香山微结构模拟器 – XS-GEM5

周耀阳
北京开源芯片研究院
2023年8月24日

香山微结构模拟器 – XS-GEM5

Features：

- Full-system simulation
  - **RVGCpt: Xiangshan’s RISC-V generic checkpoint**
  - **Online Difftest**
- Estimated SPECCPU 06 overall score (with SimPoint, 16MB L3)
  - GCC12 o3, DDR4 2400 dual-channel: **30.35@2GHz**
- Frontend: micro-architecture calibration
- Backend/Cache: latency/throughput calibration
- Memory system optimization
  - Prefetchers
  - Parallel PTW、 L2 TLB、 TLB prefetching
  - Membus bandwidth and memory controller calibration

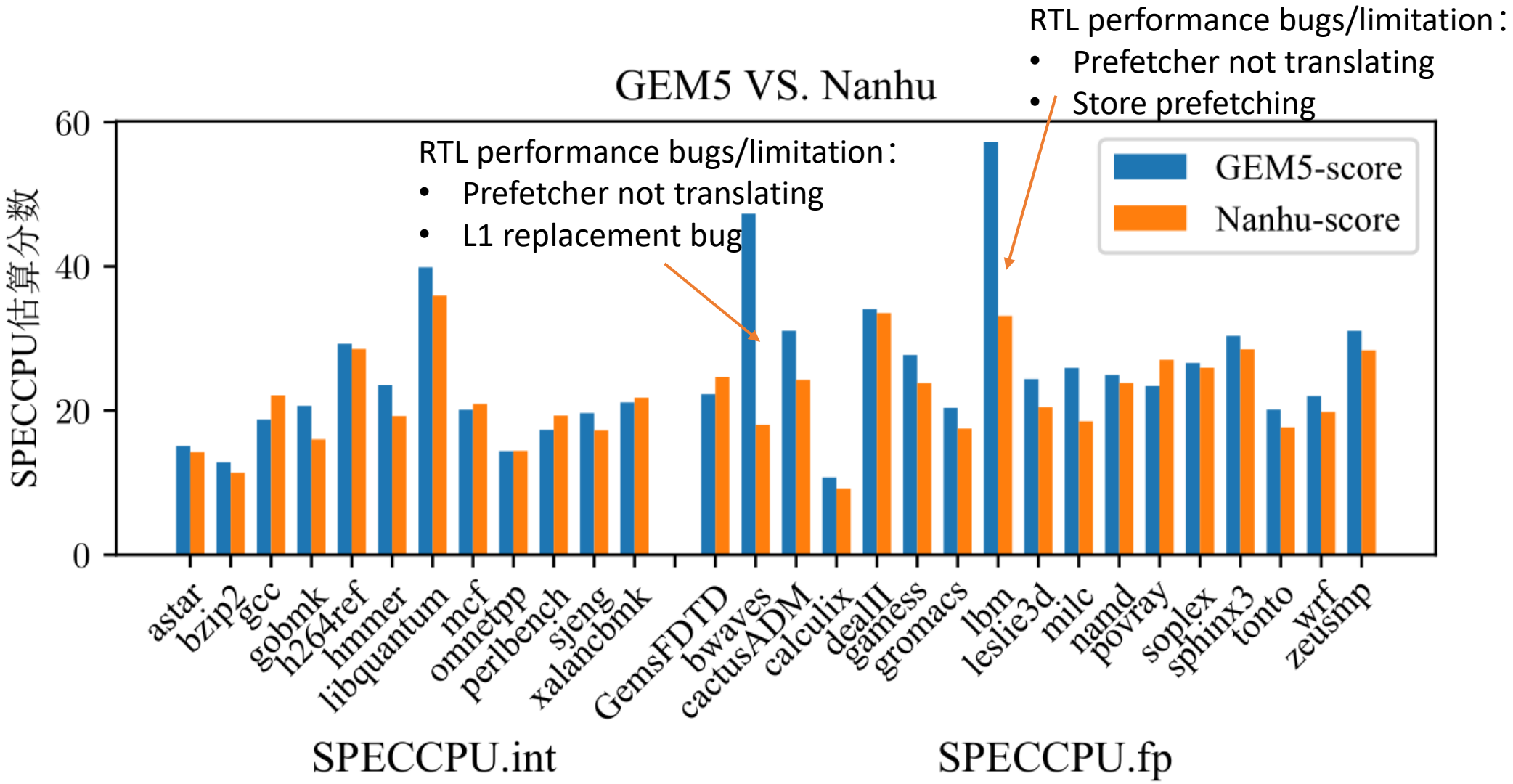
Agenda

- Basic information
- Why u-arch simulator for Xiangshan
- How to 刷分 optimize from 24@2GHz to 30@2GHz

