The TEXTAROSSA Approach to Thermal Control of Future HPC Systems

William Fornaciari[1(✉)], Federico Terraneo[1], Giovanni Agosta[1], Zummo Giuseppe[2], Luca Saraceno[2], Giorgia Lancione[2], Daniele Gregori[3], and Massimo Celino[4]

[1] Politecnico di Milano, Milan, Italy
{william.fornaciari,federico.terraneo,giovanni.agosta}@polimi.it
[2] In Quattro srl, Roma, Italy
{giuseppe.zummo,luca.saraceno,giorgia.lancione}@in-quattro.com
[3] E4 Computer Engineering SpA, Scandiano, Italy
daniele.gregori@e4company.com
[4] Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (ENEA), Roma, Italy
massimo.celino@enea.it
https://heaplab.deib.polimi.it, https://www.in-quattro.com, https://www.e4company.com/, https://www.enea.it

Abstract. Thermal control is a key aspect of large-scale HPC centers, where a large number of computing elements is employed. Temperature is directly related to both reliability, as excessing heating of components leads to a shorter lifespan and increased fault probability, and power efficiency, since a large fragment of power is used in the cooling system itself. In this paper, we introduce the TEXTAROSSA approach to thermal control, which couples innovative two-phase cooling with multi-level thermal control strategies able to address thermal issues at system and node level.

Keywords: High Performance Computing · 2-phase cooling · Thermal modeling and control

1 Introduction

High Performance Computing is a strategic asset for countries and large companies alike. Such infrastructures are of key importance to support a variety of applications in domains such as oil & gas, finance, and weather forecasting. Recently, emerging domains have been gaining traction, such as bioinformatics, medicine, security and surveillance. These newer applications tend to fall in

This work is supported in part by the EuroHPC JU and the Italian Ministry for Economic Development (MiSE) under GA 956831 "TEXTAROSSA".

A. Orailoglu et al. (Eds.): SAMOS 2022, LNCS 13511, pp. 420–433, 2022.
https://doi.org/10.1007/978-3-031-15074-6_27

the classes of High Performance Data Analytics (HPDA) and High Performance Computing for Artificial Intelligence (HPC-AI). The trend in the design of such infrastructures is more and more exploiting heterogeneous hardware architectures to cope with the request of peak performance and to meet the need of achieving a "Green HPC". This paths have prompted Europe to align its research priorities in HPC along a Strategic Research Agenda (SRA[1]) resulting from wide consultations within the European Technology Platform for HPC (ETP4HPC), the PRACE initiative[2], and the PlanetHPC[3] initiative.

The need to achieve high efficiency while remaining within reasonable power and energy bounds, is extensively discussed in the SRA, focusing also on the main technology challenges posed by these objectives. Such challenging goal can only be addressed with an holistic approach that takes into account multiple factors across the HPC hardware/software stack, including the use of application-specific, extremely efficient hardware accelerators, efficient software management of resources, data and applications, and efficient cooling systems. Together, these components can provide the desired computational power while keeping under control the power consumption of the supercomputer.

1.1 Strategic Goals

TEXTAROSSA aims at contributing to the strategic goals of the EuroHPC Strategic Research and Innovation Agenda [11] and the ETP4HPC Strategic Research Agenda 4 [16], remaining aligned with other European and national initiatives in the context of HPC and computing architectures, including in particular the European Processor Initiative (EPI) and the EuroHPC Pilot projects [4].

Thermal control is a key goal of TEXTAROSSA, together with energy efficiency, performance, and ease of integration of new accelerators based on reconfigurable fabrics. Thermal control is achieved via innovative two-phase cooling technology at node and rack level, fully integrated in an optimized multi-level runtime resource management driven by power, energy, and thermal models fed by on-board sensor data. The aspects related to the development of new IPs, including mixed-precision computing, data compression, security, scheduling and power monitoring go beyond the scope of this work, as they address a wide range of different contributions at software and hardware level, including the integration of electronic design automation tools. We note here that these technologies will be key to opening new usage domains, including High Performance Data Analytics (HPDA) and High Performance Artificial Intelligence (HPC-AI) [3].

