

Server Level Liquid Cooling: Do Higher System Temperatures Improve Energy Efficiency?

*Egor A. Druzhinin*¹, *Alexey B. Shmelev*¹, *Alexander A. Moskovsky*¹, *Vladimir V. Mironov*², *Andrey Semin*³

© The Authors 2017. This paper is published with open access at SuperFri.org

Liquid cooling is now a mainstream approach to boost energy efficiency for high performance computing systems. Higher coolant temperature is usually considered to be an advantage, since it allows heat reuse/recuperation and simplifies datacenter infrastructure by eliminating the need of chiller machine. However, the use of hot coolant imposes high requirements for cooling equipment. A promising approach is to utilize coldplates with channel structure and liquid circulation for heat removal from semiconductor components. We have designed a coldplate with low heat-resistance that ensures effective cooling with only 20-30° temperature difference between the coolant and electronic parts of a server. Under the stress-test conditions the coolant temperature rose up to 65 °C while server operation remained. We also studied power efficiency (expressed in floating point operations per watt) dependence on the coolant temperature (19-65 °C) on the individual server level (based on Intel Grantley platform with dual Intel Xeon E5-2697 v3 processors). The power performance ratio shows moderate ($\approx 10\%$) efficiency drop from 19 to 65 °C due to increase of leakage current in chipset components and reduction of processor frequency which resulted into proportional reduction of DGEMM benchmark performance. It must be taken into account by datacenter designers, that the amount of recuperated energy from 65 °C should be at least $\approx 10\%$ to justify the choice of high temperature coolant solution.

Keywords: hot liquid cooling, cold plates, energy efficiency.

Introduction

Current power usage levels of top supercomputers place a heavy burden on system owners and maintainers. For example, the power usage of the number one supercomputer (position №1 in the rating, Tianhe-2, $R_{max} = 33.86 \cdot 10^{15}$ floating point operations per second or FLOPS) in November 2015 Top500 list [1] is 17.8 MW under HPL [2] benchmark. More than 40 machines on the Top 500 report power dissipation of more than 1 MW. Even a 78 kW, which is an estimate for power consumption for the machine with the same HPL result as 500 system on the Top500 rating, is a substantial amount that requires sophisticated power supply and cooling equipment in place. With these figures, the energy efficiency that can be measured in FLOPS per watt is now referred to as a key characteristic of modern high performance computing (HPC) systems. While Koomey's law [3] gives $>50\%$ energy efficiency improvement every year with the new generation of hardware platform, its also important to address the infrastructure components, like cooling subsystem, that can dramatically affect overall system efficiency. Addressing infrastructure efficiency would be also critical for attaining exascale level of performance for the future generation supercomputers, especially given power limits available in existing computing centers (e.g. notorious 20 MW power limit for an exaFLOPS-level machine [4, 5]).

A traditional quantitate measure of how efficiently energy is utilized in datacenter is power usage effectiveness (PUE) developed by The Green Grid consortium [6]. PUE is the ratio of total amount of energy used by a computer data center facility to the energy delivered to computing equipment. The ideal PUE is 1.0 when all energy is consumed by computing equipment and the datacenter infrastructure (such as cooling) takes no energy. In a real datacenter, a lot of energy is used by infrastructure equipment and PUE value is higher than 1.0. One of the most important energy consumer in datacenter is the cooling

¹ZAO "RSC Technologies", Moscow, Russia

²Lomonosov Moscow State University, Moscow, Russia

³Intel Deutschland GmbH, Munich, Germany

system: even in the optimized liquid cooled datacenters it could utilize as much as 30% of total energy supply [7]. PUE is significantly improved in modern datacenters with the advent of so-called free cooling [8]. They are specially designed for effective use of outdoor environment to remove the heat from the servers. However, free cooling is most effective if the datacenter is able to operate at high temperatures. Compactness is uniquely important for HPC design. To be efficient, the parallel processor components require as fast communication as possible. So, at the end, these components are to be placed as close to each other as possible to minimize the communication length. These rules out the use of free airflow cooling methods, sometimes applied in the datacenter design (e.g. in [9]). Nevertheless, supercomputers must be energy efficient due to the huge amount of electricity consumed by them. Additional savings are possible by reusing some of the heat, but they may require coolant temperatures to be even higher than a human could tolerate. Higher temperature coolant is usually considered to be a more energy efficient option due to the absence of chiller equipment, which reduces capital expenditure for system construction. The coldplate-based design enables the compact system setup, which is important for HPC. However, the semiconductors operating on higher temperatures may have higher leakage current resulting in degradation of energy efficiency. The interference of these two factors remains nontrivial, especially for the most recent hardware generation and multitude of semiconductors used in server design.

