IRON MOUNTAIN®
DATA CENTERS

AI at the carbon crossroads

3 sustainable principles for AI/HPC-ready infrastructure

IRON MOUNTAIN®
DATA CENTERS

AI at the carbon crossroads

3 sustainable principles for AI-ready infrastructure

While more traditional cloud services continue to grow fast, higher density configurations are becoming more common as GPU-driven High Performance Compute (HPC) and Machine Learning (ML) come online. And the sharpest growth and highest densities of all is coming from Deep Learning and Generative AI. According to the International Energy Authority (IEA), in 2023, NVIDIA shipped 100,000 units consuming an average of 7.3 TWh of electricity a year. By 2026, the AI industry is forecast to grow to consume at least ten times this demand. Spending in the global AI infrastructure market – including data centers, as well as networks and other hardware that supports the use of AI applications – is expected to reach $423 billion by 2029, growing at a compound annual rate of 44% for the next six years.

The growth and power demands of AI put the data center industry at a carbon crossroads, where high-density power requirements have the potential to clash with emission reduction targets, and where digital transformation threatens to derail the energy transition

Sustainability targets loom, and new regulations have been launched, making GHG reporting obligatory. These include the Corporate Sustainability Reporting Directive (CSRD) which covers over 75% of EU company turnover as well the EU Energy Efficiency Directive and evolving DOE and SEC requirements in the US.

Even more importantly, data center users want to report their emissions performance, with more than 23,000 companies - representing $67 trillion in market capitalization – disclosing their emissions through CDP. According to Accenture, 37% of companies have now set net zero targets.

At the same time, Rystad Energy research estimates that over $3 trillion in investment is required before 2030 to update outdated and inadequate power grids around the world if we are to limit global warming to 1.8 degrees Celsius above pre-industrial levels. These constrained power grids cannot support the shift to AI at the same time as wholesale electrification takes place.

Bearing these factors in mind, this e-book outlines three critical sustainable principles that data center users of all sizes should consider when planning their HPC or AI infrastructure needs:

1. Redefine renewables
2. Transform operational efficiency
3. Tackle embodied impact

1 Redefine renewables

Rigorous carbon reduction measures are the key to a genuinely green digital future. AI in the current climate makes the use of low-carbon energy a critical consideration.

Most hyperscalers and a growing number of colocation providers have been growing the green grid and eliminating carbon. Data center operators are already the biggest buyers of renewables in the world, and the more sustainable colocation providers are up there with them. Iron Mountain is now one of the world's top 20 renewable buyers. The vast majority of renewable power is purchased via Power Purchase Agreements (PPAs) which offset actual power used against renewable generation, encouraging the growth of new renewable generation projects. Results so far have been impressive, but there is much more to do.

Wind, solar and hydro are not enough. With the steep growth in power demand from growing cloud services and AI, other low-carbon power sources which are not strictly renewable are also being deployed. These include hydrogen, fission and nuclear power, which can deliver constant power feeds round the clock. Today, 18.2% of US data centers use nuclear energy, according to the US Energy Information Administration, and major new nuclear contracts such as Microsoft's on Three Mile Island will power new AI architecture.

IRON MOUNTAIN DATA CENTERS CUE 2023: 0.01

CUE (Carbon Usage Effectiveness) tracks progress toward achieving net zero carbon. This is a measure of total reported greenhouse gas emissions divided by IT load. The goal is to get this number to zero. The IMDC score is very low because we cover 100% of our electricity with renewable PPAs. The only material emissions that we have are from the testing of our backup power generators. Customers can report their share of these low-carbon figures as their Scope 3 emissions using IMDC's accredited Green Power Pass solution

United Nations

"24/7 Carbon-free Energy (CFE) means that every kilowatt-hour of electricity consumption is met with carbon free electricity sources, every hour of every day. It is both the end state of a fully decarbonized electricity system, and a transformative approach to energy procurement, supply, and policy design that is critical to accelerating its arrival."

