OCP OAI System Liquid Cooling Guidelines

Authors:

Cheng Chen, Meta Platforms Inc.

Dennis Trieu, Microsoft Corporation

Tejas Shah, Intel Corporation

Allen Guo, CoolerMaster

Jaylen Cheng, Wiwynn Corp

Christopher Chapman, Boyd Corporation

Sukhvinder Kang, Boyd Corporation

Eran Dagan, Habana Labs

Assaf Dinstag, Habana Labs

Jane Yao, Enflame Technology Co., Ltd

Executive Summary

This document is not a specification for OAI/OAM products. It is a set of guidelines on the design, validation, and implementation of liquid cooling solutions for AI Training Systems with 8x OAM products or others alike. Contents of the document would help a user/designer/supplier of OAI/OAM products understand the basics around those topics/questions related to liquid cooling:

- What are the typical boundary conditions and expectations on the customer side
- What are the possible thermal challenges for OAM products
- Basic considerations around cold plate solution
- How to develop a passive cold plate loop solution for UBB based systems and others alike
- How to validate the liquid-cooled system
- Best practices and case studies

For most engineering topics/questions discussed in this document, we (the OAI Cooling workstream members) are contributing what we believed to be best practices as of today. However, for each product, there would be more than one way to design/validate/use it, not to mention potential technology evolvement or changes of dependencies down the road. Please keep open-minded while reading this document, and do not hesitate to contact us directly for feedback and further discussion.

Table of Contents

Introduction

In OAM (Open Accelerator Module) Spec 1.1 [1], we've provided insights into the air-cooling capabilities of OAM products and characterized an air-cooling limit of 450W per module based on a certain set of assumptions and boundary conditions. There are still ways to further extend air cooling for OAM products at higher power levels, however, restrictions may apply. More advanced cooling approaches are needed to support OAM product roadmap development up to 700W or even 1000W level [2], with enough confidence and fewer dependencies on product, system, and infrastructure designs.

Liquid cooling (with cold plate) is one of the most promising technologies on the horizon, as the eco-system is more prepared and extensively studied in many hardware product spaces. Yet there are a variety of design parameters and risks to consider for each platform design. OAI/OAM-based systems are one of the most challenging scenarios from the thermal/mechanical design perspective, due to the high module power and dense component layout.

In this document, we will provide a set of basic guidance, technical requirements, and best practice for OAI/OAM products using liquid cooling solutions. It aims at setting a foundation of common understandings to design and implement OAI/OAM-based platforms or similar products, for hyperscale users. Multiple studies showing the case of the passive cold plate loop (PCL) design and cooling limit analysis will also be provided.

This document is not intended to define a common specification for OAI/OAM-based products using liquid cooling. For design options not demonstrated in this document, or data out of the range listed, those application scenarios may also be valid under specific conditions. Actual cooling performance and optimum design can differ significantly from case studies demonstrated in this document, due to variations in assumptions/conditions such as package layout, mechanical tolerance, environment conditions, etc.

1. User requirement

1.1. System Layout

A typical OAI training platform has 8x accelerator modules (OAM) on the board (refer to UBB Spec [3] for more details), placed in a 4 (width) x2 (length) matrix When the OAM power or environmental condition results in demand of liquid cooling, typically all 8 OAMs will be liquid-cooled with cold plate solutions, where other components could also be included in coolant loop, depending on system architecture design. Such solution is typically capable of cooling 700W OAMs (dependencies apply), whereas the OAI group already foresees the potential need of supporting up to 1000W OAMs in a few generations.

1.2. Coolant Loop

Several types of coolant loop networks can be applied to support 8x OAMs. The most common practice is 4x parallel paths, every 2x OAM cold plates in serial. Such layout would reach a balance between preheat and coolant distribution complexity. It would also lead to a sufficient flow rate to reach the optimum operation range of every single cold plate.

Another loop design is to have all 8x OAM cold plates in parallel. Such layout would remove preheat, potentially maximizing cooling capability. However, it requires the cold plate able to deliver decent performance at a lower flow rate (speed) and has the extra challenge of coolant distribution and serviceability.

In some rare cases, 2x4 cold plate network (every 4x OAM cold plates in serial) could also be adopted if flow impedance and preheat can be kept within acceptable level. Such layout is more suitable for scenarios where liquid flow rate is low, supply coolant temperature is low and OAM power is not high.

1.3. Coolant Flow Rate

The typical operation range for OAM cold plate is expected to range between 1 and 2 liters per minute, resulting in chassis-level flow rate between 4 and 16 liters per minute, depending on the coolant delivery loop and OAM power consumptions.

Considering heat exchanger efficiency at the coolant supply side, either CDU HX or facility HX, it's recommended to keep the coolant temperature rise between 7.5°C and 12°C, which is equivalent to a flow rate/heat dissipation ratio of 1.25 LPM/kW~2.0 LPM/kW, based on the properties of PG25 based coolants. A typical design target is 10°C temperature rise, i.e. 1.5 LPM/kW. Higher temperature rise would benefit PUE, at the cost of cooling performance penalty, and vice versa for lower temperature rise.

1.4. Coolant Supply Temperature

Table 1. ASHRAE definition of air and liquid supply temperature categories [4]

Class	Typical Infrastructure Design		Facility Water Supply Temperature °C
	Primary Facilities	Secondary Facilities	
W17	Chiller/Cooling Tower	Water-side economizer	17
W27			27
W32	Cooling Tower	Chiller / District heating system	32
W40			40
W45	Cooling Tower	District heating system	45
W+			> 45

ASHRAE has defined a wide range of liquid side supply temperatures, depending on choices of facility equipment and operation efficiency target. At the current stage, we observe quite diverse coolant supply temperature choices (in deployment or desire to have) across hyper-scale users and they can be grouped as follows:

Group 1: High coolant temperature, 40 ~ 45 °C secondary supply

- Maximize efficiency and environmental goals
- Lower cooling capability

Such facility could be enabled with liquid-to-air hybrid solutions such as rear doors and sidecar heat exchangers. For facility water supply, it could also be enabled around the world with very low water usage and no chiller.

Group 2: Medium coolant temperature, 30 ~ 37 °C secondary supply [5]

- Balance cooling capability with efficiency & environmental goals

Such facility would require facility-level coolant supply, with light utilization of chillers to trim coolant temperature during certain seasons of the year, or year round for some data center locations.

Group 3: Low coolant temperature, 15 ~ 25 °C secondary supply

- Maximize cooling capability

Such facility would require facility-level coolant supply, heavily relying on chillers and evaporative cooling to maximize cooling performance. Air supply in such facility also needs to stay lower to avoid condensation.

1.5. Reliability Expectation

The consequence of liquid cooling solution failures could be more severe or even catastrophic comparing to air cooling solution failures. Some batches of liquid cooling racks may sit in the data center longer than the designed lifetime (typically 3~4 years per generation). Therefore, users are setting more strict reliability expectations.

For cold plates and coolant loops, an annual failure rate of 0.3% or lower is desired.

For the entire Level 10 system assembly (chassis/server level), MTBF is targeted at more than 300K hours at a coolant supply of 40 °C, and requires minimum 6 year product life performed in Reliability Demonstration Test (RDT).

Leakage detection combined with pumping unit (CDU or RPU) response mechanism is a typical approach to monitor and stop coolant leakage for deployment at scale. However, the design and operation trade offs are such a debating topic that it's not part of hard requirements for liquid cooled platforms yet.

1.6. Serviceability & Maintenance Expectation

It's impractical to expect liquid-cooled systems being equivalently serviceable as air-cooled systems, yet efforts can be done to reduce the complexity in this procedure. By design, users shall have the flexibility to swap OAMs in the form of:

- Replace entire board (UBB) + all OAMs + cold plate loops (PCL) as an FRU, where connector and QC disengagement/engagement would happen
- Replace multiple OAMs + their PCL as an FRU, where OAM Mezz connector and QC disengagement/engagement would happen
- Replace single OAM as an FRU, where cold plate uninstallation/installation and TIM removal/re-apply would happen

At a Level 10 System Assembly, they shall not require regular maintenance over the lifetime of the product (>4years). I.e. shall not expect to shut down the system in order to perform maintenance works (such as TIM replenishment, coolant treatment/replacement, etc.) unless if failure happens.

2. OAM Thermal Challenges

2.1. TDP Growth

Design power of OAM products is trending up, far beyond 450W in foreseeable future. OAM Spec 1.1 recommended implementation of liquid cooling as module power exceeds 450W, up to 700W which is max module power specified. In the foreseeable future, however, OAM product power may easily cross the 700W line as 54V input being adopted. OAM Spec 2.0, which is still in progress, is targeting at OAM TDP up to 1000W.
For a 1000W OAM product, approximately 70%~80% of power could come from the main chip, which can be translated to an average power density of 60~80W/cm^2. Such power density is well below the theoretical limit of liquid cooling technologies, however in reality the thermal management could be very challenging due to various thermal/mechanical restrictions.

