



Data Centres Optimization for Energy-Efficient and Environmentally Friendly INternet

Funding scheme: Specific Targeted Research Projects – STREP
Co-funded by the European Commission within the Seventh Framework Programme
Project no. 609140
Strategic objective: FP7-SMARTCITIES-2013 (ICT-2013.6.2)
Start date of project: October 1st, 2013 (36 months duration)



Deliverable D3.2

Water cooling server module

Due date: 31/10/2015
Submission date: 09/11/2015
Deliverable leader: SYNELIXIS
Author list: Kostas Pramataris, Theodore Zahariadis

Dissemination Level

-
- | | |
|-------------------------------------|---|
| <input type="checkbox"/> | PU: Public |
| <input type="checkbox"/> | PP: Restricted to other programme participants (including the Commission Services) |
| <input type="checkbox"/> | RE: Restricted to a group specified by the consortium (including the Commission Services) |
| <input checked="" type="checkbox"/> | CO: Confidential, only for members of the consortium (including the Commission Services) |
-

List of Contributors

Participant	Contributor
SYNELIXIS	K. Pramataris, Th. Zahariadis, C. Tsiopoulou

Amendment History

Version	Date	Partners	Description/Comments
0.1	30/09/2015	SYNELIXIS	Initial draft
0.2	15/10/2015	SYNELIXIS	Slight modifications and additions
1.0	09/11/2015	SYNELIXIS	Added implementation and conclusion sections

Table of Contents

List of Contributors	2
Amendment History	3
Table of Contents	4
Figures Summary	5
Abbreviations	6
Executive Summary	7
1. Introduction	8
2. Architecture	9
3. Detailed Description	10
3.1. <i>First-level water loop</i>	10
3.2. <i>Second-level water loops</i>	11
3.3. <i>Third-level water loops</i>	12
3.4. <i>Control and Monitoring</i>	12
4. Implementation	14
4.1. <i>Heat loads</i>	14
4.2. <i>“Heat” servers</i>	14
4.3. <i>The rack</i>	15
4.4. <i>The controller</i>	15
4.5. <i>Control software</i>	16
4.6. <i>Graphical User Interface (GUI)</i>	17
5. Conclusion	23

Figures Summary

Figure 1. Water-Cooling Subsystem Architecture.....	9
Figure 2. Topology of 1 st level water-loop	11
Figure 3. Topology of 2 nd level water-loops	11
Figure 4. Topology of 3 rd level water-loops	12
Figure 5. Control and monitoring related infrastructure.....	13
Figure 6. Dummy heat-loads.....	14
Figure 7. Dummy heat-loads inside a server case.	14
Figure 8. A server rack.	15
Figure 9. The controller.....	16
Figure 10. Control & Monitoring GUI: Main screen.....	17
Figure 11. Control & Monitoring GUI: Status screen	18
Figure 12. Control & Monitoring GUI: Statistics screen.....	18
Figure 13. Control & Monitoring GUI: Rules screen	19
Figure 14. Control & Monitoring GUI: “Systems” management tab	19
Figure 15. Control & Monitoring GUI: “Gateways – aka controllers” management tab	20
Figure 16. Control & Monitoring GUI: “Sensors” management tab	20
Figure 17. Control & Monitoring GUI: “Services” management tab.....	21
Figure 18. Control & Monitoring GUI: “Automations – aka actuators” management tab	21
Figure 19. Control & Monitoring GUI: “Advanced” management tab.....	22

Abbreviations

CPU	CPU Central Processing Unit
CRAC	Computer Room Air Conditioning units
DC	Data Centre
DoW	Description of Work
GPU	Graphical Processing Unit
GUI	Graphical User Interface
HVAC	Heating, Ventilation and Air Conditioning
PC	Personal Computer
PWM	Pulse Width Modulation
REST	Representational State Transfer
ToC	Table of Contents

Executive Summary

This document describes Deliverable D3.2 “Water cooling server module” of Work Package 3 “Data Centre energy consumption optimization platform (eCOP)” within the FP7-SMARTCITIES-2013 DOLFIN project.

Deliverable D3.2 is an implementation deliverable, meaning that this deliverable is actually the implementation of a water cooling prototype subsystem tailored for the Data Center (DC) server ecosystem. Within Task 3.3, whose outcome is deliverable D3.2, the following sub-tasks have been performed:

- exploration of DC cooling requirements
- identification and evaluation of possible solutions
- implementation and verification of the most prominent solution

The scope of this document, as stated above, is actually to describe this water based, cooling solution for the DC servers’ realm.

1. Introduction

As global demand for data exchange grows, so does the size and load placed on data centres, placing unprecedented constraints on space and power available to the data centre operator. Tightly packed server racks are generating exponentially increasing computer power and most of this is converted into heat, gracefully rejected. These excessive heat loads need to be removed from the computer rooms. Traditional data centres utilize computer room air conditioning units (CRACs) that are continuously driving cooled air through perforated plates in a raised floor to maintain a suitable temperature for all the racks in the room. But this, straightforward, approach wastes a lot of energy (hence money), particularly when the density of a computer room is increased.

DOLFIN aims to study alternatives to the traditional CRACs, especially in terms of power saving. Water, that has enormous thermal capacity when compared to air, is the most prominent means for removing these spatial thermal loads out of the computer rooms.

Water-cooling has been used for several years, even in high-end personal computers (PC). Many, specialized, of-the-self components have been developed for that cause. In these cases, stand alone, small water loops are used for conveying the generated heat energy of the high power components out or in the edge of the PC where the heat is usually expelled in the air through the use of a water-to-air heat exchanger. In the context of DCs, heat loads, need not only to be removed from the inside of the servers and racks but also from inside the computer rooms as well. Thus, the currently adopted water-cooling approaches need to be expanded.

