High-Performance Reconfigurable Computer Systems with Immersion Cooling

Ilya Levin, Alexey Dordopulo(✉), Alexander Fedorov, and Yuriy Doronchenko

Scientific Research Center of Supercomputers and Neurocomputers (LLC),
106, Italyanskiy alley, Taganrog, Russia 347900
{levin,doronchenko}@superevm.ru, scorpio@mvs.tsure.ru, ss24@mail.ru

Abstract. In the paper, we review the design principles and architecture of reconfigurable computer systems with immersion cooling. We prove that systems with immersion cooling are the most promising for the design of high-performance computer complexes. We give selection criteria and design results for the principal components of the immersion cooling system. We demonstrate the design of our computational module prototype, based on advanced Xilinx UltraScale FPGAs and give testing results for the principal technical solutions. We prove that the designed immersion cooling system has a high power efficiency and power reserve for designing advanced reconfigurable computer systems on the basis of new UltraScale+ FPGAs and other next-generation FPGAs. We suggest new design solutions for the case of our computational module, as well as for the layout of the main computational board and other components of the computational module for use of Xilinx UltraScale+ FPGAs.

Keywords: Immersion cooling system · Liquid cooling
Reconfigurable computer systems · FPGAs
High-performance computer systems · Energy efficiency

1 Introduction

Having considerable advantages in real performance and energetic efficiency in comparison with cluster-like multiprocessor computer systems, reconfigurable computer systems (RCS) containing an FPGA computational field of large logic capacity are used for the implementation of computationally laborious tasks from various domains of science and technique. An RCS provides adaptation of its architecture to the structure of any task. In this case, a special-purpose computer device is created. It hardwarily implements all the computational operations of the information graph of the task with the minimum delays. Here we have a contradiction between the implementation of the special-purpose device and its general-purpose use for solving tasks from various problem areas. It is possible to solve this contradiction by combining the creation of a special-purpose computer

L. Sokolinsky and M. Zymbler (Eds.): PCT 2018, CCIS 910, pp. 62–76, 2018.
https://doi.org/10.1007/978-3-319-99673-8_5

device with a wide range of solvable tasks within a concept of reconfigurable computer systems based on FPGAs which are used as principal computational resource [1].

A practical experience of maintenance of large RCS-based computer complexes proves that air cooling systems have reached their heat limit. The continuous increase of both the circuit complexity and the clock rate of each new FPGA family leads to a considerable growth of the power consumption and maximal operating temperature of the chip. So, for the XC6VLX240T-1FFG1759C FPGAs of a computational module (CM) Rigel-2, the maximum overheat of the FPGAs relative to an environment temperature of 25 °C in operating mode, and with a power of 1255 W consumed by the CM, is 33.1 °C, i.e. the maximum temperature of the FPGA chip in the CM Rigel-2 is 58.1 °C. For the XC7VX485T-1FFG1761C FPGAs of the CM Taygeta, the maximum overheat of the FPGAs relative to an environment temperature of 25 °C in operating mode, and with a power of 1661 W consumed by the CM, is 47.9 °C, i.e. the maximum temperature of the FPGA of the CM Taygeta is 72.9 °C. If we take into account that the permissible temperature of an FPGA functioning, providing high reliability of the equipment during a long operation period, is 65...70 °C, then it is evident that the CM Taygeta maintenance requires a decrease in environment temperature.

According to the obtained experimental data, the conversion from the FPGA family Virtex-6 to the next family, Virtex-7, leads to an increase of the FPGA maximum temperature by 11...15 °C. Thus, further development of FPGA production technologies and conversion to the next FPGA family, Virtex UltraScale (with a power consumption of up to 100 W for each chip), will lead to an additional increase in FPGA overheat by 10...15 °C. This will shift the range of their operating temperature limit (80...85 °C), which has a negative influence on their reliability when the workload on the chips reaches up to 85–95% of the available hardware resource. This circumstance requires a quite different cooling method which provides for keeping the performance growth rates of advanced RCS.

