Thermal transient characterization methodology for single-die and stacked structures

Oliver Steffens\textsuperscript{2}, Péter Szabó\textsuperscript{1}, Michael Lenz\textsuperscript{2}, Gábor Farkas\textsuperscript{1}

\textsuperscript{1}MicReD Ltd., Budapest XI, Gulyás u. 27, H-1112 Hungary, e-mail: <szabop | farkas>@micred.com

\textsuperscript{2} Infineon Technologies AG, Wernerwerkstrasse. 2, P.O. Box 10 09 44, D-93009 Regensburg, Germany

\textsuperscript{3} Infineon Technologies AG, Balanstrasse 73, D-81541 München, Germany

Abstract

High-power semiconductor packages typically exhibit a three-dimensional heat flow, resulting in large lateral changes in chip and case surface temperature. For single-chip devices we propose to use an unambiguous definition for the junction-to-case thermal resistance as a key parameter, based on a transient measurement technique with much higher repeatability also for very low thermal resistances compared to a two-point thermal resistance measurement. The technique is illustrated on thermal transient measurements of power MOSFETs. A comparison between different thermal coupling to the ambient is used to demonstrate the method’s capability to reveal even subtle internal details of the package. The concept is extended to multi-chip and stacked-chip structures, where transfer impedances have to be introduced. Here, the dynamic properties of the package are important and complex impedance mapping is the proper way to characterize the package.

1. Introduction

For devices with a nearly perfect one-dimensional (1D) heat flow the junction-to-case thermal resistance is defined as

\[ R_{\text{thJC}} = (T_j - T_c)/P \]  

where \( T_j \) means the temperature of the junction at a certain powering, \( T_c \) the temperature of the case and \( P \) the power dissipated at the junction. This quantity is the preeminent figure in product datasheets concerning the steady-state thermal properties.

Had we this simplicity of the one-dimensional heat flow, the concept could be extended to stacked die structures (Figure 1) with some effort. Unfortunately, this simple picture no longer holds if we consider realistic structures with a three-dimensional (3D) heat flow. No single junction temperature can be defined if the temperature changes laterally on the chip, which is the typical case for multiple heat sources in multichip assemblies. The case temperature may exhibit even larger lateral variation at up-to-date packages, which are designed for high power and are usually rather flat. For example, dual cold plate assemblies and power packages with extended heat transfer features involving leads [1]-[3] lie beyond the scope of this linear thermal chain concept. In general, the junction-to-case thermal resistance concept is only applicable if – using the terminology of thermal modeling – an unambiguous “case thermal node” can be defined which finally absorbs all heat from the junction.

In this paper we shall first define how the concept of \( R_{\text{th}} \) thermal resistance can be extended to multichip and dynamic problems introducing a matrix of self- and transfer impedances. The concept will be demonstrated on measurements of lateral multichip structures where transfer effects are strong. Then we outline inherent ambiguities of existing \( R_{\text{thJC}} \) measuring concepts based on Eq. (1) arising from the 3D heat flow, especially at the case surface. Afterwards we propose a transient characterization technique based on structure functions which allow to identify internal structures like die-attach and individual sample differences very clearly.

2. Multiport modeling

A simple linear model of a multichip structure with a parallel \( Z_H \) heat conductance path is shown in Figure 2. Applying power \( P_1 \) and \( P_2 \) we experience \( T_1 \) and \( T_2 \) temperatures (relative to the ambient temperature) at the nodes J1 and J2.

Figure 1. One dimensional model of the heat conductance path.

Figure 2. Simple model of a multichip structure

This multiport structure can be characterized by applying \( P_1 \) and \( P_2 \) powers one-by-one. Exciting only J1 we can define:

\[ T_1 = Z_{11}P_1 \quad , \quad T_2 = Z_{12}P_1 \]  

\( Z_{11} \) is the self-impedance of the J1-ambient port, \( Z_{12} \) is the transfer impedance from the J1-ambient to the J2-ambient.
port. With the shorthand notation $A \times B$ for $A-B/(A+B)$ we can say

\[ T_1 = Z_{11} \cdot P_1 = Z_{21} \cdot (Z_{A+Z_B}) \cdot P_1 \] (3)

\[ T_2 = T_1, Z_{02}(Z_{A+Z_B}) = Z_{41} \cdot (Z_{A+Z_B}) \cdot P_1, Z_{02}/(Z_{A+Z_B}) \] (4)

