A 2-D Numerical Study of Microscale Phase Change Material Thermal Storage for GaN Transistor Thermal Management

Xudong Tang, Richard Bonner, Tapan Desai, Angie Fan
Advanced Cooling Technologies, Inc
1046 New Holland Ave, Lancaster, PA 17601
xudong.tang@1-act.com

Abstract

A novel thermal management technology was explored to lower the peak temperature associated with high power GaN transistors in pulse application. The technology involves the use of an embedded microscale PCM heat storage device within the chip (near the active channels of the GaN device), which effectively increases the heat capacity of the material by taking advantage of the latent heat of the PCM. In this study, 2-D transient thermal models were developed to characterize the thermal behavior of GaN transistors with micro-scale PCM heat storage device. The model is capable of computing the spatial-temporal temperature distribution of the GaN transistor as it is rapidly pulsed and captures the formation and evolution of hot spots that form within the device. The model also captures the PCM melting behavior and latent heat absorption during the transient.

The use of a PCM can effectively control the hot spot temperature by absorbing a significant portion of the transient heat input. As shown in this modeling study, the use of PCM heat storage in GaN transistors reduces the GaN hot spot temperature for a given heat input. Alternatively, the maximum allowable GaN heat input can be increased with the use of PCM. At a given heat input flux of 5×10^5 W/cm^2, for example, the use of PCM heat storage can lower the peak temperature by 21~22°C, relative to transistors without PCM (baseline), regardless of the duty cycle ratio. In addition, a transistor with PCM heat storage can accommodate much higher joule heat generation without exceeding the maximum allowable temperature limit, 180°C. In this study, the modeling results show that by integrating a PCM that has a 140°C melting point in a 5μm×6μm groove configuration, the critical heat flux can be increased from 13.34×10^5 W/cm^2(baseline) to 16.8×10^5 W/cm^2(with PCM), a 26% improvement.

Key PCM design parameters were identified in this modeling study: (1) PCM amount; (2) PCM melting point; and (3) PCM groove structure. Their coupling and the impact on design optimization require further investigation.

Keywords

GaN Transistor, Thermal Modeling, PCM, Heat Storage

Nomenclature

AR	Interfacial thermal resistance
Cp	Specific heat
E	Internal energy in a control volume
k	Thermal conductivity
Q	Heat flux
t	Time
T	Temperature
u	Velocity

Symbols

α	Two phase ratio
ρ	Density
Δh	Latent heat
Δt	Time step
Δx	Mesh size in X axis
Δy	Mesh size in Y axis

Subscripts

|	
MF	PCM melt front
PCM, l	PCM liquid phase
PCM, s	PCM solid phase
E, W, N, S	East, West, North, South mesh nodes
$GaN - Si$	GaN-to-Si interface
$PCM - Sub$	PCM-to-Substrate interface
$Si - PCM$	Si-to-PCM interface
$Si - Sub$	Si-to-Substrate interface

1. Background
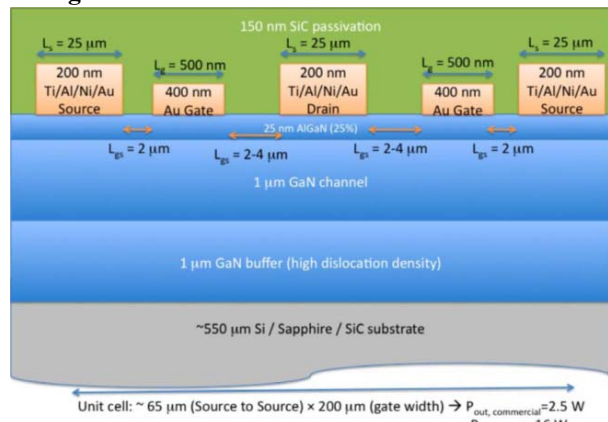
Figure 1: Structure of the unit cell of a GaN power amplifier. (The field plate structures are not shown)

In recent years, wide bandgap transistors (SiC MESFETs and GaN HEMTs) have appeared on the market for high power RF/microwave transistors [1, 2]. They offer higher power density and higher voltage operation, which are associated with lower parasitic capacitances and higher load-line dynamic resistance, and hence can be used in wider bandwidth applications. However, heat removal from these semiconductor devices and electronic packages remains a critical issue in the chips' electrical performance and life cycle [3, 4]. With a continuing increase in the levels of integration and the introduction of new chip and interconnect architectures, the background heat fluxes have become very high and must be rejected. The heat removal challenge is further exacerbated by the "pulsed" operation, where rapid

temperature transients are continuously experienced in chips of many communication applications.

As shown in Figure 1, a typical state-of-the-art GaN power amplifier for communication and radar applications, with a output power of 450W and an efficiency of 65%), operates in a pulsed mode of 2 μs pulse width with 100 μs period at 3.4 GHz, which has an average heat dissipation of ~30-50 W/mm^2. Figure 2 illustrates the heat generation within a typical semiconductor transistor. The heat flux generated from these individual transistors is then conducted through multiple layers of various metals, interface materials, etc. in what is commonly known as the thermal stack-up. Each layer adds to the overall thermal resistance of the device, resulting in high peak junction temperatures.
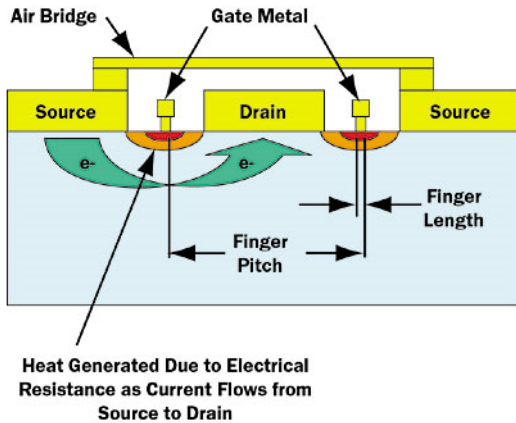