Two architecturally different, heterogeneous Integrated Development Vehicles (IDVs) will be developed: IDV-A by ATOS, X86/64 and GPUs, and IDV-E by E4, featuring ARM and FPGA. These IDVs will be used as testbed and workhorse by TEXTAROSSA's developers. The IDVs will be a single-node platform, easy to configure and reconfigure, extensible in terms of components,

[1] https://www.etp4hpc.eu/sra.html (last accessed June 2022).
[2] https://prace-ri.eu (last accessed June 2022).
[3] https://cordis.europa.eu/project/id/248749 (last accessed June 2022).

devices, peripherals, flexible in terms of supported SW (OS, utilities, drivers, run-time libraries), instrumented with thermal sensors, electric probes, thermally-induced mechanical stress sensors.

The developers will use the IDVs to test their codes, algorithms, drivers without having the constraints of a large system and having the advantage to be able to test their developments on different components through a very quick reconfiguration process.

1.2 Thermal Control in TEXTAROSSA

Since the advent of dark silicon, computing architecture needs both optimized heat dissipation solutions and run-time thermal control policies to operate reliably and efficiently despite the steady increase in power density and related issues such as hot spots [12]. Thermal control policies in the state of the art include: Linear Quadratic Regulators (LQR) [21]; Model Predictive Control (MPC) [20]; and Event-based control solutions [15].

Through detailed system modeling [17] and a multilevel thermal control strategy TEXTAROSSA aims to overcome the complexity of controlling an HPC platform from node to system level with minimal overhead. As the fastest temperature gradients occur at the silicon active layer, we will use fast event-based control loops [15] acting on DVFS [23] to limit the maximum operating temperature of compute elements. A key differentiating feature of the TEXTAROSSA approach to thermal control is that the inner control loop will in turn interact with higher level control loops operating the two phase cooling infrastructure of the node, which is comparatively slower and has higher overheads but has the capability to increase the heat transfer coefficient on-demand, thus allowing to relieve the need to reduce frequency using DVFS, in turn improving performance. A further supervisory control layer will allow to set the desired temperatures at the rack level based on reliability metrics. Multilevel control allows thus to partition the system level control problem into multiple interacting control loops, each optimized for the specific thermal dynamics to control.

1.3 The TEXTAROSSA Consortium

TEXTAROSSA started in April 2021, and will last for three years. It is supported by joint funding from the European High Performance Computing (EuroHPC) Joint Undertaking and the national governments of Italy, France, Poland, and Spain. The project is led by the Italian national agency for new technologies, energy and the sustainable economic development (ENEA) with technical leadership provided by CINI, an Italian consortium grouping together three leading universities, Politecnico di Milano, Università degli studi di Torino, and Università di Pisa. The three Italian universities are part of the lab of CINI[4], created in 2021, that is grouping together the main academic and research entities

[4] https://www.consorzio-cini.it/index.php/it/laboratori-nazionali/hpc-key-technologies-and-tools (last accessed March 2022).

working in the field of high-performance and Exascale computing in Italy. In Quattro, an Italian startup company, provides the innovative 2-phase cooling system described in this paper, while E4 Computer Engineering is in charge of the Integrated Development Vehicle. Fraunhofer (Germany), INRIA (France), ATOS (France), BSC (Spain), PSNC (Poland), INFN (Italy), CNR (Italy), Université de Bordeaux (France), CINECA (Italy) and Universitat Politècnica de Catalunya (UPC) complete the consortium, providing contributions on software, hardware, development platforms, and applications that are outside the scope of this paper. More information on the activities carried out during the execution of TEXTAROSSA can be found in the project website[5].

2 The TEXTAROSSA Platform

TEXTAROSSA leverages an Innovative Integrated Development Vehicles (IDV-E) platform developed by E4 Computer Engineering according to an open architecture model and exploiting a heterogeneous architecture (nodes using ARM64 plus FPGA solutions). The IDV-E platform will implement a multi-node HPC platform, and will allow prototyping and benchmarking all innovations addressed in the following points:

1. *New integrated heterogeneous architecture at node level:* extending the experience of core partners in the European Processor Initiative (development of processor and accelerators IPs, and integrated heterogeneous HPC platforms exploiting both ARM64 and RISC-V cores) to boost the EuroHPC roadmap in terms of energy-efficiency, high-performance and secure HPC services.
2. *High-efficiency cooling system at node and system levels:* innovative and high-efficiency cooling mechanism (based on a two-phase cooling technology) for HPC platforms at node and system levels and power monitoring and controller IP exploiting new models of the thermal behavior and of a multi-level control strategy. The data collected in this project as well as the know-how developed by the partners with respect to the two-phase cooling technology will be representative of the working environment of the EPI-based nodes.
3. *Innovative tools for seamless integration of reconfigurable accelerators:* such tools, targeting the AI/DNN computing paradigm, include compilers, memory hierarchy optimization and runtime systems, scaling over multiple interconnected reconfigurable devices, and SW header-only based on Fast Flow and memory hierarchy optimization in an EPI-like HPC architecture, compiler tools for mixed-precision, all in heterogeneous HPC platform and in future EPI tool chain. Automatic instrumentation of the accelerators with energy/power models to enhance a global (fine grain) power monitoring and control [8,22].

The node designed by E4 will apply the two-phase thermal management solution developed by InQuattro to the most thermal critical components, i.e. the CPUs.

[5] https://textarossa.eu (last accessed March 2022).

During the first year of the project the opportunity to apply these components to the accelerators (FPGA) has been evaluated within a co-design approach, and the decision to apply the cooling system also to the FPGA device was made. It was considered that the cooling of the accelerator does not constitute a substantial increase in complexity in terms of mechanical complication such as increase in the number of pipes, increase in heat exchangers, increase of the delivery liquid cooling flow rate.

After a careful evaluation of the possible commercial platforms based on ARM technology, the choice of the system to which to apply the two-phase cooling system fell on the Ampere Mt.Collins 2U system with Ampere Altra Max processor; the main reasons are: (i) it supports a number of PCIe slots providing the possibility of adding FPGA boards (up to 3) and/or other boards if needed, (ii) it has the physical space for adding the cooling system, (iii) it presents a good match between the amount of heat to be removed and the design point of the cooling system developed in the project, (iv) it has an architecture (ARM) compatible with that of the EPI project, (v) the possibility of receiving the system in times compatible with the project (an aspect not taken for granted given the current state of shortage worldwide).

As for the FPGA, the choice fell on the U280 Xilinx Passive Model, it is able to provide significant computing power and the flexibility of memory access via HBM2 or DDR protocol with a maximum consumption of 225W. This device also guarantees the use of the VITIS software stack, widely used in research by the various TEXTAROSSA partners.

The Mt.Collins system integrated in E4 laboratory is shown in Fig. 1. It is a dual socket configuration in a 2U form-factor with Ampere Altra Max Processors, is an excellent fit for Android in the Cloud, AI/ML and HPC usages.

The Ampere Altra Max processor-based Mt.Collins dual socket rack servers provide high performance with industry leading power efficiency per core.

The versatile platform offers 160 PCIe Gen4 lanes for flexible I/O connectivity via PCIe slots and another 16 PCIe Gen4 lanes for OCP 3.0 networking. Mt.Collins supports thirty-two DDR4 3200 MT/s DIMMS with a maximum memory capacity of 8 TB.

It also supports OCP NIC 3.0 connector with multi-host support to capitalize on the mechanical, thermal, manageability, and security benefits.

In addition, Mt. Collins includes one internal M.2 NVMe storage interface for ultra-fast reads/writes, eliminating PCIe switch adapters.

Mt. Collins includes MegaRAC®, BMC, and Aptio®V BIOS support. Key features include dynamic fan control, temperature monitoring, and TPM 2.0 for security. The platform includes two redundant power supplies providing the reliability required for datacenters. BMC includes support for IPMI and Redfish protocols for remote management.

The dimensions are: 33.36 in. (L) x 17.63 in. (W) x 3.425 (H).

Ampere Altra Max processor (based on ARMv.8.2+) offers up to 128 cores operating at a maximum of 3.0 GHz. Each core is single threaded by design with its own 64 kB L1 I-cache, 64 kB L1 D-cache, and a huge 1 MB L2 cache,

Fig. 1. Mt.Collins System from the top, on the left the space for the disks (covered by the metal plate of the case), in the center you can see the two sockets with the memories around, on the right at the top the two power supplies of the system, on the left center a black container for one of the possible accelerator boards.

delivering predictable performance 100% of the time by eliminating the noisy neighbor challenge within each core.

