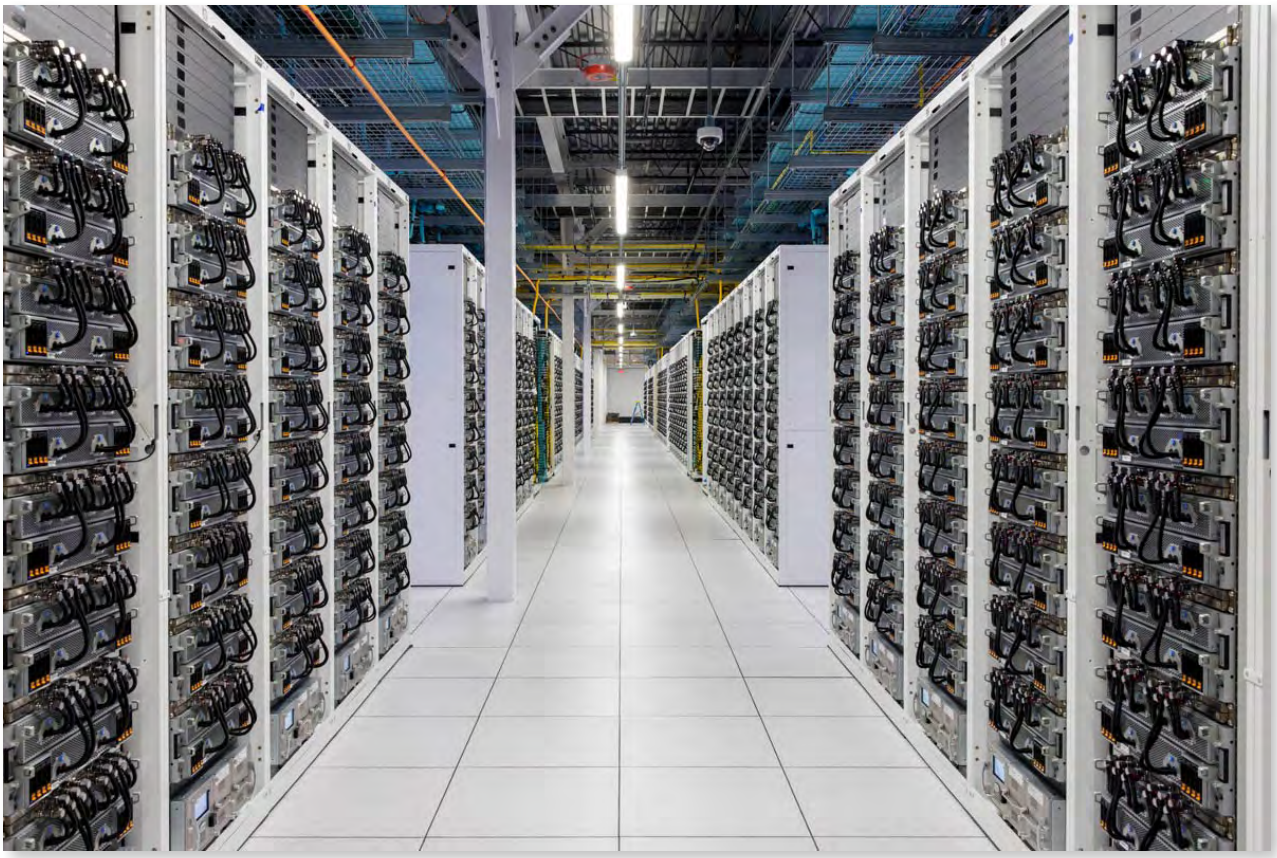
Inside the 100K GPU xAI Colossus Cluster that Supermicro Helped Build for Elon Musk

Authored by Patrick Kennedy, ServeTheHome



Today, we are releasing our tour of the xAI Colossus Supercomputer. For those who have heard stories of Elon Musk's xAI building a giant AI supercomputer in Memphis, this is that cluster. With 100,000 NVIDIA H100 GPUs, this multi-billion-dollar AI cluster is notable not just for its size but also for the speed at which it was built. In only 122 days, the teams built this giant cluster. Today, we get to show you inside the building.
Of course, we have a video for this one that you can find on X or on YouTube:

https://youtu.be/Jf8EPSBZU7Y

Normally, on STH, we do everything entirely independently. This was different. Supermicro is sponsoring this because it is easily the most costly piece for us to do this year. Also, some things will be blurred out, or I will be intentionally vague due to the sensitivity behind building the largest AI cluster in the world. We received special approval by Elon Musk and his team in order to show this.

