

# 응용분야에 따른 직접사상캐시의 성능향상과 전력소모 절감을 위한 재구성 가능한 백업캐시 제안

최병창<sup>o</sup> 서효중  
가톨릭대학교 컴퓨터공학과  
{hellocbc<sup>o</sup>, hjsuh}@catholic.ac.kr

## An Application-Specific Configurable Backup Cache for High Performance and Power Reduction in Direct-Mapped Cache

Byeong-Chang Choi<sup>o</sup>, Hyo-Joong Suh  
Dept of Computer Science and Engineering, The Catholic University of Korea

### 요 약

반도체 공정의 발달로 인해 하나의 칩에 많은 양의 소자를 넣는 것이 가능해지면서 상대적으로 넓어진 공간에서 캐시 메모리가 차지하는 공간의 비중이 증가하고 있다. 상대적으로 비중이 커진 캐시 메모리는 CPU가 소모하는 전력의 50%에 상당하는 전력을 소모하는 등 시스템의 성능뿐만 아니라 전력 소모에도 큰 영향을 주고 있다. 현재 시스템 성능 향상과 전력 소모 절감을 위하여 캐시 메모리의 논리적 구조를 개선하기 위한 많은 연구가 진행 중이다. 본 논문에서는 다양한 용도로 사용되는 범용 시스템이 아닌 특정 응용분야에 최적화되어 사용되는 소규모 임베디드 시스템에 적합한 직접사상캐시를 위한 재구성 가능한 백업캐시를 제안하려고 한다. 제안하는 백업 캐시는 특정 레지스터 값을 이용해서 백업캐시를 재구성 가능하게 하여 응용분야에 따라 직접사상캐시의 성능 향상과 전력소모를 절감하도록 하여 시스템의 성능 향상과 전력소모를 절감시키는 역할을 할 것이다.

### 1. 서 론

반도체 공정의 발달로 칩 하나에 넣을 수 있는 소자의 양이 증가하면서 칩의 면적이 상대적으로 증가하였다. 칩의 면적이 상대적으로 증가하면서 캐시 메모리의 크기 도 함께 증가하여 캐시 메모리가 시스템의 성능에 큰 영향을 주고 있으며, 칩에서 소모되는 에너지의 50% 가량을 소모하는 등 전력소모에도 큰 영향을 주고 있다. 현재 많은 시스템 디자이너들이 시스템의 성능 및 에너지 소모의 관점에서 캐시 메모리의 중요성을 인식하고 연구를 진행하고 있다. 본 논문에서는 특정 응용분야에 최적화 되어 사용되는 임베디드 시스템에 적합한, 직접사상 캐시의 성능향상과 전력소모절감을 위한 재구성 가능한 백업캐시를 제안하려고 한다. 캐시 메모리는 대표적으로 직접사상캐시와 완전연관캐시 그리고 집합연관캐시 등이 있는데, 이중 직접사상캐시를 선택한 이유는 나머지 두 캐시 메모리 구조에 비해 상대적으로 빠른 접근 속도와 적은 전력소모 때문이다. 그러나 직접사상캐시의 경우 완전연관캐시와 집합연관캐시에 비해 특정 프로그램에서 급격하게 적응률이 떨어지는 단점이 있다. 직접사상캐시의 시간 참조 지역성과 공간 참조 지역성을 보완하기 위하여 Victim캐시와 스트림 버퍼가 제안되었다. 그러나 Victim캐시와 스트림 버퍼는 직접사상캐시에 적합 프로그램에서는 그 존재 자체가 추가적인 정적 에너지를 소

모하는 시스템의 오버헤드로 작용하게 된다. 따라서 본 논문에서는 재구성 가능한 백업캐시 제안하고, 레지스터 값을 통해서 백업캐시를 재구성하여 Victim 캐시와 스트림 버퍼의 사용여부를 결정해서 응용분야에 따라 최적화된 캐시 메모리의 구조를 사용할 수 있도록 하였다.

