CMOS-Compatible Fluidic Chip Cooling Using Buried Channel Technology

by

Daniel M. Gaugel

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in

Electrical & Computer Engineering

Carnegie Mellon University
Pittsburgh, Pennsylvania
USA

Advisor: Professor Kaigham J. Gabriel

Acknowledgements

The author would like to thank the entire MEMS Laboratory at Carnegie Mellon University, Ken Gabriel, Mary Moore, Drew Danielson, Elaine Lawrence, Lynn Philibin, the staff of the CMU Nanofabrication Facility (Chris Bowman, Carsen Kline, and Tim Fisher), the staff of the ECE Machine Shop (Jim Schubert, Dave Belotti, and Steve Hoffman), Hasnain Lakdawala, Xu Zhu, George Lopez, John Neumann, Tamal Mukherjee, Keith Rebello, Kyle Lebouitz, Suresh Santhanam, Joe Suhan, Jeff Zahn, Lars Erdmann, and Chi-Fu Wu for their assistance and advice throughout the duration of this project.

Table of Contents

I.	Introduction	1
II.	Background	2
III.	Fabrication	5
IV.	Thermal Analysis	
	i. Simulation Model	14
	ii. Experimental Setup	20
	iii. Modeled, Experimental Results	21
V.	CMOS Compatibility	27
VI.	Conclusions	29
VII.	References	32
VIII.	Appendix A (Processing recipe)	33
IX.	Appendix B (Data acquisition system photographs)	34

I. Introduction

Electronic system power dissipation has increased exponentially since 1980, driven primarily by the increase in transistor switching frequency. While power per gate continues to fall (0.64-0.15 μW/ MHz from 1995-99) due to voltage reduction, overall power dissipation has increased in virtually every electronic product due to increasing transistor density and frequency [13]. For example, Intel microprocessor thermal design power has risen from 40.0 W for the Pentium II [14] to 75.3 W for the current Pentium 4 [15]. Should this trend continue to rise, forced air convection will no longer offer adequate electronic thermal management and Moore's Law will come to an end. Microfluidic convective heat transfer offers an alternative to the conventional heat sink air-cooling systems. The superior fluid properties of water compared to air (approximately 4 and 1,000 times greater specific heat and density, respectively [9]) offer more efficient cooling within a compact volume. Designs featuring fluidic heat transfer integrated into the silicon substrate offer high cooling capacity due to the high heat transfer characteristics of micron-order diameter cooling channels. Significant fabrication challenges are present with this approach however, including the need for system level and CMOS processing integration.

We present a CMOS-compatible method of fabricating buried channels within bulk single-crystal silicon serving as the substrate of CMOS circuitry. The channels were designed to provide a compact, forced convection heat transfer liquid cooling approach to microprocessor and integrated circuit thermal management. The baseline fabrication process consumes less than 8 % of surface area and serves as a foundation for further area reduction, potentially leading to incorporation of buried-channel cooling with CMOS electronics. A microchannel heat transfer solution was designed, simulated, processed, and tested with a liquid cooling data acquisition system. Compatibility with CMOS was characterized by evaluating the post-process transistor

performance of a simple oscillator circuit.

II. Background

The goal of integrated forced convection fluidic microprocessor cooling has provided a fertile area of research into microchannel development.

D.B. Tuckerman [1] achieved heat removal between 181-790 W/cm^2 with high aspect ratio rectangular channels in 1981. Internal heat transfer enhancement through the reduction of thermal resistance between fluid and substrate serves as the primary objective of this research. 50 μm wide by 300 μm deep vertical channels were processed within a silicon die with an anisotropic KOH (potassium hydroxide) wet etch. The channels were spaced 50 μm apart over the 1 cm^2 surface area. The rectangular channel height/width aspect ratio ranged from 5-6 and offered increased surface area with relatively small hydraulic diameters. A cover plate anodically bonded to the fin array contained the fluid, with heat provided via thin film WSi_2 resistors on the chip surface. Experiments conducted with water at flow rates from 4.7-8.6 mL/min yielded fluid pressures from 15-31 psig. The major disadvantage to this approach as a microprocessor cooling solution lies in the wafer bonding, which will add processing steps and decrease yield.

Significant work on the development of optimum heat exchanger design has been reported. By taking advantage of the high heat transfer coefficients of sub-cooled boiling, Bergles et al [2] explored the design parameters to achieve optimum heat exchanger design as well as summarized seminal two-phase heat exchange microchannel research. The authors present a number of design recommendations for surface heat source, rectangular channel bulk silicon substrate cooling die formations. These include locating the onset of incipient boiling (requiring microscopic viewing of exiting fluid), establishing uniform flow distribution through the channels, including header orifices to reduce the effect of pressure drop due to channel wall friction, and spacing channel location for uniform substrate heat distribution.

de Boer [3] et al have reported buried channel fabrication using a single mask process without the need for wafer bonding, eliminating alignment error. The buried channel technology

(BCT) involves a multi-step process in which access trenches are deep etched into the silicon substrate. Microchannels are formed by isotropically etching the exposed silicon on the bottom of the trench with the use of masked sidewalls. Two sidewall-masking methods (thermal SiO_2 and LPCVD Si_xN_y) successfully protected the trench sidewalls from the remaining etching steps. An anisotropic etch cleared the SiO_2 and Si_xN_y at the base of the trenches, allowing channel formation from isotropic silicon etching. A number of sidewall protection geometries are presented, as well as a method of sealing and releasing the channels (Figure 1). Due to the high temperature oxide growth, integration of CMOS electronics must be added after the formation of the channels.

