



Netzanschluss für KI

Ein Einblick in die Energieversorgung für den Ausbau hochperformanter Rechenzentren in Deutschland für die massive Nutzung generativer KI

Rechenzentren bilden das Rückgrat der digitalen Infrastruktur. Sie ermöglichen Cloud-Dienste, KI-Anwendungen, Datenverarbeitung, Speicherung und Kommunikation. Doch vor allem bei der massiven Nutzung zunehmend populärer Anwendungen der generativen Künstlichen Intelligenz steigt der Bedarf an Computingleistung in modernen Rechenzentren. Um diesem Bedarf gerecht zu werden und um digitale Souveränität zu gewährleisten, werden große Investitionen in sogenannte High-Performance Computing Rechenzentren (HPC RZ) - im US-amerikanischen Raum sogenannte Hyperscaler - vorgenommen. Mit wachsender Digitalisierung steigt auch deren Energiebedarf und damit die Bedeutung energieeffizienter Technologien. In Deutschland gab es 2025 über 1.600 unternehmenseigene Rechenzentren, über 300 sogenannte Colocation-Rechenzentren (von Drittanbietern). Dazu gehören einige, die sich in der Fertigstellung befinden wie: CyrusOne FRA VII (72 MW IT), STACK FRAL1 in Liederbach (80 MW IT), Cyrus One FRA V Hanau (45 MW IT), Vantage Offenbach (24 MW IT), Digital Reality Fechenheim (140 MW IT)).

Zur Verbesserung der Energieeffizienz auf Computingebene sind zwar die Möglichkeiten, wie In-Network-Computing, Optical-Switching-Technologien, Workload-Optimierung und fortschrittliche Kühlmethoden noch nicht vollkommen ausgeschöpft, doch die Energieversorgung bleibt ein zentraler Bestandteil der Systemoptimierung.

Der Betrieb von HPC-RZ stellt höchste Anforderungen an Zuverlässigkeit und Versorgungssicherheit. Um eine Verfügbarkeit von mindestens 99,999 % zu gewährleisten, ist insbesondere die kontinuierliche und leistungsfähige Kühlung der Server essentiell, da bereits kurzfristige Überhitzungen zu kritischen Ausfällen führen können. Dies bezeichnet man als „Five-Nines-Konzept“ der Optimierung.

Daraus ergeben sich erhöhte Anforderungen sowohl an die elektrotechnische Infrastruktur als auch an die Kühlsysteme: Beide müssen redundant ausgelegt sein, um auch bei Teilausfällen einen unterbrechungsfreien Betrieb sicherzustellen. Zusätzlich sind ausreichende Leistungsreserven im Netzanschluss vorzuhalten, um auch bei extremen Wetterbedingungen, etwa hohen Außentemperaturen in Kombination mit ungünstigen Windverhältnissen, jederzeit die notwendige Kühlleistung - und damit den sicheren Betrieb - gewährleisten zu können.

Kapazitätsengpässe, insbesondere in Ballungsräumen, Systemsicherheit und Planungsprioritäten sind Themen, die sowohl auf technischer Ebene, aber eben auch im wirtschaftspolitischen Kontext einer sich neu aufbauenden Digitalindustrie im Fokus der gesellschaftlichen Diskussion stehen.

Mit dieser VDE Kurzinformation möchten wir die aktuellen technischen und wirtschaftlichen Gegebenheiten darstellen und wagen eine Analyse insbesondere im Vergleich mit den Aktivitäten und Rahmenbedingungen in den USA.

Anforderungen an den Netzanschluss von HPC RZ

Mit dem Stromnetz in Deutschland muss man in Folge der Energiewende zahlreiche dezentrale, erneuerbare Energiequellen einbinden, bidirektionale Energieflüsse bewältigen und dabei Angebot und Nachfrage in Echtzeit ausgleichen.

Der Netzausbau kommt jedoch oft nur langsam voran. Häufig sind es planungsrechtliche und regulatorische Herausforderungen, die zu Verzögerungen führen.

Nicht zuletzt ist damit der Stromverbrauch der KI von HPC RZ längst keine alleinige technische Planungsaufgabe mehr, sondern bestimmt den gesellschaftlichen und vor allem den wirtschaftlichen Diskurs. Dies wird insbesondere motiviert durch massive Investitionen, die in den USA¹ und Asien vorgenommen werden. Bei dieser hohen Dynamik möchte man wirtschaftlich nicht ins Hintertreffen geraten und die bisherigen Abhängigkeiten in der Digitalisierung, insbesondere von den USA, lösen. Dies liegt daran, dass das Thema zum zentralen Hebel für die technische Unabhängigkeit hinsichtlich der Verfügbarkeit von digitalen Services wie Computingressourcen, Kommunikationsinfrastruktur und Datenspeicher und letztlich für die Reputation der Rechenzentrenbetreiber geworden ist.² Daher ist die Sichtweise auf den aktuellen Ausbau von Rechenzentren komplexer geworden und stellt die Investorinnen und Investoren sowie die Entscheidungstragenden vor Herausforderungen, die wir hier kurz skizzieren.