Supermicro Liquid Cooled Racks at xAI

The basic building block for Colossus is the Supermicro liquid-cooled rack. This comprises eight 4U servers each with eight NVIDIA H100's for a total of 64 GPUs per rack. Eight of these GPU servers plus a Supermicro Coolant Distribution Unit (CDU) and associated hardware make up one of the GPU compute racks.

These racks are arranged in groups of eight for 512 GPUs, plus networking to provide mini clusters within the much larger system.

XAI Colossus Data Center Supermicro Liquid Cooled Nodes Low Angle

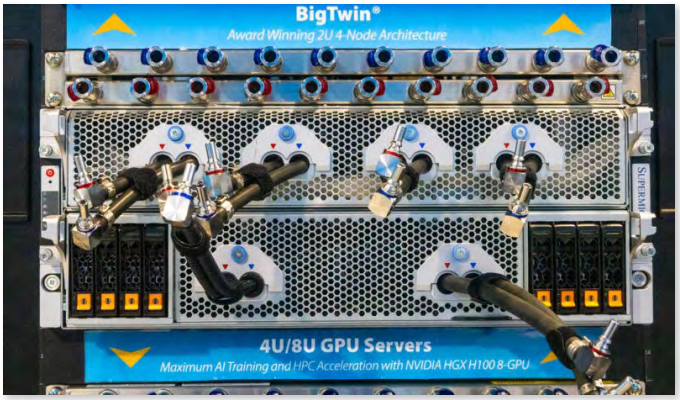
XAI Colossus Data Center Supermicro 4U Universal GPU Liquid Cooled Servers

Here, xAI is using the Supermicro 4U Universal GPU system. These are the most advanced AI servers on the market right now, for a few reasons. One is the degree of liquid cooling. The other is how serviceable they are.

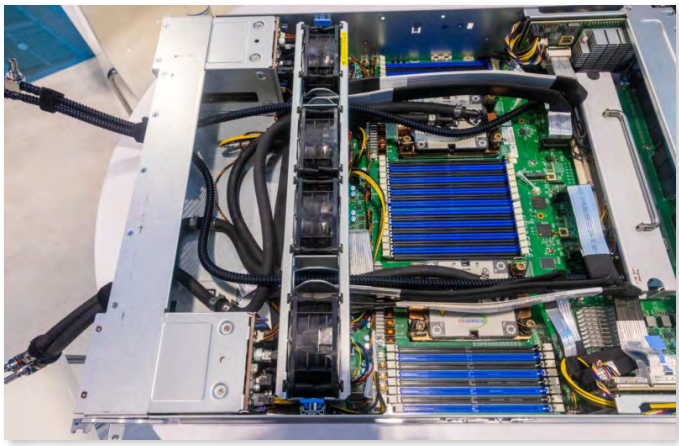
We first saw the prototype for these systems at Supercomputing 2023 (SC23) in Denver about a year ago. We were not able to open one of these systems in Memphis because they were busy running training jobs while we were there. One example of this is how the system is on trays that are serviceable without removing systems from the rack. The 1U rack manifold helps usher in cool liquid and out warmed liquid for each system. Quick disconnects make it fast to get the liquid cooling out of the way, and we showed last year how these can be removed and installed one-handed. Once these are removed, the trays can be pulled out for service.

XAI Colossus Data Center Supermicro 4U Universal GPU Liquid Cooled Server Close

Luckily, we have images of the prototype for this server so we can show you what is inside these systems. Aside from the 8 GPU NVIDIA HGX tray that uses custom Supermicro liquid cooling blocks, the CPU tray shows why these are a next-level design that is unmatched in the industry.



