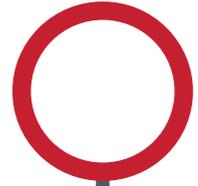


A SEQUENCE OF EVENTS

David Crooks

A/C & POWER INCIDENT

- Context
- Timeline
- Post mortem
- Follow up
- Hardware loss
- Reflections
- The Future



Machine rooms

243d

Old Mainframe room

Power locks off

141

Reclaimed workshop

Cooled by chillers + DFUs

Grid + School research groups

Timeline



FRIDAY 1 JANUARY



15:49:40 and 16:23:16

Scottish Power had a trip on one of the circuits which supplies the university. The campus didn't lose supply, but will have experienced a voltage dip for under 400ms.



About midnight

School admin receives a user email timed at 21:18, noting that their servers had disappeared; resolves to go in on Saturday.

NB: We had an expectation that 3 monitoring systems would have triggered by this point.

Scotgrid

Chiller (A/C Contractor)

Campus Estates Building Management System

Timeline

SATURDAY 2 JANUARY

NOON

Admin arrives on site to investigate unresponsive servers, discovers 141 very hot with door handle too hot to touch, a smell of plastic and a haze in the corridor, and plastic light fittings in the room melting.

Powers off the room at the wall. Informs other admins and security, and props open the room door. Admins and security set up fans to try to cool the room down. Room heating for ~20 hours

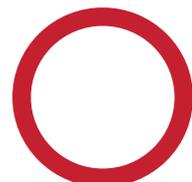
243d mostly off apart from servers powered by UPS

EVENING

Gordon and Gareth in attendance. 243d checked, and the power-supply reset. On power up, the non-grid systems seem to be OK.

The building (and room) entry system had troubles during the day.

Timeline

 SUNDAY 3 JANUARY

 1700

Research IT Manager in attendance. Fans still going, and the room at a normal temperature.

○ MONDAY 4 JANUARY

○ MORNING

Building Superintendent, Gareth, David, School Admin, Research IT Manager and A/C engineers in attendance. Building access working again. Room 141 powered up.

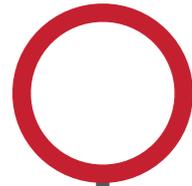
○ 1500

Chillers back on, fire panel inspected by engineer

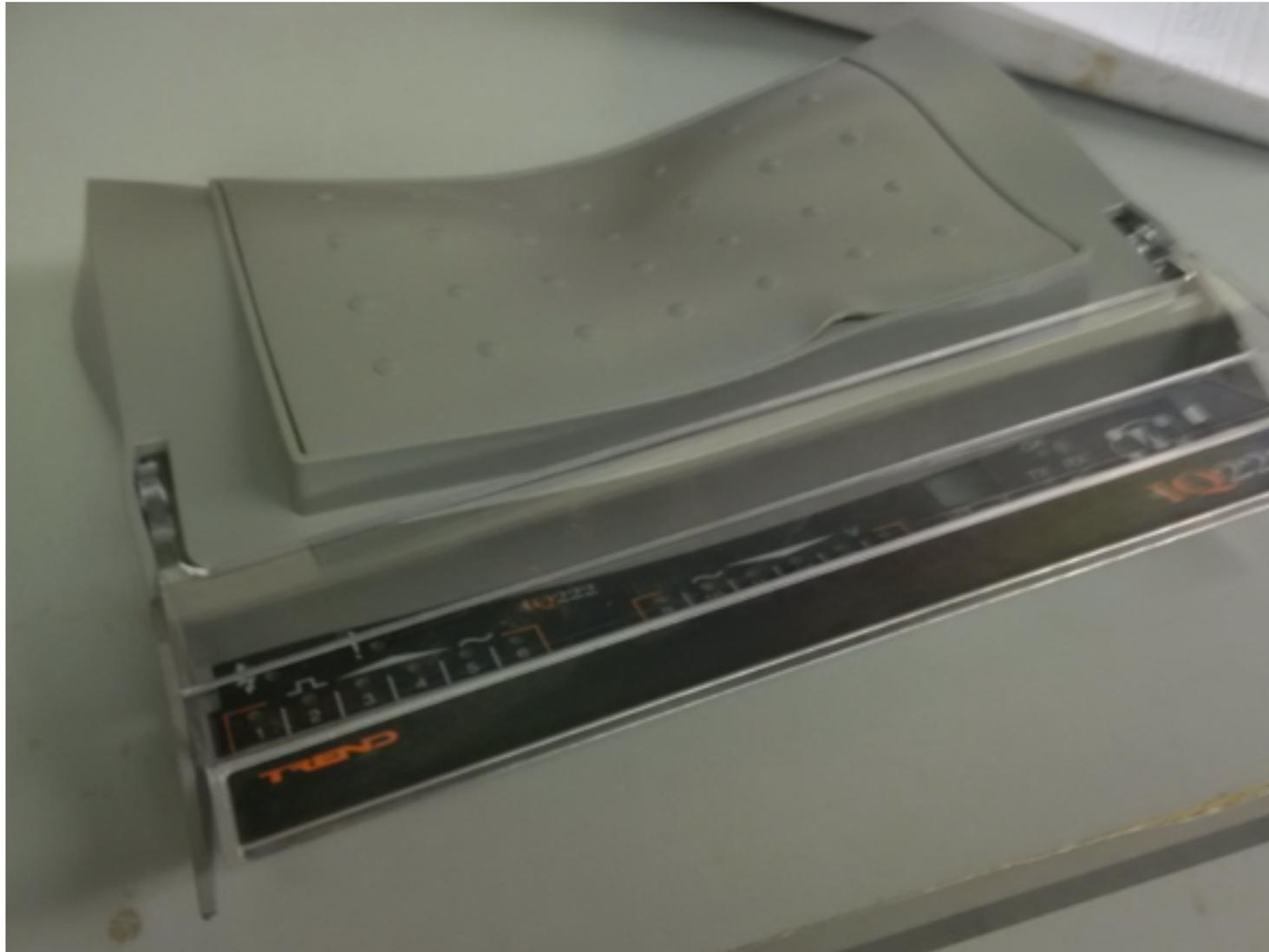
○ Cluster recovery

- What happened?
- Working theory is that power spike tripped the A/C but not the machines. The chillers then failed to restart.
- Logs from a central services switch show temperature passed 65°C.
- High temperature also interfered with building access control system.

