

캐쉬 구성에 따른 3차원 쿼드코어 프로세서의 성능 및 온도 분석

손동오*, 안진우*, 최홍준*, 김종면**, 김철홍*

Analysis on the Performance and Temperature of the 3D Quad-core Processor according to Cache Organization

Dong-Oh Son*, Jin-Woo Ahn*, Hong-Jun Choi*, Jong-Myon Kim**, Cheol-Hong Kim*

요약

공정기술이 지속적으로 발달함에 따라 멀티코어 프로세서는 성능 향상이라는 장점과 함께 내부 연결망의 긴 지연 시간, 높은 전력 소모, 그리고 발열 현상 등의 문제점들을 내포하고 있다. 이와 같은 2차원 멀티코어 프로세서의 문제점들을 해결하기 위한 방안 중 하나로 3차원 멀티코어 프로세서 구조가 주목을 받고 있다. 3차원 멀티코어 프로세서는 TSV를 이용하여 수직으로 쌓은 여러 개의 레이어들을 연결함으로써 2차원 멀티코어 프로세서와 비교하여 배선 길이를 크게 줄일 수 있다. 하지만, 3차원 멀티코어 프로세서에서는 여러 개의 코어들이 수직으로 적층되므로 전력밀도가 증가하고, 이로 인해 발열문제가 발생하여 높은 냉각 비용과 함께 신뢰성에 부정적인 영향을 유발한다. 따라서 3차원 멀티코어 프로세서를 설계할 때에는 성능과 함께 온도를 반드시 고려하여야 한다. 본 논문에서는 캐쉬 구성에 따른 3차원 쿼드코어 프로세서의 온도를 상세히 분석하고, 이를 기반으로 발열문제를 해결하기 위해 저온도 캐쉬 구성 방식을 제안하고자 한다. 실험결과, 명령어 캐쉬는 최고온도가 임계값보다 낮고 데이터 캐쉬는 많은 웨이를 가지는 구성을 적용할 때 최고온도가 임계값보다 높아짐을 알 수 있다. 또한, 본 논문에서 제안하는 캐쉬 구성은 쿼드코어 프로세서를 사용하는 3차원 구조에서 캐쉬의 온도 감소에 효과적일 뿐만 아니라 성능 저하 또한 거의 없음을 알 수 있다.

▶ Keywords : 쿼드코어 프로세서, 3차원 집적회로, 온도, 캐쉬 메모리, 사상방식

Abstract

As the process technology scales down, multi-core processors cause serious problems such as

• 제1저자 : 손동오 • 교신저자 : 김철홍

• 투고일 : 2012. 01. 10, 심사일 : 2012. 03. 14, 게재확정일 : 2012. 05. 03.

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

** 울산대학교 컴퓨터정보통신공학부(School of Computer Engineering and Information Technology, University of Ulsan)

※ 본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터(NIPA-2012-H0301-12-3005)의 지원과 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단 기초연구사업의 지원을 받아 수행된 연구임(2011-0003350)

increased interconnection delay, high power consumption and thermal problems. To solve the problems in 2D multi-core processors, researchers have focused on the 3D multi-core processor architecture. Compared to the 2D multi-core processor, the 3D multi-core processor decreases interconnection delay by reducing wire length significantly, since each core on different layers is connected using vertical through-silicon via(TSV). However, the power density in the 3D multi-core processor is increased dramatically compared to that in the 2D multi-core processor, because multiple cores are stacked vertically. Unfortunately, increased power density causes thermal problems, resulting in high cooling cost, negative impact on the reliability. Therefore, temperature should be considered together with performance in designing 3D multi-core processors. In this work, we analyze the temperature of the cache in quad-core processors varying cache organization. Then, we propose the low-temperature cache organization to overcome the thermal problems. Our evaluation shows that peak temperature of the instruction cache is lower than threshold. The peak temperature of the data cache is higher than threshold when the cache is composed of many ways. According to the results, our proposed cache organization not only efficiently reduces the peak temperature but also reduces the performance degradation for 3D quad-core processors.

▶ Keywords : Quad-core processor, 3D integrated circuits, Temperature, Cache memory, Mapping method

1. 서 론

마이크로프로세서의 동작 주파수는 공정기술의 발달에 힘입어 지속적으로 향상되고 있다. 동작 주파수의 증가는 마이크로프로세서의 성능을 크게 향상시키는 역할을 하지만, 이로 인해 전력소모가 증가하고 심한 발열이 발생하는 문제의 원인이 되기도 한다. 이와 같은 문제점들로 인한 주파수 증가의 한계를 극복하기 위해, 한 개의 코어를 사용하는 싱글코어 프로세서에서 두 개 이상의 코어를 사용하는 멀티코어 프로세서로 마이크로프로세서의 패러다임이 변화하게 되었다[1][2][3][4]. 멀티코어 프로세서는 싱글코어 프로세서와 비교하여 상대적으로 낮은 동작 주파수를 가지는 여러 개의 코어들을 하나의 칩에 집적함으로써 기존의 싱글코어 프로세서에서 발생하는 높은 전력소모와 발열의 문제를 해결할 수 있다는 장점을 지닌다.

멀티코어 프로세서는 코어들과 구성요소들 사이의 통신을 위해 내부 연결망(Interconnection)을 사용하는데, 공정기술의 발달에 따라 집적도가 향상되면서 프로세서 내부 구성요소들의 크기는 줄어들고 내부 연결망을 구성하는 배선(Wire)의 길이는 상대적으로 늘어나는 문제점이 발생하였다. 배선

길이의 증가는 내부 연결망에서 소모되는 지연 시간(Interconnection Delay)이 커지는 현상으로 이어졌다. 또한, 내부 연결망에서의 높은 전력소모와 발열 현상과 같은 문제점들을 유발하게 된다[5]. 최근에는, 멀티코어 프로세서에서 발생하는 이와 같은 문제점들을 해결하기 위한 기법 중 하나로 3차원 적층구조를 활용하는 3차원 멀티코어 프로세서 설계 기법에 관한 연구가 많은 주목을 받고 있다 [6][7][8][9][10][11].