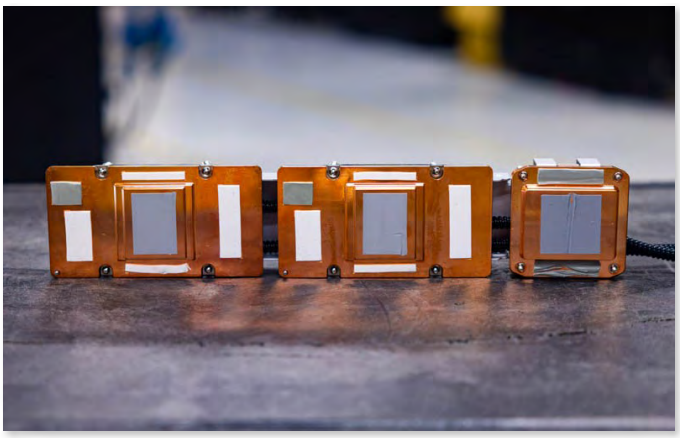
Supermicro 4U Universal GPU System For Liquid Cooled NVIDIA HGX H100 And HGX 200 At SC23 3



Supermicro 4U Universal GPU System For Liquid Cooled NVIDIA HGX H100 And HGX 200 At SC23

The two x86 CPU liquid cooling blocks in the SC23 prototype above are fairly common. What is unique is on the right-hand side. Supermicro's motherboard integrates the four Broadcom PCIe switches used in almost every HGX AI server today instead of putting them on a separate board. Supermicro then has a custom liquid cooling block to cool these four PCIe switches. Other AI servers in the industry are built, and then liquid cooling is added to an air-cooled design. Supermicro's design is from the ground up to be liquid-cooled, and all from one vendor.

At the bottom of the rack, we have the CDUs or coolant distribution units. These CDUs are like giant heat exchangers. In each rack, there is a fluid loop that feeds all of the GPU servers. We are saying fluid, not water, here because usually, these loops need fluid tuned to the materials found in the liquid cooling blocks, tubes, manifolds, and so forth. We have articles and videos on how data center liquid cooling works if you want to learn more about the details of CDUs and fluids.



XAI Colossus Data Center Supermicro 4U Universal GPU Liquid Cooled Servers Rear

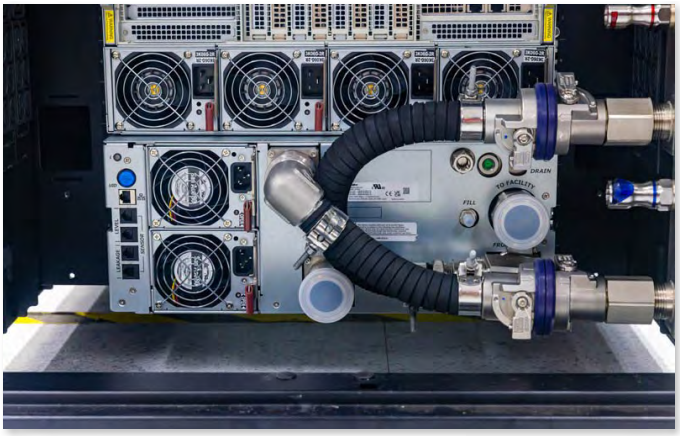
Each CDU has redundant pumps and power supplies so that if one of either fails, it can be replaced in the field without shutting down the entire rack. Since I had replaced a pump in one of these before, I thought about doing it at Colossus. Then I thought that might not be the wisest idea since we already had footage of me replacing a pump last year.

