



# Cooling the Digital Future: The Strategic Role of Fluids in High-Density Data Centres

How advanced cooling architectures and intelligent fluid management are shaping efficiency, resilience, and sustainability in the AI era

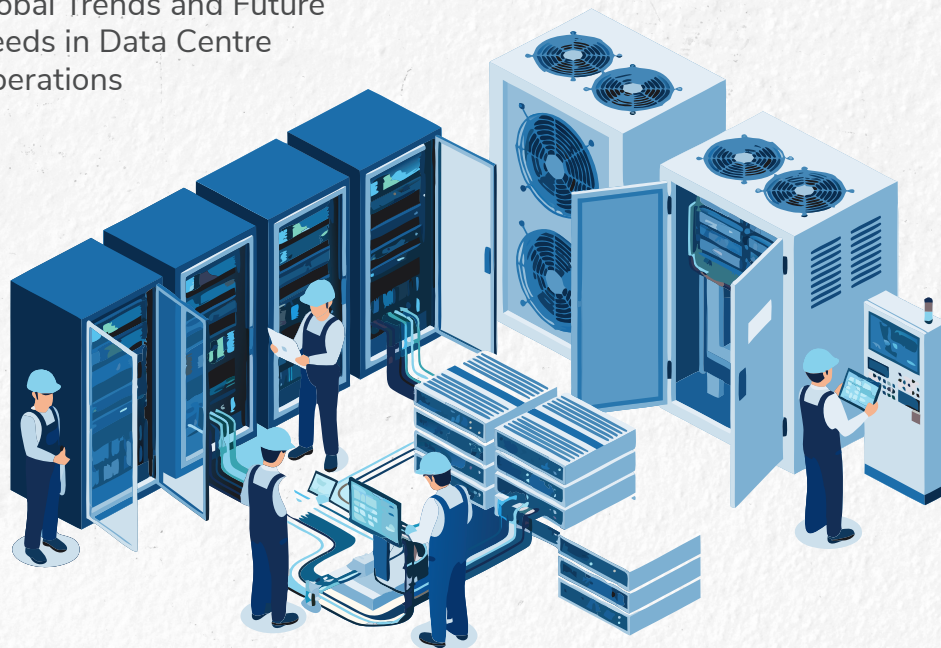
FROST & SULLIVAN WHITEPAPER

The contents of these pages are copyright © Frost & Sullivan. All rights reserved.

[frost.com](https://www.frost.com)

# CONTENTS

- 3** Why Coolants Matter in Data Centre Infrastructure
- 7** Emerging Pressures on Data Centre Uptime and Energy Efficiency
- 10** Market Context and Regional Relevance
- 14** Maintenance Risk in Servers – An Underrated Vulnerability
- 18** ESG & Efficiency – Coolant Role in Carbon and Energy Savings
- 21** Global Trends and Future Needs in Data Centre Operations
- 23** Strategic Considerations for Operators, OEMs, and Infrastructure Partners
- 24** The Fluid Supplier’s Role: From Vendor to Technical Partner
- 26** Conclusion: A Call for Deeper Collaboration
- 27** About Gulf
- 28** Sources



# Why Coolants Matter in Data Centre Infrastructure

Data centres have become a critical pillar of the global digital economy. They underpin cloud computing, artificial intelligence, digital services, financial systems, healthcare platforms, and public infrastructure. As demand for data processing and storage accelerates, the energy footprint of data centres has moved from a background concern to a strategic issue.

## Escalating Electricity Share

A decade ago, the share was closer to the **1 percent** range, as efficiency gains broadly kept pace with workload growth.<sup>[1]</sup>

Today, data centres account for approximately **1.5 to 2 percent** of global electricity consumption.<sup>[1]</sup>

That balance is now shifting.

Forecasts indicate that by 2030, data centre electricity demand could represent between **2.5 and 4 percent** of global consumption, depending on the pace of AI adoption and hardware efficiency improvements.<sup>[2]</sup>

This growth trajectory has sharpened attention on every component that influences energy efficiency, reliability, cost and sustainability within data centre operations.



## System-Level Consequences of Growth

The scale of projected growth has system-level consequences.



Global data centre electricity demand, currently estimated at roughly **460 to 500 terawatt-hours annually**,<sup>[1]</sup> could approach **800 to 1,000 terawatt-hours by 2030**<sup>[2]</sup> under high-AI adoption scenarios, representing a potential increase of **70 to 100 percent within a decade**.



Individual AI campuses are increasingly being designed at **300 to 1,000 megawatts**,<sup>[3]</sup> equivalent to the annual consumption of a mid-sized city, placing material pressure on generation and transmission infrastructure in constrained markets.



Global data centre water withdrawal is estimated at approximately **220 to 250 billion litres per year**,<sup>[4]</sup> and water usage intensity typically ranges between **0.2 and 1.8 litres per kilowatt-hour**<sup>[5]</sup> depending on climate and cooling configuration.



Higher-density AI deployments can increase facility-level cooling energy requirements by **10 to 20 percent**<sup>[6]</sup> if thermal infrastructure is not redesigned.



Retrofitting legacy facilities to support these loads frequently requires reinforcement of power distribution and cooling systems, with capital expenditure premiums commonly in the **15 to 30 percent**<sup>[6.1]</sup> range compared with purpose-built, AI-ready developments.



## Cooling as an Operational Constraint

Within this context, cooling has become a defining operational constraint rather than a background utility.

In many facilities:



Cooling accounts for roughly **30 to 40 percent** of total energy consumption.<sup>[7]</sup>



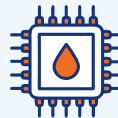
Rack densities now commonly exceed **30 kilowatts**, reaching **80 to 100 kilowatts or more** in AI deployments.<sup>[8]</sup>

## Cooling Architectures in Use

Data centres today employ:



Advanced air cooling



Direct-to-chip liquid cooling



Full immersion cooling

These architectures influence not only thermal efficiency but also facility layout.

In high-density deployments, liquid and immersion configurations can improve space utilisation per megawatt by approximately **20 to 40 percent**<sup>[9]</sup> by reducing dependence on wide aisles and large air-handling infrastructure, subject to power distribution and design constraints.

## Immersion Cooling Performance Characteristics

Immersion cooling has demonstrated measurable performance advantages in appropriate AI-intensive deployments. Compared with conventional air-cooled configurations, well-optimised immersion systems can:

▶ Reduce cooling energy consumption by approximately **30 to 50 percent**.<sup>[10]</sup>

▶ Achieve power usage effectiveness levels in the **1.03 to 1.08** range.<sup>[11]</sup>

▶ Provide more stable thermal conditions at sustained high loads.