### 2. 관련 연구

캐시 메모리의 전력소모를 절감하거나 성능을 향상시키기 위한 여러 가지 연구가 진행되고 있다. [1]에서는 작은 회로를 추가하여 레지스터 값을 이용해 캐시 메모리를 직접사상캐시나 집합연관캐시로 재구성 가능하게 하여 응용프로그램에 따라 적합한 캐시 메모리 구조를 제공할 수 있는 기법을 제안하였고, [2]에서는 사용하지 않은 캐시 메모리를 Gated-V<sub>dd</sub>를 이용하여 전력 공급을 끊어서 정적 에너지 소모를 절감할 수 있는 기법을 제안하였다. 전력 공급 여부를 동적으로 결정할 수 있어 정적 에너지 소모를 줄일 수 있으나, 캐시 메모리의 상태를 전환(활성상태/휴면상태)할 때 시스템에 많은 보상을 주게 된다. [3]에서는 캐시 메모리를 집합연관캐시로 구성하는데, 캐시 집합 중 하나는 정적에너지 소모는 상대적으로 크지만 속도가 빠른 낮은 한계전압을 필요로 하는 캐시 집합으로 구성하고 나머지는 정적에너지 소모는 상대적으로 작지만 속도가 느린 높은 한계전압을 필요로 하는 캐시 집합으로 구성하였다. 많이 사용되는 데이터

(Critical-Path)를 낮은 한계전압을 필요로 하는 캐시 집합에 위치시키기 위해 한번 참조된 데이터는 캐시 집합 사이의 교환(Way-Exchange)을 통해 낮은 한계전압을 필요로 하는 캐시 집합으로 옮긴다. 하지만 캐시 집합 사이의 교환이 자주 발생할 경우 시스템에 큰 부하를 주게 되는 단점이 있다. [4]에서는 직접사상캐시를 보완하기 위하여 시간 참조 지역성을 보완하기 위한 Victim캐시와 공간 참조 지역성을 보완하기 위한 스트림 버퍼를 제안하였다. Victim캐시는 L1 캐시에서 나중 쓰기(Write-Back) 되는 데이터를 Victim캐시에 기록하였다가 데이터를 필요로 할 때 빠르게 읽어올 수 있도록 하여 충돌 실패(Conflict-Miss)를 줄여주며, 스트림 버퍼는 프로그램이 연속적으로 실행되며 메모리에 연속적으로 저장되는 경우가 많기 때문에 외부 메인 메모리로부터 데이터를 읽어올 때 연속된 데이터를 한 번에 읽어와서 버퍼에 기록하였다가 필요로 하는 데이터가 있을 경우 빠르게 읽어올 수 있도록 해서 필수 실패(Compulsory-Miss)가 발생하는 것을 줄여준다. 그러나 Victim캐시나 스트림 버퍼가 필요하지 않는 경우에는 그 존재자체가 정적 에너지를 소모하게 되어 시스템에 부하를 주게 된다. 본 논문에서는 Victim캐시와 스트림 버퍼를 동적으로 선택하여 사용할 수 있는 구조를 제안한다.

### 3. 백업캐시의 구조

#### 3.1 캐시 구조 및 동작

그림 1은 백업캐시를 포함하는 캐시 메모리의 논리적인 구조이다. 우선 참조 데이터를 직접사상캐시에서 찾고 성공하면 바로 값을 리턴하고, 실패하면 백업캐시에서 다시 검색을 한다. 백업캐시는 그림 2와 같이 구성하였고, 크기가 작은 완전연관캐시로 이루어져 있다. 직접사상캐시의 경우 한계전압이 낮게 구성되어 정적 에너지는 커지지만 접근 속도를 높게 하고, 백업캐시의 경우 한계 전압을 높게 구성하여 접근 속도는 조금 늦지만 정적 에너지 소모가 작게 구성되어 있다.

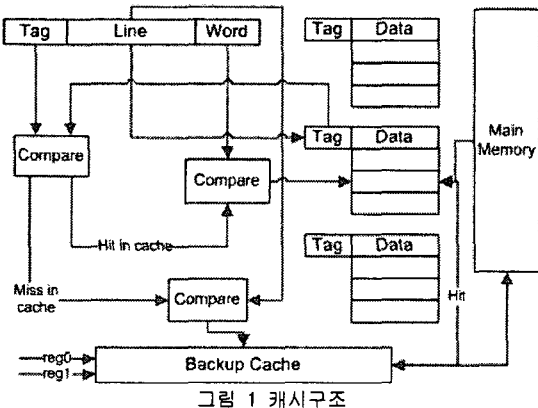


그림 1 캐시구조

백업캐시는 레지스터 값에 따라 재구성 가능한 구조로 reg=0, reg=0 일 경우 Victim캐시와 스트림 버퍼의 전력

공급을 차단하여 꺼버리고, reg0=0, reg=1 일 경우 스트림 버퍼를 reg0=1, reg1=0 일 경우 Victim캐시를 사용할 수 있는 상태로 만든다. reg0=1, reg=1 일 경우 Victim 캐시와 스트림 버퍼 모두 사용가능하다. 백업캐시에서 사용하는 Victim캐시는 기존의 Victim캐시와는 달리 적중 시 해당 데이터를 캐시로 옮기지 않고 계속 Victim캐시에 보관하여 데이터의 중복성을 줄일 수 있었다.

