

GREEN DATA CENTRE TECHNOLOGY PRIMER: A SUMMARY

BACKGROUND

Given Singapore's reputation as a trusted location for data hosting, management and analytics, the data centre industry has witnessed fast growth. While data centres are critical in supporting the development and operations of nearly every sector of the economy, they are energy-intensive facilities that contribute significantly to carbon emissions. Internationally, energy-related costs account for approximately 12% of overall data centre expenditure and are the fastest rising cost component. Hence, energy efficiency has evolved as a top priority for data centre operations, together with availability, performance and security. The focus of this technology primer is to conduct an extensive review of the existing and emerging technologies for improving overall data centre energy efficiency. Promising areas for R&D would also be highlighted.

ENERGY CONSUMPTION IN DATA CENTRES

Energy consumption in a data centre can be categorised under *IT equipment* (servers, storage and network) and *facility infrastructure*, such as chillers, humidifiers, computer room air conditioners (CRAC), power distribution units (PDU), uninterruptible power supplies (UPS), lights, and power distribution. [Figure 1](#) shows the electricity flow in a typical data centre.

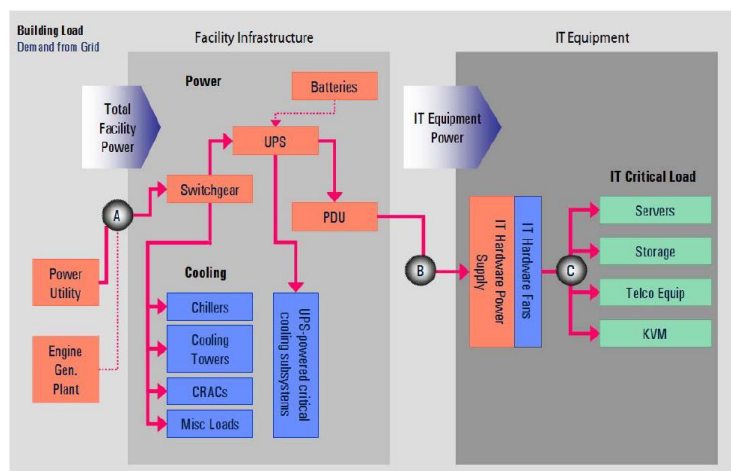


Figure 1: Electricity flow and measurement points in a typical data centre

On average, IT equipment uses 45% of the energy in a data centre and the facility infrastructure that supports the functioning of the IT equipment uses the other 55%. However, for many of the legacy data centres in Singapore, IT equipment consumes about 30% of energy in the data centre, with the remainder being consumed by facility infrastructure. This can be attributed to our tropical climate conditions as well as inappropriate cooling system or air management design and operation.

OPPORTUNITIES TO IMPROVE ENERGY EFFICIENCY IN DATA CENTRES

ICT EQUIPMENT

The authors assessed that energy efficiency for the ICT subsystem at data centres can be improved significantly, by *around 20%*. Best practices or optimisation strategies can be categorised under four hierarchical levels.

(1) Energy Efficient Embedded Components

Energy efficient embedded modules, such as energy efficient Central Processing Unit (CPU), motherboard and memory can be incorporated into the three components for the ICT subsystem (server, storage and network). In this design-based approach, the focus is the design of embedded modules to reduce energy usage and make energy usage proportional to the load. Emerging technologies that can improve energy efficiency include photonic interconnection for intra-and inter-chip communication as well as emerging memory devices such as phase changing memory and spin-transfer torque random-access memory. These increase energy efficiency through better overall server performance by improving the data storage capability and the cache access success rate.

(2) System (Server/Storage/Network) Optimisation

For server optimisation, existing technologies and design strategies have shown potential to reduce the energy use of a typical server by 25% or more.¹ For example, multi-core, Desktop Virtualisation Solutions (DVS), and power management technology (e.g. dynamic power capping) improves the efficiency of individual server units, resulting in higher performance per watt.

For storage optimisation, potential strategies include the use of tiered storage to reduce the load on electrical and cooling systems and the adoption of hardware and coding designs such as deduplication, thin provisioning, Massive Array of Idle Disks (MAID), solid state storage and alternative storage architecture.

¹ Broyles, M., C. Francois, A. Geissler, G. Grout, M. Hollinger, T. Rosedahl, G. J. Silva, M. Vanderwiel, J.V. Heuklon and B. Veale (2010): *IBM EnergyScale for POWER7 Processor-based Systems*. IBM White Paper, Feb 2010.

For network optimisation, MIT researchers have recently demonstrated a new way of organising optical networks called Optical Flow Switching (OFS) that promises to eliminate the need for converting optical signals to electrical ones for processing and converting them back again for transmission. As a result, it could significantly reduce the amount of energy consumed by the network infrastructure of a data centre.²

(3) Application Optimisation

The authors assessed that within existing ICT subsystems, implementing best energy management practices and consolidating applications from many servers to one server, enabled by virtualisation, could reduce current data centre energy usage by around 20%. Consolidation, virtualisation, efficient resource utilisation, and load balancing the way applications run on servers can improve the overall energy efficiency of a data centre.

Application optimisations are also guided by the hardware design. While some components in the computer can be designed to withstand higher temperatures, it is envisioned that servers in the long term have a mix of hardware with different temperature tolerances. Hybrid hardware design and software reliability are key challenge issues. Data replication, error handling and reliable software engineering improve the fault tolerance, and thus reduce the energy consumption of system failures and their associated recovery operations.

(4) Inter-Domain Optimisation

The interface between ICT equipment and data centre infrastructure opens an additional venue for energy savings. Dynamic Power Capping technology allows systems administrators to safely limit the amount of power consumed by one or more servers without impacting server performance. Advanced cool technologies can react proportionally to the amount of load at the rack level to save energy by providing cooling only as required. In the close-coupled architecture, cooling systems are placed near ICT equipment so that cooling capacity can follow ICT loads that move with virtualisation and server power management. An integrated management and monitoring system at the interface should be in place for green data centres, to optimise the use of power, cooling and IT resources.