Resolving the $\times$ operations we get

\[ T_2 = (Z_{02}Z_{A+Z_B})(Z_{A+Z_B}) \cdot P_1 \] (5)

and so from (3) and (5)

\[ Z_{11} = Z_{41} \cdot (Z_{A+Z_B})/((Z_{A+Z_B})) \] (6)

\[ Z_{12} = (Z_{01}Z_{A+Z_B})/((Z_{A+Z_B})) \] (7)

For very high $Z_B$ values we get the obvious results

\[ Z_{11} = Z_A + Z_B, \quad Z_{12} = Z_B \] (8)

The graph of Figure 2 is symmetric, if we excite only J2 we get $Z_{22}$ and $Z_{21}$ swapping $Z_{01}$ and $Z_{02}$ in (6) and (7):

\[ Z_{22} = Z_{21} \cdot (Z_{A+Z_B})/((Z_{A+Z_B})) \] (9)

\[ Z_{21} = (Z_{01}Z_{A+Z_B})/((Z_{A+Z_B})) \] (10)

The network complies the reciprocity principle, $Z_{12} = Z_{21}$.

For very high $Z_B$ values we get now

\[ Z_{22} = Z_{A+Z_B}, \quad Z_{12} = Z_B \] (11)

It has to be noted that we can mathematically compose a $Z_{ODD} = Z_{11} - Z_{12}$ amount. $Z_{ODD}$ has no physical content, and does not yield the value of $Z_B$ as sometimes stated. $Z_H$ can be neglected very rarely, even not in cold-plate setups. In many cases (e.g. wire bond PBGA package with heat sink) it is in the same order as $Z_B$. How we can get rid of the effect of unwanted parallel impedances is well written in [13].

3. Multi-chip measurements in lateral structures

In stacked structures self- and transfer impedances can be of different magnitude, so we selected a lateral multichip structure initially, where these effects are very expressed.

Example 1. Infineon’s BTS 7810 K quad DMOS switch device is a trilithic chip assembly in a P-TO263-15-1 package with one double high-side switch in the center and two low-side switches at the sides. The three vertical DMOS chips are mounted on separated leadframes. Figure 3 shows the package and its footprint outline with heatslugs H1, H2, and H3.

The first measurement was done on a JEDEC standard board in still-air chamber. After switching off a constant power dissipation $P_1$ at the device mounted on H1 we captured temperature transients of all devices.

In Figure 4 we show junction temperature curves fitted at their cold end. The curves are also scaled in $Z_{th}$:

\[ Z_{11}(t) = -[T_1(t) - T_A]/P_1, \quad Z_{12}(t) = -[T_1(t) - T_A]/P_1 \]

For short times until 3–5 s (junction 2) or 12 s (junction 3), we see that the remote positions do not yet respond to the stopping of heat dissipation at junction 1. In the 100 second range all junctions have almost identical temperature and behave as a single heat source.
Figure 6. Self-impedance and transfer impedance curves for the three-chip module in still-air setup, complex loci.

Figure 7. Self-impedance and transfer impedance curves for the three-chip module in cold-plate setup, complex loci.

the mass of the own heatslug and does not reach the other chips in the package.

For the cold-plate setup, we find values of 8, 1.2, and 0.6 K/W for the low-frequency limits of $Z(\omega)$ corresponding to the strong decoupling of the chips. This is also reflected in the transient $Z_{th}$ plots (Figure 4, Figure 5).

Figure 8 repeats the complex loci of Figure 6 with junction 1 over H1 driven ($Z_{11}, Z_{12}$) and adds two measured complex curves with junction 2 over H2 driven ($Z_{22}, Z_{21}$). The expected reciprocity of transfer impedances can be well observed, although H1 and H2 are not symmetric at all (Figure 3).