Joo et al [4] present a single-sided, low temperature, single mask microchannel fabrication process. Rectangular cavities are formed by spinning a 10-µm layer of photoresist onto a seed layer of nickel and chromium. After development, nickel-electroplating formed channel walls in the areas not covered with the photoresist pattern. A subsequent electroplating step after stripping the photoresist sealed the openings, leaving an enclosed microchannel on top of the silicon substrate (Figure 2). The ability to grow the channels on top of electronics, with thermal resistance considerably smaller than bonded heat sinks, present the major advantage of

this approach. The overall dimensions of the formed channels were 10 μm in width and 8 μm high, with wall thickness of 10 μm.

III. Fabrication

The primary goal of this research effort lies in the development of a CMOS compatible microchannel baseline fabrication process with the following features: 1) minimal mask area to facilitate integration with electronics 2) channels located sufficiently deep into the substrate to preserve electronics functionality 3) low temperature sidewall passivation (non-SiO_2 growth).

Because polymers are resistant to etching inXeF_2, we set out to create deep channels using the passivation growth of C_4F_8. Anisotropic trenches could be etched using the Bosch etch/passivation ICP (inductively coupled plasma) process, followed by an extended polymer deposition. A directional removal of the polymer would expose the bulk silicon at the bottom of the trench, allowing a spherical, bulbous pattern to be produced at the base using isotropic silicon XeF_2 etching. A series of these features patterned optimally close together would produce connected spherical cavities resulting in a scalloped-wall buried sub-circuit channel. Due to the size of the "access" mask area relative to the channel diameter, we could seal the access ports using an extended polymer deposition. Figure 3 presents a conceptual drawing of the target microchannel process.

While the eventual application for the microchannel process lies in integration with CMOS electronics, silicon wafer samples were used to establish the baseline process without the need for expensive and long-lead time microelectronics fabrication. A series of 4" masks were created

with various access-hole feature sizes and configurations to determine the final fabrication baseline process.

We developed an eleven (11) step process to create buried microchannels in bulk silicon. The procedure is presented schematically in Figure 4, with Appendix A listing the process steps in greater detail, including plasma pressures and flow rates.

Figure 4: Processing recipe schematic

The processing steps presented in Figure 4 are as follows:

Steps 1-2) Photolithography, mask selection

Channel access patterns were formed on 1.5 cm x 1.5 cm SiO_2-coated silicon chips with a single spin, exposure, and development procedure. AZ 4620 photoresist was selected due to the relatively thick film (~13 μm) required to maintain the mask pattern. This thickness restricted the minimum feature sizes to just below 10 μm, but these feature sizes were later determined to be sufficient for channel creation.

A sufficiently stable mask was needed to maintain the structural integrity of the pattern. The low etch selectivity of photoresist to silicon (50-75/1) led to the exploration of stronger films capable of withstanding the subsequent processing steps. We attempted to take advantage of the high aluminum/silicon etch selectivity by patterning a single layer CMOS mask (0.44 μm Al + 1.25 μm SiO_2). This was eliminated when excessive aluminum-fluoride reaction led to polymer redeposition and micromasking during standard silicon etching (cross section presented in Figure 5). We selected a 4 μm SiO_2 etch mask due to the successful results achieved with 1.25 μm SiO_2 samples.

Step 3) Dry RIE SiO_2 etch

A dry SiO_2 etch was then used to develop the channel mask. Non-uniform SiO_2 thickness and over-etching caused the mask features to expand during this step, but recipe optimization controlled the growth to 1-2 μm at most.

Step 4) Anisotropic etch

Next, the access trenches were created using the Bosch DRIE (deep reactive ion etch) process. The trench depth dictated the maximum channel diameter, due to the isotropic expansion of the channel. The trenches were etched 110-150 μm deep to leave a minimum of 50 μm of silicon between the top edge of the channel and the die surface.

Step 5) Isotropic RIE etch

An isotropic etch undercut the walls of the trenches 4-5 μm underneath the SiO_2 mask (cleaved cross section presented in Figure 6). We added this step to protect the upper portion of the trenches during subsequent processing. The upper sections of the trenches are vulnerable in the final channel etch (Step 9) due to the weakening of the sidewall protection during passivation removal (Step 7). We hypothesized that the polymer along the top edges of the trenches weakened due to the sidewalls lying inline to the path of the plasma ions. The plasma would remove the polymer at the base of the trench, but weaken the protection along the upper sidewalls in the process. During the XeF_2 etch, the weakened polymer mask experienced undesirable

silicon sidewall etching (Figure 7), which diluted the process gases and increased the sacrificial mask area.