2.2. Extend Air Cooling

As for air cooling, the limit of 450W was a nominal value derived from a set of boundary conditions and assumptions, as described in OAM Spec 1.1. On the other hand, a few most recent applications are asking for extending air cooling limit to approximately 600W. Within current technologies, this may be achieved by:

- Select more 'advanced' heatsink type, such as 3D-VC and EVAC
- Taller and heavier heatsink
- Compensate with chassis space and airflow delivery
- Package improvements, such as die size increase or lidless

The extra OAM cooling limit that can be claimed depends on a variety of factors associated with the chip and the system designs. It has been demonstrated that 500W, 600W, or even higher power can be supported, by changing one or multiple factors together.

2.3. Power Modules

With such design power, the chip package is undoubtedly the most critical component to address, but power modules would also require equal attention. Each OAM product may have its unique scheme and placement of power module(s), where a few universal challenges have been observed:

- High power density due to TDP increase and PCB space limitation
- High junction-to-case thermal resistance
- Thick TIM layers due to complicated/uneven surface(s)
- Long heat conduction path
- Challenge of contact pressure management

It's not uncommon to have OAM's power modules hotter than the main package in operation. With proper characterization in the design stage, those challenges are typically solvable by customized cold plate designs and TIM combinations.

2.4. Interface with Cooling Solution

The thermal/mechanical interface between OAM components and the cooling solution contributes to a significant portion of the temperature gap, especially if not managed properly. Depending on the characteristics of each OAM product, the factors that could impact include, but are not limited to:

- Package Surface flatness
- Mounting pressure limit
- TIM choice and curing temperature
- Stiffener material
- Keep out zone
- Etc.

2.5. Thermal resistance Elements

For an OAM product in a liquid cooling environment, its overall thermal resistance from junction to coolant consists of a variety of elements, including and not limited to:

- Cold plate
- TIM2
- Lid & TIM1
- Lateral heat crosstalk
- Die Stack

In addition to cold plate design, TIM and preheat management, which can be optimized after OAM products are released, a large portion of the temperature built up could originate from the surface of and within the package. In many scenarios, the cooling limit of an OAM product would depend more on its package design and quality control, compared to its cooling solution. We speculate that collaborative efforts from OAM suppliers, solution providers, system integrators and infrastructure owners are needed to enable max utilization of the product.

3. OAM Cold plate

3.1. Flow rate & Pressure drop

The cold plate design would be based on the coolant flow rate and pressure drop requirements, and the thermal enhancement structure could be a straight fin, pin fin, and offset fin. The fin dimension, i.e., thickness, pitch, diameter, and height, will obviously impact flow resistance inside the cold plate. Besides, the flow rate can affect the flow distribution, which means the pressure distribution would be better with a higher flow rate, but it will cause more flow resistance. Therefore, a flow distribution unit may be needed for a cold plate to get a better flow pattern, and it will provide a lower pressure drop and higher fin efficiency.

3.2. Cooling Performance

The cooling performance will be directly related to the thermal enhancement structure and the flow distribution inside the cold plate. The designer can use some software to simulate the cold plate performance and determine what kind of flow characterization, i.e., laminar or turbulent, will get the direction to improve the performance. Moreover, using the temperature profile of fin and coolant from simulation can determine the fin efficiency and prevent fully developed flow in the microchannel. So far, increasing the heat transfer area should be the easiest way to have a lower thermal junction temperature; however, the pressure drop still needs to be considered.

3.3. Cold plate material

The cold plate material should be compatible with the whole coolant loop network, especially wetted materials. Corrosion damage is the main phenomenon when using unsuitable components, and it relates to the coolant used. Considering the thermal performance, copper is the most used material in the cold plate, and what kinds of surface treatment will relate to the manufacturing process. Moreover, the whole coolant network material may include copper, aluminum, stainless steel, PPS GF-40, and EPDM. All material choices should be based on the reliability tests.

3.4. Coolant Choice

DI Water, aqueous ethylene glycol solution, and aqueous propylene glycol solution are the most popular coolants used in the cold plate loop. Generally, the coolant properties directly affect thermal performance and the CDU's (or RPU) pump selection. And the volume percent of aqueous ethylene/propylene glycol solution should refer to environmental conditions. For lower volume percent, the aqueous solution can get better specific heat. Also, the kinematic viscosity would be lower as well, which can reduce the pump loading. For high power computing, the coolant's specific heat significantly affects the performance.

Although the coolant supplier has provided a compatibility material list for reference, the corrosion performance will depend on operating temperature, the volume percent of aqueous ethylene/propylene glycol solution, and even the micro channel's flow velocity. The reliability test should be based on the conditions noted above, and the over-strict test may not be necessary.

3.5. Interface Recommendation

Contact resistance between the heat source and cold plate will significantly affect the core temperature, especially in high heat-flux chip regions. Besides, it will hold more percentage in the thermal resistance network. Here are some parameters to ensure better contact conditions. Some of them may be difficult to meet due to certain dependencies but provided for reference.

- Cold plate surface flatness: < 0.1mm
- Package surface warpage: < 0.2mm
- TIM effective thermal conductivity: > 5W/mK
- Cold plate mounting pressure: > 40psi
- Package surface temperature limit: > 70°C

3.6. Structure Strength

Cold plate stiffness needs to satisfy different platform requirements. The designer can use simulation to find a way to improve the stiffness, e.g., by increasing the copper base thickness, adding sheet metal, or some fixture on the cold plate.

4. PCL in Chassis

4.1. PCL Design Considerations

At the passive cold plate loop (PCL) level, although the thermal performance & pressure drop depends heavily on the internal structure of the cold plate, the topology and assembly method will also have a significant impact on the performance and other user experience.

4.1.1. Topology

For a system with 8x OAM cards, the traditional layout of a 4x2 cold plate loop would reach a good balance between thermal performance and design complexity. On the other hand, an 8x1 parallel loop would bring better OAM temperature uniformity, and reduce flow impedance significantly. With such layout, however, flow distribution management would be a critical topic.

4.1.2. Quick connect (QC) Selection

QC size is one of the important portion of PCL design, lager QCs would benefit flow impedance, however higher cost and space occupation would be concerning, especially as it may limit other aspects of system design.

QC connection type can be either manual-mated or blind-mated. Manually-mate QCs are less convenient but more mature and reliable, while blind-mate QCs are more convenient at the cost of structure complexity (tolerance, mating force, etc.). Those trade-offs need to be evaluated for different use cases.

4.1.3. Tube/Hose type and routing

Tube/Hose selection and layout is another important portion of PCL design. The hard tube is better on reliability and leakage proof, but the hose is better on tolerance requirement and flexibility, user could make the choice base on the design and serviceability requirement. The thickness of the tube/pipe also needs to be considered. A thicker tube/pipe could resist higher pressure, but flexibility will become worse than a thinner one.

Although the OAMs are cooled by liquid, the remaining components in the system (such as VR, QSFPs, switches, and other components) may still be air-cooled. To avoid impact on their cooling, and also give flexibility to system design, it's recommended to minimize the space occupation of PCL parts (hose, tube, clamped connection, manifold, etc.).

4.2. Chassis Design Considerations

For the liquid cooling chassis design, as the same reason for air-cooled component cooling, so reserve enough space for airflow and fan space is necessary. On the other hand, high-power OAMs require higher coolantflow rate, the larger or more quick connectors will also occupy more space, those are the considerations for chassis outside dimension design.

Inside the chassis, as there are 8x cold plate+OAM assemblies, it's expected to use hard pipe/hose for cold plates connection. To reduce the shock and vibration impact, pipe clamp for hose holding is suggested. The PCL design for 8x OAMs usually use chassis manifold to direct coolant into each flow path (each with single or 2 cold plates in serial), therefore support feature with anti-vibration design for chassis manifold is also recommended, in order to support the weight and reduce impact by shock and vibration.

4.3. 4x2 PCL Practice

To practice PCL on an UBB-type system with high power OAMs, we built a prototype system with thermal test vehicles (TTV) up to 1000W each, to validate its cooling performance, flow distribution, and serviceability. In this practice, the passive cold plate loop has 2 pcs chassis manifold to connect the cold plates as 4 (in parallel) by 2 (series) layout, and also implemented 4x 6mm QCs (2 for coolant inlet and 2 for coolant outlet) to reduce flow impedance with smaller width occupation, so as to fit 4x 80mm fans for airflow delivery need.

