Cooling to keep your Key Chips Alive

Guy Diemunsch

Aavid Thermalloy
70 Commercial Street, Suite 200
Concord, NH 03301 USA

Abstract: plenty of cooling technologies are available, they can be used to spread the heat, conduct it to or exchange it with the cold source. We will review the key parameters, adapt them to the automotive constraints and see how to solve some critical problems.

Keywords: Cooling, Heat, Power, Dissipation, Heat Sink.

1. Introduction

Heat generated by electronic chips must be extracted and dissipated inside the cold source: the air surrounding the car. To do this efficiently we have to be sure to always reduce the heat density from the heat source to the air. We also have to pay careful attention to automotive constraints like balancing the chip temperature to increase reliability and designing solutions able to cool even in emergency situations.

2. Cooling to keep your Key Chips Alive

2.1 From the Chip to the Fluid

2.1.1 Heat Density & Heat Transfer.

The heat generated by a chip is often only characterized by the power dissipated. This is of course important but only key to define the size of the heat exchanger with the air and the air mass flow rate needed to extract this power correctly.

The equation used to extract the surface of the heat exchanger:

$$P\ [W] = H\ [W/m^2\ °C]\ S\ [m^2]\ \Delta T_{solid/fluid}\ [°C] \qquad [1]$$

P: power

H: heat transfer coefficient

S: heat exchange surface

$\Delta T_{solid/fluid}$: temperature difference between the solid and the fluid

The equation used to extract the mass flow rate:

$$P\ [J/s\ or\ W] = Hc\ [J/kg\ °C]\ M\ [kg/s]\ \Delta T_{in/out}\ [°C] \qquad [2]$$

Hc: heat capacitance of the fluid

M: mass flow rate of the fluid

$\Delta T_{in/out}$: temperature difference between the input and the output of the fluid

There is another very important parameter which is important to calculate & design all the thermal interfaces, select the materials or solutions to transfer the heat: the heat density. The equation of the conductivity is:

$$C\ [W/m\ °C] = P\ [W]\ T\ [m]\ /\ S\ [m^2]\ \Delta T_{in/out}\ [°C] \qquad [3]$$

From which you can see easily that when the power density decreases, $\Delta T_{in/out}$ decreases as well.

$$\Delta T_{in/out} = P/S * T/C \qquad [3b]$$

In simple terms, the heat density should always decrease from the chip to the air. In other words we must spread the heat.

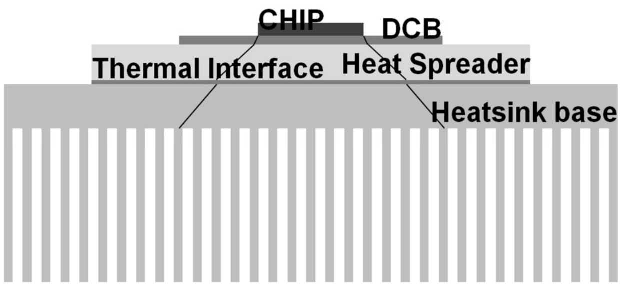

Figure 1: usual heat spreading inside a cooling solution

This looks obvious to most of the Engineers however the decrease of the heat density may not happen efficiently when using pipes (water pipe or heat pipes) or may be created by using bad thermal interface that will affect the available temperature difference.

This available temperature difference might be very small in hybrid cars.

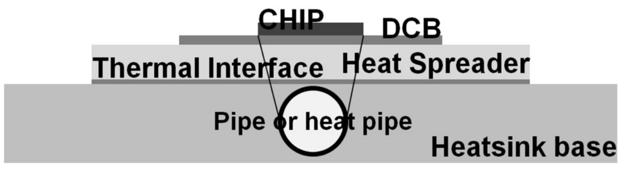

Figure 2: possible heat concentration inside a cooling solution using pipes

There are great tricks to increase the ΔT in equation [1] & [3]: using Peltier cell or AC system (Air Conditioning system) of the car. At least these techniques are used to compensate Thermal

Interface resistances, sometimes to lower the chip temperature close or even below the ambient temperature.

These techniques do not improve the working condition of the final heat exchanger because they add power to dissipate.

