High Performance Air Cooled Heat Sinks for Integrated Circuits

CLAUDE HILBERT, MEMBER, IEEE, SCOTT SOMMERFELDT, OMKARNATH GUPTA, MEMBER, IEEE, AND DENNIS J. HERRELL, SENIOR MEMBER, IEEE

Abstract—**A simple, compact, and high performance air cooling technology has been developed. It combines laminar flow with novel heat sink designs in order to achieve a cooling performance of less than 2°C/W/cm^2 at 2 ft^3/min. We have demonstrated the capability to remove in excess of 600 W from a multichip module (25 chips at 25 W and 16% chip-to-substrate coverage) with a quiet 10-W tubeaxial fan. The volume of the complete module including the plenum and the fan is only about 1 L. The chips are held to a maximum temperature rise of 55°C. Even higher cooling performance at greater chip-to-substrate coverage was shown to be possible with larger air movers. This technology extends the applicability of air cooling to power levels traditionally considered to lie in the domain of complex liquid cooled modules.**

I. INTRODUCTION

THE TREND to larger chips and higher power density on each chip has made thermal management an important aspect of semiconductor packaging technology. Chips with power dissipation up to 10 W have been in existence for a while and 1-cm^2 chips with power dissipations of 20 W are now becoming available. In addition these chips are frequently packaged in close proximity to each other in multichip modules in order to minimize interchip propagation delays. The total package power dissipation can easily reach several hundred watts.

Liquid cooling will generally be able to outperform air cooling, because of the higher thermal conductivity and density of the liquids. However, air cooling is less expensive, more portable, and generally evokes less apprehension among users. Traditionally, air cooling has been thought to be fairly inefficient. However, this is largely due to the manner in which the air cooling systems have been designed. They usually make only partial use of the cooling potential of air by simply blowing it around large heat sinks with an open fin structure. By containing the air flow and forcing the air to flow through the heat sink we manage to make all the air participate in the heat exchange and create a highly efficient air cooling system. The price to be paid for this tremendous increase in cooling efficiency is only a slight increase in pressure requirements, as can be seen from the result of this paper.

Manuscript received March 9, 1990; revised July 30, 1990. This paper was presented at the Sixth Annual IEEE Semiconductor Thermal and Temperature Measurement Symposium, Phoenix, AZ, February 6–8, 1990.

C. Hilbert, S. Sommerfeldt and D. J. Herrell are with the Microelectronics and Computer Technology Corporation, Austin, TX 78727.

O. Gupta is with Gupta Packaging Technologies, Fremont, CA 94539.

IEEE Log Number 9039046.

The basic technique used in this paper is forced convection through finned heat exchangers. Conventional forced convection heat exchangers achieve their high heat transfer by employing turbulent flow around finned structures. On the other hand Tuckerman and Pease [1], [2] demonstrated that a very efficient and compact liquid heat exchanger could be made using laminar flow through microchannels whose dimensions were on the order of a few tens of micrometers. The thin boundary layers provided a high heat transfer coefficient, and the laminar character of the flow kept the pressure requirements low in comparison with those of turbulent flow. We are aware of only two reports [3], [4] on air-cooled microchannels applied to integrated circuits. Our work combines the principles demonstrated in these papers [1]–[4] with novel fin geometries and easily manufacturable heat sink designs in order to achieve a high performance and compact air cooling system which extends the applicability of air cooling to power levels traditionally considered to lie in the domain of liquid cooling. This paper focuses on the experimental feasibility demonstration of employing ducted air for cooling of high powered chips and packages.

II. HEAT SINK DESCRIPTION

All air cooled heat sinks described in this paper combine two important features. They take advantage of the high heat transfer capability of laminar flow at low pressure drops through narrow channels formed by two interleaved set of fins. At the same time they are designed to duct the air flow from center top to lateral vents, thus changing its direction from vertical incidence (impingement in some cases) to horizontal exit (or vice versa). This concept of a vertical air inlet and a horizontal outlet enables us to supply cool inlet air to a multitude of heat sinks connected in parallel into the same plenum. Hence, we eliminate all downstream effects in the cooling of chip arrays. Having the cool air enter at the center of the heat sink and exit at the sides rather than forcing it across the whole heat sink from one side to the other reduces the required pressure drop and decreases the average velocity through the channels. This produces less turbulence at the entrance and exit of the heat sinks, which in turn generates less noise and lower pressure losses. Finally center feeding the heat sink decreases temperature gradients across the heat sink and the chip.

One geometry for air cooled heat sinks that we have studied extensively, is shown schematically in Fig. 1. The heat sink is fabricated from triangular shaped fins which are

0148-6411/90/1200-1022$01.00 © 1990 IEEE

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 1
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

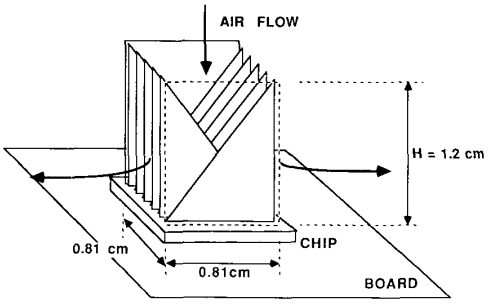

Fig. 1. Schematic perspective view of heat sink Design #1. The top of the heat sink can be connected into an air distribution plenum or submitted to impinging air. The air enters the heat sink from the top and exits at the sides. The bottom of the heat sink is attached to the chip or package. The tip of the triangular fins can be folded to provide constant fin spacing. The prototype dimensions are indicated on this figure.