Configuration

- 192 Int PRF + 192 FP PRF, 256-entry ROB, 80-entry LQ, 64-entry SQ
- 32 ITLB + 64 DTLB + L2 TLB , nextline prefetcher
- 64KB L1 Instruction Cache
- 64KB L1 Data Cache with 32 MSHR
  - Default prefetchers: **Stream + Stride + SMS + BOP**
  - Optional prefetchers: IPCP, SPP, and temporal prefetcher
  - Prefetching request **offloading**: passive & active
- 1MB **inclusive** L2 Cache with 64 MSHR
- **16MB** non-inclusive L3 Cache with 128 MSHR

Performance comparison @ April 2023



Current SPECCPU 2006 Score

- Estimated with RVGCpt + SimPoint
- Memory simulated with DRAMSim3:
  - DDR4 2400 Dual-channel
- Cache: 16MB L3
- Prefetcher:
  - **Stream + Stride + SMS + BOP**
  - With IPCP/SPP/Temporal **off**
- **SPEC int 2006 score: 14.68/GHz**
  - Compilation: **GCC 12 o3 base**

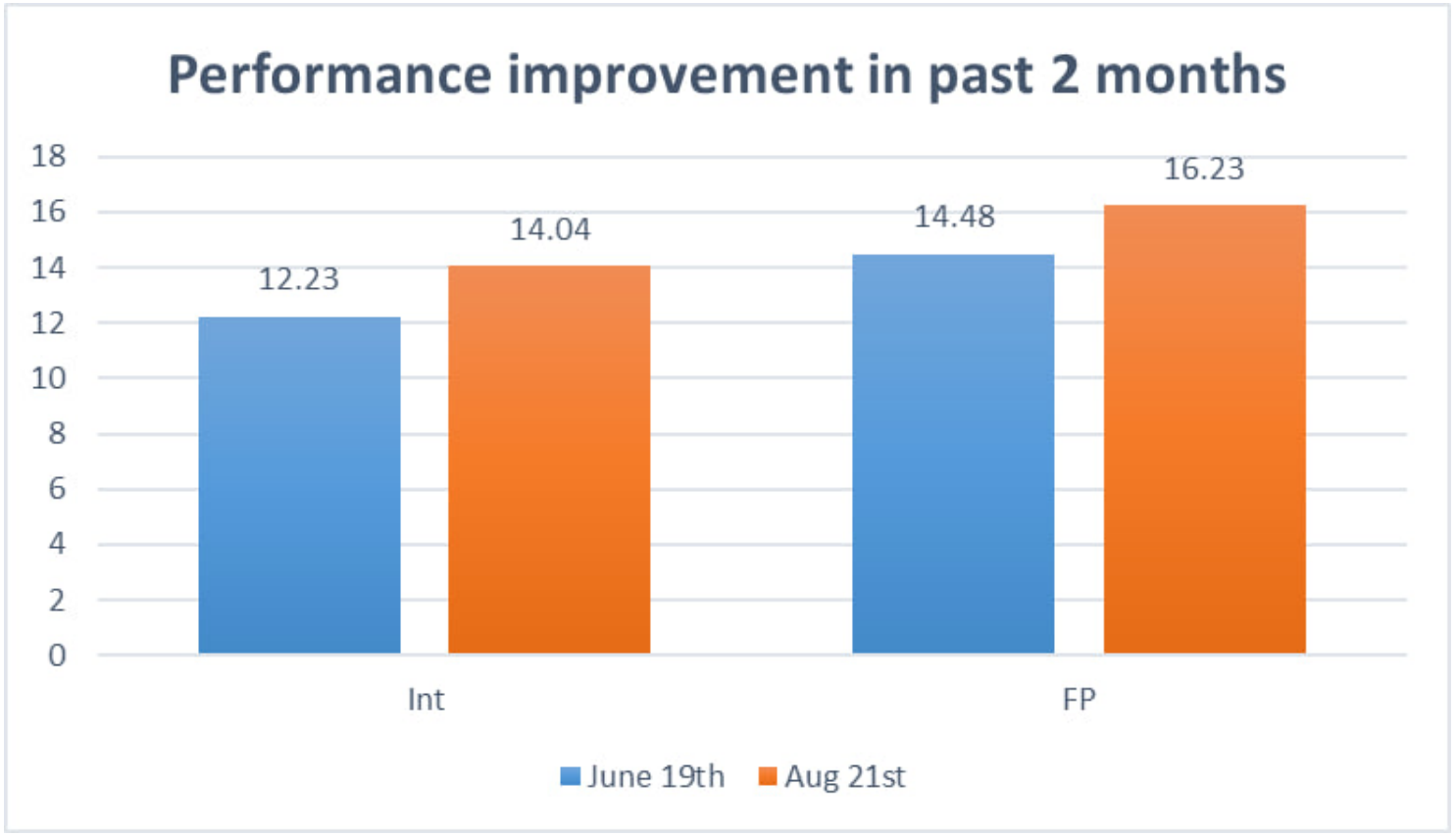
	benchmarks	gcc 10 o2 **Dual-channel**	gcc 12 o3 **Dual-channel**	gcc 12 o3 **Single-channel**
Integer	perlbench	22.40	22.37	22.27
	bzip2	15.58	15.98	15.94
	gcc	27.77	29.80	29.00
	mcf	43.94	46.92	41.04
	gobmk	24.32	27.09	27.04
	hmmer	28.41	24.85	24.86
	sjeng	23.70	25.03	24.85
	libquantum	83.09	103.59	65.92
	h264ref	33.91	35.43	35.41
	omnetpp	20.80	24.41	23.59
	astar	20.78	22.41	22.34
	xalancbmk	28.39	24.24	23.71
	Estimated @2GHz	28.09	29.36	27.72
	Estimated per GHz	**14.04**	**14.68**	**13.86**
Float	**Estimated per GHz**	**16.23**	**15.54**	**14.85**

