Liquid-cooling on the Agilent 93000 SOC Series

The Advantages for Testing and Cost of Test

Abstract

Semiconductor devices emit heat during operation. The operating temperature of the device affects circuit performance and more importantly, device reliability. This is a significant issue for semiconductor test because if the temperature of the tester electronics cannot be stabilized to a target level, yield will decrease and repeatability will degrade. If the target level cannot be kept relatively low, then system reliability is significantly reduced.

Automated Test Equipment (ATE) systems employ cooling techniques based on either air or liquid cooling mediums. Liquid cooling systems offer greater temperature stability over air. The higher heat transfer efficiency of liquid cooling systems also means that ATE operating temperatures can be reduced. This leads to greater system reliability, reduced tester operating costs, improved throughput and protection of the investment in the tester.

This paper explores the effects of temperature on semiconductor devices and evaluates the benefits of liquid cooling over air cooling in ATE systems.

1.0 Introduction

This document examines the issues associated with the cooling of ATE. It starts by looking at the fundamentals of cooling semiconductor devices, the importance of cooling efficiency, and a comparison of air and liquid-cooling in ATE systems. It discusses the design issues in developing liquid cooled ATE, and how Agilent Technologies' years of experience and consistent cooling strategy contribute to the effectiveness of the Agilent 93000 architecture. Finally, it considers the fundamental technical and financial requirements of semiconductor ATE, and how meeting these requirements creates challenges that are best met by liquid cooling.

2.0 Fundamentals of Cooling Electronic Circuits

Semiconductor devices produce heat due to leakage currents (steady state) and the switching action of transistors. The amount of power (heat) to be dissipated depends upon the number of circuits in the device, their switching, speed and the load on the circuit. Moore's Law has driven the power density of semiconductor devices to very high levels. Today's state-of-the-art CMOS devices can produce up to 50 Watts of heat or more for a silicon die that is $2cm^2$ in area. To understand more about the challenges of cooling semiconductor devices and ATE systems (which consist of many semiconductor devices) we need to consider a few important concepts.

Junction temperature, T_j, is the operating temperature of the circuit. The switching speed of CMOS circuitry depends partly on T_j – the lower the operating temperature, the faster the circuit will switch. Reliability is also directly affected by T_j. The failure rate for the component approximately doubles for about every 10°C[1] rise in operating temperature. An effective cooling system maintains T_j at an optimum, low level. The thermal resistance ($\theta_{j\text{-}a}$) of the cooling system can be measured by how well it can maintain a constant T_j as the amount of heat to be dissipated increases. That is:

$$\textbf{\textit{Thermal Resistance}}\ (\theta_{j\text{-}a}) = \frac{\Delta T_j}{\Delta\ \text{(Dissipated Heat)}} \quad in\,°C/W$$

$\theta_{j\text{-}a}$ denotes the thermal resistance from the device to the ambient reference.

For example, in a system with a thermal resistance of 5°C/W, and a device dissipating 1W at standby and 10W in operation, the operating temperature of the device will rise by 45°C (5°C/W * 9 W) when the device goes from standby mode to operating mode under steady-state operating conditions.

Sidebar 1

The thermal capacity of a material is the amount of energy (kilowatt hours –kWh) required to raise the temperature of 1kg of the material by 1°C. The thermal capacity of water is high, taking 1.16Wh to heat 1kg. The following table shows the thermal capacities of some common materials.

Material	Specific Thermal Capacity	
	J	kWh
Helium	5.20	1.44
Air	**1.005**	**0.28**
Transformer oil	1.88	0.52
Water	**4.18**	**1.16**
Glass	0.84	0.23
Aluminum	0.897	0.250
Iron	0.449	0.125

---

[1] Though some conventions use Kelvin for relative temperatures, in this document, both relative and absolute temperatures are given in Celsius.

2.1 The importance of coolant thermal capacity

Because T_j affects the performance of the circuit, it is important for the cooling system to keep the temperature stable and independent of environmental and operational factors such as air pressure and circuit loading. This has direct implications for the repeatability and stability of the circuit, which impacts the DUT (device under test) yield, test repeatability and diagnostics capability of semiconductor ATE equipment.

