DELL'S LIQUID COOLING INNOVATION FOR SCALE-OUT DATACENTER ENVIRONMENTS

DELL'S 'TRITON' LIQUID COOLING OFFERS LEADING-EDGE PERFORMANCE, COOLING CAPACITY, & DESIGN SIMPLICITY THAT COULD TRANSLATE TO SIGNIFICANT TCO SAVINGS

EXECUTIVE SUMMARY

Achieving the best total cost of ownership (TCO) requires delivery of the highest possible performance at the lowest possible cost. For CPU-intensive workloads, this generally means operating the CPU at the highest achievable frequency. Heat generated by the CPU is usually the key limiter. Dell's Extreme Scale Infrastructure (ESI) group has been working with hyperscale customers on customized liquid cooling for more than 6 years via their Data Center Solutions (DCS) team. Under the codename of 'Triton', Dell's third generation liquid cooling innovation has the ability to cool CPUs with performance and power that has been almost impossible to achieve before.

'Triton's' unique cooling capabilities provided inspiration for the development of a custom Intel CPU, which has the potential to offer a large TCO benefit for customers who have workloads that scale well with CPU frequency and core count. While results may vary, the combination of 'Triton' and this special CPU (3.3GHz Turbo, 20C, 200W TDP) can provide a double-digit performance increase over the highest performing Xeon in the market today (E5-2699 v4, 3.6GHz, 22C, 145W TDP) and up to 59% more performance than the popular E5-2680 v4 (3.3GHz, 14C, 120W TDP) on CPU-intensive workloads.[1] This is achieved at a cost that is only slightly more than a conventional (non-water-cooled, standard CPU-based) solution. In addition, 'Triton' actually overcools the CPU to well below the required temperature, virtually eliminating the need for it to exit the Turbo Boost mode, thereby delivering optimized performance.

This paper explores various cooling system approaches used by IT equipment vendors and examines the unique innovations the Dell 'Triton' approach provides.

WHAT IS OLD IS NEW AGAIN: THE RETURN OF LIQUID COOLING

Hyperscale web giants and many high performance computing (HPC) organizations purchase datacenter infrastructure that is specifically optimized for each of their workloads. For large-scale datacenters where IT is a central element of their business, every amount of performance gained, dollar saved, and efficiency realized in their IT infrastructure can have a significant impact on the bottom line. To serve their diverse needs, industry leaders and specialty vendors have developed custom and semi-custom businesses to create optimized solutions for this group of customers.

One area of growing interest over the past decade has been to improve the energy efficiency of datacenters. IBM first introduced water-cooled mainframes in 1964, but the transition to lower-power CMOS technology (that could be cooled with less expensive ambient air) and the availability of lower cost air conditioning drove datacenters to adopt air cooling as their primary cooling system model. However, vendors are again looking at liquid cooling as an approach that may be a good fit for very specific environments.

Non-conventional cooling approaches like liquid cooling can be especially helpful for customers or regions with high energy costs (like those based in Europe) or for customers who have specific green initiatives as a part of their business. In addition, customers who are pushing the performance envelope may benefit from using liquid cooling for their high power, high performance IT equipment that is almost impossible to cool using traditional methods. Workloads today that can benefit from liquid cooling include HPC verticals such as oil and gas, research labs, gaming, and financial services organizations doing high frequency trading. Other web-scale workloads that can experience potential benefit include certain web-scale query / search centric applications and some machine learning applications that use neural networks. However, liquid cooling does not make sense for all workloads, and it is important to ensure the potential gains can be realized in the target environment.

Dell's DCS group has been a pioneer in building datacenter infrastructure for some of the largest datacenters in the world with a primary focus on large web-scale. In 2010, Dell DCS started looking at liquid cooling primarily for energy efficiency benefits and referred to their liquid cooling efforts with the codename 'Triton'. The potential performance gains and cost effective ability to cool caused Dell DCS to see the opportunity for millions of dollars of savings per year for its customers. In addition, Dell DCS found an opportunity to innovate using 'Triton' technology to offer unmatched performance for workloads that scale well with cores and frequency.