The prototype system is designed to be compatible with UBB1.0, using TTV tray as place holder, in an 3OU chassis. Inside the chassis, a support bracket with sponge is in place to hold the manifold and absorb the impact of shock and vibration. Pipe clamp were also used in front of the QC, to hold the hose and keep it steady. At front side, a labor-saving ejector is implemented to help the user overcome the friction force between the chassis and rack, and also the reaction force from blind-mate QC.

Although still room to improve/optimize, this design demonstrated sufficient performance to support 700~1000W OAM products (with dependencies on boundary conditions and package characteristics). It serves as baseline example of liquid cooling solution design for OAI system and other products alike. Result of this practice are provided in the case study section.

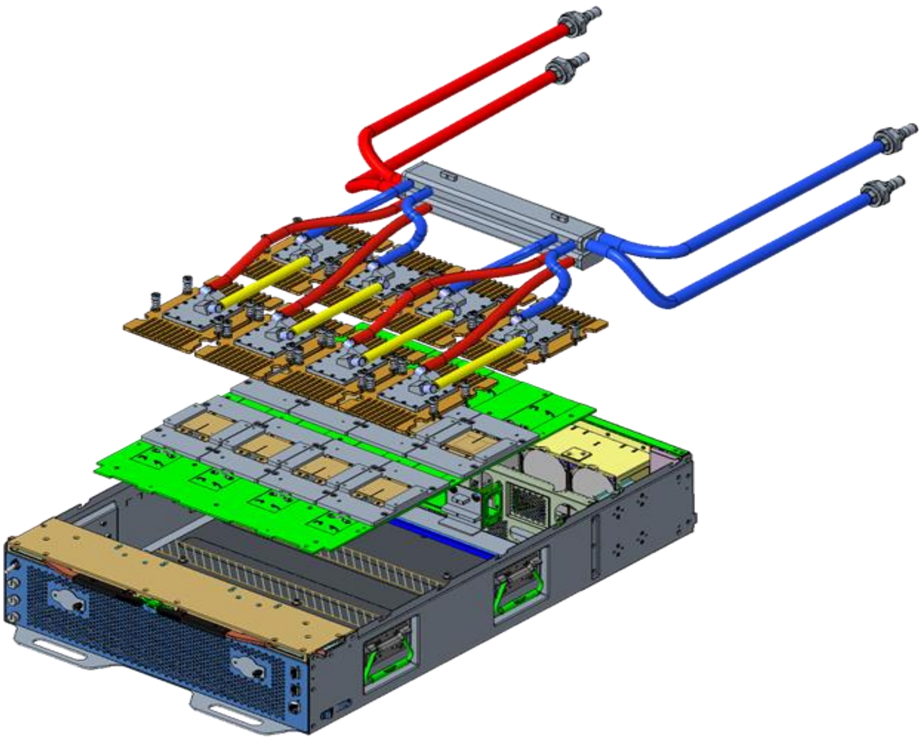


Figure 1. Prototype Chassis Design with TTV tray representing UBB1.0 and the 4x2 cold plate loop, using 2x pairs of 6mm Quick Connects for coolant supply/return.

5. Validation and Reliability

5.1. Thermal Performance Validation

For system level thermal validation, hereunder are the suggested validation matrix parameter and data to be recorded:

- Coolant inlet temperature
- Coolant outlet temperature
- Total coolant flowrate
- Coolant flow impedance
- T_{case} of each OAM
- $T_{junction}$ of each OAM

Thermal resistance is often used to define the performance of heatsink, it is used to define the performance of cold plate.

- $R_{th,cold\ plate}$: $(T_{case} - T_{in})/W_{ASIC}$, thermal resistance of cold plate
- Q_{ASIC} : ASIC power which is cooled by liquid
- T_{case}: temperature on ASIC case
- T_i: coolant temperature at cold plate inlet

To record T_{case} data, the temperature sensor should be placed on top of the TTV. With thermal resistance of TIM in mind, the temperature sensor can instead be implemented on the bottom side of the cold plate for real chip application.

For Level 10 system assembly to pass the worst condition, 8x OAMs need to meet the thermal requirement under the lowest coolant flowrate and the highest coolant inlet temperature (client dependent). Refer to typical test items from air cooling thermal validation for other component item tests.

5.2. Shock & Vibration validation

If the fluid is retained in the conduit during the shipping of the Level 10 system assembly, the corresponding system test(s) should include the fluid in the simulation to mimic the whole transportation process. It is recommended that compressed air of the system is examined before tests to capture any potential leak from integrating the sub-system. The parameters, such as temperature and pressure, are recommended to be chosen based on client’s criteria and monitored throughout the shock and vibration tests. Inspections in the following areas are recommended after the sub-system assembly:

- Mechanical fitment checks on SKU-dependent QDs
- Tube/pipe & connection locations

- Any other valves or connection joints

If the fluid is not retained in the conduit during the shipping of the Level 10 system assembly, the corresponding system test(s) should not include the fluid in the simulation to mimic the whole transportation process. The client should choose the parameters, such as temperature and pressure, based on the criteria and should be monitored throughout the shock and vibration tests. It is recommended that the system is also tested with compressed air or nitrogen pressure decay method, or hydrogen or helium sensor monitoring method, depending on vender's suggestion and end-user requirement. The leakage from the following sub-system integration areas can therefore be captured:

- Mechanical fitment checks on SKU-dependent QDs
- Tube/pipe & connection locations
- Any other valves or connectors

For the operational validation, the corresponding system tests should include the fluid in the simulation that mimics the environment throughout the operation. The parameters, such as temperature and pressure, are recommended to be chosen based on client's criteria and monitored throughout the shock and vibration tests. It is recommended that the tested system include a Level 10.5 system assembly test fixture, CDU and manifold.

- Mechanical fitment checks on SKU-dependent QDs
- Tube/pipe & connection locations
- Any other valves or connectors

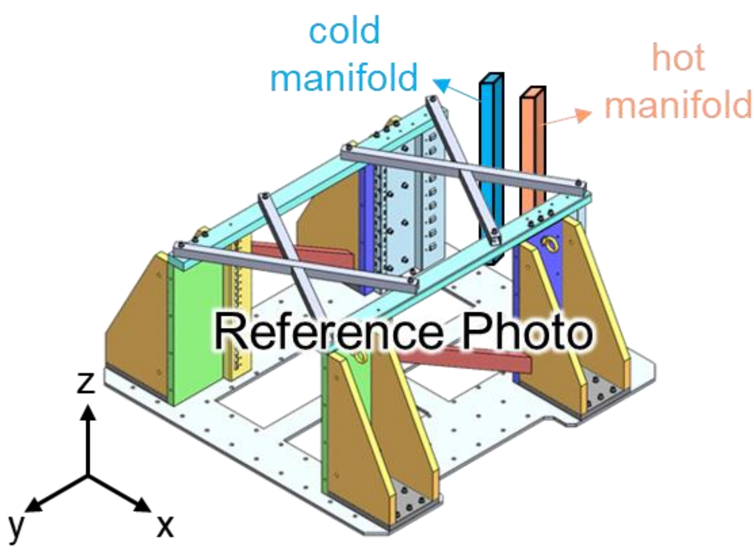


Figure 2. Mechanical jig for shock and vibration testing

Any leaks after the S&V testing should be identified, traced to the root cause and rectified. In some cases, engineering change order may be required. For other inspection items, refer to the typical test items of devices cooled by air.

