



GRIDONE

I N F I N I T E C O M P U T E

GPU Cloud Infrastructure for the UAE's AI Decade

Company & Deployment Whitepaper

GRIDONE · DUBAI, UNITED ARAB EMIRATES

gridone.ae · support@gridone.ae

VERSION 1.1 · AUGUST 2026

Notice

This whitepaper is prepared by GridOne, Dubai, United Arab Emirates, and describes the company's business focus, target market, technology platform, and phased deployment plan. It is provided for business information purposes to prospective customers, partners, and suppliers.

The deployment phases described in this document represent GridOne's current plan and are indicative; they are subject to hardware availability, facility readiness, and customer demand. Hardware specifications describe GridOne's selected platform as confirmed with its supplier. Statements about third parties and public initiatives are drawn from the public sources listed in the References section.

Company	GridOne — “Infinite Compute”
Registered domain	gridone.ae
Location	Dubai, United Arab Emirates
Chief Executive	Abdulla Alfalasi
Contact	support@gridone.ae
Document	Version 1.1 · August 2026

CONTENTS

Table of contents

01	Executive summary	3
02	The problem: compute that can't leave the country	4
03	Market context — the UAE's AI decade	4
04	Target market: priority verticals and the wider buyer set	5
05	The GridOne platform & how it works	7
06	Deployment roadmap	10
07	Risk management & sustainability	11
08	Governance & compliance	12
09	About GridOne · closing · references	13

Executive summary

Who this document is for: institutional decision-makers in UAE education and healthcare, technology partners, and hardware suppliers evaluating GridOne as a deployment partner.

GridOne is a Dubai-based GPU cloud company building dedicated, in-country AI compute for the United Arab Emirates. The platform is standardised on NVIDIA HGX B300 systems — KERNO UQLEAD K03-8220 Ultra NVIDIA-Certified Systems, each with eight B300 SXM6 GPUs — to be deployed in UAE facilities and delivered as managed GPU capacity to organisations that need serious compute without building their own data-centre footprint. The first phase is in deployment; this document describes the platform and the approach, not an operating fleet.

GridOne's two priority verticals at launch are education and healthcare — two sectors named as priorities in the UAE National Strategy for Artificial Intelligence 2031, and two sectors where data residency requirements make locally hosted compute a practical necessity rather than a preference. They are where the go-to-market effort is concentrated, not the limit of who the platform serves: the same standardised fleet is sold to regional model builders and research groups, enterprise AI teams, government and regulated buyers, inference platforms and startups, and operators who need a UAE footprint without building one.

Deployment is phased and deliberately conservative. Capacity is added in whole servers, and each block is deployed, burned in and brought to production utilisation before the next is ordered. Expansion is gated on committed demand rather than run ahead of it — which keeps utilisation high, operations disciplined, and every accelerator accounted for from delivery to decommissioning.

288 GB	1.8 TB/s	2
HBM3e memory per B300 GPU	NVLink 5 bandwidth per GPU	priority verticals at launch: education & healthcare

What GridOne offers is narrow on purpose: one standardised NVIDIA HGX B300 server platform, operated in the UAE by a Dubai team, sold as reserved capacity on terms that run from a single month to twenty-four, so that an institutional budget can meet it either way. Customers get dedicated allocations rather than contention for shared pools, their data stays inside the country, and the people who run the hardware are reachable in the same timezone. The sections that follow set out the market position, the platform, the deployment approach and the governance that surrounds it.

02 • THE PROBLEM

Compute that can't leave the country

Every serious AI initiative in the UAE runs into the same three walls:

- Regulated data cannot travel. Health information must, by law, stay in the UAE, and much education and government data is bound by similar policy. Overseas cloud regions are legally out of bounds for exactly the workloads that matter most.
- Global GPU capacity is scarce and distant. Flagship accelerators are allocated to hyperscale buyers on long lead times; an individual university or hospital queues behind the world.
- Owning infrastructure is impractical. High-density GPU racks demand specialised power, cooling and around-the-clock operations that teaching and clinical institutions cannot justify building for themselves.