In the current work we have studied a traditional homogeneous server operating with hot liquid cooling setup. The liquid coolants have much better thermal properties than gases: the typical PUE values for water cooled datacenters utilizing free cooling could be less than 1.1. We have designed a coldplate with low thermal resistance: the temperature difference between the CPU top cover and liquid coolant is minimized. That design enabled us to explore server energy efficiency in a wide coolant temperature range.

The goal of this work is to measure the actual performance values of the practically important benchmark such as DGEMM for the real-world HPC solution that utilizes hot liquid cooling technology. There are many studies of hot liquid cooling systems; however, they cover mostly the heat reuse issue but not the efficiency of the computation. We have conducted a series of tests, which are described in the sections below. In the tests we studied not only the cooling efficiency (which can be considered among the best ones for high temperature coolant system), but also the impact of the coolant temperature on the server performance and the power consumption ratio.

1. Details of the experiment

The coolant temperature impact on performance and energy consumption has been studied on an individual server level. The server was thermally insulated from the environment to eliminate its impact on the experiment and to mimic a very high density packaging that is traditionally found within the blade server systems racks. The thoroughness of the isolation was examined with Fluke Ti32 thermal imager. The resulting experimental setup is presented on Figure 1.

1.1. Hardware specifications

A single RSC Tornado server was used to study the performance and energy efficiency of our liquid cooling solution. The server was based on the dual socket Intel ServerBoard S2600KPF with two Intel Xeon E5-2697 v3 (14 cores, 2.6 GHz, 145 W TDP) installed. A latest production BIOS was used. The server was also equipped with 64 GiB (8×8 GiB) of DDR4-2133 registered memory with ECC support and a single 120 GB solid state drive.

The coolant temperature was measured by thermocouples, installed on inlet and outlet of the server. The flow rate sensor was used to measure liquid flow rate through the system. The liquid-to-air heat exchanger was activated when the coolant temperature reached an experiment threshold; its efficiency was optimized to avoid the liquid temperature decline on its fan activation. The average rate of water flow in cooling system was 30 ml/s; the average difference between inlet and outlet water temperature was 3.5 °C.

1.2. Benchmark details

The DGEMM matrix multiplication benchmark was used to simulate stress-test conditions. Intel MKL (version 12.2.3) DGEMM implementation was used in this test. The dimensions of the matrices were selected to be 87936×192×87936 that fit in 58 GiB of memory and provides the stress level similar to the hot phase of the HPL benchmark.

The DGEMM kernel was running continuously for at least 1 hour for each of the different inlet temperatures of the coolant ranging from 19 to 70 °C. The performance information was collected for every DGEMM iteration during the benchmark. The power usage and temperatures of the CPU, memory as well as the whole system was monitored using out-of-band telemetry provided by Intel NodeManager. The state of liquid cooling system (temperature and inlet and outlet coolant flow rate) was also controlled during all benchmarks.

The data from the first 20-30 minutes of all benchmarks are omitted from to ensure stationary coolant temperature is reached and the temperature levels of all node components are stabilized.

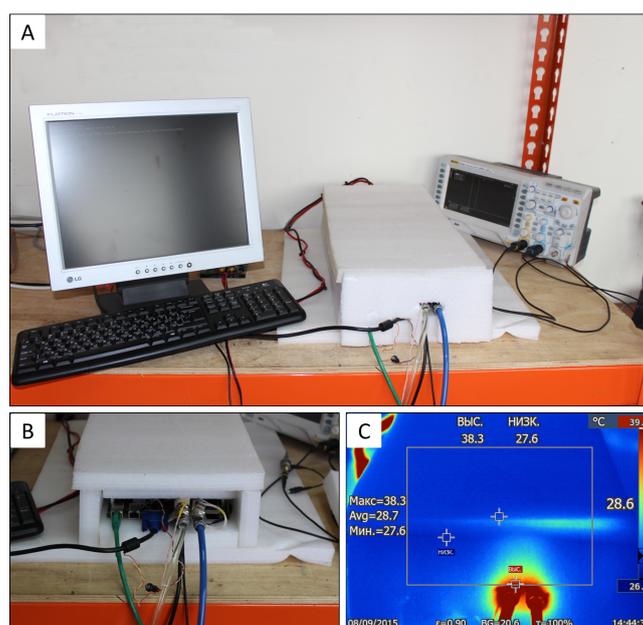


Figure 1. a) Overview of the experimental setup; b) liquid cooling and system connections; c) Fluke Ti32 view ($T_{\max} = 38.3$ °C, $T_{\min} = 27.6$ °C)

