64

1999 Cryogenic Engineering Conference
Paper CEC-3 : Cryogenic Electronics

PERFORMANCE OF A MIXED-REFRIGERANT SYSTEM DESIGNED FOR COMPUTER COOLING

Michael J. Ellsworth, Jr.,[1] Eric K. Moser,[2] Ajay Khatri,[3] and Mikhail Boiarski[3]

[1] IBM Corporation, Poughkeepsie, NY, USA.
[2] IGC-Technology Development, Latham, NY, USA.
[3] IGC-APD Cryogenics, Allentown, PA, USA.

ABSTRACT

CMOS-based computer processors will operate at increased frequencies when cooled to sub-ambient temperatures by a refrigeration system. This benefit must be weighed against many refrigerator considerations, the most noteworthy being thermal performance, efficiency, reliability, and cost. With this in mind a throttle-cycle, mixed-refrigerant cryocooler using a single-stage, oil-lubricated compressor was modified for increased capacity at 173K (–100°C) and its thermal performance was characterized. This mixed-refrigerant system cooled 119 Watts (with an average heat flux of better than 5 Watts/cm^2) at an interface temperature of 173K, with a Coefficient of Performance (COP) of more than 22%. The maximum in cooling capacity was 140 Watts at 183K (–90°C). The ratio of the COP relative to Carnot (ideal) efficiency held very constant at 16% over the temperature range of 153 to 183K (–120 to –90°C). This performance (relative to Carnot) is comparable to that of refrigeration systems with similar capacity operating at much higher temperatures. These results establish the good efficiency obtainable using mixed-refrigerant technology, making cooled CMOS applications attractive. Based on experimental data and computer simulations, we anticipate that better than 20% of Carnot efficiency can be achieved with a more optimal design.

INTRODUCTION

Computing performance levels continue to rise at a frenzied pace. This trend is expected to continue in the future, despite all technological obstacles. In this context, thermal management becomes an ever more challenging problem, and active refrigeration is an increasingly attractive technique for enhancing performance.

The paper reviews the benefits of computer cooling and presents an introduction on mixed-refrigerant systems, with a brief discussion of the challenges to be addressed in refrigerators specifically made for computer cooling applications, and a historical perspective on computer cooling. The mixed refrigerant system designed for 173K (–100°C) operation is presented, and its performance described and discussed.

BENEFITS OF COOLING COMPUTERS

Cooling a CMOS processor does not automatically increase performance, but there are benefits associated with cooling that enable performance increases. These benefits can be subdivided into device level improvements (improvements in the characteristics of the transistor devices themselves) and interconnect improvements (improvements in the characteristics of the on-chip "wiring" connecting the devices). Carrier mobility is increased, and devices feature reduced threshold voltages, steeper transitions, and lower leakage currents. Lower interconnect resistances reduce the interconnect delay times. The thermal conductivity of silicon also improves, aiding in heat transfer.

The Semiconductor Industry Association (SIA) provides a well-know "roadmap" projecting state-of-the-art processor technology into the future. The roadmap is exemplary of the rapid rate of advancement in computer performance.

Table 1: 1997 SIA Roadmap Projections.

Year	1997	1999	2001	2003	2006	2009	2012
Minimum Feature Size (µm)	0.25	0.18	0.15	0.13	0.10	0.07	0.05
Processor Speed (MHz)	400	600	800	1000	1200	1400	1500
Chip Size (mm^2)	480	800	850	900	1000	1100	1300
Device Voltage (Vdd, Volts)	1.8-2.5	1.5-1.8	1.2-1.5	1.2-1.5	0.9-1.2	0.6-0.9	0.5-0.6
Number of Interconnect Levels	6	6-7	7	7	7-8	8-9	9
Transistors Per Chip	11M	21M	40M	76M	200M	520M	>1G

The traditional method of increasing CMOS performance is "scaling," or the reduction of all device dimensions, shrinking the circuit and providing higher circuit density. Scaling approaches are the basis for Moore's law (the well-known rule for doubling of transistors per processor every 18 months), as well as the aggressive SIA roadmap projections. However, as feature sizes decrease below the 0.1µm level, scaling becomes much more difficult, as the dimensions become comparable to characteristic lengths for diffusion layers and device field zones. At reduced temperatures, CMOS can be operated at reduced supply voltages, thus providing some relief. A unique feature of low-temperature operation is that scaling *can* be applied in circuit designs with sub-0.1µm feature sizes.

