



From the Trenches: The Physics, Politics, and Practice of Data Centre Optimisation

An Insider's Perspective on Balancing SLAs, Unlocking Efficiency, and the Real Science of Energy Savings



Peter Simon
Head of Technology APAC and ANZ, EkkoSense





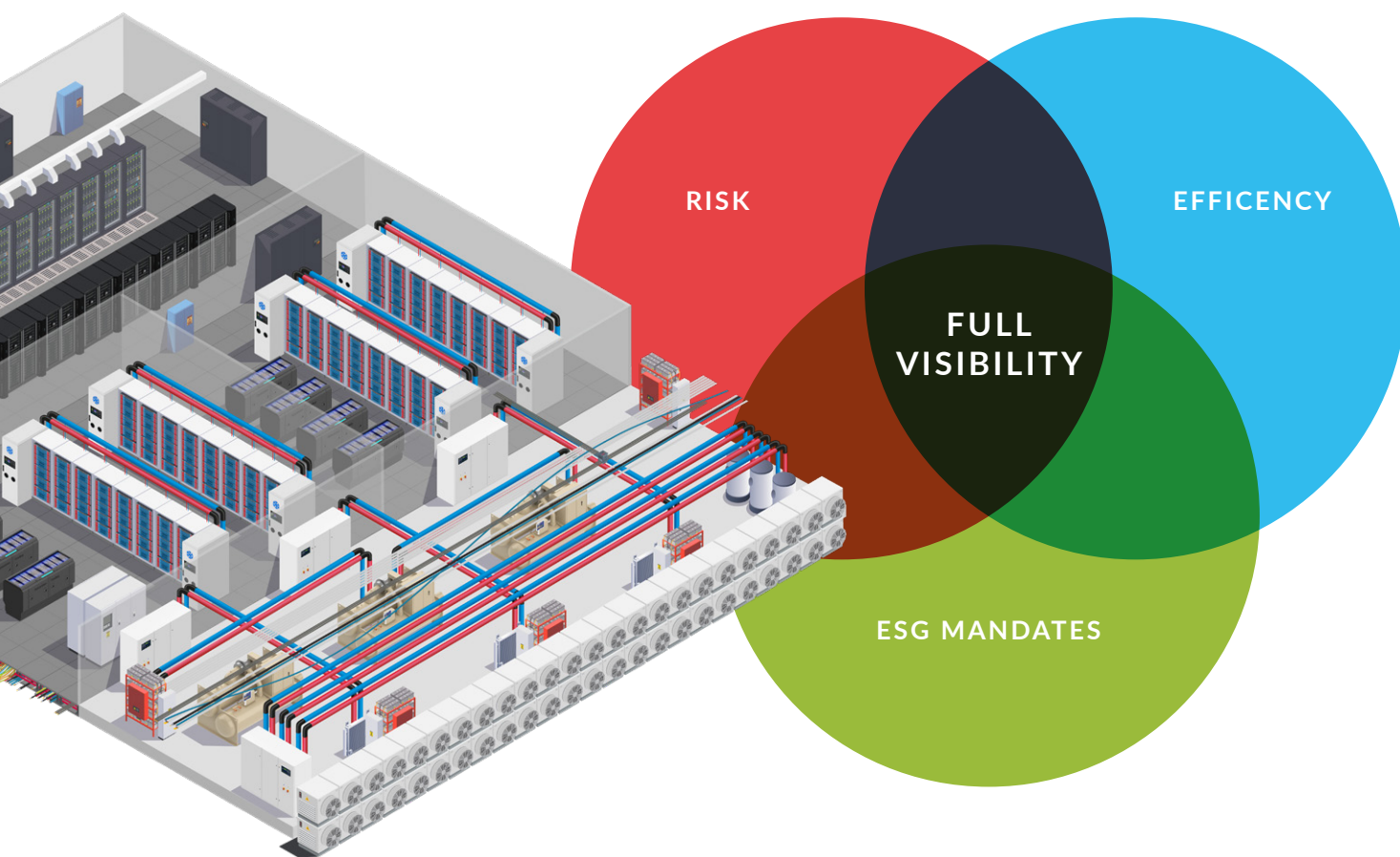
Introduction

Over the last 25 years, I have lived and breathed data centres. I've watched the industry evolve from the days of over-provisioned "meat lockers" to the complex high-density, AI-driven architectures of today. But despite all the technological leaps, the fundamental tension on the data centre floor remains unchanged: ***the battle between risk and efficiency.***

Data centre engineers are tasked with an incredibly difficult balancing act. We must maintain 100% SLA achievement, manage escalating power use, maximise IT equipment load in the data halls, and still somehow find time to optimise efficiency. And this all while navigating an industry-wide skills shortage.

