

ManagedAir Cooling for PowerStream 7000 RapidIO-Based Multicomputers

Randall Banton and Richard Jaenicke
Mercury Computer Systems, Inc.



MERCURY

WHITE

PAPER



Computer Systems, Inc.
MERCURY

The PowerStream® 7000 RapidIO-based multicomputer system was developed to provide a TeraFLOPS or more of computational performance in a single chassis suited for deployment in military aircraft, ships, and other mobile platforms. The size constraints for processing systems in these platforms require a minimum computational density of 150 MFLOPS/cu. ft.

To meet this processing density goal, the PowerStream 7000 system was designed to support processors drawing up to 20W each, deployed in close proximity, creating a dense heat environment. The fully loaded unit is designed to dissipate about 4 kW.

Managing the heat in this compact system is a significant challenge that required innovative approaches to air cooling. Mercury addressed this challenge by developing a patented* technique known as ManagedAir™ cooling. In this method, airflow is strategically directed throughout the system and over boards and components to maximize its effectiveness. Other considerations, including board size, shape, and layout, are also managed to satisfy the cooling requirements of components with high thermal flux.

Multiple Heat Sources

Processors are not the only components in the system that produce heat as a byproduct. Faster processors require faster memory, faster interconnect, and more I/O. SRAM, onboard power supply FETs, integrated AC/DC converters, and various ASICs also generate heat in significant amounts. Whereas processors alone accounted for 70% or more of the power consumption in previous systems, today processors typically account for only 40 to 50%. Many of the other components are also very small, but some of these components, such as power supply FETs and host bridges, can pose heat challenges of a magnitude similar to the processors.

The ability to dissipate the heat is hindered by the shrinking area of the silicon die. For example, even if a processor consumes less power after a die shrink, the energy concentrated in a smaller area may be harder to cool. As a result, these smaller, hotter chips may actually be operating at or near their maximum junction temperatures. The primary measurement of the problem is power flux, which quantifies the amount of energy dissipated over the surface area of the die. As an example, Figure 1 below illustrates the power flux of several recent and future PowerPC microprocessors.

The advanced switch fabric needed to support the PowerStream 7000's data communications requirement also adds heat to the system as a whole. These components add still more complexity to the thermal management challenge.

System-Level Perspectives

The trend of escalating power consumption of both processors and infrastructure chips presents an increasing challenge for deploying high-performance processor systems, especially in military environments. From a system perspective, the challenges of more components, smaller packaging, and greater power requirements demand innovative cooling solutions. However, most programs would prefer to retain forced air as the cooling method, due to its lower cost and greater simplicity. The challenge thus becomes making the air work more efficiently.

Meeting this challenge requires a system-level approach to manage the airflow through the chassis and across the components. The PowerStream 7000 system utilizes ManagedAir cooling, a technique that includes elements at the component, board, card cage, and chassis levels. Each element, and the system as a whole, is designed to direct all the available air across the hottest components instead of around them, thereby extracting the maximum thermal efficiency.

One of the key elements of the ManagedAir technique is a cover that encloses each printed circuit board assembly and manages the airflow for that board. Each board's cover is tailored to the profile, placement, and airflow requirements of its components. Such a technique is most effective when mated with the corresponding infrastructure in the card cage and chassis to shape the airflow from the system's air intake to the system's air exhaust.

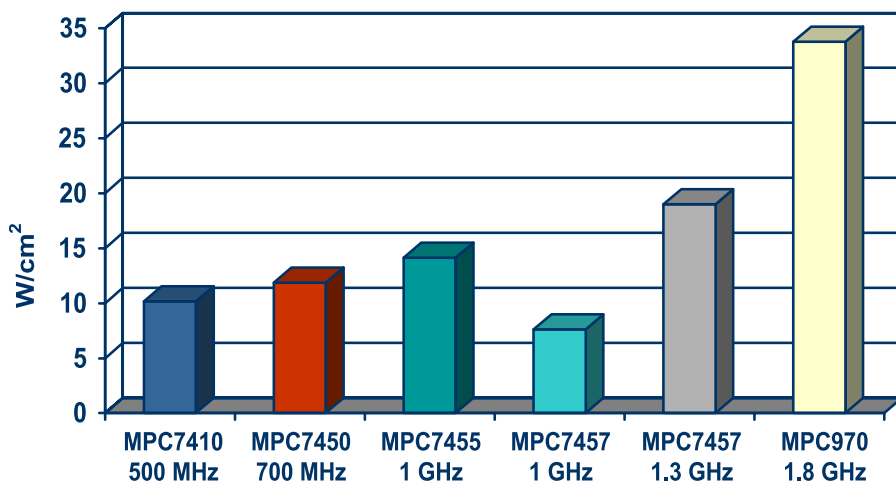


Figure 1 Power flux of recent PowerPC processors.