The authors assessed that holistic ICT transformation initiatives can deliver a greater impact than site improvements. It is believed that there exists room for companies to more aggressively reduce unnecessary or redundant ICT infrastructure of their data centres to reduce their energy consumption, without affecting performance.

² Chan, V. W. S. (2010): *Optical Flow Switching*. Invited Paper, IEEE Optical Network Design and Modeling, Kyoto Japan, Jan 2010.

DATA CENTRE VIRTUALISATION

Innovation in servers, storage, and networking – especially server and storage virtualisation³ – has paved a new way of designing energy-efficient data centres, while accommodating the need for scaling performance and capacity.

Data centres often have underutilised servers where huge energy savings can be reaped using virtualisation technologies. Consolidation of workloads onto fewer servers using virtualisation is a powerful tool that can be applied to many data centres to increase server efficiency and reduce energy consumption. Decreasing the number of physical servers not only reduces the energy consumption of the servers, but also has a positive impact on the efficiency of the entire data centre. For example, server consolidation can result in reduced cooling load (the energy used to remove heat from the data centre), increased UPS backup time, and longer generator backup times.

Two typical approaches for server consolidation are:

- a) Combining applications onto a single server and a single operating system, and
- b) Virtualising application workloads in virtual machines and hosting them on a hypervisor.

For certain applications such as mission-critical applications, server consolidation may not be suitable as the protection of service levels is more important than improving utilisation, achieving energy savings and minimising risk for downtime.

ICT POWER MANAGEMENT AND DISTRIBUTION

Given the rapid growth in server-based ICT services, there is an increased demand for energy efficient power supply architectures. This is important as excessive heat that is generated by the nonlinear and unbalanced loads may result in harmonic currents as well as neutral currents overloading the distribution cables, transformers and ICT equipment, and in extreme cases, lead to fire. Therefore, cooling is needed in data centres to ensure safety of servers and information as well as prolonging the lifespan of equipment. One solution to reduce the heat generation is through *efficient design of power supply architecture* so as to reduce the cooling requirements and thereby increase the efficiency of the data centres.^{4,5}

Apart from improvements in power supply architectures, energy savings can come from improvements to the power distribution systems through the use of direct current (DC) distribution systems or hybrid AC/DC systems. Typically, a data centre is supplied alternating current (AC)

³ Virtualisation refers to the creation of a virtual (rather than actual/physical) version of something, in this case, a virtual server and storage.

⁴ Emerson Network Power (2011): *Maximising Data Centre Efficiency, Capacity and Availability through Integrated Infrastructure: Efficiencies and Cost Benefits Intrinsic to Integrated Infrastructure Solutions*. Emerson Network Power White Paper.

⁵ Rasmussen, N. (2007): *A Quantitative Comparison of High Efficiency AC vs. DC Power Distribution for Data Centres*. APC.

power from the grid, which is distributed throughout the centre's infrastructure. However, most of the electrical equipment, such as servers and batteries storing backup power in the UPS system, require DC power. Hence, having a DC power distribution system would improve energy efficiency through the elimination of power conversion required for AC systems. Currently, DC systems are not prevalent in Singapore. However, should the application of DC power generators increase, there is potential in the expansion of this technology in the future. See [Appendix A](#) for more details on AC/DC systems.

COOLING TECHNOLOGIES IN DATA CENTRES

In Singapore, approximately 55% to 70% of total energy inputs into data centres are utilised by the facilities infrastructure where most of it is consumed by the cooling system. While the high cooling requirement in data centres is often attributed to our tropical climate conditions, another major reason is the inappropriate design and operation of the cooling and air management system. There is vast potential for retrofits and implementation of best practices as the first level in improving overall data centre energy efficiency. As illustrated in [Figures 2 and 3](#), heat given out by sources inside a data centre goes through many stages before it reaches the external environment.

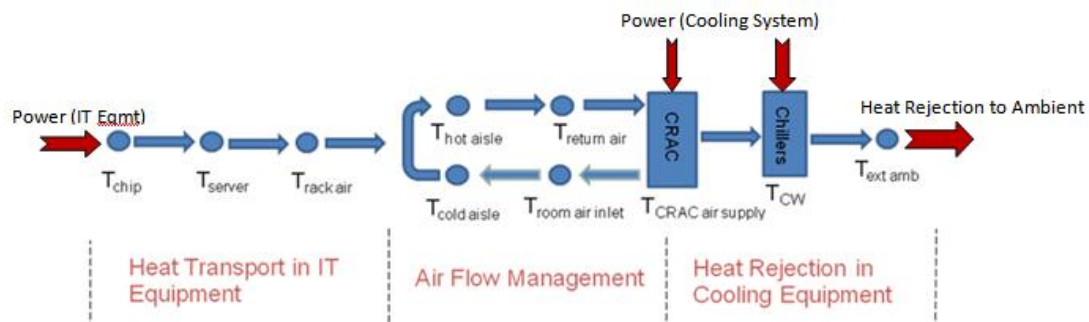


Figure 2: Heat flow paths in the data centre

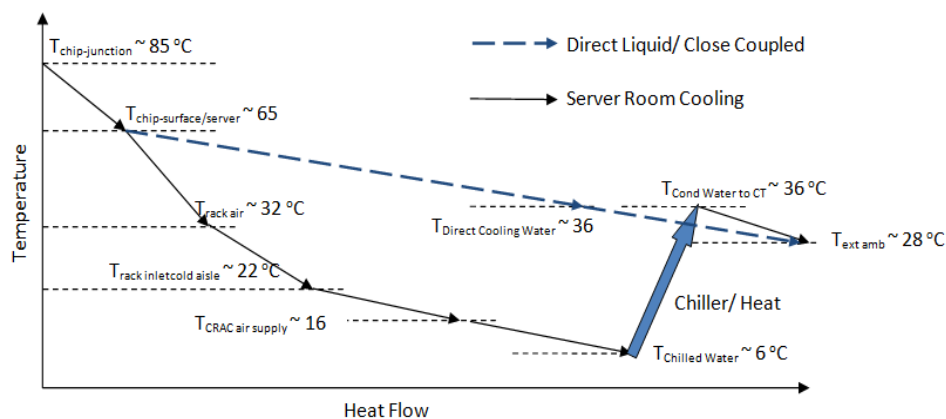


Figure 3: Typical temperature profiles in a data centre

In a typical data centre, the heat that is generated from electronic devices is conducted through interconnected surfaces of the components inside the server, which is then picked up by cooler air entering the rack in which the server is placed.