3차원 멀티코어 프로세서는 코어를 수평적으로 배치하는 기존의 2차원 멀티코어 프로세서와 달리 코어를 수직으로 적층한다. 평면적으로 생산되는 기존의 2차원 멀티코어 프로세서보다 적은 면적을 사용하기 때문에 칩의 생산 비용을 감소시키는 장점이 있다[12]. 또한, 2차원 구조에서의 내부 연결망과 달리 TSV(Through-Silicon Via)를 통하여 수직으로 각 코어들을 연결하기 때문에, 내부 연결망 지연시간을 크게 감소시킬 수 있다[13]. 또한, 내부 연결망의 길이가 줄어들면서 배선에서의 전력 소모 또한 감소하게 된다. 하지만, 코어들이 수직으로 적층됨에 따라 전력밀도가 상승하게 되고, 이로 인해 열섬현상(Hotspot)이 발생하는 단점이 있다[14]. 열섬현상은 칩의 변형을 일으키고 신뢰성에 악영향을 미치는 심각한 문제점으로 인식되고 있다. 따라서 3차원 멀티코어 프로세서를 효율적으로 설계하기 위해서는 온도 문제에 대한 고

려가 반드시 이루어져야 한다. 본 논문에서는 최근 주로 사용되는 쿼드코어 프로세서를 3차원 구조로 개발 시 반드시 고려되어야 하는 온도 문제를 상세하게 분석하기 위한 방안 중 하나로 3차원으로 적층된 캐쉬(Cache)의 구성에 따른 쿼드코어 프로세서의 성능 및 온도 변화를 살펴보고자 한다.

캐쉬는 중앙처리장치(Central Processing Unit)와 주 기억장치(DRAM)의 속도 차이를 완화해주는 고속의 버퍼 역할을 하는 장치로, 레지스터(Register)에서 사용할 데이터를 주 기억장치로부터 미리 가져와 레지스터에 데이터를 제공해주는 역할을 수행하는 마이크로프로세서에서 매우 중요한 구성요소이다. 캐쉬의 구조는 사상방식(Associativity)에 따라 가장 간단한 직접사상(Direct Mapping)과 가장 복잡한 구조의 완전연관사상(Fully-Associative Mapping), 그리고 이 두 가지를 혼합한 세트연관사상(Set-Associative Mapping)으로 구분된다. 2차원 평면구조 마이크로프로세서에서는 캐쉬의 사상방식에 따라 프로세서의 성능과 온도가 달라지는데 웨이(Way) 수가 작고 단순한 캐쉬는 상대적으로 낮은 성능과 온도를 보이고 웨이 수가 많고 복잡한 캐쉬는 상대적으로 높은 성능과 온도를 보인다. 본 논문에서는 3차원 쿼드코어 프로세서 설계 시 발생하는 온도 문제를 해결하기 위한 방안을 찾기 위해, 캐쉬 구성에 따른 성능 및 온도 변화를 상세하게 분석하고자 한다.

캐쉬의 구성에 따라 변화하는 성능과 온도 변화에 대한 연구는 현재까지 많은 연구가 진행되어 왔다[15][16][17]. 하지만, 3차원 멀티코어 프로세서에서 캐쉬 구성에 따른 특성을 분석하는 논문은 활발하게 이루어지지 않았다. 따라서 본 논문에서는 3차원 멀티코어 프로세서에 적합한 캐쉬 구성을 알아보기 위해, 캐쉬 구성에 따른 3차원 쿼드코어 프로세서의 성능과 온도를 분석해 보고자 한다. 실험의 다양성을 위해서는 완전연관사상 캐쉬 구성을 가지는 프로세서에 대한 분석 또한 수행하는 것이 가능하지만, 현재 상용화된 일반적인 고성능 마이크로프로세서들은 대부분 8-way에서 16-way 세트연관사상 캐쉬 구조를 적용하고 있으며[18], 웨이 수가 급격하게 늘어나면 설계 복잡도와 높은 전력 소모로 인해 3차원 구조에서는 심각한 문제점들이 발생할 것으로 예상되므로 많은 수의 웨이를 가지는 세트연관사상 구조와 완전연관사상 구조에 대한 분석은 수행하지 않는다. 이와 관련된 실험은 3장에서 자세히 설명한다. 따라서 본 논문에서는 직접사상 구조(1-way)부터 16-way 세트연관사상 캐쉬 구성을 3차원 쿼드코어 프로세서 구조에 적용하면서 이로 인한 프로세서의 성능 및 온도 변화를 분석하고자 한다.

이하 본 논문의 구성은 다음과 같다. 2장에서는 관련내용

인 3차원 멀티코어 프로세서 구조에 대해서 기술하고, 3장에서는 본 논문의 실험에 적용한 프로세서 구조에 대해서 기술한다. 4장에서는 실험환경과 자세한 실험결과를 기술하고, 마지막으로 5장에서 결론을 맺고자 한다.

II. 3차원 멀티코어 프로세서

공정기술의 발달과 함께 내부 구성 요소들의 크기가 작아지면서 칩의 크기는 감소하지만, 칩 내부 연결망의 길이는 상대적으로 늘어나게 된다. 이에 따라 내부 연결망을 이용하는 데이터 전송 지연시간과 내부 연결망에서 소모되는 전력은 증가하게 된다. 이는 최신 멀티코어 프로세서를 개발하는데 있어서 큰 제약조건이 되고 있다. 이와 같은 내부 연결망 문제를 해결하기 위한 기법 중에서 최근 가장 주목을 받고 있는 방안이 적층구조를 통해 3차원 멀티코어 프로세서를 설계하는 것이다.

3차원 멀티코어 프로세서는 기존의 2차원 평면구조 멀티코어 프로세서의 문제점인 연결망 길이를 감소시키기 위해 두 개 이상의 레이어(Layer)들을 수직으로 적층하여 하나의 칩에 통합하는 기술이다. 3차원 멀티코어 프로세서 기술은 레이어의 숫자를 n 이라 할 때 2차원 구조와 비교하여 내부 연결망의 길이를 \sqrt{n} 만큼 줄여주는 장점이 있다[19]. 이를 통해 데이터 전송 지연 시간과 연결망에서 소모되는 전력을 감소시킨다. 3차원 멀티코어 프로세서는 TSV를 통하여 레이어 사이에 통신이 이루어진다. 즉, 3차원으로 적층된 레이어 사이에 통로를 뚫어 그 사이를 구리를 사용하여 전극을 형성시킴으로써 내부 연결망을 구축하는 것이다. 최근까지는 TSV를 활용한 3차원 멀티코어 프로세서 개발이 미미한 수준이지만, 향후 3차원 멀티코어 프로세서를 개발할 때에는 TSV를 이용한 기술이 주로 사용될 것으로 예상된다[6].