DOLFIN aims at developing an end-to-end water-cooling approach, where heat loads from the CPUs (and other high power components) are transferred out of the servers, out of the racks and finally out of the computer rooms. Furthermore, an additional goal is to reuse this heat energy, which is carried out of the computer room, for heating purposes.

2. Architecture

The water-cooling subsystem that was developed within DOLFIN, follows a three-level hierarchical approach:

- at the lower-level (3rd level), a water-loop is utilized per server for conveying the heat of the high power components (CPUs, GPUs, RAM, hard-disk) out of the server
- at the second level, a water loop is utilized per rack for interconnecting all the 3rd level loops within the rack and thus enabling the heat energy of all the local loops to escape out of the rack.
- lastly, a top level loop is utilized for interconnecting the 2nd level water-loops of all the racks and thus enabling to convey the heat energy of all the racks out of the computer room.

Furthermore, the top-level loop is interconnected to the conventional HVAC system of the DC and thus the heat-energy is re-used for conventional heating purposes (i.e. air and water heating).

The following figure depicts the aforementioned approach.

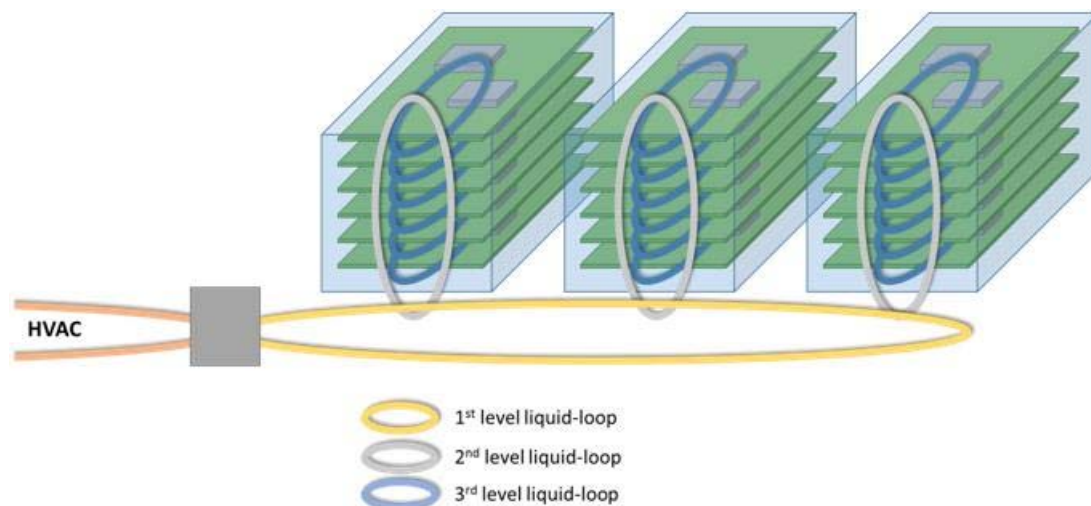


Figure 1. Water-Cooling Subsystem Architecture

In cases of very large and very crowded computer rooms, this hierarchical approach could be expanded to include an extra level of water-loops for interconnecting groups of racks together and thus enabling better and more fair heat dispersal.

3. Detailed Description

In general, a water-loop consists of:

- An appropriate liquid, in most of the cases this is distilled water with appropriate agents that protect against corrosion and build-up of algae.
- Tubes and pipes that are used for implementing the loops and interconnecting all engaged components
- One or more active components that are utilized for the circulation of the water - typically a pump.
- And passive components that are used for transferring heat in and out of the loop. Typically, appropriate components, called water blocks, are used for alleviating the generated heat from high power components into the water loop. Also, water-to-air heat exchangers (radiators) and or water-to-water heat exchangers can be used to convey heat energy in and out of the loop.
- Finally, water reservoirs are usually required to buffer extra water coolant and help remove air-bubbles out of the loop.

Below we describe each of the three types of loops utilized in the DOLFIN cooling system and the way these loops are interconnected to each other.

3.1. First-level water loop

The role of the top level loop is to interconnect the water loops of all the racks of a computer room into a single loop and thus enabling to handle all the generated heat as a whole. In the general case, this top-level loop can be interconnected to the traditional HVAC of the DC and thus enabling the reuse of the generated heat energy. The following figure depicts the topology of the top level loop.

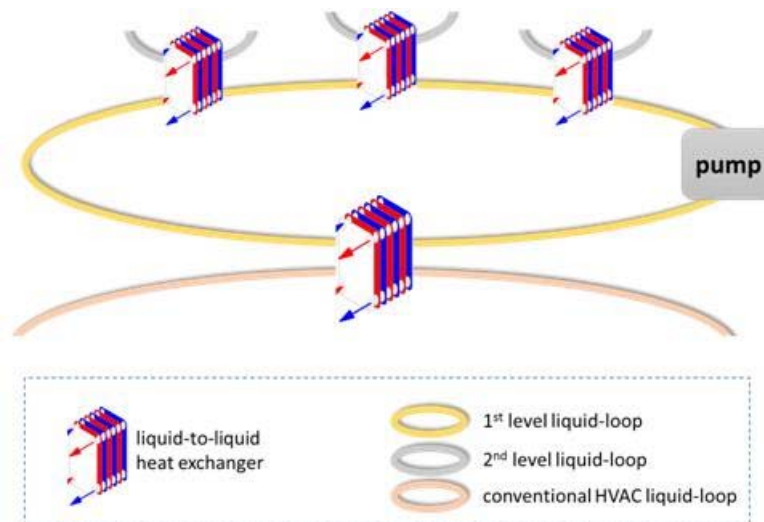


Figure 2. Topology of 1st level water-loop

The heat exchangers utilized for the interconnection of the distinct water-loops are plate-type heat-exchangers.

3.2. Second-level water loops

The second level water-loops are at the rack level, this means that their role is to interconnect all 3rd level water loops that are housed inside the rack into a single loop that in its turn is interconnected into the top level loop. The interconnection of the 3rd level water loops to the 2nd level loop is realized by simple T-joins. The 2nd level water loops also incorporate a reservoir that is used to buffer water coolant for all the water loops of the rack. The following figure depicts the topology of a 2nd level water loop.