Zur Gewährleistung von Ausfallsicherheit sind typischerweise 2N oder N+1 Redundanzkonzepte auf Einspeise- und Netzseite vorzusehen. Eine redundante und diversifizierte Einspeisung erfordert zwei vollständig voneinander getrennte Mittelspannungseinspeisungen, idealerweise von unterschiedlichen Umspannwerken, welche die physische Trennung von Kabeltrassen, Schaltanlagen und Transformatorstationen sowie Selektivitätskonzepte zur Vermeidung von Kaskadenausfällen sicherstellen.

Da der Betrieb von HPC RZ kurze Umschaltzeiten und hohe Wirkungsgrade erfordert, ist eine USV-Architektur (unterbrechungsfreie Stromversorgung) mit hoher Effizienz und geringer Latenz notwendig. In Frage kommen hier z. B. online Doppelwandler USV, multimodale USV oder Lithium-Ionen-Batterien. Die USV hat nicht nur die Aufgabe, die Infrastruktur der Energieversorgung im Rechenzentrum aufrechtzuerhalten, sie muss auch empfindliche elektronische Geräte vor Störungen in der Energieversorgung schützen, z. B. vor Spannungseinbrüchen oder Spannungsspitzen. Ohne diesen Schutz können solche elektrischen Ereignisse zu Dienstunterbrechungen, Datenverlust oder Schäden an elektronischen Geräten führen.

Moderne KI-Anwendungen, insbesondere die Nutzung generativer KI, verursachen dynamische und unvorhersehbare Lastwechsel zwischen Rechenoperationen (Training / Inferenz) und Datenübertragungen. Für diese hochdynamischen Lasten in HPC RZ, bedingt durch Computing-Lastwechsel unterschiedlicher KI-Workloads, muss der Netzanschluss eine hohe Kurzschlussleistung bereitstellen, um Spannungsstabilität bei schnellen Laständerungen zu gewährleisten. Insbesondere ist die Energieaufnahme von Kühleinheiten bei schwankenden Umgebungstemperaturen ein Unsicherheitsfaktor, weswegen man die Kapazitätsauslegung höher plant, als im Durchschnitt benötigt wird. Gerade hier bietet sich Optimierungspotential, das wiederum durch den Einsatz physikalischer KI in schnellen Simulationen der technischen Umgebungen gehoben werden kann. Bereits jetzt nutzen Unternehmen KI für die Optimierung von Kühlkreisläufen in der Planung sowie im Betrieb von HPC RZ und erzielen damit höhere Wirkungsgrade, als bei einer konventionellen Planung.

Die enge Kopplung von Stromversorgung und Kühlstrategie ist ein zentraler Optimierungsgegenstand für mehr Energieeffizienz und thermische Stabilität in hochverdichteten HPC RZ.

Hohe Leistungsdichten erfordern eine optimierte Power Usage Effectiveness (PUE), kurze Strompfade und verlustarme Transformatoren. Die PUE ist eine technische Kennzahl, mit der sich die Energieeffizienz eines Rechenzentrums darstellen lässt. Sie bestimmt das Verhältnis zwischen der in Summe konsumierten elektrischen Energie (Gesamtenergieverbrauch) und der von der IT-Infrastruktur aufgenommenen elektrischen Energie (Server, Storage, Switches etc.) innerhalb einer definierten Periode.

Es entfallen ca. 60% der eingespeisten Primärenergie auf das Computing und auf die Kommunikationsinfrastruktur. Ein großer Skalierungsfaktor ist die benötigte Kühlung, deren Optimierungsmaßnahmen relevant sind, da die Schwankungen des zusätzlichen Energiebedarfs für die Kühlung von 7 % bis 30% betragen.³ Eine positive Entwicklung ist auch von dem Trend abzuleiten, dass zwar HPC RZ mehr Energiebedarf als konventionelle Cloud-Rechenzentren haben - und zwar bis zu viermal höher - jedoch steigt die Effizienz im Computing deutlich an und liegt je nach verbauter Technologie deutlich über dem Faktor vier.

