Controlled Direct Liquid Cooling of Blade Servers

Riccardo Lucchese, Damiano Varagnolo, Andreas Johansson

***Abstract*—We formulate a flexible modelling and control framework for direct liquid cooling systems in blade servers. We present our modelling derivations in full detail and then describe how to manipulate the coolant's flow while 1) regulating the temperatures of the blade's on-board components within safe operational conditions and 2) minimizing the amount of supplied coolant while increasing the temperature of the coolant at the blade's outlet. The latter is indeed a performance index of paramount importance when coupling liquid cooling strategies with heat recovery systems. We quantify in silico the benefits of the proposed liquid cooling control strategy over twelve realistic scenarios corresponding to different inlet coolant temperatures and computational loads.**

***Index Terms*—Controlled Liquid Cooling, Thermal Networks, Thermal Management, Energy Recovery, Energy Reuse, Direct Liquid Cooling, Blade Servers, Polynomial Optimization, Model Predictive Control.**

I. Introduction

Modern data centers are large scale, energy intensive processes that can accommodate millions of computing cores and hundreds of thousands of blade servers. While the industry's trend has been to double the energy efficiency of computing units every 18 months (a fact known as Koomey's law [1]), deployments have seen an increase in both total power loads (up to 120MW) and power densities (up to 30KW per square meter). Higher power densities present a number of technological challenges concerning the design of the electronic equipment, its packaging and its thermal management [2], [3]. In particular, the heat loads are approaching the limits of traditional air-cooling solutions, exacerbating the power consumption and reliability issues. Indeed, air-cooling operates with high temperature gradients between air and the active components: On one hand, it becomes necessary to pre-cool the air affecting directly the overall energy efficiency; On the other hand, the wide temperature gap translates into low exergetic gains at the outlet which hinder the repurposing of waste heat [4].

Liquid cooled Computer Room Air Handlers (CRAHs) and air-cooled blade servers are, de facto, the standard cooling solution in existing data center designs [5], [6]. In these systems, the thermal state of the data center is typically managed in a decoupled fashion: Cooling of the computer room is performed through air-conditioners/air-handlers while the thermal state of the blade servers is managed by actuating a varying number of local fans. Air-cooled server enclosures have been subject to a substantial modelling effort, see [7], [8], [9], [10] and references therein. State of the art Model Predictive Control (MPC) strategies for air-based cooling are then based on minimum cost, polynomial, optimal control problems [8], [9], [10].

With acquisition costs being overwhelmed by the running costs, data center operators are turning to more energy efficient solutions based on liquid cooling [11], [12], [13], [14], [15]. A compelling alternative is then to let each blade server reject its heat-load directly into a liquid cooling loop using cold plate heat exchangers. Indeed, the higher thermal capacitance and lower thermal resistance of liquids enable compact designs that match higher power density specifications and moreover allow direct liquid cooling implementations to operate with a smaller temperature gradient between the chips and the coolant compared to conventional air-cooling solutions. This results in a smaller rate of exergy destruction and enables heat recovery systems with efficiencies (up to 85 percent) that are impossible to be obtained with lower thermal capacity coolants such as air [4]. Furthermore, liquid cooled blades can run on hot-water opening to adaptive, free-cooling, implementations and higher quality heat harvests which enable, in their turn, a multiplicity of secondary reuse scenarios such as supplying the basic heat load needs to indoor complexes, greenhouses [16], district heating [12], desalination and refrigeration processes [4], preheating of boiler feed water in power plants [17].

Direct ("on-chip") liquid cooling solutions have not, to our current best knowledge, been investigated thoroughly from a control perspective. In particular, there is a lack of studies evaluating the benefits of dynamical provisioning of the cooling resources in a direct liquid cooled setting. A large part of the existing body of works focuses on the data center level. For instance, [12], [18], [19], [20] among others, pursue to quantify the prospective efficiency gains of direct liquid cooling over other liquid cooling and air-cooling technologies. In hybrid liquid cooling problems, a portion of the deployed air-cooled servers are retrofitted using direct liquid cooling in order to ameliorate hot-spots due to unwanted air recirculation. Reducing the overall cooling costs corresponds then to 1) an off-line, optimal, selection of the blades to be retrofitted and 2) the on-line, optimal, job allocation over the mix of air-cooled and liquid cooled platforms [21], [22]. We stress that in the above studies the liquid cooling provisioning is not adaptive and rather it is matched to the peak heat load.

The modelling of liquid cooling convective heat ex-

This work is supported by the Celtic Plus project SENDATE-Extend (C2015/3-3), by the Norrbottens Forskningsråd project DISTRACT, and by the "Smart Machines and Materials" area of excellence at Luleå University of Technology, Luleå, Sweden. All the authors are with the Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Luleå, Sweden. Emails: { riccardo.lucchese | damiano.varagnolo | andreas.johansson }@ltu.se.

changes at the chip level has been investigated in [23], [24], [25] for 3D stacked/integrated architectures where using air as the coolant medium becomes inadequate due to the manifold increase in power consumption and thermal resistance between the layers. Notice, however, that these approaches focus on the chip level and disregard any dynamical interactions (such as heat exchanges) with other heating elements.

A. Statement of contributions

Our contributions can be summarized in:

1) Developing a novel, flexible, modelling framework of the thermal dynamics of blade servers that extends naturally to the rack and the data center levels;
2) Studying a minimum coolant-supply strategy for the direct liquid cooling of blade servers that aims at reducing the control effort and enabling heat reuse applications.

We propose control-oriented models of the temperature dynamics of the different physical devices that participate in the heat processes within data centers. These devices are thus organized in a *thermal network* whose nodes are thermal entities that can locally transfer heat to, from, and within the network. A graph theoretic formalism concisely summarizes which devices interact by exchanging heat and by which mechanism between conduction, convection, or both. We specialize our treatment to the thermal networks that arise at the server level since in a (direct) liquid cooled data center the bulk of the heat is harvested at the servers.

We then consider that implementing cost-effective heat recovery systems pends on the data center's ability to systematically act as a stable source of high quality heat, that is, on the ability of the infrastructure to sustain outlet flows with high temperature [19], [17]. Since the computational workload of blade servers (and the corresponding heat load) is typically time-varying, static provisioning of the cooling resources based on nominal, worst-case, considerations leads to over-use of the liquid coolant, time varying and overall lower outlet temperatures and thus lower quality heat yields. Instead, we propose a dynamic provisioning strategy, in terms of a feedback law, that aims at minimizing the volumetric flow of the coolant servicing the blade and that is able to attain higher output coolant temperatures, making it more suitable for heat reuse scenarios.

Remark 1: In this work we focus on the modelling and control issues at the server level and do not treat in an explicit and detailed manner other technological components that are necessary to operate the cooling loops such as purifiers, pre-heaters or the maintenance and operation of the Coolant Distribution Units (CDUs).

Remark 2: Modelling of the computational aspects such as the compute, memory, storage and networking loads is accomplished in an aggregate manner by mapping these quantities into the corresponding electrical power consumption at the chip level.

B. Organization of this manuscript

Section II introduces our thermal modelling framework in terms of networks of devices with local thermal dynamics and heat exchange interactions described by graph overlays. Section III specializes our thermal model into a library of reusable nodes aimed at the server level. Section IV introduces an MPC strategy for regulating the volumetric flow of the liquid coolant. Section V shows numerical investigations of the benefits of dynamical control laws against static provisioning laws. Finally, Section VI collects concluding remarks and future directions.

