ANNALS of Faculty Engineering Hunedoara – International Journal of Engineering
Tome XX [2022] | Fascicule 1 [February]

[1.] Stanko P. STANKOV

COOLING TECHNOLOGIES OF DATA CENTERS

[1.] University of Niš, Faculty of Electronic Engineering, Department of Automatic Control, Niš, SERBIA

Abstract: The installation of more and more powerful processors and the constant striving to reduce the size of the hardware architecture have conditioned the cooling of data centers to become a very complex task. Modern servers generate ten times more heat energy than systems built just a decade ago. The consequence of the installation of powerful servers leads to the formation of the so–called hot zones with high heat load. Efficient cooling of these zones in traditional way is often impossible. In order to ensure the reliable operation of modern data centers, in which the infrastructure is installed with high density, it is necessary to apply new strategies and technologies to ensure adequate cooling of the equipment. Modern server racks often consume more than 5 kW of electrical energy, with a constant tendency to increase consumption. Practice shows that conventional cooling systems are already inefficient for these powers. A quality cooling solution for data centers will significantly contribute to the overall energy efficiency.

Keywords: servers, rack, cooling, aisles, energy efficiency

1. INTRODUCTION

Data centers are sets of servers and other computer equipment on which numerous services required by a large number of users are carried out. Modern companies have a need for services, which must be fast, reliable and available at any moment. This has imposed the need to develop data centers in a growing number of companies [1–3]. Virtual servers reduced the need for a large number of physical servers, which at the same time allowed more services to run on the same servers through the virtualization process. It encouraged the development of information systems in small and medium enterprises and added momentum to their improvement in the enterprise segment [3], [4]. The development of data centers requires the environment in which the servers are installed to be improved in order to ensure stable, reliable and uninterrupted operation of all servers in the data center in order to provide users with constant access to data and mechanisms for their processing. These requirements can be met by keeping track of changes in technology, safety and other components of the environment such as climate conditions, communications and rules of conduct. Therefore, it is necessary to transform the data centers step by step until the final consolidation [5].

Data center should provide:

— reliable internet, provided with several internet operators;
— safe power supply of devices and equipment;
— generators that will enable normal mode in case of mains supply failure;
— room for air monitoring system (temperature, humidity);
— reliable air conditioning system to maintain the required temperature and humidity in the premises; • appropriate fire protection system;
— technical support, in order to take timely and adequate measures in case of failure of a particular device;
— physical security of the facility;
— limited access to the premises [1], [6].

The role of cooling systems in computer centers consists of maintaining appropriate conditions that are optimal for the exploitation of computer resources (Information technology equipment – ITE). These systems must operate reliably and continuously. It is recommended that the room temperature is in the range of 21–24 °C with a deviation of ± 3 °C, and humidity of 50 % with a deviation of ± 10 %. For the security and functionality of the data center, it is necessary to provide: authorized persons with access to the center, monitoring of the data center by video surveillance and access control, supervision of the operation of complete equipment, installation of fire protect and alarm systems. Maintaining the optimal temperature is extremely important for the smooth operation of computer machines and it is a priority in data centers in order to provide quality and reliable services to customers. The energy required to cool data centers during operation can be quite high if the necessary measures to install a reliable and energy efficient cooling system are not taken [7]. In addition to cooling, power supply is crucial for data centers. Power outages can cause major problems, because even after the power is restored, it is not certain that all data has been saved due to improper shutdown of the server, nor that all services are functional. Servers need some time to boot operating systems, whereby power consumption reaches peaks, so in case of inadequate sizing of the electrical installation, the data center may crash. The organization of power supply of the data center is shown in Figure 1. [1], [7].

