



30°C COOLANT - A DURABLE ROADMAP FOR THE FUTURE

Revision: 1.0

Steve Mills, META

Yin Hang, META

Andrea Moscheni, Vertiv

Jaechoon Kim, Samsung

Paul Artman, AMD

John Fernandes, META

Mark Steinke, NVIDIA

Chris Malone, META

Casey Carte, Intel

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Executive Summary

The Open Compute Project is responsible for fostering and serving the OCP Community to meet the market and shape the future, taking hyperscale-led innovations to everyone. To this end, this document seeks to provide an understanding behind why both data center operators and silicon providers are aligning on a durable coolant temperature. Coolant provided by the data center of 30°C at the silicon provides a target that extends the life of the data center investment and provides a large volume of users for silicon suppliers. While aligning on a specific coolant temperature is valuable to the industry, the industry will still need innovation across a wide variety of cooling technologies to maintain performance.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Table of Contents

1. Introduction	4
2. Compliance with Open Compute Project Tenets	5
2.1. Openness	5
2.2. Efficiency	5
2.3. Impact	5
2.4. Scale	5
2.5. Sustainability	5
3. Component Power Trends	6
3.1 GPU and CPU Power Trends	6
3.2 Thermal Stack in Modern Silicon	7
3.2.1 Various Junction Temperature Requirements	9
3.2.2 Various Stack Heights	9
3.2.3 Buried Components	10
3.2.4 Heat Crosstalk	10
3.3 HBM	10
4. Why Does Durability Matter?	13
4.1 From a Silicon Manufacturer's Point of View	13
4.2 From a Data Center Operators' Point of View	13
5. Why 30°C is best for Durability	15
5.1 30°C or above for the coolant temperature at the silicon	15
5.2 Data Center Energy Efficiency	17
6. How do we Maintain Durability in the Future?	26
6.1 High Bandwidth Memory Silicon Improvements	26
6.2 Improving IT Hardware-level Cooling Solutions	27
6.3 Improved Data Center Heat Exchangers	28
6.4 Silicon packaging innovations	31
5. Conclusion	35
6. Glossary	36
7. References	37
8. License	37
8.1. Creative Commons	37
9. About Open Compute Foundation	38

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

1. Introduction

While liquid cooling within computing systems has been around for many decades, the demand for cooling higher power densities is driving liquid cooling from boutique to the hyperscale. AI solutions must evolve quickly, but data centers have very long timelines to design, build and depreciate. So, it is valuable for both data center operators and silicon providers to agree on a coolant temperature where it is delivered to the silicon that is durable. A minimum coolant temperature of 30°C for the Technical Cooling System (TCS) loop offers a setpoint that provides both parties a valuable operating point. This document will provide the reasons behind why 30°C is the preferred temperature for cooling AI silicon in the future and identifies paths for new technological investments towards increased performance in the future.

The 30°C coolant temperature limit is not intended to discourage development of a wide range of future solutions. While there is a significant value in aligning on a coolant temperature of 30°C or greater, there is no desire to limit the types of technology used to deliver the cooling capability to the silicon. Freedom of innovation will be needed across many different cooling technologies to provide ideal solutions across the industry including immersion, cold plate, thermal interface materials and many others. There will also be demand for silicon solutions across a range of coolant temperatures including some opportunities below the 30°C target, however, the goal of this initiative is to target silicon with the best price per performance ratio above the 30°C coolant limit.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

2. Compliance with Open Compute Project Tenets

2.1. Openness

This initiative is actively working to educate the larger community about cooling trends in the marketplace, and encourage a larger, public discussion about future investments around cooling technologies and data center designs.

2.2. Efficiency

Selecting a durable coolant temperature provides an efficient operating environment for data centers and silicon designers alike. Choosing a temperature that is too high when designing a data center will result in a very short DC lifespan for high volume accelerator solutions in the coming years.

2.3. Impact

This paper is intended to impact the thinking of both silicon suppliers and data center operators in the coming years with the goal of optimizing the investments of industry and focusing their investments in cooling technologies in the future to provide improved performance.

2.4. Scale

Without aligning the roadmaps of silicon providers and data center operators, there will be a mismatch across the industry between silicon needs and infrastructure capability. While smaller operators will certainly ask for different silicon skus for data centers with both higher and lower coolant temperatures, this paper seeks alignment for mainstream operators.

2.5. Sustainability

While warmer coolant temperatures normally allow for more efficient heat transfer and opportunities for reusing waste data center heat, the goal of this paper is to pick a coolant temperature that is both useful and durable long term for mainstream users and the future demands of silicon.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

3. Component Power Trends

3.1 GPU and CPU Power Trends

The adoption of liquid cooling has greatly increased over the last few years due to the rapid increase in GPU/ASIC power consumption for AI/ML workloads. Liquid cooling has generally been a “niche” market primarily served by OEMs for HPC workloads. However, Cloud Service Providers (CSPs) and OEM adoption has greatly increased as CPU power has steadily increased from being flat at around 100W for a decade to greater than 400W over the last 5 years.

As shown in Figure 1 below, GPU/ASIC power growth has been more dramatic than CPUs with power increasing hundreds of watts on a close to annual basis. Silicon vendors such as NVIDIA, AMD, and Intel have published GPU powers greater than 1KW. In addition, CSPs such as MSFT, Google, Amazon, and Meta are designing their own silicon for AI/ML workloads with their future silicon requiring liquid cooling.

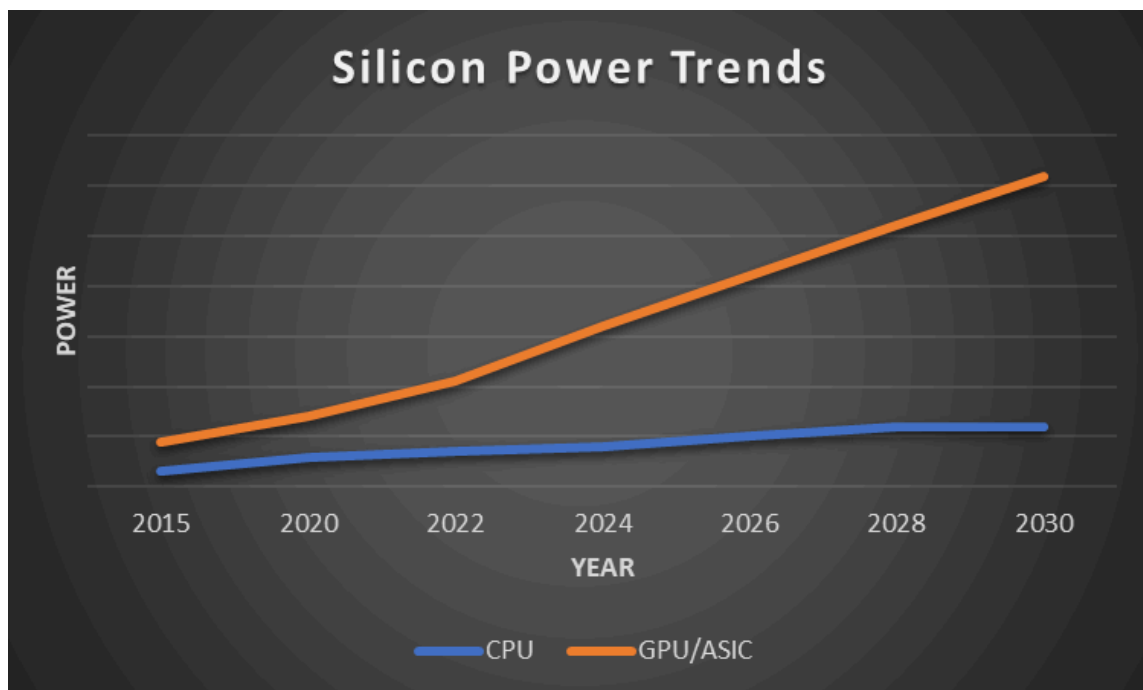


Figure 1. Silicon Component Power is Rapidly Increasing

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

3.2 Thermal Stack in Modern Silicon

The construction of silicon packages for typical CPUs has remained largely unchanged over the course of several generations. They are still primarily constructed using monolithic silicon. Die shrinks (reduction in feature critical dimensions) following Moore's Law have allowed for moderate increases in performance and power over time.

In contrast, the construction of GPUs has seen a dramatic shift in their silicon package construction. The use of multi-chiplet stacking in the form of 2.5 dimensional (2.5D) construction using Chip-on-Wafer-on-Substrate (CoWoS) has allowed for a more complex silicon package to be designed and built. Modern data center GPUs rely upon this process to tightly package the System-on-Chip (SOC) with High Bandwidth Memory (HBM). This allows for a different type and number of chiplets to be combined to create unique and more performant GPU packages. The improvements of process technology, package assembly techniques, processing performance requirements, and co-location of memory have largely driven the large power increases seen in GPUs.