II. A thermal modelling framework for controlled liquid cooling

Our control-oriented modelling framework abstracts a data center as a network of interacting devices endowed with local, lumped, thermal properties and temperature dynamics. We propose to use a graph-theoretic formalism where $\mathcal{N} = \{1, 2, \dots, n\}$ denotes the set of participating nodes (with $n \doteq |\mathcal{N}|$ being the network size) and where the topology of the heat transfers is described by two distinct directed and static graph overlays $\mathcal{E}_{cd}, \mathcal{E}_{cv} \subset \mathcal{N} \times \mathcal{N}$: $\mathcal{E}_{cd}$ ($\mathcal{E}_{cv}$ respectively) encodes which devices interact by exchanging heat through the mechanism of conduction (convection). Notice that the semantic meaning of a link in the two overlays differs (see the detailed discussion later in Section II-A and Section II-B): A link $(j,h) \in \mathcal{E}_{cd}$ indicates a directed *heat flow* from node j to node h; A link $(j,h) \in \mathcal{E}_{cv}$ establishes instead a liquid cooling interconnection between node j and node h, indicating the flow of *both heat and mass* between the two nodes. Given a generic graph overlay $\mathcal{E}$ on $\mathcal{N} \times \mathcal{N}$, we define the in-neighborhood and out-neighborhood sets of node $j \in \mathcal{N}$ respectively as

$$\delta^{\mathcal{E}}_{-}(j) \doteq \{h : (h,j) \in \mathcal{E}\}, \quad \delta^{\mathcal{E}}_{+}(j) \doteq \{h : (j,h) \in \mathcal{E}\}. \quad (1)$$

A. The heat conduction overlay

The nodes that exchange heat through conduction are connected in the heat conduction overlay $\mathcal{E}_{cd}$. If $(j,h) \in \mathcal{E}_{cd}$, then nodes j and h exchange heat energy at some rate that depends on the physical properties of the system and the temperature gap between them.

Example 3: Consider the simple case of a thermal network with two electrical components connected through a solid thermal bridge. Let j, h be two generic and fixed indices and define $\mathcal{N} = \{j, h\}$; The thermal bridge induces two links in the heat conduction overlay, that is, $\mathcal{E}_{cd} = \{(j,h),(h,j)\}$. Let $t \mapsto x^c_j(t)$ and $t \mapsto x^c_h(t)$ be the continuous time trajectories of the proxy temperatures for the first and the second component respectively. Then the rate at which heat is absorbed by j through the mechanism of conduction can be approximated by Fourier's law as

$$q^{cd}_j(t) = -k_{hj}\big(x^c_j(t) - x^c_h(t)\big) \quad (2)$$

where the constant $k_{hj} \in \mathbb{R}_{>0}$ is the lumped thermal conductivity of the bridge.

For a generic thermal network with size $n = |\mathcal{N}|$ we collect the thermal conductivity parameters in the matrix $K \in \mathbb{R}^{n\times n}_{>0}$ and set $k_{jh} = 0$ whenever $(j,h) \notin \mathcal{E}_{cd}$. The rate of heat transfer due to conduction at the generic node $j \in \mathcal{N}$ can then be written as

$$q^{cd}_j(t) = -\sum_{h\in\delta^{\mathcal{E}_{cd}}_-(j)} k_{hj}\big(x^c_j(t) - x^c_h(t)\big). \tag{3}$$

We stress that the conduction overlay $\mathcal{E}_{cd}$ is directed and that in general $(j,h) \in \mathcal{E}_{cd}$ does not imply $(h,j) \in \mathcal{E}_{cd}$. This asymmetry reflects those situations where a device rejects heat to an *environmental node*, that is, a specific node type describing heat reservoirs and thus characterized by a constant temperature. In particular, for an environmental node $j \in \mathcal{N}$ we have $\delta^{\mathcal{E}_{cd}}_-(j) = \emptyset$ and thus $q^{cd}_j(t) = 0$ at all times t.

Not all nodes participate in the conduction overlay. For instance, *supply* and *collector* nodes (discussed in the next section) model the entry and exit points for the coolant in the liquid cooling circuit and have empty in- and out-neighborhoods in $\mathcal{E}_{cd}$.

B. The heat convection overlay

Nodes can exchange heat through a convective mechanism either with the surrounding air (for instance, in the presence of forced air flow) or with a liquid coolant through a cold-plate. In this manuscript, we explicitly assume a quasi-static regimen for the air flows and disregard any heat terms due to air convection[1].

The heat convection overlay serves as a means to describe the flow of the coolant within the network: A directed edge $(i,j) \in \mathcal{E}_{cv}$ models an actual interconnection (or pipe) where the liquid coolant can flow. We let $\varphi : t\times\mathcal{N}\times\mathcal{N} \to \mathbb{R}$ be the continuous time trajectory of the volumetric coolant flow rates over $\mathcal{E}_{cv}$. $\varphi(t,j,h)$ is then the (volumetric) flow rate of liquid coolant through link (j,h) at time t if $(j,h) \in \mathcal{E}_{cv}$ and is equal to zero otherwise. In the following, for the sake of a compact notation, we identify $\varphi_{jh}(t) \doteq \varphi(t,j,h)$.

The coolant enters the network at the supply nodes $\mathcal{S} \subset \mathcal{N}$ and exits it at the collector (or drain) nodes $\mathcal{C} \subset \mathcal{N}$. We consider networks with a generic number $m \geq 1$ of supply nodes

$$\mathcal{S} = \{s_1, \dots, s_m\}, \quad |\mathcal{S}| = m \tag{4}$$

and a generic number $p \geq 1$ of drain nodes

$$\mathcal{C} = \{c_1, \dots, c_p\}, \quad |\mathcal{C}| = p. \tag{5}$$

Notice that supply nodes cannot act as collector nodes and vice versa: $\mathcal{S}\cap\mathcal{C} = \emptyset$. Moreover, supply nodes do not have inflows and, specularly, collector nodes have no outflows:

$$\delta^{\mathcal{E}_{cv}}_-(j) = \emptyset \quad \forall j \in \mathcal{S}, \quad \delta^{\mathcal{E}_{cv}}_+(j) = \emptyset \quad \forall j \in \mathcal{C}. \tag{6}$$

Instead, each supply node $j \in \mathcal{S}$ is univocally associated to a control variable $u_{I_s(j)}$ through the bijection $I_s : \mathcal{S} \mapsto \{1, \dots, m\}$. The order on $\mathcal{N}$ induces an order on the control variables such that they can be orderly stacked to form the control vector

$$\boldsymbol{u} \doteq \begin{bmatrix} u_{I_s(s_1)} & \dots & u_{I_s(s_m)} \end{bmatrix}^T \in \mathbb{R}^m_{\geq 0}. \tag{7}$$

The control input $t \mapsto \boldsymbol{u}(t)$ sets the supply flow rates to the liquid cooling network through

$$\sum_{h\in\delta^{\mathcal{E}_{cv}}_+(j)} \varphi_{jh}(t) = u_{I_s(j)}(t), \quad \forall\, j \in \mathcal{S}. \tag{8}$$

We assume that the volumetric flow rates are conserved at all nodes except for the supply and collector nodes:

$$\sum_{h\in\delta^{\mathcal{E}_{cv}}_-(j)} \varphi_{hj}(t) = \sum_{h\in\delta^{\mathcal{E}_{cv}}_+(j)} \varphi_{jh}(t), \quad \forall\, j \in \mathcal{N}\backslash(\mathcal{S}\cup\mathcal{C}) \tag{9}$$

In light of (6) and (9) we define the total flow crossing the j-th node at time t as

$$f_j(t) \doteq \begin{cases} \displaystyle\sum_{h\in\delta^{\mathcal{E}_{cv}}_+(j)} \varphi_{jh}(t) & \text{if } j \in \mathcal{S} \\ \displaystyle\sum_{h\in\delta^{\mathcal{E}_{cv}}_-(j)} \varphi_{hj}(t) & \text{otherwise} \end{cases} \tag{10}$$

and define the network's flow vector by stacking the individual flows:

$$\boldsymbol{f}(t) \doteq \begin{bmatrix} f_1(t), \dots, f_n(t) \end{bmatrix}^T. \tag{11}$$

We assume that the flow splitting ratios at nodes with multiple outflows are time-constant parameters of the physical system and independent from the flow rate. This independency from the working condition is motivated on a practical basis: A typical server platform accommodates the components (Central Processing Units (CPUs), Dual In-line Memory Modules (DIMMs), companion chips) in pairs leading to liquid cooling circuits with purposefully symmetric designs for which the above assumption holds. These ratios are formally encoded in the matrix $\Lambda \doteq (\lambda_{jh}) \in \mathbb{R}^{n\times n}_{\geq 0}$ defined through

$$\sum_{h\in\delta^{\mathcal{E}_{cv}}_+(j)} \lambda_{jh} = 1, \quad \varphi_{jh}(t) = \lambda_{jh} f_j(t), \quad \forall\, j \in \mathcal{N}. \tag{12}$$

The following proposition relates the control input in (7) and the flow vector in (11). For its proof (see the Appendix) we require two additional assumptions of both a practical and technical nature:

- $\mathcal{E}_{cv}$ has no self-loops, that is, $(j,j) \notin \mathcal{E}_{cv}$ for all $j \in \mathcal{N}$. In other words, the liquid coolant exiting component j never recirculates back into j through a direct interconnection;
- For each node $j \in \mathcal{N}$ such that $\delta^{\mathcal{E}_{cv}}_-(j) \neq \emptyset$ or $\delta^{\mathcal{E}_{cv}}_+(j) \neq \emptyset$ there exists a directed path over $\mathcal{E}_{cv}$ starting from j and reaching a collector node in $\mathcal{C}$. That is, the flow through each liquid cooled node must be able to reach a collector and exit the circuit.