Cooling System Evolution on the Path to Efficiency

It is useful to compare various datacenter cooling system approaches to understand how liquid cooling can be beneficial for some environments. A datacenter cooling system removes the heat generated by datacenter equipment to maintain the proper operating conditions. This complex system serves as a collection of heat exchanges (places where heat is transferred from one medium to another) with the cascading efficiency of each determining the overall efficiency of the system. Efficiencies lost at each exchange point have a multiplicative impact on operating costs.

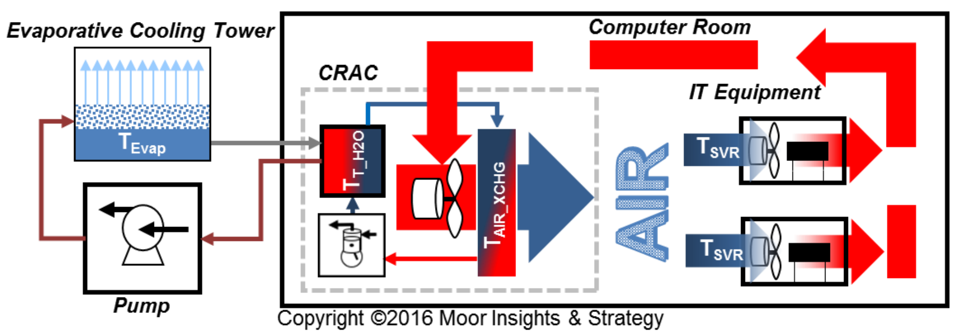
For example, if you have 1MW of IT load and your cooling system is 95% efficient, then you must remove not only the 1MW of energy (heat) but also overcool by 1MW / 95% to make up for the inefficiencies. If you have four serial heat exchanges that are each 95% efficient, then their inefficiencies quickly accumulate to an overall of 95% x 95% x 95% x 95%, for a total efficiency of ~81%. For the same 1MW of IT equipment load, the actual cooling requirement is 1MW / 81% or 1.235MW, requiring about 235KW of extra cooling to make up for the inefficiencies. This example illustrates the golden rule of thermal design which says, "**Make each interface as effective as possible, and keep their total number as small as possible**".

STANDARD COOLING SYSTEMS TODAY

The most common cooling systems use a combination of water, refrigeration, and air. The system shown in Figure 1A will serve as the basis for this discussion with four major areas of interest.

1. Computer room IT equipment which typically uses air to remove heat (Tsvr)
2. Computer room air conditioners (CRAC) using cool air
3. Conventional air conditioning (Tair_xchg) which transfers the heat from the refrigerant to water (Tt_H_2O) and moves the air
4. A cooling tower which transfers heat from water to air using evaporation (Tevap)

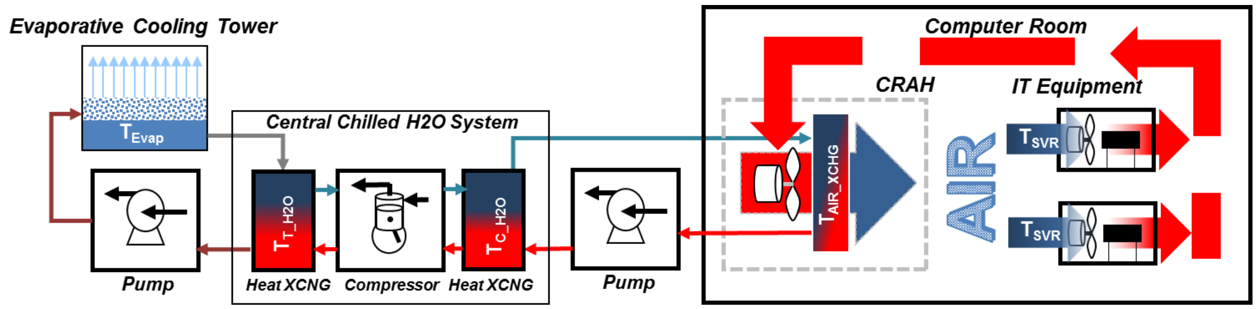
FIGURE 1A: TYPICAL IT COOLING USING A COMPUTER ROOM AIR CONDITIONER (CRAC)



