Efficient Thermal Management of Power MOSFETs,

Vital to Improve the Reliability

Of Your Power Conversion Application

Dr. Axel Kleinitz

ON Semiconductor
Zamdorferstr. 100
81677 Munich
Germany

☎ +49-89-93 08 08-0

axel.kleinitz@onsemi.com

ABSTRACT

When it comes to the design of a power converter, the thermal characteristics of the used components - and of MOSFETs in specific - do not seem to represent the major focus. Traditionally, the main concern is to solve a given problem and leave the thermal management to the end of the design. In many cases the neglect of the component's thermal behavior leads to reduced system reliability.

The purpose of this paper will therefore be to explain the main thermal implications and show how they would have an impact on the reliability of your system. The paper will therefore be split in three main sections:

1. THE SELECTION OF THE RIGHT MOSFET BASED UPON THE PROCESS TECHNOLOGY, which will describe the two main process technologies, it's pro's and con's and then move on with a phenomenon that may endanger your system reliability, that is, the thermal run-away.

2. THE SELECTION OF THE RIGHT MOSFET BASED UPON ITS THERMAL RESISTANCE; based upon the conclusions taken under 1., this section will get more detail and focus on the thermal resistance only and how this parameter can question the reliability of your whole system.

3. THERMALLY SELF-PROTECTED MOSFET; the objective of this section will be to demonstrate one way to prevent your system from unwanted damages by using self-protected MOSFETs.

HISTORICAL REVIEW

The first patents based upon the physical principles of a field effect transistor were registered in 1928 by Julius Edgar Lilienfeld, Germany. In 1947 John Bardeen and Walter Houser Brattain succeeded in building the first bipolar transistor, a so called point-contact transistor. Its principles of operation were quite similar to those used nowadays, yet by far much larger than conventional ones.

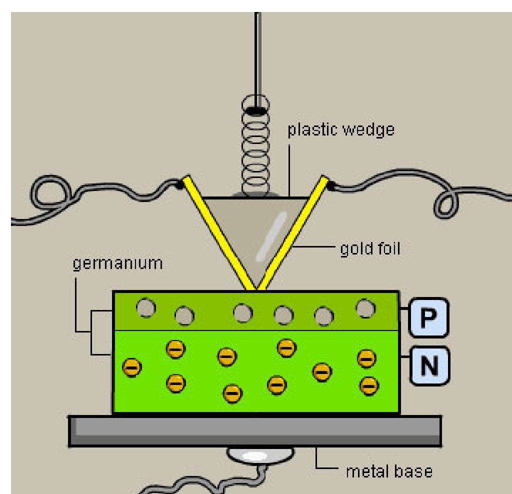

Fig. 1 - First "Point-Contact Transistor"[1]

[1] http://www.nyas.org/programs/nobel/essay5.html

Although very well used in switching applications, the Bipolar Power Transistor eventually got replaced by the Field Effect Transistor (FET) due to its physical limitations.

1. THE SELECTION OF THE RIGHT MOSFET BASED UPON THE PROCESS TECHNOLOGY

Two traditional technologies currently used are the planar and the trench process technology. The differences between both are pretty clear, starting with the geometrical structures.

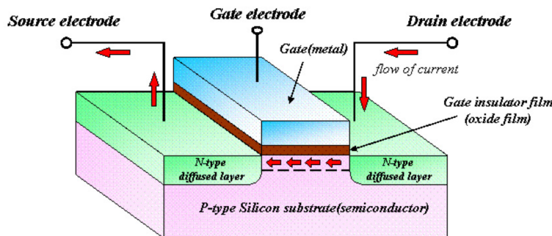

Fig. 2 - Planar MOSFET Structure[2]

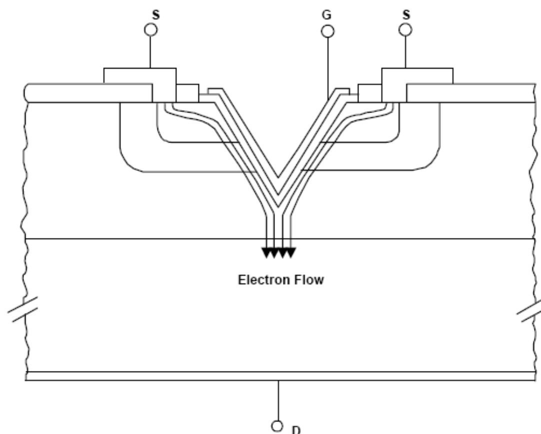

Fig. 3 - Trench MOSFET Structure[3]

As figure 2 and 3 clearly demonstrate the substantial difference between planar and trench technology is the way the Gate had been implemented: planar process technology puts planar layers above each other, trench technology generates a trench inside the Epi-Layer. However, one of the reasons why the trench technology got introduced was the higher cell density as compared to the traditional planar technology reducing the die size of the device and making it more suitable for portable, small form factor applications. Modern Trench MOSFET technologies allow cell densities of 200 Mio cells/inch and more.[4]

However, with the reduction of the physical dimension of the MOSFET the Gate-Oxide Capacitance will automatically increase, according to the following, very well known trivial relationship

$$C_{OX} = \frac{\varepsilon_{OX}}{t_{OX}}, \quad (1)$$

with

C_{OX}	-	Gate Oxide Capacitance
ε_{OX}	-	Dielectric Constant
t_{OX}	-	Thickness of the Gate-Oxide

It is important to keep in mind, that the overall capacitance is still a function of the voltage, following a behavior like in the curve shown below:

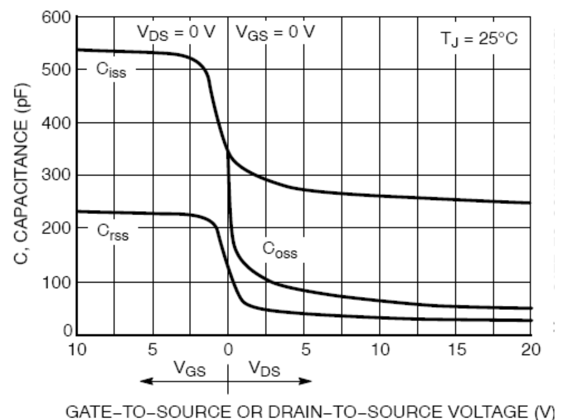

Fig. 4 - Capacitance Variations[5]

Another phenomenon that goes along with the different physical structures is the so called Drain-Source Resistance "$\boldsymbol{R_{DS(ON)}}$", which is a purely parasitic parameter.

