BOSTON

Servers | Storage | Solutions

A DIVISION OF SOURCECODE

FROM VISION TO REALITY
POWERING YOUR AI BREAKTHROUGHS

CONTENTS

OVERVIEW

Whether you are a hyperscaler building an AI Factory or an SME developing an initial AI Strategy, you are going on a transformative journey that presents both exciting opportunities and complex challenges. From navigating technical obstacles to balancing innovation with efficiency, modern organisations require a reliable partner who understands their goals and can provide tailored support at every step.

At Boston Limited, we combine decades of expertise with a collaborative, end-to-end approach, ensuring your vision becomes reality without the hurdles of managing intricate, multi-vendor ecosystems.

Embarking on a journey from zero to fully operational AI Factory can be an incredibly daunting and complex task.

Who can you rely on to truly understand every aspect of the AI ecosystem? From the data centre infrastructure and advanced liquid cooling to compute accelerators, high performance networking, HPC, AI, MLOps and Cloud software solutions?

Boston Limited is that partner; ready to engage on complex, multi-vendor deployments, project managing and delivering all the elements, ensuring interoperability and smooth deployment.

Modern organisations face a constant demand to innovate while maintaining efficiency and reliability. Our decades of expertise, paired with an innovative approach, enable us to design, deliver and maintain tailored IT solutions that go beyond expectations.

Whether you're building AI-driven systems, scaling existing workloads or exploring new technological frontiers, our services are crafted to help you create your fully fledged AI factory. The Boston AI services offer more than technical solutions—it's a partnership committed to ensuring your infrastructure is a foundation for growth and transformation.

BOSTON AI SOLUTION STACK SELECTOR

BOSTON SERVICES	Solution Design & Consultancy	Installation	Warranty & Support	Maintenance & Managed Services
AI & MLOPS SOFTWARE	NVIDIA AI Enterprise	Run:ai	NVIDIA NIM™	AMD ROCm Suite
INFRASTRUCTURE MANAGEMENT	Base Command™	TrinityX	Supermicro DCMS	OpenHPC
COMPUTE	Supermicro HGX	NVIDIA DGX SuperPOD™	AMD Instinct ™	Intel® Gaudi® 3
NETWORKING	NVIDIA Spectrum™-X Ethernet	NVIDIA Quantum InfiniBand	Ultra Ethernet	SONiC Ethernet
STORAGE	Igloo AI+	VAST	DDN	WEKA
COOLING	Direct-to-Chip Liquid Cooling	Liquid Immersion	Rear Door Heat Exchanger (RDHx)	Air Cooling
HOSTING	Colocation	Self-Hosted	Managed Services	Cloud-Hosted

NVIDIA Elite Partner

CONSULTATION, SOLUTION DESIGN & PLANNING

Successful IT projects begin with smart, forward-thinking design. **Boston's Solution Design and Planning** services combine technical precision with strategic vision, ensuring your infrastructure not only meets today's requirements but also prepares you for future challenges.

Instead of one-size-fits-all designs, we focus on precision—creating scalable network architectures, optimising power usage and visualising workflows. This approach is particularly critical for AI factories and HPC environments, where performance and efficiency must align perfectly. By anticipating future growth and potential challenges, we create designs that evolve with your business.

Boston also offer comprehensive technical qualifications and consultation services for AI and accelerated computing requirements, all of which can be found on our website: www.boston.co.uk.

CONSULTANCY COULD INCLUDE :

- **Compute, Storage & Networking Design:** Reference architecture and custom POD designs to match performance requirements, model training profile, training time scales or ExaFlop performance or other key metrics.

- **Rack Design & Layout:** According to thermal design requirements, alternative cooling technologies and data centre capabilities (PUE, peak power, thermal mapping, etc).

- **Software Architecture:** The management and deployment of GPUs to train AI at scale, offering both a range of both open source and commercially managed cluster management and MLOps packages.

- **Specialised Deployment Services:** Custom designed software development, APIs to further enable customer requirements i.e. GPU marketplace offerings or other cloud integration services.