▶ Materially reduce or eliminate onsite cooling water consumption, depending on heat rejection design.<sup>[12]</sup>

Suitability, however, remains dependent on workload profile, integration complexity, and site economics.

### Fluids as Strategic Enablers

In this environment, fluids are strategic enablers.

01

Conventional coolants support air and liquid systems.

02

Lubricants underpin mechanical resilience.

03

Immersion fluids integrate thermal control directly at the server level.

As computing intensity rises, fluid performance and stability become integral to operational, spatial, and efficiency outcomes.



# Emerging Pressures on Data Centre Uptime and Energy Efficiency

## Thermal Density Escalation

Modern AI infrastructure relies heavily on GPUs and other accelerators that generate far higher heat loads than traditional CPU-based systems.

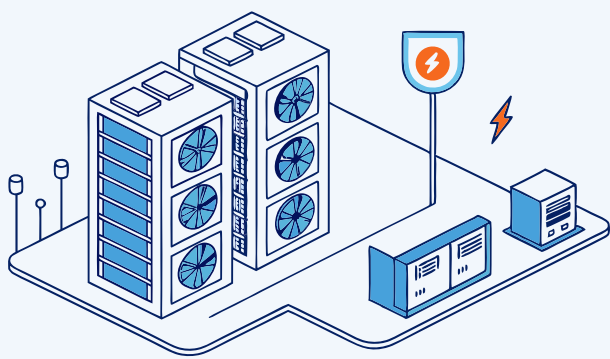
These components concentrate heat in smaller physical footprints, creating localised thermal hotspots that are difficult to manage using conventional air-based cooling alone.

As rack densities increase, even minor inefficiencies in heat removal can translate into performance throttling, or unplanned shutdowns.

## Rising Uptime Expectations

At the same time, data centres are increasingly expected to operate without interruption.

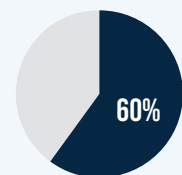
Digital services now underpin essential economic and social functions, leaving little tolerance for downtime.



Industry estimates place the average cost of a data centre outage at

approximately  
**USD 5,000 to 9,000**  
per minute, with more than

60 percent of major incidents exceeding USD 100,000 and a growing proportion surpassing USD 1 million.<sup>[13]</sup>



Planned maintenance windows are shrinking, while expectations for resilience and availability continue to rise.

Cooling system failures can directly disrupt service delivery and erode customer trust.

Continuous operation amplifies the importance of stable thermal conditions, as cooling systems must perform consistently under sustained load rather than fluctuating demand patterns.

## Geographic and Infrastructure Constraints

Geography adds another layer of complexity.

Much of the current and future growth in data centre capacity is occurring in regions with challenging climates, including hot, humid, or arid environments.

Facilities in tropical Asia, the Middle East, and other warm regions face higher ambient temperatures that reduce the effectiveness of traditional air-based cooling and increase energy consumption.

In parallel, the expansion of edge and remote data centres introduces sites with limited infrastructure, constrained power availability, and fewer options for large-scale mechanical cooling.

These conditions make thermal efficiency harder to achieve and increase sensitivity to cooling system performance.

## Regulatory and Disclosure Pressures

Overlaying these operational pressures are formal regulatory and disclosure requirements.

In the European Union, data centres above 500 kilowatts of installed IT capacity are required under the revised Energy Efficiency Directive to report annual performance metrics, including power usage effectiveness (PUE) and water usage effectiveness (WUE), into a central sustainability database.<sup>[14]</sup>

The Climate Neutral Data Centre Pact sets benchmark PUE targets for new facilities and reinforces expectations around renewable energy sourcing and water stewardship,<sup>[15]</sup> while the Corporate Sustainability Reporting Directive requires large operators to disclose energy, carbon and water performance in statutory ESG filings.<sup>[16]</sup>



## Comparative Performance of Data Centre Cooling Architectures<sup>[17]</sup>

Cooling Architecture	Typical Rack Density Range [17]	Indicative PUE Range [17]	Indicative WUE Range (L/kWh IT) [17]	Spatial Impact [17]	Water Dependency
Conventional Air Cooling	5–25 kW	1.4–1.8 (legacy), ~1.3–1.5 (modern optimised)	0.5–1.8 (climate dependent)	Wide aisles, significant air-handling infrastructure	Moderate to High (if evaporative)
Direct-to-Chip Liquid Cooling	30–80 kW	~1.2–1.4	0.2–1.2 (depends on heat rejection method)	Reduced aisle width, partial airflow reduction	Low to Moderate
Full Immersion Cooling	80–100+ kW	~1.03–1.2 (optimised deployments)	~0–0.5 (depending on rejection design)	High rack compaction, reduced airflow infrastructure	Low to Minimal (can eliminate onsite evaporative cooling)

As a result, PUE and WUE have shifted from voluntary efficiency indicators to externally scrutinised compliance metrics.<sup>[14]</sup>

Cooling performance is therefore assessed not only on operational grounds, but against regulatory benchmarks and investor reporting standards, as illustrated in the comparative performance ranges below.

### Strategic Implication

Together, these forces are redefining the problem statement for data centre cooling. Higher heat densities, continuous operation, challenging climates, and ESG pressures are converging to expose the limitations of legacy approaches.

The question for operators is no longer whether cooling needs to evolve, **but how quickly and in what direction.**

Understanding these pressures is a necessary first step before evaluating alternative cooling architectures and the role that advanced coolants may play in addressing them.



## Market Context and Regional Relevance

### Uneven Regional Pressure Landscape

These pressures are not uniform across markets.

Data centre expansion is increasingly concentrated in regions where rapid capacity growth coincides with climate stress, grid constraints, or tightening sustainability requirements.

Asia-Pacific, the Middle East, and selected emerging markets illustrate how regional growth dynamics directly shape cooling strategy and coolant deployment decisions.

### Asia-Pacific: Scale Under Climate Stress

Asia-Pacific is witnessing some of the fastest rates of data centre expansion globally, driven by cloud and AI demand.

Many markets operate in persistently hot and humid climates, where:

- ▶ Cooling energy demand can rise by approximately **15 to 30 percent** compared with temperate benchmarks.<sup>[18]</sup>
- ▶ Cooling can account for **40 to 50 percent** of total facility electricity use.<sup>[18]</sup>

Industry surveys indicate that weather-related events contribute to roughly **10 to 15 percent** of major data centre incidents globally,<sup>[19]</sup> underscoring heightened operational sensitivity during extreme heat and grid stress.

