doitnow
HPC Services

White paper

Real ROI: Delivering Scalable, Reproducible, High-Performance Software Environments for Scientific Computing.

Unified, portable, reproducible, architecture-optimised software framework delivered at global scale.

EXECUTIVE OVERVIEW

As scientific computing keeps getting more complex - with new hardware, software, and hybrid deployments emerging at an unprecedented pace - organizations face a challenge: how to keep scientific software installed, updated, optimized, and reproducible across increasingly heterogeneous systems. The traditional model of deploying and maintaining separate software stacks on every cluster, cloud, and workstation wastes time, introduces inconsistencies, and undermines both performance and reproducibility.

This white paper examines how the **European Environment for Scientific Software Installations (EESSI)** delivers a unified, performance-optimized software framework that seamlessly spans on-premises clusters, cloud-bursting workflows, containerized platforms, and workstation environments. With a single, reproducible, version-controlled stack, EESSI cuts operational overhead, accelerates time-to-solution, and ensures consistent performance across heterogeneous compute environments.

Drawing from real client data and operational experience, this paper highlights how EESSI transforms scientific software into a strategic enabler, maximizing ROI and unlocking faster, more reliable scientific outcomes.

01.

Understanding the Economics of Software Management Models

TRADITIONAL DECENTRALIZED MODEL

Traditionally, HPC centers built, tuned, and maintained software stacks separately for each cluster. This leads to:

- Redundant labor across teams and facilities
- Divergent compiler and library versions across systems
- Difficulty reproducing results across systems
- Longer onboarding for new users
- Performance reproducibility due to inconsistent compilation options
- Difficult migration when hardware is replaced

The hidden cost of this model can exceed hardware expenditures when factoring in staff time, user errors, and lost productivity.

MODERN CENTRALIZED MODEL

As environments become more complex and heterogeneous, the traditional model of managing software has emerged as one of the most persistent barriers to efficiency, performance, and reproducibility. Traditional approaches no longer scale. A new strategic and unified solution is needed: one that leverages a community-driven, architecture-aware, and continuously optimized software stack.

02.

Enabling Unified Environments

EESSI introduces a unified, collaborative, globally maintained, continuously optimized, architecture-aware software infrastructure designed to run anywhere: HPC clusters, cloud VMs, containers, and laptops. Instead of maintaining separate environments for each platform, organizations gain a unified, version-controlled software framework that behaves predictably everywhere.

Key advantages:

- A consistent software framework across HPC clusters, clouds, containers, laptops
- Built-In reproducibility
- Architecture-optimized performance
- Centralized maintenance
- Seamless hybrid and cloud scaling
- Faster user onboarding
- Community-driven and shared global expertise

The economic impact is immediate:

- One build pipeline serves thousands of nodes across many institutions
- Every environment provides the same user experience, reducing user errors and portability efforts
- Performance is optimized once and reused everywhere
- Software remains usable across hardware generations

EESSI transforms software management from a costly necessity to a powerful, shared value.