Figure 2: Schematic detailing heat generation within a transistor [5]

Simulations of the cross-sectional profile of the lattice temperature distribution in standard AlGaN/GaN HEMTs using Silvaco ATLAS were performed [6]. A representative simulation shows in Figure 3 for a standard AlGaN/GaN HEMT heat is generated in the GaN channel due to the joule heating and the hottest temperature occurs in the AlGaN and GaN layers around the corner of the gate contact close to the drain. This is due to the high electric field and the large current density in the region. During operation, GaN device hot spot temperatures in transistors can oscillate between 180°C and 80°C during the rapid transients. Continuous thermal cycling and high maximum junction temperatures cause cyclical thermal stresses that may lead to reduced reliability. Directly dissipating such a large amount of waste heat during the short pulse operation while maintaining the GaN device temperature within the limit is becoming impracticable, and therefore requires advanced chip/die level thermal management.

A micro-scale thermal storage approach was therefore investigated in this work and takes advantage of the transient pulse behavior common to high power transistors in communication applications. The idea is to increase the effective heat capacity of the material near the active channels of the GaN devices by embedding a phase change material (PCM, material that changes phase between liquid and solid at a given temperature) in close proximity to the active channels. During the heat generating pulses of each duty cycle, the heat generated in the GaN devices can be temporarily stored in the heat of fusion of the solid phase PCM as it melts. During inactive times, heat can be dissipated as the liquid phase PCM re-solidifies. By using PCM with an appropriate melting point, the peak junction temperature can be reduced and the transient temperature fluctuation can be stabilized, which can result in improved device reliability.
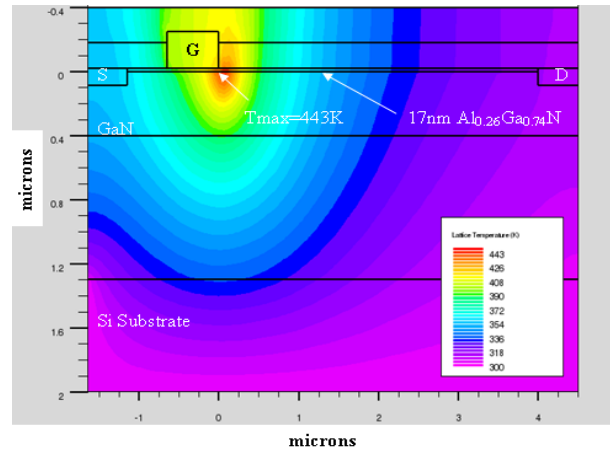

Figure 3: Lattice temperature distribution in standard AlGaN/GaN HEMTs using Silvaco ATLAS (V_{ds}=30V, V_{gs}=0V, 30nm thick GaN buffer).

This paper focuses on a 2-D transient thermal model for GaN transistors needed to evaluate the effectiveness of the microscale PCM thermal storage mechanism for reducing peak junction temperatures and improving thermal stability. This modeling effort was supported by parallel research efforts devoted to (1) a transistor level molecular dynamics (MD) simulations to simulate the heat flow amongst various nanometer sized transistor materials, and (2) a Boltzmann Transport Equations (BTE) model to calculate heat generation and transport in GaN by taking into account the non-equilibrium nature of electrons and phonons. Key parameters, such as interfacial boundary resistances from MD simulation and hot spot joule heating from BTE modeling, were incorporated to improve the model fidelity. The thermal model is therefore a useful tool to characterize the overall transient thermal behavior of high power transistors in pulse operation, and also provide preliminary guidelines to microscale PCM thermal storage design for GaN transistor thermal management.

2. Numerical Methodology

A typical unit cell of a GaN transistor is shown in Figure 1. It includes two sources, two gates, and one drain. When an electric field is applied to the transistor, electric current flows in the GaN channel from the sources to the drain, which results in joule heating due to the energy transfer from the electrons to the lattices. The heat is eventually dissipated through the GaN buffer and substrate to a heat sink. As indicated in Figure 3, most of the heat is generated at two locations around the corner of the gate contact close to the drain.
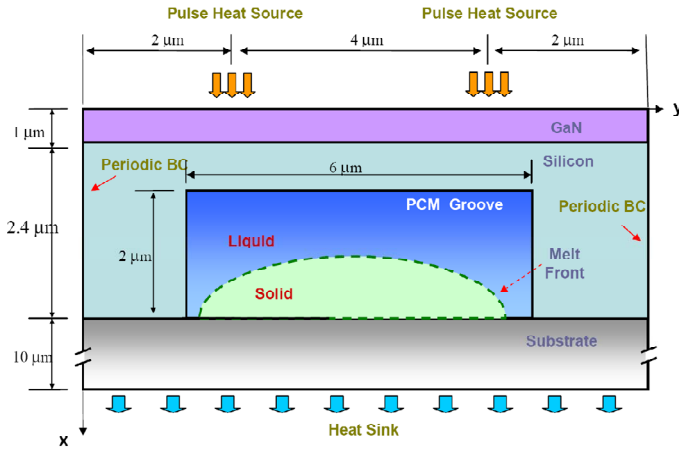
Figure 4: Thermal model of a GaN Transistor Unit Cell with Micro-Scale PCM Groove.