Operators therefore face the dual challenge of scaling capacity while managing structurally higher baseline cooling intensity.

## India: High-Growth Market Under Structural Constraint

India stands out as one of the fastest-growing large data centre markets globally, with installed capacity exceeding **1,200 megawatts** and a development pipeline that could more than double this over the next three to five years.<sup>[20]</sup>

Growth is being driven by:

- ▶ Hyperscale cloud expansion
- ▶ AI workloads
- ▶ Financial services digitisation
- ▶ Regulatory drivers such as data localisation requirements under India's Digital Personal Data Protection Act and sector-specific compliance mandates<sup>[21]</sup>



At the same time, operators must navigate:

- ▶ Structural power constraints in key metropolitan clusters
- ▶ Evolving state-level electricity tariffs
- ▶ Increasing scrutiny on energy efficiency from agencies such as the Bureau of Energy Efficiency<sup>[22]</sup>

Several states, including Maharashtra and Tamil Nadu, have introduced dedicated data centre policies offering incentives tied to infrastructure investment and renewable energy integration.

India's climate diversity, ranging from high-humidity coastal zones to prolonged summer temperatures exceeding **40°C** in northern regions, creates materially different thermal operating conditions across markets.

As capacity scales toward multi-hundred-megawatt campuses, cooling strategy and coolant selection are becoming central to ensuring energy intensity remains manageable under both regulatory and grid constraints.



## Singapore: Efficiency-Constrained Expansion

Singapore provides a contrasting example of tightly managed, efficiency-constrained growth.

Following the temporary moratorium on new data centre approvals between 2019 and 2022, the Infocomm Media Development Authority reintroduced capacity allocations through a competitive selection process that prioritises energy efficiency, innovation, and resource optimisation.<sup>[23]</sup>

New facilities are expected to achieve best-in-class performance benchmarks, with target PUE levels typically around or below **1.3**<sup>[23]</sup> and increasing emphasis on water efficiency and tropicalised cooling design.

Given land and grid limitations, operators are required to demonstrate high compute output per unit of power and space.

In this environment, advanced cooling architectures, including liquid and immersion systems, are being evaluated not as expansion tools, but as mechanisms to maximise performance within fixed infrastructure envelopes.

## Other Markets: Policy-Led Strategic Alignment

Other markets illustrate additional facets of regional diversity.

Malaysia is positioning itself as a regional data centre hub under the Malaysia Digital (MD) initiative and the Malaysia Digital Economy Blueprint,<sup>[24]</sup> with federal and state-level incentives in Johor and Selangor supporting hyperscale development alongside renewable energy procurement frameworks and corporate green power programmes.

The United Arab Emirates is advancing digital infrastructure under the UAE Net Zero 2050 strategy and Dubai Clean Energy Strategy 2050,<sup>[25]</sup> with regulators increasingly emphasising energy efficiency, district cooling integration, and sustainable water management in large-scale developments in Dubai and Abu Dhabi.

China continues to expand hyperscale capacity under the “Eastern Data, Western Computing” initiative,<sup>[26]</sup> which directs new data centre construction toward renewable-rich western provinces and mandates stricter green data centre standards, including efficiency benchmarks and renewable energy utilisation targets.

Each of these markets reflects a distinct regulatory and energy policy context, reinforcing the importance of cooling architectures that can align with national sustainability objectives while supporting rapid capacity expansion.

### Regional capacity, pipeline, and policy indicators based on industry reports and government disclosures.<sup>[27]</sup>

Region	Capacity (MW)	Growth Pipeline	ESG / Regulatory Pressure Intensity	Operational Cooling Impact	Strategic Capability Requirement	Strategic Cooling Implication
India	~1,200 MW	Up to ~2.5–4.5 GW by 2030	Moderate → tightening (DPDP Act, BEE, state DC policies)	Very high thermal stress; grid constraint	High-efficiency + high-temp capable fluids	Immersion & liquid tech strategic for AI clusters
Singapore	~1,400 MW	~300–700 MW contingent	High (IMDA allocation model, PUE targets, water controls)	Persistent heat & humidity stress; Land + power cap	Compact, ultra-efficient thermal systems	Performance-per-MW optimisation
UAE	~350 MW	~950 MW (2028)	High (Net Zero 2050, Clean Energy Strategy, district cooling)	Extreme heat; water scarcity pressure	Water-minimal, high-temp stable tech	Strong relevance for immersion + water-independent cooling
Malaysia (Johor)	~370 MW	~900 MW planned	Moderate (Malaysia Digital, green energy frameworks)	Rapid growth; land + energy stress	Renewable-aligned cooling	Hybrid liquid + low-water systems
China	~6,000–7,000 MW	10,000–12,000 MW by 2030–31	High (Eastern Data, Western Computing; green DC standards)	Scale + compliance pressures	Hyperscale cooling optimisation	Liquid + immersion for high CO <sub>2</sub> and efficiency programmes
APAC (Aggregate)	~13–14 GW	~16–19 GW pipeline	Rising disclosure & grid scrutiny	High cooling energy share; year-round load	Scalable coolant systems; variable density	Prioritise water-efficient + high-density solutions

Taken together, these regional dynamics underscore why cooling and coolant considerations cannot be treated as uniform or secondary decisions. Growth patterns, environmental conditions, and policy frameworks vary widely, shaping both the feasibility and attractiveness of different cooling approaches. As operators navigate this complexity, regional context becomes a critical input into decisions about cooling architecture, fluid selection, and long-term operational strategy.

# Maintenance Risk in Servers – An Underrated Vulnerability

## Hidden Sources of Operational Failure

Discussions around data centre resilience often focus on capacity, redundancy, and design standards, yet some of the most consequential risks arise from routine maintenance and overlooked systems.

In highly automated, continuously operating environments, failure is rarely the result of a single dramatic event.

More often, it emerges from gradual degradation in components that are assumed to be reliable because they are idle, redundant, or infrequently used.

Industry incident analyses indicate that between **40 and 50 percent** of significant data centre outages involve human error or process-related factors, with a substantial proportion occurring during routine maintenance or testing activities.<sup>[28]</sup>

## Standby and “Idle” Systems: Latent Vulnerability

Standby and “idle” systems illustrate this risk clearly.

Backup cooling loops, emergency power systems, and secondary mechanical components may operate only occasionally, but they are expected to perform flawlessly when called upon.