In case of the planar structure, the parasitic components that will be responsible for this resistance are shown in the following simple schematic:

[2] http://www.masuoka.riec.tohoku.ac.jp/english/kenkyusyoukai/kenkyusyoukai-sgt.htm

[3] http://www.irf.com/technical-info/appnotes/mosfet.pdf

[4] Erroneously, sometimes the perception in terms of how to differentiate planar from trench technology implies the current path as a key criterion; however, as shown above, the criterion is not the current flow itself, but the way the Gate had been implemented, whether in a planar way, as an additional layer on top of the Epi, or "trenched" into the Epi.

[5] http://www.onsemi.com/PowerSolutions/product.do?id=NTLJF3118NTAG

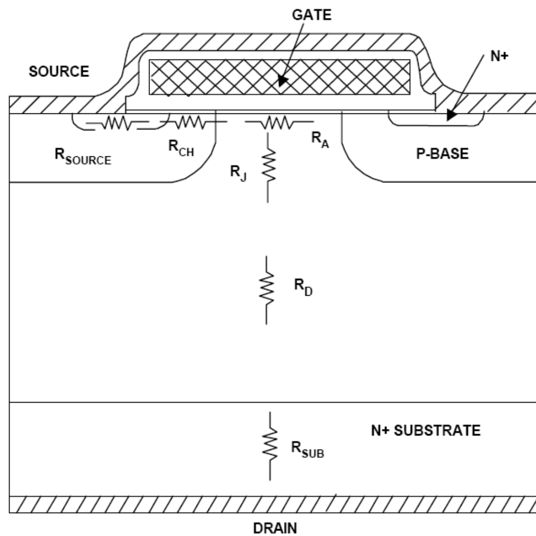

Fig. 5 - ***$R_{DS(ON)}$*** Contributors, Planar Structure[6]

Following above picture, the Drain-Source Resistance would be a sum of all contributors, leading to the following equation:

$$R_{DS(ON)} = R_{Source} + R_{CH} + R_A + R_J + R_D + R_{Sub} + R_{WCL},$$

(2)

with

R_{Source}	-	Source diffusion resistance
R_{CH}	-	Channel resistance
R_A	-	Accumulation resistance
R_J	-	"JFET" component-resistance of the region between the two body regions
R_D	-	Drift region resistance
R_{Sub}	-	Substrate resistance
R_{WCL}	-	Sum of Bond Wire resistance

Since the physical structure of the Trench FET varies significantly from that of a Planar FET, above relationship is no longer valid. The parasitic effects in this case can be shown through the following drawing:

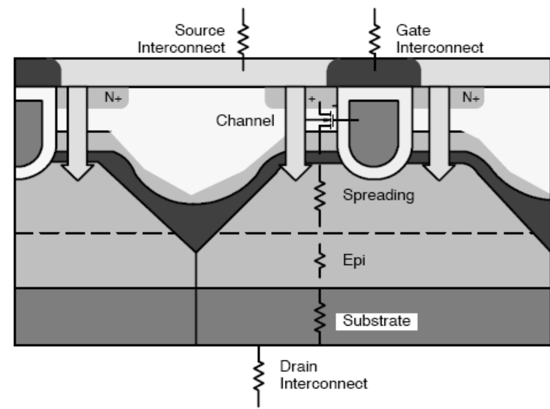

Fig. 6 - ***$R_{DS(ON)}$*** Contributors, Trench Structure[7]

Taking into consideration above picture, the ***$R_{DS(ON)}$*** will now be the sum of following parameters:

$$R_{DS(ON)} = R_{Source} + R_{CH} + R_A + R_D + R_{Sub} + R_{WCL}$$

(3)

As compared to equation (2) the JFET parameter is missing in case of the Trench FET structure, providing a much smaller ***$R_{DS(ON)}$*** than the one in case of the planar technology.

As a result of above considerations, planar technology provides MOSFET solutions with much smaller Gate-Oxide Capacitances, while Trench Technology provides MOSFET solutions with much smaller Drain-Source Resistances.

A comparison of two typical devices demonstrates the differences:

ON CHARACTERISTICS (Note 5)						
Gate Threshold Voltage	$V_{GS(TH)}$	$V_{GS} = V_{DS}$, $I_D = 250\ \mu A$	0.45		1.2	V
Negative Threshold Temperature Coefficient	$V_{GS(TH)}/T_J$			2.4		mV/°C
Drain−to−Source On Resistance	$R_{DS(on)}$	$V_{GS} = 4.5$ V, $I_D = 890$ mA		0.20	0.35	Ω
		$V_{GS} = 2.5$ V, $I_D = 780$ mA		0.26	0.45	
		$V_{GS} = 1.8$ V, $I_D = 700$ mA		0.43	0.65	
		$V_{GS} = 1.5$ V, $I_D = 200$ mA		0.56	1.2	
Forward Transconductance	g_{FS}	$V_{DS} = 10$ V, $I_D = 800$ mA		1.6		S
CHARGES, CAPACITANCES AND GATE RESISTANCE						
Input Capacitance	C_{ISS}	$V_{GS} = 0$ V, f = 1 MHz, $V_{DS} = 16$ V		79	120	pF
Output Capacitance	C_{OSS}			13	20	
Reverse Transfer Capacitance	C_{RSS}			9.0	15	

Fig. 7 - NTK3134N, **Planar** 20V, n-channel FET[8]

