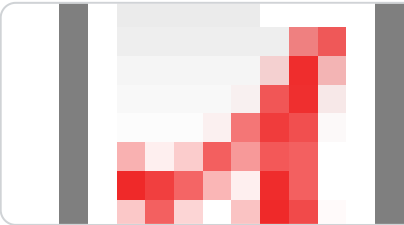


文献研读报告作业

实验报告

- 文献研读报告作业：Managing Memory Tiers with CXL in Virtualized Environments
- 姓名：梁嘉嘉
- 学号：23125240
- 课程：高级操作系统
- 提交日期：2024/11/05

Managing Memory Tiers with CXL in Virtualized Environments



 <https://www.usenix.org/conference/osdi24/presentation/zhong-yuhong>

This website uses cookies

Preferences 0 Preference cookies enable a website to remember information that changes the way the website behaves or looks, like your preferred language or t...

问题产生

随着云计算的增长，数据中心中对内存的需求急剧增加。传统的 DRAM 成本高，且其容量难以满足现代云环境中的虚拟机（VM）需求。为解决这个问题，云供应商引入了基于 CXL（Compute Express Link）的存储扩展技术，将 CXL 设备中的持久内存作为第二层内存以降低成本和能耗。然而，CXL 内存相比本地 DRAM 有较高的访问延迟，且现有的基于软件管理的分层方法在虚拟化环境下性能表现不佳，面临一系列的挑战：

1. 指令采样技术并不适用于虚拟机；
2. 细粒度的页表操作会消耗额外的 CPU；
3. 软件管理的方法只能在 page 的粒度管理内存，会导致一个 page 内同时放置了冷热数据，这一问题在虚拟机普遍使用大页（2M 或 1GB）的情况下变得更加严重。

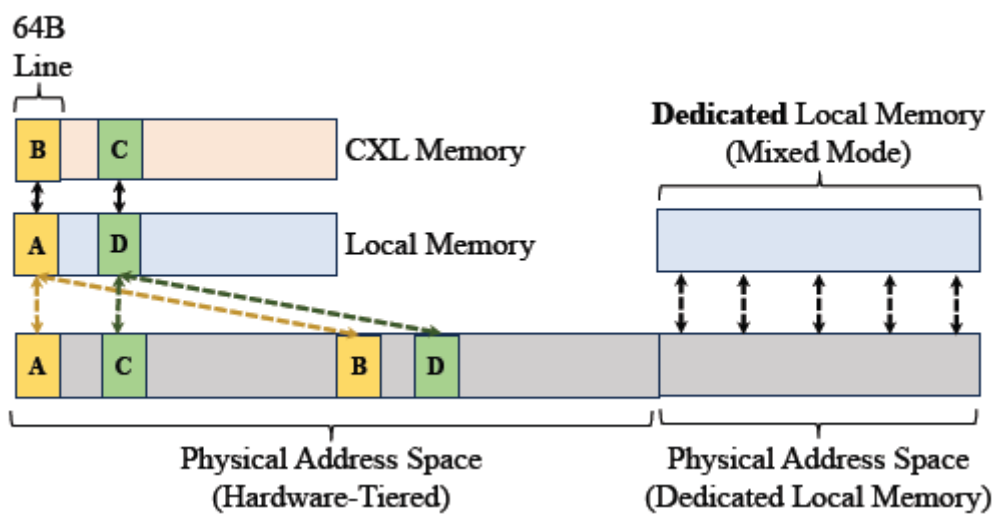
方法/模型

论文提出了一种结合硬件管理和软件多租户隔离的内存分层方法，以应对虚拟化环境中的内存管理挑战：

硬件管理：Intel® Flat Memory Mode

这是第一个支持 CXL 内存的、基于缓存行粒度的硬件管理内存分层方法。其特点是在处理器内存控制器内实现透明的两层内存管理。如下图所示：

- 硬件管理的地址空间（图中左侧的“Hardware-Tiered Memory”部分）中，数据在本地 DRAM 和 CXL 内存之间动态调度，以最小化对主机 CPU 资源的依赖。
- 专用地址空间（图中右侧的“Dedicated Local Memory”部分）则被完全放置在本地 DRAM 中，从而确保访问频繁的数据获得更高性能