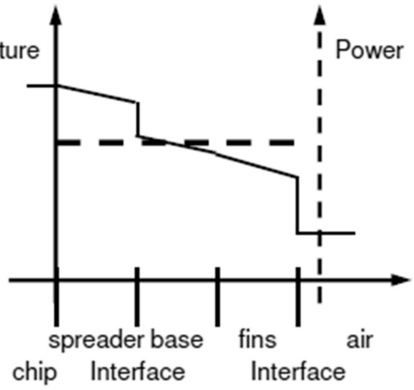

Figure 3: temperature & power graph without Peltier cell

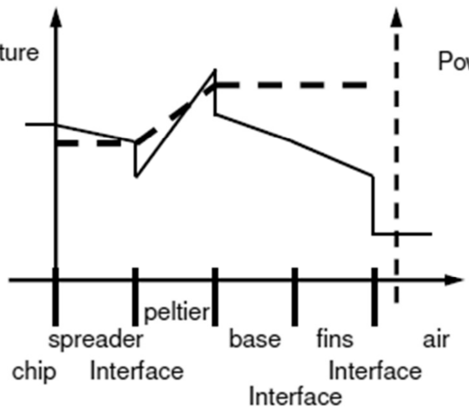

Figure 4: Temperature & power graph with Peltier cell

With Peltier Cell or AC system it is a must to increase the surface of the final heat exchanger to take care of the additional power to dissipate.

From equation [1] we understand as well that the Heat Transfer Coefficient H is very important.

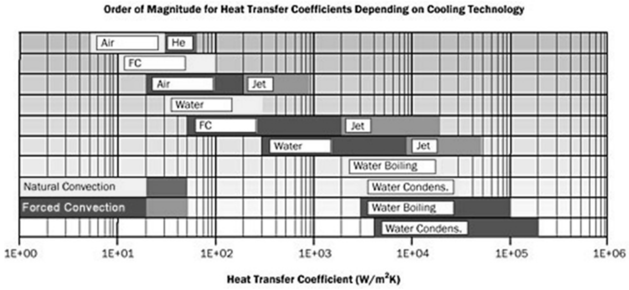

Figure 5: Heat Transfer Coefficient (Electronic Cooling)

Designing cooling for a car to dissipate the heat in the surrounding air imposes the use of the second worst Heat Transfer Coefficient: forced air convection, 20 to 200 W/m² °C. The surface of the heat exchanger is going to be huge. To reduce the volume of this heat exchanger we reduce the space for the forced air: it creates high pressure drop. The pressure drop is the parameter generating high power consumption, often critical for the range of electrical vehicles. More important, the pressure drop reduces the life expectancy of the fan and therefore creates failures.

The automotive industry created one of the best heat exchangers: huge heat exchange surface in a reduced space with an acceptable pressure drop. The trick was to create very thin fins with high fin efficiency by dispatching competently the heat to dissipate on the fin surface.

Figure 6: Automotive Heat Exchanger (T-rad)

Currently other industries are learning from this design.

When we are close to the chip with reduced heat spreading and high heat density, it is critical to select a Heat Transfer Coefficient a lot higher than forced air convection. Currently forced water convection is usable in many applications with an H of 300 to 8,000 W/m² °C. However more and more often we have to deal with chips reaching more than 150 W/cm² (1.5 MW/m²). Here we have to use the best available H: jet impingement, up to 50,000 W/m² °C, and boiling water, up to 100,000 W/m² °C. However even with this high Heat transfer Coefficient, $\Delta T_{solid/fluid}$ is going to reach 15 °C, too high for most of the designs. Heat spreading to reduce the Heat Density is still necessary.

At the end Cooling in a car is managing the Heat Density well with the Heat Transfer Coefficient.

2.1.2 Automotive Constraints

The main constraint is cleanliness: a car is not moving in a clean environment which is of course impacting the air but also the fluid we may use.

At the moment the liquid of a car ages quite rapidly creating clogging and therefore precluding most of the high density designs. Finally we need volume where we do not have space!

This is more obvious for hybrid cars in which in the objective of saving cost, we may decide to use the same cooling fluid between the electronic and the thermal engine. This fluid gets dirty more rapidly, is very hot (100°C) and therefore limits the $\Delta T_{solid/fluid}$ a lot.

2.1.3 Solutions

Heat Pipes are using boiling water and therefore have the highest Heat Transfer Coefficient. The amount of water is quite small and therefore this solution is very light as well. The fact that it is pipes allows creating heat exchange with air copying automotive solutions.

