The LHC PC Rack Project

Fred Wickens

On behalf of the LHC PC-Rack Study Group

10th LECC Workshop

Boston

Sep 2004

The LHC PC rack study group

- The LHC PC rack study group started late in 2002
- Initially was LHC PC Rack Cooling Project
- Has since studied other issues with PC racks
- Many people involved from the 4 experiments, PH-ESS and TS
  - Alice: A.Augustinus, S.Philippin.
  - Atlas: N.Elias, Y.Ermoline, J.Godlewski, O.Jonsson, B.Martin, F.Wickens.
  - CMS: A.Gaddi, F.Glege, L.Pollet, A.Racz.
  - LHCb: L.Brarda, B.Chadaj, G.Decreuse, D.Gasser, Ph.Gavillet, R.Lindner, D.Ruffinoni, Ph.Vannerem.
  - PH-ESS: P.Maley, V.Pittin, Ch.Parkman.
  - TS-CV: M.Santos.
  - TS-EL: M.Delidais, J. Pierlot, K.Kahle, A.Funken
    - Note this is a low-level activity for essentially all of the above
- We acknowledge the generous support and contributions from others from the experiments and the technical sector.

The cooling problem

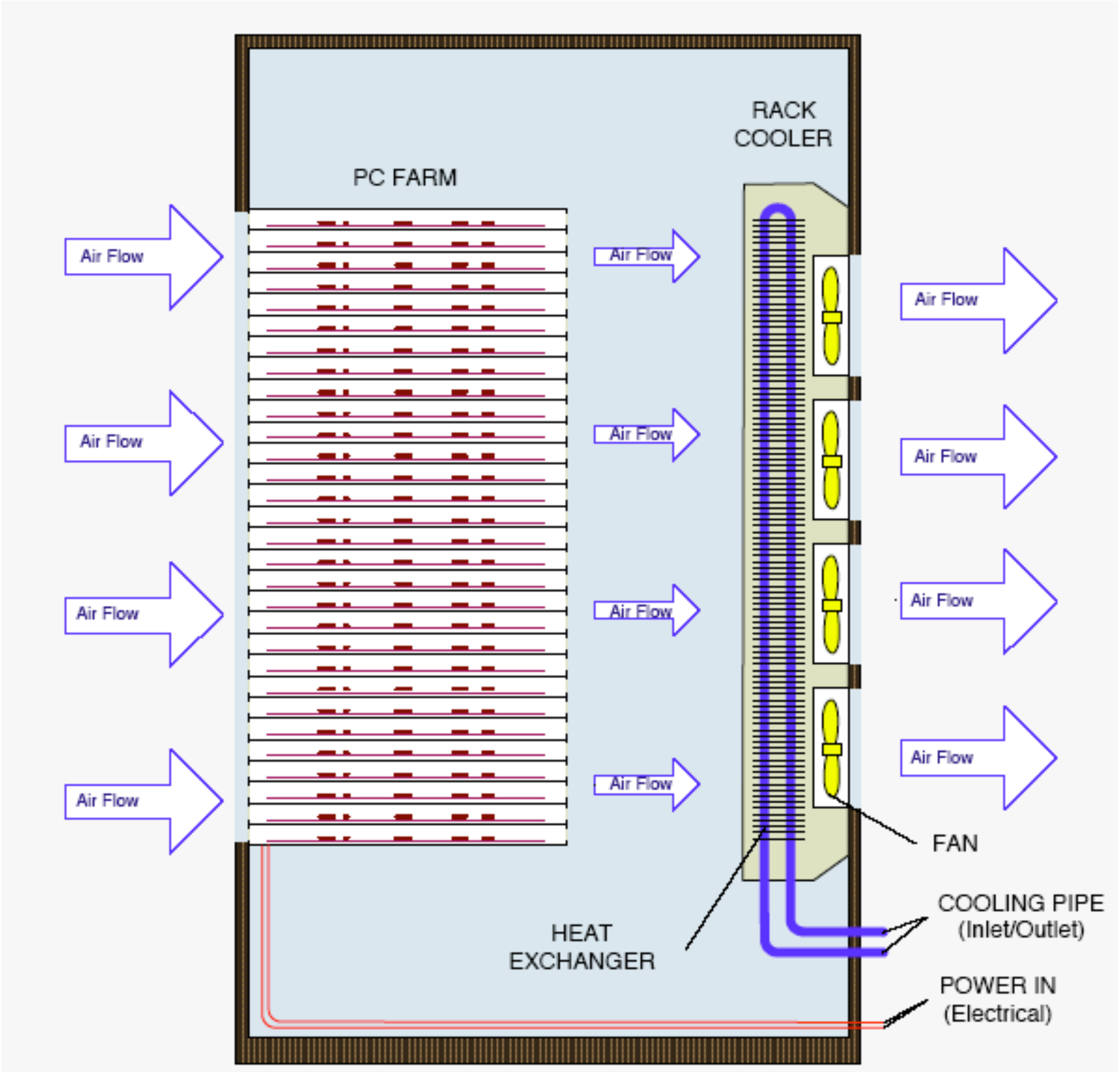
- **All of the LHC Experiments will have large numbers of PCs in their Trigger/DAQ systems**
- **Assumption is that these will be "Rack-Mounted" servers (generally mixture of 1U and 4U)**
- **By late 2002 a solution of water-cooled racks was established for LHC electronics - with vertical air flow**
- **But PC's require horizontal (front to back) air flow**
- **What does CERN Computer Centre do ?**
  - **Lots of space so machines not closely packed**
  - **Lots of air conditioning with high ceilings (so with care can obtain cooling ~2 kW/m^2)**
- **Looked to IT industry**
  - **Rapidly increasing power densities were recognised as a problem (Watts/CPU doubling in ~5 years. Density increasing at a similar rate)**
  - **Almost no installations yet at power density we need**