The concept of a GaN transistor embedded with microscale PCM thermal storage device is described in Figure 4. The channel length (~100μm) is usually much longer than the channel width. In this 2-D numerical model, it is therefore assume to have a uniform temperature distribution along the channel length. The transistor was simplified to an 8 μm wide periodic unit cell wi th two localized pulse heat sources on the GaN surface. The transistor can be seen as consisting of many identical unit cells. The whole device was divided into three layers: (1) a 1 μm GaN layer that acquires the heat, (2) a 4 μm silicon layer with embedded PCM as the thermal storage device, and (3) a 10 μm substrate attached to a heat sink. The PCM was stored in the silicon layer groove with a dimension of 2 μm in thickness and 6 μm in width.
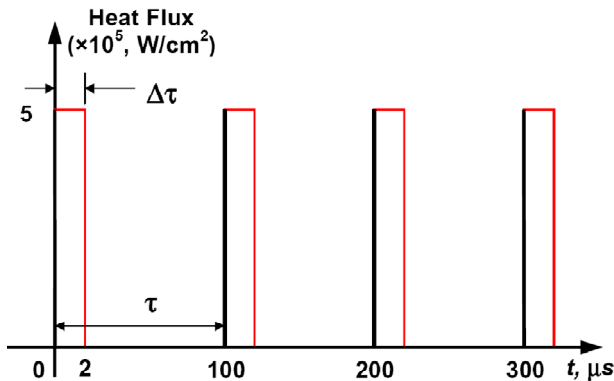
Figure 5: Heat Flux Boundary Condition Applied on GaN Surface Hot Spots.

In the transient thermal model, it was assumed that localized heat fluxes into the unit cell resulted in two hot spots. The heats sources were applied over 0.1μm section to simulate concentrated joule heat generation. Identical pulse heating profiles were applied at these hot spots on the GaN layer surface, as shown in Figure 5. The duty cycle of the heating profile was τ = 100μs. It included τ = 2~10 μs "ON" time followed by an "OFF" time for the remainder of the cycle. During the pulse "ON' time, a constant heat flux (~5×10^5 W/cm^2) boundary conditions were applied at the hot spots. During the "OFF" period, zero heat flux (adiabatic) boundary conditions were applied at the heat sources. Periodic boundary conditions were applied on the left and right sides of

the unit cell. At the bottom of the unit cell, the substrate surface evenly dissipated the heat to a 25°C heat sink. Interfacial boundary thermal resistances were provided for every two adjacent layers. The Si-to-PCM interfacial thermal resistance, AR_{Si-PCM}= 2.7×10^{-8} K·m^2/W, was obtained by MD analysis. The initial temperature in the entire unit cell was set as the same as the heat sink temperature.

Regarding the selection of the PCM, desirable properties for the chip-level transient thermal storage and release include: (1) an adequate melting temperature, (2) a high thermal conductivity, and (3) commercially available. Based on these criteria, low-melting temperature metals and binary alloys stand out. Solder PCM, such as Indium or Indium/Tin alloy, were therefore chosen in this research over normal wax based PCM, because of their higher thermal conductivity (34-81W/m-K), adequate melting point (118-160°C), substantial fusion heat (~29 kJ/kg), and commercial availability. A summary of the material used in the transistor unit cell model are listed in Table 1.

The PCM volume expansion/contraction due to the density difference (about 2.6%) between the solid and liquid phase was not considered in this study. The void formed in the PCM during the thaw/freeze cycle is most likely to happen in the center of the PCM groove, where is the last location to be solidified. Therefore, the impact of the voids on the interfacial thermal resistance is not significant.

Table 1: Material bulk properties [7-9].

	ρ (kg/m^3)	Cp (J/kg-K)	k (W/m-K)	α (×10^{-5}m^2/s)
GaN	6150	441.3	130	4.790
Si	2329	705.4	149	9.070
In (solid)	7214	258.6	80	4.289
In (liquid)	7026	255.7	36	2.004
In$_{52}$Sn48	-	-	34	-

The numerical methodology used to solve 2-D time dependant energy conservation equation was the finite volume method with a five-point stencil explicit discretization scheme, which was 2nd order special accuracy and 1st order temporal accuracy. The whole calculation domain was discretized with a fixed mesh. The mesh size was set to be $\Delta x = \Delta y = 0.1 \mu m$. Therefore, according to the explicit scheme stability condition,

$$\Delta t < \frac{\rho Cp \Delta x^2}{2k} \approx 2.2 \times 10^{-10} \mu s$$

$$\frac{\partial E}{\partial t} = \frac{\partial Q_x}{\partial x} + \frac{\partial Q_y}{\partial t}$$

