

# Can Data Center Become Water Self-Sufficient?\*

Kishwar Ahmed, Mohammad A. Islam, Shaolei Ren, Gang Quan  
Florida International University

## Abstract

To curtail data centers’ huge cooling power consumption and water demand (for cooling), air-side economizer has been increasingly adopted to cool down servers. Recently, another sustainability practice, rainwater harvesting, has also seen a growing adoption in data centers, potentially leading to water self-sufficiency without connecting to water utilities to supply cooling water. Nonetheless, various factors, e.g., unpredictable rain falls and limitations on water harvesting area, make water self-sufficiency challenging. In this paper, we present a first-of-its-kind study to evaluate whether it is feasible to achieve water self-sufficiency in data centers. We find that although water self-sufficiency depends on non-controllable factors such as weather; *improving power proportionality* (via power management) and *increasing water tank size* will increase the feasibility, relieving the requirement on water harvesting area and making water self-sufficiency a reality in different locations.

## 1 Introduction

As critical assets for digital economy, data centers have collectively accounted for over 2% of global electricity usage [13]. Meanwhile, data centers are also very “thirsty”, consuming an enormous amount of water (e.g., through evaporation in cooling towers to exhaust server heat into the environment) [18]. For example, U.S. National Security Agency (NSA) data center in Utah, when fully operational, would consume up to 1.7 million gallons of fresh water [14]. Amid the anticipation that water demand worldwide exceeds the supply by 40% in 2030 and also urged by prevalent severe droughts [15], reducing data centers’ water consumption has become increasingly imperative, as attested by LBNL’s recent guideline which places water and energy as two major considerations for U.S. federal data center consolidation [20].

While there have been numerous efforts to improve data center energy efficiency [12], attention has also been recently paid to reduce water consumption in data centers. For example, Google is using recycled/industry water or seawater instead of potable water in some of its data centers [10], AT&T is partnering with EDF to improve water efficiency in cooling towers through water treatment [5], and recent research also investigated saving water by routing more workloads to data centers with higher water efficiency [18].

Among various water-saving approaches for data center, a very appealing, and perhaps the most attractive one is using air-side economizer (a.k.a. “free” outside air cooling), which is highly power efficient compared to traditional cooling systems and has a significantly lower water footprint compared to systems with cooling towers [17]. In this paper, we consider a data center using air-side economizer with evaporation-assist as the main cooling method, where evaporation-assist is employed to reduce the outside air temperature and extend the operating hours for economization. While sustainability leaders (e.g., Google and Apple [1, 11]) have pledged to become “net-zero” (i.e., electricity usage completely offset by renewable energy) and/or power their data centers entirely by renewable energy without grid power, our paper investigates another distinctly different sustainability practice: **Can a data center become water self-sufficient?** Here, the data center cooling system does not rely on any fresh water tapped from water utilities;<sup>1</sup> instead, all the cooling water (in evaporation assist [17]) is harvested from available natural resources — rainwater. While the limited rainfall cannot possibly sustain water self-sufficiency for conventional data centers that have huge water footprints due to water-consuming cooling towers, the low water footprint of increasingly popular air-side economizers could potentially make water self-

<sup>1</sup>Data center also needs water for other purposes such as employees’ drinking, but such non-cooling water comes from utilities through a different pipe and is out of the scope of this paper [10].

\*This work was supported in part by NSF CNS-1423137.

sufficiency a reality and hence motivate our study.

*What is rainwater harvesting and why does it matter for data centers?* Rainwater harvesting is a clean and natural water resource processing, providing various benefits such as reduced rainwater pollution (e.g., runoff rainwater carries significant amount of debris) and decreased pressure on sewage systems. With increasingly frequent droughts and water scarcity, many states in the U.S. are encouraging rainwater harvesting through programs such as tax rebates or incentives on rainwater harvesting equipment [22]. In addition, there are other benefits that motivate data center operators to harvest rainwater including, among others: green certification (e.g., LEED program), social responsibility, lowering the rapidly-increasing water cost, and business risk concerns (e.g., in drought-prone areas where compliance codes on water usage are tightening [5]). Further, rainwater harvesting is cost-effective and easy to maintain, since rainwater only needs basic processing (e.g., filtration) before it can be used for evaporation-assist in air-side economizer. Industry leaders, such as Google and Equinix (a global colocation data center provider), have all implemented rainwater harvesting in (some of) their data centers.