The cooling problem continued

- **1U servers are very compact, can fit 40 (or more) in a rack**
- **Typical dual 1U PC uses ~200W (PSU rated at ~400W), but increasing, so expect 250W (or more)**
- **Thus require 10-12 kW per rack**
- **Power density far too high for air-conditioning**
- **Best solution found was a water-cooled heat exchanger which could be fixed to the rear door of the rack**
  - **But only rated at 8 kW**
  - **Not initially available in Europe (made to USA specs)**
- **Some custom racks for this power were starting to become available - but expensive and generally much larger**

Cooling with horizontal air flow

- **Basic layout as shown**
- **Rack with ~40 x 1U PC's**
- **Water cooled heat exchanger fitted inside the rear door**



Cooling with horizontal air-flow

- Tests run with 30-48 PCs
- 30 single Xeon 2.4 GHz PCs
- 18 dual Xeon 2.4 GHz PCs

- Single CPU PCs used ~ 90W (at full load - ~60% if CPU idle)
- Dual CPU PCs ~170W
- Power factor typically over 90% (ratio W/KV)

- Max power in rack ~5.8 kW

Cooling with horizontal air-flow

- **Liebert RackCooler mounted inside rack (Max 8kW)**

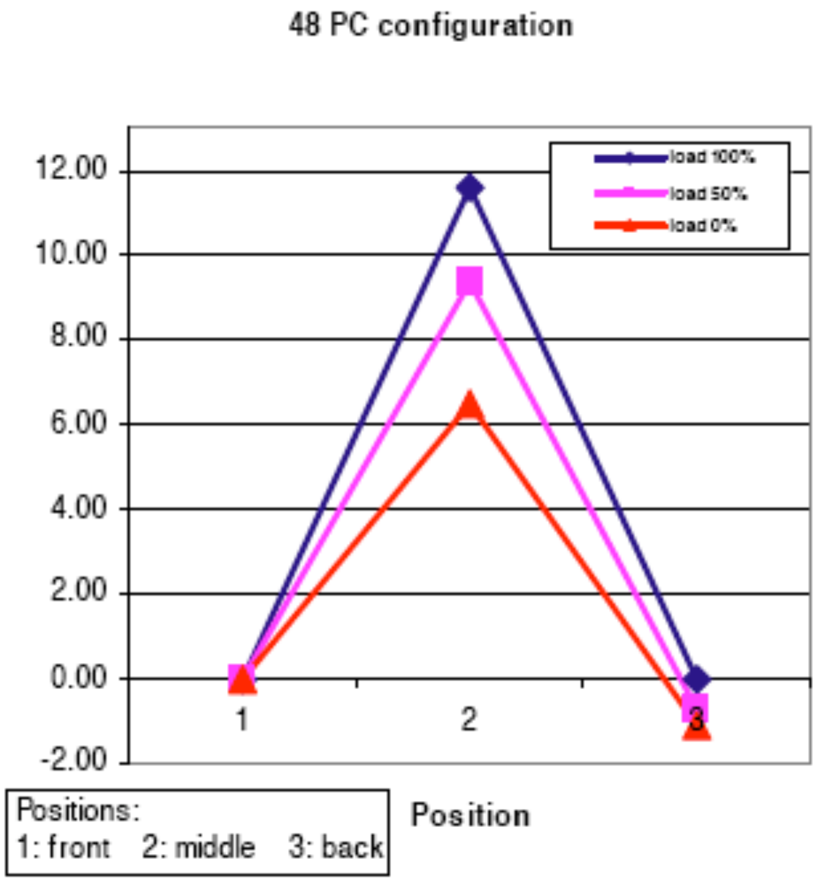
- **Rack sealed with door**
  - **Holes cut just round fans**
- **Extensive measurements made**

Cooling with horizontal air-flow

- Measure the CPU temperature and power consumption of individual PCs and power consumption of a complete rack
  - under full load
- Measure air temperatures at various positions inside and outside the rack
- Test cooling behaviour in different failure scenarios
  - Fan failure or water flow stop
- Results described in a technical note - published within each experiment
  - ALICE-INT-2004-014
  - ATLAS-DAQ-2004-9
  - CMS-IC-EN-0001
  - LHCb-2004-035 DAQ