위에서 기술한 장점들에도 불구하고 3차원 멀티코어 프로세서는 설계 시 주요한 제약 조건을 내포하고 있다. 코어들이 수직으로 적층되기 때문에 전력밀도가 상승하게 되고, 이로 인해 높은 발열이 발생하는 것이다. 발열으로 인한 열섬현상은 칩에 변형을 일으켜 신뢰성 문제와 타이밍 오류(Timing Error)를 발생시킨다[19][20]. 구성요소들이 수직으로 적층됨에 따라 고온이 발생하는 특정 장치에서는 이러한 문제가 더욱 심화된다[21]. 기존의 연구결과에 따르면 Alpha 계열 마이크로프로세서는 코어 2개를 수직으로 쌓으면 기존의 2차원 멀티코어 프로세서보다 17~20°C의 온도가 증가할 수 있다고 한다[20]. 3차원 멀티코어 프로세서에서의 온도 문제는

데이터 연결 통로로 사용되는 TSV에서도 발생한다. 65nm 공정에서 TSV가 칩에서 차지하는 비율은 1~2%밖에 되지 않아 칩에 미치는 온도 영향이 미미하지만, 공정이 발달함에 따라 칩 면적이 감소하게 되면 TSV가 차지하는 면적 비율이 증가하게 되어 칩의 온도에 큰 영향을 미치게 된다[12]. 그러므로 3차원 멀티코어 프로세서를 개발할 때에는 온도 문제를 해결하기 위해 전력밀도와 칩 면적당 TSV 면적을 고려하여 설계를 진행해야 한다.

일반적으로 마이크로프로세서의 최대온도는 85~110°C 사이가 한계이다[22]. 따라서 본 논문에서는 칩에 문제를 발생시킬 수 있는 임계값 온도를 90°C로 지정하고, 3차원 쿼드코어 프로세서에서 90°C를 넘지 않는 최적의 성능을 도출하는 캐쉬 구성을 살펴보고자 한다.

III. 적용된 캐쉬 구성 및 기초 분석

그림 1은 본 논문에서 사용한 3차원 쿼드코어 프로세서 구조를 보이고 있다. 4개의 레이어가 3차원으로 적층되어 있고 아래부터 레이어1, 레이어2, 레이어3 그리고 가장 상위에 레이어4가 적층된다. 하나의 레이어는 크게 L2 캐쉬와 코어로 나눌 수 있다. L2 캐쉬는 코어의 주변에 배치되어 있고, 코어 내부의 하단 부분은 명령어 캐쉬와 데이터 캐쉬가 배치되어 있다. 레이어1의 아래에는 방열판 역할을 담당하는 히트싱크(Heat Sink)가 배치된다.



그림 1. 3차원 쿼드코어 프로세서 구조
Figure 1. Architecture of 3D Quad-core processor

본 논문에서는 명령어 캐쉬와 데이터 캐쉬를 각각 변경하

여 3차원 쿼드코어 프로세서에서 캐쉬의 구성 변화가 프로세서의 성능과 온도에 어떤 영향을 미치는지 알아보려고 한다. 앞에서 기술한 바와 같이, 캐쉬 구성을 직접사상부터 16-way 세트연관사상까지 변화시키면서 실험을 수행한다. 일반적인 고성능 마이크로프로세서는 주로 8-way 세트연관사상부터 16-way 세트연관사상을 사용하므로, 본 논문에서는 16-way 세트연관사상까지만 실험을 수행한다.

1. 사상 방식에 따른 캐쉬의 기본 전력 소모 분석

3차원 쿼드코어 프로세서에서 캐쉬 구성에 따른 성능 및 온도를 분석하기에 앞서 캐쉬의 온도에 많은 영향을 미치는 캐쉬의 기본 전력을 사상 방식에 따라 분석해 보고자 한다. 본 연구를 위한 예비 실험으로써 사상 방식별 캐쉬의 기본 소모 전력을 측정해 보았다.

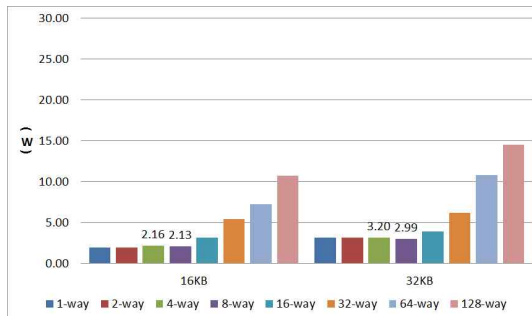


그림 2. 명령어 캐쉬 기본 전력 소모
Figure 2. Power consumption of instruction cache

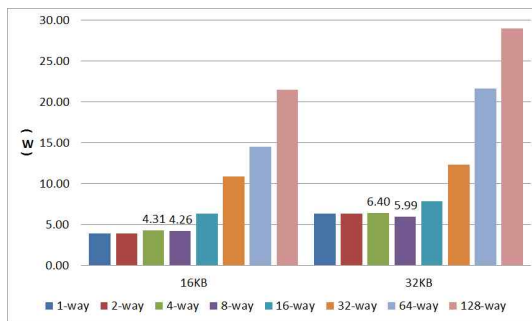


그림 3. 데이터 캐쉬 기본 전력 소모
Figure 3. Power consumption of data cache

그림 2와 그림 3은 16KB와 32KB의 크기를 가지며 직접 사상부터 16-way 세트연관사상까지의 구성에 따른 캐쉬의 기본 소모 전력을 보인다. 캐쉬에서 소모되는 전력은 디코더,

