



AMD Barcelona and Nested Paging Support in Xen

Elsie Wahlig
Wei Huang
Xen Summit Spring 2007

Introducing "Barcelona"...

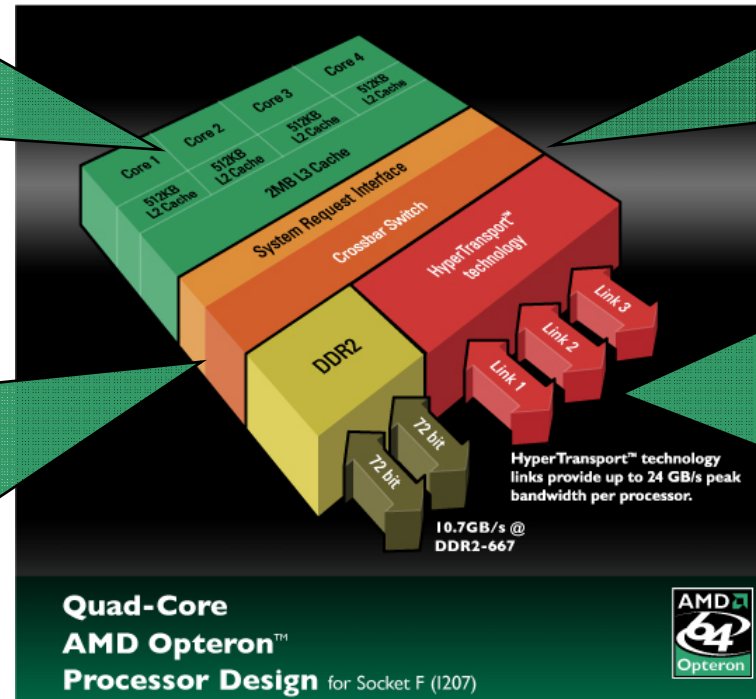
Native Quad-Core Processor

To increase performance-per-watt efficiencies using the same Thermal Design Power.

Advanced Process Technology

65nm Silicon-on Insulator Process

Fast transistors with low power leakage to reduce power and heat.



Platform Compatibility

Socket and thermal compatible with "Socket F".

Direct Connect Architecture

- Integrated memory controller designed for reduced memory latency and increased system bandwidth
 - Memory directly connected
- Provides fast CPU-to-CPU communication
 - CPUs directly connected
- Glueless SMP up to 8 sockets