The result: the organisations the national AI strategy most depends on have the least practical access to compute. GridOne exists to remove that bottleneck — in-country capacity, institutional-scale procurement, operated by a local team.

03 • MARKET CONTEXT

The UAE's AI decade

The UAE was the first country in the world to appoint a Minister of State for Artificial Intelligence (2017), and its National Strategy for Artificial Intelligence 2031 commits the country to becoming a global AI leader by 2031, with education, healthcare and government services among its priority sectors [1][2].

National investment in AI infrastructure has accelerated sharply. Stargate UAE — the multi-gigawatt AI campus announced in Abu Dhabi by G42, OpenAI and partners including NVIDIA — is reported at more than US\$30 billion for its initial phases, with the first 200 MW targeted for 2026 [3][4]. Frontier-scale projects of this kind anchor the national compute agenda, but they serve hyperscale training workloads. A parallel need exists beneath them: right-sized, locally operated GPU capacity that a university lab, a hospital group, or a ministry department can procure in weeks, keep inside the country, and pay for at institutional scale. That is the gap GridOne addresses.

Why in-country compute matters

- Health data must stay in the UAE. Federal Law No. 2 of 2019 on the use of ICT in healthcare requires health information to be stored and processed inside the UAE unless an exception is approved — cloud regions abroad are not an option for most patient-data workloads [5][6].
- Personal data is federally regulated. The UAE Personal Data Protection Law (Federal Decree-Law No. 45 of 2021) governs processing and cross-border transfer of personal data [6].
- Institutions want local accountability. Procurement in education, healthcare and government favours suppliers with UAE presence, UAE support hours, and UAE data handling.

The institutional options, compared

Consideration	Overseas cloud region	Self-built cluster	GridOne
Regulated-data residency	Not permitted for UAE health data	In-country	In-country
Time to capacity	Subject to global allocation queues	Extended design-and-build programme	Weeks against deployed capacity; a new block is ordered and brought to production in about five months
Procurement fit	Foreign entity, consumption billing	Major capital project	UAE supplier, reserved terms from one to twenty-four months
Operations burden	Carried by the customer's team	Full facility and platform burden	Managed by GridOne
Support	Remote, ticket-based	In-house only	Dubai-based engineers, UAE hours

04 • TARGET MARKET

Priority verticals at launch, and the wider buyer set

Education

The UAE has embedded AI across its entire education system: from the 2025/26 academic year, artificial intelligence is a curriculum subject in government schools from kindergarten to Grade 12 [7][8], and Abu Dhabi's Mohamed bin Zayed University of Artificial Intelligence (MBZUAI) is the world's first graduate research university dedicated to AI [9]. Universities, research groups and school networks now need hands-on compute for model training, fine-tuning, and applied research — at a scale and price accessible to an academic budget.

- Reserved GPU allocations for university AI/CS departments and research labs
- Burst capacity for semester projects, hackathons and thesis work
- Fine-tuning and inference environments for Arabic-language and domain models

Healthcare

UAE health authorities are investing heavily in AI — from the Department of Health Abu Dhabi's genomics and AI partnerships to AI-assisted diagnostics and digital-twin initiatives [10][11]. These workloads sit directly on regulated patient data, which under Federal Law No. 2 of 2019 must remain in-country [5]. GridOne's UAE-resident clusters let hospital groups, imaging centres and research institutes train and serve models without data ever leaving the Emirates.

- Medical imaging model training and inference on locally resident data
- Genomics and life-sciences pipelines aligned with national programmes
- Clinical NLP and administrative automation for hospital operators

The wider buyer set

Education and healthcare are where the go-to-market effort is concentrated, not a restriction on who can buy. The same standardised fleet is sold to regional model builders and research groups taking whole clusters on twelve- to twenty-four-month terms, enterprise AI teams taking whole servers on shorter ones, government and regulated buyers who need an isolated environment inside the jurisdiction, inference platforms and startups that want to start small and burst without an allocation queue, and operators reselling regional capacity who need a UAE footprint without building one. Every one of them buys the same product on the same residency and support model; what differs is the size of the commitment and its length.