Source: Moor Insights & Strategy

A variation of this system (shown in Figure 1BThe systems illustrated in Figures 1A and 1B have been and continue to be the industry workhorses.) uses a centralized refrigeration system, distributes chilled water, and replaces the in-room air conditioner with a simple cooling coil-based air handler.

FIGURE 1B: TYPICAL IT COOLING WITH CHILLED WATER DISTRIBUTION USING A COMPUTER ROOM AIR HANDLER (CRAH)



Copyright ©2016 Moor Insights & Strategy

Source: Moor Insights & Strategy

The systems illustrated in Figures 1A and 1B have been and continue to be the industry workhorses.

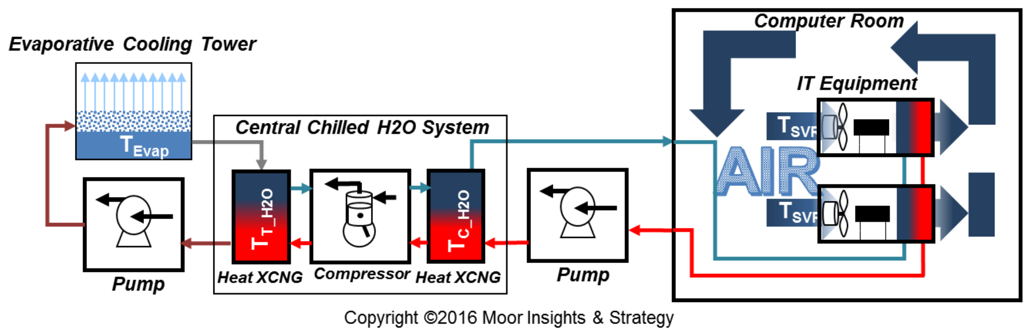
A well-published metric describing the effective use of power as a part of the operational cost of a datacenter is **power usage effectiveness (PUE)**. PUE is the ratio of the total power consumed by a datacenter to the power used by IT equipment. Even today when using the methods in Figures 1A and 1B, it is not uncommon to find an older datacenter with a PUE of about 2. This means for every watt of power the IT equipment consumes, an additional watt of power is needed to distribute the power itself, operate support equipment, and cool the datacenter. With this type of system, the power required to operate the example 1MW datacenter is actually 2MW, resulting in significant ongoing operating expense.

Large hyperscale installations have achieved major PUE improvements in recent years by the use of outside air as “free cooling”. Some have been able to reach a PUE in the 1.035 range very cost effectively. However, there are practical limits to the amount of heat that can be removed by simply using air, as the heat carrying capacity of water is significantly higher than air. In addition, using only outside air is not useful in some environments due to air quality, and it is difficult to add outside air cooling to an existing datacenter. Using best-of-breed conventional techniques with air cooling can generate a PUE of 1.3 to 1.4, which is a significant improvement over historical approaches.

Liquid Cooling Provides Advantages for Some Environments

Recent solutions have returned to using water distributed directly to the rack as an alternative means of cooling. Figure 1C shows a solution where chilled water is circulated through the rear door of the rack and the hot air exiting the IT equipment is cooled before it re-enters the computer room. This approach decentralizes the CRAH (shown in Figure 1B, above) and can improve PUE due to the efficiency of rear door cooling. It is also a useful solution when expanded capacity is needed.

