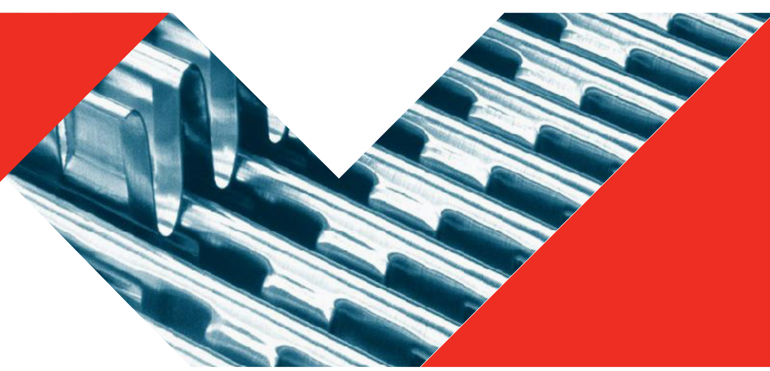
The Hidden Cost Driver in Data Center Cooling

And How to Eliminate It



The Rising Cost of Keeping Data Centers Cool

The data center industry is in the middle of a historical transformation. With the increasing use of AI platforms, the data centers behind some of the world's most popular technologies are consuming substantial amounts of electricity. According to the International Energy Agency, global data centers consumed about 415 TWh of electricity in 2024, with this number expected to more than double by 2030. In the United States alone, data centers consumed an estimated 176 TWh in 2023, equal to 4.4% of the total national electricity use.

Cooling drives nearly 40% of data center power consumption, second only to the power used during computing. By some estimates, data center energy demands could consume as much as 9% of the U.S.'s annual electricity generation by 2030. **The result is that cooling loads are rising faster than most facility designs anticipated**, putting pressure on every system that moves heat from the data center to the outdoor environment. For the engineers designing, building, or upgrading these facilities, cooling infrastructure is becoming the primary determinant of operating cost.

Every watt of heat generated must ultimately be directed to the outdoor environment, and the heat exchanger is the heart of this system, serving as the critical link between the data center's indoor cooling loop and the outdoor cooling loop. Their performance determines how much of that cooling comes cheap and how much comes at full chiller cost.

Why Data Center Cooling Costs So Much

Free Cooling Versus Mechanical Cooling

Data centers that use water-cooled systems—the standard approach at hyperscale and colocation scale—face a continuous operational challenge: rejecting heat generated by computing equipment to the outdoor environment, 24 hours a day, seven days a week, 365 days a year. There are two ways to accomplish this.

Free Cooling

Also called economizer or water-side economizer operation, free cooling uses a cooling tower or dry cooler to transfer heat directly from the facility water loop to the ambient air. The only energy inputs are pumps and fans. When the outdoor conditions allow, free cooling is much less expensive to operate than the alternative because you are only paying for moving air or water.

Mechanical Cooling

Mechanical cooling uses a refrigeration compressor, also called a chiller, to force heat transfer when ambient conditions cannot do the work passively. Chillers are energy-intensive and expensive to operate. One study found that a plant relying primarily on mechanical cooling consumed approximately 4.7 million kWh annually for a 1,500 kW steady load, nearly three times the 1.6 million kWh consumed by an equivalent plant operating predominantly in free cooling mode.

In modern data centers, hybrid systems that combine free and mechanical cooling have become the new normal. Because of the cost differential between the two systems, the goal is to maximize the number of hours per year in free cooling mode, minimizing chiller runtime. **Depending on the size, load profile, and local conditions, free cooling can achieve energy savings of up to 90% compared to mechanical cooling.**

The Temperature Approach:

Why a Single Degree Can Change Your Operating Economics

Wet-bulb temperature—the lowest temperature ambient air can reach through evaporative cooling at current humidity conditions—sets a hard limit on how cold a cooling tower can make its water. When the heat exchanger requires a large temperature difference between the two loops to operate, the system hits that limit sooner, forcing the facility onto the chiller earlier and more often.