When a platform like EkkoSense enters the conversation, executives often ask a very practical question: *"How does this actually generate energy savings? I understand all the benefits of moving floor tiles and adjusting set points, but what is the actual science behind the financial return?"*

This white paper answers that question. Not from the perspective of a software inventor, but from the viewpoint of a veteran engineer. It explores the operational realities we face every day, the physics of data centre cooling, and how giving engineers the right visibility allows us to turn thermodynamic theory into verifiable, ongoing energy savings, risk mitigation and ESG mandates. What was once a trade-off becomes a disciplined control loop with measurable business value.





Part 1: The Engineer's Dilemma

Before we can discuss *how* to save energy, we have to understand *why* it is so difficult to do so in practice. Data centre engineers operate typically in an environment where the rules of the game often work against optimisation.

1. The "Optimisation vs. Growth" Paradox

In the data centre industry, it's growth that's grabbing the headlines – with little bandwidth left over for smart optimisation. Executives and investors love cutting ribbons on shiny new 200MW+ facilities. Building capacity drives headlines. Conversely, spending months tweaking an existing legacy facility to squeeze out a 15% efficiency gain is viewed as necessary, but unglamorous, "plumbing." Because optimisation isn't top of the agenda, it rarely gets the budget or the boardroom attention it deserves. That's until the facility and the power grid maxes out.

2. The Power Pass-Through Problem

In many colocation models, there is a fundamental disconnect regarding energy costs. If a facility operates with a "power pass-through" model, the end-user (the tenant) pays the utility bill. Consequently, the facility operator lacks a compelling financial incentive to invest heavily in optimisation. The cost of inefficiency is simply pushed down the line. It becomes everyone's problem, and therefore, nobody's priority. This is short-sighted at best.

3. The Skills Shortage and its impact

There is a critical shortage of experienced data centre engineers. The veterans who truly understand the complex fluid dynamics of a white space are few and far between. We cannot afford to have our best minds walking the floor with a handheld thermometer trying to guess where the airflow is going. We need to deploy these highly skilled individuals where they achieve the biggest "bang for buck" - making strategic decisions, not manually chasing hot spots and capacity threats.

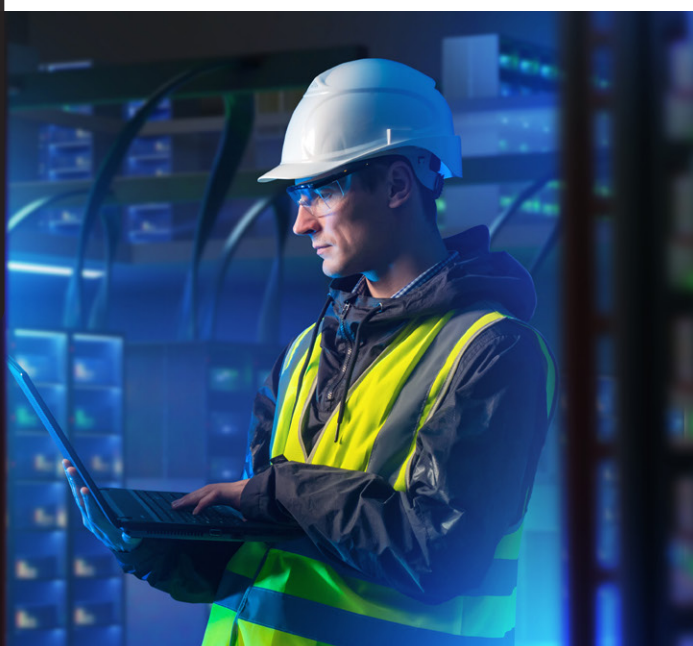
Here's where applying AI to a vastly increased telemetry density can deliver both the new insights and risk management optimisation that's needed to unlock significant efficiency improvements.

4. SLAs are non-negotiable

Every data centre operator knows the golden rule: **Never break the SLA**. If a server overheats, jobs are lost and heads roll. Because we traditionally lack real-time, granular visibility into the thermal health of every single rack, engineers do the only logical thing they can to protect the SLA: **we overcool**.

We flood the room with cold air and drop the set points to 18°C or 19°C. It is a massive waste of energy, but it creates a thermal "safety buffer." To an engineer flying blind, overcooling is the best insurance policy against an outage.

There is a critical shortage of experienced data centre engineers. The veterans who truly understand the complex fluid dynamics of a white space are few and far between.





Part 2: The Science of Savings

So, how do we remove that safety buffer and ensure that SLAs and risk can improve at the same time? When we propose using EkkoSense AI and Machine Learning to increase Air Handling Unit (AHU) set points and balance airflow, it is not just about "tweaking the AC." It is about going back to engineering first principles and leveraging the laws of thermodynamics.

