BOSTON

Servers | Storage | Solutions

A DIVISION OF SOURCECODE

FROM VISION TO REALITY

POWERING YOUR AI BREAKTHROUGHS

AI SOLUTION GUIDE

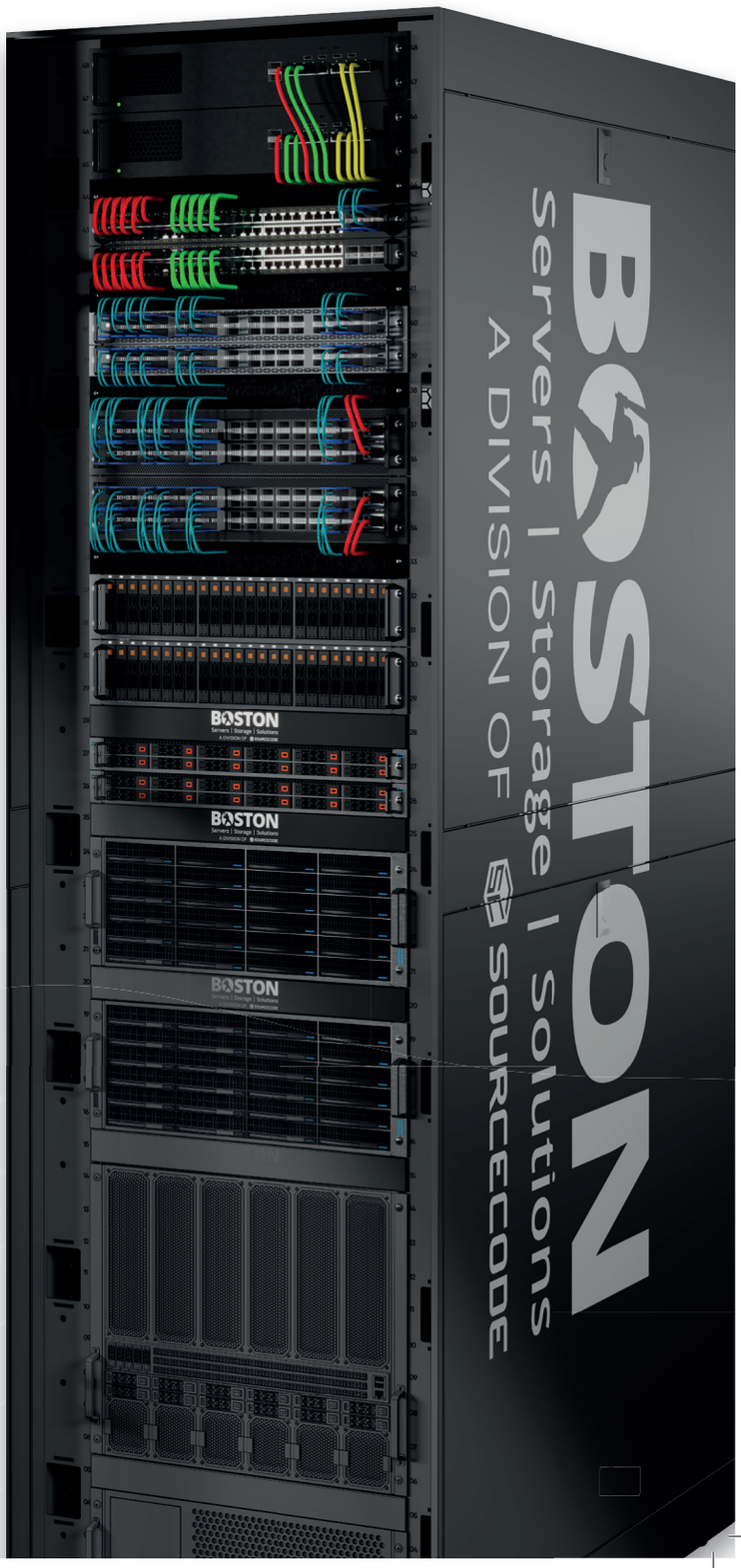



BOSTON
Servers | Storage | Solutions
Powered by SUPERMICRO
CAUTION!

CONTENTS

WHAT'S INSIDE?

Boston: Powering Innovation For Over 30 Years	04
What Sets Boston Apart?	05
Beginning your AI Journey	06
An Introduction To AI Factories	08
Designing The Right Architecture	10
From Concept To Deployment	12
Your Journey With Boston	13
Boston AI Solution Stack Selector	14
AI Factory Managed Services & Maintenance	16
Leasing & Finance	17
Enterprise AI Applications	18
AI & Mlops Software	20
Infrastructure Management For AI	22
Accelerated Compute Performance	24
AI Networking Architecture	26
Storage	28
Cooling	30
Hosting & Data centre Services	32
Boston Training Academy	34
The Genetic Code: Nature's Ultimate Dataset	36
Our Global Partner Network	40
Boston Around The World	42



BOSTON: POWERING INNOVATION FOR OVER 30 YEARS

Boston Limited has spent over three decades redefining what's possible in high-performance computing. From our early days pioneering bespoke server solutions to becoming a global leader in AI, cloud and advanced computing, our journey has been shaped by a relentless drive to push technology forward.

AI FACTORY: FROM ZERO TO FULLY OPERATIONAL.

Whether you're a hyperscaler building an AI Factory or an SME shaping your first AI strategy, the path is transformational - and complex. Success depends on having a partner who understands your goals and delivers the right support at every stage.

CUSTOM HPC & AI INFRASTRUCTURE

We design tailor-made HPC clusters and end-to-end platforms for the most demanding workloads, including AI and machine learning, CFD, molecular dynamics, particle physics, cloud infrastructure, high-frequency trading and large-scale storage.

ENGINEERED FOR PERFORMANCE

At Boston, we don't just build systems - we engineer solutions designed for efficiency, reliability and uncompromising performance.

FULL ECOSYSTEM EXPERTISE

From data centre infrastructure and advanced liquid cooling to compute accelerators, high-performance networking, HPC, AI, MLOps and cloud software, we understand every layer of the AI ecosystem.

INTEGRATION, TESTING & DELIVERY

Boston Limited is built for complex, multi-vendor deployments. We project manage and deliver every element - ensuring interoperability, smooth implementation and dependable outcomes.

SUPPORT BEYOND DEPLOYMENT

Every solution is backed by expert integration, rigorous testing and deep technical insight. We offer branded custom-built platforms, global lifecycle management, training, advisory services and hands-on support long after deployment.

A PARTNER FOR GROWTH

As demand for high-performance technology accelerates, Boston continues to expand, supporting more industries, more markets and more innovators who refuse to settle for off-the-shelf solutions.

READY WHEN YOU ARE

If performance, precision and partnership matter to you, Boston is ready to power your success.



WHAT SETS BOSTON APART?



SOLUTION CO-DESIGN & CUSTOMISATION

Complex, multi-vendor and end-to-end solutions built to meet specific customer needs and workloads



EDGE TO EXASCALE

From edge computing to the most demanding exascale environments - solutions for any scale



INDEPENDENT & VENDOR-NEUTRAL GUIDANCE

Our agnostic approach ensures the best-fit solution for every customer, appraising options from industry respected vendors



LABS, TESTING & PROOF-OF-CONCEPT

In-house labs deliver rapid verification, certification, benchmarking and value demonstration - accelerating decision making and reducing risk



ELITE PARTNERSHIPS & INDUSTRY RECOGNITION

Accredited and recognised partnerships with industry leaders like NVIDIA (Elite Partner, SuperPOD) - giving you early access to innovative, ground-breaking products and technologies



AI & HPC CONSULTANCY + TRAINING

Strategic guidance paired with practical training (incl. NVIDIA DLI), helping customers adopt and scale to match their complex computing challenges



GLOBAL REACH, LOCAL SUPPORT

Our international presence ensures responsive support - wherever your operations are



SUSTAINABILITY THROUGH INNOVATION

We integrate energy-efficient, future-facing technology into every solution, supporting greener IT and reducing TCO while simultaneously improving performance and density



FULL-SERVICE SUPPORT: DESIGN, DEPLOYMENT, OPERATION

Professional services that span through design, build, testing, integration, ongoing managed services and support

BEGINNING YOUR AI JOURNEY

Boston's Future State AI Blueprint and Roadmap helps organisations turn AI ambition into a clear, actionable plan. Through a focused stakeholder workshop and targeted interviews, we define your current capabilities, future goals and the highest-impact AI opportunities.

The engagement concludes with a practical roadmap outlining priority initiatives, platform and infrastructure recommendations, and a phased implementation plan-giving you the confidence to move from concept to delivery with Boston as your trusted partner.

INITIAL WORKSHOP

- Definition of "Current State" of the business with a view to how the business wants to evolve to a "Future State".
- Half day workshop to bring together stakeholders in the business and IT to pull this detail onto the whiteboard.

INTERVIEW PHASE

- Stakeholders identified in the workshop will be interviewed on an individual basis to extract more detailed information pertinent to the transformational need.
- Information from the first two phases will be brought together to Boston's team of technical specialists to build out the Future State Blueprint and Roadmap.

DELIBERATION PHASE

- The team will select a limited number of AI initiatives to develop working on the options available to develop a best of breed platform for delivery of the AI project.
- Boston will work with industry specialists on a roadmap and plan to implement a working AI Solution.