2 Liquid Cooling Systems for Reconfigurable Computer Systems

The development of computer technologies leads to the design of computer technique providing higher performance and, hence, more heat. Dissipation of released heat is provided by a system of electronic element cooling which transfers heat from the more heated object (the cooled object) to the less heated one (the cooling system). If the cooled object is constantly heated, then the temperature of the cooling system grows and, for some period of time, will be equal to the temperature of the cooled object. So heat transfer stops and the cooled object will get overheated. The cooling system is protected from overheat with the help of a cooling medium (a heat-transfer agent). Cooling efficiency of the heat-transfer agent is characterized by the heat capacity and heat dissipation. As a rule, heat transfer is based either on the principles of heat conduction, which requires a physical contact of the heat-transfer agent with the cooled object, or on the principles of convective heat exchange with the heat-transfer agent, which consists in the physical transfer of the freely circulating heat-transfer agent. To organize heat transfer to the heat-transfer agent, it is necessary to provide heat contact between the cooling system and the heat-transfer agent. Various *radiators* – facilities for heat dissipation in the heat-transfer agent are used for this purpose. Radiators are set on the most heated components of computer systems. To increase efficiency of heat transfer from an electronic component to a radiator, a *heat interface* is set between them. The heat interface is a layer of heat-conducing medium (usually multicomponent) between the cooled surface and the heat dissipating facility, used for reduction of heat resistance between two contacting surfaces. Modern processors and FPGAs need cooling facilities with as low as possible heat resistance, because at present even the most advanced radiators and heat interfaces cannot provide necessary cooling if an air cooling system is used.

To organize heat transfer to the heat-transfer agent, it is necessary to provide heat contact between the cooling system and the heat-transfer agent. Various *heat-sinks*-facilities for heat dissipation in the heat-transfer agent – are used for this purpose. Heat-sinks are set on the most heated components of computer systems. To increase the efficiency of heat transfer from an electronic component to a heat-sink, a *heat interface* is set between them. The heat interface is a layer of heat-conducing medium (usually multicomponent) between the cooled surface and the heat dissipating facility, used for reduction of heat resistance between two contacting surfaces. Modern processors and FPGAs need cooling facilities with as low as possible heat resistance, because at present even the most advanced heat-sinks and heat interfaces cannot provide the required cooling if an air cooling system is used.

Before 2013, air cooling systems were used quite successfully for cooling supercomputers. But due to a growth of performance and circuit complexity of microprocessors and FGAs, used as components in supercomputer systems, air cooling systems have practically reached their limits for advanced supercomputers of that time, including hybrid computer systems. The majority of vendors of computer technique therefore consider liquid cooling systems as an alternative solution to the cooling problem. Today, liquid cooling systems constitute the most promising design area for cooling modern intensively operating electronic components in computer systems.

A considerable advantage of all liquid cooling systems is the heat capacity of liquids, which is better than that of air (from 1500 to 4000 times), and a higher heat-transfer coefficient (up to 100 times higher). To cool one modern FPGA chip, 1 m^3 of air or 0.00025 m^3 (250 ml) of water per minute is required. Much less electric energy is required to transfer 250 ml of water than to transfer 1 m^3 of air. Heat flow, transferred by similar surfaces at the conventional velocity of the heat-transfer agent, is 70 times more intensive in the case of liquid cooling than in the case of air cooling. An additional advantage is the use of traditional, rather reliable and cheap, components such as pumps, heat exchangers, valves, control devices, etc. For corporations and companies dealing with equipment with high packing density of components operating at high temperatures, liquid cooling is in fact the only possible solution to the problem of cooling modern computer systems. One more option to increase liquid cooling efficiency consists in improving the initial parameters of the heat-transfer agent: increasing velocity, decreasing temperature, creating turbulent flow, increasing heat capacity, reducing viscosity.

The heat-transfer agent in liquid cooling systems used in computer technique is a liquid such as water, or any dielectric liquid. Heated electronic components transfer heat to the permanently circulating heat-transfer agent - a liquid which, after being cooled in the external heat exchanger, is used again for cooling heated electronic components. There are several types of liquid cooling systems. In closed-loop liquid cooling systems, there is no direct contact between liquid and electronic components of the printed circuit boards [6,7]. In open-loop cooling systems (liquid immersion cooling systems), electronic components are immersed directly into the cooling liquid [8,9]. Each type of liquid cooling systems has its own advantages and disadvantages.

In closed-loop liquid cooling systems all heat-generating elements of the printed circuit board are enclosed by one or several flat plates with a channel for liquid pumping [10,11]. So, for example, the cooling system of the SKIF-Avrora supercomputer [12] is based on the principle “one cooling plate, one printed circuit board”. The plate, of course, has a complex surface relief to provide tight heat contact with each chip. In the IBM Aquasar supercomputer, cooling is based on the principle“one cooling plate, one (heated) chip”. In each case, the channels of the plates are joined by manifolds into a single loop connected to a common heat-sink (or another heat exchanger), usually placed outside the computer case and/or rack, or even the computer room. With the help of the pump, the heat-transfer agent is pumped through the plates and dissipates, by means of the heat exchanger, the heat generated by the computational elements. In such systems, it is necessary to provide access of the heat-transfer agent to each heat-generating element of the calculator, which means a rather complex “piping system” and a large number of pressure-tight connections. Besides, if it is necessary to provide maintenance of the printed circuit boards without any significant demounting, then the cooling system must be equipped with special liquid connectors providing pressure-tight connections and simple mounting/demounting of the system.

