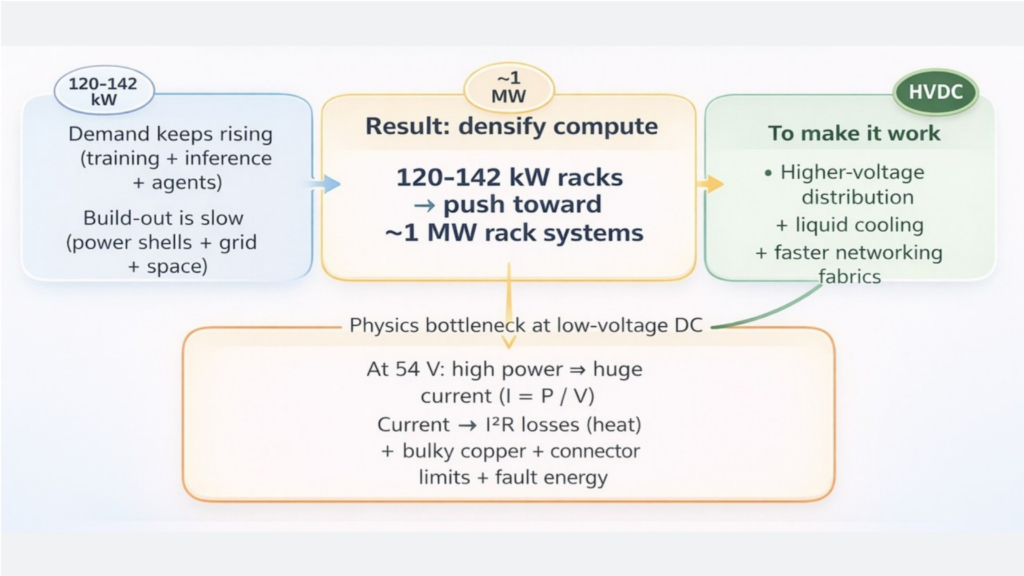


World of AI Factory: Episode 1 — From 120 kW Racks to Megawatt Racks: The Economics, Physics, and “AI Factory” Trajectory

Tamoghna Sengupta • Published 2026-01-14 • Source: https://www.linkedin.com/pulse/world-ai-factory-episode-1-from-120-kw-racks-megawatt-sengupta-9wnxc/ • PDF generated: 2026-01-15

The Key Question:

At this point it is no longer about selling more GPUs alone. It is now about exploring optimizing and shaping an architectural idea: AI factories. The key aspect which shapes up the idea of AI factory is, how do data centers optimize power, cooling and memory to produce tokens (inference) and model updates (training) at industrial scale? How can the “AI Factory” digest megawatt (MW)-scale racks, which currently operates at 120-142 kW?

Current Baseline:

To begin with, let us anchor around what we have today:

- ~120 kW-class (GB200 NVL72): Oracle describes NVIDIA GB200 NVL72 racks drawing “over 120 kW at peak.” Oracle Blog
- · Up to ~142 kW-class (GB300 NVL72): Vertiv’s reference architecture for NVIDIA GB300 NVL72 explicitly supports rack densities up to 142 kW. Vertiv news blog

Now, let us work upwards with the above numbers as anchors.

What is a “frontier model”?

Capability-Based

The UK AI Safety Summit materials describe "frontier AI" as highly capable general-purpose models that match or exceed the most advanced current systems. UK Government AI Safety Summit

Compute-based (regulatory/threshold language)

A commonly used policy threshold is > 10^{26} integer or floating-point operations (FLOPs) as a "frontier model" criterion. Institute For Law and AI

None of the above definitions are perfect, but, if you consider the two definitions with 'OR' criteria, then it gets very near to the concept of frontier model.

A logical way to think is - from the parlance of frontal lobe of your brain. A frontier model is the base intelligence which serves as an underlying basic premise of the tasks and thinking which can be performed using the model.

The widely recognized frontier models which exist so far are : GPT 5, Claude 4, Gemini 2.x, Llama 4, Grok-4, Mistral Large 2, DeepSeek V3.x, Qwen2.5-Max, GLM-4.5, Command R+ etc.

Case Study: Grok-4

Epoch AI estimated Grok-4's training resources at 246 million H100-hours , assuming H100 SXM with 700 watts peak, plus 2.03× for non-GPU IT hardware and 1.2× for Power Usage Effectiveness (PUE). Epoch AI

Let us deduce this step by step:

Step 1: GPU energy at 100% power (upper bound)

- Peak power per H100: 700 W = 0.700 kW Epoch AI
- GPU energy = 246,000,000 hours × 0.700 kW = 172,200,000 kWh = 172.2 GWh (gigawatt-hours)

Step 2: GPU energy at ~75% average draw (more realistic)

Epoch uses ~75% of average training draw. Epoch AI

- Average power per H100 ≈ 0.700 × 0.75 = 0.525 kW
- GPU energy = 246,000,000 × 0.525 = 129,150,000 kWh = 129.15 GWh

Step 3: Plus, non-GPU IT hardware (networking, CPUs, memory, etc.)