AMD Virtualization™ Update

Virtualization Updates

- Nested Paging
- Improved switch time

IOMMU

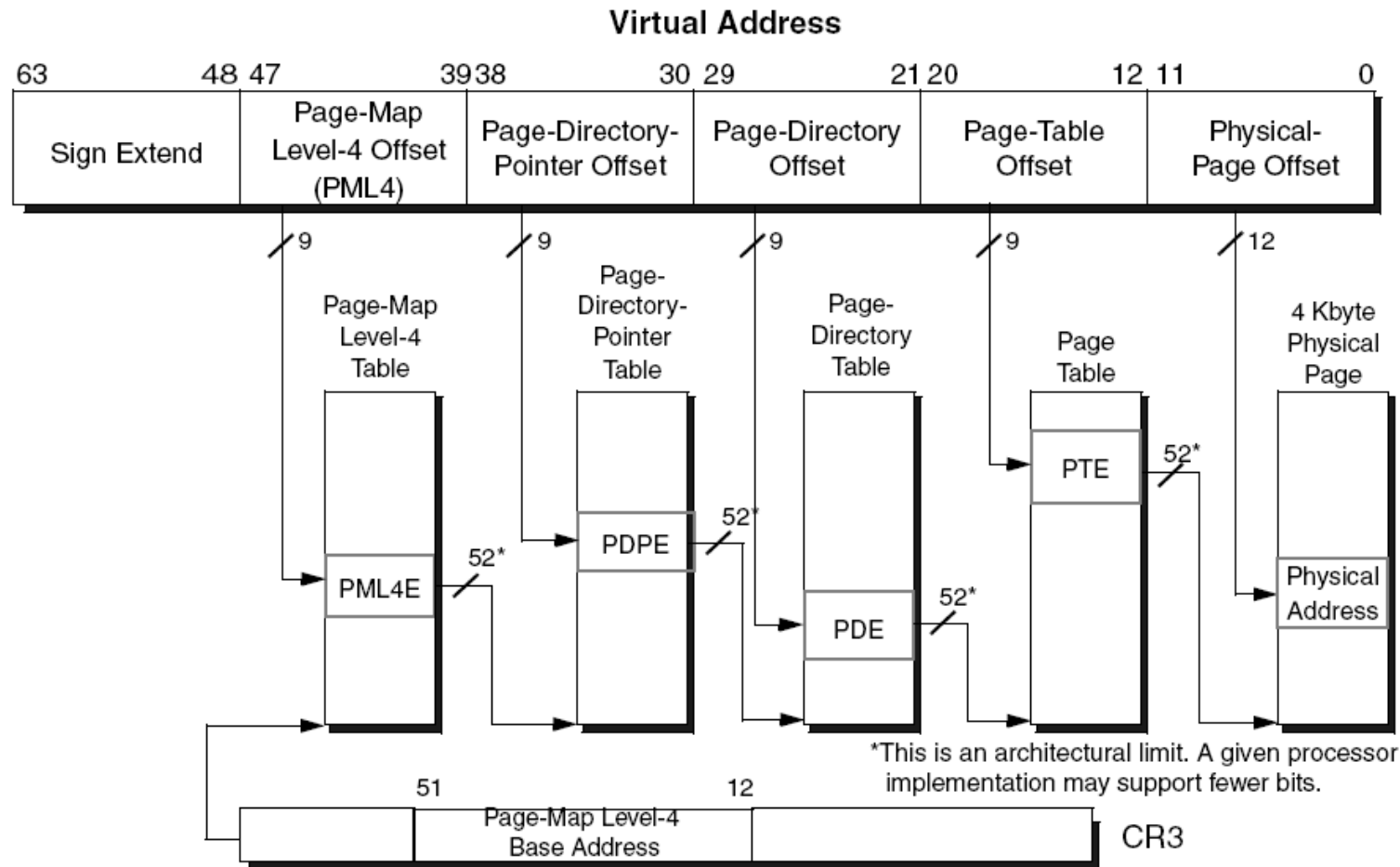
- Update posted

http://www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/34434.pdf

Outline

- Nested paging design
- Performance comparison with SPT
- Current status of Xen nested paging support