It is worthwhile composing a full impedance matrix of all $Z_{kn}$ elements. This is the metric that corresponds to the concept of $R_{th}$ for multichip packages when the matrix contains DC values only. Giving full set of $Z_{kn}(\omega)$ loci we extend the $R_{th}$ concept to dynamic modeling.

4. The junction-to-case problem

In the previous section we implicitly supposed that the junctions have uniform temperature. We can measure this temperature (or some average of the actual distribution) by the tester equipment and so we get the $Z_{kn}$ values towards ambient or between temperature sensitive structures directly.

At packages having a large cooling surface (tab, heatslug) assuming uniform temperature is obviously unacceptable. On the other hand, the case surface is not temperature sensitive. Applying a large number of sensors on the case surface is impractical, and would distort the temperature field.

Below we present a simple and accurate approach based on temperature transients and demonstrated on simulation and measurements.

Example II. We selected an SPP80N06S2L-11 high power MOSFET in TO-220 package. For junction-to-case measurements one mounts the device on a cold plate and applies thermal grease on all surfaces to ensure good thermal contact.

We simulated the device on a cold plate at 1 W powering by a thermal simulator [4] to estimate the ambiguity of definition (1). As Figure 9 shows we obtain a total junction-to-ambient resistance $R_{thJA}$ value of 0.89 K/W (interpreting "ambient" as the end of the cold plate). However, the temperature of the copper base shows much variation.

Figure 9. Simulation results of a MOSFET device on cold plate. Temperature distribution on the bottom of the copper base.

Plotting temperatures along different cross sections at the "case" plane (Figure 10) we experience changes between 0.1 and 0.28 °C. i.e. we cannot deduce a well-defined $R_{thJC}$ ac-
according to (1). With $T_J = 0.89 \, ^\circ\text{C}$ we can state that $R_{\text{thJC}}$ varies between 0.61 and 0.79 K/W – a 30% variation, which is also expected in a real experiment depending on where $T_C$ is measured.

Figure 10. Plot of the temperature distribution along different cross sections at the bottom of the copper base.

From the theoretical point of view, the proper way to characterize the package would be to define concentric thermal nodes instead of a single case point. Also a multiport compact model with thermal resistances connecting the junction node to the bottom nodes and the latter to each other would give an acceptable result [5].

In practice, however, a device datasheet without a reproducible key parameter like $R_{\text{thJC}}$ is not acceptable. As it turns out in the next section it is still possible to define such a parameter based on a transient measurement technique. Such a simple amount also offers advantages when comparing different devices.

5. Transient measurements for single MOSFET devices

In the steady state technique, $R_{\text{thJC}}$ is calculated from the endpoints of a cooling curve by taking the difference between the initial temperature at the junction ($T_J$) and the final temperature of the case surface ($T_C$), after switching off a constant power dissipation at the junction. $T_J$ is determined from a calibrated temperature-sensitive parameter as the forward voltage of a diode, $T_C$ may be measured by a thermocouple located in the cold plate beneath the device. Figure 11 illustrates a cooling curve, showing the junction and cold-plate temperatures vs. time.

This method adds the grease interface (and some of the cold plate) to the measured thermal resistance. In the transient and frequency domain we face additional problems. Figure 12 shows an RC model of the package, with a single junction node and more nodes corresponding to different case locations, one of which is used for temperature probing. Comparing this figure to Figure 2 besides $Z_A$ and $Z_B$ we can also identify $Z_{\text{III}}$ as the parallel impedance composed of all threads except the one where probing occurs. Dividing the temperature curves by $P_I$ in Figure 11 $T_J$ can be scaled to $Z_{\text{III}}$, $T_C$ to $Z_{\text{II}}$. The curve composed as $- (T_J - T_C)/P_I$ will correspond to $Z_{\text{ODD}}$ defined in (12). $Z_{\text{ODD}}$ has no physical meaning and is not monotonic. This is discussed shortly in [14] and more detailed in [6] but only an ad-hoc solution is proposed.

Figure 11. Cooling curves from traditional $R_{\text{thJC}}$ measurements.

Moreover, depending on the probing position we always get different temperature values due to the lateral temperature variation on the case surface. Small lateral displacements can cause 15% variation in the obtained $R_{\text{thJC}}$ value.

