IMPERIAL

Our Datacentre Journey

ACIT-HubTraining CoDesign Day 2025 - Bristol

Dugan Witherick
09/07/2025

Agenda

Imperial College London

09/07/2025

Before we begin

In this presentation I am showing Imperial's datacentre journey.

Every institution has its own requirements and constraints and therefore will have a different journey to ours.

However, hopefully some of the learnings from our journey will prove to be useful to your own.

Where we started
JISC Shared Data Centre

Otherwise known as VIRTUS London 4

   Located in Slough

Approximately 600kW of Research Computing IT load for Tier 3 only

   400+ nodes of high throughput compute
      incl. 144 GPUs (L40s & RTX 6000)
   300+ nodes of capability compute
      incl. 60 A100 GPUs

Rack density tops out at ~26kW

Energy supplied is low carbon

Mostly managed for us by ICT Datacentre team

Why Direct Liquid Cooling

CPUs are getting hot
> **Granite Rapids 6960P has a TDP of 500W**

GPUs are already hot and are getting hotter
> **A 4-GPU HGX H200 board has a total power requirement of 2.8 kW**

It is getting extremely difficult (or impossible in some circumstances) to cool these systems with air.

Or the servers have such a high-power density we can only fit a few in a rack.
> **75 kW for a rack of Granite Rapids**
> **110 kW for a rack of H200s**

Water cooling should be more energy efficient
> **Typically, much lower Power Usage Effectiveness (PUE) than air-cooled**
> **Good for Imperial's sustainability ambitions**

Location
Can we use Imperial's buildings

South Kensington
> **Space constraints**
> **Combined Heat and Power**

Hammersmith and St Mary's
> **Not the best use of constrained hospital land**

White City
> **Lots of new buildings under-construction**
> **Nothing ready on a viable timescale**

Silwood Park
> **No suitable buildings**
> **But plenty of space**

What about containerised/modular datacentres

Essentially a datacentre in a "box"

Silwood Park has plenty of space.
> **It also had the UK's last civilian nuclear reactor**
> **A nuclear history**

You still need
> **Somewhere suitable to put it like a large concrete slab**
>
> **Probably planning permission**
>
> **Power**

May still be possible in the future

Co-located Data Centres
Space, power, cooling

Water-cooling makes already complicated discussions with co-lo's even more complicated

Do they have experience with DLC (either generally or with your chosen supplier).

Are they just retro-fitting a low-density air-cooled data-centre
> **Will you end up just paying for lots of floor space you can't use**

The density of DLC racks means they are quite heavy.
> **Is the floor up to weight (our heaviest racks in the new system are 1,700 kg)**

How do you cool the air-cooled racks (management and storage) and take away the radiative heat
> **Rear-door heat exchangers (RDHX), connected to where?**
> **Hot or cold air containment of the whole system?**

CDU
Or coolant distribution unit

Transfers heat from the secondary coolant (the loop that goes around the servers) to the primary coolant (the datacentre loop).
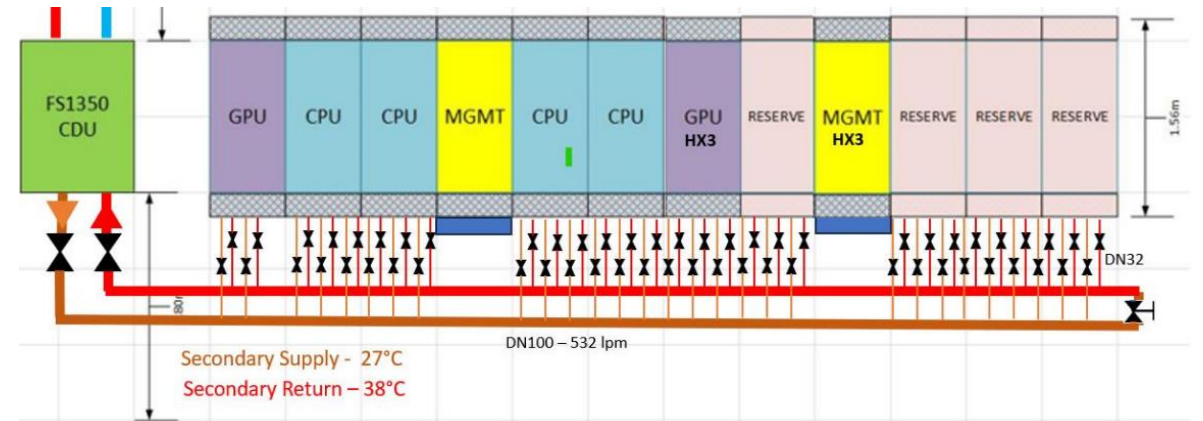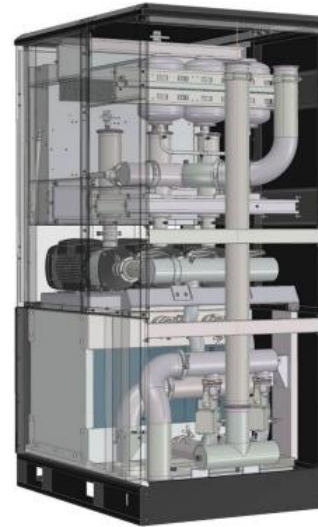
In-rack or standalone CDUs. Advantages and disadvantages of each.

