



# Enhancing KVM/IA64

Xiantao Zhang / Anthony Xu

[Xiantao.zhang@intel.com](mailto:Xiantao.zhang@intel.com) [Anthony.Xu@intel.com](mailto:Anthony.Xu@intel.com)

Intel Open Source Technology Center



# Agenda

Architecture Review

Status

New IA-64 Hardware Feature

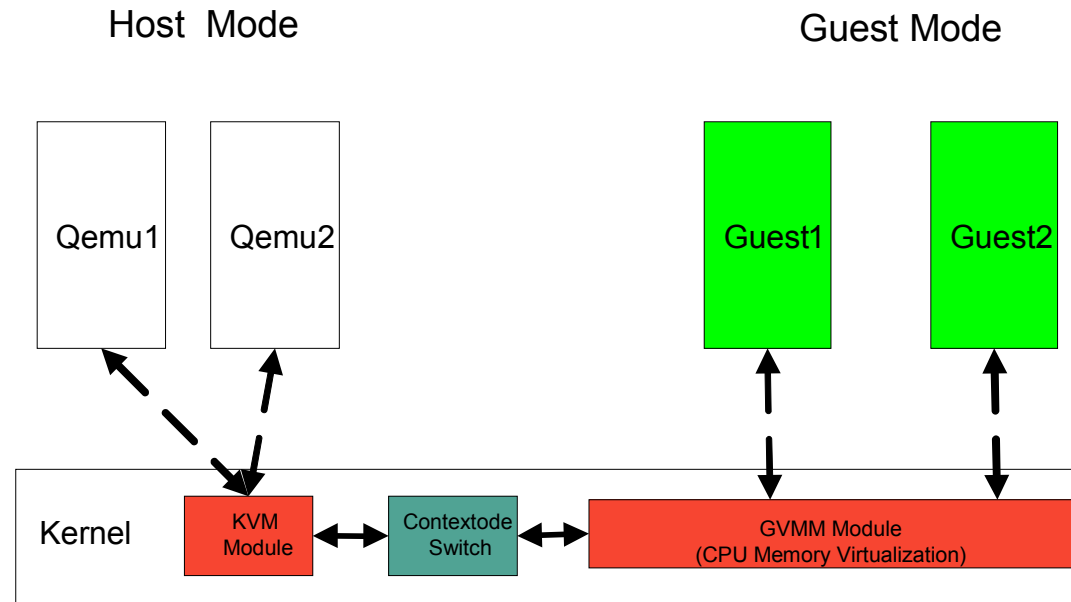
Future Enhancement

Current Performance

Call to Action



# Architecture Review



## Context Switch

- Switch host and guest contexts
- Through well-defined exit/entry interfaces
- Lightweight/heavyweight exits

# Status

Merged into Linux-2.6.26 Development tree

- Subscribe the mailing list through <http://vger.kernel.org/vger-lists.html#kvm-ia64>

Support Guests

- Linux SMP Guests
- Windows 2003 Server SMP Guests

Comparable performance with XEN/IA-64

Guest Firmware

- Using Same Open Source Guest firmware as Xen/IA-64 side
- <http://xenbits.xensource.com/ext/efi-vfirmware.hg>



# Status – Cont.

## Save/Restore

- Done

## Live Migration

- A temporary solution w/o incremental memory migration

## PV Driver for virtual I/O Devices

- Not supported

## Host swapping

- Not supported

## Large Page support

- Not supported



# New IA-64 Hardware Feature

## VTi2

- Available in Tukwila
- Some frequent instructions are executed directly by the processor
  - Move from interrupt control register
  - Disable interrupt
  - Enable interrupt unless virtual interrupt can deliver
  - Read TPR
  - Write TPR unless virtual interrupt can deliver

# New IA-64 Hardware Feature

## Multiple Global TLB purges

- Single Global TLB purge currently
  - SW guarantees only one global purge instruction is executed
    - Spin lock among CPUs
  - Heavyweight exit for Global TLB purge
- Tukwila support multiple Global TLB purges
  - Capable of executing multiple global purge instruction concurrently
  - Max support number comes from PAL\_VM\_SUMMARY

# KVM/IA-64 Enhancement

Large page support

- Used extensively
- More performance gain compare to IA32
  - Larger page size up to 256M
  - No hardware page table