Two methods were tested to alleviate the upper trench degradation. We attempted an undercut demonstrated by de Boer [3] prior to the DRIE trench etch (Figure 8). After passivation, the undercut area would thus be protected from the plasma by the mask. We witnessed sidewall etching results similar to Figure 7. The second method, an isotropic etch between DRIE and polymer deposition, was the most successful. The undercut mask blocked the subsequent plasma ions from the upper trench area, protecting the polymer layer.

Steps 6-7) Polymer deposition, bottom trench polymer removal

Next, an 8.5-long minute deposition coated the trenches with a ~0.8-1.0 μm of C_4F_8 polymer. We then explored four methods to remove the polymer at the base of the trenches. The first attempt involved an etch cycle of SF_6 and C_4F_8 similar to the Bosch process, with etch/passivation steps of 20-30 seconds. The second method involved a similar cycle, substituting O_2 for SF_6 [18]. Both procedures failed to build a polymer sidewall mask that would survive the remaining steps of the recipe. This is most likely due to the isotropic degradation of the polymer with SF_6 and O_2. We thus attempted a similar etch/passivation cycle with Ar and C_4F_8. While the Ar did not react with the polymer, the low power (30 W) plasma produced scattering of ions, which weakened the polymer buildup along the sidewalls (Figure 9). Stronger argon plasmas were therefore needed to produce anisotropic ion milling. Due to the control restrictions our DRIE machine (STS) recipe, high power argon plasma could not be produced using the etch/passivate cycle method. We found the most successful procedure for this step to be an extended passivation followed immediately by an extended high power Ar plasma [19]. Increasing the platen power in the STS, which increased the impinging ion energy into the polymer film, optimized the process. We found additional improvements by lowering the plasma pressure, which increased the mean free path of the argon ions. A ramping procedure was needed to build the argon plasma from 30 W platen power, 7 mT pressure at the onset to the optimal 150 W and ~1.5 mT conditions.

The Ar passivation removal step was difficult to control and reproducibility was poor. This is most likely attributed to polymer deposition rate variations and curing. For an extended 8.5 minute-deposition, measured polymer thickness varied up to 20%. The polymer also cured under vacuum, causing a range of argon milling times to effectively remove the polymer at the base of the trench. Due to curing, we were unable to sufficiently clean the chamber before the remaining STS recipe steps were completed. We observed that polymer removal at the base of the trenches was best determined by aborting the process at the point where the bias voltage of

the platen would rapidly increase. We successfully estimated this point by recording the bias voltage of a photoresist film silicon wafer and comparing to the ion milling process conditions.

Step 8) RIE etch

A 20 second-SF_6 etch ensured the exposure of silicon at the base of the trench [17]. Non-uniformity observed in the channel diameters led to the inclusion of this step. Without this step, sputtered polymer particles present at the base of the etch trench caused significant masking during XeF_2 etching.

Step 9) XeF_2 etch

After a 30 minute-200 ^{o}C bake to cure the sidewall polymer [20], the channels were formed at the base of the trench with an isotropic XeF_2 etching process. Channel diameter was controlled by adjusting the etch duration. Polymer delamination was evident at the base of the trench walls, but the upper wall integrity remained intact. Figures 10-15 present cross sectional and surface views of three samples processed for this analysis.

Steps 10-11) Oxygen polymer clearing, sealing

Oxygen plasma cleared the remaining polymer from the walls of the trenches.

At this point, there were two options for sealing the microchannels. An extended polymer deposition could seal the access trenches at the mask surface if the access trenches were reasonably small. Due to the 22-25 μm mask widths of the samples, a wafer bonding process was implemented. We sealed the samples by bonding silicon sections of similar size to the openings of the channels with a high temperature, water resistant epoxy.

With the process recipe yielding repeatable results, the final challenge lay in the determination of the smallest successful feature size. We established the process with 300 μm

open areas and adjusted parameters for smaller trench areas. A 20 μm-wide trench proved to be the smallest access feature size achieved during this work. Attempts to process channels through smaller access holes were unsuccessful due a number of possible reasons, including insufficient deposition growth and an inability to clear the polymer at the base of the holes.

An additional trend occurred during the XeF_2 etch step that may qualify the worthiness of the C_4F_8 polymer as an etch mask. Delamination was apparent on every test sample, with an undesirable undercut occurring at the base of sidewalls. Figure 16 presents an open area delamination indicative of this occurrence. The polymer may be permeable at weakened or thin mask areas, which present the need for further processing analysis to ensure uniform, thick sidewalls on the access trenches.

IV. Thermal Analysis

i. Simulation Model

The analysis of the heat sink capability of the processed buried channel samples presents the second research objective. Performance models based on fundamental heat transfer principles were created to examine the effects of channel geometry, flow rate, and chip substrate temperature on heat transfer performance. While significant cooling advantages exist by operating the system in dual phase (heat transfer due to change from liquid-gaseous state under constant temperature), this analysis will focus on single-phase, sensible (change in coolant temperature) heating of the coolant. The following analysis is derived from Chapters 3 and 8 of [9].