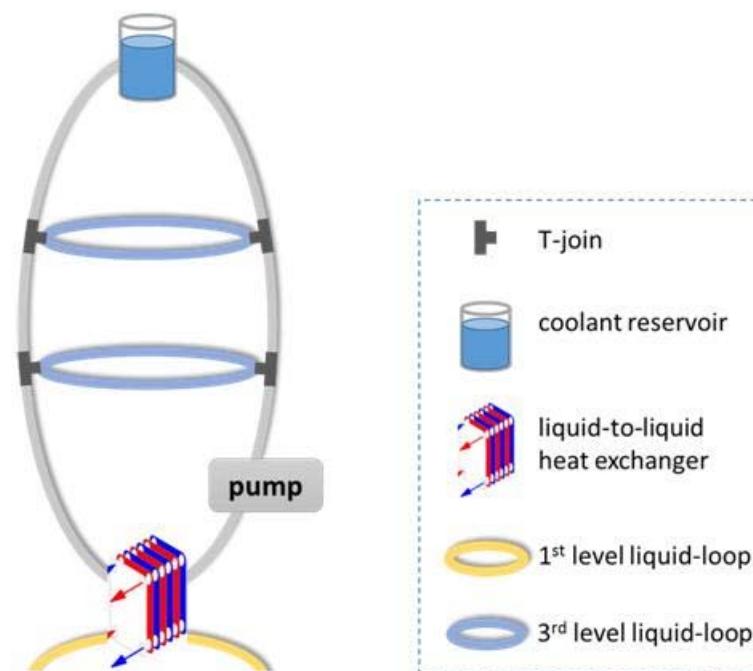


Figure 3. Topology of 2nd level water-loops

3.3. Third-level water loops

Finally, the 3rd level water loops are at the server level. Their use is to convey the heat generated by high power server components out of the server and into the 2nd level water loop of the respective rack. As mentioned in the previous section, these server-level loops are interconnected by means of simple T-joins to the 2nd level water loops of the rack. Except from a low-power pump, these loops also contain appropriate passive components that are attached to the high power components (i.e. CPUs, GPUs, RAM, hard disks) in order to carry away the generated heat from the respective component and into the water loop. Typically, these passive components, called “water blocks” in the water cooling terminology, contain a metal plate that attaches to the surface of the high power component. Depending on the type and socket of the respective component different types of water blocks are required. The following figure depicts the topology of a 3rd level loop.

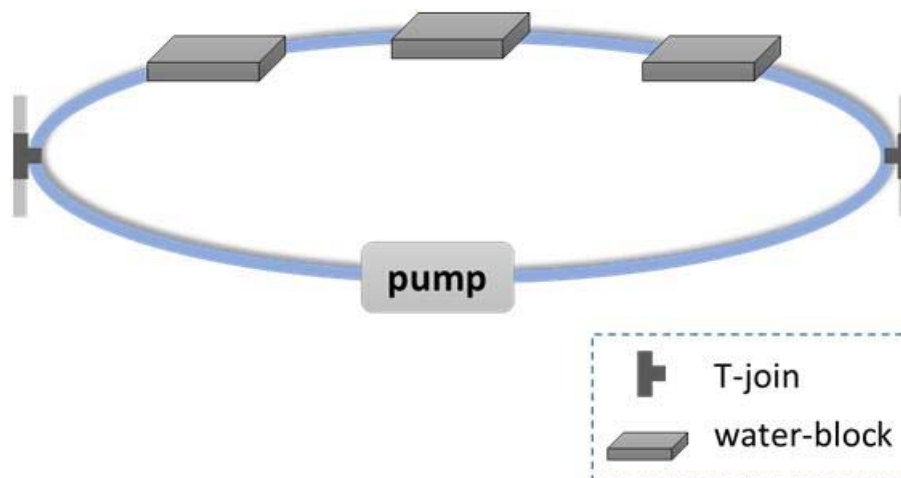


Figure 4. Topology of 3rd level water-loops

3.4. Control and Monitoring

On top of the water-cooling ecosystem, described above, a control and monitor mechanism has been developed that is responsible for the correct operation of the subsystem.

The **monitoring part** of this mechanism, continuously observes the following environmental parameters:

- the coolant temperature of each of the water loops of the subsystem
- optionally, the coolant flow for some/all of the water loops
- the ambient temperature and relative humidity of the computer room

All of the above environmental parameters are logged in a database so that the DC operator except from the current system status can also view graphs and statistics of these parameters.

The **control part**, takes into account the monitored environmental parameters and is responsible for the following tasks:

- Control the pumps of the water cooling subsystem so that heat energy is conveyed effectively and “fairly” from all the end components to out of the computer room. The pumps are controlled through the means of pulse width modulation (PWM). There is a plethora of of-the-self, PWM controlled pumps.

- Control coolant temperatures so that air condensation is avoided for the particular environmental conditions (ambient temperature and relative humidity).
- In case a parameter of the water-cooling subsystem (i.e. the coolant temperature of a water-loop) crosses a predefined threshold, to inform immediately the DC operator (i.e. through email or SMS) of this situation. Optionally, it can also automatically perform corrective actions by itself, for example it can automatically migrate virtual-machines that are hosted to a server that is overheated to other servers with lighter loads.

In order to realize the control and monitoring mechanism a **customized controller** has been developed that is able to sample environmental parameters through the use of specialized sensors and to control a pump through the use of a PWM interface. This controller also incorporates a wireless network interface (IEEE 802.11) that is utilized to interconnect the controller to a server in order to upload sensed data and receive appropriate commands for setting the power level of the water-loop pumps. A proprietary RESTful API has been developed for the interconnection of the controllers with the server. These controllers are incorporated one per each rack.