워드라인, 비트라인, 감지증폭기, 태그배열에서 소모되는 전력의 합으로 이뤄진다. 그림 2와 그림 3을 보면 기본적으로 16KB 크기의 캐쉬보다 32KB 크기의 캐쉬가 상대적으로 기본 전력 소모가 많음을 알 수 있다. 이는 캐쉬의 크기가 늘어남에 따라 웨이 개수 또는 라인의 수가 늘어나 각 유닛이 증가하기 때문이다. 또한, 데이터 캐쉬가 명령어 캐쉬보다 기본 전력 소모가 높은 것을 알 수 있다. 명령어 캐쉬는 읽기 동작만을 수행하기 때문에 읽기/쓰기 동작을 모두 하는 데이터 캐쉬에 비해 기본 전력 소모가 낮다. 따라서 3차원 쿼드코어 프로세서에서는 명령어 캐쉬보다 데이터 캐쉬에서 온도 문제가 발생할 가능성이 높음을 알 수 있다. 특이점으로는 웨이 수가 늘어남에 따라 기본적으로 소모되는 전력이 일정하게 증가할 것이라는 예상과 달리 8-way 세트연관사상에서 기본 전력 소모가 감소하는 것을 알 수 있다. 이러한 이유는 다음과 같다. 캐쉬의 기본 전력 소모는 캐쉬의 크기, 블록 크기, 웨이 수, 태그 수, 워드라인 개수, 비트라인 개수, 세트 개수 등의 다양한 유닛의 구성에 따라 기본 전력 소모 값이 정해진다. 이러한 유닛들은 최적화된 조합을 통해 캐쉬가 구성된다. 따라서 8-way 이하의 세트연관사상에서는 캐쉬 최적화에 따라 캐쉬의 기본 전력 소모가 웨이 수에 비례하여 증가하지 않는다. 16KB, 32KB 크기의 캐쉬는 캐쉬 최적화에 따라 8-way 세트연관사상에서 가장 적은 전력 소모를 보여준다. 하지만 8-way 이상으로 웨이 수가 증가하게 되면 최적화에도 불구하고, 복잡도의 증가 또는 유닛의 증가로 인해 전력이 급격히 증가하기 때문에 캐쉬 최적화의 효과를 볼 수 없다. 이에 따라 직접사상부터 8-way 세트연관사상까지는 캐쉬 최적화 따라 약간의 소모전력의 변화가 있지만 16-way 세트연관사상부터는 급격한 소모전력의 증가를 볼 수 있다. 실험결과 128-way 세트연관사상은 직접사상보다 4배 이상의 기본 소모전력이 증가하는 것을 볼 수 있다. 또한, 16-way 이상의 많은 웨이를 가지는 세트연관사상에서는 기본 전력 소모의 급격한 증가로 인해 3차원 쿼드코어 프로세서에 적용 시 심각한 온도 문제가 예상된다.

IV. 실험 결과 및 분석

1. 모의실험 환경

3차원 쿼드코어 프로세서에서 캐쉬의 구성에 따라 변화하는 성능과 온도를 분석하는 모의실험을 위해 본 논문에서는 SimpleScalar[23] 시뮬레이터를 사용하였다.

표 1. 프로세서 구성 변수
Table 1. System parameters

실험 인자	값
Functional Units	4 integer ALUs, 4 FP ALUs, 1 integer multiplier/divider, 1 FP multiplier/divider
L1 I-Cache	16KB to 32KB, 4-way, 32byte lines, 1 cycle latency
L1 D-Cache	16KB to 32KB, 4-way, 32byte lines, 1 cycle latency
L2 Cache	2MB, 8-way, 128byte lines, 12 cycle latency
Mapping Method	Direct Mapping(1-way) Set-Associative Mapping (2-way ~ 16-way)

캐쉬의 크기와 구성에 따라 소모되는 기본 전력을 측정하기 위해서는 CACTI[24]를 사용하고, 캐쉬에서 소모되는 전체 전력을 구하기 위해서는 Wattch[25]를 사용하였다. Wattch를 통하여 얻은 전력 값을 이용하여 3차원 쿼드코어 프로세서에서 발생하는 온도를 예측하기 위해서는 Hotspot5.0[26]을 사용하였다. 실험에 적용된 기본적인 코어프로세서 구조는 Alpha 21264(EV6)[27][28]를 사용하였다.

벤치마크 프로그램으로는 SPEC CPU2000[29] 벤치마크 프로그램들 중 웨이 변화에 따라 성능이 가장 크게 변하는 4개(정수 2, 실수 2)의 프로그램들인 crafty, equake, mcf, mesa를 선정하여 실험을 수행하였다. 실험에 사용한 시스템 구성요소들의 설정 변수는 표 1과 같다. 명령어 캐쉬와 데이터 캐쉬는 하나씩 변경해 가며 실험하기 때문에 변경하지 않는 캐쉬는 기본 설정 값인 16KB 4-way 세트연관사상을 사용하였다.

2. 실험 결과

벤치마크별로 다른 성능과 온도 특성을 보이기 때문에 효과적인 분석을 위해 본 논문에서는 4개의 벤치마크 결과의 평균값을 이용하였다.

2.1 16KB 캐쉬에 대한 성능 및 온도 분석

그림 4는 16KB 크기를 가지는 캐쉬의 구성 변화에 따른 성능과 온도 그래프이다.

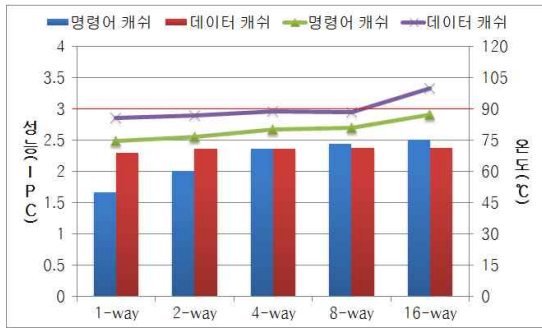


그림 4. 캐쉬 구성 변화에 따른 성능 및 최고온도(16KB)
Figure 4. Performance and peak temperature according to cache organization(16KB)

명령어 캐쉬는 웨이 수가 증가함에 따라 성능이 꾸준히 향상되었다. 직접사상의 성능은 1.66(IPC)로 기본구조인 4-way 세트연관사상에 비해 29.46%의 성능이 감소하였다. 또한 가장 높은 성능을 보이는 16-way 세트연관사상의 성능은 2.49(IPC)로 기본구조에 비해 5.65%의 성능이 향상되었다. 온도의 경우, 직접사상은 최고온도가 74.56°C로 기본구조보다 5.64°C 감소하였다. 16-way 세트연관사상에서는 87.37°C로 기본구조보다 7.16°C의 온도가 증가하였다. 데이터 캐쉬의 성능을 보면 직접사상은 기본구조에 비해 3.01%만의 성능 감소가 있고 16-way 세트연관사상에서는 기본구조보다 0.14% 성능이 향상되었다. 온도의 경우 직접사상은 85.87°C로 기본 구조보다 2.86°C 감소하고 16-way 세트연관사상에서는 99.89°C로 11.17°C 증가하였다.