# Future Enhancement

## Previous Lightweight Path

- Switch backing store
- Save guest context to memory stack

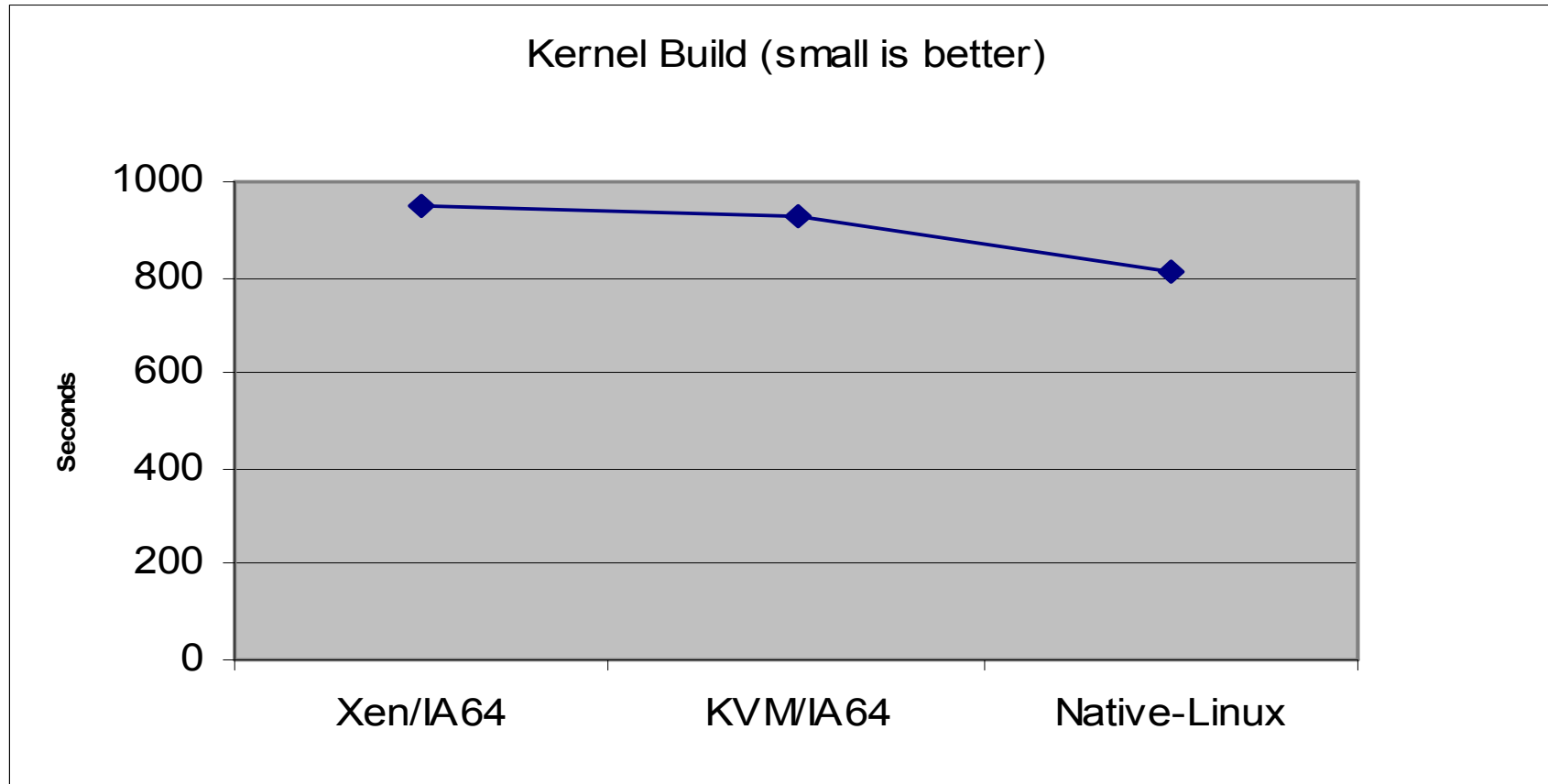
## Proposed Lightweight Path

- Switch backing store
- Save guest context to register stack

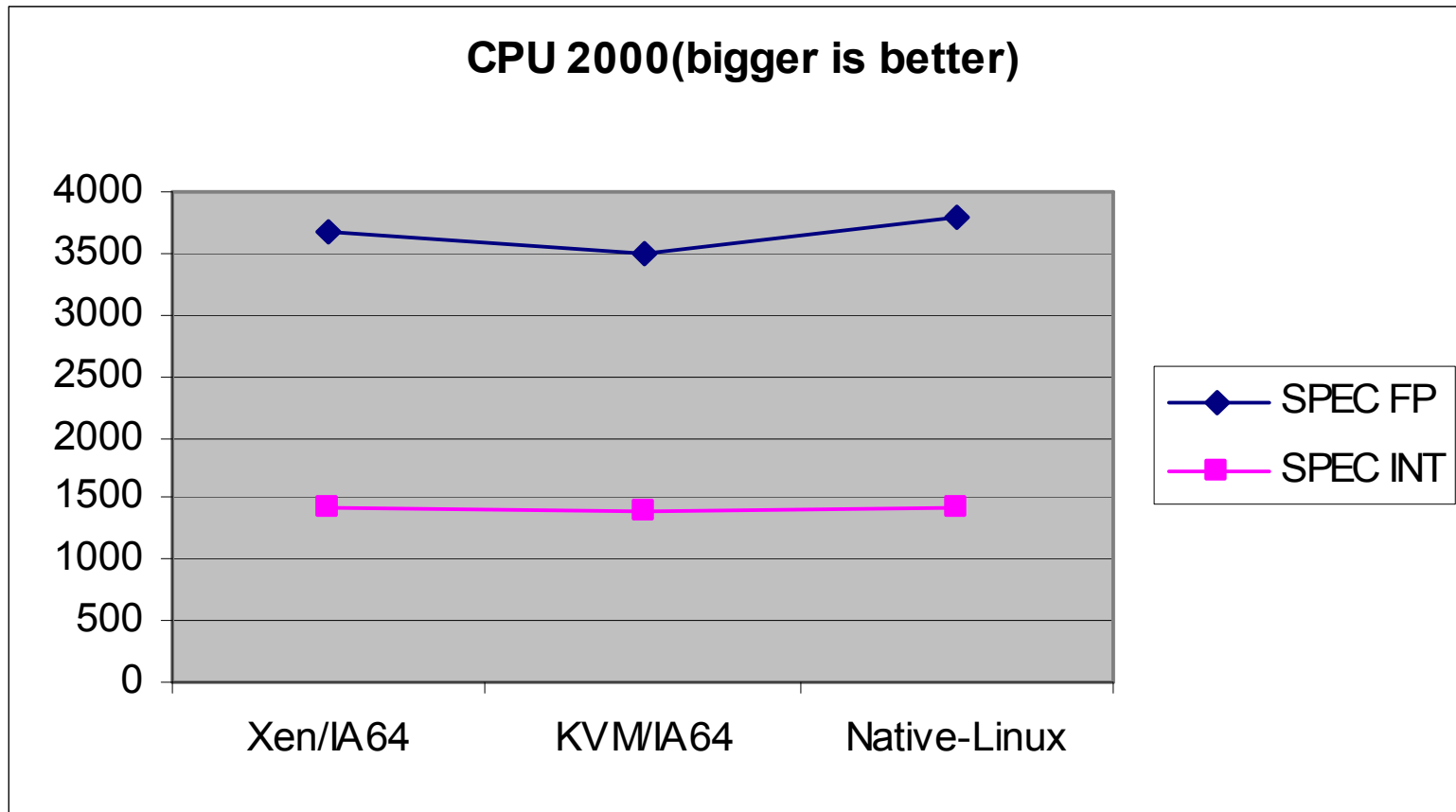
## Pro

- Take advantage of RSE
- Less memory access

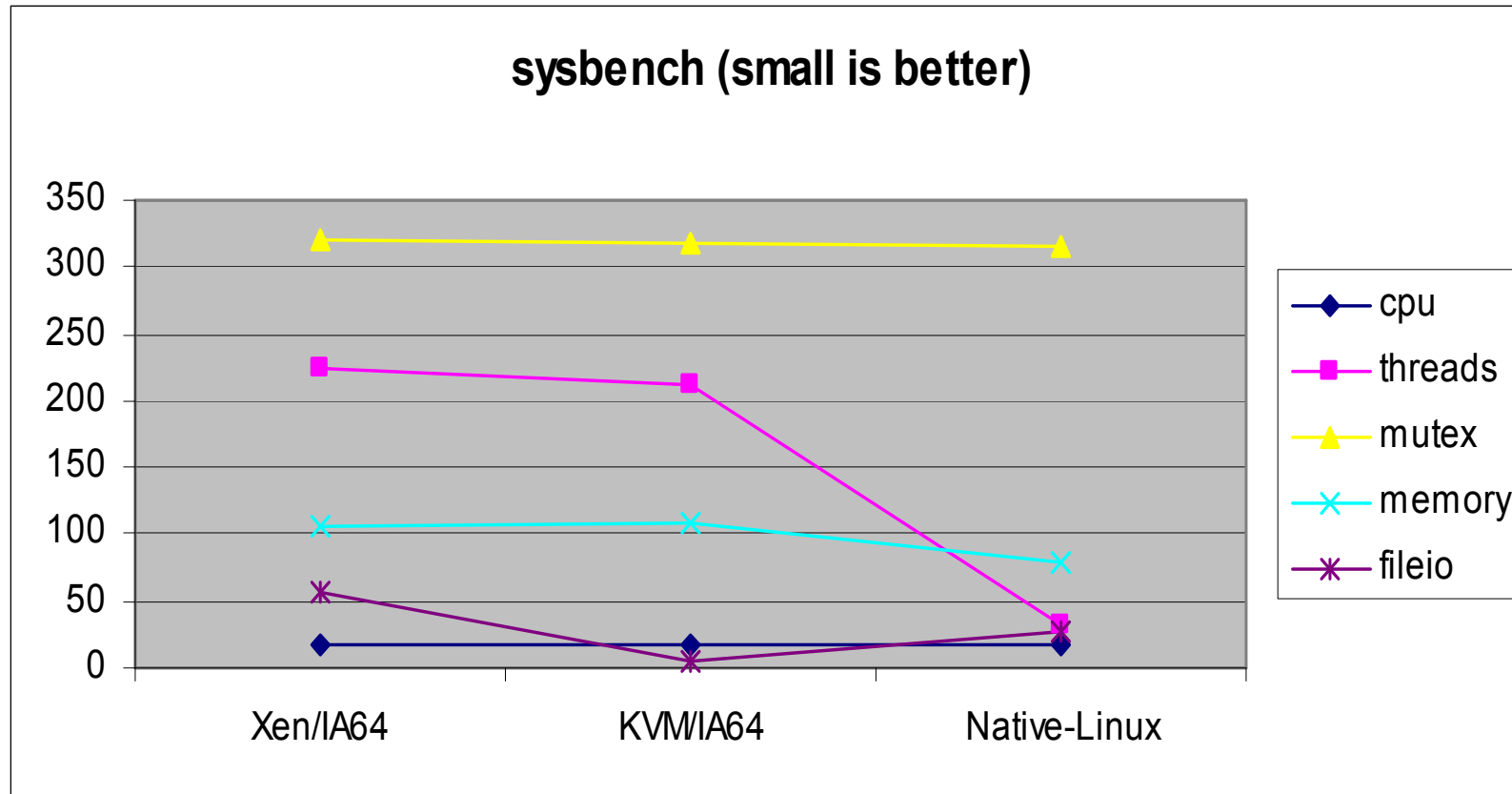