Epoch multiplies by 2.03× to include non-GPU hardware power. Epoch AI

- IT energy = 129.15 GWh × 2.03 = 262.17 GWh

Step 4 — Add facility overhead (cooling + power conversion losses)

Epoch uses Power Usage Effectiveness (PUE) = 1.2× . Epoch AI

- Facility energy = 262.17 GWh × 1.2 = 314.61 GWh (~315 GWh)

But, to calculate the IT rack power consumption, let us exclude the facility overhead, as that is not directly part of what the racks contribute to.

Turning energy into rack-hours implied number of racks

262.17 GWh = 262,170,000 kWh

Rack-hours for 120 kW racks

- Rack-hours = 262,170,000 kWh ÷ 120 kW = 2,184,750 rack-hours

Rack-hours for 142 kW racks

- Rack-hours = 262,170,000 ÷ 142 = 1,846,268 rack-hours

Thus, based on assumed training duration, below are the number of racks used:

Assumed Training Duration	Hours	Number of 120 kW Racks	Number of 142 kW Racks
30 days	720	~3,035 (2,184,750/720)	~2564 (1,846,268/720)
90 days	2,160	~1,012 (2,184,750/2160)	~855 (1,846,268/2160)

Now, if the same racks would have 1 MW capacity, the number of racks required will come down significantly:

Rack-hours = 262,170,000 ÷ 1,000 = 262,170 rack-hours

Assumed Training Duration	Hours	Number of 1 MW Racks
30 days	720	~364
90 days	2160	~121

This is the first economic argument for megawatt-class systems: same work, fewer rack-units to house, cool, cable, and network.

Why 142 kW --> 200 kW --> megawatt-class racks journey is possible?

The market consistently pressures providers to improve latency (response time) while also improving quality, especially in paid/enterprise settings.

OpenAI has explicitly reported that performance of its reasoning model improves both with train-time compute (more reinforcement learning) and with more test-time compute (more time spent thinking at inference).

Thus, users want lower latency and higher reasoning quality, which forces providers to spend more computing per second (more parallelism, more capacity) to deliver better answers without waiting.

Additionally, the below factors are working together with significant force:

a.) More AI minutes per person per day (copilots and agents embedded into work)

b.) Heavier inference (reasoning/agentic workflows = more iterative steps per task; vendors explicitly framing around “AI reasoning”)

c.) Capacity pressure (demand outrunning supply makes “more compute per rack” economically attractive)

Thus,

When AI becomes embedded in daily work, total demand often rises in two ways:

i.) More sessions (more people using it)

ii.) More simultaneity (more people using it at the same time)

If you have unlimited space and power delivery, you can meet that by adding racks (scale-out).

But the industry is increasingly constrained by power availability, grid interconnect timelines and physical footprint.

Microsoft has explicitly pointed to build-out limits and power grid constraints affecting capacity.

Thus, moving from 142 kW => > 200 kW rack directionally provides the following benefits:

1.) You can cramp higher compute into lesser space, which addresses the physical footprint limitation.

2.) When you cramp higher computer into a single rack, it means fewer top-of-rack switches, fewer optics, shorter cables and lesser scale out (costlier) network hops

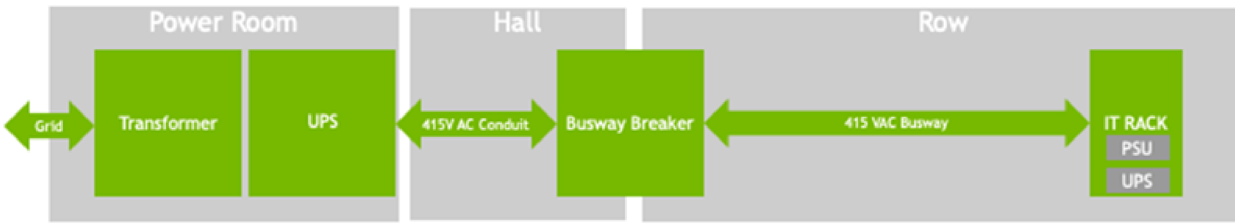
3.) When you compare NVL72 with 2xNVL36, NVL72 involves less NV Switch and involves less cross-rack interconnect complexity, thus improving the compute delivered per MW at cluster level.

4.) Big training runs involve collective communication among racks, involving all reduce, parameter sync. When the number of racks is reduced, you reduce collective communication and sync overhead, thus improving overall efficiency.

The hidden roadblocks: current, copper, and why 54 volts hit a wall

At 54 volts direct current (DC), 142 kW implies:

I = P / V = 142,000 / 54 ≈ 2,630 amps



Current data center power architecture (courtesy : NVIDIA_Blog)

The above picture depicts:

1. The building feeds power at high voltage (hundreds of volts, usually alternating current — AC) via grid and transformers.

2. Near the rack (or in the rack), power supplies convert that to low-voltage direct current (DC) like ~48/54 V to run electronics.