Proposition 4: The instantaneous volumetric flow rate vector $\boldsymbol{f}(t)$ in (11) can be written explicitly as a linear

[1] In practice, one aims to minimize both conductive and convective heat exchanges with air since heat losses to the environment reduce the overall heat recovery efficiency [4].

function of the control vector $\boldsymbol{u}(t)$. In particular there exists a constant matrix $\Phi \in \mathbb{R}^{n\times m}$, function of the liquid cooling network topology and Λ, such that

$$\boldsymbol{f}(t) = \Phi \boldsymbol{u}(t). \tag{13}$$

Finally, we stress that not all nodes in the network are connected in the heat convection overlay. Moreover, not all nodes connected in the heat convection overlay will exchange heat with the liquid coolant (and thus influence its temperature). For instance, pure manifold nodes, devoid of local temperature dynamics, are used to model piping joints and splitters in the liquid cooling circuit.

C. The local thermal dynamics of nodes

The thermal dynamics of each node are modelled using low order state space representations: We do not explicitly account for complex three dimensional geometries and non-heterogeneous thermal properties. Instead, we consider lumped models that explain the average effect of the distinct heat contributions.

The dynamics of the generic j-th node has three states:

- $x_j^i(t)$: The temperature of the coolant entering the node at time t;
- $x_j^c(t)$: The local temperature of the node at time t (for instance, the temperature of a CPU);
- $x_j^o(t)$: The temperature of the coolant leaving the node at time t.

The flow of the coolant through the interconnections is assumed adiabatic and its temperature is thus determined only at the entry and exit points of each node, disregarding the explicit description of the in-transit thermal dynamics. The coolant flows into j, collects the heat produced within the node due to electrical dissipation, and exits at a higher temperature. During normal operation we have then $x_j^i(t) \leq x_j^o(t) \leq x_j^c(t)$ at all times.

Remark 5: We stress that thermal phenomena affecting the in-transit coolant can still be described in a control-oriented, lumped, fashion by introducing opportune, thermal, Resistor-Capacitor (RC) sub-networks modelling, for example, local thermal inertias and parasitic resistances to the environment.

1) Dynamics of the inflow temperature $x_j^i(t)$: Recall that the flow rates at the supply nodes are manipulable variables set through (8). The liquid coolant enters then the thermal network at the generic supply node $s_i \in \mathcal{S}$ at rate $u_{I_s(s_i)}(t)$ and given temperature $x_{s_i}^i(t)$. The latter temperature should be understood as an exogenous input since in practice the Coolant Distribution Unit (CDU) acts as a heat reservoir with a large thermal capacitance and a slow varying temperature that is too costly to regulate directly. Using again the bijection $I_s(\cdot)$ in (14) we introduce the vector of input temperatures by orderly stacking the supply coolant temperature of each supply node

$$\boldsymbol{x}^i(t) \doteq \begin{bmatrix} x^i_{I_s(s_1)}(t) & \dots & x^i_{I_s(s_m)}(t) \end{bmatrix}^T \in \mathbb{R}^m_{\geq 0}. \tag{14}$$

Those nodes that have more than one inflow act as mixing manifolds and correspond to points in the cooling circuit where multiple interconnections are channeled into a single pipe. We write the temperature of the total flow crossing the j-th node as the following average

$$x_j^i(t) \doteq \frac{1}{f_j(t)} \sum_{h\in\delta_-^{\mathcal{E}_{cv}}(j)} \varphi_{hj}(t) x_h^o(t), \quad j \in \mathcal{N}\setminus\mathcal{S}. \tag{15}$$

We notice that (15) weights the instantaneous temperature contribution of each incoming flow by the corresponding flow rate. This corresponds to consider the conservation law (9) together with two additional assumptions:

i) The coolant is perfectly mixed at the manifolds;
ii) The heat energy of the coolant is conserved during the mixing.

We notice that *ii*) above is motivated by the low flow rates while *i*) is supported by practical considerations: mixing and heat-exchange sites do not coincide in the hardware, and this allows flows to be mixed before they enter the active to-be-cooled parts.

Taking the derivative with respect to time on both sides of (15) yields the temperature dynamics of the coolant reaching node j:

$$\begin{aligned} \dot{x}_j^i(t) \doteq \frac{1}{f_j(t)} \sum_{h\in\delta_-^{\mathcal{E}_{cv}}(j)} \dot{\varphi}_{hj}(t) x_h^o(t) + \varphi_{hj}(t)\dot{x}_h^o(t) \ \dots \\ - \frac{\dot{f}_j(t)}{f_j^2(t)} \sum_{h\in\delta_-^{\mathcal{E}_{cv}}(j)} \varphi_{hj}(t) x_h^o(t), \quad j \in \mathcal{N}\setminus\mathcal{S}. \end{aligned} \tag{16}$$

Remark 6: For the sake of simplicity, we have up to now disregarded transport delays in our discussion. These phenomena are inherent in a liquid cooling setting and we treat them in a formal way later in Section III.

2) Dynamics of the local temperature $x_j^c(t)$: As the coolant flows in the cooling circuits it traverses the components that need to be cooled harvesting heat from these local sources. Tracking the temperatures of these components is central to our modelling effort since the safe operation of the network pends on being able to regulate these temperatures below specified thresholds (this aspect is further discussed and formalized in Section IV).

To this aim, the generic j-th node is also seen as a thermal subsystem with heterogeneous thermal characteristics. To model the local temperature dynamics we thus consider: $x_j^c(t), j \in \mathcal{N}$, the proxy temperature of the component; $\dot{x}_j^c(t)$, the derivative with respect to time of the node's proxy temperature; and d_j, the lumped heat capacity of the node relating the local heat exchanges and $x_j^c(t)$. The corresponding dynamics is thus

$$d_j \dot{x}_j^c(t) = q_j^{cd}(t) + q_j^{cv}(t) + p_j(t), \tag{17}$$

and builds on top of the following three contributions:

- $q_j^{cd}(t)$ is the rate at which heat energy is transferred to/from the node purely through conduction mechanisms;

- $q_j^{cv}(t)$ is the rate of heat exchanges with the coolant in the liquid cooling circuit (through the mechanism of convection);
- $p_j(t)$ is the rate at which electrical energy is converted into thermal energy locally at the node.

The functional structure of q_j^{cd} follows from (2) in Section II-A. Under constant flow conditions, the term q_j^{cv} in (17) can be approximated using Newton's law of cooling as

$$q_j^{cv}(t) \propto -\left(x_j^c(t) - x_j^l(t)\right), \tag{18}$$

where the effective temperature of the coolant $x_j^l(t)$ is, in first approximation, a function of $x_j^i(t)$, $x_j^o(t)$ and the physical parameters of the specific heat exchanger. The exact structure of q_j^{cv} is given for a relevant class of heat exchangers in the following Section III. Finally, we notice that the local self-heating phenomena are modelled by considering that the node converts all the electrical power that it drains, that is, $p_j(t) \in \mathbb{R}_{\geq 0}$, into heat with a first order effect on its proxy temperature.

Remark 7: In order to consider (17), the scalar state $x_j^c(t)$ has to be a meaningful descriptor of the temperature of the device or component that is modelled. This is fitting in the considered application where the heat generated by the electrical components is stored in a small portion of space and heat exchange phenomena are predominantly localized in the same volume.