The management of cooling in a data centre can be categorised into three areas:

(1) Heat Transport in ICT Equipment

The primary impetus is to direct the cooling effort closer to its heat source, to reduce the heat transport path and hence thermal resistance. Most of the techniques involve the direct use of cooling water (either chilled or ambient temperature water from water-side economisers⁶ within the chip, server or rack. It effectively removes the air flow management stage of the energy/heat flow.

- IBM's Zurich Research Laboratory and ETH Zurich have demonstrated a chip-level water⁷ cooling system for a 10 Teraflop supercomputer configuration. The system uses water running at 60°C to maintain chip temperatures below 85°C, with the excess heat removed for reuse in office or other heating systems.
- Server-level cooling involves the heat from the processors spreading first to the boards or backplanes of the server, assisted by heat pipes with low thermal resistance.
- Rack-level cooling systems typically have a liquid cooling loop integrated into the cabinet. They allay concerns on the use of water, as the liquid is not in such close proximity with the processor. Rear door heat exchangers can be installed on racks to cool the hot air being expelled, reducing the burden on the air distribution system in the data hall.

A variant of rack, server and chip-level cooling is immersion cooling where the server boards or entire racks are submerged in baths of an inert liquid coolant.

(2) Data Centre Air Flow Management

Conventional data centres use cold air to remove heat from ICT equipment. Inappropriate placement of the ICT equipment may result in mixing of cool and warm air streams, resulting in inefficient use of available cooling capacity. Developments in the last decade have shown that careful air flow management in the computer room can reduce cooling requirements. The principles of air management can be summarised as follows:

⁶ Economisers are cooling devices that take advantage of outdoor conditions to reduce the need for refrigeration in data centres. They come in two forms – air driven or water driven. The former utilises cool air from the outside for direct cooling or passed through heat exchanges prior to cooling; while the latter uses cool outside air to help chill the water that is being used to cool the air inside the data centre.

⁷ Water in general is a better heat conductor than air. It has a larger specific heat capacity and a much higher convective heat transfer coefficient for flow over a heated surface.

- Arrangement of ICT equipment in hot aisle/cold aisle configuration (Figure 4) to channel cold air to the inlet side of the racks and hot air from the exhaust side to minimise air mixing.

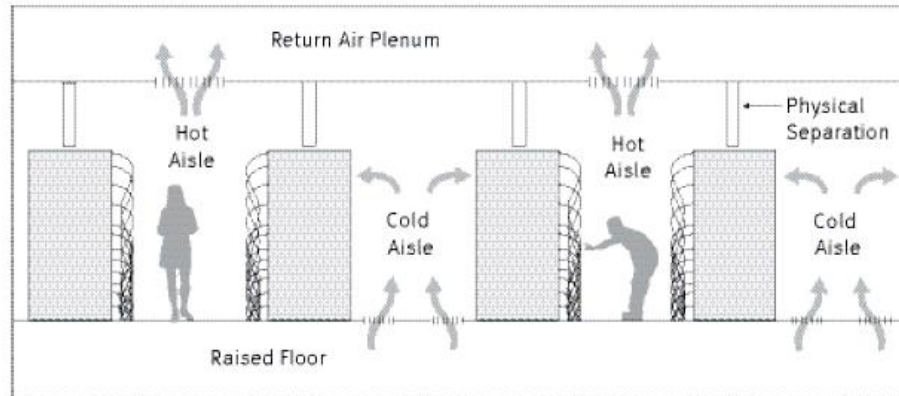


Figure 4: Hot aisle/cold aisle configuration⁸

- Immediate removal of hot exhaust air from the racks to minimise air mixing to prevent short circuiting to the rack inlet side and improve cooling.
- Allowing a higher difference between supply and return air temperatures can increase ICT load density and also reduce the size of cooling equipment.
- Preventing degrading the cooling capacity of cooling equipment by good airflow management practices such as sealing leaks in raised floor tiles and cable openings, proper placement of inlet and exhaust air openings, and reducing obstructions in raised floor air plenums.
- Use of air economiser when outside ambient conditions permit to reduce cooling energy consumption.
- Unlike human occupied space such as offices where large amounts of ventilation air need to be dehumidified and cooled for thermal comfort, ICT equipment space needs only minimal air ventilation and the heat load emitted by the equipment does not add moisture to the space. The regulation of humidity within the cooling space is only relevant if air economisers are used. In such a situation, it is possible to pre-design a tight building envelope or optimise the pressurisation of the data centre to minimise humidity issues and therefore, energy consumption.

In temperate climates, data centres with no space humidity control requirements may be able to use an economiser operated based on the dry-bulb temperature. Energy is saved by eliminating the compressor work used to cool the data centre by mechanical refrigeration processes. However, in Singapore's climate, the refrigeration process cannot be entirely eliminated due to the higher outdoor air temperature. Instead, an integrated economiser could only take in outside air that has fallen below the economiser set point, which is around the data centre return air temperature. The

⁸ Greenberg S., E. Mills, B. Tschudi, P. Rumsey, B. Myatt (2006): *Best Practices for Data Centres: Lessons Learned from Benchmarking 22 Data Centres*. ACEEE Summer Study on Energy Efficiency in Buildings, 2006.

compressors in the mechanical refrigeration plant are still required to cool the outside air down to the supply air temperature set point. The potential to reduce cooling energy increases as the temperature difference between the outdoor air and warmer return air widens.

(3) *Heat rejection by cooling equipment*

In data centres, the heat from the ICT equipment space is rejected to the external environment using a refrigerating or cooling system. There are different centralised and distributed cooling technology options⁹ available to achieve this.

Therefore, careful selection and design of cooling equipment and capacity should be made to ensure that efficient cooling is achieved, including at part load conditions.

