

## **PIC Tier-1 Service Incident Report 20091219 - Cooling problem**

**Incident Start:** 19/12/2009 at 10:00 local time (UTC+1)

**Incident End:** 19/12/2009 at 14:30 local time (UTC+1)

### **Description**

A failure in the cooling system of the building in the morning of Saturday 19th December 2009 caused a fast rise of the temperature in the PIC machine room. The Manager on Duty (MoD) and the Infrastructure Responsible (IR) were alerted at around 10:20 local time, and started an ordered fast shutdown of the most critical services at PIC: LCG-3D, CE, SE ... Then they went to PIC to check the situation on site. After checking that the cooling problem had been solved, they started an ordered start up of the services. The Tier-1 was fully back up and running at around 14:30 local time.

### **Impact**

Most of the core Tier-1 services were impacted during the roughly 4 hours that the incident lasted. Due to the cooling problem, they had to be shutdown and then started up later.

### **Time line of the incident** (local times, UTC+1)

- 10:08 the Nagios alarm reporting about temperature problems (TEMP\_SALA) goes CRITICAL (the MoD should have received SMS notification but there was a problem with the SMS generation - under investigation)
- 10:20 the Infrastructure Responsible (IR) gets a phone call from the maintenance company notifying the cooling problem.
- 10:24 IR starts remote shutdown of main services.
- 10:40 An UNSCHEDULED Downtime is declared in the GOCDB to inform the VOs.
- 11:00 IR goes physically to PIC and checks that the cooling problem has been solved and checks that the temperature in the room is back to normal.
- 11:45 IR and MoD start bringing up services in an ordered manner. The process is completed at around 15:00.

### **Analysis**

The cause of the original cooling problem is still not known at the time of writing this report (22nd Dec). We are waiting for a detailed report from the maintenance company. There are hints that point to a faulty compressor in one of the cooling units, but the details of what caused the complete system go off are still not known.

The lesson learnt is that we need to make sure we have in place the proper sensors for an early detection of a temperature problem and also the procedures for a fast shutdown and ordered start up.

After the start up all of the services came back correctly but one exception: the Oracle Streams of the LHCb LFC. After starting the Oracle backend and LFC frontend, the streams had some problem which was correctly detected by the SAM tests of the LHCb LFC. PIC Oracle experts got in contact with CERN Streams responsables to try and fix the problem. By the time of writing this report this activity is still ongoing and the Streams for the LHCb LFC are still not properly working.

### **Actions**

1. Follow up with the building maintenance team of the cooling problem primary cause and the derived actions to make the system more robust and avoid future similar episodes.
2. Debugging of the SMS trigger after Nagios temperature alarm.
3. Understanding and fixing the problem in the LFC LHCb Oracle Streams together with the 3D team at CERN.

4. Study the room for improvement in the sensitivity of our current temperature alarms. Goal is to provide to the MoD with as much information as possible for taking fast decision on whether to start site poweroff procedure. Study the need and feasibility of a fully automatised site poweroff when a temperature alarm is triggered.
5. Improve the fast poweroff procedure and automatise (script) it as much as possible.
6. Improve the ordered start up procedure and documentation for each of the Tier-1 services.