Why there was no u-arch simulator for Xiangshan before

- Why not u-arch simulator in the past?
  - Limited manpower
  - Our past partner suggests not to develop a u-arch simulator
    - V-core （微核芯）
  - Some believe that architectural simulators are useless
    - “From paper to paper” – 龙芯
    - “Considered harmful”
- Why u-arch simulator now?
  - Best practice from industry
  - Faster Design Space Exploration (DSE)
  - Our current partners (on simulator) urge us to develop a u-arch simulator
    - Tencent Penglai （腾讯蓬莱）and other partners

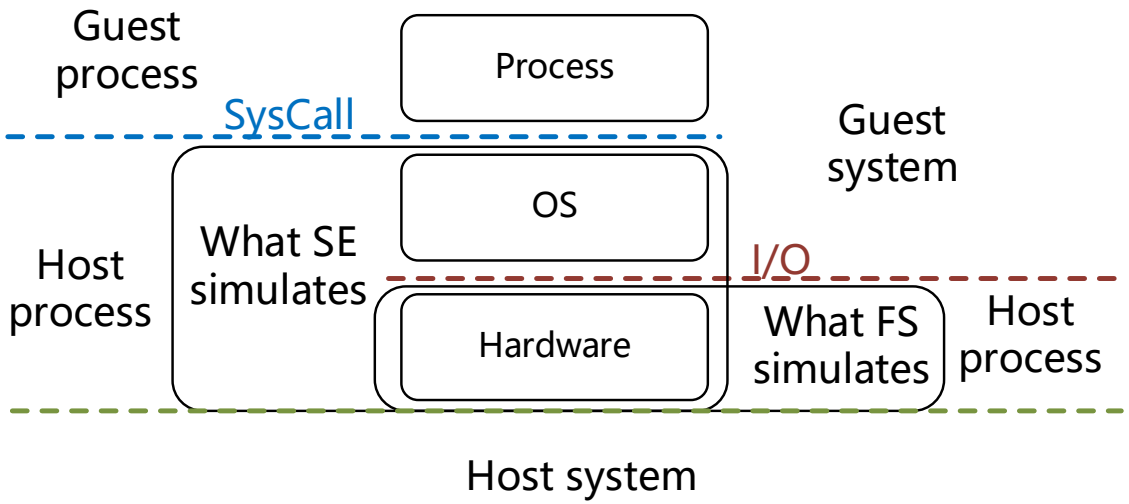
On design space exploration

- SPECCPU score on XS-GEM5 increases about 1.8/GHz in past two months
  - With GCC 10 o2 and 16 MB L3



