NVIDIA DGX SuperPOD Data Center Design

Reference Guide

Featuring NVIDIA DGX H100 Systems

DG-11301-001 | April 2023

Document History

DG-11301-001

Version	Date	Authors	Description of Change
01	2023-03-29	Steven Hambruch (DCA), Dennis O'Brien, Srikanth Cherukuri, and Robert Sohigian	Initial release
02	2023-04-06	Steven Hambruch (DCA) and Robert Sohigian	Updates to Table 4

Abstract

The NVIDIA DGX SuperPOD™ with NVIDIA DGX™ H100 system provides the computational power necessary to train today's state-of-the-art deep learning (DL) models and to fuel innovation well into the future. The DGX SuperPOD delivers groundbreaking performance, deploys in weeks as a fully integrated system, and is designed to solve the world's most challenging computational problems.

This document provides guidelines for selecting or configuring the right data center to deploy a DGX SuperPOD, and is the result of co-design between DL scientists, application performance engineers, system architects, and data center architects to build a system capable of supporting the widest range of DL and High Performance Computing (HPC) workloads.

This guide provides an overview of scalable options for DGX H100 installations and covers near-term and long-term data center deployment considerations.

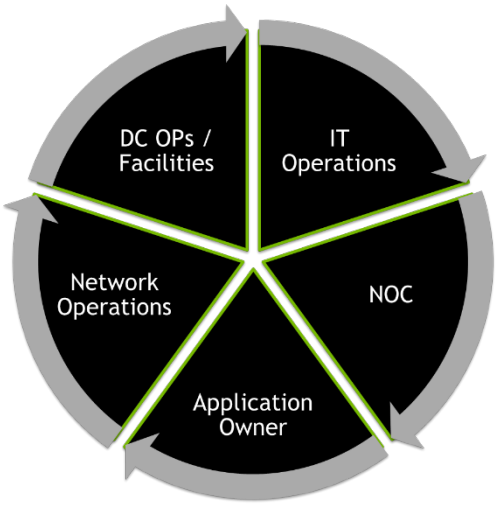
The information contained in this guide is intended for IT and data center professionals who are generally familiar with the network, power, space, and cooling technologies that are inherent to data center deployments.

Contents

Chapter 1. Planning a Data Center Deployment

1.1 Coordination

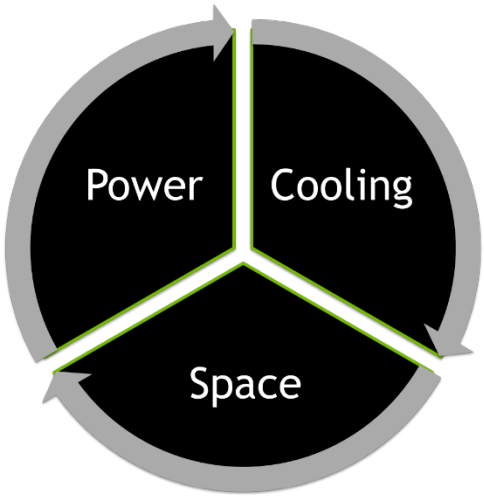
The planning of a DGX SuperPOD deployment requires coordination and alignment of multiple constituencies within an organization and may impact third-party vendors that provide critical services to those teams. Teams that can be impacted include the application owner, end-users, data center operations, facilities teams, security, various information technology and networking teams, and network operations center support teams. The various teams must have alignment to ensure a well planned and executed installation of a DGX SuperPOD.



During the planning and implementation, changes that occur in one technical domain can impact on another domain. For example, if there is a power constraint that limits rack density and causes the implementation to be distributed across a higher number of rack footprints, not only will that impact the data center facility's floor layout plan, but it will likely also impact the network—necessitating a recalculation of cable lengths and possibly introducing latency-related performance impacts. Good alignment, coordination, and communication among the various domain experts at every phase of the design and implementation of a DGX SuperPOD deployment will create the best result.

1.2 The Economy of Data Center Resources

In every data center environment, there are finite resources. Just as in an economy, the constraints or limitations in the supply of these resources creates scarcity in the face of demand. This drives the need to optimize resource utilization to achieve overall efficiency and maximum return on the investments made in those data center facilities.



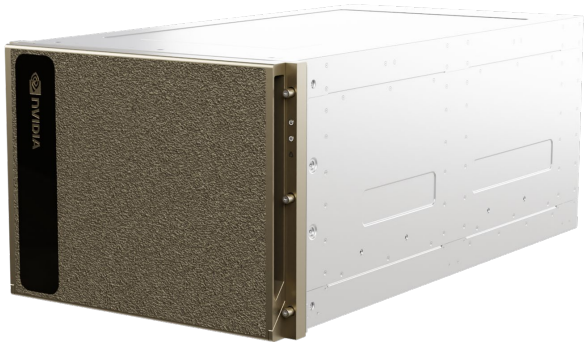
The three main resource constraints in an air-cooled data center environment are power, cooling, and space.

The interrelationship between these resources is such that scarcity of one resource could create increased demand for another resource. For example, limitations in cooling capacity could constrain rack power density, resulting in a need to occupy a larger number of racks to house a given number of servers. In this example, the resource constraint in cooling is “paid for” using floor space. This enables the deployment to take place, although it is not an optimized use of floor space in the data center.

1.3 DGX H100 Key Specifications

The NVIDIA DGX SuperPOD with NVIDIA DGX H100 systems is an optimized system for multi-node DL and HPC. It consists of between 31 and 127 DGX H100 systems (Figure 1), with a total of 1,016 NVIDIA Hopper GPUs. It is built using the *NVIDIA DGX SuperPOD Reference Architecture* and is configured to be a scalable and balanced system providing maximum performance.

Figure 1. DGX H100 system



Key specifications of the DGX H100 system are in Table 1.

Table 1. DGX H100 system key specifications

Specification	Value
System power consumption	11.3 kW max
System weight	287.6 lb (130.45 kg)
System dimensions	14 × 19.0 × 35.3 in (356 × 482.3 × 897.1 mm)
Rack units	8
Cooling	Air
Operating temperature	5–30 °C (41–86 ˚F)
Operating altitude	3,048m (10,000 ft) maximum 5–30 °C: altitude 0–1000 ft 5–25 °C: altitude 1000–5000 ft 5–20 °C: altitude 5000–10000 ft
Acoustic noise operating	Acoustic Power (LWA,m) 25 ˚C /77 ˚F = 97 dB – 30 ˚C / 86 ˚F = 98.7 dB
PSU redundancy	N+2 > Loss of one or two PSUs: • System continues to operate at full performance > Loss of three or more PSUs • System not operational

Caution: Due to the weight and size, high power requirements, high heat rejection, and the sound power level of this equipment, operator safety is of major importance. Personal Protective Equipment (PPE) must be worn, and safety procedures must always be observed when working on or near DGX H100 systems.

1.4 Density of Compute Racks

Every data center has its own unique constraints regarding power, cooling, and space resources. NVIDIA has developed DGX SuperPOD configurations to address the most common deployment patterns. However, should customization be necessary to address specific data center parameters in a particular deployment, NVIDIA can typically accommodate. It is important to work with NVIDIA and communicate any data center design constraints, so that a performance optimized deployment can be achieved. Altering the deployment pattern without such consultation can lead to serious performance, operational, support, or scalability challenges.

The building block of a DGX SuperPOD configuration is a scalable unit (SU). Each scalable unit consists of up to 32 DGX H100 systems plus associated InfiniBand leaf connectivity infrastructure. A DGX SuperPOD can contain up to 4 SU that are interconnected using a rail optimized InfiniBand leaf and spine fabric. A pair of NVIDIA Unified Fabric Manager (UFM) appliances displaces one DGX H100 system in the DGX SuperPOD deployment pattern, resulting in a maximum of 127 DGX H100 systems per full DGX SuperPOD.

DGX H100 systems are optimally deployed at a rack density of four systems per rack. However, rack densities can be customized to fit within the available power and cooling capacities at the data center. Naturally, reducing rack density increases the total number of required racks.

Some specifications are shown in Table 2.

Table 2. One, two, and four DGX H100 systems per rack

Number of DGX Systems per Rack	Number of DGX System Racks	Total SU Server Rack Power Requirement	Total Power Per Rack Footprint
1	32	361.6 kW	11.3 kW
2	16	361.6 kW	22.6 kW
4	8	361.6 kW	45.2 kW

In addition to these racks, every SU also requires two racks for the InfiniBand leaf and spine infrastructure, management servers, and storage infrastructure. Use the values provided in Table 4 to assist in calculating the power requirements for these racks.

The design of the InfiniBand fabric within the DGX SuperPOD architecture presents constraints regarding cable path distance (the total distance of travel must not exceed 50m for any InfiniBand cable). Therefore, the deployment patterns are modeled with careful attention to cable length.

1.5 Safe System Delivery

The following considerations are essential for the safe delivery of equipment to the data center location:

- The facility should have a loading dock or permit a liftgate equipped delivery vehicle.
- If there is no loading dock, a hydraulic lift or ramp must be provided to safely offload pallets.
- There must be a secure receiving room or staging area separate from the data hall to store equipment before installation.
- There must be clear access between the loading dock and the receiving room.
- There must be adequate space in the receiving room to remove equipment components from pallets before transferring to the data hall. All shrink-wrap, cardboard, and packing material should remain in the receiving room.
- Conveyances must be available to safely move equipment from the receiving room to the data hall.

NVIDIA components will be put on pallets for shipping by common carriers. Pallet information for the DGX H100 system is provided in Table 3.

Table 3. DGX H100 pallet information

Specification	Value
Units/Pallet	1
Actual Product Weight	287.6 lb (130.45 kg)
Chargeable Product Weight	376 lb (170.45 kg)
Pallet Weight	421 lb. (191 kg)
Product Box Dimensions	38.2 × 28 × 46.5 in (970 × 711 × 1,178 mm)
Overpack Material (Crate/Corrugated Box)	Corrugated box

Depending upon the size of the DGX SuperPOD configuration, up to 127 DGX system pallets along with numerous network switches, management server appliances, and cables will be shipped. Procurement teams and suppliers should coordinate with onsite data center personnel to ensure that the material can be received and stored in a secure location before installation.

1.6 Power and Heat Dissipation

Management racks contain network infrastructure, storage, and management servers for the DGX SuperPOD, in varying quantities based on the number of SUs being deployed. Each system component has an expected average power (EAP) and an expected peak power (EPP).

EAP, EPP, and heat dissipation values for key components of a full DGX SuperPOD are shown in Table 4.

Table 4. Component power and heat dissipation of a 127-node DGX SuperPOD

		Servers				Switches			
		Compute	Storage	Mgmt	Fabric	Compute	Storage	In-band Mgmt	OOB Mgmt
Model		DGX H100	Varies[1]	PowerEdge R750[1]	NVIDIA UFM 3.1	QM9700	QM9700	SN4600C	SN2201
Qty		127	Varies[1]	5	4	48	16	8	8
EAP (Watts)	Each	11,300[2]	2880	704	600	1,376	1,376	466	98
	Subtotal	1,435,100	17,280	3,520	2,400	66,048	22,016	3,728	784
EPP (Watts)	Each	11,300	3,600	880	750	1,720	1,720	820	135
	Subtotal	1,435,100	21,600	4,400	3,000	82,560	27,520	6,560	1080
Peak Heat Load (BTU/h)	Each	38,557	12,284	3,003	2,559	5,869	5,869	2,798	461
	Subtotal	4,896,764	73,702	15,013	10,236	281,706	93,902	22,384	3,685
Percent of System Total		90.72%	1.37%	0.28%	0.19%	5.22%	1.74%	0.41%	0.07%

1. See *NVIDIA DGX SuperPOD Reference Architecture*
2. DGX H100 systems operate at or near peak utilization continuously when running AI workloads

In general, the design requires a minimum airflow of 157 ft^3/min (4.445 m^3/min) per kilowatt. However, the actual requirements can vary based on the environmental conditions and altitude of each specific data center, as well as the Delta T of each component.