XAI Colossus Data Center Supermicro CDU On the Bottom Of The Rack

Patrick Removing A Supermicro CDU Pump

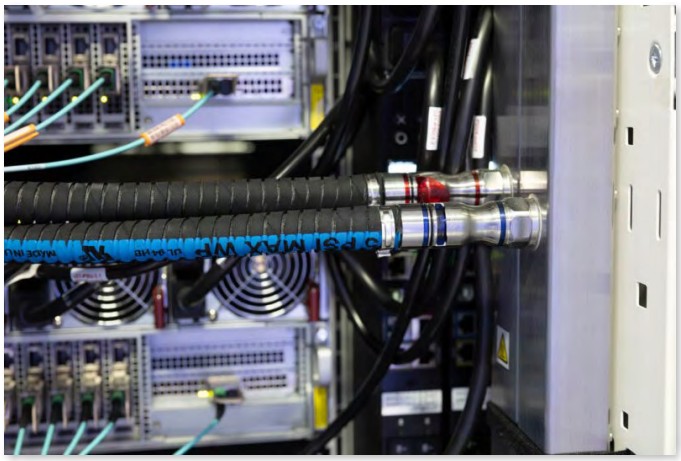
The xAI racks have a lot going on, but while filming the 2023 piece, we had a clearer shot of the Supermicro CDU. Here, you can see the input and output to facility water and to the rack manifold. You can also see the hot-swappable redundant power supplies for each CDU.

Here is the CDU in a Colossus rack hidden by various tubes and cables.



Supermicro CDU 2023 Rear 1

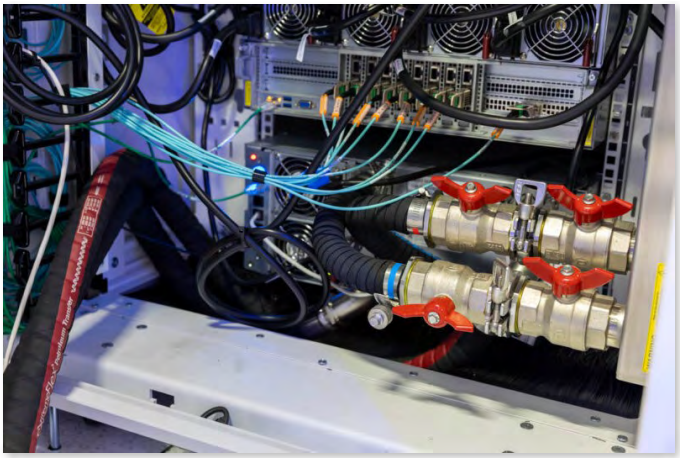
On each side of the Colossus racks, we have the 3-phase PDUs as well as the rack manifolds. Each of the front mounted 1U manifolds that feed the 4U Universal GPU systems, is in turn fed by the rack manifold that connects to the CDU. All of these components are labeled with red and blue fittings. Luckily, this is a familiar color coding scheme with red for warm and blue for cooler portions of the loop.



XAI Colossus Data Center Supermicro Rack Manifold Hoses

While the rear door heat exchangers may sound fancy, they are very analogous to a radiator in a car. They take exhaust air from the rack and pass it through a finned heat exchanger/radiator. That heat exchanger has liquid flowing through it, just like the servers, and the heat can then be exchanged to facility water loops. Air is pulled through via fans on the back of the units. Unlike most car radiators, these have a really slick trick. In normal operation, these light up blue. They can also light up in other colors, such as red if there is an issue requiring service. When I visited the site under construction, I certainly did not turn on a few of these racks, but it was neat to see these heat exchangers, as they were turned on, go through different colors as the racks came online.

These rear door heat exchangers serve another important design purpose in the data halls. Not only can they remove the miscellaneous heat from Supermicro's liquid cooled GPU servers, but they can also remove heat from the storage, CPU compute clusters, and networking components as well.



XAI Colossus Data Center Supermicro CDU Rear

Something you are likely to have noticed from these photos is that there are still fans here. Fans are used in many liquid-cooled servers to cool components like the DIMMs, power supplies, low-power baseboard management controllers, NICs, and so forth. At Colossus, each rack needs to be cooling neutral to the data hall to avoid installing massive air handlers. The fans in the servers pull cooler air from the front of the rack, and exhaust the air at the rear of the server. From there, the air is pulled through rear door heat exchangers.

XAI Data Center Tour Rear Door Heat Exchangers

XAI Data Center Tour Rear Door Heat Exchanger

Supermicro-based Storage at xAI

Storage was really interesting. In AI clusters, you generally see large storage arrays. Here, we had storage software from different vendors running, but almost every storage server we saw was Supermicro as well. That should not be a surprise. Supermicro is the OEM for many storage vendors.

One aspect that was very neat to see while we toured the facility was how similar some of the storage servers look to the CPU compute servers.