16KB 크기를 가지는 명령어 캐쉬를 3차원 쿼드코어 프로세서에 적용할 경우 성능은 웨이 수 증가에 따라 향상됨을 알 수 있다. 따라서 많은 웨이를 사용할수록 좋은 성능을 얻을 수 있다. 또한, 온도는 가장 성능이 높고 복잡한 16-way 세트연관사상에서도 90°C이하의 온도를 보이기 때문에, 3차원 쿼드코어 프로세서에서는 가장 높은 성능을 보이는 16-way 세트연관사상을 가지는 명령어 캐쉬를 사용하는 것이 최적임을 알 수 있다. 반면에, 데이터 캐쉬의 성능은 웨이 수 변화에 따라 성능 변화가 적으므로 온도를 우선적으로 고려하여야 한다. 온도는 8-way 세트연관사상에서 90°C를 넘지 않기 때문에 온도를 고려한다면 8-way 세트연관사상 이하로 캐쉬 구성을 해야 할 것으로 판단된다.

2.2 32KB 캐쉬에 대한 온도 및 성능 분석

그림 5은 32KB 크기를 가지는 캐쉬의 구성 변화에 따른 성능과 온도 그래프이다.

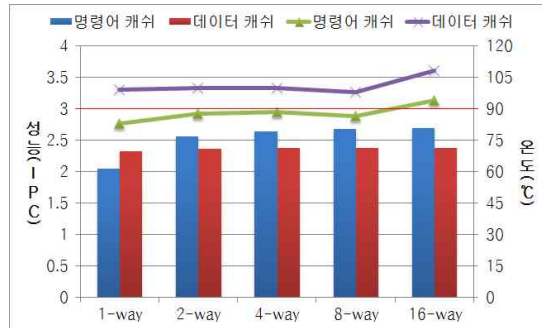


그림 5. 캐쉬 구성 변화에 따른 성능 및 최고온도(32KB)
Figure 5. Performance and peak temperature according to cache organization(32KB)

명령어 캐쉬는 16KB와 마찬가지로 웨이 수가 많아짐에 따라 성능이 향상되었다. 직접사상은 2.04(IPC)로 기본구조보다 13.61% 성능이 감소하여 16KB보다 성능 감소 폭이 줄었고 16-way 세트연관사상에서는 2.69(IPC)로 기본구조보다 13.77% 증가하여 16KB보다 성능 증가 폭이 증가하였다. 온도는 직접사상에서는 83.04°C로 기본구조보다 2.83°C 감소하고 16-way 세트연관사상에서는 94.11°C로 기본구조보다 13.91°C도 증가하였다. 데이터 캐쉬의 성능은 직접사상에서는 기본구조보다 1.76% 감소하고 16-way 세트연관사상에서는 기본구조보다 0.27% 증가하였다. 온도의 경우 직접사상은 98.89°C로 기본구조보다 10.17°C 증가하고 16-way 세트연관사상은 108.29°C로 기본구조보다 19.57°C 증가하였다.

32KB 크기를 가지는 명령어 캐쉬를 3차원 쿼드코어 프로세서에 적용하는 경우, 성능은 16KB와 마찬가지로 웨이 수 증가에 따라 향상되었다. 온도의 경우 가장 높은 성능을 보이는 16-way 세트연관사상에서는 90°C가 넘어 온도 문제가 발생할 수 있으므로 8-way 세트연관사상 이하의 웨이를 가진 구조를 구성하는 것이 적절할 것으로 판단된다. 데이터 캐쉬의 성능은 16KB와 마찬가지로 웨이 수 변화에 따른 성능 변화가 크지 않으므로 온도를 우선적으로 고려해야 한다. 하지만 온도가 모든 사상방식에서 90°C를 넘기 때문에 32KB 크기의 데이터 캐쉬에서는 온도 문제를 해결하지 않을 경우 3차원 쿼드코어 프로세서에 그대로 적용하기는 힘들 것임을 알 수 있다.

2.3 저온도 캐쉬 구성 방식

기존의 실험결과를 보면 가장 복잡한 캐쉬인 16-way 세트연관사상에서 대부분의 경우(명령어 캐쉬(32KB), 데이터 캐쉬(16KB, 32KB)) 임계값 온도인 90°C를 넘는 것을 알

수 있다. 본 논문에서는 기존 실험결과를 기반으로 발열문제를 해결하기 위해 저온도 캐쉬 구성 방식을 제안하였다. 자세한 분석을 위해 4개의 벤치마크의 평균값이 아닌 각 벤치마크 (정수 2, 실수 2) crafty, equake, mcf, mesa를 사용하여 3차원 쿼드코어 프로세서에 16-way 세트연관사상을 적용시킬 때 어떠한 온도 분포를 보이는지 16KB와 32KB로 캐쉬 크기를 설정하여 실험을 수행한 결과는 그림 6에서 알 수 있다. 그림 6를 보면 16KB, 32KB 크기의 캐쉬는 대부분의 벤치마크에서 칩에 문제가 발생할 수 있는 온도인 90°C를 넘는 것을 볼 수 있다. 따라서 3차원 쿼드코어 프로세서에서는 16-way 세트연관사상보다 많은 웨이를 가지는 캐쉬 구성을 사용할 경우 온도 문제가 발생할 수 있음을 알 수 있다.

