Methods for Reducing Energy Consumption, Optimization in Operational Data Centers

Catalin Dumitrescu, Adrian Plesca, Maricel Adam, Costica Nituca, Alin Dragomir
"Gheorghe Asachi" Technical University of Iasi
21-23 Dimitrie Mangeron Blvd.
Iasi, Romania
e-mail: catalin.dumitrescu@tuiasi.ro

***Abstract*—This paper proposes a methods to reduce energy consumption in an operational data center facility. Data centers are the places where Information Technology and Communications equipment are hosted. From the day when the first computer was developed, the room that houses the mainframe faced with challenges regarding power and cooling requested, redundancy and availability for the systems to avoid any interruption, also with capacity planning for future expansions. All back-up and hot stand-by system are more often over-sized the focus are on availability nor on efficiency. The Information and Communication Technology (ICT) evolution, the need to compute and store more information lead to exponential increase of ICT equipment. The technology evolution lead to changes of equipment hosted in Data Centers with more performant systems but with high density of energy consumed. The virtualization of services is a way to increase computing performance with increasing utilization load of processors. Overall IT systems providing, the same functionality more efficient, consuming less power. Usually the evolution of ICT infrastructure (consolidation) is not correlated with facility changes that will impact the performance of existing data center. Increasing the power density in an operational computer room will produce a major impact to cooling system capacity and efficiency. The challenge is to avoid overheat areas (hot spots) and to provide same environment condition without changing the space geometry or design, increasing the efficiency and optimize energy consumption.**

Keywords—energy efficiency, data center; metrics; airflow optimization.

I. Introduction

A data center is the facility used to house Information Technology infrastructure and associated components as telecommunications, storages and support systems. Availability of the services are assured with redundant or back-up components regarding power, environment controls (cooling), security (access, fire suppression). Telecommunications Infrastructure Standards issued in 2005 the first standard (TIA-942) for data centers [1]. Data center is a complex facility with: rooms with ICT infrastructure, dedicated spaces for power systems (transport, back-up, and conversion), and cooling equipment to maintain the proper temperature and humidity conditions within facility [2]. The concept of data centers came into existence in 1940s in University of Pennsylvania with birth of the mainframe computer ENIAC (Electronic Numerical Integrator and Computer) [3]. Early data centers were mainly limited to space and technology, defense and other government institution and had highly restricted access. During the boom of microcomputer industry and internet, during 1980s, the IT operations started to grow in complexity. Companies needed to host servers, storages and networking equipment in specific rooms – enterprise data centers. The boom of data center came during the dot-com bubble (2000s), when the communications bandwidth increase and large data centers (Internet Data Centers) facilities were deployed, capable to provide non-stop operation and high availability services [4].

Data centers are energy-intensive facilities, being tens to hundreds more energy intensive than an average office building, with a continuous evolution and increasing energy demand. The virtualization and cloud technologies will determine a continuous evolution of data centers. Increasing of server's computing performance, storage capacity and bandwidth of communications will lead to an increase of energy density consumed in data center, a high demand of power in the same facility maintaining high availability and reliability of services [5]. Due to technology evolution, power density in computer room increasing from 10 W/m^2 in 1990 to 1000 W/m^2 in 2014 [6], some server racks being now designed for more than 30kW/cabinet (~10kW/m^2). It is estimated that in 2020 the electricity consumed in data centers (IT equipment, cooling and power systems, security etc.) will be +104TWh, double than energy for the year 2007 (56TWh) [7].

In order to provide high availability of services, data centers are inefficient with an architecture with many systems in parallel supplying the load, redundant paths, systems and sources [8]. The data centers initial design considers the assurance for enough capacity operation and future needs. On startup the power consumption will be low and inefficient, and a small percent from the grid power is needed by the IT systems. Only the cooling system maintaining environment condition of operations may account up to 40% average of the energy demands in a data center [9] and up to 25% are infrastructure losses [10].