---

[6] http://www.irf.com/technical-info/appnotes/mosfet.pdf

[7] http://www.vishay.com/docs/71933/71933.pdf

[8] http://www.onsemi.com/PowerSolutions/product.do?id=NTK3134NT1G

ON CHARACTERISTICS (Note 5)						
Gate Threshold Voltage	$V_{GS(TH)}$	$V_{GS} = V_{DS}$, $I_D = 250\ \mu A$	0.4	0.7	1.0	V
Gate Threshold Temperature Coefficient	$V_{GS(TH)}/T_J$			–3.0		mV/°C
Drain–to–Source On–Resistance	$R_{DS(on)}$	$V_{GS} = 4.5$, $I_D = 3.8$ A		37	65	mΩ
		$V_{GS} = 2.5$, $I_D = 2.0$ A		46	75	
		$V_{GS} = 1.8$, $I_D = 1.7$ A		65	120	
Forward Transconductance	g_{FS}	$V_{DS} = 10$ V, $I_D = 1.7$ A		4.2		S
CHARGES, CAPACITANCES AND GATE RESISTANCE						
Input Capacitance	C_{ISS}	$V_{GS} = 0$ V, f = 1 MHz, $V_{DS} = 10$ V		271		pF
Output Capacitance	C_{OSS}			72		
Reverse Transfer Capacitance	C_{RSS}			43		
Total Gate Charge	$Q_{G(TOT)}$	$V_{GS} = 4.5$ V, $V_{DS} = 10$ V, $I_D = 3.8$ A		3.7		nC
Threshold Gate Charge	$Q_{G(TH)}$			0.3		
Gate–to–Source Charge	Q_{GS}			0.6		
Gate–to–Drain Charge	Q_{GD}			1.0		

Fig. 8 - NTLJF3118N, **Trench** 20V, n-channel FET[9]

Both devices, the NTK3134N and the NTLJF3118N are comparable, n-channel, 20V MOSFETs, with the last one being one of ON Semiconductor's latest Trench MOSFET developments and the first one being a representative of the traditional planar technology.

As the tables in figure 7 and 8 clearly demonstrate the planar technology provides much lower Gate Capacitance values (120 pF vs. 271 pF) and this although the test conditions are not 100% the same: the Drain-Source Voltage in case of the first one is V_{DS}=16V, in case of the second it is V_{DS}=10V. Still the differences are big.

In regards to the $\boldsymbol{R_{DS(ON)}}$ the behavior goes exactly vice versa: 350 mΩ @ V_{GS}= 4.5V as far as the NTK3134N is concerned and 65 mΩ @ V_{GS}= 4.5V with respect to the NTLJF3118N.

With above observation a traditional way to select "the right device" for a given application would be by taking into consideration the so called Figure-Of-Merit parameter, FOM[10], which compares gate charge Q_G against $\boldsymbol{R_{DS(ON)}}$. It follows the trivial relationship

$$FOM = R_{DS(ON)} \cdot Q_G . \qquad (4)$$

Therefore, the smaller the FOM, the better the MOSFET's performance as far as these two parameters are concerned.

A typical application where the $\boldsymbol{R_{DS(ON)}}$ value becomes crucial to the system's performance would be charging circuits for cell phone applications. A traditional cell-phone charger circuit could look as follows:

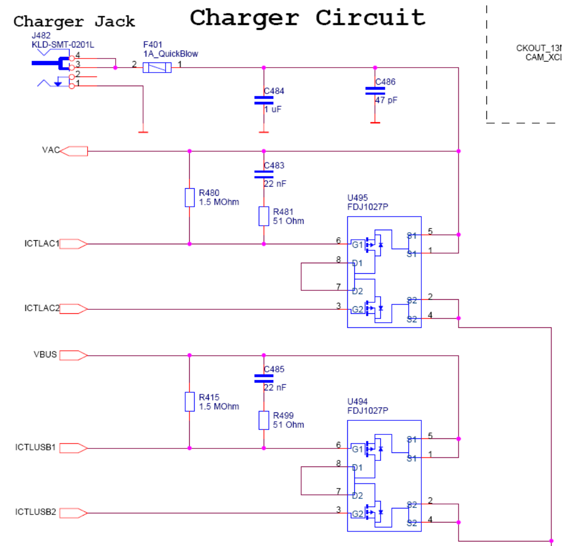

Fig. 9 - Typical Cell Phone Charging Circuit[11]

As far as charging procedures are concerned, the MOSFETs are driven in a linear way, so the switching characteristic is not critical at all. However, both FETs are in a so called Common-Drain configuration, summing up the Drain-Source Resistances of both FETs to twice the value of one.

Keeping in mind the overall efficiency, the $\boldsymbol{R_{DS(ON)}}$ needs to be as low as possible, especially when one of the power paths is being used in a bi-directional way, for example to drive a subsystem out of the cell-phone battery. In above example a value of $\boldsymbol{R_{DS(ON),\,max}}$ = 200 mΩ seems to be acceptable for the complete path.

Above requirements go along with the specific need for small form factor packages (in above case SC75, 1.7x2.15 mm²), which is the reason why the technology selected for this MOSFET had to be a Trench process.

Another example where the trade-off between trench vs. planar becomes very obvious is the integration of Power MOSFETs in planar Mixed Signal Chip Solutions, like the Cell Phone's PMU: the requirement for higher integration on one side is being limited by the MOSFET's physical limitations in terms of its $\boldsymbol{R_{DS(ON)}}$ converting the whole approach to a technical challenge.

---

9 http://www.onsemi.com/PowerSolutions/product.do?id=NTLJF3118NTAG

10 Other alternative methods would be the **BHFOM**, Baliga High-Frequency FOM, or the **NHFFOM**, New High-Frequency FOM

11 LOCOSTO Reference Design (TI)

Besides these two parameters, there is another phenomenon very often forgotten, but which is crucial in terms of the MOSFET's reliability, that is its tendency for a thermal-runaway.

Figure 10 demonstrates the root cause for a thermal runaway:

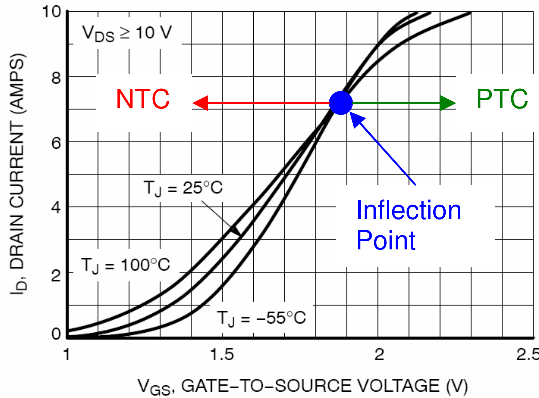

Fig. 10 - NTLJF3118N, Transfer Characteristics[12]

Above transconductance characteristic shows a very typical behavior of a MOSFET: for Gate-Source voltages $V_{GS} < V_{Inflection\ Point}$ the Drain Current I_D increases with temperature at a given set point, while for voltages $V_{GS} > V_{Inflection\ Point}$ this behavior goes the other way around.