1.7 Environmental Thermal Guidelines

Table 5 illustrates the general ASHRAE temperature and humidity standards for the cooling of IT and telecommunications equipment. To meet the cooling demands of DGX SuperPOD, data center facilities should satisfy the **Recommended** requirements, and at the very least must satisfy the **Class A1** requirements up to the limits noted.

Table 5. ASHRAE specifications

Range	Class	Dry-Bulb Temperature	Humidity Range, Non-Condensing	Maximum Dew Point
Recommended	All A	64.4–80.6 °F 18–27 °C	41.9 °F to 60% RH and 59 °F DP 5.5 °C to 60% RH and 15 °C DP	59 °F 15 °C
Allowable up to 30 °C for DGX H100 Systems	A1	59–89.6 °F 15–32 °C	20–80% RH	62.6 °F 17 °C
Allowable per ASHRAE for various other classes of data center and telecom environments	A2	50–95 °F 10–35 °C	20–80% RH	69.8 °F 21 °C
	A3	41–104 °F 5–40 °C	10.4 °F DP and 8–85% RH -12 °C DP and 8–85% RH	75.2 °F 24 °C
	A4	41–113 °F 5–45 °C	10.4 °F DP and 8–90% RH -12 °C DP and 8–90% RH	75.2 °F 24 °C
	B	41–95 °F 5–35 °C	8–80% RH	82.4 °F 28 °C
	C	41–104 °F 5–40 °C	8–80% RH	82.4 °F 28 °C

Source air contamination (such as smoke, dust, pollution, or other types of contamination) must be mitigated through filtration.

Table 6 describes the ISO 14644-1 maximum particle sizes for different classes of air cleanliness. In an Airside Economizer Mode, or any other environment where source air may be contaminated, the air must be filtered by a minimum MERV 13 (or EN779-2012 M6/F7, or ISO 16890 ePM1-50%) rated filter, and the rack area must meet the cleanliness level of ISO 14644-1 Class-8 standard with maximum particle counts not to exceed 3,520,000 @ 0.5 µm/m^3 for no longer than 15 minutes.

Table 6. ISO 14644-1 standard for air cleanliness classifications

Class	Particle Size[1]					
	> 0.1 µm	> 0.2 µm	> 0.3 µm	> 0.5 µm	> 1 µm	> 5 µm
1	10	2				
2	100	24	10	4		
3	1,000	237	102	35	8	
4	10,0000	2,370	1,020	352	83	
5	100,000	23,700	10,200	3,520	832	29
6	1,000,000	237,000	102,000	35,200	8,320	293
7				352,000	83,200	2,930
8				3,520,000	832,000	29,300
9				35,200,000	8,320,000	293,000

1. Uncertainties related to the measurement process require that data with no more than three significant figures be used in determining the classification level.

Table 7 provides a general comparison of filtration standards as a cross-reference.

Table 7. Comparison of filter types by standard

ASHRAE Standard 52.2-2007 Minimum Efficiency Reporting Value (MERV)[1]	EN 779-2012	ISO 16890
MERV 1, 2, 3, 4	G1, G2	--
MERV 5	G3	--
MERV 6, 7, 8	G4	Coarse 90%
MERV 8, 9, 10	M5	ePM10–60%
MERV 9, 10, 11, 12, 13	M6	ePM2.5–50%
MERV 13, 14	F7	ePM1–50%
MERV 14, 15	F8	ePM1–75%
MERV 16	F9, E10, E11, E12, H13, H14, U15, U16	--

1. The testing and evaluation procedures for the ASHRAE 52.2, EN 779-2012, and ISO 16890 are significantly different, making direct comparisons of filter efficacy and efficiency difficult and potentially misleading. The values in this table are provided as a general reference but should not be considered scientifically precise.

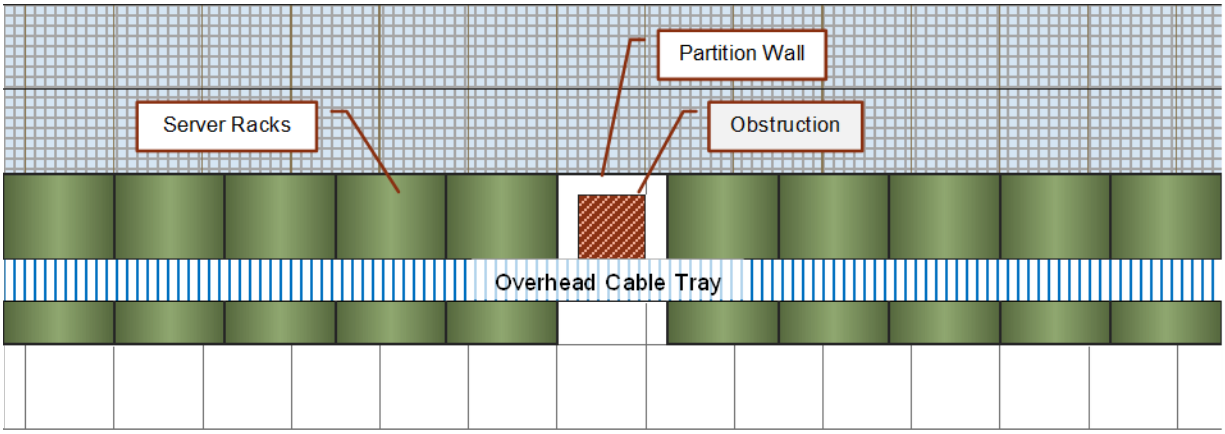
Chapter 2. Interrupted Rack Layouts

This section details some options for addressing rack layout challenges caused by data center architectural features. In all cases, overhead ladders or trays must span the racks on either side of the obstruction to facilitate cabling across the gap.

2.1 Overcoming Obstructions to Contiguous Racks

Figure 2 shows a top view of how a site-specific obstruction, in this case a structural column within a row, can be accommodated.

Figure 2. Obstructions to contiguous DGX SuperPOD deployments

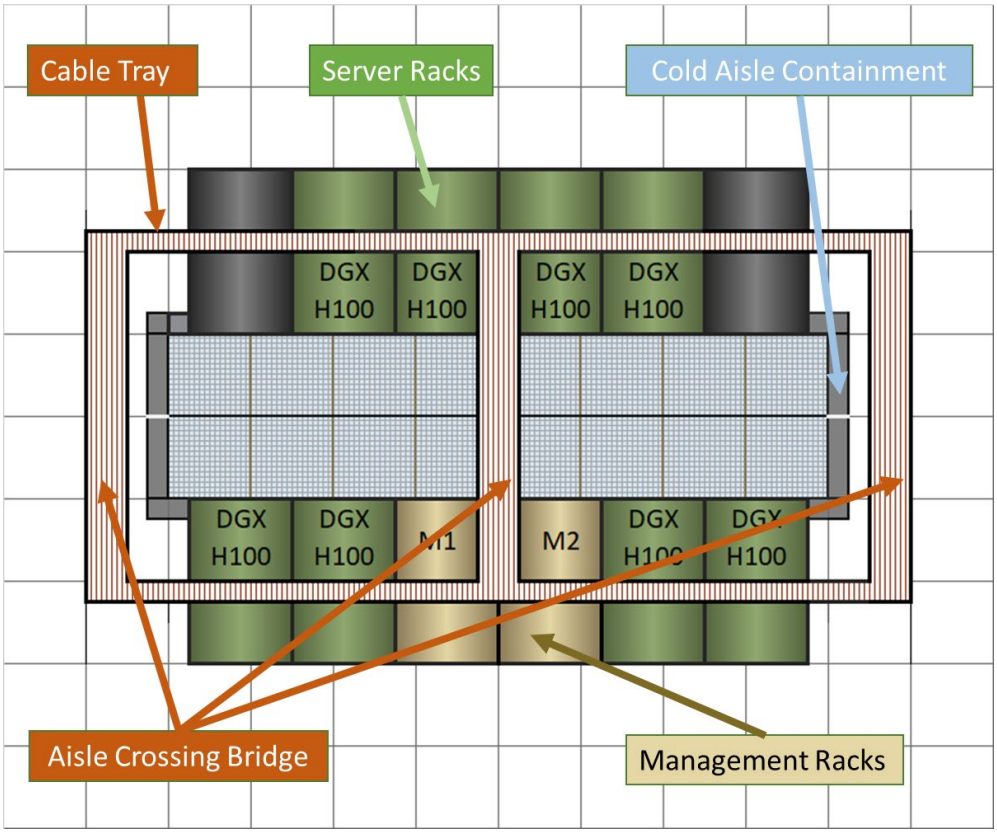


Note: In this example, the tray has been located closer to the rear of the rack to clear the obstruction. If the obstruction was toward the rear of the rack, it might have been necessary to move the cable tray toward the front of the racks. This is not ideal but preferred over bending the tray around the obstacle. An aisle containment partition takes the place of the missing rack, to ensure that hot aisle/cold aisle integrity is maintained.

2.2 Spanning Deployments Across Two Rows of Racks

Figure 3 shows how the standard layout can be modified to place racks in two rows. When spanning more than one row, cable length is an important factor. The total distance of travel must not exceed 50m for any InfiniBand cable that connects a DGX H100 system to a leaf switch, and for optimum system performance, cable lengths should be limited to 30m for AOC cables.

Figure 3. Spanning across two rows of racks (top view of racks)



The maximum distance for switch to switch interconnectivity is 500m. These distances cannot be extended using intermediary repeaters or switches. For overhead or underfloor cable trays that provide more than one bridge to adjacent rows, cables should be routed to the bridge that provides the shortest path to the destination rack.

Chapter 3. Electrical Specifications

3.1 Data Center Power Configuration

The DGX SuperPOD is typically deployed with a rack density of four DGX H100 systems per rack, although deployments with lower rack densities are possible. Combining international norms on voltages and circuit protection yields common power provisioning patterns for data centers. A DGX H100 power supply system using components certified for 200–240 VAC can be deployed world-wide. Connectors, distribution boxes, fuses, circuit breakers, and wire gauges selected at compatible steps ease certification and installation. Rack power distribution units (rPDUs) typically derive 200–240 VAC single phase power by dividing a three-phase input power circuit into three individual single-phase circuits.

Table 8 identifies the most common supply/distribution voltages and currents that can support the defined SU deployment patterns.