The definition of control volumes were described in Figure 6. The temperature and material properties on a mesh node represented the temperature and properties of the surrounding control volume. For a control volume in the bulk material, the energy balance equation can be discretized into:

$$\left(E_p^{n+1} - E_p^n\right)\Delta x \Delta y = Q_W \Delta y - Q_E \Delta y + Q_N \Delta x - Q_S \Delta$$

where Q_W , Q_E Q_N and Q_S are heat flows from four adjacent control volumes.
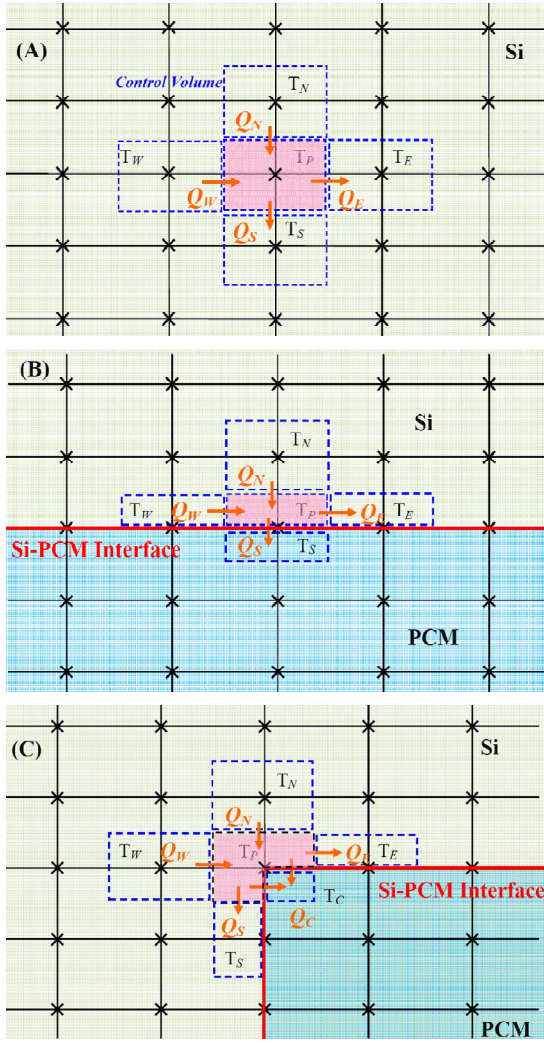
Figure 6: Five-point stencil discretization of finite volume method in bulk material and at material interface.

To accurately capture the temperature gradient across the material interface, the control volume on the material interface was split into two (as indicated in Figure 6(b). For a control volume on a straight material interface, the energy equation can be discretized into:

$$\left(E_p^{n+1} - E_p^n\right)\Delta x \frac{\Delta y}{2} = Q_W \frac{\Delta y}{2} - Q_E \frac{\Delta y}{2} + Q_N \Delta x - Q_S \Delta x$$

where $Q_S = \frac{T_P - T_S}{AR_{Si-PCM}}$, AR_{Si-PCM} was the interfacial boundary resistance between the Si layer and PCM material. The discretized energy equation on the edge of the PCM groove interface can be described as Figure 6(c):

$$\left(E_p^{n+1} - E_p^n\right)\frac{3}{4}\Delta x \Delta y$$
$$= Q_W \Delta y - Q_E \frac{\Delta y}{2} + Q_N \Delta x - Q_S \frac{\Delta x}{2}$$
$$- Q_C \left(\frac{\Delta x}{2} + \frac{\Delta y}{2}\right)$$

where $Q_C = \frac{T_P - T_C}{AR_{Si-PCM}}$

Two-phase ratio α was used for control volumes in the PCM groove, indicating the percentage of solid phase PCM

that turned into liquid. When $\alpha = 0$, it means the PCM in the control volume is all solid. When $\alpha = 1$, it means the PCM in the control volume is all liquid. When $0 < \alpha < 1$, it means the control volume is in two phase state. The internal energy of control volumes in PCM region can be written as:

$$E = (1 - \alpha)\rho_{PCM}^{solid} Cv_{PCM}^{solid} T + \alpha\left[\rho_{PCM}^{solid} Cv_{PCM}^{liquid} T + \rho_{PCM}^{solid} H_{PCM}\right]$$

where H_{PCM} is the PCM fusion heat. The effective thermal conductivity in the control volume is:

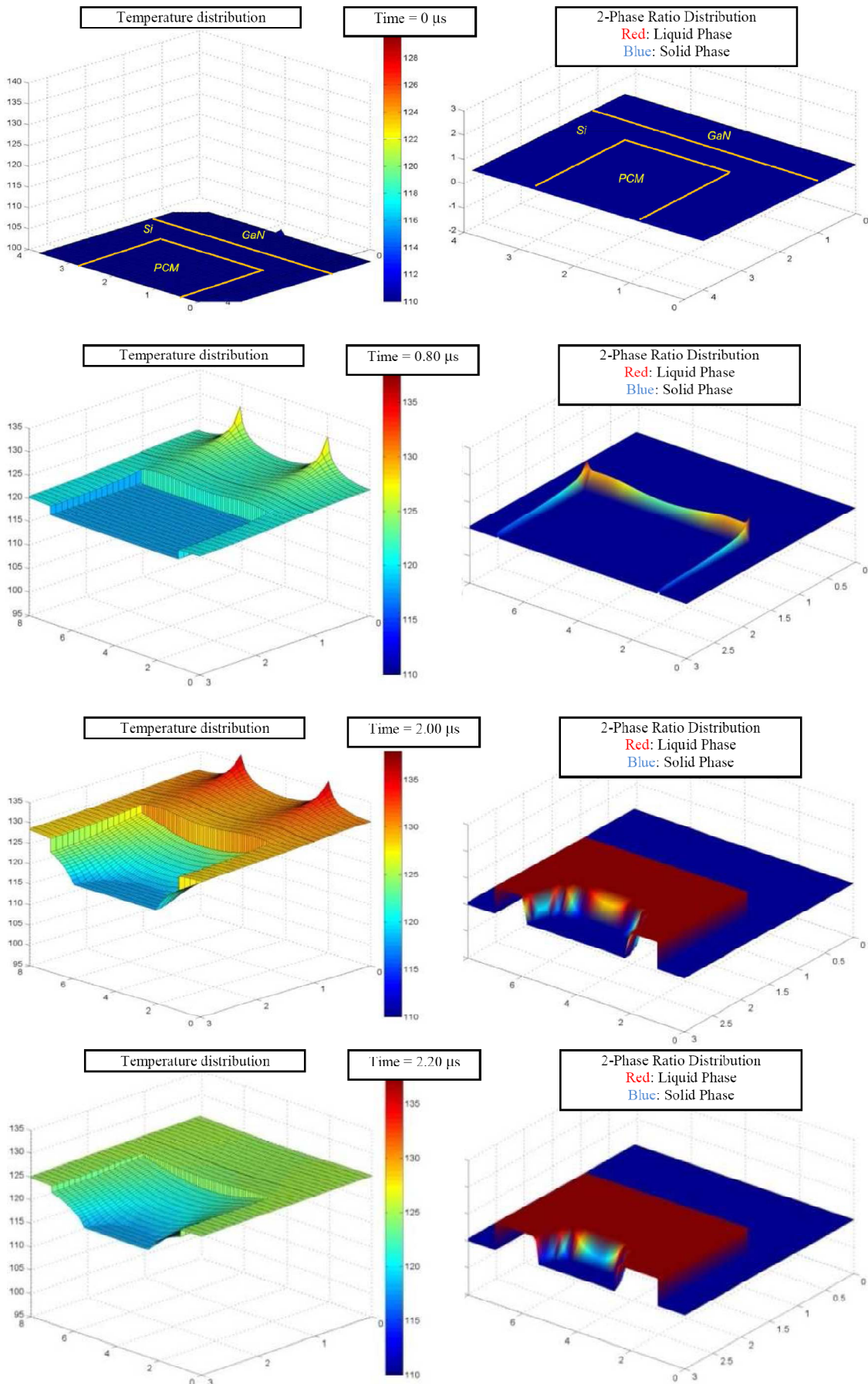$$k_{PCM} = \alpha k_{PCM}^{liquid} + (1 - \alpha)k_{PCM}^{solid}$$

The temperature of a control volume in PCM two-phase state equaled to the melting point. The change of the two-phase ratio α in a control volume can be deduced by the change of the internal energy and the latent heat.
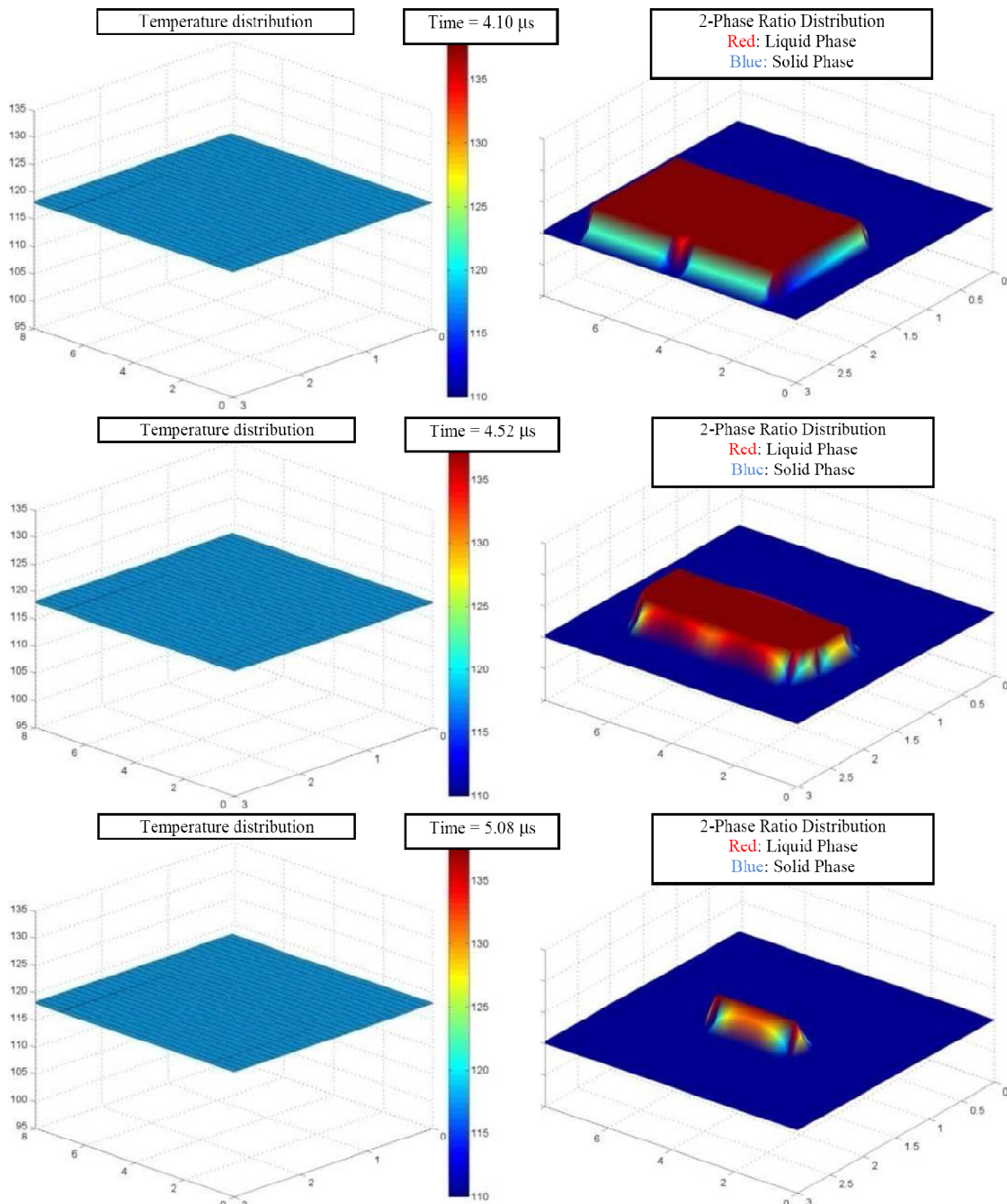
3. Results and Discussion

