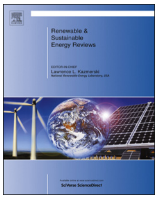
Renewable and Sustainable Energy Reviews 31 (2014) 622–638

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser



A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities

Khosrow Ebrahimi, Gerard F. Jones, Amy S. Fleischer*

Department of Mechanical Engineering, Villanova University, Villanova, 800 Lancaster Avenue, PA 19085, USA

ARTICLE INFO

Article history:
Received 22 May 2013
Received in revised form
22 November 2013
Accepted 20 December 2013

Keywords:
Data center
Waste heat recovery
Waste energy reuse
Absorption refrigeration
Organic Rankine cycle
Thermoelectric

ABSTRACT

The depletion of the world's limited reservoirs of fossil fuels, the worldwide impact of global warming and the high cost of energy are among the primary issues driving a renewed interest in the capture and reuse of waste energy. A major source of waste energy is being created by data centers through the increasing demand for cloud based connectivity and performance. In fact, recent figures show that data centers are responsible for more than 2% of the US total electricity usage. Almost half of this power is used for cooling the electronics, creating a significant stream of waste heat. The difficulty associated with recovering and reusing this stream of waste heat is that the heat is of low quality. In this paper, the most promising methods and technologies for recovering data center low-grade waste heat in an effective and economically reasonable way are identified and discussed.

A number of currently available and developmental low-grade waste heat recovery techniques including district/plant/water heating, absorption cooling, direct power generation (piezoelectric and thermoelectric), indirect power generation (steam and organic Rankine cycle), biomass co-location, and desalination/clean water are reviewed along with their operational requirements in order to assess the suitability and effectiveness of each technology for data center applications. Based on a comparison between data centers' operational thermodynamic conditions and the operational requirements of the discussed waste heat recovery techniques, absorption cooling and organic Rankine cycle are found to be among the most promising technologies for data center waste heat reuse.

Contents

---

* Corresponding author. Tel.: +1 610 519 4996.
E-mail address: amy.fleischer@villanova.edu (A.S. Fleischer).

1364-0321/$ - see front matter
http://dx.doi.org/10.1016/j.rser.2013.12.007

1. Introduction

During the past two decades, the increasing demand for data processing, data storage systems and digital telecommunications coupled with the simultaneous advances in computer and electronic technology have resulted in a dramatic growth rate in the data center industry. In United States, during the years 1998–2010, the number of data centers increased from 432 to 2094 [1] and the overall power usage of data centers experienced a doubling between the years 2000 and 2006. Recent energy statistics indicate that the data center industry is responsible for 1.3% of the world and 2% of the United States electricity consumption [1]. In fact, between 25% and 35% of the worldwide power consumption of data centers (30 Billion Watts) is consumed by US data centers [1].

The growth has been not only in the number of data centers, but also in the size, floor area and/or computing density of datacenters [2]. The data center industry operates on many different scales, with the construction of huge datacenters with floor areas as large as ~9000 m^2 containing thousands of server racks and several megawatts power usage [2], yet also with the design of compact data centers with more and more computing power packed into smaller and smaller spaces. The construction and annual operating costs for a typical data center are ~$15,000/$m^2$ and $1500/$m^2$, respectively [2].

Considering the ongoing growth of the data center market coupled with the development of higher power density server components, it is expected that the share of electricity consumption by data centers will continue to increase for the foreseeable future. The predictions for annual increases in data center power demand are as high as 15–20% [3].

Despite increasing efforts to integrate renewable energy sources into the overall US power grid, fossil fuels are still the major energy source for electricity production. Considering the limited supply of fossil fuels and the consequent air pollution and global warming of using them to produce electricity, not to mention the increasing cost of electricity in many markets, there is a growing effort to capture and reuse waste heat in all types of energy conversion systems. In many cases, this effort will need to include substantial work into improving capture and transport of the waste heat, and into increasing the efficiency of the recovery system.

Large data center operators, including familiar companies such as Google (with ~300 million watts of power usage) and Facebook (with ~60 million watts of power usage) are taking action to increase the fuel efficiency and decrease their wasted power [1]. In fact, most data centers rarely operate at their maximum rated load. The majority of servers operate at or below 20% of their maximum capacity most of the time yet even when the system is idle, 60–100% of the maximum power is still drawn from the grid [4,5]. Almost all the electrical power supplied to the server is dissipated into heat, necessitating the use of large scale cooling systems to keep the server rack temperatures in a safe operational range. There are significant research efforts ongoing to develop dynamic need-based resource allocation as one way to reduce energy dissipated by data centers [6,7]. In combination with recovery and reuse of the waste heat, these efforts can lead to more efficient and cost-effective operation of data centers.

Many data centers feature servers with power densities in excess of 100 W/cm^2 and even up to as high as 200 W/cm^2, which means that a rack with a 0.65 m^2 footprint has heat dissipation requirements as high as 30 kW, or roughly 30 times higher than the amount of energy dissipated by a typical rack with the same footprint in 1990 [8]. This increased heat dissipation demand [9,10], increases the costs associated with powering and cooling datacenters. Thus, the recovery and reuse of waste heat energy has the potential to significantly reduce data center operational costs.

The main barrier to the implementation of waste heat recovery and reuse systems into operational data centers in that in contrast to many industrial waste heat recovery systems, the heat, although plentiful, is of low quality. The capture temperature is limited by the temperature limits of the electronics, which remain below 85 °C in most cases. This low quality temperature makes it quite challenging to be reuse the heat through conventional thermodynamic cycles and processes.

This study provides a much needed comprehensive review of all commonly available waste heat reuse techniques with a specific focus on their particular application to data center operating conditions. Each technique will be examined through the lens of a data center operator with consideration of both retrofit to legacy air cooled data centers and to integration into newer water cooled data centers. Through a comparison between the technology's operational requirements with the specifications of waste heat sources and streams in different type of data centers, the most promising solutions for waste heat reuse are identified.

This review paper is presented in four sections. In the first section, the significance and necessity of research on waste heat recovery in data centers was explained through a review of the past, current, and projected future growth trend of data center industry and its impact on energy consumption and environment. In the second section, the energy generating components of data centers are introduced and their physical arrangement is described. A comprehensive literature review is given in which thermal loads and temperature limits in different components are categorized and tabulated depending upon the data center cooling type. This section also presents a detailed review of heat sources and streams in different data center designs which provides the reader with a broad and detailed background on data center thermodynamic operating conditions. In the third section, eight potential; low-grade waste heat recovery technologies are reviewed along with their operational requirements. The benefits and challenges associated with the implementation of each technology for data center waste heat recovery are discussed and listed. In the final section, the suitability of each technology for different data center designs is discussed and the most promising options for data center applications are identified for further investigation.

2. Overview of data centers

2.1. Physical organization

A data center is a space allocated to house most of the ICT (Information and Communication Technology) modular assets such as servers, switches, and storage facilities and to control the environmental conditions (temperature, humidity, and dust) to ensure that the ICT systems operate reliably and in a safe and efficient manner [2,11]. Depending upon the scale of the business, a data center might include a single rack of equipment or a few or even many racks and cabinets. A rack is a standardized metal frame or enclosure in which the ICT modular assets inserted horizontally. A standard rack has dimensions of 78 in. in height, 23–25 in. in width and 26–30 in. in depth. In a typical data center, the standard racks are arranged in rows at a pitch of approximately 2 m [5]. The height or thickness of the modular asset mounted in a rack is described by a unit of measurement which is named "U". U is approximately equivalent to 1.8 in. A majority of servers (usually with single or dual socket processors) have 1 U thickness, however, some larger ones (such as servers with 4-socket multiprocessors) may have thicknesses larger than 2 U [11]. A typical full size rack can take a total of 42 of 1 U modular assets [11]. Other modular assets which might be placed in racks are power distribution units, built-in KVM switches (keyboard, video, mouse), rack-level air or liquid cooling, and in some cases, the rack-level management unit. An alternate method of server arrangement through which greater compactness and functionality can be provided features server blades housed in self-contained enclosures called chassis [11]. Each chassis has its own power supply, fans, backplane interconnect and management infrastructure. Although the height of a typical chassis is 13 in., models with height as high as 10 U (18 in.) have been introduced to the market. A chassis can house at least 8 and perhaps as many as 16 server blades [11]. A full-height rack can be filled up to 64 [12], 84 [13] or even 96 [11] server blades depending upon the size and the chassis capacity. Therefore, this design vastly increases per-rack power.

2.2. *Data center thermal loads and temperature limits*

The increasing demand for the ICT services from one side and the direct proportionality between data center costs and floor area from the other side, have led manufacturers to design and produce more compact and higher power modules. While the energy flux dissipated by traditional data centers is in the range of 430–861 W/m^2, the energy flux dissipated by the newer generations of data centers has been increased at least by 10 times (6458–10,764 W/m^2) [14]. Comparing the range of data center energy dissipation with the capacity of conventional HVAC systems for similar size rooms (40–86 W/m^2 [2]), it can be seen that the design and manufacture of thermal management systems is one of the most challenging aspects of data center design. The thermal management system must be capable of handling the increasing thermal loads while maintaining the temperature of electronic components at a safe operational level. In order to design such a system, it is necessary to have accurate and reliable information about the maximum thermal loads and temperature limits in each component of a data center. Knowledge of the thermal loads and temperature limits is also essential for waste heat recovery purposes.

A server, which can be considered as the smallest data processing unit in a data center, typically features highly integrated microprocessors, additional memory (or DIMMs) and auxiliary components including input-output (I/O) devices, mass storage (disk drive) and a power supply. The microprocessors contain at least one CPU core as well as memory, network, graphics controller and cache memory. Microprocessor chips are the major power dissipation components in servers. A typical server with two processors consumes almost 50% of the total server power through the microprocessors. Table 1 shows the reported heat load and size of the microprocessors and CPU cores in recent studies [5,12,15–20].

However, microprocessors are not the only power dissipation components in a typical server. An individual hard disk can dissipate powers as high as 12 W and up to 20–30% of the total power supply can be consumed by mass storage devices [11]. Memory components also draw power and a DIMM′s power consumption in standby mode is approximately 3 W while the power consumption in full load mode increases to around 9.8 W [11]. This is of course a range, as Iyengar et al. [20] considered a slightly lower full load power consumption (~6 W) for each DIMM in their study, and Campbell and Tuma [19] assumed 8 W. According to Marcinichen et al. [12], the auxiliary equipment dissipates as much as 55.6 W per half blade server.

Table 2 summarizes the total heat load of standard and blade servers as recently reported [5,12,16–18,20]. The table indicates that for standard servers the total power consumption is typically in the range between 300 and 400 W, however, for highly populated servers the power consumption can reach up to 525 W. For blade servers, the power consumption was around 250 W each. With the current growth rate in the amount of power dissipated by ICT equipment, it is expected that the power densities for computer servers and highly integrated communication devices will increase to 8 kW/m^2 and 15 kW/m^2, in the next few years [21].

While in legacy data centers, per-rack power consumption is about 7 kW [11], the power consumption of a high performance fully utilized rack is on the order of 10–15 kW [12], and the power dissipated by racks loaded with blade servers can approach 21 kW [11]. ASHRAE anticipates power consumptions of 60 kW for a single rack filled with extreme density communication equipment and 35 kW for a single rack filled with extreme density computer servers [22].

Table 2
Heat load of servers/blade servers in recent literature.

Investigator(s)	Total server/blade power consumption (W)
Patel (2003) [5]	400 (Standard) 250 (blade)
Marcinichen et al. (2012) [12]	300 (Blade)
Samadiani et al. (2008) [16]	525 (Standard)
Shah and Patel (2009) [17]	250 (Standard)
Marcinichen et al. (2010) [18]	300 (Blade)
Iyengar et al. (2012) [20]	400 (Standard)

Table 1
Heat load and physical size of microprocessors/cores in recent literature.

