Thermal Management and Modeling of GaAs-based HBTs

Robert Anholt

Gateway Modeling, Inc., 1604 East River Terrace, Minneapolis, Minnesota 55414
anholt @ieee.org, 612-339-4239 | fax: 612-339-3995, www.gateway-modeling.com

Abstract

Thermal impedance models are described and results are compared for HBTs with varying emitter areas, cooled using thermal shunts, flip-chip mounting, or thinned wafers.

INTRODUCTION

HBTs can deliver high power using the smallest possible chip area. But with high dissipated power, the chip temperature will rise, so chips must be designed for the smallest possible thermal impedances (TIs) in order to maximize reliability, power added efficiency, and to prevent current collapse (or current hogging in multi-finger transistors [1]). The purpose of this paper is to compare different methods of minimizing chip or on-wafer thermal impedances: thermal shunt technology, wafer thinning, versus flip chip mounting. Analytical and numerical models are first presented to provide a basis for comparison.

THERMAL MODELS

We begin with some analytical formulas. We are interested TIs averaged over the emitter area at the base-emitter (BE) junction depth for HBTs. In the one-dimensional case, if one has a GaAs bar (thermal conductivity k=0.44W/cm/K) of thickness h and area A, where the power is applied and temperature is measured near the top and the other end is grounded, the TI is given by θ_A=h/k=1420x10^{-4} K-cm^2/W. Now if we arrange for the power source to be round and make the slab area infinite (or much greater than h), the average TI is given by θ_A=D, where D is the diameter of the heat source. Thus for 4μm diameter dots, one gets 4x10^{-4}K-cm^2/W. Notice that we quote the area-averaged thermal impedance in K-cm^2/W. Since dissipated power in a real HBT can be quoted in W per cm^2 of emitter area, the product gives the temperature rise in Kelvin or Centigrade for any emitter area.

The Green's function method and the method of images can be used to compute the average TI for an array of n LxW$_f$ heat sources arranged g apart, on an infinite slab of thickness h:

$$\theta_W=7.23\ [(\ln(2W/L)+1)R(g,W,L)-0.346W/h]\ \text{W-mm/K} \qquad (1)$$

where W=nW$_f$ is the total width and R is of the order of 0.88 to 3. This formula is suitable for on-wafer TI measurements for FETs and HBTs with fingers that are usually longer than they are wide (W>L). The units here are suitable for FETs where power is often quoted in W per mm of width. The former units can be regained by multiplying by 0.1L, where L is in cm.

Finally let us consider the TI in a flip-chip mount. Here the BE junction is between the power source and the ground. Assuming a perfect thermal ground at the emitter contact, θ_A is given by the sum of t_i/k_i, where t_i and k_i are the thickness and thermal conductivity of each layer between the BE junction and sink. For an AlGaAs/GaAs HBT with 300A of $In_{0.5}GaAs$, 300A of graded In_xGaAs, 1900A GaAs, 600A $Al_{0.3}GaAs$, 700A Al_xGaAs, we get 2.54x10^{-4}K-cm^2/W, where ~40%, 20%, and 40% comes from the InGaAs, GaAs, and AlGaAs layers. Flip-chip TIs can be no lower than this. The thickness of the ternary layers should be reduced as much as possible because they are thermally highly resistive.

Beyond these analytical models there are three approaches [2]:

(1) The Green's functions method is suitable for uniform media where the lateral chip dimensions are much greater than the thickness [3,4,5].

(2) The Fourier transform method is suitable for uniform media where the chip area is not much greater than the transistor area [6]. A Fortran computer program, TXYZ, can be found on the Internet [7].

(3) Finite elements are useful when one wants to include metal layers on top of the GaAs chip, vias, trenches, or the actual power dissipation profile [8,9]. Many commercial programs are available, but they are most useful for computing the peak temperature. Post processing must be used to get an average TI.

Figure 1 shows R values in Eq.(1) obtained from the Green's function method. This figure is similar to Fig.4.4 of ref.[4], except that formula quoted W/3h instead of Wln2/2h [5] and it did not average over L; at large g, those results went to 0.94, but these go to 0.88. This formula is inaccurate for h/W<~0.4 [5]; in some cases it will even predict negative TIs for lower h values.