Table 1: Performance and power characteristics at different coolant temperatures

T_{LC} (°C)	Performance (GFLOPS)	System power (W)	Efficiency (GFLOPS/W)	Efficiency decrease (%)	Possible energy reuse ¹ (%)
19	974	359	2.72	–	-1.8
45	985	388	2.54	7.0	7.8
50	985	390	2.53	7.5	9.7
55	981	395	2.49	9.4	10.8
60	976	396	2.46	10.3	12.3
65	969	398	2.44	11.5	14.9

¹ assuming environment temperature $T_0 = 25$ °C

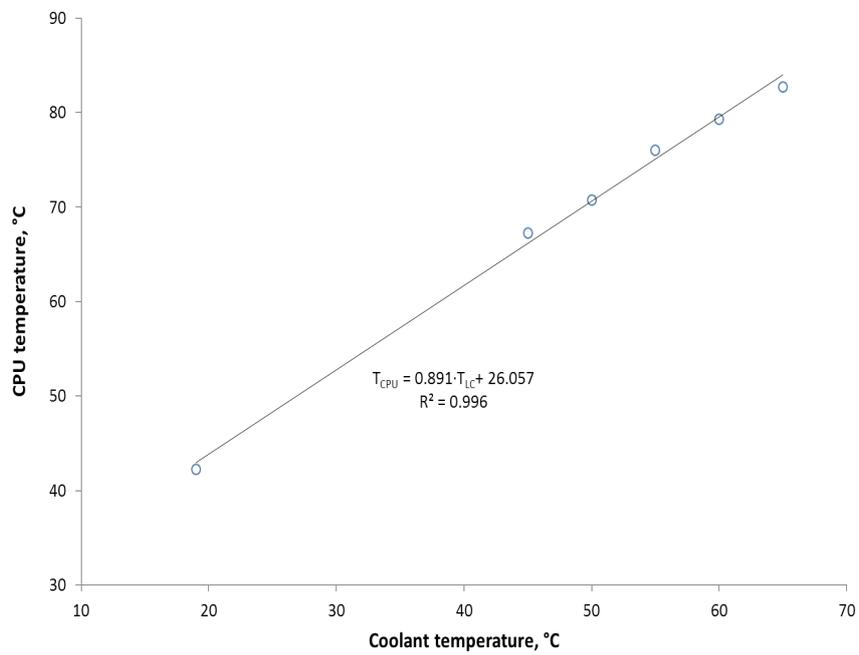


Figure 2. Dependence of the average measured CPU sensor temperature on coolant inlet temperature.

T_{CPU} - CPU temperature, T_{LC} - coolant temperature

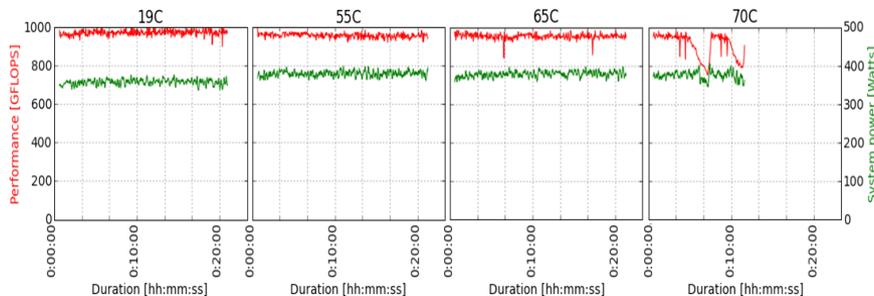


Figure 3. Illustration of system performance and power profiles at $T_{LC} = 19, 55, 65$ and 70 °C

2. Results

2.1. Liquid cooling performance

As one can see on the Figure 2, the liquid cooling solution that was used in this study has relatively low thermal resistance. The average difference between CPU (T_{CPU}) and liquid coolant (T_{LC}) temper-

atures is about 26 °C even at $T_{LC} = 60-65$ °C. The actual thermal resistance value was estimated to be 0.74 °C·in²·W⁻¹ that is on-par with commercially available coldplates (eq. Lytron plates [10]). We also monitored the data from the distributed temperature sensing (DTS) devices integrated in CPUs. These sensors show the difference between current and critical CPU temperature, at which it is considered to be overheated [11]. In this case DTS sensors return to zero. In current work we never get zero DTS value up to $T_{LC} = 65$ °C. However, at this coolant temperature the average DTS value corresponds to the temperature, which is only 5° below critical.