Service model & technology

What GridOne delivers

- Dedicated GPU capacity — reserved allocations (by GPU, node, or cluster) with predictable monthly pricing suited to institutional procurement
- Managed environment — provisioning, scheduling, monitoring and support handled by GridOne; customers focus on their models, not on infrastructure
- UAE data residency — compute and storage inside the United Arab Emirates, aligned with health-data localisation and PDPL requirements
- Local support — Dubai-based team, UAE business hours and timezone, direct engineering contact rather than ticket queues

Hardware platform

GridOne standardises on a single server platform: the KERNO UQLEAD K03-8220 Ultra, an NVIDIA-Certified System built on the NVIDIA HGX B300 baseboard. Every server is identically configured — eight B300 SXM6 GPUs with 288 GB of HBM3e each, connected by NVLink 5 at 1.8 TB/s per GPU, on a dual-socket Intel Xeon 6 host, in an 8U air-cooled chassis with fully redundant 80+ Titanium power [12][13]. Blackwell Ultra supports the precision formats (FP4/FP8/FP16) used by modern training and inference workloads. One platform across the whole fleet means every allocation behaves the same way. Servers are deployed in tier-appropriate UAE data-centre facilities with power and cooling engineered for high-density racks.

Layer	Approach
Server	KERNO UQLEAD K03-8220 Ultra · NVIDIA-Certified System · 8U air-cooled
Accelerators	NVIDIA HGX B300 — eight B300 SXM6 GPUs per server, 288 GB HBM3e each
Interconnect	NVLink 5 within the server; InfiniBand XDR compute fabric across servers
Facilities	UAE data centres, high-density racks, redundant power & cooling
Access	Reserved bare-metal and managed cluster allocations
Data	In-country storage and processing; customer data never leaves the UAE

How customers come on board

1 · Engage	2 · Contract	3 · Allocate	4 · Run
Workload review with GridOne engineering; sizing and residency assessment	Verification and onboarding (see Governance); reservation agreement with predictable monthly pricing	Capacity provisioned from deployed servers; environments configured to the customer's stack	Production use with monitoring, UAE-hours support with emergency escalation, and quarterly capacity reviews

Commercial model

GridOne sells reserved capacity — by server, by GPU or by cluster, on terms running from one month to twenty-four — rather than anonymous spot compute. The smallest unit a customer can take is a whole server; smaller units are available inside a customer's own allocation, so that no two customers ever share one machine. Longer terms carry a lower rate, which is what makes the model work in both directions: reservations map cleanly to academic-year and clinical-programme budgeting, and they give GridOne the demand signal that gates each deployment phase.

