BOSTON
Servers | Storage | Solutions

HIGH PERFORMANCE COMPUTING

HPC

CONTENTS

WHAT'S INSIDE?

Boston Limited your trusted partner 05
HPC cluster architecture and workflow 06
Compute Solutions 08
- Supermicro Compute Solutions for HPC 09
- AMD 4th gen EPYC™ 9004 series processors 10
- AmpereOne 11
- Intel® 6th gen Xeon scalable processors 12
- NVIDIA Grace Blackwell 13

Accelerated Solutions 14
- AMD Instinct™ MI300A series platform 15
- AMD Instinct™ MI300X series platform 16
- Intel® Habana® Gaudi2® 17
- NVIDIA® NVL72 SuperPod 18

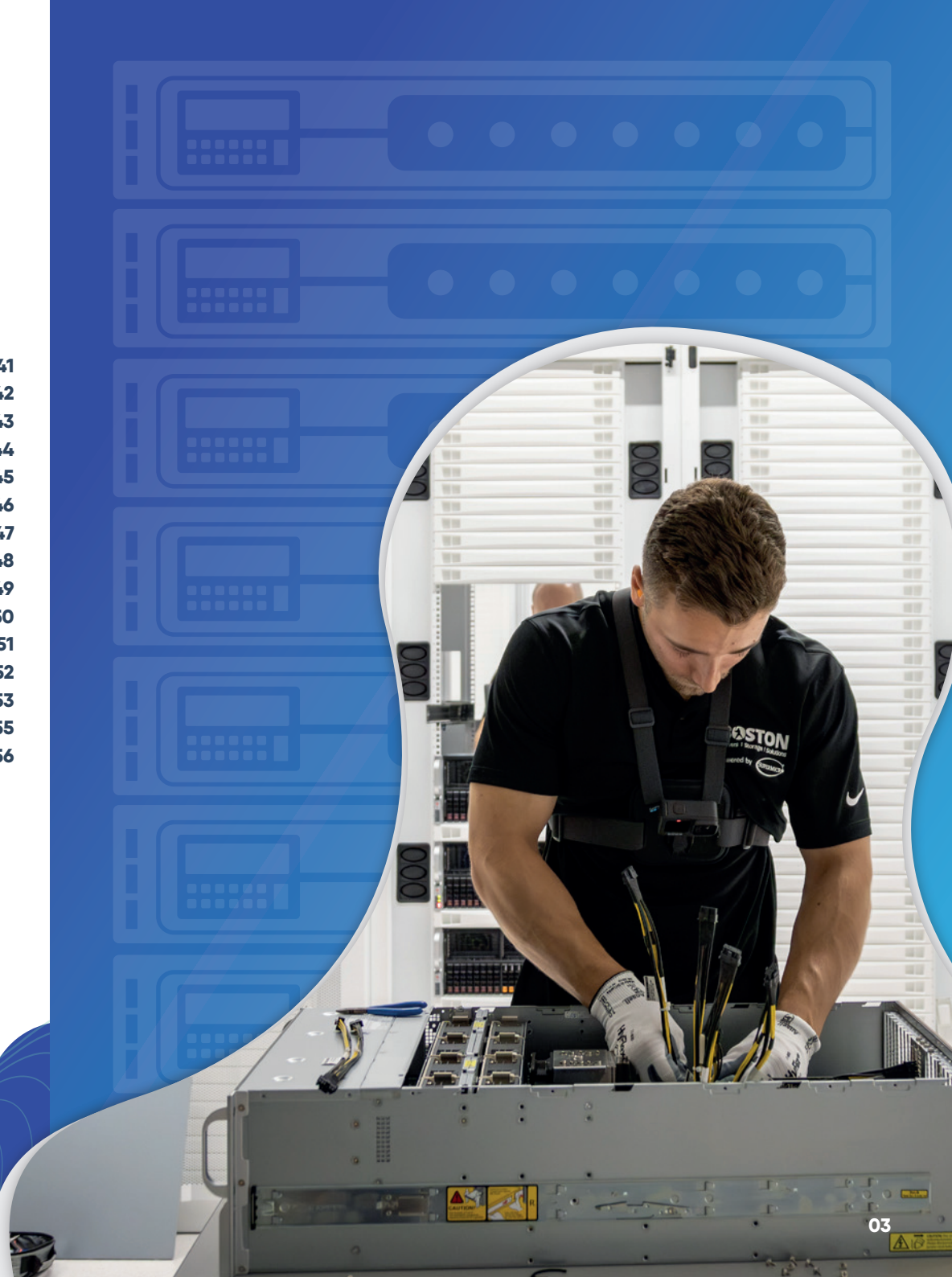
Storage solutions 19
- Boston Igloo AI+ 20
- DDN Exascaler 22
- IBM® Spectrum Storage™ 24
- Weka IO 26

Liquid Cooling 28
- Types of Liquid Cooling 28
- Asperitas 33
- Castrol 34
- The GlacierCore Series 36
- Submer 38
- Supermicro 39
- Zutacore 40

Networking Solutions 41
- NVIDIA Ethernet Spectrum®-4 42
- NVIDIA Infiniband Quantum™-2 43
- NVIDIA Spectrum™-X 44

Software 45
- DKube 46
- NVIDIA Base Command Manager 47
- Warewulf 48

Boston Services 49
- Boston Technology Consulting 50
- Boston Training Academy 51
- Boston Labs 52
- Professional Services 53
- Support and onsite warranty 55
- Co-Design in HPC 56



WHAT SETS US APART?

INNOVATIVE SOLUTIONS

We innovate tech solutions for diverse sectors, adapting to evolving business needs

BESPOKE CUSTOMISATION

Tailored, bespoke solutions harmonise with customer objectives, defining our excellence in technological delivery

GLOBAL REACH

Our global presence assures swift support, regardless of location, backed by a history of successful collaborations

TRACK RECORD

A proven track record with diverse projects instils confidence in us as a dependable partner

MANUFACTURING AND TESTING

Our in-house design, manufacturing and testing in the Boston Labs guarantees quality and quick turnaround

RESEARCH AND DEVELOPMENT

Ongoing R&D equips us with expertise for the latest tech advancements, providing customers a competitive edge

SUSTAINABILITY

We are environmentally conscious and prioritise eco-friendly, sustainable solutions

ONE-STOP SHOP

Our diverse tech offerings, from computing to AI, position us as a one-stop shop for all your needs

CUSTOMER-CENTRIC

Customer-centricity is our cornerstone, understanding challenges and exceeding expectations in the solutions we create

BOSTON LIMITED
YOUR TRUSTED PARTNER

WHO ARE WE?

BOSTON: THE POWER OF HIGH-PERFORMANCE

Boston Limited is a leading provider of high-performance, power-optimised technologies for various markets. We offer bespoke award-winning solutions that are designed to help our customers stay ahead of the curve.

Our products include servers, storage, workstations and clustered solutions. We also provide a range of value-added services, such as global on-site maintenance packages and life cycle management for IT equipment.

What sets Boston apart is our expertise in integrating, testing and validating high-performance custom solutions. We can fully brand and package our solutions to meet the specific needs of our customers.

We also have extensive experience in AI and data. We offer end-to-end services, from training and advisory services to implementation, to help organisations navigate the complexities of digitalisation successfully.

Our expertise in AI and ML allows us to provide solutions that automate processes, optimise operations and provide valuable insights from data.

Our broad range of capabilities, commitment to customised solutions and value-added services have made us a trusted partner for organisations in various industries and markets worldwide. We recently celebrated our 30th anniversary, and we are expanding to meet the growing demand for high-performance systems.

If you are looking for a high-performance technology partner that can help you stay ahead of the curve, then Boston is the right choice for you.

HPC CLUSTER ARCHITECTURE AND WORKFLOW

When it comes to the High Performance Computing (HPC) market, clusters are rapidly reshaping it, driven in the most part by price/performance and the open source Linux OS.

Even though there are endless opportunities that clusters offer, and clusters, comprised of commodity server hardware and software are gaining acceptance...getting a cluster running, and learning how to use it, requires skilled resources, services and time.

If not installed correctly, this can lead to badly planned clusters with software that hasn't been chosen appropriately.

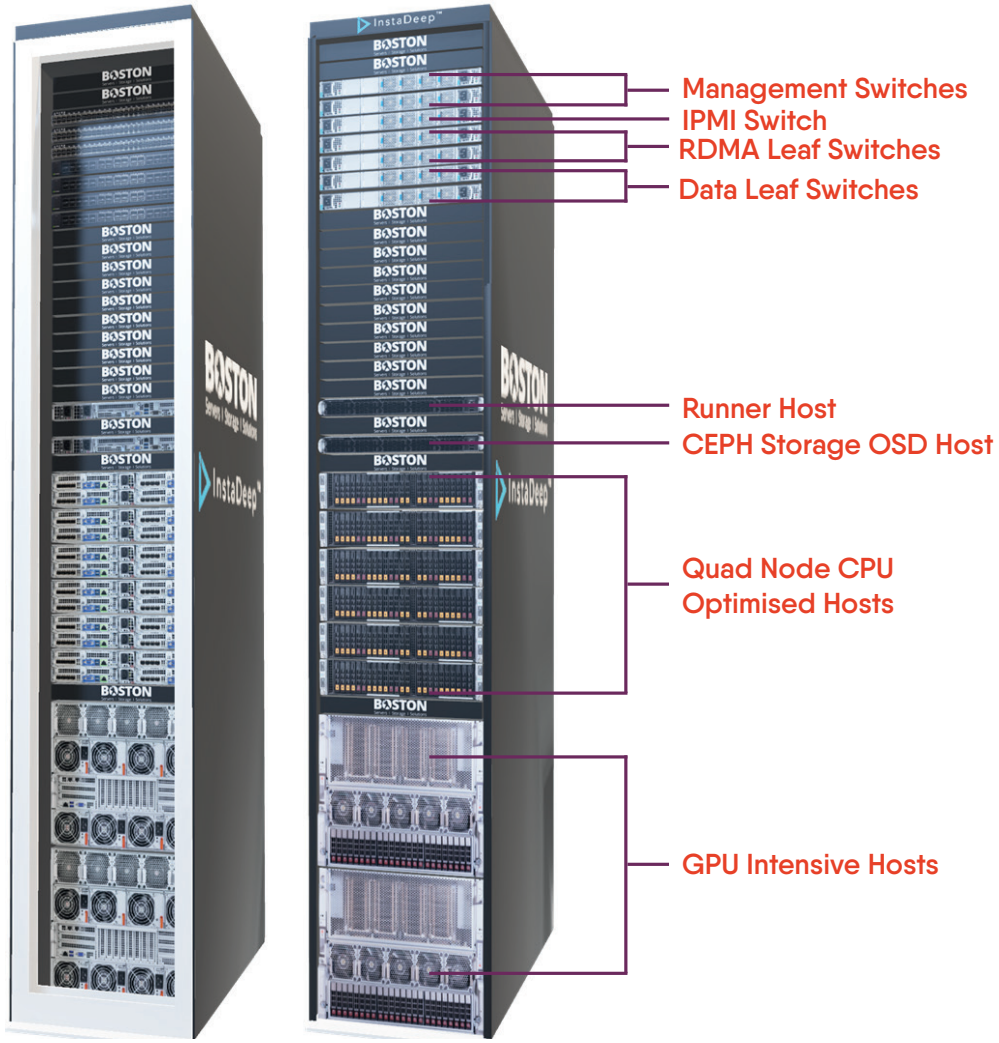
The result? IT departments turn to costly SMP alternatives because of the standardisation shortcomings of cluster computing.