PRESENTATION AND DELIVERY

- Boston will present a document which captures the "Current State" and desired "Future State". The documentation will present how the existing IT environment can be augmented to support the needs of a select number of identified AI projects. The document presents a roadmap outlining the steps needed to deliver a working AI solution. TCO analysis will be presented where available.

FUTURE SUPPORT

- Boston will provide follow up support to help with delivery of these projects.

AMOS ANKRAH

Solution Specialist

"BEFORE YOU BUY INFRASTRUCTURE, DEFINE OUTCOMES"

Amos Ankrah is a consultant at Boston Limited, specialising in AI-driven full stack infrastructure for mid-market enterprises. He helps organisations design, deploy, and scale intelligent systems that integrate seamlessly with existing operations.

With a focus on performance and innovation, Amos delivers practical solutions that accelerate digital transformation and long-term business value.

E: amos.ankrah@boston.co.uk
T: +44 (0) 1727 876 100

SCAN TO FIND OUT

MORE INFORMATION



AN INTRODUCTION TO AI FACTORIES

WHAT IS AN AI FACTORY

An AI Factory is an integrated computing environment engineered to operationalise the development, training and deployment of artificial intelligence at scale. It consolidates high-performance computing resources, low-latency networking, accelerated storage systems and optimised software frameworks into a unified architecture.

A CLOSED-LOOP AI PRODUCTION ENVIRONMENT

Within this environment, data is systematically ingested, processed and prepared for model training. Models are then iteratively developed, validated and deployed, creating a closed-loop workflow that is repeatable, measurable and optimised for throughput.

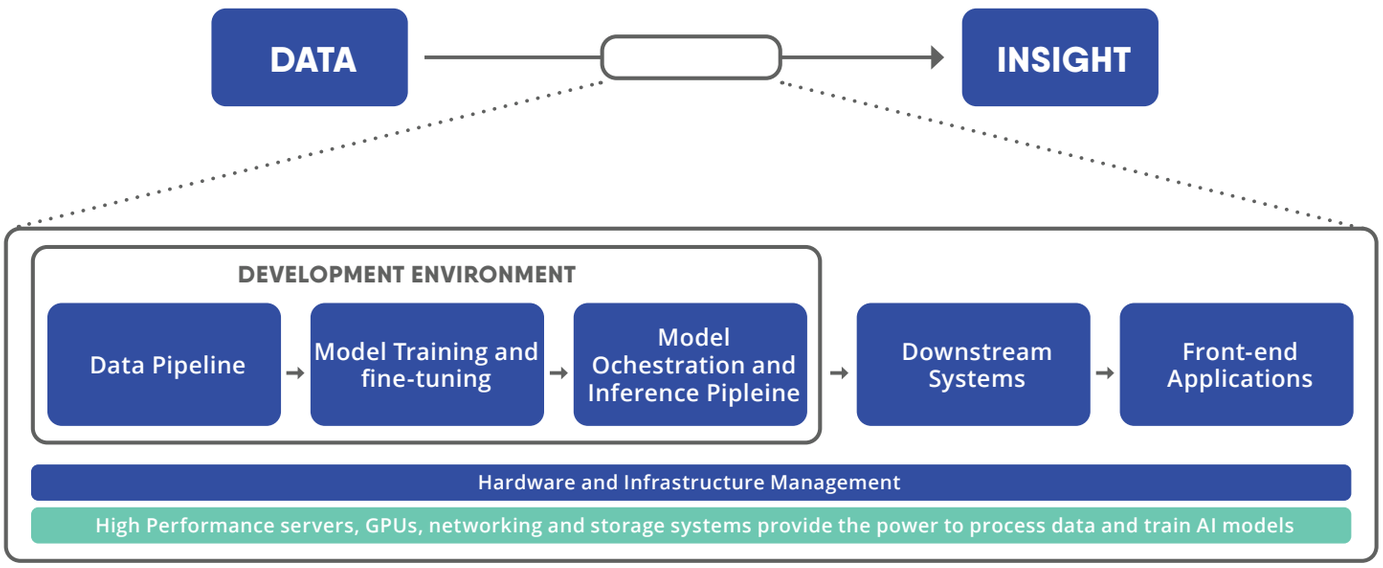
By standardising and industrialising these processes, AI Factories transform AI from ad-hoc experimentation into a sustained, high-efficiency capability.

SCALING AI WITH CONSISTENCY AND CONTROL

This approach enables organisations to execute multiple AI initiatives in parallel, reduce development cycles and maintain consistency and quality across AI-driven operations.

FROM RAW DATA TO TOKENS

To understand the outputs of this process, consider the analogy of a brick factory transforming raw clay into uniform bricks. In an AI Factory, an AI model transforms raw data into its fundamental unit of production: the token. Each token is a discrete digital building block.



MANUFACTURING INTELLIGENCE

The model's internal machinery learns patterns and relationships within massive datasets, assembling tokens one by one into coherent outputs, whether that be an essay, a line of code or a spoken sentence. In this way, the AI Factory manufactures intelligence, piece by piece, from data.

BOSTON LIMITED'S AI FACTORY EXPERTISE

Boston Limited designs and delivers end-to-end, application-defined AI Factories. Our accredited architects and specialists manage the entire project lifecycle, from requirements analysis, design and system configuration to installation, software integration, optimisation and long-term maintenance.

Every solution is engineered around specific workload demands, ensuring optimal performance and scalability.



DESIGNING THE RIGHT ARCHITECTURE

Designing an AI Factory 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.

AI CLUSTERS BUILT FOR REAL-WORLD RESULTS

AI clusters are enabling breakthroughs in scientific research, financial modelling, AI and beyond. But building a truly optimised cluster isn't about stacking servers, it's about engineering an ecosystem where compute, networking, storage and software work seamlessly together.

AI SUCCESS ISN'T DEFINED BY HARDWARE ALONE; IT'S DEFINED BY BALANCE, INTEGRATION AND OPTIMISATION.

ENGINEERING THE RIGHT ARCHITECTURE

At the heart of every AI cluster is a carefully balanced design, where every component plays a critical role.

COMPUTE: CPU AND GPU WORKING IN HARMONY

CPUs orchestrate serial workloads and manage system processes, while GPUs accelerate parallel computation, delivering the performance required for AI training, simulations and complex modelling.

NETWORKING: ELIMINATING BOTTLENECKS

High-speed, low-latency networking ensures data moves efficiently between nodes, preventing performance constraints and enabling real-time processing at scale.



STORAGE: MATCHED TO THE WORKLOAD

From high-capacity disk arrays to ultra-fast NVMe, storage must be aligned with the application. This supports fast ingest, throughput and access to large datasets.

SOFTWARE: THE LAYER THAT UNIFIES EVERYTHING

A robust software stack ensures every element operates as one, from job scheduling and resource management to monitoring, orchestration and optimisation.

A CLUSTER IS ONLY AS FAST AS IT'S SLOWEST BOTTLENECK.

THE RISK OF GETTING IT WRONG

Designing and deploying an AI environment requires specialist expertise. Poorly planned clusters often lead to inefficiencies, underutilised resources and escalating costs.

Many IT teams face challenges such as incompatible software environments, networking misconfigurations, poor workload performance, or systems that don't scale effectively over time.

POOR CLUSTER DESIGN DOESN'T JUST LIMIT PERFORMANCE; IT INCREASES COST, COMPLEXITY AND DOWNTIME.

BOSTON: EXPERTISE THAT REMOVES COMPLEXITY

With decades of experience in advanced infrastructure, Boston takes the complexity out of cluster deployment.

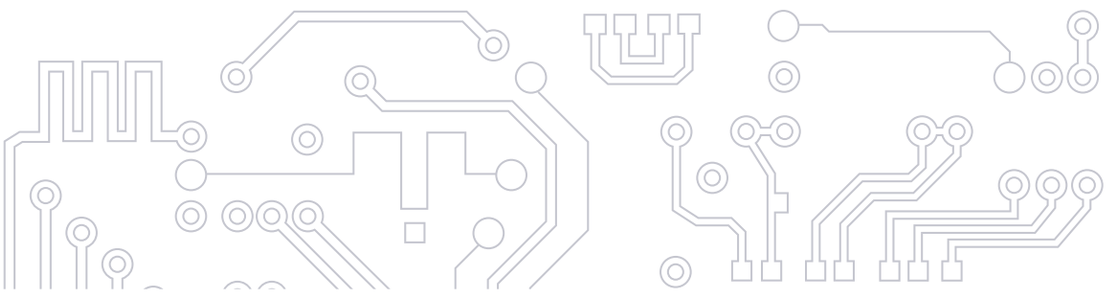
Every solution is pre-engineered, tested and performance-tuned in Boston Labs, ensuring seamless integration of hardware and software before it reaches the customer site. Whether using open-source or commercial cluster management tools, we tailor each system to match your workflows, environment and performance objectives.

BUILT TO DEPLOY, MANAGE AND SCALE

The result is a cluster that delivers more than raw power. Boston provides an AI ecosystem that is efficient, scalable and operationally robust from day one.

With Boston, you don't just receive a collection of servers, you gain a fully optimised platform designed to accelerate innovation, improve productivity and support long-term growth.

WITH BOSTON, AI IS READY FROM DAY ONE.



FROM CONCEPT TO DEPLOYMENT

1. PRESALES CONSULTANCY

Align your business objectives with mission-critical technologies through expert guidance from Boston's integrated sales and engineering teams.

Budget Alignment & Finance

Control costs and accelerate procurement with flexible financing and budget-conscious solution planning.