Figure 12. RC model of the junction-to-case problem

Here we propose a transient characterization method, which yields more appropriate information about the thermal properties (and the device structure in detail), by comparing the measured and calculated curves belonging to two different setups, i.e. boundary conditions. To this end we have measured the transient junction temperature of power MOSFET devices in a TO-220 package with two boundary conditions (cf. Figure 13), without explicit measurement of the temperature at any place at the case surface.

Figure 13. TO-220 package on cold plate: “nc” - directly attached; “cp” - with ceramic inlet
• Boundary condition "nc": Package pressed directly against cold plate;
• Boundary condition "cp": Package pressed against a thin ceramics substrate placed on the cold plate.

In Figure 14 we can compare three transients of one sample in both setups, fit at their cold end. The repeatability is very high, the difference among the curves belonging to one setup is less than 2%. In Figure 15 we fit the curves at their hot ends. We observe that for the first 30 ms different boundary conditions yield the same curve, the heat propagates still in the package. Making the difference between the hot point and the temperature where the curves depart we can already say that $R_{thJC}$ is slightly more than 0.6 K/W.

![Figure 14. Cooling curves of a packaged power MOS in "nc" and "cp" setup, fit at the cold point.](image1)

![Figure 15. Cooling curves of Figure 14, fit at the hot point.](image2)

6. Structure functions

For accurate description of thermal self-impedances we often calculate an equivalent 1D thermal RC network. At a large number of RC stages it is convenient to use a graphic representation, the structure functions ([7][8]), where the cumulative thermal capacitance $C_\Sigma$ is plotted as a function of the cumulative resistance $R_\Sigma$ (Figure 16).

Figure 17 shows the cumulative structure functions derived from the transients in Figure 14 and Figure 15. Steep sections correspond to material regions of high thermal conductivity or large cross-sectional area.

It is easier to identify the interface between the sections by looking directly at the $K(R_\Sigma) = d C_\Sigma / d R_\Sigma$ differential structure function (Figure 18) [11].

![Figure 16. Cumulative structure function: the graphic representation of the thermal RC equivalent of the system.](image3)

![Figure 17. Cumulative structure functions of a device in cold plate setup calculated from the curves in Figure 15; with ceramics ("cp") and without ceramics ("nc").](image4)

![Figure 18. Differential structure functions of the device in cold plate setup.](image5)
numbers only between isotherms, along trajectories. The interface planes show lateral $R_{th}$ value variations, as demonstrated by simulation results in Figure 20.

Structure functions of 1D structures show sudden changes at the material interfaces. If we sum up thermal capacitances ($C_T$) in a narrow tube along a trajectory between two isotherms we get different cumulated $C_T$ values belonging to the same $R_T$. In a thread near the center, the trajectory contains less copper and more grease or ceramics in a certain $R_{thJC}$ interval thus the contribution to $C_T$ is less than around the edges.

In our present example the last peaks in the differential structure functions (Figure 18) at 0.62 K/W correspond most likely to the bulky copper base region. Afterwards we see the transition from copper to ceramics between 0.6 and 0.8 K/W along a flat curve. In the case of grease interface between copper base and cold plate we can’t see pronounced peaks or valleys as in Figure 18 but a rolling curve with inflection points at the interfaces copper–grease and grease–cold plate at 0.67 K/W and 0.75 K/W (for details see also Figure 19).

As proposed previously [12], we derive a parameter $R_{thJC}$ determined from the point of maximum curvature in the cumulative structure function, or the inflection point of $K(R_{\Sigma})$, in the region of the supposed interface from copper base to cooling mount (grease or ceramics in our example). This gives back exactly the geometrical interface planes in the 1D model of Figure 2 and can be considered as best approach for 3D heat flow. We obtain $R_{thJC}$ values around 0.66 K/W with 2–3% repeatability when comparing different samples and setups (Figure 19).

It has to be noted that from the $R_{th}$ values in “nc” and “cp” setups also the heat flux through the ceramics can be calculated knowing the case area.