Why full-system simulation

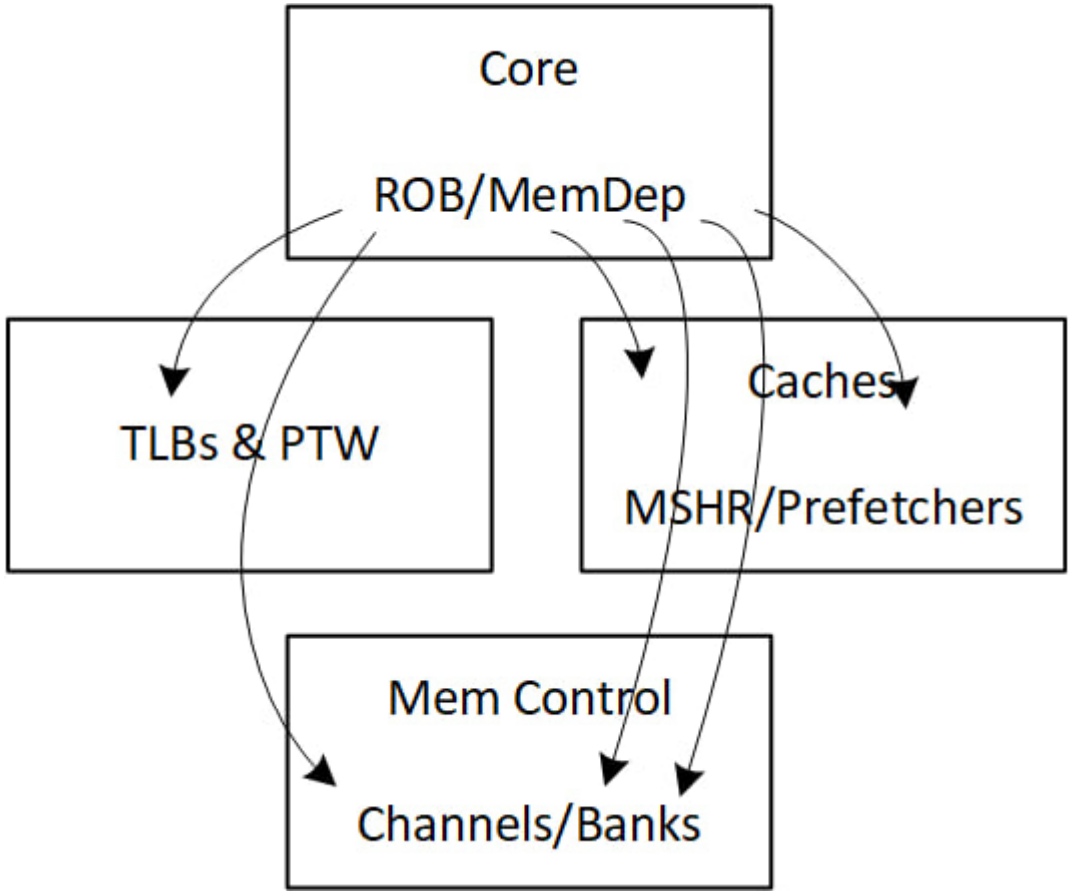
- Misconception: FS simulation is slow
  - QEMU-system runs at near-native speed
- Why full-system
  - Easier
    - Avoid hacky system calls in SE (system call emulation)
    - Checkpointing is easier
  - Accurate system call modeling
    - Performance of serializing instructions
  - Accurate virtual memory behavior
    - Random VA → PA mapping
    - Accurate TLB/PTW modelling



How: Memory-level parallelism is all you need

- MLP in Core
  - False violation in GEM5's LSU
  - False positive prediction in Storeset
- MLP in MMU
  - Parallel PTW
- MLP in **Caches & Prefetchers:**

  Top-1 contribution to performance
- Parallelism in Memory Controller
  - Channel-level parallelism
  - Bank-level parallelism