软件多租户隔离的内存分层：Memstrata

在虚拟机环境下，不同 VM 之间共享 DRAM 或 LLC 时，可能会出现资源竞争，导致性能下降。为解决这一问题，Memstrata 利用页面着色（page coloring）技术识别并隔离冲突页面，确保相同内存页不在不同 VM 之间共享，从而减少内存访问的冲突，提高整体系统性能和稳定性。

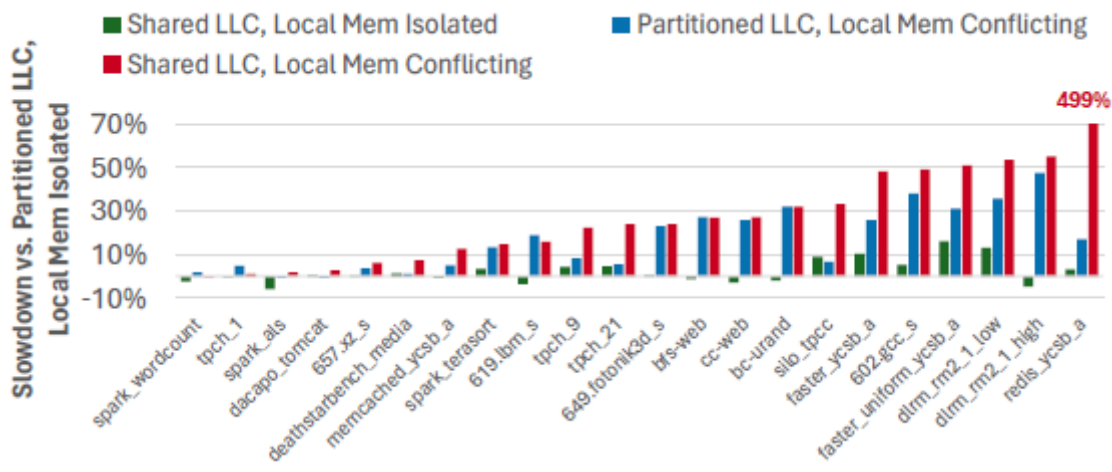


Figure 7: Inter-VM interference caused by LLC contention and local DRAM contention.

测试方案

实验装置：使用了一台预生产的 Intel® Xeon® 处理器，支持 Intel® Flat Memory Mode，配备 128GB 的本地 DRAM 和 128GB 的 CXL 存储扩展内存。测试环境中运行 115 个不同类型的工作负载，包括 Web、数据库、大数据、机器学习等应用，以涵盖多种内存访问模式。

测试在四种内存配置下进行：

1. 本地 DRAM，
2. CXL 内存
3. 硬件分层内存（使用 Intel® Flat Memory Mode）
4. 混合模式（33% 的专用 DRAM 和 67% 的硬件分层内存）

为比较不同方法，研究了硬件分层内存配置下是否采用 Memstrata 带来的性能差异。同时，也对比了其他软件分层方法如 TPP。

结果分析

在四种配置下测试结果显示，使用 Intel® Flat Memory Mode 的硬件分层内存配置在保持相似性能的同时显著降低了 DRAM 需求。通过 Memstrata 的优化，约 20% 的工作负载在硬件分层模式下表现优于完全依赖本地 DRAM 的设置。Memstrata 在多租户环境中，能够通过页面着色隔离虚拟机间的内存访问，避免竞争带来的性能下降。在测试中，对于占主导的 Web 和数据密集型工作负载，Memstrata 使这些应用的性能下降维持在 5% 以内，相比其他方案表现更稳定。这些测试和结果表明，Intel® Flat Memory Mode 结合 Memstrata 在不牺牲性能的情况下有效地扩展了服务器的内存容量，并减少了多租户环境中的资源竞争带来的性能波动。