In closed loop liquid cooling systems all heat-generating elements of the printed circuit board are covered by one or several flat plates with a channel for liquid pumping. So, for example, cooling of a supercomputer SKIF-Avrora is based on a principle “one cooling plate for one printed circuit board”. The plate, of course, had a complex surface relief to provide tight heat contact with each chip. Cooling of a supercomputer IBM Aquasar is based on a principle “one cooling plate for one (heated) chip”. In each case the channels of the plates are united by collectors into a single loop connected to a common radiator (or another heat exchanger), usually placed outside the computer case and/or rack or even the computer room. With the help of the pump the heat transfer agent is pumped through the plates and dissipates heat, generated by the computational elements, by means of the heat exchanger. In such system it is necessary to provide access of the heat transfer agent to each heat-generating element of the calculator, what means a rather complex “piping system” and a large number of pressure-tight connections. Besides, if it is necessary to provide maintenance of the printed circuit boards without any serious demounting, then the cooling system must be equipped with special liquid connectors which provide pressure-tight connections and simple mounting/demounting of the system.

In closed-loop liquid cooling systems, it is possible to use water or glycol solutions as the heat-transfer agent. However, leak of the heat-transfer agent can lead to possible ingress of electrically conducting liquid to unprotected contacts of the printed circuit boards of the cooled computer, and this, in its turn, can be fatal for both separate electronic components and the whole computer system. To eliminate failures, the whole complex must be stopped, and the power supply system must be tested and dried up. The control and monitoring systems of such computers always contain many internal humidity and leak sensors. Cooling systems with liquid at negative pressure are frequently used to solve the leak problem. In these systems, water is not pumped in under pressure but instead is pumped out, thus practically excluding leaks of liquid. If the air-tightness of the cooling systems is damaged, then air enters the system but no leak of liquid happens. Special sensors are used for detection of leaks, while modular design allows maintenance without stoppages of the whole system. However, all these capabilities considerably complicate the design of the hydraulic system.

Another issue affecting closed-loop liquid cooling systems is the dew point problem. In the section of data processing, the air is in contact with the cooling plates. It means that if some parts of these plates are too cold and the air in the section of data processing is warmer and not very dry, then moisture can condense out of the air on the plates. The consequences of this process are similar to leaks. This problem can be solved either by hot-water cooling, which is not effective, or by controlling and keeping at the required level the temperature and humidity parameters of the air in the section of data processing, which is complicated and expensive.

The design becomes even more complicated when it is necessary to cool several components with a water flow that should be proportional to their heat generation. In addition to branched pipes, it is necessary to use complex control devices (simple T-branches and four-ways are not enough). An alternative approach is to use an industrial device with flow control, but in this case, the user cannot considerably change the configuration of the cooled computational modules.

In open-loop liquid cooling systems, the principal component is the heat-transfer agent, which is a dielectric liquid based, as a rule, on a white mineral oil that provides a much higher heat storage capacity than the air does in the same volume. According to their design, these systems contain printed circuit boards, servers of computational equipment, and a bath that is filled with heat-transfer liquid and placed into a computer rack. The heat generated by the electronic components is dissipated by the heat-transfer agent, which flows within the whole bath. We can mention here some advantages of immersion liquid cooling systems: simple design and capability of adaptation to the changing geometry of printed circuit boards, simplicity of mani-folds and liquid connectors, no problems with control of liquid flows, no dew point problem, high reliability and low cost of the product.

The main problem of open-loop liquid cooling systems is the chemical composition of the used heat-transfer liquid which must fulfil strict requirements of heat transfer capacity, electrical conduction, viscosity, toxicity, fire safety, stability of the main parameters and reasonable cost.

Considering the given advantages and disadvantages of the two types of liquid cooling systems, we can affirm that open-loop cooling systems for electronic components of computer systems have more weighty advantages. In this connection, when dealing with advanced RCS, it is reasonable to use direct immersion of heat-generating system components into a mineral oil-based liquid heat-transfer agent.

At present, the technology for liquid cooling of servers and separate computational modules is being developed by many vendors, and some of them have achieved success in this direction [9–11]. These technologies, however, are intended for cooling computational modules containing only one or two microprocessors. All attempts to adapt this technology for cooling computational modules, which contain a large number of heat-generating components (an FPGA field of eight chips), failed due to a number of shortcomings.