Neben den genannten technischen Anforderungen, die in ein Pflichtenheft bei der Planung der Anbindung von Rechenzentren an den Netzanschluss eingehen, kommen organisatorische Faktoren hinzu, die eine wichtige Rolle für den Betrieb von HPC RZ spielen, wie z. B. das Monitoring von Lastprofilen und der Netzqualität, die Predictive Maintenance (Alterung von Komponenten) und zunehmend auch die KI-gestützte Vorhersage für Batterie- und USV-Zustände sowie KI-gestützte Betriebsstrategien der Gebäudetechnik.

Weitere zu beachtende Aspekte sind z. B. Lastmanagement und Demand Response Fähigkeit, die Integration erneuerbarer Energien, aber eben auch und gerade im deutschen Raum die Möglichkeit von Inselnetzbetrieb, durch die Realisierung von Microgrids.

Sofern Workloads im Verantwortungsbereich der RZ-Betreiber sind, können durch zeitliche Lastverschiebung z. B. rechenintensive Prozesse, wie Batch-Verarbeitungen, Backups oder KI-Trainings, gezielt in Zeiten hoher Stromverfügbarkeit verschoben werden. So können überschüssige Energiekapazitäten aus erneuerbaren Energiequellen sinnvoll genutzt und damit Einspeisespitzen im Netz kompensiert werden. Umgekehrt kann durch die Netzplanung in Rechenzentren eine dynamische Energieverteilung effizienter umgesetzt werden. Dies betrifft Lastspitzenzeiten sowie flexible Prozesse zum Abregeln oder Pausieren. Darüber hinaus können installierte Batteriekapazitäten und Generatoren netzdienlich eingesetzt werden.

Eine Listung der Parameter, wie z.B. Leistungs- und Energiebedarf, findet sich in Tabelle 1 für die unterschiedlichen Leistungsklassen von Rechencluster-Racks.

Kategorie	5–10 MW AI-Cluster	20–50 MW AI-Cluster	>100 MW Hyperscale AI
typische Leistungsdichte pro Rack ^{4,5,6, 7}	40–80 kW pro Rack	80–120 kW pro Rack	120–150+ kW pro Rack
Einspeisung (MS*) ⁸	1–2 Einspeisungen, 10–20 kV	2 unabhängige Einspeisungen, 20–30 kV	2–3 diversifizierte Einspeisungen, 110 kV-Anbindung (üblich)
Redundanz	N+1 oder 2N für kritische Pfade	2N für IT-Last, N+1 für Kühlung	vollständige 2N-Architektur bis zum Rack
USV-System ^{9,10,11}	modular, 1–3 MW Blöcke, Li-Ion	3–6 MW Blöcke, Doppelwandler, Flywheels	10+ MW USV-Cluster, hybride Systeme (BESS** + Flywheel)

Kategorie	5–10 MW AI-Cluster	20–50 MW AI-Cluster	>100 MW Hyperscale AI
On-Site-Power-Lösungen ^{12,13,14}	Diesel, 5–10 MW	Geplante Gas-Turbinen, Gas-Motoren + BESS**	Geplante Gas-Turbinen, Gas-Motoren Wasserstoff-Ready, 20–100 MW
Netzinteraktion	geringe Netzdienlichkeit	Demand-Response möglich	aktive Netzteilnehmer, Peak-Shaving, Inselnetzbetrieb
Typische Nutzung	Forschung, kleinere KI-Teams, Edge KI-Inferenz	Unternehmens-KI, KI-Inferenz, LLM-Training***	KI-Supercomputer, großskalige KI-Inferenz, Foundation-Model-Training

* MS – Mittelspannung; ** BESS – Battery Energy Storage System; *** LLM – Large-Language-Model

Tabelle 1: Parameterübersicht in Abhängigkeit von Leistungsklassen

Die Standortfrage unter Berücksichtigung des Netzanschlusses

Während Metropolregionen, wie das Rhein-Main-Gebiet und Berlin-Brandenburg, zunehmend durch begrenzte Bauflächen und Netzengpässe limitiert sind, geraten mittelgroße Städte in den Fokus von Auftraggebern von Rechenzentren und deren Betreibern. Als Vorteil erweist sich hier oft die Realisierung einer hybriden Versorgungsstrategie aus erneuerbaren Energiequellen einerseits und einem klassischen, gesicherten Netzanschluss andererseits. Dieser garantiert den Betreibern auch bei Dunkelflaute die erforderliche Versorgungssicherheit. Ein weiterer Vorteil dieser Standorte ist, dass die seit 2026 vom Energieeffizienzgesetz (EnEg) geforderte Abwärmenutzung für neue Rechenzentren leichter mit benachbarten Gewerbeparks oder über kommunale Nahwärmenetze realisiert werden kann.