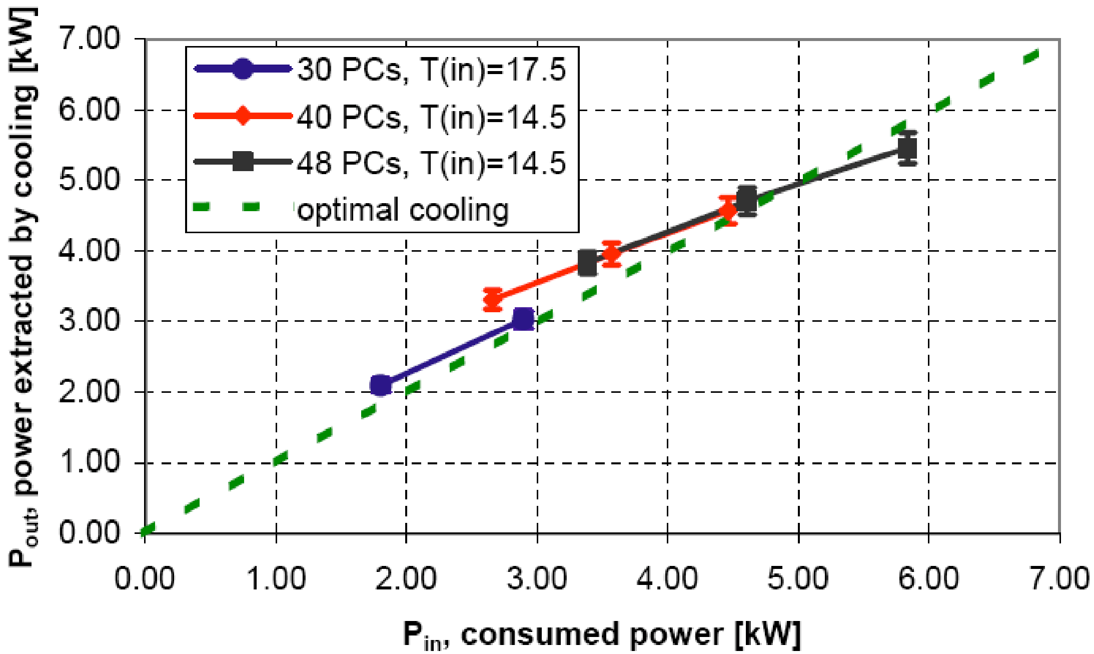
Cooling with horizontal air-flow

- **PC temperatures reasonably uniform over the whole rack**
  - **Even though rack taller than cooler**
- **PC's slightly cooler with RackCooler than in open rack**
  - **Improved air flow (~20%)**
- **Air temperature at outlet ~ as at inlet to rack**



Cooling with horizontal air-flow

- **At total load of 5.8 kW over 90% of the heat is removed by the cooler**



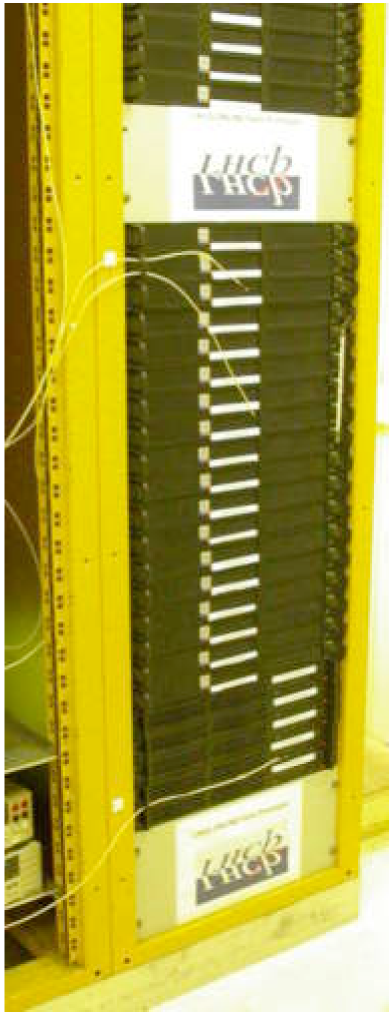
- **Failure of the rack-cooler fans or the cooling water led to a gradual rise in temperature - but in worst case only became critical for the CPU temperatures over ~2 hours**