Conventional X86 Page Translation



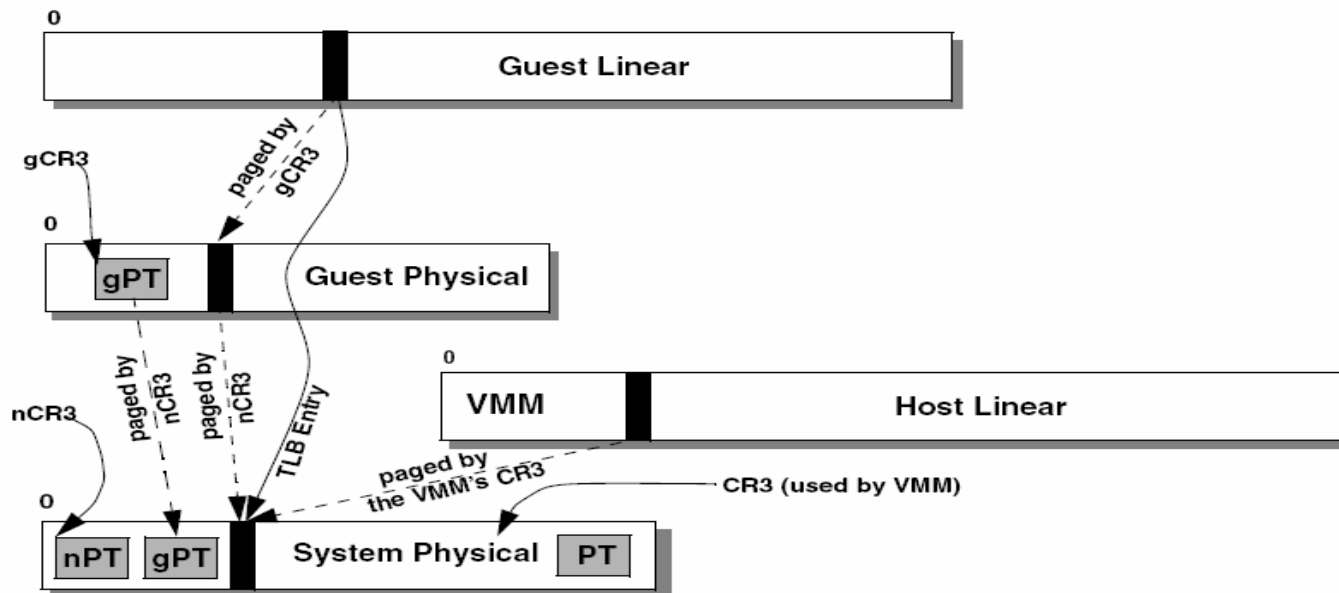
Background of Nested Paging

- Purpose
 - Provides support to translate guest physical address to machine physical address
- Benefits
 - Helps reduce the complexity of memory management in virtualized environment