3) Dynamics of the outflow temperature $x_j^o(t)$: Applying the first law of thermodynamics to a control volume containing only the generic node j yields the balance equation

$$q_j^o(t) - q_j^i(t) = -q_j^{cv}(t), \tag{19}$$

where $q_j^{cv}(t)$ is the heat rate in (17) and $q_j^i(t)$ ($q_j^o(t)$, respectively) is the rate at which heat energy, transported by the coolant, enters (exits) node j. Expanding these rates in function of the volumetric flow and the physical properties of the coolant, assuming that the pressures in the cooling circuit are in first approximation constant in time, yields

$$c_p \rho f_j(t) x_j^o(t) = c_p \rho f_j(t) x_j^i(t) - q_j^{cv}(t), \tag{20}$$

where ρ is the density of the coolant and c_p its specific heat capacity at constant pressure. By taking time derivatives and rearranging terms we obtain the dynamics of the outflow temperatures $x_j^o(t)$ as

$$\dot{x}_j^o(t) = \dot{x}_j^i(t) - \frac{\dot{q}_j^{cv}(t)}{c_p \rho f_j(t)} + \frac{q_j^{cv}(t)}{c_p \rho f_j^2(t)} \dot{f}_j(t), \quad f_j(t) \neq 0. \tag{21}$$

(13), (16), (17) and (21) are the salient ingredients of our control-oriented framework for liquid cooling applications. In the following section we show how they can be specialized to model the characteristic thermal networks of data centers at the server level.

III. A library of standard models at the server level

In our liquid cooling framework each node of the thermal network is an instance of the model of Section II. We now specialize this generic description into a library of reusable node models. To this aim, we start by categorizing nodes into thermal nodes and transport nodes (see Figures 1 and 2). *Thermal nodes* participate in the thermal dynamics by exchanging heat with their neighbors and by acting as local heat sources. For instance, heat reservoirs and electrical components such as CPUs are thermal nodes. *Transport nodes*, instead, have an infrastructure character: they do not participate directly in the thermal dynamics but support the liquid cooling operations, for example, by modelling piping manifolds and transport delay nodes.

A. The transport nodes

In practice, the liquid cooling overlay $\mathcal{E}_{cv}$ is implemented using pipes, joints and flow splitters. In our framework, these elements are modelled as *transport nodes*, that is, mathematical constraints describing how the coolant can flow in and out of each manifold and how the flow temperature propagates along the cooling circuits. We stress that transport nodes are a means to describe the topology of the cooling circuit while no heat is absorbed or rejected within these nodes. Therefore

$$\left|\delta_-^{\mathcal{E}_{cd}}(j)\right| = \left|\delta_+^{\mathcal{E}_{cd}}(j)\right| = 0 \tag{22}$$

for all transport nodes $j \in \mathcal{N}$.

1) Delay nodes: Transport delay nodes are fictitious nodes, each associated to an actual pipe interconnect in the liquid cooling circuit, that are employed to account for mass transport delays.

The generic transport delay node j has a single inflow and a single outflow:

$$\left|\delta_-^{\mathcal{E}_{cv}}(j)\right| = 1, \quad \left|\delta_+^{\mathcal{E}_{cv}}(j)\right| = 1. \tag{23}$$

There is then a unique tubular inflow interconnect $(h,j) \in \mathcal{E}_{cv}$ with length $\ell_{(h,j)}$ and constant area section $a_{(h,j)}$. There exist moreover a scalar $\lambda_{hj} \in (0,1]$ indicating the fraction of the coolant that first crosses node h and then flows into node j such that $\varphi_{hj}(t) = \lambda_{hj} f_h(t)$. The temperature dynamics of the inflow is modelled using (16) where the sum is taken over the single edge (h,j). Since j does not exchange heat with the coolant (that is, heat energy is conserved) we omit tracking the node's temperature dynamics (17). We account for the mass transport delay by shifting in time the temperature dynamics of the coolant entering node j

$$\dot{x}_j^i(t) = \dot{x}_h^o(t), \tag{24}$$

$$\dot{x}_j^o(t) = \dot{x}_j^i\Big(t - \tau_j\{t, \varphi_{hj}\}\Big)\left(1 - \frac{\partial}{\partial t}\tau_j\{t, \varphi_{hj}\}\right), \tag{25}$$

where $\tau_j\{t, \varphi_{hj}\}$ is a functional mapping corresponding to the mass transport delay at time t given the past trajectory of the flow rate $t \mapsto \varphi_{hj}(t)$. In other words, it is the transit time of a parcel of coolant along the link $(h,j) \in \mathcal{E}_{cv}$ reaching node j at time t given the history of $\varphi_{hj}(t)$. For

a straight interconnection, the transport delay at time t is formally the root of

$$\tau_j \mapsto a_{(h,j)}\ell_{(h,j)} - \int_{t-\tau_j}^{t} \varphi_{hj}(s)ds, \tag{26}$$

where the current time t, the past flow rate $\varphi_{hj}(s)$, the interconnection length $\ell_{(h,j)}$ and section area $a_{(h,j)}$ are determinate variables while τ_j is indeterminate.

2) Joint nodes: These nodes model actual joints in the liquid cooling circuit. The generic joint node j has multiple inflows and one outflow:

$$\left|\delta_-^{\mathcal{E}_{cv}}(j)\right| \geq 2, \quad \left|\delta_+^{\mathcal{E}_{cv}}(j)\right| = 1. \tag{27}$$

The temperature dynamics of the coolant entering j is given in (16), the dynamics $x_j^c(t)$ is again omitted and, moreover, the entire flow crossing j exits it instantaneously at its input temperature:

$$\dot{x}_j^o(t) = \dot{x}_j^i(t). \tag{28}$$

3) Splitter nodes: Splitter nodes describe constant-ratio flow splitters. The generic splitter node j has one inflow and at least two outflows:

$$\left|\delta_-^{\mathcal{E}_{cv}}(j)\right| = 1, \quad \left|\delta_+^{\mathcal{E}_{cv}}(j)\right| \geq 2. \tag{29}$$

The inflow dynamics is given in (16) with the summation reduced to the single inflow link, the dynamics $x_j^c(t)$ is omitted and the outflow dynamics is the same as that of joint nodes in (28). Finally, splitter nodes are characterized by the splitting ratios Λ defined in (12).

4) Supply nodes: Let $j \in \mathcal{S}$ indicate a generic supply node in the network. j is characterized by the manipulable flow $u_{I(j)}(t)$ and the coolant temperature $x_j^i(t)$ (considered here an exogenous input). Supply nodes have no inflows and one outflow:

$$\left|\delta_-^{\mathcal{E}_{cv}}(j)\right| = 0, \quad \left|\delta_+^{\mathcal{E}_{cv}}(j)\right| = 1. \tag{30}$$

In this case the output and input temperatures coincide as in (28).

5) Collector nodes: A collector node $j \in \mathcal{C}$ has one inflow and no outflows:

$$\left|\delta_-^{\mathcal{E}_{cv}}(j)\right| = 1, \quad \left|\delta_+^{\mathcal{E}_{cv}}(j)\right| = 0. \tag{31}$$

The temperature dynamics of the inflow is (16) while both the dynamics of the node and outflow temperatures are disregarded.

B. The thermal nodes

Nodes endowed with a local temperature state are called thermal nodes. They are specialized in *environmental nodes*, that model heat reservoirs, and *active nodes*, that model self-heating components.

1) Environmental nodes: In our set-up a generic environmental node j is not connected to the liquid cooling circuit:

$$\left|\delta_-^{\mathcal{E}_{cv}}(j)\right| = 0, \quad \left|\delta_+^{\mathcal{E}_{cv}}(j)\right| = 0. \tag{32}$$

In particular the temperature dynamics of inflows and outflows are disregarded. Rather, the node acts as a heat reservoir being connected in the heat conduction overlay $\mathcal{E}_{cd}$:

$$\left|\delta_-^{\mathcal{E}_{cd}}(j)\right| = 0, \quad \left|\delta_+^{\mathcal{E}_{cd}}(j)\right| \geq 1. \tag{33}$$

Environmental nodes are thus assumed to have constant temperature in time,

$$\dot{x}_j^c(t) = 0, \quad x_j^c(0) = \overline{x}_j^c, \tag{34}$$

for some given temperature $\overline{x}_j^c \in \mathbb{R}_{\geq 0}$ of the reservoir.

