Five Data Center Predictions 2024

June 13, 2024

uptime INSTITUTE

The Digital Economy

A robust digital economy can serve as a catalyst for promoting non-oil sectors that will contribute to economic growth and diversity.

A modern and flourishing digital economy should be underpinned by a resilient and sustainable digital infrastructure.

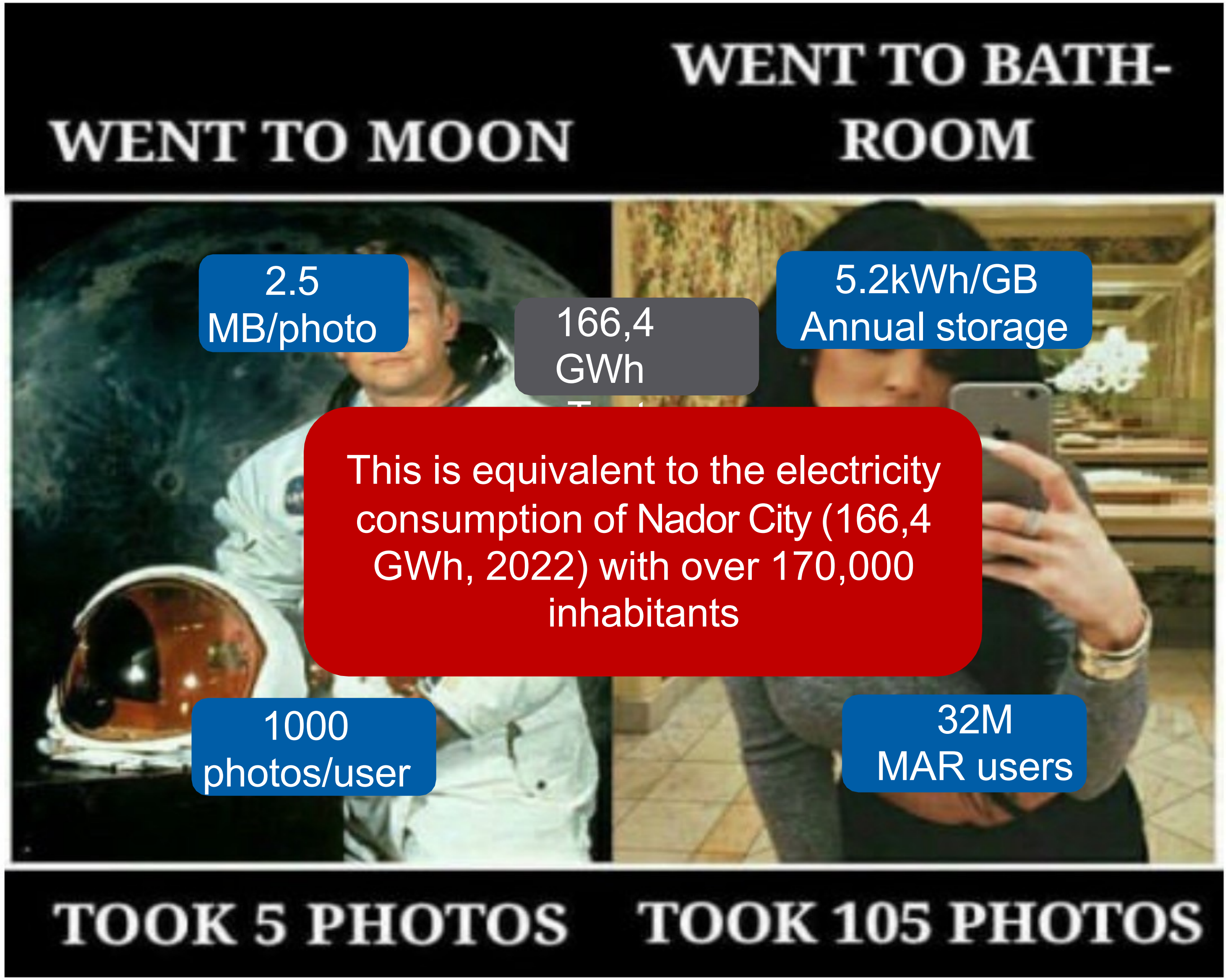
Data Centers and the Digital Infrastructure have Wider Economic Impacts

Digital Ecosystem

Clustering Effect

Knowledge Building


While data centers have helped many other sectors, their environmental impact has become a key cause for concern



(Image Source: https://wanna-joke.com/speaking-of-selfies/)

Sources
Carbon and the cloud - stanfordmag.org
Cisco Annual Internet Report - cisco.com
The Megawatts behind Your Megabytes - aceee.org

Intelligence Predictions for the Year (and Years) Ahead

We aim to highlight crucial, yet often overlooked topics

Many challenges facing the industry result from the ongoing success of the IT sector

6

2023 predictions remain valid

Effects likely to continue beyond 2024

2023 prediction	Summary
Geopolitics deepens supply chain worries	Data center supply chains are still reeling from long delays, and current geopolitical dynamics present new threats. Uptime sees particularly high risks around the advanced semiconductors that are vital to IT hardware and data center equipment, and around subsea cable systems.
Too hot to handle? Operators to struggle with new chips	Operators face various trade-offs for handling new-generation IT technologies. Driven primarily by silicon technology, data center capacity planning will need to accommodate a potentially fast-shifting balance between power, cooling and space.
Cloud migrations to face closer scrutiny	The cost of migration and the threat of spiraling cloud costs deter some mission-critical migrations in the years ahead. Regulators and executives try to understand and limit the risks associated with a high dependency on the public cloud, causing some organizations to proceed more cautiously than before.
Energy efficiency focus to shift to IT — at last	More stringent sustainability regulation and reporting requirements will force IT to deliver improved performance in energy efficiency. Fewer higher-performance, highly utilized servers could deliver major energy gains. Refusing to deploy these improvements will be increasingly difficult to justify.
Data center costs set to rise and rise	The costs of critical digital infrastructure are set to rise: a trend likely to persist in the medium term. Supply chain issues and higher cost of capital, energy and labor, have all contributed to rising prices.