The ICT sector, including data centers, generates up to 2% of the global CO_2 emissions [11], and data centers are estimated to have the fastest growing carbon footprint from

978-1-5386-5062-2/18/$31.00 ©2018 European Union

across the whole ICT sector [12]. The ICT sector consumes currently +7% of worldwide electricity and it will increase to 13% in 2030 [13]. Content traffic from on-line television or video-on-demand (VoD), internet broadcast, live streaming even traffic on mobile devices already reach 60% from all data traffic, and is estimated to be +80% in 2020 [14]. Only the data centers are estimated to consume +1.4% of the global electricity [15], and the compound annual growth rate (CAGR) of this consumption between 2007 - 2012 has been estimated as 4.4% (two times higher than the projected 2.1% increase in global demand from 2012 – 2040) [16].

Monitoring of all systems in the data center using objective metrics consists the key to improve the energy efficiency [17]. The efficiency metrics should be considered for upgrades, comparison of different data center facilities, for design and operation evaluation and costs calculation for services. Infrastructure efficiency metrics have to evaluate the energy consumption correlated with effective utilization for ICT:

PUE (Power usage effectiveness) is ratio between *Total data center energy* (total energy consumed from the source, including loses in power conversion and distribution, energy for cooling and support areas, lighting, security and ICT) and *Total energy used for ICT.*

DCiE (Data center infrastructure efficiency) is the inverse of PUE: *IT equipment energy/Total data center energy x100%*

The performance indicators quantified by efficiency metrics for a data center give the overall characteristics regarding quality of services and facility infrastructure efficiency.

II. Methods for reducing power consumption

Data centers are very dynamic due to rapid growth and change of technology. Data centers can now host more processing power into less space, high density computing environments can be an important impact on operational costs due to expanding of power demands, increasing overall power costs (secure) and excessive heat. The demand of data center energy is justified by IT system evolution, the computing performance increased by 25x but energy efficiency 8x only [18]. Considering the life span of ICT infrastructure and also the fast evolution and demand, the upgrade of the infrastructure is mandatory to take into account the energy efficiency and consumption for high performant equipment.

Methods to reduce power consumption in data centers are:

(a) IT equipment and infrastructure:

- *Refresh the oldest equipment with new high-efficient;*

New equipment are more powerful has more computing per watt, allow virtualization and integrate better power management software; Energy and software costs saving typically justifies a faster refresh rate of hardware.

- *Virtualize the application to increase server utilization;*

Most servers operate with very low utilization (5-15% on average), the consolidation can achieve the optimal load (~75%), better efficiency and saving space in data center.

- *Turn off unused servers, start them on-demand;*

Important part of infrastructure is for disaster recovery purpose, the servers are on line but unused for hot-standby (5-15%).

- *Improve servers' power management, avoid idle time;*

Better management for new IT systems, integrate systems with DCIM (Data center infrastructure management) software, stop unused infrastructure and start backup only when needed.

(b) Infrastructure / facility components:

- *Increase temperature set-point toward the high end of the ASHRAE range;*

IT inlet temperature is what matters, according with ASHRAE the range is 18-26ºC. Ensuring a good air management, this measure will save energy consumed from Chillers or CRAC Compressors. (Metrics to be monitored: RTI = return temperature index and RCI = rack cooling index).

- *Minimize UPS requirements and optimize utilization of elements in electrical and cooling chain to maximize efficiency (nominal/optimal load);*

High availability of infrastructure leads to double the paths and systems for increasing redundancy. Unnecessary redundancy leads to inefficiency, moreover due to many systems in parallel the load for the UPSs is low, even the best efficiency is over 75%. The recommendation is to minimize number and size.

- *Turn off continuous environment humidity control;*

The humidity control in a data center has to be measured and calibrated but often there are redundant system uncorrelated working inefficient.

