Physical Infrastructure Issues In A Large Centre

July 8th 2003

Tony.Cass@CERN.ch

Physical Infrastructure Issues In A Large HEP Centre

July 8th 2003

Tony.Cass@CERN.ch

Why the change of title?

- I only have experience with an HEP centre…
- Non-commercial nature of the task we support influences choices
  - CERN does not lose money if the centre is not working.
    - » At worst, accelerator exploitation stops—but failures elsewhere are much more likely.
  - We can't write costs off against (increased) profits
    - » Pressure to minimise investment
  - We know we are planning for the long term
    - » B513 upgrade planning started in 2000. LHC operation starts in 2007 and continues for 10-15 years.

Fundamental requirements

- Flexibility
- Adaptability
- Futureproof
- Able to cope with change

Fundamental requirements

- Flexibility
- Flexibility
- Flexibility
- Reliability & Redundancy

Plan for the rest of the talk

- Equipment
  - needs an electrical supply
  - turns electrical power to heat at 100% efficiency
  - takes up space
  - can start a fire

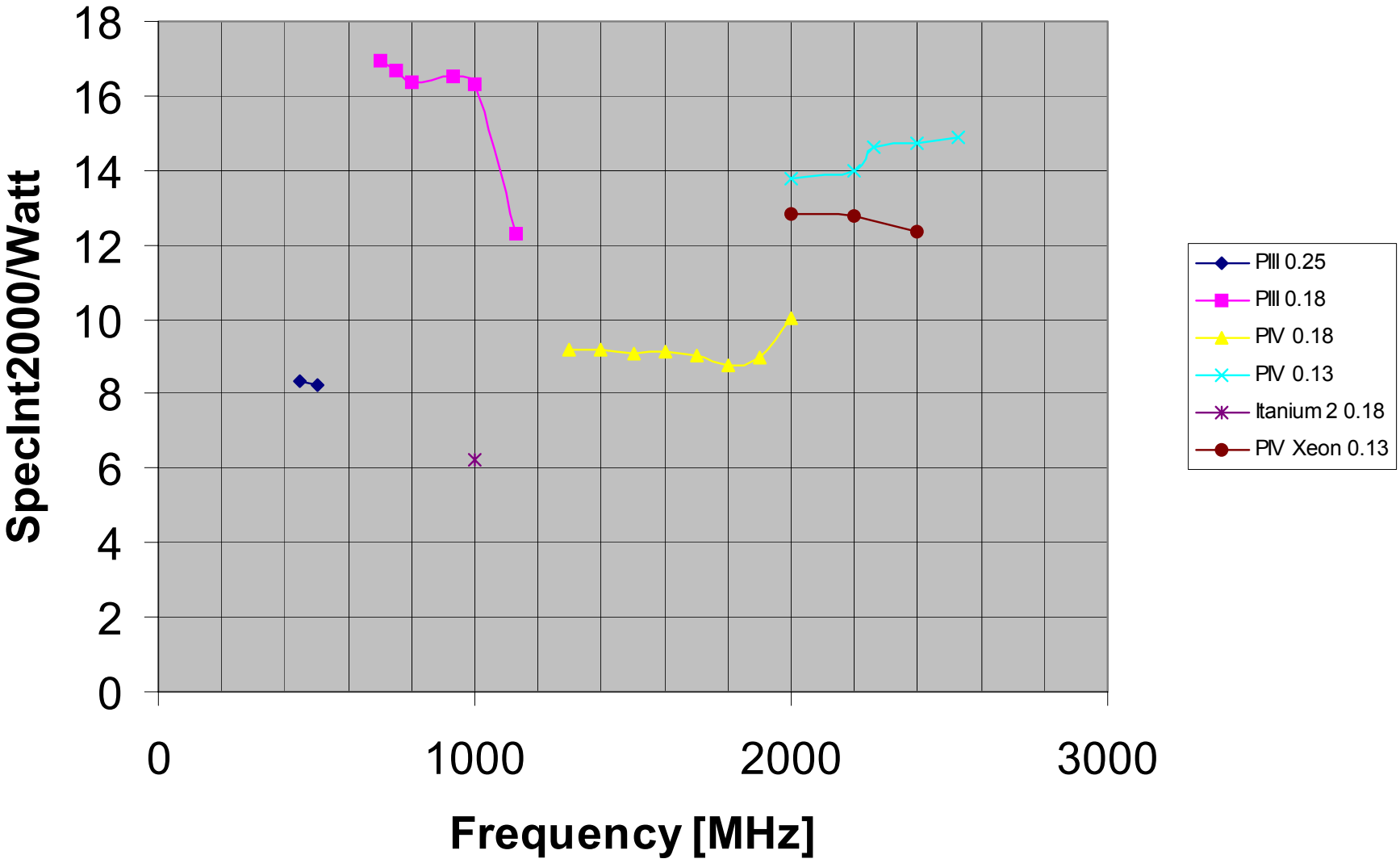
Electrical Supply Issues

- Demand
- Multi-level redundancy
- Backup
- Low voltage distribution network
- Recovery from failure
- Adapted to the load

Electrical Supply Issues

- Demand
  - How will this grow? Impact of
    - » demand for CPU capacity
    - » **Power efficiency of computing systems**
  - Future unclear
    - » Plausible historical evidence that W/SpecInt is constant
    - » Vendor statements about power budget for system type
- Multi-level redundancy
- Backup
- Low voltage distribution network
- Recovery from failure
- Adapted to the load

Processor performance (SpecInt2000) per Watt



Electrical Supply Issues

- Demand
  - How will this grow? Impact of
    - » demand for CPU capacity
    - » Power efficiency of computing systems
  - Future unclear
    - » Plausible historical evidence that W/SpecInt is constant
    - » Vendor statements about power budget for system type
- Multi-level redundancy
- Backup
- Low voltage distribution network
- Recovery from failure