XAI Colossus Data Center Supermicro 1U NVMe Storage Node

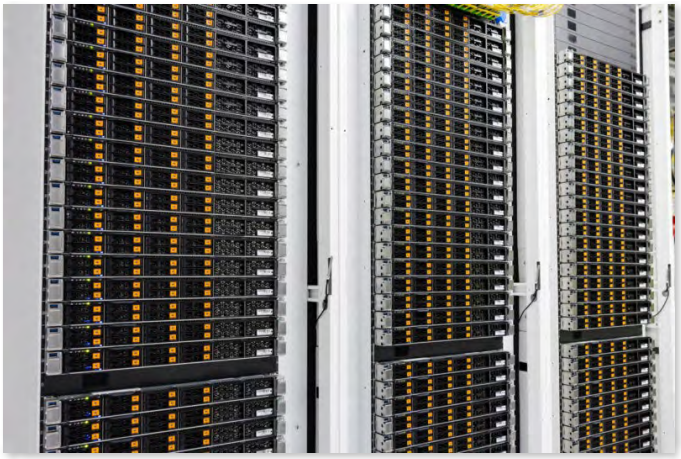
XAI Colossus Data Center Supermicro 1U NVMe Storage Node 2

In either case, you will see a lot of 2.5" NVMe storage bays in our photos and video. Something we have covered on our Substack is that large AI clusters have been moving away from disk-based storage to flash because it can save significant amounts of power while offering more performance and more density. Flash can cost more per petabyte, but in clusters of this scale, flash tends to win on a TCO basis.

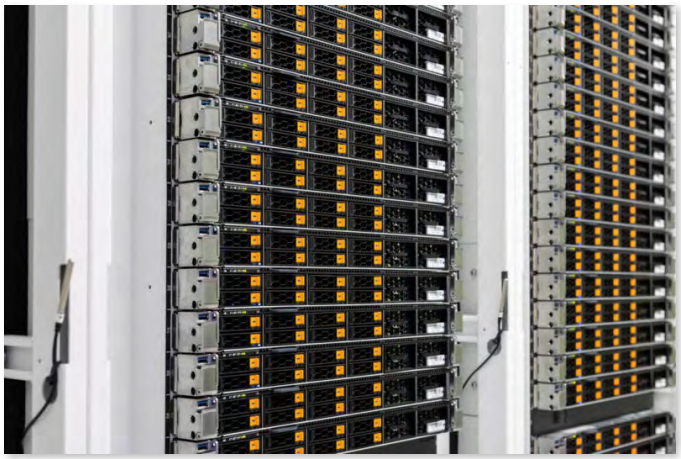
Supermicro-based CPU Compute at xAI

With all of these clusters, you generally see a solid number of traditional CPU compute nodes. Processing and data manipulation tasks still run very well on CPUs versus GPUs. You may also want to keep the GPUs running AI training or inference workloads instead of other tasks.

Here, we see racks of 1U servers. Each of the servers is designed to balance compute density with the heat being generated. A great example of this is that we can see the orange tabs for NVMe storage bays on front but also about a third of the faceplate being dedicated to drawing cool air into the system.



XAI Colossus Data Center CPU Compute Rack



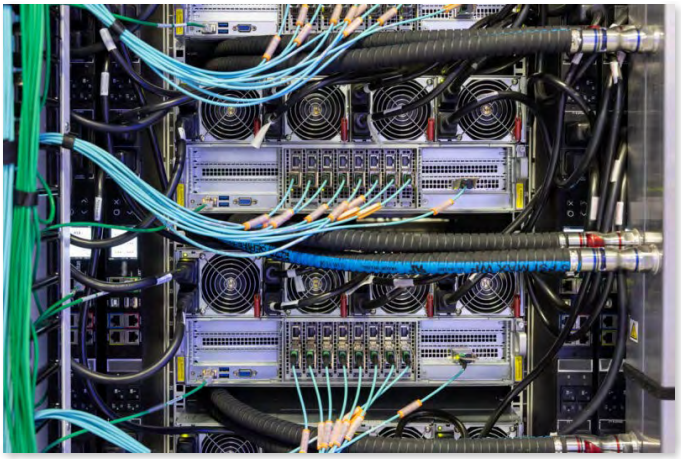
XAI Colossus Data Center CPU Compute Rack 2

These 1U compute servers can be cooled by fans and then a rear door heat exchanger can remove heat and exchange it with the facility water loops. Due to the data center design with rear door heat exchangers, xAI can handle both liquid-cooled gear and air-cooled gear.

