NVIDIA Requirements for AI Clouds

Technical Reference Manual

TRM-12815-001_v2.2 | March 2026

Document History

TRM-12815-001

Version	Date	Authors	Description of Change
2.1	Feb 26, 2026		Initial version
2.2	April 10, 2026		Update to v2.2

Table of Contents

Purpose and Intent

This document serves three main purposes:

1. **Setting requirements for NVIDIA Cloud Partners (NCPs) delivering GPU capacity to NVIDIA**
   This is the primary requirements document from NVIDIA to any NCP providing NVIDIA GPU/AI compute and software services. These requirements cover the full stack of AI cloud infrastructure services and operations needed to run NVIDIA DGX Cloud, expanding on the NVIDIA hardware reference design.
2. **Providing a reference set of requirements for the industry**
   NVIDIA is publishing this document openly so that NCPs, GPU datacenter operators, and AI practitioners can use it as a reference for the capabilities a large GPU consumer requires
3. **Defining NVIDIA's service delivery expectations**
   NVIDIA expects services to be delivered as Generally Available to all — not bespoke implementations built for NVIDIA alone. NVIDIA expects operational excellence, transparent communication, and proactive engagement from all partners.

NVIDIA will consider additional services that an NCP offers or plans to offer beyond what is described here.

Service Delivery SLAs

NCPs should be able to demonstrate ability to meet below SLA by category and operational requirements to be considered for offtake.

Services Delivery Timelines

The NCP must demonstrate API readiness, transport establishment at least 12 weeks ahead of GPU delivery, and the ability to provide Dev capacity (CPU nodes only) with the API integrated 6 weeks prior to GPU and cluster delivery.

One key request is for early access to ancillary compute nodes to act as the Data Mover function. This will help us pre-position data into the data center for use when GPUs are available. Access to Data Mover compute (and target storage) should be available ~2 weeks ahead of GPU cluster delivery.

SLA and SLO

Managed K8s

- Control Plane SLA target: Financially-backed 99.95%+ uptime for production.

Storage

- Performance (QoS): Must provision needed throughput requested for minimum bandwidth and IOPS.
- Home Directory Storage:
    - Availability: Over 99% availability for unplanned incidents. Exclusive of scheduled maintenance.
    - Durability: Over 99.99% for any FS less than 1 PB
- High Speed Storage Service Requirements:
    - Availability (SLO): Must meet 99.99% availability in a 30-day rolling SLO exclusive of maintenance
- High-Speed Storage Filesystem Requirements
    - End to End Availability: Over 99.5% uptime per PB
    - Durability: Over 99.999% durability per PB annually

Operational Requirements

- Dedicated Technical specialist/engineer available to NVIDIA
- Slack channel monitored by technical specialist / engineer
- 24x7 support available per partner standard incident severity procedures
- Service impacting incidents, planned, and unplanned maintenance events are communicated to NVIDIA.
- For planned maintenance, NVIDIA can schedule maintenance windows via APIs / console tools - avoiding unexpected outages + the ability for NVIDIA to provide feedback.
- NCP to remediate critical vulnerabilities in a timely manner while providing transparent disclosures of any issues

Telemetry Delivery Method

NCP shall deliver all required telemetry, including metrics and logs, in a manner that allows for ingestion into DGX Cloud systems. The preferred methodology is natively via the OpenTelemetry Protocol with a latency of no longer than 120 seconds.

Exemplar Cloud Workload Performance

NVIDIA Exemplar Cloud seeks to improve performance per TCO with hardware and software recipes, references, tools, and capabilities. Run the latest publicly available release from https://github.com/NVIDIA/dgxc-benchmarking **(Always pick the latest release version from the GH repo)** to be successfully completed on 1 uniform HW cluster type. Please run all the workloads for a given release and share the results in the template below.

Test ID	Feature	Min Size	Description
BM01	Benchmarking for exemplar cloud	512 GPU cluster	Achieve within 5% of an NVIDIA provided target performance number

Compute and Network Provisioning

This section outlines the requirements for provisioning compute and network. Compute instances can be provided as either Bare Metal instances (via BMaaS) or Virtual Machines (via VMaaS) to support the NVIDIA DGX Cloud engagement. All operations must be controlled via a fully documented and secure API, gRPC or REST preferred. All systems are expected to scale and perform at scale.

General, Compute and Lifecycle Management

Req ID	Test Details (Legend)	Requirement Area	Description
CNP01	BM: #49, #52, #53, #57, #106, #108, #105, VM: #42, #45, #46, #47, #50, #110, #111, TBD-topo	API/CLI Access	DGXC must have API or CLI access to the NCP provisioning system for: (1) Node lifecycle management (create, update, delete, list, or manage power states (reboot, on/off power cycle); (2) Network configuration; (3) Inventory and topology discovery; (4) security configuration (users, service accounts, groups, roles); (5) Maintenance and operations (see later section) (storage discussed in storage section)
CNP02	INFO	Declarative Resource Interfaces	For resources requiring multiple steps and a workflow, please provide the appropriate mechanism. A terraform provider is preferred. E.g. automating filesystem provisioning
CNP03	TBD-topo	NVLink-Aware Allocation	For NVL72 the API must support NVLink domain-aware allocation
CNP04	BM: #51 VM: #45	Resource States	Must support clear resource (e.g. instance, network) states where applicable. For example, provisioning, running, degraded, maintenance required, stopping, stopped, terminating, terminated.
CNP05	VM: #112 TBD-tag	Tagging	Support for user-defined tags/labels and cloud-init metadata on instances.
CNP06	TBD-console	Console Access	Serial console access is required (read-only sufficient, interactive preferred).Serial console output shall be logged and be available for historic queries (at least 1 month retention).
CNP07	INFO	If VMaaS present: # VMs/Node	GPU Nodes: no more than one VM per Node General Purpose CPU Nodes: More than one per node, with ability to select via memory/core count shape.
CNP08	add	Stable Identifiers	All resources (e.g. nodes, switches) must have a stable and persistent ID that does not change during the lifespan, even when it goes offline for a service event. VMs must also have a stable identifier.
CNP09	TBD	Firmware	Between tenants, all firmware must be brought to a known good state, all firmware must be cryptographically signed and attested during boot.

