Numerical and experimental study on surge current limitations of wire-bonded power diodes

Thomas Hunger[1], Oliver Schilling[1], Frank Wolter[2]
[1]Infineon Technologies AG, Max-Planck-Straße 5, 59581 Warstein, Germany
[2]Infineon Technologies AG, Am Campeon 1-12, 85579 Neubiberg, Germany

Abstract

Forward surge currents are considered both experimentally and numerically for wire bonded silicon free wheeling diodes. Measurements indicate that the destruction threshold is not given by the onset of intrinsic conductivity. Instead the failure signature indicates a purely temperature driven destruction mechanism caused by the limited ruggedness of the front side contact against excessive heating. The dependence of the surge current limit on pulse width is investigated for different diode types both experimentally and in theoretical models. The transient thermal behaviour is important for a quantitative analysis of the surge current leading to a good understanding of experimental data. A simple analytical formula for the maximum surge current is derived. The failure mode is studied in more detail via a finite element simulation. Finally the influence of higher junction temperatures on the surge current is discussed.

1 Introduction

The ability of silicon power diodes to withstand a rapid forward current load is an important criterion that has to be regarded in the selection of the proper device for a power electronic application. The possible failure mechanisms and the physics that dominate the regime of high surge current densities have been discussed for Si and SiC diodes [1,2]: In general it has to be taken into account that the current-voltage (I-V) characteristic can be drastically influenced by the dependence of mobility μ and intrinsic carrier concentration n_i on temperature when certain current densities are exceeded. An analysis of the I-V characteristic for typical wire bonded Si-diodes is part of this work to differentiate whether the regime of strong $\mu(T)$ or $n_i(T)$ dependence has to be regarded (μ: mobility, n_i: intrinsic carrier concentration). Further specific focus is laid upon the influence of pulse width on the surge current capability. Historically, diodes have been employed in passive input rectifier circuits prior to their extensive use as free wheeling diodes (FWDs) in inverter stages. In a rectifier circuit the time dependence of on-state current is closely linked to the frequency f=50Hz of the supply voltage. Therefore 10ms sinusoidal half wave pulses are typically applied to characterise the surge current capability. But there are working conditions that demand other pulse durations t_P. If the 16⅔ Hz power line used in railway power supply is considered t_P=30ms may occur at the input rectifier. Furthermore a rectifying action of the FWDs may also happen if energy is regenerated from the motor through the inverter into the DC link capacity at frequencies below 50Hz giving rise to t_P~100ms. On the other hand short current pulses have to be withstood by the diodes if the energy in the DC link capacitor discharges during a bridge short circuit. The circuit consisting of the stray inductances and the capacitance of the inverter gives rise to damped current oscillations characterised by small pulse widths far below 1ms.

We focus our research on fast silicon pin diodes (Infineon EmCon™ technology) usually used as FWD in standard IGBT modules. The paper is organized as follows. First, experimental results are discussed. Typical failure signatures are shown. Then a scaling of all datasets onto a single curve is discussed based on considerations of thermal impedance. Finite element simulations are performed to analyse the microscopic failure mechanism leading to the observed failure signatures.

2 Surge current experiments

The surge current is usually measured via half-sine-current of a certain pulse length t_P. Mostly, t_P=10ms from a transferred 50Hz line is used. Fig. 1 and 2 show typical I-V characteristics taken for a succession of pulses.

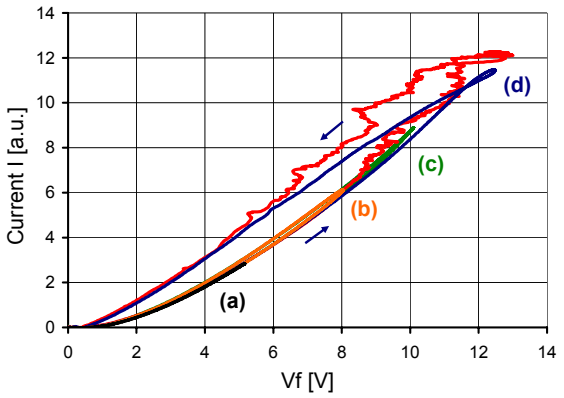

Fig. 1 I-V characteristics of 10ms half sine pulses subsequently taken at different peak currents (curves (a) – (d)). Curve (d) is slightly above the destruction threshold, the red curve far above the destruction.