5.3. Reliability Testing

For system level reliability validation, the targets are liquid cooling commodities. These may include cold plate, wetted material, tubing, fitting and coolant. Thermal Performance and Leakage risks are the 2 key points to see during the validation. The following three related tests in addition to the S&V test are recommended:

5.3.1. Thermal Cycling Life Test

- Setup a long-term heat load testing or an Accelerated life testing
- Considering the total system thermal cap and then setting the Temperature and Humidity range conditions.
- Target: Joints and Pipelines
- Test criteria:
  - I. Visual inspection for cracking detection: No deformation, No crack
  - II. System can work normally. (No hang up, no shut down)
  - III. No component or module failure or damage
  - IV. No thermal performance degradation
- Reference test process:
  - I. Temperature range: 10°C-40°C
  - II. Temperature holding time: 24hrs for each temperature
  - III. Temperature change time: 2hrs
  - IV. Cycles: 7 cycles at least
  - V. Total test time: 15 days at least include check point with every 5 days (back to room temp)
- Reference specifications:
  - I. MIL STD 810G – Test Method 501.5
  - II. JEDEC JESD22 – A101/A102
  - III. IEC 62368-1

5.3.2. Pressure/Leak Test

- The components are considered comply if they pass at the completion of the test and if they do not rupture, burst, or leak. It is recommended that the tested system to be investigated with compressed air or nitrogen pressure decay method, or hydrogen or helium sensor monitoring method, etc. As to which method should be utilized and which standard should be followed, that’s a decision to be made between the manufacturer and the end-users.
- Target: Joints, connectors and pipelines
- Test criteria:
  - I. Leak rate should at least meet water-tight level requirement
  - II. Visual inspection: No rupture, No burst, No leak
  - III. No component or module failure or damage
  - IV. Validate thermal performance

- Reference test process
    I. Suggest taking 1.5 times of design pressure for testing
    II. Temperature range: Normal ambient temperature
    III. Test time per cycle: Depends on test equipment capability (e.g. pressure sensor resolution)
    IV. Total test time: at least 7 Days and includes check points.
- Reference specification:
    I. IEC 62638-1
    II. ASME B31.n

5.3.3. Unit Freeze/Thaw Test (Storage test)

- Freezing-point validation for water-based solutions
- Target: Joints between Cold Plates, connector, pipeline, MBs
- Test criteria:
    I. Visual inspection for cracking detection: No deformation, No crack
    II. System works normally. (No hang up, no shut down)
    III. No component or module failure or damage
    IV. Validate thermal performance
    V. Re-perform leakage test
- Reference test process
    I. Temperature range: -10°C-70°C (take PG25 as example, can be modified based on coolant property)
    II. Temperature holding time: 24hrs for each temperature
    III. Temperature change time: 2hrs
    IV. Cycles: 7 cycles at least
    V. Total test time: at least 15Days and includes check points with every 5 days (back to room temp)
- Reference specification:
    I. MIL STD 810G – Test Method 501.5
    II. JEDEC JESD22 – A104/A106/A119

6. Case Studies

6.1. Parallel (8x1) PCL Analysis [Boyd]

In addition to cold plate development, careful design works are required to establish the flow network. Traditionally 4 parallel x 2 serial layout has been adopted by a variety of platforms with 8x accelerator modules, bringing the benefits of:

- Higher flow rate per cold plate
- Lower manifold and hose routing complexity
- Manageable preheat towards downstream

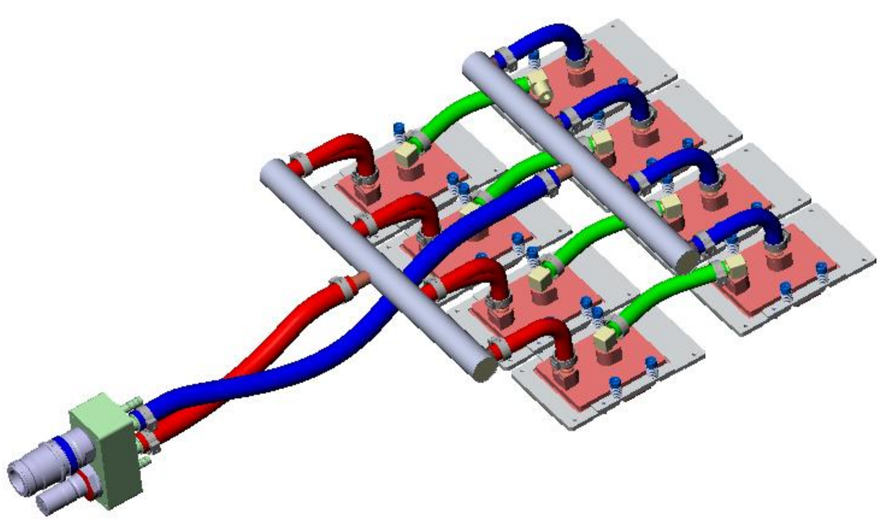
As accelerators' power continue growing, however, the need of 8 x 1 network to further stretch cooling capability shows up. It would remove the preheat from upstream module, extending single phase liquid cooling capability by a few hundreds of watts theoretically. Challenges would be managing the flow distribution and cold plate operation at low flow rates. The flow distribution management would have impact on OAM temperature margin, system operation point and cooling power consumption.

Using Boyd's Smart CFD which has an integrated flow network tool built in, we were able to develop a fully parallel 8X1 cooling loop within the same space claim as a 4x2 loop. Calibrated with test data from an existing 4x2 loop prototype, we were able to validate and refine the network model around pressure drop and use this for the 8x1 loop. Results are presented in the following sections.

In this study, all cold plates are Meso-Channel with internal fin features between 0.1 mm – 1.0 mm

Two layout options are considered:

- 4 x 2 (4 parallel x 2 series)



- 8 x 1 (8 parallel x 1 series)

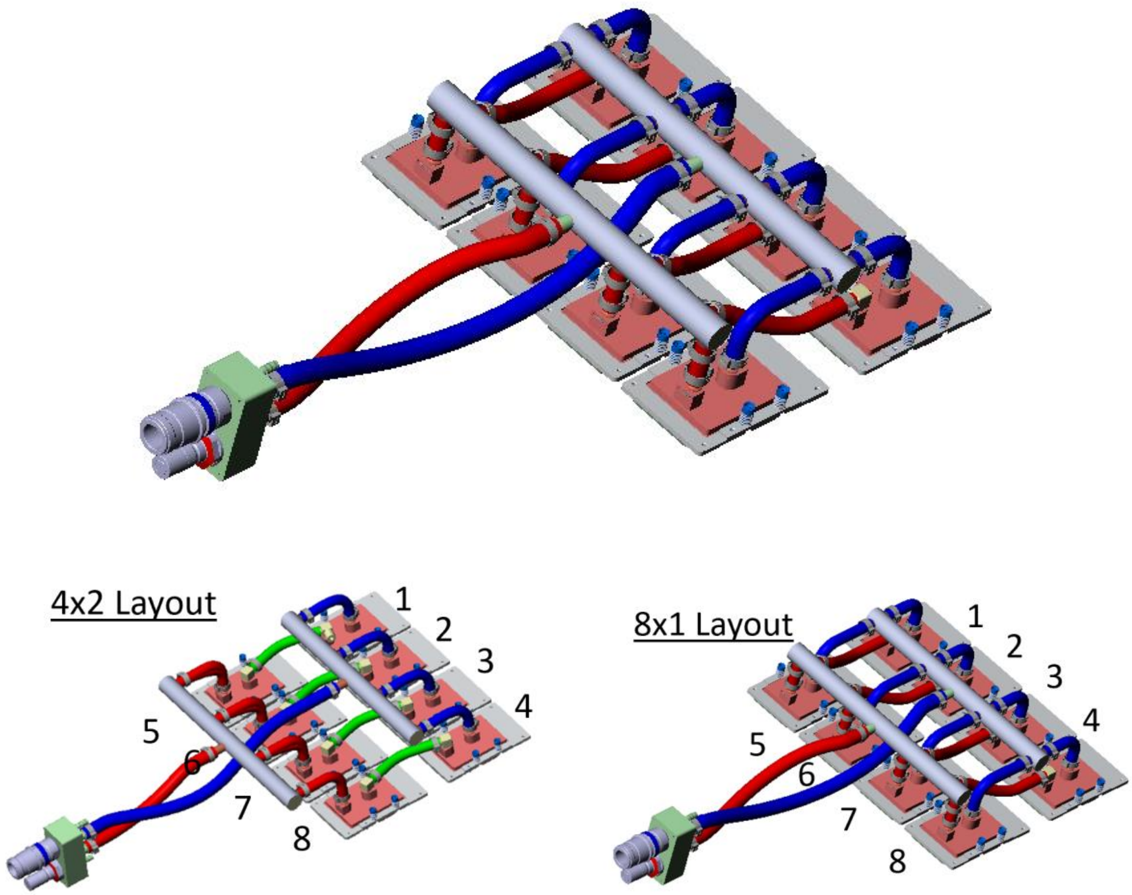


Figure 3. Baseline example PCLs for 8x1 and 4x2 flow networks

The pressure drop of each loop was determined using analytical flow network software that has been correlated with test data.   Below is the “refined” PQ curve of each loop.

