Multiple Passes of the FreeBSD Device Tree

May 9, 2009

John Baldwin
jhb@FreeBSD.org
Overview

- FreeBSD's Existing Device Driver Framework
- Some Tricky Problems
- Adding Multiple Passes
- Problems Revisited
“new-bus”

- Device Tree
- Dynamic
- Single-pass Depth-first Traversals
Tricky Problems

- Low-level Hardware
- Resource Discovery and Allocation
- Boot vs Non-Boot Probing
Low-level Hardware

- Interrupt Controllers, Timers, etc. are Devices, Too
- Probed Too Late
- Not Always Managed Via “new-bus”
- ACPI Hacks to Enforce Ordering
Resource Discovery and Allocation

- Reserve All Fixed Resources Before Assigning Dynamic Resources
- PCI-PCI Bridge Problem

- PCI-PCI Bridge
  - 0xe000 - 0xe3ff

- Device C
  - 0xe400 - 0xe440

- Device A
  - 0xe000 - 0xe3ff

- Device B
  - N - N+0x100
Boot vs Non-Boot Probing

- Boot Probes Run Without Full Scheduler
- `config_intrhook(9)` Helps Some
- Scheduler Needs Timeout Which Needs Timer
Common Themes

- Devices Have Prerequisites
- Not All Are Consistent With Depth-First Traversal
- How to Fix?
Multiple Passes

- Multiple Scans of the Device Tree
- “Early” Drivers Probe During Initial Scans
- Most Drivers Probe Final Pass
Possible Passes

- **BUS_PASS_ROOT** – marker for root0
- **BUS_PASS_BUS** – populate the tree
- **BUS_PASS_CPU** – add CPU devices
- **BUS_PASS_RESOURCE** – resource discovery
- **BUS_PASS_INTERRUPT** – PICs
- **BUS_PASS_TIMER**
- **BUS_PASS_SCHEDULER**
- **BUS_PASS_DEFAULT**
Implementation Details

- Driver Attachments Have Pass Level
- **New** `BUS_NEW_PASS()` Bus Interface Method
- `bus_generic_new_pass()`
- Pass Levels Are Sparse and Dynamic – like `SYSINIT()` subsystems
- `bus_set_pass()` Raises Pass Level
Writing an “Early” Driver

- EARLY_DRIVER_MODULE()
- Device Attach Routines May Need to Be Revised
- Attach Routines May Need to Handle Boot vs Non-Boot
Problems Revisited

- Low-level Hardware Drivers
- Resource Discovery and Management
  - Perform Actions After Pass Traversal
  - Hooking Passes vs. New Bus Methods
- Boot vs Non-Boot Probing
- Suspend and Resume
  - *Not In The Paper*
  - Similar Prequisite and Ordering Problems
Conclusion

- Code Currently Lives in FreeBSD p4 in
  //depot/projects/multipass/...
- Plan to Commit Framework to 8.0
- Future Work in 9.0
Q&A

- Questions?