An energy balance between the heat source (microprocessor), and the heat sink (fluid passing through the channels) determines how the mean coolant temperature varies with position along the channel. This balance also relates the heat transfer rate to the coolant inlet and outlet temperatures. For the entire channel, irrespective of fluid transport properties or surface conditions, this balance is expressed as:

$$Q''_{chip} A_{chip} = \dot{m} cp \Delta T_{coolant} \text{ (W)}$$

with Q''_{chip} = thermal heat load density (W/cm^2), A_{chip} = chip area (cm^2), $\dot{m}$ = coolant mass flow (kg/s), cp = coolant specific heat (J/kg K), and $\Delta T_{coolant} = (T_{coolant,\ exit} - T_{coolant,\ entrance})$.

It is necessary to relate the thermal heat density produced by the chip to the heat flux from the substrate into the individual channels. The chip surface area relative to the cooling channel surface area is therefore presented:

$$\mathrm{A_{chip} = Length_{chip} Width_{chip}\ (cm^2)}$$

$$\mathrm{A_{channels} = (\#channels) Perimeter_{channel} Length_{channel}\ (cm^2)}$$

$$Q''_{chip} A_{chip} = Q''_{channels} A_{channels}$$

The quantification of the forced convection between the chip substrate and the coolant is expressed through Newton's Law of Cooling for internal flow:

$$Q''_{channels} = h(T_{substrate} - T_{m,coolant}) \quad \text{(W/cm}^2\text{)}$$

With $T_{substrate}$ = substrate temperature (K), $T_{m,coolant}$ = mean coolant temperature (K), and h = local internal heat transfer coefficient (W/ m^2 K), a parameter which encompasses all of the convection heat transfer factors. These include boundary layer conditions (reliant on channel geometry), and fluid thermal and transport properties. The convection heat flux is termed as positive when heat is transferred from the surface to the fluid. With heat transfer occurring across the channel, the mean coolant temperature will increase in the direction of the moving fluid.

Newton's Law of Cooling applies to the case of a constant surface heat flux, with the surface and fluid temperatures changing along the path of the channel. A variation exists for the constant surface temperature, variable heat flux scenario:

$$q_{channels} = \bar{h} A_{channels} \Delta T_{lm} \quad \text{(W)}$$

with $\bar{h}$ = average internal heat transfer coefficient and T_{lm} = log mean temperature difference, which expresses the average temperature difference along the length of the channel. This is defined as:

$$\Delta T_{lm} = \frac{\Delta T_o - \Delta T_i}{\ln(\Delta T_o / \Delta T_i)} \quad \text{(K)}$$

with $T_o = T_{surface} - T_{mean\ coolant\ outlet}$ and $T_i = T_{surface} - T_{mean\ coolant\ inlet}$. While a constant heat flux would apply for the microprocessor model, the constant temperature model will be used for the following analysis. The heat sink/ temperature controller used to simulate transistor thermal heat generation will produce experimental results more accurately represented by the constant surface temperature model. Figure 17 presents a simulated temperature profile along the length of the channel for each case.

Figure 17: Axial temperature variations for heat transfer in a channel: a) constant surface heat flux, b) constant surface temperature. Derived from [9].

In order to estimate the internal flow heat transfer coefficient and pumping power, we analyzed the thermal and transport properties of the fluid. For the transport properties, the Reynolds # represents the ratio of the inertial to viscous forces in the velocity boundary layer. For internal flow through a circular channel, this parameter is expressed as:

$$\mathrm{Re}\# = \frac{\rho V L}{\mu}$$

with μ = fluid absolute viscosity (N s/m^2), ρ = fluid density (kg/m^3), V = mean fluid velocity (m/s), L = channel length (m). For laminar, or highly ordered and streamlined flow, Re # < 2,300.

The Nusselt # provides a measure of the convection heat transfer occurring at the surface. This is expressed as a dimensionless temperature gradient:

$$Nu\# = \frac{\partial \dfrac{(T - T_{surface})}{(T\infty - T_{surface})}}{\partial \dfrac{y}{L}}$$

with T = fluid temperature within the boundary layer, and T_∞ = fluid constant temperature outside of the thermal boundary layer. For constant surface temperature and laminar, fully developed conditions, the Nusselt # is constant and independent of channel location:

$$Nu\# = \frac{hD}{k_{fluid}} = 3.66$$

with D = channel diameter (m) and k_{fluid} = fluid thermal conductivity (W/ m K). The heat transfer coefficient and channel diameter are therefore inversely proportional, with greater heat transfer occurring at smaller diameters.