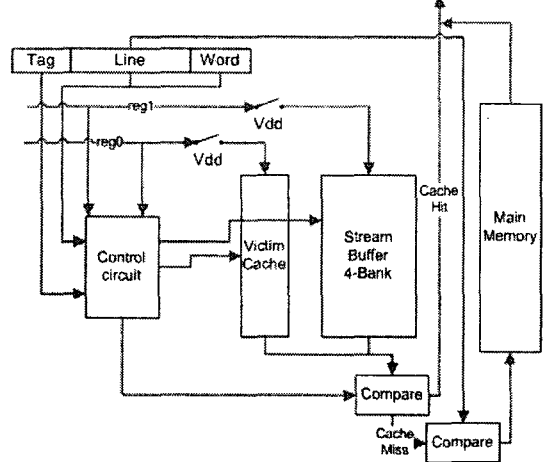


그림 2 재구성 가능한 백업캐시

### 4. 성능분석

#### 4.1 에너지 모델

캐시 메모리에서 소모되는 에너지를 측정하기 위해서 아래와 같은 에너지 모델을 사용한다[?].

$$E_{total} = E_{dynamic} + E_{static}$$

$$E_{dynamic} = C_{hit} \times E_{hit} + C_{miss} \times E_{miss}$$

$$E_{static} = C_{cycle} \times E_{static\ per\ cycle}$$

$$E_{miss} = E_{uP\ stall} + E_{cache\ block\ fill} + E_{offchip\ access}$$

에너지 소모는 동적 에너지 소모와 정적 에너지 소모의 합으로 구할 수 있다. 동적 에너지는 캐시 메모리에 적중했을 때 캐시 셀에 전하를 채우기 위해 사용되는 에너지와 적중 실패 했을 때의 에너지의 합으로 구할 수 있다. 적중 실패 했을 때 해당 데이터를 메모리로부터 읽어오는 동작을 하게 된다. 이때 읽어 오는 시간동안 CPU의 동작이 중지되면서 소모되는 에너지와 읽어오 메모리를 캐시 메모리에 기록하는데 사용되는 에너지와 외부의 메인 메모리에 접근하기 위해 소모 되는 에너지가 소모 되는데, 이 세 가지 에너지의 합으로 캐시 적중 실패시의 에너지를 구할 수 있다. 정적 에너지는 프로그램을 실행하기 위해 수행되는 총 사이클 수와 한 사이클에 소모되는 정적 에너지 소모량으로 구할 수 있다. 이때 일반적으로 캐시 적중 실패 시 소모되는 에너지는 캐시 적중 시 소모되는 에너지의 50배에서 크게는 200배까지

소모가 되는 것으로 알려져 있으며, 미세공정이 발달할 수록 그 차는 더 커질 것으로 예상된다. 정적 에너지의

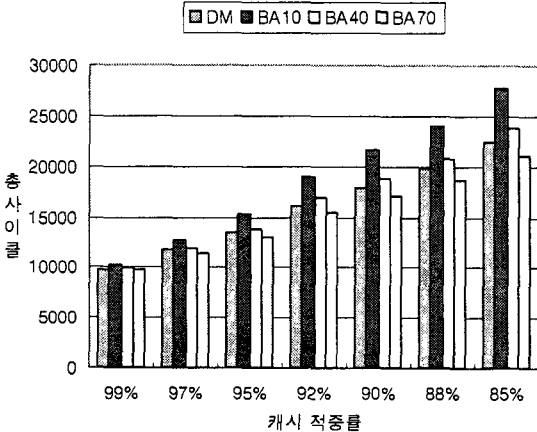


그림 3 총 수행 사이클

경우 캐시 메모리에서 총 소모되는 에너지의 30%에서 크게는 50%까지 차지하는 것으로 알려져 있으며 이것 역시 미세공정이 발달할수록 차지하는 비율이 커질 것으로 예상된다. 본 논문에서는 성능 평가를 위하여 다음과 같은 가정을 하였다. 캐시에서 적중하였을 때 1사이클로 수행되며 상대적인 동적 에너지 1이 소모된다. 외부 메모리에 접근하기 위해서는 10사이클이 소모되고 에너지는 캐시에 접근하는 동적 에너지의 50배 정도의 에너지가 소모된다. 백업 캐시에 접근하기 위해서는 5사이클이 소모되며 완전연관캐시에서의 추가적인 태그 비교로 인해 적중 성공 시에 비해 15배 정도가 소모되는 것으로 가정하였다.