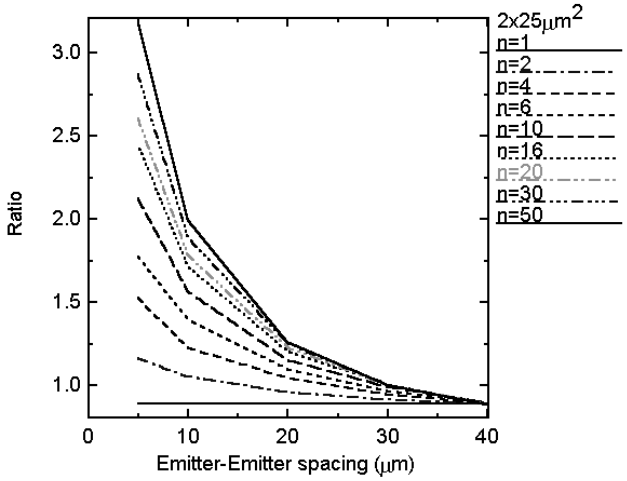

Fig.1. R values in Eq.(1).

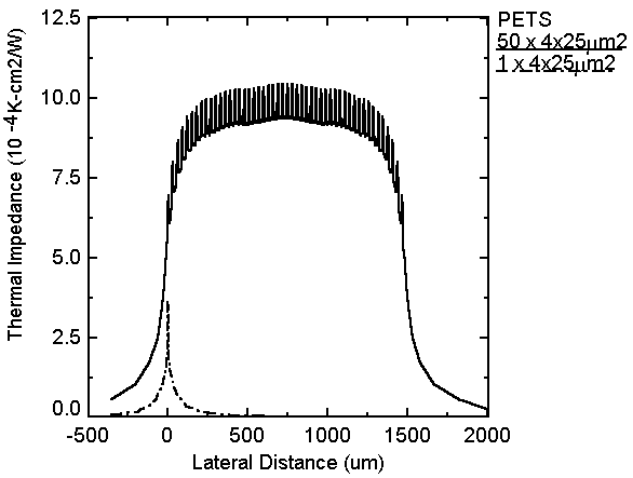

Fig.3. Temperature profiles in the X direction (Z=0)

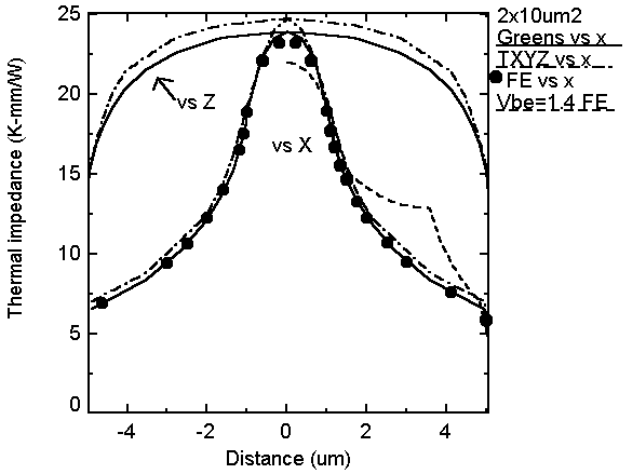

Fig.2. Temperature profiles in the X and Z directions (normalized to the power dissipation in W/mm)

Figure 2 compares temperature profiles (normalized to power so the units are those of a TI) for a 2x10μm^2 emitter element centered at X=0, Z=0 (Z being the long direction). The Green's and Finite element (FE) simulations assume infinite or >(500μm)^2 chip areas whereas the Fourier transform (TXYZ) program assumes a 202x210μm^2 chip area, thus has a slightly higher TI. The first three simulations assume the heat is applied uniformly and temperature is measured at the surface. The last FE simulation used the actual power dissipation profile for an InGaP/GaAs HBT [9] where the temperature is measured at the BE junction depth. The peak is slightly lower. Laterally heat is conducted away from the emitter by the base contacts and it reflects off the edge of the base-mesa (at ~3.5μm).

