PCI Interrupts for x86 Machines under FreeBSD

May 18, 2007

John Baldwin
jhb@FreeBSD.org
Introduction

• Hardware for PCI INTx interrupts
  – x86 CPU interrupts
  – PCI INTx signals
  – x86 interrupt controllers

• Interrupt Routing
  – PCI-PCI bridge swizzle
  – Tables: $PIR, MP Table, ACPI _PRT

• Message Signaled Interrupts
x86 CPU Interrupts

- OS provides array of handlers called Interrupt Descriptor Table (IDT)
- Each interrupt contains vector which indexes table
- Vectors 0-31 reserved for Faults/Exceptions
- Vectors 32-255 available for device interrupts, IPIs, etc.
PCI INTx

- PCI slots have 4 interrupts
  - INTA#, INTB#, INTC#, INTD#
- PCI functions use 1 interrupt from parent slot
x86 Interrupt Controllers

- 8259A PICs (PC-AT)
- Programmable Interrupt Router
- APICs
Interrupt Routing

- How everything is hooked together
- Maps PCI (bus, slot, pin) to input pin on 8259A or I/O APIC
  - May have to detour through a PIR
- PCI-PCI bridge swizzle
- Various tables on x86
  - $PIR for 8259A & PIR
  - MP Table for APIC
  - ACPI for both
PCI-PCI Bridge Swizzle

- Routes interrupts across bridge from “downstream” bus to “upstream” bus
- For PCI-PCI bridges in add-on cards
- new_pin = (child_slot + child_pin) % 4
## $PIR Table

<table>
<thead>
<tr>
<th>Entry</th>
<th>Location</th>
<th>Bus</th>
<th>Device</th>
<th>Pin</th>
<th>Link</th>
<th>IRQs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>embedded</td>
<td>0</td>
<td>2</td>
<td>A</td>
<td>0x60</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>1</td>
<td>embedded</td>
<td>1</td>
<td>0</td>
<td>A</td>
<td>0x60</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>1</td>
<td>embedded</td>
<td>1</td>
<td>0</td>
<td>B</td>
<td>0x61</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>1</td>
<td>embedded</td>
<td>1</td>
<td>0</td>
<td>C</td>
<td>0x62</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>1</td>
<td>embedded</td>
<td>1</td>
<td>0</td>
<td>D</td>
<td>0x63</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>2</td>
<td>embedded</td>
<td>0</td>
<td>29</td>
<td>A</td>
<td>0x60</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>2</td>
<td>embedded</td>
<td>0</td>
<td>29</td>
<td>B</td>
<td>0x63</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>2</td>
<td>embedded</td>
<td>0</td>
<td>29</td>
<td>C</td>
<td>0x62</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>2</td>
<td>embedded</td>
<td>0</td>
<td>29</td>
<td>D</td>
<td>0x6b</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>3</td>
<td>embedded</td>
<td>0</td>
<td>31</td>
<td>A</td>
<td>0x62</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>4</td>
<td>embedded</td>
<td>4</td>
<td>13</td>
<td>A</td>
<td>0x61</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>5</td>
<td>embedded</td>
<td>2</td>
<td>4</td>
<td>A</td>
<td>0x60</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>6</td>
<td>embedded</td>
<td>4</td>
<td>3</td>
<td>A</td>
<td>0x68</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>7</td>
<td>slot 1</td>
<td>3</td>
<td>7</td>
<td>A</td>
<td>0x62</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>7</td>
<td>slot 1</td>
<td>3</td>
<td>7</td>
<td>B</td>
<td>0x63</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>7</td>
<td>slot 1</td>
<td>3</td>
<td>7</td>
<td>C</td>
<td>0x60</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
<tr>
<td>7</td>
<td>slot 1</td>
<td>3</td>
<td>7</td>
<td>D</td>
<td>0x61</td>
<td>3 4 5 6 10 11 14 15</td>
</tr>
</tbody>
</table>
## MP Table

<table>
<thead>
<tr>
<th>I/O Ints:</th>
<th>Type</th>
<th>Polarity</th>
<th>Trigger</th>
<th>Bus ID</th>
<th>IRQ</th>
<th>APIC ID</th>
<th>PIN#</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExtINT</td>
<td>active-hi</td>
<td>edge</td>
<td></td>
<td>5</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>5</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>5</td>
<td>0</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>5</td>
<td>9</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>5</td>
<td>12</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>0</td>
<td>2:A</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>0</td>
<td>29:A</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>0</td>
<td>29:B</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>0</td>
<td>29:D</td>
<td>8</td>
<td>23</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>0</td>
<td>31:A</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>4</td>
<td>13:A</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>4</td>
<td>3:A</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>2</td>
<td>4:A</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>3</td>
<td>7:A</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>3</td>
<td>7:B</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>3</td>
<td>7:C</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>INT</td>
<td>conforms</td>
<td>conforms</td>
<td></td>
<td>3</td>
<td>7:D</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>
ACPI

- Global System Interrupts
  - 0-15 are ISA IRQs
  - 16+ are I/O APIC pins

- OS calls _PIC method to tell BIOS 8259A vs APIC

- Each PCI bus contains _PRT table

- Pins on Programmable Interrupt Router are Link Devices
<table>
<thead>
<tr>
<th>Dual 8259As</th>
<th>APIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>15</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>72</td>
</tr>
</tbody>
</table>
ACPI_PIC Method

Scope (\)
{
    Name (PICF, 0x00)
    Method (_PIC, 1, NotSerialized)
    {
        Store (Arg0, PICF)
    }
}
Device (PXHB)
{
...
Name (PIC3, Package (0x04))
{
  Package (0x04)
  {
    0x0007FFFF,
    0x00,
    LNKC,
    0x00
  },
...
}
Name (APC3, Package (0x04))
{
  Package (0x04)
  {
    0x0007FFFF,
    0x00,
    0x00,
    0x42
  },
  ...
}

Method (_PRT, 0, NotSerialized)
{
  If (LNot (PICF))
  {
    Store (PIC3, Local0)
  }
  Else
  {
    Store (APC3, Local0)
  }
  Return (Local0)
}
FreeBSD's PCI Interrupt Routing

- PCI bus asks parent bridge to route interrupt
- Requests propagate up device tree until answered
Mapping IRQs to IDT Vectors

- Interrupt routing routes PCI interrupt to interrupt controller pin (IRQ)
- FreeBSD uses GSI-like scheme to map interrupt controller pins to IRQs
- Allocate IDT vectors on-demand to IRQs
Message Signaled Interrupts

- Problems with INTx Interrupts
  - Single interrupt per function
  - Interrupt is a sideband signal
  - Interrupt routing is a pain

- MSI Addresses These Problems
  - Each function can have multiple messages
  - Interrupt triggered via memory write
  - Memory data contains IDT vector on x86
INTx Sideband Signals are Bad

CPU

PIC

Host-PCI Bridge

PCI-PCI Bridge

PCI Slot

RAM
FreeBSD's MSI Implementation

- PCI bus asks parent bridge for MSI IRQs
- x86 “nexus” driver creates MSI IRQs on the fly and assigns IDT vectors on-demand
Conclusion

- PCI INTx interrupts are messy
  - Interrupt Routing can be painful, esp. on x86
- MSI will hopefully be better
  - PCI-express mandates it for PCI-express devices
Q&A

- Paper and slides are available online
- Questions?