Study of Hot-Spot Phenomena in Cellular Power Transistors by Analytical Electro-Thermal Simulation

Paolo Emilio Bagnoli
Dept. of Information Engineering, Univ. of Pisa
p.bagnoli@iet.unipi.it

Stefano Di Pascoli
Dept. of Information Engineering, Univ. of Pisa
s.dipascoli@iet.unipi.it

Giovanni Breglio
Dept. of Electronic Engineering, Univ. of Naples "Federico II"
breglio@.unina.it

Abstract.

The cellular power transistors are affected by the so-called hot spot phenomenon, a current crowding within few cells occurring for high power conditions, which decreases the device operating range. The onset of this phenomenon was studied under a wide range of electrical and thermal boundary conditions by means of a fast electro-thermal simulation procedure operating in the steady-state regime and using explicit analytical relationships for the temperature mapping. Furthermore, the results of experimental electrical characterizations performed on some samples were in agreement with the foreseen data.

1. Introduction

The so-called "hot spot" is a phenomenon occurring in power transistors with a cellular structure, bipolar [1,2] and also MOS [3] working in high conduction regimes. It consists of a particular situation in which, due to the interaction among geometrical, electrical and thermal factors, a small part of the device accumulates most of the total current. This causes, of course, a local dramatic temperature uprising that can, but not necessarily, provokes the device failure. This phenomenon, whether damages the transistor or not, is anyway dangerous for the device reliability and causes a serious limitation for the operating range and the maximum power ratings.

The aim of the present work exposed in this paper is the study of the hot-spot phenomenon in cellular power transistors by means of an electro-thermal simulation whose thermal solver is based on explicit analytical relationships and working in the steady-state regime. These two facilities drastically decreased the computing time hence extending the range of operating conditions in which the hot spot could be observed. After a brief description of the simulation tools and of the structure of the device under investigation, the onset and the behaviour of hot-spot are reported for several working conditions and as a function of the package thermal resistance. The results of electrical measurements performed on real samples were also compared with the foreseen data.

2. Thermal Solver

The thermal solver, i.e. the part of the simulation program which calculates the temperature maps, was implemented starting from the basic relationships of heat conduction solved under the hypothesis of homogeneous and isotropic thermal conductivity *ks*, for a steady state condition and with a thermal power having any geometrical distribution on the top surface of the chip.

The basic configuration of the solid sample is a rectangular slab of homogeneous material whose size is defined by the lengths Lx, Ly and Ls and where the power is supposed localized in a grid of rectangular cells. Within the single cell the power density was supposed to be uniformly distributed, which is a good approximation in the present case where the device is composed by the parallel of a large number of small area cells.

The top and lateral surfaces of the slab were considered as adiabatic while the bottom one was in contact with a constant temperature heat sink, directly or through a convective coefficient h. The reciprocal of *h* can be retained as the contact thermal resistance at the bottom of the die per unit area and addressed as R_{pk}. This last parameter can represent, from the steady-state point of view, the contribution of the packaging assembling structure to the total thermal resistance.

The solution for the three-dimensional temperature map function T(x,y,z) is reported below in a compact form:

$$T(x,y,z) = T_O + S_0(z) + S_1 \cdot \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} D(n,m) \cdot F(n,m) \cdot \cos(\hat{a}_n x) \cdot \cos(\hat{\imath}_m y) \qquad (1)$$

where To is the heat sink, *n* and *m* are two infinite sets of eigenvalues and the other functions are defined by the following relationships:

$$S_o(z) = \frac{P_o(Ls + R_{PK} \cdot ks - z)}{ksL_xL_y} \qquad S_1 = \frac{2P_o}{ksL_xL_y} \tag{2}$$

$$D(n,m) = (2 - \ddot{a}n - \ddot{a}m)\frac{U_{n,m}}{\acute{i}_{n,m}} \tag{3}$$

$$\acute{i}_{n,m} = \sqrt{\hat{a}_n^2 + \grave{i}_m^2} = \eth\sqrt{\frac{n^2}{L_x^2} + \frac{m^2}{L_y^2}} \tag{4}$$

$$F(n,m) = \frac{\sinh[\acute{i}_{n,m}(L_S - z)] + R_{PK} \cdot ks \cdot \acute{i}_{n,m} \cosh[\acute{i}_{n,m}(L_S - z)]}{\sinh(\acute{i}_{n,m}L_S) + R_{PK} \cdot ks \cdot \acute{i}_{n,m} \cosh(\acute{i}_{n,m}L_S)} \tag{5}$$

Po is the total dissipating power while δn and δm in equation (3) are the Kroneker delta functions.