Figure 2 shows a representative cross-section of a modern GPU stack construction. In this example, a GPU package is attached to the system PCB via BGA. There is a package substrate that carries the structure of the assembly. Typically, there is an active interposer die that handles power and signal distribution. The SOC and HBM are attached to the interposer die and something like silicon dioxide is added to fill any voids of height differences. A thermal interference material (TIM 1.5), often referred to TIM 1.5, is added to the assembly to allow thermal transfer between the bare dies and cold plate. Finally, a cold plate or heat sink is attached to the GPU to complete the full assembly.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

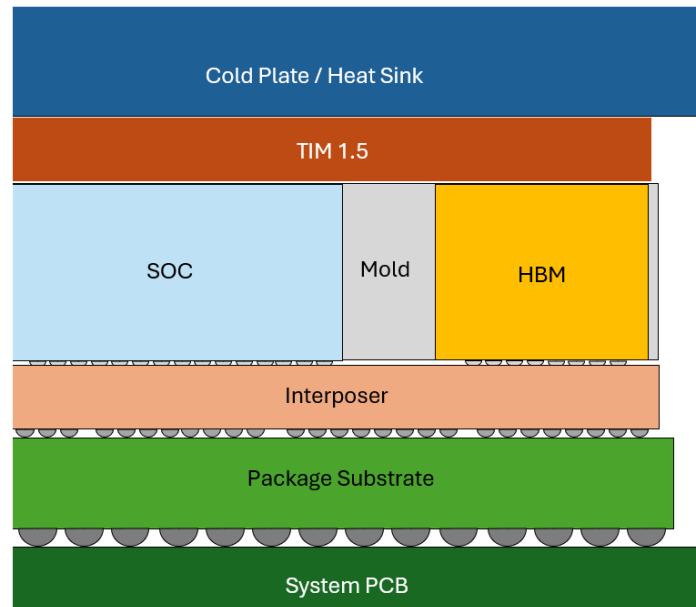


Figure 2. CoWoS Based Construction of a Modern GPU

The complexity of the package construction for these GPUs is much higher than CPUs and creates some unique thermal challenges in the package. Firstly, the different chiplets or components may require a different maximum junction temperature. Secondly, the different chiplets may have a different stack height. Next, the assembly can place heat generating devices in the middle of the stack without a direct path to the cooling solution. Finally, the close proximity of the different chiplet to each other can cause cross-talk heating that would not exist in monolithic silicon.

Figure 3 shows the different thermal paths contained within the GPU stack. Unlike other CPU based package assemblies, all of the heat is dissipated up into the top mounted cooling solution and no heat makes its way into the package substrate and into the system PCB. Due to the differing junction temperature requirements of the individual components, it is beneficial to track the thermal resistance of the SOC, $Q_{j,SOC-C,SOC}$, and HBM, $Q_{j,HBM-C,HBM}$, junction to case separately. In this context the case temperature would refer to the top of the silicon assembly due to not having an integrated heat spreader or lid.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

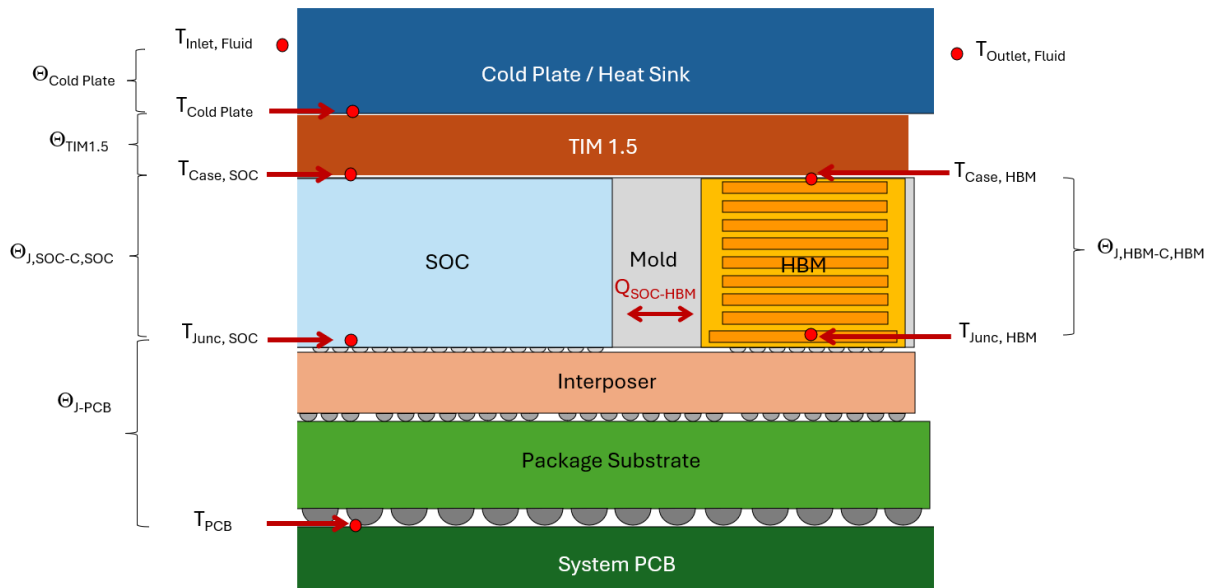


Figure 3. Thermal Paths for a Modern GPU using 2.5D Stacking.

3.2.1 Various Junction Temperature Requirements

The individual chiplets in the GPU package may be on different technology nodes, serve different functions, and operate with different junction temperature requirements. For example, the SOC chiplet is typically operating with a maximum junction temperature of 105°C. Whereas, the HBM will require a lower junction temperature of 85°C for single-refresh and 95°C for double-refresh operation in the early generations. Later generations of HBM have increased that to 95°C and 105°C, respectively, to better match with SOC junction temperature requirements. One must be mindful of the variation of junction temperature requirements in the package and plan the cooling solutions accordingly.

3.2.2 Various Stack Heights

The different components in the stack will have different individual heights. For example, a typical SOC will have a stack height of 450 micron whereas the HBM stack will have a height of 720 micron. Therefore, some other carrier silicon must be added to fill voids created by the various height differences. The additional material is added so that a uniform surface height is

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

seen by the cooling solution. The height difference will become worse over time as new HBM4 stacks are increasing the number of HBM die and hence will increase in height.

3.2.3 Buried Components

The assembly of these active components naturally leads to some of them being buried beneath other active heat generating components. Take the example of the Active Interposer. It can be passive or active and as such can produce some amount of heat that needs to be managed. The integrated devices in that device will have its own junction temperature requirements and must dissipate through other chiplets, like the SOC or HBM, in order to be cooled.

3.2.4 Heat Crosstalk

The heating of components from neighboring components creates a heat crosstalk situation within the stack. The SOC and HBM can create a heat flow between the devices, $Q_{\text{SOC-HBM}}$, that is not seen in CPU packages. This effect can be very workload dependent and generate a non-zero impact on the individual components. The impact of heat crosstalk should be studied using several different workloads and use cases to ensure it is accounted for in the overall cooling solution design.

3.3 HBM

The integration of High Bandwidth Memory (HBM) is another factor in the rapid increase in GPU power. The HBM devices have greatly increased in power consumption of the successive generations. Furthermore, the construction and addition of more HBM die in the total stack contribute to both the power increase and thermal resistance of the overall stack.

Figure 4 shows the typical construction of an HBM stack. The HBM stack starts with a logic die at the base of the assembly. The logic die routes signals, power, and connection to the discrete HBM die within the stack. The number of HBM die included in the stack is dependent upon the generation. The first generation of HBM supported two and four HBM die, referred to as 2H and 4H respectively. The third generation of HBM (HBM3) supports eight and 12 high stacks, 8H & 12H. Interestingly, the HBM vendor figured out a method to keep the overall stack height to 755 micron and still fit in the additional 4 HBM stacks. Finally, HBM is projected to support up to

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

sixteen HBM die, 16H. However, the overall stack height will have to increase to 810 micron to fit the additional HBM die.

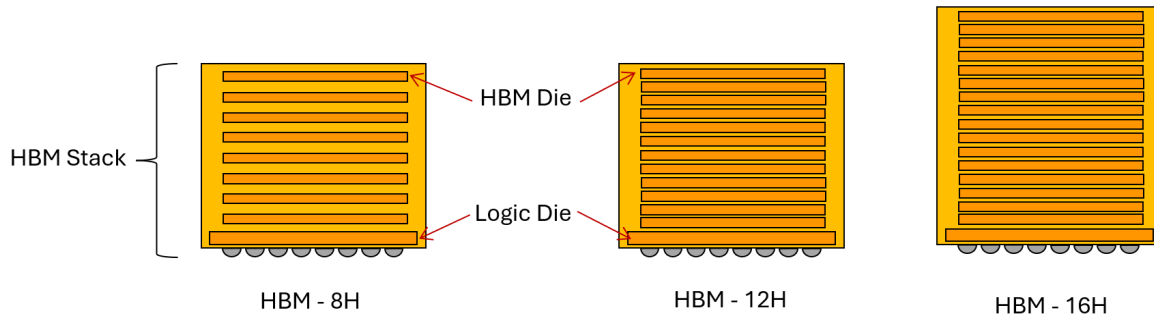


Figure 4. HBM with Multiple DRAM Stacks

The HBM power is also rapidly increasing from generation to generation. For example, the power has doubled from HBM2 to HBM3. The AI/ML workloads are driving this need for increased capacity and bandwidth and the only way to keep pace with that demand is to increase frequency (power) and memory size (number of dies). So, this trend is predicted to continue as HBM4 is close to doubling the HBM3 power. Figure 5 shows the progression of both the average stack power and stack height over the generations and year of introduction.

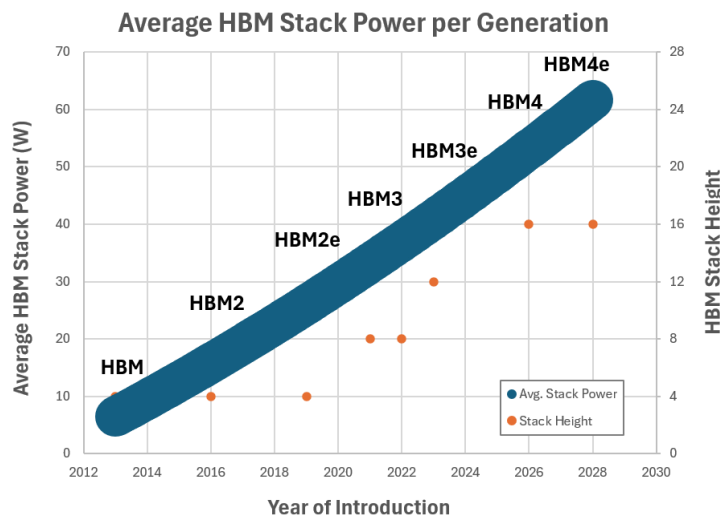


Figure 5. Average HBM Stack Power and Height per Generation

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

The power and stacking trends add to the thermal stack complexity of the GPU packaging. First, the raw increase in power consumption is driving a large increase in overall package power and neighbor heating effects to the SOC chiplets. Second, the additional HBM dies in the stack are adding to the 1D conduction problem within the stack. Lastly, the added height to accommodate the additional HBM dies creates more height differences within the package assembly.