The reason for it is the ***$R_{DS(ON)}$***'s temperature coefficient: for voltages $V_{GS} > V_{Infliction\ Point}$ this factor behaves in a positive way with temperature, hence it will increase reducing the Drain current at a certain set-point. Accordingly, this characteristic goes vice versa for voltages $V_{GS} < V_{Infliction\ Point}$.

In case that the application requires an active control of the current during the power-on-procedure for a significant period of time, above behavior might become critical, especially if the application requires significantly small ***$R_{DS(ON)}$*** values, since this need automatically leads to the selection of a MOSFET device with a higher MOSFET cell density/mm².

Due to the smaller physical distance to each other, those cells with a higher current capability (so called "hot spots") will start to increase their temperature heating up not only themselves but all other cells in their surrounding; a thermal-runaway can therefore occur with disastrous results for the MOSFET and affect the reliability of the application.

[12] http://www.onsemi.com/PowerSolutions/product.do?id=NTLJF3118NTAG

A typical application where such a constellation may occur could be an inrush current limitation circuit within a "hot-pluggable" application. Such a circuit could look as follows:

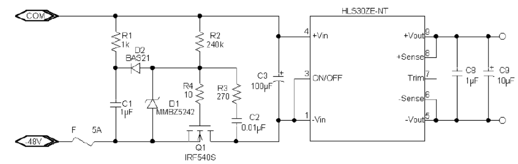

Fig. 11 - Inrush Current Limit Circuit[13]

In above circuit, C_3 represents the input capacitance of the DC-DC converter. If plugged into a system with C_3 totally discharged, this capacitance will initially act as a short circuit draining such a high amount of current that, if not actively controlled, will damage the system.

The MOSFET's (Q_1) job will be to limit the current surge by driving the Gate voltage bellow the critical current value, with a time delay reflected by

$$\tau = R_1 \cdot C_1 = 1k\Omega \cdot 1\mu F = 10^{-3}s = 1ms,$$

the low pass filter's time constant.

This delay may be sufficient under normal operating conditions; however, if for some reasons an over-load / over-current condition happens, the MOSFET can quickly start to operate in thermal-runaway mode, as shown in figure 12:

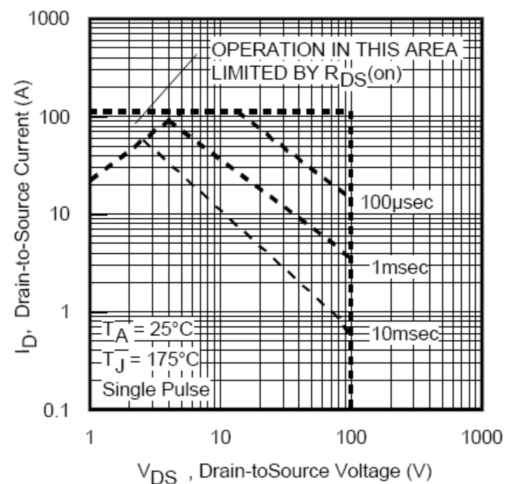

Fig. 12 - **IRF540N** – Safe Operating Area[14]

The IRF540N's point of infliction would be of $I_D = 13A$ @ $V_{GS} = 5.5V$ & $V_{DS} = 50V$ for 20µs; obviously, there is not

[13] http://www.power-one.com/technical/articles/dc-dc_1-app.pdf

[14] http://www.irf.com/product-info/datasheets/data/irf540n.pdf

too much room for safety in case of an unpredicted event that requires higher power dissipation for a significantly longer period of time than those 20µs to 1ms.

In conclusion, the tendency to thermally runaway, should be a parameter to be looked closer at, besides the traditional parameters ***$R_{DS(ON)}$*** and Gate Charge, reflected by the FOM.

2. THE SELECTION OF THE RIGHT MOSFET BASED UPON ITS THERMAL RESISTANCE

The key to a solution always depends upon the method to understand the root cause of a problem. Section 1 explained the root cause for a MOSFET's thermal-runaway based upon the observations taken from figure 10. Section 2 will introduce a basic, generic model to analyze the implicated parameters.

In general terms a thermal open system could look as follows, taking into consideration a thermal to electrical analogy:

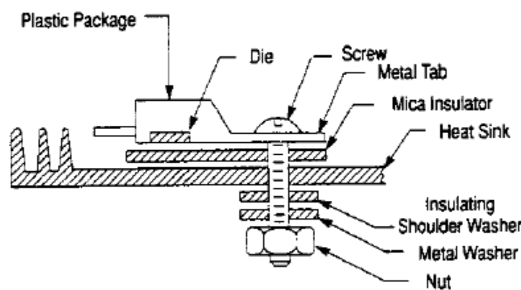

Fig. 13 - Mechanical Structure on Board Level[15]

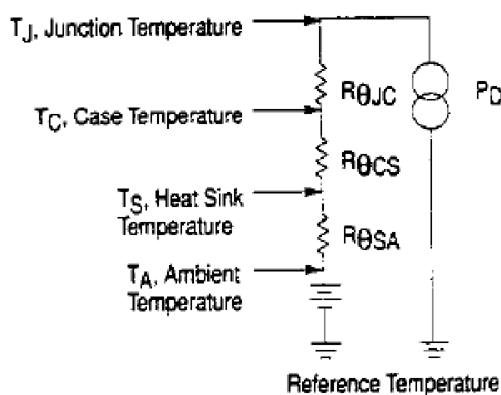

Fig. 14 - Equivalent Thermal Model[16]

[15] http://www.onsemi.com/pub/Collateral/AN1083-D.PDF

In above model, the physical relationship between the Junction Temperature and the Thermal Resistance is being expressed through the following mathematical term

$$T_J = P_D \cdot (R_{\theta JC} + R_{\theta CS} + R_{\theta SA}) + T_A, \qquad (5)$$

with

T_J	-	Junction Temperature
P_D	-	Power Dissipated at the Junction
$R_{\theta JC}$	-	Junction-Case Therm. Resistance
$R_{\theta CS}$	-	Case-Sink Therm. Resistance
$R_{\theta SA}$	-	Sink-Ambient Therm. Resistance

The coefficient $(R_{\theta JC} + R_{\theta CS} + R_{\theta SA})$ of equation (5) is usually being simplified by providing a thermal resistor value for the complete thermal path, hence $R_{\theta JA}$, from Junction to Ambient.