CNP10	add	Remote Management	Platform management solutions (e.g., BMC) must support Redfish over TLS (Disable IPMI).

Boot Process and Disks

Req ID	Test Details (Legend)	Requirement Area	Description
BOOT01	#38, #39, #4 , #43, #44, #58, add-k8s	Image Deployment & Updates	API-driven workflow allowing DGXC to deploy, update, and manage vendor-provided or custom disk images via bare-metal, VM, or k8s node pool provisioning.
BOOT02	BM: #107 VM: #113	Access to Instance Metadata from Guest OS	Support for cloud-init and instance metadata discovery via link-local addresses or virtual devices.
BOOT03	#104, #40	Custom Disk Images	Support for tenant created custom OS images (either of: raw, qcow2, etc). API calls: get, list, create, delete. Images should be accessible across all tenant projects/clusters/environments.
BOOT04	TBD	Node Local Storage	GPU and CPU nodes support access to node local storage (NVMe / SSD) for use as scratch storage or for caching services

SDN and Virtual Networking

This section covers the virtual networking requirements. Physical transport and network are discussed later in the document.

Req ID	Test Details (Legend)	Requirement Area	Description
SDN01	#11, #12, #13, #14: #115 #116	Virtual Networking	Full API/CLI lifecycle management (Create, Read, Update, Delete, List) for software-defined private networks. Must support non-conflicting **BYOIP** (including **7.0.0.0/8**) and stable private IP allocations. Applicable to all types of resource nodes (CPU, GPU, Storage, etc).
SDN02	TBD	Security Groups	Support for VPC-style security groups (or equivalent), including IP/CIDR-based allow and deny rules. Must define scope/application at workload, node, service (e.g. K8s API Service) and subnet/tenant levels.
SDN03	add	Security Operations	Full API/CLI capabilities to Create, Read, Update, and Delete security groups, including defined audit processes
SDN04	#16, #17, #18	Tenant Isolation	Hard logical or physical network segmentation for out-of-band management (BMC), user traffic, and storage-specific operations.
SDN05	#117	Floating/Movable IP	Ability to automatically or API-driven switch a floating private IP between nodes via API within <10 seconds without requiring an instance reboot.
SDN06	#118	Localized DNS	Support for tenant-defined localized DNS configuration to enable internal domain resolution to private endpoints (e.g. storage endpoints)

SDN07	#119	VPC Peering	Support for cross-virtual-network connectivity with full bandwidth and no "hairpin" routing.
SDN08	INFO	Storage Mesh Connectivity	The virtual network (from SDN01) must provide unrestricted L3 routing between all storage hosts, enabling full-mesh, all-to-all communication across different subnet (w/o going thru a gateway)
SDN09	INFO	Observability	The platform shall provide comprehensive logging for network infrastructure, including hardware faults, latency/performance fluctuations, and a detailed audit trail of all configuration changes to network filtering rules.

Kubernetes As a Service (KaaS) Requirements

Kubernetes Conformance, Versioning, & Compliance

Req ID	Test Details (Legend)	Requirement Area	Description
K8S01	TBD	Certified Versions	Certified Upstream Versions: Official CNCF-certified versions only; no proprietary forks.
K8S02	INFO	Version Updates	Support the three most recent minor releases (in the maintenance window); new minor versions must be available within 4-6 weeks of the upstream release; automated control plane security patching.
K8S03	INFO	EOL Policy	Defined notification periods for version deprecation.
K8S04	INFO	Kubernetes Security Response	Must participate in the Kubernetes Security Response Committee (SRC) process. Must be able to: 1. Responsibly disclose any discovered vulnerabilities to the Kubernetes SRC 2. Receive and respond to embargo notifications from the SRC 3. Patch disclosed vulnerabilities in the managed service during embargo prior to public disclosure and in compliance with direction provided from the Kubernetes SRC ensuring that the patching process does not violate embargo or SRC guidance.

Kubernetes Operational Excellence

Req ID	Test Details (Legend)	Requirement Area	Description
K8S05	#2, #5, #6	Lifecycle Management - Control Plane	API/CLI for CRUD provisioning; <30 min control plane bring-up. Strong preference for terraform provider.