2) Active nodes: In a blade server the active nodes coincide with the electrical components that necessitate cooling such as CPUs, DIMMs and low-power chips implementing Input/Output (I/O) functions to the memory, to the disks, and to the network. During operation, the generic component converts electrical power into heat which is rejected to the liquid cooling circuit through a surface mounted heat exchanger.

The generic active node j has one inflow and one outflow:

$$\left|\delta_-^{\mathcal{E}_{cv}}(j)\right| = 1, \quad \left|\delta_+^{\mathcal{E}_{cv}}(j)\right| = 1. \tag{35}$$

The temperature dynamics of the inflow and the outflow are given by (16) and (21), respectively. The local temperature dynamics is modelled by specializing (17).

In particular, the dynamical contribution due to thermal conduction is given by (3) while the heat rate $q_j^{cv}(t)$ due to convection is approximated by the following one-dimensional resistive thermal model

$$q_j^{cv}(t) = -\frac{x_j^c(t) - x_j^l(t)}{R_j\big(f_j(t)\big)}, \tag{36}$$

where

- $x_j^l(t) - x_j^c(t)$ is the effective temperature difference between the liquid coolant and the component;
- $f_j \mapsto R_j(f_j)$ is the lumped thermal resistance of the heat exchanger in function of the (volumetric) flow rate f_j.

Finally, self-heating phenomena are modelled by assuming that the node converts all the electrical power that it drains (the time-varying quantity $p_j(t) \in \mathbb{R}_{\geq 0}$) locally into heat.

Here we specifically consider Manifold Micro-Channel (MMC) heat exchangers, a specific liquid cooling technology that has been investigated extensively, both outside [26], [27], [28], [29], [30] and inside data centers [31], [32]. We thus model the flow dependence of the thermal resistance R_j in (36) through the following rational form

$$f_j \mapsto R_j(f_j) \doteq R_j^p + R_j^s + \frac{R_j^b}{f_j}, \tag{37}$$

where R_j^p, R_j^s, R_j^b are positive parameters defining the heat transfer performance of the physical device. The thermal

resistance R_j corresponds then to the series connection of three thermal resistances:

- R_j^p: the thermal resistance given by the component's package and the copper plate at the base of the heat sink;
- R_j^s: the resistance of the heat transfer structure, that is, the channel's fins in a MMC design;
- R_j^b: the bulk resistance between the copper and the liquid coolant.

We stress that the one-dimensional model (37) has been shown to capture accurately the heat exchange profiles of MMC devices and that its parameters can be estimated from first principles [29], [30], [32].

As for the effective coolant temperature $x_j^l(t)$ in (36), in the case of MMCs we propose to use the identity

$$x_j^l(t) = x_j^i(t), \tag{38}$$

namely, we set $x_j^l(t)$ equal to the input coolant temperature as done in [30], [32].

Equations (36), (38) and (37) yield then the convection heat rate

$$q_j^{cv}(t) = -\frac{f_j(t)}{R_j^b + (R_j^p + R_j^s) f_j(t)} \big(x_j^c(t) - x_j^i(t)\big). \tag{39}$$

By inserting (3) and (39) in (17) we obtain the full continuous time temperature dynamics of the node:

$$\begin{aligned} \dot{x}_j^c(t) = &-\frac{f_j(t)}{d_j \left(R_j^b + (R_j^p + R_j^s) f_j(t)\right)} \big(x_j^c(t) - x_j^i(t)\big) + \dots \\ &- \sum_{h \in \delta_-^{\mathcal{E}_{cd}}(j)} \frac{k_{hj}}{d_j} \big(x_j^c(t) - x_h^c(t)\big) + \frac{p_j(t)}{d_j}. \end{aligned} \tag{40}$$

Remark 8: Specializing (40) further to a specific family of vendor chips is outside the scope of this manuscript. We thus disregard the thermal networks that arise within the self-heating component. Instead, we assume that the single temperature $x_j^c(t)$ is a meaningful descriptor of the component's state. We notice, however, that accurate RC thermal networks capturing hot-spots within a single electronic chip have been considered in [33] and that those models find exact equivalents in our framework.

IV. Controlled liquid cooling

We propose a controlled liquid cooling strategy that aims at simultaneously minimizing the supply flow rates (to decrease the actuation costs) while increasing the temperature of the coolant at the outlet (to improve its quality in heat-recovery applications). To this aim, we formulate a polynomial, optimal, control problem in a Receding Horizon Control (RHC) fashion where the cost to minimize is measured by the volume of the coolant flowing through the blade and the constraints follow from the underlying thermal network model.

A. Discretization

We assume a uniform sampling schedule with period Δ. The dynamics of the thermal network is discretized using Euler's forward rule and propagated over a horizon of length H sampling periods. With a slight abuse of notation, we let $x_j^i(k)$ denote the inflow temperature of node j at time $k\Delta$ and adopt the same convention for all the time-varying quantities. All manipulable and exogenous inputs are assumed zero-order held.

We assume time-scale separation of the temperature dynamics of the blade server and that of the storage Coolant Distribution Unit (CDU). In particular, the temperature of the supply inflows $x_j^i(k), j \in \mathcal{S}$, are measured at time k, that is, at the beginning of the receding horizon and assumed to remain constant over it. We assume moreover that the power consumption of the blade's active nodes is unknown, that the computational loads are also unknown, and that they are difficult to forecast. To cope with this minimal-information setting we consider the following worst-case scenario where each component dissipates the highest plausible power:

$$p_j(k) = p_{j,\max}, \quad \forall\, k \geq 0 \tag{41}$$

for all active nodes $j \in \mathcal{N}$. Considering (41) leads then to a feedback law that satisfies the operation constraints over the specified horizon irrespective of the unknown future computational loads. Clearly, given supplementary information one may reformulate the optimization problem accordingly.

B. Static and dynamical constraints

The safe operation of the blade requires to keep the temperature of the main components below established thresholds:

$$x_j^c(k) \leq x_{j,\max}^c, \quad \forall\, k \geq 0 \tag{42}$$

for all active nodes $j \in \mathcal{N}$. Moreover, the supply flow rates must satisfy box constraints of the form

$$\boldsymbol{u}_{\min} \preceq \boldsymbol{u}(k) \preceq \boldsymbol{u}_{\max} \quad \forall\, k \geq 0, \tag{43}$$

with $\boldsymbol{u}_{\min}, \boldsymbol{u}_{\max} \in \mathbb{R}_{\geq 0}^m$.

The continuous time thermal dynamics of the network are approximated under the assumption of piece-wise constant flows, discretized using Euler's forward rule and then rewritten as polynomial constraints. For instance, discretization of (40) yields

$$\begin{aligned} \frac{x_j^c(k+1) - x_j^c(k)}{\Delta} = &\dots \\ &- \frac{f_j(k)\big(x_j^c(k) - x_j^i(k)\big)}{d_j \left(R_j^b + (R_j^p + R_j^s) f_j(k)\right)} \dots \\ &- \sum_{h \in \delta_-^{\mathcal{E}_{cd}}(j)} \frac{k_{hj}}{d_j} \big(x_j^c(k) - x_h^c(k)\big) + \frac{p_j(k)}{d_j}, \end{aligned} \tag{44}$$

which can be rewritten into an equivalent polynomial constraint by multiplying both its left and right hands by the positive affine term $R_j^b + (R_j^p + R_j^s) f_j(k)$.