In order to keep the following analysis as simple as possible, we will introduce the following terminology:

T_J - Junction Temperature, [°C]
T_X - Junction Temperature at a reference point, [°C]
Q - System's Power Dissipation, [W]
P - Device's Power Dissipation, [W]
θ_{JX} - Steady State Thermal Resistance of a system with respect to a reference temperature "X", [°C/W]

With above simplification, equation (5) can be expressed in the following way:

$$T_J = Q \cdot \theta_{JX} + T_X, \qquad (6)$$

which leads to the equivalent relationships (7) and (8)

$$\frac{Q}{T_J - T_X} = \frac{1}{\theta_{JX}}, \qquad (7)$$

$$\frac{dQ}{dT} = \frac{1}{\theta_{JX}}. \qquad (8)$$

While equation (7) reflects the total power dissipated with respect to a reference temperature T_A, equation (8) represents the *changes* of *power dissipation* with respect to a corresponding *thermal change*. In other words, a system (reflected by the θ_{JX}) is capable to handle slightly more power in case of a small thermal increase.

For the following discussion, the system's thermal behavior (θ_{JX}), hence its capability to cool will be

[16] http://www.onsemi.com/pub/Collateral/AN1083-D.PDF

assumed to be constant which leads us to the following linear relationship

$$Q_{(T)} = \frac{1}{\theta_{JX}} T - \frac{1}{\theta_{JX}} T_X , \quad (7)$$

While $-\frac{1}{\theta_{JX}} T_X$ represents a constant offset, $\frac{1}{\theta_{JX}} T$ reflects the system's linear thermal behavior as a function of temperature T.

Case 1:

A device (blue line) generates a fixed amount of power P, regardless of its temperature, hence $P_{(T)} = P$. The relationship between the power generated by the device and system's capability to dissipate it can be illustrated through the following graph:

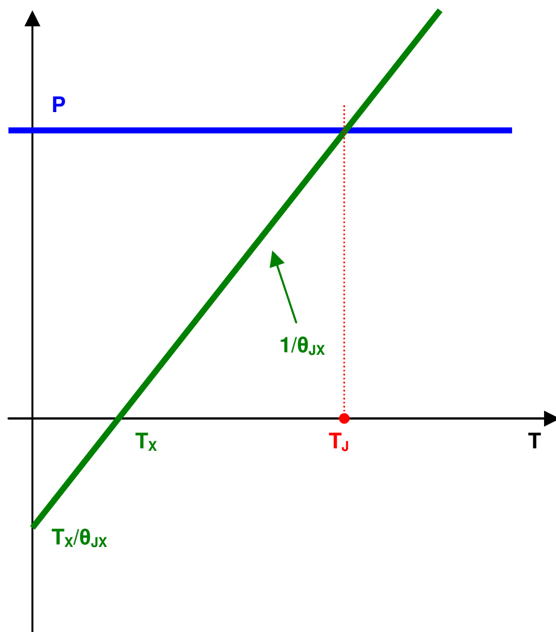

Fig. 15 - Operating Point with Temperature Independent Power

As long as the device's graph (blue) is above the system's graph (green) the complete system will heat up since the device is generating more power, than the system can dissipate until this open system has reached a point of thermal equilibrium (T_J). In case of slight temperature increases, the system will be in the position to dissipate it and fall back to the original equilibrium, since the device is generating less power than the system is capable to dissipate.

Thus in above case the system will experience a tendency to fall back to the original thermal equilibrium T_J in case of small temperature perturbations due to the system's capability to dissipate more power than the device can generate. This open system therefore acts in a stable way.

Case 2:

The device's generated power is not a constant, fixed value, but a function of temperature $P(T) = aT + P$. "a" is a constant quotient that represents the power change over temperature that might be positive or negative (Obviously, "a" had been a = 0 in "**Case 1**"!!).

In practice such a condition turns out to be more complicated than a simple linear approach. Following the discussion in the context of figure 10 the $\boldsymbol{R_{DS(ON)}}$'s temperature coefficient turns out to be a function of the Gate voltage V_{GS}, causing an impact in the Drain current I_D. However, in order to demonstrate the principles of this method, the device's Power behavior over temperature will be assumed to be linear (!).

The relationship between generated power and the system's capability to dissipate can be illustrated in the following way:

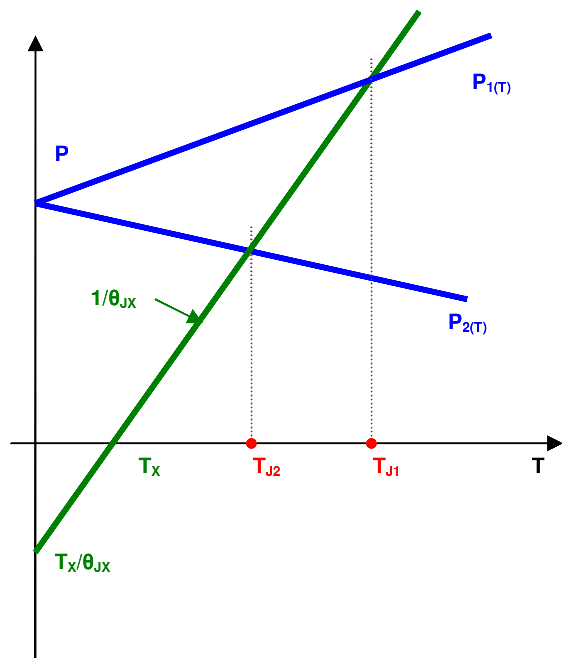

Fig. 16 - Operating Point with Temperature Dependent Power

In essence, the conclusions taken in Case 1 are still applicable: although the power generated by the device is no longer constant, but a function of temperature, small temperature increases above the thermal

equilibrium T_{J1} and T_{J2} will easily be handled due to the system's capability to dissipate more power than the device can generate; the system will therefore fall back to the original thermal equilibrium T_{J1} and T_{J2} correspondingly.

Again, these two open systems are working in a stable balance too.