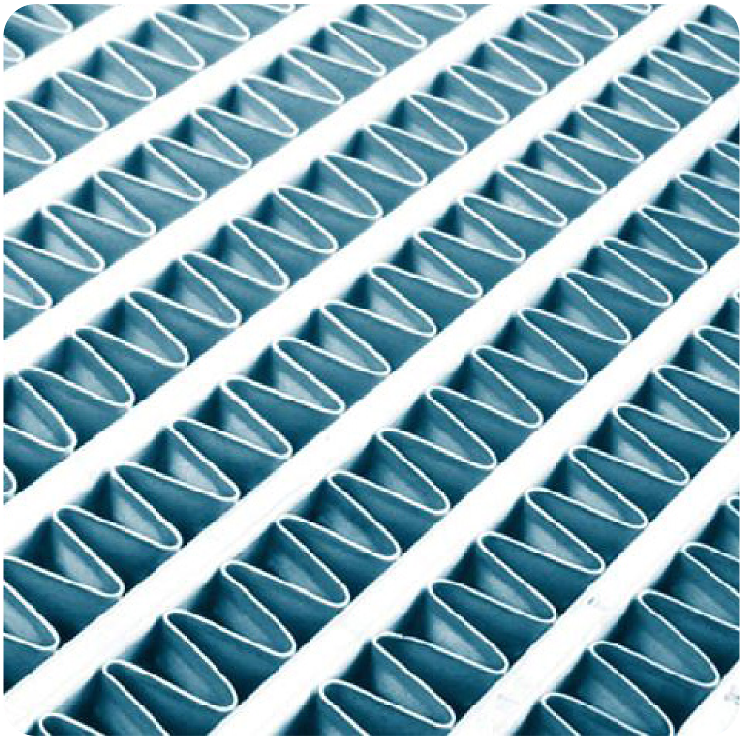
For example, a standard heat exchanger typically operates with a 2 °C temperature approach, the gap between the two fluid streams at the heat exchanger outlet. The total temperature difference between ambient air and the cooled circuit reflects both the cooling tower's wet-bulb margin and the heat exchanger's temperature approach combined. So, at an ambient temperature of 30 °C (86 °F), a cooling tower can typically cool the circuit down to 25 °C (77 °F). If the ambient temperature rises above this threshold, free cooling is no longer sufficient and the chiller must run. A heat exchanger with a tighter temperature approach of 1 °C extends this range because now you can have a 33 °C (91.4 °F) ambient temperature and still cool the circuit down to 26 °C (78.8 °F), expanding the range of free cooling days.

This is not a marginal efficiency consideration; it is the primary variable governing cooling plant operating cost, and it is determined almost entirely by the engineering of the heat exchanger.

The SIGMA M186 from API Heat Transfer achieves a validated temperature approach of 0.5 °C, well below the 2 °C industry standard and the 1 °C best-practice design target. The engineering behind that performance, and what it means for data center operating economics at any scale, is examined in detail below.

What to Evaluate When Specifying a Data Center Heat Exchanger

Rated capacity is a useful starting point, but the variables that actually drive operating cost—temperature approach, pressure drop, fluid compatibility, installation complexity, and lifecycle durability—vary significantly across manufacturers and products. The following criteria matter most.



The Three Variables That Determine Real-World Performance

As discussed above, temperature approach is the primary performance metric. In plate heat exchangers, the temperature approach is largely dictated by the plate corrugation geometry: the corrugation pattern, depth, angle, and spacing control how turbulent the flow becomes in the channels. Increased turbulence directly enhances the heat transfer coefficient and enables a tighter temperature approach.

Pressure drop is the unavoidable companion to heat transfer. Greater turbulence transfers more heat but also increases resistance to flow, requiring more pump energy to circulate fluid through the exchanger. The engineering objective, minimizing pressure drop while maximizing heat transfer, requires precise optimization of the corrugation geometry and cannot be achieved by adjusting a single variable in isolation.

Pressure resistance adds a third constraint. Plates must maintain structural integrity at operating pressures, commonly up to 16 bar in data center applications, while being stamped from thin sheet metal. Thinner plates improve thermal conductivity but reduce structural margin. Achieving the correct balance requires both rigorous engineering analysis and proven manufacturing process control, and it is one of the primary reasons that performance varies substantially across heat exchanger products with similar specifications on paper.

Flow Distribution: The Determinant of Realized Versus Rated Performance

Within the plate pack, the stacked set of thin plates that form the heat-transfer channels, the uniformity of fluid distribution across each plate's full cross-section is the largest single factor determining whether a heat exchanger achieves its theoretical performance in the field. Uneven distribution where fluid bypasses portions of the heat transfer area or stagnates in others simultaneously reduces heat transfer and increases pressure drop.