Networking at xAI Colossus

Networking is one of the fascinating parts. If your computer uses an Ethernet cable, that is the same base technology as the networking here. Except, that this is 400GbE or 400 times faster, per optical connection than the common 1GbE networking we see elsewhere. There are also nine of these links per system which means that we have about 3.6Tbps of bandwidth per GPU compute server.

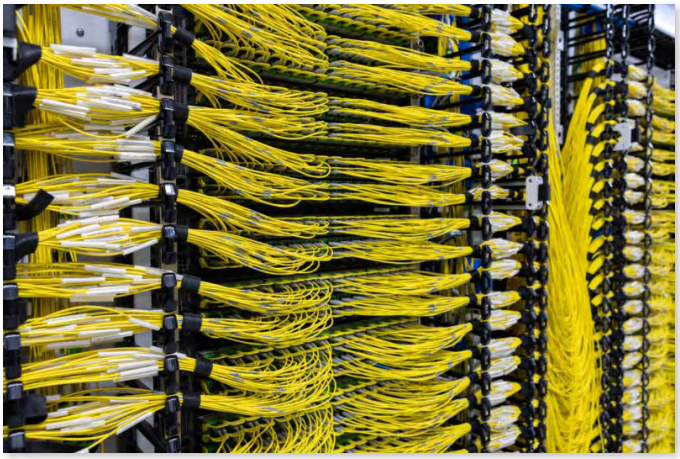
The RDMA network for the GPUs makes up the majority of this bandwidth. Each GPU gets its own NIC. Here, xAI is using NVIDIA BlueField-3 SuperNICs and Spectrum-X networking. NVIDIA has some special sauce in their network stack that helps ensure the right data gets to the right place navigating around bottlenecks in the cluster.



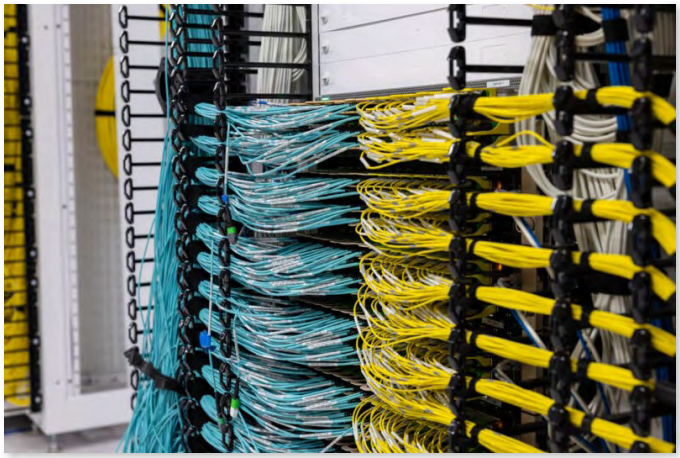
XAI Colossus Data Center NICs

That is a big deal. Many supercomputer networks use InfiniBand or other technologies, but this is Ethernet. Ethernet means it can scale. Everyone reading this on STH will have the page delivered over an Ethernet network at some point. Ethernet is the backbone of the Internet. As a result, it is a technology that is immensely scalable. These enormous AI clusters are scaling to the point where some of the more exotic technologies have not touched in terms of scale. This is a really bold move by the xAI team.

Beyond the GPU RDMA network, the CPUs also get a 400GbE connection, which uses a different switch fabric entirely. xAI is running a network for its GPUs and one for the rest of the cluster, which is a very common design point in high-performance computing clusters.