At the low temperature the performance is very stable for a long period of time and only at $T_{LC} = 65$ °C a very few and spiky performance drops are observed (Figure 3). Further growth of T_{LC} would result in scaling down of the CPU clock rate [12]. The performance of DGEMM is known to be linearly dependent on the CPU clock [13] and performance degradation would be observed on higher temperatures. Actually, our tests at $T_{LC} = 70$ °C show significant drop in performance. Almost half of the benchmark time the performance at 70 °C is 100-200 GFLOPS lower than the one at 65 °C. At the same time the power consumption changes only a little and a decrease of efficiency is also observed. For these reasons we don't refer to 70 °C results further.

2.2. Computational performance and energy efficiency

As it was discussed in previous section, the studied liquid cooling solution is very effective up to the 65 °C coolant temperature (Table 1). Indeed, upon this temperature we didn't observe any significant decrease of the average performance of the system (actual variations are $\approx 1-2\%$). At the same time, the average system power consumption is almost linearly dependent on temperature. These trends lead to efficiency decrease of $\approx 2.2\%$ at every 10 °C growth of coolant temperature. Actually, the observed total change in efficiency is 11.5% upon growth of coolant temperature from 19 to 65 °C.

We also provided basic exergy analysis of possible hot water energy reuse. Exergy is the thermodynamic quantity that corresponds to the amount of energy which could be utilized for useful work. The following equation was used for exergy production rate (E_Q) calculation [14]:

$$\dot{E}_Q = \dot{m}C_p \left((T_{out} - T_{in}) - T_0 \ln \frac{T_{out}}{T_{in}} \right) \quad (1)$$

where \dot{m} stands for water flow; C_p is heat capacity at constant pressure; T_0 , T_{in} , T_{out} are environment, inlet and outlet coolant temperatures respectively. The estimates on energy reuse at $T_0 = 25$ °C are presented in Table 1. One could see that theoretical amount of energy, which could be reused would overcompensate the decrease of efficiency. However, the amount of reused energy would always be lower in real applications.

2.3. Power consumption of system components analysis

Modern Intel hardware allows fine-grain power profiling for many system components. In this study, we collected data on memory and CPU power consumption. As expected, the power consumption of CPU (Figure 4) and memory (21 W over all tests) modules depend only slightly ($\approx 1\%$) on temperature. The tests were conducted with Turbo Boost feature enabled that adapts CPU performance to fit in the power limit. Moreover Haswell uses an improved version of Intel's 22-nm process where leakage current is relatively small [15].

The contribution to power consumption from other system components is much more sensitive to the temperature (Figure 5). The largest contribution to residual power consumption is made by the chipset components and voltage regulators. They don't have such power control and their power con-

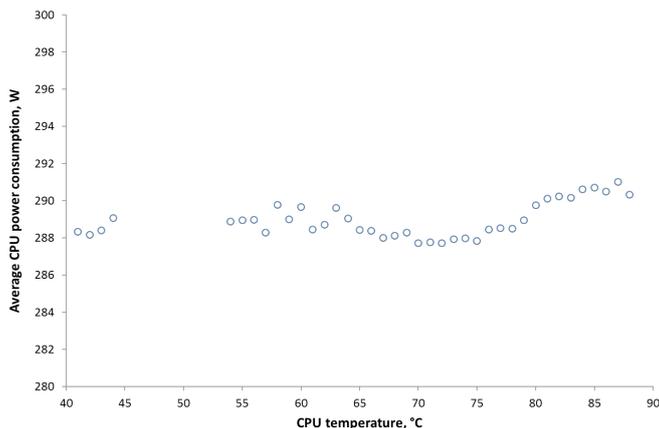


Figure 4. Dependence of CPU power consumption on its temperature. Gap at 45-55 °C is due to the lack of the data for this temperature interval