Solution Architecture

Design high-performance, bespoke infrastructure tailored to your workload, scalability, and operational goals.

2. PROOF OF CONCEPT

BOSTON LABS

Validate performance, reduce deployment risk, and fine-tune solutions in Boston's dedicated test environment before production rollout.

Proof-of-Concept & Solution Validation

Test complex, multi-layered solutions in a controlled lab environment to validate performance before live deployment.

Early Access to Emerging Technologies

Explore the latest hardware and software innovations - often years ahead of general availability - through Boston's vendor ecosystem.

Collaborative Development & Testing

Work directly with Boston engineers and technology partners to optimise architectures, refine workloads, and ensure deployment readiness.

Expert Support & Debugging

Accelerate troubleshooting and minimise downtime with hands-on technical expertise spanning testing, integration, and production support.

3. COMMISSIONING

System Integration & Customisation

Deploy fully configured, interoperable systems engineered to integrate seamlessly within your existing infrastructure.

Delivery & On-site Installation

Ensure smooth deployment with secure logistics, professional installation, and certified On-site engineering support.

4. DEPLOYMENT & ONGOING SERVICE

Transition systems into production with firmware alignment, network provisioning, and operating system deployment.

Handover & Acceptance Testing

Verify performance and operational readiness through rigorous benchmarking and validation testing.

Maintenance & Managed Services

Maximise uptime and operational efficiency with proactive monitoring, remote management, and expert support services.

Warranty & Support

Protect critical infrastructure with responsive technical support and enterprise-grade SLA coverage.

5. DECOMMISSIONING & RECYCLING

Retire infrastructure securely and sustainably with certified data destruction and compliant e-waste recycling.

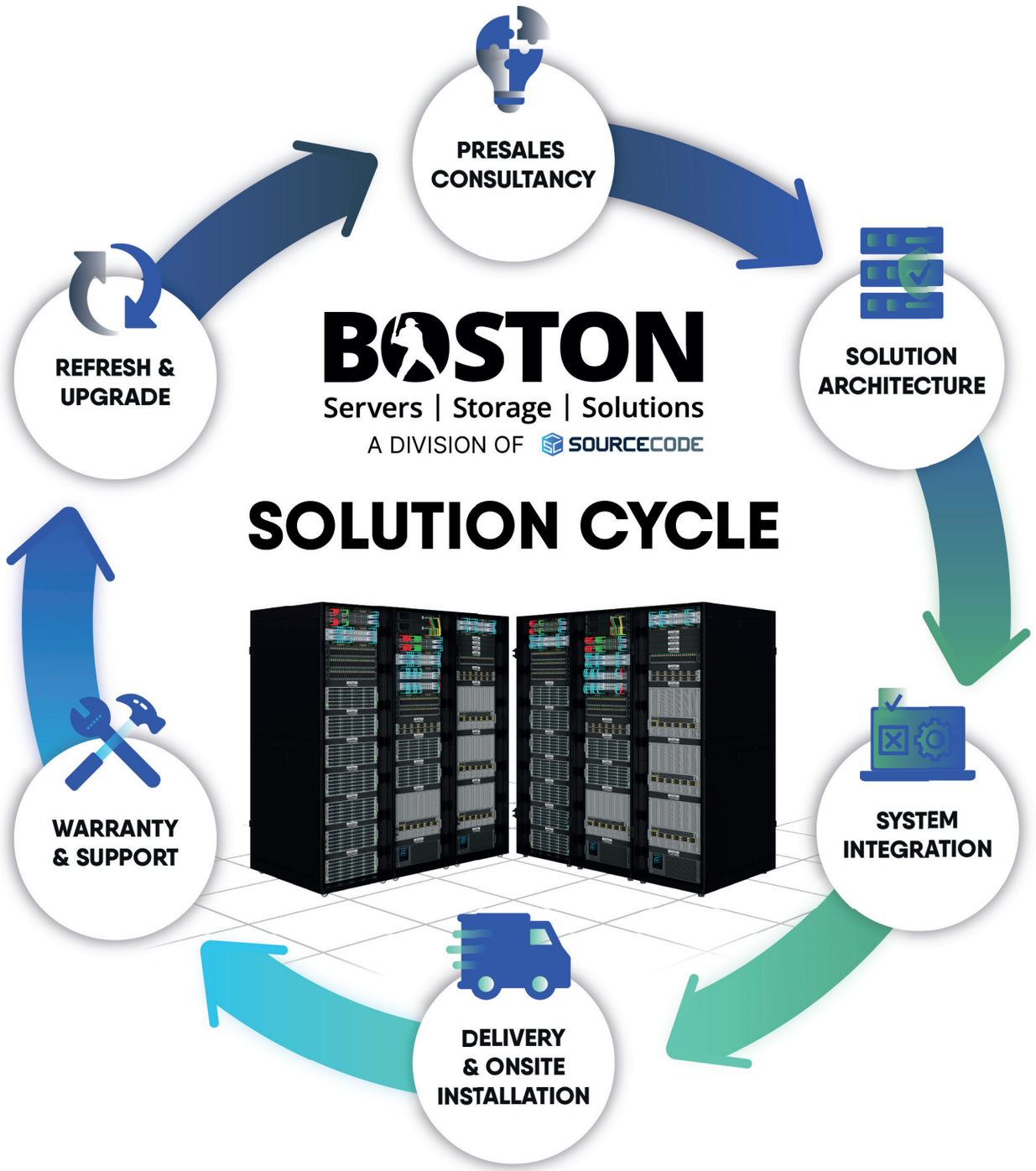
Refresh & Upgrade

Stay ahead of evolving demands with technology refreshes, performance optimisation, and scalable infrastructure planning.

YOUR JOURNEY WITH BOSTON

TEST DRIVES AVAILABLE

CHECK OUT THE LABS



BOSTON AI SOLUTION STACK SELECTOR

BOSTON SERVICES	Solution Design & Consultancy	Installation	Warranty & Support	Maintenance & Managed Services
ENTERPRISE AI APPLICATIONS	KVARK Enterprise Agentic AI	NVIDIA AI Data Platform	NVIDIA NIM Models & Blueprints	RAG
AI & MLOPS SOFTWARE	DKube	Lablup Backend Ai	NVIDIA AI Enterprise	NVIDIA Run:ai
INFRASTRUCTURE MANAGEMENT	NVIDIA Base Command™ & Mission Control™	OpenHPC	Supermicro Cloud Composer	TrinityX
COMPUTE	AMD Instinct Accelerators	NVIDIA DGX SuperPOD™	NVIDIA RTX Pro Servers	Supermicro HGX Servers
NETWORKING	NVIDIA Quantum InfiniBand	NVIDIA Spectrum™-X Ethernet	SONiC Ethernet	Ultra Ethernet
STORAGE	DDN	Igloo AI+	VAST	WEKA
COOLING	Single-Phase DLC	Two-Phase DLC	Immersion Cooling	RDHX Cooling
DATA CENTRE SERVICES	Colocation	Self-Hosted	Managed Services	Cloud-Hosted

DESIGNED TO SIMPLIFY
COMPLEXITY AND ACCELERATE SUCCESS

From planning and design to deployment and optimisation, our in-house specialists and trusted partners support you throughout the entire project lifecycle, ensuring your systems run securely, reliably and at maximum efficiency.



AI FACTORY 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.

CHOOSE FROM A COMPREHENSIVE RANGE OF SERVICES TAILORED TO YOUR PRIORITIES, INCLUDING:

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



LEASING & FINANCE

INVEST IN AI WITHOUT LOCKING UP CAPITAL

Establishing an AI Factory is a major long-term investment that can tie up significant cash resources for extended periods. Yet many organisations now favour OPEX-driven financial strategies to protect cash flow and maintain flexibility. That shift doesn't have to limit access to cutting-edge infrastructure.

FLEXIBLE LEASING & FINANCE OPTIONS

Boston's Leasing & Finance Solutions (BLFS) make it easier to acquire the technology you need to drive innovation and growth. With flexible leasing models and tailored payment plans, you can access next-generation platforms – including AI infrastructure clusters and full AI Factory deployments – without placing unnecessary strain on budgets.

ACCELERATE INNOVATION WITH FINANCIAL CONFIDENCE

By reducing upfront costs and removing financial barriers, Boston enables organisations to adopt advanced technology sooner, scale faster and invest with confidence.

FINANCING SOLUTIONS PROGRAMME

Access to Best-in-Class Technology and Comprehensive Support

Our BLFS Programme provides businesses with access to the latest technology and a full range of support services. This ensures that companies can stay ahead of technological advancements without the burden of large upfront investments.

- **Flexible Acquisition Models:** The programme offers various acquisition models beyond standard upfront purchases, including:
- **Enhanced Cash Flow:** FMV (Fair Market Value) lease structures reduce overall costs and improve free cash flow.
- **Increased Purchasing Power:** 100% financing options preserve working capital, allowing for greater investment capacity.
- **Optimised Asset Management:** Fixed payments and bundled services help manage risks and residual value effectively.
- **Continuous Access to New Technology:** Lease terms can be aligned with hardware refresh cycles, ensuring ongoing access to the latest technology.
- **Scalable Solutions:** Flexible leasing structures enable businesses to expand with additional lease schedules, supporting growth with predictable monthly payments

ENTERPRISE AI APPLICATIONS

CHOOSING THE RIGHT AI SOLUTION FOR YOUR BUSINESS

Deploying AI successfully is about more than selecting a model or investing in infrastructure. Organisations must identify the right applications for their objectives, data, users, and operational requirements to ensure AI delivers measurable business value.