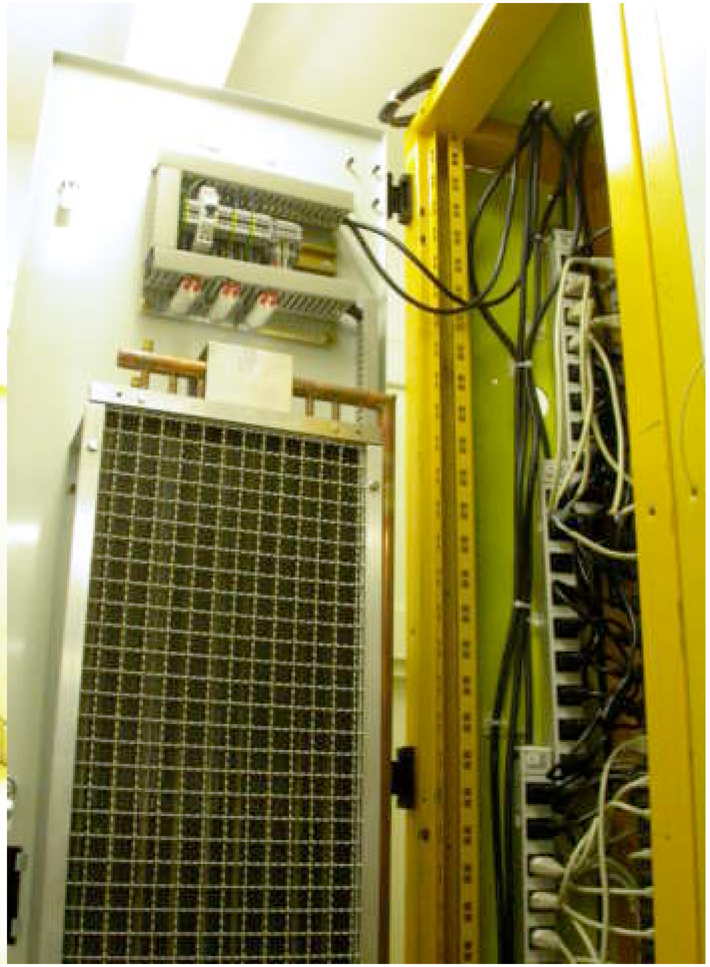
Cooling with horizontal air-flow

- **Following the measurements, discussed with various companies and an improved version was developed (CIAT)**
- **ATLAS, CMS and LHCb now have prototype racks with this cooler and ALICE plans to do so**
  - **Rated at ~10 kW**
  - **Can be mounted inside or outside rack**
  - **Dimensioned to fit 600 mm wide racks**
  - **Requires a rack of height ~2200 mm or more**
  - **Air is taken from the room and returned at ~ same temperature**
    - I.e. not closed circuit