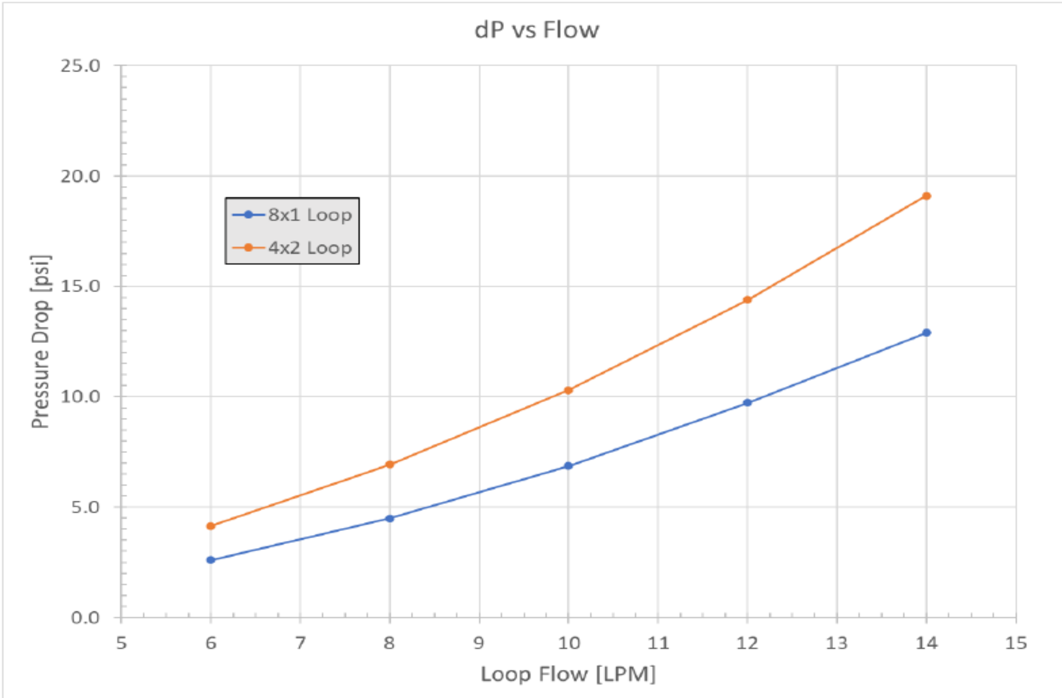


Figure 4. Pressure drop at flow rate of 6~14 LPM per system

At a fixed flow rate of 10 LPM, with 25% PGW, the pressure drops are:

- 4x2 Layout - 10 psi
- 8x1 Layout - 7 psi

At a fixed pressure drop of 10 psi, with 25% PGW, the resulting flow rates are:

- 4x2 Layout - 10.0 L/min
- 8x1 Layout - 12.5 L/min

Additionally, CFD was used to generate RQ curves for each cold plate and this is incorporated into the flow network model. When the entire System IT loop is characterized with a flow network, we can see how the different designs can directly affect the system operating points.

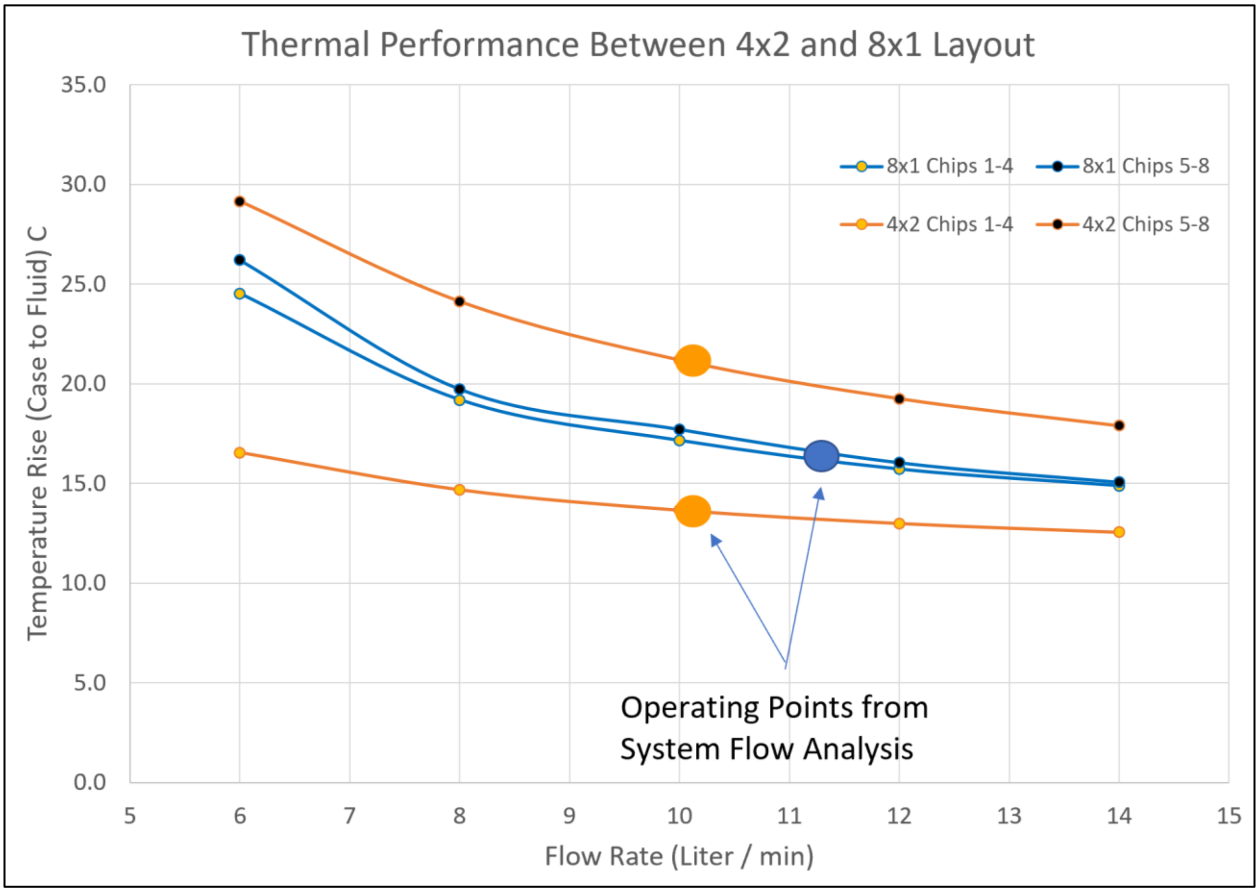


Figure 5. Case temperature increase comparing to coolant supply temperature, at flow rate 6~14 LPM

The 8x1 layout has two advantages over the 4x2 in that:

1) All chips within the loop are within 1°C of each other with improves the reliability of the electronics on which this loop cools – as compared to 7°C for the 4x2 configuration

2) The lower pressure drop allows for a multitude of options:
    a. As depicted above, a shift in the operating point can be realized
    b. The pump could run at a lower speed increasing its reliability
    c. Other areas of the system could tap into this overhead to drive performance, serviceability, or cost reduction

6.2. 1kW OAM TTV design [CoolerMaster]

It's recommended that all cooling modules being tested with a thermal test vehicle (TTV), to validate the thermal performance and reliability. An OAM TTV with 1kW stress capability is introduced in this case study. Such TTV can be used to validate air cooling or liquid cooling solutions. It can also be used to quantify the airflow/liquid flow rate through the cooling module.

6.2.1. Design Guidelines

Heating Element - To ensure the compatibility of TTV with chassis, the TTV's height shall be close to the height of real OAM product as much as possible. Foil heater is recommended to serve as primary heat source representing the main package of OAM product. It can also be designed to simulate various power maps. In comparison, cartridge heaters are typically too thick and unable to represent the power map well.

Material – Copper is the most feasible material as spreader directly above and below the foil heater. In order to prevent heat spreading through the base, Bakelite can serve as a good insulation surrounding the heated area. It can also provide enough mechanical strength for cooling module mounting.

Inspection – test result on high power TTVs could vary dramatically due to small differences across units. Following items need extra inspection to ensure the consistency:

- Surface Flatness
- Electrical Resistance
- Wire length/material/connector
- Circuit diagram (multi-TTV board layout)

6.2.2. Baseline Design

Following picture shows a baseline design for OAM TTV using thin foil heater (thickness < 0.5mm). The foil heater is sandwiched by copper top & bottom layers, with thermal grease on both contact surfaces. Noted that the foil heater wire is relatively fragile and would break easily if steady or transient temperature becomes too high.

This baseline TTV is capable of generating 1kW through 30mm x 30mm thin foil heater, up to 111 W/cm^2. Case temperature is monitored at the top center of the copper lid, which is recommended to maintain below 75C through test. For each OAM product, the dimensions of the thin foil heater, copper lid, and Bakelite base shall be adjusted to represent the product's characteristics.

Such design does not consider complicated heat map inside package or heat dissipation of VRMs yet. Our future TTV design(s) will aim at simulating those with heat map flexibilities.