Cooling effectiveness and efficiency depend on factors such as heat sink design, the properties of the fluid (liquid or air) that is used to transport the heat away from the device and the heat transfer characteristics between the heatsink and the cooling fluid. A major role is played by the thermal capacity (see Sidebar 1.) per unit volume of the coolant. With a much higher thermal capacity, water is better than air at achieving low device operating temperatures and lower variability in the operating temperature.

2.2 A comparison of air and liquid as ATE coolants

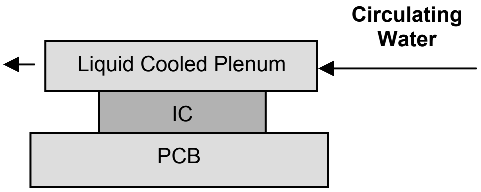


Figure 1

In a water-cooled test system, the temperature of the liquid cooled plenum is controlled directly by the liquid circulating through it. With good thermal contact to the device, T_j can easily be kept to within 1°C of the target temperature. For example, in the Agilent 93000 SOC Series tester, a small plenum (see Figure 1) with water circulating through it is sufficient to meet the cooling requirements of even the high performance cards and analog instruments. With the water temperature at 30°C, T_j can be kept at 35°C ± 1°C, guaranteeing performance, reliability, stability, and repeatability.

In contrast: the efficiency of an air-cooled system is limited by its heat sink design (see figure 2), and the speed, direction and uniformity of the air flow. Stability is limited by the formation of "dead spots" or "hot spots" in the air flow. The need for heat sinks and adequate space for air to flow around the components results in lower packing density for the pin electronics. Lower packing density can be a major issue in configuring a test head for a particular applications. Since the airflow in the test head has limited cooling capability, the full range of instruments and tester electronics may be limited because only alternating slots may be available in the testhead in order to allow enough cooling air to flow around the pin electronics.

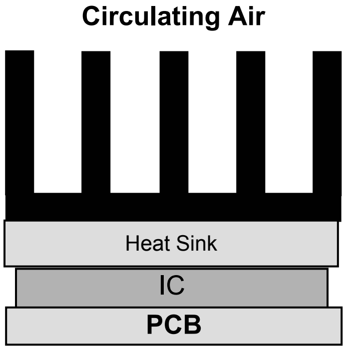


Figure 2

Lower packing density also limits top end speeds and precision because longer propagation delays and larger parasitics from longer signal lines degrade the signal under test. Lower thermal efficiency makes it harder to maintain the pin electronics at a stable operating temperature, which affects the stability of the tester, and is made worse by the presence of pressure and temperature gradients in the air flow. Lower stability translates directly into lower device yields (due to wider guardbanding) and poorer repeatability.

The physical size, vibration and noise of an air-cooled SOC test system can also be an issue. Air-cooled test systems can be as much as six times the size of a comparable liquid-cooled system, and even then they will not keep T_j within the specifications that can be achieved by liquid cooling. In many cases the integration of the fans in the system creates issues in balancing performance versus reliability.

In an air-cooled semiconductor ATE system, when channel performance and the corresponding heat dissipation per channel increase, the volume of air passing by the components must increase to keep T_j constant. In some cases, improved ducting and exotic heat sink options (such as PinFin) are needed to keep the circuits within operating specifications. This takes larger or more powerful fans, and increases the surface area of the heat sinks and boards to allow enough thermal interface area between the heatsink and the air. This further increases the noise level and the overall size of the tester. As the warmer air emerges from the tester, keeping the ambient temperature constant places an increased load on the heating, venting and air conditioning (HVAC) system, which can be expensive to replace or upgrade. Liquid cooled systems use efficient heat exchangers to dump heat into facility water cooling networks or to connect to remote chillers that dump heat into outside ambient air.

3.0 The Design of Liquid Cooled ATE

3.1 The limitations of air cooling

Few system designers would dispute that liquid-cooling is more efficient. In fact, the efficiency of liquid in absorbing and transporting heat is the reason liquids are an established feature of many types of equipment. This characteristic of liquid is also the reason that liquid is the cooling architecture of choice for many high performance ATE systems.

Many manufacturers of semiconductor test equipment continue to rely on air-cooling because they did not design their original platforms for the type of high performance that requires liquid-cooling. As the performance requirements of semiconductor ATE systems continue to increase, air cooled systems are finding themselves with a dead end cooling architecture. This is forcing a choice between moving to a liquid cooled scalable architecture or maintaining different architectures for low performance systems and for high performance systems. The latter choice is temporary at best, since Moore's Law will continue to put pressure on air cooled ATE.