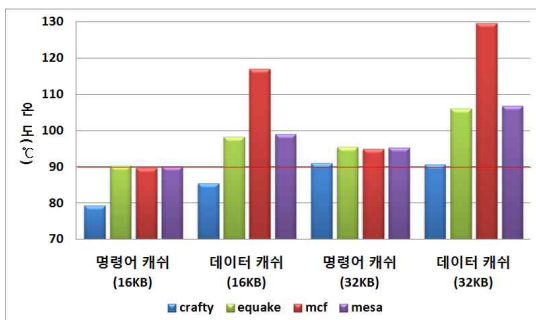


그림 6. 16-way 세트연관사상의 최고온도
Figure 6. Peak temperature of 16-way set-associative

본 논문에서는 앞서 기술한 바와 같이, 직접사상부터 16-way 세트연관사상까지 캐쉬 구성을 변화하여 3차원 쿼드코어 프로세서에 적용해 보았다. 하지만 16-way 세트연관사상에서 높은 발열로 인한 온도문제 발생이 예상되므로 16-way 세트연관사상에서 발생하는 높은 온도 문제를 해결하기 위해 16-way 세트연관사상만을 적층하는 3차원 쿼드코어 프로세서 대신 낮은 기본 전력 소모로 인해 온도가 낮을 것으로 예상되는 8-way 세트연관사상을 16-way 세트연관사상과 레이어별로 혼합하여 사용함으로써 온도 문제를 해결해보고자 한다. 표 2는 8-way 세트연관사상과 16-way 세트연관사상을 혼합한 3차원 쿼드코어 프로세서의 캐쉬 구성변화를 나타내고 있다. 먼저 히트싱크에서 가장 멀리있는 레이어 4에서 온도 문제가 발생할 확률이 가장 높으므로, 레이어 4에 낮은 온도가 예상되는 8-way 세트연관사상을 적용하고 나머지 레이어들에는 16-way 세트연관사상을 적용하였다. 이러한 방식을 Level-1이라고 지정하였다. 또한, Level-1에 8-way 세트연관사상을 온도 문제가 발생하기 쉬운 상위 레

이어에 하나씩 추가하여 사용하는 구성들은 8-way 세트연관사상의 개수에 따라 Level-2, Level-3로 정의하였다. 이처럼 8-way 세트연관사상과 16-way 세트연관사상을 혼합하여 실험한 이유는 16-way 세트연관사상부터는 급격한 온도 상승이 발생하여 3차원 쿼드코어 프로세서에서 발열에 의해 문제가 발생하는 온도인 90°C를 초과할 가능성이 높기 때문에 상대적으로 낮은 온도를 갖는 8-way 세트연관사상을 16-way 세트연관사상과 혼합하여 사용함으로써 온도문제를 해결함과 동시에 웨이 감소로 인한 시스템의 성능 감소를 최소화할 수 있기 때문이다.

표 2. Level에 따른 각 레이어의 캐쉬 구성
Table 2. Cache organization of each layer according to Level

	Level-1	Level-2	Level-3
레이어4	8-Way	8-Way	8-Way
레이어3	16-Way	8-Way	8-Way
레이어2	16-Way	16-Way	8-Way
레이어1	16-Way	16-Way	16-Way

2.4 온도 최적화에 따른 캐쉬 성능 분석

3차원 쿼드코어 프로세서에 캐쉬 구성에 변화를 줄 경우 웨이 수가 증가하게 되면 복잡도의 증가에 따라 온도 문제가 발생할 확률이 높아진다. 온도를 고려하여 직접사상과 같이 적은 웨이를 가지는 캐쉬를 구성하게 되면 캐쉬의 적중률 감소로 시스템의 성능이 저하된다. 따라서 온도 문제가 발생하지 않는 한도 내에서 캐쉬의 웨이 수를 증가시켜 성능을 최적화시켜야 한다. 그림 7과 그림 8은 앞에서 설명한 16-way 세트연관사상과 8-way 세트연관사상의 혼합에 대한 16KB, 32KB 크기 캐쉬의 성능 및 온도를 보인다.

그림 7은 16KB 크기의 캐쉬에 대한 최적화 실험의 결과이다. 명령어 캐쉬는 모든 경우에서 90°C를 넘지 않으므로 가장 높은 성능을 보이는 16-way 세트연관사상을 사용하는 것이 가장 효과적이다. 데이터 캐쉬는 16-way 세트연관사상에서 90°C가 넘게 되므로 16-way 세트연관사상과 8-way 세트연관사상을 혼합하여 온도를 감소시켜야 한다. 실험결과 Level-3에서 90°C를 넘지 않고 최적의 성능(2.36(IPC))과 온도(89.64°C)를 보였다. 그림 8은 32KB 크기의 캐쉬에 대한 최적화 실험의 결과이다. 명령어 캐쉬는 캐쉬 크기가 증가함에 따라 전체적으로 온도가 증가하게 되는데 가장 높은 성능을 보이는 16-way 세트연관사상에서는 90°C가 넘는다. 따

라서 8-way 세트연관사상과 혼합한 결과 Level-2에서 가장 최적화된 성능(2.68(IPC))과 온도(88.80°C)를 보였다. 데이터 캐쉬는 캐쉬 크기가 32KB가 되면 16KB와 비교하여 전체적으로 온도가 상승하여 모든 경우에서 90°C가 넘어 온도 문제가 발생하였다. 따라서 데이터 캐쉬는 3차원 쿼드코어 프로세서에서 온도 문제가 발생하기 때문에 캐쉬의 구성변화가 아닌 다른 온도 관리 기법이 필요함을 알 수 있다.