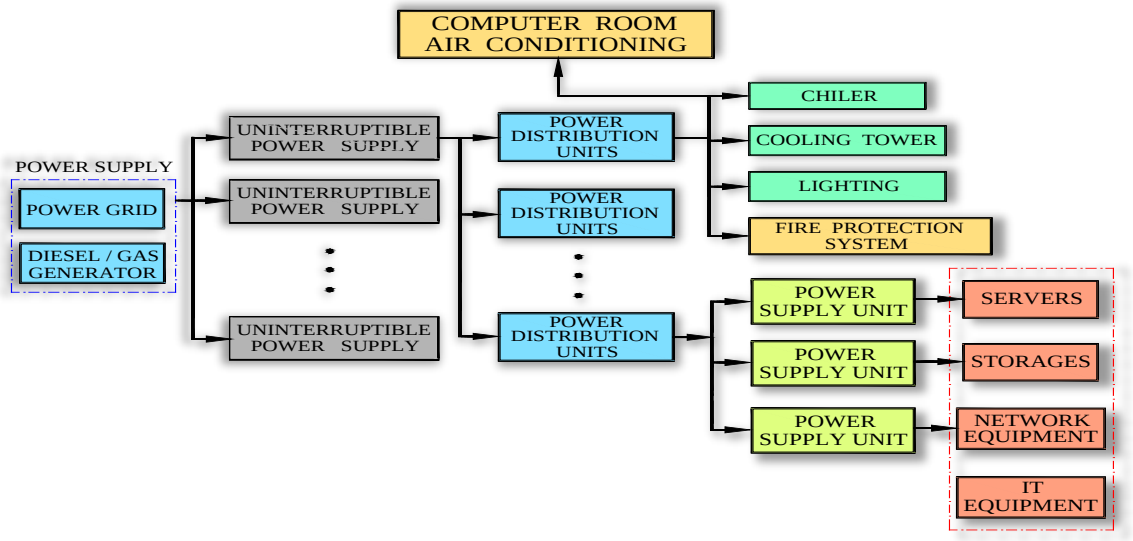

Figure 1. Power supply of the data center

The installation of powerful servers in computer centers has of creating the so–called as the consequence hot spots – zones with a higher thermal load compared to zones with older generation equipment. Efficient cooling of such zones is practically impossible to achieve by applying conservative techniques. Reliable operation of modern data processing centers, which are characterized by a high density of embedded computer resources, uses a new strategy and technology that can provide adequate cooling of equipment. Efficient cooling is ensured by installing cooling systems to the heat source. The use of technological fluids that absorb heat during the change of its physical state can also achieve cooling efficiency. When the refrigerant is in a gaseous state, the risk of liquid leakage is eliminated, which can lead to equipment failure in the data center [8]. In order to ensure adequate heat spreading, racks with servers are usually placed at a certain distance from each other, which practically limits the number of racks per unit area and increases the maintenance costs of the entire system. For example, in a room of 1000 m2, with adequate maintenance, no more than 50 rack cabinets with a heat load of 10 kW per rack can be installed, provided that a raised floor is used. With additional cooling, 386 racks can function safely and efficiently on that surface, which is practically eight times more equipment on the same surface. Data centers generate heat with less humidity compared to ordinary rooms where people live or the outside air enters, which is why such facilities use air conditioners that consume less energy when absorbing moisture from the air, and more to absorb heat, compared with air conditioners intended for households. Such air conditioners enable additional humidification of the air, which maintains the relative humidity in the room to the allowed value of 50 % [9]. Optimization of the spatial arrangement of server racks that enables the formation of alternating hot and cold aisles – Figure 2 can increase the cooling efficiency with a raised floor system.