Due to the non-uniform profiles of the buried channels, a hydraulic, or effective, radius is needed to calculate the hydrodynamic and thermodynamic properties of the coolant. This is

defined as the ratio between the perimeter and the area of the channel. Effective diameter would therefore be expressed as:

$$D_{eff} = 4\frac{Area_{channel}}{Perimeter_{channel}} \text{ (m)}$$

An iterative analysis can now be performed by equating the energy balance with the constant temperature internal flow model:

$$q_{conv} = \bar{h}A_{channels}\Delta T_{lm} = \dot{m}cp(T_{coolant,out} - T_{coolant,in}) \text{ (W) (1)}$$

The thermal heat load can therefore be estimated by specifying a range of surface temperatures, channel diameters, and flow rates, and iterating each case with the coolant outlet temperature.

These relations are strictly confined for the fully developed regions of the internal boundary layer region. The fully developed region (Figure 18) is defined as the point in the channel where the velocity profile of the boundary layer no longer varies with increase in channel distance. This point, referred to as the hydrodynamic entry length, can be estimated from the Reynolds # and channel diameter for the laminar case:

$$\left(\frac{x_{fd,h}}{D}\right)_{laminar} \approx 0.05\,\mathrm{Re}$$

Figure 18: Laminar, hydrodynamic boundary layer development in a circular channel. Derived from [9].

A thermally fully developed condition is likewise represented as the point in the channel where the fluid temperatures along the thermal boundary layer no longer change with channel length. The thermal entry length can be estimated as:

$$\left(\frac{x_{fd,t}}{D}\right)_{laminar} \approx 0.05\,\mathrm{Re}\,\mathrm{Pr}$$

with Pr = Prandtl #, defined as a ratio of momentum and thermal diffusivities. With water at 22-90 ^{o}C, the Prandtl # varies from ~7-2, respectively. Therefore, hydrodynamic fully developed conditions will arise closer to the channel entry than the thermal fully developed case. Figure 19 displays the cross section of the thermal boundary layer.

Figure 19: Thermal boundary layer development in a heated circular tube. Derived from [9].

The coolant flow pumping power presents an additional design variable for the microchannel heat exchanger. Fluidic pressure differential across the chip is primarily attributed to friction loss through the channel and entrance losses from the transition of fluid from the header to the individual pathways. Channel pressure loss is expressed as:

$$\Delta P = \frac{f\rho V^2 L}{2D_{eff}} \text{ (Pa)}$$

With f = Moody (or Darcy) friction factor, a dimensionless measure of pressure drop for internal flow. Similar to a voltage potential across parallel branches of an electrical circuit, the pressure drop across the parallel channels will be identical for even flow distribution. The Moody friction factor is a function of the surface roughness relative to the channel diameter and the Reynolds #:

$$f = \frac{-(dp/dx)2D}{\rho V^2} = F(\mathrm{Re}, \frac{\varepsilon}{d})$$

with dp/dx = pressure gradient along the channel wall, and ε = wall roughness height (m). For laminar flow, regardless of relative surface roughness, the Moody friction factor is expressed as:

$$f = \frac{64}{\mathrm{Re}}$$

Due to the inclusion of a constant temperature, electrical resistance heat sink in the experimental setup, an estimate of the temperature drop across the epoxy was needed to complete the model. This was accomplished by examining a derivation of Fourier's law for one-dimensional, steady state conduction:

$$q_x = -k\nabla T$$

$$q_x = -kA\frac{dT}{dx} = \frac{kA}{L}(T_{s,1} - T_{s,2}) \text{ (W)}$$

with k = material thermal conductivity (W/m K), A = area (m^2), L = material length (m), and $T_{s,1}$ - $T_{s,2}$ = surface temperature difference (K). For a composite wall, as in the heat sink/buried channel chip interface, an overall heat transfer coefficient is implemented to estimate the thermal resistance:

$$U = \frac{1}{R_{total}A} = \frac{1}{\left[\left(\frac{L_{heat\sin k}}{k_{heat\sin k}}\right) + R_{t,c} + \left(\frac{L_{chip}}{k_{chip}}\right) + \left(\frac{1}{h}\right)\right]} \quad \text{(W/m}^2\text{ K)}$$

with $R_{t,c}$ = thermal contact resistance (0.2-0.9 m^2 K/ W for 20 μm aluminum/silicon epoxy interface [9]) and h = internal heat transfer coefficient (W/m^2 K). The temperature drop across the interface can be found from the input electric resistance heat flux, measured interface lengths, and material properties with an expression analogous to Newton's Law of Cooling:

$$q_{conductive} = UA_{chip}\Delta T \text{ (W)}$$

The temperature drop from the temperature controller to the internal channel surface was estimated at 6.8-11.2 ^{o}C for the three processed channel diameters.

ii. Experimental Setup

A fluidic data acquisition system was designed and constructed to test the heat sink capability of the buried channels. Figure 20 presents a diagram of the system, with photographs presented in Appendix B. The system features a constant displacement gear pump, two gauges to capture pressure drop across the chip, thermocouples inline to the flow to record fluid temperature, two turbine flow meters to measure coolant flow, and a constant temperature bath to regulate the inlet coolant temperature to the sample. The flow meters (25-350 mL/min sensitivity range) were operated in series to measure the majority of the coolant flow. A portion of the coolant through the first meter was directed to the chip through metering valves. Although the

use of an additional flow meter increases the error of the measurement, this configuration offered the greatest flexibility of flow conditions for the initial testing phase. A polyetherimide manifold served as an interface for the coolant flow system, aluminum heat source, and sample chip. A 17.1 W/cm^2 mica flexible heater with temperature control provided thermal heat to the aluminum heat sink, and thus the substrate of the buried channel sample. The manifold was designed to provide a uniform fluid path to the cooling channels while blocking undesired flow around the sample.