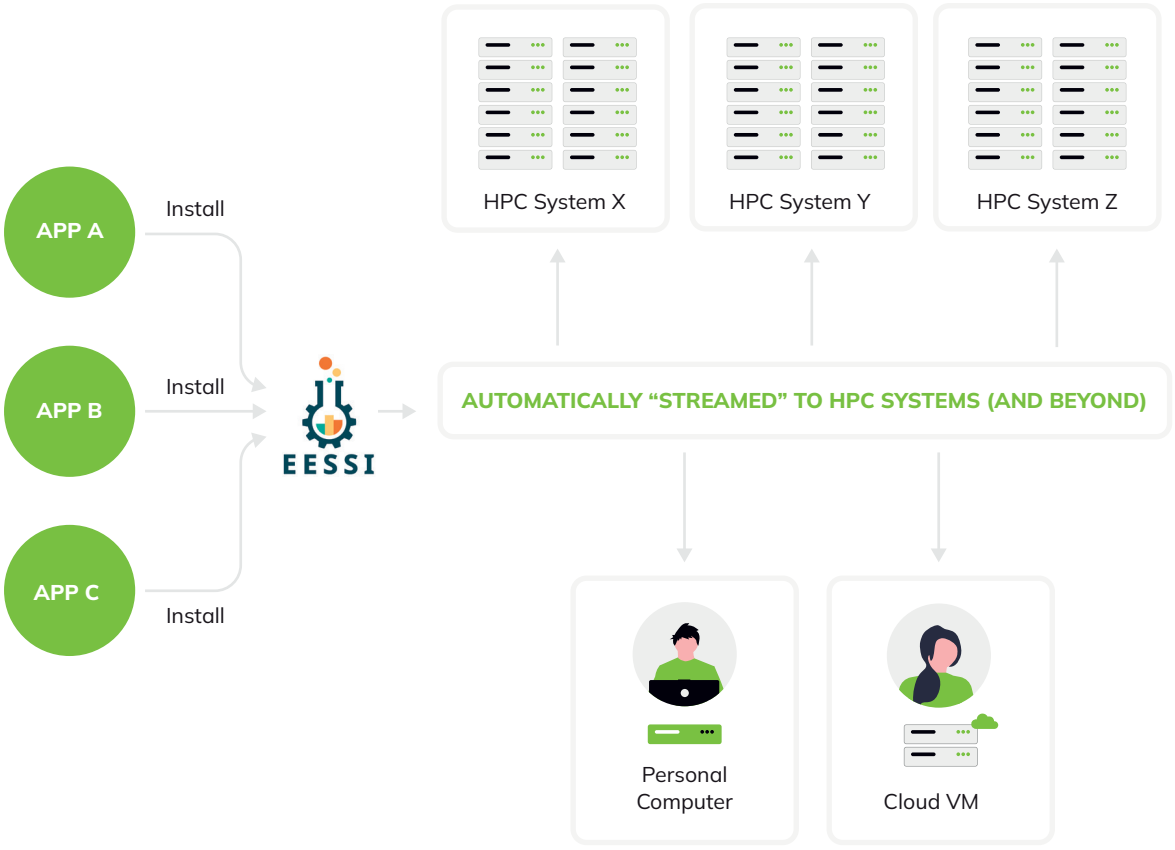

Figure 1. EESSI as a unified, collaborative and consistent software framework across HPC clusters, clouds, containers, and laptops.

03.

Foundational Challenges and How EESSI Solves Them

EESSI provides a technological foundation that directly addresses the common challenges of managing scientific software across heterogeneous environments.

3.1 ELIMINATING ENVIRONMENT INCONSISTENCY

Consistency is the cornerstone of reproducibility. Version mismatches, incompatible dependencies or divergent behavior are common issues of an inconsistent environment.

EESSI removes these risks by distributing a unified globally maintained, continuously optimized software framework. By standardizing compilers, libraries, and dependencies, EESSI ensures that workflows are portable, reliable, and reproducible. It creates a “single source of truth” for scientific software, enabling researchers to execute jobs confidently, knowing that software behavior is predictable regardless of the underlying infrastructure.

Impact on ROI:

- Eliminates duplicated software installation efforts across clusters and sites
- Reduces user errors and troubleshooting time
- Simplifies hybrid HPC and cloud bursting initiatives by providing a consistent baseline

3.2 ELIMINATING OPERATIONAL COMPLEXITY

Installing, updating, validating, and patching hundreds of scientific packages on multiple systems is labor-intensive, error-prone, and often inconsistent across teams. Manual processes can lead to configuration errors, failed builds, or incompatible libraries that compromise research workflows and increase operational overhead.

EESSI introduces a centralized management model: packages are maintained, validated, and version-controlled within the EESSI framework. Automated build pipelines and dependency resolution

reduce manual intervention, while continuous integration ensures software quality and stability across all supported architectures.

Impact on ROI:

- Significantly reduces maintenance labor and operational cost
- Minimizes human errors and failed builds
- Frees HPC administrators and researchers to focus on value-generating tasks

3.3 ENABLING REPRODUCIBILITY AND VALIDATION

Scientific workflows must be reproducible across hardware generations, sites, and time. Divergent software environments can lead to inconsistent results, hinder collaboration, and compromise research integrity. Validating experiments becomes challenging if software stacks differ even slightly between the systems.

EESSI enforces reproducible builds with strict version control and continuous validation. Packages are tested and certified across architectures as a community effort, ensuring that workflows can be reliably repeated. This capability underpins rigorous scientific collaboration, facilitates compliance, and enables reproducible AI and HPC research at scale.