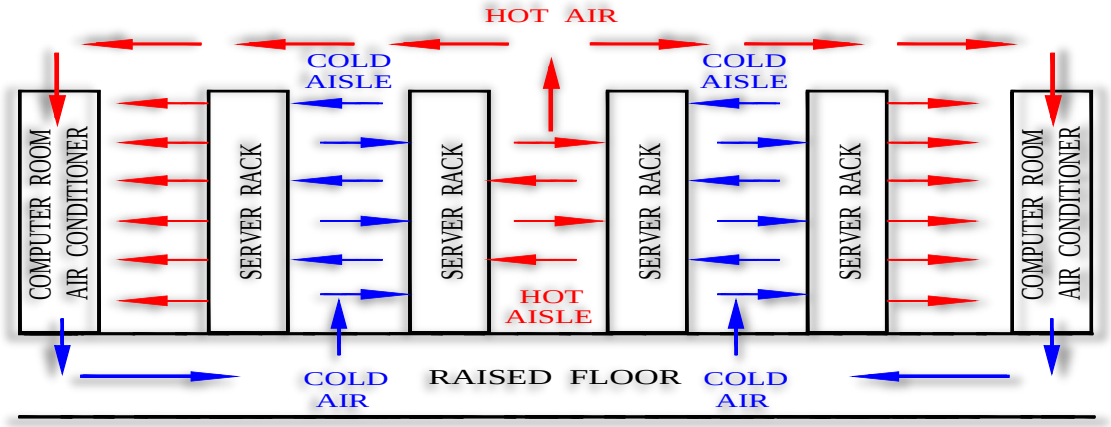

Figure 2. Circulation with raised floor and hot–cold aisles

Hot and cold aisles can be lined up face to face, with a cold stream of air passing past the cold aisles and entering the server cabinets to exit through the back of the hot aisle. There is also the possibility of partitioning so as not to mix cold and warm air. A partition in combination with a double floor can be an efficient way of cooling. Contrary to this, the insulation of warm aisles is done by means of a system on the back of the server cabinets

from where warm air comes out. In this way, certain parts of the room become a warm aisle. These systems can be directly connected to the air duct to the air chamber in the computer room and can work with a pre–installed water economizer [10].

2. ENERGY EFFICIENCY

In the era of IoT (Internet of Things) technology, trends in the use of cloud infrastructure, remote services, training and remote workplaces are becoming increasingly popular, and the Covid 19 pandemic has contributed to this being used in less developed countries as well. As already mentioned, data centers around the world use more than 4 % of the world's total electrical energy consumption.

Forecasts say that the consumption will double every four years. Electrical energy is used in data centers for equipment operation (50 %) and for cooling (30 %). It is obvious that in the current configuration of data centers, electrical energy consumption will increase enormously in the future. There is a need to find environmentally friendly solutions for cooling computer resources. It is believed that the free cooling method will save 80 % on cooling data centers. Energy efficiency is the main focus when solving the problem of cooling data centers. Having in mind the profitability of the facility, it is recommended to choose modern efficient technologies regardless of the current high price, which pays off in the long run. An efficient solution in every sense of the data centers will significantly contribute to the overall energy efficiency [1], [8]. ServeCool combines three different cooling methods. Free cooling takes place through the outside air that enters the room indirectly, passing first through a plastic exchanger. Thanks to efficient recuperation, the air is cooled and led into the interior of the room, without creating dust and retaining moisture in sensitive server blocks. When cooling based on recuperation is not enough, an adiabatic cooling system is automatically switched on – an efficient way of using renewable resources. These two cooling modes – recuperation with a plastic exchanger and adiabatic cooling cover 97 % of the operating time. In the remaining time of 3 %, the mechanical system (chiller) is switched on. This concept enables maximum reliability and minimum energy consumption [11]. This concept provides maximum reliability and minimum energy consumption compared to conventional systems that use a combination of passive (50 %) and mechanical cooling in the amount of 25 %.

