2차원 수직적 배치의 듀얼코어 프로세서에서 각 코어의 최고온도를 보여주고 있다. 그림 7에서 HORZ-2D, VERT-2D는 2차원 수평적 배치의 듀얼코어 프로세서와 2차원 수직적 배치의 듀얼

코어 프로세서를 각각 나타낸다. 같은 애플리케이션을 수행한 경우에는 (gg, mm) 두 코어 모두 비슷한 온도 패턴을 나타내었으며, 다른 애플리케이션을 수행한 경우에는 (gm, mg) 메모리 연계 프로그램인 mcf를 수행하는 코어의 온도가 조금 더 높게 발생한 것을 확인 할 수 있다. mcf 애플리케이션은 gcc 애플리케이션에 비해 작업량이 많기 때문에 mcf 애플리케이션을 수행하는 코어의 온도가 평균 76.67°C로 69.84°C인 gcc 애플리케이션을 수행한 코어에 비해 약 7°C 가량 높게 나타난다. 수평적 배치와 수직적 배치에 따른 멀티코어 프로세서의 온도 변화는 1°C 이하로 거의 발생하지 않음을 알 수 있다.

2. 방열판과의 거리

그림 8은 3차원 듀얼코어 프로세서에서 각 코어의 최고온도를 보여주고 있다. 그림 8에서 2die-3D는 2층으로 적층된 3차원 듀얼코어 프로세서를 나타낸다.

그래프에서 보이는 바와 같이, 3차원 듀얼코어 프로세서에서는 방열판에서 거리가 먼 core1의 온도가 방열판과 가까운 core2의 온도에 비해 gg를 수행한 경우에는 7.66°C, mg를 수행한 경우에는 10.46°C, gm을 수행한 경우에는 9.11°C, 그리고 mm을 수행한 경우에는 9.65°C 높게 발생한다. 그리고, 3차원 듀얼코어 프로세서에서 gm을 수행한 경우보다 mg를 수행한 경우의 core1의 온도가 더 높게 나타난다. gm의 경우에는 방열판과 멀리 있는 core1에서 gcc를 수행한 반면 mg는 core1에서 mcf를 수행한다는 사실에서 3차원 듀얼코어 프로세서에서의 코어 온도는 방열판과의 거리가 주된 원인이기는 하지만 2차원 듀얼코어 프로세서에서 보았듯이 수행되는 애플리케이션의 작업량도 영향을 미친다는 것을 알 수 있다.

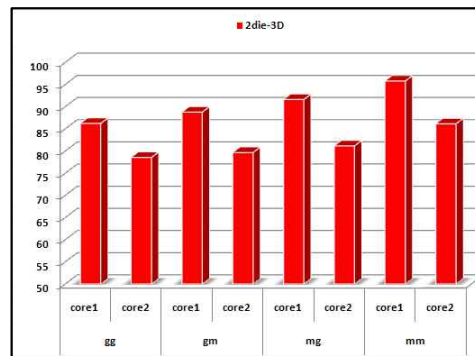


그림 8. 3차원 듀얼코어 프로세서의 각 코어 최고온도
Fig 8. Peak temperature on each core of 3D dual-core processor

그림 7의 2차원 듀얼코어 프로세서와 비교하여 3차원 듀얼코어 프로세서에서는 각 코어간의 온도차이가 크게 나타난다.

2차원 구조 멀티코어 프로세서에서는 층이 하나이기 때문에 모든 코어가 동일한 냉각 효과를 얻을 수 있는 반면 3차원 적층 구조 멀티코어 프로세서에서는 가장 상위 계층에서 아래로 내려갈수록 방열판과의 거리가 멀어져서 냉각 효과를 덜 받기 때문에 온도 차가 크게 나타난다.

3. 적층 다이의 수

그림 9와 그림 10은 프로세서 구성에 따른 듀얼코어 프로세서와 쿼드코어 프로세서의 최고 온도를 나타내고 있다. 그림 10에서 4die-3D는 4다이 3차원 쿼드코어 프로세서를 나타내고, Quad-2D는 2차원 쿼드코어 프로세서를 나타내고 있다. 가로 축은 코어의 순서대로 수행된 애플리케이션을 나타내기 때문에 gmgm은 core1은 gcc, core2는 mcf, core3는 gcc, 그리고 core4는 mcf를 수행하는 경우를 의미한다.