Table 8. Common distribution schemes compatible with DGX H100 racks

Phase	Distribution Voltage	Line Voltage	Amps	Breaker Derating	Circuit Capacity kW[1]	Maximum Supported DGX H100 Systems per Rack[2]	Peak Server Demand per Circuit kW[2]	Stranded Capacity at Peak Demand kW[2]
1Φ	230	230	63	100%	13.7	2	11.3	2.4
3Φ Delta	208	208	60	80%	32.8	4	22.6	10.2
3Φ Wye	400	230	32	100%	21	2	11.3	9.7
3Φ Wye	415	240	32	100%	21.8	2	11.3	10.5
3Φ Wye	415	240	60	80%	32.7	4	22.6	10.1

1. 0.95 power factor.
2. Based on a three circuit N+1 power provisioning scheme where no circuit carries more than 50% of the load.

The preferred power for high-density deployment patterns is 415 VAC, 60A, three-phase, N+1. The design can be modified to support other supply voltage schemes, depending on the number of servers per rack. Power supplied to each rPDU must originate from separate data center floor-mounted or busway PDUs. All power feeds must be supported by facility-level UPS and generator back-up power to mitigate the risk of power loss.

3.2 Power Redundancy

Generally, the data center should meet or exceed Uptime Institute Tier 3 design standards, or alternatively the TIA942-B Rated 3 or EN50600 Availability Class 3 design standards, including concurrent maintainability and no single point of failure.

In addition to those foundational standards, the DGX H100 system has additional requirements regarding power redundancy and resiliency. The system includes six internal power supply units. Four power supply units must be energized for the server to operate.

ⓘ	**Caution:** Four of the six power supplies must be energized for the system to operate. This is a critical data center design consideration.

A failure of a single system in a multi-node AI workload will cause the entire job to stop on all the systems. In environments where system availability is paramount, and work would not be recoverable (for example, from a checkpoint), a minimum of three power sources (rPDUs fed by discrete upstream power distribution paths), must be provisioned to each rack. Each of those sources will connect to two of the six system power supplies on each system, guaranteeing that a failure or maintenance event on any one of those sources will leave a minimum of four system power supplies energized.

Due to this requirement, the data center must minimally provide $N+1$ power, where N equals two power sources. Each power source must be sized to support 50% of the total peak load. This requirement applies to DGX H100 systems racks only. Management racks may be powered with traditional $2N$ redundancy using two power feeds.

The following illustrations and tables describe three power provisioning design concepts, each with their own advantages and disadvantages. Other power provisioning solutions are also possible, depending on the unique power system architecture of a given data center site. Consult with NVIDIA to determine if an alternate solution will meet the DGX SuperPOD availability requirements.

3.2.1 Traditional Redundant Power

For a data center supplied with two utility feeds or two UPSs supplying power to the racks, Figure 4 and Table 9 describe traditional power provisioning of each DGX system.

Figure 4. Traditional redundant power provisioning pattern

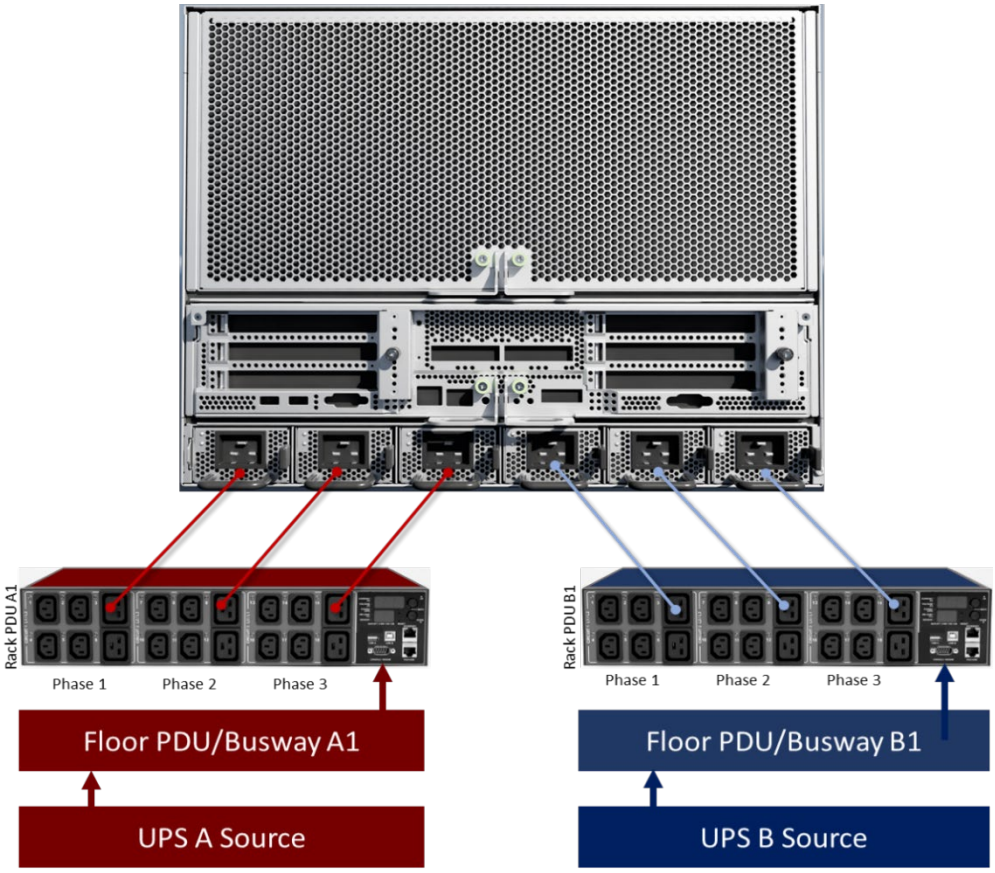


Table 9. Traditional redundant power provisioning advantages/disadvantages

Advantages	Provides basic 2N power redundancy for typical I.T and network devices. Sufficient for management racks, but not for DGX H100 system racks. Compatible with nearly all data centers.
Disadvantages	During a failure of one power source, the number of energized PSUs on a DGX H100 system would fall below four, resulting in a shutdown of that system. Any active AI workloads running on that system at the time of failure will cease, resulting in a disruption to that job on all systems.
Grade	Not acceptable for DGX H100 systems

3.2.2 N+1 Configuration

Figure 5 and Table 10 illustrate a typical configuration for a data center with two UPSs supplying three power paths to the racks. Wherever possible, the load should be distributed so that one rack features two feeds from UPS B, and the next rack features two feeds from UPS A, in a repeating pattern. This minimizes dependency on any given UPS source and balances the load across them.

Figure 5. N+1 power provisioning pattern

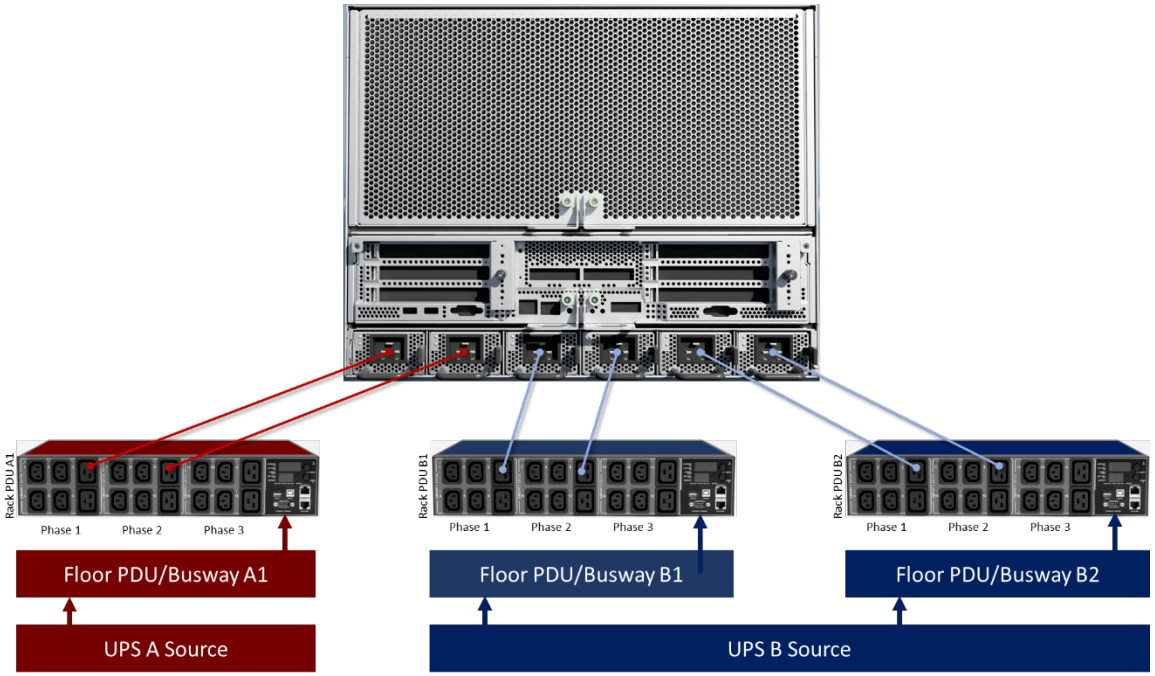


Table 10. N+1 power provisioning pattern—advantages/disadvantages

Advantages	Provides basic power redundancy, and the ability to support AI workloads during a local power loss or maintenance event caused by a system PSU, a single rPDU, or a Floor PDU/RPP breaker. Compatible with most data centers.
Disadvantages	Adds complexity and cost. Two of the three rPDUs (powering a total of four system PSUs) are supplied by the same upstream UPS power source. Therefore, a failure or maintenance event affecting that upstream UPS would cause the system to power off.
Grade	Acceptable. Fault tolerant for most common failure modes, but some risks remain unmitigated

3.2.3 Enhanced N+1 Configuration

Figure 6 and Table 11 illustrate a power provisioning scheme using three discrete UPS systems, providing three discrete power distribution paths.

Figure 6. Enhanced N+1 power provisioning pattern

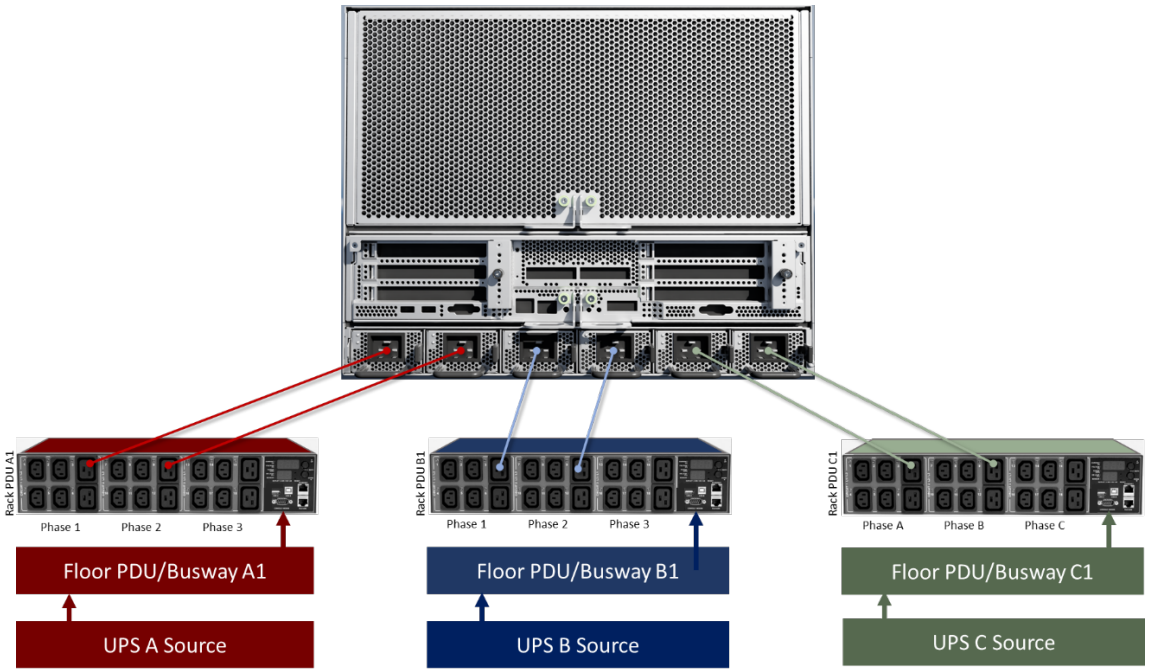


Table 11. Enhanced N+1 power provisioning pattern

Advantages	Provides both fault tolerance and the ability to support AI workloads during a power loss or maintenance event that affects a single upstream power distribution path. Each rPDU is powered by a discrete upstream UPS/Power Distribution Path.
Disadvantages	Many data centers are not designed to provide power from three discrete upstream UPS/power distribution paths.
Grade	Acceptable. Optimal provisioning pattern for maximum performance and reliability.