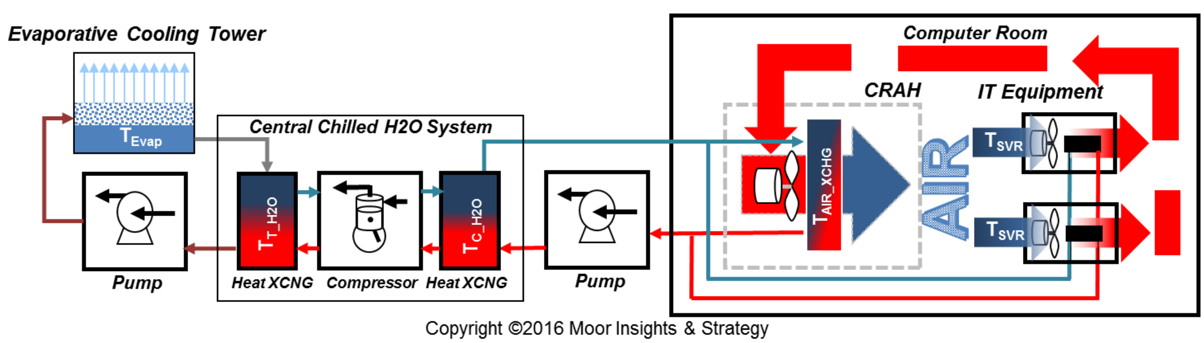
Figure 1C: Centralized Chilled Water System with Rear Cooling Doors



Source: Moor Insights & Strategy

Hybrid systems have emerged as an alternate solution to improve cooling. Figure 1D shows a hybrid system using water from a distributed chilled water supply as the primary means of cooling the server.

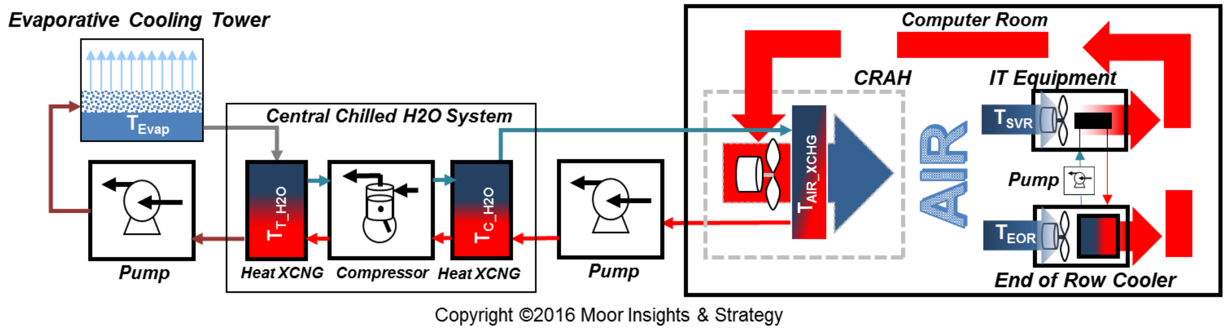
Figure 1D: Centralized Chilled Water Hybrid System Sharing Chilled Water Supply



Source: Moor Insights & Strategy

Figure 1E shows a hybrid system using water from a "private" loop maintained within the computer room as the primary means of cooling the server.