Each HBM manufacturer is trying to address this trend based upon their individual expertise. Some HBM manufacturers are more power efficient in their design and are working to optimize that aspect to minimize the power consumption. Some HBM manufacturers are better at the assembly and can produce a stack with lower thermal resistance and work to decrease the amount of temperature loss in order to reach the logic die junction temperature. Others are working to include new packaging techniques, such as Hybrid-Compression Bonding (HCB) to get more thermal conductance within the stack to decrease the overall thermal resistance of the full stack. There is no single best approach and it will likely take a combination of these features in order to minimize the impact of this rapid increase in power and stack height.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

4. Why Does Durability Matter?

4.1 From a Silicon Manufacturer's Point of View

It is in the best interest of silicon manufacturers to have a durable datacenter such that the temperature requirements for silicon will not change with each generation. Next generation AI products have thermal design powers that require significantly cooler fluid temperatures than what some data center designers are currently planning for. By settling on an agreed-upon coolant standard, silicon manufacturers can design their products with the assurance that they can be sufficiently cooled. Additionally, a set temperature provides a good design direction for silicon manufacturers rather than having them develop products that:

1. may be less capable but operate in warmer environments
2. are highly capable but not widely adopted due to cooling limitations from datacenter operators.

A durable datacenter coolant temperature ensures there is a large market to sell into for silicon tailored to operate in a liquid cooled environment. While there will be demand for some silicon solutions across a range of coolant temperatures including some opportunities below the 30°C target, the goal of this initiative is to target silicon with the best price per performance ratio above the 30°C coolant limit.

4.2 From a Data Center Operators' Point of View

Data center operators have a vested interest in aligning on a durable coolant temperature for several reasons.

While it takes many years to plan, design, build and commission a data center, AI silicon is rapidly evolving. The industry benefits from quick iterations from silicon, but the lower limit of the coolant needed to deploy AI silicon were to change rapidly, then new data centers could be obsolete before they are even built. It is also relatively easy and efficient to operate a TCS loop at the hottest point of its temperature range at least up to some temperature value, however, it is

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

very expensive and time consuming to modify a TCS loop to operate below their existing design range once built. Therefore, defining the lower temperature limit of a data center's TCS loop temperature is very important to long term viability and investment in data center infrastructure.

Furthermore, as shown in [section 5.2](#), the annual energy consumed to cool a data center increases significantly when the TCS coolant temperature decreases from 30°C to 20°C in both hot and warm regions. 30°C coolant, while not ideal, also provides at least some options for heat re-use. Moving to coolant temperatures below 30°C are technically possible and available from industry, but are hindered by lower efficiencies.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](#)

5. Why 30°C is best for Durability

5.1 30°C or above for the coolant temperature at the silicon

Data centers must be designed for a specific coolant operating temperature and flow rate. If the temperature requirement for a particular generation of IT hardware is higher than design temperature, the data center operational set points may be adjusted to raise the coolant temperature. This offers the opportunity for improved efficiency. If the temperature requirement is lower than the data center design temperature, expensive and time-consuming modifications to the physical design may be required. For example, lower temperatures may require additional chiller capacity or a new type of chiller. Given the consequences of such a change, it is important to set the coolant temperature requirement such that it will not change through multiple generations of IT hardware.

This is especially challenging for AI hardware, given the rapid changes in silicon power and thermal requirements [1]. AI hardware is driving the need for liquid cooling. GPU power is increasing much faster than CPU power. The chart shows the associated fluid temperature requirements for GPUs over time. The asymptotic temperature requirement is 30°C.

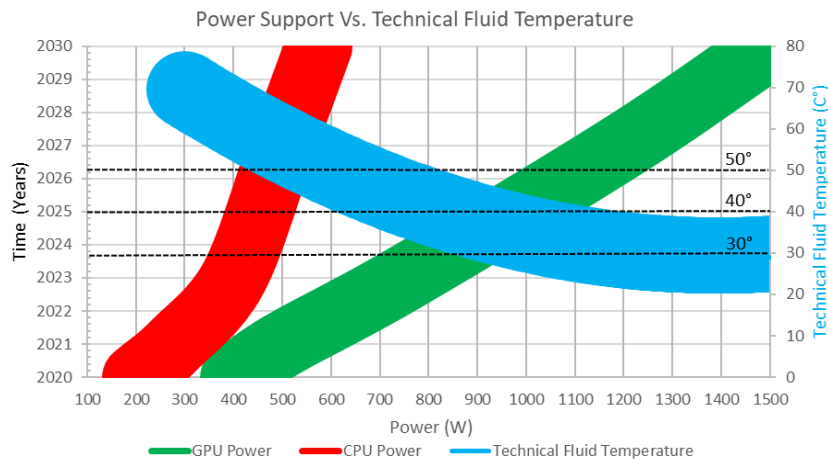


Figure 6. GPU and CPU power trends and associated technical fluid temperature requirements at 1.5 lpm per 1KW. [1]

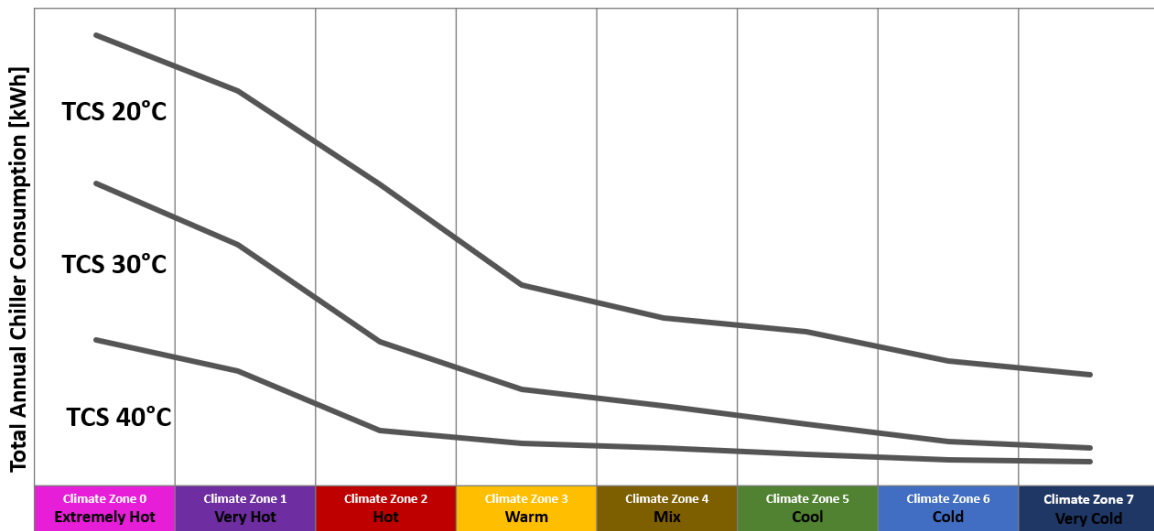
Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

We have aligned on 30°C as the minimum fluid temperature for hardware and data center design. The line thickness represents prediction uncertainty. Investments in advanced silicon packaging and liquid cooling performance are needed to maintain 30°C as a long-term interface specification.

The choice of 30°C matches a common minimum air temperature specification for large-scale data center design. Choosing the same value for air-and liquid-cooled IT hardware allows the industry to maintain the significant PUE improvements achieved over the past 15 years. Figure 7 illustrates the total annual power consumption for air-cooled free-cooling systems vs. climate zones for a range of technical fluid temperatures. The consequences of lowering the fluid temperature from 30°C to 20°C are evident. In hotter climate zones, the number of free-cooling hours in a year are reduced. Another study [1] showed a 20% reduction in annual chiller power consumption for the majority of climate zones as the fluid temperature was increased from 21°C to 26°C and another 20% for 26°C to 30°C.



Data Center in configuration N+1 for the cooling units and loaded at 80% of the maximum IT load (Temperature approach of 6°C)

Figure 7. Total annual consumption for an air-cooled free-cooling vs. climate zones for a range of technical fluid temperatures.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

5.2 Data Center Energy Efficiency

Traditional air-cooling methods are often insufficient to handle the thermal load efficiently, especially in high-density environments. This has led to the adoption of liquid cooling as an alternative or supplement to air cooling in data centers. The American Society of Heating and Air-Conditioning Engineers (ASHRAE) in 2011 set out broad classes for the fluid temperatures - W1, W2, W3, W4, and W5, based on the cooling temperature. Originally those classes were 17°C, 27°C, 32°C, 45°C and over 45°C, respectively [2]. When the work was updated in 2022, new temperature refinements were required, including a temperature of 40°C, and ASHRAE moved to new class definitions: W17, W27, W32, W40, W45, and W+. These temperatures are correlated by the following factors: the type of liquid cooling, the type of silicon, and use. The current high-end processors can be “effectively” cooled at high liquid coolant temperatures, allowing the facility water supply at the silicon to run as high as W40 and W45. But, as already shown at the beginning of this paper the silicon market is evolving very rapidly and specifically for GPU/ASIC is reaching very high power density, which will require lower temperatures for future silicon applications. It’s essential to choose the right fluid temperature by considering both current silicon specification and future power trends in the industry. This ensures that data center designs remain viable for at least 10 years, accommodating multiple generations of silicon technology.