Component-Level Techniques

The classic heat dissipation technique at the component level is the heat sink; and faster components have always challenged heat-sink designers. The same die shrink that increases clock frequency decreases the heat-generating area and often increases total power dissipation. This effect applies to all types of components, from processors and FPGAs to power components.

When size and weight are not major constraints, the size of the heat sink used can be many times the area of the processor itself and several inches thick. To ensure the processor gets sufficient air, a fan is mounted directly to, or even embedded in, the heat sink.

For high-performance systems deployed in size-constrained environments, such as the PowerStream 7000, placing even moderately sized heat sinks on a large percentage of the board's components is impractical. The mounting hardware for a heat sink consumes precious area on the printed circuit board. This forces a tradeoff between the cooling effectiveness and the performance density of the system. The first technique of ManagedAir cooling avoids this tradeoff by mounting the heat sinks not to the PCB, but to a cover that encloses the board. The number of mounting points for the cover is much less than the sum of the mounting points for all of the heat sinks, thus minimizing the impact of the through-board mounting on signal routing and component layout.

Apart from EMI shielding gained with a board cover, this approach has some similarities to a technique that creates a monolithic heat sink that covers an entire board and conforms to a height profile of the components below. The differences become apparent at the high-level elements.

Board-Level Techniques

At the board level, maximum thermal efficiency requires careful placement of the hot components. Ideally, each processor or high-power component should get an independent flow of air.

This goal conflicts with current designs for high-density, high-end, deployed commercial off-the-shelf (COTS) systems that use large area boards, such as VME 9U x 400mm, to maximize the number of processors per given volume using both the available height and depth. New commercial form factors have similar approaches, such as PICMG 3.0 (AdvancedTCA) with a board 8U high and 280mm deep, or the proposed VITA-34 with a board 8U high and 220mm deep. In those tall-board designs, there can be as many as three to six rows of high-power components in an attempt to maximize processing physically in the layout.

Figure 2 shows a deployed VME 9U board with 18 processors in three rows. With traditional cooling, components not on the inlet edge of the board will see air that was already heated by processors or other high-power components lower on the board, along with a pressure drop from the impedance of those components and their heat sinks.

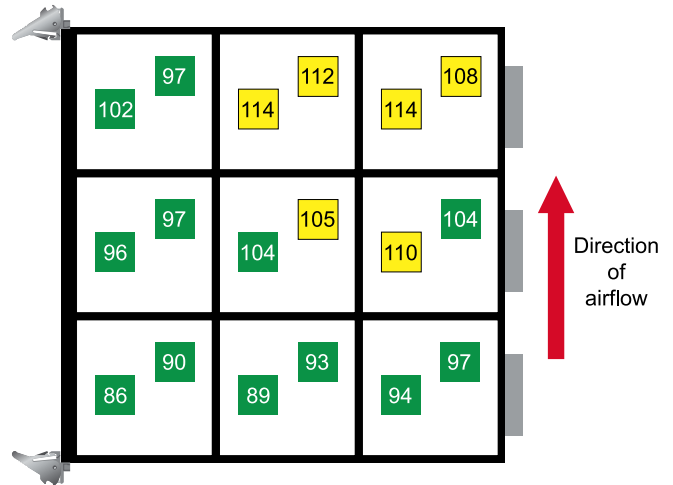


Figure 2 Example 9U VME board with three rows of 8W processors. Numbers indicate processor temperatures with air inlet temperature of 55°C with 18 CFM at sea level.

PowerStream 7000 designers addressed this problem by selecting the M155 board's 5.7" high x 15.7" long form factor, and locating high-power components along the bottom of the board on the leading edge of the airflow. Air moving vertically across this short, wide board provides the maximum thermal dissipation by combining this longer leading edge with a lower flow resistance.

Yet even if all the hot components are on the leading edge of airflow, there is still the challenge of getting the air to flow through the heat sinks. In conventional designs, the path of least resistance causes most of the air to flow around the high-impedance heat sinks instead of through them. With the ManagedAir technique, the same board cover that provides the mounting for the component heat sinks can contain features that direct air through the heat sinks, increasing their cooling capacity (See Figure 3).