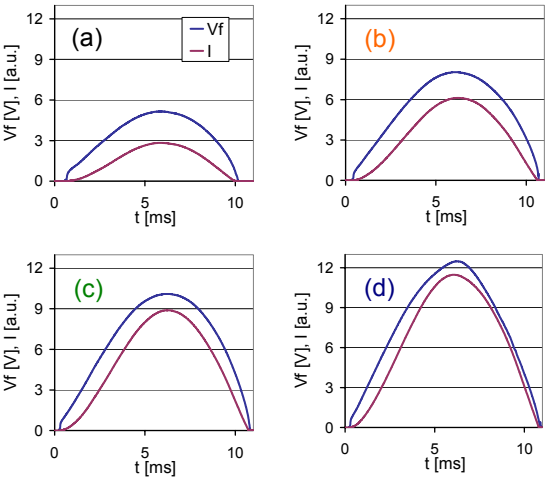

Fig. 2 Time dependencies of the current *I* and the forward voltage V_f for the pulses (a) – (d) of fig. 1.

The curves (a) – (c) in fig. 1 depict I-V curves which are typical for diodes both in the up- and downward branch of the sinusoidal current. This is also observable in fig. 2 where the time dependencies of the current I and the forward voltage drop V_f are shown. Curve (d), however, shows deviations between the branches in fig. 1. This is seen in fig. 2 as a deviation of *V(t)* from the sinusoidal behaviour.

The measured on-state characteristics at high current density do not show signs of current limitation by *μ(T)* as reported by Silber and Robertson [1] for the case of diodes with full surface contact. This indicates that the failure mechanisms of the specific wired bonded FWDs under investigation are not based on *μ(T)* or $n_i(T)$-dependence. The wire bonded FWDs seem to be destroyed at an earlier stage. A closer understanding of the destruction mode is gathered by taking the failure signature into account. A FWD after destruction is shown in fig. 3.

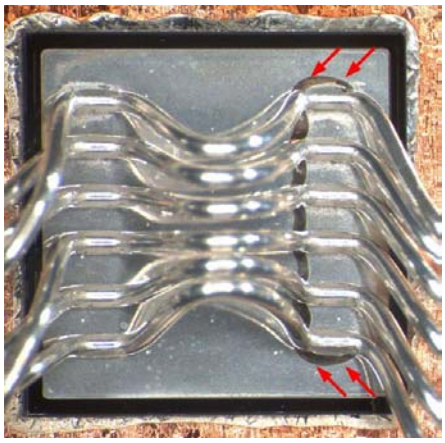

Fig. 3 Failure signature of a FWD destroyed by a surge current overload. There are molten areas located around the bond wedges (red arrows).

The failure signature indicates a local heating up to the destruction of the aluminium metallization layer. Possibly, the underlying silicon diffuses into the aluminium and vice versa and hence spike formation occurs. The destruction of the upper pn junction is a consequence. Indeed, the loss of the blocking capability is the most sensitive indicator to determine the destruction threshold during surge current experiments.

The failure signatures point to a thermal limitation of the surge current capability. Therefore, we consider the amount of energy *E* applied to the silicon

$$E = \int P(t')dt' = \int V(t')I(t')dt' \qquad (1)$$

during a surge pulse. We assume a linear I-V characteristic

$$V(t) = V_0 + R \cdot I(t) \qquad (2)$$

Furthermore, we neglect V_0 and, by using the current, obtain

$$I(t) = I_{FSM} \sin(\pi t / t_p)$$

$$E = R\int I^2(t')dt' = \frac{\pi}{2} R \cdot I_{FSM}^2 t_p \qquad (3)$$

Assuming that the energy $E = \rho \cdot V \cdot c \cdot \Delta T$ (ρ is the mass density, *c* the specific heat) is stored in a certain volume *V* we get a temperature increase

$$\Delta T \propto I_{FSM}^2 \cdot t_p = const. \qquad (4)$$

In eqn. (4) the constant value represents the critical temperature rise which will destroy the device irreversibly. From this point of view I^2t should be independent of the pulse widh t_p.