3.3 Planning and Deploying Power Connections

Follow these best practice guidelines when connecting AC power to the racks and systems:

- Validate AC power redundancy at each server rack. An outage could occur if these requirements are not met.
- Complete power provisioning within the data center before connecting power to the rPDUs and system deployment.
- Have an electrician or qualified facilities representative verify that the kVA supplied is within specification at each of the floor-mounted PDUs and individual circuits that feed the racks.
- Label all the power connections to indicate the source of power (PDU #) and the specific circuit breaker numbers used within each PDU.
- Color code the power cables (and associated rPDUs) to help ensure that redundancy is maintained.
- Clearly label the equipment served by each circuit breaker within the PDU.
- Earth/bond the data center racks to the telecommunications ground that in turn will be connected to the facility ground system.
- Have an electrician or qualified facilities representative verify that there are three or more power connections fed from separate redundant PDUs before turning on the system.
- Have an electrician perform an AC verification test by turning off the individual circuit breakers feeding each rack power strip to verify that power redundancy has been achieved in each rack.

3.4 Rack Power Distribution Unit (rPDU) Selection

This section describes different options for providing redundant power. Each of the three required power input paths must support one half of the expected peak power of the rack. Keep in mind however, that for the typical N+1 provisioning pattern, two of the three power paths will eventually converge at some upstream junction (Such as a Room PDU or UPS), so that junction point must still be sized to manage each downstream rack's full peak load.

rPDU features should include remote power monitoring Rest API capability for automation, and rack temperature/humidity monitoring.

For optimal integration into the DGX SuperPOD architecture, NVIDIA recommends using Raritan, Vertiv/Geist, or ServerTech rPDUs whenever possible.

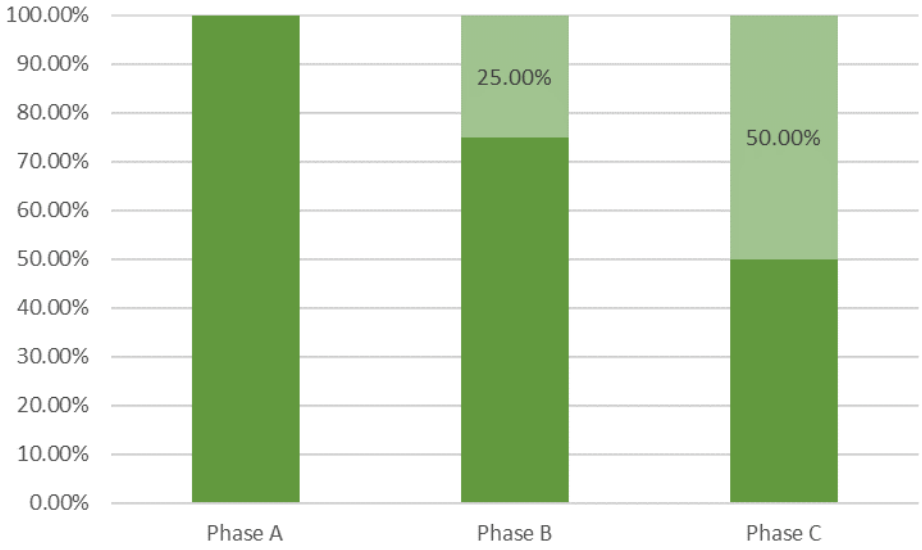
Due to cable management and cabinet depth limitations, in addition to the potential quantity of rPDUs to be deployed, horizontal rPDUs may be required. Vertical (0U) rPDUs are an option only in racks of sufficient width and depth where they would not block access to any portion of the back of the system chassis, and they shall be mounted at the rearmost mounting points at the back of the rack. A maximum of two vertical PDUs are possible, therefore the remaining rPDU must be horizontal. NVIDIA can be consulted to provide rPDU recommendations based on the target data center's power provisioning specifications. NVIDIA can also provide recommendations for the power outlet mapping that is specific to the selected rPDUs, to ensure proper electrical phase/load balancing across multiple PDUs.

3.5 Phase Balancing

The power draw across the phases of a three-phase circuit should be as balanced as possible. In simple terms, a three-phase circuit is said to be unbalanced when the load on one of its phases is drawing more current than the average drawn by all three. This has several negative implications, including possible thermal derating of the conductors, deviation of proper electrical phase angle, possible damage to upstream transformers, unanticipated breaker trips during failover events, and most noticeably, stranding of the power capacity of the other phases should one phase reach 100% utilization before the others.

While it is typically not possible to balance the utilization on all three phases perfectly, minimizing the delta between phases is highly advantageous. Figure 7 shows that Phase A has reached 100% utilization, while Phases B and C have 25% and 50% respectively, of their capacity still available. That unused capacity is effectively stranded due to the unbalanced utilization pattern.

Figure 7. Unbalanced phase utilization



For this reason, each PSU of each system should be connected to a different “leg” (or phase) on the rPDUs. The onboard metering function of the rPDU provides an indication of power draw per phase or circuit, to assist in evaluating phase balancing. In a system with potentially complex power provisioning schemes, such as the DGX SuperPOD, phase balancing is especially important.

It takes two racks of systems to balance the phases while maintaining the availability and performance characteristics of the N+1 design. Rack 1 is described in Figure 8 and Table 12, with rack 2 being described in Figure 9 and Table 13. The PSUs on the systems are grouped in pairs, with each pair feeding from defined phases on specific rPDUs. Note that for each PDU that is sharing a common upstream UPS source, the feeder circuit for it is coming from a different floor PDU or busway, to maximize upstream diversity. For clarity of illustration, only two systems are depicted for each rack. However, the same patterns are employed with higher rack densities.

Figure 8. N+1 Phase balancing scheme—rack 1

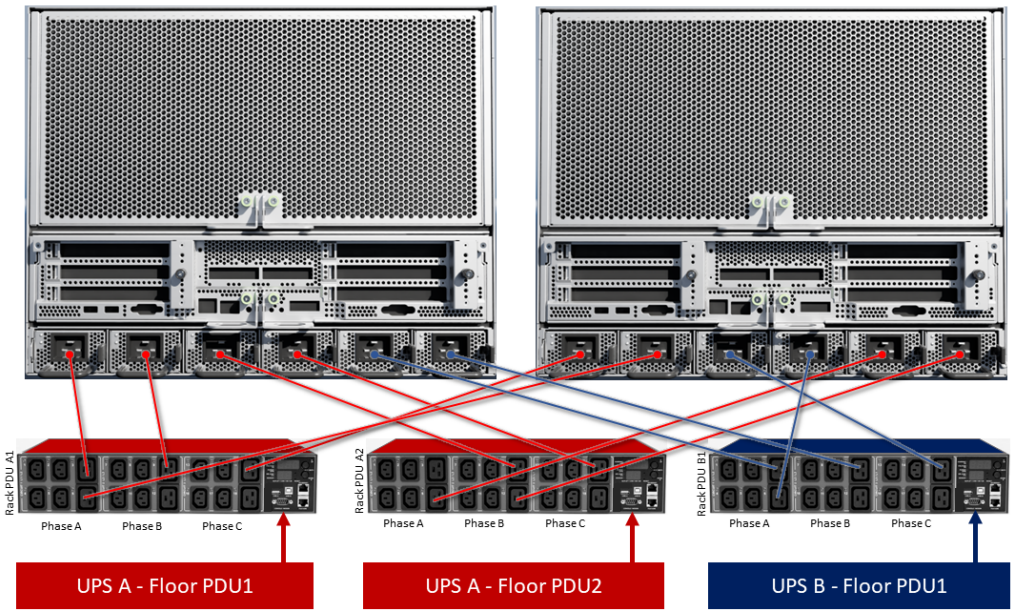


Table 12. Logical phase balancing table for N+1 power—rack 1

UPS/Gens	Floor PDU	Rack PDU	PSU Phase Assignments											
			DGX H100 #1						DGX H100 #2					
			PSU1	PSU2	PSU3	PSU4	PSU5	PSU6	PSU1	PSU2	PSU3	PSU4	PSU5	PSU6
Power Path A	PDU1	rPDU A1	A	B					C	A				
Power Path A	PDU2	rPDU A2			B	C							A	B
Power Path B	PDU1	rPDU B1					A	B			C	A		

Figure 9. N+1 Phase balancing scheme—rack 2

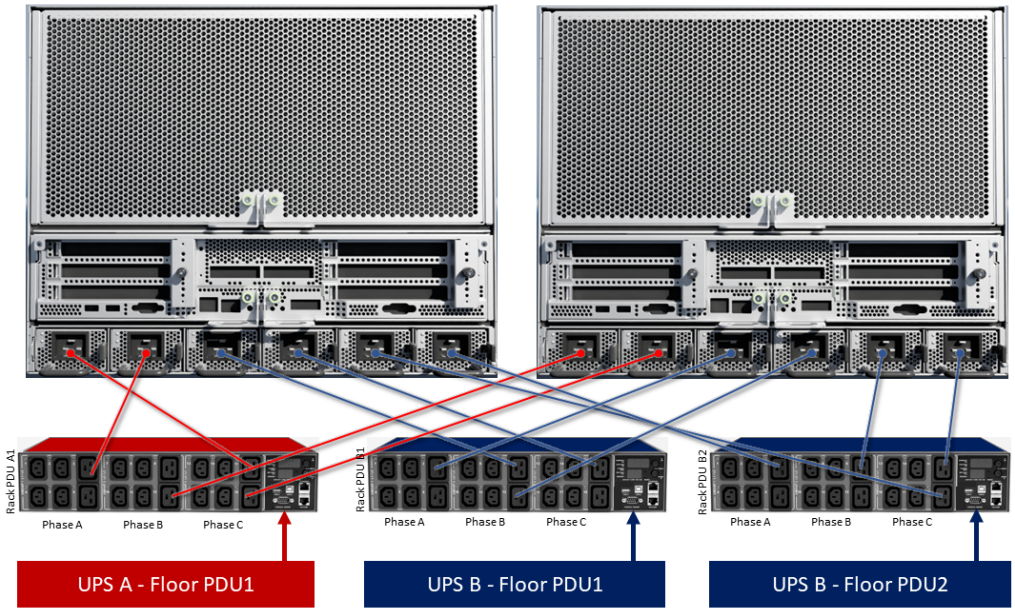


Table 13. Logical phase balancing table for N+1—rack 2

UPS/Gens	Floor PDU	Rack PDU	PSU Phase Assignments											
			DGX H100 #1						DGX H100 #2					
			PSU1	PSU2	PSU3	PSU4	PSU5	PSU6	PSU1	PSU2	PSU3	PSU4	PSU5	PSU6
Power Path A	PDU1	rPDU A1	C	A					B	C				
Power Path B	PDU1	rPDU B1			B	C					A	B		
Power Path B	PDU2	rPDU B2					C	A					B	C

For Enhanced N+1 power, the phase balancing scheme is slightly less complex. As with N+1, it takes two racks (Figure 10 and Table 14 for rack 1, Figure 10 and Table 14 for rack 2) to complete the phase balancing pattern.

