COMMERCE DEPARTMENT
ENERGY RESOURCES

Liquid Cooling of Data Center Servers

Energy Impacts of Liquid Cooling Servers in Small- and Medium-Sized Data Centers in Minnesota

12/22/2017

Contract # 105475

Conservation Applied Research and Development (CARD) FINAL Report

Prepared for: Minnesota Department of Commerce, Division of Energy Resources
Prepared by: GDS Associates

Prepared by:
Travis Hinck

GDS Associates,
440 Science Drive, Suite 400
Madison, WI, 53711
Phone: 612-916-3052

© 2017 GDS Associates. All rights reserved.

Contract Number: 105475

Prepared for Minnesota Department of Commerce, Division of Energy Resources:
Jessica Looman, Commissioner, Department of Commerce
Bill Grant, Deputy Commissioner, Department of Commerce, Division of Energy Resources
Mark Garofano, Project Manager
651-539-1864
Mark.Garofano@state.mn.us

ACKNOWLEGEMENTS

This project was supported in part by a grant from the Minnesota Department of Commerce, Division of Energy Resources, through the Conservation Applied Research and Development (CARD) program, which is funded by Minnesota ratepayers.

The authors would also like to acknowledge the following (individuals, utilities, and/or organizations and/or other entities) for their financial, in-kind, or other contributions to the project:

LiquidCool Solutions, LLC
Dr. Brad Erickson and team, Mayo Clinic
Ebullient Cooling

DISCLAIMER

This report does not necessarily represent the view(s), opinion(s), or position(s) of the Minnesota Department of Commerce (Commerce), its employees or the State of Minnesota (State). When applicable, the State will evaluate the results of this research for inclusion in Conservation Improvement Program (CIP) portfolios and communicate its recommendations in separate document(s).

Commerce, the State, its employees, contractors, subcontractors, project participants, the organizations listed herein, or any person on behalf of any of the organizations mentioned herein make no warranty, express or implied, with respect to the use of any information, apparatus, method, or process disclosed in this document. Furthermore, the aforementioned parties assume no liability for the information in this report with respect to the use of, or damages resulting from the use of, any information, apparatus, method, or process disclosed in this document; nor does any party represent that the use of this information will not infringe upon privately owned rights.

Table of Contents

Table of Contents **1**

List of Figures **2**

List of Tables **2**

Executive Summary **3**

Background **4**

Technology 4

Promising Potential 6

Barriers to Adoption 8

Methodology **9**

Site Selection and System Design 9

Equipment Definitions 11

Variable Definitions 12

Test Methodology 14

Evaluation Metrics 17

Results **20**

Summary of Data 20

Summary of Results 22

Discussion of Results **23**

Field Demonstration 23

Conclusions and Recommendations **29**

Data Center Operator Recommendations 29

References **31**

Appendix A: Liquid Cooling Product Specifications **32**

LCS Product Specifications 32

Ebullient Product Specifications 33

Appendix B: Project Marketing Materials **35**

Appendix C: Statement From Ebullient Cooling **37**

List of Figures

Figure 1: LCS LCC200 servers fully submerged in liquid coolant 5
Figure 2: Ebullient direct-to-chip liquid cooling heat sink 6
Figure 3: Ebullient DirectJet heat exchanger installed in a server rack – spec drawing 6
Figure 4: Hardwick building data center 9
Figure 5: University of Northwestern candidate site 10

List of Tables

Table 1: LCS Test Equipment Descriptions 11
Table 2: LCS Test Case Definitions 16
Table 3: Format of Summary Test Result Evaluation 19
Table 4: Summary of Collected LCS Test Data 20
Table 5: Summary of Calculated Values from LCS Test 21
Table 6: Summary of LCS Test Results 22

Executive Summary

GDS Associates was commissioned by the Minnesota Department of Commerce to investigate the viability and energy impacts of two new data center liquid cooling technologies. The products are a full-submersion technology developed by LiquidCool Solutions (LCS) out of Rochester, MN and a direct-to-chip retrofit technology manufactured by Ebullient Cooling in Madison, WI. The emerging technologies offer promising potential to reduce energy consumption at data center facilities, specifically small- and medium-sized sites that are often overlooked for energy conservation due to unique site characteristics and lack of reliable, predicable savings measures available.

The study produced mixed overall results. It was not possible to test the Ebullient technology at all due to underestimated costs of the planned installation. This suggests the technology is not yet mature enough to recommend deployment to Minnesota data centers at this time. The LCS product completed a successful field test to demonstrate the viability of the technology and verify the predicted energy consumption trends that indicate conservation potential. The precise energy impacts were difficult to characterize at a high level of confidence due to complications arising at the test site. Therefore, findings are presented as a range of possible outcomes rather than precise energy savings predictions.

The LCS technology was observed to deliver approximately 38%-75% reduction in cooling energy required. This corresponds to a 4.6%-10.1% reduction in overall energy consumption (counting server and cooling combined consumption). With appropriate vetting of sites and planning to optimize data center operations to take advantage of liquid cooling, most similar data center sites should be capable of achieving energy conservation closer to the high end of the observed savings range.

Several criteria indicate whether liquid cooling may be a cost-effective option at a site. At this time, only sites planning new construction or major renovation, where choosing liquid cooling can avoid the cost of purchasing a new air-cooling CRAC unit, should be considered for deployment. Without that savings, sites are unlikely to benefit enough to justify the costs of a liquid cooling strategy. Additionally, the greater a data center's server processor utilization rate, the more it will benefit from liquid cooling. Specific sites may be able to improve the value of liquid cooling technology by capturing additional benefits such as waste heat recovery, reduced maintenance costs, smaller equipment footprint, quieter operations, and improved server performance under high process loads.

The study results suggest that data center operators should characterize energy consumption at their site and review the criteria outlined above to consider whether liquid cooling may be a cost-effective option for them. Utility implementers should familiarize themselves with the technology and be on the lookout for ideal candidate sites where their customers may benefit from liquid cooling. A conservative deemed savings is possible to develop for specific sites drawing on the study's findings, but at this point it is not yet possible to develop an accurate prescriptive method of estimating savings across a wide range of potential sites due to the small sample size and variability of savings observed over the course of this study.

Background

Technology

Data centers are designed to house energy-intensive computing equipment - typically servers used to process and store data. A significant portion of the energy consumption at data centers is used for cooling. Exact percentages vary significantly across data centers, but typically at least 15% of total energy consumed is used for cooling air-cooled sites with some using as much as 40%. In particular, smaller data centers tend to be less efficient. Most of these data centers have dedicated Computer Room Air Conditioner or Computer Room Air Handler units (CRAC or CRAH) to cool the ambient air in the data center and then pass it over heat sinks with high-speed fans mounted near, or directly on, each server.

Using liquid as the cooling medium instead of air has several advantages that improve energy efficiency at data centers:

- Liquid has a much higher heat capacity than air, so it can provide the same cooling with much lower volume.
- Liquid can be applied to high power density components directly rather than cooling the ambient temperature of the entire data center, which reduces the overall cooling capacity required.
- Waste heat can be rejected from liquid coolant to an existing cooling loop, or even recovered for useful purposes, much easier than air-cooled heat rejection.
- There is the potential for data centers designed for liquid cooling to eliminate dedicated air cooling units and local high-speed fans completely.

This project focuses on two specific products that use liquid cooling technology. Both can be installed in small- and medium-sized data centers and can be used alongside traditional air-cooled equipment or without air cooling. Existing servers can continue running as usual throughout the test period. The goal is to demonstrate a successful implementation in functional data centers in Minnesota without interrupting operations and to characterize energy savings potential.

LCS Technology

The technology evaluated is an innovative rack-mounted, total-liquid-submersion-cooled server manufactured by Liquid Cool Solutions, a Minnesota-based company. The specific product tested is a LSS200 server. The server is completely submerged in CoreCoolant™, which is a non-hazardous, dielectric fluid with 1400 times the heat capacity of air. Using a dielectric (electrically insulating) coolant means that the electronic components of the server can operate normally while in contact with, or even fully submerged in, the fluid. Coolant is circulated directly to the components with the highest power density, which efficiently sinks heat with as little fluid flow as possible, which in turn allows the coolant temperature setpoint to operate as high as possible without impacting the performance of the server.