Radical advances are implicit in the SIA roadmap. Recent years have seen advances in lithography and interconnect material. Reduced temperature has not been fully exploited as an avenue to higher performance, though it is now mentioned in the roadmap[1] as a "potential solution."

As a practical matter, overall processor speed is limited by either the devices *or* the interconnects. For a given processor, there is a maximum allowable clock frequency, F_{max}, and decreasing the processor temperature results in an increase in F_{max}. Increasing the clock frequency above "designed for" levels is termed "overclocking," especially as it applies to personal computing. While overclocking is a specialized operating option of limited use, it provides an indication of the speed enhancements possible with actively cooled computers.

The percentage increase in F_{max} rises monotonically with decreasing temperature, a bit faster than linearly, the slope depending on the extent to which the circuit is designed to operate at reduced temperatures[2]. For instance, generic CMOS circuits cooled to 173K might feature a 50% increase in F_{max}, while CMOS with optimized doping could be operated at 80% higher clock frequency when cooled to 173K.

HISTORICAL PERSPECTIVE ON COMPUTER COOLING

Low-temperature operation has historically been a holy grail for high-performance computing[3]. Many attempts have been made in the past, with no sustained commercial success. IBM and Carrier Corp. collaborated on a computer cooling system using Stirling technology in the early 1980s[4], but the refrigerator proved too cumbersome. The Cray ETA-10 was a true cryogenic computer, utilizing liquid nitrogen (77 Kelvin) for cooling[5]. However, liquid nitrogen-cooled computers are no longer commercial. RSFQ superconducting circuits operated at liquid helium temperature (4 to 5 Kelvin) show promise and are presently at the early development stage[6].

Kryotech, Inc., Columbia, SC, is a unique company dedicated to the commercialization of cooled computing[7]. Other competing companies have also arisen. Kryotech's current cooling systems utilize vapor-phase refrigeration technology based on Joule-Thomson expansion to actively cool processors to -40°C (233K).

The most appropriate temperature range for cooled computers is a matter of opinion, since a number of complicated factors apply. The potential computing benefits are enhanced by operating at very low temperatures, but the refrigerators are limited by fundamental performance limits. Carnot's law[8] places an upper bound on refrigerator efficiency. Coefficient of performance is often expressed as a fraction of Carnot efficiency. Our rationale for building a refrigerator in the 173K range is that there may be a "sweet spot" where performance increases are fairly balanced with the increased input power and added complexity associated with active cooling.

MIXED-GAS REFRIGERATION SYSTEMS

Refrigerators based on the principle of Joule-Thomson (JT) expansion provide a proven, reliable method for localized cooling. Two common examples of "vapor phase" JT systems are air conditioners and kitchen refrigerators. Similar systems using mixed gas refrigerants[9] enable JT refrigerators to operate at lower, including cryogenic[10], temperatures. A generic diagram for a JT refrigerator for cryogenic cooling is shown in Figure 1.

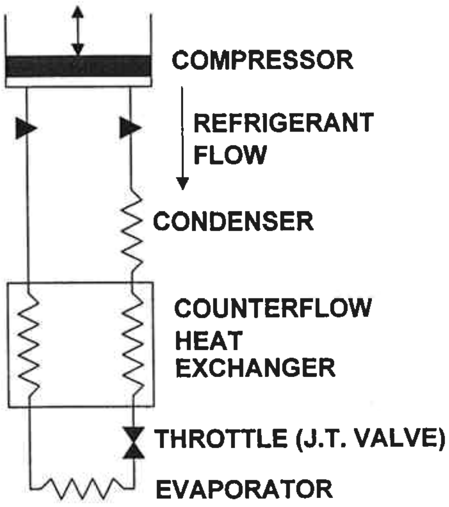

Figure 1. The Joule-Thomson Cycle.

The system is recuperative, with refrigerant flowing continuously in one direction. The particular components required and the configuration vary somewhat with factors including the desired temperature of operation, capacity, and efficiency. For example, the counterflow heat exchanger and oil separator (not pictured) would not be needed at relatively high temperatures.

The IGC-APD CRYOTIGER® system is a commercial low-temperature JT system with blended refrigerant, designed to provide temperatures in the range of 70 to 120K. The CRYOTIGER® cools to these temperatures using an innovative, patented blend of refrigerants. The compressor uses 450 Watts of input power and it can provide over 2 Watts of cooling at 70K, over 20 Watts at 125K, or over 80 Watts at 200K by charging the system with properly selected refrigerant blends. The CRYOTIGER® provides a valuable benchmark for computer cooling, though it is designed for lower-temperature operation.