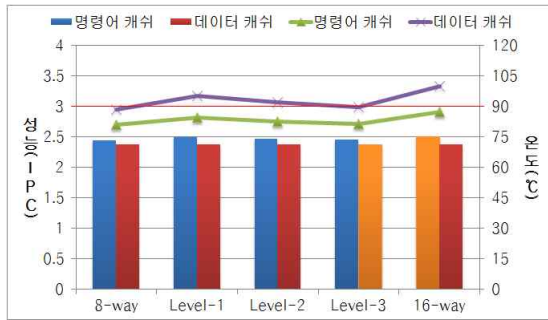


그림 7. 온도 최적화에 따른 캐쉬 성능(16KB)
Figure 7. Cache performance according to temperature optimization(16KB)

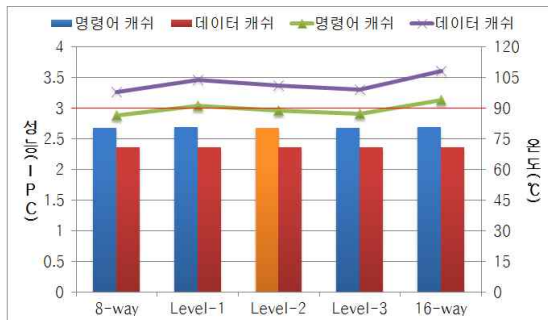


그림 8. 온도 최적화에 따른 캐쉬 성능(32KB)
Figure 8. Cache performance according to temperature optimization(32KB)

V. 결론

본 논문에서는 3차원 쿼드코어 프로세서를 대상으로 캐쉬 구성 변화에 따른 프로세서의 온도와 성능을 상세하게 분석하였다. 또한, 3차원 구조 프로세서의 가장 큰 문제점인 온도를 고려하는 최적의 캐쉬 구성을 찾기 위하여 실험 및 분석을 수행하였다. 실험 결과에 따르면, 일반적인 캐쉬의 특성에 따라 웨이 수가 증가하면 성능은 향상되고 온도 또한 웨이 수에 비

례하여 증가하는 것을 확인할 수 있었다. 다만, 성능의 경우 명령어 캐쉬는 크기에 따른 성능 향상 폭이 큰 반면, 데이터 캐쉬의 성능 변화는 크지 않았다. 온도는 전체적으로 명령어 캐쉬가 낮게 발생하였고 데이터 캐쉬에서는 명령어 캐쉬에 비해 훨씬 높은 온도가 발생하였다. 그러므로 3차원 쿼드코어 프로세서의 캐쉬를 설계하고자 한다면, 명령어 캐쉬는 신뢰성을 확보할 수 있는 온도를 넘지 않는 범위 내에서 최대한의 웨이를 사용하여 성능을 높이고, 데이터 캐쉬는 캐쉬 구성에 따른 성능 변화가 작기 때문에 적은 웨이 수를 사용하여 온도 문제를 우선적으로 해결하는 구조가 가장 효율적인 구조라고 판단된다. 본 논문에서 분석한 캐쉬 구성에 따른 특성을 잘 활용한다면 3차원 마이크로프로세서를 더욱 효율적으로 설계할 수 있을 것이라 기대된다. 본 논문에서는 캐쉬 구성에 따라 데이터 요구량이 급증할 경우 발생하는 병목현상에 대해 고려되지 않았다. 따라서 향후 연구로써 데이터 요구량으로 인한 병목현상을 해결하기 위한 캐쉬 구성에 대한 연구를 진행하고자 한다.

참고문헌

- [1] R. Kalla, S. Balaran, and J. M. Tendler, "IBM Power5 chip: a dual-core multithreaded processor," *IEEE MICRO*, Vol. 24, Issue 2, pp. 40-47, Mar-Apr. 2004.
- [2] P. Kongetira, K. Aingaran, and K. Olukotun, "Niagara: a 32-Way multithreaded Sparc processor," *IEEE MICRO*, Vol. 25, Issue 2, pp. 21-29, Mar-Apr. 2005.
- [3] Y. J. Kwon, C. D. Kim, S. R. Maeng, and J. H. Huh, "Virtualizing performance asymmetric multi-core systems," In *Proceedings of 33th International Symposium on Computer Architecture*, pp. 45-56, San Jose, USA, June. 2011.
- [4] M. DeVuyst, A. Venkat, and D. M. Tullsen, "Execution migration in a heterogeneous-ISA chip multiprocessor," In *Proceedings of 17th International Conference on Architectural*

- Support for Programming Languages and Operating Systems, pp. 261-272, London, UK, Mar. 2012.
- [5] C. Zhu, Z. Gu, L. Shang, R. P. Dick, R. Joseph, "Three-dimensional chip-multiprocessor run-time thermal management," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 27, No. 8, pp. 1479-1492, Aug, 2008.
- [6] 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, SanFrancisco, USA, Sep. 2009.
- [7] A. W. Topol, D. C. L. Tulipe, L. Shi, D. J. Frank, K. Bernstein, S. E. Steen, A. Kumar, G. U. Singco, A. M. Young, K. W. Guarini, and M. Jeong, "Three-dimensional integrated circuits," *IBM Journal of Research and Development*, Vol. 50, Issue. 4-5, pp. 491-506, Jul. 2006.
- [8] B. Black, M. Annavaram, N. Brekelbaum, J. DeVale, L. Jiang, G. H. Loh, D. McCaule, P. Morrow, D. W. Nelson, D. Pantuso, P. Reed, J. Rupley, S. Shankar, J. Shen, and C. Webb, "Die Stacking (3D) Microarchitecture," In *Proceedings of 39th International Symposium on Microarchitecture*, pp. 469-479, Orlando, USA, Dec. 2006.
- [9] D. O. Son, J. W. Ahn, J. H. Park, J. M. Kim, and C. H. Kim, "Analysis on the Temperature of 3D Multi-core Processors according to Vertical Placement of Core and L2 Cache," *Journal of The Korea Society of Computer and Information*, Vol. 16, No. 6, pp. 1-10, June. 2011.
- [10] D. H. Woo, N. H. Seong, D. L. Lewis, and H. H. S. Lee, "An optimized 3D-stacked memory architecture by exploiting excessive, high-density TSV bandwidth," In *Proceedings of 16th International Conference on High-Performance Computer Architecture*, pp. 1-12, Bangalore, India, Jan. 2010.
- [11] H. J. Choi, J. W. Ahn, H. B. Jang, J. M. Kim and C. H. Kim, "Thermal Pattern Comparison between 2D Multicore Processors and 3D Multicore Processors," *Journal of The Korea Society of Computer and Information*, Vol. 16, No. 9, pp. 1-10, Sep. 2011.
- [12] A. K. Coskun, A. B. Kahng, and T. S. Rosing, "Temperature- and Cost-Aware Design of 3D Multiprocessor Architectures," In *Proceedings of 12th Euromicro conference on Digital System Design, Architectures, Methods and Tools*, pp. 183-190, Patras, Greece, Aug. 2009.
- [13] E. Wong, and S. K. Lim, "3D Floorplanning with Thermal Vias," In *Proceedings of the conference on Design, Automation and Test in Europe*, pp. 878-883, Munich, Germany, Mar. 2006.
- [14] A. K. Coskun, J. L. Ayala, D. Atienza, T. S. Rosing, and Y. Leblebici, "Dynamic Thermal Management in 3D Multicore Architectures," In *Proceedings of Design, Automation & Test in Europe Conference & Exhibition*, pp. 1410-1415, Nice, France, Apr. 2009.
- [15] H. Homayoun, M. Rahmatian, V. Kontorinis, S. Golshan, and D. M. Tullsen, "Hot Peripheral Thermal Management to Mitigate Cache Temperature Variation," In *Proceedings of International Symposium on Quality Electronic Design*, pp. 755-763, Santa Clara, USA, Mar. 2012.
- [16] M. Paul, and P. Petrov, "I-Cache Configurability for Temperature Reduction through Replicated Cache Partitioning," In *Proceedings of Symposium on Application Specific Processors*, pp. 81-86, Anaheim, USA, June. 2010.
- [17] M. A. Z. Alves, H. C. Freitas, and P. O. A. Navaux, "Investigation of Shared L2 Cache on Many-Core Processors," In *Proceedings of Architecture of Computing Systems*, pp. 1-10, Delft, Netherlands, Mar. 2009.
- [18] A. Fog, "The microarchitecture of Intel, AMD and VIA CPUs : An optimization guide for assembly programmers and compiler makers," Copenhagen University College of Engineering,