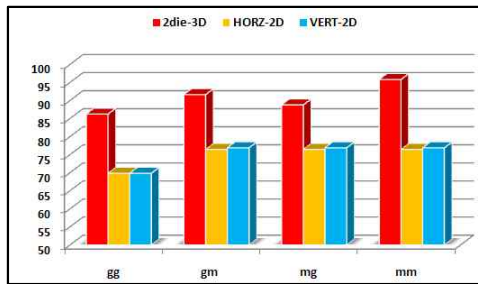


그림 9. 듀얼코어 프로세서의 최고온도
Fig. 9. Peak temperature of dual-core processor

그림 9에서 보이는 바와 같이, 3차원 듀얼코어 프로세서의 최고온도는 95.8°C(평균:90.60°C)로 76.45°C(평균:74.95°C)인 2차원 듀얼코어 프로세서의 최고온도에 비해 19.35°C(평균:15.81°C)나 높게 측정된다.

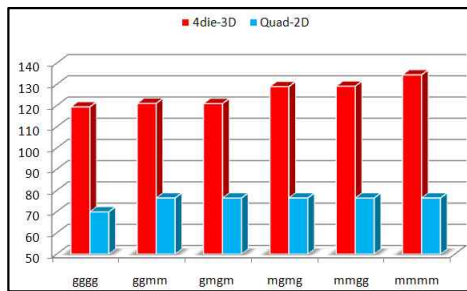


그림 10. 쿼드코어 프로세서의 최고온도
Fig. 10. Peak temperature of quad-core processor

쿼드코어 프로세서를 3차원 구조로 설계하는 경우(최대:134.48°C, 평균:125.65°C)에는 2차원 구조로 설계하는 경우(최대:76.48°C, 평균 75.43°C)보다 온도 차(최대:58.03°C, 평

균:50.30°C)가 더욱 크게 나타남을 그림 10을 통해 확인 할 수 있다. 그림 9와 그림 10을 통해 멀티코어 프로세서를 2차원 구조로 설계하는 경우와 비교하여 3차원 구조로 설계하는 경우에는 매우 높은 온도가 발생한다는 사실을 확인할 수 있다.

2차원 구조 멀티코어 프로세서 구조의 경우 듀얼코어와 쿼드코어에서의 최고 온도는 약 1°C(수평적 듀얼 코어 프로세서: 76.45°C, 수직적 듀얼코어 프로세서: 76.88°C, 쿼드코어 프로세서: 76.45°C)로 차이가 크지 않는 반면 3차원 멀티코어 프로세서 구조에서는 4다이의 쿼드코어 프로세서와 2다이 듀얼코어 프로세서의 최고온도 차가 48.33°C(듀얼코어 프로세서: 86.15°C, 쿼드코어 프로세서: 134.48°C)로 매우 크게 나타난다. 3차원 멀티코어 프로세서에서는 적층되는 다이의 수에 비례하여 높아진 전력밀도로 인해 온도가 급격히 높게 발생하는 것이다. 또한, 듀얼코어 프로세서와 마찬가지로 쿼드코어 프로세서에서도 애플리케이션과 방열판과의 거리에 따라 각 코어에서 발생하는 온도는 영향을 받는다.

4. 성능

동적 온도 제어 기법(DTM: Dynamic Thermal Management)은 프로세서의 온도를 제어하기 위해서 사용되는 대표적인 기법이다. 동적 온도 제어 기법은 가변 전압 주파수 조절(DVFS: Dynamic Voltage and Frequency Scaling)이나 인출 지연(fetch throttling) 기법을 통해 구현되고 있다[20-22]. 동적 온도 제어 기법은 온도를 효과적으로 제어하기는 하지만 이를 위해서 전압이나 주파수 등을 조절하기 때문에 성능을 희생시킨다는 단점이 존재한다. 그래서 많은 연구자들은 성능 희생을 최소화하면서 온도를 제어하는 동적 온도 제어 기법을 연구하고 있다. 3차원 적층 구조 멀티코어 프로세서에서의 온도 문제를 해결하기 위한 동적 온도 제어 기법 적용 연구 또한 최근 이루어지고 있다[23]. 본 논문에서도 멀티코어 프로세서에서 발생하는 온도 문제를 해결하기 위해 동적 온도 제어 기법을 적용하였다.

