Solving Cooling Problems in Today's Power-Dense Data Centers

BY JEFF GRAU; HARISH KRISHNASWAMY; JOHN WITHOUSE, MEMBER ASHRAE

In May 2024, Goldman Sachs estimated that data center power demand will grow an unprecedented 160% by 2030, largely due to skyrocketing demand for data and more sophisticated data apps (including artificial intelligence) such as ChatGPT. On average, a ChatGPT query needs nearly 10 times as much electricity to process as a Google search.[1] More energy is needed to both power the technologies and to cool the systems. This is because the power used for IT equipment is electricity, which is released as heat. The growth of sophisticated AI apps drives the need for more efficient cooling alternatives.

As servers become more powerful, they produce more heat per square foot. Not so long ago, average power densities would sit around 4 kW per information technology (IT) rack.[2] Today, that requirement has jumped to the 12 kW to 30 kW range and higher. Managing heat removal is one primary challenge within data centers today as traditional cooling methods, such as air cooling, struggle to keep up with this increased cooling load. That's why many of the newer chips can only be installed with liquid cooling systems.

Liquid Cooling Takes Center Stage

Liquid is a far more efficient coolant than air, with the ability to carry away more than a thousand times more heat. Other benefits of liquid cooling include more efficient heat transfer, lower cost and less environmental impact. Liquid cooling significantly reduces energy and water consumption, which reduces total operating expenses and creates a more sustainable data center. Additional benefits of liquid cooling include both noise reduction and extended life of computer hardware.

Direct-To-Chip Liquid Cooling Options: Single Phase vs. Two-Phase

Liquid cooling can be categorized into two main types: single-phase (*Figure 1*) and two-phase (*Figure 2*). Both use a cold plate heat exchanger mounted directly onto the energy-dense components (central processing unit [CPU], graphics processing unit [GPU], etc.).

Jeff Grau is division engineering manager for Parker Hannifin's Parflex Division. Harish Krishnaswamy is division engineering manager and John Withouse is senior principal engineer for Parker Hannifin's Sporlan Division.

With single-phase cooling, a coolant such as water glycol (a mixture of water and propylene glycol) or deionized (DI) water with additives used to inhibit biological growth and limit corrosion is used and circulated inside the cold plate by cooling distribution units. The coolant absorbs heat as it passes through the cold plate heat exchanger, which is in direct contact with the GPUs and CPUs to promote effective heat transfer.

In two-phase cooling, heat absorption primarily occurs through latent heat during the phase change of the refrigerant. This efficient cooling process continuously cycles the refrigerant within a sealed, closed-loop system using a small pump to deliver just enough liquid refrigerant to the evaporator—usually a series of one or more cold plates optimized to acquire the heat from the IT gear. By doing so, the liquid refrigerant begins to boil, which maintains a cool, uniform temperature on the surface of the chips. The two-phase or vapor-phase refrigerant is then transferred to a heat exchanger where it rejects the heat to the facility cooling loop.

FIGURE 1 Data center piping diagram—single-phase setup.