Feb. 2012.

[19] J. W. Joyner, P. Zarkesh-Ha, and J. D. Meindl, "A Stochastic Global Net-length Distribution for a Three-Dimensional System on Chip (3D-SoC)," In Proceedings of the 14th IEEE International ASIC/SOC Conference, pp. 147-151, Arlington, USA, Sep. 2001.

[20] K. Puttaswamy, G. H. Loh, "Thermal Analysis of a 3D Die Stacked High Performance Microprocessor," In Proceedings of ACM Great Lakes Symposium on VLSI, pp. 19-24, Philadelphia, USA, Apr-May. 2006.

[21] W.-L. Hung, G. M. Link, X. Yuan, N. Vijaykrishnan, and M. J. Irwin, "Interconnect and Thermal-aware Floorplanning for 3D Microprocessors," In Proceedings of the 7th ISQED, pp. 98-104, San Jose, USA, March. 2006.

[22] K. Skadron, M. Stan, K. Sankaranarayana, W. Huang, S. Velusamy, and D. Tarjan, "Temperature-Aware Microarchitecture: Modeling and Implementation," ACM Transactions on Architecture and Code Optimization, Vol. 1, No. 1, pp. 94-125, Mar. 2004.

[23] D. C. Burger, and T. M. Austin, "The SimpleScalar tool set, version 2.0," ACM SIGARCH CAN, Vol. 25, No. 3, pp. 13-25, Jun. 1997.

[24] N. Muralimanohar, R. Balasubramonian, and N. P. Jouppi. "Optimizing NUCA Organizations and Wiring Alternatives for Large Caches With CACTI 6.0," In Proceedings of 40th International Symposium on Microarchitecture, pp. 3-14, Chicago, USA, Dec. 2007.

[25] D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: a framework for architectural-level power analysis and optimizations," in Proceedings of the 27th International Symposium on Computer Architecture, pp. 83-94, Vancouver, Canada, Jun. 2000.

[26] 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, Vol 14, No 5, pp. 501-513, May. 2006.

[27] R. E. Kessler, E. J. McLellan, and D. A. Webb, "The Alpha 21264 Microprocessor Architecture," In Proceedings of the ICCD '98, pp. 90-95, Austin, USA, Aug. 2002.

[28] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan, "Temperature-Aware Microarchitecture," In Proceedings of the 30th annual international symposium on Computer Architecture, pp. 2-13, San Diego, USA, June. 2003.

[29] SPEC CPU2000 Benchmarks, available at <http://www.specbench.org>

저 자 소개



손 동 오
 2010: 전남대학교 전자컴퓨터공학부 공
 학사
 2010-현재: 전남대학교 전자컴퓨터공학
 과 석사과정
 관심분야: 컴퓨터구조, 임베디드시스템
 Email : sdo1127@gmail.com



안 진 우
 2010: 전남대학교 전자컴퓨터공학부 공
 학사
 2012: 전남대학교 전자컴퓨터공학과
 석사
 관심분야: 컴퓨터구조, 임베디드 하드웨
 어 설계
 Email : ajw0411@gmail.com



최 흥 준
 2009: 전남대학교 전자컴퓨터공학부 공
 학사
 2011 : 전남대학교 전자컴퓨터공학과
 석사
 2011-현재 : 전남대학교 전자컴퓨터공
 학과 박사
 관심분야: 저전력 설계, 컴퓨터 구조
 Email : chj6083@gmail.com



김 종 면

1995: 명지대학교 전기공학사
2000: University of Florida ECE 석사
2005: Georgia Institute of
Technology ECE 박사
2005 - 2007: 삼성종합기술원 전임연구원
2007 - 현재: 울산대학교 컴퓨터정보통신공학부 교수
관심분야 : 임베디드 SoC, 컴퓨터구조,
프로세서 설계, 병렬처리
Email : jnkim07@ulsan.ac.kr



김 철 홍

1998: 서울대학교 컴퓨터공학사
2000: 서울대학교 대학원 컴퓨터공학부
석사
2006: 서울대학교 대학원 전기컴퓨터공학
부 박사
2005 - 2007: 삼성전자 반도체총괄
SYS.LSI사업부 책임 연
구원
2007 - 현재: 전남대학교 전자컴퓨터공학
부 교수
관심분야 : 임베디드시스템, 컴퓨터구조,
SoC 설계, 저전력 설계
Email : chkim22@chonnam.ac.kr