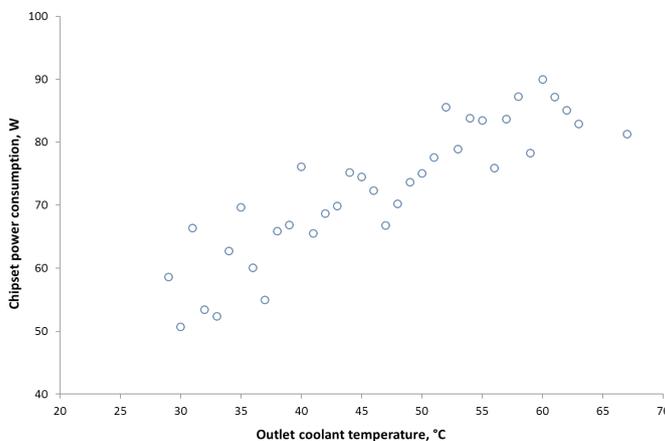


Figure 5. Dependence of the chipset and voltage regulators power consumption on temperature. The system temperature was estimated as the outlet coolant temperature. Its real value could be 10-20 °C higher

sumption grows with temperature due to increase of leakage current [16]. In our setup the residual power consumption almost doubles when the temperature rises from 30 to 65 °C.

3. Related work

An important problem of hot liquid cooling systems is the possibility of energy reuse. It should be stressed, that the gain from free cooling and energy recuperation/reuse should compensate computational efficiency decrease. Switching to the free cooling already provides notable energy efficiency growth. The problem of the heat reuse is much more complex. The most efficient way is to use hot water for heating [17, 18], however, it is very climate and country specific and not always possible. In many cases HPC datacenter designers have to think of possible energy conversion in useful work, however its mount is limited by thermodynamics second law. Energy recovery is most effective at very high coolant temperatures, but there are restrictions imposed by the hardware.

We observe a modest decrease (about 10%) of power efficiency when coolant temperature increases from 19 to 65 °C. The similar results have been obtained for other contemporary hot water cooling solutions, namely Aquasar (7% power consumption increase from 30 to 60 °C), CoolMUC (5% power

consumption increase from 30 to 50 °C) and iDataCool (7% of efficiency decrease from 49 to 70 °C) [18–20]. In the study of Aquasar hot liquid cooling system [18] the exergy analysis showed about 10% of possible energy reuse at 50-60 °C water temperature. Our estimates also showed that the heat reuse could overcompensate the performance loss. Some additional gain is also possible due to free cooling. Thus the use of hot coolant is reasonable in datacenters and would result into reduction of their PUE value. One of such success stories of hot liquid cooling usage was presented by Eurotech [21]. They reported PUE value of 1.05 at 50 °C coolant temperature, however they didn't perform a detailed performance analysis that includes FLOPS/W metrics.

4. Conclusion

In this paper we explored the performance and power profiles of liquid cooling equipment with hot water, designed for RSC Tornado supercomputer architecture. This hardware exemplifies modern server platform with high temperature coolant option available due to appropriate design. Since the commodity semiconductor components are used (CPU, memory modules, server board), our results provide the insight on the modern hardware in general.

The power efficiency of our benchmark server is about 2.5 Gflops/W that is comparable to top 40 devices of current (November 2015) Green500 list [22]. Taking into account the possible energy recuperation will further increase efficiency of HPC system. We also demonstrated that the power efficiency decrease upon the coolant temperature growth should be compensated by possible heat reuse. A number of options for such reuse exist, starting from facility and building heating to adsorption by chiller machines. Climate-wise, 40-60 °C coolant temperatures enable free cooling on most of the Earth 24×7, except desert areas.

We are looking for the future to study behavior of the larger cluster with high temperature coolant, which is expected to be available in 2016, with respect to efficiency, reliability and ease of maintenance.

This paper is distributed under the terms of the Creative Commons Attribution-Non Commercial 3.0 License which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is properly cited.

References

1. TOP500 supercomputer sites. [Online]. Available: <http://www.top500.org>
2. A. Petitet, R. C. Whaley, J. Dongarra, and A. Cleary. (2008) HPL - a portable implementation of the high-performance Linpack benchmark for distributed-memory computers. [Online]. Available: <http://www.netlib.org/benchmark/hpl/>
3. J. Koomey, S. Berard, M. Sanchez, and H. Wong, "Implications of historical trends in the electrical efficiency of computing," *IEEE Annals of the History of Computing*, vol. 33, no. 3, pp. 46–54, 2011.
4. P. Kogge, K. Bergman, S. Borkar, D. Campbell, W. Carson, W. Dally, M. Denneau, P. Franzon, W. Harrod, K. Hill, J. Hiller, M. Richards, and A. Snavelly, "ExaScale computing study: technology challenges in achieving exascale systems," *Government PROcurement*, vol. TR-2008-13, p. 278, 2008.
5. D. Wade, "ASC business plan (NA-ASC-104R-15-Vol.1-Rev.0)," Tech. Rep., 2015.
6. J. Haas, J. Froedge, J. Pflueger, and D. Azevedo, *Usage and public reporting guidelines for The Green Grid's infrastructure metrics (PUE/DCiE)*, The Green Grid, 2009. [Online]. Available: <http://>