K8S06	add	Lifecycle Management - Node Pool	API/CLI for CRUD provisioning (e.g., create node pool, update node pool, delete node pool, scale a node pool to a target count). Strong preference for terraform provider. ○ Must be able to specify node type (specific CPU or GPU instance type) Ability to specify default node labels and node taints within a node pool when a node joins the cluster.
K8S07		API Server Metrics	Share API Server metrics in a Prometheus scrapable format to allow NVIDIA to measure API Server SLO
K8S8	INFO	Versioning	Provider-managed control plane upgrade processes.
K8S9	INFO	Zero-Downtime Upgrades	Minor version control plane updates without app downtime or maintenance windows.
K8S10	TBD	Node Upgrades	User-initiated rolling updates respecting disruption budgets.
K8S11	INFO	HA Control Plane	Redundant architecture with etcd separation.
K8S12	INFO	Backup & Disaster Recovery	Supported recovery within defined RPO/RTO; needs to be auditable & testable
K8S13	INFO	Kubernetes Security Response	Participate in Kubernetes security response & disclosure process; provide CVE patches prior to public disclosure (https://github.com/kubernetes/committee-security-response)

Robust K8s Security

Req ID	Test Details (Legend)	Requirement Area	Description
K8S14	TBD	Control Plane Isolation	Per-tenant k8s control plane nodes must be separate from worker nodes and outside of the tenant cluster/VPC.
K8S15	TBD	Access Controls	Cluster endpoint must provide network access controls.
K8S16	TBD	IAM Integration	Kubernetes Service Accounts shall integrate with the platform IAM system to enable workloads to assume platform-managed identities and roles with appropriate scopes.
K8S17	TBD	Service Accounts	Kubernetes shall support standard Service Accounts and projected tokens as the workload identity mechanism, including an OIDC issuer for federation.
K8S18		Public OIDC Endpoint	The cluster must be able to support OIDC-based workload identity via a cluster-specific OIDC Issuer endpoint
K8S19	INFO	Encryption	At-rest encryption for etcd and secrets
K8S20	add	Logging	Ability to view or export Kubernetes control plane logs (apiserver, kcm).

Kubernetes Component and Extension Requirements

Req ID	Test Details (Legend)	Requirement Area	Description
K8S21	add	API Extensions	Mandatory support for CRDs and Validating/Mutating Admission Controllers.
K8S22	add	CNI	Standard compliance; supports Network Policies; IPv4/IPv6 dual-stack desired. Preference for Calico.
K8S23	add	CSI	• NCP provides CSI Driver installable by NVIDIA (Helm or Kustomize) for Block, shared FS, and NFS • Support for static and dynamic provisioning, snapshots, and resizing via PVs and PVCs. • CSI credentials are tenant cluster scoped (no cross cluster) • APIs to query storage usage against overall cluster quota with per PVC/Volume usage to manage utilization across PVCs • Vendor provided storage kernel modules and tools provided via (1) installed by CSI driver, (2) pre-installed in NCP provided machine image or (3) installable packages provided
K8S24	TBD	DRA	Enabled Dynamic Resource Allocation (DRA) regardless of upstream feature status (Beta/GA). Some DRA features require enabling feature gates for the control plane, in case our customers want to run AI workload with new DRA features
K8S25	add	Operator Support	Support standard operator-based management of hardware accelerators and associated drivers. Provider-default accelerator operators and drivers shall be replaceable or overridable to allow installation of tenant-required operator and driver versions (e.g., GPU Operator, Network Operator).

Kubernetes Functionality

Req ID	Test Details (Legend)	Requirement Area	Description
K8S26	add	Clusters	Support multiple clusters in the same tenancy; support multiple clusters in the same VPC.
K8S27	add	Managed CP (CP Pinning)	Pin Control Plane instances to handle a particular load-limit
K8S28		Performance	Meet the standard Kubernetes performance test certified up to 5000 nodes (or to the maximum size of the cluster, whichever is smaller). Managed Kubernetes Control Plane SLO and Performance meets or better than the Kubernetes standards results.

Security and Identity Management

Identity & Access Management (IAM)

Req ID	Test Details (Legend)	Requirement Area	Description
SEC01	add	Authentication	**Users:** Support standards-based user authentication via OIDC for platform and tenant-facing services, and validate OIDC-issued tokens including signature, issuer, audience, expiration, and required claims for identity and authorization decisions.
SEC02	add	Authentication	**In-Cluster Workloads/Nodes**: Support authenticated in-cluster identities for workloads and nodes, using short-lived credentials or tokens
SEC03	add	Authentication	**External Services**: Support authentication of out of cluster service accounts for service-to-service access. Must support credential-based access, including long-lived credentials where required.
SEC04	add | INFO	Authorization (RBAC)	The platform shall enforce least-privilege RBAC for all managed services and infrastructure, featuring granular API actions (e.g. CRUD), scopes (e.g. dev vs staging vs prod), and function (e.g. image builder, provisioner, auditor).
SEC05	INFO	Identity / Directory Services	The platform shall integrate with an LDAP (RFC2307bis) directory service such that users identities and group membership can be resolved by dependent services for authentication and authorization decisions (e.g. storage - POSIX-based access control)
SEC06	INFO	Workload/Service Identity	Support standard workload, service, and node security identities, including OIDC-based identity federation and Kubernetes Service Accounts where applicable.
SEC07	INFO	Admin Interfaces	All administrative interfaces—whether UI, CLI, or API—must be protected by Multi-Factor Authentication (e.g. kubectl)
SEC08	add	Audit Logs	Audit logs must be generated and retained for all security-relevant events, including management and control plane API calls, authentication events, and authorization decisions. Audit logs shall be retained for a minimum of 30 days and accessible to authorized platform operators.

Cryptography and Key Management

Req ID	Test Details (Legend)	Requirement Area	Description

SEC09	add | INFO	Key & Certificate Lifecycle	The platform shall support secure issuance, distribution, storage, rotation, and revocation of cryptographic keys and certificates used across platform services. It shall support automated rotation of provider-managed and customer-managed keys and certificates, with configurable rotation intervals.
SEC10	add	Key Usage	The platform shall support use of managed keys and certificates across platform services for encryption, authentication, and signing.