Investigator(s)	Microprocessor				Cores		
	Size	Heat load (W)	Heat flux (W/cm^2)	Number of cores	Size	Heat flux (W/cm^2)	Heat load (W)
Patel (2003) [5]	20 mm × 20 mm	100–125	NA	1	5 mm × 5 mm	200	50
Marcinichen et al. (2012) [12]	NA	150	100	2	NA	NA	NA
Trutassanawin et al. (2006) [15]	NA	NA	NA	NA	1.9 cm^2	40–75	NA
Samadiani et al. (2008) [16]	30 mm × 30 mm	900	100	NA	NA	NA	NA
Shah and Patel (2009) [17]	NA	5, 50, 125, 175	NA	NA	NA	NA	NA
Marcinichen et al. (2010) [18]	2.5 cm^2	162.5	65	2	NA	NA	NA
Campbell and Tuma (2012) [19]	NA	150	NA	2	0.51 cm^2	NA	40.5
Iyengar et al. (2012) [20]	NA	130	NA	NA	NA	NA	NA

Recent literature reports various estimates for future rack loading. For instance, Schmidt et al. [23] investigated the effectiveness of different cooling strategies in a data center considering three different rack power levels including 4, 8 and 12 kW. Rambo and Joshi [2] considered high power racks with heat dissipations of 57 kW in a model for data center airflow and heat transfer. While acknowledging a single rack can dissipate powers as high as 40 kW, Hamann et al. [24] analyzed racks with 25–30 kW heat load. Samadiani et al. [16] investigated a situation in which each rack featured 6 server chassis, each with a power consumption of 5 kW, for a rack power dissipation of 30 kW. McAllister et al. [25] studied thermal inefficiencies in the cooling system of a typical legacy data center, by developing an exergy-based model in which two high and low heat loads of 3583 W/m^2 and 717 W/m^2 were taken into account for the racks. Considering the footprint area of the studied racks (0.6 m by 0.6 m), the amount of heat dissipated by a single rack was between 0.26 and 1.3 kW. Campbell and Tuma [19] considered racks with 250 kW power dissipation for liquid cooled systems and 60 kW for racks with air-forced cooling systems. Marcinichen et al. [12] stated that in designing cooling system for today's data centers the assumed heat capacity for the racks is in the range of 10–15 kW, however, if rack is filled with supercomputer servers, it can generate in excess of 60 kW of heat. Table 3 summarizes these reported levels of heat dissipated by a single rack as considered in the thermal analysis of data centers performed by different investigators between 2001 and 2012 [2,12,16,18–20,23–39].

If attention is turned from the power loading to the temperature limits in ICT equipment, the majority of the electronics thermal management research considers 85 °C as the maximum allowable junction temperature for the safe and effective operation of microprocessors [5,13,15,18,40]. However, there are few other references in the literature which recommend slightly higher or lower limits as the junction temperature limit. For instance, Schmidt and Notohardjono [41] and Ohadi et al. [42] considered 100 °C and 78 °C, respectively, as the maximum microprocessor operating temperature. The same temperature limit of 85 °C is also typically applied to DIMMs [43,44]. However, the temperature limit for hard disk drives is noticeably lower than for microprocessors and DIMMs. In fact, working temperatures above 40–45 °C for long periods of time, increase the failure probability in HDDs [45]. Despite this, some manufacturers have set higher temperature limits up to 60 °C [46,47] for their product disk drives.

Considering the collated data center operating characteristics and temperature limits in Table 3, the authors of this paper have selected a range of thermal loads and temperature limits to represent a "typical" data center characteristic (Table 4) and will be used in the thermodynamic analysis and technical assessment of waste heat recovery techniques in data center applications presented here.

2.3. *Management of waste heat sources and streams in data center cooling systems*

The Computer Room Air Conditioning unit (CRAC) is responsible for controlling the temperature and humidity in the data center such that all the electronic equipment operates in a safe and efficient way. The design of this unit for energy efficiency is one of the major challenges in the construction and operation of data centers. Comparing the high level of dissipated heat flux (as high as ~ 100 W/cm^2) in newly designed high power density data centers with the much lower heat removal capacity of air (~ 37 W/cm^2) implies that the forced-air cooling systems, widely used in legacy data centers, will not be able to handle the thermal loads in new data center designs [12]. Therefore, thermal management systems are shifting from traditional air cooling to liquid or two-phase cooling.

Capturing and reusing the resultant data center waste heat is the major focus of this review. As the waste heat is captured through the coolant stream flow for most heat recovery techniques, the quantity and quality of waste heat strongly depends on the type and specification of thermal management system. In this subsection, current and developmental cooling technologies for data centers are reviewed, their corresponding coolant flow rates and temperatures are surveyed and the best locations for capturing the waste heat are discussed.

Table 4
Summary of "typical" data center thermal loads and temperature limits.

Power loads	
Component	**Values**
Processors	60–75 W each (2 per server)
DIMM	6 W each
Auxiliary power per server	150–250 W
Total power per server	300–400 W
Rack capacity	1 U servers, up to 42 per rack Blade servers at 10 U, up to 64 per rack
Total rack power	13–26 kW
Racks per data center	250
Total power per data center	3.2–6.5 MW
Temperature limits	
Component	**Values (°C)**
Processor	85
DIMM	85
Disk drive	45

Table 3
Single rack levels of heat dissipation.

Investigator(s)	Rack power (kW)	Investigator (s)	Rack power (kW)
Patel et al. (2001) [26]	14.4	Li et al. (2007) [37]	30
Schmidt and Cruz (2002) [27]	2–4	Rambo and Joshi (2007) [2]	57
Patel et al. (2002) [28]	12	Hamann et al. (2008) [24]	25–40
Sharma et al. (2002) [29]	12	McAllister et al. (2008) [25]	0.26–1.3
Rambo and Joshi (2003) [30]	4.23	Samadiani et al. (2008) [16]	30
Schmidt et al. (2005) [23]	4, 8, 12	Marcinichen et al. (2010) [18]	10
Shrivastava et al. (2005) [31]	12	Joshi and Samadiani (2011) [38]	35
Bhopte et al. (2005) [32]	4.5	Ahuja (2012) [39]	5
Iyengar et al. (2005) [33]	4–36	Campbell and Tuma (2012) [19]	60–250
Sharma et al. (2005) [34]	15.75	Marcinichen et al. (2012) [12]	10–60
Rambo and Joshi (2006) [35]	3.2	Iyengar et al. (2012) [20]	15
Beitelmal and Patel (2007) [36]	10		

Table 5
Air-cooled data center heat sources and streams in recent literature.

Investigator(s)	Cold aisle temp., °C	Hot aisle temp., °C; max temp. rise, °C	Rack power (kW)	Air flow rate per rack (CFM)
Patel (2003) [5]	25	40/15	NA	NA
Schmidt et al. (2005) [23]	Ideal: 10–32; actual: 10–15 rack bottom; 30–40 rack top	NA/NA	10	1500
Crippen et al. (2005) [51]	NA	NA/NA	NA	200–500 (standard) 1000 (max)
Rambo and Joshi (2007) [2]	15 (at CRAC outlet)	NA/40	57	2400
Beitelmal and Patel (2007) [36]	20 (standard) 35 (max)	NA/NA	NA	NA
Li et al. (2007) [37]	15	NA/NA	NA	310 (per server)
McAllister et al. (2008) [25]	NA	NA/NA	727–3583 W/m^2	0.589–289 kg/s per 0.6 m × 0.6 m tile
Hamann et al. (2008) [24]	13 (at perforated tiles)	50/NA	NA	NA
Samadiani et al. (2008) [16]	15 (at CRAC outlet) 15–32 in aisle	NA/NA	3–30 (525 W per server)	750–2730 (125 per server)
Shah et al. (2008) [40]	16	NA/NA	NA	750 cfm per 0.6 × 0.6 m tile
Kant (2009) [11]	21	40/NA	NA	NA
Samadiani et al. (2010) [13]	15–32	NA/NA	NA	NA
Wang et al. (2010) [50]	28 (max)	NA/NA	NA	NA
Kahalifa and Demetriou (2010) [52]	27 (max)	NA/10 (ideal) ~20 (max)	NA	NA
Ohadi et al. (2012) [49]	5	NA/NA	85 W (Single CPU)	26.7

2.3.1. *Air-cooled systems*

In most air-cooled or legacy data centers, which constitute the majority of existing data centers, server racks are arranged into cold and hot aisles. In cold aisles, the front sides of the server racks face each other and this aisle provides the cool intake air to each server. In hot aisles the rear sides of the racks line the aisles and the hot exhaust air exits each server here. The chilled air produced by the CRAC unit is driven into the cold aisles, either through the floor plenum and perforated tiles (in raised floor design) or through diffusers in the ceiling (in non-raised floor design). The warm air in hot aisles is captured and returned to the intake of the CRAC. The heat in the CRAC is absorbed into a chiller/cooling tower loop for ultimate dissipation to ambient. According to Patel [5], the air conditioning system in a legacy data center is designed based on the racks′ maximum power dissipation and a typical temperature rise of 15 °C for the air flow passing through the high power density servers. Patel [5] and Bash et al. [48] point out that for a data center with an efficient air cooling system, the cold air is typically supplied at 25 °C and the exhaust air leaves the room and returns to CRAC at 40 °C.

Schmidt et al. [23] studied a number of parameters affecting the performance of cooling in legacy data centers. The parameters included: ceiling height (the space between the rack top surface and ceiling where heat traps and hot air stratification may occur), raised floor/dropped ceiling height (which is important for proper distribution of air between the racks and servers), and airflow direction in the room. The airflow recirculation, created by the turbulent mixing of hot and cold airs over the racks, was identified as the major cooling problem in raised floor data center designs. Because of this recirculation and resultant permeation of hot air into the cold aisle, the temperature of air in the cold aisle is not uniform along the height of the rack. The inlet air temperature at the top of the rack can be as high as 30–40 °C [23]. Schmidt et al. [23] showed that the hot air recirculation decreases the rack load capacity. The complete separation of hot and cold aisles air circulation is difficult as only a fraction of the required rack airflow rate is provided by chilled air supply and the remaining fraction is supplied by ambient room air circulation. In addition to hot air-cold air mixing, it is pointed out by Ohadi et al. [49] that the bypass of cold air passing over the servers is an important source of exergy losses and inefficiencies in air-cooled systems.

Air flow rates are found to vary from data center to data center. For instance, the 0.5 MW raised floor data center referred by Schmidt et al. in the above mentioned study [23] housed 50 racks with a cooling flow rate per rack of ~1500 CFM and average 10.5 °C temperature increase from bottom to top of every rack. Rambo and Joshi [2] in their research on improving data center energy efficiency using variable speed server fans studied a data center with 28 extremely high power density racks with a maximum air flow rate for each rack of 2400 CFM, cold air temperature of 15 °C and a maximum temperature rise of 40 °C. McAllister et al. [25] used air flow rates per tile of 0.589 kg/s and 289 kg/s correspondingly for the high and low rack heat loads of 3583 W/m^2 and 717 W/m^2. Table 5 summarizes recent literature regarding hot and cold aisle temperatures as well as air flow rates in air-cooled data centers [2,5,11,13,16,23–25,36,37,40,49–52].

In CRAC units the heat from the hot air returning from the racks is rejected to the outdoor atmosphere typically using a chiller and cooling tower loop. The chiller temperature will vary for different data centers. However, in the report published by the Silicon Valley Leadership Group [53] temperature ranges of 10–13 °C and 15.5–18.4 °C were reported as standard for the supply and return water temperatures, respectively. The low temperature of the return water limits the effectiveness of capturing waste heat at the chiller water return. So although it might be logistically challenging, the optimum points to capture the heat in air-cooled data centers for maximum energy capture (~35–45 °C) will be directly at the rack exhaust prior to room air mixing and exergy losses. Alternately, it is easier to capture the waste heat at the return to CRAC (~30–40 °C), or at the chiller water return (~16–18 °C), however the lower temperatures available here limit usefulness. Considering the collated information on air cooled data center operation in Table 5, the authors of this paper have selected a range of operating conditions to represent a “typical” legacy air cooled data center characteristic as seen in Table 6. This data will be used in the thermodynamic analysis and technical assessment of waste heat recovery techniques.

2.3.2. *Water-cooled systems*

Many newer data center designs have power loadings to levels that are difficult to remove with CRACs alone. Therefore other cooling techniques, such as single phase forced liquid flow or phase change liquid systems, are now being adopted [23,38]. As pointed out by Greenberg et al. [54] using liquid cooling systems can lead to substantial savings in the total cooling energy requirement. In fact, through the use of liquid cooling at the server and/or rack level along with using water/air side economizers with ambient free-air cooling, the need for both CRAC units and chiller plants can be reduced or in some cases potentially eliminated, resulting in considerable energy savings.

Brunschwiler et al. [3] show that due to more efficient heat transfer, low temperature differences across the system (as low as ~10 °C from package to water supply) occur in water cooled systems. According to their research, a water flow-rate of 0.7 L/min with an input temperature of 60 °C is more than sufficient to keep the microprocessor junction temperature well below 85 °C and if the chip is allowed to approach 85 °C, the inlet water temperature can be as high as 75 °C [3].

These higher temperatures for liquid cooling systems will provide a higher quality waste heat and the energy capture is also easier, using the direct discharge liquid stream. The higher temperature coolant can eliminate the need for chillers, which in turns reduces the system energy requirement. In this method, cold plates are located as close as possible to the heat generating components [3]. The thermal resistance of these systems, which have heat removal capacities of more than 200 W/cm^2, is less than 20% of thermal resistance in air-cooling systems [3].