The information concerning the power distribution on the top surface are contained within the functions $U_{n,m}$. In fact, being u(x,y) the surface power distribution normalized with respect to the total value Po, the functions $U_{n,m}$ are generally defined as follows:

$$U_{n,m} = \int_0^{Lx}\int_0^{Ly} u(x,y) \cdot \cos(\hat{a}_n x) \cdot \cos(\grave{i}_m y)\, dxdy \tag{6}$$

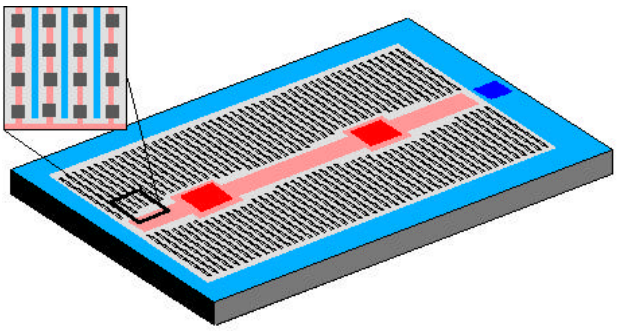

Figure 1. Cellular transistor layout showing the emitter (red) and base (blue) metal and pads.

3. Device configuration

The physical structure of the cellular device (the die size is 9 x 6 mm^2) is shown in fig. 1. It has two symmetrical areas occupied by the grid of square emitter cells, each having an area of 50x50 μm^2 area. The emitter has two pads located in the chip center and a wide metal for current distribution to the columns of cells. The base has a single lateral pad and a wide distribution metal running along the die perimeter. The collector contact is on the bottom side of the chip that has a 0.5mm thickness. The inset of fig. 1 shows the interdigitated structure of the metal fingers of base and emitter which distribute the current to the emitter cells.

Fig. 2 shows the equivalent electrical circuit of a single row of cells, connected to the same emitter supply line and to the same base supply line. In the circuit the resistances R_e and R_b are referred to the metal step distances between two adjacent cells and are the same for all the columns. Instead R_{pb} and R_{pe}, which are different for the various columns, are the resistances due to the large metals of base and emitter respectively which depend on the distance from the nearest pads.

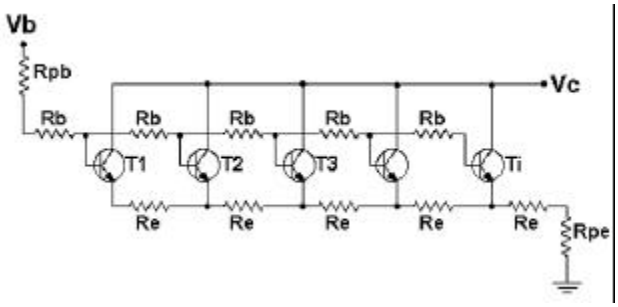

Figure 2. Equivalent electrical circuit for a single row of cells, including the metal resistances.

5. Electrical and Electro-Thermal solver

The transistor cells were electrically modelled by means of the well-known relationships, where the temperature dependence of the reverse saturation current and of the voltage drop of the base-emitter junction was taken into account.

The transistor parameters and h_{FE} have been modelled by means of a purposely carried out set of measurements. For h_{FE} both temperature and current dependence has been derived from experimental results obtaining an empirical $h_{FE}(T, I'_B)$ relationship.

