Purpose, potential and realization of chip-attached micro-pin fin heat sinks

M. Conrad a,⁎, A. Diatlov b, R.W. De Doncker a

a Institute for Power Electronics and Electrical Drives (ISEA), RWTH Aachen University, Germany
b Fraunhofer Institute for Laser Technology (ILT), Germany

A R T I C L E   I N F O

Article history:
Received 25 May 2015
Received in revised form 27 June 2015
Accepted 2 July 2015
Available online xxxx

Keywords:
Thermal management
Thermo-mechanical stress
Power electronics
Pin fin
Heat sink
Power semiconductor packaging

A B S T R A C T

This work presents the design process of an actively cooled power semiconductor package that makes use of a structured chip contact that is preferably produced by Selective Laser Melting (SLM). The concept has a drastically reduced amount of material transitions within the cooling path and is designed for low thermo-mechanical stress. It is shown how to design and build the geometry of the integrated heat sink in such a way that plastic deformation of the applied materials is kept as low as possible. The comparison with state-of-the-art actively cooled power modules shows that a similar or even smaller thermal resistance can be achieved in a much smaller volume.

1. Introduction

With the upcoming of new wide bandgap materials, the possible power loss density and maximum junction temperature will increase. Even for Si-based devices this trend exists since years [1]. Hence, the thermal cooling performance as well as the induced thermo-mechanical stress of the package poses a key criterion for the overall performance of the final device.

There are various life-time models which show that the aging effects caused by thermal cycles are caused by the plastic deformation of the applied materials. E.g., [2] showed for common Sn60Pb40 solder that the relation between plastic shear strain \( \gamma_p \) and number of cycles to failure \( N_f \) obeys the Coffin–Manson law:

\[
\gamma_p \cdot N_f^{0.51} = 1.14.
\]

In bond wire contacts plastic deformation is responsible for life-time degradation and depends strongly on the geometry, as [3] showed in their investigation on cycled bond wires with different diameters. The latter aspect enables a design, which accounts for minimum plastic deformation under operational conditions.

Besides high temperature stability, the thermal resistance of the package is aimed to be small. For higher power ratings, active cooling becomes inevitable. A shorter cooling path does not automatically result in a lower thermal resistance, as it implicates less available heat spreading. Consequently, the heat transfer coefficient \( h \) of the heat sink has to be increased in order to benefit from a short cooling path. This condition has also been shown in [4]. The most consequent way of decreasing the length of the cooling path is to integrate cooling channels into the chip itself. In [5] it is shown that with chip integrated microstructures, enormously high heats transfer rates could be achieved. However, the concept implicates limited height of the cooling channels and a high realization effort.

Recently, in [6] it has been shown that metallic contacts can be printed directly on top of a power electronic diode, without violating electrical functionality. These printed pins can be directly designed as heat sink. Compared to the chip integrated cooling channels, on-chip printed heat sinks offer a higher cross section area for the incoming fluid. Thus, higher flow rates can be achieved.

2. Concept

Similar to [6], the chip is attached to a current carrying heat sink on both sides as depicted in Fig. 1.

Ideally, this is realized by Selective Laser Melting (SLM). However, the configuration offers similar advantages in case that the heat sink is soldered onto the chip. In both cases, the heat sink has a pin fin configuration with very fine pitch — even below 1 mm. In contrast to existing approaches, in this work the heat sink is in direct contact to the chip with the pin fins pointing towards the chip. Hence, parts of the chip
are directly exposed to the cooling fluid. This implicates some requirements for the cooling fluid.

The cooling fluid must be electrically insulating and it must not cause any corrosion. That is why oil cooling is regarded in this work, although water cooling would be much more efficient regarding thermal point of view. Compared to actively cooled state-of-the-art modules, like the Infineon Hybridpack 2, the length of the cooling path as well as the amount of material transitions is reduced drastically. The presented configuration enables double sided cooling and flip-chip arrangements. The latter guarantees a low commutation inductance.

The micro-pin fin heat sink is designed in such a way, that the temperature on top, respectively bottom of the heat sink, where the single pins are joining each other, has decreased to the temperature of the fluid. This ensures that there is no relevant difference in the thermally caused expansion of the joining terminal and the chip itself because the coefficient of thermal expansion (cte) of Si is much smaller than the cte of Cu or Al. Consequently, the induced stress is independent of the chip size and can be investigated by the simulation of a single pin.