Prototype Racks

- **Mounted inside a DELPHI rack for LHCb**



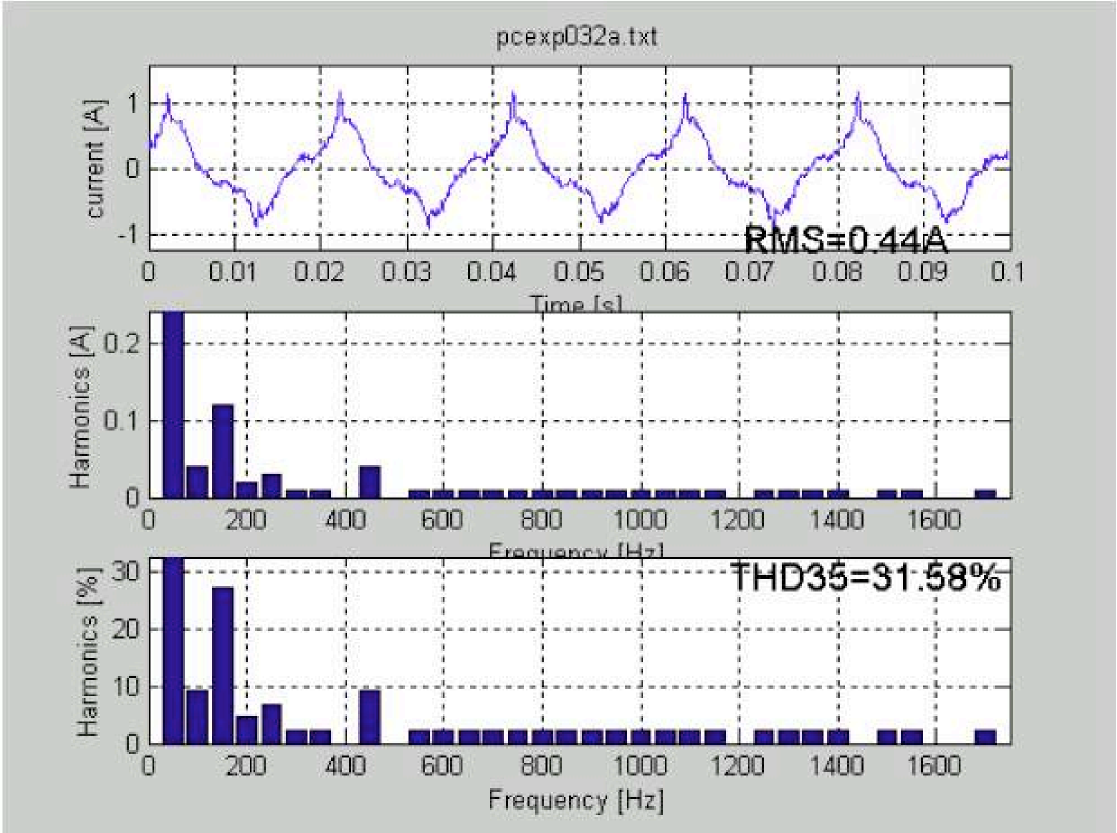


Prototype Racks

- **Mounted outside a new server rack for ATLAS & CMS**

Electrical Issues - 3rd Harmonic

- **Measured for several PC's - example: 1U 2.4GHz dual Xeon**
- **Current 1U PCs with Active pfc are not that well corrected**
  - **In principle can improve - but significant price**
- **Remove differential trip**
- **Reinforced neutral (phases add, do not cancel)**
  - **changes power distribution network + main transformers**



Electrical Issues - continued

- **Harmonic currents:**
  - **rate the neutral for same current (or larger) than each phase**
  - **Include a circuit breaker on the neutral**
  - **Size the transformer accordingly i.e *1.5-1.7**
- **These precautions, are not free, but may be less than special Active PFC for all PCs.**
- **Comparison with Bat 513 (EDMS 413142)**
  - **3rd Harmonics smaller in Bat 513**
    - Towers tend to have better correction (less constraint on size)
    - Greater mix of PC types
    - UPS systems help (although more on in-rush - see below)