REFRIGERATOR ISSUES

Generally speaking, refrigerator issues include evaporator temperature and heat capacity, efficiency, reliability, size, and cost. When considering the use of refrigeration in a computer cooling application, evaporator temperature distribution, temperature stability, and initial cool down time also come into play. Furthermore, overall system availability is critical. To assure continuous system operation, IBM has incorporated a redundant design where two independent refrigerant loops pass through a single evaporator[11].

Packaging the refrigeration unit in the computer system is also important. The refrigeration system must be transparent to the end user (i.e., the end user shouldn't have to know or care that refrigeration is taking place.) Refrigeration system heat must therefore be transferred to the ambient air in the same manor as do conventional computer cooling systems today. Condensation must also be prevented from forming on components within the computer as condensation on electronics components is detrimental to reliability and functionality. This is a major issue[12], but further consideration is beyond the scope of this paper.

A wide variety of different approaches to building a refrigerator for computer cooling are possible. Active cooling systems for computers must compete with passive alternatives such as air cooling and heat pipes as well as an array of refrigeration technologies including thermoelectrics, Gifford-McMahon systems, pulse tubes, and Stirling Engines. Commercially successful refrigerators and cryocoolers are generally highly adapted to meet specific applications. Computer coolers have context in the broader classification of spot-cooling electronic components such as filters, detectors, lasers, and superconducting devices. No "perfect" refrigeration technology exists, so expansion of refrigerators into new markets depends on successful compromise between the benefits of cooling and the hassles of incorporating the refrigerator. This having been said, we return our focus to mixed-gas systems as computer coolers.

THE 173K (–100°C) PROTOTYPE MIXED-GAS REFRIGERATOR

Several modifications were made to an "off-the-shelf" CRYOTIGER® system to achieve 173K cooling with good efficiency. A custom blended refrigerant was used, the oil separator within the compressor unit was modified, and both the heat exchanger and evaporator area were increased in size to enable cooling capacity above 100 Watts. A larger diameter capillary tube (throttle) was also used to accommodate the higher refrigerant mass flow rate.

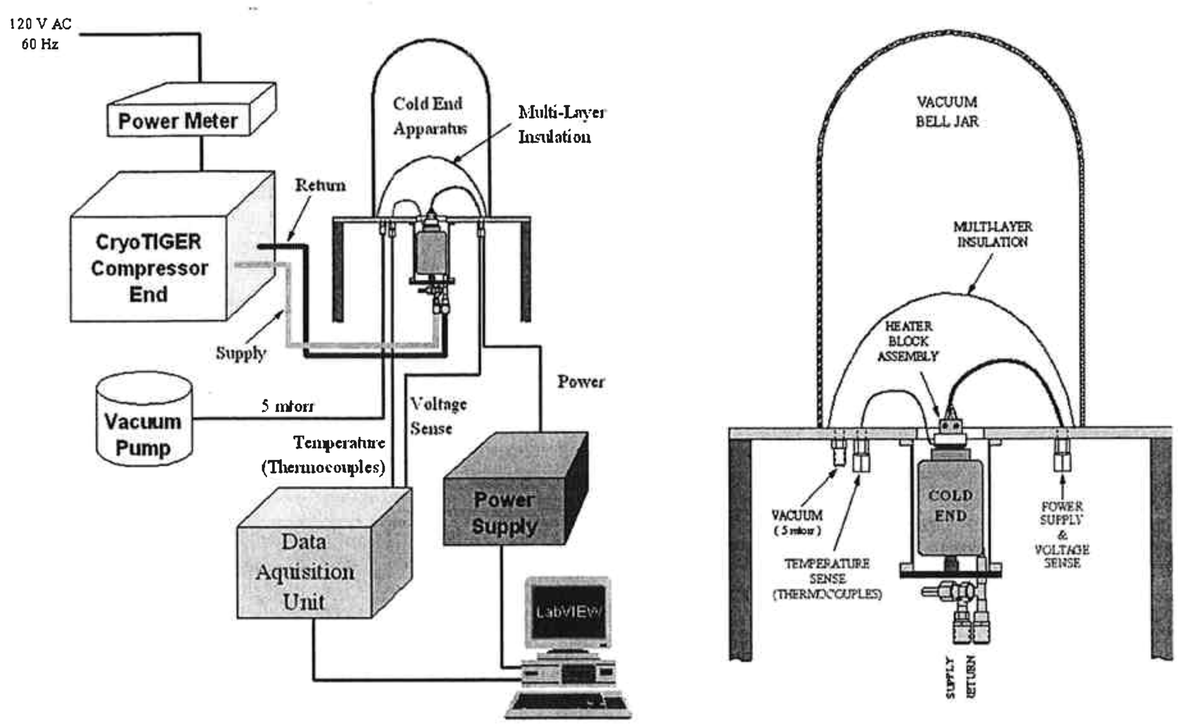

Figure 2. Experimental apparatus.