7

Prediction 1

Operators – Prepare for a Sustainability Reckoning

Prediction One: Prepare for a Sustainability Reckoning

The data center industry will continue to expand strongly, using ever greater amounts of power (and sometimes water) – while reporting requirements will tighten.

Use of many types of carbon offsets will become unacceptable, forcing operators to seek locally sourced renewable energy – and cut consumption.

Uptime Intelligence is predicting a challenging period for the sector from 2024 to 2030 as organizations struggle to meet publicly stated and increasingly binding sustainability goals and reporting requirements.

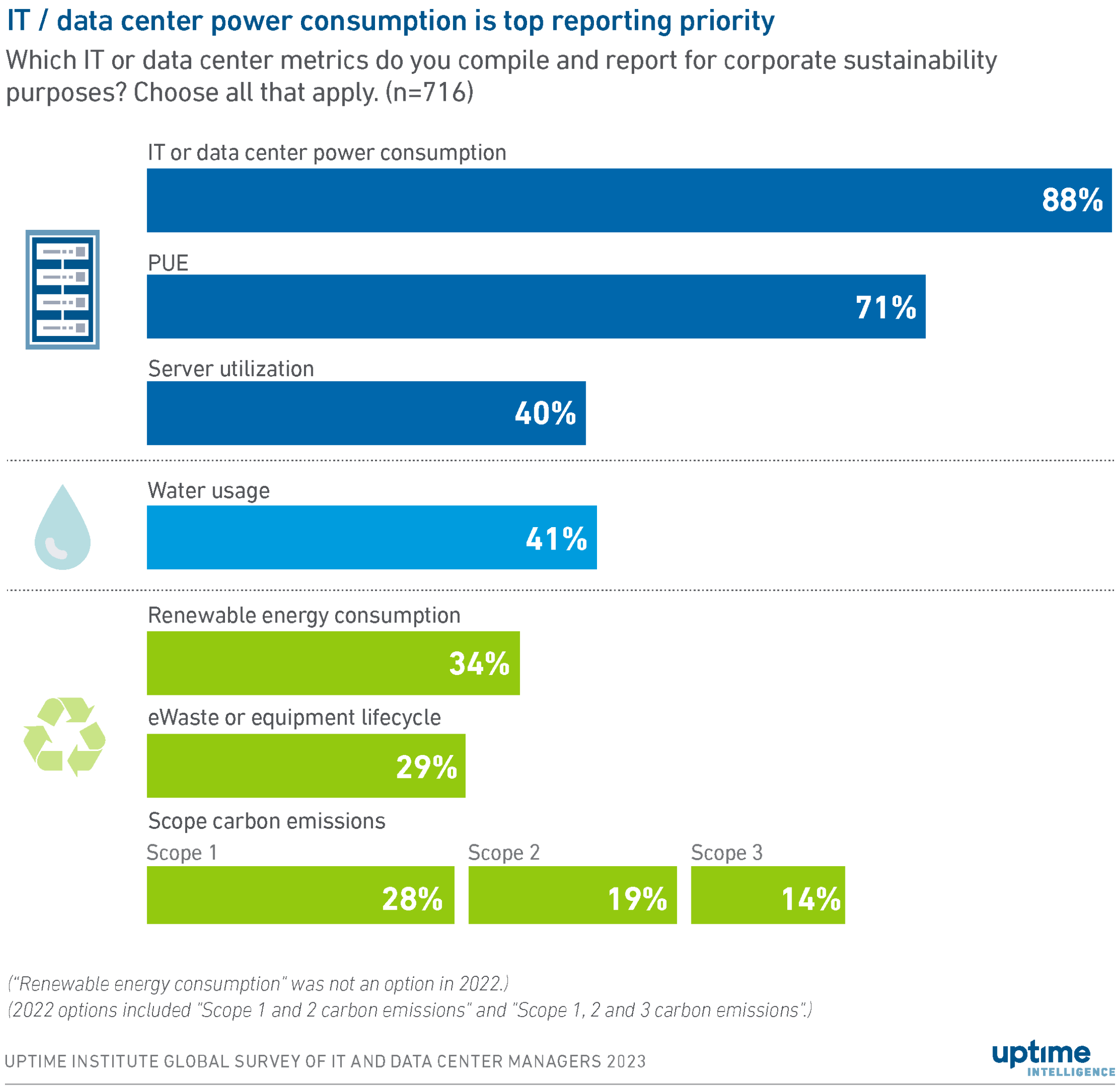
uptime® INSTITUTE

Operators are Not Ready to Report Data

Sustainability related data collection remains low and patchy.

Many companies have made Net Zero promises – but do not have the systems in place to meet reporting commitments.

Many operators are highly dependent on their local grid for achieving low carbon targets.



Prediction 2

Demand for AI Will Have Limited Impact on Most Operators

The New AI Hype Cycle

Transformers architecture (2017) is the breakthrough that started the most recent 'AI hype cycle'.

These can scale up to huge data sets and more complex neural networks, taking advantage of large compute clusters.

This created runaway expectations of demand for data center capacity and high-density racks.