Over time:

- ▶ Fluids can degrade.
- ▶ Contaminants can accumulate.
- ▶ Mechanical tolerances can shift without obvious external indicators.

When these systems are finally activated, often during periods of peak stress or grid disruption, latent issues can surface suddenly, leading to failure on demand.

In a data centre context, such failures can escalate rapidly, compromising uptime and triggering wider operational consequences.



## Shift Toward Condition-Based Maintenance

This vulnerability has led many operators in other critical infrastructure sectors to adopt condition-based maintenance approaches.<sup>[30]</sup>

Rather than relying solely on fixed schedules or visual inspections, they use monitoring techniques to track the health of fluids and components over time.

Fluid analysis and condition monitoring provide early warning signals

- ▶ Wear
- ▶ Contamination
- ▶ Chemical degradation

Allowing interventions to be planned before failures occur.

In data centres, applying similar principles to cooling systems and associated mechanical assets can reduce unplanned downtime and improve operational predictability, particularly as thermal loads and operating hours increase.

## Where Lubricants Fit into the Data Centre Ecosystem

While advanced cooling technologies are reshaping how heat is managed within data centres, traditional mechanical systems remain fundamental to operational resilience. Lubricants play a critical role in ensuring the reliability and efficiency of these systems, many of which operate continuously or are expected to perform flawlessly during rare but high-impact events. Understanding where lubricants fit within the broader data centre ecosystem reinforces the importance of disciplined fluid management alongside emerging cooling strategies.

Diesel generators represent the ultimate line of defence against power interruptions. Although they may run infrequently, their performance is mission-critical when grid failures occur. Lubricant condition is a key determinant of generator reliability, particularly during long idle periods when oil degradation, contamination, or moisture ingress can go unnoticed. When generators are called upon under emergency conditions, any latent lubrication issues can lead to failure at the moment of greatest need.

HVAC systems and cooling compressors form another essential layer of infrastructure. These assets often operate under sustained thermal and mechanical stress, particularly in warm climates or high-density facilities. Stable lubrication supports smoother operation, reduces frictional losses, and can contribute to more predictable maintenance schedules.

Pumps, motors, blowers and fans are equally critical, yet they are frequently overlooked because of their relative simplicity and ubiquity. These components underpin fluid circulation, airflow management, and heat rejection throughout the facility. Failures in pumps or fans can quickly cascade, compromising cooling performance and triggering protective shutdowns of IT equipment.

Consistent lubrication practices, combined with condition monitoring where appropriate, help ensure these assets operate reliably under continuous duty cycles. Preventive attention to such components reduces the likelihood of minor issues escalating into major operational disruptions. In large-scale facilities, incremental improvements in performance and reliability can translate into meaningful energy savings and reduced downtime risk over time.

Taken together, these systems illustrate that lubrication strategy remains a cornerstone of data centre resilience. Integrating lubrication management into broader maintenance and monitoring frameworks supports both reliability and efficiency objectives.

### Fluid Selection and Service Life Implications

The choice of fluids has measurable implications for maintenance risk.

Conventional mineral-based oils can oxidise / age more rapidly under sustained high-temperature operation.

Engineered synthetic and dielectric fluids are formulated for continuous bulk temperatures above **50–60°C** and demonstrate:

- ▶ Significantly improved oxidation resistance
- ▶ Lower deposit formation
- ▶ Broader material compatibility with seals, polymers, and electronic substrates

Field and laboratory data in comparable high-load environments indicate that advanced formulations can extend effective service life by approximately 50 to 100 percent while maintaining stable dielectric strength<sup>[29]</sup> throughout its life.



## Environmental and Handling Considerations

From an environmental perspective, many modern fluids offer:

- ▶ Higher flash points
- ▶ Reduced volatility
- ▶ Improved biodegradability profiles

This lowers handling risk, simplifies disposal relative to legacy formulations and reduces emissions during operation.

As thermal density increases, these characteristics directly influence maintenance predictability and sustainability alignment.

## Strategic Maintenance Framing

Recognising maintenance risk as a strategic issue shifts the focus from reactive fixes to proactive management.

As facility portfolios expand and workloads intensify, the cost of unplanned outages increases, while the tolerance for disruption declines.

Integrating condition monitoring, fluid analysis, and informed fluid selection into maintenance strategies can help operators manage this risk more effectively.

Rather than being a background activity, maintenance becomes an integral part of resilience planning, supporting both reliability and long-term operational efficiency.



# ESG & Efficiency – Coolant Role in Carbon and Energy Savings

## Efficiency Beyond Operations

As data centre operators face rising energy costs and intensifying sustainability expectations, efficiency is no longer viewed solely through an operational lens.

Cooling and coolant strategies now sit at the intersection of:

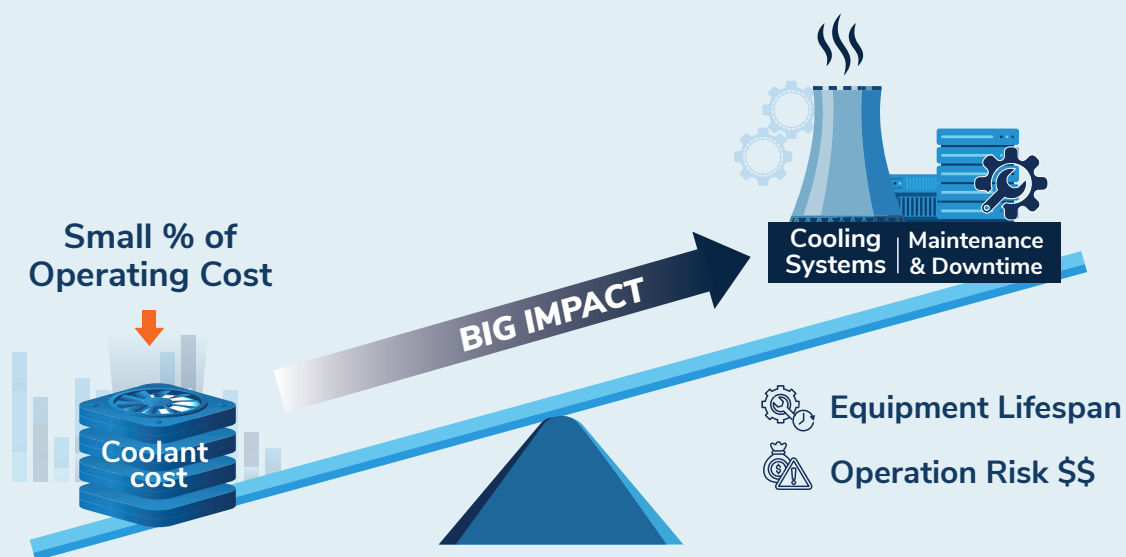
- ▶ Energy performance
- ▶ Environmental impact
- ▶ Long-term cost management

Decisions made in this area can materially influence:

- ▶ Carbon emissions
- ▶ Water use
- ▶ Maintenance intensity
- ▶ Overall total cost of ownership

Coolant and fluid expenditure typically represents a relatively small share of overall data centre operating cost.