The **server** can actually be hosted at one of the servers of a rack. The server incorporates a web-server for providing a graphical interface to the DC operator to control and monitor the water cooling ecosystem and also implement a RESTful API that enables the controllers to send and receive data. Furthermore, the server hosts a database that is used for storing historic data for all environmental parameters and other information (i.e. configuration data). Finally, the on the server particular software modules are executed that are responsible for controlling the overall subsystem. The following figure depicts the control and monitoring mechanism of the system.

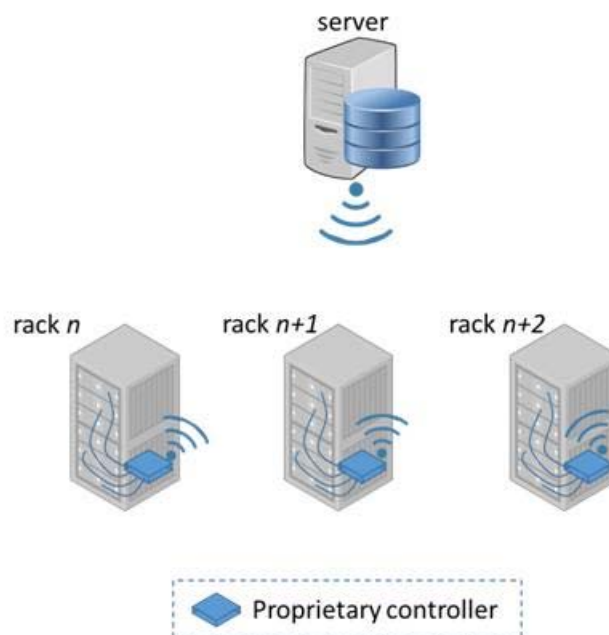


Figure 5. Control and monitoring related infrastructure

4. Implementation

4.1. Heat loads

In order to be able to “heat-up” the whole system, a simple MOSFET-based dummy load has been developed with programmable wattage up to 200W each. This simple load is cooled using regular CPU water blocks and are the basic and only heat load for our test system.



Figure 6. Dummy heat-loads

4.2. “Heat” servers

The heat loads mentioned in the previous section are housed inside a blade server case that provide the power to the loads. The loads are cooled using regular water blocks as can be seen in the following figure.



Figure 7. Dummy heat-loads inside a server case.

4.3. The rack

The heart of the water cooling system is the rack. The rack houses:

- all the servers (in our case the dummy-load servers)
- the water cooling related infrastructure (pumps, heat-exchangers, tubing)
- the controller (along with its sensors)

The following figure shows a basic implementation of the rack that was developed in order to test and validate the overall system.

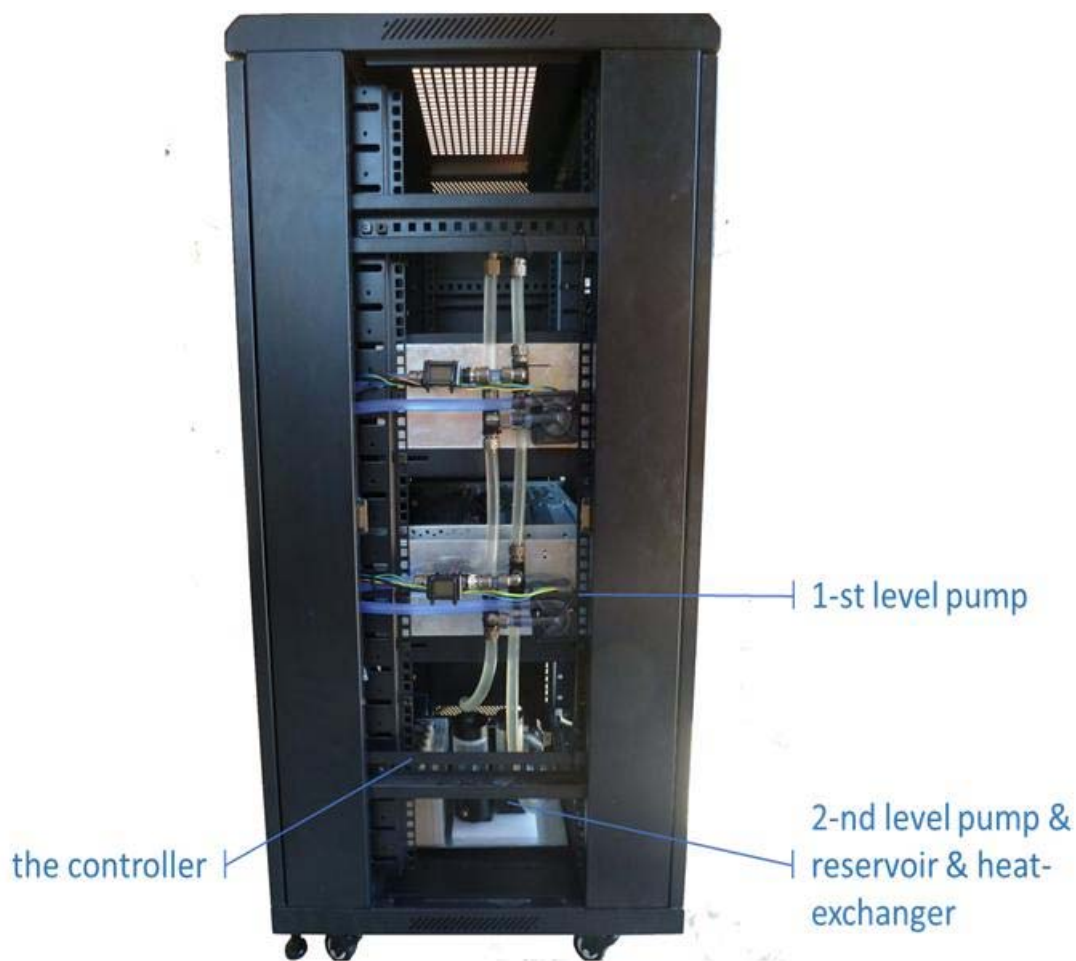


Figure 8. A server rack.