The prototype unit was tested and characterized in the IBM Advanced Thermal Laboratory. The experimental setup, with some detail on the test apparatus, is depicted in Figure 2.

Care was taken in the design of the experimental apparatus to minimize the infiltration of heat into the system by virtue of the temperature difference that exists between the cold end and the environment during operation. This parasitic heat load is minimized so that the unit could be accurately characterized in terms of its heat capacity-evaporator surface temperature behavior. The cold end, comprised of the heat exchanger, capillary tube, and evaporator, are maintained at a vacuum environment of approximately 5 mtorr to minimize the infiltration of heat by convection. Multilayer insulation is placed over the evaporator end to minimize radiative heat infiltration. The majority of parasitic heat load, therefore, comes from conduction through the mechanical support of the components. By characterizing the thermal capacitance of the system (referenced to the evaporator surface temperature) and measuring the rate at which the system increased in temperature with no heating or cooling load applied, it was estimated that the parasitic heat load at 173K was approximately 2.5 Watts, which constitutes roughly 2.1% of the applied heat load at said temperature.

Electric resistance heaters were used to apply the heat load. The voltage drop across the heaters (measured as close to the heater as practicable) and the voltage drop across a precision resistor (i.e. a known resistance) was used to determine the power dissipated by the heaters. It can be assumed that 100% of the heat dissipated within the heaters is absorbed by the evaporator. Temperature was measured using type E (Nickel-Chromium vs. Copper Nickel) thermocouples. Type E thermocouples are ideally suited for low temperature measurements because of their high Seebeck coefficient (58 μV/K). Five thermocouples are placed in contact with the evaporator surface: three located within the center; one located towards one corner, and the last located in the corner diagonally opposed from the previous thermocouple. Voltage and thermocouple measurements were made using an HP3852 data acquisition system. The data acquisition and the application of electrical heater power were controlled using a personal computer. Finally, the electrical power supplied to the compressor end was measured so that coefficient of performance could be determined.

A number of performance characteristics were measured. First was the time required to bring the system down from room temperature to operating temperature. The system was not able to absorb the requisite heat load until it was at operating temperature. This translates in a computer application to not being able to turn on the computer electronics until the cooling system is down to operating temperature. For large system servers with expectations of 99.9999% or better availability this time delay is critical. It took the prototype unit over 50 minutes on average to get down to operating temperature (see Figure 3).

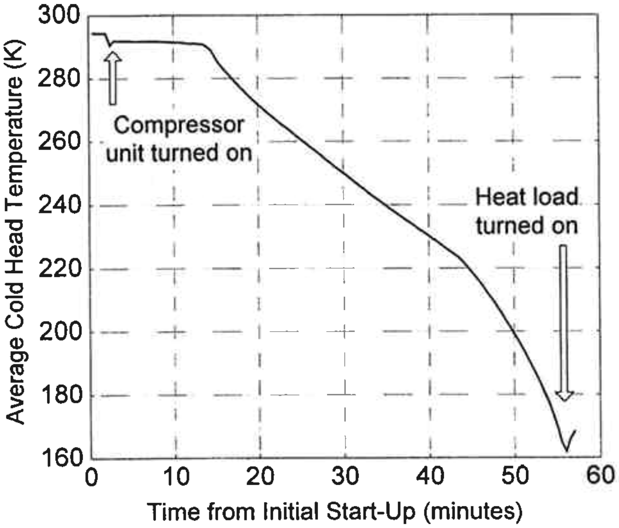

Figure 3. Cooldown (initial configuration).