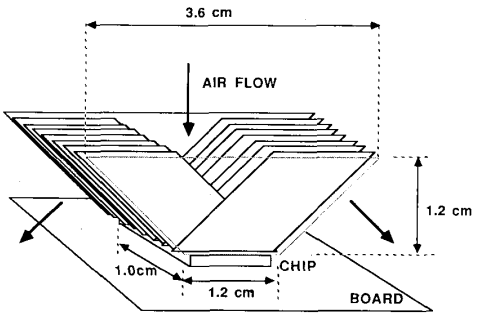

Fig. 2. Schematic perspective view of heat sink Design #2. The top of the heat sink can be connected into an air distribution plenum or submitted to impinging air. The air enters the heat sink from the top and exits at the sides. The bottom of the heat sink is attached to the chip or package. The prototype dimensions are indicated on this figure.

interleaved at the bottom and alternatively point in opposite directions at the top. At each end a rectangular fin is added to form a prism shaped heat sink (Design #1). The bottom part of the heat sink is thus solid material whereas at the top narrow channels alternate with thin fins on each side. Each triangular fin acts as a spacer for the microchannel on the opposite side, yielding a one-to-one ratio of fin thickness to channel width. The fin tips are folded over to provide structural support, keep the fins equally spaced, and force the air to travel a substantial distance along the fins, thus favoring an increased heat transfer. The top of the heat sink appears as a rectangular orifice that ducts air into the intersecting fins. The thickness and the lateral dimensions of the fins can be varied to optimize the heat exchange and minimize the air flow resistance. Once assembled the bottom surface of the heat sink is attached to the backside of the chip or package by a number of methods: soldering, thermal grease, pressure contact, microgrooves, etc. As proposed, Design #1 is of the counterflow type: the entering air flows from the cooler end of the fin to warmer parts. This increases the effectiveness of the heat transfer over that of parallel flow [5]. The top of the heat sink can be connected into a plenum supplying the air or exposed to vertically impinging air.

The heat sink can be designed with other fin geometries. Another design we have focused our efforts on is illustrated in Fig. 2. It again combines a multitude of thin fins spaced by narrow channels with the center feeding in which the air enters vertically at the top, is ducted by the heat sink, and exits horizontally through the channels. Some of the compactness of the previous prism shaped design is sacrificed by expanding the fins as they rise above the chip. The greater surface area of each individual parallelogram shaped fin enables us to go to a wider fin spacing and thicker fins, i.e., a smaller number of fins per heat sink and still keep the same thermal performance. The end pieces are trapezoid shaped. In this design we departed from the one-to-one ratio of fin thickness to channel width in Design #1 in order to get wider channels and not waste any space with unnecessary fin material. In this trapezoid-shaped heat sink (Design #2) the channels are twice as wide as the fins are thick. This can be achieved with spacers at the top and/or at the bottom or by appropriate assembly methods. In addition to providing an increased fin surface area the expanding fins create a much larger inlet and outlet cross section for the air flow. This decreases the air velocity substantially, and hence, reduces the amount of noise generated by several orders of magnitude. Design #2, therefore, trades off compactness for lower pressure drop and noise.

The actual dimensions of the tested heat sinks are indicated in Fig. 1 and 2. For Design #1 the base was chosen to be 0.81-cm square wide in order to fit our thermal test dies (base surface area 0.66 cm^2). The triangular fin height is 1.4 cm before bending and the fins are folded at about 1.5 mm below their tip (not shown in Fig. 1). This particular size ratio yields an approximately equilateral inlet triangle for the air and a constant total cross-sectional area for the air flow. The goal is to keep the air speed relatively constant throughout the entry ducting and reduce pressure losses. Design #2 was fabricated with a 1.2 cm by 1.0 cm base, attached to a 1-cm^2 die. The individual fins are parallelogram shaped with a 45° angle between its sides. The fin height is 1.2 cm.

III. Theoretical Model

The following simple analysis is based on the theory of convective heat transfer applied to microchannels [1], [5]. The thermal resistance R_{th} of a heat sink has a heat term R_{heat} due to the heating of the cooling air and a convective term R_{con} representing the temperature rise across the thermal boundary layer. For high aspect ratio channels (infinite parallel plates) the maximum thermal resistance can be written approximately as

$$R_{th} = \frac{\Delta T}{P} \approx R_{con} + R_{heat} \approx \frac{2w}{\eta Nuk_a A} + \frac{1}{\rho C_p f} \quad (1)$$

where

- ΔT maximum temperature rise (Kelvins) of the heat sink base over the entrant air temperature,
- P total power dissipation (watts),
- η dimensionless fin efficiency,
- Nu dimensionless Nusselt number
- k_a thermal conductivity of air (0.00026 W/K · cm),

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 2
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

w channel width (centimeter), equal to half the hydraulic diameter,
A total exposed fin area (square centimeters),
ρ air density (0.0012 g/cm^3),
C_p heat capacity per unit mass (1.0 J/g · K)
f volumetric air flow (cubic centimeters per second).

The corresponding pressure drop Δp due to frictional losses across the channels is given by:

$$\Delta p = \frac{12\mu L v}{w^2} \tag{2}$$

where

μ air viscosity (0.00020 g/cm · s),
L channel length (centimeters),
v mean air velocity between the infinite plates (centimeters per second).

The fin efficiency for the triangular and parallelogram shaped fins can be computed analytically assuming one-dimensional heat flow [5]. For a parallelogram shaped fin (constant cross section) it is given by

$$\eta = \frac{\tanh(mH)}{mH} \tag{3}$$

and for a triangular shaped fin:

$$\eta = \frac{2}{mH}\frac{I_1(mH)}{I_0(mH)} \tag{4}$$

where H is the fin height (centimeters), I_0 and I_1 are modified Bessel functions, and

$$m = \sqrt{\frac{Nuk_a}{twk_f}} \tag{5}$$

where t is the fin thickness (centimeters) and k_f the thermal conductivity (watts per Kelvins times centimeters) of the fin material.

Using (1)–(5) we compute the thermal resistance of a Design #1 heat sink versus fin thickness for given supply pressures, neglecting any pressure losses in the manifold, entrance, and exit regions. For our geometries the flow distribution will be nonuniform due to the variable channel length. We assume the air velocity in the heat sink channels to be parallel to the hypotenuse of the triangular fins. For the purpose of this analysis we neglect shear forces existing between adjacent air elements traveling at different speeds. The local air velocity in the channel can then be obtained from (2). We also assume the Nusselt number to be equal to 8, which is its approximate value for fully developed flow between parallel plates. In Design #1 the fin thickness equals the channel width. By integrating the variable air velocity along the triangular fins for the parameters indicated in Fig. 1, we obtain:

$$R_{\text{th}} \approx \frac{2t^2}{\eta Nuk_a A_f W} + \frac{9.1\mu}{\rho C_p \Delta p\, t^2 W} \tag{6}$$

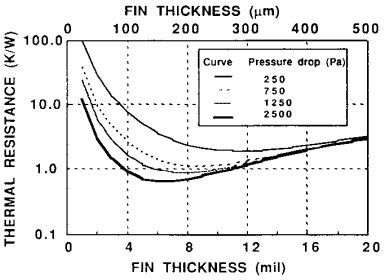

Fig. 3. Theoretical optimization of fin thickness at different operating pressures for heat sink Design #1. The geometry parameters used are those indicated in Fig. 1. The optimal fin thickness depends on the available pressure drop.

where A_f is the exposed area of a single fin and W is the total heat sink width.