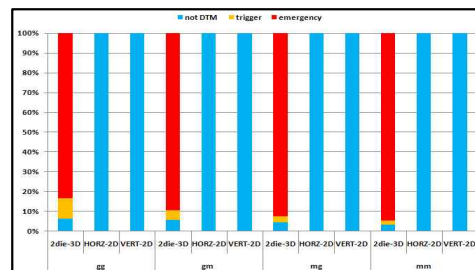


그림 11. 듀얼코어 프로세서의 동적 온도 제어 기법 사용률
Fig. 11. DTM utilization ratio of dual-core processor

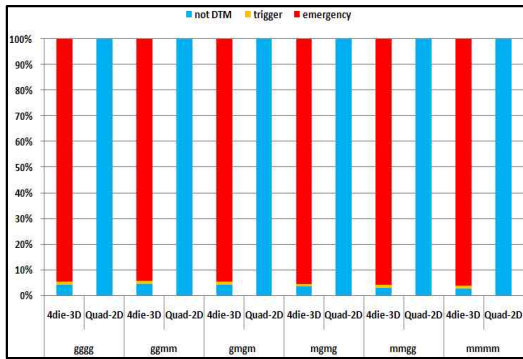


그림 12. 쿼드코어 프로세서의 동적 온도 제어 기법 사용률
Fig 12. DTM utilization ratio of quad-core processor

적용된 동적 온도 제어 기법은 2단계로 동작한다. 1단계는 구동단계(trigger threshold)로 명령어의 인출 지연 방법으로 온도를 조절한다. 2단계는 긴급단계(emergency threshold)로 가변전압 및 주파수 조절 방법으로 온도를 제어한다. 국제 반도체 기술 로드맵(ITRS : International Technology Roadmap for Semiconductors)에 따르면, 반도체 공정이 0.13um 이하에서는 칩의 적합 온도는 90°C 보다 낮아야 한다. 그러므로 칩의 신뢰성을 안정적으로 확보하기 위하여, 모의실험 시 구동단계 온도는 80°C로 설정하였고, 긴급단계온도는 85°C로 설정하였다.

그림 11과 그림 12에서 not DTM은동적온도제어기법이작동하지않은상태로어떤성능저하도일어나지않은경우를나타내고, trigger는 1단계 구동단계의 동적 온도 제어 기법이 작동한 상태로 약간의 성능 저하가 일어난 경우를 의미한다. 마지막으로 emergency는 2단계 긴급단계의 동적 온도 제어 기법이 작동한 상태로 온도 제어를 위해 큰 성능 저하를 감수하고 있는 상태이다. 좌측은 프로세서가 동작하고 있는 동안의 전체 시간 중에서 동적 온도 제어 기법이 적용된 시간의 비율을 나타낸다.

2차원 멀티코어 프로세서의 경우 듀얼코어와 쿼드코어 모두에서 최고 온도가 80°C가 넘지 않았기 때문에 동적 온도 제어 기법이 작동하지 않았다. 결과적으로 본 논문의 실험에서 2차원 멀티코어 프로세서는 온도 제어를 위한 성능 저하는 발생하지 않았음을 확인할 수 있다. 반면에, 3차원 멀티코어 프로세서 경우 듀얼코어 프로세서는 프로그램 실행시간의 약 95.2%, 쿼드코어 프로세서는 실행시간의 약 96.5%에서 동적 온도 제어 기법이 적용되었다. 그래프에서 볼 수 있듯이, 3차원 멀티코어 프로세서에서는 발생하는 온도를 제어하기 위하여 동적 온도 제어 기법이 프로그램이 수행하는 대부분의 시간에 적용이 되고 있다. 프로그램이 수행되는 대부분의 시간에 동적 온도 제어 기법이 적용된다는 것은 전반적인 프로그

램 수행 시간이 증가하여 프로세서의 전체적인 성능이 저하된다는 것을 의미한다. 특히 쿼드코어 프로세서의 경우 듀얼코어 프로세서에 비해서 구동 단계의 시간(듀얼코어 프로세서: 4.93%, 쿼드코어 프로세서: 1.29%)이 짧아 더 많은 성능이 저하될 것으로 예상된다.