A brief description of the cooling system is mandatory to understand W temperatures and what they refer to. Considering a simplified view of a cooling architecture in a data center, there are at least two cooling loops in a liquid-cooling system from the Technology Cooling System (TCS) to the Facility Water System (FWS) as shown in Figure 8.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

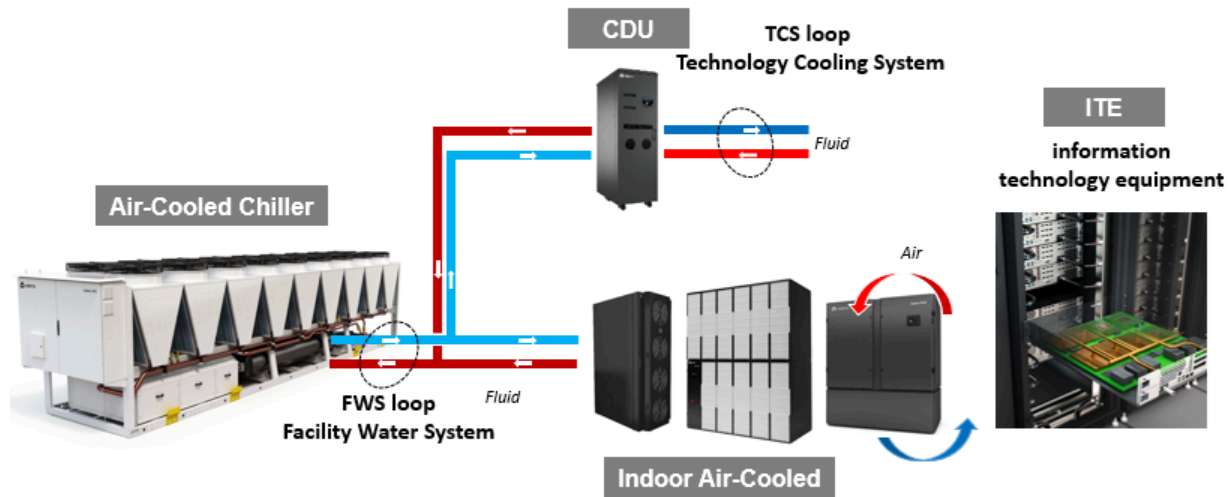


FIGURE 8. Cooling System Components

This schematic also shows the three main elements in a liquid cooling system: the heat capture system within the Information Technology Equipment (ITE), the Cooling Distribution Unit (CDU), and the heat-rejecting system that in this case is an air-cooled chiller. As reported from this schematic, the liquid cooling system might not remove the total heat load, and a percentage of heat is dissipated by air and removed using an indoor air-cooling unit.

The water temperatures W17, W27, W32, W40, W45, and W+ refer to FWS and they are exactly the inlet water temperatures of the CDU.

The CDU is a system that isolates the Technology Cooling System (TCS) fluid loop from FWS. It provides essential functions such as temperature control, flow control, pressure control, fluid treatment, and heat exchange to isolate the system.

The liquid heat capture system is an essential element of the liquid cooling architecture, it may use a dielectric in direct contact with the components, or a refrigerant or water pumped through cold plates attached to the heat-generating components. While the method of heat capture from ITE using a liquid is not covered in this document, it is important for understanding TCS. In this

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

setup, TCS temperatures are always higher than FWS because CDU uses a heat exchanger to separate two fluid loops.

Exploring the trade-offs between environmental impact and performance for a cooling system with varying fluid temperatures is important. It's also crucial to consider these temperatures in the context of the next generation data center' lifespan, rather than just the current silicon generation. Instead of beginning with the Facility Water System loop (FWS), we would start from the silicon side, focusing on the Technology Cooling System loop (TCS). ASHRAE can provide valuable guidelines for this approach. During the ASHRAE TC0909 2024 meetings, the maximum TCS supply temperatures were identified to aid in designing of the TCS loops for multiple generations of ITE [3]. Although these values are still under discussion, they provide industry direction. Among the various TCS temperatures mentioned, S20, S30 and S40 (20°C, 30°C and 40°C) were highlighted as the most suitable for future data centers considering current and future hardware power trends. With S20, S30 and S40 established as reference TCS temperatures, certain assumptions need to be made.

First, let's define the data center size: 80 MW of IT equipment has been considered for all the calculations, excluding any additional heat loss such as UPS or electrical distribution losses. The data center is divided into sixteen pods of 5 MW. Each pod contains 40 racks, and each rack is designed for 125 kW, with 80% cooled by liquid and 20% by air. A chilled water system has been chosen as the reference cooling system, composed of: external air-cooled free-cooling chillers, CDUs and indoor air-cooled units. The method used for capturing the heat (air/liquid) is not within the scope of this document. Redundancy is considered at the cooling unit level, with an N+1 configuration.

For the annual calculations a data center operating at 80% maximum IT load has been considered, as this represents a significant working condition. All data halls are assumed to be loaded at the same percentage.

These analyses have been conducted in dry mode, excluding any adiabatic and evaporative cooling methods to ensure the analysis is versatile and applicable worldwide. Although combining these systems can reduce overall energy consumption compared to using a single

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

cooling method, it's important to note that these methods are not suitable everywhere. In the data center market, mitigating water usage is crucial, especially in regions with scarcity.

The location of the data center significantly influences the performance of the cooling system since, as the efficiency of the chillers is affected by the climate. Chillers perform better in colder environments versus the hotter ones. To gain a global perspective on cooling system performance for different TCS temperatures, several climate zones from the ASHRAE 2021 climate design condition table have been considered. ASHRAE categorizes these into 9 zones based on average annual temperatures: extremely hot, very hot, hot, warm, mixed, cool, cold, very cold, subarctic/arctic [4]. A location has been selected for each zone, except for the subarctic/arctic zone since the different TCS temperatures have minimal impact on system efficiency in that region. For each chosen location, the performance of the air-cooled chiller has been evaluated based on the three different TCS temperatures selected.

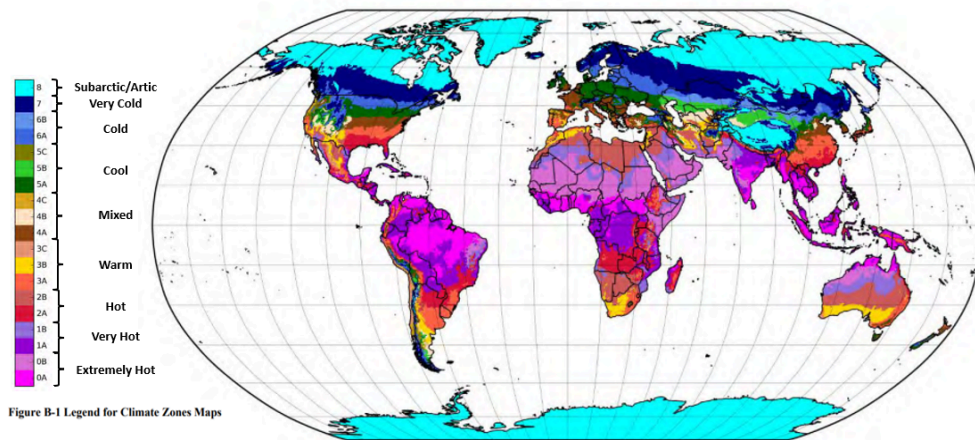


Figure 9. World climate zones [5]

Selecting a location site for each climate zone enables the proper sizing of free-air cooling chillers, as their capacity and efficiency are indeed affected by external temperatures. ASHRAE provides the maximum and the minimum extreme annual design temperatures. For each location an additional +4°C has been added to the maximum ambient temperature declared by ASHRAE over the last 20 years to define the size of the chiller, knowing that heat rejection always occurs at a higher temperature than the ambient one. The minimum temperature

Date: Oct 1, 2024



declared by ASHRAE guides the selection of the fluid type for the FWS loop (pure water or a different glycol mixture).

The analysis was performed using the meteorological data from the real hourly temperatures acquired in the relevant localities. To avoid anomalies, the frequency data represent the average values over ten years of acquisition to avoid a singular phenomena.

The chiller size is influenced not only by the maximum external temperature but also by the water temperatures selected. Lower chilled water temperatures require a larger chiller to handle the load, whereas higher chilled water temperatures may allow for a smaller chiller to manage the same heat load. The three different TCS temperatures selected 20°C, 30°C, and 40°C impact both the chiller size and dimensions.

For an indoor cooling unit, increasing TCS temperatures from 20°C to 30°C or 40°C determines leads to higher supply temperatures for both the CDU and the air-cooled units. On the liquid side, the CSU's design remains unaffected. However, on the air-cooling side, this change can impact the cooling units and even the data center design and operation, especially when the return air temperature reaches very high levels.

Since, the TCS and FWS temperatures are correlated, defining one automatically determines the other given the size and model of CDU. A large CDU unit usually has an approach temperature of around 4-8°C between the cold fluid temperatures of the FWS and TCS loops. In this analysis, an average approach temperature of 6°C has been selected for the CDU. Consequently, this defines supply temperatures of the chiller for the FWS loop as: 14°C, 24°C and 34°C.

A common ΔT temperature of the TCS cooling system loop for cooling a GPU is typically around 10°C and usually uses 25% propylene glycol as fluid. The same ΔT value can be assumed even for the FWS cooling system loop since 10°C is a common value for today's chilled water system. For the indoor air-cooled unit, a temperature approach of 6°C has been selected, which is the difference between the inlet water temperature and the supply air temperature. However, the ΔT on the air side between the return and the supply air has been considered around 12°C.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Figure 10 summarizes the main parameters and assumptions listed and the result in terms of temperature working conditions for the different TCS loops S20, S30 and S40 which are set at 20°C, 30°C and 40°C respectively.