Water self-sufficiency, without tapping into water utilities, naturally represents the *ultimate goal* of rainwater harvesting for sustainability, which brightens corporate image and brings tax benefits. Nonetheless, it is very *challenging* to achieve water self-sufficiency. First, and foremost, rainwater is scarce in supply and demonstrates a high unpredictability. Second, water harvesting area is often limited (due to physical area constraint and/or government regulations [16]), and so is water tank size for storing rainwater. In the following, we present an evaluation model for water self-sufficiency. We present a first-of-its-kind study on whether water self-sufficiency is feasible, via trace-based evaluations for data centers in six different places in the U.S. Our results show that water self-sufficiency is indeed highly feasible, under a variety of weather conditions.

As a stepping stone, understanding the feasibility at the first place can help data center operators decide *whether* and *what* additional efforts are needed for water self-sufficiency.

## 2 Rainwater Harvest and Cooling Model

This section provides some background information on rainwater harvesting in the context of data centers, and also presents the operating modes and water consumption for air-side economizer with evaporation-assist.

### 2.1 Rainwater harvesting

Rainwater harvesting area is surface area in direct contact to rainfall (e.g., rooftop, parking area), while storage tank/cistern represents the place where rainwater is stored. With large space, data centers can harvest rainwater from rooftop and/or other plane surfaces (e.g., parking lot, terraces). Besides, applicable area for data center water storage can also be explored for rainwater harvesting (e.g., Google's rainwater retention pond in its South Carolina data center [10]). In Section 3, we demonstrate optimal harvesting area and water tank size for data centers located in places with volatile weather pattern.

### 2.2 Cooling system

Air-side economizer directs outside cold air into the data center to cool down servers and is mostly suitable in regions with cool/cold climate. This type of cooling systems is becoming popular for energy savings over traditional cooling systems [4, 7, 17]. In our study, we consider the data center uses air-side economizer with evaporation-assist as the cooling system. In this cooling system, outside air first enters the air-mixing chamber, where it is mixed with the return air from the server room. Controllable openings are used to control the amount of outside air and return air that enters the air-mixing chamber. The air from the mixing chamber then passes through the evaporation chamber where water is sprayed from sprinklers onto the air. The air gets cooled down through water evaporation. The cold air (i.e., supply air) is then directed into the server room to cool down the servers. Humidity and temperature of the supply air are conditioned by controlling the flow rate of air and evaporation (amount of water spray), such that the supply air falls within the recommended operating range by ASHRAE:  $64.4^{\circ}F(18^{\circ}C)$  to  $80.6^{\circ}F(27^{\circ}C)$ , and relative humidity of 40% to 60% [2].

Based on the environmental condition, the cooling system may run in different operating modes. Here, we briefly present the operating conditions of these modes.

**Direct-air:** The cooling system runs in this mode when the outside air temperature and humidity falls within the ASHRAE specified range [2]. In this mode, the outside air is directly sent to the server room without any conditioning other than filtering for particles. This operating mode has no water footprint and a minimum/negligible power consumption.

**Evaporative:** When the outside air temperature is above the allowed range but the humidity is low, evaporation is used to cool down the outside air. The temperature drop during evaporation depends on the difference between dry-bulb and wet-bulb temperatures of the outside air. Humidity of the air increases during the evaporation process, and therefore the evaporation rate is controlled to keep the supply air humidity within specification. This

operating mode has a high water footprint because of the evaporation, but still low power consumption.

**Mechanical:** The cooling system uses mechanical cooling when the outside air cannot be conditioned for the data center. This happens when the temperature and humidity of the outside air are both higher than the allowed range. In this mode, the return air is cooled down using mechanical cooling and then used in the server room. This operating mode has no water footprint, but has the highest power consumption due to the high power requirement of the mechanical cooling.

**MIX:** When the outside air temperature is low but the humidity is high, this operating mode is used to reduce the humidity of the outside air by mixing it with the server room return air which has low humidity. In this process, the temperature of the mixed air also increases. The supply air temperature and humidity is controlled by adjusting mixing ratio of the outside air and return air. This operating mode also has no water footprint, and has low power consumption.