Card-Cage Level

Traditional forced-air-cooled systems often employ coarse-grain air-management techniques to equalize airflow across slots in a chassis. This typically takes the form of plenums that restrict the flow of air to some slots. The plenums are generally developed using the same backpressure for each slot, and the resulting static pattern provides a reasonable solution for an unknown set of boards. If the board set is

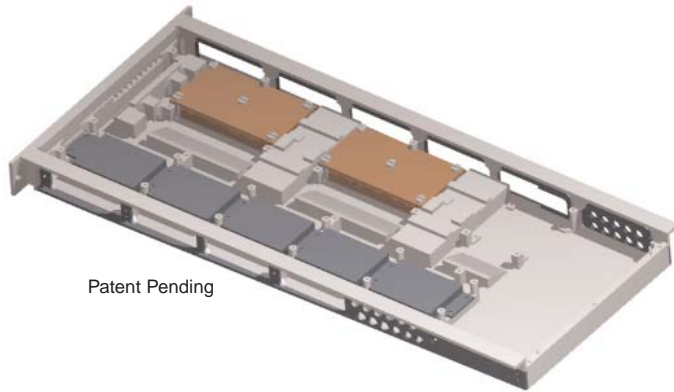


Figure 3 Example of a board cover with attached heat sinks and air-shaping features. Behind the processor heat sinks (shown in brown), a ramp feature pushes the air to the lower profile memory chips instead of allowing it to flow over the top.

known, the static plenum can be customized for maximum efficiency on that system configuration. If the configuration subsequently needs to change, because the requirements changed or new technology needs to be inserted, a new plenum must be created.

In the PowerStream 7000's ManagedAir technique, the plenum is part of the board cover. Boards with the hottest components have the least blocking plenum. Since the plenum travels with the board, new configurations automatically direct the appropriate portion of airflow to each board. Similarly, because unmanaged high-velocity airflows tend to ride up the backplane and starve the front of the board, ManagedAir balances the airflow from front to back within individual slots to maximize cooling efficiency. For example, the board cover shown in Figure 3 contains a plenum on both the board's air intake and exhaust that reduces the airflow to the right side (rear) of the board.

Chassis Level

At the chassis level, ManagedAir techniques help to regain the component density given up at the board level with the choice to use short boards instead of conventional, tall boards. To achieve the highest density in a given vertical space, two boards are stacked vertically with an air intake in the middle. An example of such a configuration is shown in Figure 4. This arrangement achieves two leading edges of inlet air, enabling high density without resorting to "columns" of the hot components, such as processors, heating one another. As with conventional air-cooled systems, chassis-level features are also used to shape the airflow at the chassis inlet areas to distribute the cooling air volume where it is needed most. When the chassis has the air intake in the

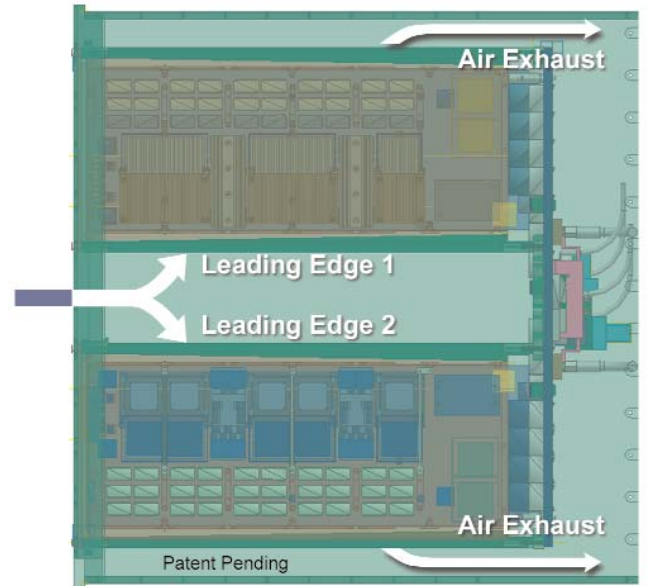


Figure 4 Stacking two boards vertically in a single chassis maximizes the functional density. A center air intake provides airflow to each of the boards.

middle of two rows of boards, the chassis intake features must also split the air without significantly reducing its velocity.