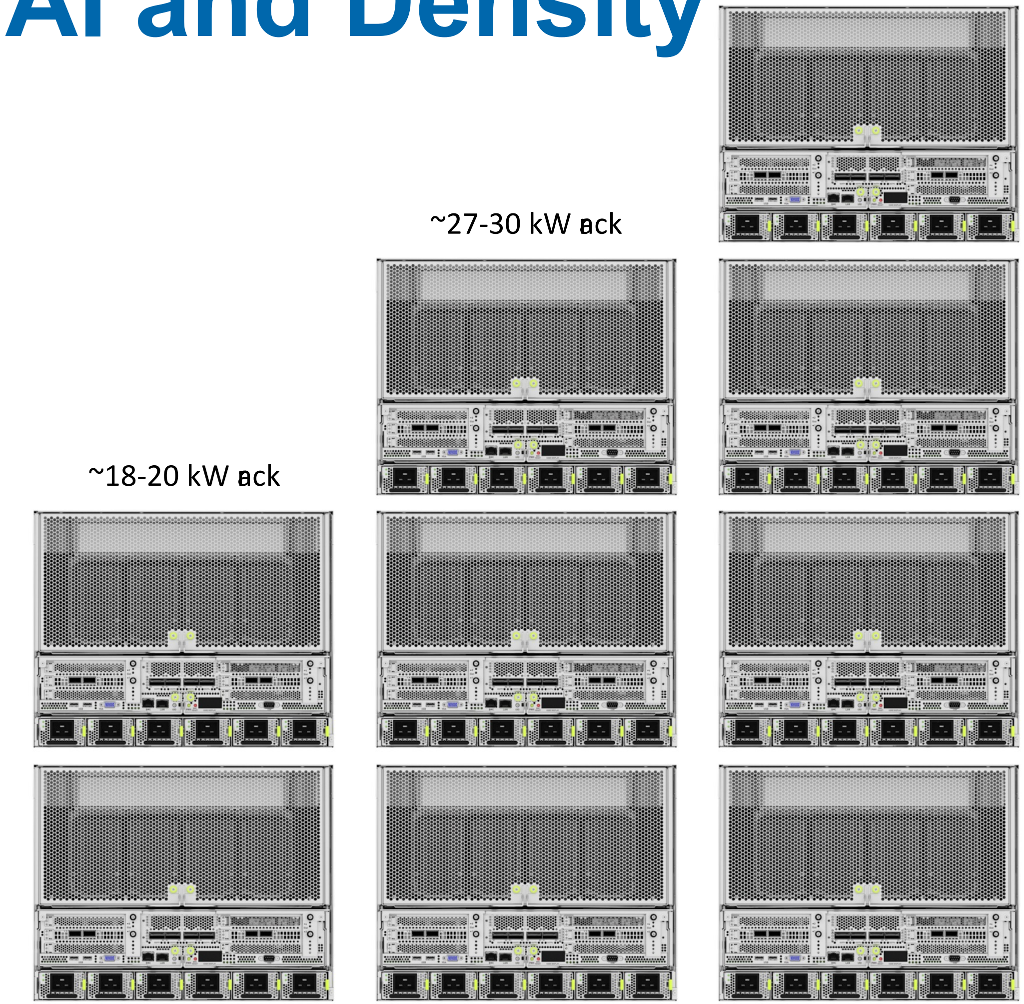
Examining Projections for AI and Density

Demand for AI-accelerators is outstripping supply. Installations are concentrated in large compute clusters in few facilities.

By the first quarter of 2025, Uptime Intelligence estimates global capacity demand of 2,000 to 2,500 MW due to the AI rush.
(Based on publicly disclosed information)

Density expectations are also overblown.

- Larger clusters tend toward high density for connectivity.
- Most facilities will spread out AI systems to manage power and cooling (DLC is optional).



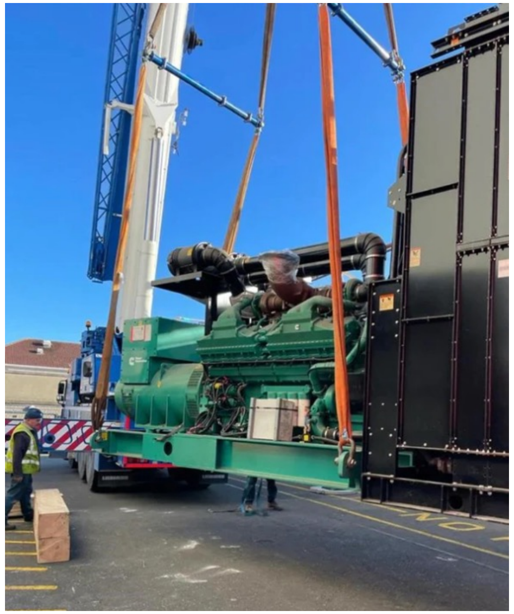
Not all Effects are Direct

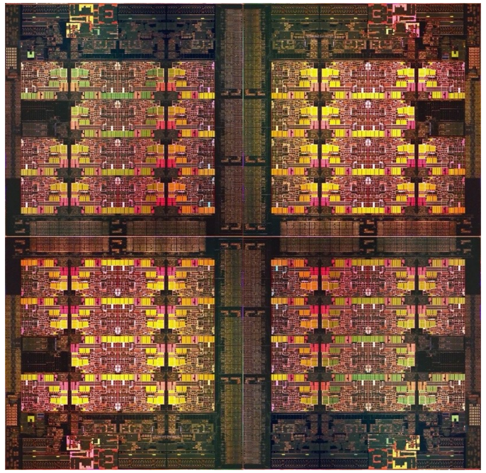
Rather than having to find the capacity and install large, extremely dense clusters of compute, for many operators, the impact on data centers will be more *indirect* through AI's effect on supply chains, chip designs and attitudes towards facility resiliency.