- Adapted to the load

Electrical Supply Issues

- Demand
- Multi-level redundancy
  - High voltage supply & switchboards
  - Transformers
  - UPS
  - Low voltage supply, switchboards and distribution
  - [Equipment power supplies]
- Backup
- Low voltage distribution network
- Recovery from failure
- Adapted to the load

Electrical Supply Issues

- Demand
- Multi-level redundancy
- Backup
  - Independent high voltage supply
    - » CERN lucky with supply from EDF and EoS
  - UPS options
    - » Rotary
    - » Static with/without diesel backup
  - **Test your backup!**
    - » Override complaints that this causes pain elsewhere.
- Low voltage distribution network
- Recovery from failure
- Adapted to the load

Electrical Supply Issues

- Demand
- Multi-level redundancy
- Backup
- Low voltage distribution network
  - Can run cables to PDUs as necessary
  - Pre-installed busbar system much better
  - Flexibility: distribution network capacity should exceed available power.
    - » B513: 2MW available, network sized to deliver 3MW
- Recovery from failure
- Adapted to the load

Electrical Supply Issues

- Demand
- Multi-level redundancy
- Backup
- Low voltage distribution network
- Recovery from failure
  - Pre-failure triage
    - » We will support 250kW on diesels. Physics load shed after 5mins
  - Power should **stay off** if it **goes off**
  - Group services using the low voltage network
    - » Easy to power on services, not simply machines
    - » Consider dedicated supply (busbar) for, e.g., network switches
- Adapted to the load

Electrical Supply Issues

- Demand
- Multi-level redundancy
- Backup
- Low voltage distribution network
- Recovery from failure
- Adapted to the load
  - Switched mode power supplies are badly behaved
    - » generate harmonics and high currents in the neutral conductor
  - Early PC systems had power factor of 0.7. Recently installed systems are better behaved (0.95) in line with EU directive. **Can we assume 0.95 in future?**
    - » If not, need to install filters or higher rated equipment upstrean (UPS, switchboards, transformers)