The temperature of input coolant can be as high as 45C (113F) and still cool components. This has two important implications. First, heat rejection does not necessarily require access to a cooling loop or outdoor condenser unit. Heat can be moved out of the data center and rejected into ambient indoor air via a dedicated Cooling Distribution Unit (CDU), also designed by LCS. The second implication is that waste heat recovery is possible even for HVAC applications. This means that besides cooling efficiently, most sites have the potential to capture waste heat from servers for useful purposes, even if they don't have a process heat load nearby.

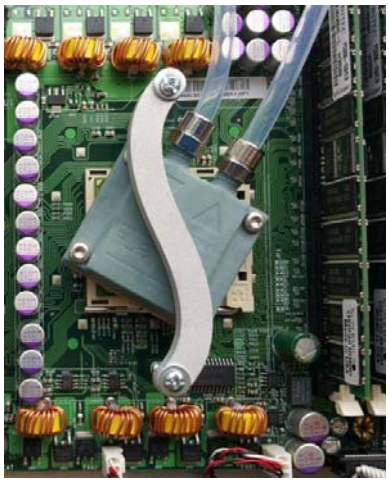
Figure 1: LCS LCC200 servers fully submerged in liquid coolant

Ebullient Technology

The second technology is a direct-to-chip liquid cooling strategy developed by Ebullient Cooling out of Madison, WI. Similar to the LCS product, the system uses a dielectric (electrically insulating) fluid as a coolant, which ensures that the system is not introducing the possibility of catastrophic failure of the electronics in the way using a water cooling loop would. Coolant is circulated via hose directly to the heat sink on the processor of each individual server. Server heat is absorbed through vaporization of the coolant, which is recirculated back to the fluid distribution unit heat exchanger and then can be rejected to a facility cooling loop or dedicated heat rejection system.

A key advantage of the Ebullient system is that it can be installed as a retrofit on existing servers by replacing existing air-cooled heat sinks with the liquid-cooled equivalent without any other required upgrades. That is, the data center doesn't have to plan to purchase new servers, their existing servers in their existing configuration can benefit from the retrofit. Further, the fluid distribution system is designed to fit in the footprint of a server rack, so no specialized equipment is necessary beyond the Ebullient system assuming that there is a facility cooling loop available.

Figure 2: Ebullient direct-to-chip liquid cooling heat sink



Two specific configurations were proposed for testing, the ER-25 and ES-60 systems. They are designed to cool up to 25 and 60 servers, respectively. The objective of testing these units was to show that small- and medium-sized data centers can benefit from retrofitting existing equipment with a package unit to convert to liquid cooling and capitalize on the energy savings potential without a complete overhaul of the data center (in fact with very little server downtime to complete the retrofit).

Figure 3: Ebullient DirectJet heat exchanger installed in a server rack – spec drawing

Promising Potential

Liquid cooling strategies in data centers, including the two specific technologies proposed for testing, show promising potential to deliver energy efficiency savings by reducing the energy required to cool while having no impact, or even a positive impact, on server processing capacity. In particular, these technologies can be installed in small- and medium-sized data centers that often are not as energy

efficient as larger facilities. A previous CARD grant estimated that there are over 3,600 such smaller data centers in the state of Minnesota. This means that if the technology reliably delivers savings, there will be significant potential for conservation statewide. One of the goals of this project is to show that potential energy savings will actually materialize in a field setting and to confirm that the technologies will function in the targeted smaller-sized data centers.

LCS Potential

An NREL study is being conducted concurrently with this study focused on calculating energy impacts of the LCS technology in a laboratory setting. Phase 1 of that study resulted in a white paper concluding in the validation of several clear advantages of Liquid Cool submersion technology compared to air cooling. The NREL report includes the following key findings:

- LCS liquid submersion technology reduced system level power-to-cool and associated costs by as much as 98% compared to air-cooling. This represents up to 18% of total server power usage (note: under extreme computational load conditions)
- Total power consumption can be reduced by up to 26% when accounting for reduction in leakage current enabled by lower temperature operation of liquid-cooled servers (note: again, under extreme computational load conditions)
- Computational performance of LCS servers was better than or equal to comparable air-cooled servers at all ambient temperatures and all computational loads
- The LCS power-to-cool remained constant at very low power usage regardless of computational load or coolant temperature between 15°C and 45°C. The power-to-cool of the air-cooled servers increased significantly as computational load and ambient air temperature increased.

The laboratory setting results are impressive. It focuses on the potential savings under extreme computational loading conditions, which illustrates an upper bound of potential energy conservation potential. The findings are complimentary to the goals of this project, which are to demonstrate viability of the technology in a field setting and to estimate energy impacts under typical operating conditions.

Phase 2 of the NREL study will focus on additional benefits of the liquid cooling technology including waste heat recovery potential, better allocation of computational assets (no longer constrained by air-cooling limitations), and optimizing total system energy flow including processor power, leakage current at different operating conditions, and cooling power. Preliminary Phase 2 results are scheduled to be available in late 2017.

Ebullient Potential

The manufacturer has been running a cooling system on a rack of 19 servers continuously since October 2013. The servers are run under typical loading conditions, so the setup simulates real-world applications of the technology. The system continues to operate, demonstrating functionality of the technology. Estimates made by the manufacturer show that Ebullient cooling technology can reduce

cooling energy (and costs) by approximately 74% compared to air-cooling the same units. The goal of this study is to provide third-party verification of these findings in a functional data center.

Barriers to Adoption

There are two main reasons liquid-cooling technology has not yet been widely adopted organically despite the potential for energy savings: concern for effects on reliability and the uncertain value proposition of the technology. The proposed research addresses both issues to some degree.

The top priority of data center operators is reliability. Any change made to their facility, even if it may have some benefits, is inherently difficult to adopt due to uncertainty of its effects on maintaining uptime. Liquid-cooling technology necessarily introduces liquid to an electronics environment, which seems counterintuitive to safely operating electronics. Often, the natural response from data center operators is to err on the side of caution and decide to avoid a liquid-based technology without looking into details of its operation.

The coolant used by both technologies is a dielectric liquid, which means that even if the liquid containment were to fail and spill into the data center, no damage would be done to electronic components. This is already well understood, but this project is an opportunity to demonstrate safe, reliable installations of the technology in a functional data center without any unintended side effects or server downtime issues.

The second barrier blocking adoption of the technology is the uncertain value proposition because both the costs and benefits were only estimates before this project. The technologies are still in early developmental stages, and they are not yet mature products with stable pricing. Over the course of the project, actual costs incurred were recorded and used to develop more accurate estimates of costs for future installations (actual, direct costs were recorded for this project, but since it is a prototype design with unique costs, an attempt was made to adjust those costs to better reflect expected costs of future projects).

Most of the benefit of the technology is the reduction in cooling energy consumption and associated energy cost savings predicted by models and lab tests. The main objective of the research is to measure the potential energy savings that can be achieved by the technology in an actual field setting. The study thereby addresses major barriers to adoption by clarifying costs and savings as well as demonstrating safe and reliable installations are possible in real applications.

Methodology

The research plan changed several times over the course of this project because many issues arose with the technologies, site selection, and cost estimates. For the LCS technology, the test plan encountered unexpected uncontrolled variables, which were dealt with as well as possible. For the Ebullient technology, finding a suitable site to install the system proved too difficult to complete. The original plan called for three separate sites, two Ebullient and one LCS, as well as a comparison between the two technologies in terms of energy conservation potential and cost-effectiveness. This section outlines the final research plan for the one site and one technology that were evaluated in the final results as well as a description of issues that led to deviations from the original plan.

Site Selection and System Design

LCS Site Selection

The product manufacturer had already identified a potential installation site when the project commenced. The site is a data center housed in the basement of the Hardwick building on the Mayo Clinic campus in Rochester, Minnesota. The manufacturer worked with the site owner to design a specific system to meet the site needs. The system included twelve LSS200 liquid-cooled servers installed in existing rack space. The system design also called for two dedicated cooling distribution units installed in the ceiling space just outside the data center room as well as the associated equipment required to pump coolant from the servers to the cooling distribution units and all miscellaneous parts and labor required to complete the install. Full system specifications are included in Appendix A.