Prolonging demand-supply imbalance

AI arms race will push chip power envelopes further

Promoting mixed-tier facilities





uptime INSTITUTE

Prediction 3

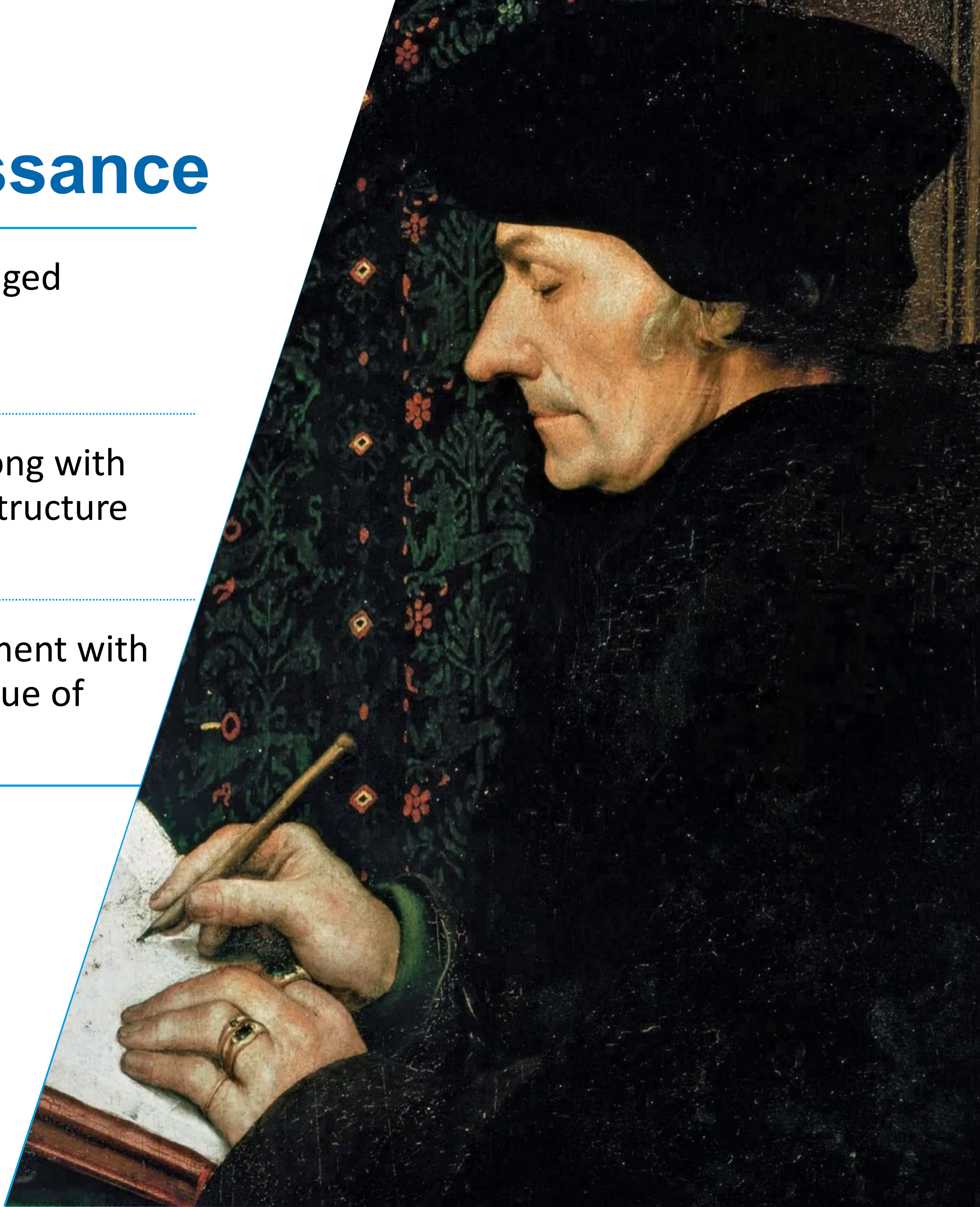
Data Center Software Gets Smarter, Leverages Data — at Last

A Data Center Software Renaissance

The design and capabilities of data center equipment have changed considerably over the past 10 years and traditional data center management tools have not kept up.

The increasing scale and complexity of modern data centers, along with industry-wide staffing shortages, create more demand for infrastructure automation.

Hype around generative AI is pushing more operators to experiment with machine learning and is changing their understanding of the value of facilities equipment data.



uptime INSTITUTE

What About BMS and DCIM?

Most data center operators use building management systems and/or data center infrastructure management software as their primary interfaces for facility operations.

These tools are important but have limited analytics and automation capabilities.

Operations teams need new software tools that can generate more value from the data produced by the facilities equipment.

Such tools are not intended to replace BMS or DCIM; they can be used to augment the functionality of existing data center management software.

New Tools Bring New Challenges

The adoption of data-centric tools for infrastructure management will require owners and operators to recognize the importance of **data quality**.

In some cases, data center operators will have to hire **analysts and data scientists** to work alongside the facilities and IT teams.

Collection of equipment data will require more networking inside the data center. This presents a potentially wider attack surface for cybercriminals.

Cybersecurity will be an important consideration for any operational AI deployment and a key risk that will need to be continuously managed.

Evolution is Inevitable

Data center management maturity model

Level	Description	Operating efficiency
Level 1	No integration of infrastructure data. Basic monitoring is provided by the equipment vendor software and the BMS.	Low
Level 2	Software installed to monitor environmental and equipment power use. Able to adjust basic controls (e.g., cooling) in line with demand.	Low
Level 3	Software can track physical data center equipment characteristics, location and operational status. Energy and environmental data are used to reduce risks and waste.	Medium
Level 4	Machine learning models are used for prediction, service management and multiple views, optimizing the data center in near real time. AI is applied to DCIM-based data lakes for advanced analytics.	Medium
Level 5	AI-driven integrated management software adjusts data center behavior and makes the best use of resources according to goals, rules and service requirements throughout its life cycle.	High

DCIM software alone will likely never evolve these capabilities. Instead, it needs to be combined with a new generation of data-centric tools.