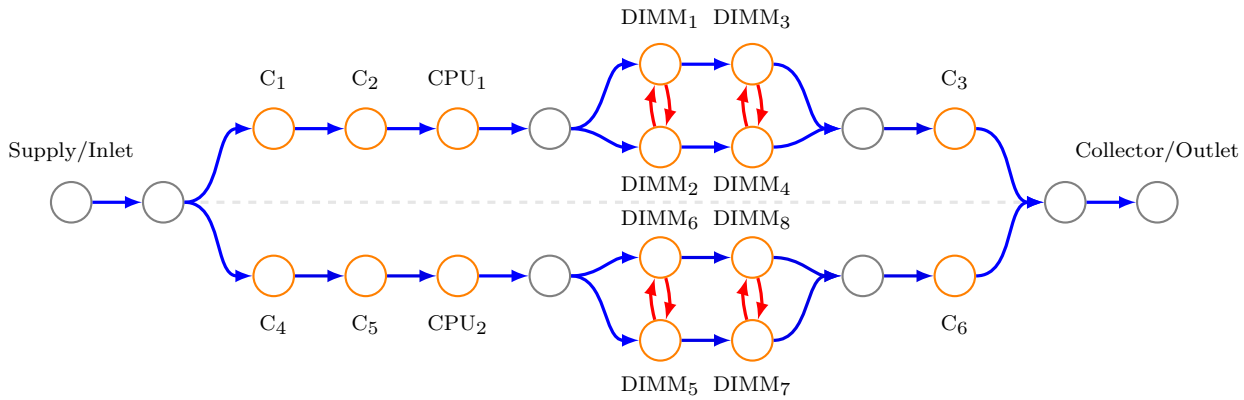

Figure 1: Schematization of a typical thermal network arising in a liquid cooled blade server. This network refers to the blade in Figure 2, with two CPUs, 8 banks of DIMMs organized in four groups and six main companion chips implementing I/O functions to the memories, storage, networking and DC/DC converters. Supply, collector, splitter and joint nodes are drawn in gray while the liquid cooled active nodes are drawn in orange; Links in the heat conduction and heat convection overlays, $\mathcal{E}_{cd}$ and $\mathcal{E}_{cv}$, are drawn in solid red and blue respectively. Notice that each blue interconnection is associated with a corresponding transport delay node not shown explicitly in the scheme.

Figure 2: A typical liquid cooled blade server. The specific platform shown here is the IBM BladeCenter® QS22, it develops on a rectangular base that is 29mm wide, 245mm high and 446mm long and has all electronics with a power consumption above 3 Watts connected in the liquid cooling circuit.

Finally, we notice that solving (26) in our discrete time setting amounts to solving a mixed integer problem. In this work, we thus choose to disregard the exact modelling of the transport delays in discrete time by considering $\tau_j\{k, \varphi_{hj}\} = 0$ in (24) for all transport delay nodes $j \in \mathcal{N}$. This choice greatly simplifies the computational MPC problem at the cost of introducing a degree of conservativeness: Indeed, an increase in the temperature of the coolant is now propagated instantly along the next link in the network. Nevertheless, we consider these effects to be negligible, especially in the more interesting case of higher flow rates (corresponding to higher computational and heat loads) characterized by transport delays with lengths equal to fractions of a second.

C. The cost function

Denote a generic candidate control sequence through

$$\boldsymbol{u}_{k:k+H-1} \doteq \big(\boldsymbol{u}(k), \boldsymbol{u}(k+1), \dots, \boldsymbol{u}(k+H-1)\big). \quad (45)$$

and define the corresponding control cost by

$$\boldsymbol{u}_{k:k+H-1} \mapsto g(\boldsymbol{u}_{k:k+H-1}) \doteq \sum_{z=k+1}^{k+H} \boldsymbol{u}(z). \quad (46)$$

The cost above penalizes the overprovision of the liquid coolant while attaining the auxiliary objective of increasing the coolant temperature (cf. (20),(39)) throughout the circuit and thus its eventual economic value when this higher quality heat is harvested at the outlet [4].

D. The RHC problem formulation

Let g be the cost function in (46), ψ be an opportune vector-field of polynomial constraints obtained by stacking all the static and dynamical polynomial constraints of Section IV-B. For instance, design ψ by first stacking the $2m$ scalar inequalities corresponding to (43), then, for all active components j, consider the static and dynamical constraints corresponding to (42) and (44). Finally, append the inflow and outflow constraints derived by first discretizing (16) and (21) and then reformulating the resulting rational form as a polynomial constraint. Let then $\boldsymbol{X}_0$ be a compatible vector corresponding to the measured state of the thermal network at time k_0. Our control policy aimed at heat recovery corresponds then to solving the polynomial optimal control problem

$$\begin{aligned} &\min_{\boldsymbol{u}_{k_0:k_0+H-1}} g(\boldsymbol{u}_{k_0:k_0+H-1}) \\ &\text{subject to:} \\ &\psi\big(\boldsymbol{x}^i(k_0), \boldsymbol{X}_0, \boldsymbol{u}_{k_0:k_0+H-1}\big) \preceq 0. \end{aligned} \quad (47)$$

Remark 9: Implementations of (47) require: *i)* efficient estimators of the state, and *ii)* an opportune estimate

of the parameters of the whole thermal model. In this manuscript we assume that both a state estimator and a suitable model parameters are available. Designing effective state observers for dealing with systems with incomplete measurements and performing system identification from real data are, nonetheless, two important future directions that require extensive treatment.

V. Numerical experiments

We compare a static provisioning policy where the coolant is supplied to the blade at a constant rate (chosen as the smallest volumetric flow rate that maintains the blade's on-board temperatures below the safety thresholds (42)) against the dynamic MPC policy proposed in (47).

The simulated server blade, inspired by the one shown in Figure 2, implements the thermal network schematized in Figure 1. The underlying thermal models have been built to mimic a dual socket Intel S2600KPF platform used in the experimental study [34]. More in detail, the model describes a symmetric design with two Intel Xeon E5-2697 v3 CPUs, $4+4$ banks of DIMMs (each group of four is dedicated to a specific CPU), $2+2$ companion chips implementing I/O functions to the memories and peripherals and $1+1$ support chips dedicated to power electronics (see Table I). The idle and maximum power consumption of each component have been retrieved from their specifications; As a whole, the board draws from 182 Watts when idle up to a 410 Watts at peak utilization. Additionally, to build a representative model we assumed that the heat-sinks are produced in copper, that their thermal insulance profile (with units $(s \cdot K \cdot cm^2)/J$) is identical to the one measured experimentally in [30], and obtained their total thermal resistance profiles through scaling (by contact area factors) to match the measurements in [34]. Finally, we estimated the volume of the cooling interconnections through graphical inspection.

We analyzed twelve simulation scenarios by crossing three different inlet coolant temperatures - cold water at 20°C degrees Celsius and warm water at 35 and 50 degrees Celsius - with four different power load traces generated by sampling a forecaster with seasonal patterns [35] to address the following relevant and characteristic computational loads:

- P_{idle}: the blade's components dissipate a constant amount of power in time corresponding to their idle/minimum power consumption;
- P_{lo}: the power consumption of each component is time-varying and corresponds to a low-to-medium computational load;
- P_{hi}: each component dissipates a time-varying amount of power corresponding to a medium-to-high computational load;
- $P_{\max}$: each component dissipates a constant amount of power corresponding to their maximum rated power.

Each simulation is run for $T = 3600$ steps with a sampling period of $\Delta = 1$ seconds. We moreover considered a horizon length H of 20 prediction steps and set the temperature thresholds in (42) as $x^c_{j,\max} = 70$ [° C] for all plausible js (see the example realizations in Figure 3). The average electrical power dissipated by the blade in each load scenario is listed in Table II.

For each scenario, we evaluate the performance of the static and dynamic provisioning policies in terms of the following two indexes: *1)* the amount of water coolant supplied to the system, and *2)* the average temperature of the coolant at the blade's outlet. The above indexes are formally defined as

$$\widehat{u} \doteq \frac{1}{u_{\text{nom}}} \sum_{k=0}^{T-1} u^*(k), \qquad \widehat{x} \doteq \frac{1}{T} \sum_{k=0}^{T-1} x^o(k) \tag{48}$$

where:

- u_{nom} is the minimum feasible volumetric flow for the static provisioning law, computed as the solution to an opportune steady-state version of (47) and equal to $u_{\text{nom}} \doteq 0.0339$ liters per second;
- $x^o(k)$ corresponds to the sampled outflow temperature of the blade's single collector node, obtained by integrating the continuous time dynamics in Section II along the given scenario.

Values of $\widehat{u}$ near one (near zero, respectively) indicate that the liquid cooling system is operating with average flow rates that are comparable to (substantially less than) those of the static provisioning policy. As for the temperature index $\widehat{x}$, higher values indicate higher quality heat harvests at the blade server' outlet which in turn relate directly to the economic value of the outflowing coolant in heat-reuse applications [4].