A pulse mode transistor with grooved PCM thermal storage structure was simulated. The transistor simulation started from its "Cold" state. The initial temperature was equal to the heat sink temperature. After 10~15 duty cycles (1500 μs), the transistor reached "steady" operation, where the temperature variation during one duty cycle was identical to the previous cycle. Figure 7 shows the snap shots of temperature and two phase ratio α contours of the unit cell during a duty cycle after the transistor reached "steady" operation.

The working principle of a transistor with PCM thermal storage device in "steady operation" mode can be illustrated as: When the transistor was turned "ON", hot spots developed in the GaN layer. Since more heat was supplied than the heat being dissipated during this period, the transistor temperature, as well as the PCM temperature, increased. The solid PCM in contact with Si layer was the first place that melted, since that was the location that had the highest temperature in the PCM. A liquid/solid phase interface (melt front) was thus formed in the groove. The PCM used the fusion heat to absorb sensible heat in the transistor and control the GaN layer temperature. As more heat was supplied, the melt front formed an arc shape and advanced towards the center of the PCM-to-Substrate interface. At the end of the "ON" period, most of the PCM melted.

When the transistor was turned "OFF" in the subsequent period, the hot spots in the GaN no longer existed. However, the temperature in the GaN and Si layer temperature still remained higher than the PCM melting point. The PCM continued to absorb sensible heat from the GaN and Si layers until the GaN, Si and PCM temperature equalized. Afterwards, the liquid phase PCM started to solidify. At this stage, the PCM became the heat source in the device. Heat flowed from the PCM through the surrounding Si and substrate layer to the heat sink. The initial location of the liquid PCM that solidified was at the Si-to-PCM and PCM-to-Substrate interfaces. The melt front retreated from the perimeter towards its center. At the end of the "OFF" period, all the liquid phase PCM turned into solid.

Temperature distribution — Time = 0 μs

2-Phase Ratio Distribution
Red: Liquid Phase
Blue: Solid Phase

Temperature distribution — Time = 0.80 μs

2-Phase Ratio Distribution
Red: Liquid Phase
Blue: Solid Phase

Temperature distribution — Time = 2.00 μs

2-Phase Ratio Distribution
Red: Liquid Phase
Blue: Solid Phase

Temperature distribution — Time = 2.20 μs

2-Phase Ratio Distribution
Red: Liquid Phase
Blue: Solid Phase

Figure 7: Temperature and PCM phase change during a duty cycle (118°C PCM melting point).

At the beginning of the "steady" duty cycle, the GaN layer, Si layer and PCM started from 97°C. The PCM in the groove was in the solid phase. A low melting point PCM, 118°C, was used in this simulation. Once the heat sources were applied, hot spots appeared on the GaN surface, with a temperature about 7°C higher than the rest of the GaN layer. The temperature distribution along the GaN-to-Si interface was quite uniform, indicating good heat spreading of the GaN layer due to its micro-scale dimension. A 3~5°C temperature drop existed on the Si-to-PCM interface with the given Si-to-PCM thermal resistance. At about t=0.8μs, it can be observed that the corners of the PCM groove in contact with the Si layer reached the melting point, 118°C. PCM started melting from those two corners. Then the melt front extended to the

entire Si-to-PCM interface to form an arc-like shape two-phase interface. As more heat continued to be applied, the melt front propagated inward to the center of the PCM-to-Substrate interface, and more heat was stored in the PCM latent heat. During this process, the solid PCM temperature remained around the melting point and the GaN hot spot temperature was effectively controlled. At t=2μs, the end of "ON" period, the device temperature reached the highest point. The hot spot temperature was 135°C. Most of PCM in the groove turned into liquid.