VI. 결론

멀티코어 프로세서를 3차원으로 설계하는 경우에는 높은 전력밀도로 인해 발생하는 높은 온도가 프로세서의 신뢰성에 악영향을 미치므로 효율적인 온도제어가 매우 중요하다. 본 논문에서는 3차원 멀티코어 프로세서에서 작업량, 방열판과의 거리, 그리고 적용되는 다이의 개수와 발생하는 온도간의 상관관계를 자세히 분석해 보았다. 분석 결과, 3차원 멀티코어 프로세서에서 온도에 미치는 영향은 적용되는 다이의 개수, 방열판과의 거리, 작업량 순으로 나타났으며, 동적 온도 제어 기법의 적용 시간이 프로그램 수행시간의 대부분을 차지하여 시스템 성능이 많이 저하시킬 것으로 예상된다. 본 논문의 실험 결과를 통해 3차원 적층 구조로 설계된 멀티코어 프로세서에서 칩의 신뢰성과 향상된 성능을 얻고자 한다면 설계 과정에서부터 온도를 반드시 고려해야 한다는 사실을 다시 한 번 확인할 수 있다. 뿐만 아니라, 3차원 구조의 멀티코어 프로세서를 설계하고자 하는 경우에 본 논문의 분석 결과는 온도 제어를 위한 방향을 제시하는데 큰 도움이 되리라 기대된다. 향후에는 본 연구 내용을 바탕으로 온도와 프로세서 내부 구조 사이의 구체적 관계 모델을 제시하여, 3차원 구조의 멀티코어 프로세서에서의 온도를 효과적으로 제어하는 기법에 대한 연구를 수행하고자 한다.

참고문헌

- [1] V. Agarwal, M.S. Hrishikesh, S.W. Keckler and D. Burger, "Clock rate versus IPC: the end of the road for conventional microarchitectures," In Proceedings of the 27th International Symposium on Computer Architecture, pp. 248-259, 2000.
- [2] K. Olukotun, L. Hammond and J. Laudon, "Chip Multiprocessor Architecture: Techniques to Improve Throughput and Latency," Morgan&Claypool, 2007.

- [3] D. M. Tullsen, S. Eggers, H. M. Levy, "Simultaneous multithreading: Maximizing on-chip parallelism," In proceedings of the 22nd international symposium on Computer Architecture, pp. 392-403, 1995.
- [4] C. Isci, A. Buyuktosunoglu, C.Y. Cher, P. Bose and M. Martonosi, "An Analysis of Efficient Multi-Core Global Power Management Policies: Maximizing Performance for a Given Power Budget", In Proceedings of 39th Annual IEEE/ACM International Symposium on Microarchitecture, pp. 347-358, 2006.
- [5] R. Kumar, V. Zyuban and D.M. Tullsen, "Interconnections in multi-core architectures: Understanding mechanisms, overheads and scaling," In Proceedings of The 32nd International Symposium on Computer Architecture, pp. 408-419, 2005.
- [6] B. Black, D.W. Nelson, C. Webb and N. Samra, "3D Processing Technology and its Impact on IA32 Microprocessors," In Proceedings of IEEE International Conference on Computer Design, pp. 316-318, 2004.
- [7] D. Nelson, C. Webb, D. McCauley, K. Raol, J. R. II, J. DeVale and B. Black, "A 3D Interconnect Methodology Applied to IA32-class Architectures for Performance Improvements through RC Mitigation," In Proceedings of the 21st International VLSI Multilevel Interconnection Conference, pp. 78-83, 2004.
- [8] S. W. Yoon, D. W. Yang, J. H. Koo, M. Padmanathan and F. Carson, "3D TSV processes and its assembly/Packaging technology," In Proceedings of IEEE International Conference on 3D System Integration, pp. 1-5, 2009.
- [9] S.Das and A.Chandrakasan, "Timing, energy, and thermal performance of three-dimensional integrated circuits," In proceedings of Great Lakes Symposium on VLSI, pp. 338-343, 2004.
- [10] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan and D. Tarjan, "Temperature-Aware Microarchitecture," In proceedings of the 30th International Symposium on Computer Architecture, pp. 2-13, 2003.
- [11] K. Skadron, M. R. Stan, K. Sankaranarayanan, W. Huang, S. Velusamy and D. Tarjan, "Temperature-Aware Microarchitecture: Modeling and Implementation," Transactions on Architecture and Code Optimization, Vol. 1, No. 1, pp. 94-125, 2004.
- [12] S. Im and K. Banerjee, "Full chip thermal analysis of planar (2-D) and vertically integrated (3-D) high performance ICs," In proceedings of International IEDM Technical Digest, pp. 727-730, 2000.
- [13] International Technology Roadmap for Semiconductors, available at <http://public.itrs.net>.
- [14] Jichuan Chang and Gurindar S. Sohi, "Cooperative Caching for Chip Multiprocessors," In Proceedings of the 33rd international symposium on Computer Architecture, pp. 264-276, 2006.
- [15] R. E. Kessler, "The Alpha 21364 microprocessor," Journal of IEEE MICRO, Vol. 19, Issue 2, pp. 24-36, 1996.
- [16] D. Brooks, V. Tiwari and M. Martonosi, "Wattch: A Framework for Architectural-level Power Analysis and Optimizations," In Proceedings of the 27th Annual International Symposium on Computer Architecture, pp. 83-94, 2000.
- [17] W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan and S. Ghosh, "HotSpot: A Compact Thermal Modeling Method for CMOS VLSI Systems," IEEE Transactions on VLSI Systems, 2006.
- [18] J. L. Henning, "SPEC CPU2000: measuring cpu performance in the new millennium," IEEE Computer, Vol. 33, no. 7, pp. 28-35, 2000.
- [19] S. Chung and K. Skadron, "Using On-Chip Event Counters For High-Resolution, Real-Time Temperature Measurement," In Proceedings of Thermal and Thermomechanical Phenomena in Electronics Systems, pp. 114-120, 2006.
- [20] D. Brooks and M. Martonosi, "Dynamic Thermal Management for High-Performance Microprocessors," In Proceedings of the 7th International Symposium on High-Performance Computer Architecture, pp. 171-182, 2001.
- [21] K. Choi, R. Soma and M. Pedram, "Dynamic voltage and frequency scaling based on workload decomposition," In Proceedings of the 2004 international symposium on Low power electronics and design, pp. 174-179, 2004.
- [22] L. Benini, G. De Micheli, E. Macii, M. Poncino and R. Scarsi, "Symbolic synthesis of clock-gating logic for power optimization of synchronous controllers," In Transactions on Design Automation of Electronic Systems (TODAES), Vol. 4, no. 4, pp. 351-375, 1999.