Figure 4: Hardwick building data center

Ebullient Site Selection

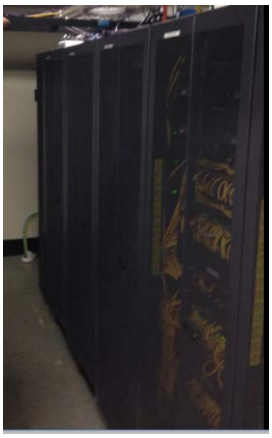
Site selection for the installation of the Ebullient product did not go well. The original proposal was to build off a list of Minnesota data centers developed by an earlier CARD grant. The list included approximately 50 sites identified as small- or medium-sized data centers that would be good candidates to participate in this study. From the 50 sites, it was assumed that at least two would be willing to participate when presented with the opportunity (at an estimated low-, or even no-, cost to the operator).

Unfortunately, there was miscommunication among project team members. The majority of data centers classified as small were actually closet-sized with less than a full rack of servers present. Of the 50 candidate sites, only two were appropriately sized for installing the retrofit cooling equipment and both of those sites had other constraints that prevented participation.

The backup plan for site identification included multiple efforts. The project team reached out to contacts at utilities who work with data center customers. A promotional flyer was designed and circulated to possible sites - this was essentially a cold-call outreach effort (see Appendix B for the final version of the flyer). These efforts had a very low hit rate resulting in a handful of slightly-interested data center operators, but none who ultimately chose to participate in the project.

Eventually the project team identified a participant (with help from MNTAP and their network of connections). The University of Northwestern, St. Paul campus provided a site that initially appeared to be nearly ideal. The data center housed eight racks of servers, which is the perfect size for the study. The room was clean and well organized with some available rack space to install the cooling unit as well as a cabinet design that could easily incorporate the liquid coolant distribution manifold. The server equipment had dedicated circuits from the electric panel that were simple to isolate for power consumption data collection. Additionally, the operators could not have been more accommodating and willing to help with the study.

Figure 5: University of Northwestern candidate site



Unfortunately, one important detail of the site was missed in the initial assessment. There is no nearby facility cooling loop available to tap into for heat rejection. The manufacturer assumed that there was

either a loop available or that a dedicated heat rejection unit could be installed for minimal cost. As the design process progressed, the lack of a cooling loop was noted and a dedicated heat rejection system was planned. The proposed system consisted of an air-cooled condenser on the roof and a refrigerant loop plumbed to the DirectJet unit in the data center. Several bids were collected from local plumbing/HVAC contractors to install the dedicated heat rejection system, but the cost quotes were much higher than estimated (approximately $40,000 rather than the assumed < $10,000). According to the manufacturer, the high cost of the heat rejection system is atypical and results from unique characteristics of this particular site. In most data centers, there will be a nearby cooling loop to tap into for very little cost or the location of the room should allow a dedicated heat rejection plan that is not cost-prohibitive. However, in this case, the planned installation was not possible within the available budget. There was also some staff turnover at the manufacturer during this time period, which complicated the planning process. Some details of the site design may have fallen through the cracks as the project was handed off to new engineers.

By the time the Northwestern site had to be rejected for cost issues, the project schedule had already stretched well past the originally planned completion date and the LCS portion of the project had been complete for several months. At this time, rather than starting from scratch to attempt to identify another participant, Ebullient withdrew from the project and the scope of the test plan had to be narrowed. Appendix C includes a statement from Ebullient explaining their decision to withdraw from the project.

Equipment Definitions

Table 1 lists the equipment definitions for the test setup of the LCS product.

Table 1: LCS Test Equipment Descriptions

Equipment Label	Description of Equipment
RIL-GPU2 server	The largest and most energy intensive existing server in the Data Center
CRAH unit	The existing air handling unit dedicated to cooling the Data Center. McQuay D010W6 unit. 460V/3φ. 5HP fan motor controlled by VFD
LCS Server	Newly installed LSS200 liquid-cooled servers. Dual Intel Xeon 5670, 2.93GHz CPUs. NVidia Tesla M2090 GPU. 48GB DDR3 Memory. 2X 250GB SATA III 6Gb/s SSDs (RAID-1) storage. 12 servers installed

Equipment Label	Description of Equipment
CDU	Cooling Distribution Units. Koolance ERM-3K3U. Used to pump liquid coolant to servers and exchange heat from coolant to ambient air. 2 units installed
Eaton Power Unit Log	The Eaton unit monitors energy consumption and switches battery backup online in cases of power failure. High level, low resolution data
Grid Engine Job Logger	Generates a record of requests for server processing jobs. Includes size of job (GB processed), server time used, and server assigned processing responsibility
HOBO dataloggers	Monitoring equipment used to collect current draw and temperature data. Used CT clamp and J-type thermocouple sensors

Variable Definitions

The following variable definitions apply to the test setup for the LCS product at the Hardwick building. Separate variables were defined for the proposed Ebullient installation at the Northwestern site, but were not used when that portion of the scope was eliminated.

Variables Collected by Direct Measurement

- Existing RIL-GPU2 server power draw. Average value calculated from measurements of current draw at the rack power strip into the server power supply units over the course of the test period. Four extension cords were modified and a CT clamp was added to each to allow a HOBO logger to collect 5-second-interval data on all four power supply input lines over the duration of the baseline test (case 1).
- CRAH unit power draw. Average value calculated by measuring current draw on one phase of the three-phase input to the dedicated McQuay Computer Room Air Handler unit over the course of the test period. A HOBO logger with CT clamp sensor was used to collect 1-second interval data throughout all test cases. A preliminary test early in the project showed that the phases are well balanced, so one phase measured is enough to calculate total power draw for the unit. Used in all test cases.
- LCS server power draw. Average value calculated from measurements of current draw at the rack power strip into our individual server power supply units over the course of the test period. Four extension cords were modified and a CT clamp was added to each to allow a HOBO logger

to collect 5-second-interval data on four power supply input lines over the duration of test cases 2, 3, and 4.

- CDU (Cooling Distribution Unit) power draw. Calculated from measurements of current draw at the unit installed in the ceiling space above the hallway outside the data center over the course of the test period. Two extension cords were modified and CT clamp sensors attached to each to allow a HOBO logger to collect 5-second interval data over the course of test cases 2, 3, and 4.
- Air-cooled total power draw (kW). Average value calculated from measurements collected from the Eaton interval data log of room power consumption drawn by servers and CRAH unit over the test period. There is one measured variable showing the total power draw of both the servers and CRAH unit for each interval. The log includes voltage, current, and power readings at 1-second intervals for the entire data center. Collected for all test cases.
- Room Temperature. Collected above the main server rack in the data center. A HOBO logger with a J-type thermocouple sensor collected 5-second interval data during all test cases.
- Grid Engine job log. A history of all data processing requests sent to the data center servers. Job size (GB processed), start time, and end time collected by the data center Grid Engine accounting file. This data is used at a high level to compare the processing load for the existing RIL-GPU2 server across test cases.

Calculated Variables

- Air-cooled Server power draw (kW). Average value calculated by subtracting the CRAH power draw (measured variable) from the total air-cooled power draw (from measured variables) at each interval over the test period. Used for all test cases.
- Air cooled % cooling energy. Calculated by dividing CRAH unit energy drawn (measured variable) by the total air-cooled system energy drawn (measured variable) over the course of the test period. Used for all test cases.
- Liquid-cooled % cooling energy. Calculated by dividing CDU energy drawn by the total liquid-cooled system energy drawn over the course of the test period. Used in test cases 2, 3, and 4.
- Potential reduction in cooling energy. Calculated by dividing the difference between air cooled and liquid cooled system energies drawn (calculated variables) by the air-cooled system energy over the course of the test period. Used in test cases 2, 3, and 4.
- Extrapolated air-cooled annual energy (kWh). Calculated by multiplying average air-cooled power draw (measured variable) over the course of the test period by an assumed full year of operation (8766 hours). Used in all test cases.
- Counter factual extrapolated annual liquid cooling annual energy (kWh). Calculated by multiplying average air-cooled server power draw (calculated variable) by [1 plus liquid-cooled % energy (calculated variable)] over the course of the test period and then multiplying by an assumed full year of operation (8766 hours). Used in test cases 2, 3, and 4. This value is the

potential energy savings achieved in the counterfactual case assuming the existing air-cooled servers were instead cooled with the liquid technology at the observed efficiencies of both systems.