Fig. 3 demonstrates the essential results for a copper heat sink (k_f = 3.88 W/cm · K), the geometry given in Fig. 1, and different applied air pressures. It shows that at any positive pressure a minimum exists for the thermal resistance. The optimal fin thickness is only a very weak function of the applied pressure. (It varies approximately as the fourth root of the air pressure.) The optimal thickness varies from 280 μm (11 mils) at 250 Pa (1 in of water) pressure to 150 μm (6 mils) at 2500 Pa (10 in of water). For channels and fins that are narrower than the optimum, the air viscosity reduces the flow rate through the heat sink drastically and thereby increases the heat term; for wider channels and fins the reduced surface area combined with the increased boundary layer thickness increases the convective thermal resistance. At the minimum the contributions of the heat and the convection terms are approximately equal. This plot also shows that for a given operating pressure, a cautious designer will probably want to choose a fin thickness that is somewhat greater than the optimum value, because of the steep increase in resistance for slightly narrower fins. Pressures of several inches of water can be achieved with centrifugal blowers. The total air flow delivered at optimal fin thickness increases with pressure whereas the thermal resistance decreases. The minimum thermal resistance achieved at the optimum fin thickness varies from 0.64 K/W at 2500 Pa with 2600 cm^3/s (5.5 ft^3/min of air flow, through 0.88 K/W at 1250 Pa and 1750 cm^3/s (3.7 ft^3/min), to 1.9 K/W at 250 Pa and 850 cm^2/s (1.8 ft^3/min). These values are based on a 0.66-cm^2 chip area and heat sink base.

The calculated local Reynolds numbers indicate that, even at the locations of highest air velocity, i.e., in the channels close to the fin tips, the air flow remains laminar in the heat sink when the fin thickness is optimized. This makes our analysis self-consistent and keeps the pressure drop to a minimum. The present discussion does not include any pressure drop corrections for turn losses, entrance, and exit effects, or for the manifold. These could turn out to be substantial.

The simple analytical model presented above is only approximate. In an exact theory the fin efficiency involves the precise knowledge of the air flow pattern and the value of the thermal resistance at each point. The Nusselt number for

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 3
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

fully developed laminar flow can vary between 8.23 for constant flux and 7.54 for constant temperature boundary conditions. Also, in the thermodynamic entrance region the local Nusselt number is higher than the assumed value of 8, and hence, the convective resistance will be somewhat reduced in the real heat sink. Finally, the solid bottom part of the heat sink will contribute a small addition to the total thermal resistance.

A similar analysis can be completed for Design #2. For the dimensions shown in Fig. 2, a fin thickness of 250 μm (10 mils) and a channel width of 500 μm (20 mils), the thermal resistance at a flow rate of 2 ft^3/min (944 cm^3/s) is predicted to be 1.58 K/W. At a reduced flow rate of 1 ft^3/min (472 cm^2/s) the resistance rises to 2.70 K/W. The fin spacing is about 3 times larger than for the prototype Design #1; this should reduce the pressure drop by a factor of 9. In addition the average cross section is roughly doubled and the average velocity about half its previous value. This should reduce the pressure drop by another factor of 2. Hence, we expect the pressure drop across this heat sink version to be only about 1/20th of the pressure drop across heat sink Design #1. This brings the pressure requirements down to a level where they are well within the possibilities of standard tubeaxial fans used for cooling electronic equipment.

As a final comment on this theoretical analysis and optimization, it should be mentioned that in practical applications the choice of a heat sink geometry is governed by the pressure flow characteristics of the available air sources. The flow and pressure that a given fan or blower can sustain are intimately related. The operating point of the whole cooling system will be determined by the intersection of the fan performance curve with the system flow impedance curve. For a given air source, changing the heat sink design will change both the operating pressure and flow. Lastly blower size, blower power dissipation, and noise generated in the blower and heat sink are all important considerations that depend on the application.

IV. Experimental Performance

A. Design #1

The prototype for Design #1 was fabricated from 5-oz copper with the other dimensions specified in Fig. 1. 5-oz copper is approximately 175 μm thick. This thickness was chosen both because of its ready availability and because it is close to the optimal thickness for pressures of a few inches of water (about 1000 Pa). With this fin thickness and channel width, a total of 46 fins (or channels) could be fitted on the back of our thermal die, yielding a total exposed fin surface area of about 24 cm^2. Equation (4) predicts a very high fin efficiency of approximately 96% for these parameters, i.e., a practically isothermal fin.

All the data for Design #1 were taken with thermal die. They contain 460 evenly spaced thin film resistors for heat generation and an array of 48 diodes spread among the resistors for die surface temperature measurements. In our measurements we used 9 diodes distributed evenly over the chip surface. They were biased with a constant current of 0.4 mA and we measured the temperature-dependent voltage across them with a Fluke 2285B data logger. The thermal dies were calibrated to $\pm 0.5°C$ in an environmental oven the temperature range from 25 to 75°C. In this temperature range the relation between diode voltage and temperature was very nearly linear with a negative slope of approximately 1 1 mV/°C. The thermal dies were tape bonded face down (flip-TAB) to a printed circuit board. The heat sinks were attached to the back side of the thermal dies with a very thin layer of thermal grease (GC silicone heat sink compound, Type Z9, #10-8109). The thermal conductivity of this grease is guaranteed to be higher than 0.004 W/K · cm.

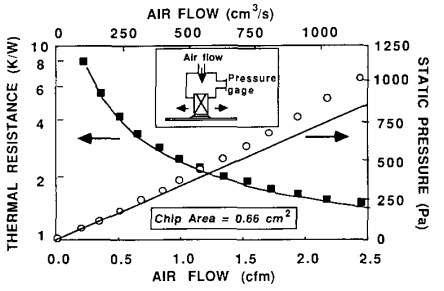

Fig. 4. Experimental performance of heat sink Design #1. Thermal resistance (on the left) and static pressure in the inlet plenum (on the right) are plotted versus air flow. Continuous lines indicate theoretical predictions based on our analytical model; discrete points show experimental results. The chip area is 0.66 cm^2. The insert shows a schematic cross section of the test setup. (1 in of water pressure = 249 Pa.)