Figure 20: Schematic of fluidic test setup

iii. Modeled/Experimental Results

Pressure drop and thermal cooling capacity were measured for the three test samples previously presented. We varied flow rate through each chip from 2-15 mL/min for four substrate temperatures. These conditions were likewise entered into the iterative energy balance model (1) to compare simulated versus experimental conditions. We performed an additional test to determine the heat removal of the cooling fluid due to conduction from the manifold. The experimental data was normalized based on thermal trends evident in the calibration. Figures 21-23 present the experimental and simulated cooling capacity versus flow rate for Samples A (D_{eff} = 60 μm), B (D_{eff} = 70 μm), and C (D_{eff} = 72 μm), respectively. Plotted data points represent an average of 40 measured values at each condition.

The experimental data clearly fall below the simulated values cooling performance at the higher set-point temperatures. This is most likely due to error in estimation of surface temperature. On-chip temperature sensors would allow an accurate quantification of thermal epoxy resistance. The experimental data does present similar trends in cooling profiles along the

surface temperature, as heat removal is directly proportional to set-point temperature. The data also presents a convergence to a steady state cooling capacity with increasing flow rate. This trend is expected due to the finite amount of heat conducted by the heat sink. The assumption of fully developed conditions along the channel length also presents a source of model error. The steady state equations were used to quantify the hydrodynamic and thermal properties in these regions due to the unavailability of closed form undeveloped estimations.

The high flow, high surface temperature data present an additional phenomenon. Heat flux values removed from the samples at these points are slightly greater than the theoretical uniform maximum heat flux (24 W/cm^2 for Sample A versus a theoretical 17 W/cm^2). A non-uniform heat distribution was most likely because the majority of the heat sink was wrapped in insulation.

To view the effect of cooling capacity with channel diameter, the experimental thermal heat flux (W/cm^2) at the 77 ^{o}C set-point is presented in Figure 24:

The 60 μm-effective channel diameter sample represents the most efficient heat sink of the three samples. This is expected due to the inverse relationship of internal heat transfer coefficient with channel diameter (Nu # = hD/k). The remaining samples, with minimal relative difference in effective diameter, yielded similar heat removal capacity.

Figures 25-27 present experimental and simulated differential pressure drop versus flow for the three channel geometries.

The experimental data roughly agrees with simulated results for the larger sample sizes. Pressure drop also decreased at similar flow rates at higher temperatures with the decrease in water density. The steep pressure curves in the small diameter sample could possibly be

explained by the estimation of the hydraulic diameter. The ratio of area to perimeter for non-circular hydraulic diameter estimation fails at geometries where the length of a profile is much larger than the width [11]. In such cases, the hydraulic diameter is best estimated by the width of the channel. This estimation could apply to Sample A due to the dimensions of the access trench etch (121 μm x 33 μm, Figures 10-11). Figure 28 presents pressure versus flow for the three experimental chips, along with the adjusted hydraulic diameter model for Sample A. The revised estimate of hydraulic diameter clearly presents a more accurate model of differential pressure drop.

The experimental results, along with theoretical thermal and hydrodynamic entry lengths, are summarized:

Sample	**Measured Channel Diameter**	**Effective Diameter**	**Channel Length**	**Modeled Hydrodynamic Entry Length**	**Modeled Thermodynamic Entry Length**	**Maximum Measured Heat Density**
	(μm)	(μm)	(μm)	(mm)	(mm)	(W/cm^2)
A	72.5	60.4	8.23	0.05-0.52	0.26-2.24	24.0
B	88.6	70.2	7.56	0.06-0.69	0.36-3.00	22.2
C	92.0	72.3	6.52	0.07-0.71	0.38-3.25	17.6

V. CMOS Compatibility

The final objective of this research involves the verification that the fabrication recipe used to create buried channels does not harm the on-chip electronics performance. We investigated the compatibility between the CMOS circuitry and the channel processing recipe by observing the performance of a simple 120 gate clock generator circuit adjacent to a z-axis accelerometer MEMS device. Trenches were machined 100 μm from the circuits using the identical process steps used to create the cooling microchannels reported on above. Figure 29 displays the clock generation input and output signals prior to any fabrication of channels. A 0.5 MHz square wave output mirrors half of 0-5.25 V, 1 MHz square wave input signal. Significant noise due to probe impedance mismatching exists in the output signal, but is not important to the tests conducted.