4.4. The controller

The controller is responsible for controlling and monitoring the overall process. The controller monitors the ambient temperature and relative humidity of the computer room along with the temperature (and optionally the flow) of all the water cooling loops. Additionally the controller can

control the flow of each pump through its PWM interface. This way, the system is able to operate in an effective way and remove heat faster from loops with higher heat loads.



Figure 9. The controller.

The controller is based on the ATMEGA1284P microcontroller and has the following features:

- Support of up to 6 analog sensors (i.e. thermocouples)
- Support of I2C sensors
- Control of up to 16 PWM channels
- WiFi connectivity
- Built-in charger and battery
- Micro-SD card support
- Control of up to 4-relays

The main control software however executes outside of the controller (typically in one of the servers) and the role of the controller is to act as a gateway between the control software and the sensors and PWM interfaces of a rack. Specifically, the controller gathers sensor data at regular intervals and transmits them, over a WiFi network, to the server-side control software. In its turn, the control software identifies if a correction action needs to be taken (i.e. alter the speed of a pump) and sends the respective command to the controller.

4.5. Control software

The basis of the control software is a **rule-engine** module that is typically executed at one of the servers. The DC operator specifies “rules” that define the control behaviour of the ecosystem. For example, a simple rule could be that “when coolant temperature of water-loop-X is greater than 50° C, then set PWM interface of pump-X to 75%”, a little more complex rule could be that “the duty-cycle of PWM interface of pump-X is 50% plus two-times the temperature difference of water-loop-X compared to the average temperature of all water-loops of rack-Y”. Except from the control rules, a

DC operator can also specify “alert” rules, for example “when coolant temperature of water-loop-X is greater than 50° C, then send notification email to the following email addresses”.

The rule-engine is executed periodically, observes all defined rules and current system status and generates appropriate commands. This approach enables the DC operator to easily update the control behaviour of the ecosystem.

4.6. Graphical User Interface (GUI)

The GUI of the **control and monitoring module** of the water-cooling subsystem is shown in the following figures. Upon logging into the control and monitoring website, the **main screen** is displayed. The layout of this main screen is shown in the following figure.

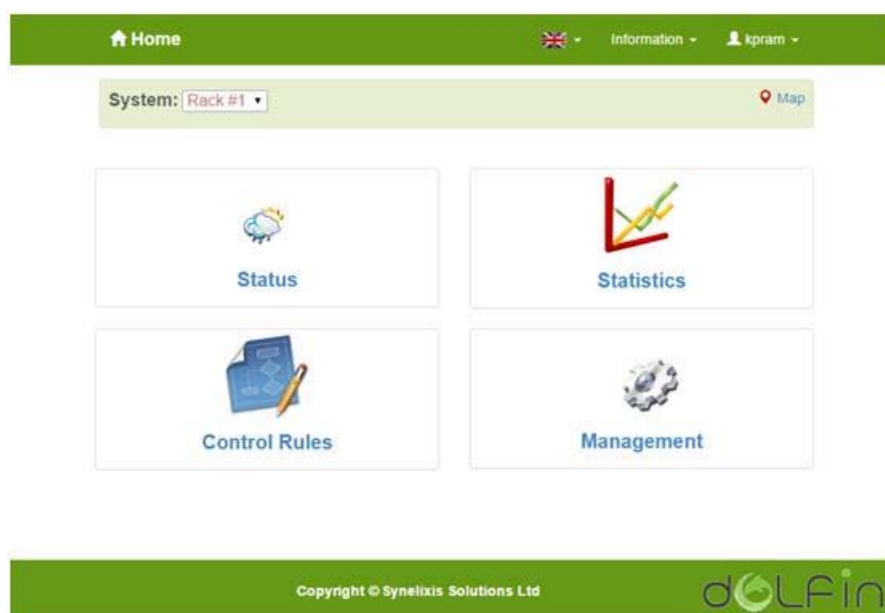


Figure 10. Control & Monitoring GUI: Main screen

Through the main screen, the user can optionally select a particular rack of the DC and can perform the following actions:

- view and alter the current status of the system/rack
- view graphs and statistical data for the various sensed environmental parameters
- observe and define the control and alert rules of the water-cooling ecosystem
- manage and configure the ecosystem

Through the **status screen**, the user can view the current values of the various sensed environmental parameters and also view and set the values of the various active components of the water cooling system (pumps mainly and optionally solenoid valves). The following figure depicts the status screen.