The top of the heat sink was mounted into a small plenum where we measured the air inlet temperature with a type *T* (copper-constantan) thermocouple. For testing of individual heat sinks we used the in-house air supply, the flow rate being measured with a Cole-Parmer variable-area flow meter. The static pressure differential between the plenum and the ambient was measured with a Dwyer magnehelic differential pressure gage.

Fig. 4 illustrates the thermal and pressure characteristics of the heat sink Design #1 shown in Fig. 1. The air flow entered vertically into the plenum through one aperture while a second horizontal aperture was used to measure the static pressure in the plenum. Fig. 4 shows the thermal resistance (plotted on a logarithmic scale) and the static pressure in the plenum plotted as a function of air flow. The thermal resistance is defined as the temperature rise of the chip surface over the temperature of the incident air in the plenum, divided by the power dissipated in the chip. The experimental data are discrete points, whereas the continuous curves indicate the predictions of our theoretical model. The model did not take into account the small contributions to the total thermal resistance of the chip/heat sink interface and of the conduction resistance of the solid bottom part of the heat sink. Similarly, pressure losses in the plenum were not included.

We first checked that the measured temperature rise was repeatable and proportional to the power dissipated in the thermal die. The thermal resistances were measured at an approximately constant temperature rise of 25°C above ambient. Nine diode temperatures on the thermal die were con-

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 4
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

stantly monitored. Variations in temperature rise across the chip were less than ±5%. The high thermal conductivity of the silicon and the solid copper bottom of the heat sink reduce the gradients across the chip. The data shown in Fig. 4 represent the average chip temperature and have been corrected for heat leakage through the TAB leads into the printed circuit board. This correction is less than 5% at the higher flow rates (> 1 ft^3/min). It is useful though to remember that in any practical application secondary heat flow paths will exist and improve the overall thermal performance beyond that one of the heat sink itself.

As expected the thermal resistance decreases as the flow increases. For Design #1 a flow rate of 1 ft^3/min (472 cm^3/s) yields a thermal resistance of about 2.5 K/W; at 2 ft^3/min (944 cm^3/s) it is reduced to 1.7 K/W. As the chip surface area is only 0.66 cm^2, this performance is equivalent to less than 1.2 K · cm^2/W, an excellent performance for an air cooled heat sink. Furthermore, the required pressure drops are reasonable, on the order of 1.5 in of water (375 Pa) at 1 ft^3/min and 3 in of water (750 Pa) at 2 ft^3/min. The pressure flow relation is fairly linear in this flow range. The agreement between model and experiment for the pressure flow relation is not as good as for the thermal performance, especially at the higher flow rates. The discrepancy is due to pressure drops at the heat sink entrance and exit. These losses increase with the square of the air velocity or flow rate. In addition the measured static pressure depends on the exact plenum geometry, the measurement configuration and the flow direction. Our data indicate that the pressure developed at a given flow rate is a sensitive function of the exact flow distribution, whereas the thermal resistance depends primarily on the total air flow.

We made several additional measurements on this heat sink. Upon reversing the air flow direction the performance deteriorated by about 25%, as the air flow changed from counterflow to parallel flow. The pressure drop across the heat sink remained the same.

We also operated the heat sink in an impingement configuration without a plenum. The required pressure drop is thus delivered under the form of dynamic velocity head instead of mostly static pressure or a combination of static and dynamic pressures. This operating mode might provide some additional flexibility in the design and assembly of single-chip packages and multichip modules. We used a tube with an inner diameter of 4.7 mm. At flow rates above 1 ft^3/min (472 cm^3/s, equivalent to an average air velocity of about 20 m/s) the difference in performance between the two modes becomes very small. The penalty for taking off the plenum at these flow rates is less than 5% change in R_{th}. This means that above 1 ft^3/min the dynamic velocity head of the air is sufficient to push most of the air coming out of the tube through the heat sink. The impingement configuration could, therefore, turn out to be an attractive alternative to a closed plenum in certain applications.

The weight of a heat sink can be reduced by more than a factor of 3 by changing the material from copper to aluminum. This reduction in weight is accompanied by a reduced thermal performance because of the lower thermal conductivity of aluminum (k_f = 2.02 W/cm · K). Indeed, the conduction resistance of the solid bottom part of the heat sink is almost doubled in this way. However, this resistance constitutes only a small part of the total resistance. The only other place where the thermal conductivity of the heat sink material comes in is the fin efficiency, which is in turn only a very weak function of it. Equation (4) predicts that the fin efficiency will drop form about 96 to 92%. Hence, there is only a 4% decrease of the fin heat transfer performance. Measurements on an aluminum heat sink proved that the weight of this heat sink can be reduced by 70% while suffering only a 10% penalty in thermal performance at flow rates around 2 ft^3/min.

After characterizing the steady-state thermal performance of these heat sinks, we briefly investigated their transient behavior. When the air flow is switched off abruptly the temperature of the chip started rising approximately linearly with time, the time scale involved being on the order of several seconds, due to the relatively large thermal mass of the heat sink. This time span would provide a safe margin during which the cooling system failure could be sensed and the chip (or the computer) could be switched off to prevent permanent damage.

Finally, we tested the sound pressure level generated by the heat sink at a distance of 1 m. For these measurements an *A* weighting filter was used. This filter simulates the frequency response of the human ear. The noise level increases with increasing flow and velocity. The highest level was generated when the heat sink was connected into the plenum and operated in the counterflow mode. At 2 ft^3/min it reaches about 65 ± 2 dBA referenced to 0.0002 μbar, which compares (and, therefore, interferes) with normal conversation level. Using the impingement mode with a very small separation between the tube orifice and the top of the heat sink reduced the noise by about 5 dBA. Using a plenum in the parallel flow mode lowered the level by another 5 dBA. However, as we mentioned earlier, this mode is also thermally less efficient. The conclusion from these data is that measures will have to be taken to absorb the noise produced by these heat sinks, especially if a multitude of them are being used to cool multichip arrays. Adding sound absorbing material to the system or operating the cooling system in closed cycle are a few of the possible solutions to this problem. Using a different low pressure, low noise heat sink based on the same principles, such as Design #2, is another alternative as we will see in the next section.