Total decarbonization

Following methodology established by Google, Iron Mountain Data Centers (IMDC) has gone a step further, targeting not just 100% renewables via PPAs but 24/7 carbon-free power use or #247CFE. To supplement on-site generation, we are now matching a growing portion of our site electricity use with local clean power generation every hour of every day.

Transmitting power over a long distance certainly works, but power is lost in transmission, and most PPAs do not guarantee that renewables are being generated on the same grid, leading to uneven long-term capacity. Having generation closer to the point of consumption is also more power- and cost-efficient. Renewable energy is becoming more and more price-competitive, and installing a power-generating facility or deploying local renewable energy-generating infrastructure can greatly reduce long-term power costs for high-power deployments.

24/7CFE will in time replace the current year-by-year renewable Power Purchase Agreement model. IMDC is currently the only global data center colocation provider to have committed to this total decarbonization model. In 2023 we made significant progress in using locally-sourced hourly carbon-free energy, achieving 186,457 MWhrs matched with locally produced carbon free energy for our initial five sites in the US and UK - almost 20% of total power overhead. Our target is to run all data center operations on local 100% carbon-free energy by 2030.

Total decarbonization means restructuring power purchasing to use local low-carbon energy. In this video Chris Pennington, IMDC Director Energy & Sustainability, talks to Sunrock, whose solar PV installations in Rotterdam and Oud Gastel are providing 4.4 MW and 1.3 MW respectively to our Amsterdam customers.