Single-threaded stream performance

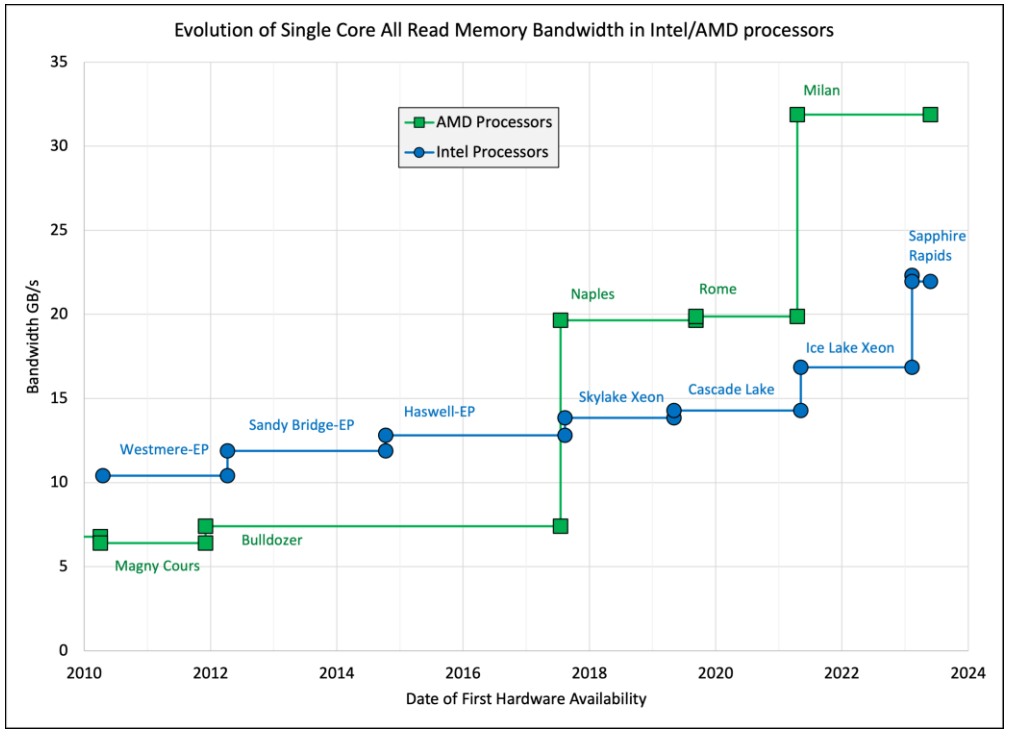
- Increasing single-threaded stream score

龙芯 3A6000 Stream（V5.10）测试数据

（单位：MB/s）

模式	单线程数据带宽	双线程数据带宽	4 线程数据带宽	8 线程数据带宽
Copy	32210.8	38858.6	42467.9	36450.4
Scale	19788.4	41964.0	42199.5	35999.7
Add	32921.8	42807.2	42151.6	34493.8
Triad	33028.5	42683.8	42020.5	34451.9

https://mp.weixin.qq.com/s/Lm_6varu0ovntPGfVzeGLw

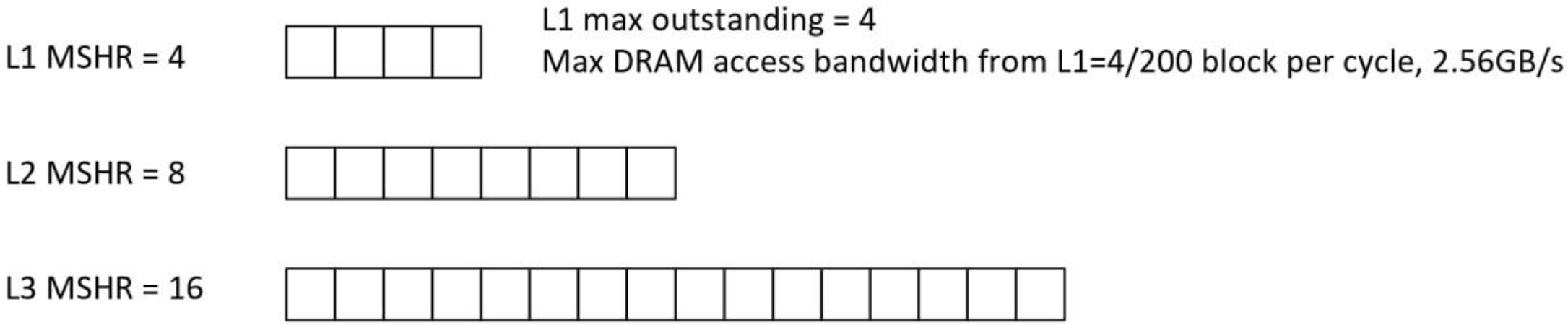


https://sites.utexas.edu/jdm4372/2023/04/25/the-evolution-of-single-core-bandwidth-in-multicore-processors/