Figure 10. Enhanced N+1 Phase balancing scheme—rack 1

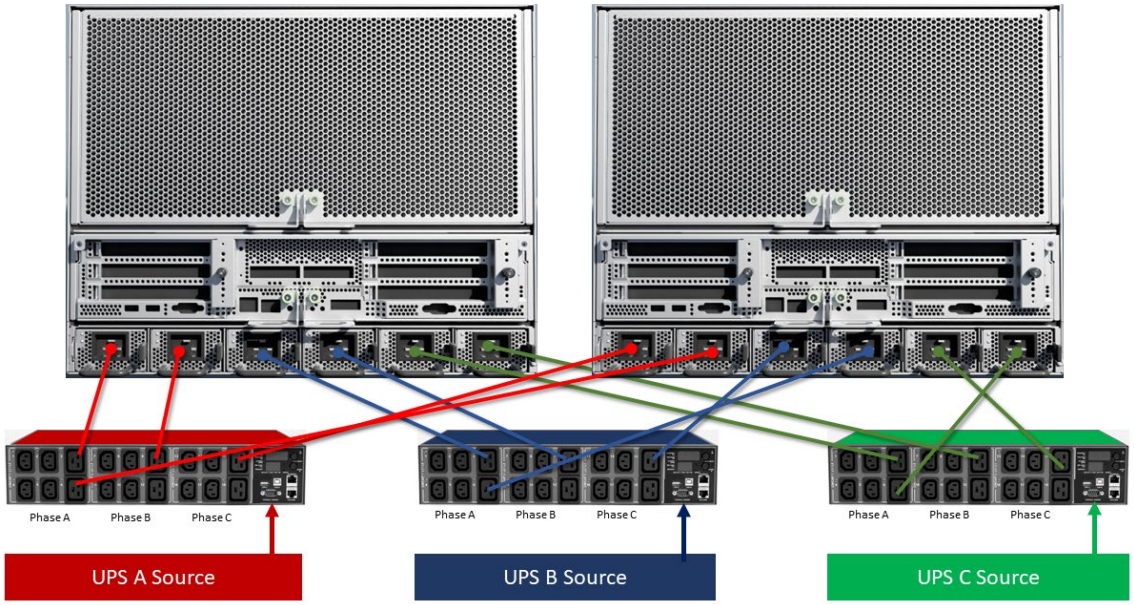


Table 14. Logical phase balancing table for Enhanced N+1 power—rack 1

UPS/Gens	Floor PDU	rPDU	PSU Phase Assignments											
			DGX H100 #1						DGX H100 #2					
			PSU1	PSU2	PSU3	PSU4	PSU5	PSU6	PSU1	PSU2	PSU3	PSU4	PSU5	PSU6
Power Path A	PDU1	rPDU A1	A	B					C	A				
Power Path A	PDU1	rPDU B1			A	B					C	A		
Power Path C	PDU1	rPDU C1					A	B					C	A

Figure 11. Enhanced N+1 Phase balancing scheme—rack 2

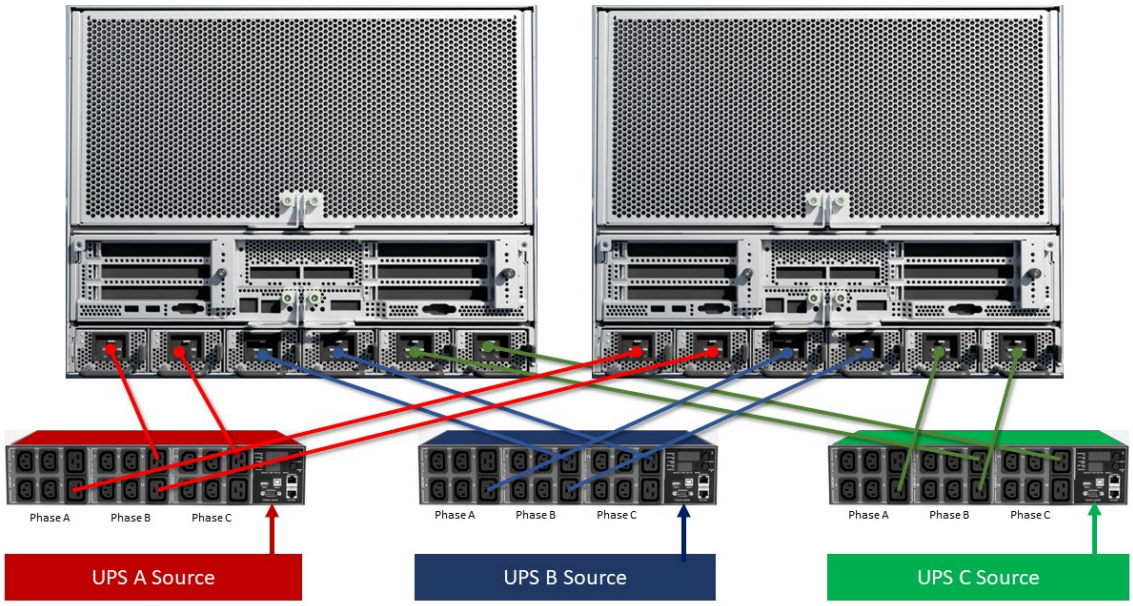


Table 15. Logical phase balancing table for Enhanced N+1 power—rack 2

UPS/Gens	Floor PDU	rPDU	PSU Phase Assignments											
			DGX H100 #1						DGX H100 #2					
			PSU1	PSU2	PSU3	PSU4	PSU5	PSU6	PSU1	PSU2	PSU3	PSU4	PSU5	PSU6
Power Path A	PDU1	rPDU A1	B	C					A	B				
Power Path B	PDU1	rPDU B1			B	C					A	B		
Power Path C	PDU1	rPDU C1					B	C					A	B

Chapter 4. White Space Infrastructure

4.1 Space Planning

When deciding where to place an SU, many important factors including power capacity, cooling capacity, cable routing and management, and adjacent equipment requirements are usually considered. While these factors have a significant impact on location selection, thought should also be given to scalability. As the DGX SuperPOD expands from one scalable unit to two or more, there must be adjacent space available for that expansion. Performance-based cable length limitations prohibit distributing the racks or scalable units too far from one another. A well-planned deployment will consider future expansion and reserve sufficient floor space for the future state of the system, not just the initial state.

4.2 Rack Standards and Requirements

Racks must conform to EIA-310 standards for enclosed racks with 19" EIA mounting. Cabinets must be at least 24" × 48" (600 mm × 1,200 mm) in size, and at least 48U tall. For proper cable management, rPDU placement, airflow management, and service clearance to the rear of the systems, NVIDIA recommends 32” × 48” (800 mm × 1,200 mm) racks. Racks must not have chimneys. Side walls must be installed in all racks.

For each rack, two temperature sensors will be connected to Ethernet ports in the PDU closest to the front of the rack. The sensors must be mounted at the front side of the rack, at the 4U position and at the 42U position. When facing the front of the rack, the sensors will be on the right side. Telcordia GR-63-CORE may also be followed for thermal rack measurements. Cable management devices are prescribed and must be used. For network switches, Air Intake ducts may be prescribed.

IT cabinets come in a variety of sizes and are often designed for specific purposes. Each cabinet OEM follows specific minimum EIA-310 standards to ensure that industry standard devices fit properly. But OEMs will enhance the EIA-310 standard with their own unique designs and features that allow them to stand out in the marketplace. These features may include:

- Air management
- Cable management
- Modular sub-components
- Removable components
- Proprietary accessories
- Supplemental security devices
- Custom manufacturing, colors, company logos, and so on

4.3 Options When Ordering Cabinets

- Cabinet Doors
  - Unless mandated by facility operations or security policy, front and rear doors are not recommended. Eliminating the front and rear doors helps improve airflow and reduces cost.
- Side panels
  - Side panels should be added and installed to maximize airflow management.
- Grounding and Bonding Kit
  - This kit is often an optional item and may not come with the cabinet. Proper grounding and bonding are essential.
- Blanking Panels
  - It is necessary to install blanking panels in each unused RU position to prevent exhaust air recirculation.
- Cabinet Top Options and Accessories
  - Many cabinet manufacturers offer a range of cabinet top options and accessories to facilitate cooling management and cable routing management.
  - Whether it is a standard feature or an optional feature, cabinets must have sufficient cable ingress ports to support the volume of cabling traversing into the rack, and all such ports shall be protected with brush grommets or similar devices to control airflow. See *Cabling Data Centers* for more information.

4.4 Cabinet Mounting

For the safety of staff members and equipment, all cabinets shall be fastened to the floor surface in accordance with manufacturer recommendations. This may involve bolting the cabinet through flanges or mount points designed for that purpose.

4.5 Seismic Considerations

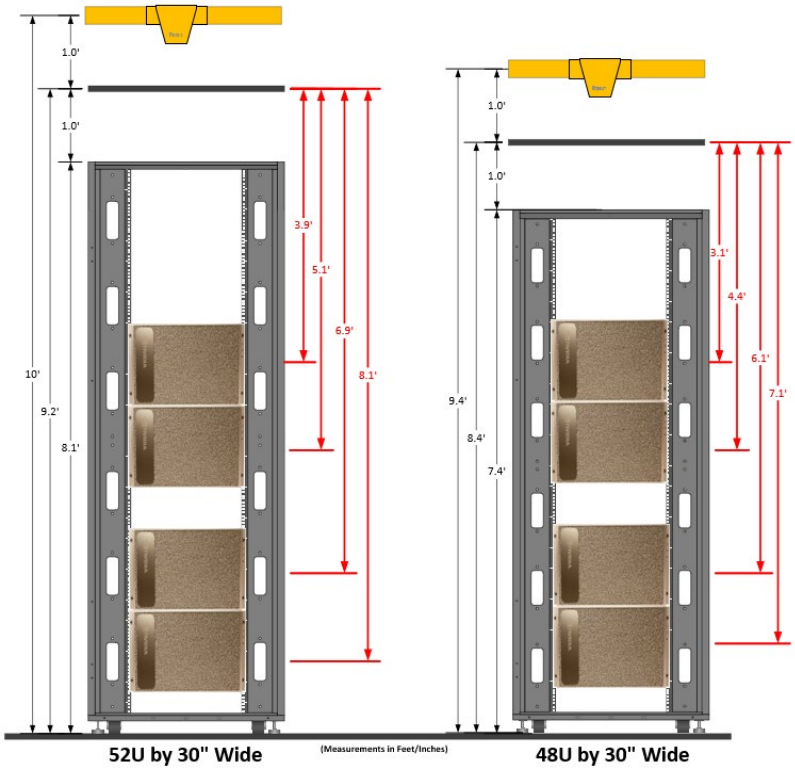
Wherever mandated by local authorities having jurisdiction (AHJs), cabinets mounted on raised floor surfaces may require seismic bracing to be installed in the subfloor space. Seismic bracing should be designed and installed by qualified licensed structural engineers specializing in seismic engineering.