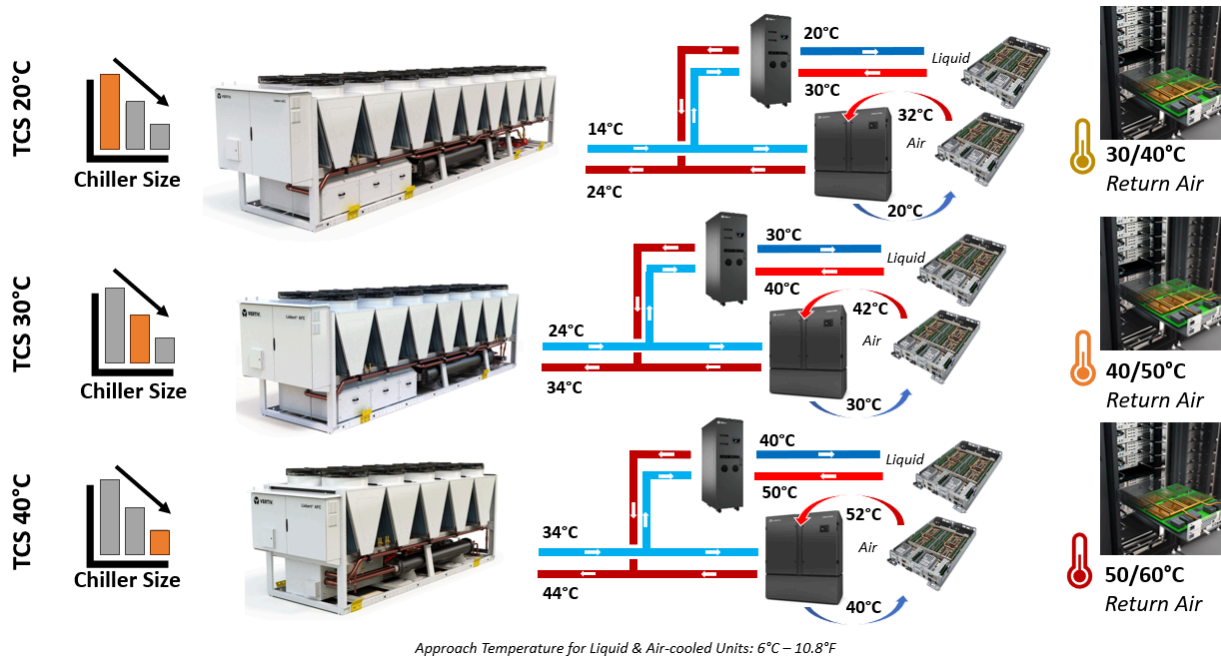


Figure 10. Summary of parameters and assumptions for the analysis

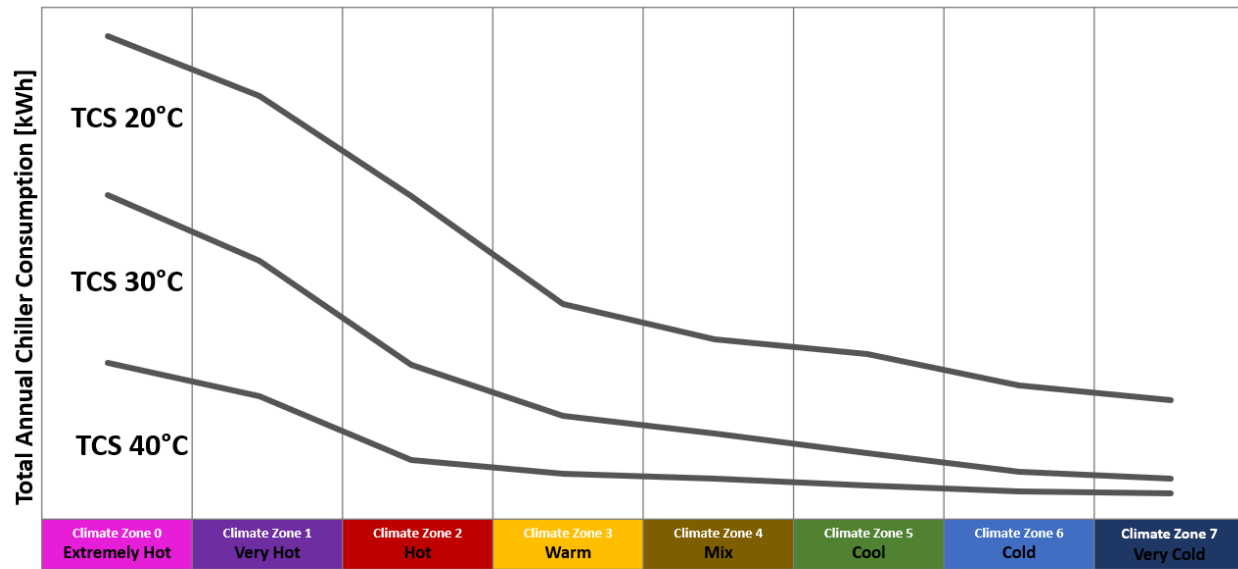
The TCS loop temperature of 20°C and 30°C do not significantly impact current data center design and operation, but the working temperatures for the TCS loop at 40°C present several challenges for standard cooling products. These conditions limit standard chiller compressor selection, as they fall outside the typical operating range. Only chillers with a specific design can reliably function at these temperatures. There are chillers on the market designed for these higher temperatures, offering a versatile solution for data centers by operating at a wider working range compared to standard products. Standard indoor units have several challenges with return air temperatures over 50°C. While there are specific air-cooling units that are capable of handling these temperatures, they come with additional considerations such as the IT rooms need to be specifically designed for high temperatures and which can shorten the lifespan of many electrical components, leading to more frequent maintenance. Furthermore, very high temperatures increase safety concerns complicating operational challenges.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Finally, Figure 11 below shows the results of the analysis for the different TCS temperatures and the different climate zones:



Data Center in configuration N+1 for the cooling units and loaded at 80% of the maximum IT load (Temperature approach of 6°C)

Figure 11. Annual chiller power consumption by climate zone

The first chart shows how different climate zones (x-axis) impact total annual consumption (y-axis) for a free-air cooled chiller. It highlights the beneficial effect of increasing the TCS temperatures to reduce data center power consumption. The chart shows the big jump in efficiency when moving from an S20 to an S30 of the TCS temperatures. However, the same level of improvement is not observed when moving from S30 to S40 except in Extremely Hot, Very Hot and Hot climate zones. In these hotter zones, further increasing of the fluid temperature allows a reduction in chiller consumption by a similar amount. In colder climate zones, the positive effect is diminished. Increasing TCS temperatures (and consequently the working chiller temperatures) helps to flatten the TCS curves, resulting in nearly the same power consumption across different regions. For instance a data center operating at TCS 20°C in climate zone 6 “COLD” and has the same power consumption as a data center operating at TCS 30°C in climate zone 2 “HOT”.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

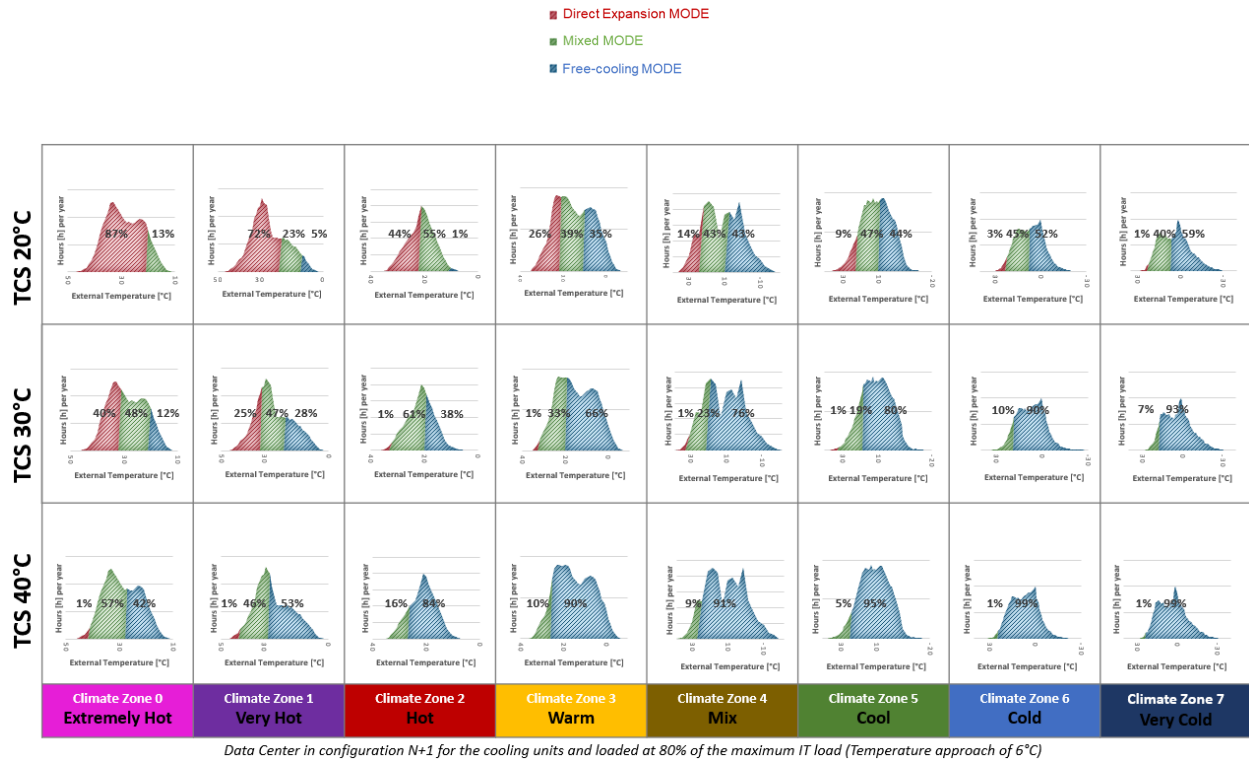


Figure 12. Time in Data Center cooling modes by TCS temperature and climate zone

Figure 12 provides a detailed analysis of the information presented in Figure 11. The reduction in total annual consumption for a free-air cooled chiller in different climate zones is driven by the capability of limiting the compressor utilization. As the climate shifts from hotter to colder, there is a significant increase in the operation of the air-cooled free-cooling chiller in MIX and FC cooling modes. Chart 1 also highlights the beneficial effect of increasing the TCS temperatures, which minimizes the use of compressors and significantly reduces cooling system operational costs. The chart also shows a notable increase in the percentage time the system operates in these efficient modes when moving from an S20 to an S30 of the TCS temperatures. The same positive effect is not achieved when increasing TCS temperatures from S30 to S40, except in Extremely Hot, Very Hot and Hot climate zones. In these zones, further increasing of the fluid temperature reduces chiller consumption by a similar amount. However, in colder climate zones, the positive effect diminishes. Increasing the TCS temperatures (and consequently the working chiller temperatures) helps to minimize compressor use (DX mode), which is then needed to cover only peak demands during the hottest periods of the year. This results in nearly the same

Date: Oct 1, 2024



working mode percentages across the different regions, especially when the TCS temperature is set higher. It's important to remember this result was achieved by simulating the data center in an N+1 configuration for the cooling units and operating at 80% of the maximum IT load.