Measurement results were cross-validated by simulation. Simulated transients resembled the measured curves well. Figure 20 shows steady-state simulation results of the device on cold plate. $R_{thJC}$ approximately corresponds to the −0.65 K/W isotherm (if the junction is considered 0 K/W). It is important to note that this isotherm is not related to the geometrical case surface in any manner.

Structure function methodology enables us to look also in the fine details of the structure. Figure 21 compares three MOSFET samples. We can attempt to find the chip region and die attach using the dimensions of the chip from Table 1. This gives a thermal capacitance of 1.95 mm$^3$ silicon of about $C_{th} = 0.0032$ Ws/K. We find this $C_T$ value after 0.1–0.2 K/W in Figure 21, depending on the sample. The capacitance does not increase much until 0.25 K/W, i.e. in this section there is a material of low thermal conductance or cross section. This applies to the lead-based die-attach with less than 20% of the silicon specific heat.

As proposed previously [12], we derive a parameter $R_{thJC}$ determined from the point of maximum curvature in the cumulative structure function, or the inflection point of $K(R_{\Sigma})$, in the region of the supposed interface from copper base to cooling mount (grease or ceramics in our example). This gives back exactly the geometrical interface planes in the 1D model of Figure 2 and can be considered as best approach for 3D heat flow. We obtain $R_{thJC}$ values around 0.66 K/W with 2–3% repeatability when comparing different samples and setups (Figure 19).

It has to be noted that from the $R_{th}$ values in “nc” and “cp” setups also the heat flux through the ceramics can be calculated knowing the case area.

Measurement results were cross-validated by simulation. Simulated transients resembled the measured curves well. Figure 20 shows steady-state simulation results of the device on cold plate. $R_{thJC}$ approximately corresponds to the −0.65 K/W isotherm (if the junction is considered 0 K/W). It is important to note that this isotherm is not related to the geometrical case surface in any manner.

Structure function methodology enables us to look also in the fine details of the structure. Figure 21 compares three MOSFET samples. We can attempt to find the chip region and die attach using the dimensions of the chip from Table 1. This gives a thermal capacitance of 1.95 mm$^3$ silicon of about $C_{th} = 0.0032$ Ws/K. We find this $C_T$ value after 0.1–0.2 K/W in Figure 21, depending on the sample. The capacitance does not increase much until 0.25 K/W, i.e. in this section there is a material of low thermal conductance or cross section. This applies to the lead-based die-attach with less than 20% of the silicon specific heat.

Dots in the structure function of the figure correspond to RC stages of Figure 16. The end of the die-attach region is clearly indicated by a sharp increase of the structure function beyond 0.25 K/W. Differences between samples develop in the die-attach region, as all curves run parallel afterwards.

This difference in die-attach quality can be verified by ultrasonic microscopy (Figure 22, using a 75 MHz transducer)
where samples no. 4 and 39 exhibit thinner solder thickness and greater die tilt. Die tilt (i.e. varying solder thickness) shows up as an interference pattern in the ultrasonic echo signal (dark–bright transitions) which is most evident for part 39. The white spots are solder voids, here considered being of minor importance regarding the thermal resistance.

Figure 22. Ultrasonic images of the lead-based die-attach of the three devices in Figure 21, no. 4 (left), 12 (center), 39 (right).

The reduced solder thickness can be verified from the return times of the echoes. It correlates with reduced thermal resistances of the die-attach compared with sample 12: the latter part has a more homogeneous and thicker solder layer yet a larger thermal resistance (Figure 21). Inhomogeneity in solder layer thickness (a 3D effect!) shows up as variation in the slope of the structure function in the die-attach region.

We can see that the cumulative structure function of all samples starts bending just after $C_{th} = 0.2$ Ws/K. Introducing the term $R_{thJC}$ for $R_{th}$ read at this fixed $C_{th}$ value we slightly underestimate the equivalent one-dimensional $R_{thJC}$ defined by the maximum curvature, but we get excellent repeatability, because here all trajectories still run in the copper. At $R_{th}$ values near to the bending some trajectories cross the case surface and actual grease width and attachment force influence the structure functions. Below the bending, we obtain the same $\Delta R_{th}$ difference within approximately $\pm 1\%$, regardless of the particular value of $C_{th}$ chosen for reading-off $R_{thJC}$.