ATE users also must accept the risk due to the lack of scalability and performance flexibility if they are using a vendor that delivers different architectures for high and low performance systems because they cannot mix and match features and instruments to provide an optimum solution.

3.2 Early adoption of liquid cooling

Agilent Technologies identified liquid-cooling as the preferred solution for its semiconductor ATE systems over ten years ago. Agilent Technologies' approach to the problem was to develop an architecture that could be scaled with IC manufacturing processes. This meant that Agilent looked at the scalability, reliability and the cost of ownership of the liquid cooled architecture before making the decision. Agilent is now on its fourth generation of tester pin electronics cards for the 83000 and 93000 Series testers. A number of patents are pending on some of Agilent's innovations in liquid-cooled electronics. Today, the full range of performance needed to test SOC devices can be provided on a single, scalable platform. Because the cooling capacity of Agilent's water-cooled system is easily increased, there is still plenty of headroom for coping with further demands in pin-counts, interface speeds, and functionality envelopes.

3.3 Water cooled card design

Cooling is an integral part of the design of the test cards for an ATE system. For cards in an Agilent 93000 SOC Series tester, the design of the plenum must ensure the correct volume of water passes each of the components to collect the heat and provide the stability and flow needed by that component. The cross-section, shape and path of the channel must guarantee an even flow without eddies or stills that could become hot spots.

Agilent decided on the use of water to maximize cooling efficiency while minimizing cost. Water is environmentally friendly, non-toxic and readily available (Agilent testers use drinking quality water, all the necessary conditioning is handled by the system). Some other manufacturers use Flourinert (FC-77 – an electrically inert liquid) as a cooling liquid due to the lack of design features to prevent spillage. This is an expensive tradeoff due to the cost of Flourinert and its environmental handling requirements. With over ten years using liquid-cooling in its ATE systems, Agilent's uncompromising approach to the design for watertight and spill proof operation has eliminated the possibility of damaging leakage from of the tester. There are well over 1,000 Agilent water-cooled testers currently in operation around the world.

4.0 Cooling and the True Cost of Semiconductor ATE

Semiconductor ATE must be able to perform the appropriate tests at the appropriate performance level, it must do this repeatedly and consistently, and most importantly, it must be able to do this reliably.

4.1 Liquid Cooling reduces the cost of test

Cooling affects tester performance in three important ways.

Smaller form factor – Liquid cooling helps determine the density of components, the operating speed, and the precision of the pin electronics. More efficient cooling enables higher component densities, which decrease propagation delays within the tester and eliminate the distortions inherent in long paths to the Device Under Test (DUT). Cooling efficiency also determines usable performance of the devices in the tester. Most devices are optimized for operation at around 90°C, though lower operating temperatures mean consistently faster slew rates. Similarly, cooling efficiency also influences precision by ensuring that changes in operating conditions and in the operating environment do not result in a significant change in operating temperature (T_j).

Improved stability – Liquid-cooling also has the thermal capacity to ensure repeatability. Often, with today's pico-second accuracies, re-calibration is required any time the temperature changes by more than 1°C since the last calibration. Because the temperature in an Agilent tester is kept within this limit, there is no need to stop production in order to recalibrate the ATE to verify repeatability over the course of a day, or even over