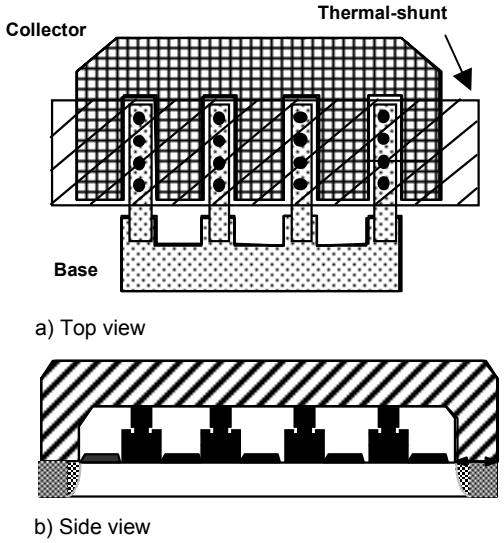

Fig.4. Thermal shunt layout

The main differences between the Green's or TXYZ simulations and the FE simulations with realistic layer structures, layouts, and realistic power dissipation profiles occur near the peak of the temperature profile. At distances g>30μm away from the peak, the FE, Green's, and TXYZ simulations predict similar profiles for the same conditions, thus the analytical models become more accurate as the transistor area increases. Figure 3 illustrates this by comparing temperature profiles for a 50 and 1-finger HBT.

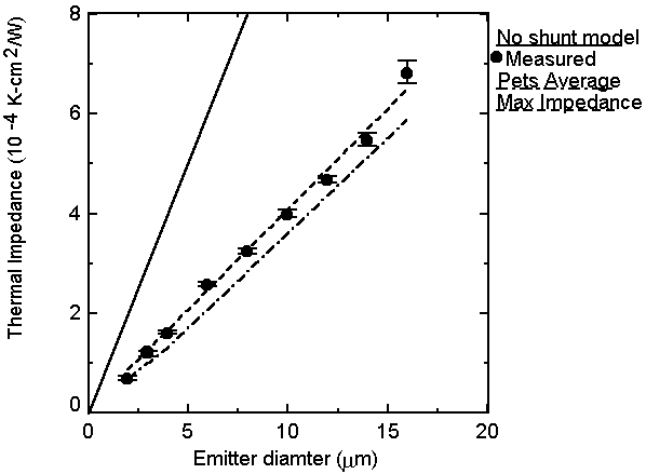

Fig.5. Measured and simulated TIs for single emitter dots

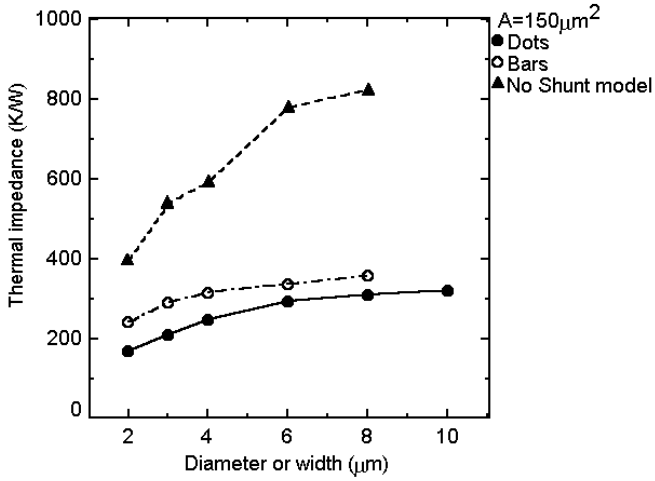

Fig.6. Measured and simulated TIs for HBTS with 150μm^2 total emitter area laid out as emitter dots or bars

TECHNOLOGY COMPARISONS

The remainder of this paper compares three technologies with the objective of deciding which is most suitable for cooling HBTs with varying emitter sizes: (1) wafer thinning, (2) thermal shunts, or (3) flip chip mounting.

Figure 4 shows the thermal shunt [10]. The emitter dots here are shown as black circles, which are surrounded by base metal fingers that are interdigitated with collector fingers. The thermal shunt is just a 10-15μm-thick air bridge that contacts the emitters, drawing heat out of the emitter, and spreading it over the substrate. Several papers describe this technology in more detail [9,10]. Figures 5 and 6 illustrate its effectiveness. In Fig.5 we compare area-averaged thermal impedances with the results of an analytical model without shunt (θ_A=D). With the thermal shunt the TI is reduced to about θ_A=0.4D. Figure 6 compares TIs for HBTs with 150μm^2 total emitter area arranged in dots or bars in multiple fingers. The dot layouts give lower TIs, in essence because the thermal elements are more spread out. Comparing the no shunt model [Eq.(1)] with the result for bar-layouts, the shunt reduces TIs by factors of 0.45 to 0.6. These TIs were measured on an unthinned wafer using the base-emitter voltage thermometer method [11] and are for an InGaP HBT.