Case 3:

The next option demonstrates the conditions for a thermal runaway:

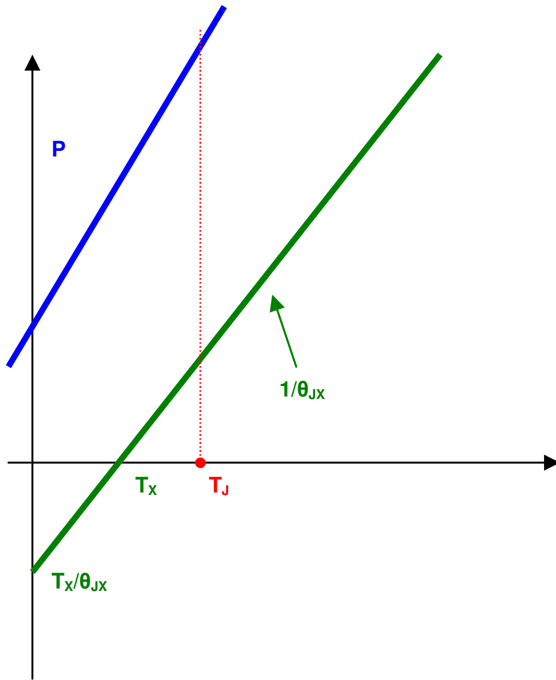

Fig. 17 - Thermal Runaway

In analogy to the discussion up to now, T_J represents an operating point without thermal equilibrium; the device generates more power than the system's capability to dissipate it, causing a thermal-runaway condition, if not actively reacted.

In practice such a condition can happen, when a MOSFET is being driven with a voltage $V_{GS} > V_{Infliction\ Point}$ and a predictable behavior like the one describe in figure 15, but due to unpredictable conditions (like an overload condition) the Gate voltage needs to be driven to a condition $V_{GS} < V_{Infliction\ Point}$ in order to limit the current, changing the MOSFET's thermal behavior similar to the one describe in figure 16.

Up to now we had been discussing thermal scenarios where the system's behavior, had been left untouched. The following analysis will leave the device's characteristic untouched and discuss different possible configurations with changing cooling systems.

$P_{(T)} = aT + P$ represents the linear power generation of a device (for example a MOSFET) as a function of temperature; this behavior is reflected by the datasheet (figure 10) and is supposed to be known at the design stage[17].

$Q_{1(T)} = \frac{1}{\theta_{JX1}}T - \frac{1}{\theta_{JX1}}T_{X1}$ reflects the cooling characteristic of a system "1", with a reference temperature T_{X1}, and a thermal resistance θ_{JX1}, with all parameters being assumptions taken at an early design stage.

$Q_{2(T)} = \frac{1}{\theta_{JX1}}T - \frac{1}{\theta_{JX1}}T_{X2}$ shows a linear cooling system with the same thermal resistor θ_{JX1} but a different reference temperature T_{X2}.

In analogy to $Q_{1(T)}$, $Q_{3(T)} = \frac{1}{\theta_{JX2}}T - \frac{1}{\theta_{JX2}}T_{X1}$ illustrates a variation in terms of the thermal resistor θ_{JX2}; the reference temperature T_{X1} would still be the same.

All 4 graphs are illustrated in figure 18:

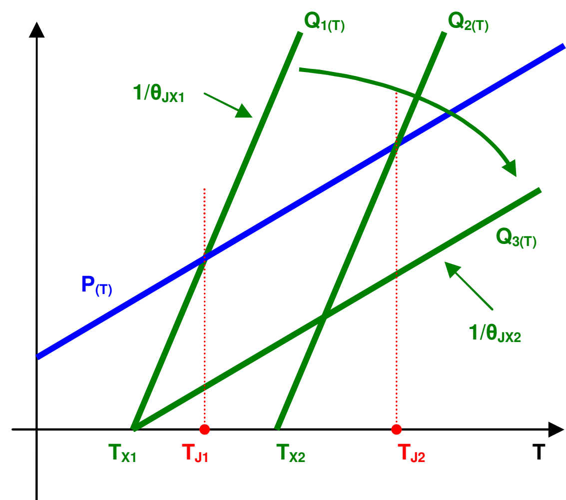

Fig. 18 - Thermal Behavior with Generic Linear Cooling Systems

[17] For more information on the discussion of the thermal behavior of device with non linear power generation, pls. refer to http://www.onsemi.com/pub/Collateral/AND8223-D.PDF

In case of $Q_{1(T)}$ the expected thermal equilibrium $T_{J1} = \frac{\frac{1}{\theta_{JX1}} T_{X1} + P}{\frac{1}{\theta_{JX1}} - a}$ [18] would be a stable operating point, equivalent to those operating points discussed in "**Case 2**". A minor thermal heat up of the system will cause the device to generate slightly more power, but due to the system's capability to dissipate it, it will eventually fall back to T_{J1}. This might be a typical scenario at the design stage, where based upon the technical specification of the chosen devices, a first pass simulation of the system's behavior will be performed and be acceptable.

$Q_{2(T)}$, however, represents a significant drift in terms of the operating thermal equilibrium T_{J2} $T_{J2} = \frac{\frac{1}{\theta_{JX1}} T_{X2} + P}{\frac{1}{\theta_{JX1}} - a}$; the reasons for it are the wrong assumption at the design stage in terms of the reference temperature. Rather then dealing with T_{X1} as originally assumed, the system ends up dealing with T_{X2}, the new reference temperature, shifting the operating thermal point to a higher level.

This situation usually happens at the end of a design cycle, where the final architecture had been defined, but due to unknown parameters, the whole system ends up working at a higher temperature then originally assumed. Depending upon how much tolerance had been left at the early design stage, this configuration might be acceptable or not; if not the traditional way would be to bring down the reference temperature, for example by placing a fan that was initially not planned to be used.

Finally, $Q_{3(T)}$ represents a worse case scenario with a cooling system that is not capable to provide a thermal equilibrium, hence the system will thermally run away. Such a situation may happen when the thermal behavior of the cooling system (θ_{JX1}) turns out to be completely different as to what the datasheet specifies, providing a characteristic totally insufficient (θ_{JX2}) for the designed system.

Since the design had already been finished at this stage the only practical way to cope with this problem would probably be by using a different cooling material (θ_{JX3}) which will be capable to generate a stable thermal equilibrium T_{J3}. [19]

While the scenario described by $Q_{2(T)}$ rather represents the day-to-day reality engineers have to deal with, $Q_{3(T)}$ represents a situation that sometimes may happen when the assumptions taken at an early design stage were to vague and not based upon facts. While engineers usually know how to successfully deal with $Q_{2(T)}$, $Q_{3(T)}$ may probably require a re-design, at least of the cooling system. However, these kinds of issues may be prevented by following the methodic approach introduced above.