- Possible annual energy savings (kWh). Calculated by subtracting the total extrapolated annual air-cooled energy from the counterfactual extrapolated annual liquid-cooled annual energy (calculated variables). Calculated for test cases 2, 3, and 4. Note that this is *not* the actual energy savings delivered by the installed system. It is the calculated estimated savings that would be possible if the existing air-cooled system were replaced with a liquid-cooled system at the site.
- Potential total energy savings (%). Calculated by dividing possible annual energy savings by extrapolated annual air-cooled energy (calculated variables). Calculated for test cases 2, 3, and 4. This is the percentage of total energy consumption by the air-cooled system (servers and CRAH unit) that could be conserved if they were instead cooled with the liquid cooling technology at the observed efficiencies of both systems under the test conditions.

Test Methodology

The general objective of the research project was to compare air-cooled server energy consumption to liquid-cooled server energy consumption. In both cases, the energy consumed by the servers themselves as well as the energy required for cooling needs to be accounted for. Each of the two tested technologies had a planned methodology to accomplish the objective. Both had major deviations from the plan.

LCS Methodology

A major goal of the test is to demonstrate that the product can operate in a functional field test setting. The product is undergoing concurrent energy impact evaluations in a laboratory setting, so one goal of this project was to prove that the product has reached a mature enough design phase to serve actual data centers in Minnesota. This portion of the project is essentially pass/fail based on the ability to meet the customer's needs and design a functional product.

The second part of the test plan is to measure energy impacts of the technology. At a very high level, energy savings can be shown if the servers are added to an operational data center and do not add to the air cooling load in the space. This high-level check is accomplished by measuring the air cooling energy consumption before and after installation of the new servers and ensuring the existing servers continue to operate in both cases.

To extract more precise energy impacts, the original test plan called for a relatively straightforward measurement of the identified variables and a simple comparison of total energy consumption using standard air cooling versus the new liquid cooling technology, collecting one week of data for each case. This was to be accomplished by installing monitoring equipment and measuring energy consumption of

the existing servers, new servers, existing CRAH cooling unit, and newly installed liquid Cooling Distribution Units (CDUs).

However, as the project proceeded, several constraints on data collection emerged and methods of addressing each were developed. The resulting final energy comparison test methodology is more complicated than the original plan. The methodology also forces results to be presented as a range of possible savings opportunity because the number of uncontrolled variables does not allow for high confidence in a single value for an overall savings claim.

The first constraint on the test design arose in conversations with the data center operator. Their operations require that the existing servers must continue to process their critical job load throughout the test period. They cannot be loaded with standard performance benchmarking tools, even for a short performance test. This prevented artificial loading conditions proposed in the preliminary test cases. Therefore, the test case designs were adjusted such that they monitor the existing servers using naturally occurring job loads, but do not alter their operation in any way. The solution limits the robustness of the results because it is not possible to test edge cases on the existing servers. That is, it is not possible to observe full capacity loading over long enough periods to change the heating load, which is unfortunate because those are the conditions predicted to produce the greatest savings. However, the test cases observed are expected to represent typical data center operations, so the most useful range of data is collected.

The second constraint discovered during test design also arose because the existing server operation could not be interrupted. This meant that the liquid-cooled servers could not operate independently; both the existing and the new servers had to operate simultaneously. The number of monitoring devices and CT clamps was limited, which meant that not all devices could be monitored in each planned test case. This was solved by requesting the data logged by the Eaton power distribution unit, which tracks total energy consumed by the data center. From the collected log, it is possible to calculate the power consumption of the servers that aren't directly monitored. This solves the problem, but does add some uncertainty to the results by adding an intermediate variable.

The last constraint on the test design is the most problematic. The hardware architecture of the existing servers is significantly dissimilar from the new LCS servers. Even processing the same job with the same cooling strategy, one would expect different performance outcomes from the different server processors, which complicates the task of isolating the effect of liquid cooling technology. The degree to which the servers are dissimilar was surprising even to the LCS team. The existing servers used in the data center are more sophisticated than typical units in similar-sized data centers and configured for a specialty medical deep learning application. This effectively means that the baseline conditions in this data center may not represent a good model for most other sites. There is not an ideal solution to correct for this issue.

One proposed solution was an attempt to normalize energy impact findings using a job log recorded by the local Grid Engine tool that manages the data center processing assignments. In principle, measuring the amount of data processed by each server could be used to calculate performance in a way that can be compared across technologies. For example, kWh per Gigaflop could be used as a metric, which

would normalize findings to correct for any efficiency advantages the existing hardware may have. In practice, the job log was far too low resolution to develop an actual savings metric and could only be used at a high level to compare processing loads (ex. high/medium/low processing load). Instead of creating a new metric, the observed findings are presented unadjusted. Processing load is included as a high-level variable to provide context for the results. Generally, this issue introduces an uncontrolled variable to the analysis and complicates interpretation of results and increases the range of possible savings opportunity predicted to exist in less-specialized data centers. Specifically, the final savings identified at this particular site are likely conservative compared to the opportunity at most similarly-sized sites.

Each of these constraints were identified at separate points during the testing process and each of them added an additional uncontrolled variable. As each was identified, the best identified solution was adopted to attempt to control the new variable as well as possible and the research progressed. However, if all constraints had been identified from the beginning, it is likely an effort would have been made to identify a site with fewer uncontrolled variables. A better apples-to-apples comparison would have allowed higher confidence in more precise savings claims. The actual observations do produce useful findings, but within careful context and the final results are presented as a range of outcomes.

Collected data points are an average of measured values over the runtime of each test case. Test cases were run for at least one week to ensure unusual operating conditions do not improperly weight collected average values. The standard deviation of each set of measured values was also checked to ensure no major perturbations occurred during the test run (there is no threshold for pass/fail, just a sanity check to screen wildly fluctuating data). Room temperature was also collected throughout all test cases to ensure there wasn't a major heat load unaccounted for in the measured data. Table 2 illustrates the variables measured and calculated in each test case.

Table 2: LCS Test Case Definitions

Test Case	1	2	3	4
Label	Baseline	Typical operation	Artificially loaded	High loading
Existing server loading	Typical job load	Typical job load	Typical job load	High load (naturally occurring)
LCS Server loading	Off	8 servers loaded to typical job capacity	4 servers loaded to full capacity. 4 servers off	8 servers loaded to high capacity
Existing RIL-GPU2 Server Power	Monitored	Unmonitored	Unmonitored	Unmonitored

Test Case	1	2	3	4
CRAH unit power	Monitored	Monitored	Monitored	Monitored
LCS Servers power	Off	Monitored	Monitored	Monitored
CDUs power	Off	Monitored	Monitored	Monitored
Eaton Log (full DC power draw)	Collected – used for calculations	Collected – used for calculations	Collected – used for calculations	Collected – used for calculations
All existing servers power	Calculated	Calculated	Calculated	Calculated
Grid Engine Job Log	Collected	Collected	Collected	Collected (to confirm high job load for existing)
Room Temperature	Monitored	Monitored	Monitored	Monitored
Date/Time	Recorded	Recorded	Recorded	Recorded

These four test cases are designed to observe a range of operating conditions for both the existing and the new servers. The range of scenarios is defined by the server loading (or utilization rates) of the technologies. The observed differences between technologies in terms of cooling energy required will illustrate the range of potential conservation opportunity at similar sites. Test case one is a baseline case showing operation of only the existing servers. Test case two compares typical operation of existing and new servers, which will likely reflect the most common situation in most data centers. Test case three is a worst-case scenario that should not correspond to any actual data center situation, but will demonstrate a theoretical floor for savings potential. Test case four is predicted to show the highest possible measured savings potential at this site because the utilization rate is the highest.

Evaluation Metrics

LCS Evaluation Metrics

Per the test methodology, there are three steps to evaluating the LCS test results. First, verify the technology operates successfully in a functional data center. Second, perform a high-level verification that the new technology does not add to the air cooling load of the existing CRAH unit. Finally, follow

the outlined test methodology to extract a detailed energy consumption comparison between the LCS technology and existing air cooling technology.

