In the past, memory was not as complex and did not warrant the same kind of thermal attention that designers paid the CPU. Because the CPU required cooling, chipsets came equipped with heat sinks as a production standard. By comparison, the memory module needed only minor airflow adjustments to keep temperatures in check. But with DDR3 and DDR4 technology increasing speeds in today’s embedded designs, memory module design is complex and requires thermal attention as well.

Clock speed is just one reason why memory is running hotter than ever. The customer environment, overall design choices such as memory modules’ location on the board, horizontal or vertical module orientation, and the amount of airflow over the system also can influence memory modules’ thermal condition.

Embedded system designers typically work with tight board layouts requiring near-perfect engineering to achieve flawless signal integrity and extreme performance. Though other design issues exist, successful system designers consider memory thermal management as a higher-level design issue, keeping abreast of evolving memory technologies and thermal management techniques for reducing heat in memory modules.

Memory designers can mitigate heat and design better memory subsystems using a range of simple but powerful thermal concepts. Similarly, system designers can enhance products by incorporating those concepts when creating their designs.

### The heat is on

Memory designers begin the process by choosing memory modules that mitigate heat and deliver the best overall heat reduction scheme. Incorporating modules that use the least amount of DRAM within the greatest number of module ranks can achieve the desired module density and manage power effectively. The more DRAM in standby mode, the less power the module consumes—frequently achieved by using DRAM with the widest data bus, as shown in Table 1. For example, a 36-chip four-rank x8 DIMM uses less power than a 36-chip two-rank x4 DIMM.

For an additional example, a 512 MB error-correcting code DIMM can be made using five 64x16 DRAM chips versus nine 64x8 DRAM, resulting in 44 percent heat reduction. The actual reduction may be slightly less due to differences in IDD values specified in the datasheets for 64x16 and 64x8 DRAM. Memory designers would typically explore whether or not the memory controller chipset can support the wider DRAM data bus width.

Overall, memory modules appropriately spaced between DRAM either nonstacked or without large, hot semiconductors will have better thermal characteristics. Small form factor memory such as stacked very low profile or stacked SODIMMs have higher power density (watts/area) and...
need special considerations for cooling. Fully buffered DIMMs also have high power densities because of onboard advanced memory buffer and may require additional cooling aids or airflow.

**System versus memory**

Thermal sensors are critical tools for memory designers. JEDEC’s standard specifies that memory modules have thermal sensors to give users monitoring and triggering mechanisms that adjust system performance according to fluctuations in temperature.

Depending on defined parameters, the system can issue an extended mode register set command, which would double the internal refresh rate on the DDR2 DRAM to a 32 millisecond period (tREFI = 3.9 microseconds) at a trigger temperature of +85 °C to extend the DRAM operating temperature to +95 °C. If that feature is not available, designers can incorporate special programming on the memory module for extended temperature operation. Alternatively, the system can use closed loop dynamic temperature throttling and fan speed control to optimize memory performance.

The key point here is that the CPU manages the memory board’s thermal sensors, demonstrating that system-level and board-level thermal issues are closely related. The system’s BIOS reads output from the sensor and evaluates performance options based on preprogrammed thresholds identifying acceptable temperature ranges. For example, if the memory runs over the limited temperature, the system thermal monitor alerts administrators about temperatures above defined thresholds, prompting them to take the necessary steps to lower the temperature, such as checking processors and chassis fans, addressing any chassis airflow vents that may be blocked, or adding another chassis fan.

**Airflow matters**

Airflow is a simple but critical issue for memory; the primary goal is to avoid blowing preheated air directly over the memory subsystem. Whenever possible, designers should place the memory subsystem on the sides of the processor and outside the flow of warm air generated by the processor’s heat sink or other hot components such as the power supply or chipset. The ambient intake air should flow evenly over the memory subsystem and other hot components such as the processor.

Too small an air gap between modules may create airflow back pressure from the physically obstructed DIMM modules within the airflow path. This could result in an airflow pressure drop along the side of the DIMMs, generating a decrease in airflow, or could divert the airflow away or around the entire memory subsystem. DIMM socket spacing should be 10 mm or greater from center to center.