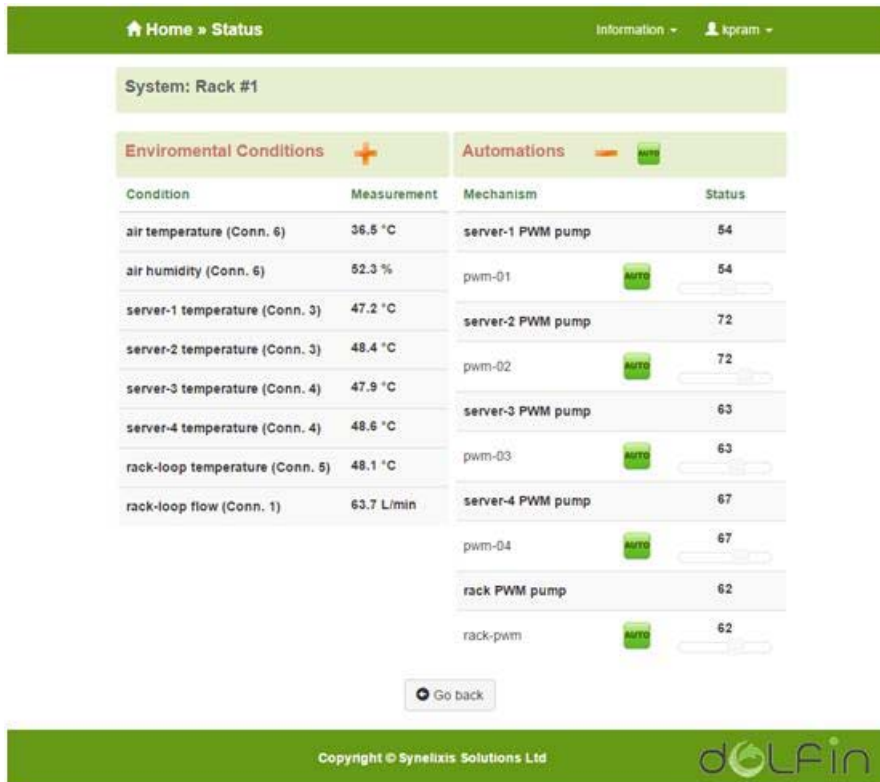


Figure 11. Control & Monitoring GUI: Status screen

The next screen is the **statistics screen** which is used for observing historical data for a specific variable(s) of the system. The following figure depicts the layout of the statistics screen.



Figure 12. Control & Monitoring GUI: Statistics screen

The control-rules screen is used for viewing, adding and altering the control rules of the system.

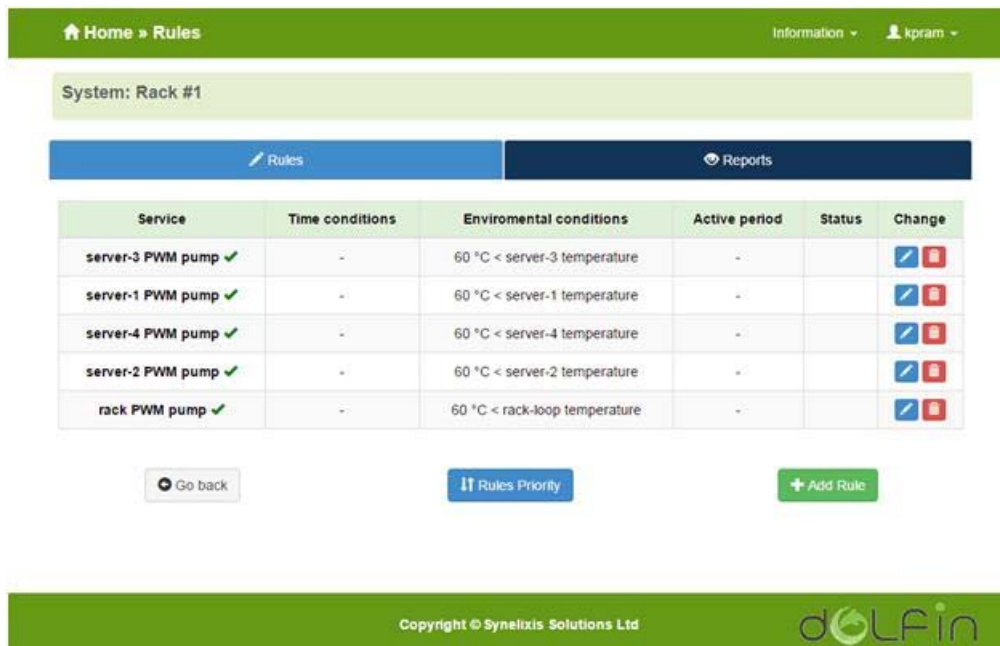


Figure 13. Control & Monitoring GUI: Rules screen

Finally, the **management screen** along with its tabs, allow the user to configure the water-cooling ecosystem. Specifically, it enables the DC operator to:

- define and alter the racks of the DC
- the controller(s) of each rack
- sensors per controller
- actuators (PWM interfaces) per controller

The following figures depict the various management tabs.

The **“systems management”** tab enables the user to define the various systems of the ecosystem – racks in our case.

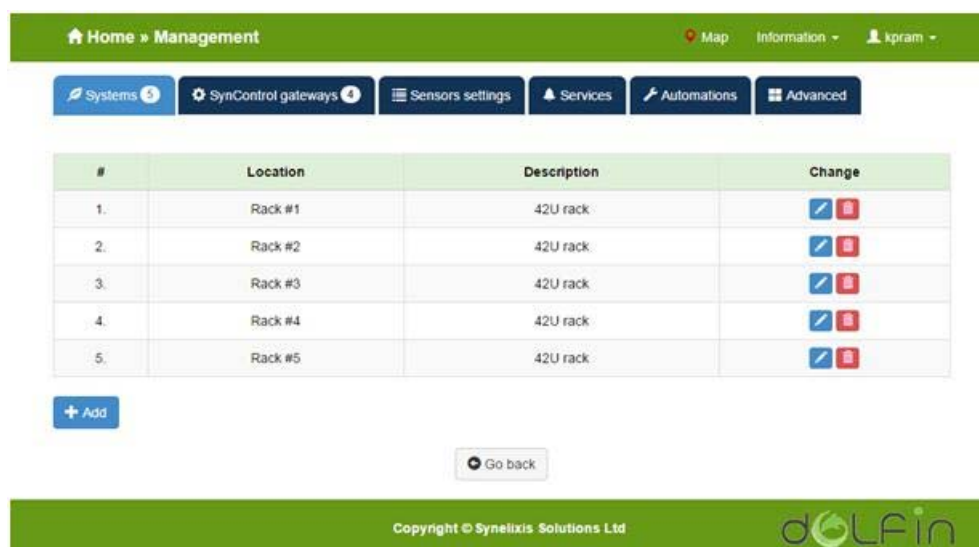


Figure 14. Control & Monitoring GUI: “Systems” management tab

The “**gateways management**” tab enables the user to define the gateways (aka controllers) for each rack.