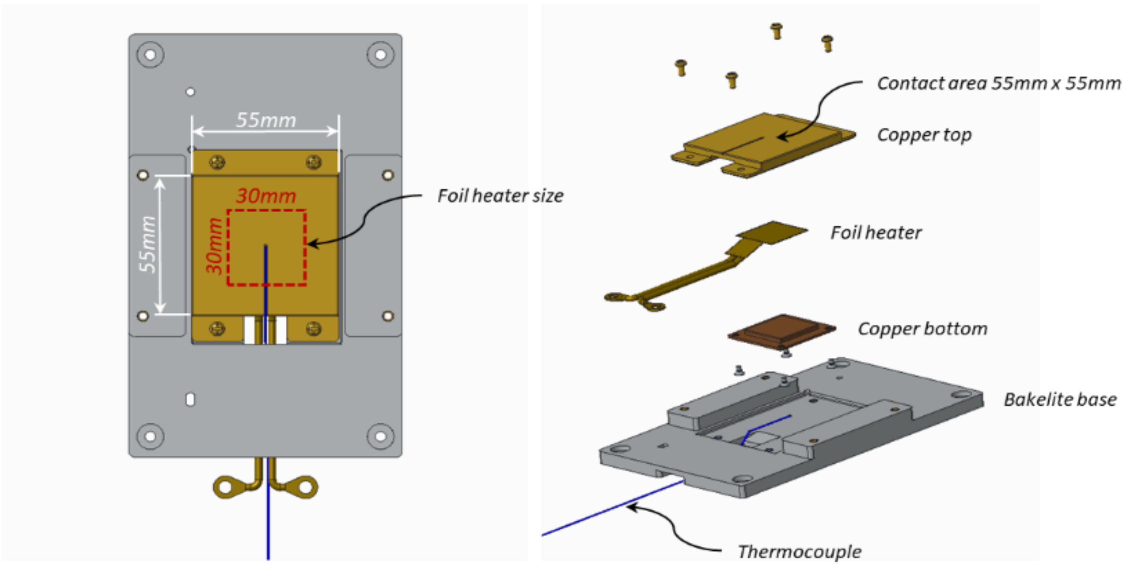


Figure 6. scheme of OAM thermal test vehicle

6.2.3. Test Result

Validation of the TTV and its cold plate solution was performed with following parameters:

- Coolant – **PG25 / DI Water**
- Flow Rate – 2~5 LPM
- Surface Flatness – below 0.05mm
- TIM2 – phase change material (cured before collecting data)
- Stress Power – 1kW

Result shows that **under ideal conditions** using PG25 as coolant, the cold plate is potentially capable of delivering thermal resistance below 0.02 C/W for an OAM package of such size and heat flux. This can be translated to: in an UBB-style system, the case temperature of a 1kW OAM product is possible to be maintained below 60C, with properly optimized solutions, at coolant supply of Group 2 category (30~37 degC).

Keep in mind that multiple factors can influence the cooling performance of real OAM products to deviate from TTV test results positively/negatively, including but not limited to:

- Coolant type
- Surface flatness/warpage
- Die size
- Power Distribution
- Package Type
- TIM selection
- Mounting Pressure
- Coolant Selection

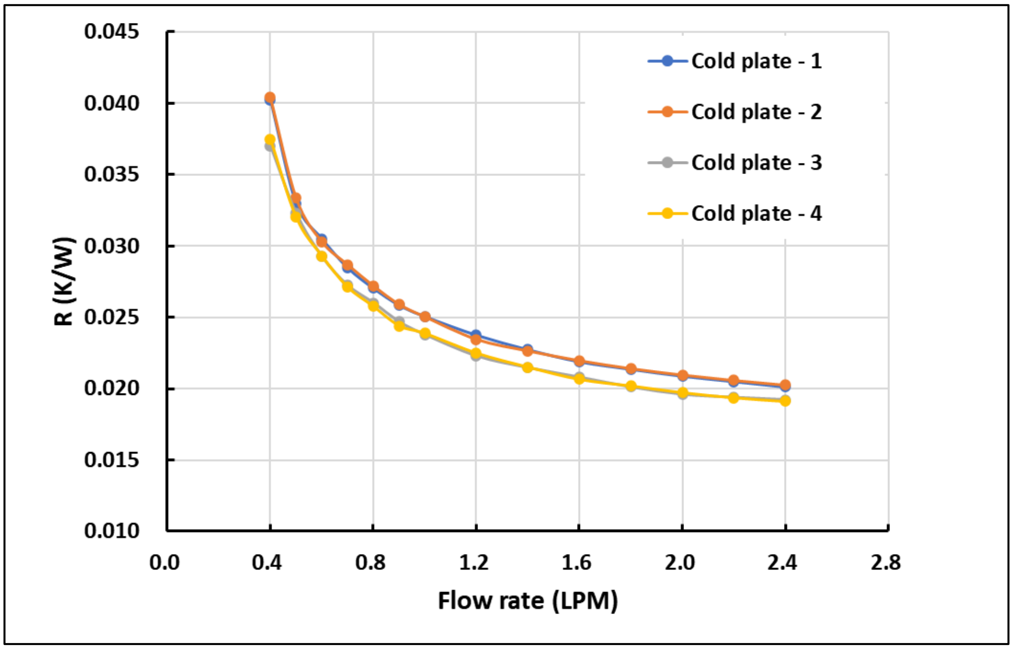


Figure 7. single cold plate validation with PG25, R is case-to-inlet thermal resistance

There's further performance increase if DI water is applied to the cold plate. Test result shows in the same setup, a cold plate can deliver as low as 0.015~0.016 C/W at high flow rates.

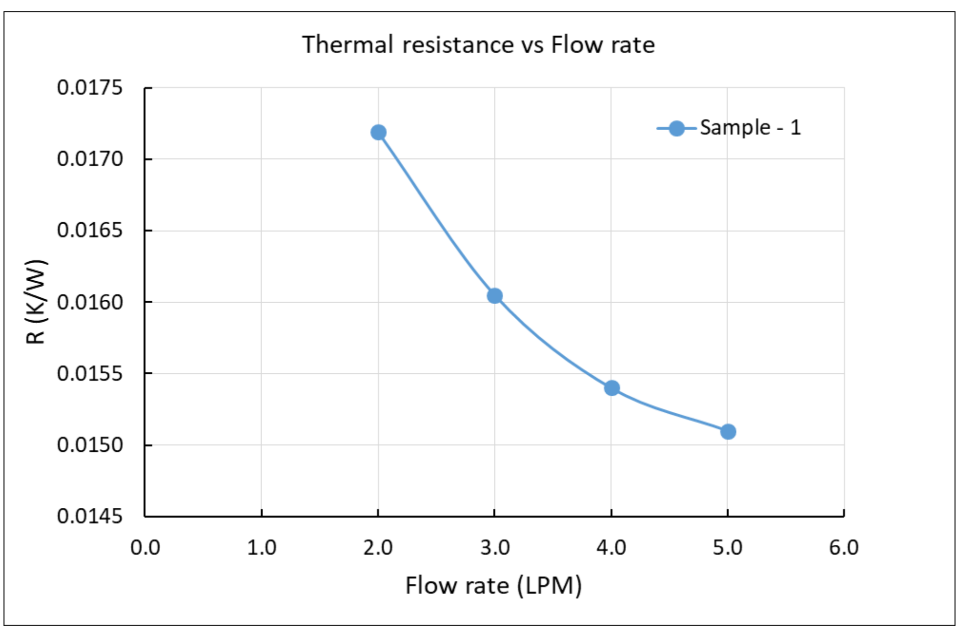


Figure 8. single cold plate validation with DI Water, R is case-to-inlet thermal resistance

6.3. – System Level Validation with 4x2 PCL [Wiwynn]

Narrative: Preliminary system-level test result is shared for design reference. According to the test result, although coolant flowrate distribution is not ideal, the current PCL design could still support TTV power at the 1000W level. Hereunder list the system layout, thermal test setup, and result analysis.

6.3.1. System Layout:

For the 3U UBB-like TTV chassis, 2 pairs of 5mm QDs were used for coolant supply/return, giving enough width to accommodate 4x 80mm fans which would deliver enough airflow for air-cooled components in a real system. The QDs are manually mated in this prototype, while the chassis design kept flexibility to accommodate blind-mated QDs as well. Two internal manifolds were positioned on rear side for coolant distribution, the flow impedance of which is yet to be further optimized.

In this prototype effort, the cold plates were connected in a 4 (parallel) x 2 (serial) network, and only OAM TTVs were stressed through the tests.