The first step is a pass/fail evaluation with a discussion of the customer's experiences with the technology. The second step requires monitoring the CRAH unit average power draw both before and after the installation of the new LCS equipment. Ensuing that there is no major increase in cooling energy required with the added compute load provides a high-level verification of energy savings potential.

Calculations can then be performed on the measured data to extrapolate more-precise energy savings potential of the liquid-cooling technology. Because the final test methodology includes workaround solutions to constraints identified during the test design process, it is not possible to perform a simple one-to-one comparison of energy consumed by the air-cooled servers to the liquid-cooled servers directly. Instead, the percentage of total energy used to cool can be compared.

The energy required for cooling compared to the total energy consumed by the servers and cooling equipment will be different for each technology and under each test case. The difference between the cooling energy percentages can be extrapolated over the data center load to determine the potential for efficiency improvement. This essentially creates a counterfactual scenario applying the liquid-cooling power consumption profile to the existing, air-cooled processing load to arrive at an estimated total energy conservation potential value.

Another possible evaluation metric that was considered was to use the processed data (GB processed or Gigaflops) as a normalizing factor. This would have resulted in a calculated kWh/GB processed metric that could be compared across technologies. Not enough precise data was available to construct that metric, but the Grid Engine job logger data can be used at a high level to compare loading conditions of the servers (high, typical, or low loading). This information can be used to evaluate how the utilization rate at a data center may impact the energy conservation effectiveness of liquid-cooling technology.

Table 3: Format of Summary Test Result Evaluation

Test Objective	Evaluation Metric	Result
Field demonstration	Does the installation successfully meet computational needs of the data center	Pass/Fail
Air cooling load	Ensure liquid-cooled servers do not add air-cooling load	Observe impact on CRAH power consumption
Energy Savings Potential	- Measure percentage change of cooling energy consumed (from air to liquid) - Calculate percentage change of total energy consumed (from air to liquid) - Calculate the counterfactual energy savings potential at this site if air-cooled servers were replaced by liquid-cooled equivalents. Extrapolate to annual savings	Comparison of air- and liquid-cooled technologies in terms of: - % change in energy used for cooling - % change in total data center energy (processing + cooling) consumption - annual energy savings potential at this site (kWh)

Results

Summary of Data

Monitoring equipment was installed and data was collected per the plan outlined in the test methodology section. Each designed test case was run for at least one week. Data was collected and processed using the HOBOware software tool. Table 4 summarizes the data collected over the course of testing. Each value in the following table is an average of the interval data collected over the course of the test case.

Table 4: Summary of Collected LCS Test Data

Test Case	1	2	3	4
Air-cooled server loading	Typical job load	Typical job load	Typical job load	High load (naturally occurring)
Liquid-cooled Server loading	Off	8 servers typical job load	4 servers loaded to full capacity. 4 Off	8 servers loaded to high capacity
Date/Time	3/7, 11:00AM to 3/24, 8:00AM	2/15, 12PM to 2/23, 7:15PM	2/24, 7:15PM to 3/7, 7:49AM	3/24, 9:30AM to 4/5, 12:00AM
CRAH power draw (kW)	0.864	0.796	0.787	0.951
LCS Server Power Draw (kW)	Off	1.250	0.860	2.718
CDU Power Draw (kW)	Off	0.110	0.111	0.107
Temp (F)	79.342	78.203	78.415	80.856
Ril-GPU2 power draw (kW)	0.410	Unmonitored	Unmonitored	Unmonitored - high volume per job log

Table 5 summarizes the values calculated over the course of testing.

Table 5: Summary of Calculated Values from LCS Test

Test Case	1	2	3	4
Air-cooled server loading	Typical job load	Typical job load	Typical job load	High load
Liquid-cooled Server loading	Off	Typical job load	4 servers loaded to full capacity. 4 Off	High load
Air-cooled total power draw (kW)	6.432	6.457	6.927	7.022
Air-cooled Server power draw (kW)	5.568	5.661	6.140	6.071
Air-cooled % cooling energy	15.5%	14.1%	12.8%	15.7%
Liquid-cooled % cooling energy	NA	8.8%	12.9%	3.9%
Potential reduction in cooling energy	NA	37.5%	-0.6%	74.8%
Extrapolated air-cooled annual energy (kWh)	56,380	56,603	60,718	61,558
Counter factual extrapolated annual liquid cooling annual energy (kWh)	NA	53,987	60,756	55,319
Possible annual energy savings (kWh)	NA	2,616	-38.5	6,238
Potential total energy savings	NA	4.6%	-0.1%	10.1%

Summary of Results

Table 6 summarizes the results of the test using the high-level metrics identified in the Methodology section. The first two metrics show a successfully installed product functioning as anticipated without adding any cooling load to the existing air cooling equipment. The final metric produces a wide range of test results that can only be interpreted generally with low confidence in the exact savings potential values.

Table 6: Summary of LCS Test Results

Test Objective	Evaluation Metric	Result
Field demonstration	Does the installation successfully meet computational needs of the data center	Pass
Air cooling load	Ensure liquid-cooled servers do not add air-cooling load	Observed no impact on CRAH power consumption
Energy Savings Potential	- Measure percentage change of cooling energy consumed (from air to liquid) - Calculate percentage change of total energy consumed (from air to liquid) - Calculate the counterfactual energy savings potential at this site if air-cooled servers were replaced by liquid-cooled equivalents. Extrapolate to annual savings	- Energy used for cooling decreases by 38% - 75% - Total energy consumed decreases by 4.6% - 10.1% - Potential for 2,616 - 6,238kWh annual savings at this site - Notes: savings potential heavily dependent on utilization rate. An artificial, unlikely worst-case scenario shows 1% cooling and 0.1% total energy *increase*

Discussion of Results

The tests performed on the LiquidCool Solutions product successfully demonstrate the viability of using full submersion liquid cooling technology in a medium-sized data center. As predicted, the liquid cooling technology was shown to use less overall energy than a typical air cooling strategy under typical operating conditions. However, test constraints prevented a measurement of exact energy savings that could be used to predict savings at future sites. Instead, the results are presented as a range of potential conservation estimates. Further, costs for the new technology are still unpredictable. The discussion of results includes an outline for evaluating possible future install sites, but such projects should be evaluated on a case-by-case basis due to the wide range of possible costs and savings.

Field Demonstration

The first major goal of the test program is to determine whether liquid-cooled servers can be safely deployed in Minnesota data centers. The risks of failure were low because the coolant itself has been tested around electronic equipment and the full submersion technology performs well in a laboratory test setting. However, until this test it had not been deployed in a functional data center to show real world applicability. The field demonstration of the technology was successful. The installed servers met their processing load with no issues and the cooling equipment reliably operated as designed.

The second goal of the test was to show that the liquid-cooled servers did not also add air cooling load to the data center. The concern is that the energy savings potential for the liquid cooling strategy assumes that liquid-cooled servers are not also leaking heat into the ambient environment that must be rejected by another means. Such leakage would reduce, or could even negate, the conservation potential for liquid cooling. The design specifications predict that no such leakage will occur, but until this test, that prediction had not been tested in a real data center. The test confirmed that no additional heat load is added to the data center. All heat generated by the servers is successfully captured and rejected by the liquid cooling technology.

High Level Energy Impact Patterns

The fact that the liquid- and air-cooling systems were shown to be separate in terms of energy and heat transfer should have made a comparison of the two relatively straight forward. The plan was to monitor the server and cooling energy consumption for both systems with each performing the same processing tasks and calculate a simple percentage of energy savings achieved by the new technology. In practice, there were several constraints on the test methodology that introduced uncontrollable variables to the test plan. The plan was adjusted and monitoring completed successfully. However, due to the additional variables, the results are presented in terms of ranges of possible outcomes rather than a single conservation potential metric as originally planned.

At a high level, the energy conservation potential of the liquid cooling technology was clearly observed. Under a range of operating conditions, the total energy consumed by the liquid-cooled system was less

than the air-cooled equivalent, with an artificial worst-case scenario of equivalent consumption. That is, even under the least optimal configuration of the liquid-cooled system, it will still perform as well as an air-cooled system in terms of energy consumption.