**MIX+Evap.:** In this operation mode, both mix and evaporation is used to condition the outside air. This mode of operation is required for very cold outside temperature, where the amount of water present in the air is also low. De-humidification process increases the temperature, and evaporation adds water to the air to get it into allowed range of operation. This operating mode has a moderate water footprint and low power consumption.

### 2.3 Water consumption

The cooling system operation modes that involve evaporation have water footprint. We calculate the water usage effectiveness (WUE) of the cooling system at different weather conditions using the formula

$$\text{WUE} = (X_s - X_a) \times \frac{3.6}{h_e - h_s}, \quad (1)$$

where WUE has unit L/kWh,  $X_a$  and  $X_s$  are the humidity ratios of outside air and supply air, respectively, with unit  $gm_w/kg_a$  (grams of water/moisture per kg of air),  $h_s$  and  $h_e$  are the enthalpy of the data center supply air and return air, respectively, with unit kJ/kg. In Equation (1),  $h_e - h_s$  represents the amount of heat each kg of air carries out of the data center, while  $X_s - X_a$  measures the amount of water added to each kg of air during evaporation to condition the outside air to be used in the server room. We can determine the humidity ratios and enthalpy from the temperature, humidity and pressure of the air at different points (e.g., outside, supply, exhaust, etc.). We consider that there is no water condensation inside the data center, i.e., the amount of water in the air does not change when it passes through the server room (i.e., humidity ratio remains constant). Although WUE is also affected by other factors (e.g., water leakage), Equation (1) gives us a reasonable estimate of water usage

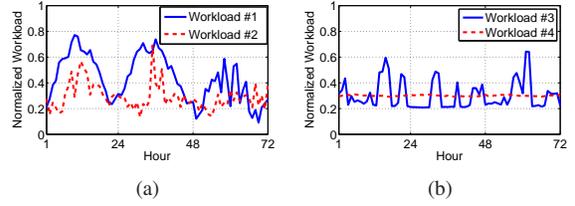


Figure 1: Workload trace.

for air-side economizer with evaporation-assist. For example, the WUE calculated for Prineville, OR, using (1) matches well with that of Facebook’s data center with a similar cooling system [3].

## 3 Case Study

In this section, we present trace-based case studies of water self-sufficiency feasibility for data centers located at six locations in the U.S. After introducing our data sets, we present experimental results and observations.

### 3.1 Data sets

We consider six geographically distributed data centers located at: Prineville (OR), Forest City (NC), Los Angeles (CA), Ashburn (VA), Chicago (IL), and New York (NY), which all have large (colocation) data centers [9], while other places can be studied using the same methodology. We conduct our simulation with a data center with 6MW static/idle power and 10MW peak IT power that has an air-side economizer.

- **IT load:** We evaluate four different workload traces, which we denote as “Workload #1”, “Workload #2”, “Workload #3” and “Workload #4”. These workloads include web service traces (e.g., Hotmail, Wikipedia), I/O log (Microsoft Research), and colocation data center UPS load (Terremark in Miami, FL). Traces are taken from prior publications [19,21] and our direct collaboration. We use Workload #1 as our default workload. Due to limited availability of trace data throughout the year, we add 30% random variation and extend the available traces to get one-year trace, and scale the traces to have a 30% average server utilization. Fig. 1 shows workload traces (normalized based on data center capacity) for four different workloads for three days.

- **Weather data:** We obtain hourly weather data from [8] at the six different cities where we consider the data centers’ locations for the year of 2013. The operation mode of the cooling system is determined based on the dry-bulb temperature and relative humidity. We consider that the recommended range of operation follows ASHRAE [2]:  $64.4^\circ F (18^\circ C)$  to  $80.6^\circ F (27^\circ C)$ , and relative humidity of 40% to 60%. The exhaust air temperature is considered to be at  $96.8^\circ F (36^\circ C)$ .

- Others: Default size of water tank is considered to be one million gallons, with an initial storage of 50%. We set the minimum water level at 5% of the total water capacity for emergency uses. Note that one-million-gallon water tank is not restrictive for a data center with 10MW IT power (e.g., Google has several 240,000 gallon water tanks in its SC data center [10]).