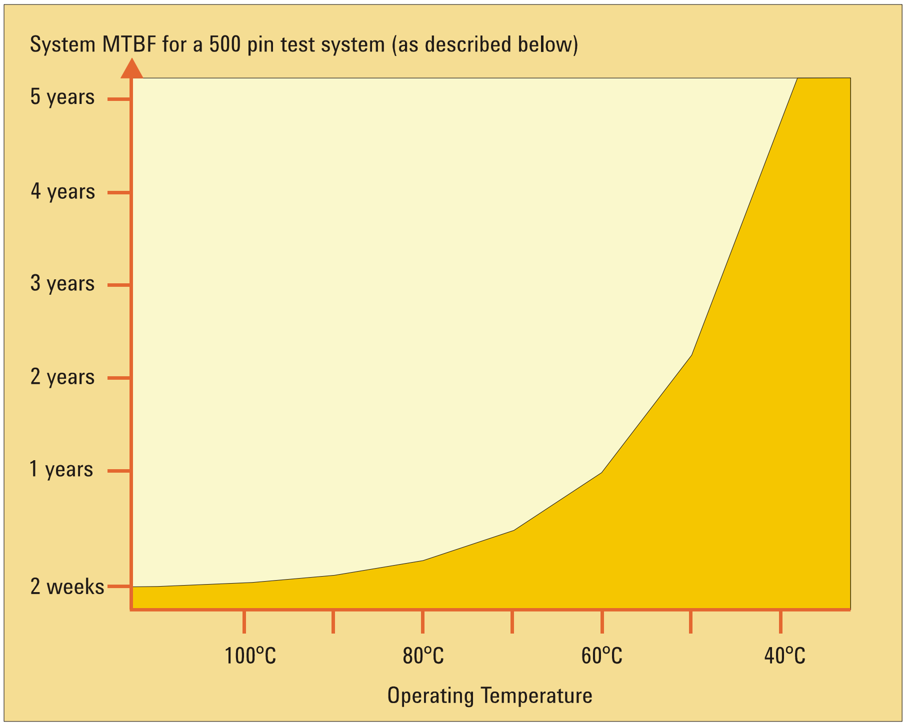


Figure 3

longer periods. All of this helps maximize test-time and throughput.

Improved reliability – The most important effect of cooling architecture is the effect on tester reliability. The Mean Time to Failure (MTTF) of a semiconductor device can be predicted based on its FIT (Failures in 10^9 device operating hours) rate. The steady state operating temperature of the device has a significant impact on the FIT rate of a device. In general, each 10°C rise in steady state operating temperature can increase the FIT rate by as much as 2X (or cut the MTTF by as much as 50%). The following example illustrates this relationship as applied to a semiconductor tester.

Assume a semiconductor test system that consists of 500 digital channels, each with one CMOS ASIC, for a total of 500 ASICs in the system.

Using FIT rate analysis we find that the MTTF for the CMOS ASIC at an operating temperature of 90 C is 40 years. With a 500 ASIC system, the MTBF (Mean Time Between Failures) of the system is roughly one month (40 years/500) This means we can expect one of the 500 devices to fail per month. Given the relationship between operating temperature and MTTF of each ASIC, every 10° C drop in operating temperature of the system approximately doubles the MTBF of the system. If the cooling system of the tester can reduce the operating temperature to 40° C, the MTBF of the system can be increased to 32 months (1 month x 2^5). This means that by cooling the system we have moved from expecting an ASIC failure every month to expecting an ASIC failure once every two and a half years.

The graph in Figure 3 illustrates this relationship. By running the pin electronics at lower temperatures, Agilent's liquid-cooling architecture directly boosts the reliability of the test system, reducing maintenance and repair costs and significantly increasing availability over comparable air cooled systems. Because testing is a bottleneck in the manufacturing process, any reduction in tester downtime has a positive effect on production efficiency and manufacturing costs.

4.2 Other cost savings of liquid cooling

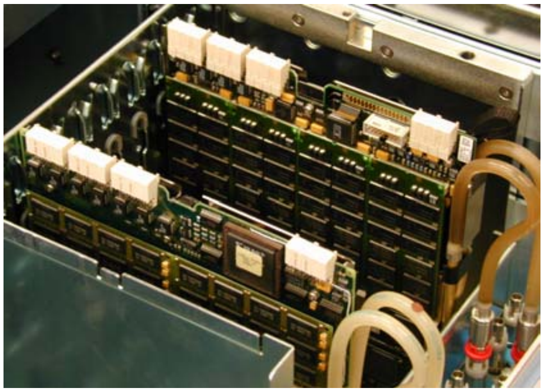
Quiet, fan-free operation of liquid cooled testers helps meet the environmental requirement (ISO 7779) for acoustic noise and eliminates special requirements for the air conditioning system. Fans also introduce electro-magnetic fields and mechanical vibration to the test system, which is an issue for some RF components, and for Micro Electro-Mechanical devices (MEMs) that are prone to microphonic effects. Though this is not a major issue currently, it is likely to be of increasing importance. Vibration can be reduced by a suspended mounting structure for the fans, rather than a rigid mount onto the testhead sheet metal. However, liquid-cooling once again provides a much better solution.