IRON MOUNTAIN®
DATA CENTERS

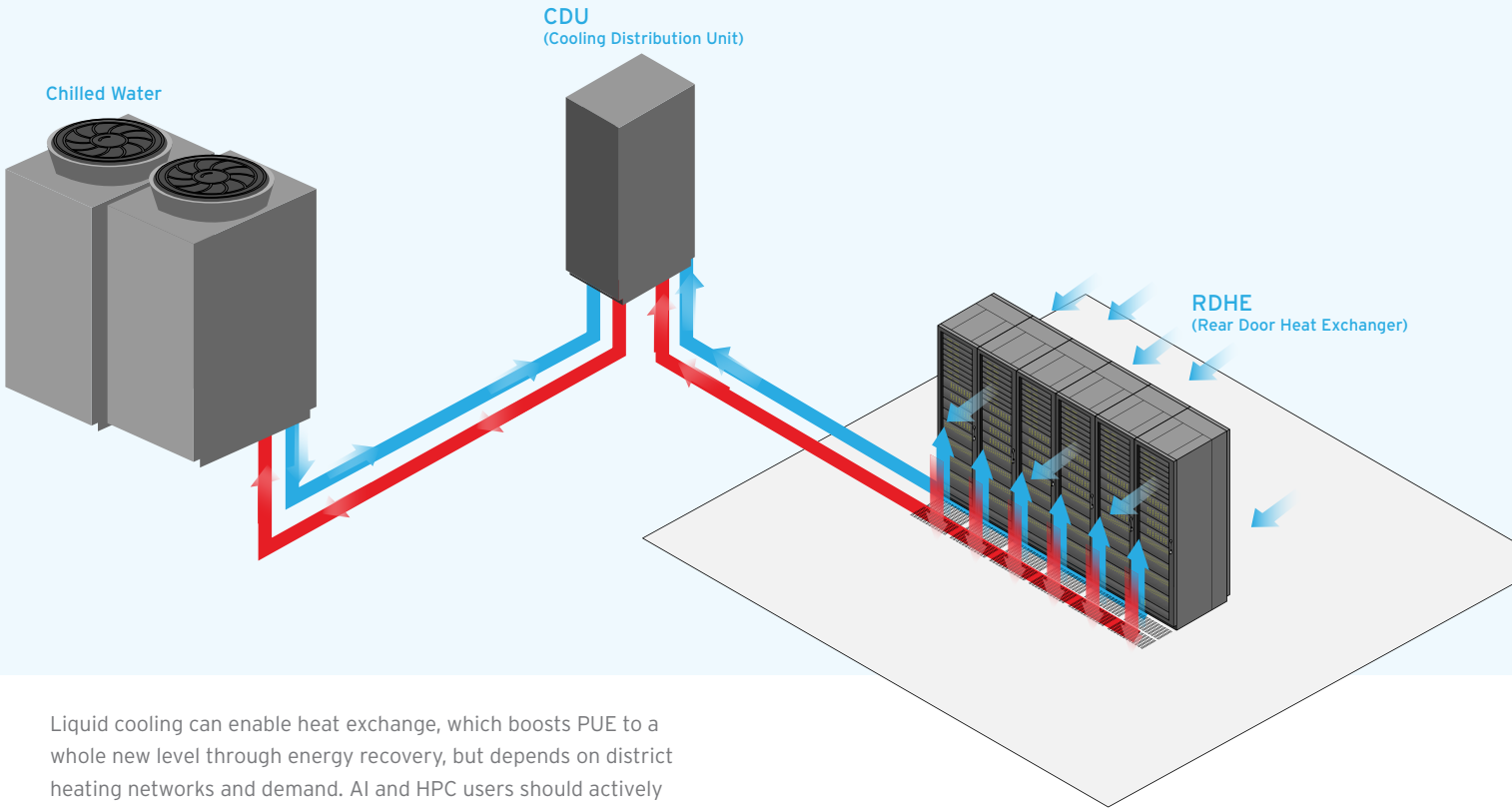
2 Transform operational efficiency

Optimizing efficiency is the baseline for sustainable operations. GPU chip redesign has the potential to improve compute efficiency hugely, and data center operators need to match these efficiencies on the ground, optimizing Power Usage Effectiveness (PUE). Again, progress is being made. Denser servers with high-powered GPUs have reduced cooling requirements compared to CPUs. These days server racks that used to handle 5 kW of power can operate at 20-30 kW with the same amount of cooling due to advances in server technology. But the huge increase in overall data center GPU density raises overall power and cooling demands.

Liquid cooling

The density of an AI or HPC deployment determines its unique cooling needs and air can only absorb a certain amount of heat. Deep learning and Generative AI require alternative solutions due to their complexity, and these will need specialized infrastructure such as direct liquid cooling (DLC), air-assistant liquid cooling (AALC), and/or a rear-door heat exchange. With liquid cooling, water or another contact fluid flows through the racks. Many consider this a key enabling technology for AI in data centers, although it is still not widespread with the vast majority of users deploying it for only a small proportion of their IT load.

Because it operates in a closed circuit, liquid cooling can lend itself to heat exchange. With Direct Liquid Cooling, the temperature of the waste heat rises, making it easier to reuse, and heat transfer between the data center and district heating becomes more energy-efficient than air cooling. With mineral oil, it also uses far less water, and wider use could put an end to traditional water-based cooling systems in data centers.



Liquid cooling can enable heat exchange, which boosts PUE to a whole new level through energy recovery, but depends on district heating networks and demand. AI and HPC users should actively seek facilities with heat exchange in place.