B. Design #2

The prototype for Design #2 was fabricated from 7-oz copper foil which is about 0.25 mm thick and the channel width between fin tips was 0.50 mm, corresponding to one fin pair per 0.75 mm. The other dimensions are shown in Fig. 2. The total number of fins per heat sink is 26, amounting to an exposed fin surface area of about 64 cm^2. Equation (3) yields a fin efficiency of approximately 83% for these parameters.

All the thermal data on the heat sink Design #2 were taken with power transistors as heat sources. More specifi-

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 5
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

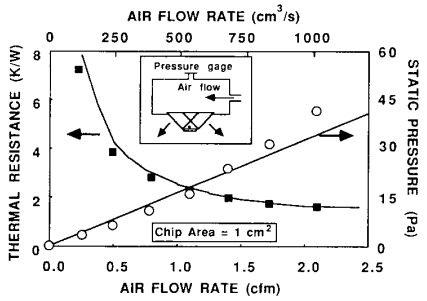

Fig. 5. Experimental performance of heat sink Design #2. Thermal resistance (on the left) and static pressure in the inlet plenum (on the right) are plotted versus air flow. Continuous lines indicate theoretical predictions based on our analytical model; discrete points show experimental results. The chip area is 1 cm^2. The insert shows a schematic cross section of the test setup. (1 in of water pressure = 249 Pa.)

cally we used MJE2955 transistors in plastic TO-127 packages which have a metal heat sinking surface. We first solder this surface to a 1-cm^2 copper plate which is 1.6 mm thick. This copper chip serves to spread the heat flux from the transistor and acts as an approximately uniform heat flux source. It is in turn soldered to the base of the heat sink. The temperature of the heat sink was monitored with a type T (copper-constantan) thermocouple located inside a hole drilled into the solder interface between the chip and the heat sink at the center of the heat sink base.

Fig. 5 shows the results of our thermal and pressure measurements. As schematically depicted on top of the graph, the air flow entered the plenum laterally and the pressure was measured on a vertical opening. The agreement of the theoretical predictions for the thermal resistance with the data is again rather good. At the lowest flow rates the experimental values seem to be a few percent lower than predicted. This is probably due to thermal leakage through the bias leads of the transistor, which is more noticeable when the heat sink cooling efficiency is lowered and has not been corrected in this case. At 2 ft^3/min (944 cm^3/s) a thermal resistance of about 1.6 K/W is achieved. This thermal performance is achieved at a very much reduced static pressure drop of about 0.17 in of water (42 Pa). As predicted the required pressure has thus been reduced roughly by a factor of 20 compared to Design #1. A pressure on the order of 0.2 in of water (50 Pa) can be achieved by small tubeaxial cooling fans.

Various measurements similar to those on Design #1 were made on this heat sink type. The thermal performance and pressure drop were unchanged when the air flow direction was changed in this case. This is probably due to the fact that we have a cross-flow situation, rather than a counterflow/parallel flow design. Again an aluminum heat sink would offer a significant weight reduction for a small performance penalty. The noise generated by a single heat sink (Design #2) was below the resolution of our sound pressure level meter (i.e., less than 50 dBA).

In summary, the two types of heat sinks described offer excellent cooling performance in a volume which is small when compared to traditional air cooled heat sinks. Design #1 with its triangular fins has a volume of less than 1 cm^3 and exhibits a thermal resistance of less than 2 K/W on a 0.66 cm^2 chip at 2 ft^3/min and 3 in of water. Design #2 trades off some of the smaller volume of Design #1 for lower pressure (0.17 in of water) and noise, while still achieving less than 2 K/W on a 1 cm^2 chip at 2 ft^3/min. Design #1 can be operated by a centrifugal blower, whereas Design #2 can work with standard tubeaxial fans. These two designs are examples of the application of laminar flow principles to air cooled heat sinks. The heat sink geometry can easily be varied to satisfy different system requirements.

V. Heat Sink Arrays

We extended our investigation to the cooling of multichip arrays in order to demonstrate experimentally the feasibility of air cooling a dense array of high powered chips with these novel heat sinks.

A. Design #1

The heat sinks were arranged in a regular 6 by 6 square array and connected into a flat plenum which was made out of sheet metal. The plenum was 15 cm^3 by 15 cm^3 and 12 mm high. Openings were punched into one of the plenum plates in order to attach the top of the heat sinks to them. The total height of the cooling package (heat sinks plus plenum) is about 2.5 cm. We made several plenum plates corresponding to different spacings of the heat sinks and yielding different chip-to-substrate coverages. The flow was metered with a Dwyer DS-100 flow sensor. The air entered the plenum laterally via a short adapter through an opening that measured approximately 15 cm by 1 cm. The static pressure in the plenum is measured with a Dwyer magnehelic differential pressure gage through an aperture which was made at a right angle to the incident flow. The air inlet temperature is measured with a type-T thermocouple inside the plenum.

In order to provide for maximum flexibility of the setup we did not use thermal dies as heat sources and temperature sensors: they need to be bonded to a board or substrate, a fact that would have made an array pitch change very time consuming and expensive. Instead we used power transistors as a heat source for these measurements. We chose a transistor in a TO-126 package with a metal heat sinking surface which was bonded to the die, and hence, to the collector of the transistor. These packages were soldered to the base of each heat sink via a copper chip which equalized the heat flux density. In order to avoid shorting all the collectors together we had to make the plenum plate out of an insulting material (FR4). All the transistors are connected in series and thus share the same current. The only disadvantage of this method is the spread in the current–voltage characteristics of these power transistors which makes it necessary to individually measure the power dissipated in each device.

In both cases we measured the temperature rise with type-T thermocouples inserted into the bottom solid part of the heat sinks close to the base. The chip/heat sink interface resistance is, therefore, not included in these measurements as opposed to the previous data on individual heat sinks. The thermocouples were connected to a Fluke 2285B data logger which scans all the inputs. The data logger was slaved to a

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 6
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

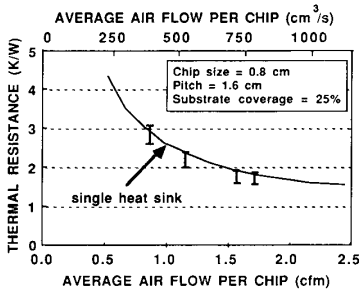

Fig. 6. Thermal performance of 6 by 6 array of heat sinks of Design #1 at 25% chip-to-substrate coverage. The average thermal resistance over the array is plotted versus average air flow per heat sink. The performance of a single heat sink is indicated by a continuous line. The bars indicate ± 1 standard deviation.