In fig. 4 we plot I^2t as a function of t_p taken with different diode types. The main difference between the diode types is their blocking voltage and hence the silicon die thickness. As it is clearly visible, the surge current integral I^2t decreases with shorter pulse length in contrast to the prediction of eqn. (4). For the explanation

heat conduction has to be taken into consideration.

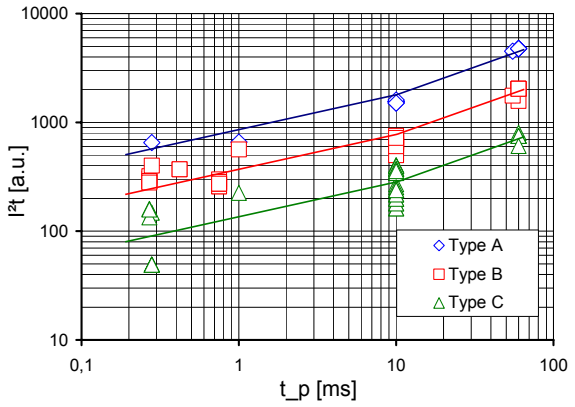

Fig. 4 The experimentally determined surge current integral I^2t as a function of the pulse width t_p for three different diode types. The lines serve only as a guide.

3 Analytical approach to the thermal limitation of surge currents

The simple approach leading to I^2t=*const*. does not reflect the real transient thermal behaviour which is dominated by both thermal conduction and the heating of thermal capacities.
The thermal impedance Z_{th} is defined by means of the equation for the temperature rise ΔT

$$\Delta T(t) = Z_{th}(t) \cdot P_0 \qquad (5)$$

for a step function

$$P = P_0 \cdot \Theta(t)\,. \qquad (6)$$

To be able to apply eqn. (5) the pulse power of a sinusoidal current (see eqn. (1) and (3)) is approximated as an effective square pulse with a temperature independent resistance R:

$$P_0 = \frac{E}{t_p} = \frac{\pi}{2} R \cdot I_{FSM}^2 = R \cdot I_{eff}^2 \qquad (7)$$

Based on the assumption of equal destruction temperatures, eqn. (5) and (7) finally result in a simple formula for the dependence of peak current I_{FSM} on pulse duration t_p:

$$\frac{I_{FSM}(t_p)}{I_{FSM}(10ms)} = \sqrt{\frac{Z_{th}(10ms)}{Z_{th}(t_p)}}\,. \qquad (8)$$

The peak current of a pulse t_p depends only on the square root of the ratio of its thermal impedance value with respect to a reference pulse (here 10 ms). Fig. 5 shows the experimental data from fig. 4 normalized to their 10ms values together with scaling curves according to eqn. (8) based on Z_{th} simulations. The datasets for all diode types under investigation scale onto one single curve. That means the surge current of all diodes is limited by the same physical mechanism. The square root behaviour predicted by eqn. (8) is also plotted in fig. 4 assuming two different models for the underlying Z_{th} calculation (Z_{th1}, Z_{th2}). The thermal impedance curves are computed via a finite element approach because of the lack of reliable experimental Z_{th}-data on such short time scales. 'Zth1' is calculated under the assumption of a homogeneous power loss in the diode volume whereas a power loss only at the diode surface was used for 'Zth2'.

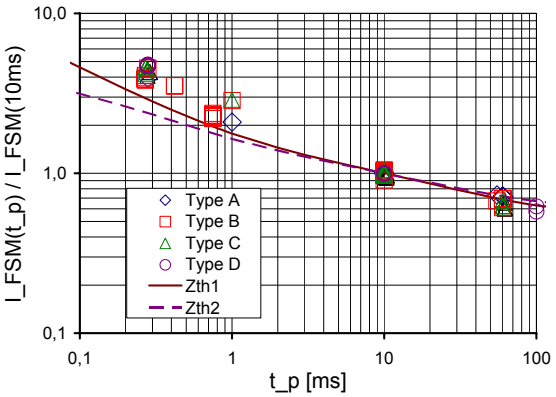

Fig. 5 Normalized peak currents I_{FSM} for different diode types. All data fall onto a single curve. 'Zth1' and 'Zth2' are calculated according to eqn. (8) based on two different Zth models.