INTEGRATED ENERGY SYSTEMS AND ALTERNATIVE ENERGY RESOURCES

(1) *Combined Heat and Power (CHP) Plants*

Combined heat and power (CHP) or co-generation is the simultaneous production of both power and heat from a single fuel source (Figure 5). By making use of the waste heat from on-site electricity production for heating or cooling, CHP reduces energy costs.

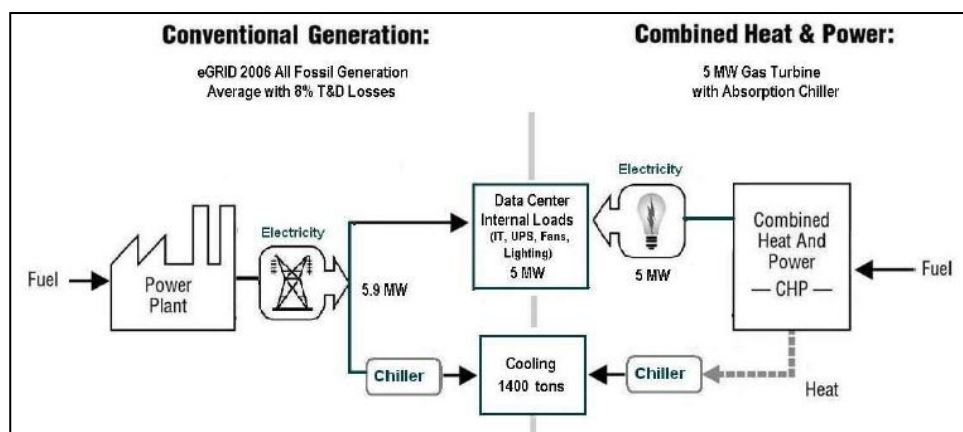


Figure 5: Concept of combined heat and power generation

Data centres are required to operate continuously (24/7) and have a high demand for electricity and cooling at a near constant load. These are conditions which are ideal for on-site generation.

⁹ These can range from large centralised chiller plants (circulating chilled water to the computer room air-handling unit (CRAU) which in turn cools the air circulated to the racks), to direct expansion (DX) units that circulate refrigerant directly to the CRAC to cool air circulated to the ICT space.

Typically, this will involve a combined heat and power (CHP) plant, where the heat recovered from the power generation is used in a thermally-activated (e.g. absorption cycle) cooling system.

On-site generation also helps to meet the redundancy requirements for power supplies by interconnection with grid supplied power. Since the CHP plants and grid power are in continuous operation, on-site generation provides a higher level of reliability than standby generators, which are commonly used as backup power. Backup for the cooling system can be via electric-powered chillers, supplied from the critical power network. A typical power and cooling infrastructure for a data centre site would look like [Figure 6](#).

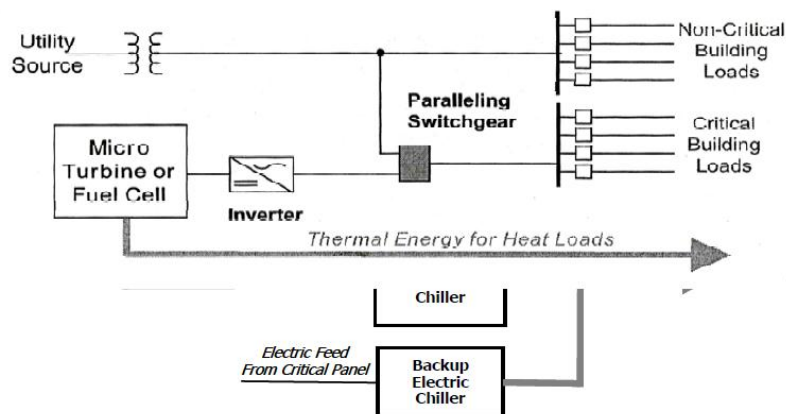


Figure 6: Typical power and cooling infrastructure for a data centre³⁸

(2) CHP and Cooling plant technologies

To maximise the energy advantages for on-site generation, the technologies for the CHP and Cooling plants must be carefully selected. For smaller data centres (< 5 MW), fuel cells, reciprocating engines or microturbines are typically used for the CHP plant while for larger capacities (> 5 MW) gas turbine powered CHP plants are more economic.

The proportion of power to heat provided by the CHP plant is critical. In general, absorption chillers have a much lower coefficient of performance (COP)¹⁰ than electric-powered chillers. Hence it is not advisable to maximise the heat production in order to generate cooling from absorption chillers. Optimisation of the proportion of power to IT equipment, heat for the absorption chillers and power for the electric chillers together, with the characteristics of the CHP plant will produce the best overall energy performance.

¹⁰ The Coefficient of Performance (COP) is the ratio of the rate of heat removal from the space to be cooled ($\text{kW}_{\text{cooling}}$) to the energy input rate required by the system (kW_{in}). Industry practice usually specifies the kW power input required per ton of refrigeration (kW/RT , $1\text{RT} = 3.516 \text{ kW}_{\text{cooling}}$), which has an inverse relationship with COP.

The type of absorption cycle plant used depends on the quality of the waste heat available from the CHP plant. A single effect absorption chiller requires low pressure steam or pressurised hot water of 100°C or less to produce cooling from phosphoric acid fuel cells, reciprocating engines and microturbine cooling systems, with a COP of about 0.7. This means that about 5kW of the waste heat can produce one ton of cooling. The exhaust of gas turbines, reciprocating engines and microturbines and heat recovered from molten carbonate or solid oxide fuel cells can produce high pressure steam to drive a double-effect absorption chiller capable of delivering a COP of 1.2, i.e. only about 3kW of the high quality waste heat is needed to produce one ton of cooling. For comparison, COPs of best performing electric chillers currently are around 6.0, i.e. less than 0.6kW of electricity is needed to produce one ton of cooling.