Platform architecture

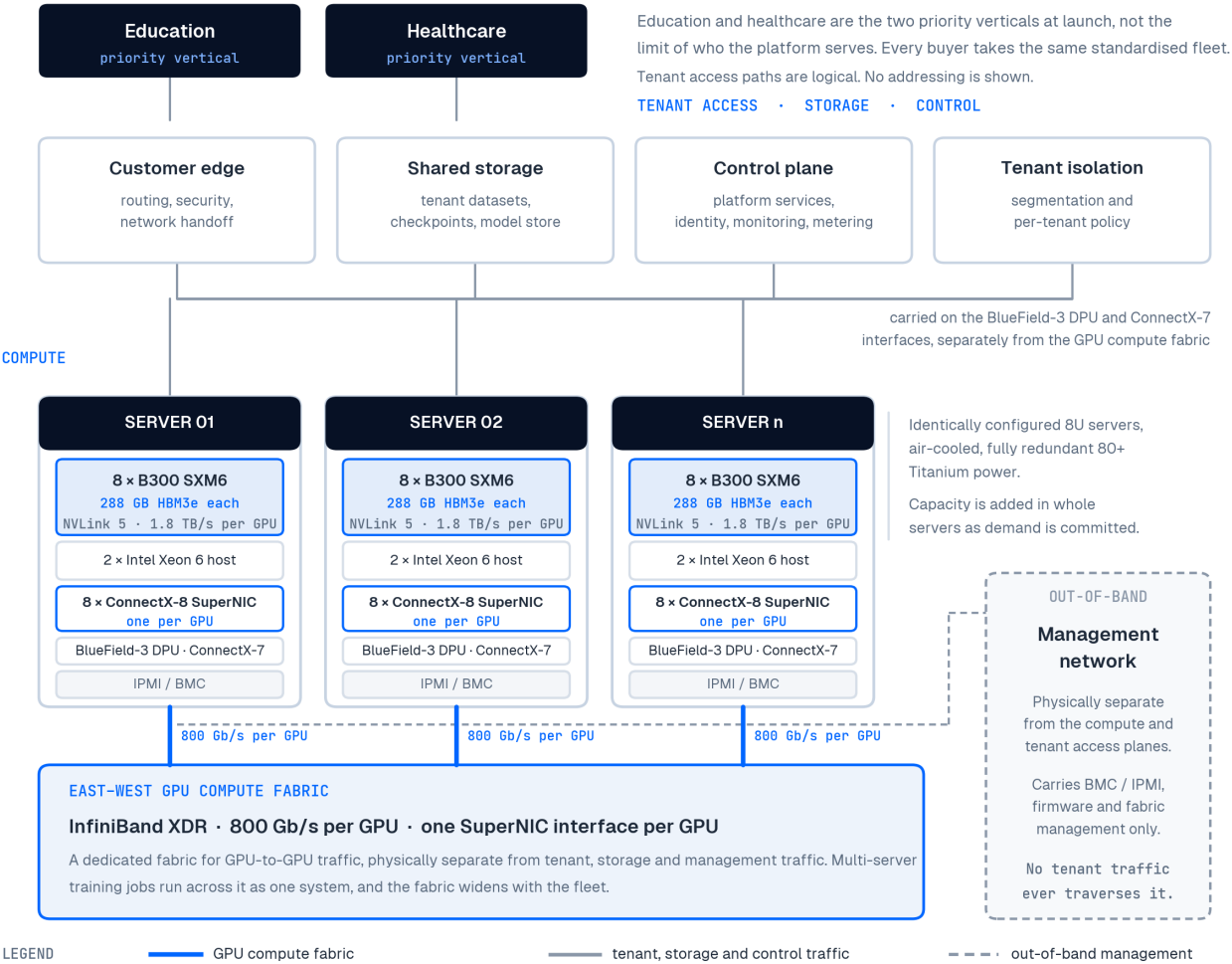
The architecture separates three kinds of traffic and keeps them apart. Inside each server, the eight B300 GPUs communicate over NVLink 5. Between servers, each GPU has its own onboard ConnectX-8 SuperNIC interface running at 800 Gb/s, feeding a dedicated East–West InfiniBand XDR fabric reserved for GPU-to-GPU traffic — this is what lets a multi-server training job behave as one system. Tenant access, shared storage and platform control traffic travel on a separate path, carried on each server's BlueField-3 DPU and ConnectX-7 interfaces. A third, physically isolated out-of-band network carries BMC and fabric management only, and never tenant traffic. Capacity is committed in whole servers, with each tenant isolated at the hardware and network layer; whole GPUs and MIG partitions are units a tenant can use to divide their own allocation, never a way of putting two tenants on one machine. Growth repeats the same building block and widens the fabric — every added server brings its own eight compute interfaces with it.

Switching, topology and port plan are set by the confirmed network design for each site and are outside the scope of this document. The diagram below is logical and carries no addressing.