Yet these materials directly influence the reliability and lifespan of high-value cooling and mechanical assets, creating a dynamic in which a minor cost component can exert a disproportionately large impact on maintenance exposure and lifecycle performance.



**Small Cost, Large Impact on Data Centre Performance**

## Energy Consumption and Performance Stability

More efficient heat removal reduces the parasitic load associated with cooling infrastructure, lowering overall energy demand and associated emissions, particularly in facilities operating in high densities or warm climates.<sup>[31]</sup>

Where thermal margins are tight, stable and efficient cooling can also prevent performance throttling, allowing IT equipment to operate closer to its intended capacity.

Rather than being incremental, these gains can compound over time as facilities scale.

## Water Intensity and Environmental Alignment

Water intensity has become a strategic concern alongside electricity consumption, particularly where evaporative or hybrid cooling systems rely on continuous water input for heat rejection.

In water-stressed or tightly regulated markets, this shifts water from a secondary metric to a material operational constraint.

While power usage effectiveness remains a core efficiency indicator, water usage effectiveness is increasingly scrutinised by regulators and investors.<sup>[32]</sup>

Traditional evaporative cooling can require sustained water withdrawal.

Closed-loop liquid and immersion-based architectures can significantly reduce or eliminate onsite evaporative demand, depending on rejection design.<sup>[33]</sup>

At the same time, interest is growing in:

- ▶ Biodegradable fluid formulations
- ▶ Lower-toxicity fluid formulations

That simplify handling and disposal, aligning cooling strategies more closely with environmental and sustainability objectives.



### Total Cost of Ownership Perspective

Total cost of ownership provides a useful framework for bringing these considerations together.<sup>[34]</sup>

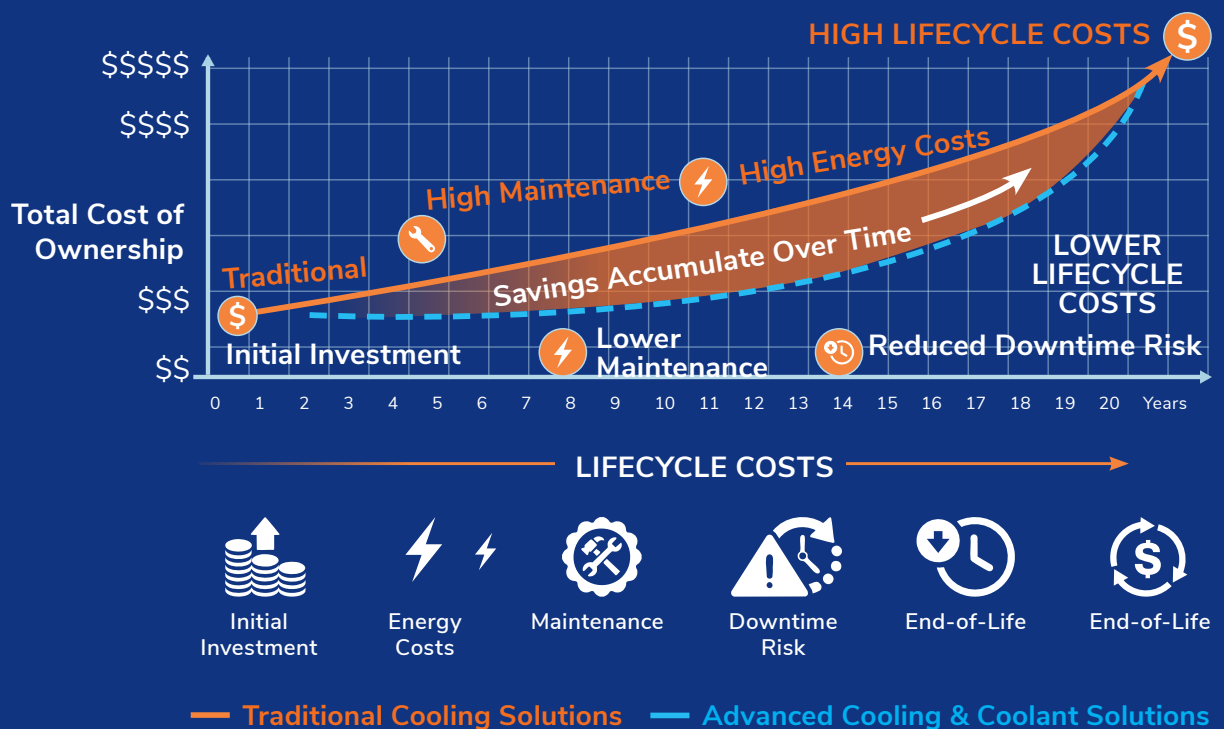
While the upfront costs associated with advanced cooling systems or specialised fluids may attract attention, they represent only one component of long-term cost.

Energy consumption, maintenance effort, downtime risk, and end-of-life disposal all contribute materially to overall economics.

When evaluated over a multi-year horizon, cooling and coolant decisions can shape cost trajectories in ways that are not immediately visible from capital expenditure alone.

Illustrative TCO comparisons highlight how improvements in efficiency and reliability can offset higher initial investments through reduced operating costs and lower disruption risk.

### Cost Trajectories of Data Centre Cooling Strategies



The diagram is for illustrative purposes only, and not based on actual figures.

# Global Trends and Future Needs in Data Centre Operations

## Portfolio Approach to Cooling Architecture

Rather than relying on a single dominant approach, operators are increasingly adopting a portfolio view of cooling technologies, selecting architectures that align with specific density requirements, climate conditions, and operational priorities.

## Architecture Spectrum

Air cooling remains viable for moderate-density deployments, but its efficiency gains are incremental and constrained by airflow physics at higher rack densities.