Under these assumptions, the calculation of the electrical configuration was reduced to the resolution of a number of circuits, equal to that in figure 2, equal to the number of columns of cells. All the transistor cell parameters were evaluated using their own temperature. The output of this electrical calculation procedure is the power distribution within the cells.

The whole electro-thermal simulation system is composed by the cyclic and mutual interaction between the electrical solver and the thermal one under the given boundary conditions which are the total input base current and the collector voltage for the first one and the bottom thermal resistance of the package R_{pk}.

In the first run the temperatures within the cells are taken equal to the heat sink temperature. Under this condition the electrical solver calculates the electrical parameters within all the cells and the power distribution within the cells. This is used as input data for the thermal solver, which calculated the new set of temperature values in the center of the cells at the top surface under which the electrical solver operated again. As the number of thermal and electrical calculations cycles increases, the solution becomes closer and closer to a convergence set of values for the thermal and electrical parameter distribution within the sample. The total electrical dissipated power Po was assumed as a general norm for the simulation process: when its relative variation becomes lower than a given accuracy (in this case set at 1.0e-6) the procedure was stopped.

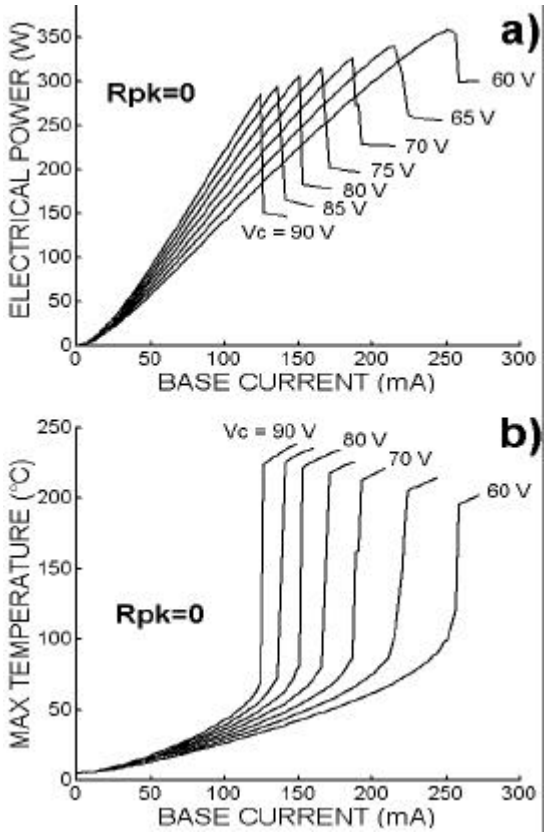

Figure 3. Total electrical power (a) and maximum temperature (b) vs. external base current for several values of the collector voltage.

This calculation method reached the convergence within few tens of cycles for bias conditions not before and far above a critical thermal situation such as the hot spot first onset, while more cycles were needed to obtain a steady state value in a bias range just above the hot-spot threshold. The whole simulation program was implemented in Matlab 5.0 environment.

6. Device simulations and experiments

The simulation program was applied to obtain the device electrical characteristics as a function of the total input base current I_b, for many values of the external collector voltage V_c and for several thermal dissipation capabilities of the die induced by different values of the bottom thermal resistance R_{pk}. All the curves were calculated from low injection currents until bias conditions above the hot-spot onset and using a heat sink temperature of 5 °C.

The figures 3a and 3b show the plots of the total electrical power and the maximum surface temperature respectively at various Vc voltage, for a sample with zero packaging resistance. The simulated curves clearly show the nature of the localized heating phenomenon, abruptly occurring once that a given threshold current was reached. The hot spot provokes an abrupt fall of the total power and current without any electrical premonitory signs. Under given and well controlled driving conditions, i.e. at constant base current at the external pins, the phenomenon can be considered as stationary. It means that a balance is established among the heating tendency of few localized cell absorbing current, the cooling of all the other cells, which are consequently less driven, and the heat removal at the bottom side of the chip.

