Cooling Methods for High-Power Electronic Systems

Andrei Blinov¹, Dmitri Vinnikov²,
Tõnu Lehtla³, ¹Tallinn University of Technology

Abstract – Thermal management is a crucial step in the design of power electronic applications, especially railroad traction and automotive systems. Mass/size parameters, robustness and reliability of the power electronic system greatly depend on the cooling system type and performance. This paper presents an approximate parameter estimation of the thermal management system required as well as different commercially available cooling solutions. Advantages and drawbacks of different designs ranging from simple passive heatsinks to complex evaporative systems are discussed.

Keywords – Electronic packaging thermal management, electronics cooling, power dissipation, power semiconductor devices.

I. INTRODUCTION

In modern power electronics reduced system volume and costs are required to stay competitive. These requirements impose the use of high switching frequencies in order to reduce passive component sizes. As a result, semiconductors are dissipating more heat with increased density. The maximum junction temperature $T_{j(max)}$ (generally 125...175°C for IGBT modules) provided by the manufacturer has to be followed at any time of operation when using power semiconductors [1]. The heat potential due to energy losses has to be dissipated by a cooling system, which may also serve as a constructional element (case, chassis etc.) [1].

There are many ways to remove heat from a device, however, nearly all of them are based on the same common principle: to move heat away from the device to the ambient medium (in most cases air) by convection, conduction and radiation.

With the drastically increased power dissipation of HV switches only high-performance cooling systems should be used to improve converter dimensions. The ability to remove heat greatly influences the overall system design and to achieve a cost-optimised solution the electrical and thermal utilization of all components should be considered very carefully [2]. Fig. 1 demonstrates the heat transfer coefficient attainable with different cooling techniques [3].

II. PARAMETER ESTIMATION

Since losses in semiconductors generally increase with temperature, it is recommended to keep it as low as possible in order to improve the efficiency of the power electronic system. Conduction losses for an actual operating temperature could be estimated by using on-state voltage $U_{f0}$ and slope resistance $r_f$ characteristics at 25°C [4]:

$$P_{cond}(T°C) = \left( f_v \frac{1}{100°C} T°C + \frac{5 - f_v}{4} \right) \cdot U_{f0}^{25°C} + \left( f_r \frac{1}{100°C} T°C + \frac{5 - f_r}{4} \right) \cdot r_f^{25°C}, \quad (1)$$

where $f_v$ is a factor for the temperature dependency of $U_{f0}$, and $f_r$ expresses the temperature dependency of on-resistance $r_f$. The factors $f_v$ and $f_r$ depend on the device technology and blocking voltage (IGBT ≈ 0.76...1.56; FWD ≈ 1...1.73).

Analogous to conduction losses, both IGBT and diode switching energies are assumed to depend linearly on the temperature according to

$$P_{sw}(T°C) = \left( f_f \frac{1}{100°C} T°C + \frac{5 - f_f}{4} \right) \cdot \frac{1}{f_f} \cdot E_{125°C}, \quad (2)$$

where $E_{125°C}$ is the energy loss at 125°C; $f_f$ is the temperature coefficient (IGBT ≈ 1.32...1.6; FWD ≈ 2.28).

The first approximation of the average junction temperature is obtained using the one-dimensional equivalent circuit and datasheet values of the device. The equivalent circuit representing the heat conduction in a transistor module is shown in Fig. 2.

Fig. 1. Heat transfer coefficient attainable with different coolants and cooling methods.
where $R_{thu}$ is the thermal resistance of the $i$-th RC-pair, $\tau_i$ is the time constant of the $i$-th RC pair, $\tau = R_iC$ and $n$ is the number of the thermal time constants. These coefficients are often listed in the datasheet of the device.

In the majority of cases, the focus of interest with respect to a periodic sequence of power pulses is the stationary temperature swing in the stationary condition. It can be derived from the general equation of the Foster network in an analytical expression:

$$\Delta T_{\text{stat}} = P_{\text{pulse}} \sum_{i=1}^{n} R_{thu} \frac{1 - \exp\left(-\frac{t_{on\text{ }i}}{\tau_i}\right)}{1 - \exp\left(-\frac{t_{on\text{ }i} + t_{off}}{\tau_i}\right)}.$$  

where $P_{\text{pulse}}$ is the single pulse power dissipation, $t_{on\text{ }i}$ and $t_{off}$ are switch on-state and off-state times, respectively.