Post mortem

 Visible effects

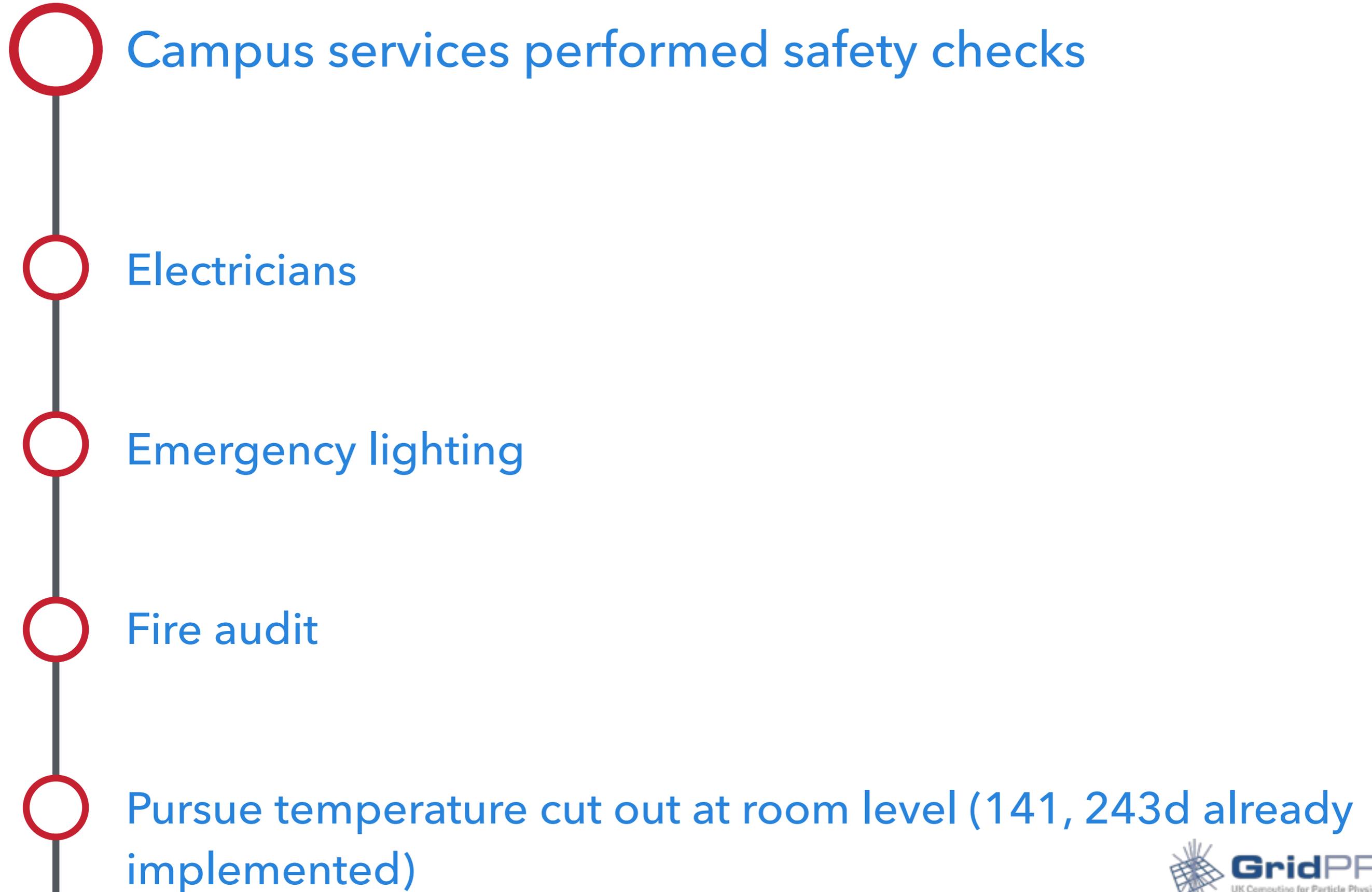
 Melted control panel



- Monitoring & Contacts
- 3 sets of monitoring failed
- Grid monitoring depended on networking which was inactive
- Chillers should have sent SMS alert which was not sent (but worked the next morning)
- Building Management System sent emails, but to inactive accounts
- Security contact list needed to be refreshed

- Chiller system
- The expansion vessel in the room 141 chiller system was undersized.
- Previously understood but fix not implemented

Follow up

- 
- Campus services performed safety checks
 - Electricians
 - Emergency lighting
 - Fire audit
 - Pursue temperature cut out at room level (141, 243d already implemented)

Hardware loss

- Grid & local PPE group
- Generation of worker nodes (~1000 cores)
- 4 PDUs + more showing signs of deterioration
- 2 UPS
- Other groups
- Visible signs of warping and heat damage internal to servers

Reflections

- Confluence of circumstances
- Not just one thing failed
- Collective responsibility
- Strong endorsement of need for purpose built facility to solve core issue
- Hope for the future

The Future