From intelligent assistants and knowledge discovery platforms to workflow automation, decision support systems, and autonomous AI agents, today's enterprise AI ecosystem offers a wide range of solutions designed to solve different business challenges. Selecting the right approach requires careful consideration of data availability, governance requirements, integration complexity, scalability, and long-term operational goals.

Boston's AI specialists work with organisations to evaluate use cases, define AI strategies, and identify the most appropriate technologies for their environment. By combining deep infrastructure expertise with practical AI implementation experience, we help customers move from experimentation to production with confidence.

GENERATIVE AI & ASSISTANTS

Generative AI applications enable users to create content, analyse information, and interact with enterprise knowledge through natural language interfaces.

- Conversational AI and enterprise chatbots
- Content creation and summarisation
- Knowledge discovery and information retrieval
- Employee productivity and support tools

RETRIEVAL-AUGMENTED AI

Retrieval-Augmented Generation (RAG) combines AI models with organisational knowledge sources to improve accuracy, relevance, and trustworthiness.

- Enterprise search and knowledge management
- Context-aware AI responses
- Reduced hallucinations and misinformation
- Secure access to business information

AGENTIC AI & AUTOMATION

Agentic AI extends beyond conversation by enabling AI systems to reason, make decisions, and complete tasks across multiple applications and workflows.

- Process automation and orchestration
- Multi-step task execution
- Intelligent decision support
- Cross-platform workflow integration

AI DATA & KNOWLEDGE PLATFORMS

AI applications rely on fast, secure access to high-quality data. Modern AI data platforms help organisations manage, organise, and prepare information for AI-driven insights.

- Unified access to enterprise data sources
- Data governance and security
- AI-ready data preparation
- Scalable support for growing AI initiatives

ECOSYSTEM PARTNERS

KVARK **KVARK ENTERPRISE AGENTIC AI**

NVIDIA **NVIDIA AI DATA PLATFORM**

NVIDIA **NVIDIA NIM MODELS & BLUEPRINTS**

NVIDIA **RAG**

AI & MLOPS 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

Dynamic cluster management tools ensure compute resources remain available, scalable, and aligned with operational demands.

- Rapid provisioning and scaling of compute resources
- Infrastructure availability and uptime management
- Proactive monitoring, reporting and alerting Capacity planning and hardware lifecycle visibility
- Early detection of system and data centre issues

JOB SCHEDULERS & ORCHESTRATION

Advanced scheduling and orchestration platforms optimise resource allocation, improve utilisation, and streamline large-scale compute operations.

- Automated scheduling for complex compute workloads
- Resource optimisation for improved efficiency and cost control
- Pre-flight validation and job dependency management
- Checkpointing and recovery for long-running workloads
- Reduced job failures and resubmission overhead

MLOPS

Integrated MLOps frameworks simplify model deployment, management, and monitoring across production AI environments.

- Automated ML deployment and workflow orchestration
- Continuous monitoring for model reliability and performance
- Scalable integration between ML and DevOps practices
- Simplified dependency and environment management
- Increased productivity for AI and data science teams.

NETWORK PERFORMANCE & CONGESTION MONITORING

Real-time network monitoring helps maintain performance across data-intensive AI and HPC environments.

- Monitoring of network throughput and congestion
- Identification of bottlenecks impacting compute performance
- Optimisation of node-to-node and storage communication
- Improved workload reliability and efficiency for distributed AI and HPC workloads

DISTRIBUTED TRAINING & PARALLEL COMPUTING

Distributed computing frameworks enable efficient scaling across GPUs and nodes for faster training and simulation performance.

- Parallel workload distribution across compute infrastructure
- Scalable training for large AI models and datasets
- Optimised inter-node communication and synchronisation
- Improved hardware utilisation and faster time-to-insight

ECOSYSTEM PARTNERS

DKUBE

LABLUP BACKEND.AI

NVIDIA AI ENTERPRISE

NVIDIA RUN:AI

INFRASTRUCTURE MANAGEMENT FOR AI

Modern AI environments require intelligent infrastructure management to ensure systems remain efficient, scalable, and operational as workloads grow.

Boston designs infrastructure management frameworks that provide centralised visibility, automation, and orchestration across complex AI environments, enabling organisations to deploy, monitor, and scale infrastructure with confidence.

MANAGEMENT ACROSS THE AI LIFECYCLE

Efficient cluster deployment ensures AI infrastructure can be provisioned quickly and consistently.

- Automated cluster provisioning and deployment
- Standardised software and configuration management
- Scalable expansion as infrastructure grows

RESOURCE SCHEDULING & WORKLOAD ORCHESTRATION

AI workloads require intelligent scheduling to maximise efficiency across users, teams, and projects.

- Centralised scheduling across GPU clusters
- Fair allocation of resources and workloads
- Policy-based prioritisation and queue management

SYSTEM MONITORING & OPERATIONAL VISIBILITY

Continuous monitoring helps maintain system health and operational efficiency.

- Real-time monitoring across compute infrastructure
- GPU, network, and storage utilisation tracking
- Health monitoring, alerting, and reporting

AUTOMATION & OPERATIONAL EFFICIENCY

Automation reduces operational complexity and accelerates infrastructure management.

- Automated configuration and software updates
- Workflow orchestration across infrastructure
- Reduced manual administration and operational overhead

ENGINEERED FOR HIGH-PERFORMANCE ENVIRONMENTS

Infrastructure management platforms help maintain peak system efficiency.

- GPU and accelerator resource optimisation
- Balanced workload distribution across clusters
- Performance tuning for compute efficiency

SCALABLE INFRASTRUCTURE OPERATIONS

As AI environments grow, management systems must scale with them.

- Multi-node cluster management
- Support for distributed AI workloads
- Integration with AI and HPC ecosystems

UNIFIED INFRASTRUCTURE CONTROL

Modern AI environments require a single management layer across compute, networking and storage.

- Centralised management of AI clusters
- Unified monitoring across compute, network, and storage
- Consistent administration across hybrid environments

BUILT FOR RELIABILITY AND CONTROL

Infrastructure management platforms are designed to ensure stability and operational confidence.

- Secure access and policy-based controls
- High availability and operational resilience
- Governance for enterprise AI environments

ECOSYSTEM PARTNERS

NVIDIA BASE COMMAND & MISSION CONTROL

OPENHPC

SUPERMICRO CLOUD COMPOSER

TRINITYX

ACCELERATED COMPUTE PERFORMANCE

AI workloads depend on high-performance accelerators that dramatically increase compute throughput compared to traditional CPU-based systems.

- Massive parallel processing capabilities
- Dedicated AI acceleration for training and inference
- High memory bandwidth for large models and datasets
- Optimised performance for AI frameworks and libraries

ENGINEERED FOR AI WORKFLOWS

Modern AI models are often trained across multiple compute nodes simultaneously.

- Multi-node distributed training environments
- High-bandwidth GPU communication
- Efficient scaling of large training workloads
- Faster time-to-solution for complex models

INFERENCE AT SCALE

AI models must also run efficiently once deployed into production environments.

- High-throughput inference workloads
- Low-latency model execution
- Scalable deployment across enterprise infrastructure
- Support for real-time AI applications

UNIFIED AI COMPUTE INFRASTRUCTURE

To support modern AI environments, compute platforms integrate seamlessly with networking and storage infrastructure.

- Integration with high-performance networking fabrics
- Efficient access to large-scale storage environments
- Scalable compute clusters for growing AI workloads
- Consistent performance across distributed infrastructure

BUILT FOR PERFORMANCE AND SCALABILITY

AI compute infrastructure is designed to deliver reliable performance at scale.

- High-density accelerator architectures
- Efficient power and thermal design for dense compute environments
- Scalable cluster architectures for expanding workloads
- Infrastructure designed to support the next generation of AI innovation

ECOSYSTEM PARTNERS

AMD INSTINCT ACCELERATORS

NVIDIA DGX SUPERPOD

NVIDIA RTX PRO SERVERS

SUPERMICRO HGX SERVERS

AI NETWORKING ARCHITECTURE

1. DEPLOYMENT LAYER

The deployment layer defines how AI infrastructure is designed, connected, and scaled to support different workloads, users, and operational requirements.

Cloud Deployments

- Ethernet networking
- RoCEv2 GPU communication
- Scalable multi-tenant infrastructure

Hybrid Training & Inference Cloud

- NVIDIA Spectrum-X Ethernet
- NVIDIA InfiniBand interconnects
- Mixed training and inference workloads

AI Factories

- NVIDIA InfiniBand networking
- NVIDIA NVLink™ GPU interconnect
- Dedicated large-scale AI training environments

2. DATA FLOW LAYER

The data flow layer enables high-speed communication between compute, storage, and networking resources to ensure AI workloads operate efficiently at scale.

East–West Traffic (Compute Communication)

- High-bandwidth node-to-node communication
- NVIDIA Magnum IO™
- GPUDirect® RDMA
- NCCL and NVSHMEM

Supports **distributed training, simulation and GPU collaboration.**

North–South Traffic (Storage Communication)

- Storage to compute data movement
- NVIDIA GPUDirect Storage
- NVMe SNAP acceleration
- High-throughput dataset access

3. INFRASTRUCTURE FABRIC LAYER

The infrastructure fabric layer provides the networking foundation that connects compute, storage, and AI resources into a unified, high-performance environment.