The inlet distribution section of the plate, the triangular zone where fluid transitions from the inlet port into the main rectangular heat transfer area, is where flow uniformity is either established or compromised. Optimizing this section requires detailed fluid dynamics simulation; it cannot be adequately characterized through analytical methods alone. The introduction of computational fluid dynamics to plate heat exchanger design fundamentally changed what was achievable in this area.

Fewer Units, Lower Installed Cost: The Port Size Advantage

The installed cost of a heat exchanger system extends well beyond the price of the unit itself. Larger port openings enable higher flow rates per unit, which reduces the total number of units required to meet a given capacity target. Fewer units means fewer isolation valves, less interconnecting pipework, reduced installation labor, and fewer components requiring ongoing maintenance.

Redundancy requirements amplify this advantage. Meeting N+1 or 2N requirements with fewer, higher-capacity units directly reduces the total installed cost, valve count, pipework, and ongoing maintenance burden.

Fluid Compatibility: One Heat Exchanger Across Every Cooling Architecture

Data center cooling loops use a range of fluid combinations depending on the application, climate, and internal cooling architecture. Water-to-water configurations are common in temperate climates while glycol-to-water is standard where freeze protection is required. Facilities deploying direct-to-chip or immersion cooling may use special thermal fluids in secondary loops.

A single plate design that performs across all of these fluid combinations simplifies specification and accommodates future changes in the facility's cooling architecture without requiring additional equipment.

Why Total Cost of Ownership Is the Right Metric

Engineers making heat exchanger purchase decisions should evaluate total cost of ownership over 30 years, not upfront capital cost alone. A unit with a lower purchase price that requires frequent maintenance, has unreliable spare parts availability, or imposes a wider temperature approach will cost significantly more over the life of the facility than a higher-quality unit purchased at a premium. The dominant long-term cost factors are energy consumption (governed by temperature approach and pressure drop performance), maintenance frequency and labor, and unplanned downtime. For a heat exchanger operating continuously for 15 to 30 years, these costs dwarf the original purchase price.

Cooling tower-connected systems introduce water quality management as a maintenance variable. Particulates, biological matter, and dissolved minerals drawn in through the tower can accumulate in the plate channels and degrade thermal performance over time. The ease with which plates and gaskets can be accessed and replaced without removing or disturbing the frame directly affects how quickly and cost-effectively this maintenance can be performed.

When spare parts are expensive, difficult to source, or of inconsistent quality, operators often find it more economical to replace the entire heat exchanger rather than service the existing unit. **The reliability of spare parts availability and the quality of the manufacturer's service organization are critical to the total cost of ownership calculation.**

The SIGMA M186 Heat Exchanger:

Where the Recooling Problem Gets Solved

The engineers at API Heat Transfer designed the SIGMA M186 to address each of the specification criteria discussed above. The result is a heat exchanger that **delivers a measurable performance advantage at the variable that matters most, temperature approach**, while also meeting the system design, fluid compatibility, and lifecycle requirements of mission-critical data center cooling applications.

The M186 is the bridge between the facility's internal cooling loop and the outdoor cooling circuit, and that position makes it compatible with any internal cooling architecture. Air-side coils, rear-door heat exchangers, direct-to-chip CDU loops, and immersion cooling secondary circuits all deliver heat load to the internal water loop. The M186 transfers that load to the outdoor recooling circuit regardless of its origin. For facilities running multiple cooling architectures simultaneously or planning to expand liquid cooling capacity as AI workload density grows, the recooling infrastructure does not need to change. The M186 serves the load in whatever form it arrives.

The financial impact of that performance scales directly with facility size. A data center moving 15 MW of heat saves a certain amount for every additional free cooling day the M186 enables. A facility moving 150 MW saves proportionally more. As AI workloads continue to drive facility heat loads upward, the value of each degree of improvement in temperature approach increases correspondingly.

A Validated 0.5 °C Temperature Approach

The M186 achieves a temperature approach between the internal and external fluid circuits of 0.5 °C compared to an industry standard of approximately 2 °C and a best-practice design target of approximately 1 °C. This performance was further validated in a recent installation with a global hyperscale AI infrastructure operator.

Operationally, 0.5°C results in approximately 10 additional free cooling days per year at a typical mid-latitude U.S. installation, relative to a unit operating at the 2 °C industry standard. Those days represent chiller energy that the facility does not consume.