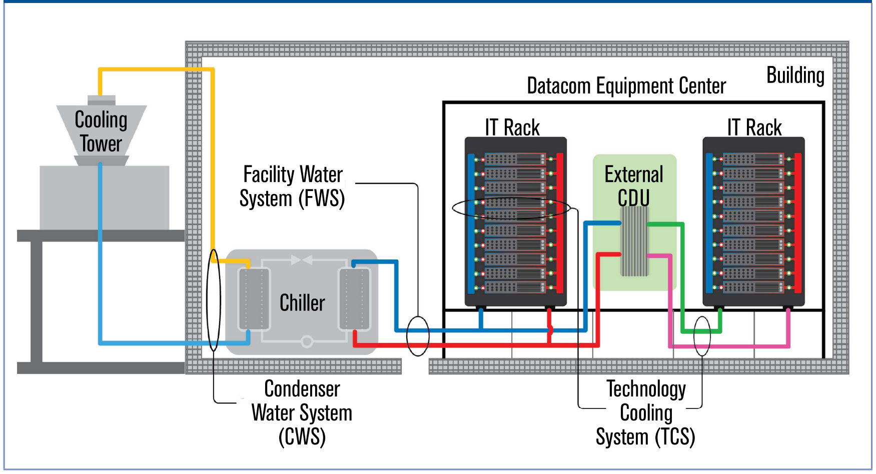
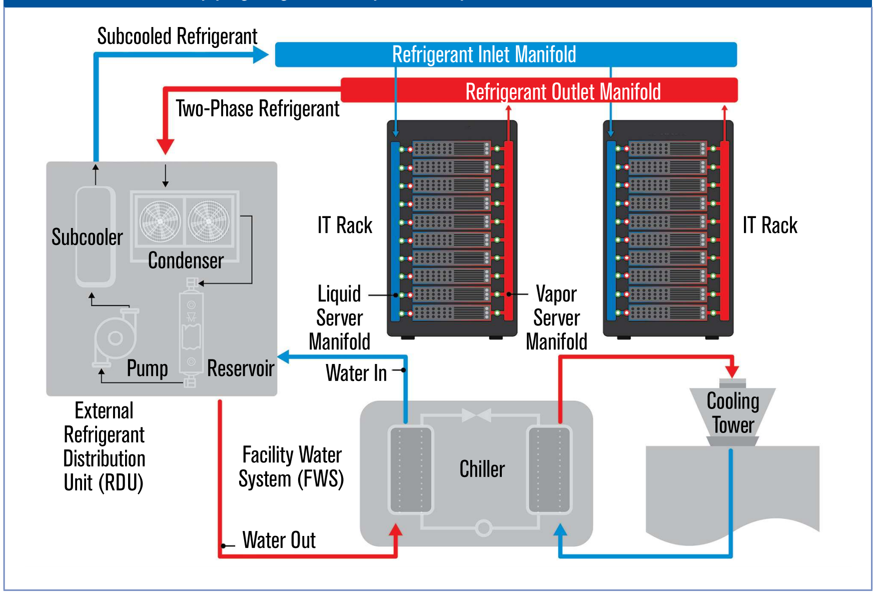


FIGURE 2 Data center piping diagram—two-phase setup.



Two-phase systems are ideal for high-power electronics systems where heat loads have moved beyond what traditional air and water (single-phase) cooling systems can effectively manage. The combination of a refrigerant with the science of heat dissipation through evaporation enables deployment of electronic components with increased power densities. The highly efficient design simplifies the plumbing and reduces the overall weight of the system, giving it an excellent thermal performance/cost ratio.

Sustainability and Safety Considerations

Data center water consumption is a global concern as some of the larger individual data centers consume billions of liters of water each year and are projected to experience double-digit year-over-year growth to keep up with newer, data-rich technologies.

A two-phase system, however, uses little to no water (depending on the region and climate), which makes it a preferred option from a water consumption standpoint.

Several viable choices of working fluids for two-phase direct-to-chip cooling exist. Factors that should be considered when making a choice include refrigerant classification, environmental impact and system design requirements.

ANSI/ASHRAE Standard 34, *Designation and Classification of Refrigerants*, designates single-molecule refrigerants and blends, including hydrofluorocarbons (HFCs) and hydrofluoroolefins (HFOs). An HFC that has been used for many applications, including two-phase electronics cooling, is R-134a. Among the HFOs of interest in this application are R-1234yf, R-1234ze(E) and R-1233zd(E).

The blends of interest for two-phase direct-to-chip cooling are combinations of HFCs and HFOs. ASHRAE Standard 34 designates azeotropic blends (that behave as a single molecule and do not exhibit temperature glide) in the R5xx series of blends. Because of the characteristics of operation of these two-phase systems, azeotropic blends will be required if a blend is used.

Standard 34 (*Table 1*) classifies refrigerants based on flammability and toxicity characteristics, creating a general hazard profile for each one. Toxicity is based on the NIOSH occupational exposure limit, (OEL) and flammability classifications go from "no flame propagation" to "mildly flammable" to "highly flammable."

TABLE 1 ASHRAE Standard 34 refrigerant.

CLASSIFICATION	OEL>400 ppm	OEL<400 ppm	
1. Flame Propagation @ 60°C, 1 atm *and* 2. LFL <0.30 kg/m^3 *or* HOC >19 000 kJ/kg	A3	B3	HIGHER FLAMMABILITY →
1. Flame Propagation @ 60°C, 1 atm *and* 2. LFL >0.30 kg/m^3 *and* 3. HOC <19 000 kJ/kg	A2	B2	
1. Flame Propagation @ 60°C, 1 atm *and* 2. LFL >0.30 kg/m^3 *and* V_t <10 cm/s and 3. HOC <19 000 kJ/kg	A2L	B2L	
No Flame Propagation @ 60°C, 1 atm	A1	B1	
	HIGHER TOXICITY →		

OEL = NIOSH occupational exposure limit
LFL = Lower flammability limit
HOC= Heat of combustion
V_t = Flame speed per ASTM E681

In response to concerns with the global warming potential (GWP) of HFCs, the American Innovation and Manufacturing (AIM) Act was created to provide a production and consumption phasedown of 18 HFCs. It bans the use of certain higher-GWP refrigerants across a broad range of applications.