> Boston understand the end-to-end process of building, implementing and managing clusters, so cost is minimised, as is the time to get a cluster fully operational.

Boston offers a range of both open source and commercial cluster management packages, that are selected in-line with our clients environment and workflow.

We are able to pre-engineer and test multiple variations of hardware and software packages in our Boston Lab facilities - making our Linux compute clusters easy to deploy, simple to use, consistent, transparent, turnkey and available.

WORKER RACK



COMPUTE SOLUTIONS

SUPERMICRO

Supermicro's compute solutions offer high-performance server systems, including rackmount and blade servers, for scalability, reliability and energy efficiency in datacentres and cloud computing.

AMD

AMD's compute solutions feature powerful Ryzen processors for gaming and content creation, as well as EPYC processors for datacentres and enterprise workloads.

AMPERE

Ampere empowers enterprise computing with their high-performance AmpereOne processors, offering exceptional core counts, clock speeds and memory capacity.

INTEL

Intel's compute solutions include Core processors for a balance of performance and power efficiency, and Xeon processors for enterprise computing with scalability and advanced security.

NVIDIA

NVIDIA's Grace Blackwell redefines large-scale high-performance computing and AI deployments, offering unrivalled computational capabilities.

SUPERMICRO COMPUTE SOLUTIONS FOR HPC

Supermicro and Boston have been providing HPC solutions for the last 30 years. Much like us, Supermicro is incredibly versatile with their range of systems and their applications. Their solutions are comprised of a wide range of building blocks to meet customers' specific needs.

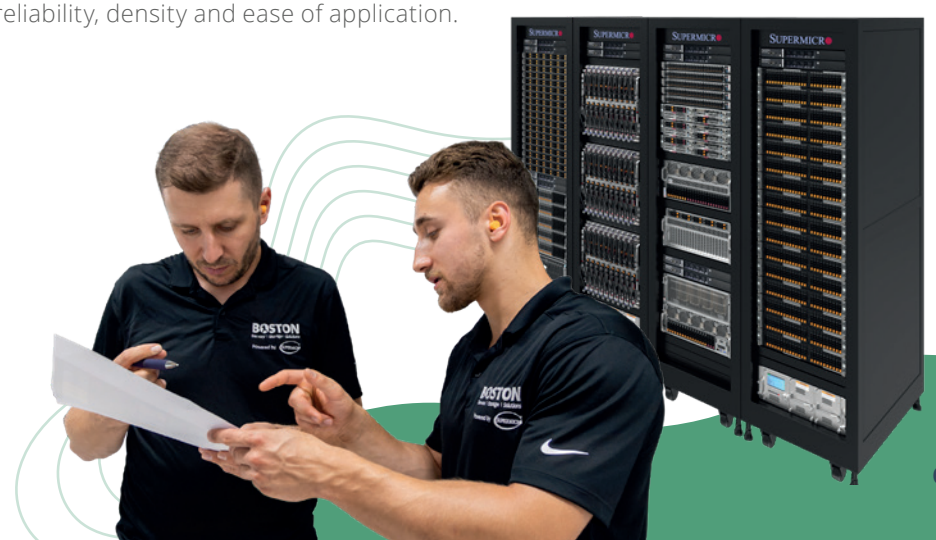
Reference designs already exist from Supermicro for enterprise HPC, scientific research, AI and deep learning – all of which can be tailored to the finest detail. If a system is certified by Supermicro, you can be sure that its components are optimised for the highest performance and the best power-efficiency.

Supermicro systems can be used in virtually any vertical. A significant portion of datacentres around the world use Supermicro systems due to their reliability, density and ease of application.

Additionally, Supermicro also provides an extremely competitive density to price ratio. These benefits have been leveraged in CFD, FEA, weather prediction, oil and gas research, automotive, aerospace, civil engineering, hardware design and naval engineering to name a few.

Finally, HPC requires vast amounts of compute power which in turn requires a significant amount of energy. This is so the workload is finished in a reasonable time span while minimising the power necessary to reach its goal.

Supermicro promotes greener computing by providing highly optimised and power efficient systems and configurations. This not only lowers TCO but also works toward sustainability goals.



AMD 4TH GEN EPYC™ 9004 SERIES PROCESSORS

The processor that will put the wind in your sails has been announced! First, AMD took you to Naples, then Rome, then Milan. Next stop: Genoa, the codename for AMD's 4th generation EPYC™ 9004 series CPU.

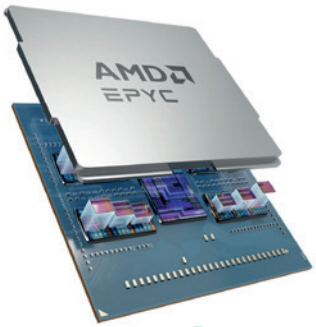
The EPYC™ family of CPUs provides great performance and security while maintaining power and cost-efficiency. AMD's 4th generation EPYC™ 9004 series CPU keeps this trend alive by doubling down on what made the previous EPYC™ CPUs so iconic.

The EPYC™ 9004 series CPU uses DDR5 memory, with 50% more channels providing 2.25 times the memory bandwidth of the previous generation.

The chip comes with twice the IO capacity and PCIe®5 making it a cornerstone of AMD's architectural leadership. All this culminates in lower TCO over three years by around 29%.

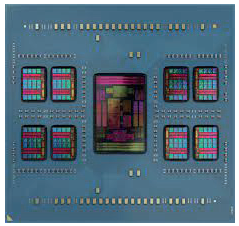
KEY FEATURES:

- Leading socket and per-core performance
  – Up to 96 "Zen 4" cores in 5nm
- Leading memory bandwidth & capacity
  – 12 channels DDR5 with up to 12TB of memory capacity



BENEFITS:

- Security features at the silicon level
- Lower TCO and energy consumption
- Excellent price-performance
- Previous AMD EPYC™ family CPUs have provided a massive upgrade for numerous businesses, and the AMD 4th generation EPYC™ 9004 series processors carries on this tradition



AMPEREONE

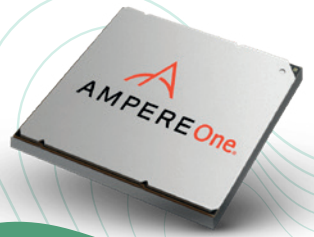
Ampere is at the forefront of enterprise computing, redefining performance with their innovative AmpereOne processors. These state-of-the-art processors have a single socket configuration with the potential for up to 192 cores, delivering exceptional clock speeds of up to 3.7 GHz and supporting from 250W to 400W Thermal Design Power (TDP), depending on the model. The AmpereOne family processors provide an unprecedented 16 DIMM slots for memory capacity, ensuring your workloads run seamlessly and efficiently.

KEY FEATURES:

- Up to 192 cores, providing exceptional processing power
- Clock speeds of up to 3.7 GHz for swift and responsive computing
- Choose from 250W to 400W TDP for tailored performance
- Support for 16 DIMM slots ensures efficient memory management

BENEFITS:

- Ideal for enterprise, high-performance computing and Telco Edge solutions
- Tailor the power consumption to your specific computing needs
- Robust memory capacity ensures seamless workload handling



INTEL® 6TH GEN XEON SCALABLE PROCESSORS

The Intel Xeon 6 Processor Family is engineered to address the complex needs of modern datacentres, offering a blend of performance, efficiency, and scalability. With its innovative hybrid architecture, it incorporates both Performance Cores (P-cores) for demanding compute tasks and Efficient Cores (E-cores) for energy-efficient processing. This makes it versatile across various workloads, including AI, high-performance computing (HPC), and cloud-native applications.

The processors also feature advanced memory bandwidth and security technologies such as Intel® SGX and Intel® TDX, ensuring robust protection of sensitive data. Additionally, their energy-efficient design reduces operational costs while maintaining high performance, making them an ideal solution for organisations aiming to optimise their total cost of ownership (TCO) and environmental impact.

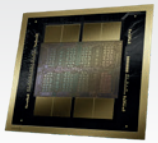
KEY FEATURES:

- Hybrid Architecture: Combines Performance-cores (P-cores) for intensive tasks and Efficient-cores (E-cores) for energy efficiency
- Intel® AMX: Enhances AI performance with advanced matrix operations
- Intel® AVX-512: Boosts vector processing, ideal for AI and HPC workloads
- MCR DIMMs: Provides 37% more memory bandwidth than standard DDR5
- Up to 288 Cores: High core count for ultra-density compute performance

BENEFITS:

- Scalability: Seamlessly scales across various infrastructure sizes, from small servers to large datacentres
- AI Optimization: Integrated AI acceleration boosts performance in machine learning and AI workloads
- Energy Efficiency: Designed to deliver high performance with reduced power consumption, supporting sustainability goals

NVIDIA GRACE BLACKWELL



Grace Blackwell serves as a cornerstone in NVIDIA's advanced technology lineup, bridging CPU, GPU and storage innovations to revolutionise computing and AI. What distinguishes Grace Blackwell is its remarkable ability to access up to 200TB of unified memory, achieved by integrating 300 Superchips.

This design circumvents the limitations of conventional networking and distributed memory systems, offering a new paradigm in efficient, large-scale computing. Grace Blackwell reimagines data processing, ensuring seamless and powerful performance for diverse computing and AI applications. With its unparalleled memory bandwidth and capacity, the Grace Blackwell Superchip enhances datacentre operations, leading to quicker insights and more agile decision-making. This cutting-edge solution marks a significant advancement in computational technology, setting new benchmarks for performance and flexibility across various industries.

KEY FEATURES:

- Vast memory capacity and bandwidth for rapid access to large datasets
- Enables seamless scalability across different workloads, from AI to scientific research
- Delivers a unified platform for AI and data-heavy tasks
- Optimises datacentre performance, streamlining operations

BENEFITS:

- Enhances computing efficiency
- Combines CPU and GPU power for integrated AI functionalities, reducing workflow inefficiencies
- Adapts to a wide range of workloads and scales effortlessly
- Boosts efficiency, facilitating quicker insights and enabling responsive, real-time decisions

ACCELERATED SOLUTIONS

AMD

AMD's MI300A is a versatile APU that combines CPU and GPU capabilities on a single chip, targeting high-performance computing and AI workloads with an emphasis on energy efficiency.

