Efficient Memory Virtualization

Reducing Dimensionality of Nested Page Walks

Jayneel Gandhi[†], Arkaprava Basu[‡], Mark D. Hill[†], Michael M. Swift[†]

[†] WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

[‡] AMD

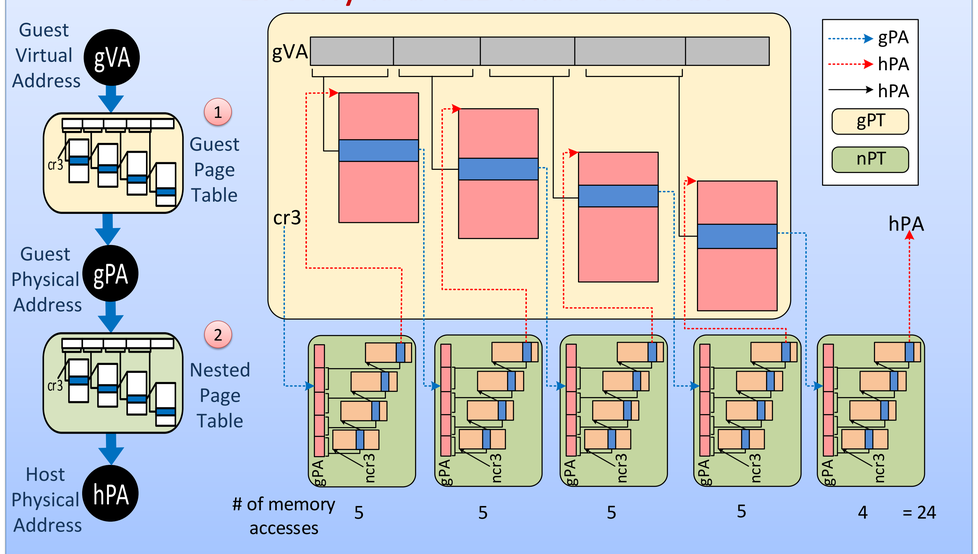
1. Problem

Hardware Virtualized MMU have high overheads

We will show that the increase in translation lookaside buffer (TLB) miss handling costs due to the hardware-assisted memory management unit (MMU) is the largest contributor to the performance gap between native and virtual servers.

—Buell, et al. VMware Technical Journal 2013

2. Why is a TLB miss costlier?

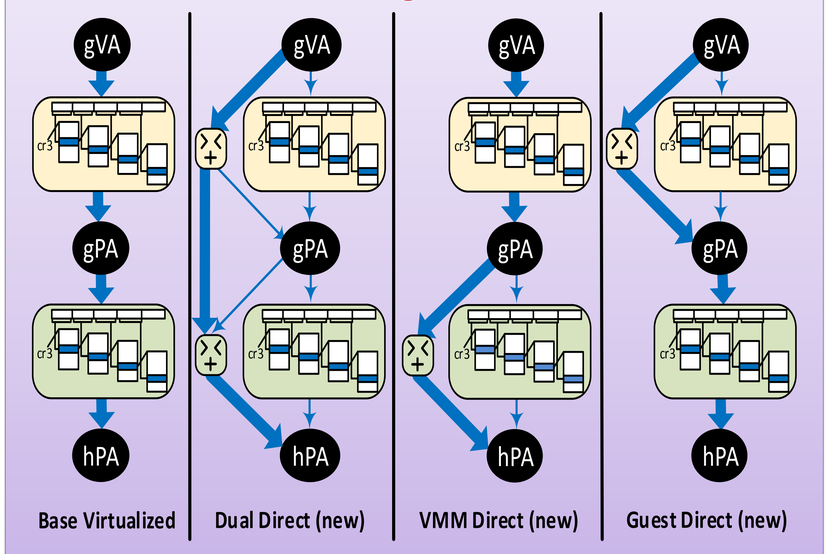


3. Solution

Segmentation to bypass paging

- ❖Extend Direct Segments for virtualization
- ❖Direct Segment at VMM, guest or both levels
- ❖Three modes with different tradeoffs

4. Configurations



5. Tradeoffs

Properties	Base Virtualized	Dual Direct	VMM Direct	Guest Direct
Page walk dimensions	2D	0D	1D	1D
# of memory accesses for most page walks	24	0	4	4
# of base-bound checks for page walks	0	1	5	1
Guest OS modifications	none	required	none	required
VMM modifications	none	required	required	none
Application category	any	big memory	any	big memory
Page sharing	unrestricted	limited	limited	unrestricted
Ballooning	unrestricted	limited	limited	unrestricted
Guest OS swapping	unrestricted	limited	unrestricted	limited
VMM swapping	unrestricted	limited	limited	unrestricted

6. Optimizations

- ❖Guest physical memory fragmentation: *Self-ballooning*
  - Balloon-out fragmented memory and provide to VMM
  - VMM hot-adds new contiguous guest physical memory
- ❖Host physical memory fragmentation: *Compaction*
  - Remap fragmented pages to create contiguous physical memory
- ❖Permanent "hard" memory faults: *Escape filter*
  - Escape filter stores few pages with permanent "hard" faults
  - Escape filter checked in parallel with VMM segment register
  - If found in escape filter, get alternate translation through paging

Acknowledgement

This work is supported in part by the National Science Foundation (CNS-1117280, CCF-1218323, CNS-1302260 and CCF-1438992), Google, and the University of Wisconsin (Kellett award and Named professorship to Hill). Arkaprava Basu's contribution to this work occurred while at UW-Madison.

7. Overheads + Results

Near- or better-than-native performance