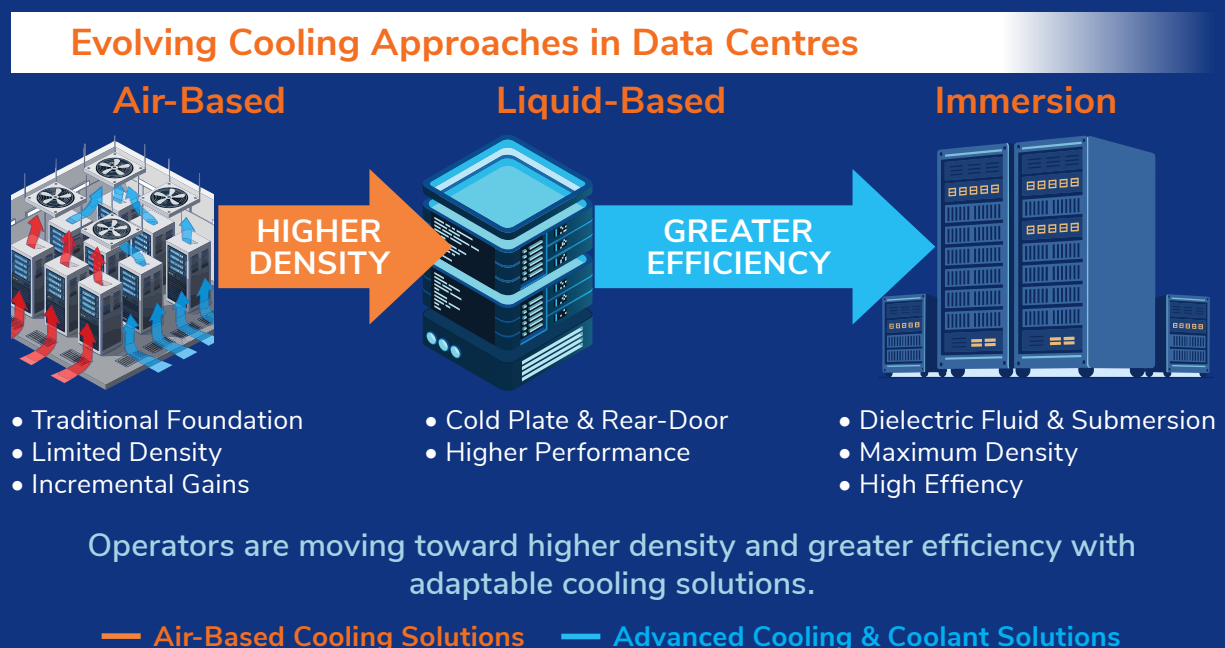
Liquid cooling extends these limits by delivering coolant directly to heat-generating components through cold plates or rear-door exchangers, enabling higher performance while retaining familiar rack architectures.

Immersion cooling represents a structural shift, submerging IT hardware in dielectric fluid to remove heat without reliance on air as the primary transfer medium.

This architecture supports:

- ▶ Very high-power densities
- ▶ Stable thermal performance
- ▶ Reduced dependence on large-scale air handling

Making it increasingly relevant as density, efficiency, and water constraints intensify.



## Standards and Regulatory Maturation

Liquid and immersion cooling deployment is increasingly shaped by formal technical and regulatory standards.

- ▶ ASHRAE thermal guidelines now define allowable operating envelopes for higher-density and liquid-cooled environments.<sup>[35]</sup>
- ▶ IEC and UL safety standards govern electrical isolation, dielectric fluid performance, and system integrity.<sup>[36]</sup>

In parallel, environmental regulation including EU REACH chemical controls, restrictions on high global warming potential substances, and tightening occupational exposure requirements are increasing scrutiny of coolant composition and lifecycle management.<sup>[37]</sup>

Together, these frameworks are formalising expectations around fluid properties, safety, and environmental impact, reducing adoption risk and accelerating structured evaluation of advanced cooling architectures.

## Forward-Looking Cooling Strategy

Looking ahead, the future of data centre cooling is unlikely to be defined by a single technology.

Instead, it will be characterised by informed choice, where operators select and combine cooling approaches based on workload demands, environmental conditions, and long-term operational goals.

In this context, coolants and fluids remain central, regardless of architecture.

Their performance, stability, and compatibility underpin the effectiveness of air, liquid, and immersion systems alike, reinforcing the need for careful consideration as data centre operations continue to evolve.



# Strategic Considerations for Operators, OEMs, and Infrastructure Partners

## Strategic Alignment of Thermal Strategy

As cooling requirements intensify, decision-makers across the data centre ecosystem must evaluate how thermal strategies align with performance, reliability, and sustainability objectives.

Immersion cooling is neither universally required nor universally optimal, but it offers distinct advantages under specific operating conditions.

The priority is determining:

- ▶ When it becomes strategically relevant
- ▶ How it should be evaluated
- ▶ How it integrates within broader infrastructure roadmaps

## Conditions for Strategic Relevance

Immersion cooling typically gains relevance when density thresholds, water constraints, or efficiency requirements begin to challenge conventional architectures. For operators, this may relate to campus expansion or AI cluster deployment. For OEMs and integrators, it may involve hardware compatibility, warranty frameworks, or system certification pathways. Evaluation should therefore extend beyond performance claims to include:

- ▶ Fluid stability
- ▶ Material compatibility
- ▶ Safety certifications
- ▶ Lifecycle management considerations

## Governance and Integration Discipline

Across stakeholders, disciplined governance is essential. This includes defined monitoring protocols, fluid sampling and condition tracking, integration with maintenance systems, and alignment with regulatory and reporting requirements. Decisions should be framed within long-term infrastructure strategy rather than short-term cost comparisons.

## Architectural Toolkit Perspective

Viewed in this way, immersion cooling becomes part of a structured architectural toolkit. Whether assessed by operators, OEMs, or system partners, its role should be defined by workload trajectory, environmental constraints, compliance context, and total lifecycle performance rather than by novelty or marketing momentum.

# The Fluid Supplier's Role: From Vendor to Technical Partner

## From Consumable to System Component

In traditional environments, coolants and lubricants have often been treated as consumables, specified once and revisited only when problems arise. In immersion and advanced liquid-cooled systems, fluids become long-life system components whose performance directly influences reliability, efficiency, and environmental outcomes.

## Fluid Performance as System Risk Variable

In immersion cooling environments, dielectric properties, thermal stability, material compatibility, and ageing characteristics all affect how effectively heat is removed and how safely equipment operates over time. Treating immersion fluids as interchangeable commodities can introduce operational risk. A partnership mindset recognises that fluid selection, monitoring, and management are integral to system design and lifecycle performance, not peripheral procurement decisions.

