While high-side MOSFET switches have several obvious advantages over mechanical relays, the ever-shrinking geometries of semiconductor devices makes the thermal design of these solid-state power components increasingly difficult. One solution is to overcome this trend with the capabilities of the process itself by providing the device with onboard intelligent driving.

With so many carmakers designing in solid-state body control modules, the demand for protected drivers is increasing quite rapidly. Specifically, the protected high side driver is used to drive everything from headlamps to rear defrosters (heated backlights). The last electronic body module I had the opportunity to examine had more than 70 loads with the vast majority of them driven by solid-state high-side switches.

When we consider the alternative, driving these loads with relays, we begin to understand the shift to solid state. A relay requires between 100 mA and 300 mA to drive, provides no understanding of the condition of the load (open or shorted), and cannot support pulse-width modulation (PWM) for efficient dimming.

In the previously mentioned 70+ load module, more than 22 of the loads were 1 A or more. If we decide to drive those 22 loads with relays (actually nine of those loads required PWM, and can’t be driven by a relay) the current needed to drive just the relay primary windings can add up to more than 6 A and the power dissipated by the primary windings alone is more than 92 W. While this may seem like an excessive power requirement just to close some switches, some things are just not feasible outside of a solid-state solution.

There is definitely an advantage to driving loads electronically. Solid-state lamp driving allows the carmaker to provide theater dimming as well as intelligent protection. Intelligent protection provides open load and shorted load detection and protection. Open load detection informs the driver of less obvious lamp outage conditions such as tail lamp, turn signal or brake lamp outages. Shorted load detection/protection can prevent an intermittent short from blowing fuses. This can keep things running for a bit longer before the vehicle needs service. These are among the clear advantages offered by solid-state switching.

One of the technologies that makes protected semiconductor switches affordable is a simple MOSFET-based process that has built into it some basic control circuitry. This simple control circuit is placed in an isolated P-Well right along side the MOSFET. This makes for a fairly inexpensive process that is die area efficient with respect to on-resistance (RDS(on)) while providing for a rather limited level...
of protection and control. In the typical solid-state switch, most of the die is devoted to the pass element. Using some die area for some simple control functions is a minimal sacrifice compared to the added value it provides.

PROTECTION

The protected high-side driver control circuit consists of a charge pump, a voltage clamp, a basic current-limiting circuit and a thermal shutdown circuit. This is a fairly straightforward circuit using as few masks as possible. In the semiconductor process each mask represents a process step. A simpler, less-expensive, process uses less masks. Each mask adds to the cost of the process, so the fewer we use the less the device will cost.

For the basic high-side driver process, masks are limited in a number of ways. One way is to use only one fairly thick metal layer for interconnecting in both the high-current power MOSFET section as well as in the low-current control section. The high-current MOSFET will require a rather “thick” metal layer. The control section does not need that thickness but uses it to save on additional masks. Another metal layer would add up to three masks and add to the cost of the device. Using the thick metal layer for the control logic section is akin to using battery cables to wire a Walkman radio. This is not considered die area efficient but it is cost effective if the control circuit is not complex.

Basically, in a shorted-load condition, the protection circuit’s goal is to limit the current, keeping the pass element within the safe operating area until the thermal shutdown kicks in and shuts the driver off. Once it has cooled below the over-temperature threshold, it turns back on. Thus, it goes back and forth from current-limit-thermal-shutdown to off-and-cooling until the input is disabled.

This kind of protection generates high physical stress on the surface of the power device. The current limit has to be fairly high to satisfy the current limit tolerance limitations of the process (remember it is a simple limited function process) while allowing for reasonable bulb inrush current requirements. For instance a 60 mΩ device is really only thermally able to handle about 4 A in automotive environments but the current limit can be as high as 15 A due to inrush requirements and process tolerances.

With a high current limit level there is a fairly fast thermal rise on the surface of the die when driving...
into a shorted load. For example, with a 15 A current limit on a 14 V battery, the power is 210 W. For a device intended to dissipate around 1 W this would lead to rapid heating. The rate of temperature change on the die is fairly rapid. At a typical 25 °C ambient there can be more than a 165 °C rise in just a few milliseconds.

This rapid change in die temperature tends to cause large thermal gradients across the surface of the die. Those thermal gradients create several problems.

**THERMAL FATIGUE**

According to the Coffin-Manson thermal fatigue model, fast thermal transients over ~60 °C impose a thermo-mechanical stress that is “remembered.” That is, the device does not return to where it was prior to the stress, and the damage is accumulated over time. With a 165 °C excursion, as seen in a shorted load condition, this would obviously apply. The silicon is only able to experience a certain amount of these transients before some serious degradation occurs.

**Coffin-Manson thermal fatigue model**

\[ N_f = A_f^{-\beta} \Delta T^{-\alpha} G(T_{MAX}) \]

*Where:*

- \( N_f \) = The number of cycles to failure.
- \( f \) = The cycling frequency.
- \( \Delta T \) = The temperature range during the cycle.
- \( G(T_{MAX}) \) = An Arrhenius term evaluated at the max temperature reached during the cycle.
- \( \alpha = 2 \) (typically).
- \( \beta = 1/3 \) (typically).