When you ask a veteran engineer how raising a set point saves money, we look past the software interface and directly at the mechanical plant. Here is the practical science behind the savings:

1. Less Temperature Difference = Less Compressor Work

Whether you are using Direct Expansion (DX) CRAC units or a Chilled Water (CW) system, cooling relies on a refrigeration cycle. The core rule here is "Lift."

- **The Physics:** The compressor/chiller's job is to move heat from a cooler area to a hotter area. The lower your room temperature set point (e.g., 18°C instead of 24°C), the larger the temperature gap between the air you are trying to cool and the refrigerant loop.
- **The Energy Impact:** A larger temperature difference forces the compressor/chiller to work much harder to achieve the necessary "lift." By increasing the set point to 24°C, we narrow this gap. The compressor doesn't have to work as hard, it cycles less frequently, and it consumes significantly less electrical power.
- **Chilled Water (CW) Impact:** In a chilled water environment, raising the room set point allows us to raise the chilled water temperatures and optimise system flow rates. This drastically improves the efficiency of the chillers (better heat exchange) and maximises the hours we can rely on "free cooling" (economisation) before the mechanical chillers even need to turn on.

2. The Secret Weapon: Fan Affinity Laws

Modern CRAC and CRAH units utilise variable-speed fans. When we balance the floor tiles and raise set points, we are ultimately reducing the volume of air required to cool the room.

- **The Physics:** Fan power consumption is governed by the Affinity Laws (specifically the *cube law*). Power consumption is proportional to the cube of the fan speed.
- **The Energy Impact:** This means the savings are not linear; they are exponential. If you reduce a fan's speed by just 20%, you don't save 20% of the energy - **you reduce the fan's power consumption by nearly 50%**. By stopping cold air bypass and raising set points, fans can spool down from 100% to 70% or 60%, resulting in massive, immediate kilowatt savings.

3. Backed by ASHRAE Standards

The numbers are real and based on fundamental physics. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) TC9.9 guidelines explicitly state that modern IT equipment is incredibly resilient.

- Most modern servers operate perfectly fine with allowable supply temperatures up to 27°C (80.6°F).
- Moving from a "traditional" legacy set point of 20°C up to a modern 25°C can save roughly 4% to 8% of cooling energy for every single degree raised.
- The average mechanical power reduction over the last 50 optimisation projects I have done is around 25%.

If you reduce a fan's speed by just 20%, you don't save 20% of the energy - **you reduce the fan's power consumption by nearly 50%.**





Part 3: Bridging the Gap with Visibility

If the science is so clear, and ASHRAE tells us it's safe, why haven't engineers been doing this all along?

Because without quantifiable, granular visibility, applying the science is a career risk for engineers.

I cannot raise the CRAC set points to 24°C+ if I don't know whether that one high-density cluster in the back corner of the data hall is suddenly going to spike to 32°C.

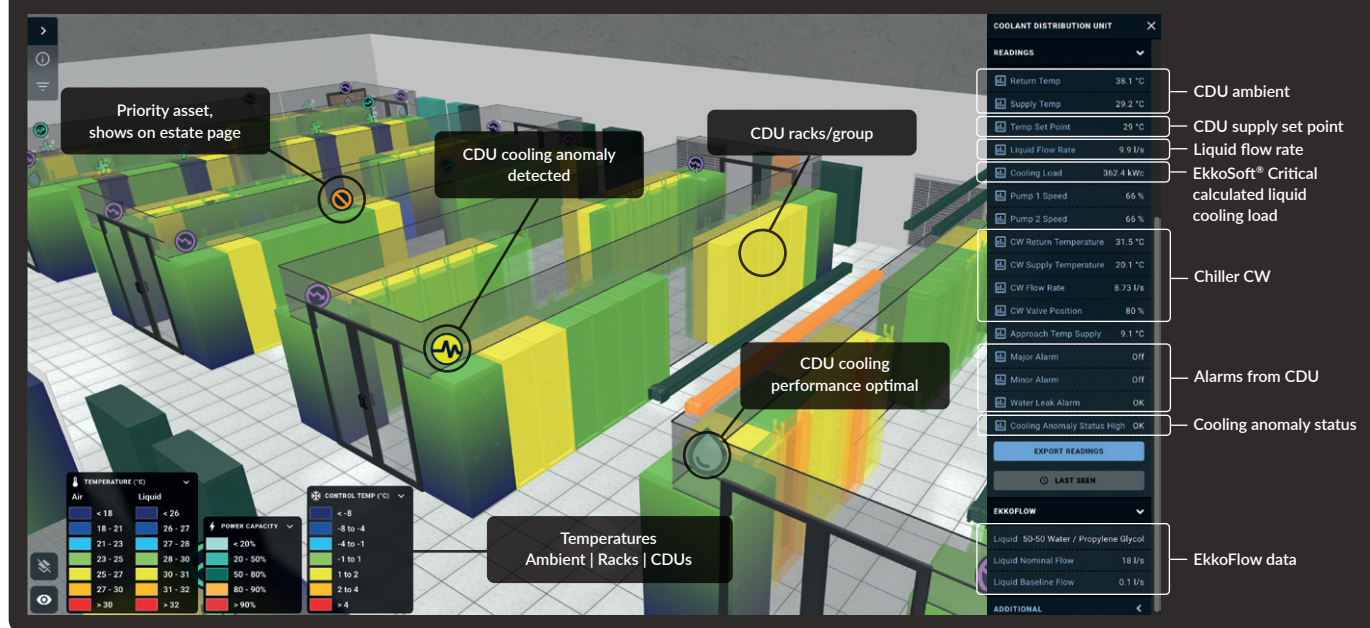
This is where EkkoSense supports engineers and operations teams by effectively making the transition from just another DCIM software tool to operating as a DC operations lifeline. EkkoSoft Cooling Advisor, for example, provides continuous, granular, real-time mapping of the exact thermal conditions of the white space.