Prediction 4

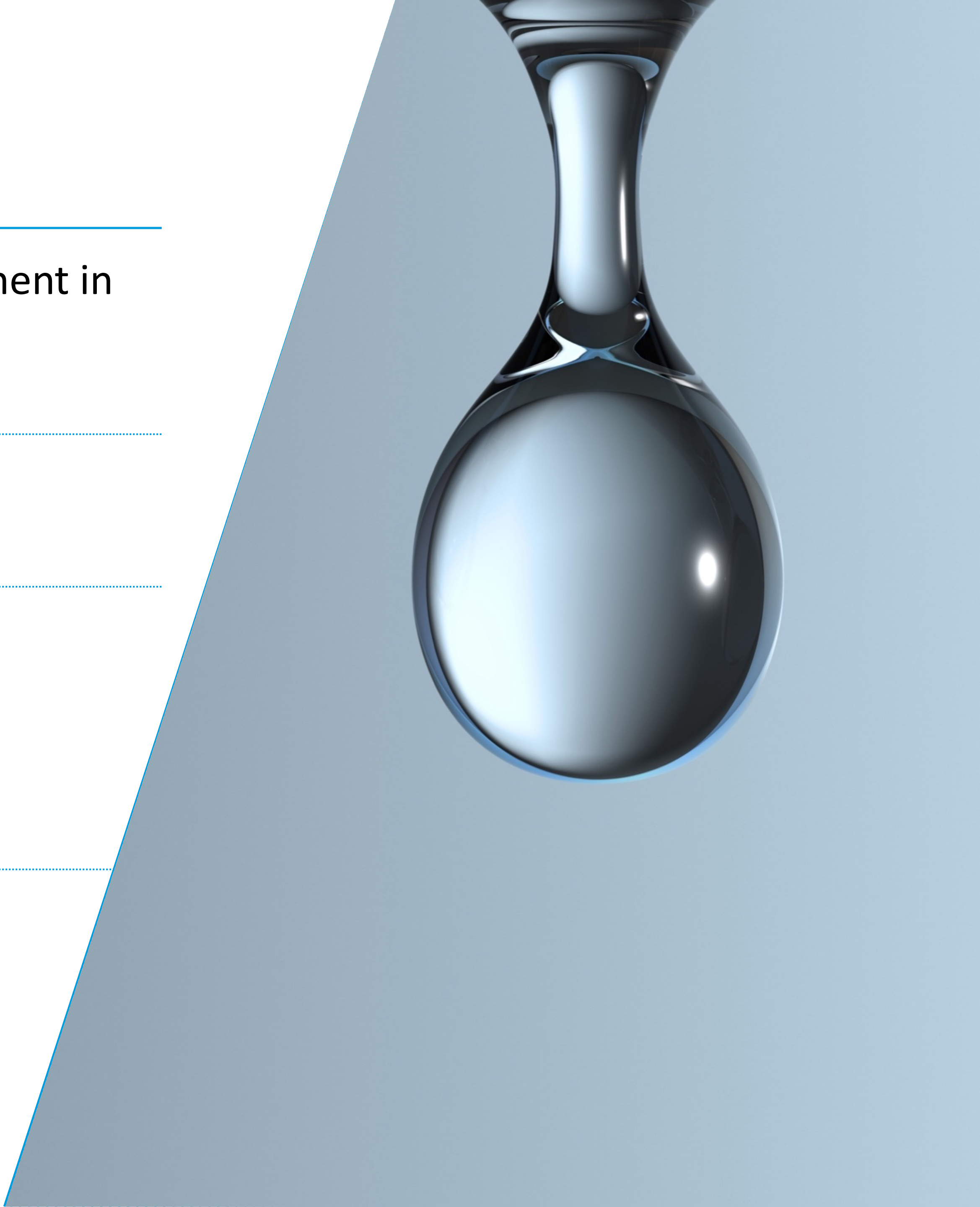
Direct Liquid Cooling will not Resolve Efficiency Challenges

Operators and vendors alike anticipate more DLC deployment in the next few years, driven by escalating cooling demands of IT.

DLC brings opportunities for energy efficiency—notably, leaner heat rejection infrastructure.

DLC real-world efficiency will be limited by:

- Gradual pace of adoption
- Hybrid cooling environments, shared infrastructure
- Trade-offs against cooling performance, capacity, interoperability



Cooling in Mixed Company

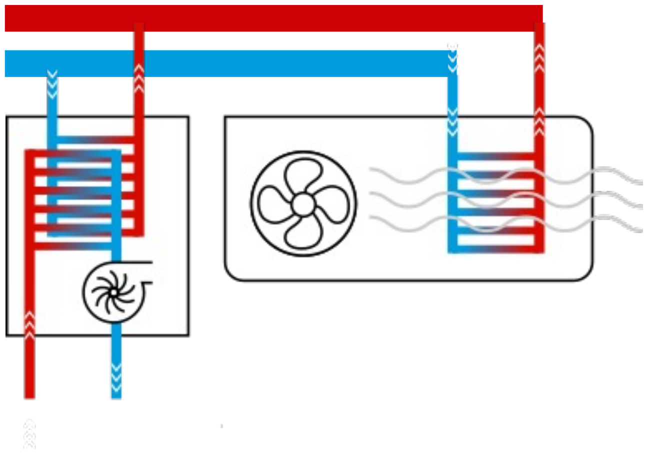
Liquid-cooled and air-cooled IT systems and infrastructure will coexist in data centers – likely for several years.

This limits opportunities for optimization and makes some major gains difficult to quantify – especially IT fan power.

It will take years for DLC installations to reach the scale where a dedicated cooling infrastructure can be justified as a standard approach in data centers.