Stream performance and MLP

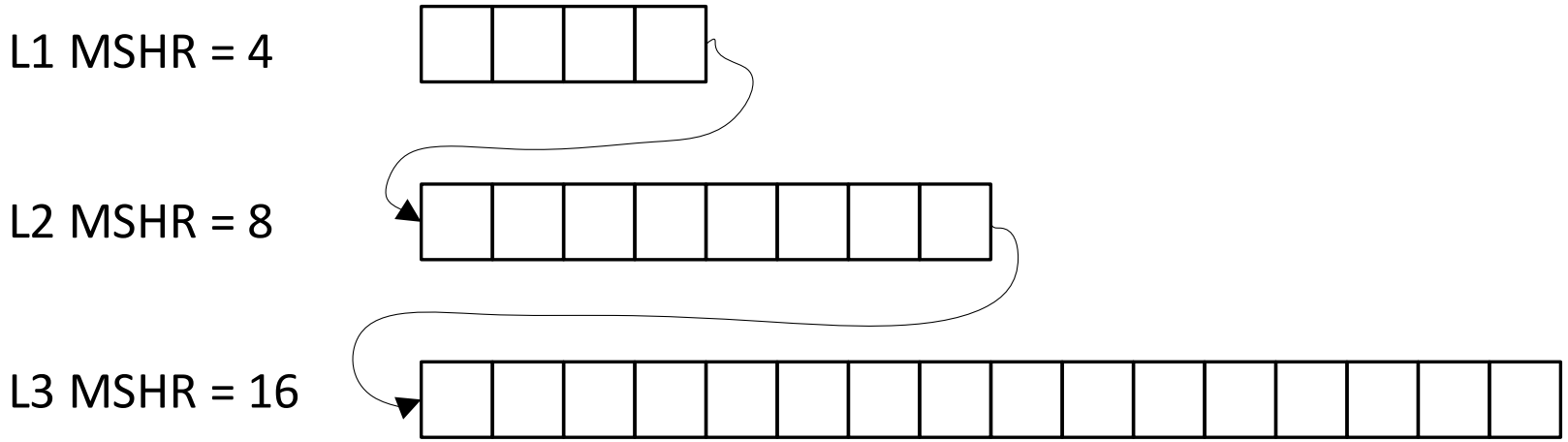
- To make use of 40GB/s DRAM bandwidth, we need at least 64 outstanding misses

Assuming DRAM latency=200 cycle,
CPU freq=2GHz (1 cycle = 0.5ns)



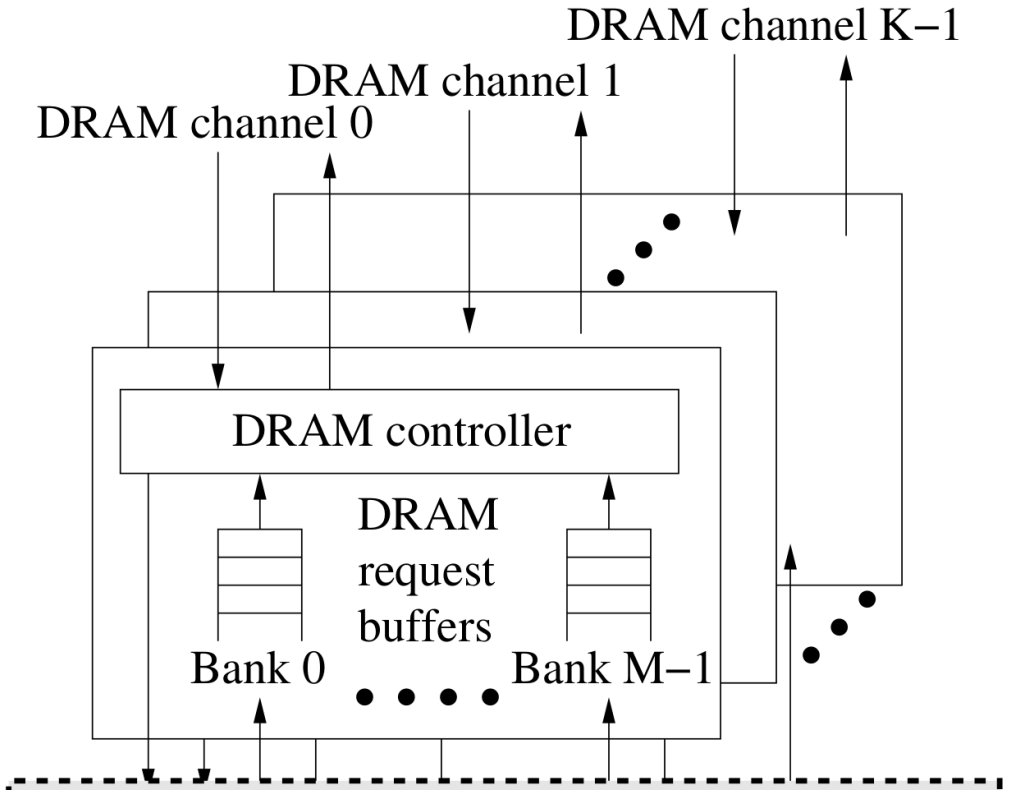
Prefetching with L2/L3 MSHR

- Generate prefetch requests and translate them in L1
- Offload them to L2/L3
  - Actively: on low confidence
  - Passively: on MSHR full