Figure 30 displays the identical signal conditions for the clock generator circuit (similar die) after the buried channel fabrication process. The clock circuit is clearly functional and unchanged after the silicon etching and ion milling removal of the polymer.

Although future channel designs will better determine the CMOS electronics/cooling channel compatibility, initial test results are promising and indicate that the buried channel process does not hinder semiconductor circuit performance. This is consistent with early results from Guillou et. al. [16], which demonstrated that released mechanical structures could be integrated 5 μm away from CMOS circuitry with no observable change in electrical characteristics and performance.

VI. Conclusions

This research presents a baseline fabrication process recipe for the creation of buried microfluidic channels in bulk single-crystal silicon. We fabricated silicon heat sinks of varied diameter using a single masking step. The samples were sealed and packaged into a fluidic data acquisition system to test the cooling capacity of the resulting geometries. Using electric resistance to simulate microprocessor thermal heat generation, cooling capacities up to 24.0 W/cm^2 were measured for channel diameters of 72.5-92.0 µm.

The significant contribution of this effort to the goal of integrated, fluidic microprocessor cooling lies in the use of low temperature fabrication processes to create microchannels. We demonstrated that polymer films can be deposited as sidewall protection and serve as an etch mask to XeF_2 processing gases. While C_4F_8 has been used previously as sidewall protection for the formation of scalloped undercuts in open areas [19], the derived recipe in this effort presents the additional advantage of machining undercut channels with the use of smaller mask features (20 µm wide trenches). Further optimization in feature reduction presents the possibility of surface integration with CMOS electronics.

The experimental thermal analysis served to observe the effects of channel geometry on cooling capacity. The clear goal of this research lay in the creation of buried microchannels of 10-50 µm in diameter. Smaller channel diameters produce fully developed conditions closer to the entrance of the channel as well as increase the heat transfer from the silicon substrate to the coolant. With greater efficiency cooling, the heat load can be satisfied with lower pumping rates, reducing the size and power consumption of the circulating pump. The microchannel diameter size in this analysis was directly attributed to the minimum mask feature size attainable during processing. Significant sidewall degradation occurred when attempting to process channels below mask sizes of 10 µm. The 20 µm mask features led to channels of 73-92 µm in diameter, which in turn required relatively high flow rates to satisfy modest heat fluxes of ~24 W/cm^2. By

optimizing the recipe to reduce feature size, higher heat capacities can be obtained at lower pumping rates.

A number of areas of improvement have been identified for future research. They include:

1) *CMOS integration*. Using either backside masking or integrating with circuitry, a high quality aluminum mask would eliminate photolithography error and feature expansion. On chip polysilicon resistance heaters would eliminate the need for thermal bonding as well as offer high thermal heat generation. CMOS temperature sensors could give an accurate substrate temperature profile during cooling.
2) *Feature reduction*. In addition to integration and cooling efficiency, smaller feature sizes offer greater flexibility in sealing the access trenches. This process was proven successfully with 20 μm trenches, but further optimization could determine the feasibility of 5 μm trenches as well as access holes.
3) *Additional low temperature passivation*. The STS ICP polymer deposition rate was measured to vary up to 15% for extended 8.5 minute processes. Due to the porous nature of the passivation, XeF_2 gas was also able to penetrate through the silicon in thin regions of the sidewalls. Future research could determine the worthiness of lower temperature growth (~400 ^oC) SiO_2 as a sidewall mask.
4) *Boiling*. Phase change heat transfer will be needed to successfully cool microprocessors with thermal heat loads up to 300 W/cm^2. This leads to the need for microscopic inspection of bubbles at the outlet header of the test chip, as well as improved packaging.
5) *Customized flow meter*. The flow divider system used in the experimental setup was selected in order to accommodate the high flow requirements of the flow meters.

Custom meters with sensitivity ranges as low as 30 mL/min would offer greater accuracy and simplicity.

6) *Custom packaging*: The manifold used to house the test chip was driven in part by the need to house the electric resistance heater. Packaging with CMOS test chips would eliminate heat transfer to the fluid through the manifold and allow wire bonding at the chip surface.