Starting in 2024, the U.S. is now at 60% of the baseline established for the production and consumption phasedown. By 2029, that will drop in half again to 30%. The phasedown is GWP-weighted, meaning lower-GWP refrigerants are favored over higher ones.

The AIM Act Technology Transition Final Rule imposes GWP limits on new equipment across a broad range of air-conditioning, refrigeration and industrial cooling applications. For data centers, computer room air conditioning (CRAC) and information technology equipment (ITE) cooling, no new systems using a refrigerant greater than 700 GWP may be installed after Jan. 1, 2027. Both the phasedown and technology transition regulations require a careful selection of refrigerants for two-phase direct-to-chip cooling systems.

With background information on refrigerant classification and environmental impact, it's easier to choose options that are available and viable. R-134a is incumbent in similar applications and serves as a good baseline.

TABLE 2 Low pressure refrigerant snapshots.

	R-134a	R-1234yf	R-1234ze(E)	R-513A	R-515B
REPLACES	R-12	R-134a	R-134a	R-134a	R-134a
COMPOUND TYPE	HFC	HFO	HFO	HFO Azeotropic Blend	HFO Azeotropic Blend
COMPONENTS	Pure	Pure	Pure	44% R-134a 56% R-1234yf	8.9% R-227ea 91.1% R-1234ze(E)
ASHRAE SAFETY CLASS	A1	A2L	A2L	A1	A1
GLOBAL WARMING POTENTIAL	1,430	<1	<1	631	299
GWP VS. R-134a	0.0%	–99.9%	–99.9%	–55.8%	–79.1%
NORMAL BOILING POINT	–14.9°F	–21.0°F	–2.2°F	–20.6°F	–1.7°F
CRITICAL POINT	213.9°F, 588.8 psia	202.5°F, 490.6 psia	228.9°F, 527.2 psia	205.8°F, 546.2 psia	227.8°F, 518.4 psia
TEMPERATURE GLIDE	0.0°F	0.0°F	0.0°F	0.0°F	0.0°F
EPA SNAP STATUS	AIM Act Eliminates Use In New Systems 2025 – 2028	Approved	Approved	Approved	Approved

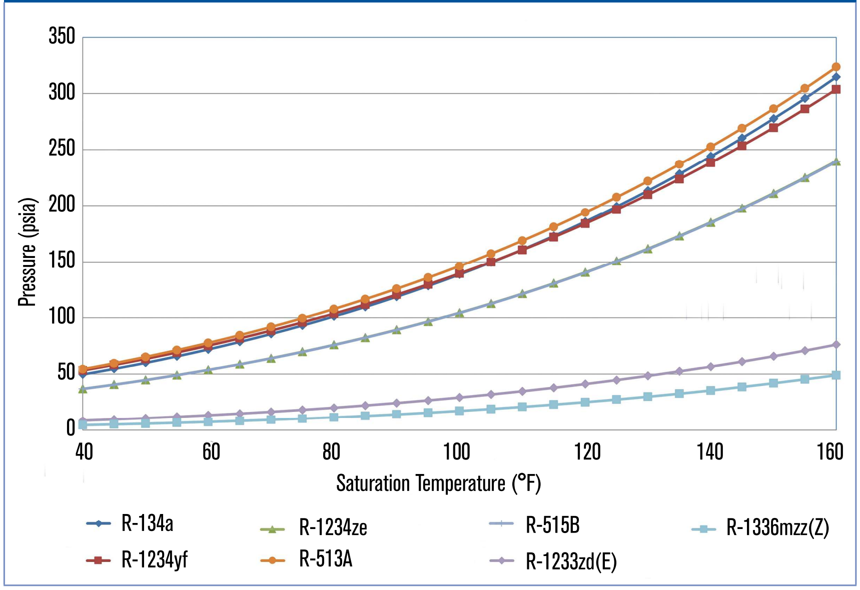
The choices in *Table 2* can all be considered low-pressure and R-134a–like, particularly in two-phase direct-to-chip systems. Among these are R-1234yf and R-1234ze(E), both pure, single-molecule HFO refrigerants with a GWP less than 1, no temperature glide and similar critical points. R-513A and R-515B are two A1-classified azeotropic blends to consider. Their GWPs are 631 and 299, respectively, both within the GWP 700 limit for data center equipment set by the AIM Act Technology Transitions rule.