## Technical Services as Risk Mitigation

A technically capable fluid partner adds value well beyond product supply.

Services such as coolant sampling, diagnostics, and trend analysis help operators understand how fluids are performing under real-world conditions.

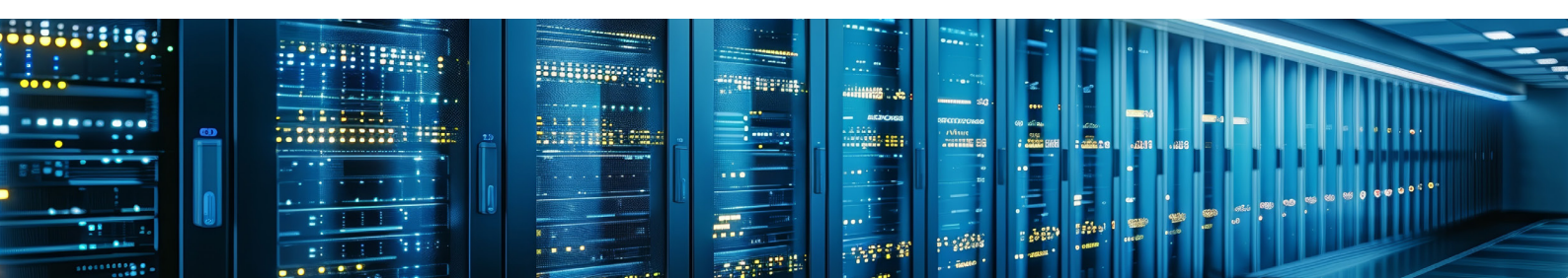
By identifying early signs of degradation or contamination, these practices support proactive maintenance and reduce the likelihood of unexpected issues.

In parallel, guidance on:

- ▶ Operating parameters
- ▶ Handling procedures
- ▶ Maintenance planning

Helps align day-to-day operations with design intent.

When delivered effectively, these services function as risk mitigation tools, supporting uptime and predictability rather than adding operational complexity.



## OEM and Design Collaboration

Collaboration with original equipment manufacturers and facility designers further strengthens this partnership model.

Early engagement during system design and deployment helps ensure that fluids are compatible with:

- ▶ Enclosures
- ▶ Materials including seals, hoses, cables and TIMs
- ▶ Thermal requirements

It also enables alignment with specific standards and certification frameworks, for example:

- ▶ Open Compute Project (OCP) immersion cooling specifications<sup>[39]</sup>
- ▶ UL 2417 safety requirements for information technology equipment using liquid cooling<sup>[40]</sup>

Reducing uncertainty during commissioning and operation.

In regions with challenging climates or unique regulatory requirements, tailored recommendations can help optimise performance while maintaining compliance.

This collaborative approach reduces downstream friction and supports smoother integration of advanced cooling systems into existing data centre environments.

## Strategic Implication

Ultimately, the shift from vendor to technical partner reflects the growing importance of fluids in modern data centre operations. As cooling strategies become more sophisticated and performance margins tighter, the quality of expertise and support surrounding fluid selection and management becomes as important as the fluid itself. Operators that recognise this dynamic are better positioned to manage risk, maintain efficiency, and adapt to evolving operational demands.



## Conclusion: A Call for Deeper Collaboration

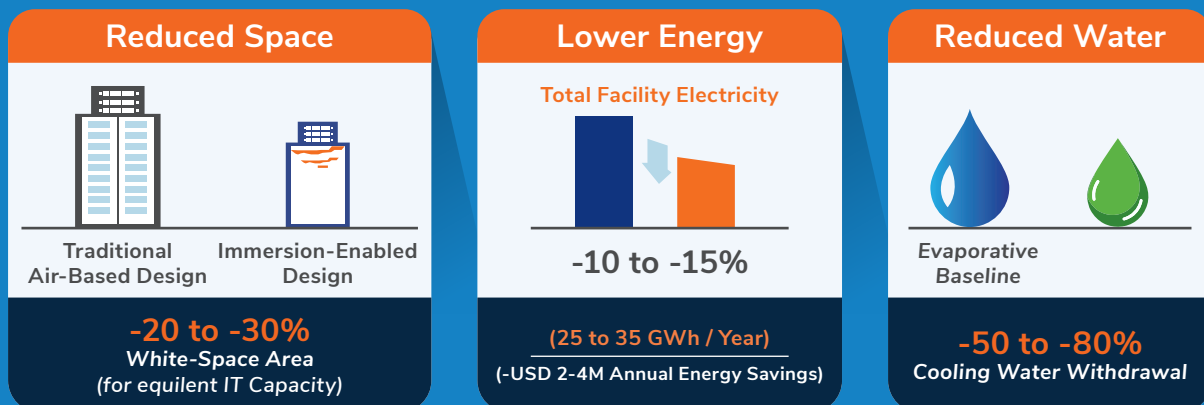
The evolution of data centre infrastructure is redefining how cooling and fluid strategies are perceived and applied. Rising computing densities, continuous operating demands, challenging climates, and intensifying sustainability expectations have elevated cooling from a background engineering function to a strategic performance variable. Within this context, coolants and lubricants are no longer passive inputs. They influence uptime, energy efficiency, maintenance intensity, and environmental outcomes across the life of a facility.

A smarter approach to coolants recognises these interdependencies. Efficient thermal management can reduce energy consumption and associated emissions, while stable operating conditions support longer maintenance intervals and lower disruption risk. At the same time, thoughtful fluid selection and monitoring contribute to safer handling, improved environmental performance, and more transparent ESG reporting. These benefits are not automatic; they emerge when cooling and fluid decisions are integrated into broader operational and sustainability strategies rather than treated in isolation.

In high-density AI deployments, these decisions can have structural financial implications. As illustrated, immersion-enabled architectures can materially reduce required white-space footprint, lower facility energy intensity, and significantly reduce cooling-related water dependency in appropriate environments. At campus scale, this combination can translate into deferred capital expansion, multi-million-dollar annual operating savings, and reduced exposure to water and carbon constraints. When evaluated holistically, cooling strategy becomes not merely a technical selection, but a lever on long-term infrastructure economics and resilience.

### Illustrative strategic impact of immersion-enabled architectures at campus scale:

#### Illustrative impact of immersion Cooling on a 20 MW High-Density Deployment



Illustrative scenario for high-density AI workloads; actual outcomes depend on design, climate, and integration.