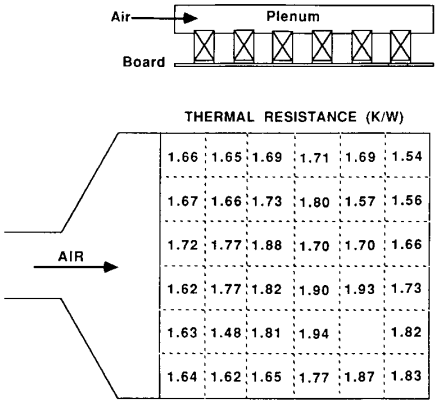

Fig. 7. Thermal resistance distribution of a 6 by 6 array of heat sinks of Design #1 at a total air flow of 57 ft³/min and a chip-to-substrate coverage of 25%. The thermal resistance of the heat sinks is fairly uniform over the array. The average thermal resistance is 1.72 ± 0.12 K/W. There is no crowding effect at the center of the array. A cross section of the experimental setup is shown schematically at the top of the figure.

personal computer which averaged and stored the data. The assembled multichip cooling array was placed on a board with the wiring pulled through the board below each heat sink so as not to encumber the air flow exiting from the heat sinks.

1) Experimental Performance: All the data described in this section were taken in the counterflow mode on the 6 by 6 array of chips and heat sinks. All the heat sinks were mounted with their fins and channels arranged in parallel. Before final assembly of the array, each heat sink/heat source unit was tested individually. Their performance agreed well with the earlier measurements using thermal dies, the variations from unit to unit not exceeding 10%.

We will describe here a series of measurements on the array with a 16-mm center-to-center pitch corresponding to a substrate coverage by the chips of about 25%. The results obtained at different flow rates are summarized in Fig. 6. We plot the thermal resistance distribution of the individual heat sinks versus average air flow per heat sink. The thermal resistance is defined here as the temperature rise of the bottom of the heat sink over ambient per unit of power dissipated in the transistor. The vertical bars on the graph correspond to the average measured thermal resistance plus or minus one standard deviation. The average air flow used on the horizontal axis is equal to the measured total air flow divided by the number of heat sinks (36). The uncertainty in flow is about 3% of full scale or roughly 0.1 ft³/min per heat sink.

The continuous line in Fig. 6 indicates the data obtained for a single heat sink unit as shown in Fig. 4. Those data included the chip/heat sink interface resistance. As can be observed in Fig. 6, the average resistance for the array agrees quite well with the single heat sink data within the experimental uncertainty. The average thermal resistance for the array is consistently a little lower than the resistance of the individual heat sink measured previously. The difference is approximately 0.2 K/W. This is probably due to the fact that for the single heat sink measurements the chip/heat sink interface resistance was included in the data whereas it is not for the array. The standard deviation is always less than plus or minus 10%, indicating that there is little spread in both the experimental measurements and the heat sink performances.

Fig. 7 takes the data set at the second highest flow rate from Fig. 6 and shows the results in detail. The air inlet adapter is also shown schematically. The air flows through a 3.8-cm diameter tube and expands into the plenum. The total air flow is 57 ft³/min or 1.6 ft³/min per heat sink. The static pressure in the plenum was measured to be 3.7 in of water. The matrix in Fig. 7 represents the array of heat sinks. The numbers indicate the thermal resistance of each unit in Kelvins per watt. (One power transistor stopped functioning during the data taking process and the corresponding value in Fig. 7 is missing). The thermal resistances are spread over a fairly narrow range (1.6–1.8 K/W). The average thermal resistance was 1.72 K/W with a standard deviation of 0.12 K/W. The thermal resistance of the heat sinks at the center or in the center rows is not significantly higher than the overall average. It agrees with the overall average within the experimental uncertainties.

We increased the chip-to-substrate coverage by decreasing the spacing between heat sinks and chips to 6 mm. This is equivalent to a 33% substrate coverage by chips. At a total flow rate of 60 ft³/min the pressure measured in the plenum was 4.2 in of water. The average thermal resistance was measured in this case to be 1.67 K/W with a standard deviation of 0.18 K/W. This is in very good agreement with the previous array data as well as the individual heat sink results. Even for this dense array, there is no sign of a systematic crowding effect at the center of the array.

These data show that we can achieve uniform thermal performance over a 6 by 6 array of heat sinks with a thermal resistance of 1.7 K/W at a total flow rate of about 60 ft³/min and chip-to-substrate coverages of up to 33%. At the module level this corresponds to a thermal resistivity of 3.4 K · cm²/W. It proves that the high cooling efficiency of these air cooled heat sinks can be used for multichip arrays featuring a very high substrate coverage by semiconductor components. Given this performance and assuming an interface resistance of 0.3 K/W, each chip could dissipate up to 30 W and only

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 7
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

experience a temperature rise above ambient of about 60°C. If operated at room temperature this would raise the chip temperature to about 85°C, still a conservative operating temperature for semiconductors. This excellent cooling performance enables us to cool a total power in excess of a kilowatt from the 6 by 6 multichip substrate we are simulating in our experiments. No existing air cooled multichip substrate can rival this performance.

One concern about Design #1 that we have pointed out in the previous section is the noise generation. The sound pressure level measured for this 6 by 6 array at a total flow rate around 60 ft^3/min and at a distance of 1 m is about 78 ± 2 dBA. This is about 13 dBA higher than for a single unit. (A factor of 36 corresponds theoretically to 15.6 dB.) This noise level will have to be attenuated if this cooling system is to be operated in an office environment or a computer room. Sound absorbing material and/or closed cycle operation are two of the possibilities.