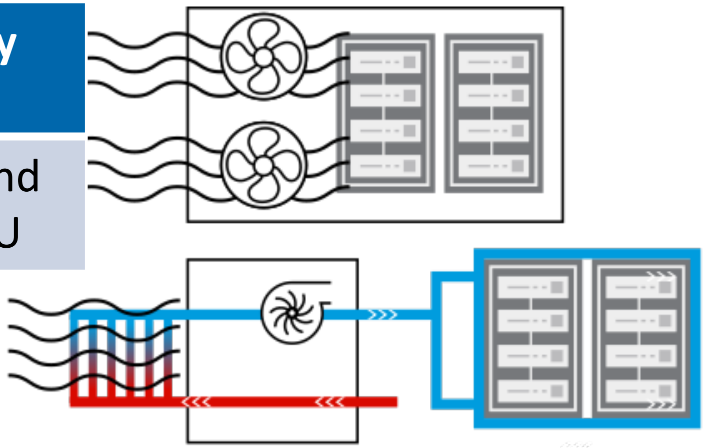
Shared facility water loop

CDU and CRAH



Shared facility cooling air

Air-cooled IT and air-cooled CDU

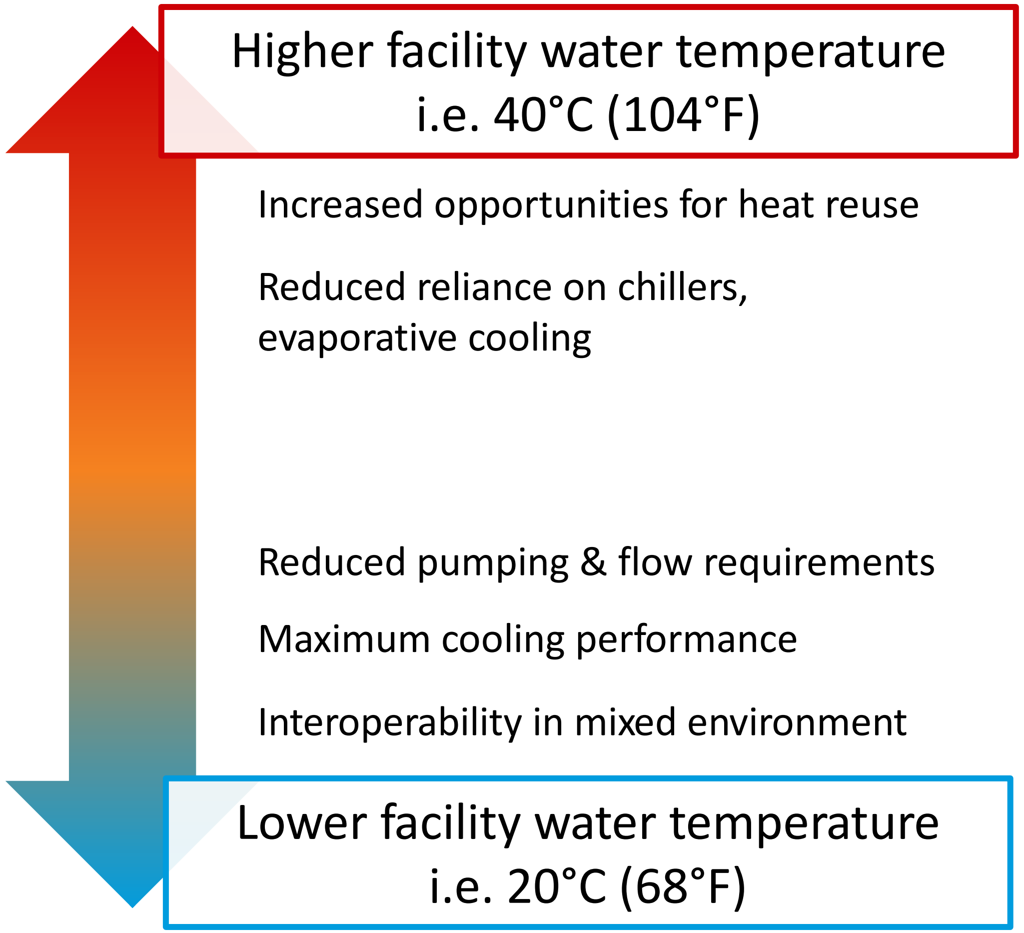


uptime INSTITUTE

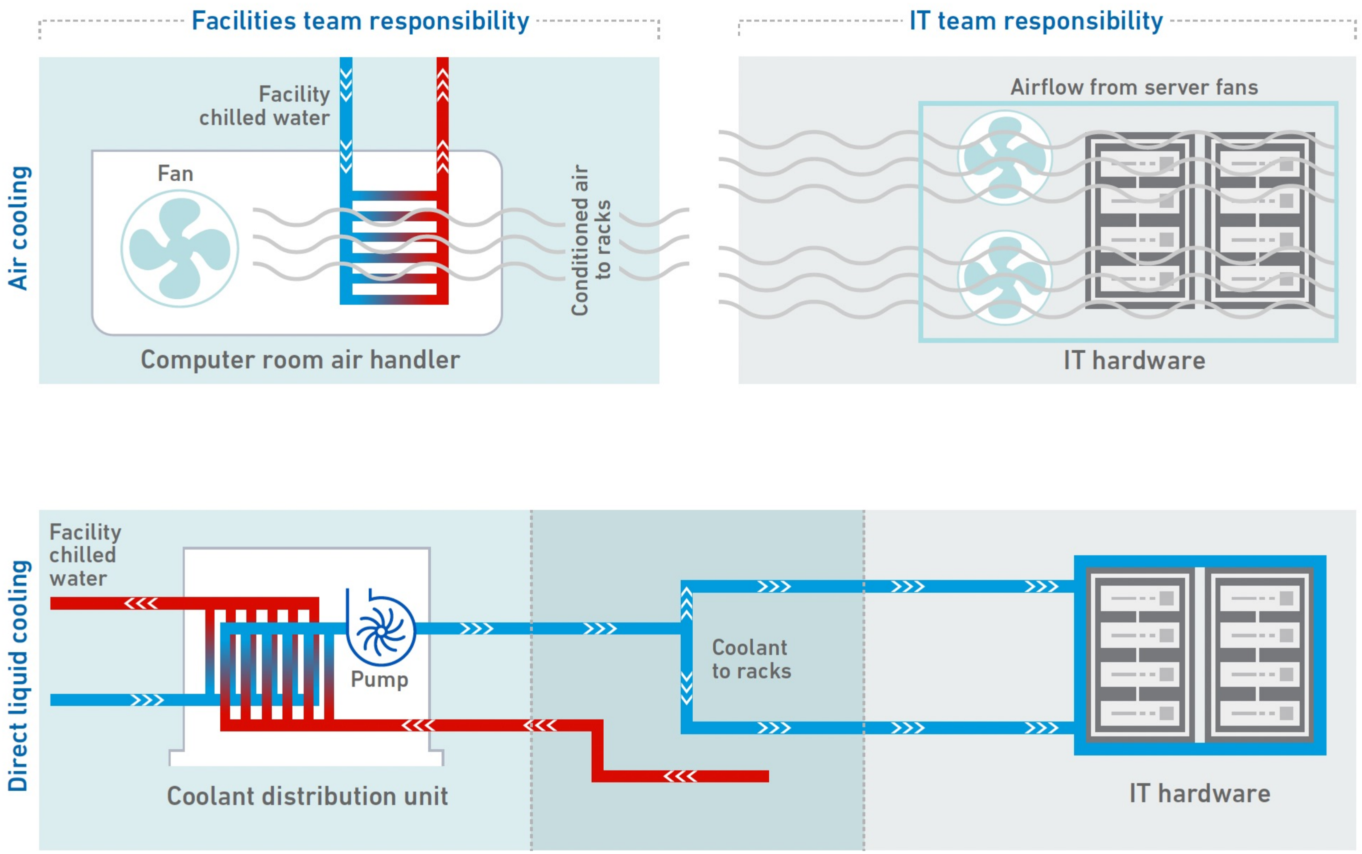
Hidden Trade-offs in Temperature

Benefits from low temperature may outweigh the attraction of a leaner heat rejection infrastructure.

A significant share of DLC adoption will likely represent an investment in cooling performance and IT capacity, rather than facility efficiency gains.



DLC Changes More Than the Coolant



DLC reshapes resiliency