7. Measurement of a vertical multichip structure

Additional features arise when we analyze the vertically stacked chip architecture (chip-on-chip systems) where the heat from the top chip must flow through the bottom chip and the mold compound with low thermal conductivity.

Example III. We studied the thermal parameters of a BTS6143 device in TO252 like package. Here a small control chip is placed on top of a large power chip. Figure 23 shows the draft geometry and the silicon structures we could use for excitation and sensing (reverse diode in the power chip, protection diode and a resistor separating the supplies in the control chip).

![Figure 23. Stacked-chip system BTS6143.](image)

We measured thermal transients and calculated complex loci in “nc” and “cp” setups of Figure 13 powering first the upper chip (Figure 24) then the lower chip (Figure 25). We can make many informative observations like:
- at pumping frequencies below 10 Hz the thermal wave does not reach the case surface, all loci run together
- the control chip-to-case $R_{th}$ is slightly more than 11 K/W, here $Z_{11}c_{p}$ and $Z_{11}n_{c}$ depart
- the power chip-to-case $R_{th}$ is approximately 0.7 K/W, here $Z_{22}c_{p}$ and $Z_{22}n_{c}$ depart
- in the “cp” setup the parallel heat conductance path cannot be neglected even when the lower chip is excited, $Z_{22}c_{p}$ and $Z_{22}c_{p}$ differ
- the structure is non-reciprocal, $Z_{12}$ and $Z_{21}$ curves differ. This can be easily explained. In case of the lateral arrangement above the heating and sensing silicon structures are far from each other and can be treated as point-like entities from this distance. In the present case $Z_{12}$ depicts the response of the large lower chip on the heatflow from the 80 $\Omega$ resistor in the small upper chip. $Z_{21}$ shows the response of the protection diode of the small upper chip on the heatflow from the whole surface of the large lower chip.

In Figure 26 we see all cumulative structure functions. We can identify the DA1 and DA2 die attach regions in the curves belonging to the excited control chip. $C_{th}$ is very low and $R_{th}$ is very high in this sections as the heat propagates in a narrow cone below the control chip only.
The curves belonging to the power chip quickly turn into the characteristic shape of spreading in a copper cone under the large chip surface at a high capacitance value. Stepping into the cold plate occurs at similar $C_{\text{copper}}$ value in all curves.

Reading the highest curvature (or inflexion of the derivative) in the curves we get 0.72 K/W for $R_{\text{thJC2}}$ and 11.7 K/W for $R_{\text{thJC1}}$.

**8. Conclusions**

An impedance matrix of DC self- and transfer impedances can be used as a metric that corresponds to the concept of $R_{\text{th}}$ for multichip packages. Extending this concept to a set of $Z_{\text{int}(\omega)}$ complex loci we get a similar metric for dynamic modeling.

Elements of this matrix can be directly gained from thermal transient measurements.

We analyzed the case surface temperature distribution of high power packages showing a characteristic three-dimensional heat flow. We found that the $R_{\text{thJC}}$ junction-to-case thermal resistance cannot be interpreted as a single value due to the lateral temperature variation on the case surface.

We proposed an unambiguous definition for the $R_{\text{thJC}}$ junction-to-case thermal resistance of such packages based on structure functions and presented a measurement technique ensuring high repeatability also at very low $R_{\text{th}}$ values.

Selecting characteristic points on the structure functions we could use the results for die-attach quality analysis and for measuring the differences of $R_{\text{th}}$ values among samples with very high resolution.

For overall thermal performance of a package, it is the average case temperature across the cooling mount that matters, rather than the local temperature at a probing position; hence by using the structure function analysis to derive $R_{\text{thJC}}$, we find more likely relevant macroscopic package differences when comparing different samples than by local probing of the case temperature.

For the future, it is of great interest to verify the measured die-attach characteristics on a larger number of measurements.

Acknowledgement

The authors would like to acknowledge fruitful discussions with Heinz Pape and Dirk Schweitzer (Infineon) and support within the European cooperation project PROFIT [1].

**References**