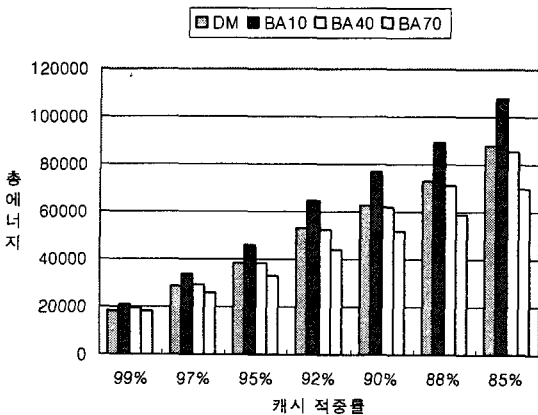


그림 4 총 소모 에너지

#### 4.2 성능 평가

성능 평가는 에너지 모델과 앞 절에서 서술한 전제 조건을 기본으로 실시하였다. 그림 3에서 캐시 적중률은

직접사상캐시에서의 캐시 적중률을 의미하며 DM은 reg0=0, reg1=0인 상태로 직접사상캐시만 사용한 경우를, BAXX는 reg0=1, reg1=1인 상태의 백업캐시에서 백업캐시에서의 캐시 적중률이 XX%인 경우를 의미한다. 수행 사이클을 살펴보면 백업캐시의 적중률이 70% 이상인 경우 직접사상캐시보다 우수한 성능을 나타내고 있다. 반면 에너지 그림 4에서와 같이 백업캐시에서의 적중률이 40%인 경우 직접사상캐시만 사용한 경우보다 수행 사이클은 길지만, 직접사상캐시의 적중률이 92%보다 낮아지면 직접사상캐시만 사용한 경우보다 에너지는 적게 소비되는 것을 알 수 있다. 따라서 응용분야에 따라 직접사상캐시의 적중률이 일정 수준이하로 낮아지는 경우, 백업캐시에서의 적중률이 일정 수준 이상 된다면 백업캐시를 사용하여 시스템의 성능향상 효과와 에너지 소모 절감 효과를 얻을 수 있다는 것을 알 수 있었다.

#### 5. 결론 및 차후 연구 과제

한계전압이 낮은 직접사상캐시를 사용하여 빠른 속도로 캐시 메모리에 접근 가능하도록 하였고, 한계전압이 높은 백업캐시를 사용하여 정적 에너지 소모를 줄였다. 그리고 응용분야의 특성에 따른 백업캐시 재구성으로 캐시 메모리 구조를 최적화하여 시스템의 성능 향상과 에너지 소모 절감 효과를 얻을 수 있었다. 차후 연구 과제로 각종 응용분야를 그 특성에 따라 분류하고 분야별로 적합한 캐시 메모리의 구조를 찾아서 제안하는 것과 백업캐시의 동작 개선 및 구조 단순화에 대한 연구를 진행 중이다.

#### 6. 참고문헌

- [1] Chuanjun Zhang, Frank Vahid, Walid Najjar, A Highly Configurable Cache Architecture for Embedded Systems, ACM SIGARCH Computer Architecture News, Proceedings of the 30th annual international symposium on Computer architecture ISCA '03, May 2003
- [2] Akihito Sakanaka, Seiichirou Fujii, Toshinori Sato, A Leakage-Energy-Reduction Technique for Highly-Associative Caches in Embedded Systems, ACM SIGARCH Computer Architecture News, Volume 32 Issue 3, June 2004
- [3] Michael Powell, Se-Hyun Yang, Babak Falsafi, Kaushik Roy, T. N. Vijaykumar, Gated-V<sub>dd</sub>: A Circuit Technique to Reduce Leakage in Deep-Submicron Cache Memories, Proceedings of the 2000 international symposium on Low power electronics and design, August 2000
- [4] Norman P. Jouppi, Improving Direct-Mapped Cache Performance by the Addition of a Small Fully-Associative Cache and Prefetch Buffers, ACM SIGARCH Computer Architecture News, Proceedings of the 17th annual international symposium on Computer Architecture ISCA '90, May 1990