The processor technology is 7 nm FinFET with a TDP of 250 W.

The Ampere Altra Max processor offers high bandwidth and memory capacity of up to 4 TB per socket.

With 128 lanes of PCIe Gen4 per socket with support for 192 PCIe Gen4 lanes in 2P configuration that can be bifurcated down to x4, Ampere Altra Max provides maximum flexibility to interface with off-chip devices, including networking cards up to 200 GbE or more, and storage/NVMe devices.

Ampere Altra Max supports cache coherent connectivity to off-chip accelerators; 64 of the 128 PCIe Gen 4 lanes support Cache Coherent Interconnect for Accelerators (CCIX), that could be used for networking, storage, or accelerator connectivity.

The Ampere Altra Max processor provides extensive enterprise server-class RAS capabilities. Data in memory is protected with advanced ECC in addition to standard DDR4 RAS features. End-to-end data poisoning ensures corrupted data is tagged and any attempt to use it is flagged as an error. The SLC is also ECC protected, and the processor supports background scrubbing of the SLC cache and DRAM to locate and correct single-bit errors.

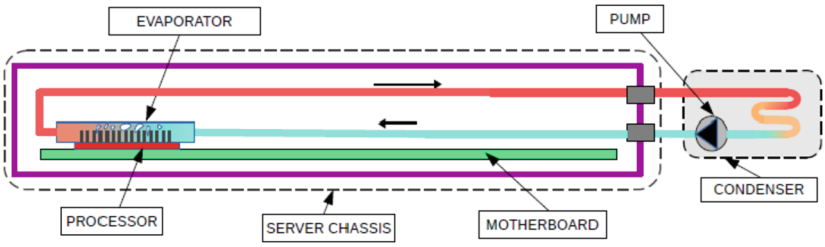

Fig. 2. Two-phase schematic loop in a server configuration.

Currently the system is available at the E4 laboratory ready to be coupled to the two-phase cooling system.

The potential problem we expect is related to the removal of the fans (even partially). The attention point consists in the system checks that may be performed at the BIOS level that could activate safety mechanisms if the checks detect the lack of fans. These mechanisms could interrupt the system startup process.

To overcome this problem, if it occurs, several strategies are possible, i.e. (i) request a BIOS change from the vendor, (ii) replace the fans with resistors, in this case consumption would be constant and it would be possible to subtract them from the total value.

While the possible problem with the BIOS is something that should be taken into account, for the time being we do not envision any roadblock for the development of the implementation of the cooling circuitry.

3 Innovative 2-Phase Cooling

3.1 Technology Description

Thermal control with two-phase cooling is a new technology for the sector of electronics cooling. This technology has been extensively studied and developed for cooling the most critical components in fusion power reactors [6]. Recently, two-phase cooling has been miniaturized matching the small dimensions of electronic components. Compared to classical cooling systems like heat pipes and liquid cooling (single phase forced convection), the new technology is able to remove higher heat fluxes, at lower pumping energy, lower mass of the entire loop, and to maintain the target surface of the electronic component isothermal.

The concept of a two-phase cooling system is quite simple, as shown in Fig. 2.

The technology uses a cold plate, (aluminum or copper), that serves as a heat sink, a direct on chip evaporative heat exchanger. The cold plate is a multi-micro-channels evaporator: a metal plate with micro fins machined on it. The evaporator is placed in direct contact with the processor (CPU or GPU) case whit the aid of a thermal paste in order to obtain stable and low thermal contact resistance. The coolant flows through micro-channels in the evaporator to capture heat from the processor by evaporation processes, and then flows on to a condenser, in which heat is dissipated to the surrounding environment via water or air. The coolant coming out of the condenser travels back through the pump, and the cycle repeats itself. The loop is a hermetically sealed closed system, so the processors and all electronic components are not in direct contact with the fluid. Other important features for two-phase cooling technology are the use of dielectric liquids that eliminates the risk of electric damages caused by an accidental leakage of the coolant, and a virtually zero maintenance that eliminates the requirement of skilled personnel. The mentioned dielectric fluids are non-flammable, non-toxic, perfectly compatible with the environment with extremely low GWP (Global Warming Potential) and Ozone Depletion Potential (ODP).