Unified AI Fabric

- NVIDIA BlueField DPUs
- 100 / 200 / 400Gb InfiniBand
- RoCEv2 Ethernet fabrics
- Sub-microsecond NIC latency

Provides:

- High throughput
- Low latency
- Scalable AI infrastructure

ECOSYSTEM PARTNERS

NVIDIA ETHERNET SPECTRUM-X

NVIDIA QUANTUM INFINIBAND

SONIC ETHERNET

ULTRA ETHERNET

STORAGE

Data sits at the centre of modern AI environments, making storage performance, resilience, and scalability critical to overall system success.

Boston designs storage and data protection solutions that ensure information remains accessible, secure, and ready to support growing AI and HPC workloads. By aligning storage architecture with compute and application requirements, we help organisations maximise performance while maintaining reliability, resilience, and business continuity.

FILE STORAGE

File storage organises data within a familiar hierarchy of files and folders, making it ideal for shared access and collaborative workflows.

- Centralised file and directory management
- Network-accessible storage using standard protocols
- Scalable capacity and performance as requirements grow

BLOCK STORAGE

Block storage provides low-latency access to data, making it well suited to performance-sensitive workloads and transactional environments.

- High-performance storage for databases and virtual machines
- Low-latency access for critical applications
- Distributed architecture for resilience and availability

OBJECT STORAGE

Object storage is designed to manage large volumes of data efficiently, providing the scalability required for modern AI and cloud environments.

- Massive scalability for growing datasets
- Efficient management of unstructured data and archives
- Ideal for cloud storage and long-term retention

SOFTWARE-DEFINED & SCALE-OUT STORAGE

Software-defined storage separates hardware from software, enabling organisations to scale infrastructure efficiently and cost-effectively.

- Flexible deployment on industry-standard hardware
- Unified support for file, block, and object storage
- High scalability with simplified management

BACKUP & DATA PROTECTION

Comprehensive data protection strategies ensure information can be recovered quickly following failures, outages, or cyber incidents.

- Automated backups, snapshots, and replication
- Support for On-site and Off-site protection strategies
- Rapid recovery to maintain business continuity

ECOSYSTEM PARTNERS

BOSTON IGLOO AI+

DDN

VAST

WEKA

COOLING

As compute density rises, effective cooling is essential for sustaining performance and managing costs. Our solutions support high-power workloads with reliable thermal management across multiple deployment models. From precision air cooling to advanced liquid and immersion technologies, we help organisations reduce energy consumption, extend hardware lifespan and maximise performance from CPU- and GPU-dense infrastructure in modern high-density data centre and edge environments.

REVOLUTIONISE YOUR COOLING SYSTEM WITH BOSTON'S LIQUID COOLING SOLUTIONS

Upgrade your cooling system and enjoy unparalleled efficiency and performance. Our innovative liquid cooling solutions reduce energy costs, enhance cooling capacity and extend the lifespan of your equipment. Discover the future of cooling technology today.

SINGLE-PHASE DLC

- **Targeted component cooling:** Single-phase Direct Liquid Cooling (DLC) cools CPUs, GPUs, and memory modules directly.
- **Cold plate heat transfer:** Liquid coolant flows through cold plates attached to high-heat components to absorb thermal energy.
- **Efficient heat management:** Captured heat is transferred to a heat exchanger for removal or potential reuse.
- **Continuous liquid process:** The coolant remains in liquid form throughout the entire cooling cycle.

TWO-PHASE DLC

- **Ultra-high-performance cooling:** Two-phase DLC is designed for extremely high-density and AI-driven computing environments.
- **Low boiling-point dielectric fluid:** The system uses a specialised dielectric liquid that boils at approximately 40–50°C.
- **Phase-change heat removal:** The liquid evaporates inside the cold plate, removing heat efficiently and evenly across components.
- **Closed-loop vapor recovery:** Vapor is condensed back into liquid form and recirculated through the cooling system.

IMMERSION COOLING

- **Full-system liquid immersion:** Servers are completely submerged in a dielectric, non-conductive liquid within a sealed immersion tank.
- **Enhanced component protection:** The dielectric fluid prevents short circuits, oxidation, and exposure to airborne contaminants.
- **Highly efficient heat capture:** Immersion cooling removes heat from nearly all server components, not just concentrated hot spots. Heat recovery potential – Captured heat is transferred through a heat exchanger for cooling, reuse, or energy recovery applications.

RDHX COOLING

- **Ideal for retrofit environments:** RDHx systems improve cooling performance without requiring major infrastructure changes or full D2C deployment.
- **Targeted high-density cooling:** A specialised rear door cools hot exhaust air directly from high-density server racks.
- **Efficient heat removal:** Integrated fans and coolant coils absorb heat and return cooler air into the data hall.
- **Supports higher rack densities:** RDHx enables data centres to handle growing AI and HPC workloads more efficiently.

ECOSYSTEM PARTNERS

LM TEK
SUPERMICRO
SINGLE-PHASE DLC

ZUTACORE **TWO-PHASE DLC**

Submer
Castrol
ON
IMMERSION COOLING

SUPERMICRO **RDHX COOLING**

HOSTING & DATA CENTRE SERVICES

BUILT FOR ORGANISATIONS THAT NEED MORE THAN SPACE

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.

CAPACITY READY FOR HIGH-DENSITY AI

Across the globe, high-density data centre capacity is becoming increasingly difficult to secure due to the explosive growth of AI deployments. Boston Limited has partnered with select data centres to reserve capacity and support new deployments, with availability to install infrastructure in as little as one month.

SUSTAINABLE BY DESIGN

Sustainability is at the core of our partner data centre services, with renewable green energy sources and low PUE options available across a diverse global portfolio.

BUILT FOR AI FROM DAY ONE

Our facilities are designed specifically for AI deployments, with Tier 3 and Tier 4 capabilities for maximum uptime and resilience. They include high-density power configurations, air, hybrid and liquid cooling, high-speed IP and dark fibre connectivity to global PoPs, and round-the-clock security.

COOLING FOR THE NEXT GENERATION OF ACCELERATORS

Advanced liquid cooling technologies from our long-standing partners enable rack densities of up to 120kW, while air-cooled options support deployments up to 75kW with 100% heat removal.

HOSTING THAT SCALES WITH YOUR WORKLOADS

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


site

BOSTON TRAINING ACADEMY

START YOUR JOURNEY, TODAY

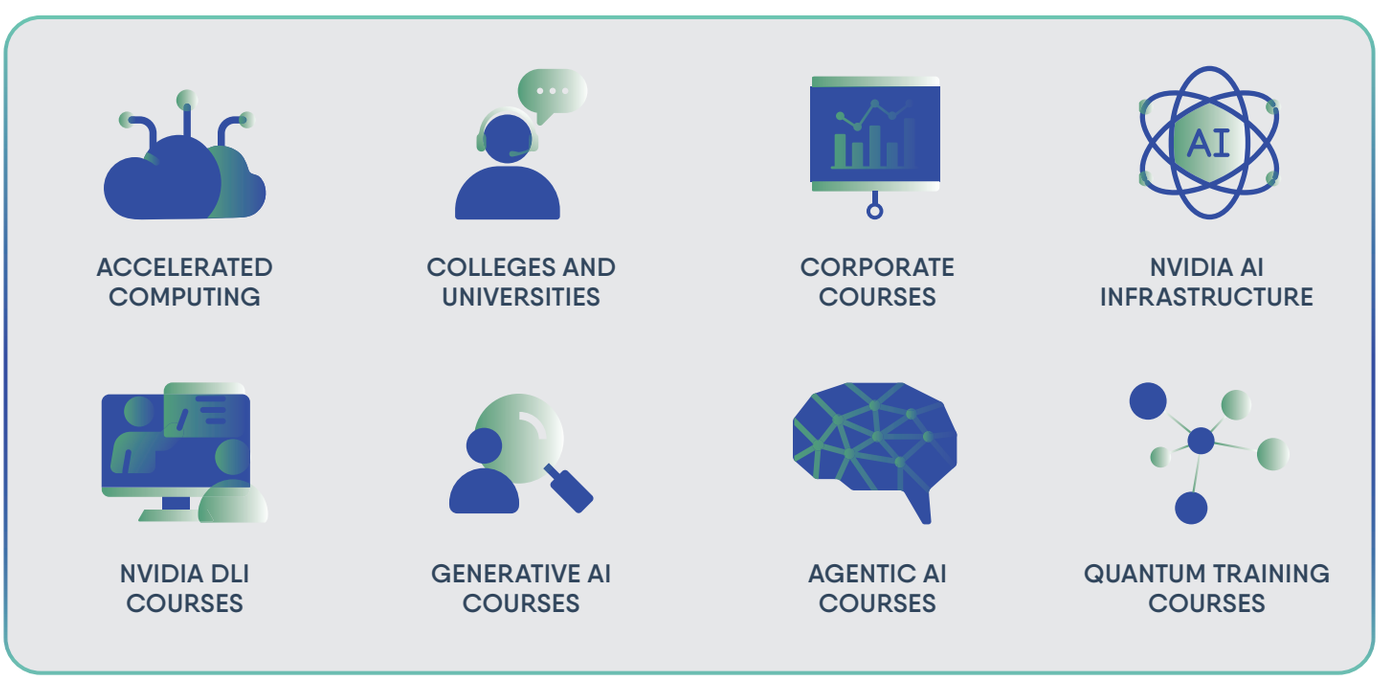
Boston Training Academy (BTA) is the learning and development division of Boston Limited, helping organisations build the technical capabilities needed to succeed in today's data-driven, AI-powered world.