(3) District integrated energy management system and Data Centre Parks

When there is a high concentration of data centres in a district, e.g. in data centre parks or financial districts, there is additional impetus for consolidation of the power and cooling infrastructure to provide the same level of reliability with reduced redundancy requirements. For the power supply infrastructure, increasing the power density will improve the economics of on-site generation coupled with reduced transmission and distribution losses. However, a major issue limiting the attractiveness of on-site generation is the need for on-site fuel storage for on-site generators. Currently, the preferred fuel choice of most efficient CHP plants is piped natural gas which is not widely available due to the lack of infrastructure in most parts of Singapore.

For the cooling infrastructure, while centralising the cooling plant into a district cooling system (DCS) may bring about savings in the space and maintenance requirements, this is balanced by the need to construct and maintain the chilled water distribution network. Moreover, the cooling demand profile for data centre parks is relatively flat all year round, without the diversity needed to gain economic benefits from peak-load shaving or thermal storage by cooling during off-peak hours. What it stands to gain instead is the centralisation of back-up cooling capacity, which if coupled with alternative pipe network paths can provide a similar level of reliability while enabling the cooling plants to work at a higher part-load condition.

Another critical issue in data centre park planning is that data centre facilities often start out at low loads and grow into the site gradually. The CHP system design must therefore be based on realistic load estimation and provide a modular solution for future growth.

(4) Use of renewables

The renewable energy technologies that are useful for data centre applications in Singapore will most likely be sourced from solar photovoltaic (PVs) and solar thermal energy.

Solar PVs are currently the most common source of renewable energy for buildings in Singapore. However, data centres are 10-100 times more energy intensive than an office building,¹¹ so the contribution of solar PVs to data centres is unlikely to be significant.

Using solar thermal energy to supplement the waste heat for the absorption cooling plants in a centralised CHP plant, e.g. Data Centre Park, may help to improve the economics as the cooling energy increases in direct proportion to solar availability. Flat-plate collectors can achieve the temperatures required for single-effect absorption system, while concentrated evacuated tube collectors will be needed for double-effect absorption plants.

In general, data centres are mission-critical, energy-hungry facilities requiring highly reliable, uninterruptible power sources. As such, there are limitations in the use of renewable energy, considering their unpredictable and intermittent nature.

(5) *Energy recovery and reuse*

Temperatures in most data centre hot aisles range from 27°C to 46°C, and are considered low compared to most heat recovery strategies. However, a growing number of data centres are re-directing the heat from their hot aisles to nearby homes, offices, greenhouses and even swimming pools. The ability to re-use excess heat from the servers in new data centres helps to improve the energy efficiency profile of these facilities. [Figure 7](#) shows a typical set up for such applications.

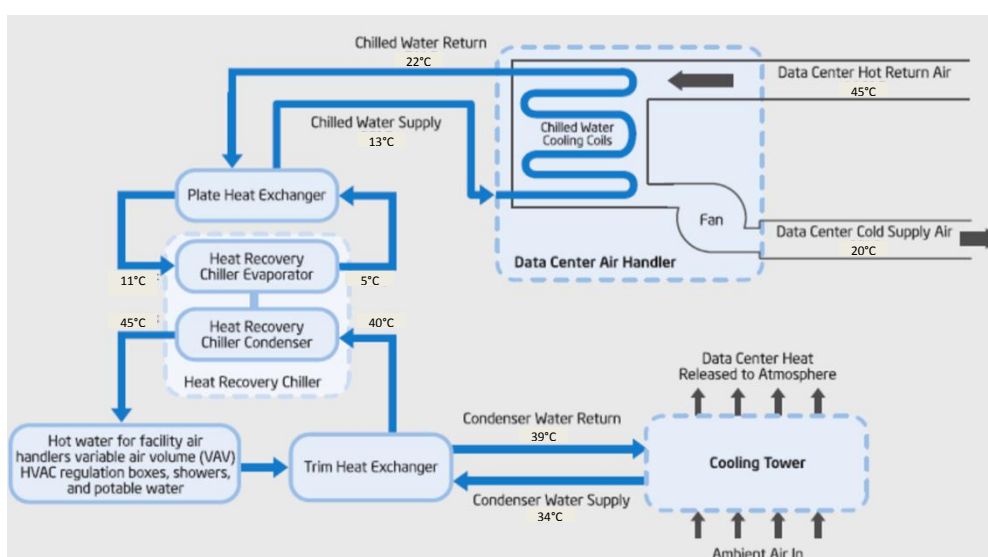


Figure 7: Energy recovery from data centres for water heating¹²

¹¹ LBNL (2010): *Environmental Performance Criteria (EPC): Guide for Green Building Operations & Maintenance of Existing Data Centres/ New Constructions*.

¹² Intel (2007): *Data Centre Heat Recovery Helps Intel Create Green Facility*. Intel White Paper.

ASSESSMENT OF GREEN DATA CENTRE PERFORMANCE

PERFORMANCE METRICS

Performance metrics for green data centres are a set of measurements that can evaluate the energy and environmental effects of operating a data centre.¹³

There are three categories of metrics for data centres:

- Overall Data Centre Efficiency
- IT (or Server-level) Efficiency
- Facility infrastructure system-level Efficiencies, covering systems such as cooling, air flow management, power transformation and distribution

(1) **Overall Data Centre Efficiency**

The most widely recognised metric of overall data centre efficiency is the Power Usage Effectiveness (PUE).¹⁴ It is the total energy consumption of the data centre facility divided by the IT energy consumption with a lower PUE regarded as being more energy efficient. Values of PUE can range widely and are dependent on climatic conditions because of the demands on cooling. They are also affected by the utilisation level of data centre. A study commissioned by the National Environment Agency (NEA) in Singapore involving a wide range of data centres, showed the average PUE was comparable to values obtained from studies made in the US and Europe, despite Singapore being located in a hot, humid climate (Figure 8). It is expected that currently a well-managed data centre facility in Singapore at utilisation of 33% or more should be able to achieve a PUE between 1.5 to 2.0.