Increasing the TCS loop temperatures enhances the efficiency of data center cooling systems across all climate zones. The analysis shows a significant efficiency boost when moving from a TCS of 20°C to 30°C. However, the improvement from 30°C to 40°C is only substantial in hot climates.

Reducing annual consumption for free-air cooled chillers is achieved by minimizing compressor use and maximizing free cooling mode. These results are based on a data center configured in N+1 with an 80% IT load. A TCS loop at 40°C is the most efficient for all climates, especially in hotter zones, compared to 30°C. However, 30°C provides the best overall efficiency improvement, even in colder climates, compared to 20°C. High TCS temperatures allow the cooling system to work in free cooling mode nearly 100% of the time in many climate zones, this does not guarantee that a chiller-less solution is practical in every situation. It is crucial to ensure that the system can sustain the required cooling capacity even under full load and during failure scenarios. Furthermore, High TCS temperatures present challenges for standard cooling products, necessitating careful evaluation and potentially dedicated products.

Designing a durable cooling system ensures efficiency, proper temperature control, lower energy consumption, and operational costs. It also reduces the likelihood of leaks, malfunctions, and failures, providing data center reliability. It is crucial to select the right TCS temperature, considering current and future IT equipment trends. A TCS at 30°C offers a balanced solution for global efficiency.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

6. How do we Maintain Durability in the Future?

With the rapid advancement of AI silicon, maintaining a 30°C coolant temperature requires continued technological advancement across the entire technology stack from the die to the data center infrastructure. The following section looks at some of the opportunities for innovation that are possible

6.1 High Bandwidth Memory Silicon Improvements

Semiconductor devices typically operate within specified maximum temperatures to ensure reliable performance such as 125°C for logic silicon and 95°C for memory silicon. In configurations like 2.5D package platforms where devices with varying thermal thresholds coexist, silicon with lower thermal limits may overheat due to heat transfer from hotter neighboring silicon. Moreover, logic silicon, which consumes more power during peak operations, often causing significant thermal crosstalk to adjacent memory silicon, quantitatively illustrated through power envelope graphs.

In High Bandwidth Memory (HBM) 3D silicon stacks (as shown in Figure 13), the hottest die is positioned closest to the substrate due to factors like I/O count and signal integrity and is therefore the most difficult to cool. Each HBM die in the stack dissipates heat and acts as a thermal resistance which affects the thermal profiles of neighboring dies. For effective thermal management, placing the hottest die near the cooling system or establishing a direct thermal pathway to it is crucial in maintaining a stable coolant temperature into the future.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

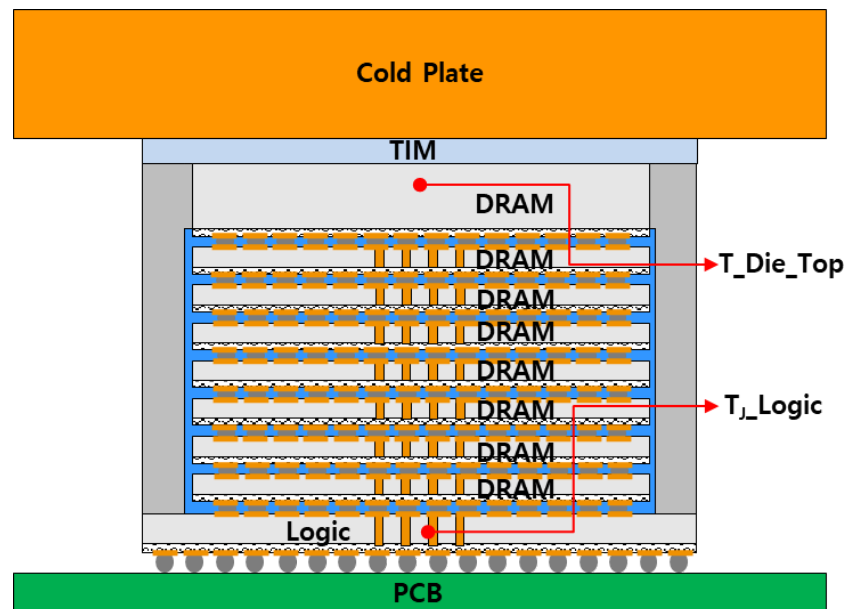


Figure 13. Example of a 3D stack of eight HBM dies

Silicon bonding substantially increases vertical thermal resistance in HBM 3D silicon stacks. The joint layer in these 3D stacks are comprised of microbumps and adhesive material which have thermal conductivities of approximately $60 \text{ W}\cdot\text{m}^{-1}\text{K}^{-1}$ and $1 \text{ W}\cdot\text{m}^{-1}\text{K}^{-1}$, respectively. This is much lower than silicon's $120 \text{ W}\cdot\text{m}^{-1}\text{K}^{-1}$, which increases the stack's vertical thermal resistance. Additionally, the silicon's back end of line (BEOL) layer, which has a lower thermal conductivity compared to silicon, further amplifies the vertical thermal resistance. Furthermore, interface thermal contact resistance at junctions exacerbates these effects, leading to a rapid increase in thermal resistance with the addition of interconnection layers.

6.2 Improving IT Hardware-level Cooling Solutions

Performance of in-rack cooling solutions play a vital role in supporting longer-term, durable coolant supply temperatures (technical loop at the facility). This is predominantly due to the electronic package and the cooling solution having comparable thermal resistances (or thermal

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

budgets). With the latter, there are options for extending capabilities with both existing and future designs. Configuring the IT loop to have cold plates in parallel to eliminate thermal shadowing is one such way, but may require the fin structure and coolant circuit to be optimized to deliver performance at lower flow rates. It's generally recommended to design the IT cooling solution for long-term operation at pressures of 50 ~ 75 psig. End users must work with solution providers to ensure performance is not impacted by such requirements while pushing (minimizing) thermal parameters such as base thickness, TIM bond-line, etc. Relatively newer technologies such as flow impingement, two-phase cold plates, and higher conductivity fluids, should be studied from the perspective of end-to-end performance, and serviceability to determine viable options that could be scaled in the future.

6.3 Improved Data Center Heat Exchangers

A way to maintain the durability of the coolant temperature in the future is to improve the heat exchanger performance of CDUs and indoor air-cooling units by lowering the approach temperature. Lowering the approach temperature leads to better thermal management by allowing higher temperatures in the FWS loop without affecting the TCS coolant temperature or the overall cooling capacity of the data center. Higher FWS temperatures also enable air-cooled chillers to operate more hours in free-cooling mode for longer periods, boosting overall data center efficiency. This minimizes compressor use, limiting it to peak periods during the hottest times of the year. This approach combines the high efficiency of free-cooling with continuous cooling availability under any condition ensuring maximum reliability of the cooling system.

Improving the coil and heat exchanger performances can be achieved through various methods depending on the specific type and operational conditions. The most common strategy is to increase the surface area available for heat transfer, which can significantly improve performance. When designing the heat exchange, a balance between cost and efficiency is sought as each additional increase in heat exchanger efficiency becomes progressively more expensive.

For indoor cooling units, improving the temperature approach involves increasing the number of coil rows for indoor air-cooling units and expanding the plate heat exchanger area for the CDU.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Section 5.3 considered a temperature approach of 6°C for the indoor cooling units. In this chapter we will show the benefits of improving the temperature approach of 2°C by moving from 6°C to 4°C. This adjustment immediately translates to a supply temperature of the FWS loop increasing from 14°C, 24°C and 34°C to 16°C, 26°C and 36°C respectively. Figure 14 illustrates how the cooling system efficiency changes based on the climate zone considered.

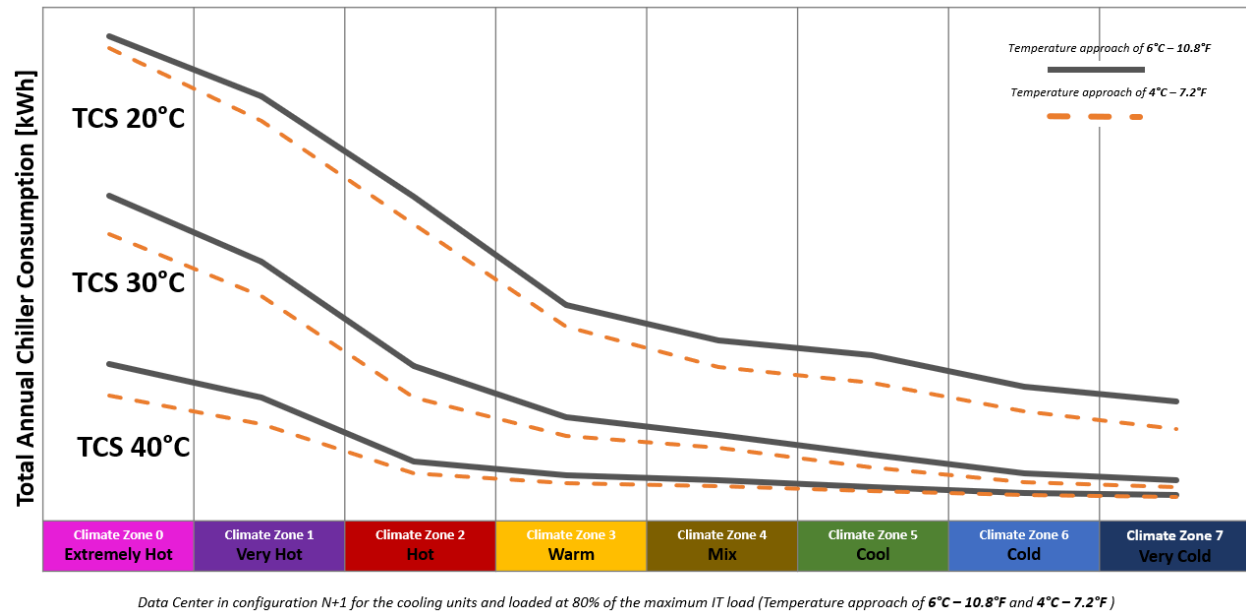


Figure 14. Plot of chiller power based on different climate zones

The chart shows how improving the temperature approach can reduce the total annual consumption for a free-air cooled chiller. A TCS loop at 20°C benefits from this change in all the different climate zones, unlike the TCS 30°C and TCS 40°C. Notably, the hottest climates show a greater reduction in power consumption compared to the coldest climates. TCS loops at 30°C and TCS 40°C perform similarly, achieving the most reduction in power consumption in hotter climates, while the reduction nearly disappears in the colder climates.