The main disadvantages of existing technologies of immersion liquid cooling [10–14] for computational modules containing FPGA computational fields are:

- poor adaptation of the cooling system for placement into standard computer racks;
- inefficiency of cooling of electronic component chips with considerable (over 50 W) heat generation;
- the thermal paste between FPGA chips and heat-sinks is washed out during long-term maintenance;
- the system of cooling-liquid circulation inside the module is designed for one or two chips but not for an FPGA field, and this fact leads to considerable thermal gradients.

In the systems based on the IMMERS technology [9], all cooling liquid is circulating within a closed loop through the chiller, and this fact leads to some problems:

- complex maintenance stoppages are necessary to remove separate components and devices;
- it is necessary to use a power specialized pump and hydraulic equipment adapted to the cooling liquid;
- a complex system for the control of cooling-liquid circulation, which causes periodic failures;
- high cost of the cooling liquid, produced by only one manufacturer.

These disadvantages can be considered as an inseparable part of other existing open-loop liquid cooling systems since the cooling of RCS computational modules containing not less than eight FPGA chips has some specific features compared with the cooling of a single microprocessor.

The special feature of the RCS produced at the Scientific Research Center of Supercomputers and Neurocomputers is the number of FPGAs, which is not less than six to eight chips on each printed circuit board, and high packing density. This considerably increases the number of heat-generating components compared with microprocessor modules, making more complicated the application of the IMMERS direct liquid cooling technology along with other end solutions of immersion systems, and requires additional technical and design solutions for an effective cooling of RCS computational modules.

The use of open liquid cooling systems is efficient owing to the heat-transfer agent characteristics and the design and specification of the used FPGA heat-sinks, pump equipment, and heat-exchangers.

The heat-transfer agent must have the best possible dielectric strength, high heat transfer capacity, the maximum possible heat capacity, and low viscosity.

The heat-sink must have the maximum possible surface of heat dissipation, must allow the circulation of the heat-transfer agent turbulent flow through itself, and manufacturability. Specialists at SRC SC & NC have performed heat engineering research and suggested a fundamentally new design of a heat-sink with original solder pins which create a local turbulent flow of the heat-transfer agent. The used thermal interface cannot be deteriorated or washed out by the heat-transfer agent. Its coefficient of heat conductivity can remain permanently high. SRC SC & NC specialists have created an effective thermal interface that fulfills all specified requirements, and additionally, its coating and removal technology has also been improved.

The pumping equipment, that is to say, is not the least of the components of a CM cooling system. The principal criteria that must be met are the following:

- performance parameters;
- overall dimensions and coordinated placement of the input and the output fittings;
- the pump must be suitable for interaction with oil products with a specified viscosity and chemical composition;
- continuous maintenance mode;
- minimal vibrations;
- the pump must have the minimal permissible positive suction head;
- the protection class of the pump electric motor must be not less than IP-55.

The heat exchanger is also an important component of the cooling system. Its design must be compact and must provide an efficient heat exchange. Research performed by the SRC SC & NC scientific team has proved that the most suitable design of the heat exchanger is a plate-type one designed for cooling mineral oil in hydraulic systems of industrial equipment.

The liquid cooling system must have a control subsystem containing sensors of level, flow, and temperature of the heat-transfer agent, and a temperature sensor for cooling components.

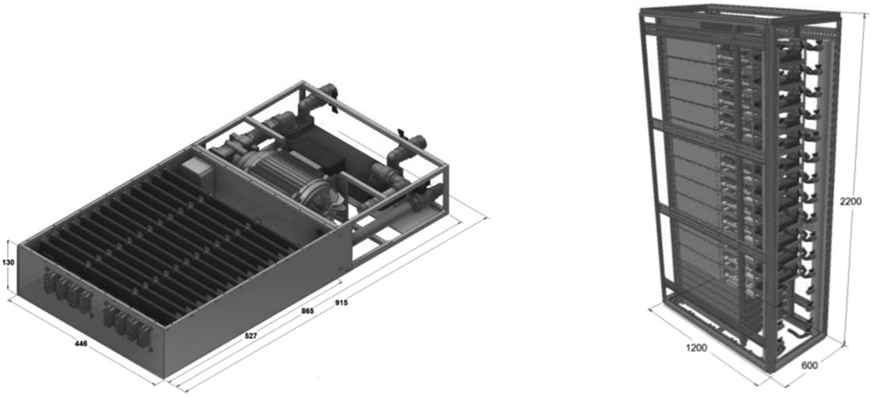
3 "SKAT" Reconfigurable Computer System Based on Xilinx UltraScale FPGAs