 - Can improve performance by avoiding a substantial amount of #VMEXIT for memory intensive guests

Design of Nested Paging

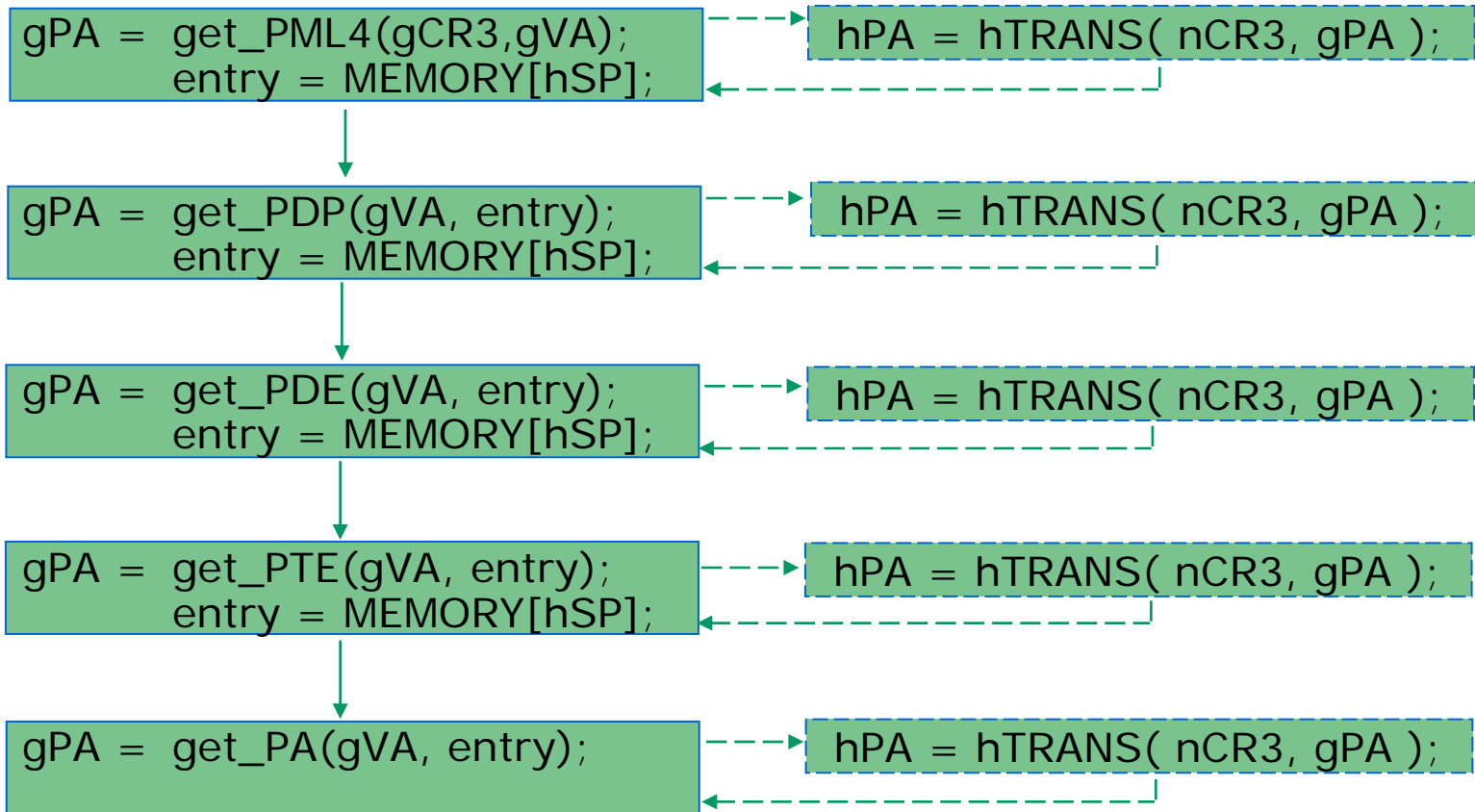
Translates a page in guest virtual address space to machine physical space through a two-level translation