4.6 Cabinet Selection vs. Cable Lengths

The cabinet height and width combined with the overhead cable tray specification and layout will affect cable lengths. Cable length in InfiniBand networks is a critical performance factor. Point to point cable runs should be limited to a maximum of 165 ft (50m). Wherever possible, cable length should be a primary design criterion when selecting cabinet dimensions and designing overhead cable routing apparatus.

When selecting specifying rack width (Figure 12), attention should be given to the impact on cable lengths and row width.

Figure 12. Dimensions for various cabinet options



While wider racks are beneficial, it may be necessary to reduce the number of racks per row, or distribute the racks across two rows, so that maximum row width and optimum cable lengths can be maintained (Figure 13).

Figure 13. Row width based on rack width

4.7 Server Mounting Requirements

All DGX SuperPOD equipment is designed to be mounted in traditional IT server cabinets that conform to the EIA-310-D standard (Figure 14), which among other factors specifies the following:

- Vertical Hole Spacing
  - Vertical hole spacing is defined as a repeating pattern of holes within 1 RU of 1.75". The hole spacing alternates at: 1/2" – 5/8" – 5/8 and repeats. The start and stop of the "U" space is in the middle of the 1/2" spaced holes.
- Horizontal Spacing
  - The horizontal spacing of the vertical rows of holes is specified at 18 5/16" (18.312) (465.1 mm).
  - Many manufacturers use equipment mounting slots instead of holes to allow for variations in this dimension.
- Rack Opening
  - The space in the rack where the equipment is placed is specified as a minimum of 17.72" (450 mm) wide.
- Front Panel Width
  - The total width of the front face of the equipment (with its rack mounting brackets) is 19" (482.6 mm)

Figure 14. DGX H100 system mounted in EIA-310-D compliant server cabinet

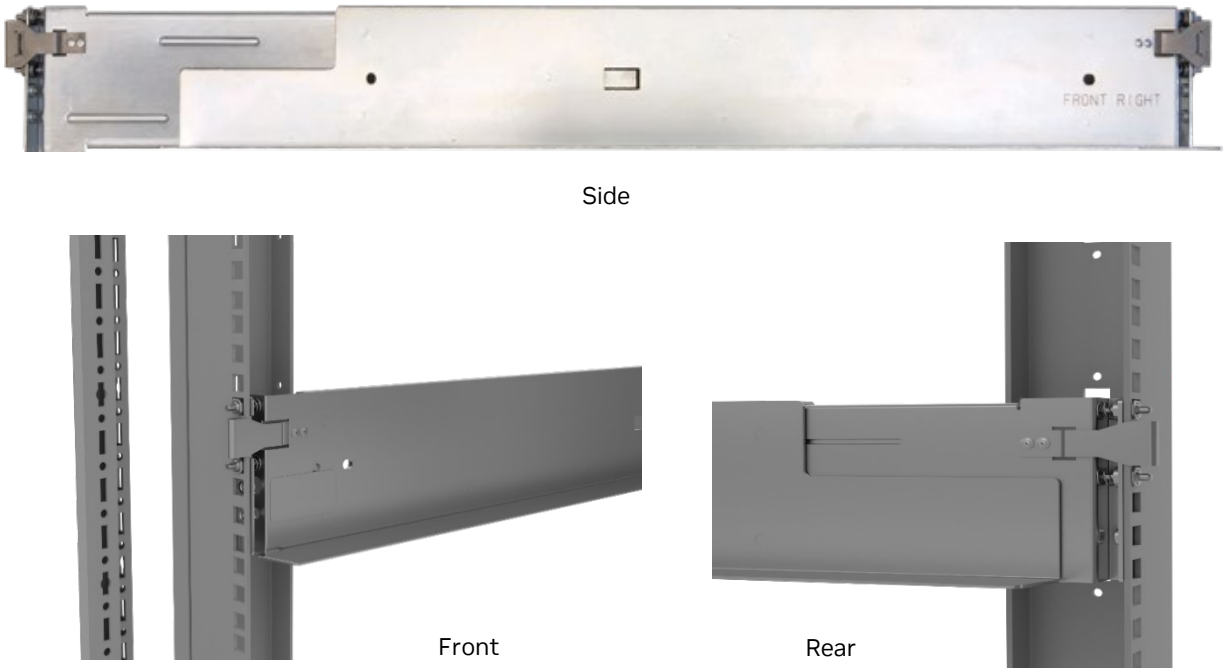
4.8 Racking Servers

Another area requiring design attention is how and where the equipment will be mounted in the cabinet. This first involves setting the proper distance for the rack rails. This will set the horizontal orientation within the cabinet.

Some of the DGX SuperPOD devices are installed using a rail kit. A rail kit is a bracket specifically designed to support the full weight of the device in the rack, by spanning from the front rack rail to the rear rack rail. Rail kits are either a stationary shelf type bracket, or a bracket that allows the device to be pulled out of the rack enclosure on retractable struts. These rail kits have an extension range of 28–32 inches. Wherever provided, rail kits must be used in accordance with manufacturer recommendations.

Rail kit components for the DGX H100 system are shown in Figure 15.

Figure 15. DGX H100 rail kit



Due to their size and weight, servers should be mounted nearest the bottom of the rack, starting with the lowest position accessible by the server lift (usually rack unit position 3). It is optional, but not necessary, to have 1 RU between each server.

A typical high-density rack profile will have two servers grouped at the bottom of the rack, a 3 RU gap in the middle to aid in rack airflow management, and then two servers grouped above that (Figure 16). Horizontal rPDUs are typically placed near the top of the rack.

Figure 16. Server deployment positions for various rack densities

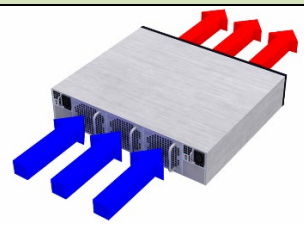
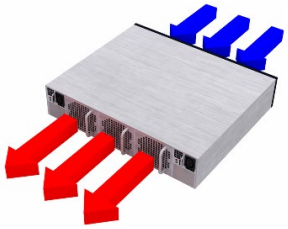
4.9 Air Flow Management

Some IT devices have the option of selecting the airflow direction through the device chassis. The determining factor for this choice is the expected hot aisle temperature. Servers and DGX H100 systems will always be oriented with front to rear airflow. However, network switches and appliances are typically configurable to align their airflow direction with front or rear rack placement. Typically network switches are mounted in the rear of the rack, with the network connectors facing the hot aisle. In this orientation, airflow enters through the back of the device chassis (the Power or “P” side) and is exhausted through the front of the device chassis (the network Connector or “C” side). This type of airflow orientation is referred to as “Power to Connector,” or “P2C”. This is the default mounting configuration and airflow orientation for data center network equipment, and the standard configuration for cable pathways into and out of server/network cabinets.

If the hot aisle will reach temperatures greater than 60 ℃ (140 ℉), cabling (including power cabling) must either be moved to the cold aisle or derated. This means that the network switches will likely be mounted in the front of the rack with the network ports facing the cold aisle. In this orientation, air is ingested on the front “network connector” or “C” side of the chassis and exhausted from the rear “Power” or “P” side of the chassis. This airflow orientation is referred to as “Connector to Power,” or “C2P”.

Examples of P2C and C2P are shown in Table 16.

Table 16. Airflow orientation

Direction	Description
	Power side inlet to Connector side outlet Designation: P2C
	Connector side inlet to Power side outlet Designation: C2P

The hot aisle temperature is directly correlated to the supply air temperature in the cold aisle, and the Delta T or expected heat rise of the equipment itself (usually 25–30˚F). For example, with a supply air temperature of 70˚F, and an expected Delta T of 30˚F, the expected hot aisle temperature would be 100˚F. Hot aisle temperatures exceeding 60˚ C/140˚F are not typical.

If a C2P orientation is necessary, this choice must be made during BoM selection, because in many cases air flow is determined by the power supply fans and is not field modifiable.

4.10 Static Weight and Point Load

A typical IT Cabinet has an average unloaded weight of 350 Pounds (158 Kg). An individual DGX H100 system weighs 287.6 Pounds (130.45 Kg). In addition to the servers and the cabinets themselves, peripheral devices such as rPDUs, blanking panels, cable management apparatus, environmental sensors, and cabling add additional weight. It is necessary to ensure that all flooring structures (including subfloor structures and slabs wherever applicable) are engineered to support the combined weight of the equipment racks they must support. It is also important to ensure that all ingress and egress pathways between the loading dock and the server room floor are engineered to support

the combined weight of the equipment plus any conveyances used to move the equipment to the rack location. Table 17 lists the estimated (rounded) weight of different rack profiles. These are only general estimates. The exact weight of a specific loaded rack would depend on the unloaded weight of the actual rack model that was selected, as well as any extraneous peripheral components and cabling.

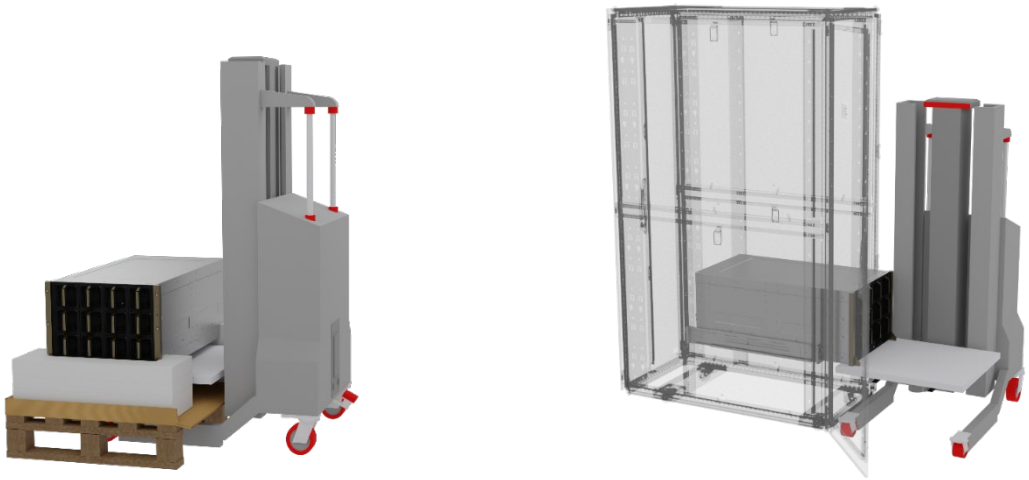
Table 17. Weight profiles of DGX H100 system racks

Number of DGX Systems per Rack	Total Rack Weight		Point Load	
	Pounds	Kilograms	Pounds	Kilograms
1	650	295	217	98
2	925	420	308	38
4	1500	680	500	226

4.11 Server Lifts

Due to the heavy weight of the DGX H100 system, it is necessary to use a server lift (Figure 17)to install and remove the devices from the racks, and to transport them to the rack location.