3. COOLING TECHNIQUES

During the decades of development of the IT system for cooling computer rooms and data centers, double floors were most commonly used, through which cold air is delivered. There is cold air from the computer room air conditioner (CRAC) or air chamber (Computer room air handler – CRAH). The use of perforated panel floors allows cold air to flow into the server room. By passing through the server, the cold air is heated and returned to the air chamber to be cooled again. The return pipe temperature is often used as a control point to which the cooling system is controlled. During a number of years of operation of data centers, the most common choice of cooling is the use of a double perforated floor, which is still a common case today. The basis of this method of cooling with a relatively small amount of air–conditioned air is mixing with the air in the room, which achieves the desired temperature. Although the comfort and efficiency of these systems is high if the server equipment is not dense, it can hardly be said that this relatively simple method is obsolete. Many companies base their data center cooling system in this way. However, more and more people are resorting to new and more efficient technologies in terms of cooling data centers. One of the popular ways to cool data centers today is to use cold and hot corridors and close empty spaces in server racks with special panels. These panels prevent the return of warm air that would enter the front of the server through the opening. In order to facilitate air exchange, it is necessary to remove all elements that would block the suction and exhaust openings of computer equipment. Besides, weakened airflow increases energy consumption, due to more intensive fan operation. If there are doors in the racks, they should be perforated and kept at least 65% open. Solid doors or those made of glass, plexiglass or other material inevitably hinder the flow of air, and thus the cooling process.

4. HEAT DRAINAGE

IT equipment is powered by electricity, which releases heat, which without proper ventilation leads to a series of failures, failures and data loss. The cooling process itself, which is necessary in data centers, causes certain problems due to the additional energy consumed by the cooling infrastructure.

There are generally two ways to dissipate heat in data centers: by air or by coolant (usually water or refrigerant). Air cooling is simpler as there is no danger to installe Information technology equipment – ITE. The favorable circumstance also lies in the distribution of hot and cold air, bearing in mind that warm air goes upwards as easier. This feature is used by a large number of cooling systems. However, this is not a particularly effective solution, as hot and cold air are mixed, which reduces efficiency [1], [10]. On the other hand, the coolant leads directly to the server cabinets and is precisely directed to the hot spots, while not cooling the entire space. However, possible leaks in this system can be a great danger for IT equipment. One of the solutions for heat

dissipation in data centers is the installation of heat exchangers in the server cabinets themselves. Exchangers extract heat from the server and eliminate it with liquid. Such cooling devices are installed in the immediate vicinity of the server. After taking heat from the server, it is usually discharged into the atmosphere, water tanks or the ground. There are several ways to dissipate heat:

- CRAH (A computer room air handler) – plants with water–cooled chillers and cooling towers,
- CRAH – air–cooled chiller plants,
- air conditioning split systems for rooms with computers,
- CRAC (Computer Room Air Conditioner) – plants with cooling towers,
- working fluids and cooling towers,
- air economizer,
- air economizer with direct cooling by evaporation,
- indirect cooling with evaporation.

