Archive Room Plan

Tom SF Haines — Revision 3 (October 26th 2023)

Summary

We're out of server room capability, and need more. Intention is to convert the archive room (1W 2.55) to be a new server room, to replace the current server room (1W 2.51) and increase capacity.

We currently have £80K of servers sitting in a store room doing nothing (compliments of the absurdity of the end of year financial rush), so this is urgent, as much as "build a server room" can be urgent. We have resorted to putting servers in the Edge server room, as an emergency measure, but its ability to take pressure off is very limited (we can maybe move another 3kW, and get another £20K of servers online, but that's it). It is causing latency issues and the Edge server room is not suitable for high end GPU servers.

As a point of clarity, we mostly buy servers for training machine learning models: they have a power/cooling density far in excess of what you find in a typical server room. Consequently, most of the assumptions usually made for classic server rooms are invalid here. This also means we're very sensitive to what we've asked for: even slight variations from the below specifications should be checked.

Status

Three limits apply to a server room (ignoring networking):

Limit	**Unit**	**Available**	**Used**
Space	U (1U = 4.445cm of vertical space)	84U	70U (36U for GPU cloud, 34U for normal servers)
Power	Watts (W)	28,750W	16,679W + unknown additional load
Cooling	Watts (W), UK sensible	14,430W	16,679W + unknown additional load

Note in all cases that over capacity is required, to handle burst usage and quantisation effects, e.g. power comes in through multiple circuits, and each PSU is on one circuit only. Power and cooling are equivalent in one sense — almost all energy consumed gets turned into heat. But they also aren't equivalent because if you break the power limit the circuits trip and everything dies, but if you break the cooling limit for a short enough period you'll be ok.

Objective

Our recent purchasing, by year:

Year	Space	Power / Cooling
2019	13U	4,395W
2020	7U	420W
2021	6U	4,752W
2022	10U	7,112W
2023	16U	~9,000W

Note that the 2023 entries can't actually be switched on, so I've guessed the power consumption (we use actual measured values, not PSU limits).

The average over that period is 10U and 5KW per year. Hence, some (horrifically simplistic) predictions:

When	Total U	Power / Cooling
Now	86U (3 racks)	25,700W
+3 years	116U (3 racks)	40,700W
+7 years	156U (4 racks)	60,700W

The initial installation will be to provide enough capacity to last three years, but the design is intended to be good for seven years with future upgrades.

Note that the planned numbers need to be higher than the above, to account for quantisation effects and because we also want to support keeping PhD student desktops in the space during their write up period, so their (remote access) desktops are not taking up valuable desk space. I'm assuming 20kW will cover the desktops:

- High end desktops are power hungry — assume 1kW each.
- We currently have ~76 PhD students (taken from space spreadsheet, so have overcounted due to postdocs)
- Assume 25% are writing up at any given time
- 25% of 80 gives us 20kW. In practice I doubt even half will use this, but the department keeps growing and Nvidia GPUs keep getting hungrier, so best to have some leeway.

With regards to quantisation effects, most server PSUs are 1-3kWs, and most circuits are 32A which means 7.36kW. We still mostly see 1kW PSUs and 1kW is 14% of 7.36kW, consequently we want to boost everything by 14% to account for quantisation. That's also good for how AC units fade over time (about 1% per year assuming proper maintenance). This gives us a new updated target:

When	Total U	Power / Cooling
+3 years	4 racks — 168U	58kW
+7 years	5 racks — 210U	92kW

Note that an extra rack has been added, because they are cheap and it avoids wasting time moving servers around to pack them in tighter. Spacing them out can help with cooling as well. In practice the watts number is good for cooling, but we'll overspec the power a bit more, just to distribute plugs around the room and avoid crazy wiring.

Floor loading

Looking at recent GPU server purchases they range from 15kg to 20kg per U. A fully loaded rack could hence weigh 840kg, plus assume about 160kg for the rack and everything else gets you to 1000kg per rack (looking at the specs of actual 42U racks they typically weigh around 100–200kg and give a static loading of 800–1000kg, so this is in alignment). A rack is, at least the ones we use, 0.6m wide and 1m deep, giving a floor loading of 1666kg per square metre, which is 16.3kN/m^2.

Google suggests the standard UK floor loading in the UK for commercial buildings is 2.5kN/m^2, though can be larger for older buildings that were over-engineered or if on the ground floor. Obviously, this may be a problem. We also have to consider that, while there is a lot of weight where the racks go, there are also large gaps where no more than a person walking needs to be supported, i.e. the exact definition given by the architects needs to be considered in context.