For low values of pulse width, the junction temperature is lower because the thermal capacity of the die, solder and package imposes various time constants on the rate at which junction temperature can rise. Therefore, for the same power level, at short durations, the thermal impedance appears to be smaller. This explains why the Safe-Operating-Area SOA is larger for short pulse widths [7].

The disadvantage of such an analytical approach is that it is based on the one-dimensional thermal model. Hence, it is impossible to estimate the influence of heat sources on each other neither is it possible to estimate the impact of non-uniform temperature distribution along the heatsink. The three-dimensional models created in computational fluid dynamics software, such as IcePack™, could provide more precise results. The comparison is shown in Fig. 3 [8]. In order to validate the precision of the model it is always recommended to compare simulation results with measurements or with the other models.

### III. STANDARD COOLING METHODS

#### A. Natural air cooling

This type of cooling is mostly used in low power range applications up to 50 W, however, it can be used as well in high power range applications if the use of fans is not possible or if extremely large cooling surfaces are available in the device [9]. Heatsinks are made of materials with high thermal conductivity such as aluminium (205 W/m K) and copper (400 W/m K). For natural air cooling, aluminium is more popular because it is cheaper, lighter and can be anodized, achieving improved power dissipation by radiation. To achieve good thermal transfer heatsinks with thick baseplate and fins with more space between them are used. Under natural convection, approximately 70% of the heat is transferred by convection and 30% by radiation [5]. To relate the $R_{thu}$ value to the size of the passive cooling system it is possible to use the graph in Fig. 4 illustrating a typical volume of heatsink required over a range of thermal resistances for natural convection. Generally, to reduce thermal resistance by 50%, the heatsink volume must be increased by a factor of 4, assuming that all the other parameters remain constant [10].
Therefore, thermal management techniques with natural convection for applications with higher power dissipation result in substantial physical volume required to handle large amounts of the heat dissipated. This drawback could be eliminated by using passive heatsinks with integrated heat pipes or by utilizing more efficient cooling methods such as forced air or liquid cooling.

**B. Forced air cooling**

With highly increased convection in contrast to natural air cooling, this way of cooling is the most widely used technique in the present power electronic systems. Highly efficient forced air cooled heatsinks have a hollow-fin structure that consists of basic extruded aluminium profiles with extruded ribs pressed into the profiles. The fin geometry is optimized in terms of flow resistance and thermal efficiency. Such a shape provides an increased surface area and improved turbulence of the air, while one or more fans provide necessary air volume to absorb the heat from the active surfaces of the heatsink. This is a relatively inexpensive thermal management solution for medium- and high-voltage IGBTs packaged in brick-style standard module formats and it is generally used for cooling converters with rated power up to 1 MW [11]. The thermal resistance of a forced air cooling system is the function of air mass flow, temperature difference, heatsink materials, and geometry. The operation parameters are selected according to the cross point between two datasheet characteristics: air flow volume and pressure drop (Fig. 5). The fan must be able to deliver suitable pressure (in order to force the cooling air through the cooling system) and should provide required airflow performance within its optimum operating range. Other considerations include noise level, lifetime, power consumption or speed adjustment. Axial and diagonal fans provide some advantages over radial ones such as smaller volume and weight, lower power consumption and easier speed control, but at increased price. An important limitation of the conventional systems is the non-uniform distribution of heat across both the heatsink base and fins in the case of high heat density semiconductors [2].

**IV. LIQUID COOLING**

The most efficient way of cooling >1 MW converters is to utilize liquid cooling. Liquid-cooling takes advantage of a liquid’s higher heat density, capacity and thermal conductivity in comparison to air. This enables cooling concentrated heat sources in remote locations, providing more freedom in the physical design of power electronic systems, allowing more compact higher power density systems to be developed.

A liquid cooling loop for contact cooling typically consists of a cold plate, a pump, a heat exchanger, and pipes or hoses. Heat generated by a component is transferred from the cold plate, completing the liquid cooling loop (Fig. 6).

Manufacturers and developers are proposing a wide range of different fluid cooling solutions, ranging from relatively simple closed loop water convection systems to very complex ones, utilizing vaporizable dielectric fluids, integrated cold plates with micro-jet arrays, air chillers etc.

