

The background is a photograph of a modern building facade. On the left, there is a large, colorful geometric graphic composed of several overlapping squares in shades of orange, blue, green, and yellow. In the center, the word "Microsoft" is visible in a large, white, sans-serif font, partially obscured by the other elements. The building's facade consists of vertical wooden slats. The overall scene is dimly lit, suggesting an evening or indoor setting with artificial lighting.

# Rethinking AI Infrastructure for Sustainable Scale

A **phased** cold-environment compute strategy for Microsoft Azure

---

Microsoft x TKS

# Executive Summary

## The Problem

AI data centres waste massive amounts of energy and water on cooling, limiting Microsoft's ability to scale AI sustainably.

→ Cooling consumes **30-40% of data centre electricity** or ~260 to 350 GWh for a 100MW data centres, adding tens of millions of dollars annually to operating costs.

→ As AI workloads scale, data centres already consuming ~1% of global electricity face a hard scalability limit, where expanding AI capacity strains power grids, cooling systems.

→ Traditional cooling systems rely heavily on freshwater and grid electricity, creating environmental strain and conflicting with Microsoft's goal of being carbon negative by 2030.

## The Solution

Build more training data centres into ultra-cold regions to leverage natural air cooling.

→ Cold climates enable **free-air cooling**, reducing cooling energy needs by ~30-40% and cutting water usage by up to 90%.

→ These locations are optimized for large-scale AI training and batch workloads, allowing Microsoft to scale compute capacity without being constrained by urban power grids, water availability, or cooling limits.

→ This approach builds on Microsoft's existing cold-region infrastructure while pushing specialization further for **large-scale AI compute**

## The Impact

Cold-climate AI data centres significantly reduce cost, emissions, and resource usage while enabling long-term scaling.

→ Building in cooler climates saves an estimated **170 million dollars USD per 100 MW** data center, which could save Microsoft **51 billion** in the US alone.

→ Eliminating most cooling-related water use saves over 100 million litres of freshwater per data centre per year.

→ Lower energy demand directly reduces carbon emissions while removing power constraints, enabling Microsoft to scale AI capacity faster in line with its carbon-negative 2030 commitment.

# Microsoft's AI Growth Is Driving Rapidly Rising Infrastructure Costs

	Per Data Center	Globally (Azure)
Average Power Load	~50–100 MW	~15–30 GW
Energy Used for Cooling	~40% of total energy	~6–12 GW
Annual Energy Use (Cooling)	~175,000–350,000 MWh	~60–120 TWh
Estimated Cost per kWh	~\$0.06–\$0.10	—
Annual Cooling Cost	~\$10–30M	\$6–12B+ annually

## Bottom Line

Incremental efficiency improvements are no longer sufficient. Microsoft's AI growth is now constrained by physical, economic, and environmental limits of Earth-based infrastructure.

## Why This Is a Problem

- AI model training and inference require high-density GPU clusters operating continuously
- As rack densities exceed 100+ kW per rack, cooling and energy costs scale non-linearly
- Cooling alone can consume up to ~40% of total data center energy
- Grid congestion, land acquisition, water usage, and permitting increasingly slow deployment

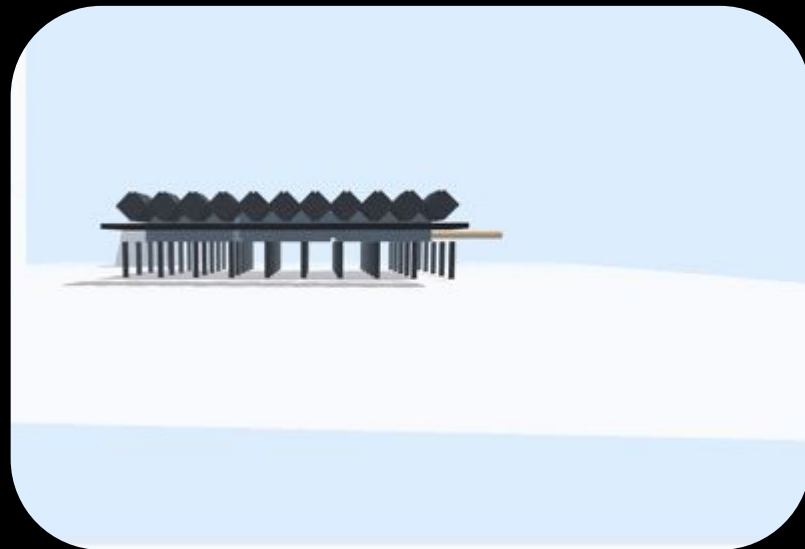
# An expansion of AI infrastructure into cold earth environments