HVAC Issues

- Water cooling
- Air cooling
- Acceptable temperature
- Redundancy

HVAC Issues

- **Water cooling is efficient**, but
  - Who has the infrastructure (anymore)?
    - » and this infrastructure is not very flexible
  - **Will vendors really ship systems requiring this?**
  - Could have water cooled 19" racks, but these still need the infrastructure and how do you ensure heat transfer from the PCs?
- Air cooling
- Acceptable temperature
- Redundancy

HVAC Issues

- Water cooling
- **Air cooling is simple and flexible**, but
  - low heat capacity
  - limited to 2kW/m² in our high computer centre and just 600W/m² in the vault.
  - Environmental problems:
    - » Noise in vault is 79dB, close to 85dB limit for continuous exposure. May be legal, but is not comfortable (so what? ☺)
    - » Will need air flow of 510,000m³/h in machine room or 60 air changes/hour.
  - Need to monitor air flow—convection may not be enough to drive adequate flow even in hot/cold aisle layout.
- Acceptable temperature
- Redundancy

HVAC Issues

- Water cooling
- Air cooling
- Acceptable temperature
  - **Fortunately, PC equipment is relatively tolerant**
  - Set point of machine room is 21°C
  - Vault set point is 25°C, Machine room will be 26°C
    - » Need 27°C for 3.5MW load across ground floor (2250m²)
  - But this is set point; **local temperatures are higher**
    - » 30°C+ in hot aisle in the vault
- Redundancy