www.thegreengrid.org/~media/WhitePapers/WhitePaper22PUEDCiEUUsageGuidelinesfinalv21.pdf

7. M. P. David, M. Iyengar, P. Parida, R. Simons, M. Schultz, M. Gaynes, R. Schmidt, and T. Chainer, "Experimental characterization of an energy efficient chiller-less data center test facility with warm water cooled servers," in *2012 28th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM)*. IEEE, 2012, pp. 232–237.
8. C. Gough, I. Steiner, and W. A. Saunders, *Energy efficient servers: blueprints for data center optimization*, 1st ed. Apress, 2015.
9. Yahoo launches second 'computing coop' data center in New York state. [Online]. Available: <http://www.datacenterknowledge.com/archives/2015/04/27/second-yahoo-data-center-comes-online-in-new-york-state/>
10. The performance of standard cold plate technologies is compared in a graph showing local thermal resistance. Lytron Inc. [Online]. Available: <http://www.lytron.com/Cold-Plates/Standard/Performance-Comparison>
11. M. Berktold and T. Tian, *CPU monitoring with DTS/PECI*, 2010. [Online]. Available: <http://www.intel.com/content/www/us/en/embedded/testing-and-validation/cpu-monitoring-dts-peci-paper.html>
12. *Intel Turbo Boost technology 2.0*, Intel. [Online]. Available: <http://www.intel.com/technology/turboboost/>
13. J. Demmel and A. Gearhart, "Instrumenting linear algebra energy consumption via on-chip energy counters," UC at Berkeley, Tech. Rep., 2012. [Online]. Available: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-168.html>
14. G. Evola and L. Marletta, "Exergy and thermoeconomic optimization of a water-cooled glazed hybrid photovoltaic/thermal (PVT) collector," *Solar Energy*, vol. 107, pp. 12–25, 2014.
15. P. Hammarlund, A. J. Martinez, A. A. Bajwa, D. L. Hill, E. Hallnor, H. Jiang, M. Dixon, M. Derr, M. Hunsaker, R. Kumar, R. B. Osborne, R. Rajwar, R. Singhal, R. D'Sa, R. Chappell, S. Kaushik, S. Chennupaty, S. Jourdan, S. Gunther, T. Piazza, and T. Burton, "Haswell: the fourth-generation Intel Core processor," *IEEE Micro*, vol. 34, no. 2, pp. 6–20, 2014.
16. Nam Sung Kim, T. Austin, D. Blaauw, T. Mudge, K. Flautner, Jie S. Hu, M. Irwin, M. Kandemir, and V. Narayanan, "Leakage current: Moore's law meets static power," *Computer*, vol. 36, no. 12, pp. 68–75, 2003.
17. S. Zimmermann, M. K. Tiwari, I. Meijer, S. Paredes, B. Michel, and D. Poulikakos, "Hot water cooled electronics: exergy analysis and waste heat reuse feasibility," *International Journal of Heat and Mass Transfer*, vol. 55, no. 23-24, pp. 6391–6399, 2012.
18. S. Zimmermann, I. Meijer, M. K. Tiwari, S. Paredes, B. Michel, and D. Poulikakos, "Aquasar: a hot water cooled data center with direct energy reuse," *Energy*, vol. 43, no. 1, pp. 237–245, 2012.
19. N. Meyer, M. Ries, S. Solbrig, and T. Wettig, "iDataCool: HPC with hot-water cooling and energy reuse," in *Supercomputing: 28th International Supercomputing Conference*, 2013, pp. 383–394.
20. A. Auweter and H. Huber, "Direct warm water cooled Linux cluster Munich: CoolMUC," *Inside*, vol. 10, no. 1, pp. 81–82, 2012.
21. Hot water cooled supercomputer. Eurotech. [Online]. Available: <http://www.eurotech.com/en/hpc/hpc+solutions/liquid+cooling>

22. The Green500 list. [Online]. Available: <http://www.green500.org>