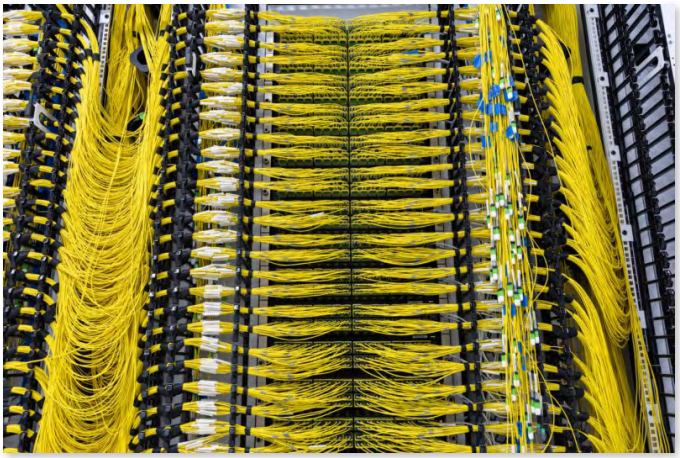
XAI Colossus Data Center Switch Fiber 1



XAI Colossus Data Center Single Mode And Multi Mode Fiber

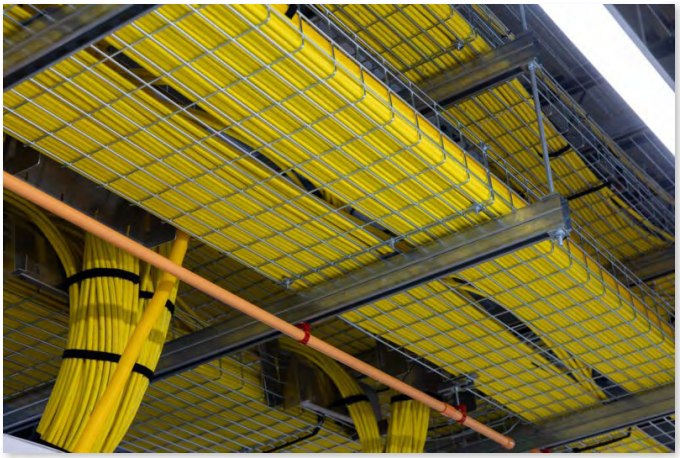
Just to give you some sense of how fast 400GbE is, it is more connectivity than a top-of-the-line early 2021 Intel Xeon server processor could handle across all of its PCIe lanes combined. That level of networking is being used nine times per server here.

All of that networking means that we have huge amounts of fiber runs. Each fiber run is cut and terminated to the correct length and labeled.

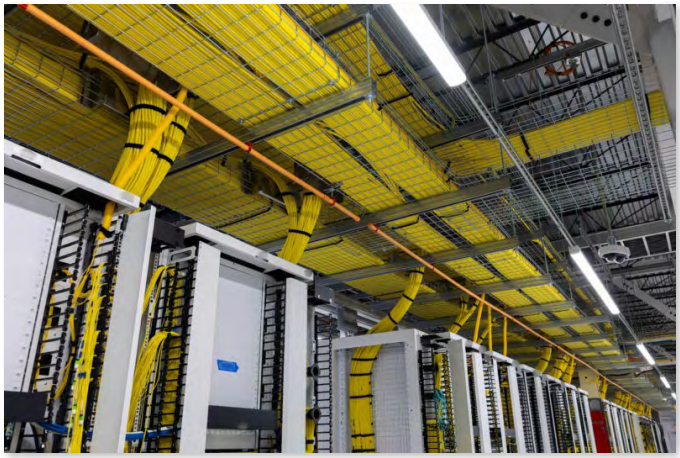


XAI Colossus Data Center Switch Stack

I had the opportunity to meet some of the folks doing this work back in August. Structured cabling is always neat to see.



XAI Colossus Data Center Fiber Runs



XAI Colossus Data Center Overhead Cabling

In addition to the high-speed cluster networking, there is lower-speed networking that is used for the various management interfaces and environmental devices that are a part of any cluster like this.