Network Isolation & Encryption

Req ID	Test Details (Legend)	Requirement Area	Description
SEC11	add	Tenancy Model	Hard physical or logical isolation for network, data, and compute. Separation of control planes and tenants is mandatory. This includes separation of storage resources.
SEC12	add	BMC Security	Out-of-band management (BMC) must be on a dedicated, restricted network (physically separate or VLAN/VRF-isolated). Direct access from the public internet or general corporate networks must be blocked, and only accessed via a hardened bastion (jumphost) server.
SEC13	add	Network Traffic Encryption	Encryption and mutual authentication (mTLS or equivalent) for all east-west and north-south network traffic

Edge Network Security

Req ID	Test Details (Legend)	Requirement Area	Description
SEC14	INFO	Private Access	No public internet access by default; all API endpoints (e.g. K8s API Server) must be restricted via firewall/private link.
SEC15	INFO	Edge Network Security Policy	All traffic must be filtered via Security Groups and/or user customizable ACLs using 5-tuple rules.
SEC16	INFO	Enforcement	NCP must specify the enforcement technology (e.g., Hardware firewalls, SDN, DPUs/SmartNICs) and its specific placement in the packet path.
SEC17	INFO	Threat Intelligence & Scale	Ability to subscribe to GeoIP threat & Embargo feeds and import them into security groups. NCP should share the max supported records/rules.
SEC18	INFO	MACSec protection links:	Protect links between NCP Data Center and NVIDIA POP & Object store.

Hardware Security & Compliance

Req ID	Test Details (Legend)	Requirement Area	Description

SEC19	INFO	SOC 2	SOC2 type 1 or better is required covering Security, Availability, and Confidentiality across all services and DC infrastructure
SEC20	INFO	At-Rest Data Protection	Mandatory encryption of all data at rest (e.g. local NVMe/SSD, network-attached storage) via Self-Encrypted Drives (SED).
SEC21	add	Data Sanitization	Data sanitization must be performed between tenants or on a hardware replacement, including cryptographic erase of all data drives between tenants; sanitization/wipe of any persistent or volatile memory including SRAM/GPU memory; resetting of TPM and BIOS
SEC22	INFO	Root of Trust + Secure Boot.	Mandatory support across all platforms for Hardware Root of Trust mechanisms (TPM 2.0). The platform must enable UEFI OS Secure Boot w/ TPM 2.0.

Breakfix Requirements

The NCP must provide a specific "Breakfix API" to support fleet reliability. Any node-level remediation must not impact other parts of the tenancy; specifically, NVLink must be re-configured properly to take a node out of the tenancy.

The API must enable the following actions:

Req ID	Test Details (Legend)	Requirement Area	Description
BFX01	add	Breakfix Lifecycle	**Compute**: Power-cycle individual nodes or reset a VM instance. GPU: Reset GPUs on an individual node (as needed - k8s) **Maintenance**: Return/Report an individual node and a rack to the Provider for maintenance **Cordon**: Mark a node as unschedulable for new workloads (but finish existing) **Replace**: Request a host-replacement when health thresholds are breached

BFX02		Breakfix Events	• Query for any upcoming/current maintenance events for a node or rack • Query for any retirement notices for a node/rack. • Query for historical / status information for equipment repair. • Event information should include: • ticket open date • ticket update date • ticket close date • Hardware Stable Identifier (e.g., node ID) • Hardware category/type impacted (e.g., GPU, fan, interconnect) • Maintenance/Error/fault description (some short description of the issue) • Action: Categorization of action (e.g. repairs done on faulty GPUs to resolve the fault) • Provider Account ID • ticket ID • Node Handover Date (Date when the node was deployed in Production)
BFX03		Diagnostics	• Identify serial numbers of installed hardware (chassis, baseboard, network adapters, CPU, GPU, etc). Obfuscated but stable identifiers are also OK. • Inspect firmware versions of compute nodes and NV switch trays.

Telemetry Requirements

The telemetry requirements are comprised of two core components that require alignment between DGX Cloud and the NCP:

1. **Delivery Method:** *How* telemetry will be delivered by NCP to DGX Cloud for ingestion
2. **Telemetry Scope:** *What* telemetry the NCP will deliver to DGX Cloud

Delivery Method

NCP shall deliver all required telemetry, including metrics and logs, in a manner that allows for ingestion into DGX Cloud systems. The preferred methodology is natively via the OpenTelemetry Protocol with a latency of no longer than 120 seconds.

Telemetry Scope

DGX Cloud will provide the NCP with a detailed specification document with the required metrics and logs. Upon receipt, the NCP shall be required to provide a formal written response detailing the following:

- Confirmation of its ability to deliver the specified metrics and logs.
- Projected timelines for delivery.
- Specific technical details, including metric names, label names, and label values.

Network Telemetry

The NCP shall provide network telemetry across the following domains:

- North-South (Front-End) Network (client-facing and external interconnects)
- East-West (Back-end) Network (GPU/GPU interconnects)
- Management Network (control plane and orchestration traffic)
- NVSwitch Fabric (intra-node GPU switching, applicable for only GB200 and beyond clusters)
- Host Network (NIC-level and server connectivity)

Logs

DGX Cloud will require the NCP to provide logs from various network technologies, including but not limited to:

1. Fabric Manager logs for the NVLink domain *(where applicable)*
2. Subnet Manager logs for the NVLink domain *(where applicable)*
3. VPC Flow logs (all ingress/egress traffic)
4. UFM Event logs
5. General Switch Logs
6. Switch syslogs
7. Switch kernel logs
8. BMC SEL logs
9. syslogs

Storage Requirements

NCP must provide shared storage solutions (where applicable) that are manageable via standard APIs and UI, including auditing rights for NVIDIA access.