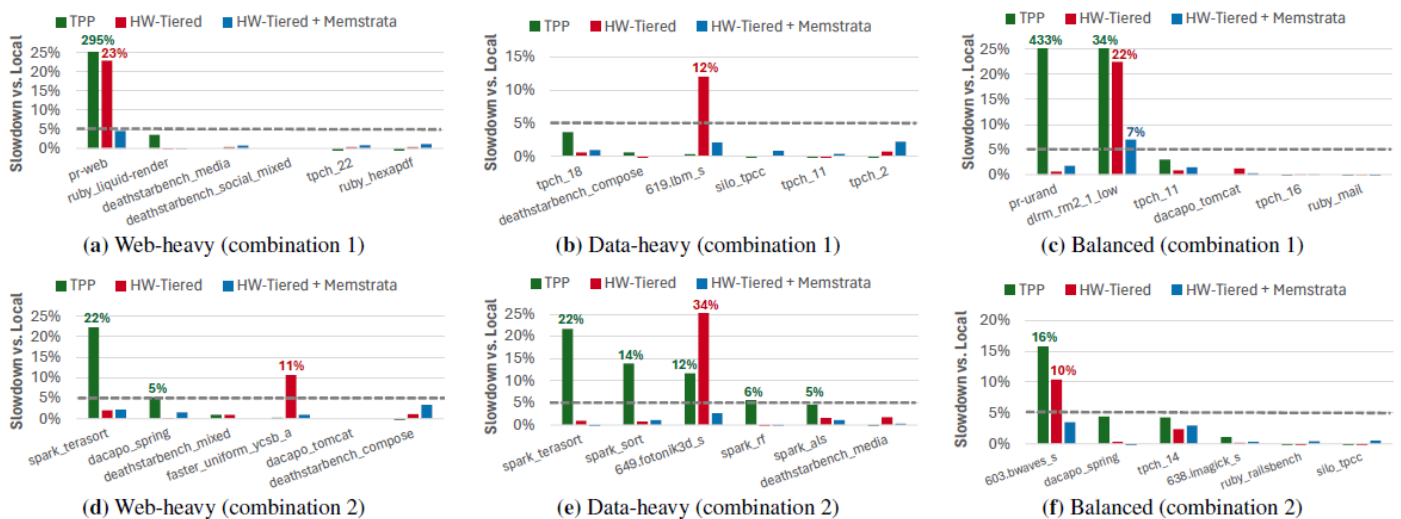


Figure 10: Application slowdown of TPP and hardware tiering with and without Memstrata using different workload combinations.

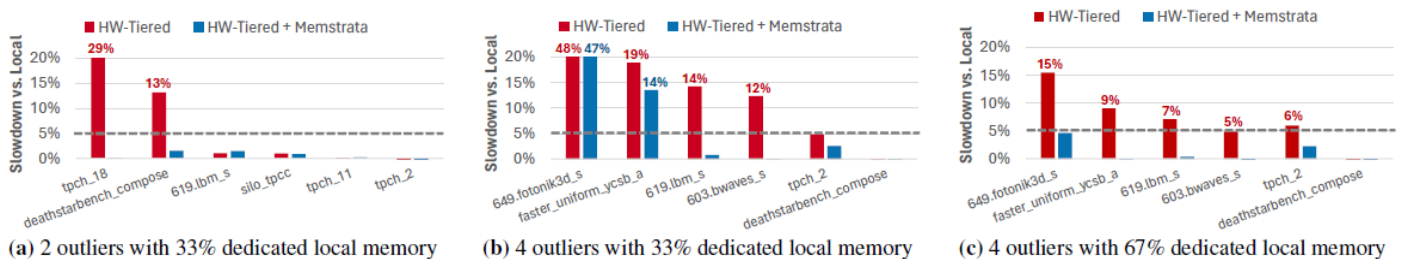


Figure 12: Application slowdown when the workload combination contains 2 or 4 outliers.