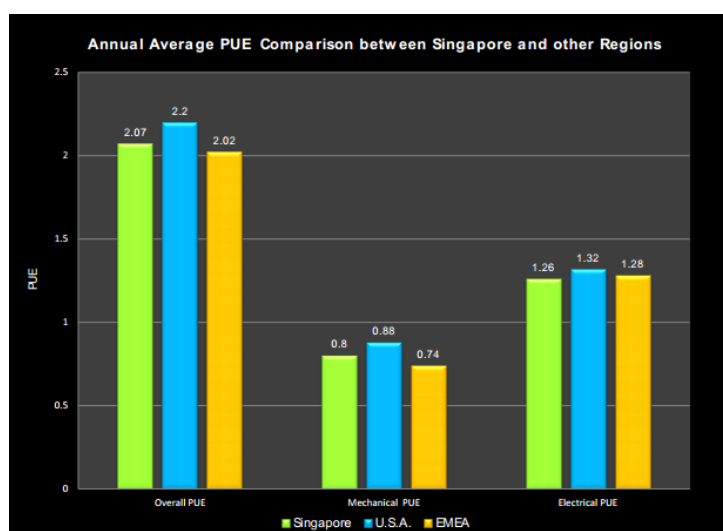


Figure 8: Annual average PUE comparison between Singapore and other regions

¹³ Wang, L. and S.U. Khan (2011): *Review of Performance Metrics for Green Data Centres: A Taxonomy Study*. Journal of Supercomputing.

¹⁴ Kaplan, J. M., W. Forrest and N. Kindler (2008): *Revolutionising Data Centre Energy Efficiency*. McKinsey & Company.

Despite being widely recognised, there are limitations of using PUE as an overall data centre metric. For example, if improvements are made in IT equipment energy consumption and efficiency, the PUE will increase despite there being energy savings. In addition, because there is currently no consensus on where measurements are taken, it is difficult to compare between different data centre's PUE performance. Hence, other performance metrics such as IT performance metrics for individual ICT components are also used.

(2) IT (or Server-level) Efficiency

Different ICT components have their own unique energy performance metrics. For servers, there are two types of energy performance metrics used:

- Power-to-Performance Ratio: GFlops/watt
 - Represents the energy efficiency of a particular computer architecture or hardware.
 - Rate of computation that can be delivered by a computer for every watt of power consumed
- Physical Server Reduction Ratio (PSRR)
 - Refers to the reduction in physical servers (e.g. through virtualisation & consolidation)
 - Ratio of installed server base from historical trends to the installed server base in the post-virtualisation scenarios. E.g. a PSRR ratio of 2:1 means that installed physical server base are halved compared with historical trends

Meanwhile, for storage, the following are the metrics used:

- Capacity metric (GB/watt)
 - Represents the energy efficiency of storing data
 - Ratio between used storage space by applications (GB) and power (Watts)
- Workload metric (IOPS/watt)
 - Represents the energy cost of the storage system while running the workload
 - Ratio between the applications Input/Output rate (IOPS) and power (Watts)
- Bandwidth metric: MBPS/watt
 - Similar to Input/Output throughput, referring to data transfer
 - Ratio between applications data transfer rate (MBPS) and power (Watts)

(3) Facility Infrastructure System-level Efficiencies

To supplement the overall data centre efficiencies, there are several metrics targeted at the major energy consuming components in the facility infrastructure. For buildings, the cooling system energy efficiency of water-cooled and air-cooled chiller systems (i.e. includes circulating pumps and cooling towers or heat rejection units) are typically specified in terms of power necessary to provide one ton of cooling, kW/RT (where 1 RT = 3.517 kW of cooling). In view of the N+1 requirement for

cooling systems in data centres to meet redundancy provisions, cooling systems for data centres will invariably have a higher kW/RT than equivalent building requirements, because an extra chiller will have to be run on hot standby.

A related metric is the Heating, Ventilation and Air-Conditioning (HVAC) System Effectiveness¹⁵. A low HVAC system effectiveness implies a relatively high HVAC system energy use per unit of IT equipment use, and therefore a high potential to improve HVAC system efficiency. Typical values range from 0.6 to 3.5, with a value of 1.4 being considered good.

Another cooling metric is the cooling system sizing factor, i.e. the ratio of installed cooling capacity to the peak cooling load. A high value here indicates the opportunity to run the right sizes of the cooling plant to meet the cooling demand under part-load requirements.

DATA CENTRE CERTIFICATION

Internationally, both the US and Europe have developed a form of performance based certification for data centres. The US has established the Leadership in Energy and Environmental Design (LEED) green building rating system as the national benchmark for Green Buildings. The LEED developed a version of the LEED rating system with adaptation credits specific to data centres¹⁶. These adaptation credits were drawn from the 2010 Guide on Environmental Performance Criteria (EPC) developed by Lawrence Berkeley National Labs (LBNL) for new constructions and existing data centres.²⁸ There is also an Energy STAR rating of data centres, which assesses how a building performs with respect to similar buildings, based on PUE data.

Europe has a voluntary scheme, the EU Code of Conduct (CoC), to bring stakeholders together to develop a set of best practices to improve energy efficiency. The Certified Energy Efficiency Data Centre Awards (CEEDA) developed by British Computer Society (BCS), the Chartered Institute of IT in the UK, complements the EU Code of Conduct. The UK Building Research Establishment's Environmental Assessment Method (BREEAM) is similar to the US LEED in covering all issues to mitigate the impact of a new or refurbished building on the environment by defining a performance target and assessment criteria that must be met. It has developed a 2010 Scheme Document (SD 5068) specific to data centres.

Locally, the IT Standards Committee (ITSC) Green Data Centres Working Group has developed a Singapore Standard (SS564: 2010) on Green Data Centres: Energy and Environmental Management System. The SS564 adopted a holistic approach. It is a world's first in integrating energy and environment management system together with performance metrics and recommended best

¹⁵ It is defined as the ratio of the annual IT equipment energy to the annual HVAC system energy.

¹⁶ U.S. Green Building Council (2010): USGBC/ LEED for Data Centres EB/NC.

practices. It is a certifiable standard that will help data centre operators to improve their overall energy and environmental performance and increase their competitive edge.