In Quattro is developing the two-phase cooling system and will integrate it at rack level for data center with liquid thermal transport infrastructure installed, as shown in Fig. 3. The two-phase cooling system will be designed for high density CPU and GPU configurations up to 5 kW per server. All the server cooling systems will be integrated at rack level with a CDU (Coolant Distributor Unit) that will transfer heat to the data center liquid thermal transport infrastructure. Each rack will be able to remove up to 90 kW using water in the liquid transport infrastructure as hot as 45 °C, eliminating the need for expensive and inefficient chillers or cooling towers.

3.2 Advantages and Innovation

Two-phase cooling can be seen as an evolution of the liquid cooling system. Therefore, the main advantages of liquid cooling compared to traditional air cooling are extended to two-phase cooling systems: higher energy efficiency, smaller footprint, lower cooling system total cost of ownership, enhanced server reliability, lower noise, etc. The main difference lies in the use of latent heat (vaporization and condensation) instead of sensible heat, depending on heat capacity, in removing thermal loads from processors. Latent heat is many times higher than thermal capacity of liquids. This allows evaporative cooling technologies to remove heat more efficiently with lower flow rates, and allows the processors to work continuously at top clock speed by removing higher thermal power densities. In particular, lower mass flow rate results in lower electrical energy spent for pumping the fluid in the cooling system with interesting economic and technical advantages.

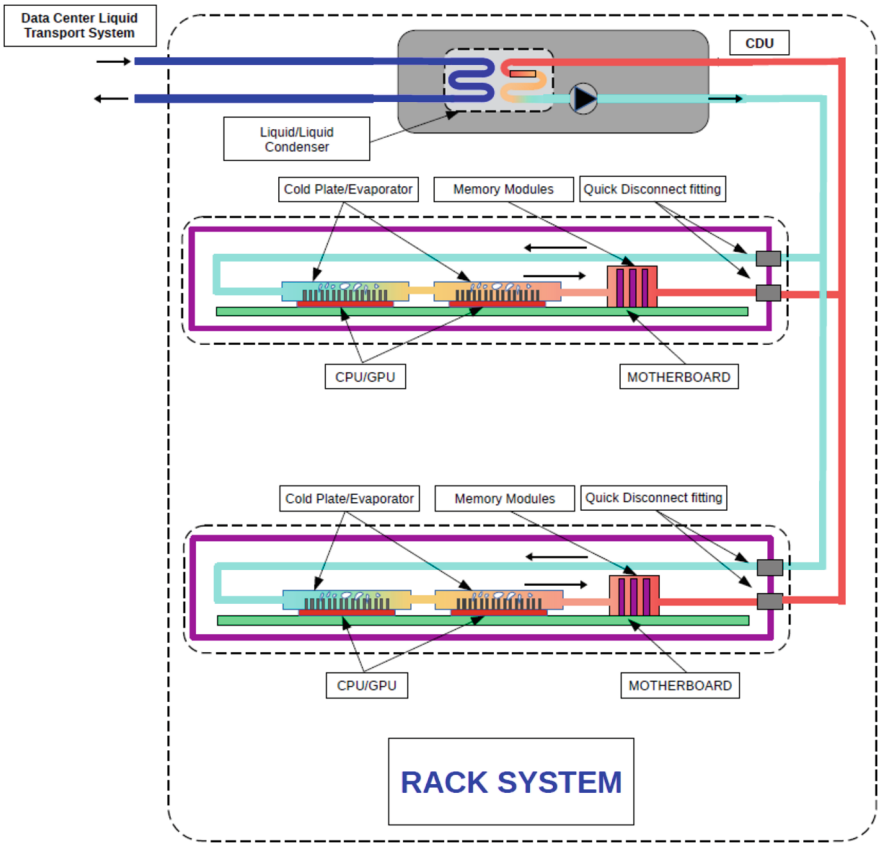

Fig. 3. Two-phase schematic loop in a rack configuration with liquid-to-liquid condensing unit.

4 Thermal Modeling

The use of evaporative cooling systems in HPC environment promises significant compelling advantages, including the possibility to improve computing systems performance thanks to the reduced need for frequency throttling, as well as improved PUE resulting in increased profitability and sustainability through improved energy efficiency. However, evaporative cooling solutions are also significantly more complex to design, and require dedicated control policies for optimal performance.