- Concurrent maintainability, fault tolerance may not be practical.
- Ride-through time may be much shorter.

Some organizations will consider software resiliency strategies.

Organizational procedures throughout the data center lifecycle are altered, as DLC disrupts the division of facility and IT infrastructure functions.

uptime INSTITUTE

Prediction 5

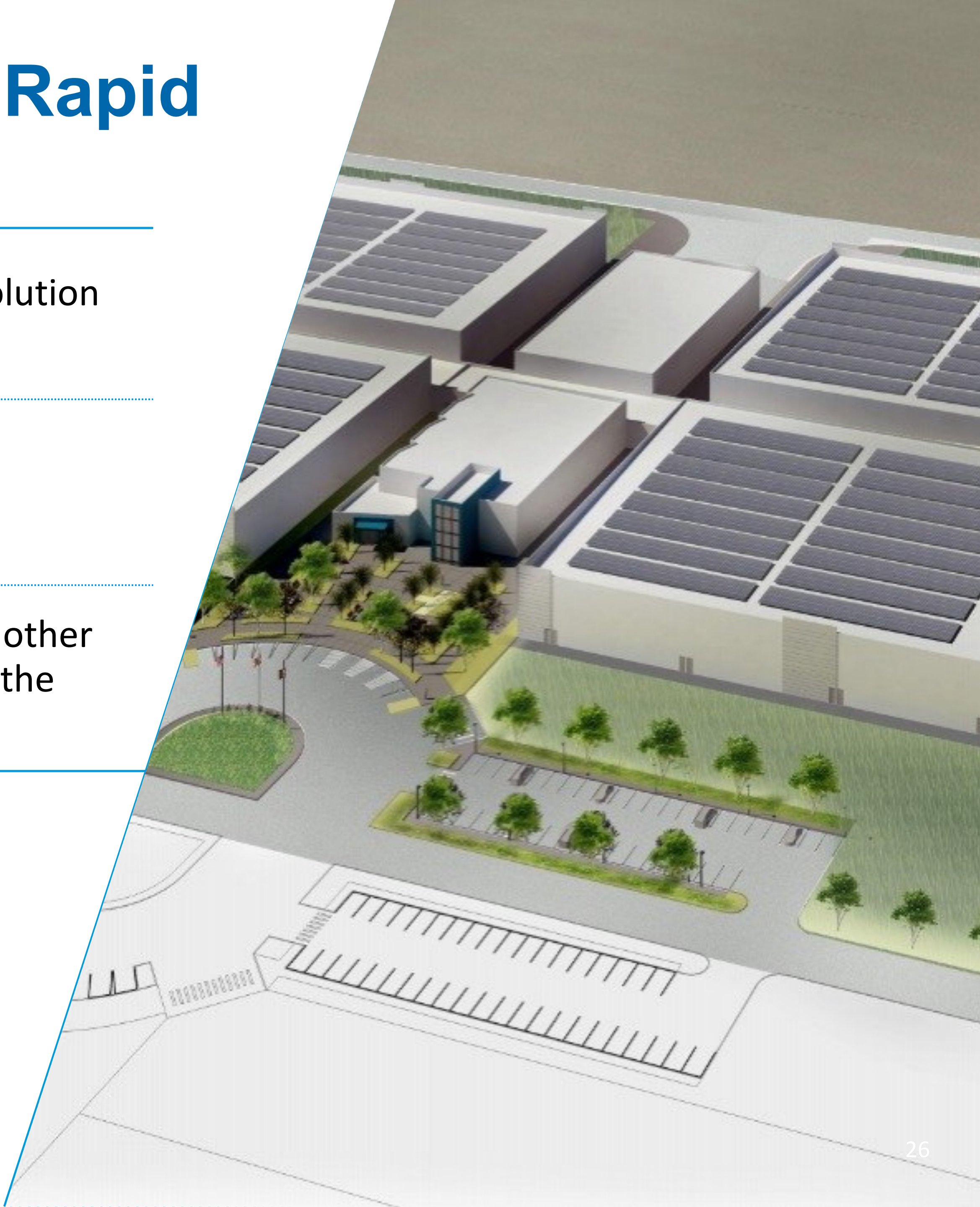
Hyperscale Campuses Begin to Redraw the Data Center Map

Hyperscale Colocation Shows Rapid Expansion

Hyperscale colocation campuses will continue to be seen as a solution to the rocketing demand for compute and storage.