Liquid-cooled test boards enable component densities that cannot be achieved in air cooled systems. The result is a smaller tester footprint, which reduces the floor space required by a factor of 3x or 4x (particularly important for personal testers, and in clean rooms where space is expensive). Compact testers also place lower demands on the infrastructure. Smaller testheads also have lower power ratings and require smaller manipulators (with lower operating and depreciation costs).

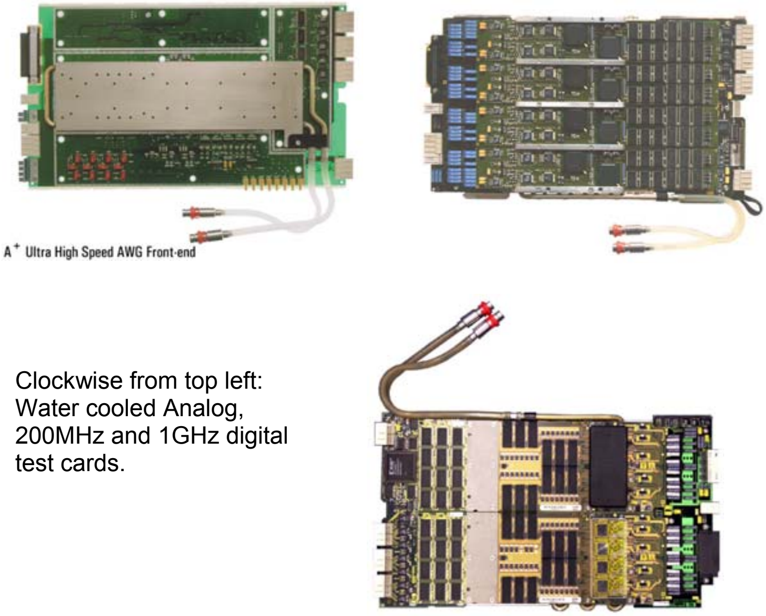
4.3 Liquid-cooling protects test investment

By minimizing operating costs, and maximizing tester throughput (through both high performance and reliability), liquid cooling plays an important role in lowering the total cost of test. Although air-cooled architectures are capable of supporting signal speeds in the 300 MHz to 500 MHz range, the cooling strategy must change for faster test speeds.

In an air cooled system, if one or more pins on the DUT require faster speeds, the test platform and the existing cards must all be replaced to be compatible with the liquid cooled requirements for the higher performance pin electronics. Because Agilent uses liquid cooling for its whole range of pin electronics, the highest to the lowest performance can be mixed in a single test head, protecting the customer's test investment. Liquid-cooling also extends the useful life of existing cards because they can be used for lower speed pins together with higher performance cards to match the DUT pin by pin. In this sense, liquid-cooling is a key enabler of Agilent's single, scalable platform. Figure 4 illustrates the common form factor and architecture of the 93000 pin electronics. This commonality of architecture and form factor is enabled by the 93000's liquid-cooled architecture.



500MHz C*e*-model and 2GHz P-model test cards can be used side-by-side in the same tester without problems.



Clockwise from top left: Water cooled Analog, 200MHz and 1GHz digital test cards.

Figure 4

5.0 Conclusion

The operating temperature of semiconductor devices is an important factor in determining their performance and reliability. As power densities increase, cooling efficiency becomes an increasingly important element in the design of semiconductor ATE, to ensure it meets performance, repeatability, reliability and operating cost requirements. These challenges are best met by liquid cooling.

While air cooled testers may be initially less expensive to design and develop than liquid cooled testers, their cost of ownership can be significantly higher than liquid cooled systems when system operating costs, reliability and performance are taken into account. The scalable liquid cooling architecture of the Agilent 93000 SOC Series reduces COT by reducing operating costs and by improving throughput.

Liquid cooling reduces operating costs through:

- Maximized configuration flexibility that enables an optimized solution
- Improved MTBF
- Higher availability
- Lower infrastructure costs

Liquid cooling improves throughput through:

- Higher performance and precision of the tester electronics, which enables tighter guardbands and higher yields
- Optimal configuration for SOC testing, to eliminate multiple insertions and enable higher degree of parallelism in test

Agilent Technologies has a strong background and proven experience in developing liquid-cooled test equipment. The excellent track record and low running costs of the Agilent 93000 SOC Series demonstrates the ways in which liquid cooling can be used to lower COT.