**A. Heat exchangers**

A heat exchanger is a device designed to efficiently transfer heat from one matter to another. Liquid-to-air heat exchangers are generally used to cool process fluids in closed loop cooling systems. Heat exchanger standard technologies include tube-fin heat exchangers, flat tube oil cooler heat exchangers and plate-fin heat exchangers. The performance of any heat exchanger depends on several factors, including tube attribute’s liquid flow rate, fin surface area and airflow available for heat rejection. Optimization of these system
elements can result in good thermal performance from low air-flow volume, which enables system fans to run at lower speeds and more quietly. Where fan noise is not a concern, higher airflow results in even a better performance (Fig. 7).

Manufacturers typically provide performance data for heat exchangers in thermal resistance described by the following equation [12]:

$$R_{thHE} = \frac{T_{Lout} - T_{Ain}}{P_{tot}}, \quad (5)$$

where $P_{tot}$ is the total heat load, $T_{Lout}$ is the exit coolant temperature, $T_{Ain}$ is the inlet air temperature.

B. Pumps

Along with a heat exchanger, a pump is the second power consuming active component in a typical closed loop liquid cooling system. The systems are designed for maximum performance at a specific flow rate. The pump must deliver the required liquid flow rate, while overcoming the pressure drop in the cooling system. Pressure drop is a result of resistance caused by friction (shear stresses) or other forces (such as gravity) acting on a fluid. It is exponentially proportional to the flow rate. In fact when the flow rate doubles, the pressure drop typically increases by a factor of four [12]. Two types of pumps are generally used in power electronics cooling systems:

- Positive Displacement Pumps (PD pumps) displace a known quantity of liquid with each revolution of the pumping elements. They provide a constant flow rate regardless of the pressure drop across the system. This makes them ideal if the system’s pressure drop is high or a constant flow rate regardless of changes in the system's pressure drop is required. It is recommended to replace PD pumps approximately every 7,000 hours of operation to ensure reliable performance.

- Centrifugal and Turbine Pump’s output varies considerably with the pressure drop across the system. If the pressure across the system changes, the flow rate from the cooling system will also change. A centrifugal pump is suitable for small pressure drops; a turbine pump operates at higher pressure drops. It is recommended to replace centrifugal and turbine pumps approximately every 28,000 hours of operation to ensure reliable performance [12].

In order to select a pump with appropriate pressure and flow rate, the pump curve provided by the pump manufacturer should be plotted over the system pressure drop curve. The system will operate at the intersection of the two curves (Fig. 8).

C. Cold plates

Cold plates are the key component of any liquid cooling system, since they are in contact with high heat flux components that must have waste heat to be removed to prevent overheating. Performance of a cold plate is dependent on a number of factors including power dissipation of the device cooled, cooling fluid, size and arrangement of the cold plate, ambient conditions along with the performance of other components in the loop. Similarly to forced air cooling the performance of a standard cold plate is typically described by a graph, indicating the dependency between thermal resistance, pressure drop and liquid flow rate. The optimum working point is selected as the cross point of two characteristic curves (Fig. 9) [13]. As shown, the dependency is not linear and after a certain flow volume the thermal resistance is not reduced by a significant amount and could be defined as [14]:

$$R_{thCP} = \frac{T_{CP} - T_{Lin}}{P_{tot}} = \frac{1}{h \cdot A}, \quad (5)$$

where $P_{tot}$ is the total heat load evenly distributed over the entire cold plate surface, $T_{Lin}$ is the inlet coolant temperature, $T_{CP}$ is the cold plate surface temperature, $h$ is the heat transfer coefficient, and $A$ is the surface area.

Hence, one way to boost the performance of a cold plate is to increase its internal surface area. Depending on the design, the internal surface area of a cold plate can be increased by increasing the number (density) of fins, pins, jets etc. However, this can also have the effect of an increase in the pressure drop across the cold plate due to viscous losses, leading to greater hydrodynamic power consumption.

Another approach is to increase the heat transfer coefficient. Primarily it depends on the liquid flow rate, type of flow, i.e. laminar or turbulent, and geometry of the cold plate.
(see Fig. 9). The type of flow is characterized by the Reynolds number [3]:

\[ \text{Re}_D = \frac{\rho U_m D_h}{\mu} \]  
(5)

where \( \rho \) is the coolant density, \( U_m \) is the coolant mean velocity, \( D_h \) is the hydraulic diameter of the cooling chamber, \( \mu \) is the coolant viscosity. A flow is considered laminar if \( \text{Re}_D < 2100 \) and turbulent if \( \text{Re}_D > 4000 \).