Figure 17. Server lifts



Proper server lifts are designed specifically for use in data center environments to lift heavy but sensitive devices such as servers and network switch chassis into server cabinets. They include a platform to hold and lift the device, not forks that are like a forklift. General purpose materials lifts, including any lift that uses forks in place of a platform, are not suitable for use as server lifts—even if they are advertised for such uses. Use the server lift in compliance with all safety precautions and protocols as specified by the manufacturer. Ensure that the server lift is rated for a minimum of 350 pounds of weight. Ensure proper clearance below any overhead obstructions before lifting the device.

4.12 Security, Noise, and Fire Prevention

In addition to the physical compute and network infrastructure, certain other site infrastructure and safety factors should also be considered when implementing a DGX SuperPOD. This section provides an overview of the top external considerations.

4.12.1 Physical Security

The DGX SuperPOD, including servers and network infrastructure, should be protected against unauthorized physical access. Access controls to the data center are required to prevent system tampering, theft of intellectual property, data copying or unauthorized removal of data. Data center site security measures should minimally be certified SOC compliant, in addition to any other authorities having jurisdiction. This includes auditable door access controls and cameras that can identify a person entering the space and record that person's actions within the space. Depending on the security policies and requirements governing the data and applications, DGX SuperPOD racks should be isolated from other unrelated IT racks using partitions or cages with access limited only to persons authorized to service the NVIDIA racks.

4.12.2 Noise

NVIDIA DGX H100 systems can generate noise levels greater than 98 dB at 1m distance. Noise reduction measures and hearing protection are the responsibility of the data center operator and should be provided in accordance with the requirements of local authorities and industry regulations. See Appendix A for further detail regarding noise risks and mitigations.

4.12.3 Fire Protection

Fire detection and fire prevention systems/equipment are required in the data center. Regulations vary from one jurisdiction to another, resulting in the use of different fire detection and suppression schemes from one data center to another. A fire detection and suppression system is the responsibility of the data center operator and should be installed in accordance with the requirements of the customer's insurance underwriter, local fire marshal, and local building inspector for the correct level of coverage and protection.

Chapter 5. Networking

A detailed examination of the network design and architecture is beyond the scope of this document. See the *NVIDIA DGX SuperPOD Reference Architecture* for an in-depth review of the network architecture, connectivity, and equipment. In the following sections, network infrastructure will be discussed only to the extent that it impacts data center design considerations.

Similarly, a detailed examination of data center cabling and cable management technology is also beyond the scope of this document. See *Cabling Data Centers* for more information.

5.1 Cable Weight

Due to the volume of cables traversing the various cable pathways, it is important to ensure that all cable pathways have sufficient size and weight bearing capacity to facilitate the load. Generally, in the planning phase, the weights and sizes shown in Table 18 can be used to calculate load-bearing requirements for cable pathway apparatus such as cable trays, ladder racks, and strain reliefs, however, these are only a general guide. Use the actual specifications of the cables being installed once they have been determined.

Table 18. Cable weight by type

Cable Type	Average Single Cable Diameter	Average Single Cable Weight
Copper	9.43 mm ± 0.4 mm	0.03 pounds per linear foot 0.045 Kg per linear meter
AOC Fiber Optic	3.0 mm ± 0.4 mm	0.006 Pounds per linear foot 0.009 Kg per linear meter

There is a large concentration of cables in the central management racks as well as in the supporting cable pathways overhead. Pre-planning for the cabling infrastructure is a critical part of the design process. Areas of analysis include:

- Top of rack cable penetration access.
- Cable tray types.
- Cable tray width and depth.
- Cable tray fill ratios.
- Structural load/weight planning.
- Cable management devices.

All cable management and pathway apparatus should be selected, installed, and maintained in accordance with applicable industry standards including ANSI/TIA, BICSI, and NEMA. Further, it should be installed in compliance with all locally applicable fire safety and electrical codes, such as National Electric Code (NEC) in the United States and Mexico, The Canadian Electric Code in Canada, The International Electrotechnical Commission (IEC) in Europe, The British Standard (BS) 7671 in the United Kingdom, and NF C 15–100 in France.

Compute fabric component and cable counts are shown in Table 19.

Table 19. Compute fabric component count

SU Count	Cluster Size # Nodes	Cluster Size # GPUs	Leaf Switch Count	Spine Switch Count	Compute + UFM Node Cable Count	Spine-Leaf Cable Count
1	31[1]	248	8	4	252	256
2	63	504	16	8	508	512
3	95	760	24	16	764	768
4	127	1016	32	16	1020	1024

1. This is a 32 node per SU design; however a DGX H100 system must be removed to accommodate for UFM connectivity.

Front and rear views of a typical overhead cable management system of SU are shown in Figure 18 and Figure 19.

Figure 18. Typical overhead cable management of a SU—front

Figure 19. Typical overhead cable management detail—rear

When selecting an overhead cable tray design, use the following formulas to calculate how many cables it will support:

- A = Inside tray area (in square inches)
- D = Cable diameter (inches)
- F = Fill ratio (percent)
- N = Number of cables

$$N = (F/100) * (A / [(D/2)2 * \pi])$$

This formula can be used to determine that a tray with a cross sectional area of 11.5 in^2 (7,419 mm^2) is the minimum size that will support the volume of cables in an SU, if the cables have an average diameter of 3 mm (such as an AOC cable) and a maximum tray fill ratio of 50% (it is not advisable to fill cable trays higher than 50%).

Care should be taken to ensure that any cable tray apparatus is structurally sound and designed for the weight of the cables that it supports. Typically, straight sections of cable tray are supported at 5 ft intervals, while curved sections may require additional supports. Any span greater than 5 ft in length should be evaluated to determine if additional supports are required. Supports should be placed within 1 ft of any splice or joint in the cable tray structure, or tray height transition. The support structures and the anchors that hold them should each be evaluated separately to ensure they will hold the combined weight of the cables, the tray apparatus, and supporting structures.

As a reference, for a single SU the structural load of the fiber cabling is 508 cables * 0.009 Kg per cable per linear meter, or 4.57 Kg (10 lbs) per linear meter. Again, the calculations illustrated here are general guidelines, and calculations based on the size and weight specifications of the specific quantities of the actual cables to be used should be performed.

Chapter 6. Cooling and Airflow Optimization

It is critical to plan for the full heat load of the rack profiles, keeping in mind that the power provisioning is based on circuits that provide only 50% of the full load. With traditional 2N redundant power provisioning schemes, the cooling capacity is generally aligned with the capacity of N, but N is usually the capacity of a single power circuit. However, with the specified N+1 power provisioning, N equals two circuits. Therefore, it is critical to align the cooling capacity with N, and not simply the capacity of a single power circuit.

Some data center designs may have constraints on cooling capacity that will require mitigation as part of the DGX SuperPOD deployment plan. The following sections reveal common strategies for mitigating these constraints.

6.1 Foundational Concepts

Before considering more drastic cooling mitigations, it is important to make sure that the airflow in the space is optimized and well managed. While the following steps may be rudimentary, their importance cannot be overstated.

6.1.1 Row Orientation

The rows of racks on a data center floor are usually arranged in such a way as to create a hot aisle and a cold aisle. These aisles are created by orienting the racks so that the backs of two opposing rows of racks face one another in one aisle, and the fronts of two rows of racks face each other in the next aisle. Supply air is delivered to the “cold aisle,” and exhaust air is evacuated from the “hot aisle.” It is important to space these rows carefully, so that the width of the cold aisle is sufficient to deliver the required volume of air for all the racks it serves, and the width of the hot aisle is sufficient to prevent racks with higher powered servers from interfering with the exhaust airflow from lower powered servers in the opposing row. At a minimum, the aisles should be at least 36 inches wide, and it is strongly recommended that the cold aisle be a minimum of 48

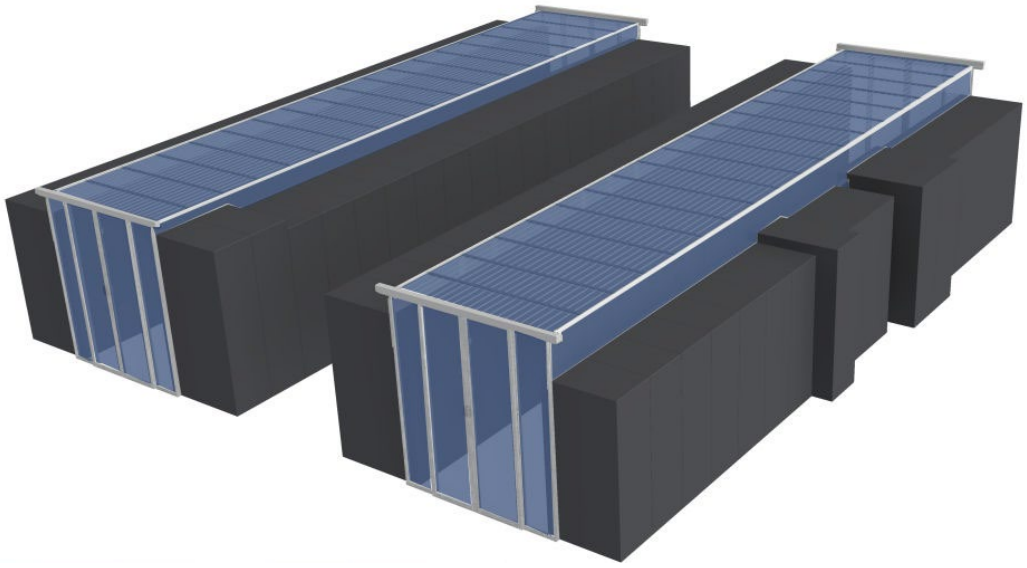
inches wide, to allow for the safe navigation and use of server lifts, technical carts, and other conveyances within the aisle.

6.1.2 Aisle Containment

Many data centers employ aisle containment strategies to help manage and optimize airflow, particularly for high-density racks. Data Center designers may choose to contain either the cold aisle or the hot aisle depending on the means of air delivery in a data center space. In either case, the main benefit of aisle containment is the prevention of air recirculation from the hot aisle to the cold aisle, which artificially increases the supply air temperature at the inlet of the servers, significantly reducing its heat exchange potential.

Figure 20 illustrates four rows of racks grouped into two cold aisle containment structures, with an uncontained hot aisle between them. These structures are typically made of some form of clear acrylic (or other similar material) partition panels encompassing the front or back of the racks, and perpendicular partitions at the ends of the rows, often featuring self-closing doors for ingress into the contained area.

Figure 20. Aisle containment apparatus



Aisle containment cannot be truly effective unless all pathways for airflow between the hot aisle and the cold aisle are blocked. For this reason, unoccupied RU spaces in the rack should be covered with blanking panels, and openings in the top, sides, or bottom of the rack, or subfloor, which are used for cable pass-throughs should be fitted with brush grommets.