5. MEGA DATA CENTERS IN THE WORLD

Huge funds in the construction of gigant data centers are invested by Google, Amazon, Mícrosoft and ÍBM. The largest number of these centers is installed in the USA (44 %), in China 8 %, 6 % are located in Japan and GB. In countries such as Australia, Germany, Singapore, Canada, India and Brazil, the percentage ranges from 3 to 5. Based on this, it is evident that the USA is the world market leader in data centers and has a dominant role in cloud and internet technologies. The consulting firm Arizton predicts that the global market of data centers will reach 174 billion $ by 2023. IoT technologies are constantly increasing the demand for data centers, with data becoming increasingly valuable. In this regard, a sustainable, efficient, adaptable and resilient data center infrastructure is needed if owners want to increase of profit. The intensive increase in investments is evidenced by the fact that in the first half of 2018. The operators of the largest data centers in the world, such as Google, Microsoft, Amazon, Apple and Facebook, have invested over $ 53 billion in the development of their infrastructures, compared to the same period in 2017. A significant part of investments in the world relates to the construction and expansion of large data centers. There are over 420 such centers installed in the world, most of which are located in the USA. These are Synergy Research analyzes based on information on the size of investments in data centers of the world's 20 largest companies providing cloud and online services, including Infrastructure as a service (IaaS), Platform–as–a–Service (Paas) and Software as a service (SaaS), search engine operators, social networks and online magazines [12]. Alibaba, Baidu, IBM, JD.com, NTT, Oracle, SAP and Tencent are among the companies that invest the most in data center development. It is interesting that these companies are not in the top 5. Capital investments in mega data centers are the biggest indicator of intensive growth and development of cloud and digital technologies, as well as on–line lifestyle. The costs of mega data center operators reach enormous values and continue to grow. They are extremely increasing their competitive advantages, with "small players" in this area no longer having a chance to race. This tendency is obvious especially in the cloud computing market where giant data centers are simply destroying competition [13]. Schneider Electric, one of the world leaders in energy management, has efficient solutions for data centers. InRow RC precision air conditioning systems significantly reduce energy consumption. In the high inlet temperature model, which is designed for optimal heat dissipation, higher cooling fluid temperatures and higher ambient air temperature increase cooling efficiency, whereby the number of hours of free cooling increases [14]. As part of Schneider Electric's InfraStruXure solution, the InRow RC reduces the distance between the heat source and the heat dissipation, thus reducing the mixing of hot and cold air currents. Variable speed fans reduce power consumption during peak reduction times or when the data center is partially loaded. Schneider Electric has also reached a high standard in UPS technology – for example the Galaxy VM, a three–phase system for uninterruptible power supply, peak reduction time or when the data center is partially loaded. Galaxy VM, is a key component in an integrated power control solution. This compact design device is easy to install, takes up little space in data centers and saves 90% of energy through the intelligent ECOnversion mode. System corrects input harmonics while combining dual online conversion with improved environmental performance and provides optimal efficiency for critical loads of various equipment such as those in data centers. Operators in data centers of demanding industrial and other facilities are constantly looking for a balance between the need for high energy efficiency and maximum access and uninterrupted operation of equipment. For these reasons, the Galaxy VM is the solution to these requirements [14].

6. COOLING OF DATA CENTERS IN THE FUTURE

Servers in data centers generate a large amount of heat, while air cooling is very expensive. For example Microsoft spends over $ 15 billion a year building and maintaining 100 data centers. This is also the main reason why this company is trying to install its cloud data centers in waterproof containers on the ocean floor.

Theoretically, ocean water should cool servers, and electricity would be generated by waves. The idea of building underwater data centers was spurred by energy savings for cooling computer resources. Data centers are large consumers of electricity (nowadays over 430 TW), with 15 to 20 % spent on computer cooling (according to CNET – Computer Network). Theoretically, the cold waters of the North Sea can contribute to significant savings. Experiments to date have shown that heat from the housing of data centers can be efficiently dissipated in the ocean. In addition, these data centers are unmanned. Air with oxygen and water vapor is extracted from the server rooms, which solves the problem of corrosion [15]. It is natural that in such conditions, equipment that is defective cannot be replaced. However, the creators of these interesting experiments are of the opinion that in such conditions of exploitation, the probability of failures and failures of computer equipment is significantly reduced [16]. Microsoft has already experimented in this regard by installing a 2.4 m diameter steel capsule (Project Natick – Figure 3) in the North Sea, which has been under water for 105 days. During testing, this module processed Azure's cloud service commercial data. Microsoft continues research with several times larger modules compared to the tested prototype [17], [18]. Microsoft conducted the first experiment with an underwater data center during 2015.