VII. References

[1] Tuckerman, D. B. and Pease, R. F. W., "High-Performance Heat Sinking for VLSI," *IEEE Electron Device Letters*, vol. EDL-2, no. 5, pp.126-129, 1981.
[2] Bergles, A. E., Kendall, G. E., Griffith, P., and Lienhard V, J. H., "Small Diameter Effects on Internal Flow Boiling," *Proceedings of ASME 2001 IMECE*, pp. 1-17, 2001.
[3] de Boer, M. J., Tjerkstra, R. W., Berenschot, J. W., Jansen, H. V., Burger, G. J., Gardeniers, J. G. E., Elwenspoeck, M., van den Berg, A., "Micromachining of Buried Micro Channels in Silicon," *Journal of Microelectromechanical Systems*, vol. 9, no. 1, pp. 94-103, 2000.
[4] Joo, Y., Dieu, K., Kim, C., "Fabrication of Monolithic Microchannels for IC Chip Cooling," *Micro Electro Mechanical Systems, 1995, MEMS '95, Proceedings IEEE*, pp. 362-367.
[5] Vieder, C., Holm, J., Forssén, L., Elderstig, H., Lindgren, S., Åhlfeldt, H., "A New Process for Combining Anisotropic Bulk Etching with Subsequent Precision Lithography,"*Transducers '97, 1997 International Conference on Solid-State Sensors and Actuators*, pp. 679-682.
[6] Perret, C., Boussey, J., Schaeffer, Ch., Coyount, M., "Integration of Cooling Devices in Silicon Technology," *5th International Workshop Thermal Investigations of ICs and Systems*, pp. 1780-1786, 1999.
[7] Zhimin, W., Fah, C., "The Optimum Thermal Design of Microchannel Heat Sinks," *1997 IEEE/CPMT Electronic Packaging Technology Conference*, pp. 123-129, 1997.
[8] Jiang, X. N., Zhou, Z. Y., Huang, X. Y., Liu, C. Y., "Laminar Flow Through Microchannels Used for Microscale Cooling Systems," *1997 IEEE/CPMT Electronic Packaging Technology Conference,* pp. 119-122, 1997.
[9] Incropera, F. P., De Witt, D. P., "Fundamentals of Heat and Mass Transfer," pp. 79-92, 312-335, 467-494, 3rd Edition, John Wiley and Sons, New York, 1990.
[10] Shames, I. H., "Mechanics of Fluids," pp. 709-712, 3rd Edition, McGraw-Hill Inc., New York, 1992.
[11] Crane, "Flow of Fluids Through Valves, Fittings, and Pipe, Technical Paper No. 410," pp. 1.1-1.7, 21st Edition, Crane Co. 1982.
[12] Azar, K, and Morabito, J., "Managing Power Requirements in the Electronics Industry,"*Electronics Cooling*, Vol. 6, No. 4, December 2000.
[13] Lucent Technology Microelectronics Databook
[14] Allen, A., "Digital Silicon Technology," NEMI Workshop, VA.
[15] Datasheet: Intel Pentium 4 Processor in the 478-Pin Package: 1.5-2 GHz.
[16] Guillou, D. E., Santhanam, S., and Carley, L. R., "Laminated, Sacrificial-Poly MEMS Technology in Standard CMOS," *Sens. Actuators A, Phys(Switzerland), Sensors and Actuators A (Physical)*, vol. A85, no. 1-3, pp. 346-55, 1999.
[17] Recommended by S. Santhanam (Carnegie Mellon University), 11/13/2001.
[18] Recommended by J. Zahn (Carnegie Mellon University MEMS Laboratory), 7/1/2001.
[19] First demonstrated by L. Erdmann (Carnegie Mellon University MEMS Laboratory) and C. Wu (Carnegie Mellon University M.E. Dept.) in the creation of bulk-silicon scallops for nozzle turbulence enhancement, 10/6/2000.
[20] Recommended by STS support personnel, 8/2001.
[21] Maluf, N., "An Introduction to Microelectromechanical Systems Engineering," pp. 66-70, Artech House, Inc., 2000.

VIII. Appendix A

Microchannel Processing Recipe

4 um SiO_2 silicon wafer, ~1.5 cm x 1.5 cm cleaved samples

Step		Time (min)	Pressure (mT)	Flow (sccm)	Power (W)	Additional
1-2	Photolithography (Karl Suess MA56)					AZ 4620, 3K rpm, 120 sec. 110 ^oC hot plate, 17 sec. exp @ 14.1 mJ/cm^2, 60 min. 120 ^oC oven hard bake
3	SiO_2 etch (PTI)	112	125	CHF_3: 22.5; N_2: 5	100	
4	DRIE (STS)	70: 12/8 sec etch/passivate cycle	Etch: 22; Pass: 12	Etch: SF_6: 130, O_2: 13; Pass.: C_4F_8: 85	Etch: 600 coil, 12 platen; Pass.: 600 coil	Etch rate: 2-2.5 µm/minute
5	Isotropic (STS)	3	32	SF_6 :130, C_4F_8: 18	600 W coil, 3 W platen	Etch rate: 2-2.5 µm/minute
6	Passivation (STS)	8.5	12	C_4F_8: 85	600 W coil	Deposition rate: 0.1-0.12 µm/min
7	Ion milling (STS)	2-3	1.4	30	700 W coil, 150 W platen	Platen bias voltage: 80-130 V
8	RIE (STS)	0.3	22	SF_6: 130, O_2: 13	600 coil, 12 platen	
9a	Oven cure	30				T = 200 ^oC
9b	Isotropic (Xactix XeF_2)	A: 20 B: 23 C: 33	Xe: 2 N_2: 10			60 sec. cycles
10	Oxygen plasma (PTI)	20	150	50	100	
11	Wafer bonding					Epotek 302-3M, 4K rpm, 2 hr. 60 ^oC cure