  - **Experiments do not have farms on UPS - too expensive**
  - **Currently major upgrade to power and cooling for Bat 513 (from 0.6 to 2.4 MW) - but note special harmonic transformers**

Electrical Issues - continued

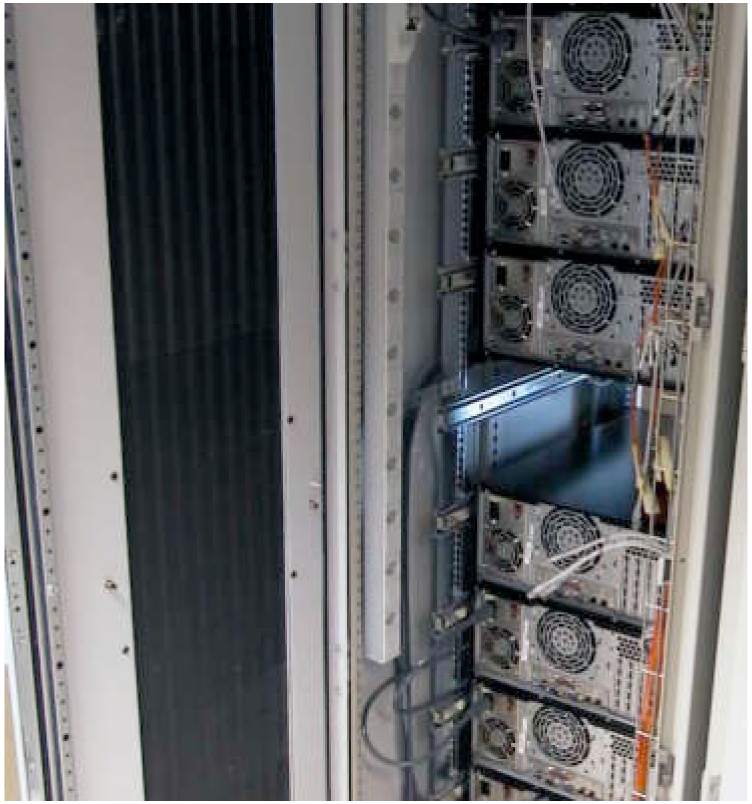
- **In-rush current (EDMS 442180)**
  - **Measurements made for different PC's**
  - **In-rush currents of 40-80 x normal current seen for ~20 ms**
    - Some supplies have much lower multiples ("Soft-Start")
- **How to distribute power to ~40 PC's in a rack**
  - **ATLAS/CMS have 3 phases each at 16 Amps in a rack**
  - **LHCb has 6 strips each at 10 Amps in a rack**
  - **Mechanical issue - 0U, where to mount**
  - **Simple barrette - cost ~10 CHF/PC**
    - Inrush current limits number of PCs per barrette
    - Use of D-Curve circuit breakers help (allow ~x10 current for 0.5s)
  - **Staggered power-on - cost ~40 CHF/PC**
    - Sockets on a barrette turn on in groups with ~.2s delay between each group
  - **Individual power-on - cost ~100 CHF/PC**
    - Optimal control, but uses more space, greater cost, limited number of suppliers and models (e.g. many have a 1A limit per outlet)