3.2D stress simulations

The stress strain relation of different materials appears roughly as depicted in Fig. 2. For small strains, the material is in the elastic region, where there remains no plastic deformation after the applied stress is released again. In material testing, usually $\varepsilon_p = 0.002$ is defined as barrier between elastic and plastic regions. For shear loads, the relation is given by:

$$\tau = G \cdot \gamma$$

Table 1

<table>
<thead>
<tr>
<th>Material</th>
<th>$E$ in GPa</th>
<th>$\sigma_{\text{yield}}$ in MPa</th>
<th>CTE in $10^{-6}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>170.0</td>
<td>300.0</td>
<td>2.6</td>
</tr>
<tr>
<td>Cu</td>
<td>110.0</td>
<td>237.0</td>
<td>17.0</td>
</tr>
<tr>
<td>AlSi10Mg</td>
<td>70.0</td>
<td>220.0</td>
<td>23.0</td>
</tr>
<tr>
<td>95.5Sn–3.8Ag–0.7Cu</td>
<td>18.8</td>
<td>16.9</td>
<td>22.8</td>
</tr>
</tbody>
</table>

where $\tau$ is the shear stress, $G$ the shear modulus and $\gamma$ the shear strain. Solders are very creepy, which means that the stress strain relation depends on the strain rate. Additionally, the stress strain relation depends extensively on the temperature.

However, the intention of the conducted 2D stress simulations is not to show the influence of different thermal loads on the thermo-mechanical stress and the resulting strain. The main aim is to show the geometry influence. Therefore, viscoplastic effects are neglected and the FEM tool supposes a sharp transition between elastic and plastic regions, where the relation between $\sigma$ and $\varepsilon$ is given by $E$ for the elastic region and perfect plasticity is assumed in the plastic region. Regarding the stress strain curves of the simulated solder 95.5Sn–3.8Ag–0.7Cu, which were measured in [7], perfect plasticity is a good assumption for higher strains. There are hardly any hardening effects visible. The corresponding parameter values used for the 2D stress simulation are listed in Table 1. Here, $E$ and $\sigma_{\text{yield}}$ of the solder are taken from [7] in the case of a strain rate of $5.6 \times 10^{-2}$.

The following 2D axis-symmetric FEM simulations were carried out with the multi physics program COMSOL.

They provide the plastic strain within a single pin caused by a temperature rise of 100 °C. It is assumed that there is no initial stress. The first two simulations suppose a Cu pin soldered onto the chip with different solder layer thicknesses. The last simulation supposes an AlSi10Mg pin printed directly onto the chip.

Due to the higher stiffness, respectively higher Young’s modulus $E$ of Si and Cu compared to solder, the solder layer is deformed according to the different expansions of the adjacent materials. Therefore, the plastic strain at the edges of the pin in Fig. 3 is much higher than in the centre of the pin. Consequently, the maximum occurring plastic strain increases with pin diameter as shown in Fig. 4.
As a conclusion smaller contact areas implicate less plastic strain. Nevertheless, even for small pin diameters, the solder layer is subjected to plastic strain.

When the heat sink is made of AlSi10Mg printed with the SLM process directly on top of the metallization, the pin can even stay in the elastic region for $\Delta T = 100$ K. Under these conditions there is nearly no plastic strain as shown in Fig. 4. The main reason is the high yield strength of AlSi10Mg, which is significantly greater than that of solder (220 MPa vs. 17 MPa). Moreover, with

$$\varepsilon_{\text{el}} = \frac{\sigma_{\text{yield}}}{E}$$

the range of the elastic strain of AlSi10Mg is roughly 3 times larger than that of the regarded solder.