After the transistor was turned "OFF", the hot spot no longer existed. However, during the time frame between t=2μs and t=4.1μs, the temperature of the GaN and Si layers were still higher than the PCM melting point. The PCM

continued to absorb sensible heat from the GaN and Si layer. At t=4.1μs, the temperature of the GaN layer, Si layer and PCM equalized. After that, the liquid PCM in the groove was turned into the heat source of the transistor. The absorbed latent heat was released, flowing from the PCM through the Si and substrate layers to the heat sink. The liquid PCM started to solidify from where it was in contact with the Si layer, substrate and solid PCM. As the PCM continued to solidify, the melt front propagated from the perimeter to the center. At the end of the duty cycle, the temperature of the GaN layer, Si layer and PCM returned to their initial value of 97°C.
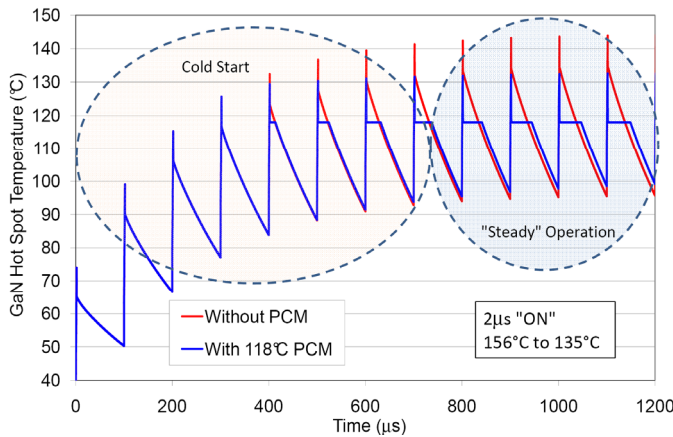

Figure 8: Comparison of GaN transistor heat source temperature variations from cold start to "steady" operation

A simulation was also done to compare the temperature profile of a transistor with or without PCM heat storage design. The transistor heat source temperature (GaN surface hot spot) variation starting from room temperature was recorded and shown in Figure 8. For the first few cycles, the transistors showed a pulse behavior with gradually increased temperature. After 10 to 15 duty cycles, both transistors were at "steady" operation. The pulse temperature variation in each duty cycle became identical.

Until the PCM reached the melting point, the hot spot temperature of both transistors reached the same peak temperatures. When the temperatures of the device reached the melting temperature of the PCM, the hot spot temperature of the transistor with PCM separated from the baseline design. It can be seen that the hot spot temperature of the transistor without PCM (baseline) went up to 156°C at the end of "ON" period, while the maximum hot spot of the transistor with PCM was 135°C. The temperature reduction was 21°C, which translates into a 20% reduction in the hot spot temperature. Furthermore, it took only another 2.1 μs for the temperature of the transistor with PCM to drop to 118°C, instead of 30 μs for the baseline transistor. Based on this numerical modeling, incorporating micro-scale PCM close to GaN channel can effectively reduce the overheating risks in transistors.

Keeping the total heat input per duty cycle constant, transistors with different duty cycle ratio, from 3 to 5 μs, were investigated in Figure 9. It was clear that with the increase of the duty cycle ratio, the hot spot temperature becomes less significant due to the lower heat flux rate applied. The benefit of using PCM heat storage device was still obvious. The hot

spot temperature reductions were 21°C and 22°C for 3 μs and 5 μs "ON" duty cycle, respectively.
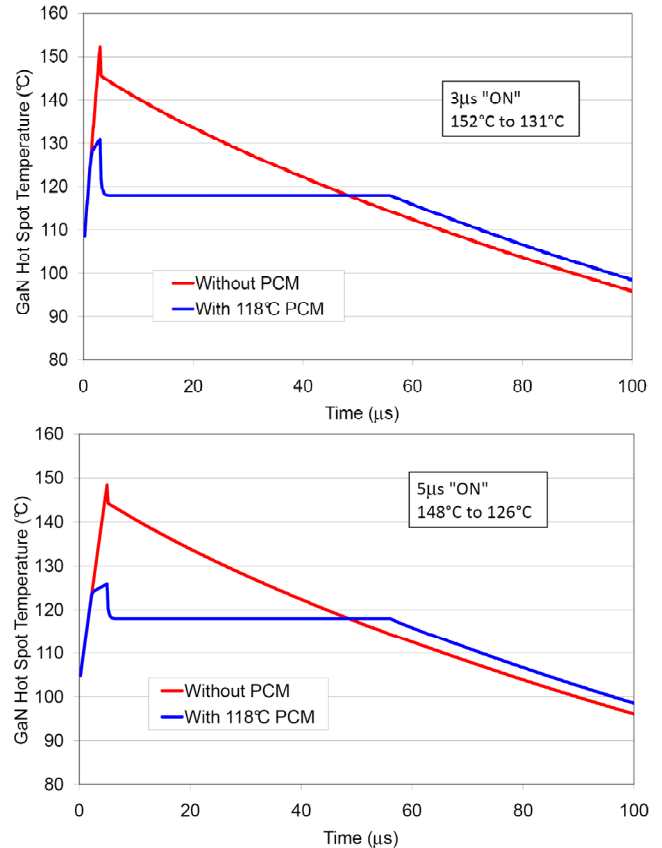
Figure 9: Effective of duty cycle ratio on GaN hot spot temperature (3μs and 5μs "ON")
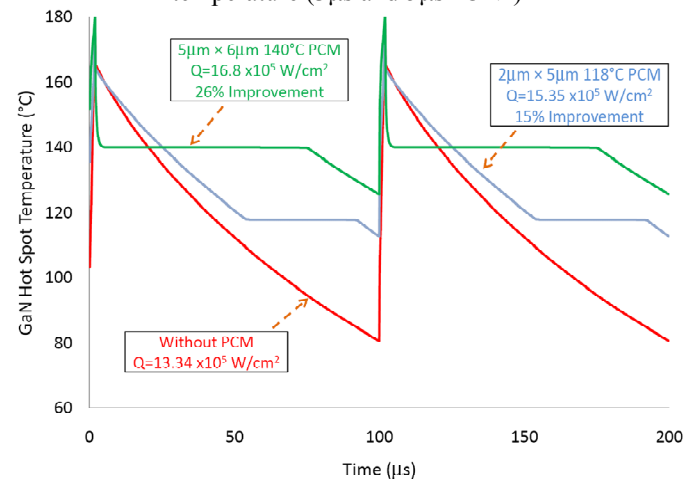
Figure 10: GaN Hot Spot Temperature Variation with Different PCM Melting point (2μs "ON" period).