It is worth noting that this stable hot spot condition can be reached only if the current gain h_{FE} of the individual transistor cell reaches a maximum and than decrease. In this case, the current concentration in a single or in few cells, leading to the hot spot creation, determines a decrease of the total dissipated power. Devices with different h_{FE} behaviors (i.e. monotonic) typically exhibit only destructive hot spots and cannot be driven in a stable hot-spot condition.

When a stable hot spot occurs and the temperature peak does not exceed the metal melting point, the presence of the hotspot does not necessarily imply the device failure. However such a condition is anyway a seriously dangerous situation because it strongly promotes the local electromigration of the metal stripes.

Not only the bias threshold of the hot-spot but also its onset localization on the die surface depend on thermal, electrical and metal layout conditions. In particular the factors which seem to determine, in different degree of importance, the heating point are the bottom thermal resistance and the distance of the cell from the base and emitter current pads. In fact, each row of cells driven with a single finger of metal have the additional electrical resistance R_{pe} and R_{pb} which increase with the distance from the current pads: therefore the far cells are driven with lower currents amounts than those localized close to current pads. The consequent current disuniformity, although it may be negligible at low bias conditions, seems to be mainly responsible in determining the over-heating site, at least for the present sample configuration.

The results of some simulation runs performed at Vc=70 V and for several values of the packaging thermal resistance R_{pk} are reported in figure 4 in terms of total power vs. total external base current.

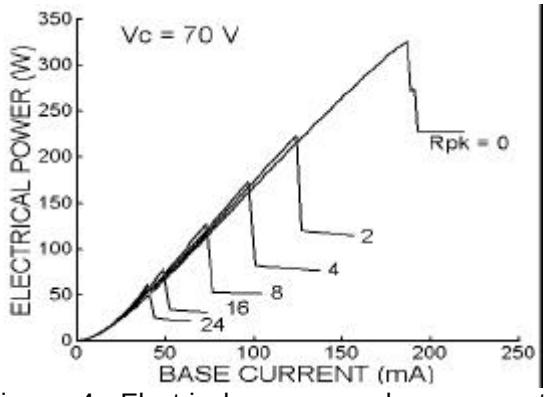

Figure 4. Electrical power vs base current for several values of the bottom thermal resistance

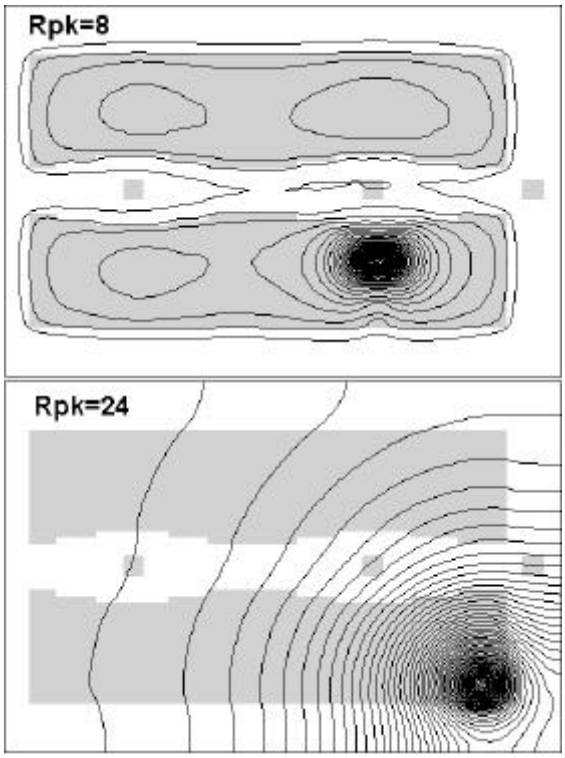

Figure 5. Device surface temperature maps at the hot-spot onset calculated for two different values of the bottom thermal resistance Rpk.