Impact on ROI:

- Reduces risk of invalid or irreproducible results
- Accelerates scientific discovery and cross-site collaboration
- Provides auditability and compliance support for sensitive workloads

3.4 ELIMINATING PERFORMANCE VARIABILITY

High-performance computing workloads are highly sensitive to the underlying hardware. Differences in CPU microarchitecture, GPU generation, or accelerator capabilities can dramatically impact execution speed, memory throughput, and parallel scaling. Manual compilation and optimization for each platform are costly, time-consuming, and difficult to maintain, often resulting in suboptimal utilization, wasted compute cycles, and inconsistent results.

EESSI provides pre-optimized, architecture-aware binaries tailored for a wide range of CPU and GPU microarchitectures. Each software package is built and validated with the specific characteristics of the target hardware in mind. Workloads automatically leverage the most efficient instruction sets and

accelerator features, without requiring manual tuning or rebuilds. This ensures consistent, predictable performance across clusters, cloud, and edge environments.

Impact on ROI:

- Unlocks peak hardware utilization and computational throughput
- Reduces runtime variability and resource waste
- Optimize energy consumption and lowers total cost of compute

3.5 FOSTERING SCIENTIFIC COLLABORATION AND INNOVATION

EESSI is more than a software distribution framework. It is a catalyst for collaboration in modern scientific research. By providing consistent, reproducible, and accessible software environments across multiple research institutions and teams, EESSI enables researchers to focus on science rather than software setup and management.

Impact on ROI:

- Eliminates delays and accelerates joint projects, directly increasing research throughput.
- Minimizes repeated software builds, maintenance, and troubleshooting across collaborating sites.
- Facilitates sharing of workflows, benchmarks, and pipelines, multiplying the value of infrastructure investments across teams and projects.

04.
Enabling Technologies Behind EESSI

4.1 HPC PACKAGE MANAGERS

EESSI uses HPC package managers such as EasyBuild to provide consistent, collaborative, reproducible, and architecture-aware software stacks. HPC package managers enable:

- Recipe-based builds for complex scientific software
- Automated management of compilers, libraries, and dependencies
- Architecture-optimised binaries
- Standardised, reproducible builds across clusters and sites
- Integration into the Global EESSI Distribution

By adopting these package managers, organisations inherit a mature, community-driven set of build recipes and optimisations, reducing local engineering effort and operational risk.

4.2 AUTOMATED BUILD PIPELINES

EESSI uses EasyBuild build pipelines that automate the execution of package manager recipes across multiple architectures, generate binaries continuously, test them, and distribute them at scale. They enable:

- Continuous integration and continuous delivery (CI/CD)
- Automated dependency resolution
- Binary caching and distribution
- Multi-architecture builds triggered automatically
- Continuous QA testing
- Automated performance benchmarking

Organizations adopting EESSI inherit a continuously validated, architecture-aware software stack built from thousands of hours of expert engineering, testing, and optimization.

4.3 MODULAR SOFTWARE LAYER

Modular Software Layer (MSL) is the core organizational principle behind EESSI, separating scientific software from underlying system dependencies. MSL ensures:

- Architecture-specific optimization paths
- Identical user experience on every system
- Seamless interaction with environment modules
- Transparent updates without system rebuilds

The result: software is both portable and performance-tuned.