What 10 Additional Free Cooling Days Are Worth

Reducing the temperature approach from 1.1 °C (2 °F) to 0.5 °C (0.9 °F) allows free cooling to absorb heat that would otherwise require the chiller, and the financial value of that shift scales with facility size. The table below, based on analysis from API Heat Transfer's engineering team, puts numbers to that relationship. Actual savings will vary by site; API engineers can assist with application-specific analysis.

Illustrative Annual Energy Savings:

M186 Temperature Approach Improvement

Assumes:

- 100 kW per rack
- Temperature approach reduced by 0.5 °C (0.9 °F)
- Chiller COP of ~3.1 (32 kW input per 100 kW cooling)
- U.S. commercial electricity rate of $0.1364/kWh (EIA)

Facility size	Rack count	Energy saved per day	Estimated daily savings
Mid-size (~5 MW)	~50 racks	~265,000 kWh	~$35,000
Hyperscale (~20 MW)	~200 racks	~1,060,000 kWh	~$140,000
AI campus (~100 MW)	~1,000 racks	~5,300,000 kWh	~$700,000

Actual savings vary by chiller efficiency, local electricity rates, facility load profile, and climate. API Heat Transfer can provide application-specific engineering analysis.

Plate Geometry Engineered for Maximum Efficiency

API Heat Transfer applies computational fluid dynamics (CFD) to the M186 plate design to model fluid behavior inside the channels and optimize performance. The primary focus is the inlet distribution section, the triangular zone where fluid transitions from the port into the main heat transfer area, so that flow is as uniform as possible across the full plate cross-section. This optimization improves heat transfer coefficients by roughly 5%, which translates to 15% less heat transfer area for the same thermal duty, enabling a **smaller footprint, lower unit weight, and lower installed cost per MW.**

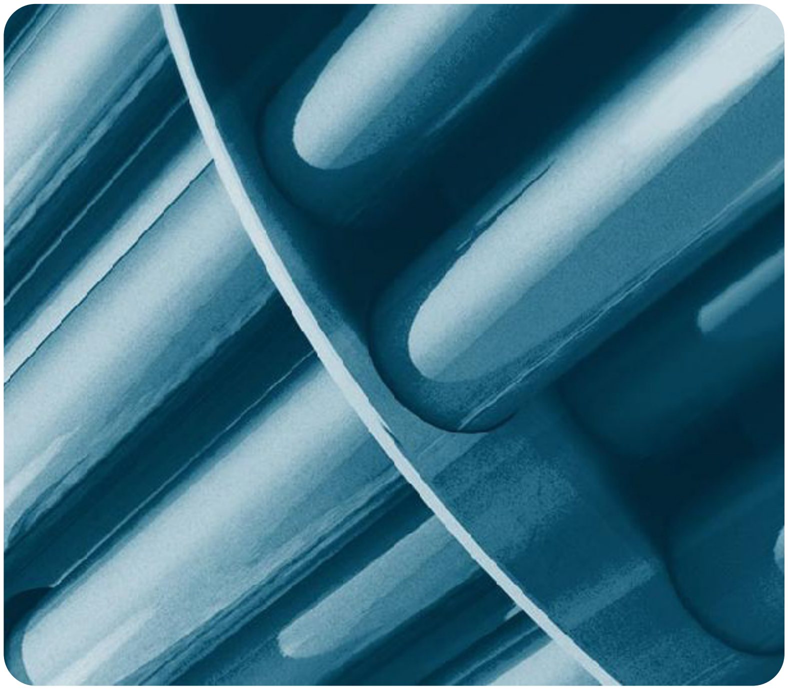
The same CFD-driven geometry also reduces pressure drop rather than increasing it, so efficiency gains in heat transfer and pumping power reinforce each other. In addition, the corrugation forming process stretches the metal, increasing the effective heat transfer surface area by approximately 20% compared with a flat sheet of the same base dimensions.

Every plate in the SIGMA platform undergoes CFD analysis of its flow distribution profile before manufacture. AI-powered visual inspection and a proprietary crack-detection system are applied to every plate in production, helping ensure the geometry validated in simulation is faithfully reproduced in the finished component.

Scalable Capacity with a Smaller Installation Footprint

A single M186 unit delivers 15+ MW of cooling capacity, with output scalable through temperature differential and flow rate configuration. Multiple units in parallel serve larger loads without constraint, with total system capacity determined by the facility's requirements and redundancy configuration.