Figure 3. Project Natick ready for deployment [15]

Then a container called Leona Philpot was sunk for 5 months. Microsoft believes that it is expedient to install data centers on the seabed, since over 90 % of these centers are now located near large cities – London, Frankfurt and other major global centers for the flow and exchange of data. In this regard, environmental associations oppose the idea, believing that the installation of underwater data centers will lead to an increase in the temperature of the world sea. But even the most approximate calculations show that the increase will be insignificant – no more than a thousandth of a degree Celsius [18], [19]. During 2018. Microsoft has installed a data center with 864 servers at the bottom of the Scottish Sea, at a depth of 36 m. This company has the idea to set up similar data centers near the shores of certain areas where these centers are needed. Practice shows that the failures of data centers installed on the ocean floor amount to only one eighth of failures in data centers installed on land. The company's next step in this direction is to find a way to extract data centers from the ocean floor for troubleshooting.

7. SUPERVISION AND CONTROL OF DATA CENTERS

The IT infrastructure is characterized by rapid changes and constant adaptation to customer needs. Cloud and as a Service technologies are very common today. Emergence of phenomena such as Big Data, Machine Learning, Artificial Intelligence, Data Mining and etc. which are related to data, has caused an increase in all segments of computer infrastructure. The fact that 90 % of all existing data has been generated in the last two years testifies to the constant increase in the amount of data. In order for the IT sector to cope with such a drastic increase in the amount of data, it is necessary, in addition to capacity flexibility, fast computer nodes and faster data transfer within the data center, to reduce downtime or interruptions to a minimum. This can be ensured by monitoring and control of the entire infrastructure [1]. In 2018. the Uptime Institute conducted research on the functionality of data centers. It was concluded that in the period 2015 – 2018, 48 respondents had work interruptions given at least once. According to Gartner's research, a one–minute data center outage costs an average of $ 5,600. Due to the interruption of the data center, negative effects occur, such as loss of sales, damage to the reputation of companies, reduction of productivity, loss of data and etc. Overcoming the aforementioned problems can be achieved by applying a central system of supervision and management of the data center, which reduces the probability of failure of data center components. Data Center Infrastructure Management (DCIM) is software that monitors the complete computer and other data center infrastructure, through a centralized control screen on which all vital components of the system can be viewed in detail. Google has promoted real–time data center monitoring and management systems using artificial intelligence. They came to the conclusion that cooling is carried out 75 % more than the required value. According to data center knowledge.com data, center cooling is designed according to the instructions of risk management organizations [20]. However, there are more and more solutions that turns to the precise operational control of data centers air conditioning, permanent monitoring of temperature and energy consumption. Based on that,

one of the world's companies in the field of cybersecurity managed to increase the average temperature of server rooms by 3°C, which leads to energy savings for cooling on an annual level of 25 % [21–23].

8. CONCLUSION

It is obvious that cooling data centers requires a very complex choice and a comprehensive solution. There are a number of techniques based on air or water cooling, which use different sources, while working on optimizing the interior of the room, controling energy consumption etc. The physical and geographical features of the location where the data center is located are considered, the existing technologies, tendencies and possibilities for optimal energy use, return on investment – ROI, power usage effectiveness (PUE), total cost of ownership – TCO. It is concluded that high–class equipment is not enough for the overall efficiency of data centers. It is necessary to have a broader view of the problem, anticipate further development and expansion of the data center, adoption of the concept by experts, design and correct execution, as well as proper maintenance of installed capacities. The efficiency of a single data center simply cannot be bought.

Note: This paper was presented at IIZS 2021 – The XI International Conference on Industrial Engineering and Environmental Protection, organized by Technical Faculty "Mihajlo Pupin" Zrenjanin, University of Novi Sad, in Zrenjanin, SERBIA, in 07–08 October, 2021.