AMD

AMD's MI300X is a datacentre GPU designed for high-performance AI and machine learning workloads, offering advanced memory bandwidth and compute capabilities optimised for LLMs and AI.

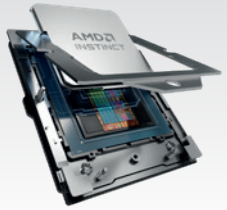
INTEL

Born for Deep Learning. Raised to a whole new level. Supermicro and Intel have partnered to bring a next-generation AI deep learning server to market.

NVIDIA

NVIDIA NVL72 SuperPod is a cutting-edge, powerful computing infrastructure that harnesses the immense processing capabilities of NVIDIA GPUs to accelerate artificial intelligence and high-performance computing workloads.

AMD INSTINCT™ MI300A SERIES PLATFORM



The AMD Instinct™ MI300A is a powerful datacentre GPU designed to handle demanding server and supercomputer workloads, including cutting-edge AI, machine learning and large-scale simulations.

Built with advanced CDNA3 architecture and leveraging 5nm process technology, the MI300A provides exceptional performance for a variety of intensive computing tasks. With an impressive 16,000 stream processors and 256 compute units, the MI300A delivers up to 450 TFLOPs of mixed-precision FP16 performance for deep learning training. It also achieves up to 60 TFLOPs of double precision (FP64) performance and 120 TFLOPs of single precision matrix (FP32) performance, making it a versatile solution for both AI and HPC applications.

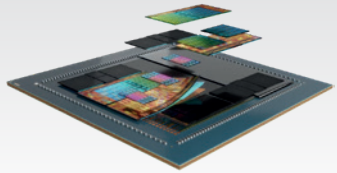
KEY FEATURES:

- The MI300A features 16,000 stream processors and 256 compute units, making it ideal for a range of workloads, from data analytics to scientific simulations
- With 160 GB of HBM3 memory, it offers significant capacity for memory-intensive applications, ensuring faster data access and reduced latency
- Built with enhanced security measures, including Secure Memory Encryption and AMD Infinity Guard, to protect sensitive data and AI models

BENEFITS:

- Capable of handling a wide variety of workloads, from AI and machine learning to traditional HPC tasks, providing flexibility in deployment
- Offers robust security features to safeguard intellectual property and sensitive data, crucial for industries requiring high levels of data protection
- Designed to deliver superior performance per watt, reducing operational costs and environmental impact, ideal for sustainable datacentre operations
- High memory capacity and Infinity Fabric Link technology minimise latency, enabling quicker insights and faster decision-making processes

AMD INSTINCT™ MI300X SERIES PLATFORM



The AMD Instinct™ MI300X is an advanced datacentre GPU specifically designed for the most demanding AI, deep learning and HPC applications. Incorporating the latest CDNA3 architecture and built on a cutting-edge 4nm process, the MI300X pushes the boundaries of GPU performance and efficiency.

With a staggering 18,000 stream processors and 300 compute units, the MI300X offers up to 500 TFLOPs of mixed-precision FP16 performance, making it ideal for deep learning training and AI model inference. The GPU also delivers up to 70 TFLOPs of double precision (FP64) performance and 140 TFLOPs of single precision matrix (FP32) performance, providing versatility and power for a wide range of computational workloads.

KEY FEATURES:

- With 18,000 stream processors and 300 compute units, the MI300X is designed for extreme performance, suitable for exascale computing and large AI models
- Equipped with 192 GB of HBM3 memory and a bandwidth of up to 4,000 GB/s, it supports data-heavy applications and extensive datasets
- Utilises the latest 4nm lithography, offering higher transistor density, which translates to improved performance and efficiency for intensive computational tasks

BENEFITS:

- Optimised for rapid deep learning training and inference, reducing time to insight and speeding up AI model development
- High memory bandwidth and capacity allow the MI300X to manage and process large datasets more effectively, crucial for big data analytics
- With massive compute power and memory bandwidth, it delivers the throughput necessary for the largest supercomputers and datacentres
- Enhanced cooling and robust build quality ensure consistent performance under heavy workloads, reducing downtime and maintenance costs

INTEL® HABANA® GAUDI2®

The original Habana® Gaudi® already boasted 40% better price performance in the AWS cloud with Amazon EC2 DL1. Habana Gaudi2®, the second generation of these processors, capitalises on the success of its predecessor by focusing on high-efficiency deep learning. Habana Gaudi2® makes the already cost-efficient previous generation even more efficient by lowering training times.

This level of efficiency will increase performance in fields such as object detection in autonomous vehicles, object detection in medical imaging and defect detection in manufacturing.

KEY FEATURES:

- High-efficiency deep learning capabilities
- Increased object detection in autonomous vehicles, medical imaging and manufacturing

BENEFITS:

- 40% better price performance in the AWC cloud with Amazon EC2 DL1
- Lowers training times to make the previous generation more cost-efficient
- Boost performance through efficiency

NVIDIA® NVL72 SUPERPOD

The NVL72 SuperPod is a highly advanced datacentre rack configuration designed to support high-performance computing and AI workloads. It comprises multiple NVL72 servers along with integrated networking switches to facilitate both single and multi-node AI model training and inference.

The NVL72 SuperPod is engineered to work seamlessly with leading storage and networking technologies, ensuring robust and scalable performance. Boston provides a range of NVL72 SuperPod reference architecture solutions in partnership with top storage vendors such as DDN, IBM Spectrum Storage, VAST Data and WEKA. These solutions follow best practices for NVL72 SuperPod design and are delivered as fully-integrated, ready-to-deploy systems to streamline and expedite your datacentre AI deployments.

KEY FEATURES:

- The NVL72 SuperPod features advanced NVL72 servers optimised for intensive AI and computing tasks
- Comes with built-in networking switches to support efficient data transfer and communication between nodes
- Designed to integrate smoothly with leading storage and networking technologies
- Supports scalability to accommodate growing data and workload demands

BENEFITS:

- Provides powerful computing capabilities to handle complex AI models and large datasets efficiently
- Fully-integrated and ready-to-deploy solutions simplify and accelerate the implementation process in datacentres
- Easily integrates with various storage and networking technologies, offering flexibility and compatibility
- Designed to scale with the needs of your organisation, ensuring long-term viability and growth support

STORAGE SOLUTIONS

BOSTON IGLOO AI+

The Boston Igloo AI+ is Storage Made For AI. It has been designed specifically for the high performance needs and low budget requirements of AI projects - without compromise.

DDN EXASCALER

DDN Exascaler offers appliances delivering fast, reliable and efficient data storage and retrieval, which makes them ideal for businesses that require high-speed data processing.

IBM

IBM offers storage solutions, including flash storage, hybrid cloud storage and data protection solutions. Their all-flash arrays deliver high-speed data transfer and low latency, making them ideal for businesses that require real-time data access.

WEKA

Weka provides storage systems that are designed for data-intensive applications and workloads, providing high-speed data access and processing that can keep up with demands.

BOSTON IGLOO AI+

Investment in accelerator technology can be critical for the success of an AI project. We have designed the Igloo AI+ with blistering performance to support AI workloads without the need to compromise investment into accelerators. The NFS protocol allows multiple DGX servers to access the storage meaning that more systems can be connected to the environment as the need for AI compute grows.

The Boston Igloo AI+, powered by PEAK:AIO, strikes the balance of cost versus blistering performance making it the perfect fit for storage needs that generally run into the low hundreds of TB storage capacity needed to feed the accelerators with data to develop a model in as short a time as possible. For smaller environments with a single GPU system, the Boston Igloo AI+ can be connected directly WITHOUT the need for a switch! This furthers investment that can be made into GPU's!

KEY FEATURES:

- Block or file access
- Designed for AI workloads
- Up to 80GB/s+ throughput
- RAID 1, 10, 5 and 6 support
- GPU direct support
- Up to 737TB capacity

BENEFITS:

- Exceptional value for money
- Packaged appliance - Built, configured and remote installation service from Boston
- NFS protocol allows additional DGX systems to be added over time
- Set it and forget it. No need for intensive administration

PEAK AIO

SPECIFICATIONS

MODEL	IGLOO AI+ 1U FORM FACTOR	IGLOO AI+ 2U FORM FACTOR
Drives	6 or 12x NVMe	6, 12 or 24x NVMe
CPU	(2x) Intel Gold 6442Y	(2x) Intel Gold 6442Y
Memory	256GB	256GB
Network Adapter	(2x) NVIDIA 200GbE card	(2x) NVIDIA 400GbE card
Installation	Remote Installation	Remote Installation
Capacity	Up to 368.6TB RAW	Up to 737.3TB RAW
Throughput	40GB/s+	80GB/s+

USE CASE:

CHALLENGE

Many organisations face a choice when buying storage to support NVIDIA DGX technology. They wish to invest as much as possible into DGX only to find that little of the budget remains for fast parallel storage. Enterprise All Flash does not deliver enough performance, while a traditional parallel storage system is too expensive.

SOLUTION

The Boston Igloo AI+ meets the need by providing blistering storage performance at a pricepoint that can support the limitations of a small budget. The NFS filesystem means that DGX systems can be added to storage over time while not compromising on performance needed to support the AI workloads.

ADVANTAGE

The Boston Igloo AI+ has been designed especially for this common use case. The low cost of this fast storage solution means that on many occasions budget can be directed into more GPUs.

DDN EXASCALER

For decades, DDN has designed, developed, deployed and optimised systems, software and storage solutions that enable enterprises, service providers, universities and government agencies to generate more value and to accelerate time to insight from their data and information on-premise and in the cloud.

Developed and optimised using the latest advances in filesystem software technology, the DDN EXAScaler storage appliance delivers extreme performance, scalability, capability, reliability and simplicity.

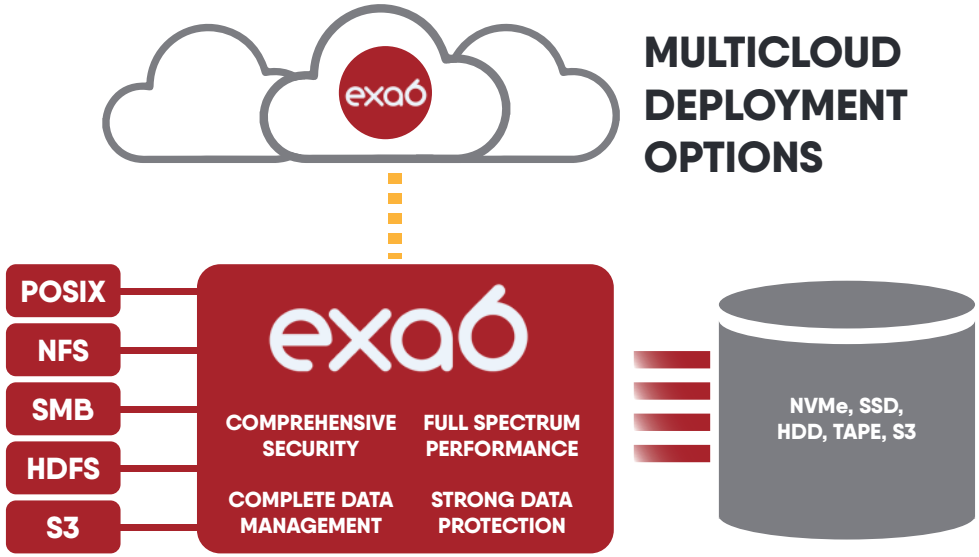
EXAScaler introduces several new data management and integrity filesystem features developed by DDN and is only available in its appliances and cloud offerings. Stratagem is a powerful data orchestration engine that gives users comprehensive data residency controls using policy-based placement.