3. THERMALLY SELF-PROTECTED MOSFET

One practical way to get around the issues mentioned under "1. **THE SELECTION OF THE RIGHT MOSFET BASED UPON THE PROCESS TECHNOLOGY**" in the context of an application like that one illustrated in figure 11 would be by using "Self Protected MOSFETs". The block diagram of such a device looks as follows:

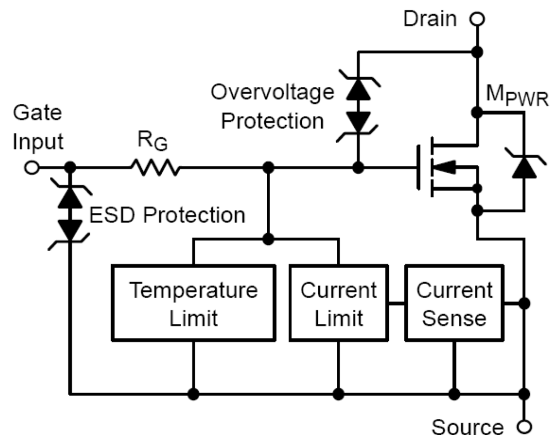

Fig. 19 - NIF5002N, Self-Protected FET, Block Diagram[20]

Besides the MOSFET, hence the "switching function", the NIF5002 (representative for a variety of devices[21])

---

[18] Note, that if a=0, hence the device generates power independent from the temperature, like in "**Case 1**", the thermal equilibrium would be $T_{J1} = T_{X1} + \theta_{JX1} P$, a relationship mathematically equivalent to equation (6).

[19] Mathematically, there is also the option to bring the reference temperature T_{X1} so much down (!) that the cooling system's curve ends up above the device's graph; however, such a configuration won't provide a thermal equilibrium either (assuming like in this case that both curves are parallel to each other). Furthermore, the commercial aspect of such an approach won't justify its implementation for obvious reasons either.

[20] http://www.onsemi.com/pub/Collateral/NIF5002N-D.PDF

[21] More information available under http://www.onsemi.com/PowerSolutions/parametrics.do?id=819

contains additional features, relevant to the system's reliability: the device integrates a current sensor combined with a current limiting function and a thermal shut-down.

Current Sensing/Current Limitation:

A traditional way to sense the current that flows through a system would be by using a sense resistor in series to the power path, like demonstrated in the following picture:

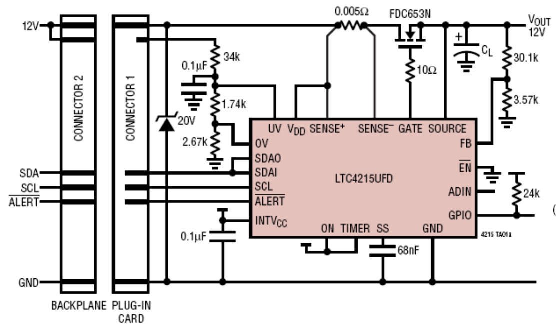

Fig. 20 - Typical HotSwap Controller Circuit[22]

Besides the cost aspect that usually goes along with a high sensing accuracy, this sense resistor , although very small (5mΩ), still represents a voltage drop at the input and therefore power dissipation in the power path.

The alternative approach would be through the implementation of a current mirror circuit like in case of all self-protected MOSFETs. A simplified drawing of the circuit used looks as follows:

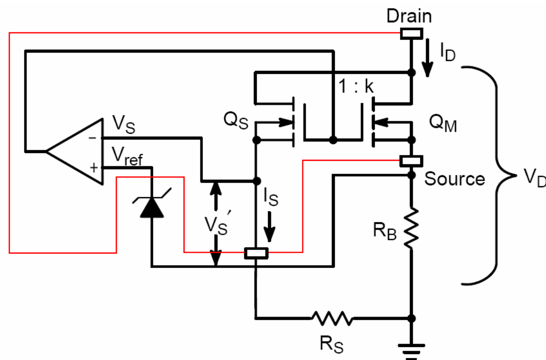

Fig. 21 - Simplified Current Sense Circuit

Given the fact that a MOSFET is not a single device but the integration of thousands of single MOSFET cells in parallel, by taking a controllable amount of cells, a trivial current mirror can be generated with a ratio of 1:k.[23]

In above circuit the reference voltage had been implemented through R_B and a Zener diode; of course there are different ways to approach that. On the other hand R_S represents an external resistor that acts as a voltage divider with the internal ***$R_{DS(ON)}$*** of Q_S. This way, the sense current threshold can be individually calibrated and controlled through the MOSFET's Gate. The relationship between V_S and R_S is therefore

$$V_S = \frac{R_S}{R_S + k \cdot R_{DS(ON)}} V_D', \quad (8)$$

with

$R_{DS(ON)}$ - MOSFET's total Drain-Source Resistance
k - Cell Ratio

Since the amount of current that flows through Q_S is just 1:k as compared to the current in the power path (Q_M), the power losses would be insignificantly small as compared to using a sense resistor in series.

The implementation of above circuit with the NIF5002N in a real system with its output short cut, (hence $V_D'=0$ => $V_S'=0$) shows a behavior like in the following graph:

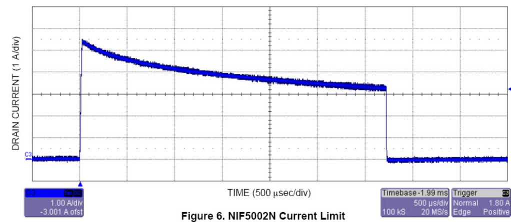

Fig. 22 - Active Inrush Current Limitation[24]

Initially, the current is limited at a value of ~5.5A; as the device heats the current limit value progressively decreases to approximately 3.3 A until it reaches the over-temperature set-point and shuts off.[25]

---

[22] http://www.linear.com/pc/downloadDocument.do?navId=H0,C1,C1003,C1006,P17572,D12697

[23] Above block diagram infers that the current mirror ratio, k, is constant, but this not the case. In fact, the current mirror ratio varies over temperature and input current as a function of V_{SENSE}, since as V_{SENSE} increases the two mirror FETs V_{GS} voltage differs by the V_{SENSE} amount. It is strongly recommended to read the application note **AND8093** (http://www.onsemi.com/pub/Collateral/AND8093-D.PDF) for further information on how to sense the current with power SENSFETs.