The SRC SC & NC scientific team has actively developed since 2013 the creation of next-generation RCS on the basis of their original liquid cooling system for computational circuit boards with high packing density and large number of heat-generating electronic components. The design criteria of computational modules (CM) of next-generation RCS with an open-loop liquid cooling system are based on the following principles:

- the RCS configuration is based on a computational module with a 3U height and 19 width, and self-contained circulation of the cooling liquid;
- one computational module can contain 12 to 16 computational circuit boards (CCB) with FPGA chips;
- each CCB must contain up to eight FPGAs, with a dissipating heat flow of about 100 W from each FPGA;
- a standard water cooling system based on industrial chillers must be used for cooling the liquid.

The principal element in the modular implementation of an open-loop immersion liquid cooling system for electronic components of computer systems is a new generation reconfigurable computational module (see design in Fig. 1-a). The new-generation CM casing consists of a computational section and a heat-exchange section. The casing, which is the base of the computational section, contains a hermetic container with dielectric cooling liquid, and electronic components with elements that generate heat during operation. The electronic components can be computational modules (not less than 12 to 16), control boards, RAM, power supply blocks, storage devices, daughter boards, etc. The computational section is closed with a cover.

The computational section adjoins the heat exchange section which contains a pump and a heat exchanger. The pump moves the heat-transfer agent in the CM through a closed loop: from the computational module, the heated heat-transfer agent passes into the heat exchanger and is cooled there. From the heat exchanger, the cooled heat-transfer agent again passes into the computational module and cools the heated electronic components there. As a result of heat dissipation, the agent becomes heated and again passes into the heat exchanger, and so on. The heat exchanger is connected to the external heat-exchange loop via fittings and is intended for cooling the heat-transfer agent with the help of a secondary cooling liquid. A plate heat exchanger in which the first and the second loops are separated can be used as a heat exchanger. So, as the secondary cooling liquid, it is possible to use water cooled by an industrial chiller. The chiller can be placed outside the server room and can be connected to the reconfigurable



a) **b)**

Fig. 1. The design of a computer system based on liquid cooling (a - design of a new generation CM, b - design of the computer rack)

computational modules by means of a stationary system of engineering services. The design of the computer rack with mounted CMs is shown in Fig. 1-b.

The computational and the heat exchange sections are mechanically interconnected into a single reconfigurable computational module. Maintenance of the reconfigurable computational module requires its connection to the source of the secondary cooling liquid (by means of valves), to a power supply block or to a hub (by means of electrical connectors).

In the casing of the computer rack, the CMs are placed one over another. Their number is limited by the dimensions of the rack, by technical capabilities of the computer room, and by the engineering services.

Each CM of the computer rack is connected to the source of secondary cooling liquid with the help of supply and return manifolds through fittings (or balanced valves) and flexible pipes; the connection to both the power supply and the hub is performed via electric connectors.

The supply of cold secondary cooling liquid and the extraction of the heated one into the stationary system of engineering services connected to the rack are made via fittings (or balanced valves).

For the purpose of testing technical and technological solutions, and determining the expected technical and economical characteristics and service performance of the designed high-performance reconfigurable computer system with liquid cooling, we designed a number of models, experimental and technological prototypes. Figure 2 shows the prototype of a new-generation “SKAT” CM. A new-design CCB with high packing density was created for this CM.

The CCB of the advanced computational module contains eight Kintex UltraScale XCKU095T FPGAs; each FPGA has a specially designed thermal interface and a low-height heatsink for heat dissipation.

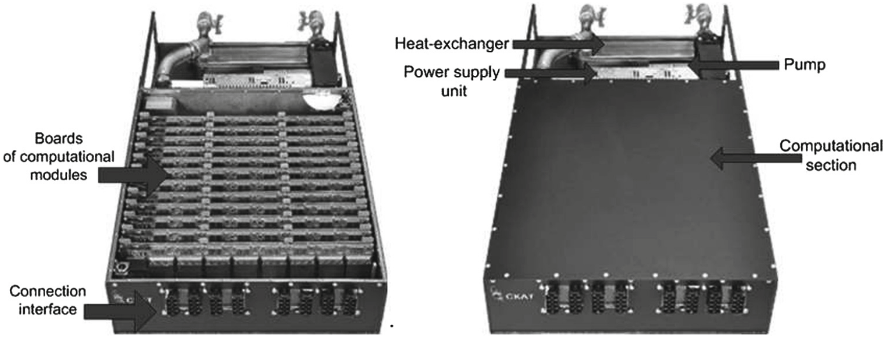


Fig. 2. The prototype of the new-generation CM

We have designed an immersion power supply unit providing DC/DC 380/12 V transducing with the power up to 4 kW for four CCBs.