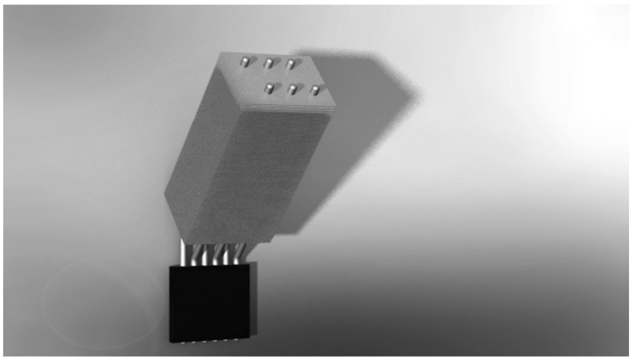

Figure 7: UPS cooling solution (Schneider improvement of a traction solution)

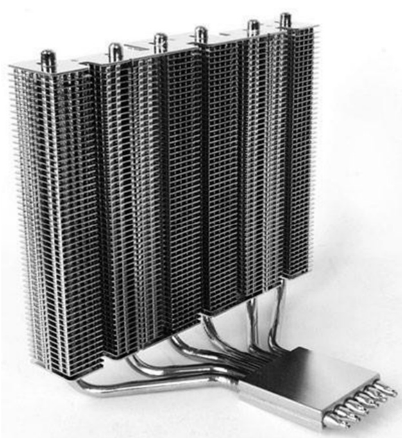

Figure 8: Heat Pipe for a server application (INPAI)

However Heat Pipes do not have only advantages:

- they are very efficient when gravity aided,
- they are quite short (600 mm, max 1m)
- they do not work well in freezing temperature
- they do not start well
- thermal interfaces are critical.

The last issue might be reduced by creating a direct contact between the component and the heat pipe:

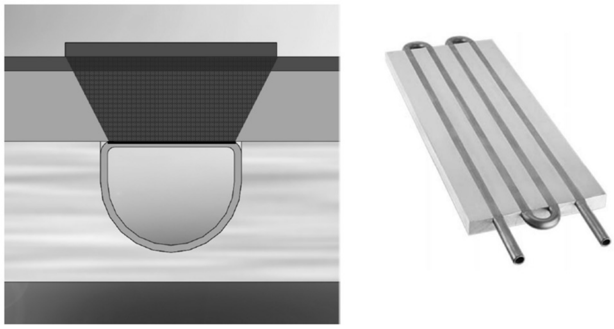

Figure 9: direct contact between a pipe/heat pipe and a module (Aavid)

Other solutions requesting more spreading can replace heat pipes. Channels with imbedded fins allow quite high surface exchange but are limited to water forced convection Heat Transfer Coefficient

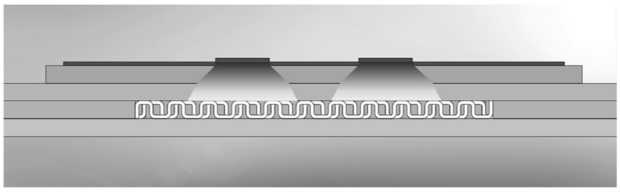

Figure 10: channel with imbedded fins (Aavid)

The drawback of this solution is that the small space between fins can easily be clogged. Jet impingement can overcome this issue because the holes can be large enough to avoid clogging.

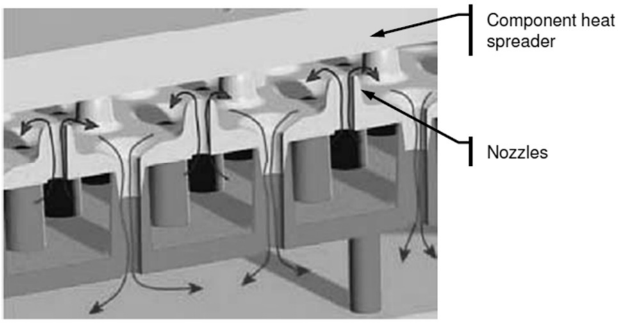

Figure 11: Jet Impingement LCP (Electronic Cooling)

2.2 Balancing Chips Temperatures

Having the same temperature for chips working in parallel is critical in power electronics. The temperature difference expected is maximum 3℃ or better <1.5℃.

