Linux Thermal Sysfs Enhancement

Srinivas Pandruvada
Motivation

• Improve performance for
  Intel® Dynamic Platform and Thermal Framework
  Linux Thermal Daemon
  Other user space tools
• Add missing interfaces
Agenda

• Thermal Management Complexity
• Linux Thermal Sysfs overview
• Limitations of thermal sysfs
• Thermal sysfs enhancements using IIO
• Example
Thermal Management Complexity

• Thermal issues
  No longer isolated conditions
  Requires Pro-active approach
• Complex thermal relationships
• User space thermal controllers
Thermal Management Example

Cooling: TFN1.0 (7/11), Fan2 (1/1), Fan3 (1/1), Fan4 (1/1), rapl_controller11 (3000000/14875000)
Linux Thermal

User space thermal controllers

Sysfs, uevents

Zones
Cooling devices
Governors

Device Drivers

HW

Thermal Core
Thermal Zone

Governors

Sensor
Trip Points

Cooling Devices 0

Cooling Devices 1

Cooling Devices N
Limitations of thermal sysfs

• Primarily developed for in-kernel monitoring and control
  Temperature read via sysfs attribute
  Performance impact due to polling
• Limited temperature change notification mechanism
  Asynchronous event setting via RW trip point
  Use kobject uevent
• Limited event support
  Trip point changed
• Only Pull Interface support
Requirements

• Binary Read I/F
• Pollable/Selectable file node
• Specify thresholds without patching trip points
• Asynchronous events with reason code
• Controllable Samples triggered from external events
Solution

- Use Thermal Sysfs 2.0
  Still can’t meet requirements
- Invent/develop new I/F
- Use Existing I/F
Linux IIO features

• Basic device registration
• Polled access via sysfs
• Software ring buffers for samples
• Events specification and notification
• Triggered sampling from external events
Thermal IIO Binding

- Each Zone Sensor as IIO Device
- Optional registration during thermal_zone_device_register
- Push Samples to buffers from user space governor notify function
- New thermal driver callback for thresholds
RFC Patches

http://marc.info/?l=linux-pm&m=143985565110987&w=2
http://marc.info/?l=linux-pm&m=143985565210989&w=2
Thermal-IIO Sensor Example

spandruvada@spandruv-mobl2:/sys/class/thermal/thermal_zone7$ tree -L

1
   ├── emul_temp
   │    └── iio:device7
   │         └── policy
   │         └── power
   │            └── subsystem -> ../../../class/thermal
   │                        └── temp
   │                               └── trigger7
   │                                    └── trip_point_0_temp
   │                                    └── trip_point_0_type
   │                                    └── trip_point_1_temp
   │                                    └── trip_point_1_type
   │                                        └── type = "x86_pkg_temp"
   │                                        └── uevent
Thermal-IIO Sensor Example

spandruvada@spandruv-mobl2:/sys/class/thermal/thermal_zone7$ tree -L 1 iio\:device7
iio:device7
    ├── buffer
    │    └── dev
    │          ├── events
    │          │    └── in_temp_raw
    │          │        └── name = “x86_pkg_temp”
    │          │        └── power
    │          │        └── scan_elements
    │          │             └── subsystem -> ../../../bus/iio
    │          │             └── trigger → current_trigger
    │          └── uevent

Device node = /dev/iio:device7
spandruvada@spandruv-mobl2:/sys/class/thermal/thermal_zone7$ tree -L 1
iio\:device7\:buffer/
iio:device7/buffer/
    ├── enable “0/1”
    │   └── length
    │       └── watermark

Thermal-IIO Sensor Example
Thermal-IIO Sensor Example

spandruvada@spandruv-mobl2:/sys/class/thermal/thermal_zone7$ tree -L 1
iio\:device7/scan_elements/
  ├── in_temp_en = “0/1”
  │    └── in_temp_index = “0”
  └── in_temp_type = “le:u32/32>>0”
Thermal-IIO Sensor Example

spandruvada@spandruv-mobl2:/sys/class/thermal/thermal_zone7$ tree -L 1
iio\device7/events
  ├── in_temp_thresh_either_en = "0/1"
  └── in_temp_thresh_either_value = "Temperature threshold value"
spandruvada@spandruv-mobl2:/sys/class/thermal/thermal_zone7$ tree -L 1
trigger7/
    ├── name = "x86_pkg_temp-dev7"
    │    └── power
    │        └── subsystem -> ../../../bus/iio
    └── uevent
Thermal-IIO Sensor Example

```
rroot@spandruv-mobl2:/sys/devices# tree -L 1
  iio_sysfs_trigger
  ├── add_trigger
  │    └── power
  │    └── remove_trigger
  │      └── subsystem -> ../..bus/iio
  │      └── trigger14
  │          └── uevent
```
Thermal-IIO Sensor Example

```
root@spandruv-mobl2:/sys/devices/iio_sysfs_trigger/trigger14# tree -L
1
  ├── name = "sysfstrig1"
  │    └── power
  │         └── subsystem -> ../../../bus/iio
  │             └── trigger_now
  │             └── trigger_poll
  └── uevent

root@spandruv-mobl2:/sys/devices/iio_sysfs_trigger/trigger14# echo 1 > trigger_now
root@spandruv-mobl2:/sys/devices/iio_sysfs_trigger/trigger14# echo 5000 > trigger_poll
```
Q & A