The computational section of the "SKAT" CM contains 12 CCBs with a power of up to 800 W each, and three power supply units. In addition, all boards are completely immersed into an electrically neutral liquid heat-transfer agent.

To achieve an effective immersion cooling system, we developed a dielectric heat-transfer agent with the best possible dielectric strength, high heat transfer capacity, the maximum possible heat capacity and low viscosity.

The heat exchange section contains pump components and the heat exchanger, both providing the effective flow and cooling of the heat-transfer agent. The design height of the CM is 3U.

The performance of a next-generation "SKAT" CM is increased in 8.7 times in comparison with the "Taygeta" CM. Original design solutions provide more than triple increasing of the system packing density. Clock frequency and logic capacity of the FPGAs are also increased. As a result, all this provides such qualitative increasing of the system specific performance.

Experimental tests of the developed solutions of the immersion liquid cooling system proved that the temperature of the heat-transfer agent does not exceed 30 °C, and the power consumed by each FPGA in operating mode equals 91 W (8736 W for the whole CM). In addition, the maximum FPGA temperature during heat experiments did not exceed 55 °C. All this proves that the designed immersion liquid cooling system has a reserve and can provide effective cooling for the designed RCS based on the advanced Xilinx UltraScale+ FPGA family.

4 "SKAT+" Advanced Reconfigurable Computer System Based on Xilinx UltraScale+ FPGAs

The use of UltraScale+ FPGAs based on the 16FinFET Plus 16-nm technology and produced by Xilinx since 2017 will provide a three time increase in computational performance due to an increase in clock frequency and FPGA circuit complexity, whereas the size of the computer system will still remain unchanged. However, despite the reduction of relative energetic consumption due to new technological standards of FPGA manufacturing and also to a certain power reserve of the designed liquid cooling system, it is expected that FPGA operating temperatures will approach again their critical values.

In addition, the new FPGAs of the UltraScale+ family have larger geometric sizes. The size of the FPGAs in the "SKAT" RCS is 42.5 × 42.5 mm. The size of the FPGAs that will be placed into the"SKAT+" RCS amounts to 45 × 45 mm. Owing to this circumstance, it is impossible to use the existing CCB design since the width of the printed circuit board will become larger and will not fit in a standard 19 rack.

In this connection, it is necessary to modify the designed open liquid cooling system and the CCB design, which will lead to a modification of the whole CM.

At present, the SRC of SC & NC scientific team is working on the design of an advanced RCS based on Xilinx UltraScale+ FPGAs. Due to these works related to the modification of the cooling system, we are going to solve the following problems:

1. Increase the effective surface of heat-exchange between FPGAs and the heat-transfer agent.
2. Increase the performance of the heat-transfer agent supply pump.
3. Increase the reliability of the liquid cooling system by means of immersed pumps.
4. Experimentally improve the heat-sink optimal design.
5. Experimentally improve the technology of thermal interface coating.

We have designed a prototype of an advanced computational module with a modified immersed cooling system (Fig. 3). Some distinctive features of the new design are immersed pumps and a considerable reliability increase of the CM due to a reduction of the number of components and simplification of the cooling system. According to our plans, the heat-exchange section will house only the heat exchanger. We are working on experimental pump equipments that can operate in the heat-exchange agent. During modification of the CCB design,

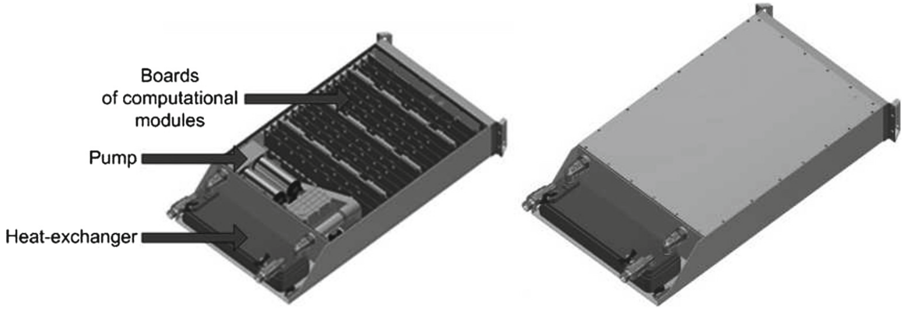


Fig. 3. A prototype of a computational module with a modified immersed cooling system

we have created a prototype of an advanced board, shown in Fig. 4. The CCB contains eight UltraScale+ FPGAs of high circuit complexity. To provide room for the new CCB into a 19 rack, it is necessary to exclude its CCB controller from its structure. The CCB controller was always implemented as a separate FPGA and used to provide access to the FPGA computational resources of the CCB, FPGA programming, and monitoring of the CCB resources.