[23] A. K. Coskun, T. S. Rosing, J. Ayala, D. Atienza and Y. Leblebici, "Dynamic thermal management in 3D multicore architectures," In proceedings of Design Automation and Test in Europe (DATE), pp. 1410-1415, 2009.

저 자 소 개



최 홍 준
 2009년 : 전남대학교 전자컴퓨터공학과 공학사
 2010년 : 전남대학교 전자컴퓨터공학부 석사
 2011년 : 전남대학교 전자컴퓨터공학부 박사과정 입학
 관심분야 : 저전력 설계, 병렬 처리, 컴퓨터 구조
 Email : chj6083@gmail.com



안 진 우
 2010년 : 전남대학교 전자컴퓨터공학과 공학사
 2010년 : 전남대학교 전자컴퓨터공학부 석사과정 입학
 관심분야 : SoC 설계, 임베디드 하드웨어 설계, 컴퓨터구조
 Email : ajw0411@gmail.com



장 형 범
 2007년 : 고려대학교 정보통신대학 컴퓨터 공학과 공학사
 2009년 : 고려대학교 컴퓨터정보통신공학부 석사
 2009년 : 고려대학교 컴퓨터정보통신공학부 박사과정 입학
 관심분야 : 저전력 설계, 저온도 설계, SSD
 Email : kuphy01@korea.ac.kr



김 종 먼
 1995년 : 명지대학교 전기공학사
 2000년 : University of Florida ECE 석사
 2005년 : Georgia Institute of Technology ECE 박사
 2005년-2007년 : 삼성종합기술원 전 입연구원
 2007년-현재 : 울산대학교 컴퓨터 정보통신공학부 교수
 관심분야 : 임베디드 SoC, 컴퓨터구조, 프로세서 설계, 병렬처리
 Email : jongmyon.kim@gmail.com



김 철 흥
 1998년 : 서울대학교 컴퓨터공학사
 2000년 : 서울대학교 컴퓨터 공학부 석사
 2006년 : 서울대학교 전기컴퓨터공학부 박사
 2005년-2007년 : 삼성전자 반도체총괄 SYS.LSI사업부 책임 연구원
 2007년-현재 : 전남대학교 전자 컴퓨터 공학부 교수
 관심분야 : 임베디드시스템, 컴퓨터 구조, SoC 설계, 저전력 설계
 Email : chkim22@chonnam.ac.kr