

GPGPU-Sim에 의한 GPGPU기억쓰기특성연구

박 상 룡

경애하는 최고령도자 김정은동지께서는 다음과 같이 말씀하시였다.

《첨단과학기술분야에서 세계적경쟁력을 가진 기술들을 개발하기 위한 투쟁을 힘있게 벌려야 합니다.》(《조선로동당 제7차대회에서 한 중앙위원회사업총화보고》 단행본 39페이지)

GPGPU-Sim3.2는 GPGPU의 성능분석과 소비전력분석을 위한 모의도구로서 GPGPU체계 연구에서 많이 리용되고있다.[2] GPGPU-Sim은 GPGPU연산조작의 주기특성을 분석하여 GPGPU의 성능과 소비전력을 평가한다. 그러나 PCM(Phase Change Memory: 상변화형기억기)에 기초한 GPGPU주기억체계의 수명평가를 위한 기능이 지원되지 못하고있다.

주기억체계의 수명을 평가하려면 주기억체계의 주소별쓰기특성을 평가하여야 한다.

본문에서는 GPGPU-Sim에 주기억체계의 쓰기특성을 평가하는 기능을 추가하여 GPGPU의 기억쓰기특성을 분석하고 PCM에 기초한 GPGPU주기억수명개선연구를 위한 실험적환경에 대하여 논의하였다.

1. GPGPU-Sim의 프로그램적구조연구

GPGPU-Sim3.x는 NVIDIA Geforce8x, 9x와 Fermi계렬의 구조와 유사한 구조를 모형화한것으로서 RODINIA시험프로그램조에 대하여 97%이상의 IPC상관도를 나타내는 정확한 모의도구로 알려져있다.[1]

GPGPU-Sim3.x는 크게 3개의 모듈 즉 CUDA-Sim, GPGPU-Sim, intersim으로 구성되는데 여기서 GPU구조는 GPGPU-Sim모듈에 의해 실현된다.

GPGPU-Sim모듈은 다시 18개의 부분모듈 즉 addrdec, delayqueue, dram_sched, dram, gpu-cache, gpu-misc, histogram, icnt_wrapper, l2cache, mem_fetch, mem_fetch_status, mem_latency_stat, scoreboard, shader, stack, stats, stat-tool, visualizer로 구성되는데 본문에서는 주기억의 시간특성과 주기억과 다른 부분의 결합을 모형화한 dram.h, dram.cc를 수정함으로써 기억기통로안에 bank와 col, row별로 계수기를 설치하여 쓰기특성을 종합적으로 분석할수 있게 하였다.

2. 수정된 GPGPU-Sim을 리용한 주기억호출특성분석

여기서는 수정된 GPGPU-Sim3.2를 리용하여 7개의 시험프로그램에 대한 row별쓰기회수를 분석하였다.(표)

표와 같은 환경밑에서 시험프로그램 STO의 row별쓰기특성은 그림 1과 같다.

표. GPU모의기의 기본파라미터 환경설정

변수이름	값
SM의 수/개	16
매개 SM이 지원하는 스레드수/개	2 048
row수/개	4 096
L2캐쉬	$\langle nsets \rangle : \langle bsize \rangle : \langle assoc \rangle = 64 : 128 : 8$ $\langle rep \rangle : \langle wr \rangle : \langle alloc \rangle : \langle wr_alloc \rangle = L : B : m : W,$ $\langle mshr \rangle : \langle N \rangle : \langle merge \rangle = A : 32 : 4$ $\langle mq \rangle = 32$
주기억용량	2GB
주기억대역폭	8B/주기
주기억관리산법	FR-RCFS
주기억조종기수	6

그림 1에서 알수 있는바와 같이 시험프로그램 STO는 특정한 row에 대하여서만 쓰기를 진행(2개의 row 1 365와 1 366에 대해서만 각각 348회와 700회 정도 진행)하고 나머지 row에 대하여서는 진행하지 않았다. 시험프로그램 NN, NQU, BFS, LIB들도 모두 STO와 비슷한 특성을 나타내었다. 즉 시험프로그램 NQU는 1개의 row인 1 367에 대해서만 8회의 쓰기를 진행하고 NN은 2개의 row인 1 366과 1 367에 대하여 각각 800, 520회정도, BFS는 2개의 row인 1 371과 1 372에 대해서만 각각 100, 800회정도 쓰기를 진행하였다.

