

Fastboot and kexec  
Linux Kernel Summit 2004  
Randy Dunlap, OSDL  
2004-July-20

- Started by Eric Biederman
- Co-maintained since 2003-OCT
- <http://www.xmission.com/~ebiederm/files/kexec/>
- <Http://developer.osdl.org/rddunlap/kexec/>

# General fastboot areas

- Bypass firmware on reboot
- Parallel driver init and device probe/discovery
- Use known-devices database unless reprobe is forced (on-demand)
- Parallel init scripts (Gooch or Midori Linux)
  - <http://www.atnf.csiro.au/people/rgooch/linux/boot-scripts/>
  - <http://sourceforge.net/projects/midori/> (CVS)
- Early userspace changes....

# CE Linux Forum Bootup-time WG

- OLS paper (Tim Bird): “Methods to Improve Bootup Time in Linux”
- Kernel XIP, “partial XIP” (hotspots in RAM)
- kernel init: reduced probing; configuration cache
- Concurrent init of some subsystems
- Fast shutdown & resume (ties to Power Mgt.)
- Timing instrumentation
- Userspace startup time reduction (scripts, app. Pre-loading, pre-linking, app. XIP)
- Now sending some patches for kernel merging

# What is kexec?

- Linux-reboots-linux patch (Eric Biederman)
- Bypasses firmware on reboot
- Alternative to replacing firmware (LinuxBIOS)
- Whitepaper on boot times:  
<http://developer.osdl.org/rddunlap/kexec/whitepaper/>

# Kexec API

- Discussed & agreed to in 2003 on lkml
- Syscall interface merged in 2.6.6
- Load: `asmlinkage long sys_kexec_load(  
void *entry,  
unsigned long nr_segments,  
struct kexec_segment *segments,  
unsigned long flags);`
- Reboot: use `sys_reboot()` with `cmd == LINUX_REBOOT_CMD_KEXEC`

# Userspace tools

- X86, with PPC32/GameCube and PPC64 added recently
- Kexec loader: `kexec -l /boot/bzImage-2.6.8-rc1`
- Kexec reboot: `kexec -e`
- Sysvinit patch: `reboot -k`

# Kexec ports for 2.6.x

- IA-32 working (relatively small patch)
- X86-64 (WIP, Eric B.)
- GameCube working (Albert Herranz) (PPC32 needs OF work)
- PPC64 (WIP, IBM; OF issues)
- IA-64 (preliminary port posted, no active work)
- Some current patch pieces are independent & will be pushed

# Kexec mods for RAS (lkcd; IBM)

- Preserve some kernel memory across a kexec reboot
- Makes the old kernel's memory accessible as a device from the new kernel (/dev/hmem)
- Can do any kind of first failure data gathering from the old kernel's memory

# Kexec problem areas

- APIC and IO-APIC shutdown (hangs)
- VM changes (some recent fixes by Hugh via Martin)
- Races in cpumask operations during shutdown/reboot (fixed around 2.6.3)
- Use in `init_mm` on PPC64 (or any non-x86)
- Firmware/kernel interactions & dependencies (PPC64 now being addressed)
- Global data structures at certain addresses (PPC64)