Thermal framework enhancements
Thermal enhancements

- User space dedicated trip points
- Debugfs for the thermal framework
- Improving the step wise governor
New trip point type

Dedicated user space trip point
Userspace dedicated trip points

- Most of the drivers have interrupt detection based notifications
- Userspace wants to be notified when a temperature threshold is crossed
- Writable trip points option
  - Passive, active, hot and critical temperature are writable
- Thermal framework is about protecting the silicon, letting userspace playing with the trip points is bad
Userspace dedicated trip points

Proposal: Create a `<user>` trip point type

- The thermal framework can clearly identify it and ignore any action except sending notification to userspace

- Takes benefit of interrupt based synchronous trip violation detection

- Writable trip point Kconfig option can be disabled

- `<user>` trip points stay writable

- Shall we create multiple `<user>` trip points?
Thermal debugfs
Debugfs for the thermal framework

- Very few information to investigate the thermal framework behavior
- Bogus cooling device stats
- We need to know how efficient are the mitigations
  - Depends on ambient temperature
  - Initial temperature
  - Heat capacity headroom
  - Temperature change speed
- Too many overshots = bad for the hardware
- Too many undershots = bad for the performances
Debugfs for the thermal framework

- Directory structure close to sysfs but simplified

```bash
thermal/
|-- cooling_devices
 | |-- 0
 |   |-- reset
 |   |-- time_in_state_ms
 |   |-- total_trans
 |   `-- trans_table
`-- thermal_zones
   |-- 0
   `-- mitigations
     `-- 1
        `-- mitigations
```
## Mitigation episodes

-Mitigation at 349988258us, duration=130136ms

<table>
<thead>
<tr>
<th>trip</th>
<th>type</th>
<th>temp(°mC)</th>
<th>hyst(°mC)</th>
<th>duration</th>
<th>avg(°mC)</th>
<th>min(°mC)</th>
<th>max(°mC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>passive</td>
<td>65000</td>
<td>2000</td>
<td>130136</td>
<td>68227</td>
<td>62500</td>
<td>75625</td>
</tr>
<tr>
<td>1</td>
<td>passive</td>
<td>75000</td>
<td>2000</td>
<td>104209</td>
<td>74857</td>
<td>71666</td>
<td>77500</td>
</tr>
</tbody>
</table>

-Mitigation at 272451637us, duration=75000ms

<table>
<thead>
<tr>
<th>trip</th>
<th>type</th>
<th>temp(°mC)</th>
<th>hyst(°mC)</th>
<th>duration</th>
<th>avg(°mC)</th>
<th>min(°mC)</th>
<th>max(°mC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>passive</td>
<td>65000</td>
<td>2000</td>
<td>75000</td>
<td>68561</td>
<td>62500</td>
<td>75000</td>
</tr>
<tr>
<td>1</td>
<td>passive</td>
<td>75000</td>
<td>2000</td>
<td>60714</td>
<td>74820</td>
<td>70555</td>
<td>77500</td>
</tr>
</tbody>
</table>

-Mitigation at 238184119us, duration=27316ms

<table>
<thead>
<tr>
<th>trip</th>
<th>type</th>
<th>temp(°mC)</th>
<th>hyst(°mC)</th>
<th>duration</th>
<th>avg(°mC)</th>
<th>min(°mC)</th>
<th>max(°mC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>passive</td>
<td>65000</td>
<td>2000</td>
<td>27316</td>
<td>73377</td>
<td>62500</td>
<td>75000</td>
</tr>
<tr>
<td>1</td>
<td>passive</td>
<td>75000</td>
<td>2000</td>
<td>19468</td>
<td>75284</td>
<td>69444</td>
<td>77500</td>
</tr>
</tbody>
</table>

-Mitigation at 39863713us, duration=136196ms

<table>
<thead>
<tr>
<th>trip</th>
<th>type</th>
<th>temp(°mC)</th>
<th>hyst(°mC)</th>
<th>duration</th>
<th>avg(°mC)</th>
<th>min(°mC)</th>
<th>max(°mC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>passive</td>
<td>65000</td>
<td>2000</td>
<td>136196</td>
<td>73922</td>
<td>62500</td>
<td>75000</td>
</tr>
<tr>
<td>1</td>
<td>passive</td>
<td>75000</td>
<td>2000</td>
<td>91721</td>
<td>74386</td>
<td>69444</td>
<td>78125</td>
</tr>
</tbody>
</table>
Debugfs for the thermal framework

- More information planned

-Mitigation at 349988258us, duration=130136ms

<table>
<thead>
<tr>
<th>trip</th>
<th>type</th>
<th>temp(°C)</th>
<th>hyst(°C)</th>
<th>duration</th>
<th>avg(°C)</th>
<th>min(°C)</th>
<th>max(°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>passive</td>
<td>65000</td>
<td>2000</td>
<td>130136</td>
<td>68227</td>
<td>62500</td>
<td>75625</td>
</tr>
<tr>
<td>1</td>
<td>passive</td>
<td>75000</td>
<td>2000</td>
<td>104209</td>
<td>74857</td>
<td>71666</td>
<td>77500</td>
</tr>
</tbody>
</table>

```
thermal/
|-- [ … ]
 `-- thermal_zones
     |-- 0
      `-- mitigations
         |-- 349988258
          |-- speed
          |-- [ … ]
         `-- stddev
```
Improving the stepwise governor
Stepwise governor improvement

- Problem: stepwise governor does not react fast enough
- Consequences: overshoots and undershoots
- Current solution: Increase sampling

Proposal:
- Temperature speed computation
- Temperature forecast
- Next trip point crossed anticipated by a timer
- We keep monitoring anyway
Temperature Monitoring intervals

- Overshot
- Hysteresis / undershot
- End of mitigation
Undershoot
overshoot

End of mitigation

Temperature

limit high

limit

limit low

Monitoring intervals

+1

-1

+1

-1

+1

limit

Temperature

Time

overshot

Undershot

End of mitigation
Thank you