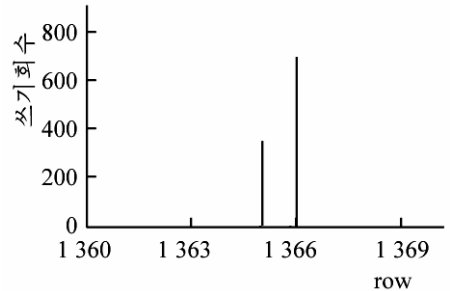


그림 1. 시험프로그램 STO의 row별쓰기특성

시험프로그램 RAY와 MUM도 위의 시험프로그램들에 비하여 쓰기를 진행하는 row들의 수가 많을뿐 기본특성은 같았다.

3. PCM을 GPGPU주기억으로 리용하는 경우 수명특성

위의 7개의 시험프로그램에 대하여 최대쓰기허용회수가 10^6 인 PCM에 기초한 주기억을 리용하는 경우 수명을 계산하면 그림 2와 같다.

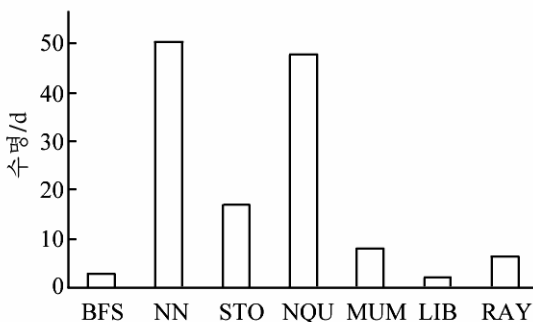


그림 2. 7개의 시험프로그램에 대한 수명특성

한편 체계의 리론상수명은 체계의 쓰기가량이 BGbps이고 매개 세포의 최대쓰기회수가 W_{max} , 기억기크기가 SGB라고 할 때 다음의 공식으로 계산할수 있다.

$$\text{수명} = W_{max} \times S/B(s)$$

실험에서 리용한 시험프로그램들에 대하여 평균쓰기회수가 초당 12GB이고 $S=4\ 096$, $W_{max}=106$ 임을 고려하면 리론상수명은 3 951이다. 결국 수명이 제일 긴 NN에 대해서도 리론상수명의 1.26%밖에 도달하지 못한것으로

된다. 이로부터 일부 특정한 row들에 집중되어있는 쓰기를 모든 row들에 대하여 평등분포시키는 방법이 필요하다.

현재 일반CPU체계에 대한 마모평균화론리에 대한 연구가 많이 진행되고있으며 이것은 GPGPU에 대해서도 그대로 리용할수 있다.

한편 GPGPU는 일부 row에 대해서만 쓰기를 진행하고 많은 row에 대해서는 읽기만 하고 쓰기는 진행하지 않는다.

이러한 특성을 고려하여 기억체계를 PCM과 DDR의 혼합형으로 설계하면 PCM의 수명을 보다 높일수 있다.

맺는 말

GPGPU-Sim3.2의 일부 모듈들의 프로그램을 수정하여 GPGPU의 기억쓰기특성을 분석함으로써 GPGPU의 주기억으로 PCM을 리용하는데서 제기되는 기술적문제들을 해명하였다.

참고 문헌

- [1] M. K. Qureshi et al.; in Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture, 14, 2009.
- [2] 穆帅 等; 计算机科学, 40, 10, 2013.

주체105(2016)년 10월 5일 원고접수

GPGPU Memory Write Behaviors using GPGPU-Sim

Pak Sang Ryong

To extend the endurance of PCM in PCM based GPGPU memory, GPGPU memory write behavior using modified GPGPU-Sim was studied. It showed that seven benchmarks had one or two extremes of memory write behaviors.

Key words: PCM, endurance, GPGPU