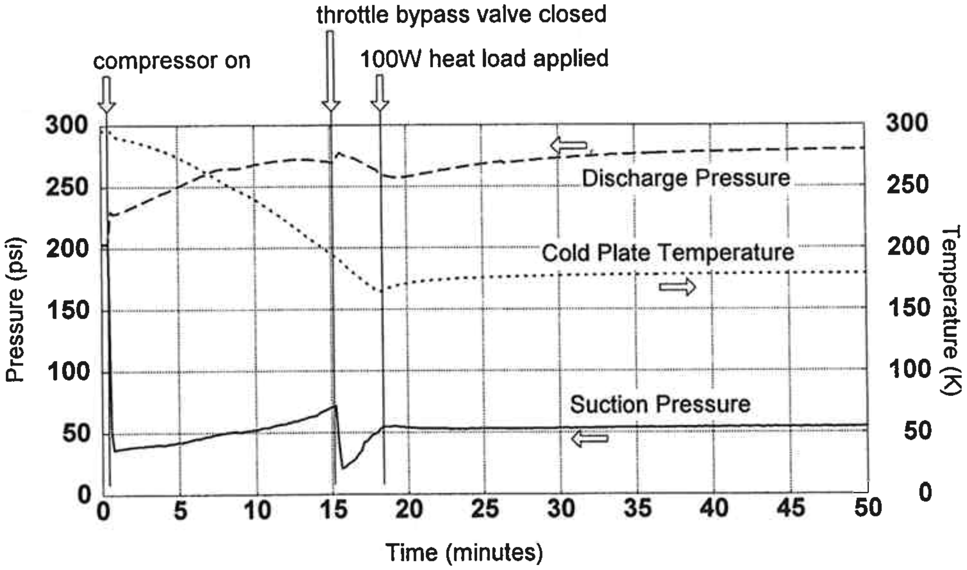

Figure 4. Improved cooldown.

Adjusting this result to take into account the thermal capacitance of a large system server, the cool down time would approach 75 minutes. This turned out to be an obvious concern. To address this concern, a second system was built utilizing a special throttle designed to enable high flow rate during the cool down period, then switch over to a continuous operation setting. Figure 4 shows that the cool down period could be improved to less than 20 minutes using this throttling scheme.

The next characteristic to evaluate was the heat load - temperature behavior of the system. Figure 5 shows how heat load varies as a function of average evaporator (cold head) temperature. The unit could absorb 119 Watts (approximately 5.3 Watts/cm^2) at 173K. The maximum load the system would sustain a stable temperature at was found to be 140.3 Watts at 182.3K. Power above this level would result in temperature runaway. The data in Figure 5 also suggests that the sensitivity in change to heat load is roughly 0.5K/Watt. The transient temperature response for a step change in heat load was observed to behave similarly to the step response of an electrical RC circuit.

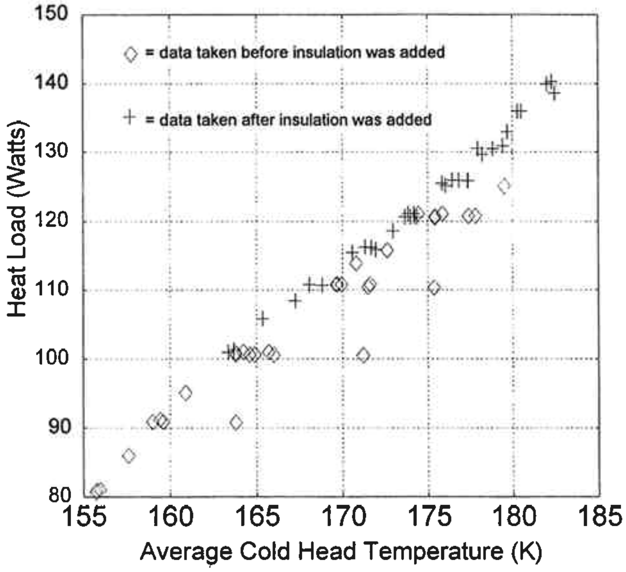

Figure 5. Capacity vs. temperature.

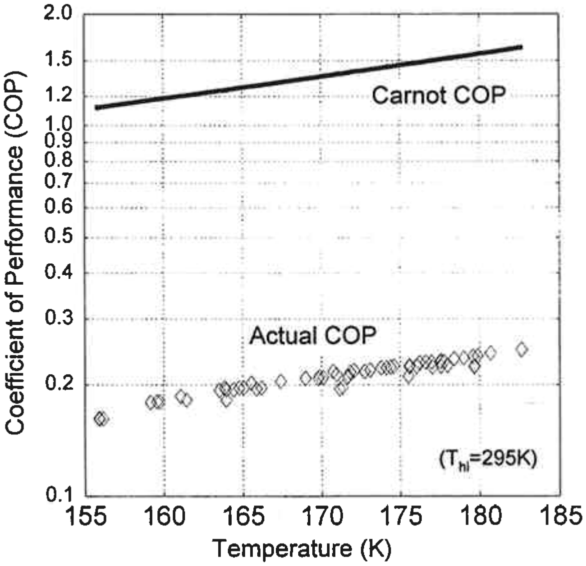

Figure 6. COP vs. temperature.