- First, map guest virtual address to guest physical address
- Then, map guest physical address to machine physical address



Nested Page Table Walk

CPU builds guest virtual to system physical TLB entries (guarded by ASID)



How to Enable Nested Paging

- Nested paging is detected via CPUID Fn8000_000A EDX, bit 0
- Constructs nested paging table (P2M table in Xen)
- VMM constructs VMCB
 - Sets np_enable bit = 1
 - Disables intercepting guest page faults
 - Sets n_CR3 to point to the base of the nested page table
 - Disables intercepting CRx registers READ/WRITE (optional)
- Nested Paging is an AMD-V™ feature to be enabled on Barcelona (Family 10h processors)

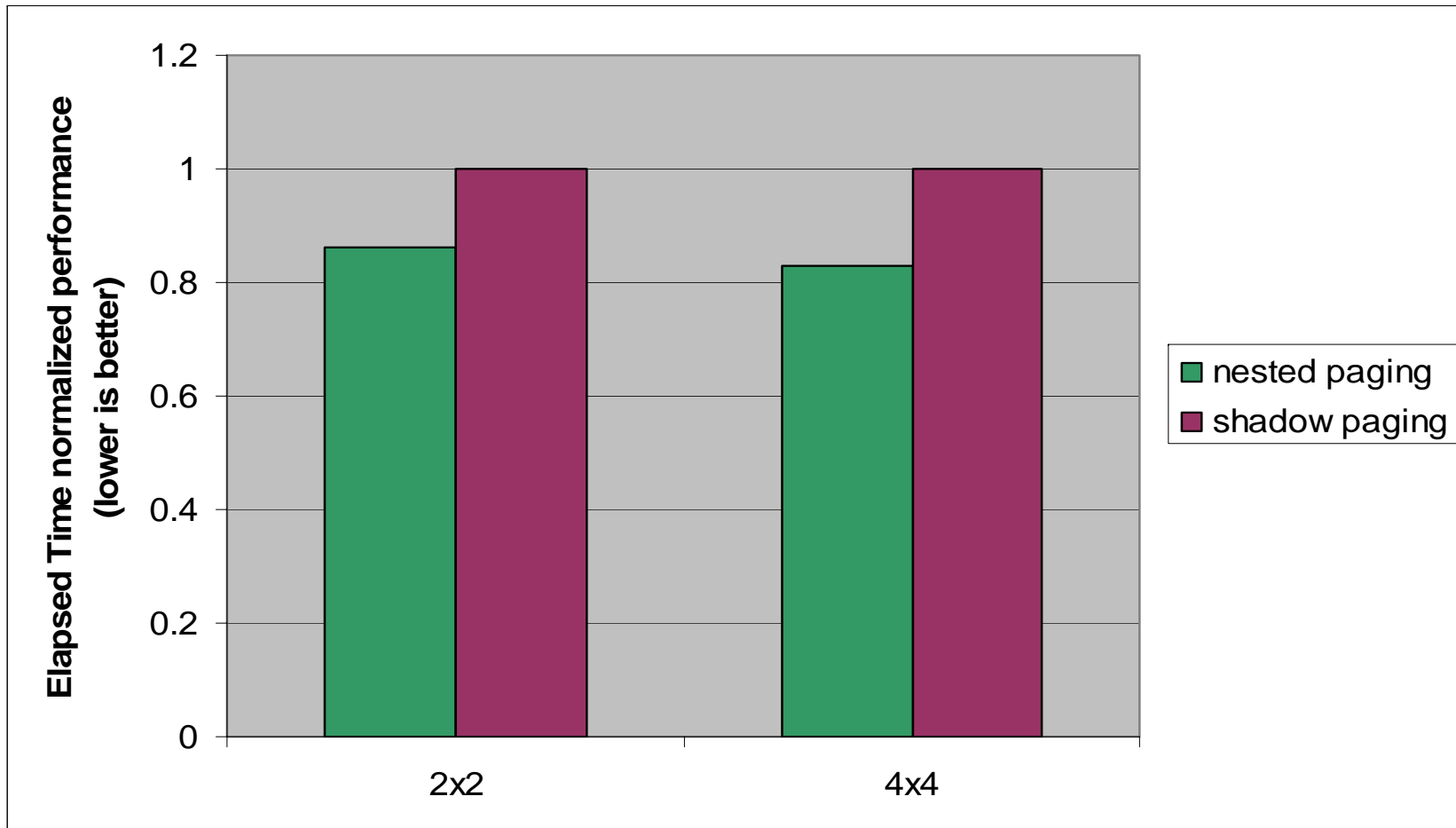
Page Faults in Nested Paging

- Guest page faults
 - Hypervisor does not need to intercept guest page faults
 - CPU can inject guest page faults automatically (no VMEXIT)
- Nested page faults
 - Behave like normal page faults
 - Two sources
 - Normal page faults, which should rarely happen
 - MMIO space

Experiment Environment

- Test cases
 - 2x2 (32bit guest on top of 32bit hypervisor)
 - 4x4 (64bit guest on top of 64bit hypervisor)
- Hypervisor
 - Xen with changeset 14679
- Guest OS
 - SUSE10 with 1GB memory, UP, all VCPUs pinned
- Benchmark
 - Kernbench (compiling Linux 2.6.16.14)
- Procedure
 - Benchmarks runs 11 times
 - Elapsed time of benchmark as measured in guest is used

Kernbench



Disclaimers:

1. *Experimental AMD processor with nested paging running experimental Xen builds. Results might vary.*
2. *Among best case improvement for nested paging, which mainly helps memory-management intensive workloads; not representative of all workloads.*

Status of Xen Nested Paging

- Supported in Xen 3.0.5
 - Shares interface with shadow paging
 - Enabled by “hap” option in Xen boot menu
 - Supports all guest/host paging modes
 - Nested paging domain save/restore
- Source code in Xen
 - xen/arch/x86/mm/hap/
- Next phase (in progress)
 - Nested paging domain live migration

Documentation

- AMD64 Architecture Tech Docs Volume 2

http://www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/24593.pdf

- AMD64 Architecture Tech Docs Volume 3

http://www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/24594.pdf

Trademark Attribution

AMD, the AMD Arrow logo, AMD-V and AMD Virtualization, and combinations thereof are trademarks of Advanced Micro Devices, Inc. in the United States and/or other jurisdictions. Other names used in this presentation are for identification purposes only and may be trademarks of their respective owners.

©2007 Advanced Micro Devices, Inc. All rights reserved.