A comprehensive comparison study between the energy efficiency of air-cooled and water-cooled high density servers, conducted by Ellsworth and Iyengar [55], demonstrates the advantages of water-cooled over air-cooled systems. Their research shows that when using water cooling, processor performance can be increased by 33%. In another water cooling study by Ellsworth et al. [56], all energy dissipating components including the processor, memory, power conversion, and I/O electronics are cooled by aluminum/copper tube cold plates. In this case the cold plate supply temperature was set at least 7 °C above dew point to avoid condensation. Campbell and Tuma [19] studied the performance of a water cooling system for an electronic module including 150 W dual-core chip and 8 W memory chip. Their model showed that to maintain junction temperature at 65 °C with a water flow-rate of 0.95 L/min, the maximum temperature of water entering the module must not exceed 28 °C.

IBM recently designed and constructed a dual loop chiller-less data center, which reduced the cooling energy requirement from an industry standard of 45% of the data center total energy consumption to only 3.5% [20,57,58]. This test scale data center consists of a single rack filled with 38 “warm water” cooled servers. In each server, the processors and memory modules are cooled with cold plates, and the remaining components are cooled using re-circulated air which is pre-cooled by the water entering the servers. Since the majority of the server heat dissipation is removed through the water loop, the need for a CRAC unit is almost eliminated. The rack′s cooling water is circulated in an internal loop with heat exchange to an external coolant loop such as a water-glycol mixture. The transport of heat from the closed external loop to the ambient air is accomplished using a dry-cooler and requires no additional make-up water such as in the wet cooling tower approach. The rack power varied from 13.4 to 14.5 kW and the water flow rate varied from 4 to 8 GPM [20,57,58].

Sharma et al. [59] determined an optimum water inlet temperature in the range of 40–47.5 °C at an optimum flow rate of 1 L/min to cool a microprocessor chip as a tradeoff between maximized waste heat recovery and the chip thermal reliability.

This literature on water cooled data centers is shown in Table 7 [3,19,20,49,56–59] including inlet and outlet water temperatures, water flow rate, and temperature difference between the CPU temperature and the cooling water supply. Considering this collated information on water cooled data center operation, the authors of this paper have selected a range of operating conditions to represent “typical” water cooled data center characteristics as seen in Table 8. This data will be used in a technical assessment of waste heat recovery techniques for these data center applications.

2.3.3. *Two-phase cooled systems*

The need to find effective cooling solutions for devices which dissipate energy loads in excess of 1000 W/cm^2 [60] has been the driving motivation behind the implementation of two-phase cooling systems [61]. By taking the advantage of the dramatically increased convection heat transfer coefficients associated with nucleate boiling and through the use of micro-channel heat sinks, two-phase cooling has been shown to remove high heat fluxes ranging from 790 W/cm^2 [62] up to 27,000 W/cm^2 [63].

In enumerating the advantages of two-phase cooling over water cooling, Leonard and Philips [64] point out to the limitations

Table 6
Summary of “typical” air-cooled data center heat sources and streams.

Parameter	Value
Cold aisle (CRAC supply) temp.	10–32 °C
Hot aisle (CRAC return) temp.	50–60 °C
Temp. rise over server	10–20 °C
Airflow per rack	200–2500 CFM
Chiller water supply to CRAC	7–10 °C
Chilled water return from CRAC	35 °C

Table 8
Summary of “typical” water-cooled data center heat sources and streams.

Parameter	**Value**
Water supply to server	20–60 °C (std) 70–75 °C (max)
Water exit from server	2–5 °C temperature rise over servers
Water flow rate per rack	5–10 GPM
ΔT from water to lid	5–18 °C
Buffer heat exchanger flow rate	5–10 GPM
Buffer heat exchanger supply temp.	3–5 °C above ambient

Table 7
Water-cooled data centers heat sources and streams in recent literature.

Investigator (s)	Inlet water temp., °C	Outlet water temp./temp. rise, °C	Junction/lid temp., °C	Heat load (kW)	Water flow rate, GPM
Brunschwiler et al. (2009) [3]	60 (std.) 75 (max)	63/NA	85/NA	NA	~0.18
Ellsworth et al. (2011) [56]	15–24 (min:7 °C above dew pt.)	NA/NA	NA/NA	72/180	15/45
Campbell and Tuma (2012) [19]	28 or less	NA/NA	NA/NA	2.5	1
Ohadi et al. (2012) [49]	62	NA/NA	78/NA	0.085	0.04
Iyengar et al. (2012) [20]	20–45	NA/2–5	NA/36–74	13.4–14.4	5–10
David et al. (2012) [57]					
Parida et al. (2012) [58]					
Sharma et al. (2012) [59]	30–60	NA/2–5	85–90/NA	0.13	0.08–0.26

associated with water cooled systems such as the need for high pumping power and reliability issues. Hannemann et al. [65] show that multiphase systems work with 4 times smaller mass flow-rates, 10 times less pumping power and 2 times smaller facility size. In addition to higher heat transfer coefficients, two-phase cooling also provides more uniform temperature distributions [66–68] when compared to water-cooled systems with lower flowrates (and thus pumping powers). Simulations performed by Thome and Bruch [69] demonstrated that with the same pumping power, two-phase cooling can maintain the temperature of the operating chip 13 °C lower than water cooling. According to the same study, the choice between pump and compressor for to drive the coolant in a two-phase system is a compromise between higher heat rejection temperatures for vapor compression and lower power consumption for liquid pumping.

Zhou et al. [70] developed an optimized steady-state model for a two-phase cooling system for electronics heat dissipation. An integrated heater-accumulator was used to completely evaporate the coolant exiting the system evaporators. Their model indicated that higher critical heat fluxes were achievable when the accumulator heat supply was increased. It also showed that there was a trade-off between the COP of the system and the accumulator heat flux to prevent device thermal failure.

Split flow configurations with one inlet and two outlets were shown to provide higher critical heat fluxes (CHF) and lower pressure drops than the standard single inlet–single outlet configuration [69,70]. Additionally, Mauro et al. [71] examined three different coolants (HFC134a, HFC236fa, and HFC245fa) and it was found that for all working fluids, CHF increases 80% when the single inlet–single outlet configuration is replaced by split flow.

Nevertheless, the success of two-phase cooling in operation without the need for a chiller requires the application of cold plates and thermal interface materials (TIMs) with very low thermal resistance. The simulations performed by Ohadi et al. [49] indicated that when using HFC245fa with an entering temperature of 76.5 °C, it is quite possible to maintain the junction temperature below 78 °C and adequately cool an 85 W processor. Therefore, it can be inferred that the quality of waste heat extracted through two-phase cooling can be considerably higher than two previously discussed cooling schemes and that there exist more opportunities for waste heat utilization with two-phase cooling.

Marcinichen et al. [18] presented a hybrid cooling cycle which was characterized by high reliability, controllability, cycle interchangeability and flexibility in heat recovery. The hybrid cycle consisted of three different two-phase flow micro-evaporator cooling cycles. The micro-evaporators were designed to remove the heat dissipated by the microprocessors. The overall design had to have such a controllability to maintain the micro-evaporator outlet vapor quality well below that of the critical vapor quality (associated with the critical heat flux), and the other heat generating components, such as the memory were cooled by the refrigerant exiting from micro-evaporators. The difference between the three cycles was the choice of prime mover. One featured a liquid pump, one featured vapor compression, and the third was a hybrid system which allowed switching between either of the first two systems.

It was found that for an evaporation temperature of 60 °C and a micro-evaporator outlet vapor quality of 30%, the predicted critical heat flux was about 2.2 times the maximum heat flux of the examined blade server. However, by increasing the vapor quality to 50% the ratio of predicted critical heat flux to the actual maximum heat flux was decreased to 1.3. The high quality absorbed heat which was then dissipated by the refrigerant in a condenser, could then be recovered in the form of space heating. This vapor compression system was characterized by high condensing temperature, more controllability, and a medium overall efficiency (compared to the liquid pumped system). The high condensing temperature made the vapor compression system a better option for waste heat recovery especially in the form of district heating during the winter season.

The corresponding liquid pumped system was characterized by low initial and maintenance costs, low vapor quality at the micro-evaporators outlet, high overall efficiency, and low condensing temperature, the low condensing temperature meant less opportunity for waste heat recovery which made the pumped liquid system a better option for summer operation. Thus a hybrid system was developed with interchangeability between liquid pumped and vapor compression cycles using shut off valves. The decision on whether to operate the liquid pumped or vapor compression cycle was dependent upon the priority of waste heat recovery.

In the continuation of their study, Marcinichen et al. [12] developed a steady-state code to design and simulate the performance of the two-phase cooling cycle. Five simulation cases were considered in which different refrigerants and cooling cycles were evaluated. The first case was a two-phase liquid pumped with HFC134a as coolant, the second case was the same as the first case but with smaller piping diameter, the third case also was the same as the first case but with different coolant (HFO1234ze), the fourth case was a single-phase cooling with liquid water, and the fifth case was the same as the first case but with vapor compression instead of a liquid pump. The analysis performed in the simulations included thermal performance, pumping power and total pressure drop. It was found that while the chip and water axial temperature rise is 2 °C for single-phase liquid water, for the two-phase cooling cycles the maximum axial temperature rise was about 1/10 of this, or 0.2 °C. The simulations also showed the mass flow rates required in the single-phase liquid water cooling is ~5 times greater than the required coolant mass flow rates in the two-phase cooling cycles and using a split flow design resulted in the lowest axial pressure drops over the processor for all the considered cases. Finally, the simulation of two-phase vapor compression system showed that although the power consumption is about 100 times more than the equivalent two-phase system with a liquid pump, there exists much higher potential for waste energy recovery. The outlet temperature of the secondary fluid in two-phase cycle is 90 °C compared to 48 °C in the liquid pumped two-phase system.

Table 9 [12,15,18,69,72] presents a summary of the literature for two phase cooling systems including the evaporator temperature, condenser inlet and outlet coolant and secondary working fluid temperatures, type of coolant, COP, coolant mass flow rate, pressure ratio and the power load.

The thermal characteristics of the hybrid system as presented by Marcinichen et al. [12] has exceptional potential for the recovery and reuse of waste heat, not just in the district or space heating, but also in other waste heat recovery systems. This system [12] will thus be used as a starting point for the integration of waste heat recovery options with two-phase systems. Tables 10 and 11 show the specific details of this system [12] for liquid pumped and vapor compression cycles. For both liquid-pumped and vapor compression systems, the optimum location for capturing the waste heat is at the condenser through the secondary working fluid (which is usually water) used to cool the primary coolant of the system.

3. Waste heat recovery technologies

In order to capture and reuse the low-temperature energy produced by server heat dissipation in data centers, eight common waste heat recovery techniques are analyzed for suitability for low temperature, high volume waste heat production. These include plant or district heating/hot water production, power plant co-location, absorption refrigeration, organic Rankine cycles, piezoelectrics, thermoelectrics, biomass co-location, and desalination/clean water production.

Table 9
Two-phase cooled data center heat sources and streams in recent literature.

Investigator	Refrigerant	Evaporator temp., °C	Junction – fluid temp. delta, °C	Condenser temp., °C	Condender secondary working fluid temp., °C		COP	Total power (W)	Pressure ratio	Refrigerant mass flow rate (kg/h)	Cooling type
					Inlet	Outlet					
Mongia et al. (2006) [72]	HFC600a	50 60	NA	90 80	NA NA	NA NA	2.25 3.7	50 50	2.4	0.936	Two-phase vapor compression
Trutassanawin et al. (2006) [15]	HFC134a	10–20	NA	40–60	NA	NA	1.9–3.2	121–268	1.9–3.2	2.88–6.12	Two-phase vapor compression
Thome and Bruch (2008) [69]	HFC134a	55–65	15	NA	NA	NA	NA	108 and 324	NA	NA	Two-phase vapor compression
Marcinichen et al. (2010) [18]	HFC245fa	60	10–15	90	NA	NA	8	146.25 to 162.5	2.19	10.82–11.90	Two-phase vapor compression
	HFC236fa	60	10–15	90	NA	NA	8.3		2.43		
	HFC134a	60	10–15	90	Water: 30	Water: 80	7		1.99		
	HFC600a	60	10–15	90	NA	NA	8.4		1.99		
	HFC134a	60	10–15	59.9	Water: 30	Water: 49.9	NA	162.5	NA	NA	Two-phase liquid pumped
Marcinichen et al. (2012) [12]	HFC134a	60	10–15	57.7–58	Water: 30	Water: 47.7–48	NA	60 per CPU/55.6 auxiliary per half server	NA	5.18 per microevaporator	Two-phase liquid pumped
	HFO1234ze	60	10–15	56.8	Water: 30	Water: 46.8	NA		NA	5.32 per microevaporator	
	HFC134a	60	10–15	95	Water: 30	Water: 90	NA		NA	5.18 per microevaporator	Two-phase vapor liquid
	Water	Inlet temperature: 60 °C, Outlet temperature: 62 °C							NA	25.81 per microprocessor	Single-phase liquid

Table 10
Summary of “typical” two-phase cooled data center heat sources and streams with liquid pump.