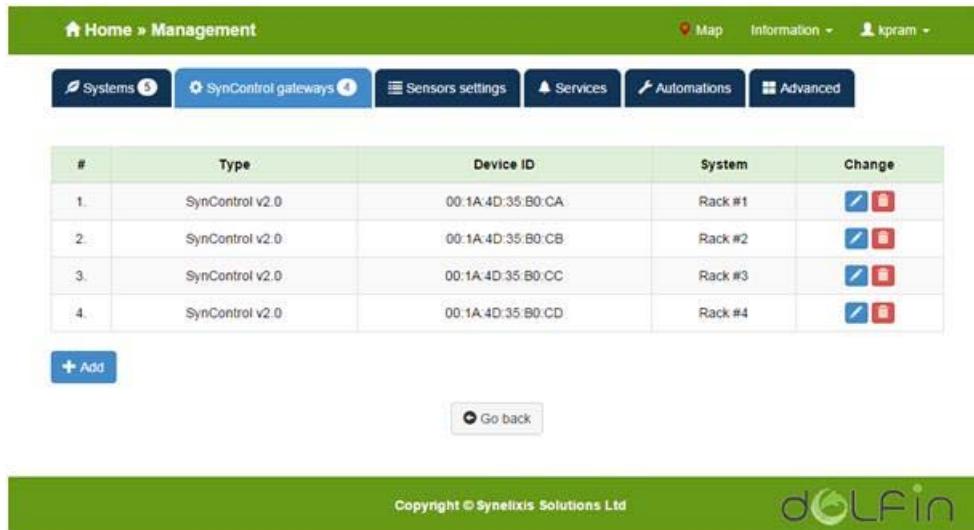


Figure 15. Control & Monitoring GUI: “Gateways – aka controllers” management tab

The “**sensors management**” tab enables the user to define the sensors for each controller.

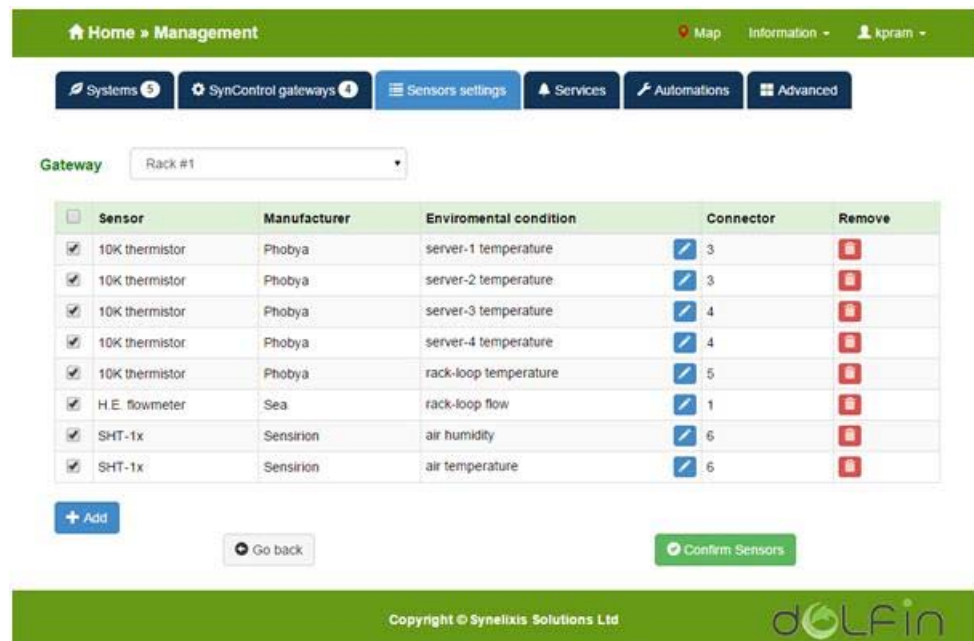


Figure 16. Control & Monitoring GUI: “Sensors” management tab

The “**services management**” tab enables the user to define the various services for each controller. A service can be associated with many actuators so that by setting the state of a particular service you can actually set the state of all the actuators associated with it.

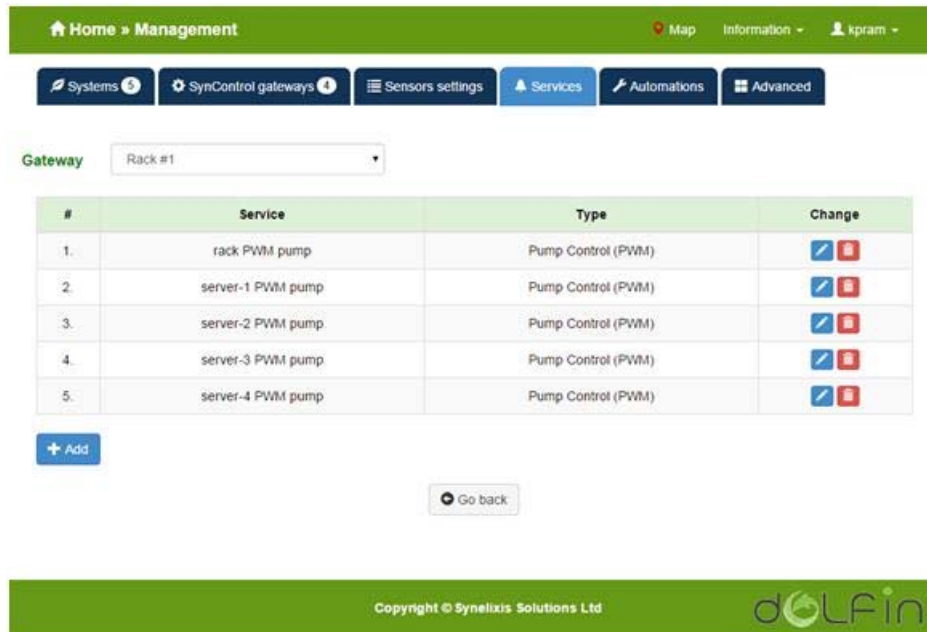


Figure 17. Control & Monitoring GUI: “Services” management tab

The “**automations management**” tab enables the user to define the automations (aka actuators) for each controller. The supported actuators are “PWM-type”, used to control the pumps and “solenoid valves” used to open/close particular water pathways.

