Xen-PPC

Hollis Blanchard, Christian Ehrhardt, Jerone Young, Jimi Xenidis
Overview

- Current Status
- Porting Difficulties
- Porting Issues
- How you can help?
Current Status

- **Ready for 3.0.5**
  - Virtual IO
  - SMP Xen and Domains
    - <= 4 CPUs box
    - <= 32 VCPUs per domain
  - XM and XM Test

- **(well sort of)**
  - Still running older, separate Linux
  - Save and resume in unreleased Linux Kernel
  - No Ballooning
  - No live Migration
  - Oprofile soon
Porting Difficulties -- Real Mode Area

- Domain takes exceptions in real mode
- 970: RMA must be minimum 64MiB in size and must be naturally aligned
- Dictates minimum domain size
- Special pages (shared, console) are placed in this region so they are available while xlate is off
- Cannot balloon this area… ever!
- Same problems as large page mappings…
Porting Difficulties -- Large Pages

- **Kernel linear map always uses largest page size**
  - Performance gain is too much to pass up
  - Mapped “pinned” at startup.
- **970: Capable of 16MiB large page size**
- **Makes ballooning difficult (if not impossible)**
- **Grant Mappings are always 4K**
  - All IO requires real page descriptors
  - Use memory hot-plug to define a “foreign” memory area
  - Manage area separately in Xen specific paths
  - Grants manage foreign PFN->MFN
  - Pre-grant mappings and Post-grant unmappings

http://wiki.xensource.com/xenwiki/XenPPC

4/17/07 © 2007 IBM Corporation
Porting Difficulties -- Auto Translate

- Always on
- PTE hcalls are (VFN to PFN)
- Xen responsible for PFN to MFN
- hcalls referring to MFN directly are flagged
  - MFN mappings are for dom0
- Always requires IOMMU hcalls
  - Even Dom0 does not know PFN to MFN
Porting Difficulties -- start_info

- Uses “flattened devtree”, from Open Firmware
- /xen node indicates you are in a Xen VM
  - Node contains properties describing relevant start_info like information
    - Console device
    - Special pages (shared, console, foreign)
- Other standard nodes and properties used for and machine specific stuff
  - CPUs
  - Memory
  - Host bridges
Porting Issues

- **Atomic operations**
  - Only performed on 32-bit and 64-bit quantities
    - Bitmaps are long arrays in Linux:
      - Even on x86, but !x86_64 :-(
      - Different in XM (32-bit) than in Kernel and Xen
    - Cannot cmpxchg() and xchg() a char or short
      - Requires “align down”
    - Load-Link/Store-Conditional
      - As do Alpha, MIPS, and ARM (are we misfits!?)
      - Vs. Compare and Swap

- **Xencomm**
  - Used by IA64 but forked, needs to “commonize” again
How can you help? Always….

- Use DECLARE_BITMAP() for bit operations
- Use atomic_t counts (even for boolean)
- Stricter interfaces to restrict with 8/16 bit atomics
- Consider Endian, don’t pack manually
- Consider type safety
- Consider alignment
- Avoid using __packed__ (add -Wpacked)
- [http://kegel.com/crosstool/](http://kegel.com/crosstool/)
  - Run it and gain cross compilers to find easy issues
- I use x86 :-)

http://wiki.xensource.com/xenwiki/XenPPC