The test also confirmed a clear pattern that shows much greater energy efficiency opportunity as the server data processing load increases. This pattern confirms the observations of an NREL study conducted under laboratory conditions. The more the servers are used and the harder they are driven, the more beneficial liquid cooling becomes in terms of energy conservation potential. This suggests that data centers with a high utilization rate of their processing capacity will benefit more from liquid cooling than those with lower utilization rates.

Range of Energy Impacts

In the worst-case test scenario, it was observed that the liquid-cooling technology can be operated in a way that does not deliver energy savings. In fact, there was a negligible penalty of approximately 0.1% of total server-plus-cooling energy consumption associated with test case three. This scenario is unlikely to be observed in a real data center, but it illustrates the fact that energy savings are not automatically achieved. Failing to consider the server application and system design from an energy perspective could result in a liquid-cooled installation that does not deliver energy savings.

On the other end of the range of impacts, under test case four where servers are processing very high loads, observations show up to 75% of the required cooling energy (10% of total energy consumed) can be conserved using full submersion liquid cooling. The trend of higher savings at higher processing loads corroborates the laboratory-setting results. However, the laboratory tests were able to produce much higher savings (up to 98% of cooling energy and 26% of total energy conserved) that were not reproducible in the field.

There are two main likely reasons for the difference in maximum observed savings between the field test and the laboratory test. First, it was not possible to interrupt the operation of the existing servers to run benchmarking server test programs. Instead naturally occurring job loads were used to compare the systems. This produces useful results that demonstrate savings potential. However, less control of the inputs limits the range of operating conditions to perform tests on. The extreme loading conditions that would likely produce the highest conservation potential were not observed. This limitation is unfortunate from the standpoint of completely characterizing the operation of the technology and identifying a maximum potential savings opportunity under ideal operating conditions. However, typical servers likely would not be operated under ideal conditions for significant periods of time. The range of observed conditions reflect more-likely savings potential at typical data centers even if they don’t characterize a theoretical maximum savings.

The second reason laboratory tests showed different results from the field test is that the existing, air-cooled servers at the site have an advanced architecture and are used for a highly specialized deep learning medical application. The degree to which the existing server application differs from typical data centers was not well understood before the test. It was assumed that the specific operations

performed by the servers would not significantly impact the energy analysis, no matter how unique. In reality, the server application is unique enough that a comparison with new servers is not quite apples-to-apples. The effect of this obstacle is that the observed energy savings may undercount those that could be present in a less-specialized application. This is because the existing air-cooled servers already used advanced hardware and a more-optimized configuration for their application than those found in a typical data center.

In effect, the baseline condition at the test site may already be more efficient than would be expected in a typical data center, which reduces the opportunity to achieve improved efficiency. There is no way to directly correct for the inaccuracy in measured savings. At a *very* high level, a comparison can be made between the observed percentage of energy dedicated to cooling and an estimated value found in the literature. At this site, only 13-15% of the total energy consumed is used for cooling while a review of the literature suggests that a value as high as 40% can be expected in at least some similar sites. This suggests that adjusting the assumed baseline cooling consumption to match the literature *could* result in as much as approximately three times the observed energy savings potential. Adjusting the observed savings potential in this way comes close to matching savings achieved in a laboratory setting. This complication unfortunately significantly widens the range of possible savings potential and reduces confidence in the final findings. Further discussion of the results does not include this adjustment factor and focuses only on observed energy impacts.

Under the test scenario most similar to typical operational conditions (test case two), approximately 38% of cooling energy and 4.6% of total energy (server plus cooling) was conserved by the full submersion liquid cooling technology compared to the existing air cooling strategy. This represents a conservative estimate of the energy impacts that would be expected if a typical data center were to replace a typical CRAC unit with the full submersion liquid cooling system and continue to operate the data center per usual. If the new technology is accompanied by a plan to optimize the operation of the data center to capitalize on liquid-cooling performance, greater savings are likely achievable.

As an example of an optimization opportunity, the liquid cooling energy consumption did not change significantly across the test cases (0.11kW in all cases). Increased computational load and heat generated did not increase the Cooling Distribution Unit power draw. This suggests that the CDUs were operating at a baseload energy consumption and could have absorbed more waste heat in test cases 2 and 3 without drawing more power. The manufacturer confirmed that this was likely. This effect resulted in lower efficiency opportunity calculated for those test cases than could have been achieved if they were operated optimally.

Another opportunity for optimization is shown by the savings achieved in the high computational load test case. This suggests that data centers with high enough processing volume and the capability to optimize job loads to drive liquid-cooled servers at a high utilization rate will achieve greater savings than a site that doesn't do so. Achieving the greater savings from processing load allocation may be inherently easier at some sites, but all sites would likely benefit from considering cooling impacts of job allocation.

These two examples suggest a greater opportunity for savings if the data center is optimized for liquid cooling, but also illustrate the possibility of missing out on available savings if the technology is installed without an accompanying optimization plan. Such a plan should, at a minimum, include balancing the cooling distribution unit capacity with the projected server heat load and optimizing the job allocation software to consider cooling implications.

Likely Typical Energy Impacts

The savings observed under typical operational conditions (test case two, resulting in 38% of cooling energy and 4.6% of total energy conserved) represents a reasonable floor for the energy savings that should be expected from installing the full submersion liquid cooling technology. The savings observed under high processing load (test case four, resulting in 75% of cooling energy and 10% of total energy conserved) represents a reasonable ceiling for possible energy savings that could be achieved in a typical medium-sized data center. The worst-case scenario (test case three, resulting in a penalty of 1% of cooling energy and 0.1% total energy) is unlikely to be observed in real applications, but does illustrate the fact that a liquid cooling strategy requires additional planning beyond simply installing the technology in order to capture possible savings.

Due to changes in the test methodology in response to unforeseen constraints, the findings are presented as a range of possible savings. There is too much variability in the results to calculate a single, high-confidence value for the savings potential of the liquid cooling technology. The ultimate result of the energy monitoring test concludes that: full submersion liquid cooling of data processing servers in typical, medium-sized data centers can achieve between 38-75% reduction in cooling energy (4.6-10.1% of total energy) compared to typical air cooling technology. With appropriate vetting of sites and planning to optimize operations for liquid cooling, most sites should be capable of achieving energy conservation closer to the high end of observed savings during the test. In particular, the greater a data center's server processor utilization rate, the more it will benefit from liquid cooling.

Additional Benefits

Beyond direct conservation of cooling energy, full submersion liquid cooling has several possible additional benefits. Phase two of an NREL study is examining these benefits in more detail and a report should be available by the end of 2017. Evaluating these benefits was not part of the scope of this study, but they are worth considering conceptually. The additional benefits are claimed by the manufacturer. Observations over the test period cannot confirm the existence of the benefits, but they are reasonable claims.

Additional benefits may include:

- Secondary energy savings from a reduction in leakage current drawn by servers. This is enabled by operating the servers as lower temperatures.

- Greater performance by servers under high utilization rates. Server capacity is not limited by air cooling capability. Temperature responsiveness and faster heat rejection allow greater loads on servers for longer periods of time without any loss of performance.
- Reduction in maintenance costs. Servers are sealed in a submersion chassis, which does not allow particulate contamination of electronic components.
- Increased lifetime of server equipment. Reduced particulate contamination extends the useful life of servers.
- It is likely that additional hardware and software developments will increase the value of liquid cooling technology. For example, specialized software for virtualizing data center servers is likely to incorporate new criteria to take advantage of energy and job processing benefits enabled by liquid cooling.
- Due to the enabled higher utilization rates, it may be possible to design data centers with fewer servers to meet the same computational load, saving upfront costs.
- Heat rejected from the liquid cooling units can be captured by waste heat recovery systems relatively easily. Because the input temperature of the coolant does not need to be extremely cold (can be as high as 45C/113F), the waste heat can be used even for HVAC applications, which means it may be possible to capture significant additional energy benefits at many sites.