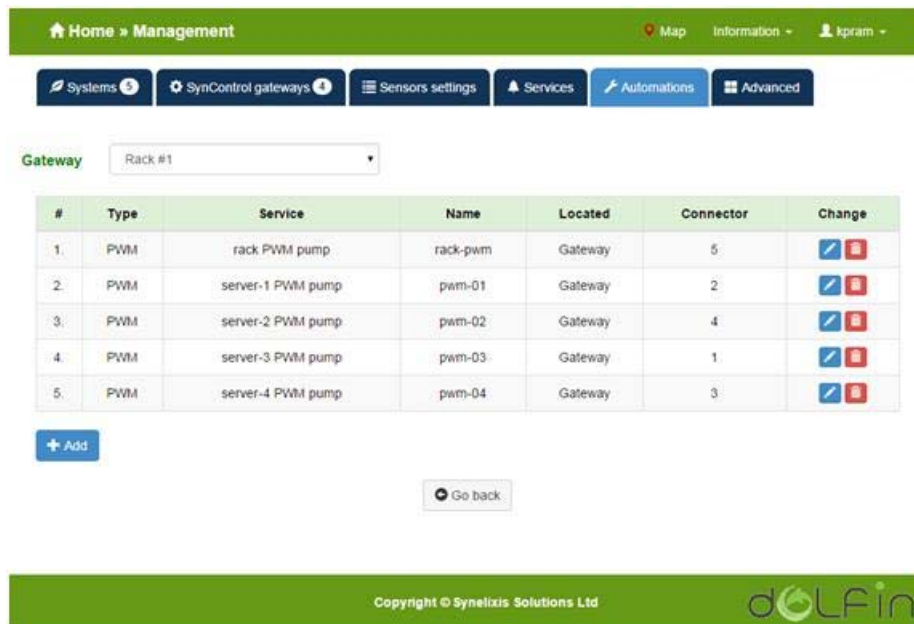


Figure 18. Control & Monitoring GUI: “Automations – aka actuators” management tab

Finally, the “**advanced management**” tab enables the user to define particular parameters of each controller. Specifically, the sensor sample rate, wireless-LAN settings and web-server settings.

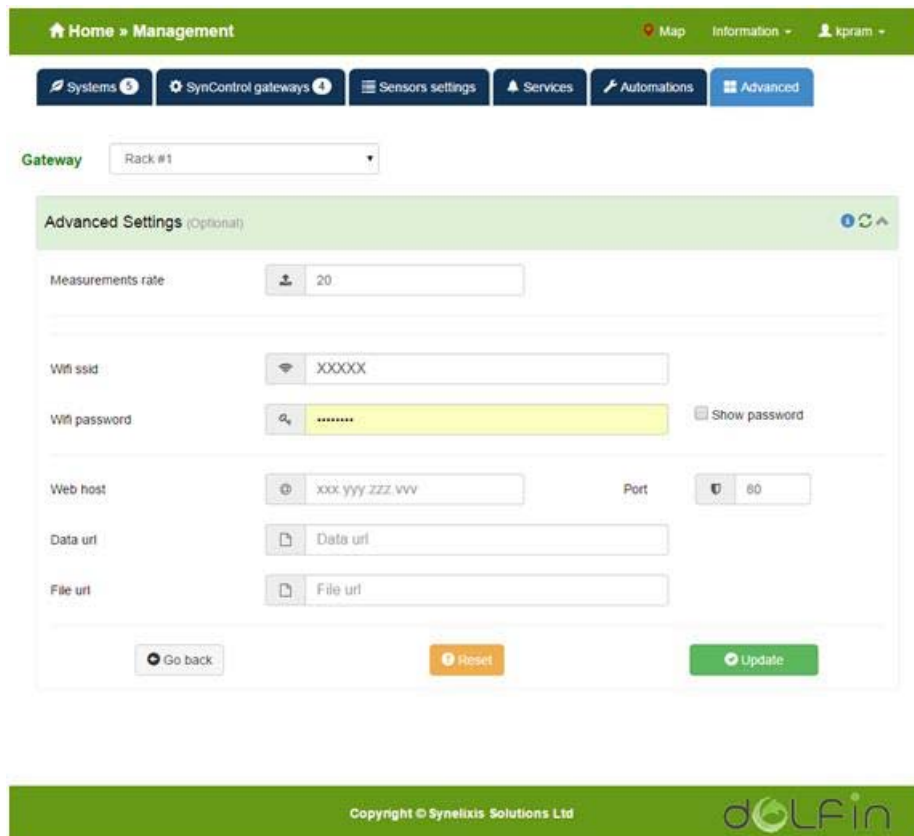


Figure 19. Control & Monitoring GUI: “Advanced” management tab

5. Conclusion

Water based cooling is a highly efficient and a widely used solution for the cooling of high-end PCs and stand-alone servers. But applying this solution in a DC computer room, generates some new hurdles that need to be solved – namely, a distribution scheme for conveying the heat load of all the water cooling loops out of the computer room and an overall control and monitoring system of the water cooling subsystem.

Within the context of Task 3.3 we have identified, implemented and validated a water-based cooling solution that can be applied for the cooling requirements of today's DCs. The solution utilizes of-the-self components (water blocks, pumps, heat-exchangers, etc.) for implementing the water cooling infrastructure and proposes a hierarchical approach for the water cooling interconnection and a control and monitoring system based on a central software module and distributed controllers.

A test system has been built for validating the solution and rigorous tests have been performed. The results were very encouraging and a bigger scale validation needs to be performed.