4. Pin fin design

The corresponding heat transfer coefficient of an in-line pin heat sink can be calculated according to [8]. With equally spaced pins, it follows for the heat transfer coefficient of the bottom exposed to the cooling fluid:

$$h_{\text{bot}} = \frac{0.75 \lambda_{\text{fluid}}}{D} \sqrt{\frac{\left(w_{\text{b}}/d\right)}{N_L \cdot \left(1 + w_{\text{b}}/d\right)^2}} \cdot Re^{0.5} \cdot Pr^{1/3}$$

and for the pins itself:

$$h_{\text{pin}} = \frac{C_1 \cdot \lambda_{\text{fluid}}}{D} \cdot Re^{0.5} \cdot Pr^{1/3}$$

with

$$C_1 = 0.2 + e^{(-0.55 \left(1 + w_{\text{b}}/d\right))} \cdot \left(1 + w_{\text{b}}/D\right)^{0.487}$$

Here, $w_{\text{b}}$ marks the wall to wall distance between the pins and $D$ the pin diameter. The fin efficiency of one pin of a pin fin heat sink is defined as

$$\eta_{\text{pin}} = \frac{R_{\text{th, pin}} \left(\lambda_{\text{pin}} \rightarrow \infty\right)}{R_{\text{th, pin}} \left(\lambda_{\text{pin}}\right)}$$

Due to the limited thermal conductivity of a pin, the efficiency is high for small pin heights. However, in this approach the pin height $H$ should be high enough in order to let the temperature of the pin further decrease. Assuming that the temperature difference between pin surface and fluid is nearly zero at the end of the pin, there is no further vertical heat transport. Under this condition, it can be shown that the temperature distribution within a single pin along the pin height is given by:

$$T(z) - T_{\text{fluid}} = (T(z = 0) - T_{\text{fluid}}) \frac{\cosh(m \cdot (H - z))}{\cosh(m \cdot H)}$$

where $m$ is defined as:

$$m = \frac{4 \cdot h_{\text{pin}}}{D \cdot \lambda_{\text{pin}}}$$

As stated before, the temperature at $z = H$ should have decreased close to the fluid temperature, which corresponds to following condition:

$$\frac{T(z = H) - T_{\text{fluid}}}{T(z = 0) - T_{\text{fluid}}} = \frac{1}{\cosh(m \cdot H)} \leq 0.1$$

This is fulfilled with $m \cdot H_{\text{min}} = 3$. The corresponding minimal height of the pins is plotted in Fig. 5. Obviously, small structure sizes minimize the required pin height. The resulting total thermal resistance is composed of three parts:

$$R_{\text{th}} = R_{\text{fluid}} + (R_{\text{bot}} | R_{\text{pin}} \rangle, \quad R_{\text{fluid}} = \frac{1}{Q_{\text{fluid}}} \rho \cdot C_{\text{th}}$$

$$R_{\text{bot}} = \frac{1}{h_{\text{bot}} \cdot A_{\text{bot}}}, \quad R_{\text{pin}} = \frac{1}{h_{\text{pin}} \cdot \eta_{\text{pin}} \cdot A_{\text{pin}}}$$

$R_{\text{fluid}}$ indicates the heating of the fluid and is neglected in the following.

The pin resistance $R_{\text{pin}}$ is clearly the most significant component and becomes:

$$\eta_{\text{pin}} = \frac{\tanh(m \cdot H)}{m \cdot H}$$

The corresponding heat transfer coefficient of an in-line pin heat sink can be calculated according to [8]. With equally spaced pins, it follows for the heat transfer coefficient of the bottom exposed to the cooling fluid:

$$h_{\text{bot}} = \frac{0.75 \lambda_{\text{fluid}}}{D} \sqrt{\frac{\left(w_{\text{b}}/d\right)}{N_L \cdot \left(1 + w_{\text{b}}/d\right)^2}} \cdot Re^{0.5} \cdot Pr^{1/3}$$

and for the pins itself:

$$h_{\text{pin}} = \frac{C_1 \cdot \lambda_{\text{fluid}}}{D} \cdot Re^{0.5} \cdot Pr^{1/3}$$

with

$$C_1 = 0.2 + e^{(-0.55 \left(1 + w_{\text{b}}/d\right))} \cdot \left(1 + w_{\text{b}}/D\right)^{0.487}$$

Fig. 4. Effective plastic strain in outer edge region of pin for various assemblies obtained from FEM simulations.

Fig. 5. Needed pin height $H$ in order to achieve $m \cdot H \geq 3$ for AlSi10Mg as pin material.