Refrigerants contained in a blend do not chemically react with or alter one another but simply coexist in

TABLE 3 Very low pressure refrigerant snapshots.

	R-1233zd(E)	R-1336mzz(Z)
REPLACES	R-123	R-123
COMPOUND TYPE	HFO	HFO
COMPONENTS	Pure	Pure
ASHRAE SAFETY CLASS	A1	A1
GLOBAL WARMING POTENTIAL	<1	2
GWP VS. R-123	–99.0%	–97.4%
NORMAL BOILING POINT	65.0°F	92.0°F
CRITICAL POINT	330.1°F, 518.9 psia	340.3°F, 421.0 psia
TEMPERATURE GLIDE	0.0°F	0.0°F
EPA SNAP STATUS	Approved	Approved

FIGURE 3 Vapor P-T relationships.



the same space. For this reason, establishing material compatibility with each constituent of a blend typically establishes compatibility of the blend. Compatibility of polychloroprene, NBR, HNBR, PTFE, polyamide and PPS has been established with R-513A and R-515B, including their constituents R-134a, R-1234yf, R-1234ze(E) and R-227ea.

Two very low-pressure options also should be considered (*Table 3*). R-1233zd(E) and R-1336mzz(Z) are both A1 classified and exhibit a very low GWP in the range of 1 to 2. Material compatibility of both is reasonably well understood, though the total industry database for these fluids is likely less complete than for the R-134a–like HFOs and blends. Their compatibility with polychloroprene, NBR, HNBR, PTFE, polyamide and PPS is also generally regarded as good.

The thermodynamic and pressure-temperature (P-T) characteristics of the different options must be considered in the design of system components. *Figure 3* shows the P-T curves of the refrigerants in comparison to R-134a.

System design pressures for R-513A or R-1234yf would be similar to that of an R-134a system—around 325 psig to 350 psig (2.241 MPa to 2.413 MPa). Systems using R-1234ze(E) or R-515B could be designed for the 250 psig to 275 psig (1.724 MPa to 1.896 MPa) range. Systems using R-1233zd(E) or R-1336mzz(Z) could be designed for as low as 70 psig to 80 psig (483 kPa to 552 kPa).

It is also necessary to consider the volumetric capacity (mass flow per unit cooling capacity) of each refrigerant over a range of operating temperatures. The R-134a–like HFOs and blends generally require similar, to as much as 25% greater, mass flow in comparison to R-134a. The very low-pressure HFOs require similar mass flow at lower temperature conditions but significantly less mass flow as temperatures increase.

However, pressure and mass flow characteristics alone do not represent a complete set of parameters for selecting a refrigerant and sizing system components. The refrigerants under consideration exhibit significant differences in liquid density and even greater differences in return vapor density that must be taken into account. *Table 4* shows mass flow information converted to liquid side volumetric flow using saturated liquid density.

Liquid-side component selection and pipe sizing for R-1234ze(E) and R-515B will be similar to that of R-134a. Both R-1234yf and R-513A require 15% to 25% greater volumetric flow, potentially requiring larger pumps and liquid-side piping. The very low-pressure HFOs will require lower liquid volumes to be pumped, particularly as temperatures rise.

Volumetric flow for the return side of the system is shown in *Table 5* using saturated vapor density. Return side components and pipe sizing for R-1234yf and R-513A can be similar to R-134a. Volumetric flow requirements for R-1234ze(E) and R-515B are from 13% to 19% greater. However, the lower return density of the very low pressure HFOs means their return volumetric flow rates are greatly increased—in the range of three

TABLE 4 Liquid volumetric flow per ton (ft^3/min).

SAT TEMP (°F)	R-134a	R-1234yf	R-1234ze(E)	R-513A	R-515B	R-1233zd(E)	R-1336mzz(Z)
70	0.034	0.045	0.037	0.041	0.038	0.030	0.032
90	0.037	0.050	0.041	0.045	0.041	0.032	0.034
110	0.041	0.057	0.045	0.051	0.045	0.034	0.035
130	0.047	0.066	0.050	0.059	0.051	0.036	0.038
150	0.055	0.079	0.057	0.071	0.058	0.038	0.040

TABLE 5 Return vapor volumetric flow per ton (ft^3/min).