Cost/Benefit Analysis

The actual cost to install the liquid cooling equipment at the test site was significantly greater than expected costs for future installations. There are several reasons for this. First, the technology is just now emerging from early stages of development when design costs are expected to drop dramatically. Also, as volume increases, the technology will benefit from economies of scale. Second, the site itself was unique and required specialized design work that likely would not be necessary in most commercial applications. This was due to the physical configuration of the facility (requiring specialized plumbing/electrical plans), the choice to reject heat to CDUs rather than a cooling loop (to enable energy measurements), and the specialized data processing taking place at the site (requiring an unusual server configuration).

Despite the conservation potential, the observed energy impacts are unlikely to justify deployment of the technology to existing data centers with functioning air-cooling systems at current prices. New construction sites or data centers with failing air-cooling systems are more likely to be able to justify deploying liquid cooling because it eliminates the need to purchase a CRAH unit, which improves the comparative value of the investment. However, even at these sites, the full cost of installing a liquid cooling system should be carefully considered. This advantage of liquid cooling was not possible to test at Hardwick. In fact, a direct cost/benefit analysis of the test installation is not possible.

The manufacturer makes claims comparing liquid-cooling technology to air-cooling at a new construction site that achieves an ROI of approximately 1.5 years. This claim includes estimated upfront system costs and expected operational savings including reduced energy costs. The claims were not verifiable at this test site and the costs quoted are projections rather than actual invoiced costs to actual customers. The estimates appear reasonable, but cannot be confirmed at this time.

Instead of a strong claim about the cost-effectiveness of liquid cooling, a general rule of thumb was developed to summarize the cost-effectiveness potential of the technology. At the Hardwick site, the range of observed energy savings would justify a liquid-cooling system that cost $600-$1,200 more than an air-cooled system and still deliver an ROI under two years. Per manufacturer claims, if the Hardwick site were considering a new system, the differential cost between the two technologies would fall in this range. To generalize the findings to apply to other sites, the threshold for cost-effectiveness of liquid cooling compared to air cooling is approximately an incremental cost of $100 per kW of sever max power draw.

This rule of thumb represents a high-level test for data center operators to apply to a possible liquid cooling system. If they obtain quotes for a liquid-cooled system with upfront costs less than $100 per kW of server power draw compared to an air-cooled system quote, the site is likely to deliver a return on the investment in under two years. Due to the wide range of energy savings results, the rule of thumb presented here assumes the most conservative energy savings impacts. Higher incremental costs could be justified at individual sites where greater savings are expected.

Conclusions and Recommendations

A full submersion liquid cooling system was successfully demonstrated as a viable server cooling strategy in a medium-sized data center in Minnesota. The system was observed to achieve between 38%-75% reduction in cooling energy, or approximately 4.6%-10.1% of total energy consumed (servers plus cooling) compared to typical air cooling technology. Findings are presented as a range due to uncontrollable variables in the test methodology. With appropriate vetting of sites and planning to optimize data center operations to take advantage of liquid cooling, most similar data center sites should be capable of achieving energy conservation closer to the high end of the observed savings range. Of particular importance, the greater a data center's server processor utilization rate, the more it will benefit from liquid cooling. Specific sites may be able to improve the value of liquid cooling technology by capturing additional benefits such as waste heat recovery and reduced maintenance costs.

The results of the study suggest some recommendations for data center operators and utility program implementers to ensure the benefits of the technology can be captured by Minnesota facilities.

Data Center Operator Recommendations

First, all data center operators should characterize energy consumption at their sites and track cooling energy consumed over time. Many sites already do this or have the capability, but the information is not always used to consider energy conservation opportunities. With innovative new technologies, including liquid cooling, it will be beneficial for data center operators to be aware of their energy consumption profile and ready to evaluate opportunities as they arise.

The following checklist summarizes criteria for likely candidate sites to benefit from liquid cooling technologies. If an operator's site meets these criteria, they should consider whether installing liquid cooling technology at their site may be beneficial. If the criteria are met, it is recommended to obtain a cost quote for installing both an air- and liquid-cooled strategy at the site to compare.

- ✓ New construction or major renovation where choosing liquid cooling avoids the purchase of a new CRAC unit (virtually a necessary condition at this point in the technology's maturation)
- ✓ Servers at the site are expected to run at a high utilization rate
- ✓ The site has a nearby HVAC or process heat load enabling the capture of waste heat
- ✓ The site will benefit from non-energy advantages of liquid cooling such as reduced maintenance costs, smaller footprint, quite operation, or higher performing servers under extreme loads

In the near-term, sites that do not meet these criteria may not yet see enough benefit from liquid cooling to justify installation of the technology. However, prices are likely to fall and benefits may increase, so it will be good to reevaluate the viability of liquid cooling periodically going forward.

Utility Program Implementer Recommendations

Utility conservation program implementers who work with data center customers should be aware of the potential benefits of liquid cooling. Not all customers will be ideal candidates for the technology in the near term, but being able to discuss the technology will help to raise conversations about conservation opportunities in general. Implementers should also be on the lookout for ideal candidate sites for adoption of liquid cooling technology. In particular, ideal early adoption candidate sites will likely plan to take advantage of non-energy benefits of liquid cooling. In fact, energy conservation may be a secondary consideration for the operator, so implementers should be aware of possible opportunities to help improve the benefits of a system that may already be under consideration separately from energy-related drivers. If a site meets the criteria outlined above, implementers may suggest that a data center operator consider installing liquid cooling.

Due to the wide range of observed energy impacts of liquid cooling and only one field test site, development of a prescriptive method of calculating energy conservation achieved by liquid cooling is not yet possible. Measurement of cooling energy before and after the installation is the only way to claim the full energy benefits of liquid cooling technology. However, it would be appropriate to assume a deemed savings of 4.6% of overall energy consumption at the site compared to air-cooling technology. This is a conservative estimate of the expected minimum savings that can be achieved. It can be applied to sites where energy measurement is not possible or where there is no actual existing air-cooled technology to measure (new construction site). Measuring impacts, where possible, would very likely result in greater claimable savings.

This study demonstrates the viability of liquid cooling technology in small- to medium-sized data centers, but further research is required to develop a prescriptive savings calculation methodology that doesn't require measurement to verify.

References

Results for Liquid Submerged Server Experiments at NREL's Advanced HVAC Systems Laboratory. National Renewable Energy Laboratory. Eric Kozubal. July 18, 2016.

Appendix A: Liquid Cooling Product Specifications

LCS Product Specifications

LSS200 Server (12)

GPU: NVidia Tesla M2090
Processors: Dual Intel Xeon 5670, 2.93GHz CPUs
Memory: 48GB DDR3 1333 MT/s Registered ECC DIMMs
Storage: 2X 250GB SATA III 6Gb/s SSDs (RAID-1)
I/O Connections:
- 2X 1Gb Ethernet ports
- 2X USB 2.0 Ports
- 1X Remote Management Port (KVM over LAN)

As configured, each server will draw about 300W of power at full load resulting in a total IT computing load of approximately 3.6kW. The servers will be installed in one of your existing data center racks, consuming 10U of rack space. An additional 1U to 2U of rack space may be needed for a PDU if a sufficient number of outlets are not available in the existing rack.

Koolance ERM-3K3U (2)

- The heat exchange system will be made up of two (2) Koolance ERM-3K3U liquid-to-air cooling distribution units (CDUs), which will be modified for dielectric liquid cooling, and a system of manifolds and hoses to connect the CDUs to the servers.
- In regard to the location of the heat exchange system during the CARD testing, our proposal is to use the custodial room immediately adjacent to the data center room. We would mount the two CDUs in that room and connect the CDUs to the servers using two hoses routed above the ceiling tiles between the two rooms.

Ebullient Product Specifications

Fluid Distribution Unit
ES120

Introduction

Ebullient's ES120 Fluid Distribution Unit is designed to cool up to 120 servers.

Overview

The ES120 circulates an equipment-safe fluid at low pressure to rack-level manifolds that supply module loops installed in individual servers. Server heat is absorbed into the fluid through vaporization and returns to the ES120. An integral heat exchanger rejects heat to a facility cooling circuit. The ES120 can be located adjacent to server racks or placed remotely up to 30 meters away and requires only facility water and power connections.

