

# Data Center Briefing

March 17, 2026

Global

## Key themes:

Meta five-year up to \$27B Nebius Rubin capacity; Digital Edge \$665M green loan for 500MW Bekasi campus; Dell-NVIDIA rack-scale IR9000 with 409.6 Tb/s switching; Gorilla-Yotta deploy 5,000+ B200 GPUs in Navi Mumbai

Meta just put a giant number on the board: a five-year agreement with Nebius Group worth up to \$27bn, including about \$12bn in dedicated Nvidia Rubin-based AI infrastructure that won't start showing up until early 2027. That's a long-dated capacity reservation at a moment when everyone is pretending supply chains and power queues will politely sort themselves out. Read it as a signal: the biggest buyers are getting comfortable underwriting "Rubin-era" scarcity now, not when the racks land.

## The Big Stories

[Meta signs up to \\$27B Nebius AI infrastructure deal](#) is the cleanest expression yet of where AI infrastructure is heading: fewer spot purchases, more multi-year capacity lockups with a clear GPU generation attached. The structure matters—~\$12bn of dedicated Rubin capacity plus up to \$15bn of additional purchases—because it gives Meta early access while still leaving room to flex. The tell is the timing: clusters "expected to start coming online in early 2027," which implies the real bottleneck isn't just chips, it's the end-to-end ability to stand up the full stack on schedule.

On the supply side, [Dell and NVIDIA deliver integrated AI rack-scale infrastructure](#) is Dell trying to industrialise the AI factory build: co-engineered rack-scale offerings, the PowerEdge XE9812 with "72-way GPUs," XE988x

servers with HGX Rubin NVL8, IR9000 factory racks, and PowerSwitch SN6000 switches pushing up to 409.6 Tb/s. Dell is also leaning into NVIDIA Confidential AI certification and liquid-cooled Ethernet/InfiniBand options. The competitive implication is straightforward: server OEMs want to sell outcomes (validated racks and designs) rather than boxes—because the buyer’s problem is integration speed and predictability, not component selection.

Capital is still flowing to big campuses in emerging hyperscale markets. [Digital Edge secures \\$665M green loan for 500MW CGK Campus](#) funds the first phase of a planned 500MW campus in Bekasi, Indonesia, within a broader multi-phase plan pegged at \$4.5bn. The financing is framed as “green,” and it’s backed by a heavyweight bank syndicate. The message for investors is that lenders are willing to underwrite very large, AI-ready builds—so long as the narrative around carbon neutrality (Digital Edge targets 2030) and project structure is tight.

India keeps stacking AI compute announcements—this time with a real deployment commitment. [Gorilla to deploy 5,000+ GPUs with Yotta in India](#) covers binding agreements to place ~640 NVIDIA HGX B200 servers (more than 5,000 GPUs) at Yotta’s Uptime Tier IV NM1 facility in Navi Mumbai, with Gorilla expecting the deployment to contribute more than \$500m in revenue over five years. It’s also explicitly positioned as a platform that could expand to “>5,000 additional servers,” with potential projects in Thailand. For anyone tracking regional compute hubs, the important bit is that this is framed as longer-duration infrastructure partnership serving enterprise and government—i.e., sticky demand, not just opportunistic resale.

Meanwhile, policymakers are trying to catch up with what “AI load” actually means for the grid. At a technical level, [DOE workshop addresses integrating AI data centers with grid](#) reads like a to-do list for the next build cycle: DC power architectures, flexible microgrids, real-time demand modeling, plus supply-chain and workforce development. Oak Ridge also announced its Next-Generation Data Center Institute. The subtext is that data centers are no longer “just another large customer”—the DOE is treating integration, security, and controllability as first-order national infrastructure questions.

## Behind the Headlines

Power architecture is being dragged into the GPU era whether operators like it or not. [Delta unveils 800 VDC power and liquid cooling for AI data centres](#) is notable because it's not a single product—it's an 800 VDC “ecosystem,” including a 660kW in-row rack with 480kW embedded battery backup, CDUs at 2.4MW and 3MW, and an 800 VDC microgrid package that even namechecks solid-state transformers, solid oxide fuel cells, and energy storage. If 800 VDC becomes mainstream, it will ripple through everything: electrical room design, safety standards, vendor ecosystems, and (crucially) retrofit viability. This is the kind of shift that doesn't happen overnight, but once it starts, it forces a re-think of how you deliver power at extreme rack densities.

Operations is becoming software, not just headcount. [Salute and Phaidra launch AI-scale data center operations partnership](#) explicitly targets liquid-cooled, GPU-based “AI factories (100MW+),” pairing Salute's facilities management execution with Phaidra's AI control systems, with promised pilot results in 60–90 days. The eyebrow-raiser is the cited “30% cooling energy reduction at Google” as proof point. Whether that exact result generalises is less important than the direction: as cooling loops and control complexity multiply, operators will be pressured to run facilities more like closed-loop industrial systems—with performance guarantees, not just best-efforts tuning.

Physical security is starting to sound less theoretical. [Drone attacks on data centers spotlight physical security needs](#) reports drone attacks targeting AWS data centers in Dubai and Bahrain attributed to Iran, and pushes a practical list—layered defenses, drone detection, AI video analytics, and regular threat/risk assessment. For the market, the point isn't the vendor advice; it's that geopolitical spillover is now being discussed in direct connection with named data center locations. That tends to accelerate security capex, complicate site selection, and raise hard questions about how “sovereign” and “resilient” claims hold up under real-world threats.