FIGURE 1E: CENTRALIZED CHILLED WATER HYBRID SYSTEM USING A LOCALIZED COOLING LOOP



Source: Moor Insights & Strategy

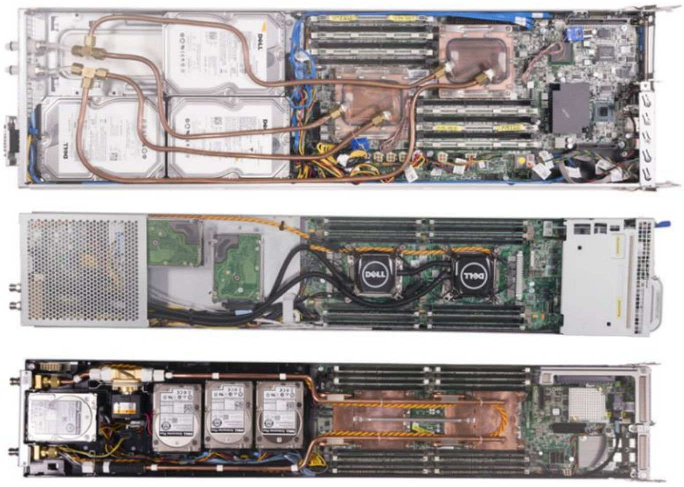
These hybrid systems can achieve a PUE of 1.036 to 1.056, but they still require cooled air to maintain the temperature for other parts of the system or other equipment in the computer room.

While PUE remains an important metric, there are other aspects to consider when evaluating cooling system solutions. The operational temperature, especially of the CPU (the package temperature or Tcase) is one of the most significant limiters of performance and directly impacts the CPU operating frequency. Unfortunately, there is a practical limit to the amount of heat that can be managed by moving air, which limits the power budgets of the components in the system. In fact, 135W CPU is generally considered the highest power that can be cooled in a density-optimized (shared infrastructure) server using air. Water, which is capable of absorbing about 4,000 times as much heat as air in a given volume, is a very effective alternative to air for cooling systems with high frequency, high power CPUs.

DELL 'TRITON': MULTIPLE GENERATIONS OF LIQUID COOLING INNOVATION

Dell has been looking at liquid cooling for a number of years and is currently on its third generation of 'Triton' technology. Figure 2 illustrates an engineering prototype of each of the three 'Triton' sleds showcasing the evolution of the three generations.

FIGURE 2: THREE GENERATIONS OF 'TRITON' LIQUID COOLING INNOVATION FROM DELL



1st Generation 'Triton'

2nd Generation 'Triton'

3rd Generation 'Triton'

Source: Dell

FIRST GENERATION OF 'TRITON'

Dell's first 'Triton' project was for a global-scale search engine provider. Dell leveraged an existing air-cooled server purchased by this customer and installed the 'Triton' liquid cooling capability on top of the infrastructure as a proof-of-concept. Dell embedded a liquid-to-liquid heat exchanger and a hot-pluggable, redundant pump system to distribute water to each of the compute blades within the chassis, thus creating a blade-mount heat exchanger.

SECOND GENERATION OF 'TRITON'

Dell's second 'Triton' project was developed as a proof-of-concept for a global web tech company. Dell, Intel, and the end customer worked together on various optimizations including CPU frequency, core count, reliability, and power to help optimize performance for the customer's specific workload and environment.

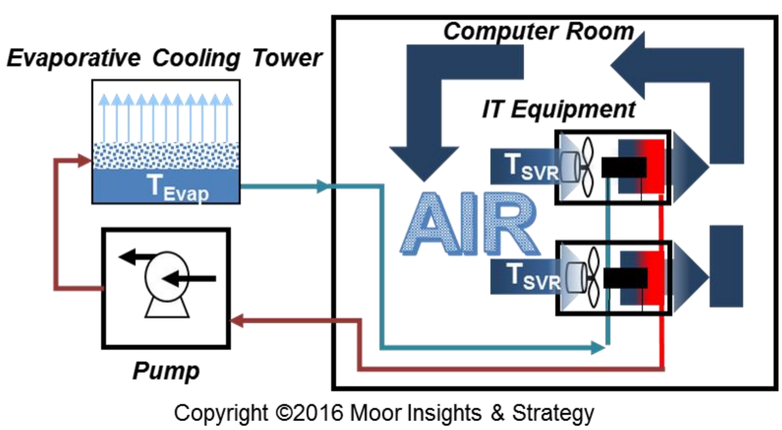
Intel has an extensive process for delivering CPUs tailored to a customer workload. This includes software optimization, SKUs with varying core counts, and different frequency and power levels. In this case, Intel worked with Dell and the customer to develop a Xeon processor optimized for their queries per second (QPS) metric. It delivered more than a 30% increase in performance over a standard Xeon processor with more than a 70% increase in thermal design power (TDP) headroom. To cool the high power processor optimally, Dell removed the centralized pump and heat exchanger unit from the previous generation and installed pumps on the cold plate modules. A chassis-mount heat exchanger system transferred CPU heat to facility water via a liquid-to-liquid heat exchanger. In 2014, the customer deployed its first proof-of-concept liquid-cooled rack based on 'Triton' technology, leading to the development of the third generation.