ANALYSE REQUIREMENTS

BOSTON LABS POC, R&D TESTING & BENCHMARKING

BUDGETING: REVIEW & COST EFFICIENCY

SOLUTION DESIGN & IMPLEMENTATION

DEPLOYMENT, TEST & CUSTOMISATION

CUSTOMER ACCEPTANCE TESTING, HANDOVER & TRAINING

PERIODIC SOLUTION REVIEW, MANAGED SERVICES & WARRANTY

UPGRADE, DECOMMISSIONING & RECYCLING

SCALE AS YOU GROW

Designing an AI Factory or HPC system isn't just about deploying hardware—it's about creating a seamless infrastructure capable of evolving with your business. At Boston, we combine technical expertise with innovative design to build scalable solutions tailored to meet your current and future needs.

OUR APPROACH INCLUDES:

- **Compute Solutions:** From GPUs and CPUs to advanced accelerators like DPUs and FPGAs, we design systems that balance raw power and energy efficiency.

- **High-Performance Networking:** Networking defines the backbone of HPC and AI environments. We integrate technologies such as InfiniBand, RoCEv2 and high-speed Ethernet to deliver low latency, high bandwidth and optimised East-West and North-South data flows.

- **Storage Integration:** Our solutions include DAS, NAS, object-based and file-based storage systems, designed for high throughput and reduced latency. These are critical for handling massive datasets while maintaining performance for read/write-intensive tasks.

- **Rack Design:** Boston's rack designs are optimised for density, cooling and power. Whether using air or liquid cooling, we ensure your infrastructure meets thermal and energy requirements while supporting rack densities up to 120kW.

- **Automation & Monitoring:** With integrated deployment and monitoring tools, we ensure your system runs efficiently from day one. Automating routine processes allows your team to focus on innovation rather than maintenance.

With Boston, your infrastructure is designed to scale as you go, ensuring you're prepared for both immediate challenges and long-term growth.

THE NETWORK DEFINES YOUR SUCCESS

In AI and HPC environments, networking is the critical foundation that defines overall system performance. We understand the unique demands of modern data centres and design networks to address diverse workloads with precision and efficiency.

OUR APPROACH INCLUDES:

- **Cloud Deployments:** For multi-tenant environments running a variety of small-scale applications, we optimise the network for agility and cost-efficiency. These setups use Ethernet and RoCEv2 technologies to provide reliable connectivity for smaller workloads.

- **Hybrid Training & Inference Cloud:** Supporting mixed workloads, including large-scale generative AI, requires robust network architectures. We design these systems with advanced Ethernet (Spectrum-X) and InfiniBand technologies to handle the high traffic demands of AI training and inference.

- **AI Factories:** For larger, private cloud or on premises deployments with single tenant operations, we recommend NVIDIA InfiniBand and NVIDIA NVLink™. These technologies deliver exceptional throughput, minimal latency and efficient data movement between GPUs and storage for large scale model training.

OPTIMISED DATA FLOW

Efficient data movement is the cornerstone of successful AI and HPC workloads. At Boston, we design data pipelines that meet the highest standards of availability, latency and scalability, ensuring seamless operations for diverse applications.

OUR SOLUTIONS ARE BUILT AROUND KEY CONSIDERATIONS:

- **East-West (E-W) Data Flows:** Critical for compute operations, E-W workflows require high bandwidth, low latency and efficient data exchange between nodes. We integrate NVIDIA technologies such as NVIDIA Magnum IO™ GPUDirect® RDMA, NCCL and NVSHMEM to optimise GPU communication for tasks like model training and simulation.

- **North-South (N-S) Data Flows:** Focused on storage operations, N-S workflows prioritise sequential read/write performance for massive datasets. We incorporate NVIDIA GPUDirect Storage and NVMe SNAP to ensure rapid data access and efficient storage utilisation.

- **Integrated Infrastructure:** Our solutions leverage advanced technologies like BlueField DPUs, 100/200/400 Gb InfiniBand and RoCEv2 to seamlessly connect compute and storage systems. These integrations support predictable performance under load and scalability as workloads grow.

- **Optimised Metrics:** We design systems with availability, throughput and security in mind. From sub-microsecond NIC latency to high throughput storage solutions, every component is tuned for maximum efficiency.

SOFTWARE

THE RIGHT TOOLS FOR THE JOB?