KEY FEATURES:

- Enterprise storage features in HPC array
- Simplify data management and orchestration with API driven data integration tool
- Snapshots, Encryption, Data Integrity, Quotas, Cloud Workload analytics (SIA), MultiCloud readiness & Multitenancy

BENEFITS:

- Multi-protocol support and features for evolving customer needs
- File (NFS & SMB) Object (S3) and Container support
- Fastest throughput per SSD, scalable metadata all drive performance



USE CASE:

CHALLENGE

A leading life sciences university required shared computational resources to accelerate genomic medicine research, while keeping costs down under a tight budget.

SOLUTION

System overhaul for improved performance and build capabilities for simulations, data analysis and sharing.

ADVANTAGE

Enabled system overhaul under limited budgets, combining 200G HDR Infiniband and 40/10 Gbps Ethernet with a DDN Storage solution to improve performance and build capabilities, with up to 90 GB/sec performance per appliance.

IBM® SPECTRUM STORAGE™

Enterprises and organisations are creating, analysing and keeping more data than ever before. In delivering those insights, an organisation's underlying information architecture must support the hybrid cloud, big data and artificial intelligence (AI) workloads along with traditional applications while ensuring security, reliability, data efficiency and high performance. IBM® Spectrum Storage™ is capable of managing petabytes of data and billions of files providing world class storage management with scalability, performance and policy-based storage tiering.

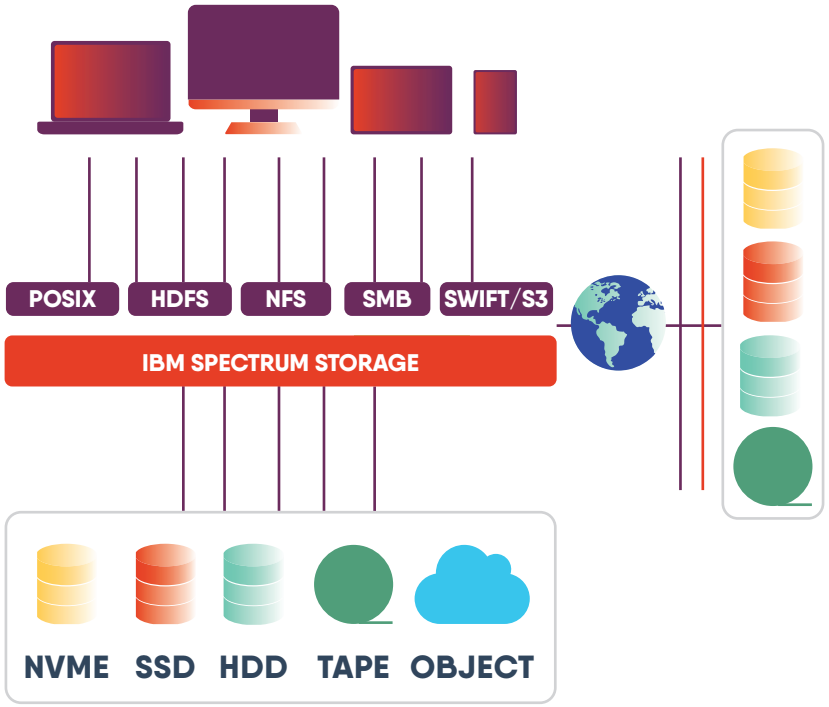
IBM® Spectrum Storage™ is a parallel file system, where the intelligence is in the client and the client spreads the load across all storage nodes in a cluster, even for individual files. The IBM® Spectrum Storage™ architecture allows it to seamlessly handle tens of thousands of clients, billions of files and yottabytes of data.

KEY FEATURES:

- The enhanced GUI supports many features: performance, capacity, cloud tiering and enhanced maintenance with IBM remote support
- File audit logging capability tracks user access to the filesystem across all nodes and protocols
- The Spectrum Storage™ client spreads load across all nodes in a cluster. This means that the architecture can scale easily while providing access to thousands of clients

BENEFITS:

- Simplified data management
- Simple administration
- Security enhancements



USE CASE:

CHALLENGE

To maintain its reputation as a premier research institution, a leading UK university needed to ensure that data is always available to a growing number of users running increasingly complex simulations.

SOLUTION

The university deployed IBM® Spectrum Storage™ and IBM Spectrum Protect™, increasing transparency around data location and who accesses it, and increasing its mobility within a diverse IT environment.

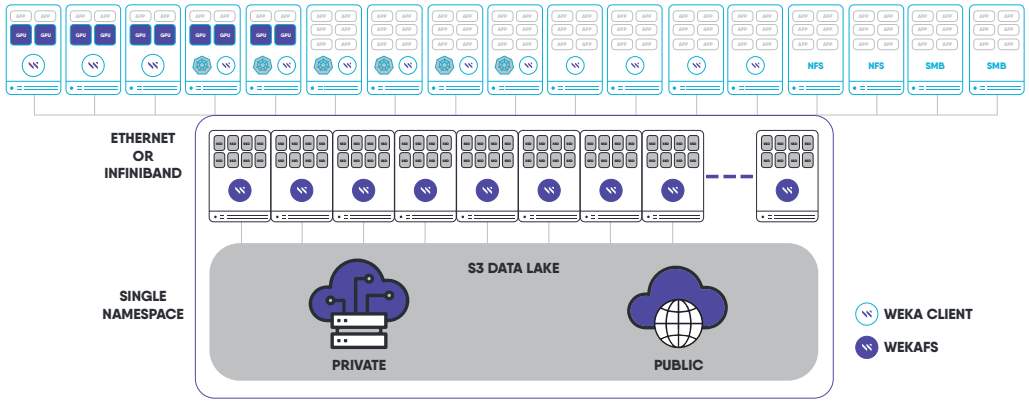
ADVANTAGE

The solution supplied supports compliance with data protection regulations at a low cost without disruption. Significant cost savings due to operational efficiency, and 5000 researchers are supported by infrastructure that helps them get to results faster.

WEKA IO

Built for NVMe flash and cloud-native, Weka is a parallel file system designed to enable organisations to maximise the full value of their high-powered IT investments - compute, networking and storage. By leveraging existing technologies in new ways and augmenting them with engineering innovations, Weka's software delivers a more powerful and simpler solution that would have traditionally required several disparate storage systems. The resulting software solution delivers high performance for all workloads (big and small files, reads and writes, random, sequential and metadata heavy).

Weka's Data platform was built from the ground up to shatter the limits imposed by legacy storage architectures by taking a unique software-only approach to address the needs of modern workloads.



KEY FEATURES:

- Manage all your file and object data on NVMe, SSD flash or disk in the system within a single global namespace
- Management is easy with a simple GUI. Rapidly provision, manage policies, change data protection levels and monitor system health intuitively
- No tuning is required once set up. WekaFS is ideal for the challenges of mixed workloads

BENEFITS:

- Manage all your data from one, centralised location
- Eliminate guesswork
- User friendly management

USE CASE:

CHALLENGE

Genomics England had previously implemented a scale-out NAS solution from a leading vendor to support the 100,000 genome project; however, it had already hit its limit on storage node scaling and performance suffered when the system was near capacity.

SOLUTION

A two-tier architecture that takes commodity flash and disk-based technologies, presenting it as a single hybrid storage solution. The primary tier consists of 1.3 Petabytes of high performing NVMe flash storage which supports the working data sets.

ADVANTAGE

GE was able to realise no limit on capacity scaling, over 10x improvement in performance, 75% reduction in storage cost per genome, as well as now embedding a full disaster recovery strategy and offering integration with public cloud for compute elasticity.

LIQUID COOLING

Liquid cooling is rapidly becoming an inevitable approach to HPC cooling. Why is that? The short answer is that air cooling will not be able to keep pace with the wattage of components. The long answer is that on top of meeting temperature requirements, it also works towards sustainability goals as liquid cooling requires less energy to maintain, and generally there is less maintenance with a liquid cooled setup.

Using liquids to cool a system can come in many shapes and sizes. There are two categories which then also break into two types; single-phase and dual-phase liquid cooling. Each approach can be applied directly on the chip or by immersing components, or even the entire system.

DIFFERENT TYPES OF LIQUID COOLING

DIRECT LIQUID COOLING (DLC)

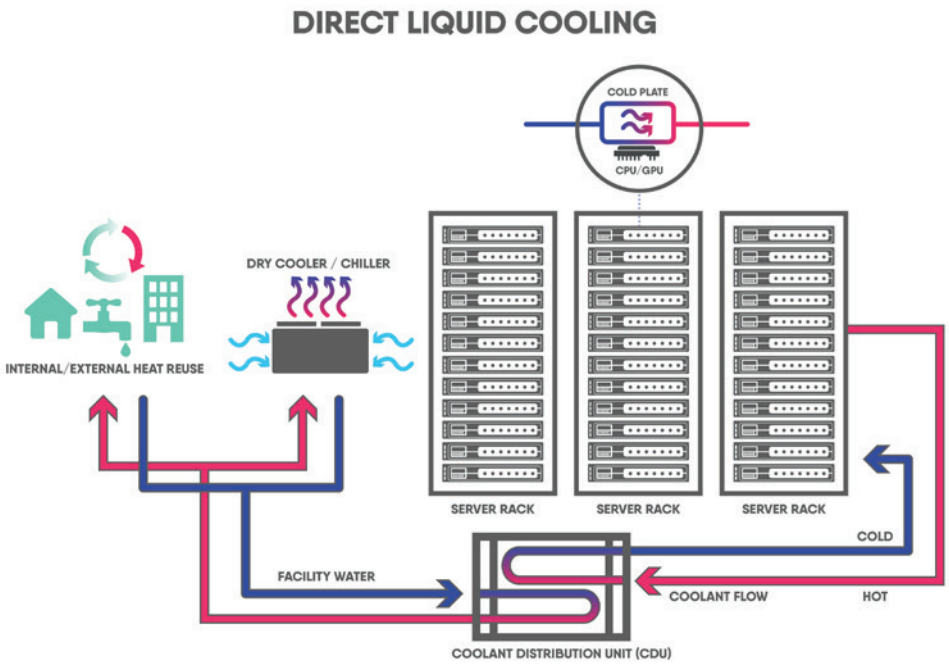
In the realm of liquid cooling, single-phase DLC maintains a consistent liquid state, efficiently absorbing heat from components and subsequently dissipating it through cooling mechanisms before being recirculated, finding applications in both direct-on-chip and immersion cooling setups.