THIRD GENERATION OF 'TRITON'

Dell's third and latest generation liquid cooling project represents a breakthrough in innovative, "complete picture" system design and operation. Built on Dell's highly flexible rack-scale infrastructure, 'Triton' now uses facility water to cool the high power CPU directly, but more importantly, it truly integrates into overall datacenter operation.

The latest version of 'Triton' keenly follows the golden rule of thermal design by having only two real heat exchanges: the heat sink at the CPU and the evaporative cooling tower. This is about half the number used in a cooling system with one of the above configurations. 'Triton' achieves a PUE of 1.026 to 1.029 and provides the maximum exchange of heat between the source and its dissipation into water vapor. Further, it includes a novel heat exchange system driven completely by a set of integrated fans. The cooling system not only provides enough cooling to handle the parts of the system that require air cooling, but it is capable of generating enough air to cool other equipment in the computer room. 'Triton' is now able to cool the CPU at 0.10°C per watt, which is almost 2x greater than solutions that use the approaches shown in Figures 1C, 1D, and 1E. Each mandates the need for chilled water or a local, in-room loop.

FIGURE 3: DELL 'TRITON' IN THE OVERALL DATACENTER



Source: Moor Insights & Strategy

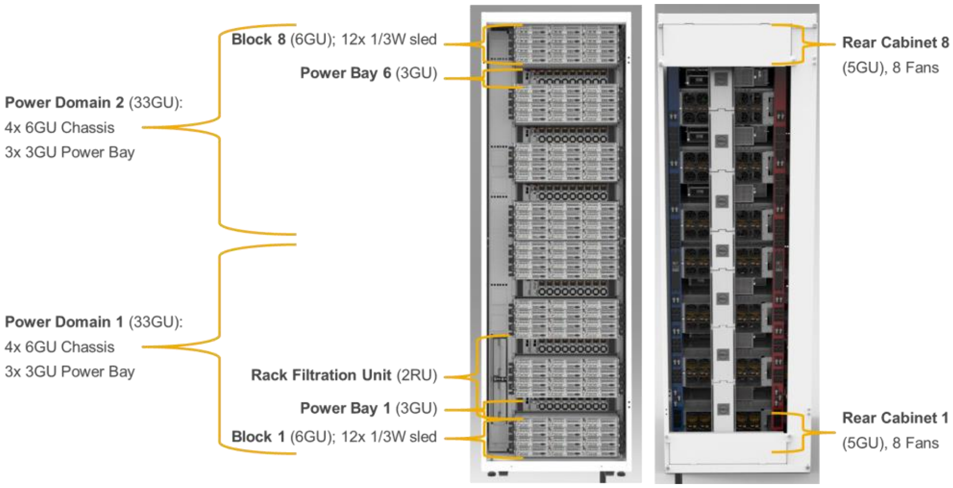
The third generation of 'Triton' can host up to 96 Dual-Socket 200W Intel Xeon Processor E5 v4 (3.3GHz, 20C, 200W TDP) CPU-based servers at slightly more than the cost of a conventional solution (non-water-cooled, standard CPU-based) with much higher performance for CPU-intensive workloads. This special processor creates a large TCO benefit for customers who have workloads that scale well with CPU frequency and core count and can adopt liquid cooling.