Parameter	Value
Coolant supply to evaporator	60 °C saturated liquid (std.) 70–75 °C (max)
Coolant exit from evaporator	62 °C at 30% quality (std.) 75–80 °C (max)
Condenser cooling fluid inlet	30 °C
Condenser cooling fluid outlet	45–90 °C

Table 11
Summary of “typical” two-phase cooled data center heat sources and streams.

Parameter	Value
Coolant supply to evaporator	60 °C saturated liquid (std.) 70–75 °C (max)
Coolant exit from evaporator	62 °C at 30% quality (std.) 75–80 °C (max)
Coolant temperature at the exit of vapor compressor	~90 °C
Condenser cooling fluid inlet	30 °C
Condenser cooling fluid outlet	~90 °C

Each of these techniques is discussed in detail considering the needs of the technology and the waste heat production of data centers of the types discussed in Section 2.

3.1. *Plant or district heating/hot water production*

One common and relatively simple application for the reuse of low quality energy is in HVAC or hot water production systems. The temperature range of the captured waste heat from air cooled servers (35–45 °C) is more than sufficient for reuse in heating needs. Depending on the size and operating conditions of the data center, the space heating provided by data centers can range from the HVAC needs of the data center itself to that of a single family home, an apartment complex, a neighborhood or even several neighborhoods (district heating). The use of remotely siting servers in individual homes to provide domestic heating is referred as a “Data Furnace” by Liu et al. [73]. They studied the opportunities and challenges associated with replacing large centralized data centers with distributed small scale “micro-data centers” which could then be used as a primary source for domestic heating in homes, office buildings and apartment complexes. According to the Energy Information Agency [74], 6% of total US energy usage goes to home heating, therefore re-using data center waste heat to directly provide domestic heating can provide considerable energy savings during the winter months. An estimated saving between $280 and $325 per server per year was calculated when using Data Furnaces in the US residential sector [73]. Possible advantages of Data Furnaces include a reduced total cost of ownership for data center operators (considering costs associated with land, construction, and air conditioning utilities) and closer proximity to users, while disadvantages include higher prices for electricity (electricity price for residential areas is 10–50% higher than that for industrial areas), bandwidth, and maintenance issues (because of the geographical distribution of the servers).

District heating is another common low-quality waste heat recovery method which is economically viable and ecologically beneficial. District heating is more suitable for the higher waste heat capture temperatures of liquid cooled servers, but does not require the remote siting of servers. The use of liquid cooling allows the capture of a slightly higher quality waste heat (up to 50–60 °C rather than 35–45 °C), which can then be used to distribute heating over a wider area. This heat can thus provide a revenue stream for the data center operator. District heating is more common in Europe than in the US, particularly in Nordic countries that suffer from cold winters that can feature temperatures of −20 °C and below [75–77]. In 2009, more than 5000 District Heating units in EU countries provided 9.7% of the EU residential heat demand [3]. Although most existing systems are powered conventionally, not using waste heat, a 2 MW data center installed in an empty WWII shelter in Helsinki provides enough water heated by waste heat to heat 500 homes or 1000 apartments [75].

The heat dissipated by data centers can also be used in preheating on-site domestic hot water (DHW) which is then upgraded using either a heat pump or heat recovery chiller. The required waste supply temperature when a heat pump is used is in the range of 22–92 °C [78] and in the case of using heat recovery chiller, is 45 °C. The benefits of using waste heat to preheat on-site DHW are energy savings and emission reductions though reduced use of fossil fuels.

One possible challenge associated with utilizing waste heat to preheating on-site DHW is the more complex piping arrangement. Intel [78] developed a heat recovery chiller to capture the waste heat in a legacy data center. The heat recovery chiller consists of an evaporator and condenser. Data center waste heat is captured through a plate heat exchanger which then transfers the heat to an evaporator. A heat pump system is then employed to transfer the heat from the lower temperature water flowing in the evaporator to the higher temperature water flowing in the condenser.

The integration of district heating/hot water production into any operating data center will require design and analyses specific to the particular data center thermal management system and operating conditions. In legacy air cooled data centers for instance, the optimum location to install a heat exchanger to extract waste heat and repurpose it in district heating or hot water production is at the return to CRAC unit or at the chiller water return. The temperature of the captured heat for this design will be as low as 35 °C which is appropriate for hot water preheating or space heating. If air side economization is used in cold weather sites, the waste heat can also be used to increase the temperature of the outside air entering the data center, preventing freezing or moisture condensation. However, if higher quality heat is required, such as for use in district heating, the temperature of the waste heat from an air cooled data center will need to be boosted using heat pumps.

In water-cooled data centers, the easiest location to capture waste heat is at the cold plate loop heat exchanger. The temperature of heat extracted from water-cooled data centers can be in the range of 60–70 °C which is quite appropriate for both district heating and hot water production. Two-phase cooling systems produce waste heat as high as 70–80 °C which is more than sufficient for any heating or hot water applications.

In summarizing the use of data center waste heat for HVAC and hot water systems, the following issues are pointed out:

- The reuse of waste heat for space and water heating requires minimum supply temperatures that vary based on the application and range from as low as 35 °C when directly using hot air for domestic heating to as high as 60–70 °C for hot water systems and 70 °C or higher for district heating.
- The benefits of using data center waste heat in space and water heating include a reduction of operating costs for on-site heating/hot water needs and potential income from selling heat to an individual user or the district grid.
- The challenges of using servers or data center waste heat in space and water heating include greater system complexity, and possible maintenance and security issues when remotely sitting servers. District heating is not generally culturally accepted for US sites but is well established in Europe.

3.2. Power plant co-location

The use of waste heat to provide heating of water in the thermal Rankine cycle of a major power plant is a common waste heat recovery technique [79]. The low-quality heat available from data centers precludes the possibility of entirely replacing the boiler, but the waste heat can be used to preheat boiler feed-water and contribute to an overall reduction in fossil fuel consumption and in pollution emission.

This waste heat recovery technique will be most beneficial when used with on-chip two-phase cooled data centers and their higher recovery temperatures, but is still possible with standard liquid cooled systems.

In a recent research study conducted by Marcinichen et al. [12], the environmental and economic benefits of recycling data center waste heat in preheating the water in a coal fired power plant were studied. The system is modeled considering the data center as a cooling cycle and the power utility as a thermal Rankine cycle. In this configuration, the data center waste heat is added into the cycle after the condenser and before the feed-water heater (see Fig. 1). The data center considered in the referred study contains 100,000 servers with each server dissipating 325 W heat. The power plant cycle is a 175 MW coal fired plant. A evaporation temperature of 60 °C is set in the microevaporators at the chip while the condensing temperature is assumed to be either 60 °C or 90 °C for the two-phase liquid pumped and two-phase vapor compression cycles, respectively.

A counter-flow heat exchanger is used as the condenser through which the data center waste heat is transferred to the water of power plant Rankine cycle. The simulations indicate that for such a system, utilizing data center waste heat to preheat the boiler feed-water in a coal fired power plant can improve the power plant efficiency by up to 2.2% under certain optimized conditions [12]. This degree of efficiency improvement can lead to huge savings in both fuel costs and in carbon tax due to the reduction in carbon footprint. Considering a value of $90/ton for coal and a carbon tax of $30/ton for CO_2, the potential savings from both capturing heat from the data center and selling it to the power plant and through reduction in carbon taxes are estimated. Table 12 shows the economics corresponding to both liquid pumped and vapor compression systems [12].

The temperature of heat extracted from water-cooled data centers can be in the range of 60–70 °C which is acceptable for boiler feedwater preheating although the 70–80 °C waste heat from two-phase cooling systems will lead to greater economic and environmental benefit. In summarizing the use of data center waste heat for preheating boiler feed-water in a power plant (power plant co-location), the following issues are pointed out:

- Power plant boiler feed-water preheating is most efficient with higher quality waste heat than is standardly available from data centers. Minimum demands are typically 60–100 °C which is at the upper end of available waste heat from water cooled data centers, and in- line with advanced two-phase flow cooling systems.
- The benefits of using data center waste heat in power plant boiler feed-water preheating include potential income from both selling heat to the power plant and from possible carbon offsets for the data center operator.
- The challenges of using data center waste heat in power plant boiler feed-water preheating are that this technology is not suitable for air cooled data centers due to low quality heat and cannot be retrofit into existing liquid cooled data centers due to the specific siting needs of the technology. As the data center waste heat will degrade significantly with distance, the data center will need to be sited directly at the power plant location

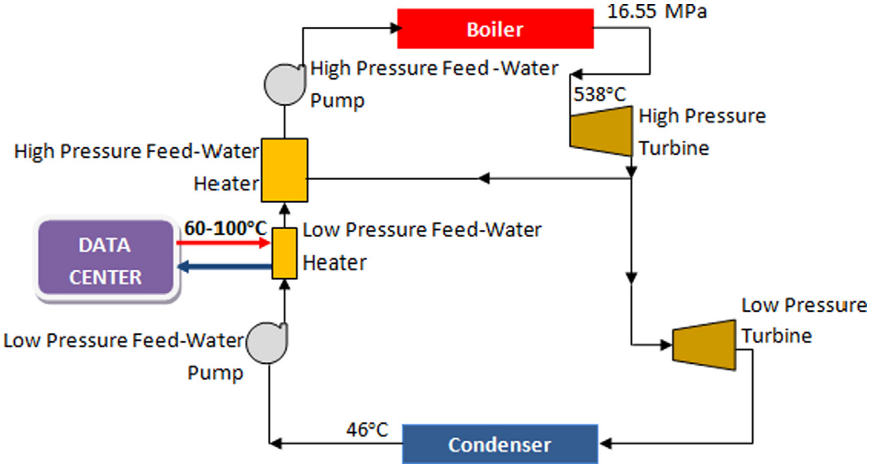


Fig. 1. A schematic of a system in which the data center waste heat is utilized in a coal power plant Rankine cycle.

Table 12
Potential savings when the heat dissipated by a 32.5 MW data center with a two phase cooling system is used to preheat boiler feed-water in a 175 MW coal power plant [12].

Data center cooling system	Annual savings in data center	Annual savigns in power plant	Annual total savings
Hybrid cooling system with liquid pumped (condensing temperature of 60 °C)	\$45,000,000	\$1,000,000	\$46,000,000
Hybrid cooling system with vapor compressed (condensing temperature of 90 °C)	\$40,000,000	\$4,500,000	\$43,000,000

3.3. *Absorption cooling*

The use of waste heat in an absorption cooling cycle is an intriguing possibility for systems that require substantial cooling, such as data centers. In a standard vapor compression refrigeration cycle, such as those conventionally used in data center CRACs, a considerable amount of power is consumed by the compressor. Since liquids have much lower specific volumes than vapors, replacing the vapor compression system with an absorption refrigeration system, in which a liquid solution of absorbent fluid and refrigerant is used, can result in considerable power savings.

Excellent comprehensive reviews are available for both absorption refrigeration technology and combined heat and power cycles (CHP) which provide both power and chilled water by using the waste heat from the power cycle in an absorption refrigeration cycle [80–82]. As such, this discussion will focus on the design aspects which affect their integration with data center waste heat recovery.

A typical absorption cycle, as shown in Fig. 2, replaces the compressor with an absorbent and a generator in absorption cycle [80–82]. The low pressure vapor refrigerant enters an absorber (state point 1) through an exothermic process where it is absorbed into a carrier liquid, which is called the absorbent. Then, the strong liquid solution of refrigerant and absorbent exiting the absorber (state point 5) is pumped to a higher pressure and enters to the solution heat exchanger (state point 6). Recall that energy costs for pumping a liquid are relatively small. By passing through the solution heat exchanger, the strong solution is preheated and then enters the generator (state point 7). Through an endothermic process, the refrigerant is separated from the absorbent. The exiting refrigerant vapor (state point 2) continues on to the condenser while the now weak liquid solution in the generator passes through an expansion valve and returns to the absorber (state points 8–10). The loop of the absorber, pump, generator and expansion valve are together referred to as a chemical compressor.

The application of absorption cooling to replace or supplement a CRAC not only reduces the power necessary to run the system, but also has the benefit of using the data center waste heat as the absorption generator heat source. Fig. 2 shows an absorption refrigeration system driven by the data center waste heat.

Haywood et al. [83,84] studied the used of data center waste heat to drive a 10-ton single-effect lithium bromide-water absorption refrigeration unit. This system transferred thermal energy from the highest power components on a server blade to the generator of an absorption cooling unit to drive the absorption unit. Excess energy was transferred to thermal storage which was used as necessary to boost the extracted waste heat when operating at non-peak conditions.

