

저전력 DRAM을 위한 리프레시 주기 설정 기법

김현구^o 신동군

성균관대학교 컴퓨터공학과

zero1201@skku.edu, dongkun@skku.edu

Refresh Period Adjustment Technique for Low-Power DRAM

Hyunkoo Kim^o Dongkun Shin

School of Information and Communication Engineering, Sungkyunkwan University

1. 서 론

컴퓨터 시스템을 사용하는 모든 장치에서 DRAM은 저장장치로서 많이 쓰이고 있다. DRAM은 휘발성 메모리로서 데이터를 저장시킨 후 일정시간이 지나면 데이터가 사라지게 된다. DRAM을 사용하는 동안 데이터를 계속 저장시켜 놓기 위해서는 지속적인 리프레시(refresh)가 필요하다. 그러므로, 사용하지 않을 때에도 리프레시를 위해서 전력을 계속 소모하게 된다. 특히, 휴대폰 같은 모바일 시스템에서 DRAM의 전력 소모는 큰 문제가 되고 있다.

2. 본 론

DRAM의 전력 소모를 줄이기 위해서 시스템 수준에서 두 가지의 기법이 소개되었다. 첫 번째는 DRAM이 제공 하고 있는 여러 가지 전력모드를 이용하는 것이다. 최근의 DRAM은 일반적으로 Powerdown이나 Self Refresh와 같은 저전력 모드를 제공하고 있다. 이러한 상태에서는 저장된 데이터가 사라지지는 않지만 셀을 읽고 쓰기가 가능한 일반적인 상태(active 모드)로 복귀시키기 위해서는 약간의 시간이 필요하게 된다. 그러므로, 셀에 대한 접근이 오랫동안 없을 시에 이러한 상태로 진입하여 전력 소모를 줄이게 된다. Huang[1]은 이러한 DRAM의 저전력 모드를 효과적으로 사용하기 위해서 저전력 모드를 적용할 수 있는 단위인 rank별로 hot 데이터와 cold 데이터를 구분하는 기법을 제시하였다. DRAM을 HOT rank와 COLD rank로 구분하여 자주 접근 되지 않는 데이터들은 COLD rank로 자주 접근 되는 데이터들은 HOT rank로 이동시켜 준다. 그러면 COLD rank에 있는 데이터들의 접근이 적으므로 COLD rank는 오랜 시간 Power-down모드나 Self Refresh 모드로 동작 할 수 있기 때문에 COLD rank에 대한 전력 소모를 줄이게 된다.

두 번째 기법은 DRAM의 리프레시 주기를 조정하는 기법이다. 현재 시중에 나와 있는 DRAM들은 64ms의 리프레시 주기를 가지고 있다. 하지만 실제 DRAM의 각 페이지들은 보통 64ms 보다 긴 데이터 보존 시간을 가지고 있다. 이러한 페이지 상태에 맞춰 리프레시 주기를 조정하여 전력 소모를 줄일 수 있다. Venkatesan[2]는 리프레시 주기를 조정하는 기법으로 DRAM의 전력 소모를 줄이고자 했다. RAPID라 이름 붙여진 이 기법에서는 데이터들을 페이지에 기록할 때 가장 긴 리프레시 시간을 요구하는 페이지부터 데이터를 기록하여 데이터가 기록되지 않은 페이지의 리프레시는 무시하고 전체 리프레시 시간을 결정함으로써 리프레시에 소모되는 전력을 줄이고 있고, DRAM 사용 도중 데이터가 페이지에서 지워졌을 때 그 페이지보다 리프레시 시간이 짧은 페이지의 데이터를 옮겨와 전체적인 리프레시 횟수를 줄임으로서 리프레시에 소모되는 전력을 줄이는 것이다.

본 논문에서 제안하는 LaRPA (Locality -aware Refresh Period Adjustment)에서는 각 페이지가 저장하고 있는 데이터의 상태를 추가로 관리한다. 페이지 데이터의 상태는 최근에 접근 된 데이터와 접근이 오랫동안 되지 않은 데이터로 구분된다. 각 데이터들에 대한 접근 시간을 구하기 위해서는 DRAM 컨트롤러에 공급되는 clock을 이용하여 카운터를 증가시킴으로서 시간을 계산해 낼 수 있다. RAPID에서와는 달리 이 논문에서 제시하는 방법은 오랫동안 접근되지 않은 데이터를 담고 있는 페이지를 식별하여

해당 페이지의 데이터를 플래시 메모리로 밀어내게 된다. 이렇게 함으로서 전체적인 리프레시 주기를 조정가능하고 이로 인해 리프레시 횟수의 감소 및 전력 소모 감소를 시킬 수 있다.