As a result, thermal simulation has a prominent role for efficient and effective cooling systems [13] and control policy [14] design. The thermal simulation of evaporative cooling systems is also significantly more complex compared to traditional cooling systems, such as air and water cooling, due to the inherent nonlinear phenomena related to evaporation.

Efficient and accurate thermal simulation of evaporative HPC cooling systems is currently an open research challenge that TEXTAROSSA aims to overcome, as well as the integration of evaporative thermal models with state-of-the-art integrated circuit thermal simulators.

4.1 Simulation Approach

The thermal design power (TDP) of HPC computing architectures is steadily increasing, and computer architectures are becoming increasingly heterogeneous. At the same time, a shift to heterogeneous solutions is being observed also in heat dissipation solutions. Traditional air cooling is progressively being replaced with liquid cooling in the server and high end desktop market, and more innovative solutions such as the evaporative cooling we are proposing in TEXTAROSSA are expected to soon become a necessity.

To effectively handle the simulation of heterogeneous heat dissipation solutions, a paradigm shift in the construction of thermal simulators for computer architectures is required. Most existing thermal simulators are in fact monolithic, in that they essentially hardcode the equations describing the physics of the problem and only provided very limited configurability, only in terms of model parameters.

The TEXTAROSSA simulation approach is instead based on co-simulation with domain-specific languages for the modeling and simulation of systems expressed in terms of differential equations. This approach, that we introduced with the design of the 3D-ICE 3.0 thermal simulator [18] also suits the need for efficient co-simulation of chip thermal models, which can be reasonably modeled using only linear differential equations, together with evaporative heat sink models, that conversely require nonlinear differential equations to be modeled.

The transient differential equation solver of 3D-ICE 3.0, unlike monolithic thermal simulators, only simulates the integrated circuit and heat spreader, and provides a co-simulation interface to a seaprate heat sink model, thus allowing arbitrary heat sinks to be simulated without the need to modify and revalidate the core of the thermal simulator. The 3D-ICE 3.0 co-simulation interface adheres to the FMI [1] standard, thus providing access to a vast set of languages and tools [2] for the modeling of heat dissipation solutions.

In TEXTAROSSA, we chose to rely on the Modelica [9] object-oriented modeling language for modeling evaporative cooling loops, and in particular the OpenModelica open source implementation [10]. Work is currently underway in extending the ExternalMedia Modelica library [7] to support co-simulation with 3D-ICE. This library provides access from Modelica models to the CoolProp [5] library for computing the thermodynamic properties of refrigerant fluids used in evaporative cooling systems. After this task is completed, a dedicated library will be developed for the simulation HPC cooling based on evaporative technologies.

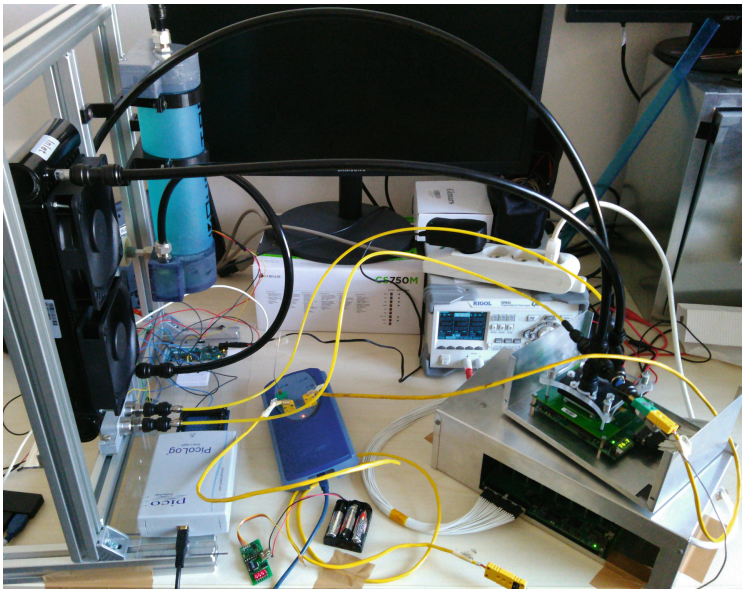

Fig. 4. Experimental setup for the characterization of the prototype evaporative cooling cycle. Thermal test chip platform fitted with evaporator (right), radiator, tank and pump completing the cooling loop (left).