Certified Green Data Centres can gain recognition and credibility from their clients and industry partners with their good energy and environment management practices. This will not only open up market opportunities for these companies but also assist them to identify gaps and areas for improvements in the energy efficiency of their data centres. Finally, it can result in significant reduction of the companies' overall operating costs and help them do their part in reducing the carbon emissions. Since the standard was launched, seven data centres have achieved a significant milestone, in getting themselves certified to the Singapore Standard, SS564.

Besides developing the Green Data Centre Standard, SS 564, there is also a need to understand the measure of energy efficiency in a data centre. Hence, the Infocomm Development Authority (IDA) and Building and Construction Authority (BCA) have also jointly developed an additional Green Mark rating system for new and existing data centres. The Green Mark label, which has been awarded to its 1000th building in 2012, is a well-recognised rating system for commercial buildings in Singapore with a strong emphasis on energy performance. The new category of Green Mark for data centres will focus on recognising the performance of data centres for their efficiency in energy use, water use, environment quality and innovations for a greener data centres.

AREAS OF R&D FOR SINGAPORE

There is considerable expertise in Singapore on each of the key technological areas impacting the energy consumption in data centres and we should continue to nurture and strength this capability through further R&D work. While research in each of the technology areas mentioned in the above sections will boost energy efficiency in data centres, Singapore should focus on the intersections of the domain areas that addresses the critical needs of the industry, particularly taking into account characteristics of our local data centres and climatic conditions, to maximise Singapore's resources and to ensure competitiveness and prioritisation of our research efforts. **The three areas the authors of this primer recommend Singapore should focus on are:**

(1) Design of ICT Systems and Operations for a Tropical Environment (Higher ambient temperature and humidity):

The aim is to minimise the penalty for cooling requirements in our climate via a combination of reduction of thermal resistance along the heat flow path and optimisation of the server load that contributes to the cooling requirements in the first place. Another approach would be to look at the power architecture and reduce the harmonic currents so as to reduce the heat generated due to the unwanted harmonic currents. Finally, dynamic thermal management techniques that work synergistically across different layers (from architecture and software to cooling solutions) are necessary to effectively conquer this issue.

(2) Design for Flexibility in Operations, to cater for Modular build-up and Multi-generation server environment:

In a typical data centre, the load requirements are gradually built-up over several years. Within the lifetime of the data centre, it is also likely that multi-generation equipment with varying requirements will be co-existing. What should be the infrastructure and systems design to cater for such needs in an efficient manner, without the need for overdesign? Further, we need to identify the architecture that is able to match the varying application requirements as closely as possible. Identifying the right architecture can achieve significant savings in power and energy. A heterogeneous architecture may prove better than a pure homogeneous architecture found in modern data centres.

(3) Design for Resiliency, including the provision of sufficient Redundancy to ensure Reliability in operation:

Currently redundancy provisions do not take into account how the system should be configured so that it can provide a higher degree of efficiency while still delivering the same level of reliability and uninterruptibility in operation. Although redundancy is desirable from reliability point of view, it adds on to energy cost due to inefficient operation of the servers at light loads. Thus, a compromise between efficient operation and reliability has to be achieved. As we approach the dark silicon era, where power and thermal constraints will only allow a fraction of the cores on chip to be switched on, it becomes increasingly important to design customisable and reconfigurable solutions that can adapt to the changing workload to provide energy-efficient reliability.

The research needed for these focus areas may be classified in [Table 1](#):

Table 1: Focus areas for research

Focus Area (a): Design for Tropical Environment	Focus Area (b): Design for Flexibility in Operations	Focus Area (c): Design for Resiliency
<i>Short Term (3 – 5 years)</i>		
- Improvement in air management	- Analysis of existing facilities and modifications to ensure better demand matching	- Analysis of existing redundancy provisions and design philosophy
- Use of outdoor air	- Server application optimisation	- Modify operations to minimise overprovision.
- Robust server design	- Novel strategies in managing data, virtual machines, accessibility and resource sharing	- Optimise current UPS and back-up power
- Monitoring and sensing for real-time control		- Optimise back-up cooling system provisions
- Modelling and software tools for visualisation and energy management		- Advanced fault detection
- Cooling plant and		

system optimisation	- Advanced controller for power electronic converters to minimise harmonics	- Implement modular design philosophy	and prognostic algorithms
<i>Medium Term (5 – 10 years)</i>			
- Close-coupled and liquid cooling technologies	- Modularised data centre design. Standardisation of equipment.	- Hybrid AC/DC networks	- Alternative cooling configurations to provide for redundancy
- Inter-domain optimisation of cooling systems and ICT equipment	- Distributed energy/cooling network (including renewables)	- Energy internet node to eliminate low-frequency transformer	- Next generations UPS and back-up systems (including fuel cells and flow batteries)
- Dynamic thermal control	- Smart grids and intelligent control	- Mobile data centre for high-stake applications	
- Server consolidation and virtualisation			
- High temperature power electronics			
<i>Longer Term (> 10 years)</i>			
- Use of emerging low power server, storage & network technologies (including server-on-chip, non-volatile memory and optical switching networks)			
- Co-design of computing, power and cooling requirements in chip/server development			
- Design of a common IP-based control network for ICT, power and cooling infrastructure.			

The authors highlight the need for dedicated research in order to place energy efficiency at the forefront of data centre operations on a national scale. This is motivated by current trends and practices in the industry, which include: (1) the risk-adverse nature of the industry, (2) a high reliability operating environment backed up by sufficient redundancy, (3) the different nature of energy requirements between data centres and buildings, and (4) the high degree of interaction between the energy-consuming assets in data centres. Singapore is well-placed to set up a centre of excellence that can be an independent facility to demonstrate viable energy-saving solutions and facilitate its adoption by the industry; study various system configurations which enable data centre equipment to meet critical operational requirements without significant overprovision of resources; address the cooling strategies in data centres which minimises its disadvantages in a tropical environment, and develop expertise for inter-domain optimisation in the technologies that enable integration of the IT and facilities requirements.