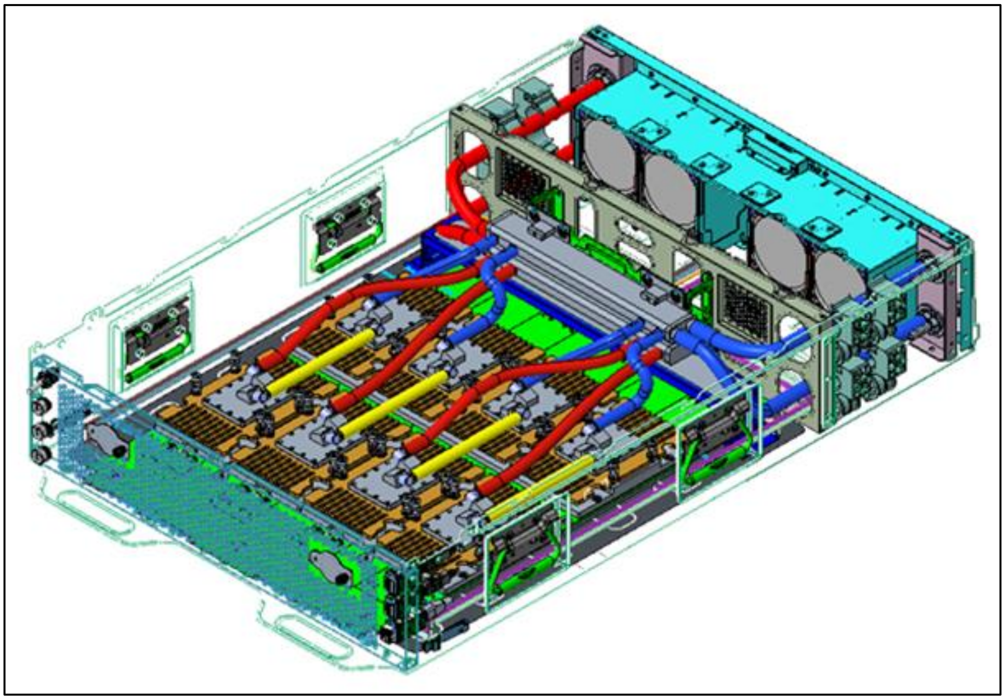


Figure 9. layout of TTV-based UBB-like dummy chassis, for PCL validation

6.3.2. Test set up for PCL thermal validation:

- OAM Cold plates 1~4 (rear row based on air direction) are upstream of the flow path and 5~8 (front row based on air direction) are downstream of the flow path
- Coolant supply (DI water) flowrate from 6~14LPM, at inlet temp = 30°C
- TTV power from DC power supply directly

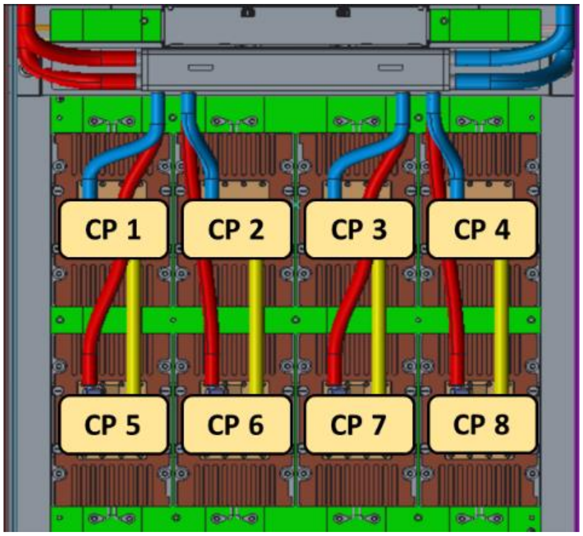


Figure 10. Cold plate numbering; noted that it does not represent any OAM interconnection topology

- Due to limited DC power supply capability in the lab, multiple test rounds stressing upstream/downstream/left side/right side cold plates separately were adopted to examine temperature uniformity and the impact of preheat. Leaving the flow field undisturbed, the resulting thermal resistance values and case temperatures were combined into results in the next section.

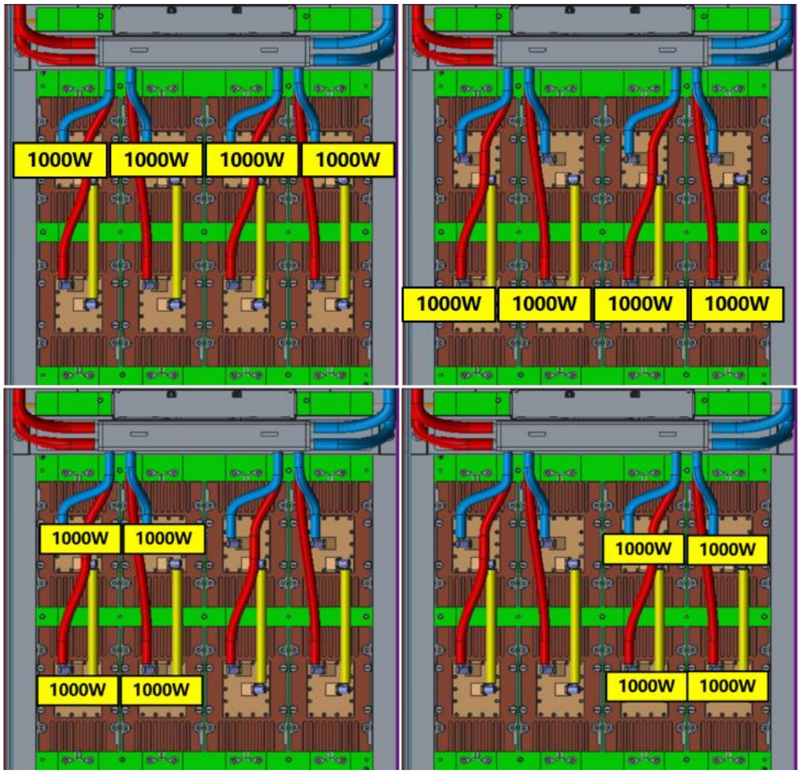


Figure 11. 4-step test to validate the PCL with limited DC power supplies; coolant supply and flow network remain the same across those tests

6.3.3. PCL Test result:

The case-to-inlet thermal resistance of a single cold plate arrives at 0.014~0.019 C/W across the designated flow rate range (1.5~3.5 LPM per cold plate). Those values were derived assuming uniform flow distribution and preheat based on TTV power/Heat Capacity.

Such performance aligns well with prior validation at a single cold plate level using DI water. Noted that up to 15% performance penalty may apply if using PG25 as coolant.

The variation of thermal resistances arrives at 0.002 C/W within the test sample size. This aligns with typical expectations where flatness of contact surfaces is properly controlled. It is observed that uncertainty control of TTV heating elements may have contributed more to the variation comparing to cold plate and TIM control. This is not surprising as empirically a few degrees variation across units of same product with same cooling solutions and boundary conditions were quite common in prior platforms.