Even the most powerful hardware can achieve nothing without complimentary software tools. Working in unison, the combination can make all the difference between achieving your goals on schedule to failing to achieve them completely.

Boston's team of AI and HPC infrastructure engineers have a long-standing pedigree of delivering compute services to customers for whom results matter. Our experience enables us to advise you on the highest quality, best of breed solutions available on the market today, including:

CLUSTER MANAGEMENT

Ensuring your infrastructure remains available and online, the right cluster management tool enables rapid, dynamic provisioning of compute resources to align with business needs. Additionally, robust monitoring and reporting capabilities help detect potential issues early, allowing for proactive measures—whether it's capacity expansion & planning, mitigating hardware failures, or addressing datacenter infrastructure outages.

JOB SCHEDULERS & ORCHESTRATION

Managing and monitoring compute jobs—whether a few or thousands—is a complex challenge. A robust job scheduler optimises resource utilisation, maximizing efficiency and cost-effectiveness. Features like pre-flight checks and checkpointing help ensure successful job execution, reducing wasted compute cycles and minimizing the need for resubmissions.

MLOPS

Streamlining the deployment, monitoring, and management of machine learning models in production. The right MLOps package automates workflows, ensures model reliability, and integrates ML with DevOps practices for scalability and efficiency. The right tools mean data scientists and AI specialists can spend their valuable time productively, generating business outcomes, not debugging package dependencies and interoperability.

NETWORK PERFORMANCE & CONGESTION MONITORING

HPC and AI workloads often involve massive data transfers between nodes, storage, and accelerators. Network bottlenecks can significantly impact performance, causing delays, inefficient resource utilisation, and failed jobs. By proactively monitoring congestion and performance, IT teams can identify issues, optimise data flow, and ensure consistent high-speed communication, ultimately improving job efficiency and system reliability.

DISTRIBUTED TRAINING & PARALLEL COMPUTING

Modern models are simply too large and computationally demanding for a single GPU or node, meaning sale is essential for AI and HPC. These tools enable efficient workload distribution across multiple GPUs and nodes, accelerating training and inference. They optimise data communication, minimize bottlenecks, and improve resource utilisation, ensuring scalability and faster time-to-insight for large AI models and simulations.

Sustainability is at the core of our partner data centre services, powered by renewable and green energy sources.

HOSTING & DATA CENTRE SERVICES

Not every business has the resources to manage an advanced IT environment in-house. That's where **Boston's Hosting & Data Centre Services** come in.

We don't just offer space and power; we provide peace of mind.

All over the world, there is a challenge to find remaining high density data centre capacity due to the explosive growth of AI deployments. Boston Limited have partnered with select data centres to reserve and plan new deployments, with availability to install your infrastructure as soon as 1 month.
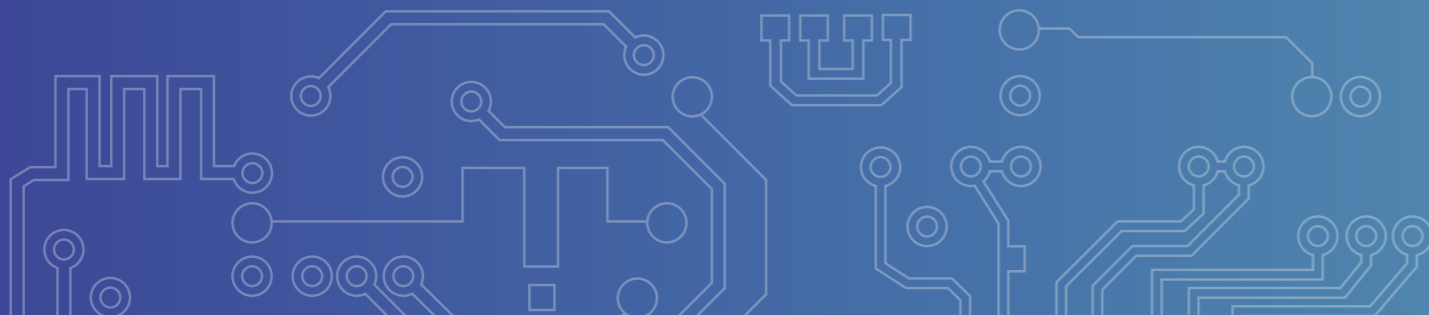
Sustainability too is at the core of our partner data centre services, with renewable, green energy sources and a low PUE from a diverse selection of data centres globally.