While results will vary for each customer's workload, the combination of 'Triton' and this special CPU (3.3GHz Turbo, 20C, 200W TDP) provides a double-digit performance increase over the highest performing Xeon in the market today (E5-2699 v4, 3.6GHz 22C, 145W TDP) and 59% more performance than the popular E5-2680 v4 (3.3GHz, 14C, 120W TDP) on the LINPACK benchmark.[2] This type of performance improvement has the potential to translate into millions of dollars in value. These customers can better respond to their business needs and reduce the number of legacy servers required to handle peak loads. In fact, Dell's global web tech customer was able to achieve an increase of 70% in search queries per second while reducing their total cost of ownership.

Dell has taken the necessary precautions (filter, filtration problem detection, redundant no-leak quick disconnection, leak detection and remediation, *etc*.) to safeguard customers from water leaks and eliminate added risk during operation. Dell has also ensured 'Triton' can operate using normal ASHRAE facility water directly from the cooling tower without special water conditioning. However, direct connection to tower water is not a requirement, as 'Triton' can operate with any traditional water system. Further, integrated redundant hot-replaceable filtration and control mechanisms not only ensure operation but also can serve as an early indicator of other problems in the facility cooling system. Dell's unique approach with 'Triton' is an achievement in how to cost effectively unlock processing power that before was not achievable, while improving overall facilities operation and integration—all without adding risk to the datacenter.

Dell's flexible rack-scale infrastructure supports a heterogeneous mix of compute / storage with both air-cooled and liquid-cooled nodes for customers who have different needs across their workloads. Figure 4 illustrates the key specifications of 'Triton' in the rack-scale solution.

FIGURE 4: DELL 'TRITON' RACK VIEW



Source: Dell

'TRITON'S' DESIGN SIMPLICITY HELPS DRIVE EFFICIENCY

While datacenters have chosen to use the metric of PUE to quantify efficiency, industrial cooling equipment manufacturers typically use the Coefficient of Performance (CoP) to describe the cooling efficiency of their products. The CoP (sometimes referred to as CP) is the ratio of the cooling provided to the work required to get that cooling. In other words, it describes how much heat (or thermal energy) can be removed per unit of electrical power input and is a common metric.

$$\text{CoP}=\frac{\text{Heat Removed}}{\text{Power Input}}$$

To bridge this gap and make technology comparison less ambiguous, the PUE impact of a cooling device can be calculated from the CoP by using the formula below, where CoP_n is representative of every subsystem that participates in the cooling system.[3]

$$\text{PUE}=1+\left[\left(\frac{1}{\text{CoP}_1}\right)+\left(\frac{1}{\text{CoP}_2}\right)+\ldots\left(\frac{1}{\text{CoP}_n}\right)\right]$$

As shown in Table 1, ''Triton' not only has excellent cooling capacity, but its simplicity eliminates many elements and their costly operation.

TABLE 1: PUE DRIVERS FOR DIFFERENT COOLING SYSTEM METHODS

Method	Figure	PUE Range	Coefficient of Performance (CoP)						
			Cooling Tower	Tower H_2O Pump	Refrigeration	Chilled H_2O Pump	Air Mover	Additional Component	Server Fans
Cooling Using CRAC	1A	>1.30	45	60-150	6	60-150	10		20
Cooling Using CRAH	1B	>1.30	45	60-150	6	60-150	10		20
Rear Cooling Doors	1C	>1.30	45	60-150	6	60-150	10	10	20
Hybrid Cooling Air & H_2O	1D	1.036 - 1.056	45	60-150				60-150	(20)*
Local Cooling Loop	1E	>1.30	45	60-150	6	60-150	10	60-150	(20)*
Dell 'Triton'	3	1.026 - 1.029	45	60-150					(20)*

*Note: CoP applied to extremely small amount of heat removal from components not cooled by H_2O

PROJECTING THE FUTURE OF DELL'S 'TRITON'

Dell's innovative 'Triton' technology is available for select Dell ESI customers today. The company is currently evaluating a "closed loop" version of 'Triton' that offers the same core liquid cooling technology and CPU support but removes the need for datacenters to have facility water at the rack. This has the potential to bring liquid cooling to an even broader set of scale-out customers.