The number of thermo-mechanical stresses the device can experience before failure is determined by the severity of the stress. The Coffin-Manson thermal fatigue model closely predicts this rate of failure in a shorted high side driver. Our modeling in Figure 3 matches closely with our measured results.

Thermo-mechanical stress tends to cause the source metallization to look like mud cracks (see Figures 4 and 5). The end result is a degradation of the interface between the source (or output) bond wires and the die surface. This increases the \( R_{ds(on)} \). Sometimes, the damage is to the point where the device will no longer pass current (an open circuit).

Unfortunately, some circuit designers think that they can extend the life of the device during a shorted load condition by letting the part cool for a few seconds then re-actuate the load. Ironically, in so doing they increase the extent of the thermal excursion and as a result the level of thermal fatigue, thus causing the device to fail sooner.
would be to sense the shorted load current and shut down the device before it experiences a rapid thermal rise or high current high temperature situation. This will extend the lifetime of the device significantly.

Many high-side drivers, like the one depicted in Figure 6, have analog current sense feedback pins. This pin mirrors some small amount of the actual output current. A sense resistor can convert that current into a voltage for the micro ADC to sense. If the microcontroller can react in time the thermal excursion can be limited in that it takes a few milliseconds to reach the thermal shutdown temperature and shut down.

Ultimately, the module engineer will have to do something in the microcontroller to prevent shorted loads from reducing the life of the high-side drivers. Many have seen OEM specifications that require a strategy that takes into account the

ELECTROMIGRATION
The next conclusion would be to turn the device on and leave it on even during a shorted load condition. This would limit the thermal excursion over time to the hysteresis of the thermal shut down circuit. Unfortunately, for semiconductor devices, high current and high temperature is a recipe for disaster. If a device is left on in the shorted load condition for an extended period of time, electromigration begins to cause its own type of wear. Electromigration tends to make the drain and the source of the pass element short out as the source metal migrates or diffuses into the junction. As more and more metal is diffused into the junction the off-state leakage current rises. The off-state leakage current can rise so much that during thermal shutdown, when the device is supposed to be cooling, the device ends up getting hotter still. The issue accelerates until something gives way. That could be the silicon, a bond wire or a circuit board.

INTELLIGENT DRIVING
The solution to these problems is intelligent driving. This translates to limiting the exposure of the high-side driver to shorted loads once it is determined that there is a problem. That exposure can be limited in a number of ways. The best method is somewhere around allowing the occupant/user to actuate the shorted load a few times per ignition cycle. Then, once a predetermined number of actuations has been exceeded, the output is disabled until the next battery cycle. Then a mechanic, after fixing the shorted load, can disconnect and then reconnect the battery, resetting the counter. This type of protection strategy can prevent catastrophic damage to the device while allowing for repeated actuations into a shorted load.

A more proactive approach would be to sense the shorted load current and shut down the device before it experiences a rapid thermal rise or high current high temperature situation. This will extend the lifetime of the device significantly.

Many high-side drivers, like the one depicted in Figure 6, have analog current sense feedback pins. This pin mirrors some small amount of the actual output current. A sense resistor can convert that current into a voltage for the micro ADC to sense. If the microcontroller can react in time the thermal excursion can be limited in that it takes a few milliseconds to reach the thermal shutdown temperature and shut down.

Ultimately, the module engineer will have to do something in the microcontroller to prevent shorted loads from reducing the life of the high-side drivers. Many have seen OEM specifications that require a strategy that takes into account the
limited lifetime of these types of drivers in shorted load conditions.

**NEW SOLUTIONS**

The next generation of intelligent drivers will have to address this problem head on. The natural progression of semiconductor devices is to get smaller, in die size and in package size. However, as die size decreases, the thermal capacitance goes down and thermal resistance goes up. In power electronics, thermal capacitance is beneficial because it slows down the rapid thermal rise that accompanies driving a shorted load. A low thermal resistance limits the die temperature to functional levels and allows the engineer to use higher Rds(on) switches. Higher Rds(on) devices are typically lower cost. So, with the increase in thermal resistance and the decrease in thermal capacitance, the next-generation high-side driver must be smarter as well as smaller.

In this new technology the end result is that the thermal rise during a shorted load actuation, if left unchecked, is orders of magnitude faster than their predecessors. If nothing is done to control the thermal rise during a short, the device will be protected for a shorted load only once, much like a fuse is useful once.

So STMicroelectronics, in its endeavor to make a more robust high-side driver, built in some fairly sophisticated thermal protection controls in its next-generation devices. Realizing that the next-generation silicon and packaging will just make things tougher with respect to thermal issues, they set out to solve the two big issues: fast thermal fatigue, and metallization diffusion into the junction.

Thermal fatigue is both a mechanical issue and a control issue. We want to even out the heat as well as slow it down to a more elastic rate. Mechanically, we can even things out a bit by adding more bond wires and using passive pads. Also, if we can slow down the rapid temperature increase by controlling the power during a shorted load we can ease the thermal stresses.