Electrical Issues - continued

- **How to remotely reboot PCs**
  - **Power cycle PC at distribution units**
    - But depends of granularity of control in power distribution
  - **Use serial connection on PC - if available**
  - **Boot on LAN signal**

Mechanical Issues

- **ATLAS/CMS propose to use industrial Server racks**
  - **Have flexibility to cope with PC mounting from different manufacturers**
  - **PC's mounted on rails supplied with PC**
    - 1U PC's vary considerably in width (+- 5mm) and depth (+- 50mm)
    - Slide rails fixed to the sides, mount front and back at 19 inch centres (Standard electronics racks do not have suitable mounting points for the rear support - cf Server racks do)
  - **Racks supplied with Cooler added to rear-door**



Mechanical Issues - continued

- **ALICE/LHCb will use recycled LEP racks**
  - **LHCb use support angle at each side of each server**
    - Add custom PTFE block to match the width of PCs
  - **Cooler attached to rear-door**
- **A rack of servers is very heavy - can easily reach 700 kg**
- **Small peripheral switches have Ethernet ports on front -PCs have Ethernet ports at back**
  - **Mount back-to-front - cannot see status LED's from front**
  - **Or bring cables from front-to-back - wastes at least 1U**

Other Issues

- **Cooling of switches**
  - **A lot of the PC racks include small switches**
  - **Some have front-to back air-flow**
    - But what if switch is mounted from back of the rack?
    - Reverse the fans or what ?
  - **Many have side-to-side air-flow**
    - Do we need additional baffles for adequate cooling?

- **Monitoring and control of rack infra-structure**
  - **DCS systems provides the tools**
  - **Overall rack power control by DCS**
    - How to integrate smart power distribution ?
  - **No smoke detection inside rack**
  - **Plan to use internal monitoring of PCs (Fans, temperatures)**
  - **Integration and correlation of farm fabric software and DCS**

Conclusions

- **A group has been working to find common solutions for the problems associated with racks of PCs at the LHC experiments.**
- **A solution has been found to provide ~10 kW of cooling with horizontal air-flow and flexibility to be used with various racks**
- **Study group (with TS) has also studied electrical and mechanical issues**
- **Continues to meet regularly to compare developments in the prototype farms now being established**

Back-up Slides

Use of Blades

- **Why not blades**
  - **Still not mature - but keeping a watch**
  - **High performance CPU blades are recent development**
  - **Currently price is more than 1U servers**
  - **Only clear advantage today is better power supplies and redundancy in cooling**
  - **For CPU intensive work power efficiency is no better**
  - **Density can be higher, but cannot use because of power/cooling and weight limits**
- **However if becomes appropriate to go to Blades**
  - **Uses 19 inch racks**
  - **Cooling is still front to back**
  - **Infra-structure largely unchanged**

Server Characteristics

- 1U servers are very compact, can fit 40 (or more) in a rack
- Typical dual 1U PC:
  - Uses ~200W (PSU rated at ~400W), but increasing, so expect 250W (or more)
  - Has large in-rush current (can be ~75 Amps for 20 ms) and significant 3rd Harmonic
  - Weighs 10-15 kg
  - Is cooled by horizontal air-flow, drawn in at front and blown out the back
  - Is ~60-70 cm deep, 1U high (no spare space for support guides), ~42-43 cm wide
  - Is normally supported by slide rails fixed to the sides, which mount front and back at 19 inch centres