The M186 heat exchanger's enlarged port diameter accommodates higher per-unit flow rates, reducing the number of units required to meet a given total capacity target, which directly reduces valve count, pipework, installation labor, and the number of components subject to maintenance. For facilities requiring N+1 or 2N redundancy, this translates to lower capital cost for equivalent coverage.



Six Rated Applications, One Heat Exchanger

The M186 accommodates the full range of fluid combinations common in data center cooling plants, making it applicable across all standard cooling architectures without the need for application-specific units. The SIGMA M186 is rated for six data center cooling applications:

- Water-side economizer cooling: Transfers heat from the data center loop to cooling towers or dry coolers, enabling free cooling and reducing chiller runtime.
- Primary-to-secondary loop isolation: Separates facility water from critical IT cooling loops to maintain cleanliness and protect sensitive equipment.
- Direct-to-chip liquid cooling support: Removes heat from coolant distribution unit (CDU) loops serving CPUs and GPUs operating at high heat flux.
- Rear-door heat exchanger integration: Supports rack-level cooling by rejecting heat from rear-door coil loops to facility water systems.
- High-temperature warm water cooling: Handles elevated return temperatures of 27–49 °C (80–120 °F) for improved energy reuse and system efficiency.
- Immersion cooling secondary loop: Cools the fluid loop connected to immersion tanks in ultra-high-density deployments.

In-House Manufacturing and Quality Assurance

Every component of the M186 is designed and manufactured in-house by API Heat Transfer. This full ownership of the manufacturing process ensures consistent performance across production runs and, critically, direct control over spare parts availability. The result is that **parts are available whenever and wherever they are needed.**

Plates are designed, tooled, and stamped in-house. Every plate undergoes AI-assisted visual inspection, and both new and reconditioned plates are inspected using API's proprietary crack detection system.

Gaskets are designed in-house with multiple material options for different fluid and temperature applications, robotically sanded, and either clipped or glued in-house.

Frames are designed, built, and assembled in-house, with multiple material configurations available for different applications.

The M186 carries third-party certification under ASME, PED (EU Pressure Equipment Directive), AS1210 (Australian standard), NR-13 (Brazilian standard), and CRN (Canadian Registration Number).

Service, Maintenance, and 30-Year Lifecycle Value

The M186 is designed for field serviceability. Plates and gaskets are replaceable without removing the frame, and the modular construction enables capacity to be increased by adding plates as the facility's heat load grows. **With a proper maintenance program in place, a gasketed plate-and-frame heat exchanger can operate reliably for 30 years or more.**

This 30-year service life is significant in the context of a data center mechanical plant, where chillers, cooling towers, and pumping systems are typically designed to last for 15 to 20 years. A heat exchanger that outlasts two generations of surrounding equipment, requiring only periodic plate and gasket replacement, represents a compounding advantage when evaluated against total facility lifecycle cost.

API Heat Transfer offers service contracts covering scheduled maintenance, field inspection, and priority access to replacement components. For facilities where 100% cooling availability is a hard operational requirement, this service model, combined with a properly sized redundant unit configuration, provides a structured approach to maintaining continuous uptime across the installation's life.

Specify with Confidence

The right heat exchanger specification pays dividends across the full operating life of a data center. Every additional free cooling day recovered through tighter temperature approach performance is chiller energy avoided. Every reduction in pressure drop is pump energy saved. Every year of reliable operation supported by a strong service relationship is downtime risk eliminated.

The SIGMA M186 was developed to meet the performance requirements of the most demanding hyperscale AI infrastructure operators and is available across the full range of data center applications, from colocation facilities to enterprise campuses to large-scale AI deployments. Its 0.5 °C temperature approach, CFD-optimized plate geometry, scalable capacity, and fully in-house manufacturing and service capability deliver the same performance advantage at any scale.

Ready to redefine your cooling infrastructure? Talk to an API Heat Transfer engineer about the SIGMA M186.

API Heat Transfer has been helping customers of all sizes and all industries solve their heat transfer needs for 150 years. As thermal transfer has evolved, so have we. For the engineers responsible for getting it right, API Heat Transfer is the manufacturing partner built to support that outcome, from first specification through the life of the installation.

The performance data and savings examples in this document are illustrative only and based on specific operating assumptions. Actual performance and energy savings will vary by installation. API Heat Transfer can provide application-specific engineering to validate performance for individual projects.

API HEAT TRANSFER