Through our corporate training programmes and academic partnerships, we deliver industry-recognised certifications and hands-on learning across Artificial Intelligence, Machine Learning, Data Science, Quantum Computing and other advanced technologies.

Our programmes are developed in collaboration with leading technology providers, including NVIDIA, and delivered by experienced instructors with real-world expertise. This ensures teams gain practical, applicable skills that can be implemented immediately within the workplace.

By investing in technical education and upskilling, businesses can reduce project costs, accelerate delivery times and improve operational efficiency. Better-equipped teams are able to handle more complex workloads internally, make stronger technology decisions and maximise the value of the tools and infrastructure already available to them.

Whether you are looking to future-proof your workforce, increase productivity, support digital transformation or develop specialist expertise in emerging technologies, BTA provides tailored training solutions designed to meet the needs of modern enterprises.



ELEVATE YOUR TEAM
TRANSFORM YOUR BUSINESS

Boston Training Academy provides targeted training designed to help you build the skills that matter most.



EXPLORE OUR

TRAINING COURSES

BOSTON
Servers | Storage | Solutions
A DIVISION OF SOURCECODE

NVIDIA

VAST

THE GENETIC CODE:
NATURE'S ULTIMATE DATASET

Beneath every forest canopy, coral reef and microscopic cell lies a language billions of years in the making, the code of life itself. Within that code are the blueprints for new medicines, sustainable materials and solutions to challenges that define our future.

By mapping the world's biodiversity into an ever-expanding knowledge graph of life, Basecamp Research is revealing how nature's molecular patterns connect across species, environments and continents. Each sample gathered from the field becomes a data point in an ecosystem of possibility, a bridge between natural discovery and human innovation.

SCIENTIFIC AMBITION BEYOND INFRASTRUCTURE

Basecamp Research's biodiversity knowledge graph is expanding at extraordinary speed, with billions of genomic connections fueling AI models that link nature's diversity to human progress. Each breakthrough broadens both insight and computational demand. Every genome sequenced adds vast new data, compounding computational complexity.

The public cloud, once ideal for experimentation, now shows its constraints. As data volumes surge into petabytes, performance becomes tied to network latency and bandwidth. Costs drift unpredictably; data mobility becomes restricted and maintaining compliance across regions adds operational overhead.

These challenges are the natural cost of scaling scientific discovery. High-throughput biological workloads demand GPUs, rapid data access and storage performance that can scale uncompromisingly.

Within this expanding scientific foundation sits EDEN, Basecamp Research's AI model family trained on one of the largest evolutionary datasets ever assembled. By learning from billions of years of biological evolution across millions of species, EDEN identifies underlying genetic design principles that can be translated into real-world applications.



Recent developments in EDEN focus on using this evolutionary intelligence to design and test programmable gene edits, including early demonstrations of precise gene insertion mechanisms that go beyond traditional CRISPR-style editing. While still in early, pre-clinical stages, these models have already shown promise in generating candidate therapies for areas such as cancer and antimicrobial resistance.

As EDEN scales alongside Basecamp's rapidly expanding "trillion gene" biodiversity dataset, it is becoming a central driver of computational demand, requiring high-performance AI infrastructure to process complex genomic relationships at unprecedented scale.

WHERE SCIENTIFIC VISION MEETS TECHNOLOGICAL PRECISION

To continue pushing the boundaries of biotechnology, Basecamp needs more than cloud compute; they need an infrastructure capable of processing nature's complexity at speed and scale.

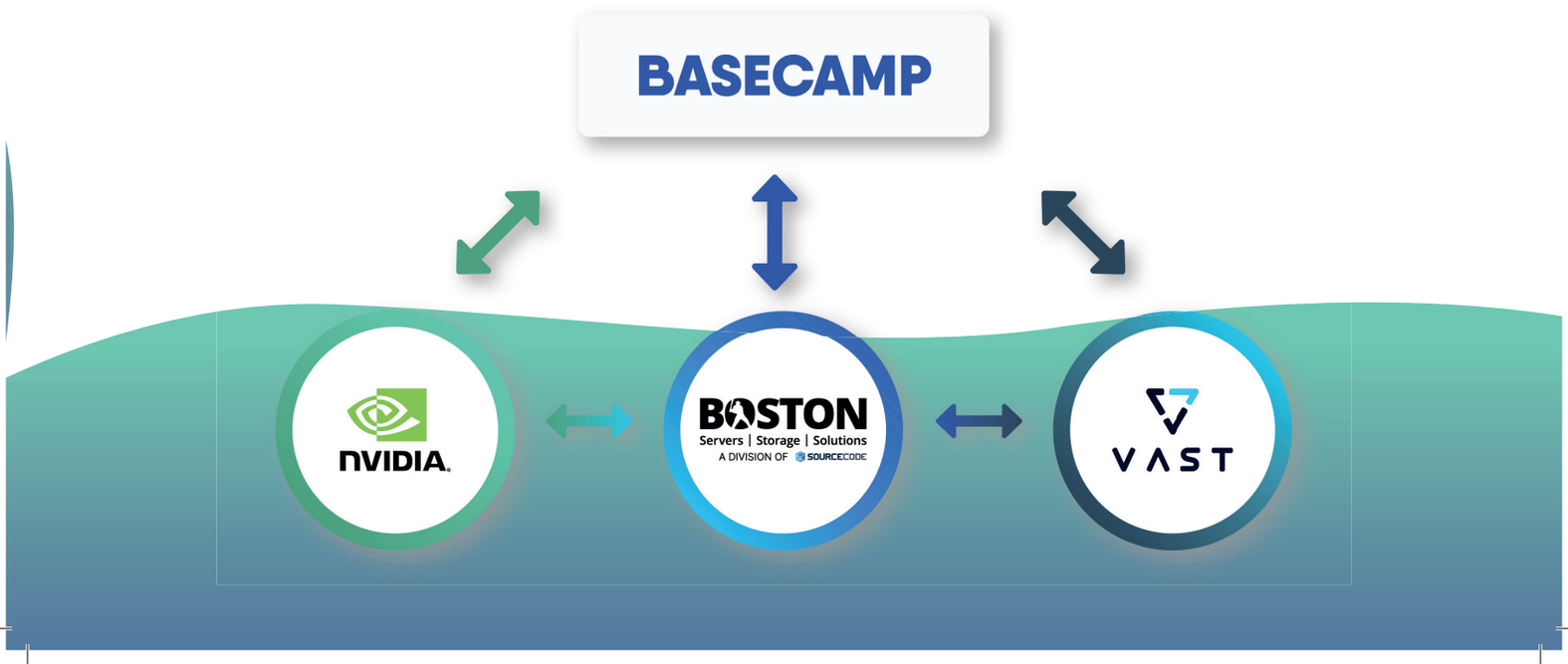
That search led to a collaboration between Basecamp Research, Boston Limited, NVIDIA, and VAST Data, each partner contributing a unique strength.

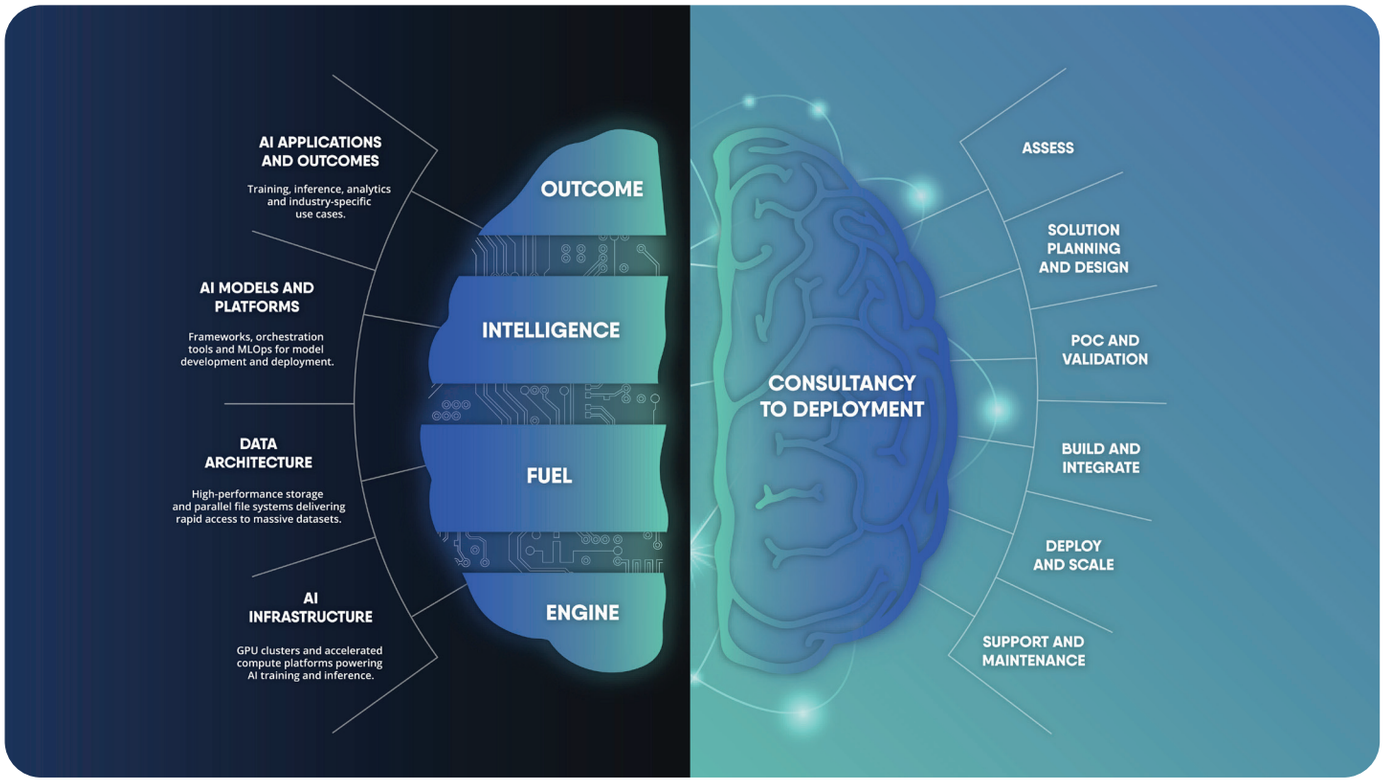
- Basecamp defines the scientific frontier, shaping workflows, data flows and performance goals.
- Boston translates those needs into architecture, designing and operating the backbone of high-performance infrastructure.
- NVIDIA provides the GPU-accelerated compute required to train and deploy AI models that map the relationships of life itself.
- VAST delivers the data platform that keeps insights flowing; fast, secure and effortlessly scalable.