Home Directory Storage

- **Quota Feature:** Configurable filesystem-wide limit, default user/gid quota settings, and per uid/gid overrides.
- **Accounting:** Usage accounting for uid/gids must be available when the feature is enabled.

Req ID	Test Details (Legend)	Requirement Area	Description
DIR01	INFO	File Service uid/gid Quota feature	Configurable filesystem-wide limit, default user/gid quota settings, and per uid/gid overrides available. Usage accounting for uid/gids when the feature is enabled.
DIR02	INFO	Must be NFS storage	• NVIDIA requires NFSv4 protocol shared storage to work • Access control based on DLs requires POSIX

High-Speed Storage Service Requirements

Req ID	Test Details (Legend)	Requirement Area	Description
HSS01	add	Provisioning APIs	Storage provisioning may be via vendor portal/API or NCP portal/API.
HSS02	INFO	Performance	Must provision needed throughput requested for minimum bandwidth and IOPS.
HSS03	INFO	Integration	K8s: CSI support Breakfix API required to report storage issues
HSS04	INFO	Quota Support	Ability for quota limits to be set on specific user workloads / volumes
HSS05	INFO	Upgrade, Maintenance	Provider / NCP initiates desired maintenance. NVIDIA can schedule actual maintenance and can defer maintenance up to 2 weeks. Upgrades should be non-disruptive.
HSS06	INFO	RDMA Memory Protection	Storage systems using RDMA must enforce memory protection via authorization keys for both local and remote access

High-Speed Storage Filesystem Requirements

Req ID	Test Details (Legend)	Requirement Area	Description

HSS07	INFO	Parallel High Speed Filesystem	Parallel or multi-path high-speed filesystem that supports scaling to thousands of simultaneous clients while sustaining requested performance.
HSS08	INFO	Single File System Size	• It must be possible to allocate a file system of at least 1 PiB even if the initial request is less. Growing to > 10PiB as cluster size increases. • • This hard requirement may be higher for a specific site and if so will be communicated via the ancillary services document.
HSS09	INFO	Multiple Filesystems (Fungible Total Capacity)	Can have >1 filesystem within our total capacity. Minimal file system size <= 50 TiB.
HSS10	INFO	Filesystem Expansion	Live file system expansion is supported, in terms of capacity, inodes, IO performance, and metadata operations performance. Performance should scale linearly with capacity.
HSS11	INFO	Client	Ability to describe your client: In-Kernel, userspace, or bare-metal client installation requirements. Support integration with client kernels / OS used by NVIDIA, as needed. DKMS-enabled packages available for Ubuntu 20.04, 22.04, and 24.04-based operating systems. ARM64 versions compatible with GB200-ready kernels are mandatory, e.g. Linux 6.8.x. Managed Storage Service Provider will provide client configuration best practices and configuration guidelines for filesystem options and kernel module configuration to reliably achieve optimal performance on ARM and x86_64-based clients.
HSS12	INFO	Quota (User, Project & Group)	Must support soft and hard quotas - uid / gid / project(directory)-id quotas with enforcement.
HSS13	INFO	Root-squash	Nvidia needs to be able to enable or disable and manage root-squash at any time.
HSS14	INFO	flock	It must be possible to mount the file system with flock.
HSS15	INFO	Ability to Audit Changes	Enable Nvidia to have access to changelog data for filesystem auditing and detailed user operations tracking. Tracking by uid/gid, create files, create dirs, rename files, rename dirs, delete files, delete dirs
HSS16	INFO	HA	All services are required to tolerate any critical component failure in the backend and provide continued client access to all storage services in such cases.
HSS17	INFO	Multi-Node Coherency	One second or less for client attribute and dentry cache updates/invalidates
HSS18	INFO	Client Multipathing	Clients must have multipathing to all storage servers

Data Movement Systems Requirements

The Data Movement system is used to copy data from an external data source (NVIDIA, other Cloud, etc) to the NCP data center.

Req ID	Test Details (Legend)	Requirement Area	Description
DMS01	add	Dedicated K8s Cluster	Provider managed k8s cluster (or ability to stand up our own) for Data Mover stack available ahead of the GPU cluster bringup to pre-stage data
DMS02	INFO	Data Mover Nodes (CPU)	Dedicated CPU nodes for running data mover - needs high performance networking (exact quantity will be communicated via ancillary services doc)
DMS03	INFO	Access to Same GPU Storage	Same filesystem as mounted on GPU nodes mounted on the Data Mover nodes (or ability to mount the same filesystem via CSI)
DMS04	INFO	Access to NVIDIA Corp Net	Dedicate link (as described in network transport) to NVIDIA corp net, preferably with vpn, but otherwise with stable IP for allowlisting
DMS05	add	Stable Egress IP	Stable IP to IP allowlist access to Nvidia services. (e.g. similar to NAT Gateway)

DGXC-Managed Storage System Deployment

For scenarios where the storage system software will be deployed and managed by DGXC rather than the NCP, the following requirements apply. These requirements enable DGXC to operate storage systems (such as high-speed parallel filesystems, capacity object storage, or block storage) using NCP-provided infrastructure while maintaining operational control.