### 3.2 Results

Now, we present our results on case studies using the aforementioned data sets.

#### 3.2.1 Cooling modes and WUE

We first study the cooling system operation modes and resulting WUE. In Fig. 2(a), we show the operation time of different cooling modes. We see that the MIX+Evap. mode is the most prominent mode of operation for air-side cooling system with evaporation-assist. In general, MIX is the second most frequently used mode. The direct-air mode is used less than 10% of the time for all locations except CA. Also, we observe that mechanical cooling is not used at all for OR and CA data centers. All the cooling modes, except the mechanical cooling, take advantage of the outside air and have low power consumption, thus demonstrating the potential of using air-economizer in these locations. Fig. 2(b) shows the resulting average annualized WUE of the cooling system. We see that, OR data center has the highest WUE, which is the result of the cooling system running in MIX+Evap. mode almost 70% of the time. Data center at CA has the lowest WUE because it operates in the water-less MIX mode most of the time. We also show the annual total rainfall of 2013 in the six data center locations in Fig. 2(c). It shows that OR and CA have very low rainfall in 2013 (less than 5 inches during the entire year), where other locations have more than 30 inches of rainfall, demonstrating that WUE is not correlated with the amount of rainfalls. Thus, geographic load balancing could potentially help achieve water self-sufficiency in multiple locations.

#### 3.2.2 Harvesting area

The harvesting area determines how much water is added to the water tanks when there is rainfall. For the same amount of rainfall, a larger harvesting area thus can collect more water. In Fig. 3, we study the harvesting area required for different locations as well as for different workload types. We show in Fig. 3(a) that, except for OR, all the other locations require a harvesting area of less than 500,000 sqft to become water self-sufficient. NC data center has the lowest surface requirement because it has a high rainfall. For CA, despite having very low amount of annual rainfall, it still can be water self-sufficient at harvesting area greater than 200,000 sqft,

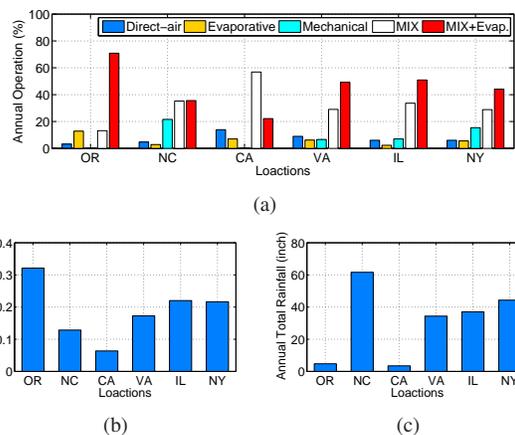


Figure 2: (a) Annual operation percentage of different cooling modes. (b) WUE of different locations. (c) Annual rainfall.

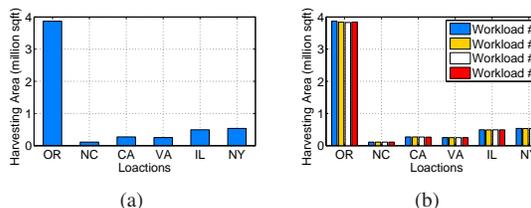


Figure 3: Harvesting area requirement for different locations and workloads.

because of the very low average WUE. OR data center requires a harvesting area of more than 1.5 million sqft to become water self-sufficient, and thus is considered infeasible for water self-sufficiency. Fig. 3(b) shows the effect of different workload types on the harvesting area requirement. We see that despite the variety in workloads, harvesting area requirement does not vary much with workload pattern because the average workload (and hence, power consumption) is kept the same.

#### 3.2.3 Impact of different factors

Here, we study the impact of different factors that affects water self-sufficiency.

