

AI-Ready GPU Cloud Compute Advisory

G Cloud 15 - LOT1A

1. Insight

Insight leads the delivery of transformative solutions, supported by a powerful ecosystem of over 6,000 global partners including Microsoft, AWS, and Cisco.

As a Fortune 500 company with a long-standing history of delivering trusted services to the UK Public Sector, we combine our own deep expertise in consulting, advisory, and transformation with partner innovation to create holistic solutions across cloud, data, modern workplace, cybersecurity, and application services.

Guided by proven methodologies and ESG principles, we work collaboratively to design and implement tailored strategies that drive measurable business impact while promoting sustainability and governance excellence.

2. Insight for the Public Sector

Your End-to-End Digital Transformation Partner

As a leading Solutions Integrator, Insight combines deep technical expertise with a vast partner ecosystem to help UK government organisations modernise, innovate, and deliver citizen-centric services. We hold places on over 85 public sector frameworks, offering a compliant, efficient route to digital transformation.

We don't just supply IT; we design, deploy, and manage the solutions that secure your data and unlock value.

Core Areas of Expertise

- **Innovation, Data & AI:** we deliver end-to-end data science and custom AI solutions. From modernising data platforms to deploying bespoke machine learning and spatial computing, we help you unlock the value of your data to drive operational efficiency.
- **Modern Infrastructure & Cloud:** Building scalable, hybrid foundations that reduce technical debt and support multi-cloud strategies.
- **Modern Workplace:** Empowering the public sector with secure, flexible tools to work from anywhere, attracting talent and improving productivity.
- **Cybersecurity:** Comprehensive managed security services, threat detection, identity protection, and compliance management to ensure your data remains sovereign and secure.
- **Modern Apps:** Agile application development creating meaningful and impactful tools that link to tangible benefits, improvements and cost savings.

Our Approach: Consulting, Lifecycle, & Managed Services

We deliver value at every stage of your technology journey, ensuring accountability from strategy to support:

- **Consulting (Architecting the Future):** We bridge internal skills gaps with data-driven strategy. From "Discover & Design" to full-scale migration planning, we ensure your investments solve real public sector challenges.
- **Lifecycle (Seamless Execution):** We take the risk out of procurement and deployment. Leveraging global buying power, we handle everything from supply chain management and imaging and implementation to sustainable asset disposal (ITAD).
- **Managed Services (Always-On Support):** Shift from "keeping the lights on" to innovation. We provide 24/7 infrastructure monitoring, cloud cost optimisation (FinOps), and security operations to ensure resilience.

Service Description

Expert consulting to evaluate, design, and optimise GPU-accelerated cloud compute environments for enterprises seeking advanced AI, HPC, or analytics workloads.

Constraints

Requires access to clients existing cloud environments and workload visibility.

Pricing Options

Pricing Option 1

Pricing Unit:

Time & Materials (T&M;)

Unit Qty Definition:

Per Resource-Day

Rationale:

Advisory and design for GPU-accelerated cloud environments are highly bespoke and iterative, with scope evolving as workloads, constraints, and stakeholder needs are better understood, making T&M; suitable for flexible engagement. It also aligns well with activities like assessments, workshops, and ongoing optimisation that vary in depth by client.

Complexity Level	Unit Price	Unit Measure
Low Complexity	£1,950	per resource-day
Medium Complexity	£2,950	per resource-day
High Complexity	£3,850	per resource-day

Pricing Assumptions:

- 22 working days per month is the standard capacity baseline for pricing across all services and complexity tiers.
- All prices are exclusive of VAT.
- This figure already includes normal public holidays, weekends, and excludes overtime unless explicitly stated.
- Seniority & Skill Mix: Low = mid level consultants/engineers; Medium = senior consultants/architects; High = principal architects/SMEs.
- Scope & Effort: Higher complexity = more heterogeneity, integrations, governance, risk, and stakeholder management.
- Volume & Scale: Sometimes tied to asset/user/API counts but often reflects qualitative complexity.

Pricing Option 2

Pricing Unit:

Fixed Price

Unit Qty Definition:

Per Assessment

Rationale:

Discrete, well-defined packages such as initial GPU readiness assessments, reference architecture design, or standard cost-optimisation reviews can be scoped and priced upfront, giving public sector clients budget predictability. This works best for clearly bounded deliverables (e.g. a one-off assessment and recommendations report) rather than ongoing advisory.

Complexity Level	Unit Price	Unit Measure
Low Complexity	£14,500	per assessment
Medium Complexity	£21,500	per assessment
High Complexity	£28,500	per assessment

Pricing Assumptions:

- 22 working days per month is the standard capacity baseline for pricing across all services and complexity tiers.
- All prices are exclusive of VAT.
- This figure already includes normal public holidays, weekends, and excludes overtime unless explicitly stated.
- Seniority & Skill Mix: Low = mid level consultants/engineers; Medium = senior consultants/architects; High = principal architects/SMEs.
- Scope & Effort: Higher complexity = more heterogeneity, integrations, governance, risk, and stakeholder management.
- Volume & Scale: Sometimes tied to asset/user/API counts but often reflects qualitative complexity.
- Well bounded scope, typically 3 to 6 weeks elapsed for assessments or up to several months for setups/migrations.

- Defined scope and deliverables per complexity tier.
- Any changes to scope of project will be impacted accordingly.

Pricing Option 3

Pricing Unit:

Dedicated Team

Unit Qty Definition:

Per Team-Month

Rationale:

For organisations running or planning large-scale AI/HPC estates, a retained advisory team provides predictable monthly cost while supporting continuous optimisation, roadmap alignment, and training over time. This model suits long-term partnerships where GPU usage, vendors, and compliance requirements evolve and need sustained expert input.

Complexity Level	Unit Price	Unit Measure
Low Complexity	£39,900	per dedicated advisory team per month
Medium Complexity	£59,900	per dedicated advisory team per month
High Complexity	£77,900	per dedicated advisory team per month

Pricing Assumptions:

- 22 working days per month is the standard capacity baseline for pricing across all services and complexity tiers.
- All prices are exclusive of VAT.
- This figure already includes normal public holidays, weekends, and excludes overtime unless explicitly stated.
- Seniority & Skill Mix: Low = mid level consultants/engineers; Medium = senior consultants/architects; High = principal architects/SMEs.
- Scope & Effort: Higher complexity = more heterogeneity, integrations, governance, risk, and stakeholder management.
- Volume & Scale: Sometimes tied to asset/user/API counts but often reflects qualitative complexity.
- 22 working days per month is the standard capacity baseline for costing a dedicated team.
- Engagements are long term and stable — typically 6 to 12+ month minimum term.

- Teams are fully utilised on the customer programme.
- Service is delivered continuously across the year (not project spikes), often with 24/7 coverage if SLA requires.

CONTACT US

For more information, including validating any assumptions or resource requirements, please contact your Insight Account Manager.

Tel: 0344 846 3333

Email: pstenderteam@insight.com

Website: <https://uk.insight.com>