Most cold plates are made of aluminium, but some new technologies use copper. Although copper has better thermal conductivity, aluminium is used more often because copper is very difficult to machine and it is more expensive. If aluminium meets the thermal performance specifications, it is generally the best material to use. Cold plates could be classified as swaged-tube, vacuum-brazed and microchannel cold plates.

Thermal performance of the cold plates is affected and limited by the thickness of coolant’s boundary layers. Typically, the liquid flow in such cold plates is laminar, in this case the boundary layers are thick and heat conducts slowly to the centre of the channel creating relatively high temperature difference between the cold plate and the liquid. The performance could be increased by greatly increasing the fluid flow rate to create a turbulent flow. Turbulence reduces the influence of a fluid’s boundary layers since a cold fluid in the middle of the channel mixes with a hot fluid next to the wall of the channel. An example is a custom-made aluminum vacuum-brazed cold plate dissipating approximately 15 kW of heat (six 160x120 mm IGBT modules), having a pressure drop of 100 kPa at 50 l/min, produced by Lytron™. However, this approach requires a powerful pump and the ability of the cooling loop to sustain high pressure and fluid flow rates.

Another way to reduce the influence of the boundary layer is to reduce channel width and increase density to create very thin layers of the fluid. These microstructures neither require nor create a turbulent flow. Instead, they take advantage of the fact that in the case of very thin channels the influence of the boundary layer decreases, resulting in a low mean temperature gradient between the cold plate fins and the liquid flow, making them ideal for applications where heat loads are concentrated. Another advantage is the requirement of significantly lower liquid flow volume (Fig. 10) [15]-[17]. Microchannels, however, are very sensitive to liquid purity, corrosion, high temperature alterations and cause high pressure drops.

![Fig. 9. Typical cold plate characteristic curves without turbulators (solid line) and with turbulators (dotted line) [13.]](image1)

![Fig. 10. Thermal resistance vs. flow rate for different commercially available cold plates.](image2)

![Fig. 11. Side-by-side comparison of IGBT module structures.](image3)

![Fig. 14. Liquid flow cooling inside of the baseplate.](image4)

![Fig. 15. Standard power module on a liquid cooled plate.](image5)
because it results in an excessive thermal resistance or produces an early failure if it causes the baseplate of the power module to deform [20]. If installed properly, the $R_{thcs}$ value depends on the thermal resistance value of the thermal interface material between the module and the heatsink. Thermal interface materials come in a wide variety of product types and price points: thermal compounds (greases), thermal pads (films) or phase change materials (PCMs) [21] and could be either electrically conducting or insulating. Impact of typical TIM thickness on the maximum temperature is shown in Fig. 13. The thermal resistance of the TIM could be estimated by

$$R_{thcs} = \frac{L}{A_m} + R_{c,t} + R_{ts},$$

(5)

where $k$ is the thermal conductivity of the thermal interface material, $L$ is the thickness of TIM, $A_n$ is the area of TIM’s layer, $R_{c}$ and $R_{ts}$ are contact resistances of the TIM at the boundary with the two surfaces.

In addition to traditional TIMs, alternatives such as thermally conductive adhesives, polymer solder hybrids, and metal composites are increasing in use. Typical properties of different TIM solutions are presented in Fig. 12. However, TIMs are still a bottleneck in many traditional thermal management systems (Fig. 13), since thermally induced stresses that arise from differences in the coefficient of the thermal expansion of the various materials used can result in thermal grease pump-out or dry-out [22] causing significant deterioration in the interface thermal performance, limiting thermal conductance, and the dissipated power coming from the power chip [23]. This barrier can be eliminated by embedding the baseplate of the module with the cold plate with either macro- or microchannel structure. In this case coolant flow is in direct contact with the baseplate (Fig. 15). Investigations [24] show that this approach has the capability of reducing the thermal resistance between the semiconductor chip junction and the ambient by about 60%. However, it requires modification of the IGBT package. Therefore, such techniques are more expensive than standard modules assembly configurations.

C. Double-side cooling

Converters ranging in power from a few MW to a few 100s of MW utilize considerable numbers of semiconductor devices. As converter voltages range from 10s or even 100s of kV, series connection of a large number of semiconductor devices is essential [25]. Most suited for series connection is the stacking of devices on top of each other, as it is well
known from press-pack thyristors and diodes. This type of housing offers several advantages compared to modules, like a higher thermal and power cycling capability, easier series connection and explosion-free failure mode. In addition to thyristors, IGBTs in press-pack housings are commercially available as well. These modules have essentially increased current capabilities (Table 1) and are designed to replace GTO thyristors in pulse power industrial and transportation applications [26][27].