These campuses will be built on huge areas of land, capable of supporting multiple tenants, including both cloud providers and enterprises expanding their digital footprint.

North America will see the largest of these campuses. However, other regions are also expanding rapidly, including Asia-Pacific, where the largest number of new developments are taking place.



uptime INSTITUTE

The Hyperscale Colocation Campus

Hyperscale colocation campus projects in progress

Campus location*	Provisioned power in megawatts (MW)	Number of projects	Average megawatts (MW)	Total spend ($ million)
North America	6,210 MW	10	621 MW	$45,000m
Asia-Pacific (excluding China)	3,832 MW	21	182 MW	$11,628m
Europe, the Middle East and Africa	750 MW	2	375 MW	$6,400m
China	500 MW	1	500 MW	$4,890m
Latin America	450 MW	1	450 MW	$400m
Total (worldwide)	**11,742 MW**	**35**	**426 MW**	**$68,318m**
Annual terawatt-hours (TWh) at 50% of provisioned power	51 TWh			

Provisioned power is likely capacity once everything in place; projects announced or identified since 2021; single location campuses (i.e., not multisite investments); public cloud vendor hyperscale projects excluded.
**Campuses 100 MW and above.*

Expensive infrastructure supported by redundant and new high-bandwidth fiber to major centers.

May involve power purchase agreements, colocation with power plants and use of on-site renewable energy sources.

uptime® INSTITUTE

The Largest Hyperscale Cloud and Colocation Campuses

Company / campus (owner)	Location	Estimated power in megawatts (projected spend if available)	Campus size in million square meters (million square feet)	Data center footprint in million square meters (million square feet)
Google	Horndal (Sweden)	1,400 MW	1.1 (11.7)	N/A
Prince William Digital Gateway (Compass Datacenters)	Prince William County (Virginia, US)	1,336 MW ($15 billion)	8.56 (92)	1.2 (12.9)
Digital Dullus (QTS Data Centers)	Loudon County (Arizona, US)	1,000 MW	1.7 (18.2)	0.69 (7.5)
Green Energy Partners	Surry Nuclear Power Plant (Virginia, US)	1,000 MW	N/A	30 data centers
Google	Skien (Norway)	850 MW	2 (21.5)	N/A
Quantum Loop Hole (Aligned Data Centers)	Frederick County (Maryland, US)	800 MW	8.4 (90.4)	1.2 (13)

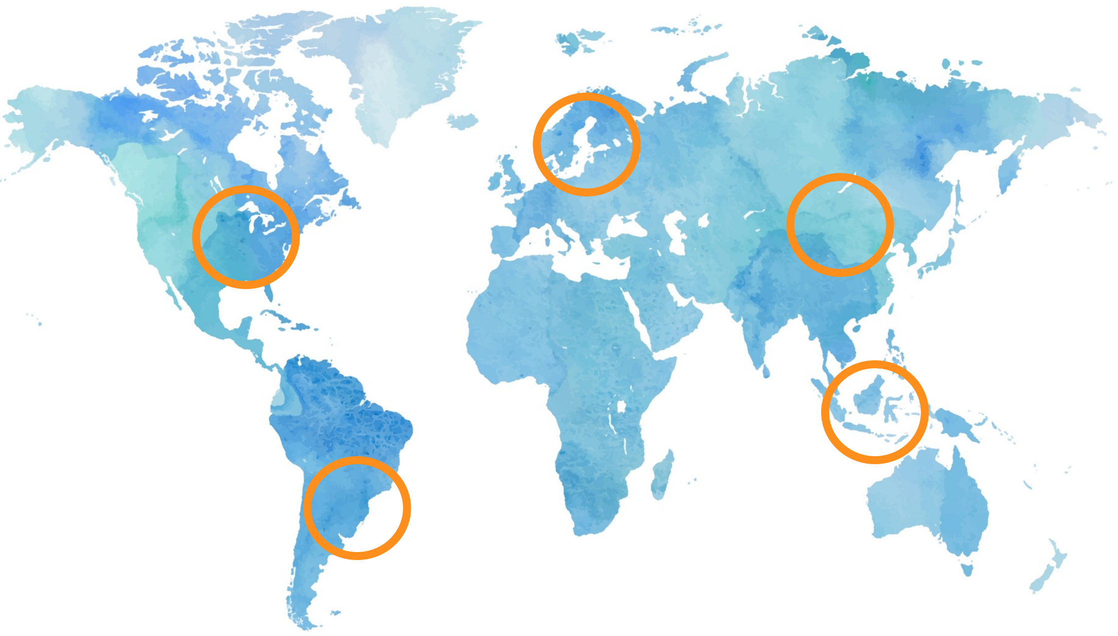
North America is home to the largest new **colocation campus** developments.

Northern Europe (Scandinavia) is witnessing the largest **cloud campus** investments.

uptime INSTITUTE

Global Map of Data Centers is Being Re-engineered

- Power availability, cost and fuel mix
- Carbon free solutions for cooling (cold air, water) and power (solar, nuclear, hydroelectric, geothermal)
- Connectivity – high speed fiber, network density
- Proximity to customers, to other data centers
- Regulatory and tax regimes
- Labor pool



uptime INSTITUTE