In Frankfurt, wo sich mit DE-CIX einer der weltweit größten Internetknoten befindet, gibt es bereits Probleme mit der Versorgungskapazität. Laut dem regionalen Energieversorger Mainova können v. a. größere, leistungsstarke Neuanschlüsse erst ab Mitte der 2030er Jahre bereitgestellt werden. Davon betroffen seien vor allem neu geplante Projekte. Somit können bereits bestehende Rechenzentren weiter betrieben und im Rahmen vorhandener Kapazitäten erweitert werden.

Neue Rechenzentren müssen ins Umland des DE-CIX-Knotens, d. h. in angrenzende Landkreise ausweichen. Bis 2030 sind hier bereits mehr als 20 Standorte in Planung. Bis zum Jahr 2037 und darüber hinaus ist nicht absehbar wie der große Bedarf an elektrischer Energie gedeckt werden kann, so mussten z.B. die Syna GmbH und Amprion GmbH Netzanschlussverträge nachträglich zurücknehmen, da die zugesicherte Leistung erst deutlich später bereitgestellt werden kann. Für den Netzanschluss schickt Syna die Rechenzentren in eine Warteschleife¹⁵.

Noch passen Netzausbau, Bedarf und der Betrieb nicht zusammen. Doch der gewollte Ausbau der Infrastruktur für die massive Nutzung von generativer KI und die damit verbundenen positiven Perspektiven für die Wirtschaft in Deutschland und Europa treiben die Abstimmung voran. So könnte das Reifegradverfahren, welches in Deutschland für BESS-Anfragen in 2026 eingeführt wurde auch ein mögliches Vorgehen zur Beschleunigung des Netzanschlusses von Rechenzentren in Ballungsräumen sein.

Resümee und Perspektiven

Deutschland hat die Möglichkeit, mit seinem hochverfügbaren, resilienten und interkontinental verknüpften Stromnetz, den wirtschaftlichen Standort mit der intelligenten Systemintegration von effizienten Rechenzentren zu stärken. Die Chance, gerade jetzt stark am wirtschaftlichen Erfolg der generativen KI und deren massiver Nutzung zu partizipieren, ist somit gegeben. Der spürbare Ausbau der Rechenzentrumsinfrastruktur eröffnet Deutschland die Chance, bei souveränen Daten-, Cloud- und KI-Kapazitäten zu den führenden Weltregionen aufzuschließen und eigene Stärken auszubauen.

Da der Bedarf an Rechenkapazitäten deutlich schneller wächst als der Stromnetzausbau Schritt halten kann, müssen intelligente Systemintegration und Effizienzinnovation nicht nur als Ziel, sondern als technische und regulatorische Gestaltungsaufgabe verstanden werden. Ein konkreter Hebel liegt in der gezielten Nutzung von Synergien zwischen On-Site-Power-Lösungen an Großverbrauchern und zentralen Kraftwerkskapazitäten, die für Versorgungssicherheit bei Dunkelflauten und Systemstabilität unerlässlich sind – eine Kombination, die sowohl Netzlast reduziert als auch Resilienz erhöht.

Das ist ein Ansatz, der sich durchaus von den in den USA oder Asien verfolgten „Hyperscaling-Strategien“ unterscheidet und langfristig sowohl ökologisch als auch wirtschaftlich überlegen sein könnte. Während internationale Tech-Konzerne auf überdimensionierte, oft redundante Infrastruktur setzen, um die schlecht prognostizierbaren Lastspitzen generativer KI abzufedern¹⁶, muss Deutschland aufgrund seines hochgradig vernetzten, aber kapazitätsoptimierten 380-/220-kV-Übertragungsnetzes einen präziseren, planbareren Ansatz verfolgen, der auf dynamischer Laststeuerung, dezentraler Erzeugung und intelligenter Netzintegration basiert.

Im Gegensatz erscheint im US-amerikanischen zentrierten „Bulk-Power-System“ der dezentrale Betrieb einfacher, wobei im europäischen Netz eine Synchronisation vorliegt und daher weniger Puffer für spontane Laständerungen vorhanden sind. Gleichzeitig zwingen die strengen deutschen Regularien (EEG, Klimaschutzgesetz, Netzausbaubeschleunigungsgesetz) dazu, den Energiebedarf durch innovative Kühlkonzepte (z. B. Flüssigkühlung mit Abwärmenutzung für Fernwärme oder Industrieprozesse), fortschrittliche KI-gestützte Betriebsstrategien, hoch effiziente PUE-Werte und die direkte Kopplung mit lokalen erneuerbaren Energiequellen zu minimieren, statt wie in den USA einfach mehr Strom, aus fossilen Quellen bereitzustellen. Während in den USA der Bau neuer Gaskraftwerke als schnelle Lösung für den Energiehunger von Hyperscalern akzeptiert wird, setzt Deutschland auf einen Energiemix aus Wind-, Solar- und Biomasseenergie, ergänzt durch hocheffiziente Gasturbinen mit Kraft-Wärme-Kopplung, die nicht nur Strom, sondern auch nutzbare Wärme liefern und so den Gesamtwirkungsgrad steigern.