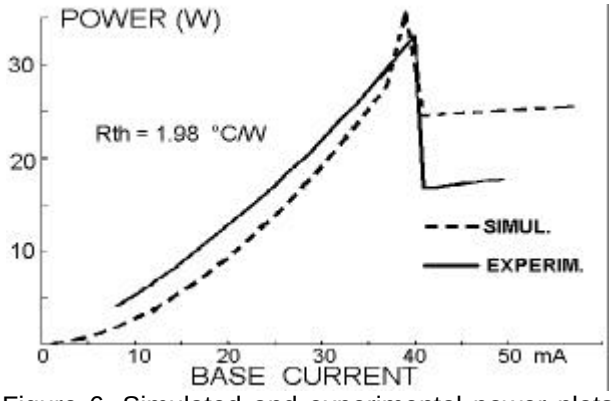

Figure 6. Simulated and experimental power plots for a sample with a thermal resistance of 1.98 C°/W.

Figure 5 shows the two temperature contour plots in hot spot regimes for two different values of the bottom thermal resistance ($R_{pk}=8x10^{-6}$, and $R_{pk}=24x10^{-6}$ °$CW^{-1}m^2$ respectively), drawn using isothermal lines with a steps of 4 °C.

From the above results it is evident that the presence of a package thermal resistance causes two main effects. At first the threshold power for the hot spot onset is a strongly decreasing function of R_{pk}, the most part of this decrease occurring just in the range of low values of R_{pk}. This clearly emphasizes the importance of package thermal design from the heat dissipation point of view in avoiding such unsuitable effects and keeping the maximum operating power as high as possible.

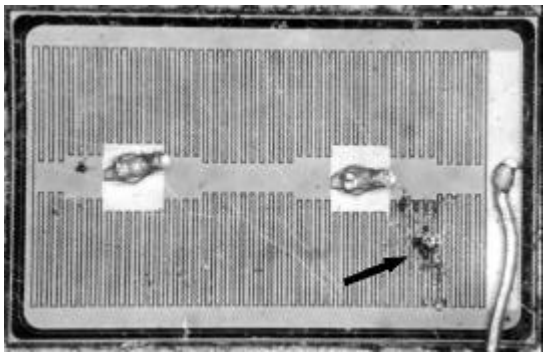

Figure 7. Photo of a device biased above the hot spot until failure. The arrow indicates the failure site.

The second effect concerns the migration of the hot spot onset site. As can be seen from figure 5, for low thermal resistance values the temperature peak is localized near the right emitter pad. In this position the additional resistance R_{pe} has its minimum value causing maximum current driving for the cells. As R_{pk} increases, the birth site progressively migrates toward the right lower corner of the chip area occupied by the cell grid. In fact the corner of the die is quite disadvantaged for the heat dissipation with respect to the central part of the surface due to the adiabatic boundary conditions at the lateral surfaces of the chip. This behaviour means that the thermal factors tend to prevail on the electrical one in determining the site. Therefore the balancing between electrical and thermal effects produces the peak progressive migration on the die surface.

Several experimental characterizations on real samples were performed in order to verify the electrical behaviour before and after the stable hot-spot. Figure 6 shows the experimental power characteristic compared with the simulated one for a sample having a bottom thermal resistance of 1.98 °C/W due to the sample holder and corresponding to $R_{pk}=20x10^{-6}$°$CW^{-1}m^2$. The data showed a good agreement, overall in the value of base current in which the hot-spot occurred. The disagreement in the plateau value after the event is probably due to non-linear thermal effects for temperatures higher than 200 °C. Also the localization of the event seemed to confirm the foreseen behaviour, as shown by the view of figure 7 of the sample biased at higher current until failure.

7. References

[1] P.L. Hower, D.L. Blackburn, F.F. Oettinger, S. Rubin, "Stable hot-spots and second breakdown in power transistors" Proceedings of PESC '76 Conference, 1976, pp.234-246.

[2] G. Breglio, P. Spirito, "Experimental detection of time dependent temperature maps in power bipolar transistors", MicroElectronic Journal, vol.31/9-10, 2000, pp. 735-739

[3] G. Breglio, N. Rinaldi, P. Spirito, "Thermal Mapping and 3D Numerical Simulation of New Cellular Power MOS Affected by Electro-Thermal Instability", MicroElectronic Journal, vol. 31/9-10, pp. 741-746, 2000