In-rack more flexible with different manufacturers per rack.

Standalone requires few DC connections

and "may" be slightly more efficient.

Cooling performance is dependent on primary datacentre supply (temperature, flow etc.).

Do they need maintaining and if so, who maintains them?
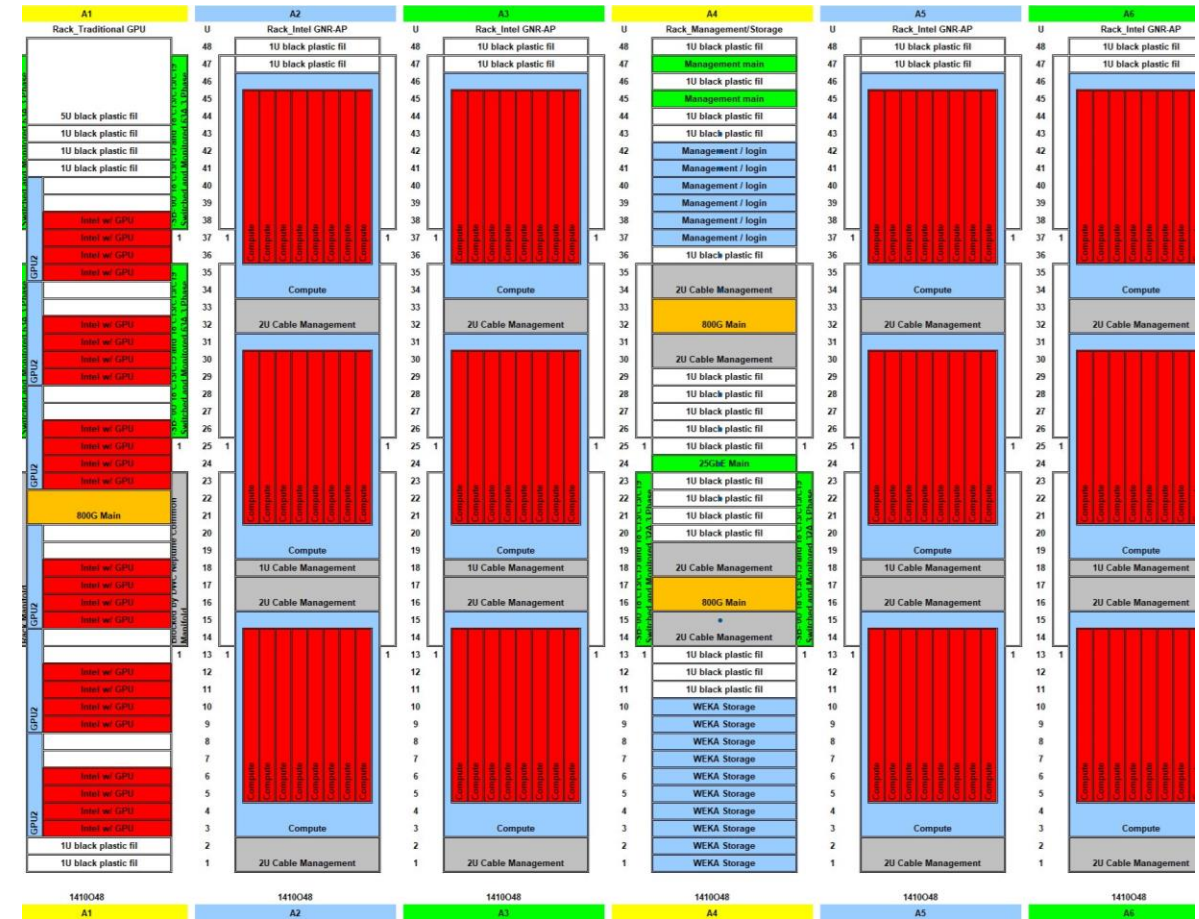
Where are we now
HX2 and HX3

HX2

New high-throughput cluster
192 Compute Nodes of Intel Granite Rapids (27,648 cores)
24 GPU Nodes providing 96 NVidia H200 GPUs

HX3 (HEX-AI)

New cluster designed for interactive AI research and teaching
16 GPU Nodes providing 64 NVidia H200 GPUs

Ordered…..

IMPERIAL

Questions?