To reach this objective the perfect solution is to use phase change. The transfer of power with a phase change fluid will occur at the same temperature when well designed, therefore all the identical chips in a power module will work at the same temperature. However the organization of the power module may create temperature difference when the chips are too close.

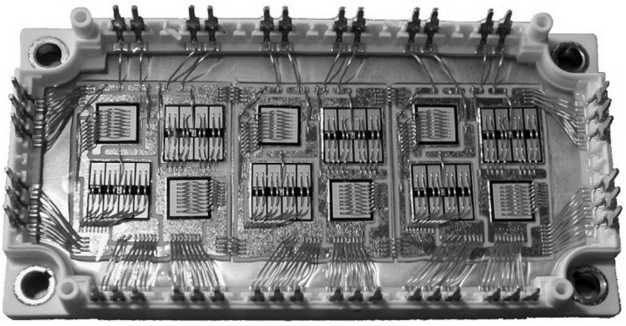

Figure 12: IGBT module (INRETS)

Central chips in figure 12 are surrounded by other chips and their temperature will be affected by these neighbours.

Phase change cannot always be used and we even see in some traction applications that designers prefer to use other solutions using air or water cooling.

We know:

$$P\ [J/s\ \text{or}\ W] = Hc\ [J/kg\ ℃]\ M\ [kg/s]\ \Delta T_{in/out}\ [℃] \quad [2]$$

And therefore:

$$\Delta T_{in/out} = (1 / M)\ (P / Hc) \quad [2b]$$

The temperature difference is proportional to the reverse of the mass flow rate.

This temperature difference will be duplicated to the chips: a chip facing cold fluid will be colder than a chip facing hot fluid (close to the output).

Of course the solution is that all chips face the same cold fluid, all chips are placed in parallel versus the fluid. This is often not feasible from an electronic point of view, it seems impossible to create a balanced module with all chips aligned.

Therefore if electronics does not allow this, the thermal engineer will have to propose solutions.

We saw previously that the low heat transfer coefficient of forced air convection does not allow to do air cooling very close to the chip. Therefore only jet impingement water cooling can be used. It is associating the advantage of a full parallel cooling and a very high heat transfer coefficient.

Figure 13: Jet impingement LCP (Aavid)

When none of these technologies are feasible, there remains the solution of adapting the heat exchange surface with the air to manage the temperature difference between the fluid and the chip. This is made by changing the number of fins along the fluid path. This can be made with air cooling and water cooling.

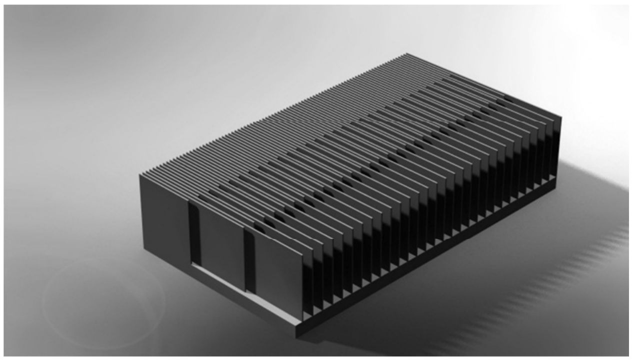

Figure 14: air cooled variable fin density heat sink (Aavid)

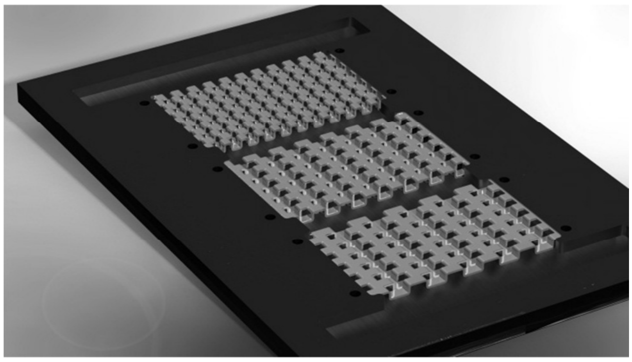

Figure 15: liquid cooled variable fin density heatsink (Aavid)

2.3 Emergency situations

When one element of the cooling solution is failing we are facing an emergency situation: even if there are thermal capacitances in the path of the heat, all of them are fully filled because in most of the cooling solutions the elements are in series and the design is not considering the possibilities of failure.

Therefore as soon as there is a failure the maximum junction temperature of the chip is no longer respected.