References

[1] Stankov, S., Perić, S.: Konfiguracija sistema za upravljanje zgradama, 41 međunarodni Kongres KGH, Zbornik radova (100–112), Beograd, 2010.
[2] HVAC Design and Cooling for data Center Efficiency Spiral–Bound – International Edition, Publisher: Strategic Media Asia Limited; 2019.
[3] Stankov, S., Mitić, D., Icić, Z.: Supervisory and Control of HVAC Systems, journal „Facta Universitatis", Series: Automatic control and Robotics, vol. 9, No 1 (pp 151 – 160), 2010.
[4] Bhatia, A.: HVAC Cooling Systems for Data Centers, Continuing Education and Development, Inc. 22 Stonewall Court Woodcliff Lake, 2020. https://www.cedengineering.com/userfiles/HVAC%20Cooling%20Systems%20for%20Data%20Centers%20R1.pdf
[5] Orsaria, A., Jang, M.: RHCSA/RHCE Red Hat Linux Certification Practice Exams with Virtual Machines, Second Edition, Jan 24, 2020.
[6] Shehabi, A., Smith, S. J., Horner, N., Azevedo, I., Brown, R., Koomey, J., Masanet, E., Sartor, D., Herrlin, M., Lintner, W.: United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, Berkeley, California, 2016.
[7] Capozzolia, A., Primiceri, G.: Cooling systems in data centers: state of art and emerging technologies, 7th International Conference on Sustainability in Energy and Buildings, Torino, Italy, 2015.
[8] Manganelli, M., Soldati, A., Martirano, L., Ramakrishna, S.: Strategies for Improving the Sustainability of Data Centers via Energy Mix, Energy Conservation, and Circular Energy, Sustainability 2021.
[9] Google Data Centers. Efficiency: How we do it. https://www.google.com/about/datacenters/efficiency/internal/
[10] Stankov, S., Mitić, D., Icić, Z.: Control of HVAC Systems, X Triennial international SAUM Conference on Systems, Automatic Control and Measurements, pp 152 – 157, Niš, Serbia, Niš, november $10^{th} \div 12^{th}$, 2010.
[11] https://www.google.com/search?client=firefox–b–d&q=ServeCool
[12] https://www.srgresearch.com/
[13] https://www.srgresearch.com/research/cloud–infrastructure–services–deep–dive
[14] Niemann, J., Bean, J., Avelar, V.: Economizer Modes of Data Center Cooling Systems, Revision 0, Schneider Electric white paper library produced by Schneider Electric's Data Center Science Center, 2011. https://download.schneider–electric.com/
[15] Microsofts Underwater Server Experiment Resurfaces After Two Years, https://www.theverge.com/2020/9/14/21436746/microsoft–project–natick–data–center–server underwater–cooling–reliability
[16] Hume, D.: Underwater Data Centers, October 24, 2017, https://theliquidgrid.com/2017/10/24/underwater–data–centers/
[17] Project Natick: Microsofts nderwater Voyage of Discovery, https://www.datacenterdynamics.com/en/analysis/project–natick–microsofts–underwater– voyage–discovery/
[18] Copeland, M., Soh, J., Puca, A., Manning, M., Gollob, D.: Microsoft Azure: Planning, Deploying, and Managing Your Data Center in the Cloud, 1st Edition, September 30, 2015.
[19] Microsft Pulls Underwater Data Centback to the Surface to Assess Benefits of Deep–Sea Cloud, 2020, https://www.geekwire.com/2020/microsoft–pulls–underwater–data–center–back–surface assess–benefits–deep–sea–cloud/
[20] China Launches Underwater Data Center, https://www.datacenterdynamics.com/en/news/china–launches–underwater–data–center/
[21] https://www.datacenterknowledge.com/
[22] Strong, L., Huttemann, R.: Managing Airflow at the Room Level, Data Center Containment best Practices that don't Your Budget – part 3, May 26, 2020, https://www.missioncriticalmagazine.com/articles/92991–managing–airflow–at–the–room level
[23] Dai, J., Ohadi, M., M., Diganta, D., Pecht, G., M.: Optimum Cooling of Data Centeres: Application of Risk Assessment and Mitigation Technicues, Springer, 2014.

ISSN 1584 – 2665 (printed version); ISSN 2601 – 2332 (online); ISSN–L 1584 – 2665