The result of this approach [83,84] was the reduction of the CRAC cooling load in two ways: through capturing and using part of the waste heat that otherwise should be cooled by the CRAC unit and through cooling produced by the absorption chiller which resulted in further reduction in the CRAC load. The 10-ton single-effect Li–Br absorption system was able to produce 35.2 kW of cooling at its design point, where the quality and quantity of the heat input were 88 °C and 50.2 kW, respectively. The corresponding COP at the design point was 0.7. However, the unit could also be powered at lower heat inputs and temperatures. The required temperature at the generator of the absorption unit was between 70 °C and 95 °C, so this system can be used in any data center cooled by liquid or two-phase flow.

The minimum required heat input and temperature for these systems were 14.1 kW and 70 °C [83,84]. Under these minimum conditions a cooling capacity of 8.8 kW (COP=0.62) is produced. The COP of the unit reaches a maximum value of 0.86 with a generator temperature of 80 °C and a resultant cooling capacity of 25 kW [83,84]. While system requirements can vary, performance will increase with waste heat temperature and so this process is generally not economically viable with the lower waste heat quality available in air cooled data centers.

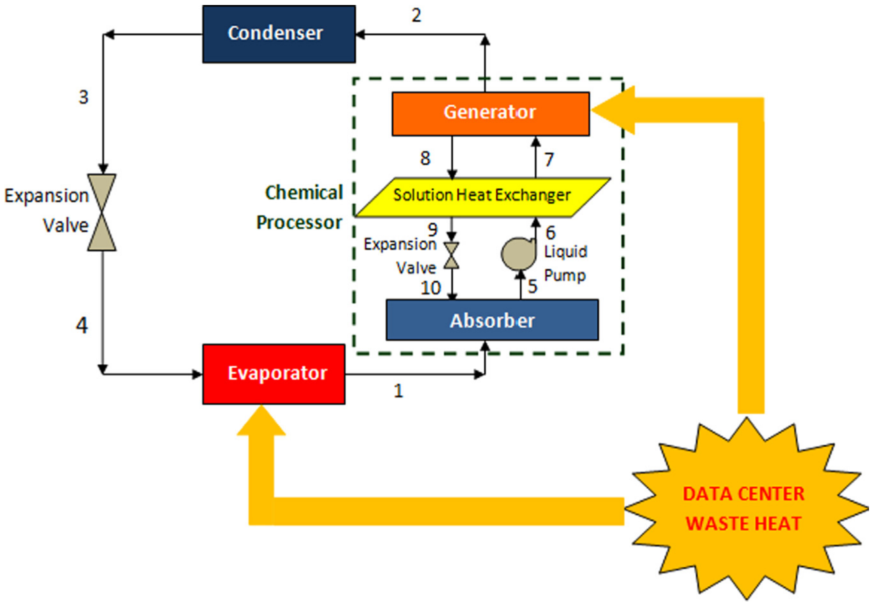


Fig. 2. A schematic of a simple absorption refrigeration system driven by data center waste heat.

The main challenge with this system [83,84] was capturing enough high-temperature heat from the two processors on each server blade, and then transporting that heat effectively and efficiently to power the absorption chiller. Two approaches were applied to capture the high quality heat and transfer it to either absorption chiller or thermal storage. One approach was the use of a heat pipe with a boiler and condenser through which 90% of the processors waste heat could be captured and utilized [83]. The other approach was exploiting a liquid cooling system through which 85% of the processors waste heat could be recovered [84]. Each server blade dissipates 294 W, 70% of which comes from the processors. Capturing 85% of the processor waste energy as heat results in 174 W per blade. Taking the most conservative case in which the chiller must be energized with a minimum of 14.1 kW at 70 °C, then 81 server blades are required to power the unit. For a higher efficiency scenario more cooling is produced and more server blades can be added to the system. For example, with a COP=0.86 at a generator temperature of 80 °C, 25 kW of cooling are produced with 144 server blades.

In addition to this focus on macroscale absorption refrigeration, a number of research studies have focused on the design of micro-scale absorption systems to directly cool a single microprocessor using the energy dissipated by the same microprocessor. Suman et al. [85] applied micro-scale absorption refrigeration to cool high performance electronics due to its compactness, easy scalability to varying cooling load, and relatively higher COP in comparison with other technologies. However, of course a major advantage of absorption refrigeration at either the macro or micro-scale is the utilization of low-quality waste heat to charge the generator [86]. Accordingly, Kim et al. [87] designed and analyzed a Li–Br Water absorption system using a dual micro-channel array evaporator. A hydrophobic membrane was used to place the generator and condenser micro-channel arrays in an intimate communication with each other. The membrane acted as a common interface between the generator and the condenser to separate the water vapor from the Li–Br solution. Depending on the microprocessor heating load, the evaporator temperature could vary between 25 °C (corresponded to 100 W heat load) and 50 °C (corresponded to 1000 W heat load), while the temperature of the generator was fixed at 90 °C. The COP of the system for a heat load of 100 W was 0.74.

Kim et al. [86] replaced the Li–Br solution in their miniature absorption design with ionic liquid (IL) to lower the toxicity of the system and to eliminate the risk of crystallization which occurs with Li–Br/Water absorption systems. They conducted a theoretical analysis to explore the effect of using different mixtures of refrigerants and imidazolium-based ILs as the working fluid pairs on the performance of the absorption system. The saturation temperatures at the evaporator and condenser were set at 25 °C and 50 °C, respectively, with a power dissipation of 100 W. The highest COP of 0.9 was achieved with a combination of water as the refrigerant and 1-ethyl-3-methylimidazolium tetra fluoroborate as the absorbent fluid at a generator temperature of 70 °C.

Kim et al. [88] studied miniaturized waste-heat driven absorption refrigeration/heat pump systems for electronics cooling in which IL and five HFC refrigerants were evaluated as the system working fluid pairs. The effects of generator and absorber temperatures, waste-heat quality, and system design on the heat pump performance were investigated. The study showed that for an evaporator temperature of 41 °C, the maximum COP is 0.35 with an evaporator cooling capability of 36 W and a generator temperature range of 50–110 °C. The addition of a solution heat exchanger increased the COP to 0.8.

Chiriac and Chiriac [89,90] designed and developed a miniaturized ammonia-water absorption system for smaller scale power electronics. However, the absorber was powered by an electric heater and not by waste heat. In the original system [89] a gravity-based thermo-syphon was used to circulate the binary solution, however, through the thermosyphon was replaced with a mini-pump in the next design [90]. This improvement increased the COP of the system from 0.4 to 0.7. The system was designed for heat removal capacity of 100 W and the temperatures at the generator, condenser and evaporator were set at 125 °C, 40 °C and 10 °C, respectively.

In an alternate implementation of absorption refrigeration for data centers, the application of combined heat and power (CHP) has gained traction in recent years [81]. Combined heat and power is proving to be an effective solution to reduce energy footprint, control energy costs, and improve the reliability of the power supplies for data centers. In this approach on-site power generation equipment powers the data center [81,21] and the waste heat from the power generation is used by absorption system to chill cold water which can then be used in server cooling [81]. Therefore through the use of CHP, the cooling is provided without any additional expenditure of primary energy [21]. Little and Garimella [21] developed a CHP design in which the data center waste heat was boosted to a higher temperature and also used in the absorption system along with the power generation waste heat. Alternately, the upgraded thermal stream can be sold for district heating applications in neighboring residential or commercial buildings, or as process heat for commercial end uses such as laundries, hospitals, and restaurants, depending on the location of the data center.

In summarizing the use of data center waste heat in absorption refrigeration the following issues are pointed out:

- Absorption refrigeration systems can operate with generator temperatures of 70–90 °C which is consistent with available waste heat from water-cooled and two-phase cooled data centers. This technology is not suitable for waste heat from air cooled data centers without an additional heat booster.
- The benefit of using data center waste heat in absorption refrigeration systems is a direct reduction of the load on data center CRAC systems, through the generation of chilled water for cooling, which is of significant economic benefit to most data center operators. Additionally, there are no issues with site co-location and the technology can be retrofit into existing data centers if space permits.
- The challenges of using data center waste heat in absorption refrigeration systems are that this technology is typically not suitable for air cooled data centers and there may be space issues when retrofitting systems to existing data centers.

3.4. *Organic Rankine cycle*

Waste heat from data centers can be used to directly produce electricity through an organic Rankine cycle (ORC). ORCs work on the same basis as the steam Rankine cycle, but using organic fluids with substantially lower boiling points as the working fluid. The lower boiling points make it possible for data center waste heat to serve as the heat source. Excellent comprehensive reviews of ORCs are available [91–95] and as such, this discussion will focus on the design aspects which affect their integration with waste heat recovery from data centers.

The thermodynamic properties of the working fluid strongly influence cycle efficiency and ORCs can operate successfully with waste heat streams of 65 °C and higher [91,94] and can even run as low as 32 °C with reduced efficiency [95]. The most appropriate temperature range for a specific cycle depends on the type of organic working fluid. Common organic fluids include R-134a, Benzene, Toluene and Propane with operating temperature ranges as wide as 65–350 °C [91]. Operation at the lower end of the temperature range, as will be necessary with data center waste heat, will lead to lower operating efficiencies. Generally, overall cycle efficiency will be between 5% and 20% [94,95]. This is mainly due to the fact that low-temperature cycles are inherently less efficient than high-temperature cycles. But considering the theoretical Carnot efficiency for the same temperature range, it is clear that actual ORC efficiency is a substantial percentage of theoretical and that ORCs should be considered as a promising data center waste heat technology.

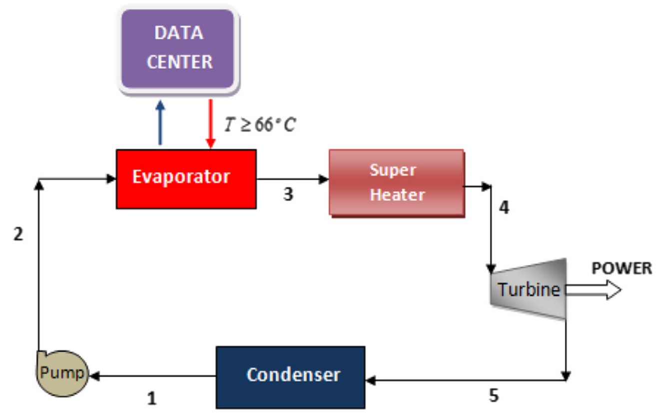


Fig. 3. Schematic diagram of organic Rankine cycle. The super heater is necessary only when the working fluid is wet.

Fig. 3 depicts an ORC consisting of an expansion turbine, a condenser, a pump, an evaporator and a super-heater. A super-heater may or may not be needed in a specific ORC cycle depending on the operating temperature and fluid [91]. The slope of the saturation vapor curve for the specific fluid is a key characteristic affecting the ORC cycle efficiency. Depending on the slope of vapor saturation curve in the temperature–entropy (T–s) diagram, organic fluids are categorized into three types: wet, dry, and isentropic. As shown in Fig. 4, a dry fluid has a positive slope ($dT/ds > 0$, e.g. isopropane), a wet fluid has a negative slope ($dT/ds < 0$, e.g. R22), and an isentropic fluid has nearly infinitely large slope ($ds/dT = 0$, e.g. R11) [91,92,96]. Wet fluids usually require superheating, while dry and isentropic fluids do not. Working fluids with high densities and high latent heats are found to provide the highest high turbine work output and isentropic and dry fluids are generally recommended for ORCs [91].

In Fig. 4 it can be seen that the saturated vapor phase of a dry fluid is slightly superheated after isentropic expansion. However, for an isentropic fluid with an almost vertical vapor saturation curve, the vapor expands along a vertical line on the T–s diagram such that the saturated vapor at the turbine inlet will remain saturated throughout the turbine exhaust without condensation [92]. This ability to remain saturated throughout the expansion process without need for a regenerator makes isentropic fluids particularly well suited for ORCs [92]. Although the thermodynamic and physical properties of the working fluid have the greatest impact on system efficiency, the working fluid stability, environmental impact, safety, compatibility, availability and cost must also be considered.

Much of the research focused on ORCs has studied the effect of working fluid type and operating conditions on ORC efficiency. Hung et al. [92] conducted an analytical study to understand the effect of turbine inlet temperature on the efficiency of an ORC for different types of organic working fluids. Saleh et al. [96] studied ORC efficiency for 31 different pure component organic working fluids. A high efficiency of 13% was achieved for dry working fluids with high boiling points. Further analysis showed that heat transfer from a 120 °C heat source to the working fluid is highest for supercritical working fluids. However, Chen et al. [91] reviewed ORCs and supercritical Rankine cycles (SRCs) for low-grade heat conversion into power and found that while SRCs do exhibit a better thermal match with their heat sources that they require much higher operating pressures, making them undesirable.