[24] http://www.onsemi.com/pub/Collateral/AND8202-D.PDF

[25] The decrease in current limit threshold is due to a decrease of the threshold voltage of the NMOS pull down transistor in the control circuit.

However, this condition leads automatically to a progressive heat up of the MOSFET, which, if not properly controlled will eventually cause its damage as explicitly discussed under "1. **THE SELECTION OF THE RIGHT MOSFET BASED UPON THE PROCESS TECHNOLOGY**" and demonstrated through figure 10. Consequently, the solution to that problem would be the implementation of a thermal closed loop control.

Temperature Sensing:

A traditional way to cope with this problem would obviously be to place a thermistor as close as possible to the MOSFET to sense its temperature and provide the value to an internal ADC integrated in the HotSwap controller, like indicated in the following drawing:

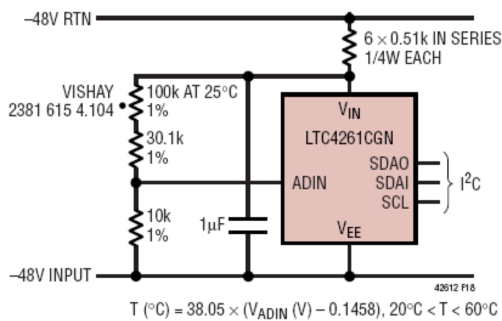

Fig. 23 - LTC4261 using Thermistor as Temperature Monitor[26]

It is clear and obvious that with such an approach the system's reliability will be depending upon the thermal migration speed from the MOSFET to the Sensor and how much tolerance (e.g. safety) had been taken into consideration on design level to make sure the protection function kicks in before the MOSFET starts to thermally run away.

Using self-protected MOSFETs this issue gets solved in an elegant way, since the complete control circuit is integrated inside the MOSFET design, as demonstrated through the following two pictures:

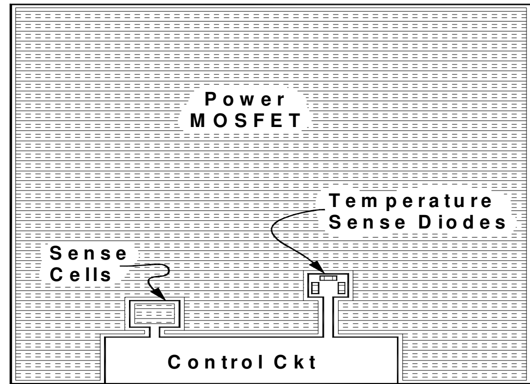

Fig. 24 - Generic Die Structure Representation

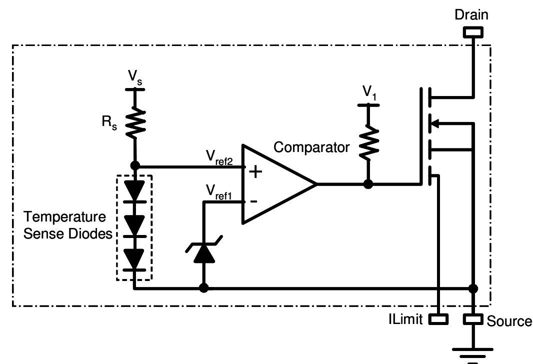

Fig. 25 - SENSE FET and Simplified Thermal Shut Down Circuit

Since the forward voltage of a diode string drifts with temperature, this phenomenon is used to sense the temperature of a MOSFET inside of the die structure, proving a "real time" value without any thermal migration delay time.

Obviously, such a simple implementation prevents the MOSFET from thermally running away, given the fact that once it starts to heat up, and hits the thermal threshold (reflected through V_{REF1}) the comparator will interrupt the power path and this way allow the MOSFET to cool off.

---

26 http://www.linear.com/pc/downloadDocument.do?navId=H0,C1,C1003,C1006,P13513,D9458

This behavior is being illustrated through a time extension of figure 22:

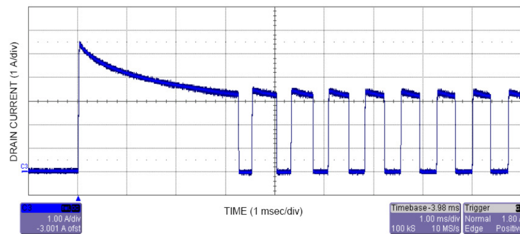

Fig. 26 - Active Inrush Current Limitation **plus** Thermal Shut Down[27]

As the NIF5002N starts to heat the thermal shut down control kicks in, allowing the device to cool by approximately 15℃. As long as the reason for the thermal heat up has not been removed yet (hence the device's output is still short cut), the NIF5002N will continue to oscillate as shown in above picture.

For some applications the requirement would be for the device to be capable to successfully withstand these conditions during a period of time of up to 6 months[28]. Tests performed by a customer in their lab under real system conditions have proven the reliability of this solution without any further damages.

In conclusion, it can be noted that the most effective way to prevent a system from unpredicted damages due to an uncontrolled thermal-runaway would be through the use of self protected MOSFETs.

Summary

In conclusion, Section **"1 The selection of the right MOSFET based upon the process technology"** described three key parameters when it comes to the selection of a MOSFET for given application: the standard parameters $R_{DS(ON)}$ and Gate Charge, both reflected through the FAM as well as the MOSFET's tendency to thermally runaway under certain conditions. The root cause for this last one has been discussed in detail and its implications with respect to the system's reliability in the context of thermal management.

Consequently, Section **"2 The selection of the right MOSFET based upon its Thermal Resistance"** has provided the methodical tool to understand and predict the thermal behavior of a MOSFET thermal runaway like the one discussed under **"1 The selection of the right MOSFET based upon the process technology"** and what would be the theoretical options to cope with the problem.

And finally, Section **"3 Thermally Self-Protected MOSFET"** provided a system solution in order to actively prevent a system from an unwanted thermal runaway where an equilibrium can theoretically not be achieved by using self-protected MOSFETs.

In essence, Section 1 analyzed a couple of MOSFET phenomena, important when selecting a specific device for a given application, Section 2 presented a usefull too how to systematically predict the thermal behavior of a system while the objective of Section 3 was to present a system solution on how to practically cope with the problems discussed in the two Sections before.

---

[27] http://www.onsemi.com/pub/Collateral/AND8202-D.PDF

[28] These 6 months represent the worse case condition between two services where the failure happens right after the last system maintenance without being noticed.