Parallelism in Memory Controller

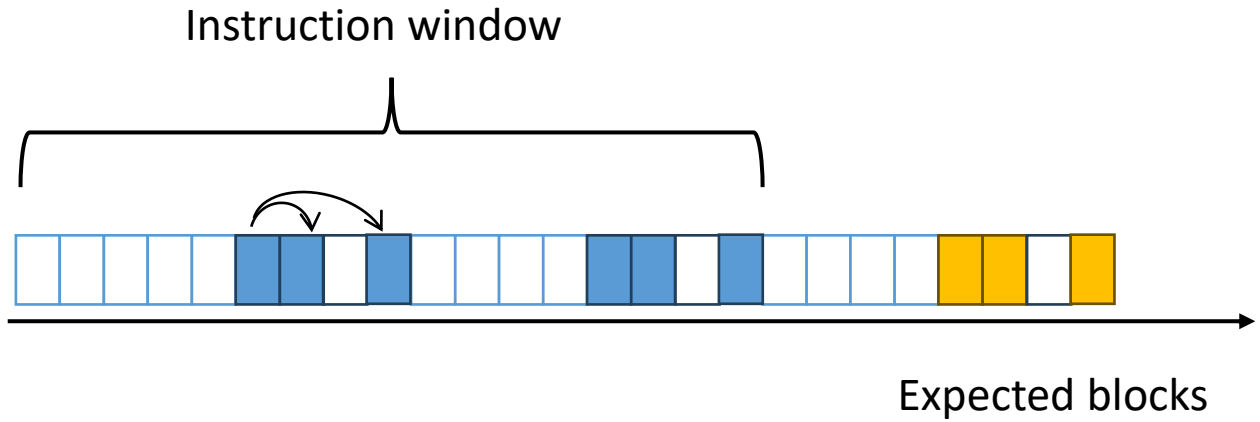
- Well-known dual channel
  - Libquantum is often bandwidth-bound
- Bank-level parallelism
  - BLP-Aware Prefetch @ MICRO 2009 by Mutlu & Patt
  - Optimize address mapping for prefetcher
- Stream with offloading + BLP + dual channel → ***libquantum=80+***
  - Confirmed by RTL



Lee, Chang Joo, et al. "Improving memory bank-level parallelism in the presence of prefetching." Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture. 2009.

Lookahead VS. MLP

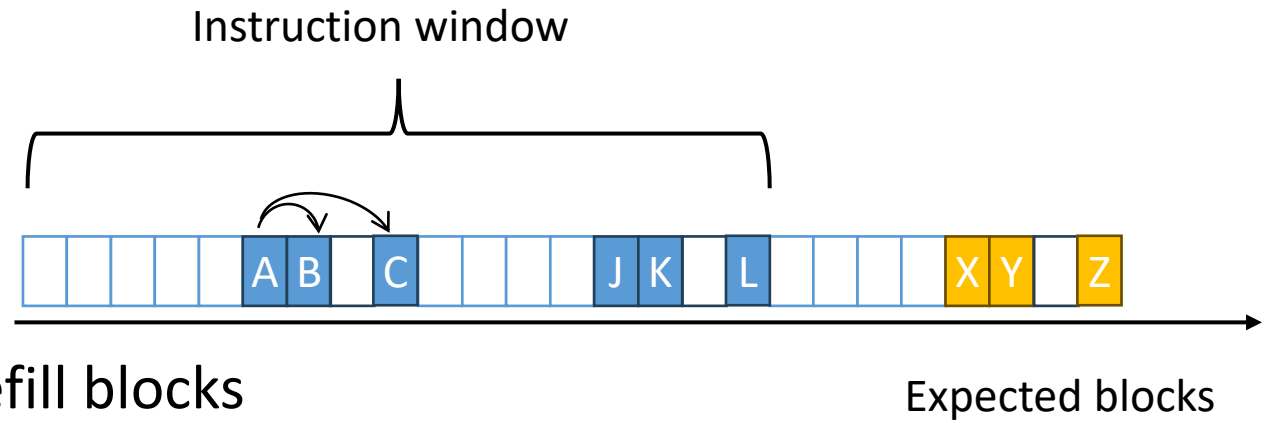
- Prefetching inside instruction window is useless
  - Timely prefetching



  - How to timely prefetch for increasingly large instruction window?
  - Adaptive lookahead depth
    - Easy to lookahead: Stream, stride, IPCP, and SPP
    - Hard to lookahead: BOP, SMS?