B. Design #2

After developing the low pressure/low noise heat sink Design #2, we also applied this design to the cooling of multichip arrays. At a pressure of 0.2 in of water a 12-cm square and 3.8-cm thick tubeaxial fan (e.g., Rotron Muffin XL DC) can deliver up to 50 ft^3/min enough to cool 25 chips. A 16-cm square and 3.8-cm thick fan (e.g., ETRI model 148 VK) will deliver about 200 ft^3/min under the same conditions. We decided to test a 5 by 5 array of chips and heat sinks in an array with 16% chip-to-substrate coverage, in combination with a closely coupled 12-cm square fan.

The heat sinks used for this array study were identical to the prototype described in the previous section for Design #2. Before the final assembly of the array all the heat sink/chip units were individually calibrated for pressure and thermal resistance. They exhibited excellent uniformity in performance. The average thermal resistance at an air flow rate of 2.1 ft^3/min was 1.59 K/W with a maximum deviation of only 0.05 K/W. The average pressure required at this flow rate was 0.17 in of water with a maximum deviation of 0.015 in. Hence, performance variations from unit to unit could be neglected in the later analysis of the array results.

The power transistors and the 1-cm^2 copper chips were arranged in a regular square array with a center-to-center pitch of 2.5 cm. In order to allow for such tight packing, the heat sinks were arranged in a "basket weave" pattern so that the direction of the fins of two neighboring heat sinks were always perpendicular (see Fig. 8). The heat sinks were connected to the bottom plate of a flat plenum which was 15-cm on a side and 1.2 cm thick. The 3.8-cm thick square Muffin XL DC fan which measured 12 cm on a side was mounted to the top plate of the plenum. It is a brushless dc fan and operates on a 12-V nominal power supply. In addition to the heat sink temperatures we monitored the temperatures of the inlet air, the air in the plenum, the outlet air, as well as the room temperature and the static pressure in the plenum. The thermocouples were connected to the Fluke data logger slaved to a personal computer. The complete array was placed on a board. All the transistors were connected in series and the thermal results were corrected for the slightly different voltages across each one of them.

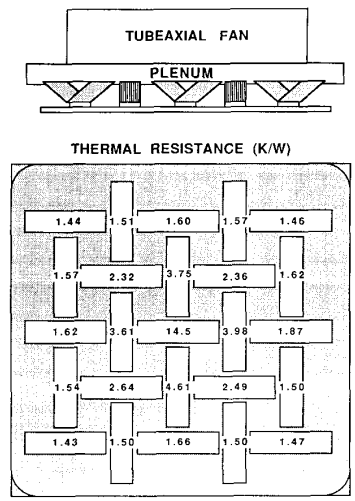

Fig. 8. Thermal resistance distribution of a 5 by 5 array of heat sinks of Design #2 in the push-mode at a chip-to-substrate coverage of 25%. The experimental setup is shown schematically at the top of the figure. The static pressure in the plenum is 0.2 in of water. The centrifugal force on the air exiting the fan forces it through the outer heat sinks, thus creating a hot spot at the center of the array. A cross section of the experimental setup is shown schematically at the top of the figure.

1) Experimental Performance: A first set of measurements was taken with the fan pushing the air through the plenum into the heat sinks. Fig. 8 shows the thermal resistances obtained as a function of position and indicates the orientation of the heat sinks. These data show that in the push-mode the air flow tends to mainly pass through the outer heat sinks. This results in an insufficient cooling performance at the center of the array where the thermal resistance reaches up to 14.5 K/W. This is due to the centrifugal force on the rotating air at the fan exit which produces a conical air flow distribution. Note though how symmetrical the resistance distribution is; it only depends on the distance of the heat sink from the center. The static pressure measured in the plenum is 0.2 in of water (50 Pa).

Several attempts were made to even out the air flow distribution in this push-mode. We tried both adding vanes in the plenum and obstructing the inlet of the outer heat sinks with wire mesh. Although significant improvements were made with these methods, the thermal resistances still varied by almost up to a factor of two over the array (from 1.48 K/W at the edge to 2.64 K/W at the center).

We then reversed the air flow direction. The results obtained in the pull-mode are displayed in Fig. 9. The air flow distribution has become much more uniform. There is no systematic dependence of flow and thermal resistance on the distance of the heat sink from the center. There remains, however, a slight difference between the heat sinks whose channels are tangential and perpendicular to the fan rotation. The average thermal resistance is 1.86 K/W with a maximum of 2.19 K/W and a minimum of 1.66 K/W. The average for the outer heat sinks tangential to the rotation is 2.02 K/W, whereas the resistance for the perpendicular ones

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 8
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

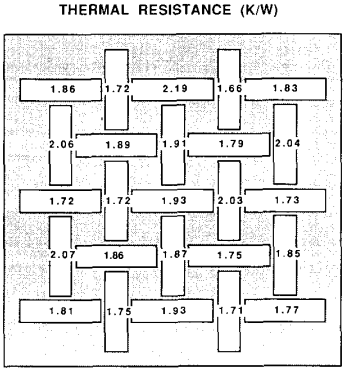

Fig. 9. Thermal resistance distribution of a 5 by 5 array of heat sinks of Design #2 in the push-mode at a chip-to-substrate coverage of 25%. The static pressure in the plenum is 0.2 in of water. The thermal performance is uniform over the array with an average thermal resistance of 1.86 K/W. This data set was taken with a total power dissipation of 612 W in the 25 chip array.

is 1.72 K/W. The pressure in the plenum is 0.2 in of water (50 Pa).

These measurements in the pull mode were taken with a total power dissipation of 612 W that is about 25 W per chip or heat sink. The power used by the fan was only 10.3 W. This is equivalent to a ratio of power used to power removed of 1.7%. This array of heat sinks (Design #2) is an efficient cooling system. At the module level, this performance corresponds to a thermal resistivity of 11.6 K · cm^2/W. In terms of volumetric efficiency, the compact setup is able to cool about 500 W per liter (1000 cm^3), this volume including the fan. In addition the noise generated by the cooling system is only 56 ± 2 dBA at a distance of 1 m, less than the background noise generated in our laboratory by the power supplies and the data logger.

It is important to note that the data presented above applies to the thermal performance of the heat sinks alone. In any system the heat sinks would need to be attached to the chips with low thermal resistance interfaces. A recent paper [6] describes the thermal resistance obtained with carefully prepared samples and thermally conductive epoxies. The results ranged from about 0.12 to 0.7 with an average performance of less than 0.3 K · cm^2/W. If this average thermal interface performance is added to the heat sink resistance, a total resistance of 2.2 K/W results. The anticipated temperature rise of a 1-cm^2 chip dissipating 25 W is then 55°C above the inlet air temperature, provided the heat sink can be directly attached to the chip. A 25 chip module of these chip/heat sink assemblies would then dissipate 625 W and remain below commonly accepted temperature limits for commercial semiconductor applications. By comparison typical liquid cooled modules [7], [8] handle similar power levels within about the same volume as this air cooled module; the water pump being excluded.