## Design Data Centers for Extreme Cold

- Standard centers fail in extreme cold
- Stilts and tanks protect hardware from melting surrounding area

## Expand to Extreme-Earth (Arctic)

- AI training + climate sims
- Natural cooling, low water, renewables
- Lowest-risk, proven tech
- Reduces energy + cuts water usage



# Case Study

# Validation & Proof

Real World Example	Cooling Strategy & Benefits	Measured Environmental Impact	Climate-Based Amplification
<ul style="list-style-type: none"><li>• META (Facebook) Data centre, located in Luleå (North Sweden)</li><li>• The area is known for its cold subArctic climate, providing low temperatures at an average of 3.1°C (varies based on seasons)</li></ul>	<ul style="list-style-type: none"><li>• The cold air from outside will serve as the cooling mechanism</li><li>• The cold climate reduces the amount of electricity required to cool the servers</li><li>• Reduction in:<ul style="list-style-type: none"><li>- ↓ Electricity consumption</li><li>- ↓ Carbon emissions</li></ul></li></ul>	<p>Air cooling can outperform other forms of cooling</p> <p>With the implementation of this cooling mechanism:</p> <ul style="list-style-type: none"><li>• ↓ <b>Power Usage Effectiveness (PUE):</b> ~1.07</li><li>• Industry Average: ~1.5-1.6</li></ul>	<p>Cold climates enable free cooling with minimal energy input.</p> <p>Lower costs, lower emissions, long-term sustainability</p> <p>As the global temperature continues to rise, cool regions will become more valuable for infrastructures</p>

# Implementation Timeline

## Phase 1

### Cold-Climate Data Center Development

May 2026 – September 2027

- Begin development of AI data centers in **cold-climate regions** (e.g., Iceland, Northern Europe)
- Average ambient temperatures of **10–13°C** during spring and fall enable **natural cooling** with minimal mechanical systems
- **Reduced dependence** on electricity-intensive cooling infrastructure
- Stable climate avoids extreme winter conditions, ensuring **operational reliability**

## Phase 2

### Performance Validation & Optimization

October 2027 – April 2028

- Measure energy efficiency using **Power Usage Effectiveness (PUE)** as the primary metric
- Analyze performance across **fall-to-spring operating conditions**
- Benchmark electricity consumption against traditional Azure data centers
- Quantify operational **cost savings** and cooling efficiency improvements

# Thank You!



**Tawfic Shnoudeh**

tawficshnoudeh@gmail.com.com



**Krish Sharma**

krishkapi100@gmail.com



**Seth John**

sethjohn058@gmail.com



**Aakarsh Jain**

aakarsh052@gmail.com



**Vihaan Bhatia**

vb.vihaan@gmail.com

We appreciate the opportunity to present our approach to using remote and extreme-environment compute as a complement to Earth-based AI infrastructure. This challenge pushed us to think carefully about long-term cost, reliability, and deployment trade-offs at scale.

We appreciated the chance to explore these ideas and learn from the real constraints Microsoft navigates when building global AI systems.

If you have any questions about our proposal, please feel free to reach out.

Thank you,

Tawfic Shnoudeh, Krish Sharma, Seth John, Aakarsh Jain, Vihaan Bhatia