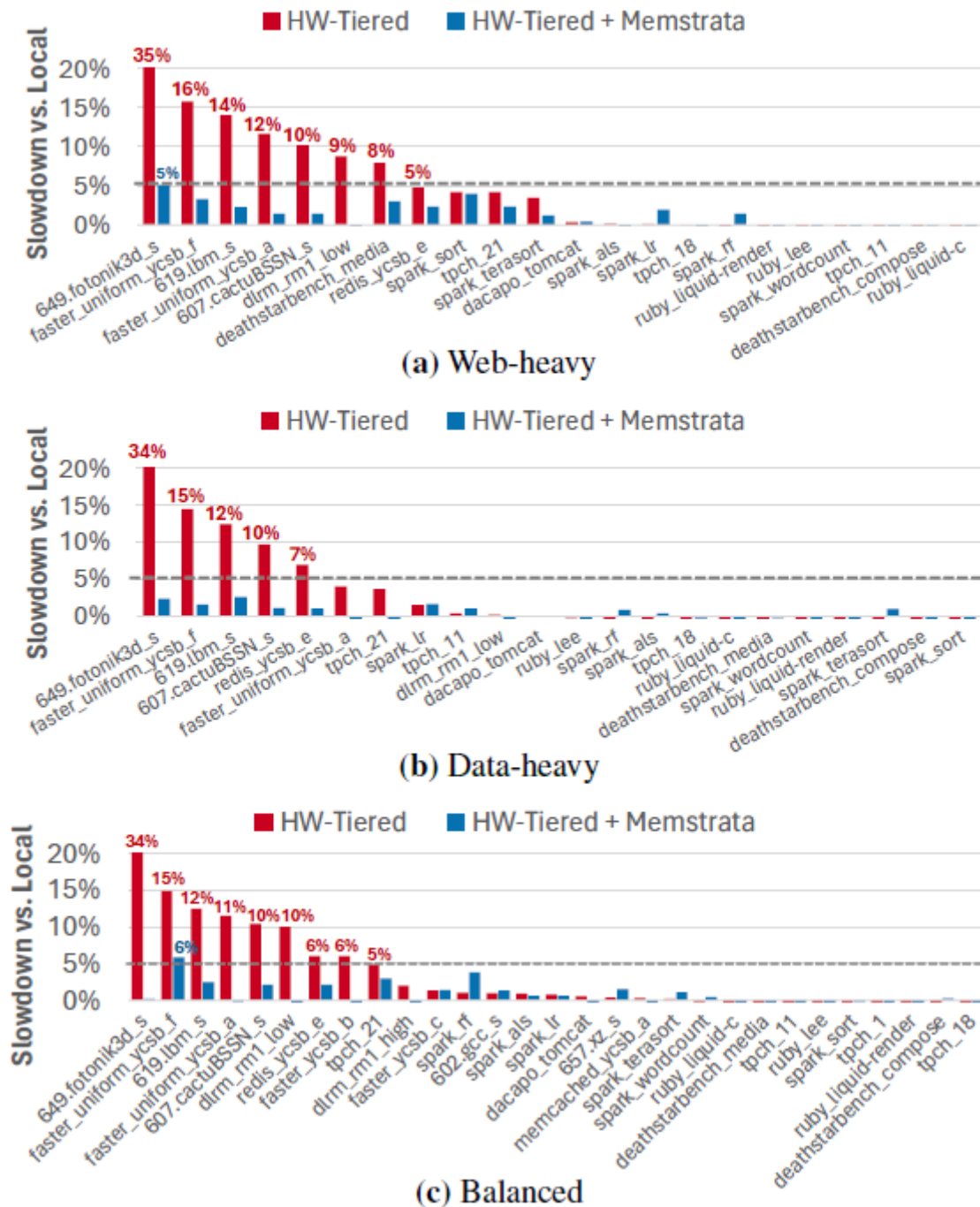


Figure 13: Application slowdown in realistic environments with three different workload mixes.

相关工作比较

传统的基于软件的内存分层方法依赖于虚拟内存页表管理和指令采样，难以在虚拟化环境中高效运行。Intel® Flat Memory Mode 相比更直接地支持内存管理，使得数据能够按缓存行粒度被管理，减少了对操作系统资源的依赖。Memstrata 相比其他多租户管理方案，能更好地在多租户环境下隔离不同租户的内存访问，防止租户之间的干扰。

个人理解

这篇论文展示了硬件与软件协同优化内存管理的潜力。通过引入 Intel® Flat Memory Mode 的硬件分层机制，CXL 内存能够以缓存行为粒度高效管理数据位置，这种硬件级管理极大降低了虚拟化环境中

的CPU开销。同时，Memstrata 的多租户隔离机制，通过页面着色技术减少了虚拟机间的资源竞争。硬件和软件的结合使内存管理更加精细化，能够动态适应不同虚拟机的需求，在保证性能的同时降低成本。

启示总结

该研究对我的启示在于，在多租户或资源受限的环境中，内存管理的优化需要兼顾性能和隔离。通过硬件和软件的协同优化，可以在减少资源冲突的同时提高资源利用率。尤其是 Memstrata 的动态页面分配和隔离策略，可以借鉴到其他资源管理中，例如处理器缓存、I/O 带宽的分配等。在未来的研究中，我可以探索类似的技术，以提升多任务环境中不同资源的分配效率。