These trends can be explained by the location profiles chosen for each climate. At a TCS loop temperature of 20°C, the improvement in heat exchanger performance does not translate in the hotter regions because the free-air cooling chiller can not switch to a more efficient mode. The behavior varies in different zones for the same TCS temperature line. Conversely, TCS loops at

Date: Oct 1, 2024



30°C and 40°C, are able to substantially improve the efficiency of the free-air cooled chillers in the hotter climates due to their higher operating temperatures. However, they have little impact on colder climates, as these regions are already achieving maximum efficiency. Figure 15 summarizes this below.

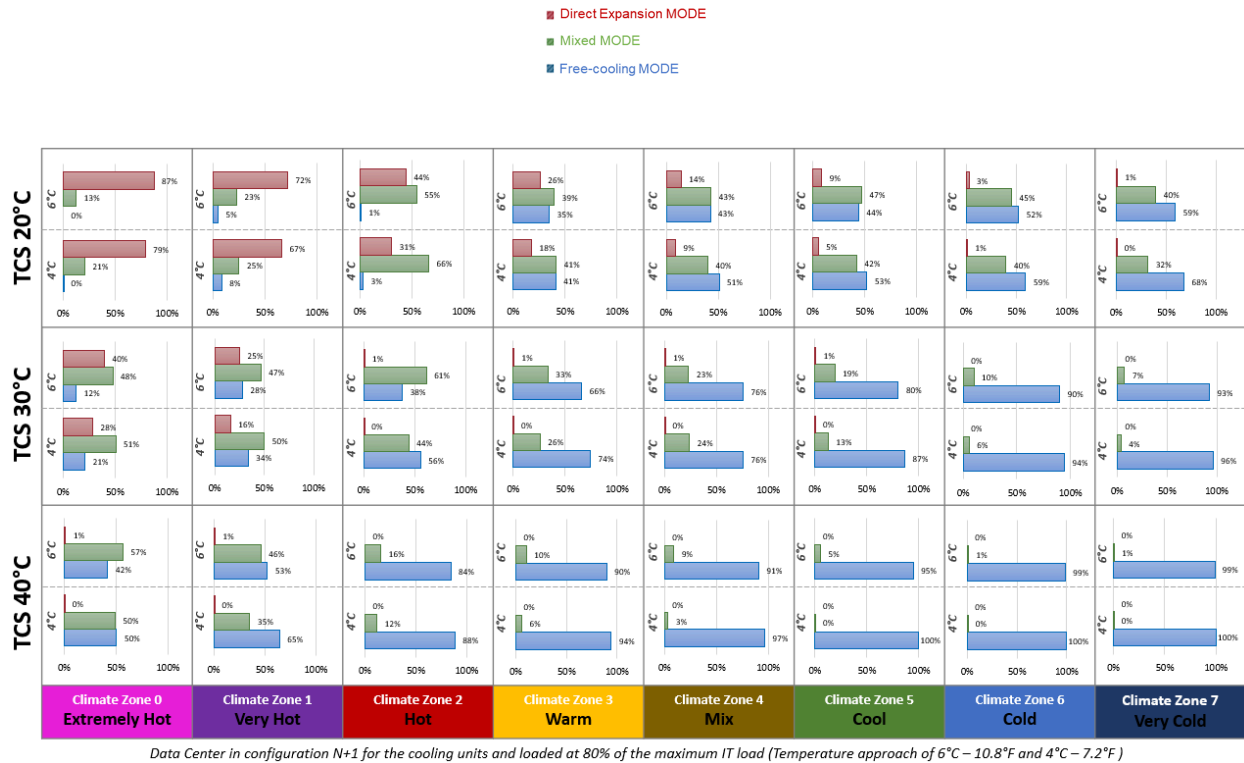


Figure 15. Plot of energy spent by data centers in different cooling modes

A better temperature approach leads to better thermal management, allowing higher temperatures in the FWS loop without impacting the TCS system and the overall cooling capacity. However, this benefit varies across different climates and is influenced by the TCS temperatures selected and the climate zone. The balance between cost and efficiency must be evaluated based on these parameters, keeping in mind that each additional increase in heat exchanger efficiency raises the cooling unit’s expense.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

6.4 Silicon packaging innovations

The figure below illustrates the thermal resistance contribution from the cold plate, TIM, internal package, and heat crosstalk compared between ASIC to HBM. Figure 16 assumes a typical 2.5D advanced package designed AI silicon without lid and with a single-phase cold plate. The data clearly shows where the biggest temperature difference (ΔT) is and therefore where the largest opportunity to improve.

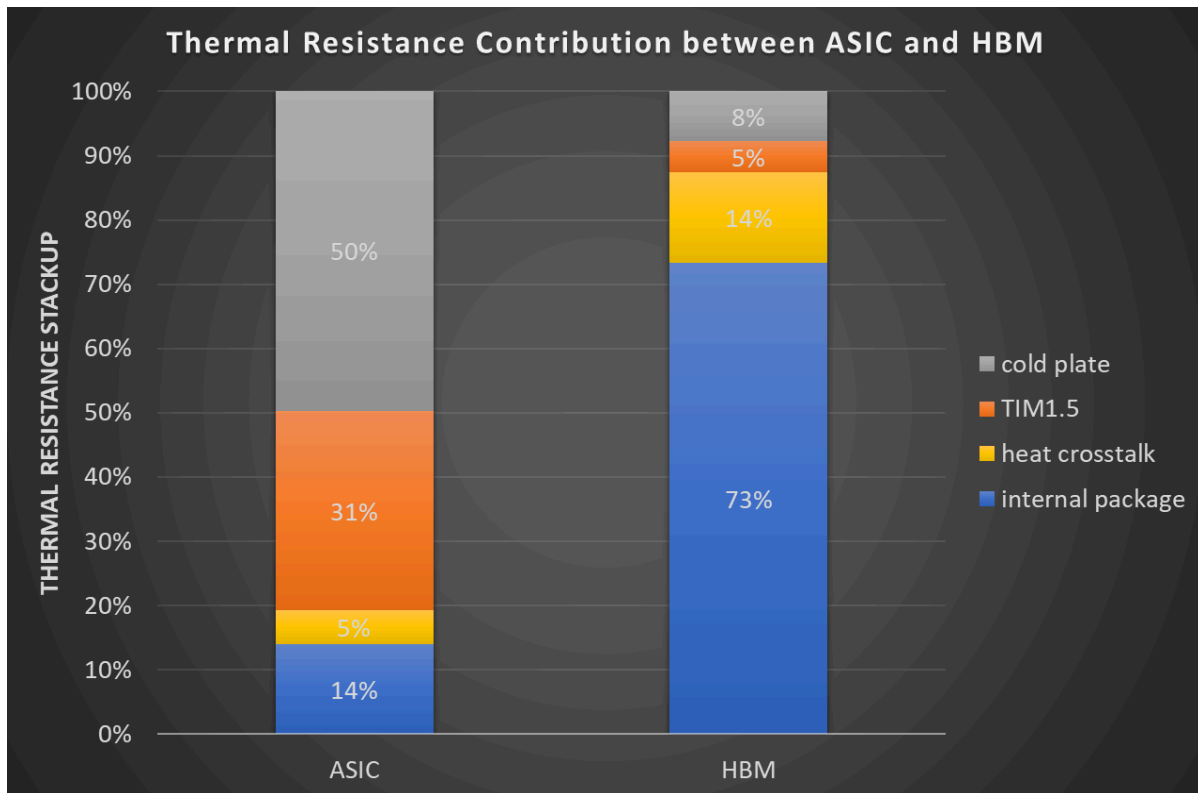


Figure 16. Plot of Thermal Resistance Contribution.

Note that the % is subject to change slightly if the package design is different, however the trend and conclusion should remain the same from package to package.

According to Figure 16, we conclude two things:

1. for ASICs, cold plate and TIM1.5 provides the big opportunities for improvement
2. for HBM, it is improving the packaging

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Based upon those conclusions, this section provides a brief introduction of the cooling technologies targeted for investment and the associated performance, availability, risks and cost assessments.

We summarized the performance gain from some of the future cooling technologies in the plot below by comparing it to the baseline cooling technology that we are planning to use in the 2025 time frame. The performance gain is represented by the percentage increase in total package power that can be cooled. Note that some technologies can significantly improve the cooling capacity of ASICs, but will not significantly improve HBM cooling capacity.

Given the potential thermal improvement with the extended cooling limit as shown in Figure 17 below, the performance gain is significant and should not be overlooked. We need other solutions to solve the HBM thermal challenges.