Our facilities are designed with HPC and AI deployments from the outset, as well as Tier 3 and 4 capabilities for maximum uptime and resilience. They feature robust high density power configurations, air, hybrid and liquid cooling, high-speed IP and dark fibre connectivity to PoPs and services around the world and round-the-clock security.

Advanced liquid cooling technologies from our long-standing partners ensure we are ready for the next generation of AI accelerators with rack densities of 120kW. For more modest deployments, air cooled racks with capabilities of up to 75kW are available with 100% heat removal.

For AI developers and data scientists, we provide the reliability needed to train sophisticated models, process massive datasets and deploy applications seamlessly. With options tailored to your workload, our hosting services scale as you do.

INSTALLATION, CONFIGURATION & BRING-UP

Building an AI Factory requires precision and expertise at every step, and that's exactly what Boston delivers through our **Installation, Configuration & Bring-Up services.**

From preparation to deployment, we handle the complexities of implementing cutting-edge infrastructure, ensuring your systems are optimised for AI workloads and ready to perform from day one.

OUR PROCESS

Pre-Installation Preparation:

At Boston, we meticulously prepare your cluster to ensure it meets the highest standards of quality and reliability. Our process includes:

- Comprehensive build procedures with peer-reviewed checks for CPU, GPU and memory.

- Burn-in testing to validate long-term stability under intensive AI workloads.

- Integration of customised configurations tailored to your specific requirements.

On-Site Installation:

Our engineers, led by an experienced HPC Installation Manager, execute deployments with precision:

- Pre-installed racks fitted with servers, switches and balanced cabling for performance and resilience.

- Structured routing and labelled cabling for streamlined servicing.

- Optimisation of thermal design with blanking panels to ensure efficient cooling.

Configuration & Validation:

To maximise AI performance, we ensure seamless hardware and software integration, including:

- Preloaded AI frameworks and libraries for model training and inference.

- Benchmarking with HPL, SPEChpc and other performance tests across CPUs and GPUs.

- Network bandwidth validation and temperature testing under load for consistent operation.

Customer Handover:

Once your AI Factory is up and running, we provide comprehensive training, documentation and hands-on guidance to ensure smooth operations. With a focus on customer satisfaction, we offer:

- Detailed installation manuals and schematics.

- Functional checks to verify stability and performance.

- Support handover from the implementation team to your dedicated account manager.

MANAGED SERVICES & MAINTENANCE

AI Factories are like living systems—they need care, attention and constant adaptation.

Our offer is simple—let us manage the details, allowing your data scientists to spend their valuable time working on data science, not IT and service maintenance.

Boston's Managed Services & Maintenance keeps your infrastructure operating at its best, so you can focus on what matters most.

Through regular monitoring, performance tuning and rapid support, we ensure your systems stay resilient. For MLOps and AI workflows, this means more than just maintenance; it's about keeping your frameworks up-to-date, workloads optimised and applications running smoothly.

Select from a comprehensive range of available packages to choose the services which matter to your business most, including:

Data Centre Services – IP Transit, Power & Cooling Administration	AI & HPC Job Scheduler Administration
Hardware Infrastructure & Services Monitoring	MLOps Platform Administration
High Performance Network Fabric Monitoring	Security & User Access Management
Kubernetes Container Management	Scale Out Storage Administration
Application Patch Management & Optimisation	Application Support & Installation

BOSTON TRAINING ACADEMY

Participants receive hands–on training sessions, with help from developers, data scientists and researchers. Attendees learn how to approach challenges using deep learning techniques like building transformer–based NLP, building conversational AI applications, building intelligent recommender systems, among other approaches.

EMPOWERING TOMORROW'S TECH EXPERTS

Boston Training Academy provides structured training by world-class trainers who deliver tailored courses based on the current knowledge of the attendees. Our program content is designed by industry experts and covers data science skills, exposure to analytical tools and corporate learning and consulting on AI projects, to name a few.

