IRQ affinity on IA64 platforms

07.01.2002, Erich Focht <<efocht@ess.nec.de>>

By writing to /proc/irq/IRQ#/smp_affinity the interrupt routing can be controlled. The behavior on IA64 platforms is slightly different from that described in Documentation/core-api/irq/irq-affinity.rst for i386 systems.

Because of the usage of SAPIC mode and physical destination mode the IRQ target is one particular CPU and cannot be a mask of several CPUs. Only the first non-zero bit is taken into account.

Usage examples

The target CPU has to be specified as a hexadecimal CPU mask. The first non-zero bit is the selected CPU. This format has been kept for compatibility reasons with i386.

Set the delivery mode of interrupt 41 to fixed and route the interrupts to CPU #3 (logical CPU number) (2^3=0x08):

```
echo "8" >/proc/irq/41/smp_affinity
```

Set the default route for IRQ number 41 to CPU 6 in lowest priority delivery mode (redirectable):

```
echo "r 40" >/proc/irq/41/smp_affinity
```

The output of the command:

```
cat /proc/irq/IRQ#/smp_affinity
```

gives the target CPU mask for the specified interrupt vector. If the CPU mask is preceded by the character "r", the interrupt is redirectable (i.e. lowest priority mode routing is used), otherwise its route is fixed.

Initialization and default behavior

If the platform features IRQ redirection (info provided by SAL) all IO-SAPIC interrupts are initialized with CPU#0 as their default target and the routing is the so called "lowest priority mode" (actually fixed SAPIC mode with hint). The XTP chipset registers are used as hints for the IRQ routing. Currently in Linux XTP registers can have three values:

- minimal for an idle task,
- normal if any other task runs,
- maximal if the CPU is going to be switched off.

The IRQ is routed to the CPU with lowest XTP register value, the search begins at the default CPU. Therefore most of the interrupts will be handled by CPU #0.

If the platform doesn't feature interrupt redirection IOSAPIC fixed routing is used. The target CPUs are distributed in a round robin manner. IRQs will be routed only to the selected target CPUs. Check with:

```
cat /proc/interrupts
```

Comments

On large (multi-node) systems it is recommended to route the IRQs to the node to which the corresponding device is connected. For systems like the NEC AzusA we get IRQ node-affinity for free. This is because usually the chipsets on each node redirect the interrupts only to their own CPUs (as they cannot see the XTP registers on the other nodes).