Performance

Heat Removal Capacity[1]	60 kW @ 40° C 17 Tons
Max Server Capacity[1]	120
Max Manifold Capacity[1]	4

Physical Attributes

Enclosure	Rolling 42U Electronics Rack
Dimensions (h x w x d)	1994 x 600 x 1055 mm 78.5 x 23.6 x 41.5 in
Weight	330 kg / 720 lbs

Fluid

Type	Novec 7000
Conductivity	Non-conductive dielectric fluid
Safety	Non-Toxic, Non-Flammable, Low GWP

[1] *Maximum performance specifications valid only at specific operating conditions. Performance will vary with application.*

100 S. Baldwin St., Suite 200 | Madison, WI 53703 608.957.7751 | www.ebullientcooling.com

Power Requirement

Power Type[2]	3φ 60 Hz
Max Power Consumption	2.7 kW
Full Load Amperage	9 A @ 208 V 8 A @ 240 V 4 A @ 480 V

Pumps

Type	End Suction Centrifugal SS
Size	2.7 kW / 3 HP
Certification	UL

Fluid Connections

W.E.G. / Facility Coolant	1" Class 150 Flanges
W.E.G. Flow Rate	148 lpm / 39 gpm
W.E.G. Liquid Cooling Class	W1 - W3+ (< 40° C)
Primary Cooling Circuit	1" Class 150 Flanges
Distribution Piping	1" Nominal
Dewpoint Control	Optional (external)

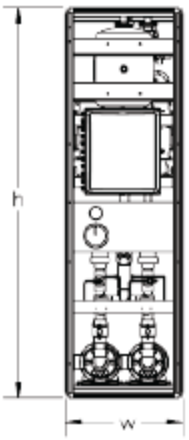
Controls

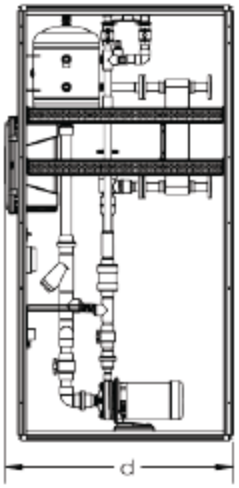
Type	Integrated Control Panel
Interface	Touchscreen Display
System Monitoring	Temperatures and Pressures, Power Consumption
Pump Monitoring	Automatic Failure Detection and Switchover
Communication Link	Ethernet
Communication Protocol	TCP/IP, IPX, Modbus TCP/IP, DHCP, HTML Configuration
Panel Certification	UL 508

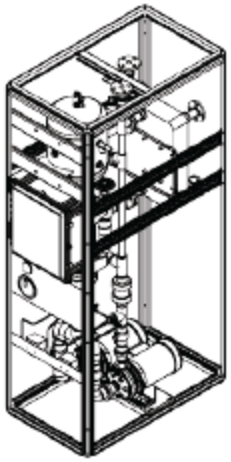
[2] *Other options available on request.*

These specifications are not intended as a warranty. In the interest of product improvement, these specifications may be changed from time to time without notice. Please consult your sales representative for details.

Drawings







Appendix B: Project Marketing Materials

GDS Associates, Inc.
Engineers and Consultants

Direct-to-Chip Liquid Cooling Technology
Seeking Medium Sized Data Centers to Participate in Energy Efficiency Study

GDS Associates is looking for data center sites to participate in a study to determine energy efficiency savings from installing state-of-the-art cooling technologies.

Overall: The goal of the project is to compare the energy consumption required to cool servers with an air-cooled strategy vs. the new direct-to-chip liquid cooling technology. To accomplish this goal, we need find two sites where we can meter HVAC cooling energy consumption and server energy consumption for approximately one week both before and after install.

What is direct-to-chip liquid cooling?

Direct-to-chip liquid cooling technology removes excess heat from Data Center servers by circulating a dielectric coolant (completely safe for electronic equipment) directly to each server. This strategy removes the need for dedicated CRAC units and can keep servers at optimal operating temperatures using far less energy than air-cooled solutions.

Why consider direct-to-chip cooling?

The main benefits of liquid-cooled servers are the significantly reduced cost of energy required to cool a Data Center and the higher reliability of precise temperature controls.

Are my servers at risk when a liquid is introduced near electronics?

No. The direct-to-chip cooling system uses a dielectric fluid as coolant. This ensures that there is no risk to server equipment. Even in a “catastrophic” failure event where coolant leaks onto the server, no electronics are damaged. In fact, electronic equipment can continue operating with no issues completely submerged in the fluid.

Are the energy savings worth the upfront costs?

That’s one of the questions this study is designed to answer. For study participants, the majority of the installation costs will be covered by the product manufacturer and Department of Commerce grant, so the minimal upfront cost to participants is projected to result in a nearly immediate return on investment.

Does liquid cooling impact server operation?

No. The operation of the servers themselves is unaffected by the cooling strategy.

GDS Associates, Inc.
Engineers and Consultants

What types of sites are we looking for?

A potential participant site must:

- Have a data center with 15-120 existing, functioning air-cooled data processing servers (not storage servers).
- Be willing to install Ebullient DirectJet liquid cooling technology on servers
- Allow energy monitoring equipment for at least one week prior to and one week after the installation of direct-to-chip liquid cooling technology.
- Process a predictable data load or have an easily measurable utilization rate (the goal is to ensure that any measured energy savings are attributable to the cooling strategy, not a lower data load).
- An ideal site will also have an existing heat rejection loop near the data center, but this condition is not absolutely necessary.

Planned cost structure:

Installed system	CARD	Manufacturer	Utility Incentive*	Customer Payment
Ebullient ER-25 (4-10 racks)	$15,200	$24,000	$2,800	$3,000
Ebullient ES-60 (10-20 racks)	$26,400	$48,000	$4,600	$6,000

*Estimated utility rebate once we can demonstrate energy efficiency results.

The proposed study will receive funding from the Minnesota Department of Commerce – Conservation Applied Research and Development program.

Interested parties should contact-
Travis Hinck, CEM
Project Manager
(612)916-3052
Travis.Hinck@GDSAssociates.com

Appendix C: Statement From Ebullient Cooling

Ebullient, Inc. ("Ebullient") applied for a grant through the Minnesota Department of Commerce, Division of Energy Resources in 2015. The purpose of the grant was to support an energy efficiency study to measure energy savings in data centers resulting from installing liquid cooling systems on servers.

As participants in the study, Ebullient sought Minnesota companies with small or medium sized data centers. To ensure adequate performance of its liquid cooling system and quality data collection, Ebullient required that eligible participants 1) have 15-120 functioning, air-cooled data processing servers, 2) allow liquid cooling system to be installed in their data center and connected to their servers, 3) allow energy monitoring equipment to be present in their data center for at least one week prior to and one week after installation of the liquid cooling system, 4) have servers that process a predictable data load or have an easily measurable utilization rate, and 5) have an existing heat rejection loop near the data center.

In short, the plan was to measure energy consumption of HVAC systems and servers in the data center prior to and after the installation of the liquid cooling system, and by comparing those measurements, determine the energy savings provided by the liquid cooling system.

Despite efforts to locate participants for the study, including direct outreach to Minnesota companies and advertising through various social media platforms, Ebullient struggled to find willing participants. In late 2016, GDS Associates, Inc. ("GDS") identified the University of Northwestern - St. Paul ("the University") as a potential participant. At that time, the term of the grant was near expiration, so GDS arranged with the state of Minnesota to extend the term to provide Ebullient time to pursue the opportunity.

In early 2017, Ebullient visited the University's data center and met with facility personnel. During the visit, Ebullient staff discovered that the University's data center lacked a heat rejection loop. (Note: The purpose of the heat rejection loop is to receive server heat that is collected by the liquid cooling system and to reject that heat to ambient air outside the data center. Therefore, a heat rejection loop is essential to the performance of the liquid cooling system.)

To make the University's data center suitable for the study, a heat rejection loop would need to be installed. Ebullient engaged a local mechanical contractor to provide a quote for installing a heat rejection loop, which required a drycooler, pump package, plumbing, and control system. With labor, the base quote was $43,709. Unfortunately, neither Ebullient or the University was able to absorb that cost, so the installation was canceled.

Due to a lack of willing participants, and a lack of a local sales personnel to identify willing participants, Ebullient decided to let the grant expire.