The AI Centre of Excellence (CoE-AI) at our Training Academy brings together experts from industry academia and Government to provide the latest infrastructure, smart tools, skilled resources and leadership for AI exploration. The CoE-AI aims to create the most sophisticated ecosystem for AI, Data Sciences, Machine Learning and Deep Learning to enable breakthrough innovations.

We also offer online starter courses for complete beginners to gain working knowledge on the fundamentals of NLP, computer vision, deep learning and Python for data science. Boston also host NVIDIA Deep Learning Institute (DLI) workshops. Participants receive hands-on training sessions, with help from developers, data scientists and researchers. Attendees learn how to approach challenges using deep learning techniques like building transformer-based NLP, building

conversational AI applications, building intelligent recommender systems, among other approaches.

We provide practical training to universities and businesses on algorithm creation, hardware development and the latest technologies. Our expert instructors empower organisations to stay competitive in a rapidly evolving digital world.

- **Expert-led Training:** We offer expert-led training programs delivered by industry professionals, ensuring participants gain practical insights and skills that are immediately applicable in their professional roles.

- **Tailored Learning Paths:** We excel in providing personalised learning paths, allowing participants to choose courses that match their skill levels and career aspirations. This tailored approach ensures that individuals receive relevant and effective training.

- **Hands-on Lab Experience:** With a strong emphasis on practical learning, we offer hands-on lab experiences that enable participants to engage with real-world scenarios, fostering better understanding and confidence in their abilities.

UPGRADE, END-OF-LIFE & DISPOSAL

Even the most advanced technology today will eventually become outdated, requiring refresh and upgrades to stay competitive.

Boston's Upgrade, End-of-Life & Disposal services help you stay ahead by managing outdated equipment responsibly and efficiently. We provide customised upgrade paths to optimise your infrastructure for evolving workloads, such as AI and data-intensive applications.

For hardware reaching the end of its lifecycle, we ensure secure data destruction and eco-friendly recycling, meeting all environmental compliance standards. Additionally, our health check services give you a comprehensive view of your system's current state, helping identify opportunities for repairs or upgrades to extend its serviceable life.

When it's time to retire old equipment, Boston makes it easy to modernise your systems sustainably, securely and with scalability in mind—preparing your organisation for the future of AI and innovation.

LEASING & FINANCE

Establishing an AI Factory is a major, long-term investment that can tie up significant cash resources for extended periods. However, modern businesses often favour OPEX-focused financial strategies to maintain flexibility and optimise cash flow. This shift doesn't have to put cutting-edge technology out of reach.

Boston's Leasing & Finance solutions simplify the process of acquiring infrastructure that drives innovation and growth. With flexible options such as leasing and tailored payment plans, you can access next-generation technologies like AI platforms, high-performance clusters or AI Factory deployments without straining your budget. We help eliminate financial barriers, empowering your organisation to confidently embrace the future.

WARRANTY & SUPPORT

Our services include responsive troubleshooting, hardware replacements and ongoing updates, keeping your systems optimised for peak performance. For AI projects, we provide specialised support for frameworks and GPU resources, giving you the confidence to pursue innovation without interruptions.

Uninterrupted performance depends on having a reliable safety net in place. Boston's Warranty & Support services provide extensive coverage and ongoing assistance to ensure your systems operate seamlessly.

When you employ Boston's services, our teams understand your environment inside and out. We support the hardware, software, storage and networking; simplifying the whole support process and reducing your burden.

Equally, tailored to meet your exact requirements, Boston offer a flexible range of predefined services with SLAs to meet your business needs.

For those mission critical applications, our 24/7 service with a 4-hour response is the perfect match. In those situations where you need a little extra care, our managed services can add an extra wrap around layer.