FPGAs are rather small, but their resource constantly grows with each new family. At the same time, the variety of functions of the CCB controller grows only slightly. As a result, the resources required at present for the implementation of all the CCB controller functions amount to only some percent of the logic capacity of the FPGAs currently used. In this connection, further implementation of the CCB controller as a separate FPGA is considered unnecessary. A single FPGA in the computation field will be able to perform all the functions of the controller. We need a system of hydraulic balancing for the heat-transfer agent flow within each hydraulic loop. Such a system will provide an equal flow of the heat-exchange agent in each computational module inside the computer rack during servicing of one or several computational modules. The additional subsystem will considerably complicate the cooling system. To simplify the hydraulic balancing of the heat-transfer system, the SRC SC & NC scientific team (LLC, Russia) has suggested an engineering solution for balancing the heat-transfer agent flow in the heat-exchange system (see Fig. 5).

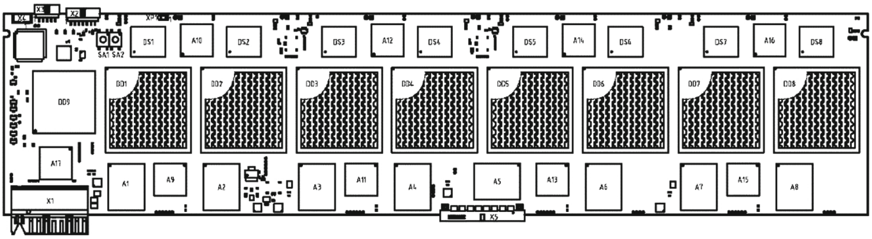


Fig. 4. The prototype of the CCB modified packing

The designed heat-exchange system includes a pump 1, a chiller 2 (that is, a cooling machine), several circulation loops 3, with heat-exchangers 15, intended for the heat exchange process between the primary heat-transfer agent (water) and the secondary one (oil MD-4.5), which is circulating in the computational modules 4. The circulation of the secondary heat-transfer agent (oil MD-4.5) in computational modules 4 and heat-exchangers 15 of each circulation loop is ensured by an additional pump (not shown in Fig. 1) connected to each circulation loop. Heat-exchangers 15 are connected by parallel tubes to supply manifold 6 and return manifold 5. The inlets and outlets of the first (No. 1), second (No. 2), third (No. 3), etc., circulation loops are arranged along the heat-transfer agent flow 7 and near the inlets of supply manifold 8 and return manifold 9. The inlets and outlets of the last circulation loop No. 6 in Fig. 1) are situated near the outlets of supply manifold 10 and return manifold 11. The return pipe 12 connects the outlet of return manifold 11 with the chiller 2, the pump 1, and the inlet of supply manifold 8. Besides, each circulation loop may be complemented with a balancing valve for finer balance-tuning. The heat-exchange system is filled with the primary heat-transfer agent (water, antifreeze, etc.), then decreased, and the pump 1 is switched on. The primary heat-transfer agent is supplied to the inlet 8 of the supply manifold 6 and then through the circulation loops 3 into the heat-exchangers 15, where heat is transferred from the primary heat-transfer agent to the secondary one, which is circulating in the computational modules 4, where the secondary heat-transfer agent (oil MD-4.5) dissipates the heat from heating electronic components. The primary heat-transfer agent gets warm and enters the return manifold. There is a return pipe 12 at outlet 11 of the return manifold. Through return pipe 12 and the chiller 2, the primary heat-transfer agent is again transferred to pipe 1, and then to inlet 8 of supply manifold 6, and then flows along the closed loop. In the chiller 2, the heated primary heat-transfer agent is chilled.