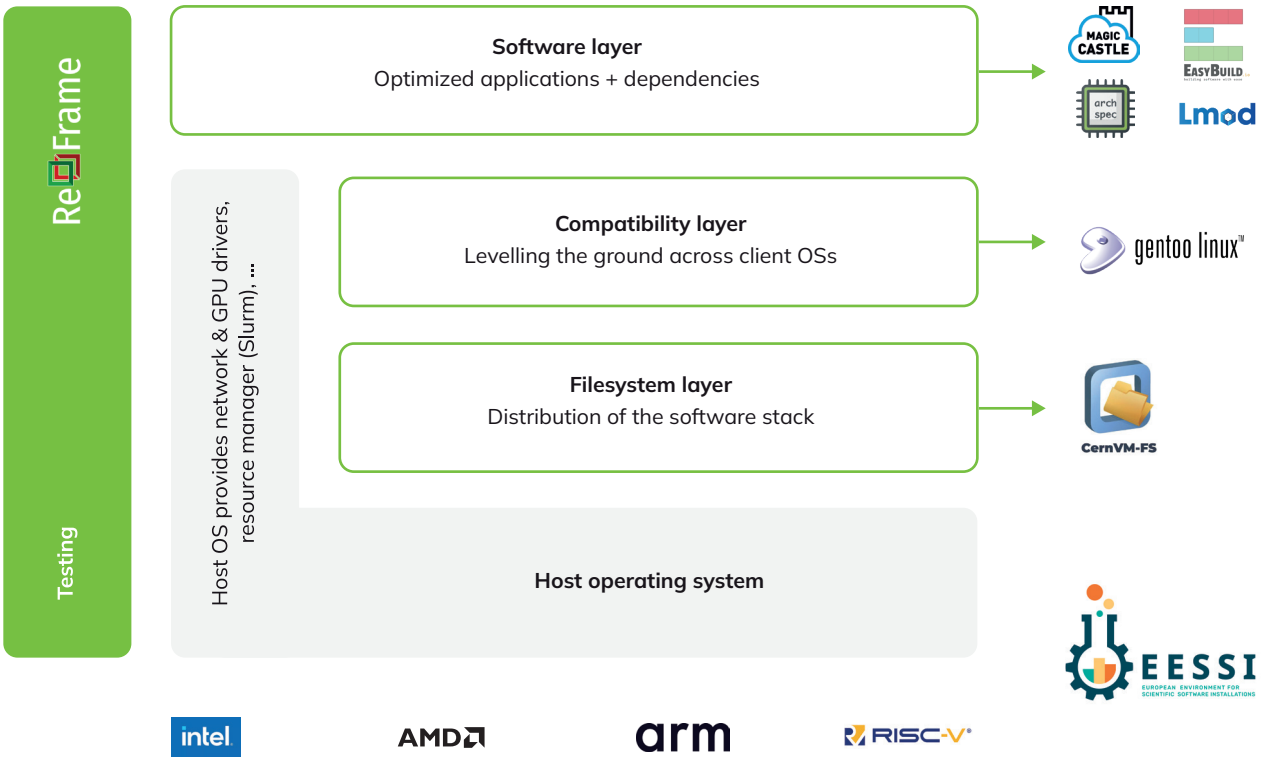

Figure 2. Overview of the layered design of EESSI. EESSI's architecture is composed of four main layers: 1) Host operating system, 2) Filesystem layer, 3) Compatibility layer and 4) Software layer. The software layer contains the actual optimized scientific software applications and their dependencies.

4.4 CROSS-PLATFORM PORTABILITY

EESSI's compatibility layer is the architectural foundation that decouples scientific applications from the underlying operating system, hardware architecture, and site-specific configuration. It is implemented through a common ABI-compatible toolchain, a CernVM-FS distributed software stack, microarchitecture-optimised builds, EasyBuild recipes and container-friendly runtimes. Together, these components form a stable, uniform runtime environment that supports:

- Containers (Singularity/Apptainer, Docker/OCI)
- Kubernetes integration for batch and ML workloads
- Cloud VM images (AWS, Azure, GCP)
- Deployments across HPC clusters, edge nodes, and laptops

EESSI's compatibility layer is important as it guarantees reproducibility across systems by delivering a versioned, identical global software stack via CernVM-FS. It enables seamless portability across HPC centres and cloud providers, supports heterogeneous and emerging architectures, and ensures environment stability for complex scientific workflows. This positions EESSI as the universal software substrate for distributed compute.