DUAL-PHASE DIRECT LIQUID COOLING (2-PHASE DLC)

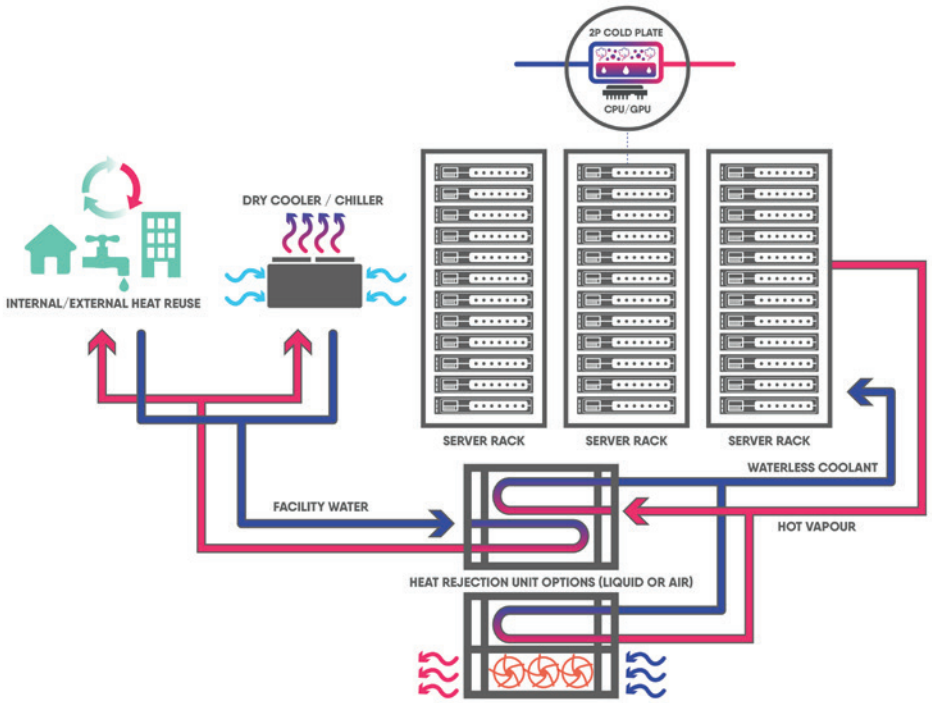
Dual-phase DLC involves an intricate process where a liquid undergoes a phase change into vapour upon contact with heat and, after vapour cooling, it condenses back to its liquid form, a method often employed in immersion cooling; it relies on a precise choice of a fluorocarbon-based liquid that transitions states around 50°C, demanding careful handling.

IMMERSION COOLING

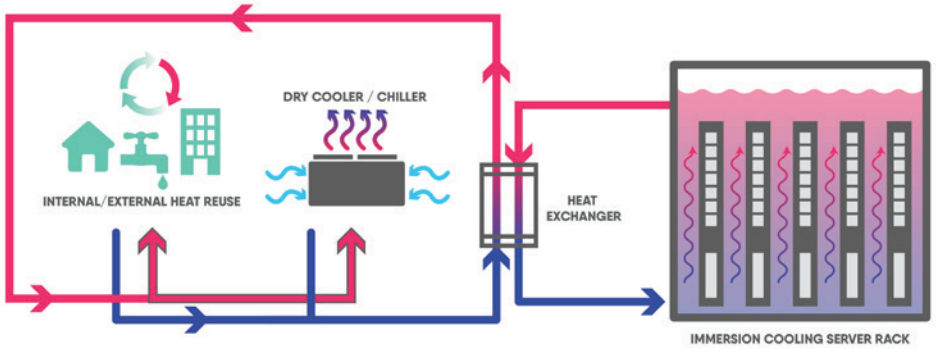
Immersion cooling submerges components within a dielectric fluid, offering the choice of both single and dual-phase cooling loops, its dielectric properties ensuring safety as it efficiently absorbs and dissipates heat from electronic parts; with a secondary heat exchanger in play, the captured heat can be redirected for dissipation or repurposing, making it particularly well-suited for high-power applications such as datacentres and industrial processes.

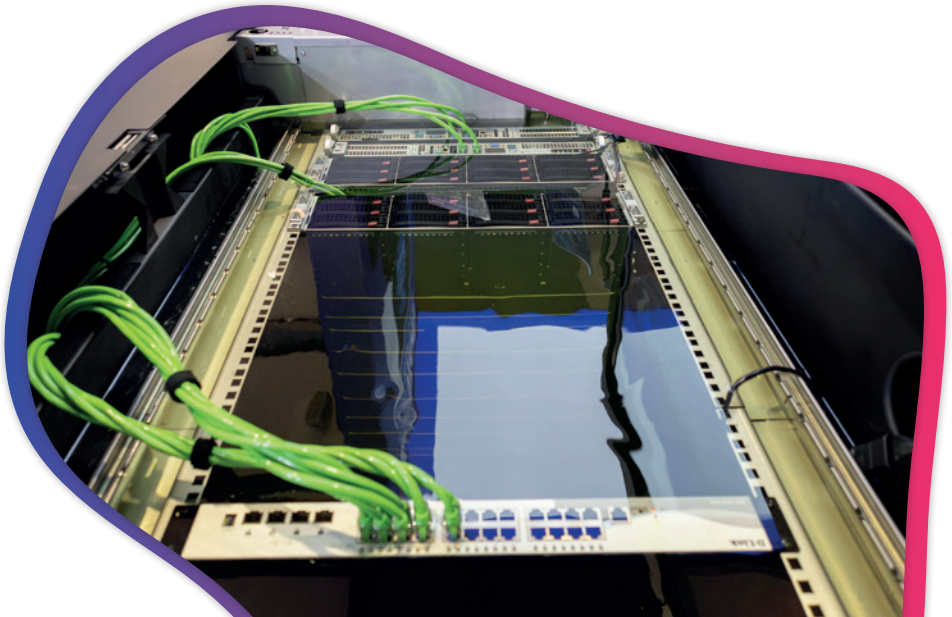


DUAL-PHASE DIRECT LIQUID COOLING



SINGLE-PHASE IMMERSION





LIQUID COOLING

ASPERITAS - SINGLE PHASE IMMERSION

Asperitas' solutions use convection to circulate the fluid making them even more sustainable as there are no pumps involved, further lowering the overall total cost of operation.

CASTROL - SINGLE PHASE IMMERSION

Boston Limited has partnered with Castrol, a global leader in lubricants and thermal management, to bring advanced cooling options to the data centre industry.

THE GLACIERCORE SERIES

The Boston GlacierCore series is crafted for businesses that require robust, reliable performance without compromise.

SUBMER - SINGLE PHASE IMMERSION

Submer are a forward-thinking company who designs solutions that address issues that may arise in the future as technology advances and our reliance on them increases.

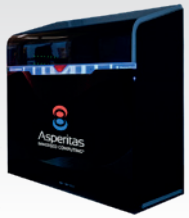
SUPERMICRO - DLC SINGLE PHASE

Supermicro offer a range of liquid cooling integrated solutions. Direct to chip liquid cooling solutions deliver superior performance and reduce datacenter electricity costs by up to 51%.

ZUTACORE - DLC DUAL-PHASE

Zutacore's solutions enable organisations to elevate their providings through compute-and data-intensive applications, ensuring their systems' best performance and longevity.

ASPERITAS



Since 2014 Asperitas has worked on validating and developing Immersed Computing® as a unique approach to the datacentre industry. Building on existing liquid immersion cooling technologies by adding integration of power and network components, improving cooling physics with a strong focus on design and engineering for usability, Asperitas has come up with a complete and integrated solution which can be effectively utilised in most, if not all situations.

Asperitas is a high-tech company who also put emphasis on sustainability. As a leading immersion cooling specialist, they provide cutting-edge solutions for energy efficient and high-density datacentres globally. Immersed Computing®, Asperitas' unique award-winning natural convection-based technology concept, enables sustainable and high-density datacentres to run your applications anywhere they are needed.

KEY FEATURES:

- Reduces CO2 emissions by up to 40%
- 98% energy reuse
- Reduces TCO by up to 40%

BENEFITS:

- Plug and play capabilities
- Modular system
- Increases density by up to 5x

CASTROL

CASTROL OIL & IMMERSION COOLING TECHNOLOGY

At the core of our liquid cooling solutions is Castrol's pioneering immersion cooling technology. Castrol's specially engineered cooling fluids provide superior thermal management, ensuring efficient heat transfer even during high loads. This results in a lower energy footprint, helping organisations minimise operational costs while extending the lifespan of their equipment.

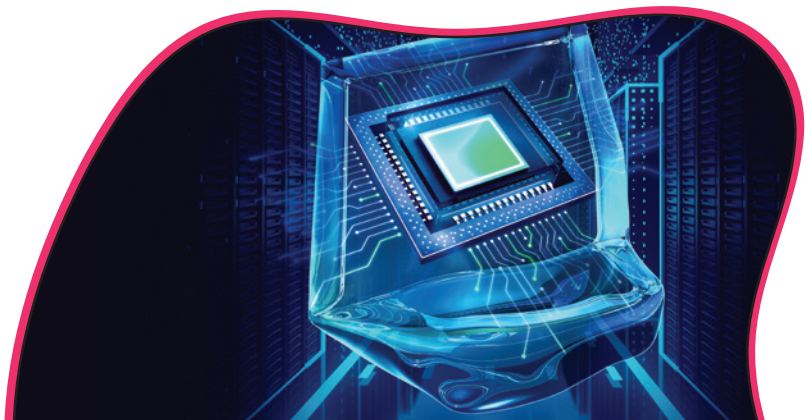
Castrol's immersion cooling fluids are engineered for superior heat transfer, optimising the cooling process and enabling data centres to operate more efficiently, this also minimises the need for traditional air conditioning which reduces energy consumption, supporting Boston's mission to deliver sustainable technology options to our customers.

KEY FEATURES:

- Enhanced Efficiency
- Environmentally Friendly
- Prolonged Equipment Life

BENEFITS:

- Custom Solutions for Every Setup
- Reduced Operational Costs
- High-Performance Reliability



LIQUID COOLING WITH BOSTON LIMITED

As data demands grow, so does the need for innovative cooling solutions that support high performance, efficiency and sustainability. At Boston, we're at the forefront of this evolution through our partnership with Castrol, a leader in thermal management and immersion cooling fluids. By combining Castrol's advanced cooling technology with Boston's expertise in high-performance computing, we're providing powerful solutions for modern data centres that require optimal cooling without compromising on environmental goals.

WHY LIQUID COOLING?

Liquid cooling provides more effective heat dissipation compared to traditional air cooling, making it ideal for data centres running intensive workloads.