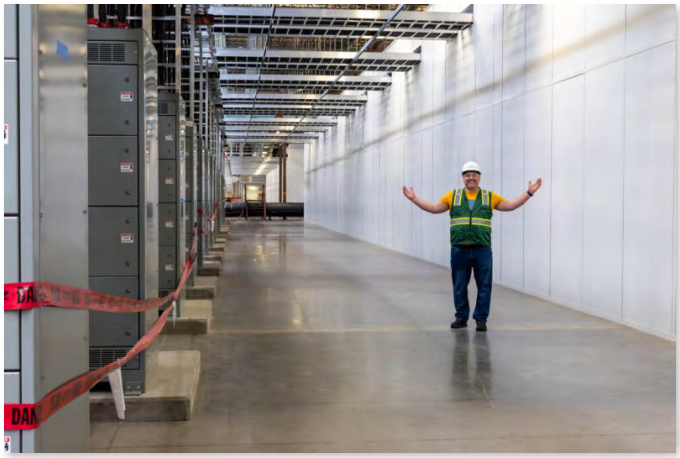
Something that was very obvious walking through this facility is that liquid-cooled network switches are desperately needed. We recently reviewed a 64-port 800GbE switch, in the same 51.2T class as the ones used in many AI clusters. Something that the industry needs to solve is cooling not just the switch chips, but also the optics that in a modern switch can use significantly more power than the switch chip. Perhaps enormous installations like these might move the industry towards co-packaged optics so that the cooling of the switches can follow the compute to liquid cooling. We have seen liquid-cooled co-packaged optic switch demos before, so hopefully a look at this installation will help those go from prototypes to production in the future.

The Colossus Facility

Since we have liquid-cooled racks of AI servers, the power and facility water is essential to the installation. Here is a look at the massive water pipes. There are sets of cooler and warmer water. Cooler water is brought into the facility and circulates through the CDU in each rack. Heat is transferred from the GPUs and rear door heat exchanger loops to the facility water loops at the CDU. The warmer water is then brought outside the facility to chillers. Of course, the chillers are not the type that will make you ice cubes. Instead, the goal is just to lower the temperature of the water enough so that it cools down enough to be recycled through the facility again.

XAI Colossus Data Center Facility Water Pipes

Power is fascinating. When we were in Memphis while the system was built, we saw the teams moving huge power cables into place.



XAI Colossus Data Center Some Electrical Infrastructure With Patrick For Scale

Outside of the facility, we saw containers with Tesla Megapacks. This is one of the really neat learning points that the teams had building this giant cluster. AI servers do not run at 100% rated power consumption 24×7. Instead, they have many peaks and valleys in power consumption. With so many GPUs on site, the power consumption fluctuates as the workload moves to the GPUs, and then results are collated, and new jobs are dispatched. The team found that the millisecond spikes and drops in power were stressful enough that putting the Tesla Megapacks in the middle to help buffer those spikes in power helped make the entire installation more reliable.

Tesla Megapacks Ready For Installation At XAI Colossus

Of course, the facility is just getting started. While the initial cluster of four 25,000 GPU data halls is up and running for around 100,000 GPUs at the time of our visit, the cluster expansion work is moving rapidly.

This seems to be the start of something truly awesome.

Final Words

One of the key themes I learned while doing this is that the xAI team needs more time for petty vendor differences. The only way this got built was a surge of experts building the systems together with a vision of building a giant AI cluster at an unheard-of speed. If I had just seen it the day we filmed the video, I would have had a different perspective on how many people were working together to build something of this scale. It was cool going on-site both times and having folks come up to me and tell me they have been avid readers or viewers of STH for so long.

XAI Colossus Data Center In Memphis Outside Expansion

XAI Data Center Tour Patrick In Compute Hall

If you want to get involved in this project or large AI installations, check out the job postings at xAI and Supermicro. I hear folks in the AI community talk about how LLMs continue scale with more compute and how they can be generally applicable than just for chatbots. As I walked around Colossus, one thought I had is that something of this scale only gets built if data-driven folks see huge value on the horizon. Grok and the xAI team's future work feels destined to be much more than a simple 2024-era chatbot. A lot of very smart people are spending a lot of money and spending their time to make that happen as fast as possible. We have come a long way since I first fielded the call on this from the hospital the day after my son was born. In the end, it was a fantastic experience to see this get built. Thank you to all of those who went out of their way to make this possible.

If you are working on a large AI cluster, let us know. It is exciting to see what will happen next.

If you want to learn more, here is the Supermicro AI link and the company's landing page for the AI Supercluster. Or, just watch the video.