For pulse times t_p > 5ms the data are very well fitted by the curve labelled 'Zth1'. Only towards shorter times the scaling underestimates the surge capability. The scaling predicted by the curve marked 'Zth2' (power generation at die surface) deviates stronger compared to the 'Zth1'-graph. This is due to the difference in temperature increases ΔT if both principles are compared, i.e. heat creation in a finite volume (case 1) and the heat transport in a semi-infinite slab with power generation only at the surface (case 2) [3]. Case 1 ($\Delta T \propto t$) leads to a stronger dependence of temperature rise on time than case 2 ($\Delta T \propto \sqrt{t}$). Hence, the slope of the thermal impedance of case 1 is steeper. Therefore, 'Zth1' describes the experimental data better than 'Zth2' especially for short pulse durations as can be seen in fig. 5. From the comparison between experiment and theory we conclude that for short heating times in the order of <10ms the Z_{th} is not sufficiently described by a surface load condition, but the instantaneous dissipation of power in a finite volume ("body load-condition") has to be regarded. For very short pulse durations a saturation of I^2t is expected since the body-load condition corresponds to the assumptions of eqn. (4).
The reason for the remaining deviation between the experimental data and the thermal impedance scaling at short times can be discussed as follows. The dynamics of the aluminium spike formation is governed by the interdiffusion of sili-

con and aluminium. This diffusion is strongly temperature dependent. But temperature alone is not sufficient since the pulse time and thus the diffusion time is also of importance. Therefore, the thermal impedance scaling, which is based only on the temperature argument, underestimates the actual surge current limit for short pulse durations because there is not enough time for a sufficient spiking.
To summarize this paragraph we highlight that the pulse width dependent surge current capability of the wire bonded FWDs can be described by their thermal impedance behaviour. Eqn. (8) can be used for approximating surge current limits for varying pulse widths if the 10ms I^2t value and the correct $Z_{th}(t)$ curve is known in the respective time frame. The approximation yields a lower bound for short pulse durations. However, it has to be pointed out, that the analysis based on the thermal impedance is not sufficient to explain the mechanism of destruction in detail since it will only provide an estimation of the average temperature rise. 3D modelling is necessary to understand the mechanism leading to the typical failure fingerprint as depicted in fig. 3.

4 Finite element simulations

Computations based on a finite element scheme are executed to investigate the failure mechanism in detail. The commercially available simulator ANSYS® [4] is used for the numerical solution of the coupled thermal-electric field. Fig. 6 shows the model setup. It consists of a diode mounted by a soldering layer onto a ceramic substrate with double sided copper layers. The substrate itself is soldered to a copper base plate, which is fixed to a heat sink. The diode's surface metallization is connected to an emitter track via one aluminium bond wire.

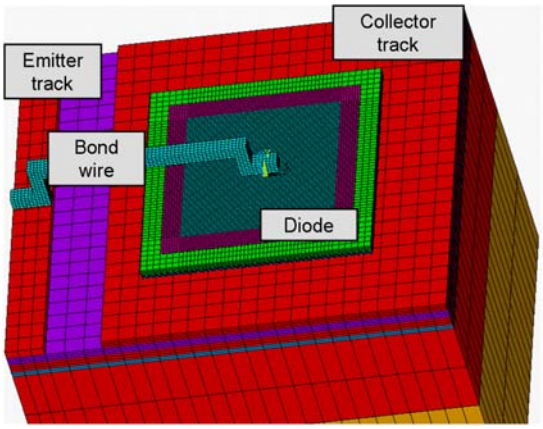

Fig. 6 Geometry and finite element mesh of the device simulated consisting of a diode body together with a bond wire and a substrate soldered to a base plate and a heat sink.