Moor Insights & Strategy believes Dell can leverage their technology to support liquid-cooled, high-powered GPUs for workloads like machine learning and other HPC verticals that benefit from GPU acceleration. Moor Insights & Strategy looks forward to seeing where Dell takes their innovations in 'Triton' in the future to address more use cases and customer segments.

CALL TO ACTION

Over the last decade, IT vendors have started to use liquid cooling to improve datacenter energy efficiency and support high-powered IT equipment that is nearly impossible to cool using traditional, air-cooled methods. Some potential benefits of liquid cooling over traditional techniques include better overall system performance, a reduction in datacenter power consumption, datacenter cost savings, and other benefits such as reuse of waste heat and potential cost savings benefits for "going green". However, liquid cooling does not make sense for all environments and is still best suited for a very specific segment of the overall IT infrastructure market.

IT organizations looking to improve overall energy efficiency or who have workloads that scale well with extremely high CPU frequency and core count may benefit from using liquid-cooled IT equipment. Specific workloads that may see benefit include HPC verticals such as oil and gas, research labs, gaming, and financial services organizations doing high frequency trading. Other web-scale workloads that can experience potential benefit include certain web-scale query / search centric applications and some machine learning applications.

Not all liquid cooling approaches provide the same level of benefit in PUE, nor do all approaches effectively cool high-powered IT components. In addition, some approaches to liquid cooling may also be difficult to implement in existing datacenters.

Dell ESI has significant experience servicing large customers with liquid cooling solutions and has been working on improvements to their approach for over six years. Their 'Triton' liquid cooling innovation not only has excellent cooling capacity, but its simplicity eliminates many unnecessary elements whose operation can be costly. If you are a customer seeking compute-dense solutions with workloads that may benefit from liquid cooling, consider adding Dell to your short list of vendors for evaluation.

NOTES

[1] CPU performance measurement by Dell DCS using LINPACK as the performance benchmark

[2] Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit: http://www.intel.com/performance/datacenter. 1. Up to 53% claim based on LINPACK MP benchmark GFLOP results using Intel® MKL 11.3.2 N=170K, memory details, and OS details as measured by Dell SPA team. Configurations: Dual-Socket 200W Intel Xeon Processor E5-2697 v4 (3.2GHz, 20C, 200W TDP) scoring up to 1578 GFLOPs compared to similarly configured E5-2699 v4 scoring up to 1416 GFLOPs and E5-2680 v4 scoring 992 GFLOPs.

[3] Technique for calculating PUE from CoP was derived by Austin Shelnutt, Thermal Architect, Dell DCS

Important Information About This Paper

Author

Gina Longoria, Senior Analyst at Moor Insights & Strategy
Jimmy Pike, Technologist at Moor Insights & Strategy

Review

Patrick Moorhead, Founder, President, & Principal Analyst at Moor Insights & Strategy

Editor

Scott McCutcheon, Director of Research at Moor Insights & Strategy

Inquiries

Contact us if you would like to discuss this report, and Moor Insights & Strategy will respond promptly.

Citations

This paper can be cited by accredited press and analysts but must be cited in-context, displaying author's name, author's title, and "Moor Insights & Strategy". Non-press and non-analysts must receive prior written permission by Moor Insights & Strategy for any citations.

Licensing

Disclosures

This paper was commissioned by Dell. Moor Insights & Strategy provides research, analysis, advising, and consulting to many high-tech companies mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

Disclaimer

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. Moor Insights & Strategy disclaims all warranties as to the accuracy, completeness, or adequacy of such information and shall have no liability for errors, omissions, or inadequacies in such information. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators and not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could cause actual results to differ materially. You are cautioned not to place undue reliance on these forecasts and forward-looking statements, which reflect our opinions only as of the date of publication for this document. Please keep in mind that we are not obligating ourselves to revise or publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.