Please cite this article as: M. Conrad, et al., Purpose, potential and realization of chip-attached micro-pin fin heat sinks, Microelectronics Reliability (2015), http://dx.doi.org/10.1016/j.microrel.2015.07.007
The most meaningful expression of the thermal capability of a package is its area related thermal resistance \( r_{th} \). This is plotted in Fig. 6 for oil cooling and in Fig. 7 for water cooling depending on the wall-to-wall distance \( w_D \) and the pin diameter \( D \). The depicted results are evaluated for an inlet velocity of \( v_{fluid} = 1 \text{ m/s} \). The oil has a comparable low viscosity (5.66 mm²/s). Obviously, the resulting resistance strongly depends on the wall-to-wall distance \( w_D \), and changes only slightly with varying \( D \). Only when the relation \( (1 + w_D/D) \) becomes too small, the resistance increases rapidly.

Theoretically, local minima do exist. However, there are production related limitations, which justify a deviation from these minima, especially as resistance merely increases. Summarizing the shown relations, fine pin fin structures are a key to realize very compact heat sinks with low \( r_{th} \).

The dependency of the required height and the thermal resistance on \( D \) and \( w_D \) appear nearly identical for Cu and AlSi10Mg. The higher conductivity of Cu, which is considered in the factor \( m \), results in a higher needed pin height and a lower thermal resistance.

### 5. Realization and measurement

Two Cu based pin fin heat sinks have been milled with pin diameters of 1 mm and 700 \( \mu \)m with an equivalent wall-to-wall distance. On top of the pins, a 1 cm² diode chip has been soldered. The whole construction is embedded into a cooling channel as depicted in Fig. 8.

In [9], the chip temperature is measured indirectly using the temperature dependent forward voltage drop, as the chips are packaged and there is no physical access to them. In this work, there is physical access and as such a type-K thermo-couple is attached to the chip surface (the green–white cable, which goes through the cooling channel of Fig. 8). A second thermocouple element measures the temperature at the opposite end of the heat sink. The voltage drop across the device is measured using two sense-wires in order to determine the applied losses. The cooling oil is cooled down to 20 °C by a process thermostat.

In both cases, the 1 mm and the 700 \( \mu \)m pin diameter heat sink, the pin height is limited to 8 mm due to the used milling tool. Hence, the realized pin height is clearly lower than the required height of 10.63 mm, 13.88 mm respectively.

Nevertheless, the temperature difference from pin to fluid at the side opposite to the chip has dropped to less than one third of the maximum temperature difference.

Table 2 shows the thermal resistance for different flow rates that were measured at a load current of 80 A. This corresponds to roughly 120 W/cm². For the comparison with the analytical approach in Table 2, a 100 \( \mu \)m solder layer is assumed with a thermal conductivity of 86.6 W/(m K). Neglecting the solder layer also results in a deviation

<table>
<thead>
<tr>
<th>Tested sample</th>
<th>Approaching fluid velocity in m/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 × 5 pins ( D = 1 \text{ mm} )</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>+2.7%</td>
</tr>
<tr>
<td>7 × 7 pins ( D = 700 \text{ ( \mu )m} )</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>-0.5%</td>
</tr>
</tbody>
</table>

Please cite this article as: M. Conrad, et al., Purpose, potential and realization of chip-attached micro-pin fin heat sinks, Microelectronics Reliability (2015), http://dx.doi.org/10.1016/j.microrel.2015.07.007
A potential disadvantage of the concept is the reduction of the thermal capacitance that comes along with the spared material that is in direct contact to the chip. This can result in a reduced transient peak current capability. However, the reduction of the thermal capacitance can be mitigated by double sided cooling together with choosing higher pin diameters and low wall-to-wall distances.

7. Conclusion

A packaging concept has been introduced, that is based on a pin fin heat sink that is directly attached to the chip. The packaging concept exposes the chip directly to the cooling fluid. FEM simulations have shown that plastic strain during operation can be significantly reduced compared to standard non-structured chip contacts.

Measurements of the area related thermal resistance of milled pin-fin Cu heat sinks, which have been soldered onto a diode chip show that they are in the expected range of analytical approximations. These approximations suggest that pin fin heat sinks based on AlSi10Mg, which are printed directly onto the power semiconductor chip, are not far away from state of the art cooling performance with today’s limitations of the production process. In summary, the approach provides heat sink miniaturization and reduced plastic strain enhancing lifetime and reliability.

Acknowledgment

The authors would like to thank the German Research Foundation DFG for supporting this project under the funding number DO 618/29-1.

References