In general, maximizing airflow extracts heat away from the memory. Designers should use a blower or dual fans to optimize airflow if acoustical noise is not an issue. Airflow with a minimum pressure drop is best achieved by extracting the hot air at the exhaust point but also can be improved by pushing air in at the intake point. Plenums, ducts, or shrouds can be used to direct and contain the airflow through the memory subsystem, flowing parallel to the longest sides of the DIMMs and on both sides. These enclosures may allow for slower fan speeds with less acoustical noise and not affect airflow.

Memory modules can be designed to allow airflow across the short side of the DIMM, eliminating heat from being dragged across the long side of the DIMM. This type of mezzanine connector technique does not expose as much DRAM to preheated air from upstream DRAM.

If the motherboard or system board is mounted flat and perpendicular to the line of gravity, the best orientation for the memory would be a vertical mount since hot air rises up along the line of gravity. A vertical DIMM orientation prevents heat from being trapped under the lower bottom side of the memory modules. If a vertical mount is not possible, then an angled-mount DIMM orientation would benefit from one-sided DIMMs with the DRAM components mounted on the top side. This would hold true for memory DIMMs placed flat over the system board as well.

Designers should choose a module with a DRAM placement that does not allow all the DRAM devices to be active on the same side at the same time. Modules with alternating DRAM placements on each side of the memory module per rank will evenly disperse the heat surrounding the DIMM. If airflow is restricted on one side of a DIMM, memory modules with DRAM placed only on the side with the maximum airflow will perform better at higher temperatures. Figure 1 illustrates how the technique of alternating DRAM ranks can reduce thermal impact.

![Figure 1](image-url)
Heat spreaders and more

Heat spreaders are metal covers placed on the surface of a memory module to disperse heat evenly across the surface and equalize surface temperature by removing localized hot spots. A heat spreader is made of a thermoconductive material such as copper or aluminum in the shape of a clamshell wrapped around the memory module.

If space permits, a heat sink placed on the side surfaces of the memory and/or top edge of the memory module will maximize heat extraction from the module. The additional surface area the heat sink adds to the memory module without affecting airflow determines its overall effectiveness.

Heat-conducting PCBs and PCB cores are also effective options. These metal or carbon composite laminated layers are embedded into the structure of the memory PCB to allow it to operate cooler than standard FR-4. The layers also equalize component temperatures by removing localized hot spots like the phase-locked loop. It is not uncommon to see numerous hot spots created via holes under hot devices to conduct heat into the core. These cores in turn conduct heat into the edge fingers of the module and can be brought to the top edge of the PCB to expose it to heat spreaders or heat sinks. The top edge of this type of PCB has the inner thermal core of the DIMM connected to an integrated heat sink at the top of the module, adding additional height to the DIMM.

During the manufacturing process, memory modules can be tested at elevated temperatures in customers’ systems running customers’ diagnostic software. This active burn-in will screen out potentially weak modules. Passive burn-in (on unpowered modules) has no effect on screening out DRAM with weak cells because DRAM cells are semiconductor-based capacitors that need to be continually recharged or refreshed to retain binary information. Some memory modules are available using DRAM screened for an extended operating temperature range of $-40 \degree C \leq T_{case} \leq +95 \degree C$. This is a specialty item, and not all DRAM suppliers offer industrial temperature DRAM as an option for commercial temperature ($0 \degree C \leq T_{case} \leq +85 \degree C$).

Thermal matters across the board

Thermal management issues are evolving with memory technology and becoming critical to embedded systems’ reliability and performance. The design dynamic between system designers and memory subsystem designers is also evolving and can impact designs built for endurance and performance. Trusted system-level and board-level partnerships plus greater understanding of current thermal concepts associated with DRAM memory modules can make all the difference in the final product’s success.

Understanding DRAM memory module thermal considerations as part of proven system design tenets can equip designers with a new level of understanding regarding ways to improve thermal performance. General design considerations and alternative thermal options can create a winning memory subsystem design, effectively meeting system requirements for the high memory bandwidth, large memory densities, small physical space, and low cost imperative in embedded environments.