What emerges is an integrated ecosystem where research, infrastructure and innovation operate in concert.

A BOSTON ENGINEERED AI FACTORY POWERED BY NVIDIA AI ENTERPRISE™, VAST DATA AND FULL-STACK SOLUTIONS

The solution takes form as a fully integrated AI Factory, an intelligent ecosystem where data, compute and algorithms operate in a continuous loop. It transforms Basecamp's biodiversity datasets into actionable insight, accelerating the path from raw sequence to discovery.





At its core runs NVIDIA AI Enterprise™, the full stack software platform powering every stage of Basecamp's AI lifecycle, from data preparation and training to deployment and scaling. Built on NVIDIA GPUs and the CUDA® parallel computing platform, it delivers the high performance foundation required to process billions of genomic relationships simultaneously.

For genomic workflows, NVIDIA Parabricks™ accelerates secondary sequencing analysis-alignment, variant calling and joint genotyping-reducing multi-day CPU workloads to minutes. NVIDIA Clara™ adds a domain-optimised framework for computational biology, enabling multimodal data analysis, biomarker discovery and integrated imaging-genomic insight.

To support this demanding environment, Boston Limited deployed a NVIDIA DGX H200 BasePOD™, integrated with NVIDIA InfiniBand/Ethernet networking and powered by AMD EPYC™ CPUs. This high-density infrastructure provides the throughput and flexibility required for modern genomic research and AI training at scale.

Beneath the compute layer, a VAST Data architecture delivers the high performance, scale out storage fabric that feeds the AI Factory. Its all flash design ensures extreme throughput and parallel access, allowing petabytes of genomic data to flow seamlessly between storage and GPUs. A high-throughput networking and orchestration layer binds the system together, dynamically allocating resources as workloads evolve.

Boston integrates all the technologies with the requirements of Basecamp to provide a full-stack, consolidated and bespoke solution that ensures continuous optimisation, monitoring and reliability across the stack.

RESULTS & IMPACT: ACCELERATING DISCOVERY

The result is more than infrastructure, it is a production system for scientific intelligence. A living AI Factory that continuously converts biodiversity data into discovery.

Unified GPU-accelerated infrastructure and adaptive orchestration make every stage of research faster, leaner and more precise. Complexity gives way to clarity. Data flows freely. Models iterate continuously. Insights emerge in real time.

THIS DELIVERS:

- Predictable performance
- Scalable growth
- A stable operational model that evolves with demand
- Lower overhead through managed optimisation
- Greater scientific freedom for Basecamp's researchers

The AI Factory becomes a catalyst for discovery, a platform where technology finally moves at the speed of science.

EXPLORE BOSTON LABS: TEST, VALIDATE AND SCALE YOUR IDEAS

Through Boston Labs, organisations can experience the future firsthand: testing proof-of-concept architectures, validating real workloads and exploring how AI, HPC and data platforms perform together.

From early exploration to full-scale deployment, Boston Labs provides the expertise, environment and technology to turn innovation into measurable outcomes.

Discover how Boston is redefining the infrastructure of discovery.

Visit Boston Labs to design, test and prove your next breakthrough.



OUR GLOBAL 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.

OUR ELITE PARTNERS

AMD Partner Network PLATINUM

NVIDIA Elite Partner

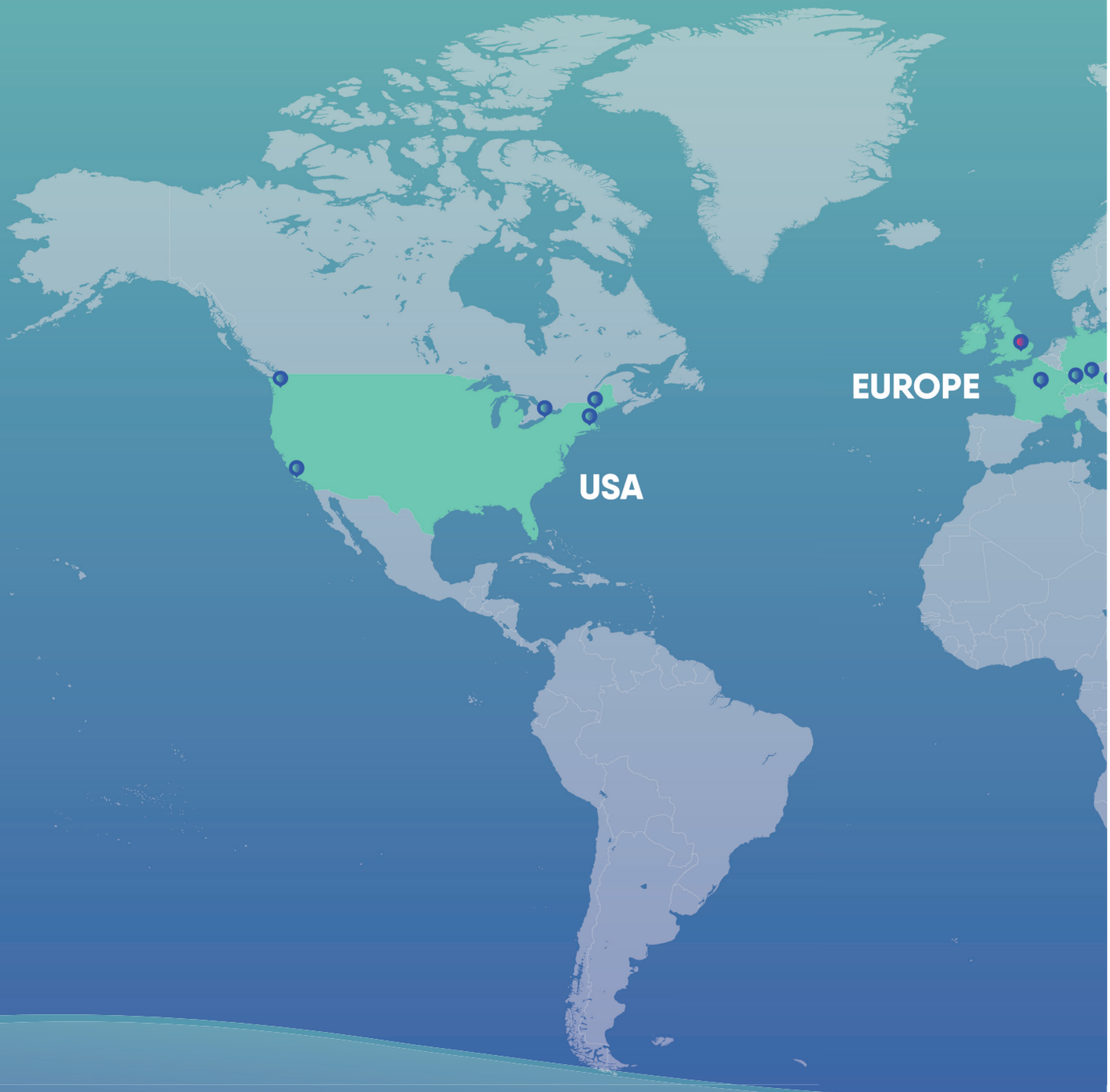
intel partner

SUPERMICRO Authorized Partner




2crsi
10ZiG
AIC
AiGency Global
AMD
AMPERE.
ASRock Rack
antillion
APC by Schneider Electric
ASUS
Asperitas
BROADCOM
Castrol CORE
CLOUDIAN
Cloudalize
ClusterVision
CLYSO
COMARK A DIVISION OF SOURCECODE
COMBACK HOCHSICHERHEIT FÜR IHRE IT
CyberPower
Data Center Group
DATACORE
Oddn
DKube
EGZAKTA GROUP
FORMERICA OE
GIGABYTE
GOLDBECK
Graid
GRAPHCORE
HUAWEI
IBM
intel.
KAYTUS
KIOXIA
KVARK
labiup
LM TEK
lustre
micron
Microsoft
MICROCHIP
MULTIVERSE COMPUTING
NetApp
NEXAVM
NUTANIX
NVIDIA.
open-e
OSNEXUS
OSS ONE STOP SYSTEMS
PEAK AIO
PioVation
PNY
ProLabs an Amphenol company
RADIANARC
radware
SAMSUNG
SCALE COMPUTING
Life Is On Schneider Electric
SEAGATE
SECQAI
SHIPSHAPE DATA
SOLIDIGM.
Submer
SUPERMICRO
STORViX
TOSHIBA
TRANQUILIT
Unframe
VAST
VDURA
Verysell AI Your Trusted AI Partner
Virtuozzo
WEKA
WD.
ZUTACORE

