

Service Incident Report for the cooling problem at PIC modular machine room on July 25

Incident Start: 24th July 2010 21:30 UTC

Incident End: 26th July 2010 07:00 UTC

Description

- Saturday 25-Jul at night there was an abrupt increase of temperature in our modular data processing center, quick increase of 15°C and peaking at 31°C. Reaching these limits triggered the immediate shutdown of the 8 blade centers (128 cores each) hosted there. All running jobs were killed. After some minutes temperatures and humidity stabilized again.
- As the increase of temperature was in a single step, everything pointed to the failure of one (out of two) of the cooling equipments. Today (Monday 26th) experts and maintenance team are investigating.

Impact

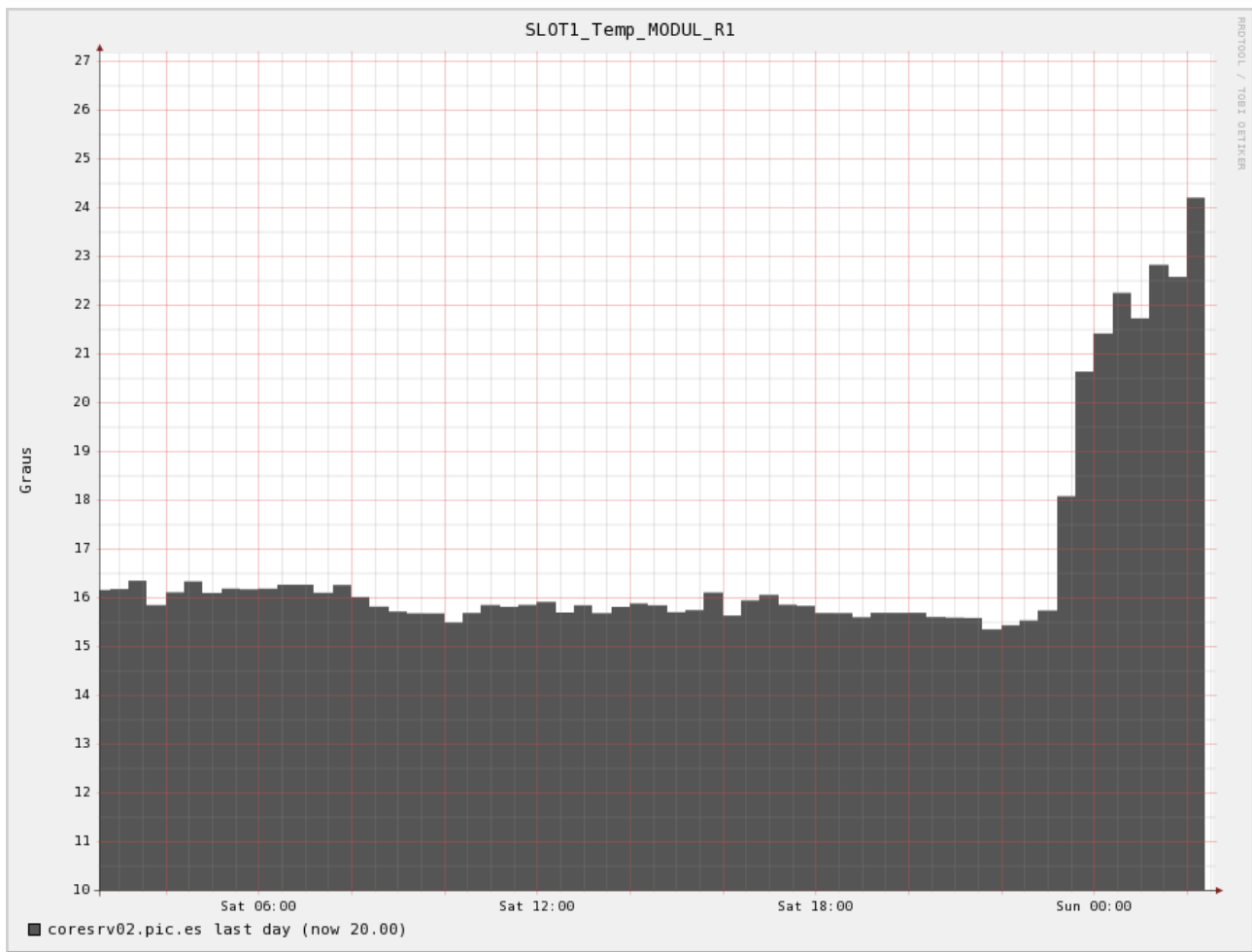
- There was no service interruption. The affected nodes were all WNs so the effect was just a degradation in the total computing power (of about 50%)
- About 800 jobs (50% of them from LHCb) were killed. Heartbeats were lost and no log files available as a consequence of the urgent shutdown, this happened at about 00:30 UTC on the 25th/July.

Time line of the incident

- 24-Jul 21:20 UTC - Temperatures started to raise from 16° to ~30°
- 25-Jul 00:30 UTC - SMS ALARM received, all probes at the modular CPD registered temperatures higher than 31°
- 25-Jul 00:30 UTC - Alarm triggered blade centers emergency stop (jobs killed)
- 26-Jul 07:00 UTC - Blade centers powered on.
- 26-Jul 08:00 UTC - Visit from maintenance and experts
- 26-Jul 14:00 UTC - Three servers affected could not be started. Still checking reason why they couldn't start automatically.

Details

The following graph shows the evolution of the temperature at the module on the 24 hours previous to the alarm.



Follow-up

- Still not clear the cause of the sudden increase in temperature, no failures observed today and experts didn't find anything relevant in alarms coming from the cooling equipments.
- Need improvements in getting relevant information from the cooling equipment (probably need an upgrade of the control card)