WARRANTY SERVICES	RETURN TO BASE (RMA)	ADVANCED WARRANTY	ONSITE WARRANTY			
			PEWTER	BRONZE	SILVER	GOLD
TECHNICAL SUPPORT DESK	YES	YES	YES	YES	YES	YES
ADVANCED PARTS SHIPPING	NO	YES	NO	YES	YES	YES
REMOTE ISSUE TRIAGE	YES	YES	YES	YES	YES	YES
ONSITE SUPPORT/ DIAGNOSIS	NO	NO	NO	NO	NO	NO
ONSITE BREAK/FIX LABOUR	NO	NO	YES	YES	YES	YES
ENGINEER TO SITE TARGET (POST DIAGNOSIS)	N/A	N/A	NBD	NBD	SBD	4 HOURS
ACCESS TO SERVICES 24 HOURS A DAY, WEEKENDS & UK BANK HOLIDAYS	NO	NO	NO	NO	NO	YES

OUR PARTNER NETWORK

CUTTING EDGE TECHNOLOGY

Introducing our esteemed network of strategic alliances and collaborations that have been fundamental in shaping our success as a leading technology solution provider. Cultivating robust partnerships is the key to innovation and delivering state-of-the-art solutions to our customers. Join us on this journey as we introduce you to our valued partners, each contributing their unique expertise and shared vision to drive unparalleled growth and excellence in the ever-evolving technology landscape.

PARTNERING WITH INSTADEEP FOR AI INFRASTRUCTURE

Boston has been instrumental in supplying and supporting InstaDeep's AI infrastructure, including the installation of 12 DGX systems and 85 compute nodes. Our collaboration extends to providing intricate engineering services and facilitating infrastructure deployment discussions with NVIDIA.

InstaDeep, renowned for its expertise in deep learning and reinforcement learning, serves various industries, from finance to healthcare. As an AI-first company, they leverage advanced technologies and accelerated compute to address complex challenges and opportunities across sectors.

InstaDeep's focus has remained on quick access to hardware and robust support for in-house data centre hardware engineering. Boston resolved memory components with a vendor-backed swap-out and facilitated a smooth hardware transition to a new data centre.

Boston's solution involved rapid build, testing, and configuration within two weeks, including the installation of 20 CPU nodes. Leveraging partnerships with vendors like NVIDIA and leveraging Boston's in-house design expertise, we ensured a smooth transition and competitive pricing for InstaDeep.

For the various projects, implementation and support has been provided by Boston. Our EU shipping capabilities and warranty support via ticketing systems ensured a customer-first experience for InstaDeep. We continue to engage with InstaDeep, providing roadmap sessions and facilitating vendor engagement to meet evolving needs.

Our collaborative efforts have resulted in optimised AI infrastructure and enhanced support for InstaDeep's growth plans. Boston remains dedicated to providing exceptional service and support to InstaDeep, ensuring they have the cutting-edge infrastructure needed to drive innovation and success in the AI landscape.

> **INSTADEEP ACCESSED AN H100 DGX (8 GPU) IN BOSTON LABS FOR BENCHMARKING. USING THE SYSTEM TO TRAIN TRANSFORMER MODELS, SPECIFICALLY BERT-STYLE, FOR GENOMICS LLM TRAINING, INSTADEEP HARNESSED MASSIVE COMPUTATIONAL POWER FOR FREE. FACILITATING A STRATEGIC APPROACH TO FUTURE AI PLATFORM PURCHASES.**
>
> **LOCATION:** UNITED KINGDOM

BOSTON AROUND THE WORLD

EUROPE

USA

ASIA

AFRICA

AUSTRALIA

1992
OVER 30 YEARS IN THE INDUSTRY

5
CO-DESIGN MANUFACTURING SITES

40
COUNTRIES SERVED

4
INTERNATIONAL LABS

IF YOU'D LIKE TO FIND OUT MORE, PLEASE GET IN TOUCH:

E: sales@boston.co.uk | **T:** +44 (0) 1727 876 100 | **W:** boston.co.uk

Follow us: Keep up to date with Boston on socials!

Facebook: facebook.com/bostonlimited
Instagram: instagram.com/boston_limited

LinkedIn: linkedin.com/company/boston-limited
X: twitter.com/bostonlimited

USA

MILFORD, MA
NORWOOD, MA
BEVERLY, MA
WESTFIELD, MA
PROVIDENCE, RI
BUFFALO, NY
SEATTLE, WA
LOS ANGELES, CA

EUROPE

LONDON
ST ALBANS - HQ
MUNICH
ANTIPOLIS
KREUZLINGEN
HAUPTPLATZ

AFRICA

JOHANNESBURG

ASIA

MUMBAI
BENGALURU
DUBAI
TAIPEI

AUSTRALIA

SYDNEY