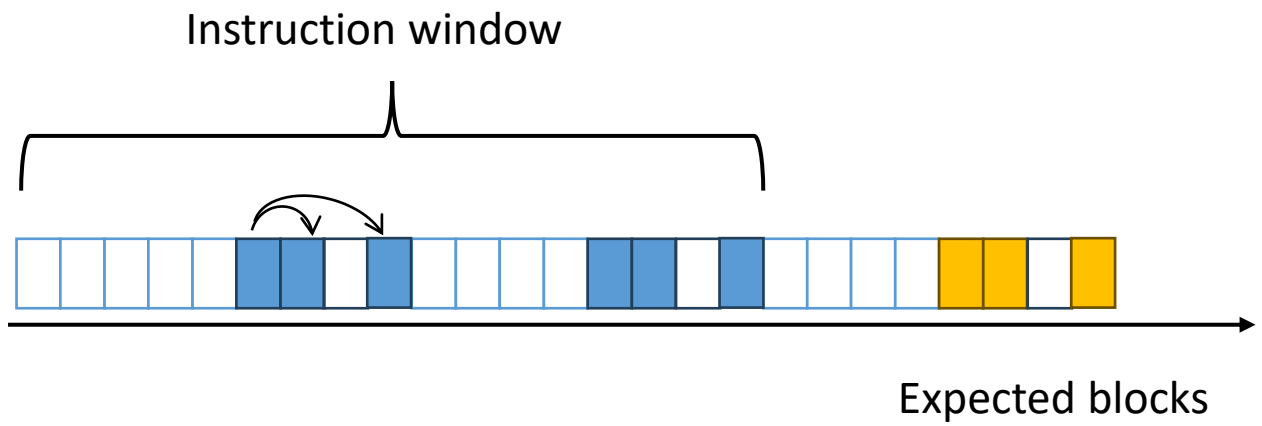
Tune BOP for timely prefetching

- Best offsets = 8, 16
- How to let BOP favor 16?
  - BOPv1: Delay queue
  - BOPv2: Update rrTable with refill blocks
- What if 16 > page size?
  - Virtual address prefetching
  - SPP-PSA @ MICRO 2022: **VA-prefetching doubles performance of *milc***
- BOP with VA: ***milc*=30~50@2GHz**
  - Partially confirmed by RTL
- What if 16 not show in offset list?
  - You guess



Spatial pattern capturing - 殊途同归

- SMS: bit vector = 11010000
- SPP: delta table = +1 +2 + 5 +1 +2 +5
- BOP: best offset = 8/16



Sum up

- XS-GEM5:
  a full-system simulator scores ~15/GHz on SPECCPU 06

- Why full-system: easy & accurate
- Why performance matters: for Xiangshan DSE & for u-arch research

- How: optimize the MLP from core to DRAM
  - MLP/prefetcher is all you need!

Acknowledgment

- Thanks for our current and past teammates：
  陈国凯 陈子航 傅腾蛟 勾凌睿 郭鸿宇 贾晓宇 蔺嘉炜
  满洋 王诲喆 王凯帆 徐岩 张传奇 张紫飞 甄好
- Thanks for **Tencent Penglai** on contributing to frontend, Storeset, SPP, and bingo
- Thanks for other individuals and companies on discussing and contributing

Future works

XS-GEM5

- Deliver features to RTL
- Calibrating caches and OoO backend against RTL
- Multi-core support

RVGCpt

- Multi-threaded full-system checkpoints
- Automatically sampling on multi-threaded workloads
- Cross-ISA region of interest matching

Other prefetcher optimizations

- Region localization VS. PC localization
  - SPP VS IPCP
  - SPP + perceptron filter
  - DSDP @ PACT 2022 by HUST

- Delta as signature VS. PC as signature VS. offset as signature
  - Offset as signature: MICRO 2022 by ISCAS

香山性能组 招人！招人！招人！HC充足！

- 工作内容：参与**15+分/GHz**的CPU架构设计！
  - 性能建模，架构对齐，Feature落地，性能gap分析
- 应届生要求（**全职**岗位 + **实习**岗位）：
  - **一生一芯**流水线CPU 和 NEMU实验（aka. PA）
  - 深入学习**UWisc ECE752**
  - 深入理解一种近年的预取/分支预测算法并重现和分析
- 种一棵树最好的时间是十年前，而后是**现在**
- 欢迎**本科生**，一、二年级就可以开始准备
  - 对工作、读研都是核心技能

个 CPU架构岗位投递

个 ECE752课程网站

个一生一芯官网