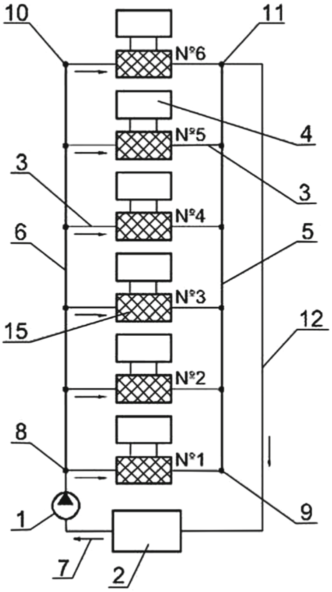


Fig. 5. The layout of hydraulic balancing of the heat-exchange system for computational modules in a computer rack

If a circulation loop in any computational module fails, then the heat-transfer agent flow is evenly changed in the rest of modules, since the closed trajectory of the heat-transfer agent flow is similar for all loops, and the distance between each loop and the pump is the same: pump – inlet of the supply manifold – supply manifold – circulation loop – return manifold – outlet of the return manifold – return pipe – chiller – pump. The described engineering solution makes it possible to balance the hydraulic resistance in all the circulation loops when the heat-transfer agent flow is pumped through them. No additional hydraulic balancing system is needed here.

Thanks to breakthrough technical solutions that we have found while designing the "SKAT" RCS with an immersed liquid cooling system, we are now able to develop this direction of high-performance RCS design, and after some design improvements, we will be able to create a computer system providing a new level of computational performance.

5 Conclusions

The use of air cooling systems in the design of supercomputers has practically reached its limit, since cooling effectiveness decreases as the rate of consumed and dissipated power grows at the same time as circuits in microprocessors and other chips become more and more complex. This explains why the use of liquid cooling in modern computer systems is considered as a priority direction for the improvement of cooling systems and has excellent prospects of further development. Liquid cooling of RCS computational modules containing not less than eight FPGAs of high circuit complexity has some specific features compared with the cooling of microprocessors and requires the development of a specialized immersion cooling system. The original liquid cooling system that has been designed for a new-generation RCS computational module provides high maintenance characteristics, such as a maximum FPGA temperature not exceeding 55 °C, while keeping the heat-transfer agent temperature below 30 °C in operating mode. Thanks to breakthrough solutions found for the immersion liquid cooling system, it is now possible to mount not less than 12 new-generation CMs, with a total performance above 1 PFlops, in a single 47U computer rack. The power reserve of the liquid cooling system for the new-generation CMs ensures an effective cooling not only for the existing but also for future FPGA families (Xilinx UltraScale+ and UltraScale 2).

FPGAs, as principal components of reconfigurable supercomputers, provide a stable, practically linear growth of the RCS performance. This makes the performance of an RCS based on Xilinx Virtex UltraScale FPGAs similar to that of the world best cluster supercomputers and opens new possibilities for the design of super-high performance supercomputers.

References

1. Perkowski, M.: FPGA computer architectures. Northcon/93. In: Conference Record, 12–14 October 1993. https://doi.org/10.1109/northc.1993.505038. ISBN: 0-7803-9972-2
2. Tripiccione, R.: Reconfigurable computing for statistical physics. The weird case of JANUS. In: IEEE 23rd International Conference on Application-Specific Systems, Architectures and Processors (ASAP) (2012). https://doi.org/10.1109/asap.2012.38
3. Baity-jesi, M., et al.: The Janus project: boosting spin-glass simulations using FPGAs. In: IFAC Proceedings Volumes, Programmable Devices and Embedded Systems, vol. 12, no. 1 (2013). https://doi.org/10.3182/20130925-3-cz-3023.00039
4. Shaw, D.E., et al.: Anton, a special-purpose machine for molecular dynamics simulation. Commun. ACM **51**(7), 91–97. https://doi.org/10.1145/1364782.1364802
5. Kalyaev, I.A., Levin, I.I., Dordopulo, A.I., Slasten, L.M.: Reconfigurable computer systems based on Virtex-6 and Virtex-7 FPGAs. IFAC Proc. Volumes Programmable Devices Embed. Syst. **12**(1), 210214 (2013). https://doi.org/10.3182/20130925-3-cz-3023.00009
6. Coolitsystems. http://www.coolitsystems.com/index.php/data-center/liquid-cooling-options.html
7. Asetek Data Center Liquid Cooling. http://www.asetek.com/data-center/oem-data-center-coolers
8. Data Center Cooling: Liquid Immersion - Green Revolution Cooling. http://www.grcooling.com/carnotjet
9. Immers.ru. http://www.immers.ru/sys/immers660
10. Eurotech HPC. http://www.eurotech.com/aurora
11. RSC Technology. http://www.rscgroup.ru
12. T-platforms. http://www.t-platforms.ru/products/hpc/a-class/cooling.html
13. Iceotope. http://www.iceotope.com/product.php
14. LiquidCool Solutions - Liquid cooled servers for a CRAC - free future. http://www.liquidcoolsolutions.com