Another EkkoSoft capability – Cooling Anomalies Detection - helps DC operators and engineers to identify cooling performance anomalies ahead of any potential equipment failure.

This level of insight enables operations teams to make fact-based incremental changes that collectively deliver against the multiple and conflicting goals of efficiency and service delivery. This works to maximise impact on facility asset value, profitability and performance.

It does not replace the engineer; it acts as a co-pilot. By providing absolute certainty about what is happening at the rack level, it allows the engineer to strip away the costly "overcooling safety buffer" without ever risking the SLA. It tells us exactly which tiles to move to stop bypass air, allowing the fans to slow down (triggering the Affinity Law savings), and tells us exactly how high we can push the set points (reducing the compressor lift).

This level of insight enables operations teams to make fact-based incremental changes that collectively deliver against the multiple and conflicting goals of efficiency and service delivery. This works to maximise impact on facility asset value, profitability and performance.





Conclusion: Delivering on the Promise of Data Center Optimization

For over 25 years, we have treated data centre cooling as a blunt instrument. We threw power at the problem because power was cheaper than downtime. Today, with rising grid constraints, sustainability mandates, and the sheer cost of electricity, that approach is no longer viable.

True optimisation might not merit a ribbon-cutting ceremony, but it extends the life of legacy facilities, reclaims stranded power capacity, and drops millions of dollars and higher operational profitability directly to the bottom line. By combining the undeniable physics of mechanical cooling with the safety of real-time thermal visibility, we can finally empower data centre engineers to achieve the ultimate goal: maximum efficiency, zero SLA risk.

This is all possible because we're living in a new paradigm where detailed telemetry can be continuously analysed by machine learning and AI through accessible Digital Twin visualisations.

This means we effectively have a 24x7 mechanical consultant that's available to us recalculating and delivering parallel insights about thermal performance, risk and efficiency in near real time. Unlike the past, where it took days to evaluate and make small risk-averse changes, we now have a PhD-level copilot that's doing a day's mechanical analysis every few minutes!

This can be unsettling at first, particularly if you're still stuck in a 20th century data centre mindset. However, it's hard to argue with the benefits.

Here are my Top Five Takeaways from recent EkkoSense optimisation projects:

- 1.** Data centres that declared they were at 100% capacity actually had another **20-25% of capacity** post optimisation with EkkoSoft Critical.
- 2.** EkkoSense optimisations delivered a compelling business case, with **PUE reducing from 1.6+ to 1.25-1.45**, concurrently enabling a much higher SLA compliance.
- 3.** Effective optimisation using AI and machine learning software **drives a data hall room design load from 100% to 115%**.
- 4.** AI-driven insights correlate cooling performance with IT workload to **reduce cooling energy usage by 10-20%**.
- 5.** And, conservatively, this AI-driven approach can consistently deliver at least 10% mechanical power savings, representing a **potential saving of 3.7 to 6.5 GW** of cooling energy reduction globally which is equivalent to 4 to 7 coal-fired grid-sized power stations.



About the Author

Peter Simon is Technical Lead for EkkoSense's APAC region operations. In this role he supports the deployment of the company's data centre optimisation software that helps unlock the real-time operational insights needed to handle growing AI workloads, increased power demands, and complex liquid cooling requirements. Peter brings some 25 years' in-depth data centre engineering experience to his role, having previously worked with organisations such as Equinix and Vertiv.