

Snake: A Variable-length Chain-based Prefetching Mechanism for GPUs

هاجر فلاحتی

پژوهشکده علوم کامپیوتر، پژوهشگاه دانش‌های بنیادی (IPM)

Abstract

Graphics Processing Units (GPUs) utilize memory hierarchy and Thread-Level Parallelism (TLP) to tolerate off-chip memory latency, which is a significant bottleneck for memory-bound applications. However, parallel threads generate a large number of memory requests, which increases the average memory latency and degrades cache performance due to high contention, resulting in suboptimal performance. Prefetching is an effective technique for reducing memory access latency, and prior research has shown the positive impact of stride-based prefetching on GPU performance. However, existing prefetching methods do not fully exploit the potential benefits of prefetching, due to their reliance on fixed strides.

In this talk, I will briefly elaborate on the prior prefetching approaches in GPUs and introduce a new prefetching technique called Snake. Snake is built upon chains of variable strides, using throttling and memory decoupling strategies. Snake prefetches 80% of memory requests with 90% accurately prefetched requests, improves GPU performance by 17%, and reduces energy consumption by 17% in memory-bound GPGPU applications.

Biography

Hajar Falahati is a Senior Postdoctoral Researcher at the School of Computer Science, Institute for Research in Fundamental Sciences (IPM). She received her BS.c from Isfahan University of Technology in 2009. She received her PhD and MS.c from the Sharif University of Technology in 2016 and 2011, respectively. She had work experience in driver development, data analysis, and ML model development at Arsh, Graph, and PANTOhealth companies. Her research interests include hardware accelerators, GPU architecture and applications, machine learning, bioinformatics algorithms, and near-data processing.

زمان: چهارشنبه ۱۴۰۲/۰۶/۲۲ - ساعت ۱۵:۰۰

ارائه به صورت مجازی انجام خواهد شد.

<https://vmeeting.ipm.ir/b/com-hrj-f9n>

*** شرکت برای عموم علاقه‌مندان آزاد است ***