Modeling the Effectiveness of the ManagedAir Technique

Software modeling of a deployed military environment demonstrates that fine-air management at the component and board level can reduce component operating temperatures by 10-20°C. Figure 5 compares an M155 board populated with 20W PowerPC processors cooled by conventional air techniques with an identical board utilizing ManagedAir cooling. In this Flotherm model, inlet air is assumed to be 55°C at 10,000 feet with 15 CFM per slot. The hottest component on the board is a processor with a case temperature of 103°C. The use of a board cover with built-in heat sinks and air shaping features reduces that temperature by 16% to 87°C.

Figure 6 illustrates the additional improvements attainable with aggressive application of ManagedAir techniques throughout the chassis. This includes shaping the airflow at the chassis air intake, air exhaust, and distribution among the boards. Shaping and managing the airflow throughout the system yields sufficient additional cooling to permit the use of hotter next-generation components in deployed military environments, which may offer only limited headroom when used with today's components and unmanaged air-cooling methods.

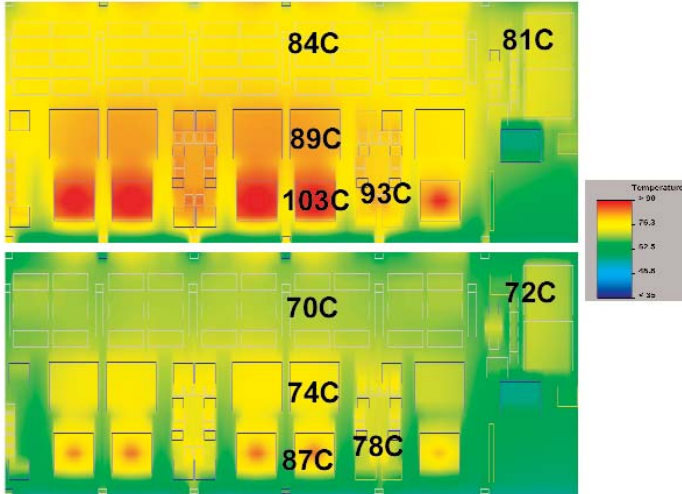


Figure 5 Flotherm images of 20W processors at 10,000 feet with 55°C inlet air. The processor that was running at 103°C with conventional air cooling (top diagram), now runs at 87°C with ManagedAir cooling techniques (bottom diagram) applied to the same inlet air temperature and flow.

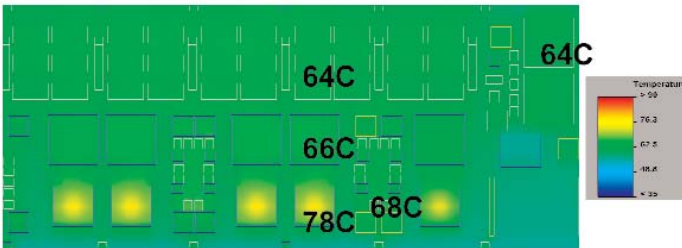


Figure 6 Effects of ManagedAir-cooling techniques when designed as an integrated chassis. The operating temperature of the hot processor is further reduced to 78°C. This example uses the same 20W processors at 10,000 feet with 55°C inlet air.

Achieving Higher Performance Density

Many innovative packaging solutions from Mercury allow PowerStream 7000 systems to support up to 25% more processing power in a small-volume chassis. Much of these gains come from the increased cooling ability of ManagedAir technology.

For a total system solution, solving this present day thermal management challenge isn't the entire issue. Chassis-level functional density must complement these solutions. The performance density of the overall system will also depend on such things as I/O throughput and interprocessor communication bandwidth, as well as reliability and serviceability. Each of these system parameters puts further demands on board space, power, and cooling. ManagedAir-cooling techniques can reduce the performance-limiting tradeoffs required to achieve the full set of system goals.

*Randall Banton is the vice president of defense electronics engineering and applications.
Richard Jaenicke is the director of product management.*

*U.S. Patent No. 6,661,657

RACE++, the RACE++ logo, and PowerStream are registered trademarks, and AdapDev and ManagedAir are trademarks of Mercury Computer Systems, Inc. PowerPC is a registered trademark of IBM Corp. RapidIO is a registered trademark of the RapidIO Trade Association. Other products may be trademarks or registered trademarks of their respective holders. Mercury Computer Systems, Inc. believes this information is accurate as of its publication date and is not responsible for any inadvertent errors. The information contained herein is subject to change without notice.
Copyright © 2004 Mercury Computer Systems, Inc. 084-1203-Powerstream 7000WP