CONCLUSION

Green Data Centre is an important area of activity for Singapore in view of the fast growth rate of the industry and the synergies it has with the other major sectors of the economy. The efforts at making the data centre, one of the highest energy intensity consumer in the country, more energy efficient can make a significant impact to reducing operating cost and carbon emissions. Singapore possesses a high proportion of professionals who are competent in this area, and has the potential to take a leading position in the research and development efforts to further push the boundaries of this field. In order to achieve this dominant position, there must be concerted investment in technologies that will provide the industry with the solutions that can enable it to stay on top of its competition. The proposed technologies that Singapore should focus on are: energy efficient data centre designs that are suitable for the tropical ambient conditions, system configurations that avoid the need for over-provisioning under flexible operating environments, and designs that achieve a good balance between resiliency and efficiency in operations.

Main Contributors:

Lead authors

Toh Kok Chuan and Dr Tseng King Jet, NTU

Dr Sanjib Kumar Panda and Dr Lee Siew Eang, NUS

Section Authors

Dr Wen Yonggang and Wong Yew Wah, NTU

Technical Writers

Dr Li Yanfei, Sahara Brahim; Energy Research Institute @NTU

Other Contributors

Dr Akash Kumar (NUS)

Dr. Akshay K. Rathore (NUS)

Dr Bharadwaj Veeravalli (NUS)

Dr Chandra Sekhar (NUS)

Choo Fook Hoong (NTU)

Dr Guan Yong Liang (NTU)

Dr He Bingsheng (NTU)

Irudhayasamy Vasantharaj Prasanna (NUS)

Dr Kaushik Dutta (NUS)

Dr Seri Lee (NTU)

Dr Mohan Guruswamy (NUS)

Dr. Souvk DasGupta (NUS)

Dr Tulika Mitra (NUS)

Dr Wang Peng (NTU)

Dr Zhang Wei (NTU)

Disclaimer, Limitation of Liability

This report represents the personal opinions of the contributors. The contributors, ERI@N, the National University of Singapore (NUS) and Nanyang Technological University (NTU) exclude any legal liability for any statement made in the report. In no event shall the contributors, ERI@N, NUS and NTU of any tier be liable in contract, tort, strict liability, warranty or otherwise, for any special, incidental or consequential damages, such as, but not limited to, delay, disruption, loss of product, loss of anticipated profits or revenue, loss of use of equipment or system, non-operation or increased expense of operation of other equipment or systems, cost of capital, or cost of purchase or replacement equipment systems or power.

Acknowledgements

The authors have benefited from comments from several colleagues from NUS and NTU as well as from the following governmental agencies: IDA, NCCS and NRF. Finally we thank Dr LI Yanfei and Ms Sahara BRAHIM (ERI@N) for their tireless effort in updating and consolidating the many versions of this Technology Primer.

The public version of this report was first published in July 2014. The contents of the primer reflect the views of the authors and not the official views of the government agencies. The publication of the primers is made possible by NCCS and NRF, and reproduction of the content is subject to the written consent of the authors, NCCS and NRF.

Hybrid AC/DC Systems

Ideally, AC sources are used for AC loads and DC sources are used for DC loads. This can be achieved by adding DC microgrids in local distribution systems, to build a hybrid DC and AC grid in distribution levels to couple DC sources with DC load and AC sources with AC loads.

In such a scheme, DC power generators such as photovoltaic (PV) panels and fuel cells are connected to DC networks through DC/DC boost converters. DC energy storages such as batteries and supercapacitors are connected to DC grid through bidirectional DC/DC converters. AC power generators such as wind turbine generators and small diesel generators are connected to AC network through transformers. AC energy storages such as flywheels are connected to AC grid through AC/AC converters and transformers (Figure A1).

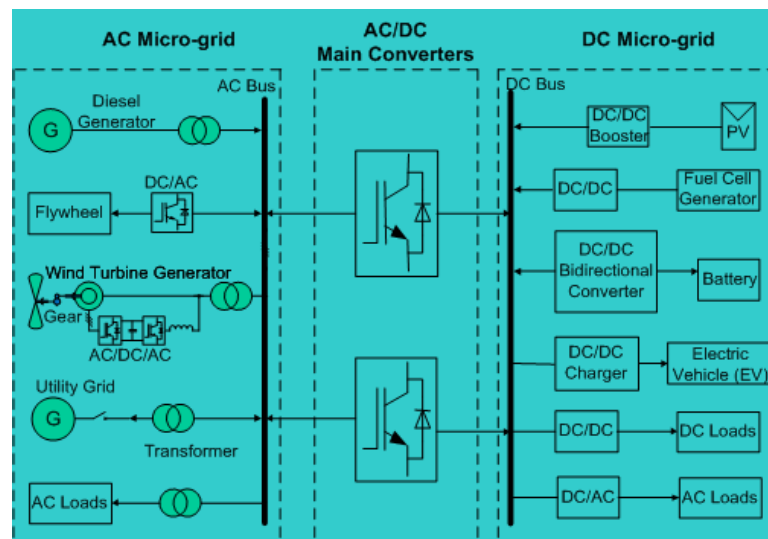


Figure A1: Schematic of an AC/DC hybrid microgrid

There are multiple advantages of a hybrid AC/DC power distribution system in improving energy efficiency:

- Elimination of unnecessary multi conversion processes to reduce total conversion loss.
- Elimination of embedded rectifiers for DC and AC loads in current AC grid to simplify equipment and cost reduction of electronic products.
- Connection of all DC loads to DC side of hybrid grid make it easy to control of harmonic into AC side through main converters, guaranteeing high quality AC in utility grid.
- DC grid can solve negative and zero sequence current problems caused by imbalance loads in AC distribution systems, and the neutral wire in sub-transmission may be eliminated and the related transmission loss reduced.

As conversions have been reduced to the minimum in the hybrid AC/DC system, a pioneering demonstration by LBNL showed energy savings of up to 10%.

Current barriers include high economic costs and the long time to upgrade or replace existing AC systems. In addition, all equipment which draws power from the system will need to be redesigned without embedded AC/DC rectifiers. New metering, protection and grounding equipments are also required to use this hybrid grid. A last technical barrier is lack of standard voltage for the DC grid.