The material models describing the heat conductivity, heat capacity and electrical resistivity of the components copper, aluminium, ceramic (Al_2O_3) and solder include temperature dependencies.

4.1 Description of diode model

ANSYS® as a multi-purpose finite element simulator does not provide the capability of solving the drift-diffusion equations for a full description of the semiconducting device. The method to simulate a surge current load in forward direction is described below.
The diode is operated in a condition of strong injection. Hence, the drift region is flooded by free carriers. A resistivity value can be formally defined under these conditions. To differentiate between the drift region and the region in which the on-state voltage is strongly influenced by the diffusion voltage of the pn-junction the diode is split into two separate layers similar to the SPICE nodes in refs. [5,6], namely anode-middle and middle-cathode. The upper volume contains the p+ anode and the p^+n^--junction. The drift region and the backward n^-n-junction form the lower volume element as sketched in fig. 7.

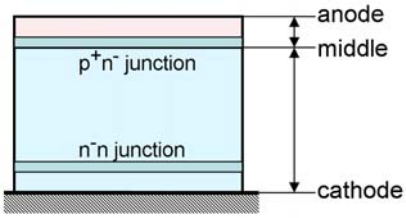

Fig. 7 Diode cross section. The nodes 'anode', 'middle' and 'cathode' are indicated. The layers of the p^+n^- and the n^-n junctions are marked acting as additional heat sources.

The temperature dependent I-V characteristics of both segments are then simplified further on via a linearization according to

$$V(I,T) = V_0(T) + R_D(T)I \tag{9}$$

The current nodes for determining R_D from the I-V curves are chosen at one tenth of and at nominal current.
The parameter $V_0(T)$ is now identified with the voltage drop at the junction itself. In the simulation it accounts for an additional source of the heat density p according to

$$p(T) = V_0(T) \cdot j / t\,, \tag{10}$$

with j the current density and t the "thickness" of the junction.
The parameter $R_D(T)$ is transferred into a resistivity as

$$\rho = R_D(T) \cdot A / D\,, \tag{11}$$

A is the area of the volume and D its thickness.
This approach is reasonable because the wire-bonded diodes fail prior to reaching the μ or even n_i limits as shown in previous sections.

4.2 Simulation results

Fig. 8 shows snapshots taken at different times during a simulation run for I_{FSM}=172A and t_p=10ms. The areas in the vicinity of the bond wedge get hottest during the sinusoidal current pulse. This corresponds very well to the failure signature shown in fig. 3 where the molten aluminium is located around the wedge.

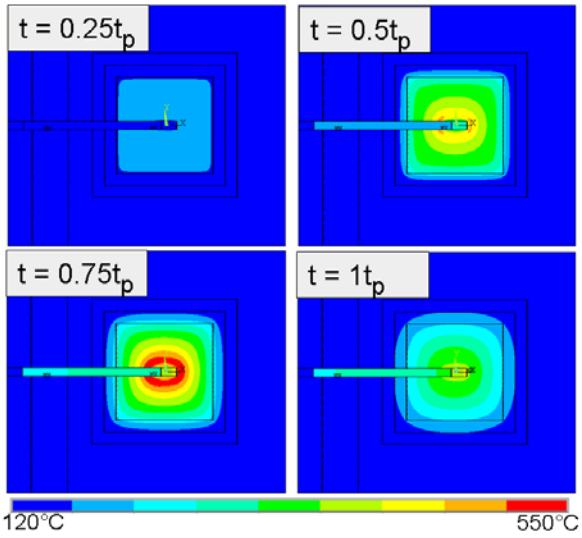

Fig. 8 Snapshots of temperature distribution taken at t=0.25t_p, 0.5t_p, 0.75t_p and t_p during the simulation for I_{FSM}=172A and t_p=10ms. The area around the bond wedge shows the highest temperature during the pulse.