Based on the calculations presented in this study, it can be deduced that a transistor with PCM heat storage can operate with a much higher joule heat generation without exceeding the maximum allowable temperature of the GaN material, 180°C. The modeling results in Figure 10 show that by integrating 140°C melting point PCM into a 5μm×6μm groove, it can raise the maximum input heat flux from $13.34×10^5$ W/cm² for the baseline device to $16.8×10^5$ W/cm², a 26% improvement. It was also obvious that the amount,

geometry and melting point of the PCM have significant impact on the performance of the PCM heat storage device. A quick comparison showed that only a 15% improvement can be achieved by incorporating a 2μm×5μm PCM with 118°C melting temperature. It was clear that the PCM groove geometry as well as the amount of the PCM had to couple with the selection of different melting point PCM in the design optimization. Further investigation on these design parameters will be conducted to thoroughly understand the design principles of PCM heat storage in transistors.

4. Conclusions

A novel thermal management technology for high power transistors in pulse applications is presented which involves the use of a micro-scale PCM heat storage device embedded near the active channel. Two-dimensional transient thermal models were developed in this work to characterize the thermal behavior of GaN transistors with and without micro-scale PCM heat storage device. Fast pulse joule heat generation and nanometer scale interfacial boundary resistances that calculated from BTE and MD models were integrated in this model. The model was capable of predicting fast transient temperature variation for microsecond duty cycles. It was also able to capture PCM melting behavior and latent heat absorption during this fast transient, and predicted the performance improvement of PCM heat storage technology, in terms of reduced GaN hot spot temperature or enhanced GaN hot spot heat input flux.

As discussed, the basic idea of the technology was to increase the effective heat capacity of the material near the active channels of GaN transistor, which absorbs transient heat input with its latent heat and reduces the junction peak temperature during "ON" period, and then dissipates the stored heat during the inactive "OFF" period. The benefit of micro-scale PCM heat storage in a GaN transistor included significantly reducing the GaN layer hot spot temperature, and preventing transistor from being overheated. At a given heat input flux of $5×10^5$ W/cm^2, the PCM heat storage device provided a 21~22°C temperature reduction of the GaN hot spots compared with baseline transistors, regardless of the duty cycle ratio. On the other hand, the transistor with PCM heat storage can operate with much higher joule heat generation without exceeding the maximum allowable temperature of GaN material, 180°C. The modeling results showed that by integrating 140°C melting point PCM in a 5μm×6μm groove, it can raise the allowable heat flux from $13.34×10^5$ W/cm^2 for the baseline device to $16.8×10^5$ W/cm^2, a 26% improvement.

A few of key PCM design parameters were identified during this model effort and can be further investigated in the future work. Specific design parameters that can be optimized include: (1) PCM charge amount; (2) PCM melting point; (3) PCM groove structure. These parameters are not independent of each other. Their coupling effects should be carefully studied to maximize the benefit of this micro-scale heat storage technology.

Acknowledgments

References

1. U. K. Mishra, L. Shen, T. E. Kazior, and Y.-F. Wu, "GaN-based RF power devices and amplifiers," Proc. Of the IEEE, vol. 96, no. 2, 287 (2008)
2. A. Koudymov, X, Hu, K. Simin, G. Simin, M. Ali, "Low-Loss High Power RF Switching Using Multifinger AlGaN/GaN MOSHFETs", IEEE Electron Device Letters, Volume 23, 2002, pages 449-451
3. H. Oprins, J. Das, W. Ruythooren, R. Vandersmissen, B. Vandevelde, M. Germain, "Thermal Modeling of Multi-Finger ALGaN/GaN HEMT's", Therminic, Belgirate, Italy, 28-30 September 2005
4. U.K. Mishra, P. Parikh, Y.F. Wu, "AlGaN/GaN HEMTs: An overview of device operation and applications", Proceedings of the IEEE, Jun 2002, Volume: 90 Issue:6, 1022 – 1031
5. S.Boslet, R.M.Morales, "Thermal Modeling Approaches for GaAs Semiconductors", Electronics Cooling, Feb 2008
6. T. Palacios and U. K. Mishra, "Modeling and Simulation of AlGaN/GaN High Electron Mobility Transistors," chapter in "Nitride Semiconductor Devices: Principles and Simulation," edited by J. Piprek. Wiley-VCH (2007).
7. F. Gronvold, J. Therm. Anal.,13, 41 (1978)
8. M. J. Duggin, J. Phys. F: Metal Phys., 2, 433 (1972)
9. I. Ziborak-Tomaszkiewicz, E. Utzig and P. Gierycz, J. Therm. Anal. Calorim., 91, 329 (2008)