PLATFORM ARCHITECTURE

KERNO UQLEAD K03-8220 Ultra • NVIDIA-Certified System • NVIDIA HGX B300

Three separated traffic planes. The minimum unit of allocation is a whole server; no server is shared between two tenants.



Logical diagram. Switching, topology and port plan are set by the confirmed network design for each site and are not represented here. No IP addressing, VLAN, subnet or partition identifiers appear anywhere in this document.

A measured, demand-led ramp

GridOne scales in step with demand rather than ahead of it. Capacity is added in whole KERN O UQLEAD servers, and every block is racked, burned in and brought to production utilisation before the next is ordered. Each phase has an explicit gate: the platform does not grow because a schedule says so, it grows because the preceding capacity is committed and running. That keeps capital efficient, keeps every unit traceable, and means a customer is never allocated onto hardware that has not been validated.

Phase	Focus	Gate before the next phase begins
Phase 1 — launch	Initial production cluster commissioned and validated end to end; first education and healthcare tenants onboarded	Platform proven in production with reference tenants running live workloads
Phase 1 — ramp	Capacity added in whole servers as reservations are committed; each block deployed, burned in and handed to production	Preceding capacity utilised and supported; reserved commitments in place for the next block
Phase 2 — scale	Demand-led expansion of the footprint; government and private-sector segments onboard on the same residency and support model	Sustained utilisation, facility power and cooling contracted ahead of hardware, operations team staffed

Operational scaling accompanies the hardware ramp: site reliability and customer engineering are hired ahead of each capacity block, and facility commitments are staged so that power and cooling are contracted before hardware arrives.

What a capacity block involves

- Facility readiness first. Rack space, power and cooling are contracted and verified at the hosting facility before hardware is shipped
- Whole servers, identically configured. Capacity is added in complete KERN O UQLEAD systems, never as partial or mismatched builds
- Burn-in and acceptance testing. Every server is stress-tested and validated — compute, memory, fabric and thermals — before any tenant workload is placed on it
- Handover to production. Assets are serialised and recorded, monitoring and support coverage are extended to the new capacity, and only then is it opened for allocation

Risk management & sustainability

A phased infrastructure build carries known risks. GridOne's plan pairs each with a specific mitigation, and the staged structure itself is the master control: no capacity block is ordered until the previous one is deployed, utilised and supported.

Risk	Mitigation
Hardware supply delays	Staged orders placed ahead of need against a single standardised server platform; roadmap gates flex without breaking customer commitments
Facility power & cooling readiness	Power and cooling contracted at the hosting facility before each hardware block arrives
Demand shortfall	Demand-led expansion: each block is gated on utilisation of the previous one and on reserved commitments
Operational capacity	Engineering and customer-operations hiring staged ahead of each capacity block
Customer concentration	Multi-segment base — education and healthcare at launch, government and private sector in expansion

Sustainability

Accelerated computing is energy-intensive, and GridOne treats efficiency as an operating principle: high-efficiency cooling appropriate to high-density racks, a utilisation-first rollout that avoids idle fleets by design, and facility selection that weighs energy performance alongside resilience. As the platform scales, efficiency metrics will be tracked and reported to customers on request.

Security, compliance & responsible deployment

GridOne is built to be a trustworthy home for advanced accelerators and for the regulated data of its customers. Governance is treated as a product feature, designed in from the first rack.

Customer & end-user assurance

- Know-your-customer verification of every tenant organisation before onboarding, including beneficial-ownership checks for private entities
- Screening of customers and counterparties against applicable international restricted-party and sanctions lists
- A binding acceptable-use policy; workloads are limited to lawful research, educational, clinical and commercial purposes

Hardware custody & trade compliance

- All accelerators are installed and operated in licensed UAE data-centre facilities under GridOne's custody, with serialised asset tracking from delivery through decommissioning
- Hardware is deployed as cloud capacity only — it is not resold, re-exported, or transferred to third parties
- GridOne conducts its procurement and operations in accordance with applicable UAE and international trade and export-control regulations, and maintains records to demonstrate it