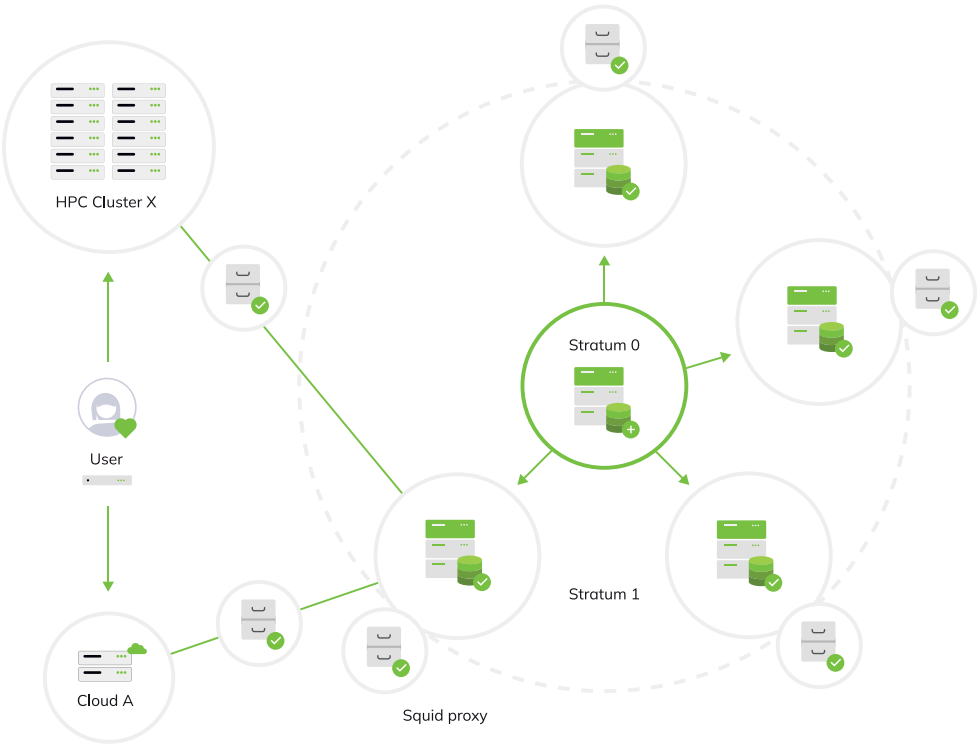

Figure 3. Overview of the CVMFS distribution architecture supporting the EESSI stack. CVMFS uses a hierarchical structure with a Stratum-0 origin, redundant Stratum-1 replicas, and local Squid caching proxies to deliver the EESSI stack efficiently at scale. Clients access the software through a read-only FUSE mount and benefit from the distributed caching layers.

4.5 VERSION CONTROL AND PROVENANCE TRACKING

- EESSI ensures all software packages, builds, and configurations are traceable and auditable:
- Strict versioning for all modules and dependencies
- Metadata and provenance records for every build
- Rollback and historical reproducibility capabilities

This traceability ensures transparency, reliable reproducibility, regulatory compliance, and confidence in workflows across all environments.

05.
Performance Optimization: The EESSI Advantage

Performance is a critical driver of cost efficiency, resource utilization, and time-to-insight. Inefficient software can underutilize hardware, increase energy consumption, prolong compute time, slow down scientific discovery, and inflate operational expenses. EESSI addresses these challenges by delivering a performance-optimized, architecture-aware software framework that maximizes the value of every CPU, GPU, and accelerator cycle.

In particular, **microarchitecture awareness**, namely compiling software for the exact CPU or GPU family, is essential to maximize throughput and minimize energy consumption. Every instruction set upgrade (e.g., AVX-512, SVE) increases FLOPs per cycle. Ignoring these features leaves new hardware underutilized and delays ROI on infrastructure investments.

Also featured out-of-the-box in the **EESSI** distribution, software compiled with microarchitecture awareness unlocks performance optimization that converts sunk cost into active productivity.

Microarchitecture	Year	Instruction Set	DP FLOPs/Cycle	Speedup vs 2008
Next Gen	?	AVX-1024	64	16
Diamond Rapids	2026	AVX-10.2	32	8
Skylake	2015	AVX-512	32	8
Haswell	2013	AVX2	16	4
Sandybridge	2011	AVX	8	2
Nehalem	2008	SSE	4	–

Table 1. Evolution of x86 microarchitectures and double-precision FLOPs per cycle.
Common software deployment strategies are killing performance, efficiency and scalability. HPC package managers ensure HPC best practices, optimisation and performance:

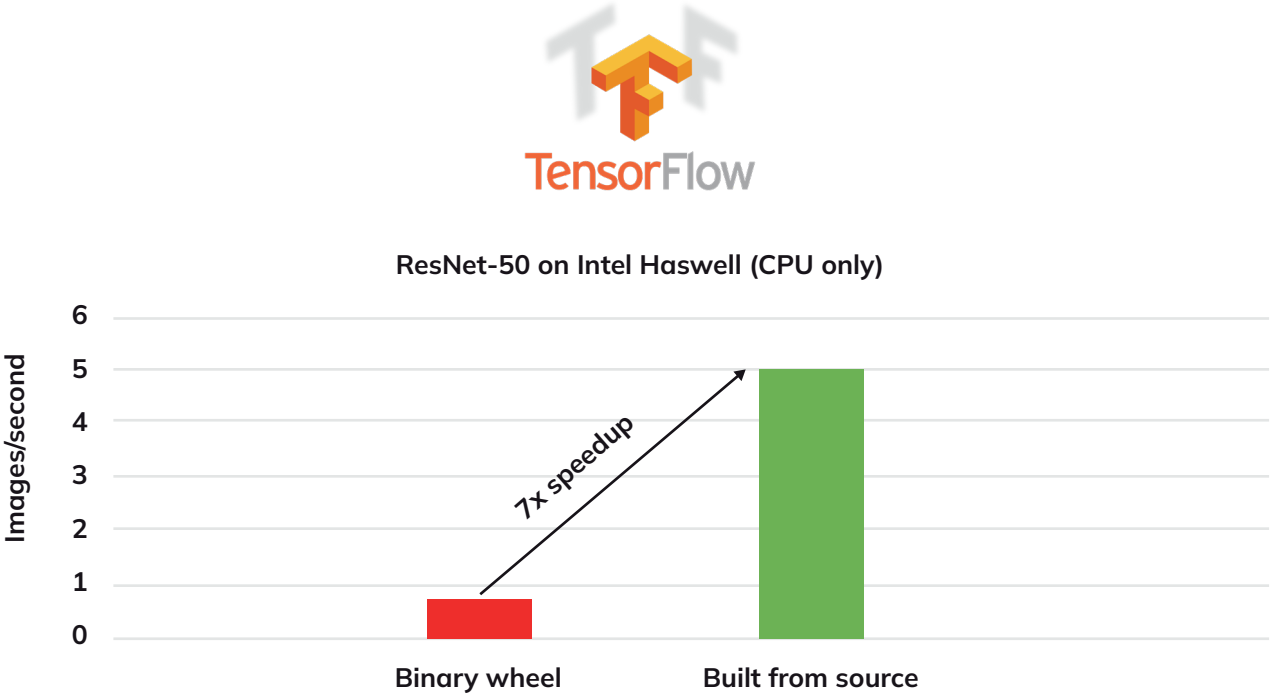

Figure 4. HPC package managers ensure HPC best practices, optimisation and performance (Tensorflow example)

	Performance Hit		
	5%	2x	8x
Lost compute capacity (A$ per year)	9,000 A$	94,400 A$	165,300 A$
Lost R&D calendar time (weeks per year)	2.5	26	46

Table 2. The long-term cost of software performance impacts in both, sustainability (cost) and time to result (R&D calendar time). Estimated cost for a 200 core server, at 0.108 A$ per core hour (AWS on-demand m7a HPC instance in Sydney).

06.
Quantifying the Real ROI

Organizations adopting EESSI achieve:

- **2–8× performance uplift** from architecture-tuned binaries
- **Significantly reduced time to scientific discovery** thanks to faster simulations
- **Drastically simplified hybrid and cloud bursting workflows**
- **Improved reproducibility and compliance** for scientific output
- **Faster onboarding** for researchers and engineers

The real ROI emerges from improved performance, saved labor, reduced compute waste, and streamlined user workflows.

07.
A Practical Roadmap to EESSI Adoption

1. **Assess:** Evaluate current software stacks, maintenance burden, and fragmentation.
2. **Deploy:** Mount the EESSI repository on on-prem, cloud, and container environments.
3. **Integrate:** Unify scientific workflows, CI pipelines, and module systems around EESSI.
4. **Optimize:** Tune scheduling and data-management policies now that environments are consistent.
5. **Scale:** Extend to laptops, cloud bursting scenarios, and multi-cluster orchestration.

EESSI adoption compounds value as more environments converge on a single, optimized software layer.

08. Conclusion

EESSI delivers more than a software stack. It provides a **strategic software framework for performance, portability, reproducibility**, and **collaboration** across the full continuum of scientific computing. By standardizing environments, optimizing binaries, and reducing operational overhead, EESSI transforms HPC & AI infrastructures into efficient, predictable, and scalable platforms.

Do IT Now helps organizations implement EESSI as the backbone of their HPC & AI strategy, enabling consistent performance, faster results, and measurable ROI across clusters, clouds, and users.

DRIVING HPC INNOVATION FORWARD

www.doitnowgroup.com

Discuss your HPC needs today

Contact us at info@doitnowgroup.com