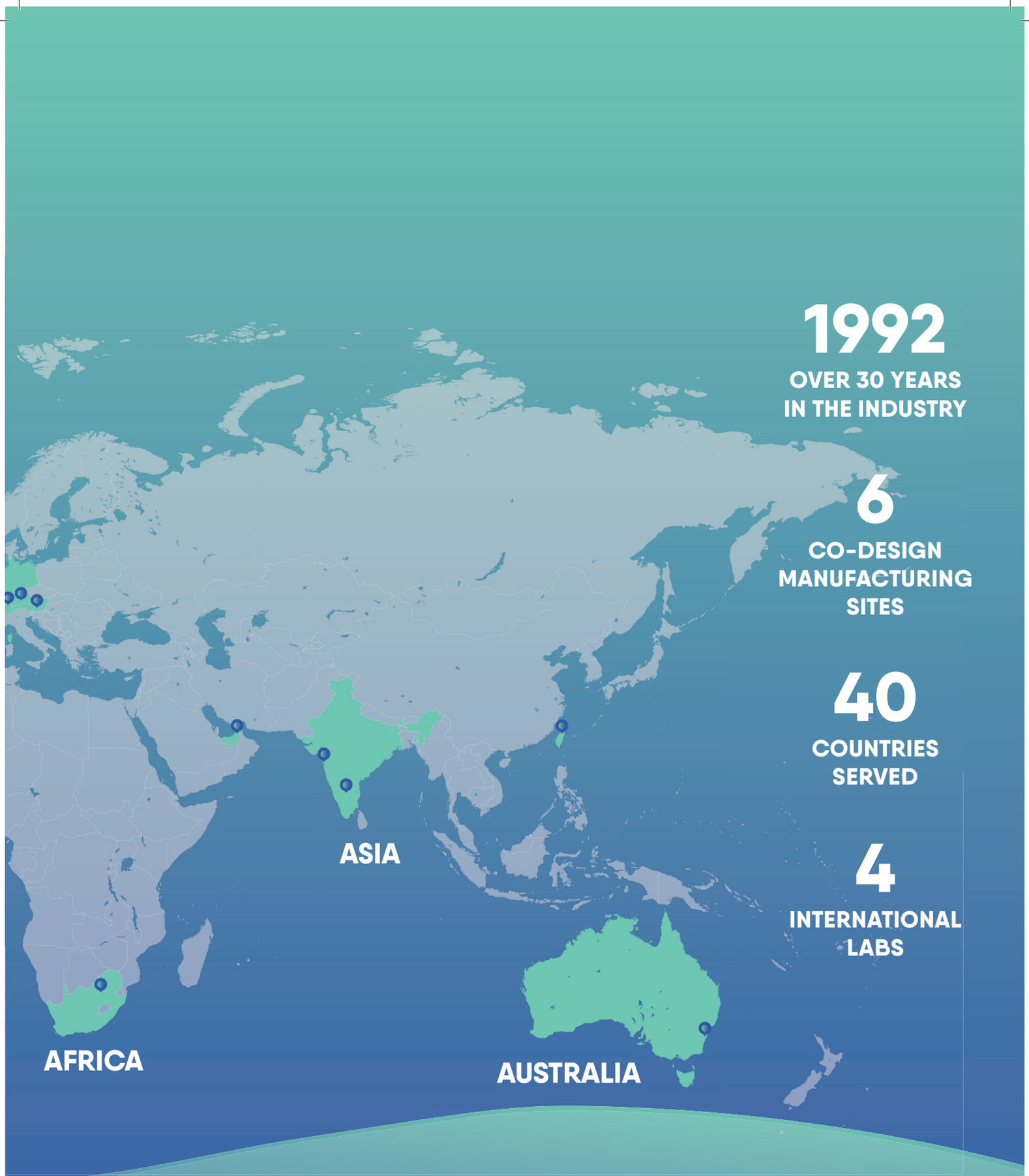
BOSTON AROUND THE WORLD



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

E: sales@boston.co.uk | **T:** (+27) 11 014 1924 | **W:** boston.co.uk

Facebook: facebook.com/bostonlimited
Instagram: instagram.com/boston_limited
LinkedIn: linkedin.com/company/boston-it-solutions-africa
Youtube: youtube.com/@BostonUk



1992
OVER 30 YEARS
IN THE INDUSTRY

6
CO-DESIGN
MANUFACTURING
SITES

40
COUNTRIES
SERVED

4
INTERNATIONAL
LABS

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

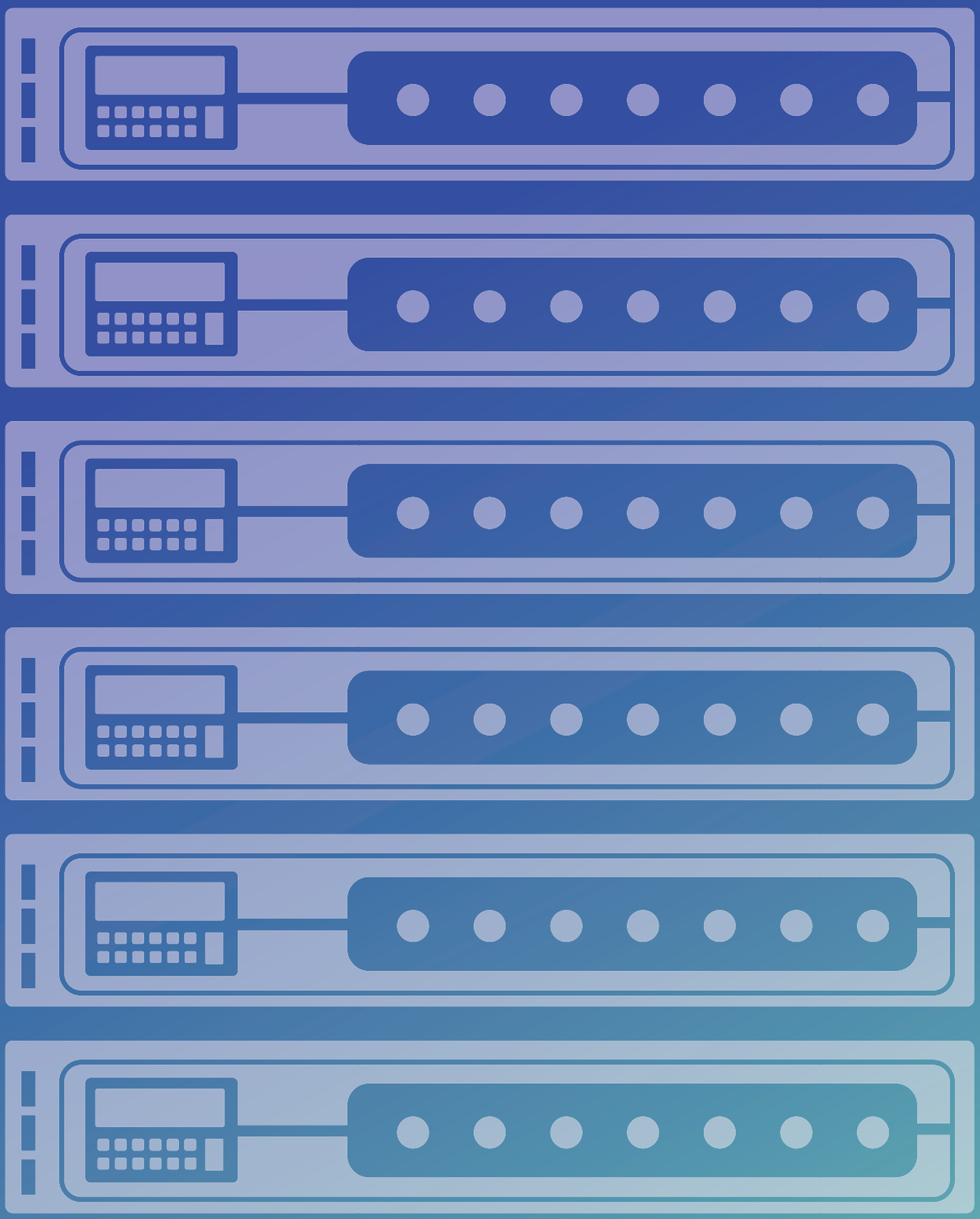
AUSTRALIA

SYDNEY

BOSTON

Servers | Storage | Solutions

A DIVISION OF SOURCECODE



sales@boston.co.uk | www.boston-it.co.za | (+27) 11 014 1924 | Wanderers Office Park, 52 Corlett Drive, Illovo, Gauteng, 2196

Austria | Australia | France | Germany | India | Taiwan | **South Africa** | Switzerland | UK | USA

October - 2026