The results are summarized in Figure 4. The upper panel, dedicated to $\widehat{u}$, highlights the savings that can be realized by considering the proposed MPC provisioning policy (47) relative to the coolant supply performance of the static policy. The lower panel, dedicated to $\widehat{x}$, quantifies the benefits of controlled liquid cooling in terms of the temperature of the coolant at the blade outlet. When the blade is idle, the static supply policy overprovisions the coolant leading to unnecessary usage of the coolant, increased actuation cost and the overall decrease of the heat quality index $\widehat{x}$. In this power load scenario, the MPC policy can reduce the average coolant supply rate up to a six-fold factor. Gains, from 10% to 80% can be achieved at a medium-high computational load depending on the coolant inlet temperature. At colder inlet coolant temperatures, the heat-quality index $\widehat{x}$ increases by 10 degrees Celsius when the dynamical policy is used instead of the static one. When the inlet coolant reaches the warm water cooling regime at 50° C, the temperature gains range from 1 to 2.5 degrees Celsius across the power load scenarios $P_{\min}, P_{\text{lo}}, P_{\text{hi}}$. Overall, controlled liquid cooling is effective at drastically reducing the actuation cost in all inlet temperature regimens. Gains in the heat-reuse index $\widehat{x}$ become marginal as the system approaches the higher inlet design temperatures.

Table I: Specifications of the liquid cooled blade adopted in this analysis and in [34].

Platform	Dual sock. Intel ServerBoard S2600KPF
CPUs	Intel Xeon E5-2697 v3
DIMMs	8×8 GiB of DDR4-2133
Chips	4 companion chips for memory, networking and storage; 2 support chips implementing DC/DC converters

Table II: Average electrical power consumption of the simulated blade server in each power consumption scenario.

	P_{idle}	P_{lo}	P_{hi}	P_{max}
Avg. power consumption [W]	182	218.9	334.9	410

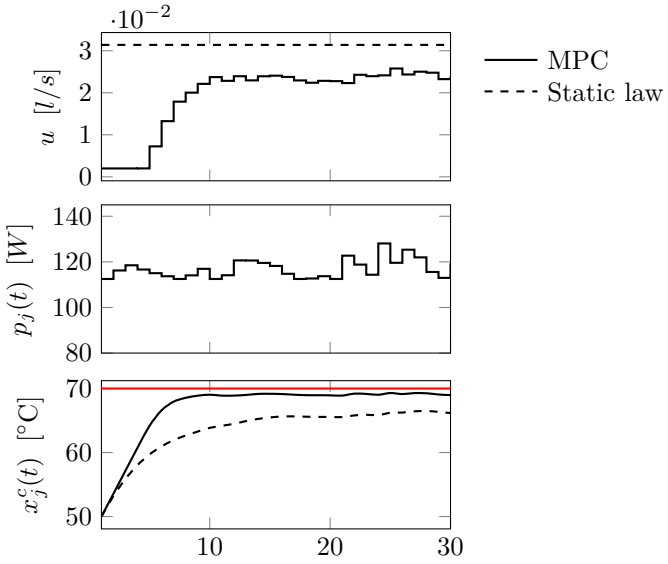

Figure 3: Example trajectories induced by the static (dashed lines) and MPC (solid lines) coolant provisioning strategies. The upper panel shows the volumetric coolant flow at the blade's supply node. The central panel shows the CPU's electrical power consumption along the P_{hi} scenario. The lower panel shows the corresponding trajectories of the package temperature of CPU_1 along the upper temperature threshold $x_j^c, \max = 70\,^\circ\text{C}$. The water inlet temperature is $50\,^\circ\text{C}$.

VI. Conclusions

This manuscript presents a control theoretic approach to energy savings in data centers that deploy direct liquid cooling systems. We have proposed a thermal modelling framework, built on first principles, that approximates the thermal dynamics of liquid cooled blade servers. Moreover, we have shown that the resulting dynamics can be discretized and reformulated into polynomial constraints that are suitable to the design of on-line, dynamic, flow-provisioning controllers. In particular, we have considered an MPC strategy based on the solution of an opportune

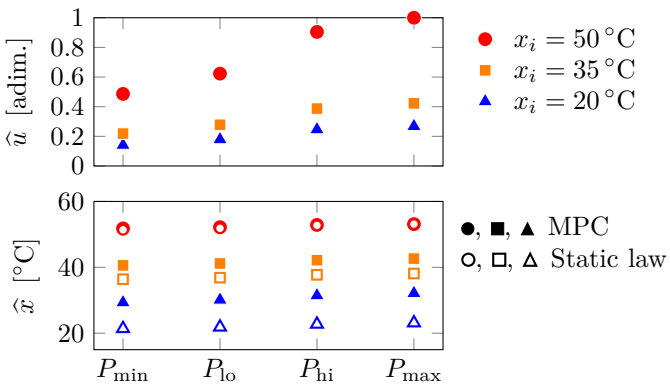

Figure 4: The relative control cost $\widehat{u}$ (upper panel) and the average outlet coolant temperature $\widehat{x}$ (lower panel) evaluated across twelve distinct experimental scenarios (cf. (48)). Each data point corresponds to a different inlet coolant temperature and computational load scenario. In the lower panel, solid lines indicate the performance of the MPC controller and dashed lines those of the static control law. The MPC strategy produces higher quality heat harvests with a lower actuation cost throughout all the considered scenarios.

polynomial optimization problem where the cost is taken to be the amount of coolant that is supplied to the blade over the receding horizon.

Numerical experiments have highlighted the effectiveness of our strategy in both minimizing the control effort and increasing the temperature of the coolant harvested at the blade's outlet, enabling thus higher quality heat at lower pumping costs, and eventually improving the prospective potential of controlled liquid cooling for heat reuse purposes.

Future directions include the design of opportune system identification algorithms exploiting measured data and the implementation and validation of the overall control strategy on real hardware.

References

[1] J. G. Koomey, S. Berard, M. Sanchez, and H. Wong, "Implications of Historical Trends in the Electrical Efficiency of Computing," *IEEE Annals of the History of Computing*, vol. 33, no. 3, 2011.

[2] B. Agostini, M. Fabbri, J. E. Park, L. Wojtan, J. R. Thome, and B. Michel, "State of the Art of High Heat Flux Cooling Technologies," *Heat Transfer Engineering*, vol. 28, no. 4, 2007.

[3] P. Kogge, K. Bergman, S. Borkar, D. Campbell, W. Carson, W. Dally, M. Denneau, P. Franzon, W. Harrod, K. Hill, J. Hiller, M. Richards, A. Scarpelli, S. Scott, A. Snavely, T. Sterling, S. R. Williams, and K. Yelick, "ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems," Tech. Rep., 2008.

[4] S. Zimmermann, I. Meijer, M. K. Tiwari, S. Paredes, B. Michel, and D. Poulikakos, "Aquasar: A hot water cooled data center with direct energy reuse," *Energy*, vol. 43, no. 1, 2012.

[5] A. Capozzoli and G. Primiceri, "Cooling Systems in Data Centers: State of Art and Emerging Technologies," *Energy Procedia*, vol. 83, 2015.

[6] A. C. Kheirabadi and D. Groulx, "Cooling of server electronics: A design review of existing technology," *Applied Thermal Engineering*, vol. 105, 2016.

[7] J. Choi, Y. Kim, A. Sivasubramaniam, J. Srebric, Q. Wang, and J. Lee, "A CFD-based tool for studying temperature in rack-mounted servers," *IEEE Transactions on Computers*, vol. 57, no. 8, 2008.

[8] Z. Wang, C. Bash, N. Tolia, M. Marwah, X. Zhu, and P. Ranganathan, "Optimal fan speed control for thermal management of servers," in *ASME 2009 InterPACK*, 2009.

[9] L. Parolini, B. Sinopoli, B. H. Krogh, and Z. K. Wang, "A cyber-physical systems approach to data center modeling and control for energy efficiency," *Proceedings of the IEEE*, vol. 100, no. 1, 2012.

[10] R. Lucchese, J. Olsson, A.-L. Ljung, W. Garcia-Gabin, and D. Varagnolo, "Energy savings in data centers: A framework for modelling and control of servers' cooling," in *IFAC World Congress*, 2017.