Facility & platform security

- Physical access control, CCTV and dual-authorisation entry at hosting facilities
- Tenant isolation at the hardware and network layer; encrypted data at rest and in transit
- Logging and auditability of administrative actions across the platform

Data protection

- Processing aligned with the UAE Personal Data Protection Law (Federal Decree-Law No. 45 of 2021)
- Health workloads operated in accordance with Federal Law No. 2 of 2019 — health data remains in the UAE
- Customer data is never used by GridOne for any purpose other than delivering the service

About GridOne

GridOne was founded in Dubai in 2026 by Abdulla Alfalasi to build the Emirates' first public GPU-as-a-service platform. The company is based in Dubai and operates under the gridone.ae domain, with engineering and customer-operations hiring staged ahead of each deployment phase. GridOne's working principle is stated in its tagline — Infinite Compute — delivered the practical way: in-country, accountable, and sized to the institutions that need it.

Compute, close to home

The UAE has set a national course for AI leadership and is building the infrastructure to match. GridOne's contribution is deliberately focused: dependable, in-country GPU capacity, operated from Dubai, sized for the institutions doing the country's teaching, healing and research — on a single proven platform, grown one validated block at a time, and gated on real committed demand at every step.

For capacity reservations, partnership enquiries, or further information:

GridOne — Infinite Compute

Abdulla Alfalasi, Chief Executive Officer

gridone.ae · support@gridone.ae · Dubai, United Arab Emirates

Glossary

GPU — graphics processing unit; the accelerator chip on which modern AI models are trained and served.	Training / fine-tuning — building a model from data, or adapting an existing model to a specific domain.
HBM — high-bandwidth memory attached to the GPU; larger HBM allows bigger models per GPU.	Inference — running a trained model to answer requests in production.
NVLink — NVIDIA's high-speed interconnect that lets multiple GPUs work as one system.	Data residency — the legal/policy requirement that data be stored and processed within a specific country.
MIG — Multi-Instance GPU; NVIDIA's facility for partitioning one physical GPU into isolated instances. GridOne uses it within a single customer's allocation, not between customers.	PDPL — UAE Personal Data Protection Law (Federal Decree-Law No. 45 of 2021).

References

[1] UAE Cabinet — National Artificial Intelligence Strategy 2031. uaecabinet.ae/en/news/uae-cabinet-adopts-national-artificial-intelligence-strategy-2031

[2] UAE Government — Artificial intelligence in government policies. uae.ae/en/about-the-uae/digital-uae/digital-technology/artificial-intelligence

[3] The National — “Stargate UAE data centre to cost more than \$30bn, AI minister says” (Jan 2026). thenationalnews.com

[4] OpenAI — “Introducing Stargate UAE”. openai.com/index/introducing-stargate-uae

[5] Federal Law No. 2 of 2019 on the use of ICT in health fields — health data localisation. See analyses at simmons-simmons.com and lexology.com

[6] DLA Piper — Data Protection Laws of the World: United Arab Emirates. dlapiperdataprotection.com

[7] Gulf Business — “UAE government schools to add AI curriculum for KG–G12” (2025). gulfbusiness.com

[8] Gulf News — “UAE schools to introduce AI curriculum in new academic year” (2025). gulfnews.com

[9] Mohamed bin Zayed University of Artificial Intelligence. mbzuai.ac.ae

[10] Abu Dhabi Media Office — DoH agreements on genomics, AI, research and healthcare investment. mediaoffice.abudhabi

[11] Gulf News — “Abu Dhabi launches AI-powered virtual human twin project”. gulfnews.com

[12] NVIDIA — Blackwell architecture / Blackwell Ultra technical materials. nvidia.com, developer.nvidia.com

[13] NVIDIA — HGX B300 platform documentation and the NVIDIA-Certified Systems programme. nvidia.com, docs.nvidia.com. Server configuration per KERN0 UQLEAD K03-8220 Ultra supplier specification.