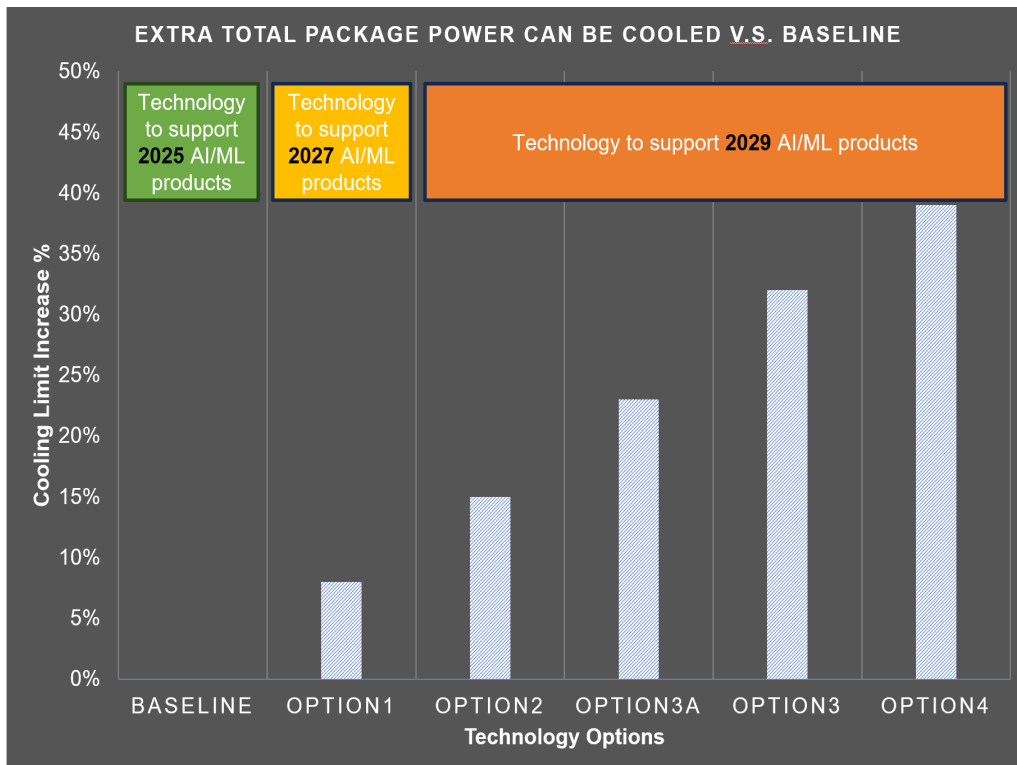


Figure 17. Performance gains compared to 2025 baseline

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Figure 18 below summarizes our assessment on the different technologies from the aspects of thermal performance gain, technology readiness, implementation risk and cost impact.

	Baseline	Option1	Option2	Option 3A	Option3	Option4
Description	Lidless package	Optimized single phase cold plate	Two phase cold plate	Single phase cold plate lid + solder TIM1/alternative TIM	Optimized single phase cold plate + solder TIM1.5/alternative TIM	Two phase cold plate solution + solder TIM1.5/alternative TIM
Thermal Performance Gain	Low	Medium	Medium*	High	High	High*
Technology Readiness	MP solution available	POC solution available	Early exploration	Early exploration	Early exploration	Not available
Implementation Risk Level	Low	Low	Medium to High	Medium to High	High	High
Cost	Low	Low	High	High	High	High

* performance and implementation of two phase cold plate is highly dependent on facility liquid cooling temperatures available.

Figure 18. Summary of different cooling technologies

With the project performance gain from 24% to 39%, Option 3, 3A and 4 will be able to solve the thermal challenges in 2029 generation for both ASIC and HBMs assuming there is no improvement of HBM internal thermal resistance and power. However, if the power and/or memory capacity demands increase in the near- or long-term (beyond 2029), the HBM package resistance needs to be addressed and it requires collaboration with the HBM vendors directly.

Option 2 through Option 4 are expected to take longer time to be mature and ready for production, and are aimed to support 2029 beyond AI/ML production generation. One thing that would like to point out is that the coolant inlet temperature (technical water temperature) of Option 2 depends on the primary loop design, the Data Center cost to keep the same coolant inlet temperature on Option 2 compared to the rest of the Options might be higher and might require retrofit of the data center in order to adopt this design. Reducing the temperature difference on the primary side impacts the plant performance and it would be an incremental change. The advantage of Option2 might come with different types of primary cooling system.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

The plot above still assumes the same coolant temperature in Option 2 compared to the rest Options.

Optimization of the single phase cold plate performance (Option 1) by reducing the cold plate base thickness and fin gap/fin thickness, etc. has a high confidence of success.

The improvements will have risks which will be evaluated. The risks include: manufacturing yield risks resulting in higher cost and reliability concerns (fouling risk and resultant system impedance increase).

There is another technology that is not included in the table above due to the uncertainty and manufacturing risks, i.e. **Embedded cooling** at the silicon level (silicon microchannel). This technology has been investigated by academic and industrial research labs for decades. The technology will yield maximum thermal performance. The benefit of implementing this technology when the power density continues to increase above $500\text{W}/\text{cm}^2$ or to address stacking of very high-power silicon. Productizing the technology will require substantial effort as it impacts the silicon design process among many other issues. We do not consider this technology necessary for the 2029 generation. It is not a current focus. We will consistently reassess this requirement by closely monitoring the industry's technology readiness.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

5. Conclusion

The technologies needed for AI silicon will continue to evolve very rapidly, however the large scale infrastructure needed to support these technologies requires many years to design and build. So, aligning on infrastructure capabilities such as 30°C coolant temperature and the investments needed to enable future technology innovation is critical to building a large, viable ecosystem in the future for AI/ML solutions.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

6. Glossary

AI/ML: Artificial Intelligence / Machine Learning is a specialized segment of computing where ultra performance, large memory footprints, and high speed bandwidth is required to analysis artificial intelligence models, and adaptive learning algorithms

ASHRAE: American Society of Heating, Refrigeration, and Air-conditioning Engineers is a professional organization that seeks to create and improve standards, designs, and best practices for any conditioned spaces

ASIC: Application-specific Integrated Circuit is a processor that is customized to complete a specific task

BGA: Ball Grid Array is a series of solder bumps used to attach a package or chip to a printed circuit board

CDU: Coolant Distribution Unit is a device that control the movement of coolant through electronic devices

CoWoS: Chip-on-Wafer-on-Substrate is a 3D stacking technology used to assemble different chiplets and devices in to a single electronic package

FWS: Facility water system is the fluidic system of a mission critical space that interacts with the outside cooling equipment

GPU: Graphics Processing Unit is a specialized electronic circuit that can perform mathematical calculations quickly to process graphical and video data

HBM: High Bandwidth Memory is a memory device that provide direct access to silicon processors and can be co-packaged with the processor

HPC: High Performance Computing is a specialized segment of computing where ultra performance is require, typically solving scientific problems

PCB: Printed Circuit Board is a board that contains multiple components to for a system or server

SOC: System-on-Chip is an integrated circuit that contains the components such as a CPU, memory, and input/output devices to form a complete, functioning system

TIM: Thermal Interface Material is a thermally conductive material that occupies the gap between the item being cooled and the cooling device

TCS: Technical Cooling System is the fluidic system that interacts with the IT equipment directly

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

7. References

1. [Open Compute Project Panel - Coolant Temperatures for Durable Data Center Designs](#)
2. [Emergence and Expansion of Liquid Cooling in Mainstream Data Centers](#)
3. [Documents | ASHRAE 9.9 Mission Critical Facilities, Data Centers, Technology Spaces and Electronic Equipment](#)
4. [ashrae meteo](#)
5. [Climatic Data for Building Design Standards](#)

8. License

8.1. Creative Commons

OCP encourages participants to share their proposals, specifications and designs with the community. This is to promote openness and encourage continuous and open feedback. It is important to remember that by providing feedback for any such documents, whether in written or verbal form, that the contributor or the contributor's organization grants OCP and its members irrevocable right to use this feedback for any purpose without any further obligation.

It is acknowledged that any such documentation and any ancillary materials that are provided to OCP in connection with this document, including without limitation any white papers, articles, photographs, studies, diagrams, contact information (together, "Materials") are made available under the Creative Commons Attribution-ShareAlike 4.0 International License found here: <https://creativecommons.org/licenses/by-sa/4.0/>, or any later version, and without limiting the foregoing, OCP may make the Materials available under such terms.

As a contributor to this document, all members represent that they have the authority to grant the rights and licenses herein. They further represent and warrant that the Materials do not and will not violate the copyrights or misappropriate the trade secret rights of any third party, including without limitation rights in intellectual property.

The contributor(s) also represent that, to the extent the Materials include materials protected by copyright or trade secret rights that are owned or created by any third-party,

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

they have obtained permission for its use consistent with the foregoing. They will provide OCP evidence of such permission upon OCP's request. This document and any "Materials" are published on the respective project's wiki page and are open to the public in accordance with OCP's Bylaws and IP Policy. This can be found at <http://www.opencompute.org/participate/legal-documents/>. If you have any questions please contact OCP.

9. About Open Compute Foundation

The Open Compute Project (OCP) is a collaborative Community of hyperscale data center operators, telecom, colocation providers and enterprise IT users, working with the product and solution vendor ecosystem to develop open innovations deployable from the cloud to the edge. The OCP Foundation is responsible for fostering and serving the OCP Community to meet the market and shape the future, taking hyperscale-led innovations to everyone. Meeting the market is accomplished through addressing challenging market obstacles with open specifications, designs and emerging market programs that showcase OCP-recognized IT equipment and data center facility best practices. Shaping the future includes investing in strategic initiatives and programs that prepare the IT ecosystem for major technology changes, such as AI & ML, optics, advanced cooling techniques, composable memory and silicon. OCP Community-developed open innovations strive to benefit all, optimized through the lens of impact, efficiency, scale and sustainability. Learn more at www.opencompute.org.

Date: Oct 1, 2024



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)