[11] C. D. Patel, "A vision of energy aware computing from chips to data centers," in *International Symposium on Micro-Mechanical Engineering*, 2003.

[12] T. Brunschwiler, B. Smith, E. Ruetsche, and B. Michel, "Toward zero-emission data centers through direct reuse of thermal energy," *IBM Journal of Research and Development*, vol. 53, no. 3, 2009.

[13] M. Iyengar, M. David, P. Parida, V. Kamath, B. Kochuparambil, D. Graybill, M. Schultz, M. Gaynes, R. Simons, R. Schmidt, and T. Chainer, "Extreme energy efficiency using water cooled servers inside a chiller-less data center," *InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITHERM*, 2012.

[14] M. K. Patterson and J. M. Walters, "On Energy Efficiency of Liquid Cooled HPC Datacenters," in *15th IEEE ITHERM Conference*, 2016.

[15] T. green grid, "White paper #70: Liquid cooling technology update," 2016.

[16] T. Brunschwiler, I. G. Meijer, S. Paredes, W. Escher, and B. Michel, "Direct Waste Heat Utilization from Liquid-cooled Supercomputers," in *Proceedings of the 14th International Heat Transfer Conference*, 2010.

[17] J. B. Marcinichen, J. A. Olivier, N. Lamaison, and R. John, "Advances in Electronics Cooling," vol. 7632, 2016.

[18] B. A. Rubenstein, H. Packard, F. Collins, R. Zeighami, R. Lankston, and E. Peterson, "Hybrid cooled data center using above ambient liquid cooling," in *IEEE ITherm*, 2010.

[19] J. B. Marcinichen, J. A. Olivier, and J. R. Thome, "On-chip two-phase cooling of datacenters: Cooling system and energy recovery evaluation," *Applied Thermal Engineering*, vol. 41, 2012.

[20] S. J. Ovaska, R. E. Dragseth, and S. A. Hanssen, "Direct-to-chip liquid cooling for reducing power consumption in a subarctic supercomputer centre," *International Journal of High Performance Computing and Networking*, 2016.

[21] L. Li, W. Zheng, X. Wang, and X. Wang, "Coordinating Liquid and Free Air Cooling with Workload Allocation for Data Center Power Minimization," *11th International Conference on Autonomic Computing*, 2014.

[22] ——, "Placement optimization of liquid-cooled servers for power minimization in data centers," *2014 International Green Computing Conference*, 2015.

[23] T. Brunschwiler, B. Michel, H. Rothuizen, U. Kloter, B. Wunderle, H. Oppermann, and H. Reichl, "Interlayer cooling potential in vertically integrated packages," *Microsystem Technologies*, vol. 15, no. 1, 2009.

[24] A. K. Coskun, J. L. Ayala, D. Atienza, and T. S. Rosing, "Modeling and dynamic management of 3D multicore systems with liquid cooling," in *17th IFIP International Conference on Very Large Scale Integration*, 2011.

[25] A. Coşkun, J. Ayala, D. Atienza, and T. Rosing, "Thermal Modeling and Management of Liquid-Cooled 3D Stacked Architectures," in *VLSI-SoC: Technologies for Systems Integration*, 2011, vol. 360.

[26] D. Copeland, M. Behnia, and W. Nakayama, "Manifold Microchannel Heat Sinks: Isothermal Analysis," *IEEE Transactions on components, packaging and manufacturing technology*, vol. 20, no. 2, 1997.

[27] J. H. Ryu, D. H. Choi, and S. J. Kim, "Three-dimensional numerical optimization of a manifold microchannel heat sink," *International Journal of Heat and Mass Transfer*, vol. 46, no. 9, 2003.

[28] T. Brunschwiler, H. Rothuizen, M. Fabbri, U. Kloter, B. Michel, R. J. Bezama, and G. Natarajan, "Direct liquid jet-impingement cooling with micronsized nozzle array and distributed return architecture," in *Thermomechanical Phenomena in Electronic Systems*, vol. 2006, 2006.

[29] W. Escher, B. Michel, and D. Poulikakos, "A novel high performance, ultra thin heat sink for electronics," *International Journal of Heat and Fluid Flow*, vol. 31, no. 4, 2010.

[30] W. Escher, T. Brunschwiler, B. Michel, and D. Poulikakos, "Experimental Investigation of an Ultrathin Manifold Microchannel Heat Sink for Liquid-Cooled Chips," *Journal of Heat Transfer*, vol. 132, 2010.

[31] R. Wälchli, T. Brunschwiler, B. Michel, and D. Poulikakos, "Combined local microchannel-scale CFD modeling and global chip scale network modeling for electronics cooling design," *International Journal of Heat and Mass Transfer*, vol. 53, 2010.

[32] P. Kasten, S. Zimmermann, M. K. Tiwari, B. Michel, and D. Poulikakos, "Hot water cooled heat sinks for efficient data center cooling: Towards electronic cooling with high exergetic utility," *Frontiers in Heat and Mass Transfer*, vol. 1, no. 2, 2010.

[33] K. Skadron, M. R. Stan, K. Sankaranarayanan, W. Huang, S. Velusamy, and D. Tarjan, "Temperature-Aware Microarchitecture: Modeling and Implementation," *ACM Transactions on Architecture and Code Optimization*, vol. 1, no. 1, 2004.

[34] E. Druzhinin, A. Shmelev, A. Moskovsky, V. Mironov, and A. Semin, "Server Level Liquid Cooling: Do Higher System Temperatures Improve Energy Efficiency?" *Supercomputing frontiers and innovations*, vol. 3, no. 1, 2016.

[35] A. M. De Livera, R. J. Hyndman, and R. D. Snyder, "Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing," *Journal of the American Statistical Association*, 2011.

Appendix

of Proposition 4: Let Λ^s be the n times m matrix encoding the flow splitting ratios at the supply nodes:

$$(\Lambda^s)_{hj} \doteq \begin{cases} \lambda_{jh} & \text{if } j \in \mathcal{S} \\ 0 & \text{otherwise.} \end{cases} \tag{49}$$

Similarly, let Λ^r be the n times n matrix encoding the flow splitting ratios at the remaining $n-m$ nodes:

$$(\Lambda^r)_{hj} \doteq \begin{cases} \lambda_{jh} & \text{if } j \in \mathcal{N} \setminus \mathcal{S} \\ 0 & \text{otherwise.} \end{cases} \tag{50}$$

Notice that, given the control input $\boldsymbol{u}(t)$, the flow vector $\boldsymbol{f}(t)$ in (11) satisfies the constraint

$$\boldsymbol{f}(t) = \Lambda^s \boldsymbol{u}(t) + \Lambda^r \boldsymbol{f}(t). \tag{51}$$

We now proceed to show that Φ in (13) can be evaluated through

$$\Phi = (I - \Lambda^r)^{-1} \Lambda^s, \tag{52}$$

where I denotes the identity matrix acting on $\mathbb{R}^n$. To this aim we only have to show that $(I - \Lambda^r)$ is invertible, or equivalently, that the spectrum of Λ^r does not contain the one. We then notice that

- Λ^r has non-negative entries in $[0, 1]$;
- The columns of Λ^r corresponding to nodes in $\mathcal{S} \cup \mathcal{C}$ are identically zero by definition (49) and since collector nodes have no outflows.

By contradiction, let then $(1, \boldsymbol{f})$ be an eigencouple of Λ^r. By the above reasoning, $f_j = 0$ for all $j \in \mathcal{S} \cup \mathcal{C}$ and there exists moreover a node $\underline{h} \in \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{S})$ such that $f_{\underline{h}} > 0$. By assumption, there exists moreover a path connecting

$\underline{h}$ and the set of collector nodes and thus a corresponding sequence of positive flow split ratios $\lambda_{\underline{h},h_1}, \lambda_{h_1,h_2}, \ldots, \lambda_{\overline{h},c}$ for some $c \in \mathcal{C}$. This implies immediately

$$f_c \geq f_{\underline{h}} \cdot \lambda_{\underline{h},h_1} \cdot \lambda_{h_1,h_2} \cdot \ldots \cdot \lambda_{\overline{h},c}.$$

It follows that $f_{\underline{h}}$ is zero for all plausible $\underline{h}$ and that $\boldsymbol{f}$ is the trivial zero vector reaching a contradiction. ■