SUSTAINABILITY

By reducing the reliance on air conditioning, liquid cooling significantly cuts energy consumption, aligning with sustainability targets. Our partnership with Castrol means Boston customers gain access to custom cooling solutions that go beyond traditional methods. Designed to support the most demanding environments, our liquid cooling systems deliver unmatched reliability, reduce energy costs and support sustainable business operations.

Experience the next generation of data centre cooling with Boston and Castrol. Connect with our team today to discover how our liquid cooling solutions can transform your infrastructure's efficiency, reliability and sustainability.

THE GLACIERCORE SERIES

PURPOSE-BUILT FOR LIQUID COOLING

Our GlacierCore range of servers is designed from the ground up for immersion cooling, ensuring maximum efficiency and performance. Built to handle the demands of modern data centres, these servers leverage Castrol's cooling fluids to optimise performance, reduce thermal stress and support sustainable operations.

GLACIERCORE MODELS OVERVIEW:

GlacierCore 1100
A compact and cost-effective solution ideal for entry-level deployments, offering efficient immersion cooling for essential tasks.

GlacierCore 1110
A powerful choice for high-density applications, designed for businesses requiring robust computing performance with reliable cooling.

GlacierCore 1200
Built for advanced workloads, this model supports intensive applications such as AI training, benefiting from high-performance cooling.

GlacierCore 1210
Designed to support data-intensive and mission-critical environments, this model ensures stability even under heavy processing demands.

GlacierCore 1220
The top-tier choice for extreme processing needs, this model is optimised for AI, ML and HPC tasks, delivering sustained performance with unparalleled cooling efficiency.

APPLICATIONS

- AI and Machine Learning Training
- High-Performance Computing
- Data Analytics and Financial Modelling

Each model can be tailored to specific use cases, from high-density tasks to intensive AI applications, ensuring that you find the ideal solution for your data centre's needs.

SUBMER

Submer's main goal is to make the operation and construction of datacentres more efficient and sustainable. Submer uses their products, platform, APIs, processes and installations to make hyperscalers and huge industries more efficient.

They put the needs of their customers first by providing solutions that are not only beneficial today but are future-proof also. Their solutions address questions like heat-reuse, net-zero water and site selection. The single-phase immersion cooling Submer uses is flexible as it can be scaled without issue, all while staying space and energy efficient.

Submer is in a unique position where they not only provide technology and solutions, but also develop them. They have expertise in thermodynamics, engineering and chemical engineering to name a few.

KEY FEATURES:

- Achieve a PUE as low as 1.03
- Up to 100kw of compute density
- Reduce latency and increase speed of deployment

BENEFITS:

- Increase hardware life-span
- Reduce hardware failure rate
- Save 50% on CAPEX building costs
- Minimise water usage

SUPERMICRO

The Supermicro liquid cooled rack solution delivers superior performance and efficiency for HPC, large scale AI and cloud scale compute infrastructure.

Servers that are application optimised for AI, HPC and Analytics require the latest in CPU and GPU technologies, which run hotter than previous generations. Multiple CPUs and GPUs per server are needed for performance intensive computing, driving up the REDUCTION electricity demands for the server as well as at the rack level.

Supermicro's liquid cooled rack solution is made of several components that are designed inhouse to achieve the highest level of performance and reliability. All the components are integrated as a rack level solution providing a true one-stop shop customer experience.

KEY FEATURES:

- Enterprise grade redundant cooling pump and power supplies, leak-proof connectors, and leak detection
- Options for Processor and GPU cooling

BENEFITS:

- Full turn-key single source solution from Supermicro
- Accelerated lead times based on in-stock inventory
- Up to 51% reduction in electricity costs in the datacenter

ZUTACORE

Having optimal server capabilities are essential for organisations engaged in compute- and data-intensive applications. These organisations must invest in efficient cooling solutions to ensure their systems' best performance and longevity.

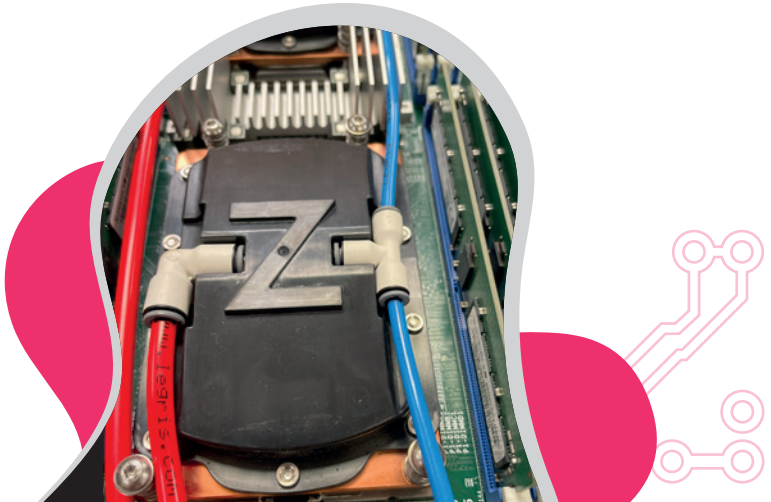
This investment is especially true as each successive generation of processors creates more heat and has more stringent cooling tolerances. The ZutaCore dielectric liquid cooling solution is one of the most effective methods for providing reliable cooling to high-end processors. It has been proven to enable CPUs to perform consistently at high utilisation rates with zero throttling.

KEY FEATURES:

- 50% less energy
- 3x Processing Capacity
- 50% less space

BENEFITS:

- >1000W of Chip Cooling
- Waterless dielectric liquid cooling
- 50% Capex Reduction



NVIDIA NETWORKING

NVIDIA ETHERNET

NVIDIA Networking Ethernet products provide scalable, low-latency and efficient fabric with simplified management, reducing overall costs and power consumption compared to traditional Ethernet.

NVIDIA INFINBAND

NVIDIA Networking's InfiniBand switches and adapters support up to 400Gb/s, enabling scalable fabric for high-performance computing and datacentres.

NVIDIA SPECTRUM™-X

The NVIDIA Spectrum™-X networking platform is the first Ethernet platform designed specifically to improve the performance and efficiency of Ethernet-based AI clouds.

NVIDIA ETHERNET SPECTRUM®-4 PLATFORM

The world's first 400Gbps end-to-end networking platform, NVIDIA Spectrum®-4 provides 4x higher switching throughput than previous generations, with 51.2 terabits per second.

It consists of the NVIDIA Spectrum®-4 switch family, NVIDIA ConnectX®-7 SmartNIC, NVIDIA BlueField®-3 DPU and the DOCA™ datacentre infrastructure software to supercharge cloud-native applications at scale. Built for AI, NVIDIA Spectrum®-4 Switch arrives as datacentres are growing exponentially and demanding extreme performance.

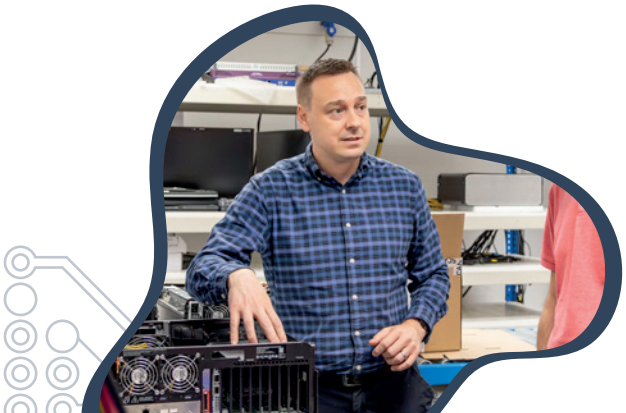
NVIDIA Spectrum®-4 features a fully-shared and monolithic packet buffer that's dynamically available to all ports. This provides excellent microburst absorption with true, port-to-port, cut-through latency.

KEY FEATURES:

- 400 Gps end-to-end networking platform
- 4x higher switching throughput
- 51.2 TB/s

BENEFITS:

- Provides extreme performance
- Enhanced efficiency
- 4x higher switching throughput than previous generations



NVIDIA INFINIBAND QUANTUM™-2 PLATFORM

The NVIDIA Quantum™-2 based QM9700 and QM9790 switch systems deliver an unprecedented 64 ports of NDR 400Gb/s InfiniBand per port in a 1U standard chassis design.

A single switch carries an aggregated bidirectional throughput of 51.2 terabits per second (Tb/s), with a landmark of more than 66.5 billion packets per second (BPPS) capacity.

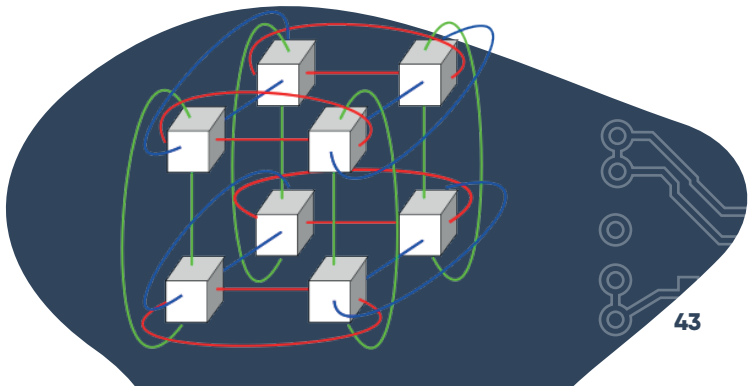
Supporting the latest NDR technology, NVIDIA Quantum™-2 brings a highspeed, extremely low-latency and scalable solution that incorporates state-of-the-art technologies.

KEY FEATURES:

- High speed, low latency and scalable
- Aggregated bidirectional throughput of 51.2 TB/s
- Delivers 64 ports of NDR 400 GB/s

BENEFITS:

- Incorporates state-of-the-art technologies
- 66.5 BPPS capacity
- 1U standard chassis design



NVIDIA SPECTRUM™-X

Spectrum™-X is a revolutionary solution created by NVIDIA for building multi-tenant, hyperscale AI clouds and AI Factories with Ethernet. With it, organisations can significantly improve the performance and power efficiency of AI clouds and gain higher predictability and consistency, resulting in faster time to market and a greater competitive edge.

NVIDIA Spectrum™-X enhances performance and power efficiency by up to 1.7X, accelerating the processing, analysis and execution of AI workloads and, in turn, the development and deployment of AI solutions. Spectrum™-X is fully standards-based Ethernet with support for open Ethernet stacks (SONiC) at cloud scale, and is tuned and validated across the full stack of NVIDIA hardware and software, creating an unmatched Ethernet solution for AI clouds.

KEY FEATURES:

- Nearly Perfect Bandwidth at Scale
- Extremely Low Latency
- End-to-End Stack Optimisation
- Advanced RoCE Extensions for Scalable AI Communications
- Deterministic Performance and Performance Isolation
- Open Network Operating System: SONiC and Cumulus

BENEFITS:

- Improved AI Cloud performance
- Reduced TCO
- Reduced time to results
- Full-stack optimisation
- Enhanced multi-tenancy features

SOFTWARE

DKUBE

DKube's software is an AI and machine learning platform that streamlines the end-to-end data science workflow, from data preparation and model training to deployment and monitoring.