SAT TEMP (°F)	R-134a	R-1234yf	R-1234ze(E)	R-513A	R-515B	R-1233zd(E)	R-1336mzz(Z)
70	1.428	1.480	1.874	1.390	1.877	6.108	9.855
90	1.087	1.147	1.418	1.073	1.422	4.351	6.854
110	0.849	0.912	1.098	0.847	1.103	3.206	4.861
130	0.678	0.742	0.873	0.687	0.876	2.396	3.552
150	0.557	0.621	0.709	0.571	0.712	1.848	2.666

times to six times those of the R-134a–like HFOs and blends.

Several trade-offs should be considered in selecting a refrigerant for two-phase direct-to-chip cooling systems. A1-classified refrigerants are desirable to avoid the restrictions and additional safety requirements placed on a system using A2L mildly flammable refrigerants. Very low GWPs are desirable to reduce climate impact. Lower-system design pressures and volumetric flow rates positively impact system cost.

The very low-pressure HFOs deliver A1 classification, very low GWP and low design pressures. They also enable use of relatively small liquid side components, such as pumps and piping, but will require significantly larger return side components and piping. R-1233zd(E) and R-1336mzz(Z) will be more sensitive to pressure drop within the system. The cost of upsizing return side components for these refrigerants is partially offset by the lower design pressure requirements. R-1234ze(E) and R-515B can offer somewhat lower design pressure with modest increases in liquid and return side components and pipe sizing.

Despite the many advantages of refrigerants, they are not without their disadvantages. Regulations limiting certain polyfluoroalkyl substances (PFAS) will have implications for both two-phase cold plates and two-phase immersion cooling. Depending on the refrigerant chosen, additional training and certifications may also be needed for the operators and maintenance personnel to ensure they are knowledgeable about best practices for handling them in a safe and environmentally responsible manner.

Single-Phase and Two-Phase System Connections

When rigid metal tubing is not preferred, flexible hose or tubing is required to connect the various fluid loop components. Hose refers to polymeric material that is reinforced with wire or synthetic fiber for the purpose of pressure containment. Tubing, on the other hand, does not use reinforcement and relies on the strength of the tubing material to resist deformation under pressure.

Tubing may be constructed with a single material (mono-wall) or comprised of multiple material layers.

Rubber hoses are constructed from a range of thermoset polymers, such as EPDM, nitrile, polychloroprene, chlorinated polyethylene, butyl, natural rubber, SBR and fluorinated rubber. Each can be custom compounded for specific properties.

Rubber hose is typically the choice that provides the greatest flexibility with a low force-to-bend. This flexibility allows for routing with a relatively tight bend radius without kinking, ripping or tearing. Lastly, due to the flexible nature of the material, rubber hose can be easily sealed with a variety of barb designs with minimal assembly force.

Thermoplastic hose and tubing can be constructed from a range of polymer types such as polyamide, polyethylene, co-polyester, polypropylene, urethane, PTFE, PFA, FEP and others. Thermoplastic hose and tubing can offer advantages in chemical compatibility, fluid extraction, permeation resistance and thinner wall sections (smaller outer diameter [OD] for less space claim). Where tight bends are required, thermoplastic hose and tubing can be formed or shaped to avoid kinking. Sealing the ends of the thermoplastic hose or tubing often requires an O-ring or assembly to a specialized barb geometry that may require higher assembly forces.

Rubber and thermoplastic hose and tubing are currently deployed in the single-phase liquid cooling application. When choosing a hose or tubing for the two-phase cooling application, it's important to consider the following characteristics:

- **Permeability.** Two-phase refrigerants can potentially migrate through the hose wall due to the small molecule size of the vapor phase refrigerant. To avoid unwanted fluid loss, the hose or tubing may use a barrier layer of dense material to resist permeation. The permeation rate is impacted by the temperature of the hose materials, so the application temperature should be considered when selecting a hose or tubing. The use of a barrier layer in a hose or tubing construction can negatively impact the flexibility of the product.
- **Pressure Rating.** Hose and tubing can be used to contain pressurized fluid up to the product's maximum rated working pressure. A design/safety factor is used so the burst pressure for the hose or tubing is higher than the maximum rated working pressure. The design factor could be 3:1, 4:1 or higher depending on the application's guidelines for safety.
- **Dimensions.** Hose and tubing are often characterized by inside diameter (ID) and outside diameter (OD). ID is an important characteristic when predicting heat transfer performance as it will influence volumetric flow rate and pressure drop. Corrugated or convoluted hose or tubing can be used to enhance the flexibility of rigid materials but will demonstrate significantly higher pressure drop when compared to smooth ID products. The OD of the hose or tubing is critical due to space claim restrictions in the server blade or rack.
- **Bend Radius.** The minimum bend radius of the chosen hose or tubing is an important consideration, so designers should try to identify the location in the routing that requires the tightest bend to avoid kinking. Variables affecting the bend radius include the flexibility of the materials, wall thickness and method of reinforcement.
- **Fluid Compatibility.** The hose or tubing materials must be able to resist chemical attack from prolonged contact with the refrigerant. Fluid compatibility is impacted by temperature, so application conditions should be considered when gauging fluid compatibility. Additionally, depending on the refrigerant choice and application temperature, ingredients in the hose or tubing liner compound could be extracted from the material and leached into the refrigerant.
- **Moisture Ingression.** If the refrigerant selected is hygroscopic (attracts moisture), hose or tubing may require materials that serve as a moisture barrier to keep unwanted moisture from passing through the wall of the hose or tubing and into the fluid stream.
- **End Connection.** Rubber hose is often terminated with a hose barb that seals against the hose's ID. This connection may use a specialized barb geometry that does not require clamping or other means of retention such as a clamp or shell on the hose OD. Thermoplastic hose may use similar methods for termination but may require the use of an O-ring to seal between the hose ID and barb. Thermoplastic tubing may be sealed similarly but can also be sealed on the outside surface using an O-ring. In this case, the entire tubing wall cross section will experience fluid exposure. When using a termination that deploys an O-ring or other elastomeric seal, the compatibility of the seal and system fluid should be considered.
- **Environmental Protection.** Some applications may require the outer surface of the hose or tubing to be resistant to the attack from ozone or ultraviolet (UV) exposure. Proper material selection will increase the ability of the hose or tubing to resist oxidation of the polymer chain, which results in spider web-like micro cracks.
- **Safety Compliance.** Some applications may require a particular flame rating of the hose or tubing material. UL 94 burn testing is often used to characterize a material's resistance to ignition, flame propagation or ability to self-extinguish.
- **Chemical Compliance.** Applications may require hose or tubing to be compliant with various restricted material lists like REACH, RoHS, Prop 65, etc.

Another critical component of data center liquid cooling systems is couplings. Quick disconnect (QD) dry break couplings enable data center operators to hot swap servers, providing an ease of adding, removing and servicing servers within an IT rack. Performance metrics for QD couplings include minimal air inclusion, minimal fluid loss, high vacuum capability, end connection adapter compatibility, corrosion resistance, wide temperature range and high-pressure robustness. The ease and speed of connection and disconnection—and whether these operations require tools—are also important to consider.

One key characteristic of a QD coupling is the ratio of flow rate performance to coupling size. Flow performance is measured by flow rate versus pressure drop for the coupling. Maintaining low pressure drop for the required flow rate while occupying minimal cabinet space is desirable for all systems but is especially important for two-phase systems. Coupling size can be defined by two measurements, diameter and length. QD couplings that are compact in length provide maximum utility and flexibility for varying server rack configurations and connection arrangements.

Conclusion

As this article has noted, air cooling systems will likely always have a place, as they still represent a viable option for less power-dense servers. However, with the explosion of sophisticated AI apps and machine learning, today's data centers are larger, hotter and more power-dense, requiring a more efficient cooling alternative. Single-phase (water/glycol) direct-to-chip systems have grown in popularity and offer many advantages over traditional air systems. Today a number of hyper-scale data centers are deploying a combination of air- and liquid-to-chip cooling systems.

When considering two-phase systems, special attention should be paid to chemical compatibility and safety issues. Further, an array of design and mechanical considerations exist when choosing the right hoses, tubing and couplings.

As the industry starts to better understand the various technologies beyond water and glycol, there will be increased pressure to produce equipment using cleaner, safer next-generation refrigerants. It is projected that technologies will continue to evolve to further improve the properties and performance of today's refrigerants, making the opportunities for a cleaner, safer and more efficient cooling solution nearly endless.

References

1. Goldman Sachs. 2024. "AI is Poised to Drive 160% Increase in Data Center Power Demand." Goldman Sachs. https://tinyurl.com/27umyk7e

2. Data Centers.com. 2018. "Understanding the Interplay Between Data Center Power Consumption/Data Center Energy Consumption and Power Density." DataCenters.com. https://tinyurl.com/yj4rtszb ■