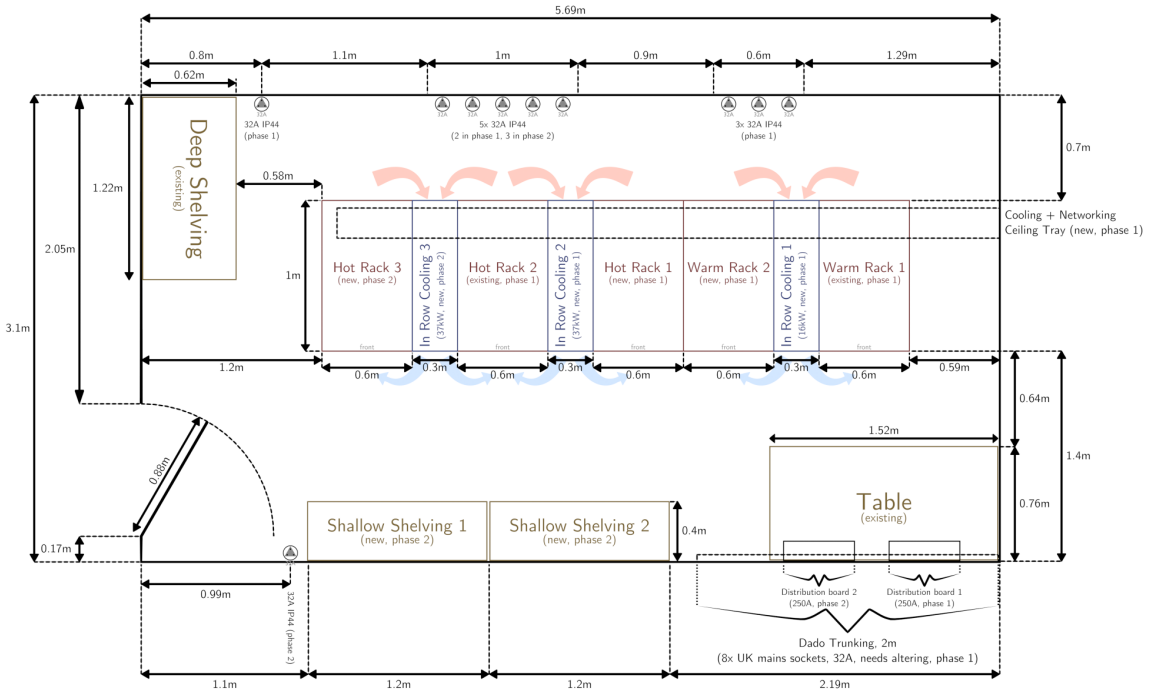
Steps

Here is what needs to happen:

1. Check weight loading of 1W 2.55 floor
2. Check if the buildings transformer has enough excess capacity
3. Clear out archives
4. Remove radiator, and the water pipe going up the wall
5. Install distribution boards and power points, including modifying dado trunking
6. Add ceiling network + cooling cage (will need to check with AC engineers first)
7. Add external networking lines
8. Add racks (hot rack 1 and warm rack 2) and in row cooling units (same time to get spacing right, AC engineers involved)
9. AC engineers to setup cooling, including external units
10. We move everything over (hot rack 2 and warm rack 1 + servers)
11. Old server room is repurposed as lab space (probably for HCI user studies)

The ordering of some steps is mutable.

Design



Part list

Phase 1 (now):

- 2x 42U racks
- 16kW in row cooling unit
- 37kW in row cooling unit
- 250A distribution board
- 6x 32A IP44 sockets (+2 if possible — overloads distribution board)
- Some dado trunking/plugs
- Cooling + networking ceiling tray

Phase 2 (3 years from now):

- 42U rack
- 2x 120x40cm industrial shelves
- 37kW in row cooling unit
- 250A distribution board
- 3x 32A IP44 sockets (+2 ideal, but not in plan as location/need uncertain)

Notes

- There isn't really enough power in either phase: both are short by 2x 32A plugs. In phase 1 this is because there isn't space on the distribution board (assuming AC also goes there, with an EER of 10). Getting a larger one gets awkward, as the next size up is a substantial jump, including in size (wall height wouldn't leave space for the table). In phase 2 this is because we can decide at the time.
- We may want to consider including a 1.2m deep rack, to support larger servers. We can take an extra 10cm in both directions to achieve this. Would probably go in the "Hot Rack 1" position, as big servers are more likely to be found as part of the GPU cloud but could also appear for other uses.
- All the 3D models in the renders (below) are representative, e.g. we're not sticking fans at the top of the racks as that would be stupid. Just to give a sense of the space and to clarify layout details that the above 2D plan cannot.
- For the commando sockets I just used what we currently have for the 3D renders, but given we're putting them at ceiling height getting a design that swaps the plug/switch would make sense.
- Have not given the exact location of ceiling tray: doesn't matter that much as on ceiling and will be driven by the need of the AC engineers. Towards the back of the servers is preferred, as that works better for networking and leaves space for baffles on the off chance that separating the front/back of the servers more explicitly will improve cooling.
- The 32A sockets have some flexibility: as long as they are roughly where indicated and won't collide with anything then should be fine.
- The plugs in the dado trunking will need 13A fuses, as will be used for server maintenance/testing. This does mean 32A isn't really enough, but shouldn't be an issue in practice. There will be a need to plug potentially broken equipment into this circuit for testing, so having two pairs of plugs with RCDs is necessary (servers have multiple PSUs, some as many as four).
- For the 250A distribution board I used as reference:
  Acti9 Isobar P B - distribution board - 8 ways TP+N (SEA9BPN8)
  https://www.se.com/uk/en/product/SEA9BPN8/acti9-isobar-p-b-distribution-board-8-ways-tp+n/
- For the in-row cooling I used
  https://www.serverroomenvironments.co.uk/in-row-cooling-units
  to get a sense of what sizes exist, but for the 3D model I used
  Mitsubishi Coolside In-Row DX 009 - 025
  https://www.melcohit.com/en/products/2571/full-inverter-air-conditioning-system-with-vrf-technology-for-it-cooling
- For the 0.4m deep shelves I used a random search result as reference:
  https://g-rack.co.uk/products/180-x-120-x-40cm-5-tier-boltless-shelving-unit-blue

Renders

The below are provided to help resolve the ambiguity of a top down plan; they include both phases. And because shiny.