NVIDIA BASE COMMAND MANAGER

NVIDIA Base Command Manager is a centralised GPU resource management platform that simplifies the allocation, deployment and utilisation of GPU resources across your HPC cluster.

WAREWULF

Warewulf empowers you to create, manage and deploy system images within your HPC cluster. This flexibility enhances your cluster's efficiency by customising node configurations to meet specific workload requirements.

DKUBE

DKube is a powerful data integration and analytics platform designed to streamline how organisations manage and analyse their data. It offers a comprehensive set of tools for aggregating data from various sources, performing advanced analyses, and visualising results through customisable dashboards.

With its capability for real-time data processing and scalable architecture, DKube supports businesses in making informed decisions and staying competitive in a dynamic market.

KEY FEATURES:

- DKube integrates seamlessly with a wide range of data sources and formats, enabling efficient data aggregation and analysis
- It provides powerful analytics tools to derive actionable insights from complex data sets
- Users can create personalised dashboards to visualise data in ways that meet their specific needs
- The platform supports real-time data processing to deliver up-to-date information and facilitate faster decision-making
- is designed to scale with the growth of data and user demands, ensuring long-term usability

BENEFITS:

- By providing comprehensive and real-time data insights, DKube helps businesses make informed decisions
- Streamlining data integration and analytics reduces the time and effort required for data management tasks
- Customisable dashboards and advanced visualisation tools enhance data accessibility and understanding
- Advanced analytics capabilities enable businesses to gain deeper insights and stay ahead of market trends

NVIDIA BASE COMMAND MANAGER

NVIDIA Base Command Manager is the cornerstone of efficient GPU resource management for your HPC cluster. This powerful platform centralises the allocation, deployment and utilisation of GPU resources, eliminating the complexities associated with GPU management. Whether you're a seasoned data scientist or a newcomer to the world of GPU computing, NVIDIA's Base Command Manager simplifies GPU workflow, making it a valuable asset for your HPC infrastructure.

The system's user-friendly interface and robust support for multiple users provide the foundation for optimising GPU-driven workloads.

KEY FEATURES:

- NVIDIA Base Command Manager offers a unified platform to manage your GPU resources across your HPC cluster
- Facilitate collaboration with support for multiple users and seamless allocation of GPU resources
- Utilise containers to streamline AI, ML and HPC workloads, ensuring efficient resource utilisation

BENEFITS:

- Bolstered efficiency
- Easy collaboration
- Simplifies workload management

WAREWULF

Warewulf is your gateway to customised system management within your HPC cluster. This versatile tool empowers you to create, manage and deploy system images tailored to your unique requirements.

HPC clusters often handle a diverse range of workloads, each with specific node configurations and resource needs. Warewulf simplifies the process by allowing you to craft custom system images that match the exact requirements of your workloads. In the ever-shifting world of HPC, flexibility is the key to success, and Warewulf offers precisely that. It's time to streamline your cluster configuration and improve your HPC cluster's efficiency.

KEY FEATURES:

- Warewulf enables the creation and management of system images for simplified cluster configuration
- Tailor node configuration to specific workload requirements, optimising performance
- User-friendly tools simplify node provisioning and system administration

BENEFITS:

- Improved efficiency
- Customisable
- User friendly

BOSTON SERVICES

At Boston, we bring together a diverse range of expertise and services under one roof to meet all your technology needs. Whether you're looking for cutting-edge solutions, in-depth training, or reliable support, we offer a complete suite of services to support your organisation's success. From our global team of technical and HPC experts who benchmark and optimise new technologies, to our Boston Training Academy that nurtures talent and education and our Boston Labs that provide custom tech solutions and testing facilities—we cover every aspect of your technology journey.

BOSTON TECHNOLOGY CONSULTING

Our technical and HPC experts, based around the world, are constantly benchmarking new technologies, the results of which we share with our customers, to deliver fully optimised solutions.

BOSTON TRAINING ACADEMY

The mission of the Boston Training Academy (BTA) is to become a renowned developmental ground for talent engagement, education and solutions across a variety of disciplines.

BOSTON LABS

Boston provides tech facilities to test emerging technologies and enterprise-grade systems from leading manufacturers. Boston Labs is customer-centric, customised solutions with off-the-shelf components to enhance business outcomes.

PROFESSIONAL SERVICES

Our senior engineers understand the technical dependencies and requirements of your organisation, we will ensure a well thought out installation is managed and completed on schedule and with the utmost professionalism.

SUPPORT AND ONSITE WARRANTY

From tailor-made SLA solutions, warranty support and spares packages – Boston offers customised services level work packages for ongoing support.

BOSTON TECHNOLOGY CONSULTING

Our technology consultancy service helps organisations to effectively create, optimise and grow their business in today's data and IT focused market. Our consultants mastermind solutions for the most complex business needs in an effort to achieve their goals, while balancing costs and performance. Our core values are leadership, integrity, collaboration, accountability and simplicity.

Our global reach enables us to provide agile and scalable solutions to everyone from the smallest players to industry heavy hitters. Our main objective is to solve current problems while minimising risks of future issues. Such all-encompassing approach to problem solving has been utilised in the banking, financial services and insurance (BFSI), government, manufacturing, oil and gas, logistics, telecoms, pharma and retail.

Our consultancy service is aimed at enterprises striving to utilise specialised services to optimise their solutions and enhance efficiency. With our deep industry knowledge, we provide valuable guidance to unlock the full potential of your Infrastructure.

- **Cross-domain expertise:** We stand out through our wide-ranging expertise across multiple technical domains, allowing us to provide holistic solutions that address complex challenges spanning various technologies.
- **Innovative problem solving:** We excel in innovative problem-solving, employing a team of seasoned consultants who approach challenges with creativity and resourcefulness, leading to unique solutions that drive customer success.
- **End-to-end project management:** We offer comprehensive project management services, overseeing projects from inception to completion. This end-to-end approach streamlines processes, reduces customer workload and ensures successful project outcomes.

BOSTON TRAINING ACADEMY

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.

BOSTON LABS

DISCOVER THE FUTURE WITH BOSTON LABS: YOUR TESTING PARTNER

Are you navigating the ever-evolving landscape of new and enhanced technologies? We understand the challenge of making informed decisions for your upcoming projects. The task becomes even more daunting when you can't experience and assess the hardware firsthand prior to making a commitment.

Through our Boston Labs, we bridge this gap for you. Our state-of-the-art, remote accessible, on-site test facilities offer the perfect environment to explore and evaluate emerging technologies. We recognise that sometimes, the true fit of a product or solution can only be gauged through hands-on experience.

Our dedicated technical team is poised to assemble comprehensive solution stacks using cloud-based lab systems or even loaned hardware that can be seamlessly integrated into your environment. This ensures that your investment aligns perfectly with your requirements. Immerse yourself in our incredible R&D lab, Boston Labs, where an array of the latest hardware and software solutions await your evaluation. Whether you choose remote assessment or prefer to collaborate closely with our seasoned field application engineers, the choice is yours.

Our veteran team excels in guiding you through the setup and optimisation of both hardware and software components, crafting solutions that harmonise performance, power efficiency and cost-effectiveness. When you reserve a session at Boston Labs, you gain access to expert advice on product selection and fine-tuning, helping you achieve unparalleled outcomes.

EXPERTS IN AI AND DATA

Boston Training Academy and Boston Technical Consulting are our in-house initiatives which represent the pinnacle of technological advancement, driving forward groundbreaking solutions that harness the power of artificial intelligence.

PROFESSIONAL SERVICES

PRESALES CONSULTATION AND SERVICES

Our Presales Consultation and Services aim to provide optimal solutions for your business needs. Our team of experts is dedicated to understanding your requirements, offering personalised consultations and delivering our wide range of technology solutions. We go beyond product knowledge, bringing a deep understanding of industry trends and emerging technologies to help you stay ahead of the competition.

WE CAN PROVIDE:

- Consultancy and Solution Design
- Inventory and Site Survey
- Technical Account Management
- Boston Labs and POC Services
- Bespoke System Design and OEM Manufacturing
- Product and Technology updates
- Environmental and Regulatory Testing

POST SALES AND IMPLEMENTATION

Our Post Sales and Implementation team ensures a smooth and successful deployment of your technology solutions. With skilled project management expertise, thorough testing, seamless integration and comprehensive training, we prioritise your success from purchase to implementation.

WE CAN PROVIDE:

- System Integration and Burn in Testing
- Racking and Installation
- Project Management
- Global delivery
- Turnkey Solutions and Rack and Roll
- Training and Education

BOSTON'S PROFESSIONAL SERVICES CONTINUED

DATACENTRE AND HOSTING SERVICES

Your business relies on a secure, reliable and high-performance infrastructure to thrive in the digital landscape. Boston offers state-of-the-art data centre and hosting solutions designed to meet your specific needs. Our industry expertise delivers robust and scalable data centre solutions, including dedicated servers, virtual private servers (VPS), cloud hosting and colocation services. With enterprise-grade infrastructure, customised solutions, 24/7 support and scalability, we empower your digital infrastructure to reach its full potential. Choose Boston's Datacentre and Hosting Services for exceptional performance and reliability.

WE CAN PROVIDE:

- Hosting
- Leasing
- Managed Services and Cloud
- Datacentre and Rack Design
- Rack and Roll

END OF LIFE AND EXTENDED LIFE CARE

As technology rapidly evolves, effectively managing end-of-life products and systems becomes crucial. Boston Limited's End of Life and Extended Life Care services provide comprehensive solutions to navigate this critical phase of your technology lifecycle. Our expert team understands the challenges associated with retiring outdated systems and mitigating risks related to unsupported or vulnerable hardware. With our services, we ensure a seamless transition to newer technology while maximising the value of your existing investments.

WE CAN PROVIDE:

- Decommissioning, Recycling and Disposal
- Out of Warranty Repair and Refurbishment

Please contact us if you are interested in discussing any of our additional service offerings.

SUPPORT AND ONSITE WARRANTY

At Boston, exceptional customer support is the foundation of our successful partnerships. Our Support and Onsite Warranty services offer peace of mind, knowing that our dedicated team of experts is ready to assist you whenever you need help.

With prompt and reliable assistance available 24/7, our experienced professionals with deep expertise across various technologies are just a phone call away. We provide onsite support to minimise downtime and ensure smooth operations, coupled with warranty coverage for eligible hardware, ensuring swift replacement or repair.

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

CO-DESIGN IN HPC: NAVIGATING A NEW ERA MARKED BY POST MOORE'S LAW, EDGE-TO-EXASCALE WORKFLOWS, AND AI

Three significant drivers are reshaping how organizations approach their computing infrastructure and bringing HPC towards the center of IT strategies: (1) the proliferation of technology choices, (2) the need for a computing "fabric" that spans from device to datacenter, and (3) the emergence of artificial intelligence (AI) as a new ecosystem.