Luckily the maximum junction temperature is not the key issue:

- for a power electronic module the stress generated inside some joints inside the electronic module, being thermal interfaces as well, is increasing and accelerating the failure of this joint. Older the joint, faster the failure of the electronic component will happen,
- for memories or µprocessors the increase of the junction temperature will increase the number of errors that may create a bug or worst an unexpected stop of the board.

Of course the closer to the junction the failure is, less the thermal capacitance will amortize the effect of it: a failure of the thermal interface between the electronic module or component is more critical than a failure of a pump or a fan.

Historically failure rates of a cooling element are in fact well adapted to the objectives:

- the thermal interface between the component and the cooling solution, at the beginning of the heat path, never usually fails suddenly, the migration of the material of the thermal interface is the failure mode,
- the equipment transferring the heat is either not failing (extruded material) or having a very low failure rate (leaks in heat pipe or water pipe, pump failure for water loop),
- the cold source (the air) is never failing in natural convection and has a medium failure rate in forced convection (fan wears).

However for automotive we may optimize this in relation with the manufacturing processes and select more effective solutions.

2.3.1 Thermal interface.

Improving the failure rate of the thermal interface between the module/component is feasible. It is simply necessary to remove uncontrollable material like grease.

We can use high mechanical pressure and/or mirror surface thermal interfaces. The failure mode is therefore mechanical: loss of contact or pressure reduction. These failures have a low occurrence but a high impact; they already exist with a slightly lower impact when using grease. Therefore these solutions are an evolution. In addition the Rth are at least 30 % improved. The drawback is that most of the components do not accept high pressure (>400 kPa) or cannot be machined to the flatness of a mirror.

To overcome these issues we can use welding, brazing or soldering between components. Soldering is the most feasible due to the lower temperature needed to apply this process. It is already often used with heat pipes or water pipes inside the cooling solution. Soldering is also used a lot inside the electronic solution therefore it should be feasible for the complete solution with a drawback: the electronics and the cooling solution become one solution. However this process will not be compatible with very high junction temperature (> 220 °C).

The thermal conductivity of this kind of thermal interface reduces the stress applicable to these joints a lot and therefore the risk of failure.

We can also fully eliminate the thermal interface to create a direct contact between the module/component and the liquid doing the heat transfer. Direct liquid or phase change cooling might be very efficient associated with the right heat spreading to obtain a heat density fully compatible with the solution. Direct liquid cooling is today the only feasible solution because currently there is no industrial solution able to imbed a component or a module inside a heat pipe or vapour chamber.

There are only two functions/elements of the cooling solution able to store heat:

- the heat transfer function,
- the cold source: air.

2.3.2 Heat transfer.

We saw that heat transfer may use solids, liquids or heat pipe.

Solids and liquids may be considered as potential heat capacitance to store heat when a failure occurs. It is quite impossible with Heat Pipes. Therefore the use of a Heat Pipe requires the implementation of back up Heat Pipe.

Solid allows for easily adding safety capacitance but is only efficient if this material is parallelized. If not it is only adding Rth.

Liquids are a lot more efficient to add thermal capacitance: it can be made by using a volume of water larger than needed and/or by using a higher mass flow rate. The first solution reduces the risk generated by pump failure as well.

2.3.3 The air.

In fact the air around the car does not have a lot of heat capacitance but the volume is infinite.

To be able to use it we should always do like for thermal engines: the movement of the car must be able to replace the fan, but we should always allow & use natural convection as well.

At the end whatever we put in place to take care of an emergency situation it is transformed in a better thermal resistance. This must be the rule for a reliable design even when a failure occurs.

3. Conclusion

There are plenty of cooling technologies available. It is important to use them in the best working environment they are developed for, to design an efficient and low cost cooling solution for the new challenges of the automotive industry.

We saw with this paper by looking at simple equations, related to the physical phenomenon we have to comply with, that we already have technologies able to overcome the automotive constraints.

The key is to look at Heat Density and to reduce it all along the heat path to be able to reach a Heat Density compatible with forced air convection, which is requested to be able to dissipate the heat in the only cold source available for a car.

Along the same path we have to implement solutions permitting the car to continue to work in case of failure which finally ends up to design with margin.

Even if solutions already exist we can still improve and develop elements and reduce constraints to develop low cost and reliable solutions.