General Description of sysfs CPUFreq Stats

information for users

Author: Venkatesh Pallipadi <venkatesh.pallipadi@intel.com>

1. Introduction

cpufreq-stats is a driver that provides CPU frequency statistics for each CPU. These statistics are provided in /sysfs as a bunch of read_only interfaces. This interface (when configured) will appear in a separate directory under cpufreq in /sysfs (<sysfs root>/devices/system/cpu/cpuX/cpufreq/stats/) for each CPU. Various statistics will form read_only files under this directory.

This driver is designed to be independent of any particular cpufreq_driver that may be running on your CPU. So, it will work with any cpufreq_driver.

2. Statistics Provided (with example)

cpufreq stats provides following statistics (explained in detail below).

- time_in_state
- total_trans
- trans_table

All the statistics will be from the time the stats driver has been inserted (or the time the stats were reset) to the time when a read of a particular statistic is done. Obviously, stats driver will not have any information about the frequency transitions before the stats driver insertion.

```
<mysystem>:/sys/devices/system/cpu/cpu0/cpufreq/stats # ls -l
total 0
drwxr-xr-x  2 root root    0 May 14 16:06 .
drwxr-xr-x  3 root root    0 May 14 15:58 ..
--w-------  1 root root 4096 May 14 16:06 reset
-r--r--r--  1 root root 4096 May 14 16:06 time_in_state
-r--r--r--  1 root root 4096 May 14 16:06 total_trans
-r--r--r--  1 root root 4096 May 14 16:06 trans_table
```

- **reset**

Write-only attribute that can be used to reset the stat counters. This can be useful for evaluating system behaviour under different governors without the need for a reboot.

- **time_in_state**

This gives the amount of time spent in each of the frequencies supported by this CPU. The cat output will have "<frequency> <time>" pair in each line, which will mean this CPU spent <time> usertime units of time at <frequency>. Output will have one line for each of the supported frequencies. usertime units here is 10mS (similar to other time exported in /proc).

```
<mysystem>:/sys/devices/system/cpu/cpu0/cpufreq/stats # cat time_in_state
3600000 2089
3400000 136
3200000 34
3000000 67
2800000 172488
```

- **total_trans**

This gives the total number of frequency transitions on this CPU. The cat output will have a single count which is the total number of frequency transitions.

```
<mysystem>:/sys/devices/system/cpu/cpu0/cpufreq/stats # cat total_trans
20
```

- **trans_table**

This will give a fine grained information about all the CPU frequency transitions. The cat output here is a two dimensional matrix, where an entry <i,j> (row i, column j) represents the count of number of transitions from Freq_i to Freq_j. Freq_i rows and Freq_j columns follow the sorting order in which the driver has provided the frequency table initially to the cpufreq core and so can be sorted (ascending or descending) or unsorted. The output here also contains the actual freq values for each row and column for better readability.

If the transition table is bigger than PAGE_SIZE, reading this will return an -EFBIG error.

```
<mysystem>:/sys/devices/system/cpu/cpu0/cpufreq/stats # cat trans_table
From  :    To
        :   3600000   3400000   3200000   3000000   2800000
```

```
3600000:         0         5         0         0         0
3400000:         4         0         2         0         0
3200000:         0         1         0         2         0
3000000:         0         0         1         0         3
2800000:         0         0         0         2         0
```

3. Configuring cpufreq-stats

To configure cpufreq-stats in your kernel:

```
Config Main Menu
        Power management options (ACPI, APM)  --->
                CPU Frequency scaling  --->
                        [*] CPU Frequency scaling
                        [*]   CPU frequency translation statistics
```

"CPU Frequency scaling" (CONFIG_CPU_FREQ) should be enabled to configure cpufreq-stats.

"CPU frequency translation statistics" (CONFIG_CPU_FREQ_STAT) provides the statistics which includes time_in_state, total_trans and trans_table.

Once this option is enabled and your CPU supports cpufrequency, you will be able to see the CPU frequency statistics in /sysfs.