데이터를 플래시 메모리로 밀어냈을 때 리프레시에 의한 전력 소모량의 줄어드는 값에서 그 데이터를 밀어내고 다시 읽어올 때 필요한 추가적인 전력 소모량 값을 뺀 결과가 0보다 큰 값이 되는 값을 알아내고 그 값만큼 접근이 안 될 데이터에 한에서 플래시 메모리로 밀어낸다면 전력 감소에 효과적인 것이다.

하지만 다음 예상 시간을 정확히 알아 낼 수가 없으므로 LaRPA에서는 마지막으로 접근된 시간부터 현재까지 흐른 시간이 다음 접근 시 까지 걸릴 시간으로 예측하고 다음 접근 예상 시간을 기준시간 T로 사용한다.

LaRPA 기법에서 또 한 가지 고려해야 할 사항은 최근에 접근이 되지 않은 데이터를 플래시 메모리로 밀어냄으로서 유발 될 수 있는 DRAM에서의 read miss로 인한 성능 저하이다. 이러한 read miss를 관찰하여 T 값을 수행시간에 조정해야 한다. 즉, read miss가 기준 값보다 많이 발생하면 T를 증가시켜서 cold로 판별되는 페이지수를 감소시키고 반대로 read miss가 기준 값보다 적을 때에는 T를 감소시켜서 cold 페이지의 수를 증가시킨다.

LaRPA 기법은 메모리 컨트롤러가 매핑 테이블을 관리한다는 가정이다. 매핑 테이블은 페이지 단위나 여러 개의 페이지 단위로 관리하면 크기를 줄일 수 있다. 하지만 매핑 테이블을 추가할 수 없거나 매핑 테이블의 저장 공간에 대한 오버헤드가 너무 클 경우엔 페이지를 이동시키지 않는 방법을 적용한다. 이 기법에서는 요청되는 주소 그대로 데이터를 DRAM에 저장시켜주고 페이지의 마지막 접근 이후의 시간 값과 그 데이터가 저장되어있는 페이지의 리프레시 주기에 대한 값을 조합하여 접근이 오랫동안 되지 않고 그 페이지의 리프레시 주기가 짧은 페이지의 데이터를 플래시 메모리로 밀어냄으로서 리프레시 주기를 조정하는 기법이다.

이 기법에서는 데이터들이 페이지 리프레시 시간에 대해 최적화 된 장소에 저장이 되지 않아 전체적인 리프레시 횟수가 많아 질 수 있으나, 데이터 이동에 따른 오버헤드와 이동으로 인한 매핑 상태를 담고 있는 테이블을 위한 저장 공간에 따른 오버헤드를 낮출 수 있다는 장점이 있다.

결과들을 종합해 보면 gcc 같은 무작위적인 주소 요청일 경우 miss 증가율이 10%를 넘지 않는 범위에서 전체 전력은 약 90%감소됨을 볼 수 있었다. 하지만 mcf 같이 특수한 주소 요청 패턴에서는 큰 전력 감소 효과가 적은데 이러한 패턴의 접근이 요청 될 시에는 다음 접근 예상 시간을 길게 잡음으로서 해결 할 수 있다.

3. 결 론

이 논문에서는 DRAM의 리프레시 전력 소모를 감소 시키는 방법을 제시하였다. 각 페이지의 서로 다른 리프레시 주기를 이용하여 유효한 페이지 중 가장 짧은 리프레시 주기로 리프레시 전력을 공급하여 전체 리프레시 횟수를 줄임으로서 전력을 아낄 수 있었다. 리프레시 주기를 조정하지 않는 기법보다 약 90%만큼 리프레시 전력 소모가 낮아짐을 볼 수 있었고 RAPID방식보다도 60%정도 더 낮은 전력 소모를 보여주었다.

이 논문의 실험에서는 DRAM에서 제공하고 있는 각 모드들을 고려하지 않았는데 향후 이것을 추가한 기법을 연구할 예정이다.

참고 문헌

- [1] H. Huang, K. G. Shin, C. Lefurgy, T. Keller. Improving Energy Efficiency by Making DRAM Less Randomly Accessed. The 2005 international symposium on Low power electronics and design, pages: 393-398, 2005.
- [2] R. K. Venkatesan, S. Herr, E. Rotenberg. Retention-Aware Placement in DRAM (RAPID): Software Methods for Quasi-Non-Volatile DRAM. The 12th Annual Symposium on High Performance Computer Architecture, pages 155-165, 2006.