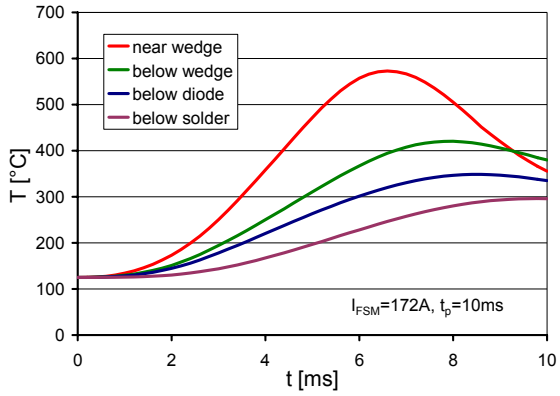

Fig. 9 Temperatures taken at different locations as a function of time. 'near wedge' refers to a position beside the wedge on the metallization.

The time-dependent temperature response taken at different locations is plotted in fig. 9 (I_{FSM}=172A, t_p=10ms). Obviously the bond wire itself has a cooling impact on the surface if a 10ms pulse is concerned, since the hottest point is always located some distance away from it.
The comparison of the temperature distribution on the diode taken for different pulse durations t_p is shown in fig. 10. For long pulses (here 100ms) the whole wedge area is heated. Here the simulation points out the bond wire itself as the hottest part. This is not surprising as one comes closer to the DC limits.
All the pulses t_p<100ms are limited by the temperature swing of diode's metallization. For shorter pulse widths a stronger localization of the hot spots near the wedge is observable.

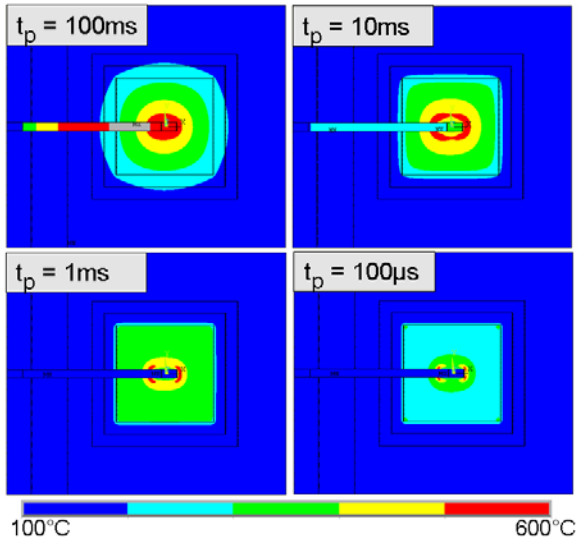

Fig. 10 Snapshots of temperature distributions taken at t=0.6t_p for different pulse durations t_p. Significant differences are visible.

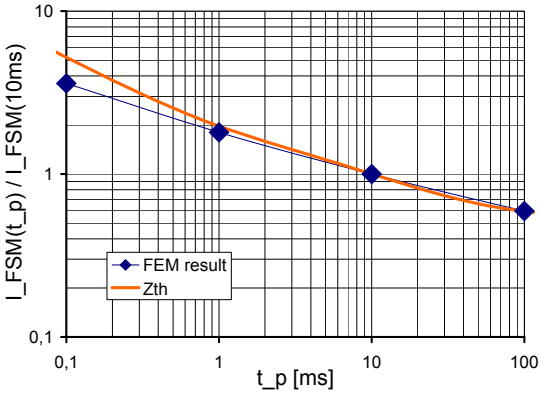

Fig. 11 Normalized peak currents from the finite element simulation ('FEM result') in comparison to the result based on the thermal impedance according to eqn. (8).

Assuming that the failure can be described by equal temperature of the hot spot, the dependence of maximum surge current on pulse width can be derived from the finite element calculations. The result is given in fig. 11 together with the prediction expected from the thermal impedance model of eqn. (8). The finite simulation results resemble the scaled thermal impedance data for pulse durations t_p>5ms. The thermal impedance used here is computed for the specific test diode with one bond wire. The deviations for

small pulse widths can be explained by the high local temperature gradients which are of the order of the finite element mesh size used. Locally improving the mesh near the wedge would enhance the match of the results.
The computed I_{FSM} are too small compared to real experiments. This is mainly due to too high R_D values used in the current simulation. Another improvement can be obtained by including the latent heat of the solder joint because the simulations suggest a solder temperature above the melting level. However, the basic simulation result of a correct resemblance of the failure signature would not be affected.