Finally, the coefficient of performance (COP) was evaluated as a function of average evaporator surface temperature (Figure 6). The unit was found to operate at a COP of 0.22 173K. Furthermore, the unit appears to operate consistently at 16% of Carnot COP over the range of 153 to 183K (-120°C to -90°C). This compares very favorably with other refrigeration technologies.

This system demonstrates capacity, efficiency, and cooldown time. Further obstacles remain before the refrigerator can truly be regarded as ready for computer cooling.

FUTURE IMPROVEMENTS

In order to attain maximal refrigerator efficiency, careful designs are needed matching refrigerant, compressor, heat exchanger, and cold plate attributes. Computer simulations have proven useful for exploring these design issues. A detailed discussion on mixed refrigeration systems operating in the temperature range of 120K to 200K is given in an accompanying paper by Boiarski, et al.[13]

SUMMARY

The benefits of computer cooling are well known. There are many challenges associated with the marriage of refrigeration technology with computer technology. Mixed-gas systems are an attractive refrigeration technology for computer cooling.

ACKNOWLEDGEMENTS

We acknowledge support from others involved in this work including George Manelski, Robert Simons, and Richard Chu at IBM, Terry Rufer, Noel Fasnacht, and Stephen Harold at APD, and Boris Yudin of Intercool. We acknowledge useful discussions with John Peeples of Kryotech and Kenneth Rose of Rensselaer Polytechnic Institute.

REFERENCES

[1] Further SIA roadmap information is available at the Sematech website, www.sematech.org/public.

[2] Discussed by Y. Taur and R. Dennard of the IBM Watson Research Center at the *IEEE Cool Electronics Workshop*, Washington DC, 15-16 October, 1998.

[3] R. Solomon, Opportunities and problems for refrigeration in small computers, *1991 IIR Proceedings*, p. 202. See also J. Sloan, Cold computing: the future of high-speed processing, *Superconductor Industry* Fall 1996, p. 30, and R. K. Kirschman, Low-temperature electronics, *IEEE Circuits and Devices Magazine* 6:12 (1990).

[4] R. E. Schwall and W. S. Harris, Packaging and cooling low temperature electronics, *Advances in Cryogenic Engineering* 37-A:587 (1991).

[5] D. M. Carlson, D. C. Sullivan, R. E. Bach, and D. R. Resnick, The ETA10 LN cooled supercomputer, *IEEE Trans on Electron Devices* ED-36:1404 (1989).

[6] Aaron Bitterman, Superconductors & cryoelectronics in the petaflops-scale computer project, *Superconductor & Cryoelectronics*, Spring 1999 Issue, p. 17.

[7] Further information is available on the Kryotech, Inc. website, www.kryotech.com.

[8] Carnot's law of refrigeration states that the efficiency of an ideal refrigerator is given by $T/\Delta T$ where T is the refrigeration temperature and ΔT is the difference between the refrigeration temperature and the ambient temperature, in Kelvin.

[9] M. Boiarski, V. Brodianski, R. Longsworth, Retrospective of mixed-refrigerant technology and modern status of cryocoolers based on one-stage, oil-lubricated compressors, *Advances in Cryogenic Engineering* 43:1701 (1998).

[10] Strictly speaking, "Cryogenic" refers to temperatures below 120K (or -153C).

[11] P. Singh, D. Becker, V. Cozzolino, M. Ellsworth, R. Schmidt, and E. Seminaro, System packaging of a CMOS mainframe, *Advanced Microelectronics*, 25:12 (1998).

[12] This topic was discussed by J. Peeples of Kryotech at the *Cryoelectronics Technologies in Microelectronics* workshop, sponsored by the Knowledge Foundation, San Francisco, 18-19 February, 1999. See also J. W. Peeples, The deskside supercomputer, *Advances in Cryogenic Engineering* 43:849 (1998).

[13] M. Boiarski, A. Khatri, and O. Podcherniaev, Enhanced refrigeration performance of throttle-cycle coolers operating with mixed refrigerants, *1999 CEC/ICMC*, Montreal, Canada.