IRON MOUNTAIN®
DATA CENTERS

Micro grids

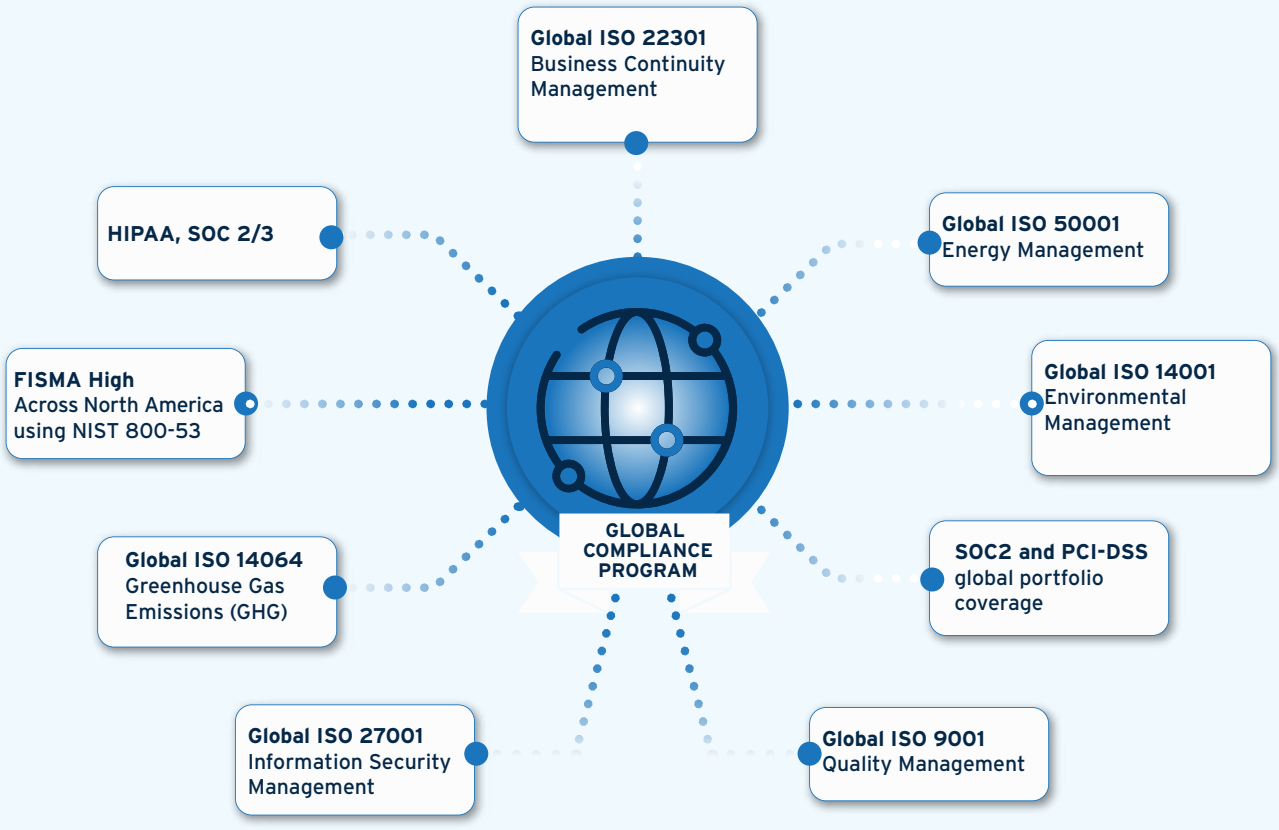
Where sizable renewable sources or backup capacity are required - often a more modern AI-ready facility - microgrids can improve efficiency and harmonize with and help stabilize constrained macro grids. Microgrids can run independently but also exchange energy with others nearby, using batteries and advanced software to enable shifting less time-sensitive tasks, such as training AI models, to periods of low power demand.

The growing role of AI

AI has a key role to play in this process due to its ability to process huge amounts of data. It can automate many aspects of microgrid operation, from routine maintenance checks to complex energy management tasks. This automation can reduce the need for manual intervention, cut operational costs, and enhance overall efficiency.
AI can also assist with HVAC efficiency. IMDC piloted the DCverse AI-Powered Cognitive Digital Twin in its Singapore facility, using advanced physics-informed machine learning algorithms and predictive analytics to deliver optimized recommendations for cooling. This delivered a PUE improvement of 3.5%. This solution has the potential to deliver similar levels of improvement at other IMDC facilities and more widely across the data center industry.

IMDC PUE 2023: 1.44

PUE (Power Usage Effectiveness) measures how much electricity is going to our customer IT equipment compared to electricity spent on cooling. The goal is to minimize energy for cooling and maximize energy for IT equipment. The Climate Neutral Data Center Pact has a goal of 1.3 to support a sustainable digital future. We construct all new facilities to achieve this.