These forces are compelling a shift toward custom, co-designed hardware — a collaborative approach to building systems that reduces cost and complexity. Co-design unites customers with the right coalition of vendors to architect and build an optimal, integrated solution tailored to specific requirements.

THE COMPLEXITY OF CHOICE

One of the most challenging aspects of modern HPC and AI is the overwhelming number of technology options available. CPUs from Intel, AMD, Ampere, and others, alongside GPUs and accelerators from Nvidia, AMD, Intel, Cerebras, Untether, Groq, SambaNova, and others. This is further complicated by interconnect and networking choices and a variety of storage options, all supporting a growing roster of software. Selecting and integrating the right component now and into the future is a difficult task.

These advances present opportunities, challenges, and tradeoffs. For example, while low-precision hardware allows for faster computations, not all applications are suited to it. Many HPC workloads, such as weather modeling or molecular simulations, may still require 64-bit precision, while AI-driven applications can often operate effectively with significantly lower-precision arithmetic. But HPC applications can come out ahead if practitioners can reformulate the algorithms to harness the strengths of newer technologies without compromising the precision or accuracy the results.

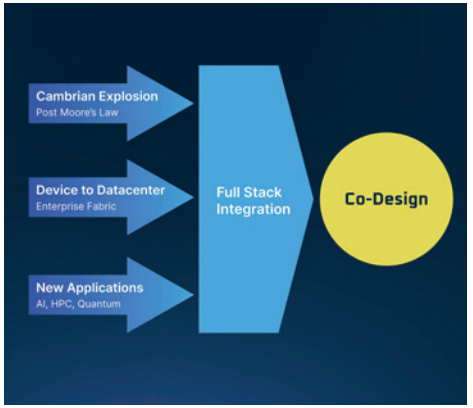
This is an example of where co-design comes into play: the ability to canvas a wide range of hardware and software options to make the best use of available technology. In such a case, it's no longer a question of selecting a pre-built platform; it's about customizing and fine-tuning the platform to serve specific needs. Co-design becomes even more essential when considering the platform's evolution over time — ensuring it adapts seamlessly as new technologies emerge while maintaining the same functionality.

BUILDING A COMPUTING FABRIC: FROM DEVICE TO DATACENTER

In a world where everything is an AI-enabled computer, all IT solutions blend into a "fabric" – a unified web of flexible infrastructure that spans from the edge (where much of the data is generated and eventually consumed) to datacenters of varying size. This distributed computing model aligns with the modern workflow, where managing data flow is as important as computational power.

Consider the landscape of a modern HPC workflow: data streams in from sensors, devices, scientific instruments, or other edge environments, and must flow seamlessly through various computing layers, from edge nodes to datacenter supercomputers and cloud resources. The ability to manage this flow, this meta system, is critical. The co-design model ensures that the computing stack is optimized both in parts and in whole, and that it interacts and integrates with the rest of its extended ecosystem.

This is especially true when speed is of the essence. In applications where real-time data analysis is critical — such as autonomous devices or fleets, or digital twins — the ability to process data close to where it is generated while also having access to centralized, powerful computing resources can be a game-changer.



THE AI REVOLUTION AND THE DEMAND FOR MORE POWER

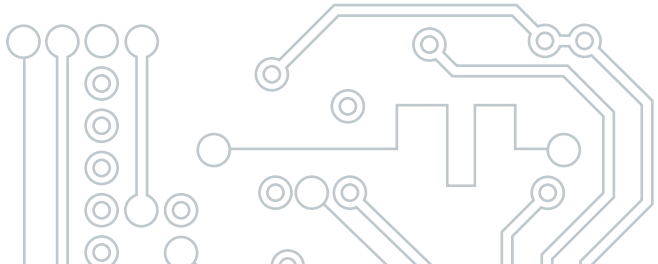
Finally, AI is creating a clean slate, removing and replacing the boundaries of what is possible, and driving the need for even more powerful machines. AI workloads differ from traditional HPC applications in that they can readily take advantage of low-precision data types and arithmetic, and they benefit from a favorable funding environment that enables massively scaled infrastructure. This shift has profound implications for the design of HPC systems because wherever there is speed, HPC is looking to exploit it.

Co-design allows HPC systems to meet these demands head-on. By aligning hardware innovation with the specific requirements of AI workloads, organizations can build systems that are not only powerful but also efficient and scalable.

THE TIME FOR CO-DESIGN IS NOW

As HPC enters this new era of technological complexity, distributed computing, and AI-driven workloads, the need for a co-design approach has never been greater. The proliferation of technology choices, the demand for a seamless computing fabric from device to datacenter, and the new AI ecosystem make it clear that only a broad, vendor-neutral, engineering-led, mission-focused approach can meet all needs. Co-design enables organizations to build HPC systems that are greater than the sum of their parts, allowing them to harness the full potential of the modern computational landscape.

Additionally, power consumption and cooling are becoming more critical as systems grow increasingly powerful. A thoughtful co-design approach can help optimize energy efficiency, ensuring that new HPC infrastructures remain sustainable while meeting performance goals.



HOW SOURCECODE CAN HELP WITH YOUR CO-DESIGN NEEDS

SourceCode's history and business model are rooted in the co-design approach. Partnering with SourceCode for your HPC and AI needs offers distinct advantages:

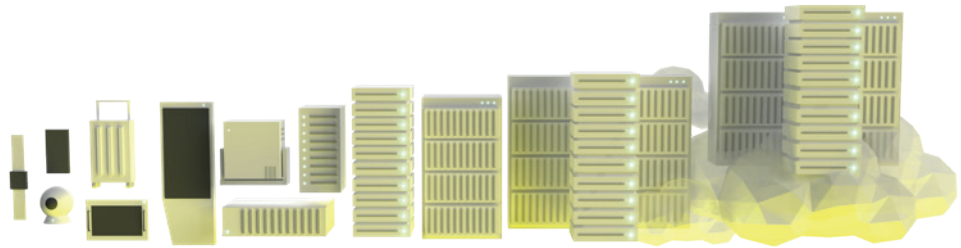
- Accelerates time to market via proven co-design methodology that brings the right expertise at the right time
- Reduces technical risk via access to best-in-class components and building blocks
- Balances design constraints: cost, performance, power, cooling
- Maximizes the potential of rapidly evolving technologies, complementing in-house expertise

SourceCode offers expertise across a broad spectrum of leading-edge technologies that span from device to datacenter, including exascale-class systems from Eviden.

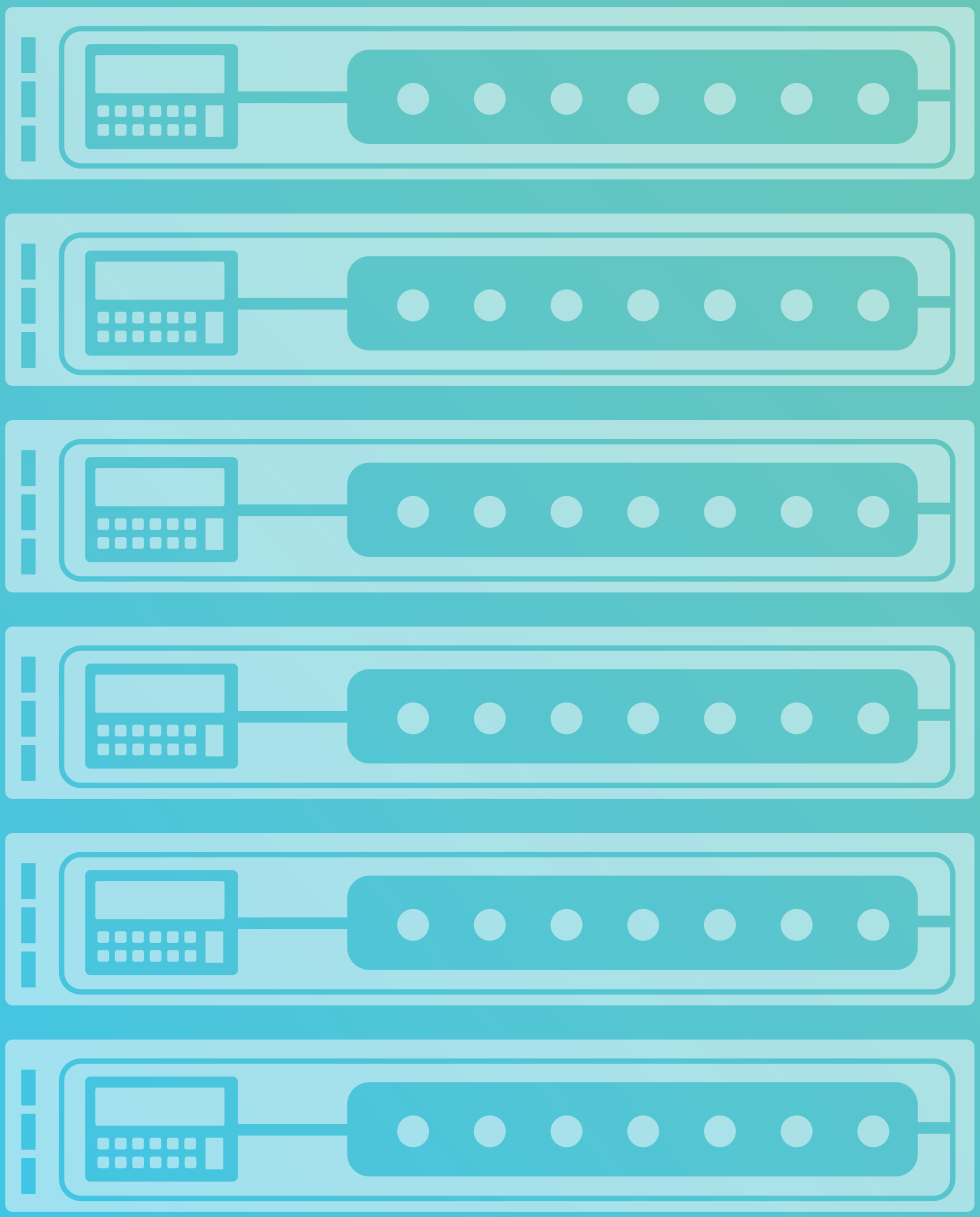
As a vendor-agnostic partner, SourceCode provides the flexibility to select the best components for optimal performance. Our streamlined co-design process enhances efficiency, saving both time and cost.

Additionally, SourceCode's in-house, U.S.-based environmental testing lab ensures that your hardware meets top-quality standards and achieves the necessary certifications, providing unmatched reliability for your most critical systems.

Learn more at www.sourcecode.com or reach out to us at info@sourcecode.com.



BOSTON
Servers | Storage | Solutions



sales@boston.co.uk | www.boston.co.uk | (+44) 1727 876 100 | Boston Limited, Unit 5 Curo Park, Frogmore, St Albans, AL2 2DD, UK
Austria | Australia | France | Germany | India | Taiwan | South Africa | Switzerland | **UK** | USA

Feb - 2025