Achieving these outcomes increasingly depends on collaboration. Data centre operators, equipment manufacturers, facility designers, and fluid specialists each bring distinct expertise to the table. When these perspectives are aligned early and maintained throughout the operational lifecycle, systems are more likely to perform as intended under real-world conditions. Shared accountability, clear standards, and continuous feedback loops help reduce uncertainty and enable ongoing optimisation as workloads and requirements evolve.

As the data centre industry continues to scale and diversify, the need for such collaboration will only intensify. By approaching cooling and coolants as shared responsibilities rather than discrete procurement choices, stakeholders can build more resilient, efficient, and sustainable digital infrastructure. In an environment defined by rapid change and growing constraints, collaboration becomes not just a best practice, but a prerequisite for long-term success.

## About Gulf



With more than 125 years of heritage, Gulf brings deep expertise in high-performance lubricants and fluid technologies that support critical industrial systems worldwide. Over the decades, we have developed advanced thermal and mechanical fluid solutions for demanding environments, building a reputation for reliability, performance, and technical partnership. Today, Gulf is applying this experience to the evolving needs of data centre infrastructure, developing next-generation lubrication and fluid management solutions designed to support efficiency, resilience, and long-term operational performance in high-density computing environments.

Find out more at [www.gulfoiltd.com](http://www.gulfoiltd.com)



## Sources

- [1] International Energy Agency, *Data Centres and Data Transmission Networks Report*, 2023–2024
- [2] McKinsey & Company, *The Internet of Things and AI Infrastructure Outlook*, 2023–2024
- [3] Hyperscaler public disclosures (Google, Microsoft, AWS infrastructure planning statements), 2023–2025
- [4] Google Environmental Report; Microsoft Sustainability Report; academic water usage studies (various), 2022–2024
- [5] Uptime Institute; academic WUE benchmarking literature, 2022–2024
- [6] Schneider Electric; Vertiv thermal efficiency studies, 2023–2024
- [6.1] Uptime Institute; Schneider Electric; McKinsey & Company, data centre retrofit vs greenfield cost analyses and AI-readiness upgrade assessments, 2022–2024
- [7] Uptime Institute; Schneider Electric energy distribution analyses
- [8] NVIDIA, Dell, and hyperscaler AI infrastructure disclosures, 2023–2025
- [9] Open Compute Project; industry deployment case studies
- [10] Immersion cooling vendor and industry performance studies (Submer, GRC, Schneider Electric), 2022–2024
- [11] Uptime Institute; OCP immersion benchmarks
- [12] Industry studies on liquid and immersion cooling water dependency (IEA, OCP, vendor disclosures), 2022–2024
- [13] Uptime Institute, *Annual Outage Analysis Report*, 2023–2024; Ponemon Institute, *Cost of Data Center Outages Report*, 2023
- [14] European Commission, *Energy Efficiency Directive (EED) – Recast*, 2023
- [15] Climate Neutral Data Centre Pact, *Self-Regulatory Initiative and Performance Targets*, 2021–2024
- [16] European Commission, *Corporate Sustainability Reporting Directive (CSRD)*, 2023

- [17] Uptime Institute; ASHRAE; Schneider Electric; Open Compute Project (OCP), data centre performance benchmarks and cooling architecture studies, 2022–2024
- [18] Uptime Institute; Schneider Electric; International Energy Agency, regional cooling energy intensity studies and climate impact assessments, 2022–2024
- [19] Uptime Institute, *Annual Outage Analysis Report*, 2023–2024
- [20] JLL; CBRE; Knight Frank, India data centre market updates and capacity outlook reports, 2023–2025
- [21] Government of India, *Digital Personal Data Protection Act*, 2023
- [22] Bureau of Energy Efficiency (India), data centre energy efficiency guidelines and policy framework, 2022–2024
- [23] Infocomm Media Development Authority (IMDA), Singapore data centre policy framework and allocation guidelines, 2022–2024
- [24] Government of Malaysia, *Malaysia Digital (MD) Initiative*; *Malaysia Digital Economy Blueprint (MyDIGITAL)*, 2021–2024
- [25] Government of United Arab Emirates, *UAE Net Zero 2050 Strategy*; *Dubai Clean Energy Strategy 2050*, 2021–2024
- [26] National Development and Reform Commission (China), *Eastern Data, Western Computing Initiative*; China green data centre policy framework, 2022–2024
- [27] JLL; CBRE; Uptime Institute; government and industry reports on regional data centre capacity, pipeline, and ESG/regulatory trends, 2023–2025
- [28] Uptime Institute, *Annual Outage Analysis Report*, 2023–2024
- [29] ASTM International; IEC standards; OEM and industrial lubricant studies on dielectric and synthetic fluid performance and service life, 2022–2024
- [30] ISO; NIST; industrial maintenance and reliability literature on condition-based maintenance practices, 2021–2024
- [31] International Energy Agency; Schneider Electric; Uptime Institute, data centre energy consumption and cooling energy share analyses, 2022–2024
- [32] European Commission; Uptime Institute; ESG disclosure frameworks and data centre sustainability reporting guidelines, 2022–2024
- [33] International Energy Agency; Open Compute Project (OCP); industry studies on water usage effectiveness and cooling water dependency, 2022–2024

- [34] McKinsey & Company; Uptime Institute; Schneider Electric, total cost of ownership (TCO) frameworks and data centre lifecycle cost analyses, 2022–2024
- [35] ASHRAE, *Thermal Guidelines for Data Processing Environments*, latest edition (2021–2024 updates)
- [36] IEC (International Electrotechnical Commission); UL (Underwriters Laboratories), safety standards for electrical systems, dielectric fluids, and liquid-cooled IT equipment, 2021–2024
- [37] European Chemicals Agency (ECHA), REACH Regulation; European Commission environmental directives on chemicals safety, global warming potential restrictions, and occupational exposure limits, 2021–2024
- [38] Open Compute Project (OCP), *Immersion Cooling Specifications and Guidelines*, latest available version, 2022–2024
- [39] UL (Underwriters Laboratories), UL 2417 Standard for Information Technology Equipment Using Liquid Cooling, latest edition, 2022–2024



## YOUR TRANSFORMATIONAL GROWTH JOURNEY STARTS HERE

Frost & Sullivan's Growth Pipeline Engine, transformational strategies and best-practice models drive the generation, evaluation, and implementation of powerful growth opportunities.

Is your company prepared to survive and thrive through the coming transformation?

Join the journey. [→](#)



FROST  SULLIVAN