**Power proportionality.** Power proportionality refers to data center’s power consumption in proportion to its workload. In the default setting, we do not consider advanced power management (e.g., dynamically turning off unused servers to reduce power) in the data center, and hence there is a low power proportionality (e.g., data center uses 6MW power even for zero workload). Fig. 4(a) shows the impact of improved power proportionality on the harvesting area requirement. We see that, when perfect power proportionality is achieved (zero static power), the harvesting area requirement becomes less

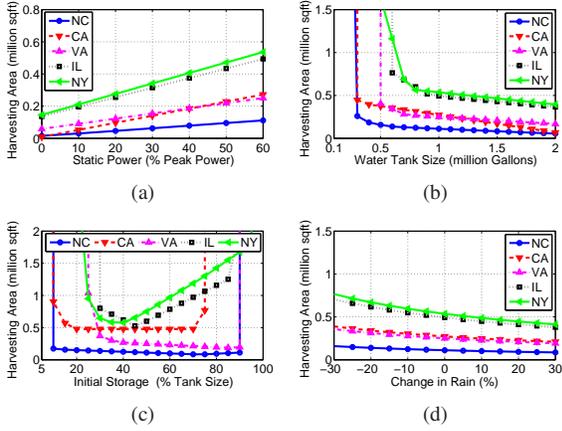


Figure 4: Impact of (a) power proportionality, (b) water tank size, (c) initial water storage, and (d) annual rainfall.

than 200,000 sqft for all the locations (except for OR, which is omitted in these figures for its very high water consumption and also harvesting area requirement).

**Water tank size.** In Fig. 4(b), we show the impact of water tank size on the minimum harvesting area requirement. We see that the harvesting area requirement rises sharply when the tank size falls short of a certain value for all the locations. This indicates that there is a constraint on the minimum tank size for a feasible harvesting area. We also observe that, the harvesting area slowly decreases with increasing tank size, showing a saturation of benefit from larger water tanks. This is due in part to inconstant rain falls, i.e., rainwater cannot be collected continuously even with very large water tanks.

**Initial water storage.** Fig. 4(c) shows the impact of initial storage size on harvesting area. We determine the minimum harvesting area with the constraint that the amount of water present in the water tank at the end of the year is not less than the initial water storage. This ensures water self-sufficiency is sustainable for subsequent years. We see that no locations can achieve water-self-sufficiency when the initial storage is too low (below 10%) or too high (above 90%). Data centers at NC, CA and VA have a wider range of initial water level for which the minimum harvesting area is not affected, while for data centers at IL and NY have the optimal initial water level around 40%.

**Rainfall.** Fig. 4(d) shows the impact of precipitation on harvesting area requirement, where the yearly average precipitation is varied by scaling the 2013-year precipitation from  $-30\%$  to  $+30\%$ . We see that the harvesting area requirement demonstrates a natural trend of increased harvesting area for lower precipitation, but with only a small variation with precipitation changes. Hence, once rainwater harvesting system is built, it is unlikely to be invalidated by normal changes in precipitation.

	Static Power (% Peak Power) - Water Tank Size			Feasibility	Harvesting Area
	60% - 1 million gallons	0% - 1 million gallons	0% - 2 million gallons		
Prineville, OR	3,868,000	1,112,500	522,550	High	<100,000 sqft
Forest City, NC	111,600	15,000	"zero"	Medium	<200,000 sqft
Los Angeles, CA	272,400	8,500	"zero"	Low	<500,000 sqft
Ashburn, VA	250,000	59,000	10,750	None	>500,000 sqft
Chicago, IL	494,000	136,000	29,150		
New York, NY	536,000	146,000	25,650		

Figure 5: Summary of water self-sufficiency feasibility.

### 3.3 Summary

Now, we summarize our findings. Considering that a 10MW data center requires 100,000 sqft for the servers alone (with the industry average power density of 100 watts/sqft [6]), we set different levels of water self-sufficiency feasibility based on the harvesting area requirement as shown in Fig. 5. For example, a harvesting area requirement of less than 100,000 sqft is highly feasible, whereas greater than 500,000 sqft is deemed as infeasible. We can see that, with 60% static/idle power (i.e., no advanced power management for energy saving) and 1 million gallons of water tank size, data center in NC achieves medium feasibility (i.e., requirement for harvesting area is feasible), while it is difficult to achieve water self-sufficiency for data centers in OR and NY. However, by reducing static/idle power (i.e., improved power proportionality), feasibility of all data centers can be improved. For example, with perfect power proportionality (0% static power), data centers at NC, CA and VA become highly feasible for achieving water self-sufficiency (i.e., low harvesting area requirement). Moreover, increasing water tank size further assists with achieving water self-sufficiency: for example, increasing water tank size to 2 million gallons improves feasibility of data centers in all locations. “zero” surface area for data centers at CA and NC indicate that, the data centers can sustain on the initial water storage (default of 50% of the tank) when the water tank size is two million gallons.