D. Two-phase cooling systems

The most widespread technology involving liquid to vapor phase change is heat pipe assisted cooling systems. The heat pipe is a passive system that combines the advantages of both phase change and thermal conductivity to achieve greatly increased heat transfer properties (5-200 kW/m K) in comparison to solid metals (0.25-1.5 kW/m K) [28]. Since heat pipes can serve only as heat conductors, they are integrated into other heatsink technologies to further increase their thermal efficiency. The modern heat pipe assisted forced air cooling solutions designed by Thermacore™ dissipating 6 kW of heat while maintaining more uniform temperature distribution along the mounting plate, occupying 27% less volume and consuming less air flow volume than traditional extruded heatsinks [29].

Another alternative for improving heatsink performance is to use flat heat pipes or heat chambers applied on the basis of the heatsink to provide more uniform temperature distribution along the surface.

Further improvements in the performance can be achieved by active two-phase cooling loops involving boiling. The evaporative coolant takes away heat losses by the use of latent heat when a liquid turns to vapor. Cooling can be direct (immersion, using dielectric fluids) or indirect (water containing liquids could be used) and can involve pool boiling, flow boiling, and jet impingement boiling. Heat generated from the module is conducted to the boiling section in the system. The vapor produced from the boiling is then condensed by going through the condenser. Finally, the condensed liquid flows back to the boiling section (Fig. 14). These cooling solutions offer impressive results [30], however, such systems experience a significant effect of module load cycles, ambient conditions and component orientation on the performance. Aside from heat pipe-based cooling systems, two-phase cooling is not typically used for cooling power electronics at present.

VI. CONCLUSIONS

An optimum power electronic system design requires many factors, such as size, noise, complexity, robustness, cost, maintenance requirements, to be evaluated. Intelligent power management is required to achieve the desired result. As a result of increased heat dissipation from electronic devices, alternative methods of cooling such as liquid cooling are becoming more attractive. Considering the wide range of available cooling solutions the following conclusions can be made:

- Forced air cooling is generally the best option for <1 MW converters.
- The performance of the natural or forced air systems could be improved by achieving more uniform temperature distribution by the help of integrating heat pipes or heat chambers while maintaining the positive properties of such systems.
- For higher power applications liquid cooling systems offer a wide range of new possibilities and could considerably improve the performance of the semiconductors.
- Despite showing better results, two-phase methods can have unpredictable performance depending on ambient conditions are costly and harder to implement and control.
- Single-phase methods are simpler, more reliable and predictable. From single-phase systems tube or fin designs are cost-effective, while jet impingement and microchannel show better performance.
- Future directions are increased application of liquid cooling systems as well as their further integration with semiconductor devices.

ACKNOWLEDGEMENTS

This research work has been supported by the Estonian Ministry of Education and Research (Project SF0140016s11), Estonian Science Foundation (Grants 7425, ETF7572) and Estonian Archimedes Foundation (project “Doctoral school of energy and geotechnology II”).

REFERENCES


Andrei Blinov received his B.Sc. and M.Sc. tech. in electrical drives and power electronics from Tallinn University of Technology, Tallinn, Estonia, in 2005 and 2008, respectively. From 2008 he pursues doctoral studies in Tallinn University of Technology. Andrei Blinov is a Research Fellow in the Department of Electrical Drives and Power Electronics, Tallinn University of Technology. His research interests are in simulation and research of switchmode power converters, semiconductor heat dissipation and different cooling systems.

Dmitri Vinnikov received Dipl.Eng., M.Sc. and Dr.Sc.techn. in electrical engineering from Tallinn University of Technology, Tallinn, Estonia, in 1999, 2001 and 2005, respectively. He is presently a Principal Research Fellow in the Department of Electrical Drives and Power Electronics, Tallinn University of Technology. He has authored more than 100 published papers on power converters design and development and is the holder of several Utility Models in this application field. His research interests include switchmode power converters, modeling and simulation of power systems, applied design of power systems and control systems and application and development of energy storage systems.

Tõnu Lehtla graduated in Tallinn Technical University in 1970. Engineer of Electrical engineering in the field of electrical drives and industry automation. Received Ph.D. degree in 1976. He is a Professor and head of the chair of Robotics in the department of electrical drives and power electronics. He has authored over 100 scientific and methodical publications in the field of electrical drive, industry automation and robotics.