4.2 Experimental Validation

Accurately modeling physical phenomena such as evaporative thermal dissipation calls for experimental data. Some of the thermal phenomena related to evaporation rely on empirical correlations, thus experiments are needed to fine tune correct parameter values.

The availability of high quality experimental data allows to perform validation not only of entire thermal models, but also of individual components that can be later rearranged to simulate a far wider range of heat dissipation solutions than can be reasonably realized for experimentation.

In TEXTAROSSA, we performed thermal experiments with a prototype evaporative cooling loop, collecting data that will be used for model validation. For performing these experiments, we relied on a Thermal Test Chip platform [19] developed at Politecnico di Milano. TTCs are a key part of the TEXTAROSSA stategy for the validation of thermal models, as the direct use of MPSoCs such as processors or GPUs to perform thermal experiments is made extremely difficult by the uncertainty in the power spatial distribution across the silicon die during computational workloads in modern processors, as well as due to the insufficient number of temperature sensors to fully reconstruct the temperature spatial distribution.

The TTC platform that we developed [19], provides a total of 16 heating elements and 16 temperature sensors, arranged as a 4×4 grid, coupled with a

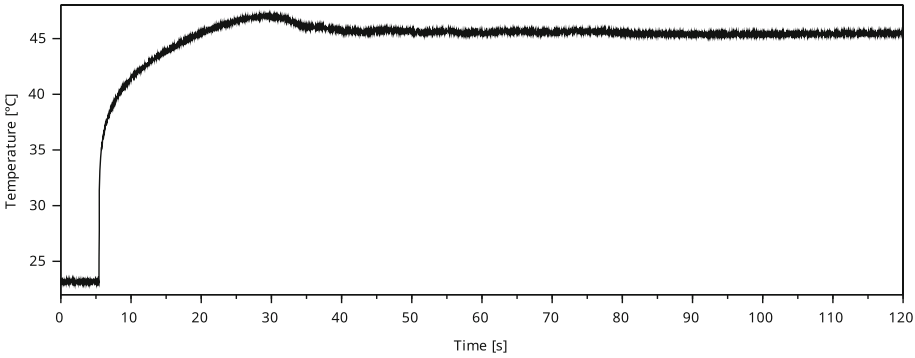

Fig. 5. Temperature of the TTC active silicon layer subject to a 30 W power step. Note the temperature decrease after the initial overshoot when evaporation starts (temperature data averaged from the 4 center temperature sensors).

fast sensing and actuation chain allowing to measure temperatures with a 0.1 °C resolution at up to 1 kHz sample rate, as well as to produce spatial (hot spots) as well as temporal temperature gradients.

Characterization experiments have already been performed with a prototype evaporative cooling cycle developed by INQUATTRO as part of TEXTAROSSA, using the setup of Fig. 4. The prototype cooling cycle uses an air radiator, that will be replaced with a water heat exchanger in a subsequent design iteration. Figure 5 shows a sample of the collected data, a transient test consisting in a 30W uniform power step applied to the TTC, reporting the active silicon temperature averaged from the four center temperature sensors. The figure clearly shows a temperature reduction after the initial transient, indicating the transition from single to two phase cooling.

5 Conclusion

In this paper, we presented the approach towards thermal control adopted in the EuroHPC TEXTAROSSA project. The approach is based on an innovative two-phase cooling system, coupled with thermal monitoring, modeling, and control policies. Initial experiments, performed on a TTC coupled with a prototype two-phase cooling system, show how the temperature can be effectively controlled, leading to a reduction of the working temperature. Future works include larger-scale experimentation on E4's integrated development vehicle, employing more advanced versions of the two-phase cooling system.