Host Provisioning and Lifecycle

Req ID	Test Details (Legend)	Requirement Area	Description
STG01	INFO	Operating System Support	NCP must support a workflow that allows DGXC storage operators to integrate vendor-provided or storage-specific operating system images via bare-metal or VM provisioning for storage servers. The workflow must: (a) Allow DGXC to deploy custom OS images (e.g., vendor-enhanced kernels for Lustre, Rocky Linux, Ubuntu 20.04/22.04/24.04).
STG02	add	Drive Sanitization Policy	Cryptographically erase data drive contents between storage system tenants with full attestation of host firmware. Must support an optional flag to skip drive sanitization during break/fix flows (e.g., power supply replacement) where tenancy does not change. Critical hardware component replacements may require sanitization without override, this is inclusive of GPU / CPU node local storage

STG03	INFO	Stable IP Assignment	Storage nodes must support static IP addressing that remains stable during host lifecycle operations and does not reset between maintenance events.
STG04	INFO	Out-of-Band Failure Detection	NCP must provide the ability to detect system failures out-of-band, including device, network, memory, and drive failures, enabling DGXC to proactively respond to hardware issues.
STG05	INFO	Topology Observability	NCP must provide visibility into failure domains to enable DGXC to provision storage nodes with physical diversity. Storage systems must be able to provision nodes that purposefully span failure domains for resilience.
STG06	INFO	BlueField/DPU Support	For storage systems utilizing BlueField-based architectures, the host provisioning system must support lifecycle management and specific configuration requirements for BlueField "JBOF" systems that export NVMe-oF to hosts.

Network Transport and Fabric Visibility

Backend Switch Fabric API

The purpose of this API is to expose sufficient information about the cluster's network topology to enable efficient scheduling, placement, and optimization of multi-node GPU workloads. Understanding the network hierarchy between compute instances and switches, as well as both intra- and inter-node NVLink domains, is essential for minimizing communication latency and maximizing throughput. Thus, this applies to North-South, East-West, and NVLink networks (not MGMT). See the appendix for a DGXC recommended reference implementation.

Req ID	Test Details (Legend)	Requirement Area	Description
NET01	INFO	Backend Switch Fabric API	For each compute node, the API must provide visibility into the backend network switches connecting the node to the core. • **Identification:** Each switch must be identified by a unique, stable identifier. A "switch" may represent a physical switch or a logical connectivity domain. • **Structure:** API may be gRPC or REST. Response structure may include multiple nodes (pagination expected). • **Topology:** Switch info can be returned as an ordered array of IDs (e.g., leaf, spine, core) or separate fields for each tier

NET02	INFO	NVLink Domain API	**Requirement:** For compute nodes supporting NVLink (e.g., GB200, GB300, Vera Rubin), the API shall return the unique identifier of the NVLink domain associated with each node. **Implementation:** Can be a separate API method or part of the Backend Switch Fabric API.

Transport and Networking requirements

Non-Conflicting IP Space Allocation for the DGXC Cluster

Purpose:
Ensure DGXC GPU clusters deployed in NCP can access the NVIDIA DGXC/CorpIT network directly via routing exchange. DGXC Cluster IP address must be non-conflicting with existing NVIDIA private IP space.

Req ID	Test Details (Legend)	Requirement Area	Description
NET03	add	Non-Conflicting IP Space Allocation for the DGXC Cluster	• **Bring Your Own IP (BYOIP):** NCP shall support the ability for NVIDIA to bring and allocate its own IP private address space for DGXC GPU clusters. • **Stable IP**: NCP shall provide a possibility to create static IP allocations that persist across instance restarts and re-creations. That includes floating IP allocations. • **DoD space:** NCP shall support allocation and use of the 7.0.0.0/8 IPv4 address space for DGXC GPU cluster deployments. This IP space shall be considered equivalent to RFC1918 addresses • **Routing Support:** NCP must support advertising and routing of BYOIP prefixes within the NCP environment and across interconnects (Private Cloud Interconnect, IPSec, etc.)

Connection to NVIDIA CorpIT Network

Purpose:
Provide connection from DGXC GPU clusters within NCP to NVIDIA CorpIT for internal Command & Control and admin access.

Req ID	Test Details (Legend)	Requirement Area	Description

NET04	INFO	Connection to NVIDIA CorpIT Network	**Bandwidth**: Low bandwidth (Up to 10Gbps). **Transport**: Private Cloud interconnect + VIF + BGP (preferred for better performance/security). DGXC will establish connectivity to NCP through a mutually agreed Point of Presence (POP) using Private Cloud Interconnect, functionally equivalent to AWS Direct Connect, GCP Dedicated Interconnect, Azure ExpressRoute, and OCI FastConnect. Connectivity will be provisioned with a Virtual Interface (VIF) and routing established via BGP. The interconnect will be used to exchange private IP space (RFC1918, as well as 7.0.0.0/8) between DGXC and NCP.