During the past three decades, ORC has been progressively adopted as a premier technology to convert low grade waste heat into power [94,95]. The market for ORCs has shown exponential growth since the 1980s when the first commercial ORC plants were installed [97]. ORCs have been used for waste heat conversion into power for many different applications including solar, geothermal, biomass, surface seawater, and other industrial processes [95]. This versatility makes ORCs particularly well suited for waste heat recovery and reuse for data center applications. The technology is generally well proven, and matches well with the waste heat recovery temperatures for both liquid cooled and two-phase cooled systems. With the use of a secondary source of heat for boosting, it can even be used with air cooled data centers. There are no specific siting needs, so ORC can be retrofit into existing plants, proving on-site electricity production from waste heat.

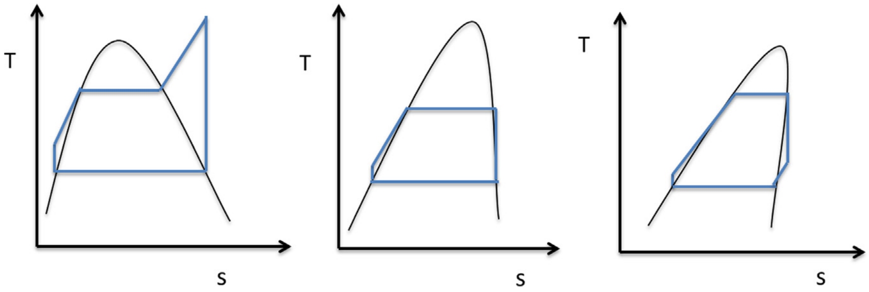


Fig. 4. *T–s* diagrams for (a) wet, (b) isentropic and (c) dry organic fluids.

Attention is now also being focused on the development of microscale ORC. Zhang and Wang [98] developed micro-scale ORC as a promising solution for waste heat capture from High-Concentration Photo-Voltaic (HCPV) units in which more than 60% of the input solar energy is lost to waste heat. In this system, the working fluid is pumped into a micro-channel heat sink above each multi-junction photo-voltaic cell. The analyses shows an 8.8% increase in the power generation efficiency when the HPCV's liquid cooling system is replaced with micro-scale ORC. This idea can be extended for data center cooling with micro-channel coolers in the server racks.

In summarizing the use of data center waste heat in organic Rankine cycle power production, the following issues are pointed out:

- Organic Rankine cycle power production requires waste heat temperatures of 65 °C and higher. These temperatures are available in the waste heat streams from water cooled or two-phase cooled data centers. With a secondary heat source, this technology is suitable for air cooled data centers.
- The benefits of using data center waste heat in organic Rankine cycle power production include the on-site electricity production from waste heat and no specific siting needs, making it appropriate for a wide range of data center designs.
- The challenges of using data center waste heat in organic Rankine cycle power production are relatively low overall system efficiencies (5–20%) associated with the low operating temperatures.

3.5. *Piezoelectrics*

Piezoelectric generation is a method to directly convert turbulent oscillations in the data center cooling air flow directly into electricity. Piezoelectrics are typically fabricated from crystalline materials with a linear electromechanical interaction between the mechanical and electrical states called the piezoelectric effect. When a piezoelectric material is subjected to a mechanical stress and/or strain, a change in its internal electric field occurs and an electric charge is produced.

Piezoelectric power generation can be considered as a heat recovery technique for converting low temperature (100–150 °C) waste heat to electricity [79]. The input mechanical energy for a typical piezoelectric device is in the form of ambient vibrations and/or oscillatory gas expansion. For data center applications, these small devices could be located in regions where the CRAC air supply creates turbulent eddies to allow the direct conversion of kinetic energy into electrical energy. The created power will be small (mW), but could potentially be used for localized energy needs such as small fans or LED lighting. Even small loads like this can add up in terms of costs savings for data center operators when the elimination of the associated wiring and support structure are considered.

However, the major technical challenges associated with piezoelectrics include very low efficiency (currently ~1%), difficulties obtaining high enough oscillatory frequencies, high internal impedance, long term reliability and durability issues, and most significantly high costs ($10,000/W) [79].

In summarizing the use of data center cooling air flowrates to power piezoelectrics for electricity generation, the following issues are pointed out:

- Piezoelectrics work by capturing oscillating frequencies induced by turbulence in the CRAC air flow. Thus performance is independent of waste heat temperature and this method is only suitable for legacy air-cooled data centers.
- The benefits associated with piezoelectric technology include direct electricity generation for small localized power needs.
- The challenges associated with piezoelectric technology include low levels of power output (in mW), low conversion efficiency, and high costs.

3.6. *Thermoelectrics*

Thermoelectric generation is a method to directly convert data center waste heat into electricity. Thermoelectric modules (TECs) operate based on a phenomenon known as the Seebek effect [99]. According to the Seebeck effect, when two different materials with different conduction energy band levels (such as semiconductors) are subjected to a temperature difference, a voltage is created. Conversely, when a voltage is applied, a temperature difference is created. Thermoelectric modules are most commonly used in the operational mode where an applied voltage generates a controlled temperature difference. This allows thermal control in challenging environments such as spaceflight, deep oil well drilling and engine compartments. However, for waste heat energy capture, they can be exposed to a high temperature difference to create a voltage. In this case the "hot side" is coupled to a waste heat source. The energy absorbed at this junction provides enough additional energy to bridge the energy gap between the N and P type semiconductor pellets and initiate electrical flow. In this way thermal energy is directly converted into electrical energy [99].

A standard single-stage thermoelectric module can work with temperature differentials of up to 70 °C with hot side temperatures of 80–175 °C [99]. This makes their usage for data center waste heat capture limited to energy capture right at the chip with close coupling of the device to the chip, or for liquid and/or two-phase flow systems with their higher waste heat temperatures.

Due to the low efficiency (between 2% and 5%) and high cost of thermoelectric modules, their application is not currently widespread. However, recent advances in materials are driving up the conversion efficiencies, defined as electrical energy produced over waste heat energy absorbed, to 15–20% [79,99]. In addition to conversion efficiency, thermoelectrics can be rated by their ZT value, a dimensionless figure of merit. Values for ZT typically range from less than 1 to just over 2. ZT values can be increased by using materials with a high Seebeck coefficient, high electrical conductivity and low thermal conductivity. The challenge is that typical materials exhibit linked electrical and thermal conductivity: either both high or both low. Currently any ZT value above 1 is considered to be good, but in order to be competitive with other energy capture techniques, ZT values will have to reach 3–4. However, progress in thin film nanobased thermoelectrics is showing promise for the future [99,100].

Another challenge in the use of thermoelectrics for data center waste heat conversion is their high cost. While thermoelectrics have the ability to directly generate electricity from waste heat, as for piezoelectrics, the produced amounts of electricity are low while costs are high. In order to economically compete with simply purchasing the required electricity, the thermoelectric packages will need to reach a value of $5/watt instead of the current value of ∼$30/watt [79].

In summarizing the use of data center waste heat to power thermoelectrics for electricity generation, the following issues are pointed out:

- Thermoelectrics are most efficient with higher quality waste heat than is standardly available from data centers. Minimum demands are typically 80–175 °C which will be available in advanced two-phase flow cooling systems
- The benefits of using data center waste heat in thermoelectric generation include small amounts of electricity generation, suitable for small localized power loads.
- The challenges of using data center waste heat in thermoelectric generation are that this technology is not suitable for air cooled data centers due to inadequate temperatures and that the thermoelectric modules currently feature low conversion efficiencies and high costs.

3.7. Desalination/clean water production

Data center waste heat can be captured and used in multiple effect distillation (MED) for clean water production from seawater. Fig. 5 shows a typical configuration for a conventional MED system using heat from a steam boiler system. MED systems were initially designed as an extension of single stage distillation systems in order to recover and use waste heat released during water vapor condensation and excellent reviews on this topic already exist [101,102]. Thus here we focus on the integration of the technology with data center waste heat. In the first stage of a conventional MED system, water is boiled creating steam which is then used as a heat source to boil salt water. The water vapor exiting the first stage then acts as a heating medium for the second stage. This procedure continues until the last stage in which the boiling salt water acts as a condensing medium for the water vapor coming from the previous stage. Thus the MED method is based on the capture of the waste heat from each stage, until the quality of the heat has dropped too low to use.

The thermal performance, operating and capital costs of a typical MED system depends on the number of stages [103]. By increasing the number of stages the energy consumption is reduced but capital costs are increased [103]. In practice, the number of stages in a typical MED system is between 4 and 21 [103].

By using waste heat, the energy costs will decrease further, particularly if the steam boiler can be eliminated. Li et al. [103] used a waste hot water stream from an industrial plant to power an MED system. However, the low quality waste heat in this case required the use of an absorption heat pump to upgrade the waste heat in order to boil the salt water [99]. The required waste heat temperature is 75 °C or higher depending on the salt content of the water and the system pressure. This may be achievable in data centers which use two-phase cooling systems, but the waste heat from air or liquid cooled data centers would require boosting with a heat pump.

In Li et al.'s system [103] the absorber/heat pump becomes the heat source for the first MED stage. The saturation temperature in the first stage in this design is 70 °C based on a salt content of 3.8% and the absorber temperature is 75 °C [103]. The MED system consisted of four stages. Industrial wastewater entered the first stage at 90 °C and exited at ∼75 °C. This steam then became the heat source for the second stage. The water was then sequentially reused for all four stages and by the end of process had cooled down to 27 °C [103]. The extraction of heat down to 27 °C is beneficial for data center waste heat recovery. In fact, if the waste heat extracted at 75–90 °C drops to 27 °C during the MED process, the need for a chiller is eliminated as all the heat is absorbed during desalination process. This solution may be particularly appealing for use in shipboard electronics.

In summarizing the use of data center waste heat in desalination/clean water production the following issues are pointed out:

- Multistage MED desalination/clean water production requires waste heat temperatures of 75 °C and higher. These temperatures are at the upper end of available waste heat from water cooled data centers, and in line with advanced two-phase flow cooling systems.
- The benefits of using data center waste heat in desalination/clean water production include the possible elimination of the need for a chiller due to complete extraction of heat during the MED process, and of course the production of clean water, which would be in demand for shipboard systems.
- The challenges of using data center waste heat in desalination/clean water production are that this technology is not suitable for air cooled data centers and that clean water is not a priority need in most data center locations.

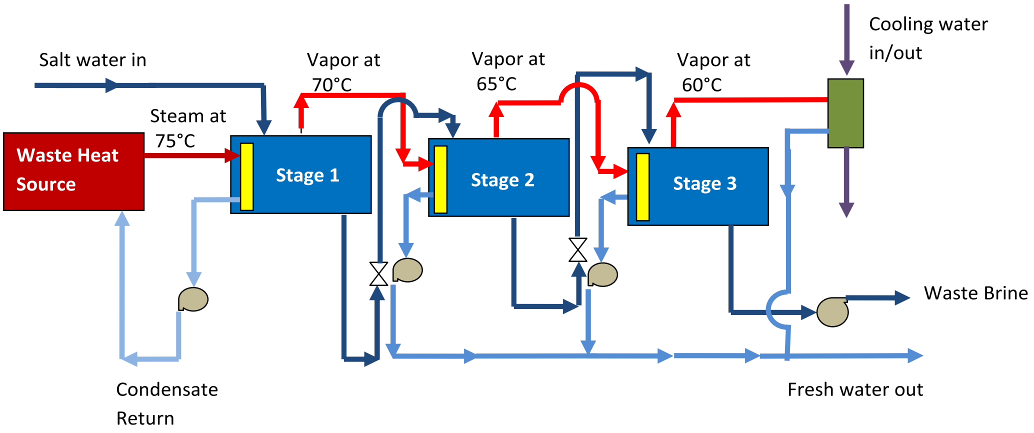


Fig. 5. A three stage MED with the first stage energized by data center waste heat.

3.8. Biomass co-location

The use of data center waste heat for processing of biomass fuels should be considered as a potential income stream and/or on-site power source for newly sited data centers. Biomass production from organic plant or animal materials provides a renewable energy source that can reduce carbon emissions. Recent reports [104,105] show growing interest in the co-location of data centers with biomass facilities including plans for a 1 MW data center powered using manure from a 10,000 cow farm and a 50 MW data center which will feature 100% renewable power from bio-solids and municipal waste [104].

There are two primary ways of producing power from plant materials or bio-solid wastes, and both require substantial sources of heat. The first is simply drying and burning either plant materials such as switch grass or algae or bio materials such as manure as the heat source to produce steam in a power plant cycle. In this case the waste heat from the data center can be used as the heat source for drying the biofuel material. Biomass drying of this type is most efficient with temperatures above 60 °C, but can be completed with temperatures as low as 45 °C.