References

1. The Functional Mock-up Interface (FMI) standard. https://fmi-standard.org/
2. Tools supporting the Functional Mock-up Interface (FMI) standard. https://fmi-standard.org/tools/

3. Agosta, G., et al.: TEXTAROSSA: Towards EXtreme scale Technologies and Accelerators for euROhpc hw/Sw Supercomputing Applications for exascale. In: 2021 24th Euromicro Conference on Digital System Design (DSD), pp. 286–294. IEEE (2021)
4. Aldinucci, M., et al.: The Italian research on hpc key technologies across eurohpc. In: Proceedings of the 18th ACM International Conference on Computing Frontiers, pp. 178–184 (2021)
5. Bell, I.H., Wronski, J., Quoilin, S., Lemort, V.: Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library coolprop. Ind. Eng. Chem. Res. **53**(6), 2498–2508 (2014)
6. Boyd, R.: Subcooled Flow Boiling Critical Heat Flux (CHF) and Its Application to Fusion Energy Components-Part I. A Review of Fundamentals of CHF and related Data Base. Fusion Technol. **7**, 7–31 (1985)
7. Casella, F., Richter, C.: ExternalMedia: A Library for Easy Re-Use of External Fluid Property Code in Modelica. In: Modelica, March 3rd-4th, 2008
8. Cremona, L., Fornaciari, W., Zoni, D.: Automatic identification and hardware implementation of a resource-constrained power model for embedded systems. Sustain. Comput. Informatics Syst. **29**(Part), 100467 (2021). https://doi.org/10.1016/j.suscom.2020.100467
9. Fritzson, P.: Principles of Object-Oriented Modeling and Simulation with Modelica 2.1. John Wiley & Sons, London, UK (2004)
10. Fritzson, P., et al.: The OpenModelica integrated environment for modeling, simulation, and model-based development. MIC—Model. Identif. Control **41**(4), 241–295 (2020). https://doi.org/10.4173/mic.2020.4.1
11. Group, E.R.I.A.: Strategic research and innovation agenda 2019. Technical report (2019)
12. Hankin, A., et al.: HotGauge: a methodology for characterizing advanced hotspots in modern and next generation processors. In: 2021 IEEE International Symposium on Workload Characterization (IISWC), pp. 163–175 (2021)
13. Iranfar, A., et al.: Thermal characterization of next-generation workloads on heterogeneous MPSoCs. In: Proceedings of 2017 SAMOS, vol. 2018, pp. 286–291, January 2018
14. Leva, A., et al.: Event-based power/performance-aware thermal management for high-density microprocessors. IEEE Trans. Control Syst. Technol. **26**(2), 535–550 (2018)
15. Leva, A., et al.: Event-based power/performance-aware thermal management for high-density microprocessors. IEEE Trans. Control Syst. Technol. **26**(2), 535–550 (2018). https://doi.org/10.1109/TCST.2017.2675841
16. Malms, M., et al.: Etp4hpc's strategic research agenda for high-performance computing in europe 4. Technical report (2020)
17. Papadopoulos, A., et al
18. Terraneo, F., et al.: 3d-ice 3.0: Efficient nonlinear mpsoc thermal simulation with pluggable heat sink models. IEEE Trans. Comput.-Aided Design Integrated Circuits Syst. **41**(4), 1062–1075 (2022). https://doi.org/10.1109/TCAD.2021.3074613
19. Terraneo, F., Leva, A., Fornaciari, W.: An open-hardware platform for MPSoC thermal modeling. In: Embedded Computer Systems: Architectures, Modeling, and Simulation, pp. 184–196 (2019)
20. Zanini, F., Atienza, D., Benini, L., De Micheli, G.: Thermal-aware system-level modeling and management for multi-processor systems-on-chip. In: 2011 IEEE International Symposium of Circuits and Systems (ISCAS), pp. 2481–2484 (2011)

21. Zanini, F., Atienza, D., De Micheli, G.: A control theory approach for thermal balancing of mpsoc. In: Proceedings of 2009 ASP-DAC, pp. 37–42. IEEE Press (2009)
22. Zoni, D., Cremona, L., Fornaciari, W.: Design of side-channel-resistant power monitors. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **41**(5), 1249–1263 (2022). https://doi.org/10.1109/TCAD.2021.3088781
23. Zoni, D., Fornaciari, W.: Modeling DVFS and power-gating actuators for cycle-accurate noc-based simulators. ACM J. Emerg. Technol. Comput. Syst. **12**(3), 27:1–27:24 (2015). https://doi.org/10.1145/2751561