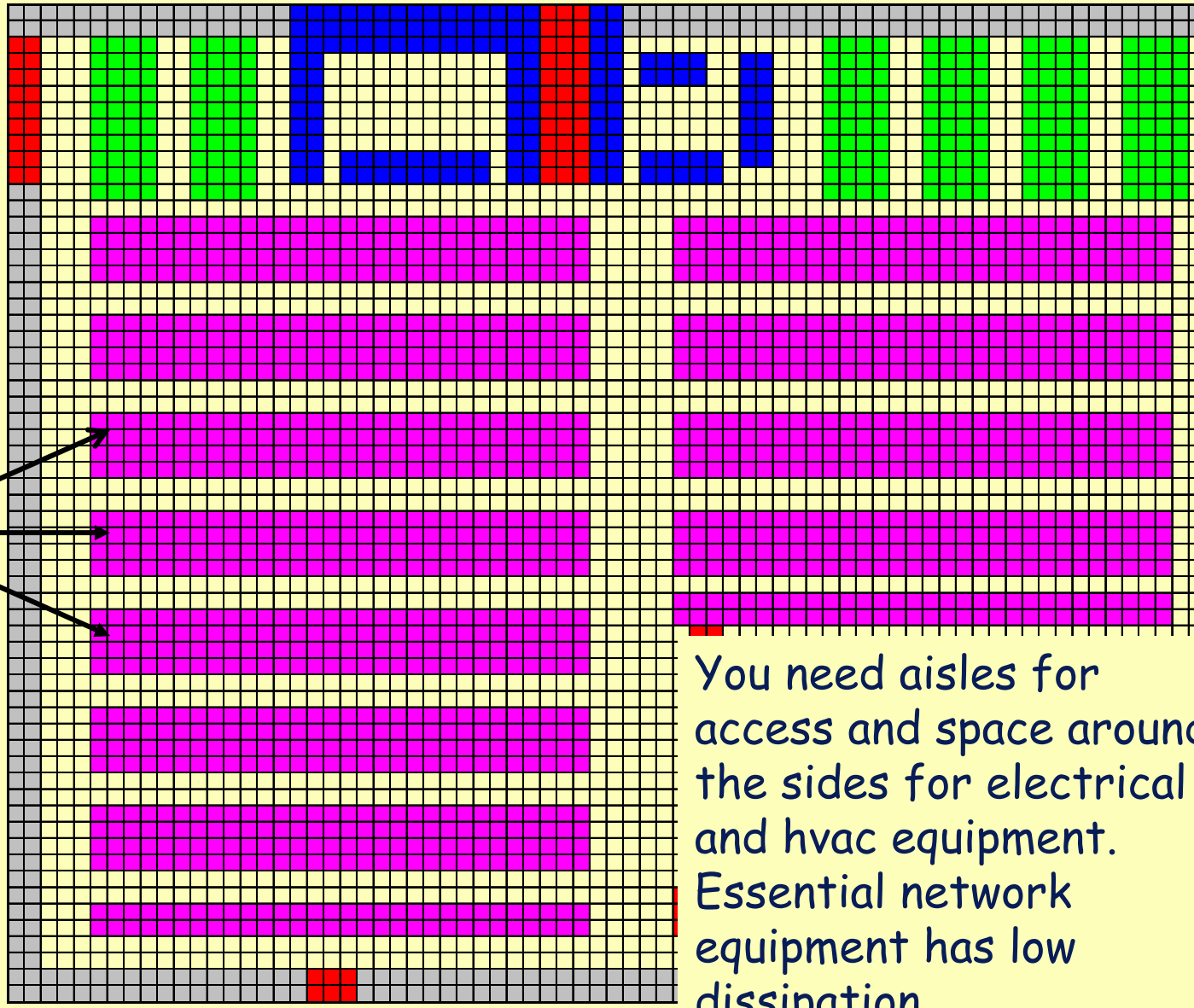
HVAC Issues

- Water cooling
- Air cooling
- Acceptable temperature
- Redundancy
  - with a 2MW load in the computer centre, **losing the air conditioning is as dramatic as losing the power supply**.
    - » heat rises above acceptable level in 10minutes
    - » and no privileged situation for the critical load
  - Large HVAC equipment (chillers, cooling stations) doesn't come with dual power supplies…
  - Still too many single points of failure in B513
    - » Human error, though, has been cause of most failures in last 5 years. Equipment is reliable, but is it coming to the end of its reasonable working life?

Space

- Is Space a problem?
  - HVAC limits heat load to 2kW/m².
  - 40x150W PCs in shelving units generate 1.25kW/m².
  - 40x150W PCs in 19" racks generate 5.5kW/m².
- Not quite the full picture:

Future Machine Room Layout



8m double
ows of racks
2 shelf units
r 36 19" racks

28 box PCs
105kW
440 1U PCs
288kW
24 disk servers
120kW(?)

You need aisles for access and space around the sides for electrical and hvac equipment. Essential network equipment has low dissipation.

Space

- Is Space a problem?
  - HVAC limits heat load to 2kW/m².
  - 40x150W PCs in shelving units generate 1.25kW/m².
  - 40x150W PCs in 19" racks generate 5.5kW/m².
- Taking into account corridors and n/w equipment
  - the machine room can hold over 6,000 white box PCs. This may be acceptable.
  - 1U PCs are required to saturate the power supply
    - » but is this the goal?
    - » We will probably have a mix of both by 2007
- Use space wisely
  - Place robots in areas with limited hvac capacity…

Fire Risks, Detection and Suppression

- **Fire is an ever present risk** for computer centres—and an area where the non-commercial business makes a difference.
  - **We can afford to shut down systems on a first alarm**
    - » or, at least, with today's level of false alarms
  - **We can't afford gas suppression** (cost & volume…)
    - » Are "hi-fog" systems really appropriate for a machine room?
- Major problem is smoke localisation
  - Given air flow, sensitive detectors will be triggered even at some distance from source of smoke.
  - Localisation of detection requires restriction of air flow which works against cooling needs.
- **We are concerned about halogenated materials.**
  - A small localised incident can produce large volumes of acrid smoke leading to widespread damage.

Questions (not a conclusion)

- **How will CPU and system power demands evolve?**
  - Will power consumption of HEP relevant systems follow the general trend?
- **Can we assume a well behaved load in future? Even for individually packaged boxes?**
  - i.e. power factor >0.95
- **Will water cooling make a comeback?**
- **Will we ever see halogen free PCs?**
  - and at what cost premium?