In sum, although water self-sufficiency depends on non-controllable factors such as weather; *improving power proportionality* (via effective IT power management) and *increasing water tank size* will significantly increase the feasibility.

## 4 Conclusions

This paper studies if it is feasible to achieve water self-sufficiency in data centers, through rainwater harvesting area and choice of air-side economizers. We show that the feasibility is strengthened by improving power proportionality and increasing water tank sizes. Further, “uncorrelation” of water efficiency and amount of rainfalls suggests that geographic load balancing may help achieve water self-sufficiency in multiple locations.

## References

- [1] Apple and the environment, <http://www.apple.com/environment/>.
- [2] 9.9 (2011) thermal guidelines for data processing environments—expanded data center classes and usage guidance. *Whitepaper prepared by ASHRAE technical committee (TC)* (2011).
- [3] Facebook Dashboard: PUE & WUE, [https://www.facebook.com/PrinevilleDataCenter/app\\_399244020173259](https://www.facebook.com/PrinevilleDataCenter/app_399244020173259).
- [4] Inside facebook’s forest city data center, <https://www.facebook.com/video.php?v=4638122307624>.
- [5] AT&T Sustainability, <http://www.att.com/gen/landing-pages?pid=24188>.
- [6] Data center density hits the wall, [http://www.computerworld.com/s/article/9144466/Data\\_center\\_density\\_hits\\_the\\_wall](http://www.computerworld.com/s/article/9144466/Data_center_density_hits_the_wall).
- [7] The evolution of facebook’s data center cooling, <http://www.datacenterknowledge.com/archives/2012/12/04/evolution-of-facebooks-cooling/>.
- [8] Historical Weather, <http://www.wunderground.com/>.
- [9] DATACENTERMAP. Colocation USA, <http://www.datacentermap.com/usa/>.
- [10] GOOGLE. Data center efficiency, <http://www.google.com/about/datacenters/>.
- [11] GOOGLE. Google’s green PPAs: What, how, and why, [http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/www.google.com/en/us/green/pdfs/renewable-energy.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/en/us/green/pdfs/renewable-energy.pdf).
- [12] HOELZLE, U., AND BARROSO, L. A. *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines*. Morgan and Claypool Publishers, 2009.
- [13] KOOMEY, J. G. Growth in data center electricity use 2005 to 2010, 2011.
- [14] KSL. New Utah NSA center requires 1.7M gallons of water daily to operate, <http://www.ksl.com/?sid=25978926>.
- [15] MCKINSEY REPORT. Charting our water future: Economic frameworks to inform decision-making, 2009.
- [16] NATIONAL CONFERENCE OF STATE LEGISLATURES. State rainwater harvesting and graywater laws and programs, 2013.
- [17] NIEMANN, J., BEAN, J., AND AVELAR, V. Economizer modes of data center cooling systems. *Schneider Electric Data Center Science Center Whitepaper* (2011).
- [18] REN, S. Optimizing water efficiency in distributed data centers. In *Cloud and Green Computing* (2013).
- [19] THERESKA, E., DONNELLY, A., AND NARAYANAN, D. Sierra: a power-proportional, distributed storage system. *Microsoft Research Ltd., Tech. Rep. MSR-TR-2009-153* (2009).
- [20] TSCHUDI, W. Guideline for water and energy considerations during federal data center consolidations, Dec. 2013, [http://www1.eere.energy.gov/femp/pdfs/consolidation\\_guidelines.pdf](http://www1.eere.energy.gov/femp/pdfs/consolidation_guidelines.pdf).
- [21] URDANETA, G., PIERRE, G., AND VAN STEEN, M. Wikipedia workload analysis for decentralized hosting. *Computer Networks* (2009).
- [22] U.S. EPA. Rainwater harvesting conservation, credit, codes, and cost: Literature review and case studies, 2013.