



Xen Linux merge status

Chris Wright

chrisw@redhat.com

Xen Summit 2006, San Jose

Xen Linux Merge Status

- Overview
- Patchset details
- Future work

Overview

- Xen Linux sparse tree
 - Contains many copied files: `arch/i386/kernel/setup.c` and `arch/i386/kernel/setup-xen.c`
 - Lags behind mainline for stability reasons
- Move to subarch: `arch/xen/i386/` -> `arch/i386/mach-xen`
- 3.0.2 release (April 2006) had 2.6.16 base with subarch
- VMI patches posted to lkml
- Xen patches posted to lkml
- Kernel Summit 2006: move to `paravirt_ops`

Xen Linux Patches

- Scope of work
 - i386 only
 - UP only
 - domU only
 - shadow mode only
 - Limited scope reduces size, complexity, and invasiveness of the patchset.
- Community response
 - Useful feedback for improving the patchset that has resulted in cleanups which are being propagated back to the xen-unstable development tree
 - Some small bits have been taken by Andrew for upstream Linux

Xen Linux – Patchset details

- Adds core Xen functionality for:
 - Hypervisor interface
 - Time
 - Reboot
 - Event channels
 - Grant tables
 - Xenbus
 - Console
 - Frontend block and net drivers

XenoLinux Merge Status – Patchset details

- ~50 patches, ~268KB (PV) + ~512KB (Xen)
- (PV) 56 files changed, 1895 insertions(+), 348 deletions(-)
- (Xen) 76 files changed, 10274 insertions(+), 85 deletions(-)
- Creates new i386 struct `paravirt_ops`: `arch/i386/kernel/paravirt.c`
- Updates infrastructure to allow a `paravirt_ops` to override default behaviour for:
 - Start-of-day
 - Segments (running in ring 1)
 - Descriptor table handling: GDT, LDT, IDT
 - Control register handling: CR0, CR1, CR2, CR3, CR4
 - CPUID
 - Interrupt handling
 - TLB handling
 - Memory and page table handling

Xen Linux: paravirt_ops

```
struct paravirt_ops
{
    void (*arch_setup)(void);
    char *(*memory_setup)(void);
    ...
    void (fastcall *cpuid)(unsigned int *eax, unsigned int
        *ebx, unsigned int *ecx, unsigned int *edx);
    void (fastcall *get_debugreg)(int regno);
    unsigned int (fastcall *read_cr2)(void);
    void (fastcall *write_cr3)(unsigned int);
    void (fastcall *irq_disable)(void);
    void (fastcall *irq_enable)(void);
    ...
};
```

Xen Linux: `paravirt_ops`

- Boot entry point for paravirt
- Hypervisor specific probe functions, defaults back to native
- Set up hypervisor specific `paravirt_ops`
- Patch call sites

XenLinux Merge Status – Future Work

- SMP support
- Writable page tables support
- dom0 support
- Other architectures (x86_64, ia64, PPC)