4.3 Extending limits

Currently, efforts are made to increase the operation temperature of IGBT modules up to a maximum junction temperature T_j=150°C. Therefore, the influence on the surge current capabilities of the diodes is explained hereafter. Fig. 12 shows the temperatures for different starting levels of T_c=125°C and 150°C under equal load conditions. The peak temperatures differ by nearly the amount of the starting difference. From that we get for a certain pulse width by applying eqn. (5) and (7):

$$I_{FSM,1} = \sqrt{\Delta T_1 / \Delta T_2} \cdot I_{FSM,2}\,, \qquad (12)$$

$$\Delta T_i = T_{fail} - T_{C,i}\,. \qquad (i=1,2)$$

Defining the upper temperature T_{fail} as 500°C results in a surge current reduction in the range of 5 per cent which fits very well to experimental findings.

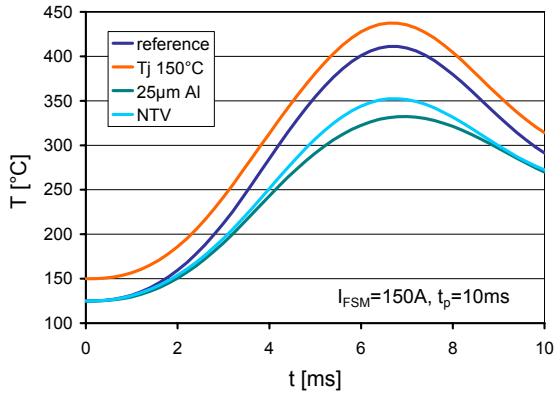

Fig. 12 Temperature responses taken for different simulation setups. 'reference': T_c=125°C, 5µm Al metallization; 'Tj 150°C': T_c=150°C; '25µm Al': T_c=125°C, 25µm Al metallization; 'NTV': T_c=125°C, 30µm silver as solder layer.

Next we discuss possible extensions of the surge current capabilities without changing the silicon of the pin diodes itself. First we consider a thicker metallization layer on top of the diode. Fig. 12 shows the impact of the top metallization on the surge current. The temperature is significantly lowered by increasing the thickness of the top metallization. This is mainly due to a reduced layer resistance and hence lower heating. However, the technological realization of such thick layers is a challenge.
Another approach is the exchange of the solder material (for "standard" simulation setup a 100µm thick Sn layer is applied) towards thinner layers and better heat conduction. Fig. 12 shows the temperature response at the top of the diode using a 30µm silver layer with reduced mass density for modelling of the low temperature sintered joint. Compared to the Sn solder a significant reduction is visible.

5 Conclusions

This paper studies the surge current limitations of wire bonded silicon free wheeling diodes both experimentally and numerically. An analytical formula describing the peak current I_{FSM} limit for different pulse widths based on the thermal impedance is derived. It is useable for the determination of a lower bound for the actual I_{FSM}.
Finite element simulations are used for the study of the destruction mode. The temperature distribution found numerical corresponds to the experimentally gathered failure signatures.
Finally, the extension to higher operation temperatures is discussed. Only a small reduction of I_{FSM} in the range of 5 per cent is expected if starting conditions are lifted from 125°C to 150°C. Further technological levers for enhancing the surge current capability are discussed.

Discussions with Peter Türkes are acknowledged.

6 Literature

[1] Silber, D. and Robertson, M.J.: Solid State Electr. 16 (1973) 1337-1346
[2] Udal, A. and Velmre, E.: Microelectr. Reliab. 10/11 (1997) 1671-1674
[3] Clemente, S.: IEEE Trans. Power Electr. 8 (1993) 337-341
[4] www.cadfem.de, www.ansys.com
[5] Kraus, R. et al.: Proc. IEEE. PESC'98, vol. 2 (1998) 1726-1731
[6] Türkes, P.: priv. comm.