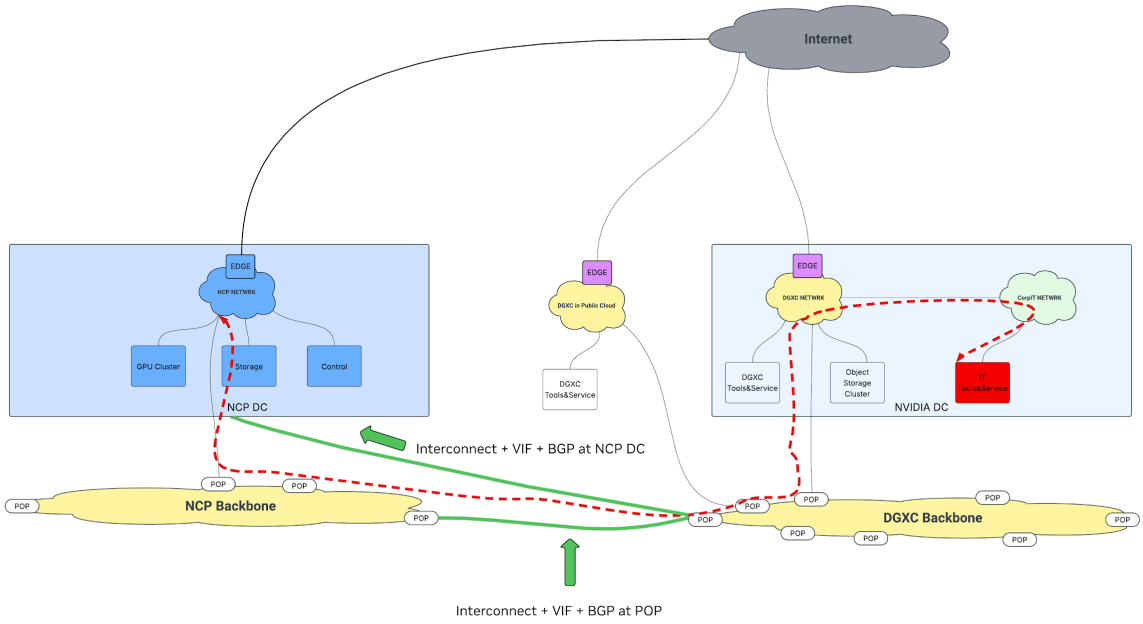


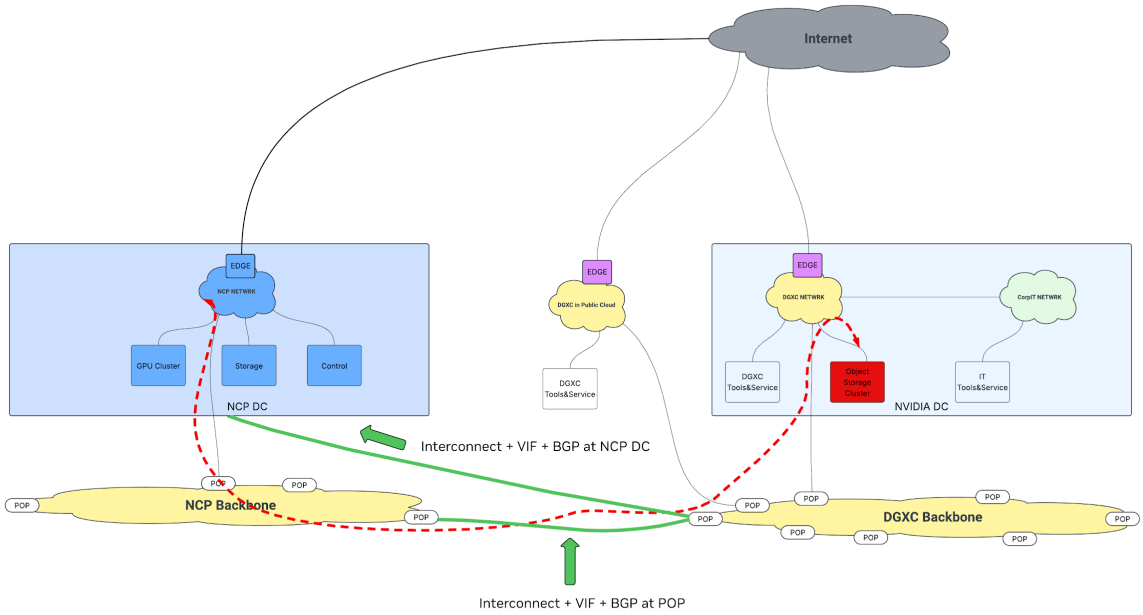
Figure: Private Cloud interconnect + VIF + BGP for CorpIT access

Connection to DGXC Storage

Purpose:
Enable high-bandwidth, end to end MACsec-encrypted (fail-closed) access between the DGXC GPU clusters within NCP and NVIDIA DGXC on-premises object storage for large-scale data movement.

Req ID	Test Details (Legend)	Requirement Area	Description
NET05	INFO	Connection to DGXC Storage	**Transport**: Private Cloud interconnect + VIF + BGP (preferred for better performance/security). DGXC will establish connectivity to NCP through a mutually agreed Point of Presence (POP) using Private Cloud Interconnect, functionally

			equivalent to AWS Direct Connect, GCP Dedicated Interconnect, Azure ExpressRoute, and OCI FastConnect. Connectivity will be provisioned with a Virtual Interface (VIF) and routing established via BGP. The interconnect will be used to exchange private IP space (RFC1918, as well as 7.0.0.0/8) between DGXC and NCP.



Cluster Local Internet Access

Purpose:
Provide general Internet access from DGXC GPU clusters within NCP to Internet, including NVIDIA DGXC hosted services on third-party public cloud services.

Req ID	Test Details (Legend)	Requirement Area	Description
NET06	INFO	Cluster Local Internet Access	**Cluster Internet access:** Egress NAT IPs should be a static pool dedicated to only Nvidia Cluster/Tenancy/VPC. These persistent IP addresses must be used exclusively for DGXC traffic and shall not be shared with or carry traffic from other NCP tenants. **Availability:** Must support redundant upstream paths to ensure connectivity under failure.