Schließlich sorgen die hohen Strompreise und die Umlagenstruktur in Deutschland dafür, dass Betreiber Anreize haben, energieeffiziente Hardware, auch mittels Photonik anstatt klassischer CPUs/GPUs und softwareseitigen Optimierungen einzusetzen.

Letztlich zeigt sich, dass der fortschrittliche, zügige und netzorientierte Ausbau der HPC RZ-Infrastruktur in Deutschland ein sinnvoller ist. Wichtig dabei ist, dass nicht nur die technische Optimierung der Systeme vorangetrieben wird – dies wird ingenieurtechnisch bereits umgesetzt und ist bedingt durch die großen finanziellen Investitionen. Ebenso entscheidend ist, dass entsprechende Regularien mit diesen dynamischen, technischen Entwicklungen schritthalten – ein Appell, der sich an politische Entscheidungstragende richtet. Damit wird man nicht nur den Herausforderungen der massiven Nutzung von generativer KI gerecht, sondern man sichert mit diesem Vorgehen eine nachhaltige Zukunft in Wirtschaft, Gesellschaft und Digitalisierung.

-
- ¹ Smith, M.S., IEEE Spectrum, April 2026. What will it take to build the world's largest data center?
- ² Elliott, R.F., Stevens, H., New York Times (April 11th/12th 2026). Some tech giants are going off grid, but at higher costs.
- ³ Global energy demands within the AI regulatory landscape. <https://www.brookings.edu/articles/global-energy-demands-within-the-ai-regulatory-landscape/> (letzter Zugriff: 22.04.2026)
- ⁴ NVIDIA Corporation, "NVIDIA H100 Tensor Core GPU Architecture," NVIDIA, 2023.
- ⁵ NVIDIA Corporation, "NVIDIA Blackwell Platform Architecture," NVIDIA, 2024.
- ⁶ ASHRAE TC 9.9, "Liquid Cooling Guidelines for Datacom Equipment Centers," ASHRAE, 2021.
- ⁷ Smith, M.S., IEEE Spectrum, April 2026. What will it take to build the world's largest data center?
- ⁸ Schneider Electric, "Electrical Infrastructure for Large Data Centers" Schneider Electric Whitepaper, 2022.
- ⁹ Schneider Electric, "Selecting a UPS System for High-Density Data Centers," Schneider Electric Whitepaper, 2023.
- ¹⁰ Vertiv, "Powering AI Workloads: UPS Requirements for GPU-Dense Environments," Vertiv Engineering Brief, 2023.
- ¹¹ Google Data Center Engineering, "High-Power UPS and Battery Systems for AI Training Clusters," Google, 2023.
- ¹² Uptime Institute, "Data Center Resiliency and Backup Power Trends," Uptime Institute, 2023.
- ¹³ Schneider Electric, "Microgrid and Onsite Generation for Large Data Centers," Schneider Electric Whitepaper, 2022.
- ¹⁴ Wärtsilä Corporation, "Gas Engines and Hydrogen-Ready Solutions for Data Centers," Wärtsilä Whitepaper, 2024.
- ¹⁵ Netzanschluss: Syna schickt Rechenzentren in die Wartschleife. <https://www.energate-messenger.de/news/257720/netzanschluss-syna-schickt-rechenzentren-in-die-wartschleife> (letzter Zugriff: 22.04.2026)
- ¹⁶ Elliott, R.F., Stevens, H., New York Times (March 18th 2026). Why Tech Giants Are Ditching the Power Grid

Autoren:

Dr.-Ing. Damian Dudek, VDE ITG

Prof. Dr.-Ing. Jutta Hanson, TU Darmstadt

Dr.-Ing. Niklas Panten, etalytics GmbH

Prof. Dr.-Ing. Christian Rehtanz, TU Dortmund

Christoph Strunck, Goodman Germany GmbH

Dr. phil. nat. Matthias Wirth, VDE ITG

Kontakt:

VDE Verband der Elektrotechnik
Elektronik Informationstechnik e.V.
Merianstraße 28
63069 Offenbach am Main
Tel. +49 69 6308-360
damian.dudek@vde.com