The second method is an anaerobic digestion process of farm wastes which produces a biogas containing 60–70% methane [105]. The methane can be used for heating, cooling or power production. The remaining effluent retains the inorganics for further reuse as fertilizer. This fertilizer material has few remaining decomposable compounds and thus little odor [105]. In this system, the waste heat from the data center can be used to keep the anaerobic digestion reactor warm and/or to reduce the moisture content in the biomaterial prior to the anaerobic process. Data center waste heat above 60 °C will be required for either process.

In summarizing the use of data center waste heat in biomass production, the following issues are pointed out:

- Biomass drying is most efficient with waste heat above 60 °C but can be used in some cases with temperatures of 45 °C. Thus biofuel production is a possible use for waste heat from all types of data centers.
- The benefits of using data center waste heat in biomass drying include the production of an energy source that is clean and renewable, can which can lead to an overall reduction in energy costs due to on-site production.
- The challenges of using data center waste heat biomass drying are that this technology is simply not suitable for air cooled data centers and cannot be retrofit into existing liquid cooled data centers due to the specific siting needs of the technology. As the data center waste heat will degrade significantly with distance, the data center will need to be sited directly at the power plant location.

4. Summary

This paper presented a comprehensive review of data center cooling technologies, operating conditions and corresponding availability of data center waste heat sources and streams. A number of currently available and developmental low-grade waste heat recovery techniques were reviewed in order to assess the suitability and effectiveness of each technology specifically for the reuse of low quality data center waste heat. The advantages and disadvantages of technology were discussed. Based on this discussion, a summary of the waste heat technologies and their suitability for integration with each of three main data center cooling classifications (air cooled, liquid cooled, two-phase) is presented in Table 13. It is clear that the higher quality waste heat stream available from liquid cooled and two-phase cooled systems leads to a greater variety of possible waste heat reuse scenarios. Several of the waste heat reuse scenarios have specific co-located siting conditions (such as biomass processing) and are thus unsuitable in most cases for retrofit into existing data centers as seen in Table 14.

Table 13
Suitability of each waste heat technology for data center applications.

Technology	Air cooled data center (waste heat 45 °C)	Water cooled data center (waste heat 60 °C)	Two phase cooled data center (waste heat 75 °C)
HVAC/domestic hot water	Yes	Yes	Yes
District heating	Yes, with booster	Yes	Yes
Boiler feedwater preheating	No	Yes	Yes
Absorption refrigeration	No	Yes	Yes
Organic rankine cycle	Yes, with booster	Yes	Yes
Piezoelectrics	Yes	No	No
Thermoelectrics	No	No	Yes
Desalination	No	Yes	Yes
Biomass processing	Yes	Yes	Yes

Table 14
Suitability of each waste heat technology for retrofit and new data center designs.

Technology	Retrofit	New siting	Co-location required
HVAC/domestic hot water	Yes	Yes	No
District heating	Maybe	Yes	Yes
Boiler feedwater preheating	No	Yes	Yes
Absorption refrigeration	Yes	Yes	No
Organic rankine cycle	Yes	Yes	No
Piezoelectrics	Yes	Yes	No
Thermoelectrics	Yes	Yes	No
Desalination	No	Yes	Yes
Biomass processing	No	Yes	Yes

Thus, based on a comparison between data centers′ operational thermodynamic conditions and the operational requirements of the waste heat recovery techniques, and based on applicability for the widest range of data center applications, absorption refrigeration and organic Rankine cycle are found to be among the most promising and economically beneficial technologies for data center waste heat reuse. Both technologies offer specific benefits that are of particular interest to data center operators. Absorption refrigeration offers a source of chilled water for additional cooling load and direct reduction of CRAC loading, while ORC provides on-site electricity generation directly from the waste heat stream. These technologies are recommended for future study and optimization for data center implementation.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant no. IIP-1134810. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] Glanz J. Power, pollution and the internet. New York Times; Sept 23, 2012, Sect. A:1.

[2] Rambo J, Joshi Y. Modeling of data center airflow and heat transfer: state of the art and future trends. Distrib Parallel Database 2007;21:193–225.

[3] Brunschwiler T, Smith B, Ruetsche E, Michel B. Toward zero-emission data centers through direct reuse of thermal energy. IBM J Res Dev 2009;53(3):1–13 (11).

[4] Best practice guide for energy-efficient data center design. U.S. Department of Energy FEMP. NREL/BR-7A40-47201. DOE/GO-102010-2956. Office of Energy Efficiency and Renewable Energy; 2011.

[5] Patel CD. A vision of energy aware computing from chips to data centers. In: Proceedings of ISMME 2003. Tsuchiura, Japan; Dec 1–3, 2003.

[6] Gandhi A, Harchol-Balter M, Das R, Lefurgy C. Optimal power allocation in server farms. In: Proceedings of SIGMETRICS/performance′09. Seattle (WA, USA); Jun 15–19, 2009.

[7] Weirman A, Andrew LLH, Tang A. Power-aware speed scaling in processor sharing systems: optimality and robustness. Perform. Eval. 2012;69:601–22.

[8] Patel CD, Bash CE, Belady C, Stahl L, Sullivan D. Computational fluid dynamics modeling of high compute density data centers to assure system inlet air specifications. In: Proceedings of IPACK′01. Kauai (HI, USA); 8–13 July, 2001.

[9] Blazek M, Chong H, Loh W, Koomey JG. Data centers revisited: assessment of the energy impact of retrofits and technology trends in a high-density computing facility. J Infrastruct Syst 2004;10:98–104.

[10] Khankari K. Rate of heating analysis of data centers during power shutdown. ASHRAE Trans 2011;117(1):212–21.

[11] Kant K. Data center evolution-A tutorial on state of the art, issues, and challenges. Comput Netw 2009;53:2939–65.

[12] Marcinichen JB, Olivier JA, Thome JR. On-chip two-phase cooling of datacenters: cooling systems and energy recovery evaluation. Appl Therm Eng 2012;41:36–51.

[13] Samadiani E, Amur H, Krishnan B, Joshi Y, Schwan K. Coordinated optimization of cooling and IT power in data centers. J Electron Packag 2010;132 (031006):1–14.

[14] Rasmussen N. Guidelines for specification of data center power density [White paper No. 120]. APC 2005.

[15] Trutassanawin S, Groll EA, Garimella SV, Cremaschi L. Experimental investigation of a miniature-scale refrigeration system for electronics cooling. IEEE Trans Component Packag Manuf Technol 2006;29(3):678–87.

[16] Samadiani E, Joshi Y, Mistree F. The thermal design of a next generation data center: a conceptual exposition. J Electron Packag 2008;130(041104):1–8.

[17] Shah AJ, Patel CD. Designing environmentally sustainable electronic cooling systems using exergo-thermo-volumes. Int J Energy Res 2009;33:1266–77.

[18] Marcinichen JB, Thome JR, Michel B. Cooling of microprocessors with microevaporation: a novel two-phase cooling cycle. Int J Refrig 2010;33(7) 1264–76.

[19] Campbell L, Tuma P. Numerical prediction of the junction-to-fluid thermal resistance of a 2-phase immersion-cooled IBM dual core POWER6 Processor. In: Proceedings of the 28th IEEE SEMI-THERM symposium. San Jose (CA, USA); 18–22 March, 2012. p. 36–45.

[20] Iyengar M, David M, Parida P, Kamath V, Kochuparambil, B, Graybill D. Server liquid cooling with chiller-less data center design to enable energy savings. In: Proceedings of the 28th IEEE SEMI-THERM symposium. San Jose (CA, USA); 18–22 March, 2012. p. 212–23.

[21] Little AB, Garimella S. Waste heat recovery in data centers using sorption systems. J Therm Sci Eng Appl 2012;4(2):1–9 (021007).

[22] Datacom equipment power trends and cooling applications. 2nd ed. Atlanta (GA, USA): ASHRAE; 2005.

[23] Schmidt RR, Cruz EE, Iyengar MK. Challenges of data center thermal management. IBM J Res Dev 2005;49(4/5):709–23.

[24] Hamann HF, Lacey JA, O'Boyle M, Schmidt RR, Iyengar M. Rapid three-dimensional thermal characterization of large-scale computing facilities. IEEE Trans Components Packag Manuf Technol 2008;31(2):444–8.

[25] McAllister S, Carey VP, Shah A, Bash C, Patel C. Strategies for effective use of exergy-based modeling of data center thermal management systems. Microelectron J 2008;39:1023–9.

[26] Patel CD, Bash CE, Belady C, Stahl L, Sullivan D. Computational fluid dynamics modeling of high compute density data centers to assure system inlet air specifications. In: Proceedings of IPACK′01. Kauai (HI, USA); July 8–13, 2001.

[27] Schmidt RR, Cruz E. Raised floor computer data center: effect on rack inlet temperatures of chilled air exiting both the hot and cold aisles. In: Proceedings of ITHERM 2002, 8th intersociety conference. San Diego (CA, USA); March 30–June 1, 2002. p. 580–94.

[28] Patel CD, Sharma R, Bash CE, Beitelmal A. Thermal considerations in cooling large scale high compute density data centers. In: Proceedings of ITHERM 2002, 8th intersociety conference. San Diego (CA, USA); March 30–June 1, 2002. p. 767–76.

[29] Sharma RK, Bash CE, Patel, CD. Dimensionless parameters for the evaluation of thermal design and performance of large-scale data centers. In: Proceedings of the 8th ASME/AIAA joint thermophysics and heat transfer conference. St. Louis (MO, USA); June 24–26, 2002.

[30] Rambo J, Joshi Y. Multi-scale modeling of high power density data centers. In: Proceedings of IPACK′03. Maui (HI, USA); July 6–11,2003.

[31] Shrivastav S, Sammakia B, Schmidt RR, Iyengar M. Comparative analysis of different data center airflow management configurations. In: Proceedings of IPACK′05. San Francisco (CA, USA); July 17–22, 2005.

[32] Bhopte S, Agonafer D, Schmidt RR, Sammakia B. Optimization of data center room layout to minimize rack inlet air temperature. In: Proceedings of IPACK′05. San Francisco (CA, USA); July 17–22, 2005.

[33] Iyengar M, Schmidt RR, Sharma A, McVicker G, Shrivastava S, Sri-Jayantha, S, et al. Thermal characterization of non-raised floor air cooled data centers using numerical modeling. In: Proceedings of IPACK′05. San Francisco (CA, USA); July 17–22, 2005.

[34] Sharma RK, Bash CE, Patel CD, Friedrich RJ, Chase SJ. Balance of power: dynamic thermal management for internet data centers. IEEE Internet Comput 2005;9:42–9.

[35] Rambo J, Joshi Y. Thermal modeling of technology infrastructure facilities: a case study of data centers. In: Minkowycz WJ, Sparrow EM, Murthy JY, editors. The handbook of numerical heat transfer, II. New York: Taylor and Francis; 2006. p. 821–49.

[36] Beitelmal AH, Patel CD. Thermo-fluids provisioning of a high performance data center. Distrib Parallel Database 2007;21:227–38.

[37] Li G, Li M, Azarm S, Rambo J, Joshi Y. Optimization of data center cabinets with a new multi-objective genetic algorithm. Distrib Parallel Database 2007;21:167–92.

[38] Joshi Y, Samadiani E. Energy efficient thermal management of data centers via open multi-scale design: a review of research questions and approaches. J Enhanc Heat Transfer 2011;18(1):15–30.

[39] Ahuja N. Datacenter power savings through high ambient datacenter operation: CFD modeling study. In: Proceedings of the 28th IEEE SEMI-THERM symposium. San Jose (CA, USA); March 18–22, 2012. p. 104–7.

[40] Shah A, Carey VP, Bash CE, Patel CD. Exergy analysis of data center thermal management system. J Heat Transfer 2008;130(021401):1–10.

[41] Schmidt RR, Notohardjono BD. High-end server low-temperature cooling. IBM J Res Dev 2002;46(6):739–51.

[42] Ohadi M, Qi J Thermal management of harsh-environment electronics. In: Proceedings of the 20th IEEE SEMI-THERM symposium. San Jose (CA, USA); March 9–11, 2004.

[43] Liu S, Leung B, Necka A, Memik SO, Memik G, Hardavellas N. Hardware/software techniques for DRAM thermal management. In: Proceedings of the 17th IEEE HPCA symposium. San Antonio (TX, USA); February 12–16, 2011.

[44] Lin CH, Yang CL, King KJ. PPT: joint performance/power/thermal management of DRAM memory for multi-core systems. In: Proceedings of the 14th ACM/IEEE ISLPED symposium. San Francisco (CA, USA); August 19–21, 2009. p. 93–8.

[45] Pinheiro E, Weber WD, Barroso LA. Failure trends in a large disk drive population. In: Proceedings of the 5th USENIX FAST conference. San Francisco (CA, USA); February 13–16, 2007. p. 17–29.