Current technologies involving MOSFET switches place big aluminum bond wires right on the active area. You can see in Figure 4 the aluminum bond wires placed right on to the MOSFET source metallization. Placing bond wires on active areas like this may have saved some die area but in the process did three things against them. First, it places this critical connection right where the highest thermal stress would occur. Second, because there are only a few aluminum bond wires on the active area there are areas on the silicon where the current will bunch up or crowd. This appears as hot spots around the bonding areas. (In Figure 7 you can see the 10 gold bond wires placed directly on the MOSFET.)
wires instead of the two or three you would see for this same $R_{ds(on)}$ device.) Third, the thermal sensor placement is forced to be far away from the hottest part of the die. This is to prevent the bonding process from damaging the sensor. The thermal gradient between the hot-spots around the bond wires and the thermal sensor during a shorted load condition can be quite high (see Figure 2). Figure 7 illustrates the placement of the thermal sensor quite well. Immediately under the fourth bond wire from the left you can see a lighter colored line. The position where the line ends near the bond pad is where the thermal sensor is located.

Bonding on passive bond pads takes up a bit of precious silicon area but in the long run helps save the device from early failure. The thermal sensor can now be close to the hottest part of the die and the thermal gradients are much less severe (see Figure 8) as the current distribution is much more even. Unfortunately, device failure still results if a shorted-load is driven indefinitely. This requires a reduction in the physical stress caused by fast thermal transients.

**SLOWING THE THERMAL RISE**

For a smaller smarter device to survive in the world of shorted loads there needs to be a slewing or a slowing of the thermal rise on the die. STMicroelectronics has found a way to slow the rate of thermal rise during a shorted load (or during a lamp inrush phase) by limiting the average power dissipated at any given moment. So when the device begins dissipating too much heat it turns off, allowing the device to cool and catch up for a moment before it turns back on. This PWMming of the shorted condition slows down the thermal rise to something more elastic in nature. So slow in fact that it satisfies the Coffin-Manson models for elastic behavior and proves out in testing.

**LIMITING TEMPERATURE AND DURATION**

But this is only half of the issue. The high-current high-temperature condition can still exist as long as the controller demands that the device drive the shorted condition indefinitely.

This can be dealt with by lowering the current limit to a less than stratospheric level once the thermal shutdown temperature is achieved. This two-stage current limit is initially high to allow for standard bulb inrushes and current-sensing tolerances. The second stage, approximately one-third of the initial current limit, is used only after a thermal shutdown has occurred. This value is still above what the device can thermally dissipate when driving a normal load. So essentially, this two-stage current limit allows the device to handle the inrush current requirements for a typical lamp load while reducing the continuous high-temperature high-current condition during a shorted load to a manageable level.

**STABLE DIAGNOSTICS**

For the current generation of protected high-side drivers there is a limitation as to how the diagnostic function is done. The technology is fairly limited in what it can do...
outside of being a switch. The limitation has to do with keeping the diagnostic pin stable when the device is in the thermal-shutdown thermal-recovery mode of operation (see Figure 1). With that said, the microcontroller has a lot of thinking to do to make sure that the high-side driver is in a faulted mode and shut it down before it’s damaged by thermal stress.

This next generation of smart switches has in them a dual reset threshold for the fault (or diagnostic) flag (Figure 9). One temperature threshold is used to shut the part down and set the diagnostic flag. A lower temperature threshold is used to restart the device while still leaving the diagnostic flag set. Then, a still lower threshold is used to reset the diagnostic flag. This multiple threshold method keeps the diagnostic flag active as long as the device is driving a load that causes thermal shutdown. As soon as the shorted load is removed the device will thermally reset, turn on, drop in temperature (as the load is not so severe) and then reset the diagnostic flag.

With 70+ plus loads to drive electronically you can imagine how a stable diagnostic flag can be fairly useful. Add to that the capability to tri-state this pin and the number of I/O signals needed to drive all of these loads gets lower.

There are limitations to what smart high-side switches can do. With improvements in technology, these limitations are being reduced if not eliminated. Truly, the switch never looked so smart as it does today. Improvements to technology not only improve their reliability but they improve functionality without adversely affecting the cost. The two mantras of automotive electronics are low cost and high reliability. We seem to be making that a possibility today. This is reflected in the rapidly increasing trend to use high-side switches in body electronics. As the sophistication of these intelligent devices grows, it may even be possible to eliminate automotive fuse protection altogether.

Moving the bond wires off of the active area of the MOSFET and onto passive bond pads may add real-estate to the silicon but in the long run saves the part from some of the stresses that occur during high current events.

The red line shows the expected failure rate with respect to thermal excursion for devices where the bond wires are on the active, thermally stressed area. The blue indicates the improvement to the expected failure rate by placing the bond wires on non-thermally stressed areas (passive bond pads) and moving the thermal sensor closer. The difference in failure rate is based on moving the bond wires to a non-thermo-mechanical stress point.

ABOUT THE AUTHOR:
David Swanson is a principal engineer in STMicroelectronic’s Automotive Business Unit. He has worked with the company since 1987 in various roles. Previously, Swanson worked for Delco Products Division of GM. He graduated with a BSEE from North Carolina State University and holds patents in several areas of automotive electronics.