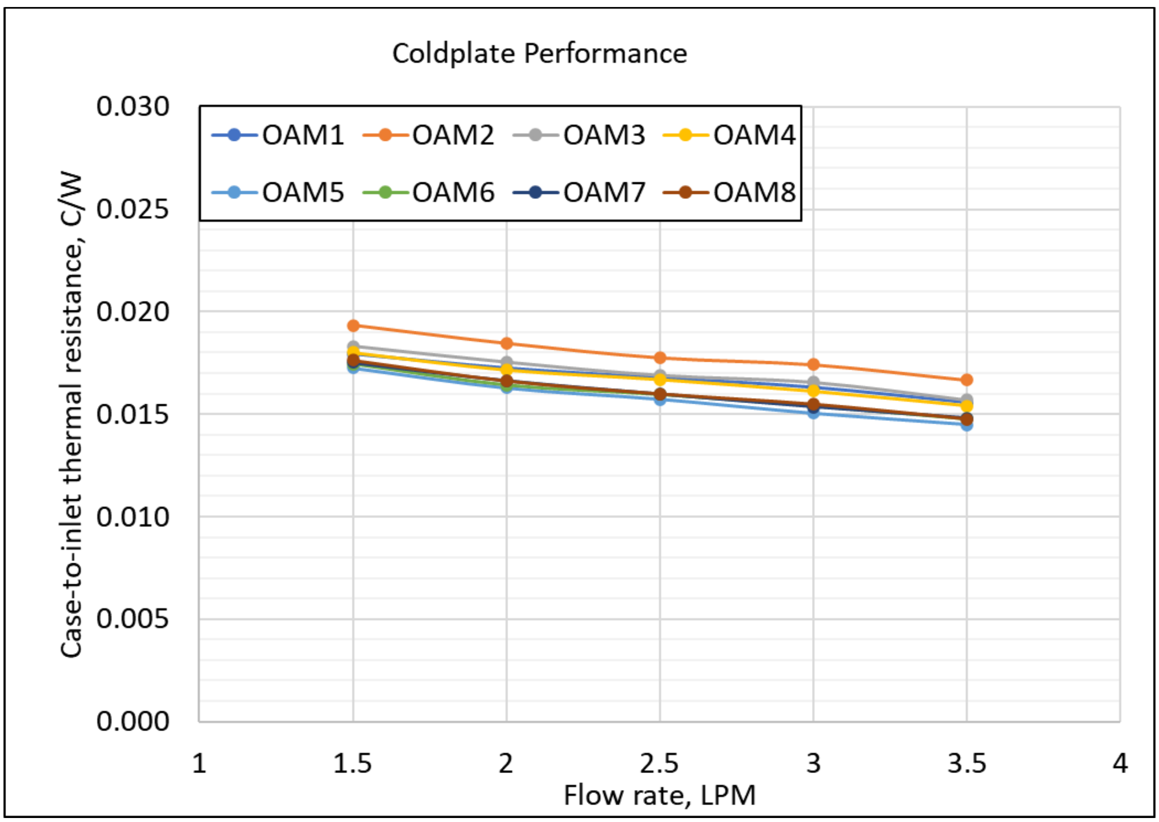


Figure 12. Calculated thermal resistance values for all cold plates in the system

Case temperatures of OAM TTVs at various system level flow rates were characterized as well, with coolant supply at 30C. At the flow rate of 10LPM, the resulting hottest case temperature is merely 52 °C, for an assembly with 8x 1kW OAM TTVs. With ‘Group 2’ coolant supply (i.e. 30 ~ 37 °C to the system), such design is potentially able to maintain case temperature of real OAM products under 60 °C, even if operate at speculated upper limit of OAM2.0 power (up to 1kW).

The uniformity across downstream OAM TTVs (5/6/7/8) were impressively well maintained, indicating uniform flow distribution across all paths.

A limiting factor of the PCL design, however, is the total pressure drop. It is observed that the system flow pressure drop (QD-to-QD) already reaches 17 psi at 8 LPM (2 LPM per cold plate path). Roughly 50% of the pressure drop come from the cold plates and remaining 50% from hose+manifold+QDs. To maximize OAM cold plate cooling capability at reasonable efficiency, especially for higher power products, it become significantly important to reduce overall flow impedance by optimizing hose routing, manifold design, and QD selection. It also indicates the need to study all parallel PCL (for example, 8x1) as described in section 7.1.

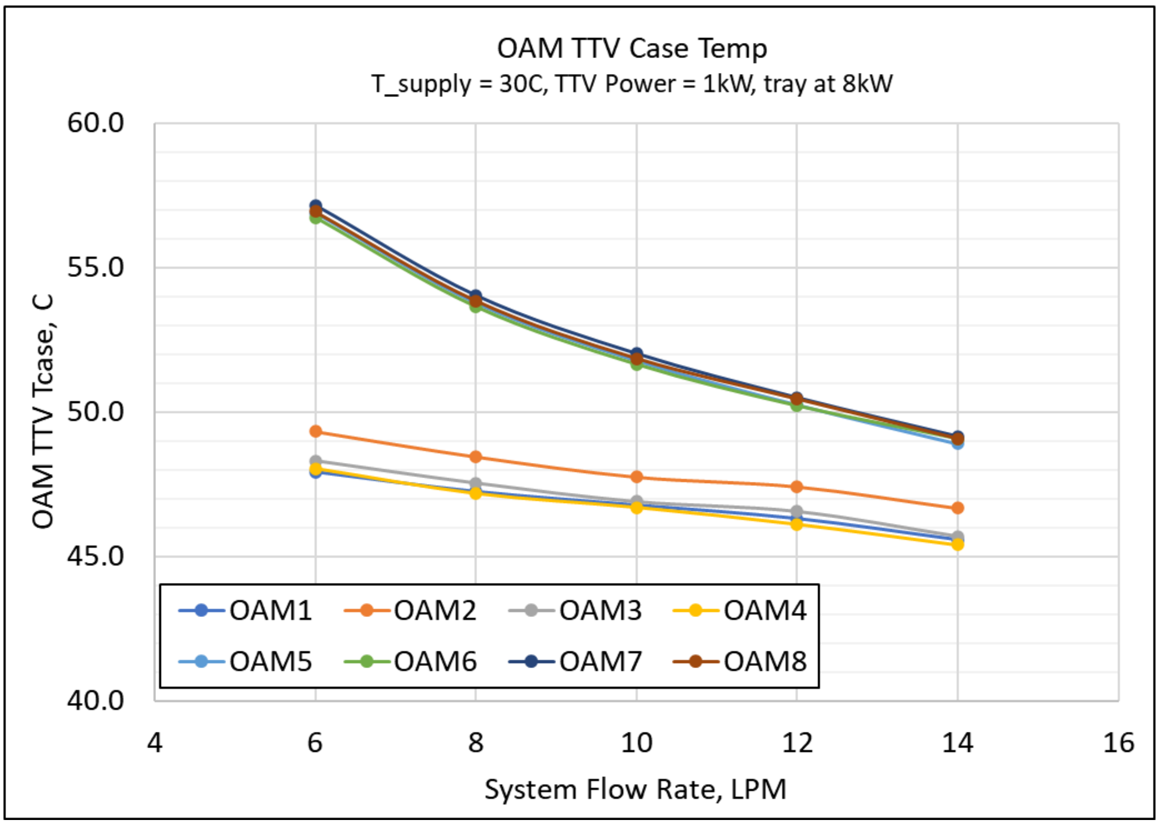


Figure 13. Case temperatures of all OAM TTVs, with 1kW load on every TTV, and 30C coolant supply to the system

6.3.4. Areas to improve

With such PCL and chassis design, it is highly recommended to have at least 2 operators to perform the PCL assembly to system. In order to enable better serviceability requiring only one operator, following are some ideas worthy looking into for future products:

- Separate the PCL into multiple FRUs with internal QCs

  As the PCL assembly includes system level QCs, chassis manifold and 8x cold plate in one piece, the FRU size is much larger than typical heatsink and difficult to install by one operator. If every single flow path (2x cold plate in series) could be designed as a FRU and add internal manual QC, serviceability would become much better. In additional, such design would require only 2x OAMs+cold plates involved if to replace any OAM card, instead of entire PCL assembly. Downside of additional internal QCs, however, is the impact on flow impedance, cost, and leakage risk.

- Middle Wall Design

  In current design, operator needs to pass QC with hose through the hole on the middle wall, although the hose has flexibility, operator might not do it smoothly because of space limitation. Design some notches on middle wall and let hose above on it may could be taken into account, but structure strength of middle wall also needs to be considered.

- Handle design on PCL

  There’s no handle on current PCL design, adding handle on it based on service requirement will have benefit on service process. Another suggestion is adding some fixing feature on PCL, and user could design specific installation fixture to improve serviceability.

7. Terminology

OCP: Open Compute Project

OAI: Open Accelerator Infrastructure

OAM: Open Accelerator Module

UBB: Universal Baseboard

PCL: Passive Cold plate Loop

QC/QD: Quick Connect / Quick Disconnect

8. References

1. OCP Accelerator Module Design Spec V1.1

   http://files.opencompute.org/oc/public.php?service=files&t=fb91baceba6c1dfaa9dc5531365b4f45

2. OCP Accelerator Module Design Spec V1.5
   https://docs.google.com/document/d/1jPwXqF6g3j6COyYMe9CD7IKNCCWl3gSj/edit#bookmark=id.1fob9te
3. OCP OAI Universal Baseboard Design Spec V1.0
   http://files.opencompute.org/oc/public.php?service=files&t=f924e6fefb20c9be651c61c0f4b1a5cc
4. ASHRAE 2021 Equipment Thermal Guidelines for Data Processing Environments
   https://www.ashrae.org/file%20library/technical%20resources/bookstore/supplemental%20files/referencecard_2021thermalguidelines.pdf
5. OCP Talk – Liquid Cooling: Drivers, timelines and a case for industrial convergence on coolant temperatures
   https://www.youtube.com/watch?v=97sYRvwhyI0

9. License

10. About Open Compute Foundation

The Open Compute Project Foundation is a 501(c)(6) organization which was founded in 2011 by Facebook, Intel, and Rackspace. Our mission is to apply the benefits of open source to hardware and rapidly increase the pace of innovation in, near and around the data center and beyond. The Open Compute Project (OCP) is a collaborative community focused on redesigning hardware technology to efficiently support the growing demands on compute infrastructure. For more information about OCP, please visit us at http://www.opencompute.org