3. Once you step down to 54 V, the current becomes enormous for the same power.

So, the “2,630 amps” pain is mostly an inside-the-rack / near-the-rack problem, after the voltage has been stepped down.

Why is “2,630 amps” at for 142 kW IT rack already a problem?

Point#1: Scaling issue of rack power interconnects

One connector path can handle a few hundred amps safely and continuously not thousands. Thus, if the IT rack needs 2630 amps at 54 V, by design the racks cannot be connected to one another using one connector or one bar unless the connector is unrealistically humungous.

In today’s world, one power path can handle ~500 amps OpenComputeBlog . Now, to connect 2630-amp IT Racks, you need 2630/500 = ~ 5.3, i.e. at least 6 parallel paths.

And in real designs you often need more than the bare minimum because of:

- Redundancy (if one path fails, the rack shouldn’t die),
- Thermal margin (things run cooler and last longer),
- Safe connectors, Safe busbars, Safe protection gear.

Point#2: Demand for Copper conductors go through the roof

Currently, a 500 kcmil (Kcmil (kilo circular mils) is a unit for measuring the cross-sectional area of large electrical conductors; 1 kcmil=0.5067 mm2) copper conductors are typically shown around ~380 A.

So, for 2630 A, one would expect:

- 2,630 / 380 ≈ 6.9 → 7 parallel conductors ,
- and another ~7 for the return
- and then add more if you want N+1 redundancy.

Roughly, it translates to 14 very thick cables just to feed power into one rack at a low voltage.

Point#3: Heat from Resistance.

Even a tiny resistance becomes big heaters when amps are huge. Think of it this way – water losses a bit of pressure as it flows through the pipe, similarly, current loses a bit of energy in the form of heat when it moves through metal conductor.

The following formula illustrates how much power is lost by amperes, while moving through a metal conductor:

```
Loss = (Current)² × Resistance
```

Thus, even if resistance is minimum, considering the metal conductor is of high quality, loss quadratically increases as current increases:

- double the current → 4× the heating
- 10× the current → 100× the heating

Point#4: Protection and fault energy

54 VDC is relatively safe for electric shock, but 2,630 A is an industrial fault-energy regime . Short-circuit currents can be enormous, and interrupting DC faults cleanly is harder than AC (DC arcs do not naturally cross zero every cycle).

Is “megawatt racks” a natural next step?

Megawatt-class racks make sense if these conditions hold:

- AI demand keeps growing (training + inference + agents) faster than new powered shells can be delivered.

When training, inference, and agents scale up, it can outpace the ability to build powered shells (utility interconnect + substation/transformers/switchgear + networking + cooling). Hence, the immediate ask goes: “can we get more useful compute per unit of delivered site capacity and floor space, faster?”, leading to packing more compute within fewer, denser rack-systems.

- Networking and cooling architectures keep improving so ultra-dense systems don’t choke on heat or communication.

Yes, the dissipation of extra heat does not go away just by reducing the amperes flowing through the racks. If racks are fed with megawatt-class power, that power gets consumed by the compute and power electronics in the rack and almost all of it ends up as heat. What does change is that higher-voltage delivery reduces I^2R (current-squared times resistance) losses in the cables/busbars/connectors, so you waste less power as heat in the delivery path (and you reduce some hotspot risk at joints). But the main heat is still very concentrated at the chips/cold plates, and that concentration can be harder to remove, which is why better liquid cooling and rack-level thermal design matter.

- Power delivery shifts (e.g., toward 800 VDC HVDC) so current don’t become absurd.

The final incentive to make the IT racks more powerful is more obvious than the rest – physical and electrical constraints. When you concentrate compute, you otherwise end up with an absurd amount of copper, busbars, connectors, and conversion stages to move power at low voltage. Moving to higher-voltage distribution (for example, 800 VDC high-voltage direct current) drops current for the same power, which reduces conductor size/weight and I^2R losses, and can simplify the conversion chain (fewer power-conversion boxes burning watts as heat). This is exactly why multiple ecosystem players are talking about 800 VDC approaches for megawatt-class rack systems.

Proposed NVIDIA 800 VDC architecture which minimizes energy conversions. (courtesy : Nvidia blog)

Further read: company filings and earnings call transcripts

Disclosure: This is a learning-oriented, primary-sources reading list to help readers understand the concepts and market structure discussed in this episode. The companies below are included only as examples of publicly traded firms that publish detailed disclosures relevant to these topics (10-K / 20-F / annual reports and earnings call transcripts). This is not a recommendation to buy, sell, or hold any security. I am not a registered investment adviser or broker-dealer, and nothing in this article is individualized investment advice. Inclusion is not an endorsement of any company’s prospects, valuation, or suitability for any investor.

1. Microsoft (MSFT) : AI factory demand + capital expenditure (CapEx) signals
2. Eaton (ETN) : switchgear, power distribution, and electrical infrastructure
3. Vertiv (VRT) : data-center power + thermal infrastructure
4. Monolithic Power Systems (MPWR) : high-density power conversion inside the rack
5. Amphenol (APH) : connectors/cabling