As one goes to larger HBTs than 150μm^2 emitter area the flip-chip method becomes more attractive because the minimum TI in units of K-cm^2/W is independent of area. The difficulty with such methods is that as one begins to add bonding layers of thickness t_i and conductivity k_i between the emitters and the sink, one adds parasitic resistances t_i/k_iA_i. It helps to make the area A_i of these layers as large as possible. One has the same problem wafer thinning, but there A_i is at least as large as the chip which is often much greater than that which can be formed on top. These factors become less of a problem with very large HBTs, because as the total emitter area increases, A_i must also increase.

Another consideration is the ability of these technologies to minimize heating in adjacent fingers [9]. This is illustrated in Fig. 6, which shows the vertical temperature profiles for a 2-finger HBT where we apply power to only the finger at x=0 and we measure the temperature also at x=30μm. The BE junction is at y=0.23μm and the emitter contact is at y=0. For transistors with shunts or top-side sinks the temperature falls rapidly across the InGaP and InGaAs layers. The shunt is not quite as good as the flip-chip or shunt on a thin wafer (h=25). The flip-chip is most effective cooling adjacent fingers, because it pulls most of the heat out of the top, leaving less heat to spread in the substrate. Thinning the wafer is the traditional way used to cool large FETs with many fingers. It focuses the heat downward, which reduces the adjacent heating. As shown in Fig.6, thinning is not quite as good as the flip-chip for doing this, even for this very thin (h=25μm) case. The thermal shunt here is somewhat counter productive, because it contributes heat to the adjacent finger, as seen from the fact that the temperature is higher above the BE junction at x=30μm.

Figure 7 attempts to compare area-normalized TIs for shunt and flip-chip technologies. Hill *et al.* [12,13] reported TIs for 105μm^2 and 8000μm^2 HBTs. Their approach used a 3+10 or 3+20μm-thick sputtered + plated gold on top of the emitters then the sink was soldered using 25μm of Au/Sn. Bayraktaroglu *et al.* [14] reported for a variety of HBTs with emitter sizes from 56 to 480μm^2 a 37% improvement in the TI for a flip chip over ones using a 12μm thermal shunt.

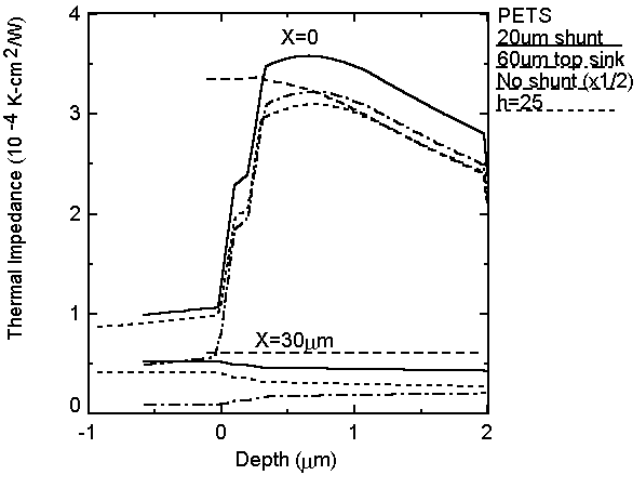

Fig.7. Temperature profiles versus depth for a 2finger HBT measured at X=0 (mid emitter where power is applied) and x=30μm (mid adjacent emitter; no power)

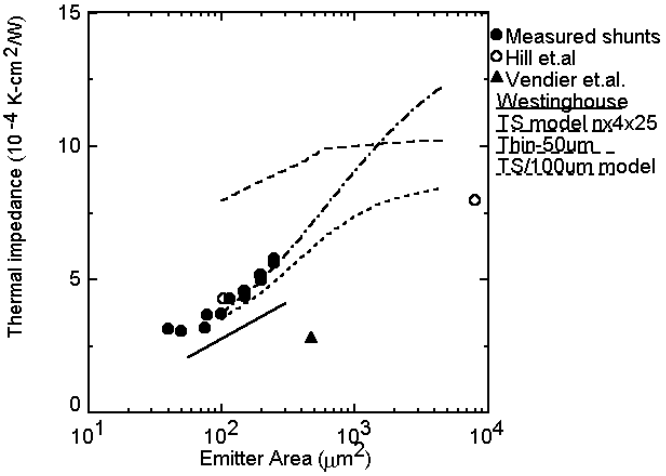

Fig.8. Measured and modeled TIs [12-16] versus total emitter area for different technologies