6.1.3 System Operation and Maintenance

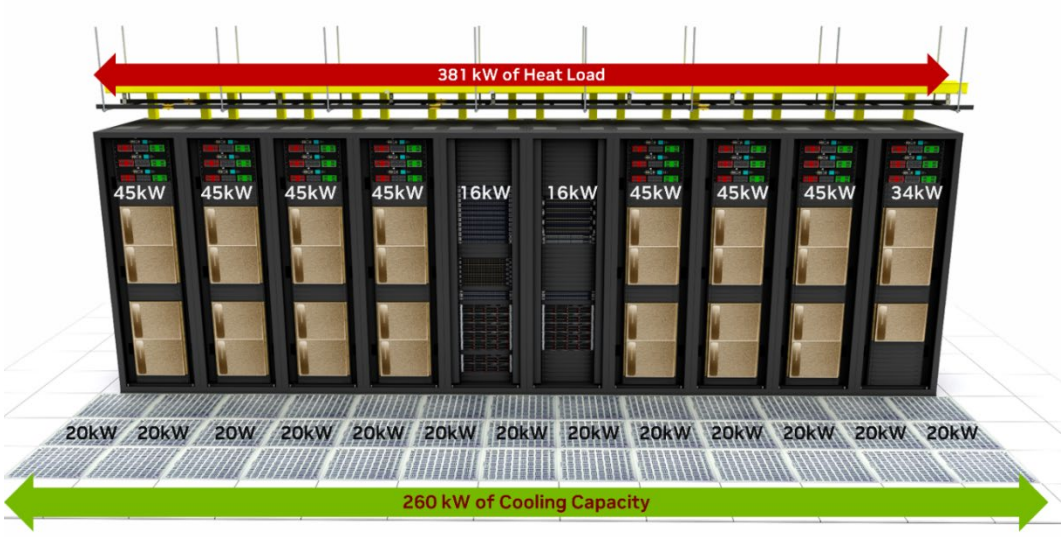
Several system maintenance steps can help ensure optimal cooling performance. Among these are routine air filter changes, using the correct filter specifications, routine measurement of humidity levels within the space (with prescribed corrective actions when tolerances are not maintained), routine audits of the flow rates of perforated tiles in a plenum floor, and routine preventative maintenance cycles on all air handlers. It is also highly advisable to have accurate Computational Fluid Dynamics models of the data center space created so that planned changes can be modeled prior to implementation to assess their potential impact on critical systems.

6.2 Cooling Oversubscription

When there is a delta between the capacity of a resource in the data center and the demand for that resource, the resource is said to be oversubscribed. Oversubscribed cooling can sometimes be mitigated by lowering the rack density thereby reducing the cooling demand per rack footprint, or by spacing the racks further apart to aggregate the cooling capacity of more than one rack footprint to each populated rack. Either solution consumes more data center space, and perhaps more importantly, requires longer cable runs to interconnect the racks. Careful attention should be given to cable length as it relates to these potential solutions.

Consider the following deployment scenario with a single SU and its management racks being deployed in a high-density deployment pattern in an area of the data center that has constrained cooling capacity. Figure 21 depicts 381 kW of heat load in a deployment pattern that supplies only 260 kW of cooling capacity.

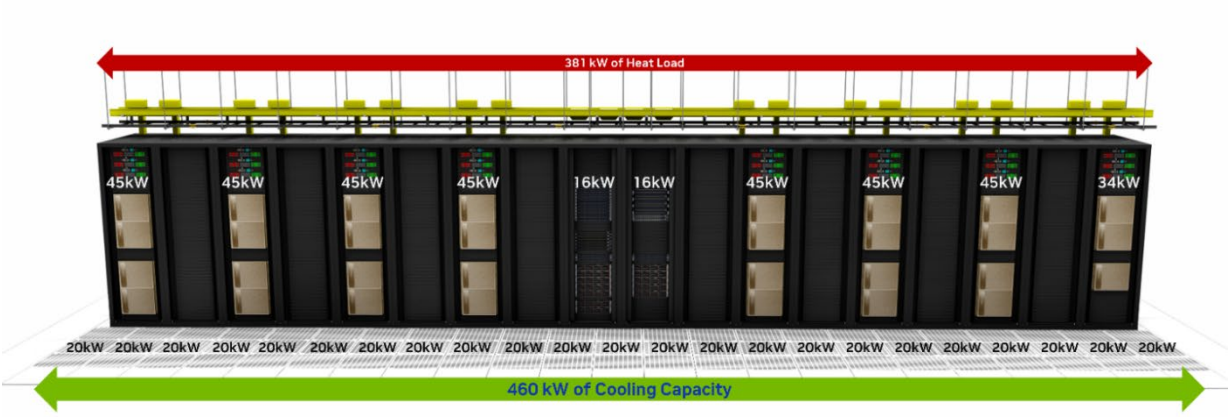
Figure 21. Cooling Oversubscription Scenario



In this scenario, if the server racks were spaced apart by one rack footprint, additional cooling capacity can be leveraged.

Figure 22 depicts the same deployment pattern, spaced apart in this manner. Now, the 381 kW of demand is supplied by 460 kW of cooling capacity.

Figure 22. Resolved cooling oversubscription



As previously noted, the cooling oversubscription could also have been resolved by reducing the rack density. This may have consumed even more additional space, depending on the actual cooling capacity per rack footprint, or it may have resulted in more stranded power capacity depending on the power provisioning options available for the racks. The objective is to resolve any oversubscription scenario using the method that consumes or wastes the fewest alternate resources (in this case, space and/or power) to arrive at the most optimal deployment pattern possible.

Chapter 7. Summary

The successful deployment of a DGX SuperPOD relies on the careful coordination and collaboration of various teams and domains of expertise across the organization. Numerous NVIDIA resources are available to assist in the planning and deployment. Emphasis should be placed on the power and cooling requirements that may impact data center space planning, InfiniBand network cabling plans, and cooling optimization strategies. Standards based methodologies, worker safety, and industry best practices for HPC installations should be a foundational element of the data center design and deployment plan.

Appendix A. Sound Mitigation

The Acoustic Power (LWA,m) of each DGX H100 system is rated at 98.7 dB at 80% Fan PWM (30 °C/86 °F). Noise levels up to106 dB per server are possible. When multiple servers are placed within proximity, such as a row of server racks, each containing multiple servers, the noise level increases logarithmically. Table 20 shows the aggregate sound level of a given quantity of servers based on 98.7 dB acoustic power of a single server. These numbers do not account for differences in distance that may result from different floor layouts of the racks.

Table 20. Sound pressure level for a configuration of multiple servers

Server Quantity	Difference in Sound Intensity (dB) relative to single server	Total System Sound Level at This Server Qty (dB)
1	0.00	98.70
2	3.01	101.71
3	4.77	103.47
4	6.02	104.72
5	6.99	105.69
6	7.78	106.48
7	8.45	107.15
8	9.03	107.73
9	9.54	108.24
10	10.00	108.70
11	10.41	109.11
12	10.79	109.49
13	11.14	109.84
14	11.46	110.16
15	11.76	110.46
16	12.04	110.74
17	12.30	111.00
18	12.55	111.25
19	12.79	111.49
20	13.01	111.71
21	13.22	111.92
22	13.42	112.12
23	13.62	112.32
24	13.80	112.50

25	13.98	112.68
26	14.15	112.85
27	14.31	113.01
28	14.47	113.17
29	14.62	113.32
30	14.77	113.47
31	14.91	113.61
32	15.05	113.75

Further, while the preceding numbers represent the noise being transmitted from the servers themselves, the acoustic properties (lack of absorbent surfaces, reflection of emitted sound, etc.) and ambient noise level of the room environment can add significantly to these measurements.

Table 21 provides a reference for these noise levels, relative to commonly experienced sounds.

Table 21. Relative sound levels

Source	Intensity	dBA
Threshold of Hearing (TOH)	$1*10^{-12}$ W/m^2	0 dB
Rustling Leaves	$1*10^{-11}$ W/m^2	10 dB
Whisper	$1*10^{-10}$ W/m^2	20 dB
Normal Conversation	$1*10^{-6}$ W/m^2	60 dB
Busy Street Traffic	$1*10^{-5}$ W/m^2	70 dB
Vacuum Cleaner	$1*10^{-4}$ W/m^2	80 dB
Large Orchestra	$1*10^{-3}$ W/m^2	98 dB
Headphones at Maximum Level	$1*10^{-2}$ W/m^2	100 dB
Front Rows of Rock Concert	$1*10^{-1}$ W/m^2	110 dB
Threshold of Pain	$1*10^{1}$ W/m^2	130 dB
Military Jet Takeoff	$1*10^{2}$ W/m^2	140 dB
Instant Perforation of Eardrum	$1*10^{4}$ W/m^2	160 dB

For reference, a difference of 10 dB(a) represents a 100% increase in perceived sound level. Using the preceding tables, it becomes apparent that a row of 32 DGX H100 systems can be louder than a major rock concert, and over four times louder than a typical vacuum cleaner, even with the fans running at only 80% PWM.

The maximum acceptable unprotected exposure limit for a standard 8-hour duty cycle is a time weighted average of 85 dBA. The purpose of PPE and engineering controls is to attenuate audible sound down to 85 dBA or lower. If that is not possible, the worker's duty cycles must be reduced to the durations allowed by the remaining unattenuated dBA. So, for example, if with maximum hearing protection the audible sound level is still 100 dBA, then the duty cycle must be reduced to 15 minutes.

The highest NRR rating for PPE earplugs is 33, and the highest available NRR rating for earmuffs is 31. These values reflect the level of noise protection available for each device when worn alone. Combining earplugs with earmuffs can offer an NRR protection level of 36.

The NRR Rating of a PPE Hearing protection device does not represent actual sound attenuation levels achieved in practical use. The formula for determining actual attenuation is NRR -7 * 0.5. When two PPE devices are used, this calculation is performed on the higher rated device, and 5 dBA is added to the result regardless of the NRR value of the second device.

Practical attenuation beyond 18 dBA requires mechanical sound mitigation or elimination (a.k.a. engineering controls) and is not attainable with PPE alone.

Data center staff should undergo routine hearing tests by a certified audiologist to establish a baseline and identify any hearing loss over time, so that appropriate and timely interventions can be made.

Table 22 illustrates the maximum exposure limits for certain sound levels, as regulated by the National Institute for Occupational Safety and Health (NIOSH).

Table 22. NIOSH Standard Exposure Limits

Sound Level (dBA)	Maximum Unprotected Exposure Duration		
	Hours	Minutes	Seconds
82	16	0	0
85	8	0	0
88	4	0	0
91	2	0	0
94	1	0	0
97	0	30	0
100	0	15	0
103	0	7	30
106	0	3	45
109	0	1	53
112	0	0	56
115	0	0	28
118	0	0	14
121	0	0	7
124	0	0	3
127	0	0	1
130	0	0	0.5

Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation (“NVIDIA”) makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgment, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer (“Terms of Sale”). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer’s sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application to avoid a default of the application or the product. Weaknesses in customer’s product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem that may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, “MATERIALS”) ARE BEING PROVIDED “AS IS.” NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA’s aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

Trademarks

NVIDIA, the NVIDIA logo, NVIDIA GeForce NOW, NVIDIA GeForce GTX, and NVIDIA SHIELD= are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright