SMP IRQ affinity

ChangeLog:

- Started by Ingo Molnar <mingo@redhat.com>
- Update by Max Krasnyansky <maxk@qualcomm.com>

/proc/irq/IRQ#/smp_affinity and /proc/irq/IRQ#/smp_affinity_list specify which target CPUs are permitted for a given IRQ source. It's a bitmask (smp_affinity) or cpu list (smp_affinity_list) of allowed CPUs. It's not allowed to turn off all CPUs, and if an IRQ controller does not support IRQ affinity then the value will not change from the default of all cpus.

/proc/irq/default_smp_affinity specifies default affinity mask that applies to all non-active IRQs. Once IRQ is allocated/activated its affinity bitmask will be set to the default mask. It can then be changed as described above. Default mask is 0xffffffff.

Here is an example of restricting IRQ44 (eth1) to CPU0-3 then restricting it to CPU4-7 (this is an 8-CPU SMP box):

```
[root@moon 44]# cd /proc/irq/44
[root@moon 44]# cat smp_affinity
ffffffff

[root@moon 44]# echo 0f > smp_affinity
[root@moon 44]# cat smp_affinity
0000000f
[root@moon 44]# ping -f h
PING hell (195.4.7.3): 56 data bytes
...
--- hell ping statistics ---
6029 packets transmitted, 6027 packets received, 0% packet loss
round-trip min/avg/max = 0.1/0.1/0.4 ms
[root@moon 44]# cat /proc/interrupts | grep 'CPU\|44:'
        CPU0       CPU1       CPU2       CPU3       CPU4       CPU5       CPU6       CPU7
44:       1068       1785       1785       1783          0          0          0          0    IO-APIC-level
```

As can be seen from the line above IRQ44 was delivered only to the first four processors (0-3). Now lets restrict that IRQ to CPU(4-7).

```
[root@moon 44]# echo f0 > smp_affinity
[root@moon 44]# cat smp_affinity
000000f0
[root@moon 44]# ping -f h
PING hell (195.4.7.3): 56 data bytes
..
--- hell ping statistics ---
2779 packets transmitted, 2777 packets received, 0% packet loss
round-trip min/avg/max = 0.1/0.5/585.4 ms
[root@moon 44]# cat /proc/interrupts |  'CPU\|44:'
        CPU0       CPU1       CPU2       CPU3       CPU4       CPU5       CPU6       CPU7
44:       1068       1785       1785       1783       1784       1069       1070       1069    IO-APIC-level
```

This time around IRQ44 was delivered only to the last four processors. i.e counters for the CPU0-3 did not change.

Here is an example of limiting that same irq (44) to cpus 1024 to 1031:

```
[root@moon 44]# echo 1024-1031 > smp_affinity_list
[root@moon 44]# cat smp_affinity_list
1024-1031
```

Note that to do this with a bitmask would require 32 bitmasks of zero to follow the pertinent one.