They fabricated 22μm-thick gold posts on top of the shunt, which were Au/Sn soldered to a sink. In Fig.7 the Westinghouse curve shows a 37% over the 3μm-bar thermal shunt results. (They gave no absolute numbers.) Vendier *et al.* [15] reported a flip-chip arrangement where the sink is connected to posts that make thermal compression to the emitters. Their TI is close to the lowest possible for a flip-chip (2-2.5x10^{-4} K-cm^2/W). Hill's TIs are 4 times greater for the large HBT.

CONCLUSIONS

With these data we can address the question: For a desired HBT emitter area, which layout gives the lowest possible TI? Table I summarizes TIs for several amplifiers, selected by emitter area. The power assumes operation at ~3V and 30000A/cm^2, which is near the peak f_t of many HBTs. I assume the fingers are L=4μm wide and they are laid out evenly 30μm apart. (Although it consumes more real-estate, a 10 finger HBT may more likely be laid out as 5 sets of 2 finger cells with intervening vias that gives lower TIs, in a bigger chip.)

TABLE 1:TEMPERATURE RISE, POWER, AND EMITTER AREA FOR SEVERAL LAYOUTS

Parameter	**LNA**	**Medium**	**High**	**Very-high**
Power (W)	0.02	.1	1	5
Emitter (μm^2)	20	100	1000	5000
Layout nxlxw	1x2x10	1x4x25	10x4x25	50x4x25
On-wafer (K)	42	88	131	152
TS (K)	25	41	99	135
Thinned (K) 50μm	38	80	100	102
Flip chip (K)	26	40	55	73
TS on thin 100μm	25	38	80	92

The on-wafer temperature rise uses Eq.(1) and the thinned wafer uses TXYZ assuming a 100μm border surrounds the fingers four directions. The TS results are based on measurements and PETS simulations. For the flip chip, the lowest possible temperature rise would be 20-25K coming from the conductivity of the InGaAs, GaAs, and AlGaAs layers. (The power density is fixed so if θ_A is constant, the temperature rise is constant.) The results shown scale as Hill's data (θ_A=[2+2(A/100)$^{0.25}$]x10^{-4}K-cm^2/W). All results used a constant k for GaAs.

Flip-chips are best for very large devices. Due to their simplicity thermal shunts are most effective for small HBTs. Wafer thinning is ineffective for small HBTs, but is useful for large ones. Shunts on thinned wafers (100μm being a conservative thinning), are not as effective as flip chips.

REFERENCES

[1] W. Liu, *1995 IEEE GaAs IC Symp.* pp.147-150.
[2] D.E. Dawson, *1994 IEEE GaAs IC Symp*. pp.285-288.
[3] D.Smith *et al*., 1986 IEEE SEMI-THERM Symp. Proc., paper A.5
[4] R. Anholt, *Electrical and thermal characterization of MESFETs, HEMTs, and HBTs* (Artech, 1995), ch.4.
[5] N. Rinaldi, Sol. State Elec. **44**, 1789 (2000).
[6] A.G. Kokkas, IEEE Trans Elect. Devices **21**, 674 (1974).
[7] http://www.decus.org/libcatalog
[8] P.W. Webb, IEEE Trans. Elect. Devices **40**, 867 (1993).
[9] R. Anholt, Sol. State. Elec. **42**, 849 (1998).
[10] R. Dettmer *et al*., *1996 IEEE MTTS,* pp.1196-1199.
[11] D. Dawson *et al*., IEEE Trans. Elect. Devices **39**, 2235 (1992).
[12] D. Hill *et al*., IEEE Microwave Guided Wave Lett. **5**, 373 (1995).
[13] D. Hill *et al*., IEEE Trans. Micro. Theory Tech. **45**, .2224 (1997).
[14] B. Bayraktaroglu *et al*, *1996 IEEE MTTS* pp.685-688.
[15] O. Vendier *et al*., *2000 IEEE GaAs IC Symp*. pp.157-160.
[16] T. Jenkins *et al*., IEEE Microwave Guided Wave Lett. **56**, 268 (1996).