# Performance Comparison



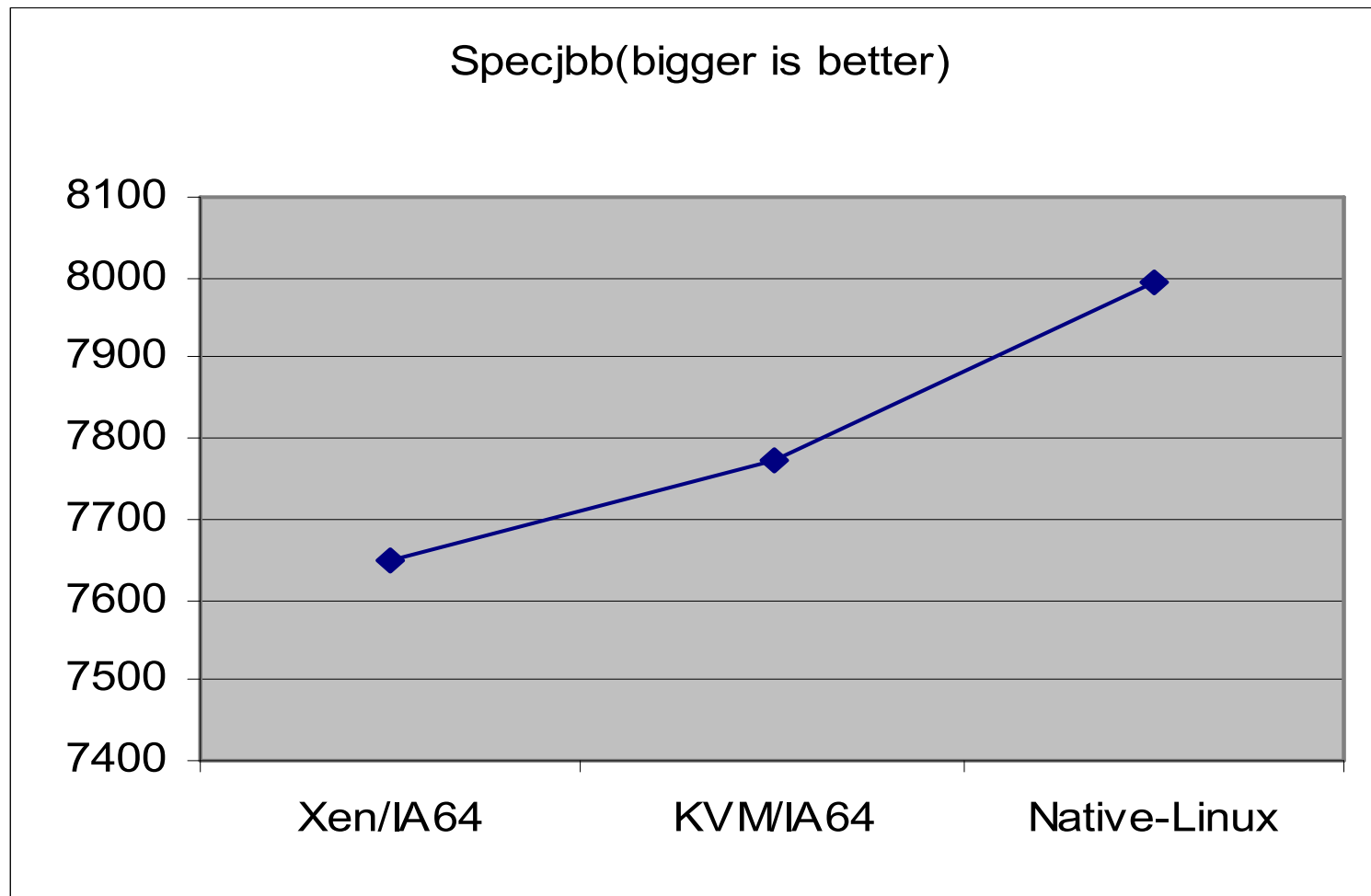
# Performance Comparison – Cont.



## Performance Comparison – Cont. 2



## Performance Comparison – Cont. 3



# Performance Comparison Summary

Good IO performance

- Short IO handler path

Similar CPU/Memory performance

- Almost same implementation



# Call to Action

Enabling live migration and save restore

PV driver for guests

VT-d Support

- Depend on X86 side

VT-i 2 Support

- Depend on the release of VT-i 2 processors

Kvm trace

# Legal Information

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY RELATING TO SALE AND/OR USE OF INTEL PRODUCTS, INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT, OR OTHER INTELLECTUAL PROPERTY RIGHT.

Intel may make changes to specifications, product descriptions, and plans at any time, without notice.

All dates provided are subject to change without notice.

Intel is a trademark of Intel Corporation in the U.S. and other countries.

\*Other names and brands may be claimed as the property of others.

Copyright © 2008, Intel Corporation. All rights are protected.