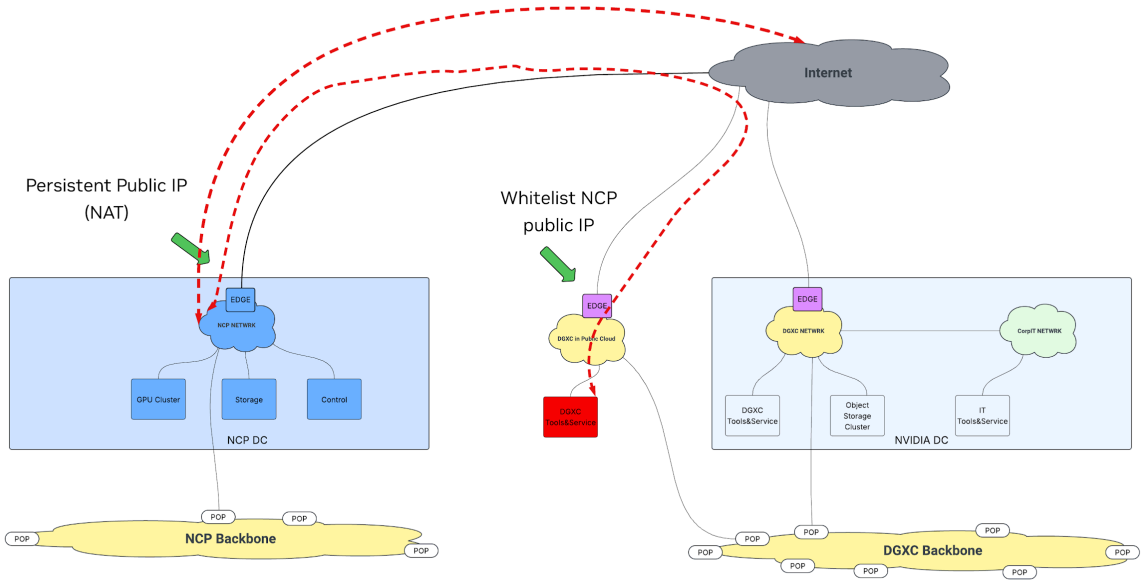


Figure: Public internet for DGXC hosted Services access

Capacity and Fleet Management

This section defines the essential metrics required for standardized monitoring and reporting of fleet health in partner engagements to support operations and contractual SLAs.

Req ID	Test Details (Legend)	Requirement Area	Description
CAP01	INFO	Governance metrics	**Required Governance Metrics** The core metrics needed to track fleet health are: **Delivered**: Nodes/GPUs provisioned and available to NVIDIA, allocated to a specific account/project/tenant. **Healthy**: Nodes/GPUs functioning and meeting SLA requirements, allocated to a specific account/project/tenant. **Reserved**: Resources allocated to a specific account/project/tenant. Total Active/In-Use: Nodes/GPUs currently in use within a specific account/project/tenant.
CAP02	add	Resource Governance API Metrics	The Resource Governance API must return the following information for each node: **Node ID** (Unique identifier for a GPU node) **Health State** (Healthy/Unhealthy classification) **Instance ID** (Identifier for virtual workload) **Creation Timestamp** (Time workload/node was created) **Hardware Type** (Descriptor for the hardware model) **GPU Count** (Number of GPUs per node)

			Top-levelAccount/ID (Identifier for the top-level organization/account) **Sub-LevelProject/ID** (Identifier for the nested project/sub-account) **In Use** (True/False status indicating if the GPU Node is turned on and in use) **Region** (Region of the data center where nodes are deployed)
CAP03	add	Resource Discovery APIs	It is not acceptable to have capacity be “handed” to DGXC through a phone, slack or email message. For example, when cluster first comes online, nodes/racks are likely being handed off weekly (or more frequently). Instead, please provide the following mechanism (and we can poll): **Programmatic Capacity Discovery**: All newly delivered capacity must be discoverable via a centralized API. This "Resource Index" must provide a resource stable identifier and some information on why it’s being provided (e.g. capacity fulfillment on gb300 project, break-fix / RMA return to cluster, etc)
CAP04	INFO	Logical Compartmentalization & Resource Isolation	To ensure performance consistency and security, the NCP must support strict logical and physical isolation of NVIDIA’s reserved capacity. **Capacity Reservations**: A mechanism to logically group and "pin" a set of resources (compute, network, storage) to accounts (or equivalent constructs) in an NVIDIA tenancy **Atomic Allocation: Support for reserving a "topology block" as a single unit, ensuring all resources in that block share identical performance characteristics and security boundaries.**
CAP05	INFO	Unified Health & Lifecycle APIs	NVIDIA requires a "single source of truth" for the health of both physical hosts and logical compute primitives. **Per-Host Health:** Real-time API access to the health bits of physical hardware (GPU state, thermal status, memory health). **Primitive-Level Status:** Health aggregation at the cluster, nodegroup, or reservation level to identify broad infrastructure failures (e.g., a spine switch failure affecting a whole block).

Appendix

This section contains links to reference documents and implementation guides to provide additional details if NCPs need them.

Test Legend

As a part of the AI Cloud-Ready Initiative, a validation suite is provided on GitHub, containing a set of tests that can be run to evaluate an environment against the provided requirements and specifications.

The Test Details column in the following sections contains four options:

- Github Test reference
TBD - item on the github roadmap
add - item to be added into the roadmap
INFO - a requirement which is either not testable or currently won't be tested from the test repo.

Implementation Guidance

Reference documents provide additional information on implementing some of the above requirements.

1. Network Topology Discovery: https://github.com/NVIDIA/topograph
2. Exemplar Cloud Website: https://www.nvidia.com/en-us/data-center/ai-cloud-performance/
3. Kubernetes Security guidance : https://github.com/kubernetes/committee-security-response

Other Feature Considerations (Not Required)

1. Disk Cloning: Disk cloning capability (for network-attached block devices). It should be possible to clone a disk even on a running instance.