New Technique to Reduce Register File Accesses in GPUs

Tech ID: 32223 / UC Case 2021-806-0

BACKGROUND

Graphics processing units (GPU) or graphics cards create images on a variety of display devices ranging from phones to gaming systems. The register file (RF) is a critical structure in GPUs. To simplify the architecture of the RF, it is organized in a multi-bank configuration with a single port for each bank. The frequent accesses to the register file during kernel execution incur a sizeable overhead in GPU power consumption, and introduce delays as accesses are serialized when port conflicts occur.

BRIEF DESCRIPTION

Prof. Nael Ghazaleh and Hodjat Asghari Esfeden from the University of California, Riverside have developed Breathing Operand Windows (BOW), an enhanced GPU pipeline and operand collector technique that supports bypassing register file accesses and instead passes values directly between instructions within the same window. While this baseline design can only bypass register reads, they also introduce an improved design capable of bypassing unnecessary write operations to the RF. Compiler optimizations help guide the write-back destination of operands depending on whether they will be reused to further reduce the write traffic.

The BOW microarchitecture reduces RF dynamic energy consumption by 55%, while at the same time increases overall performance by 11%, with a modest overhead of 12KB of additional storage which is ~4% of the RF size.
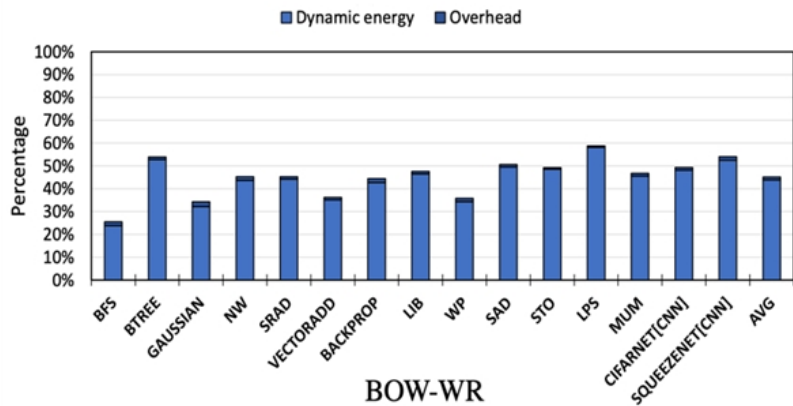

Fig 1: shows the dynamic energy normalized to the baseline GPU for BOW-WR across fifteen different benchmarks. The small segments on top of each bar represent the overheads of the structures added by the idea. Dynamic energy savings in Fig 1 are due to the reduced number of accesses to the register file as BOW-WR shields the RF from unnecessary read and write operations.

APPLICATION

New GPU microarchitecture to increase GPU performance and reduce energy consumption by reducing register file accesses.

CONTACT

Venkata S. Krishnamurty
venkata.krishnamurty@ucr.edu
tel: .

OTHER INFORMATION

KEYWORDS

operand bypassing, GPU, graphics processing unit, graphics card, register file, microarchitecture, compiler

CATEGORIZED AS

▶ **Computer**
   ▶ Other
▶ **Semiconductors**
   ▶ Design and Fabrication

RELATED CASES

2021-806-0, 2019-123-0

PATENT STATUS

Country	Type	Number	Dated	Case
United States Of America	Published Application	20230393850	12/07/2023	2021-806

RELATED MATERIALS

▶ Esfeden, Hodjat Asghari, Amirali Abdolrashidi, Shafiur Rahman, Daniel Wong, and Nael Abu-Ghazaleh. "BOW: Breathing Operand Windows to Exploit Bypassing in GPUs." In Proceedings of the 53rd IEEE/ACM International Symposium on Microarchitecture (MICRO'20). 2020. - 10/19/2020

RELATED TECHNOLOGIES

▶ Corf: Coalescing Operand Register File For Graphical Processing Units