VI. Conclusion

A novel, high performance, and compact air cooling technology has been described. The heat sinks use laminar flow through narrow channels formed by two interleaved set of fins, to achieve small thermal resistances at low pressure drops. They duct the air from a vertical top entrance to a horizontal exit. This design enables the cooling of multichip substrates by connecting the heat sinks in parallel to a plenum which supplies the cool air. Consequently the heated exhaust air does not affect any other heat sinks.

In particular, heat sinks of Design #1 have a thermal resistance of less than 1.7 K/W of power dissipated in the chip at an air flow rate of 2 ft^3/min (944 cm^3/s) and a pressure drop of 3 in of water (750 Pa). The heat sink volume is less than 1 cm^3. In an array, up to 33% chip-to-substrate coverage by chips has been achieved without a performance degradation. On the other hand, the heat sinks of Design #2 offer a similar thermal performance at a much reduced pressure of 0.2 in of water (50 Pa). Pressure requirements have been reduced by sacrificing compactness, thereby resulting in a somewhat larger heat sink of about 3 cm^3 which limits the chip-to-substrate coverage to about 20%.

With both types of heat sinks we are able to remove in excess of 600 W from a multichip array. This kind of performance was until now only achieved in liquid cooled modules [9]. We have demonstrated that an array of air cooled heat sinks (Design #2) can dissipate this amount of heat with a quiet tubeaxial fan which only uses 10 W of power. An array of heat sinks (Design #1) enables us to cool even more power (up to a kilowatt) at higher chip-to-substrate coverages with centrifugal air blowers. These two heat sink designs embody the same principles: the first one emphasizes compactness whereas the second one focuses on achieving a low pressure loss and low noise. One could conceive of many intermediate designs which would tradeoff heat sink volume against pressure requirements according to the needs of specific applications.

In addition, it is important in a multichip system to maintain all devices within a small temperature range to ensure good noise margins and tolerances. therefore, low dissipation level chips require a small cooling capability than high dissipation level chips. This goal can be achieved easily by these heat sinks because they are modular: they can be scaled to smaller or larger chip power by removing or adding fin pairs and still be optimized for the same operating pressure. Usually smaller chips dissipate less heat than larger chips, and thereby require less cooling capability to maintain the junction temperatures to within specification. Conversely larger, high dissipation level chips require a higher cooling capability to maintain the junction temperature to within the same specification. So an almost automatic scaling of cooling capability can be achieved with these air cooled heat sinks by subtracting or adding fin pairs.

Finally, we want to point out a significant difference between the cooling concepts described in this report and existing thermal conduction modules. In those systems the chips are usually connected through a series of rigid, and at least one, compliant interfaces to a cold plate which takes out the heat, either by water or by air, at the module level. On the other hand, the technology described here assumes that an individual heat sink is connected to each chip. This enables us not to have any compliant, and thereby rather poorly

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 9
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.

conducting, interface between the chips and the heat sinks where the heat flux densities are highest. The compliancy can be designed into the plenum. Also, bringing the heat sinks closer to the chips reduces the number of interfaces, and therefore, the total thermal interface resistance. These advantages will become even more important as the heat flux densities at the chip level will increase.

In summary, using novel, compact heat sink designs, we have demonstrated the feasibility of high performance air cooling, which extends the applicability of air cooling to power levels generally considered to lie in the domain of liquid cooling.

ACKNOWLEDGMENT

Special thanks go to Dr. Richard Nelson whose innumerable comments and suggestions during this work were extremely valuable. The authors to thank Sperry (Unisys) for letting us use their thermal die, Eastman-Kodak for manufacturing a series of heat sinks (Design #1), and all the MCC shareholder companies for their interest and feedback during this project. The heat sinks described in this paper and methods of fabricating them are covered by U.S. Patents 4 777 560 and 4 833 766.

REFERENCES

[1] D. B. Tuckerman and R. F. W. Pease, "High-performance heat sinking for VLSI", *IEEE Electron. Device Lett.*, vol. EDL-2, pp. 126–129, May 1981.

[2] D. B. Tuckerman, "Heat transfer microstructures for integrated circuits," Ph.D. dissertation, Stanford Univ., Feb. 1984.

[3] N. Goldberg, "Narrow channel forced air heat sink," *IEEE Trans. Comp. Hybrids, Manuf. Tech.*, vol. CHMT-7, pp. 154–159, Mar. 1984.

[4] M. Mahalingam, "Thermal management in semiconductor device packaging," *Proc. IEEE*, vol.9, pp. 1396–1404, Sep. 1985.

[5] W. M. Kays and M. E. Crawford, *Convective Heat and Mass Transfer* New York: McGraw-Hill, 1964, p. 14.

[6] G. P. Peterson and L. S. Fletcher, "Thermal resistance of silicon chip bonding materials," in *Proc. Int. Symp. on Cooling Technology for Electronic Equipment*, Mar. 1987, pp. 438–448.

[7] R. G. Biskeborn, J. L. Horvath, and E. B. Hultmark, "Integral cap heat sink assembly for the IBM 4381 module," in *Proc. 1984 Int. Electr. Pack. Soc. Conf.*, 1984, pp. 468–474.

[8] S. Oktay and H. C. Kammerer, "A conduction cooled module for high performance LSI devices," *IBM J. Res. Develop.*, vol. 26, no. 1, pp. 55–66, 1982.

[9] A. Bar-Cohen, "Thermal management of air and liquid cooled multichip modules," *IEEE Trans. Comp. Hybrids, Manuf. Tech.*, vol. CHMT-10, pp. 159–175, June 1987.

HALLIBURTON EXHIBIT 1034
Halliburton Energy Services, Inc. v. U.S. Well Services, LLC, IPR2023-00558, Page 10
Authorized licensed use limited to: Robert Durham. Downloaded on January 18,2023 at 03:19:20 UTC from IEEE Xplore. Restrictions apply.