Customers need to be able to trust the performance figures and methodologies providers use as they take these figures and put them into their own reports. Trust is created through transparency and use of common standards. For reporting GHG emissions, IMDC has aligned its efforts with the ISO 14064-1 standard.

IRON MOUNTAIN® DATA CENTERS

3 Tackle embodied impact

Data center construction comes with a high carbon price tag. Even more costly to the planet are the rare earths and metals in the servers themselves and the embodied impact of their manufacture.

Reduced carbon construction

New standards are emerging to reduce the embodied and operational impact of data center builds. These will become mainstream as the HPC and AI capacity boom gathers pace.

In 2022 IMDC was the first data center provider to earn the BREEAM (Building Research Establishment's Environmental Assessment Method) design certification in North America with its Phoenix AZP-2 data center. Certification is based on performance across a wide range of categories, from low-carbon concrete and construction methods, energy and water efficiency, ecology and the health and wellbeing of future occupants. All new multi-tenant facilities now have to achieve BREEAM certification, with active projects underway in Arizona, Chicago, Virginia, and London.

The e-waste mountain

New chips and superfast GPUs will drive the AI revolution, but what will happen to the old ones? Worldwide, the annual generation of e-waste is rising by 2.6 million tons annually and is on track to reach 82 million tons by 2030. And considering the projected growth of the AI server market, which Foxconn estimates will reach $150 billion by 2027, the e-waste crisis is poised to escalate dramatically. To put this figure in context, the total server market was $123 billion in 2022 and is poised to grow to $186 billion in 2027 (IDC). This would mean that the market for AI servers will match the market of traditional servers in just four years.

Despite the rarity and value of e-waste components and materials, according to McKinsey 89% of organizations currently recycle less than 10% of their hardware. For both efficient performance and impact reduction, AI providers will need to check that IT asset lifecycle optimization and recycling, remarketing and secure disposal are applied across their hardware base.

As with operational efficiency, AI is well placed to help solve some of the problems it creates here by logging, tracking, and disassembling components and assisting with their redeployment.

3M+
Drives sanitized per year

$1B+
Generated for clients via remarketing

16M+
Assets processed per year

Iron Mountain ALM, our global integrated hardware asset management and ITAD business complements our data center focus on sustainability at a time when large-scale solutions are vital. It will enable customers to maximize lifetime value of servers through secure data deletion, asset disposal, reselling and remarketing, to promote and drive value from increased circularity. The broader capability we can now offer - unique in the data center industry- will help customers optimize the total cost of IT asset ownership at the same time as cutting GHG impact, in some cases by up to 25%.

IRON MOUNTAIN®
DATA CENTERS

Putting your principles to work

With the rise of AI, the significance of low-carbon design build and operations has moved to the next level. It is no longer about doing one or two things well and papering over the gaps; it is about doing everything all at once, from the outset.

1 Redefine renewables

The low-carbon energy market is booming and AI investors are helping to drive it. Research the right mix of power to run your applications 24 hours a day, and make sure the facilities you use integrate seamlessly with your ESG reporting.

2 Transform operational efficiency

High-density configurations are hugely expensive to run. Cost and climate impact make optimal efficiency critical. There are advantages in high-density architectures too, but only in select locations and more modern and flexibly-designed facilities.

3 Tackle embodied impact

Low-carbon construction will become mainstream and should be non-negotiable for new builds. Perhaps even more important is secure, circular IT asset disposition for servers which contain huge amounts of embodied GHGs and are made of rare and limited natural resources.

If you are interested in finding out more about planning, locating and activating your AI-ready data center infrastructure, please get in touch:

NL: +31 800 272 4433
UK: +44 844 417 8379
DE: +49 800 408 0000
US: 833.IRM.colo
Contact Us

IRON MOUNTAIN®
DATA CENTERS