[46] 600 GByte, 3.5″, 15,000 rpm disk drive specification – 6 Gb SAS-2 interface. Sun Microsystems, Inc. 〈http://docs.oracle.com/cd/E19814-01/820-7290-10/820-7290-10.pdf〉.

[47] WD Caviar®Green™Desktop hard drives. Western Digital Technologies, Inc. 〈http://www.wdc.com/wdproducts/library/SpecSheet/ENG/2879-701229.pdf〉.

[48] Bash CE, Patel CD, Sharma RK. Efficient thermal management of data centers – immediate and long-term research needs. HVAC&R Res 2003;9(2):137–52.

[49] Ohadi MM, Dessiatoun SV, Choo K, Pecht M, Lawler JV. A comparison analysis of air, liquid, and two-phase cooling of data centers. In: Proceedings of the 28th IEEE SEMI-THERM symposium. San Jose (CA, USA); March 18–22, 2012. p. 58–63.

[50] Wang Z, Tolia N, Bash C. Opportunities and challenges to unify workload, power and cooling management in data centers. ACM SIGOPS Operating Syst Rev 2010;44(3):41–6.

[51] Crippen MJ, Alo RK, Champion D, Clemo RM, Grosser CM, Gruendler NJ, et al. Blade center packaging, power and cooling. IBM J Res Dev 2005;49(6): 887–904.

[52] Khalifa HE, Demetriou DW. Energy optimization of air-cooled data centers. J Therm Sci Eng Appl 2010;2(041005):1–13.

[53] Maximizing cooling efficiency in a concurrently maintainable and fault tolerant data center. San Jose (CA, USA): Silicon Valley Leadership Group; 2009.

[54] Greenberg S, Mill E, Tschudi B, Rumsey P, Myatt B. Best practices for data centers: lessons learned from benchmarking 22 data centers. In: Proceedings of 14th ACEEE summer study on energy efficiency in buildings. Pacific Grove (CA, USA); August 13–18, 2006.

[55] Ellsworth MJ, Iyengar, MK. Energy efficiency analyses and comparison of air and water cooled high-performance servers. In: Proceedings of IPACK′09. San Francisco (CA, USA); July 19–23, 2009.

[56] Ellsworth MJ, Goth GF, Zoodsma RJ, Arvelo A, Campbell LA, Andrel JW. An overview of the IBM power 775 supercomputer water cooling system. In: Proceedings of IPACK′11. Portland (OR, USA); 2011.

[57] David MP, Iyengar M, Parida P, Simons R, Schultz M, Schmidt R, et al. Experimental characterization of an energy efficient chiller-less data center test facility with warm water cooled servers. In: Proceedings of the 28th IEEE SEMI-THERM symposium. San Jose (CA, USA); March 18–22, 2012. p. 232–7.

[58] Parida PR, David M, Iyengar M, Schultz M, Gaynes, M, Kamath, V, et al. Experimental investigation of water cooled server microprocessors and memory devices in an energy efficient chiller-less data center. In: Proceedings of the 28th IEEE SEMI-THERM symposium. San Jose (CA, USA); March 18–22, 2012. p. 224–31.

[59] Sharma CS, Zimmermann S, Tiwari MK, Michel B, Pulikakos D. Optimal thermal operation of liquid-cooled electronic chips. Int J Heat Mass Tranfr 2012;55:1957–69.

[60] Mudawar I. Assessment of high-heat-flux thermal management schemes. IEEE Trans Components Packag Manuf Technol 2001;24:122–41.

[61] Sung MK, Mudawar I. Single-phase and two-phase hybrid cooling schemes for high-heat-flux thermal management of defense electronics. In: Proceedings of

ITHERM 2008, 11th intersociety conference. Orlando (FL, USA); May 28–31, 2008. p. 121–31.

[62] Tuckerman DB, Pease RFW. High-performance heat sinking for VLSI. IEEE Electron Device Lett 1982;2:126–9.

[63] Bowers MB, Mudawar I. High flux boiling in low flow rate, low pressure drop mini-channel and micro-channel heat sinks. Int J Heat Mass Transfer 1994;37:321–32.

[64] Leonard PL, Phillips AL. The thermal bus opportunity-A quantum leap in data center cooling potential. ASHRAE Trans 2005;3(2):732–45.

[65] Hannemann R, Marsala J, Pitasi M. Pumped liquid multiphase cooling. In: Proceedings of IMECE′04. Anheim (CA, USA); November 13–19, 2004. p. 469–73.

[66] Madhour Y, Olivier J, Costa-Patry E, Paredes S, Michel B, Thome JR. Flow boiling of R134a in a multi-microchannel heat sink hotspot heaters for energy-efficient microelectronic CPU cooling applications. IEEE Trans Components Packag Manuf Technol 2011;1(6):873–83.

[67] Agostini B, Fabbri M, Thome JR, Michel B. High heat flux two-phase cooling in silicon multi microchannels. IEEE Trans Components Packag Manuf Technol 2008;31(3):691–701.

[68] Thome JR, Agostini B, Revellin R, Park JE. Recent advances in thermal modeling of micro evaporators for cooling of microprocessors. In: Proceedings of IMECE′07. Seattle (WA, USA); November 11–15, 2007.

[69] Thome JR, Bruch A. Refrigerated cooling of microprocessors with micro-evaporation heat sinks: new developments and energy prospects for green data centers. IOR Proceedings, vol. 105; 2008/9.

[70] Zhou R, Zhang T, Catano J, Wen JT, Michna CJ, Peles Y, et al. The steady-state modeling and optimization of a refrigeration system for high heat flux removal. Appl Therm Eng 2010;30:2347–56.

[71] Mauro AW, Thome JR, Toto D, Vanoli GP. Saturated critical heat flux in a multi-microchannel heat sink fed by a split flow system. Exp Therm Fluid Sci 2010;34(1):81–92.

[72] Mongia R, Masahiro K, DiStefano E, Barry J, Chen W, Izenson M, et al. Small scale refrigeration system for electronics cooling within a notebook computer. In: Proceedings of ITHERM 2006, 10th intersociety conference. San Diego (CA, USA); May 30–June 2, 2006. p. 751–8.

[73] Liu J, Goraczko M, James S, Belady C, Lu J, Whitehouse K. The data furnace: heating up with cloud computing. In: Proceedings of 3rd USENIX workshop on hot topics in cloud computing. Portland (OR, USA); June 14–15, 2011.

[74] ⟨http://www.eia.gov/consumption/residential/⟩, Viewed October 7, 2012.

[75] Vela, J. Helsinki data centre to heat homes. The Guardian 2010 July 20. ⟨http://www.guardian.co.uk/environment/2010/jul/20/helsinki-data-centre-heat-homes⟩.

[76] Henning D, Gebremedhin A. District heating and cooling enable efficient energy resource utilisation. In: Gebremedhin A, editor. Sustainable energy – recent studies. Rijeka: InTech; 2012 (published online).

[77] Persson U, Werner S. District heating in sequential energy supply. Appl Energy 2012;95:123–31.

[78] Data center heat recovery helps Intel create green facility. Intel 2007. ⟨http://www.intel.com/content/www/us/en/green-it/energy-efficiency/intel-it-data-center-heat-recovery-helps-create-green-facility-brief.html⟩.

[79] Waste heat recovery: technology and opportunities in US industry. US Department of Energy; 2008.

[80] Srikhirin P, Aphornratana S, Chungpaibulpatana S. A review of absorption refrigeration technologies. Renew Sustain Energy Rev 2001;5(4):343–72.

[81] Ayou DS, Bruno JC, Saravanan S, Coronas A. An overview of combined absorption power and cooling cycles. Renew Sustain Energy Rev 2013;21:728–48.

[82] Sun J, Fu L, Zhang S. A review of working fluids of absorption cycles. Renew Sustain Energy Rev 2012;16(4):1899–906.

[83] Haywood A, Sherbeck J, Phelan P, Varsamopoulos G, Gupta, SKS. A sustainable data center with heat-activated cooling. In: Proceedings of ITHERM 2010, 12th intersociety conference. Las Vegas (NV, USA); June 2–5, 2010.

[84] Haywood A, Sherbeck J, Phelan P, Varsamopoulos G, Gupta SKS. Thermodynamic feasibility of harvesting data center waste heat to drive an absorption chiller. Energy Convers Manage 2012;58:26–34.

[85] Suman S, Joshi YK, Fedorov AG. Cryogenic/sub-ambient cooling of electronics: revisited. In: Proceedings of ITHERM 2004, 9th intersociety conference. Las Vegas (NV, USA); June 1–4, 2004.

[86] Kim YJ, Kim S, Joshi YK, Fedorov AG, Kohl PA. Thermodynamic analysis of an absorption refrigeration system with ionic-liquid/refrigerant mixture as a working fluid. Energy 2012;44:1005–16.

[87] Kim YJ, Joshi YK, Fedorov AG. An absorption based miniature heat pump system for electronics cooling. Int J Refrig 2008;31(1):23–33.

[88] Kim S, Kim YJ, Joshi YK, Fedorov AG, Kohl PA. Absorption heat pump/refrigeration system utilizing ionic liquid & hydrofluorocarbon refrigerants. J Electron Packag 2012;134(031009):1–9.

[89] Chiriac V, Chiriac F. Miniaturized refrigeration system with absorption application to microelectronics cooling. In: Proceedings of ASME InterPACK′07. Vancouver (Canada), 8–12 July 2007.

[90] Chiriac V, Chiriac F. Absorption refrigeration method with alternative water-ammonia solution circulation system for microelectronics cooling. In: Proceedings of ITHERM 2010, 12th intersociety conference. Las Vegas (NV, USA); June 2–5, 2010.

[91] Chen H, Goswami DY, Stefanakos EK. A review of thermodynamic cycles and working fluids for the conversion of low-grade heat. Renew Sustain Energy Rev 2010;14:3059–67.

[92] Hung TC, Shai TY, Wang SK. A review of organic Rankine cycles (ORCs) for the recovery of low-grade waste heat. Energy 1997;22(7):661–7.

[93] Bao J, Zhao L. A review of working fluid and expander selections for organic Rankine cycle. Renew Sustain Energy Rev 2013;24:325–42.

[94] Vélez F, Segovia JJ, Martín MC, Antolín G, Chejne F, Quijano A. A technical, economical and market review of organic Rankine cycles for the conversion of low-grade heat for power generation. Renew Sustain Energy Rev 2012;16(6):4175–89.

[95] Tchanche BF, Gr Lambrinos, Frangoudakis A, Papadakis G. Low-grade heat conversion into power using organic Rankine cycles – a review of various applications. Renew Sustain Energy Rev 2011;15:3963–79.

[96] Saleh B, Koglbauer G, Wendland M, Fischer J. Working fluids for low-temperature organic Rankine cycles. Energy 2007;32:1210–21.

[97] Quoilin S, Lemort V. Technological and economical survey of organic Rankine cycle systems. In: Proceedings of the 5th conference on economics and management of energy in industry. Algarve, Portugal; April 14–17, 2009.

[98] Zhang T, Wang EN. Design of a microscale organic Rankine cycle for high-concentration photovoltaics waste thermal power generation. In: Proceedings of ITHERM 2012, 13th intersociety conference. San Diego (CA, USA); May 30–June 1, 2012. p. 993–1002.

[99] Martín-González M, Caballero-Calero O, Díaz-Chao P. Nanoengineering thermoelectrics for 21st century: energy harvesting and other trends in the field. Renew Sustain Energy Rev 2013;24:288–305.

[100] Venkatasubramanian R, Siivola E, Collpitts T, O'Qinn B. Thin-film thermoelectric devices with high-room temperature figure of merit. Nature 2001;413:597–602.

[101] Saidur R, Elcevvadi ET, Mekhilef S, Safari A, Mohammed HA. An overview of different distillation methods for small scale applications. Renew Sustain Energy Rev 2011;15:4756–64.

[102] Li C, Goswami Y, Stefanakos E. Solar assisted sea water desalination: a review. Renew Sustain Energy Rev 2013;19:136–63.

[103] Li H, Russell N, Sharifi V, Swithenbank J. Techno-economic feasibility of absorption heat pumps using wastewater as the heating source for desalination. Desalination 2011;281:118–27.

[104] Pouchet J. Biomass-powered data centers: next step for green IT? Environmental and energy management news; July 28, 2008 . ⟨http://www.environmentalleader.com/2011/07/28/biomass-powered-data-centers-next-step-for-green-it/⟩.

[105] Sharma R, Christian T, Arlitt M, Bash C, Patel C. Design of farm Waste driven supply side infrastructure for data centers. In: Proceedings of the ASME 2010 4th international conference on energy sustainability. Phoenix (AZ, USA); May 17–22, 2010.