- *Improve air management/optimizing cooling;*

Is important to avoid recirculation and by-pass of air that will produce mix cold and hot air. Air management is a big challenge, the proper ventilation will avoid hot spot and allow increasing the power consumption (high density computing) in asymmetric areas. Optimizing cooling can be achieved with the following measures: using high efficient CRAC solutions with economizer, free-cooling or reutilization of heat, installing Variable–Speed Drives on cooling system fans (20% air flow reduction results in ~50% saving in fan energy), install rack or zone (in-row) cooling units, implementing containment or dedicated room cooling (room-in-room). The challenge is to take right measure for cooling optimization, for this there are three methods recommended: deterministic with measurements time variation for all temperatures in data centers and rack consumption, perform a mathematical models and CFD (Computational Fluid Dynamics) models to test different solutions and scenarios, testing with IR (infrared) thermography camera.

The list of energy efficiency actions can be complete with specific measures related to infrastructure, the most important element is to measure correct the parameters and evaluate objective the efficiency of the facility. Also an important factor to be considered is the life cycle of elements and infrastructure, obviously the new facility has to consider from the design the energy performance to be achieved. Most data center facilities are filled up incremental, infrastructure and capacity will increase in time.

III. OPTIMISATION IN OPERATIONAL DATA CENTER

For the case study will analyze a medium size data center form Bucharest, designed in 2005, operational from 2006 with 100% availability. The data collected from different systems are stored in a database for analytics. The ICT equipment growth incremental in time (racks number and equipment) that lead to power consumption increase [17].

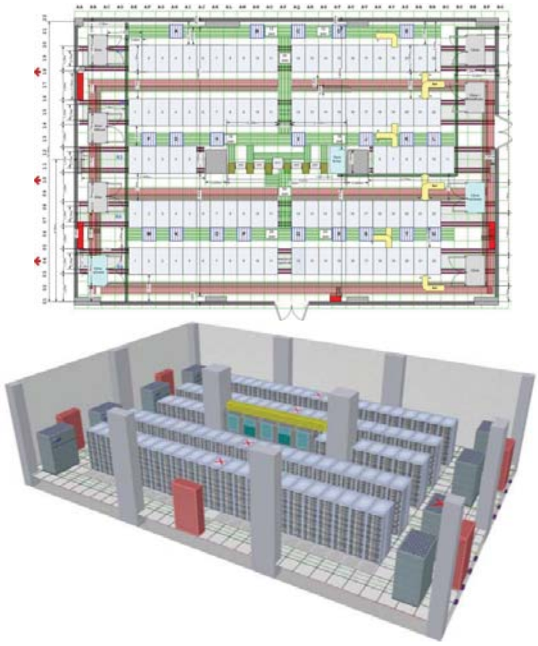

Fig. 1. Server room design layout (uniform IT rack cabinets).

The computer room design has the following details: space for hosting 100 standard IT rack cabinets and 5 open frames for network equipment, Ethernet and fiber optics interconnection (Fig. 1); room surface is 240 m^2; racks distribution (2.4 m^2/rack); average power/rack is 3.5kW/rack and power/area =2kW/m^2. The facility design is Tier 3 compliant (Uptime Institute), systems are redundant and concurrently maintainability with overall availability of 99.98% [17].

Only the central area with frames, cable interconnect and trays was installed from the beginning (Fig. 2). The computer room load increase in steps, in time with customers. In the data center are different models and size of cabinets, with unbalanced load density regarding power consumption. Even though there are difference from the initial design, the efficiency parameters value (PUE, DCiE) improve corresponding with the utilization ratio increase until the nominal capacity. The PUE desired by design is 1.6, for uniform maximum load, total (80% for UPS, 80% for cooling and for IT 75% load from full capacity (100 racks). In the initial design, energy consumption repartition is considered: 55% for ICT use, 35% cooling, and 10% others (5% infrastructure loses, lights, fire suppression, etc.).

In Fig. 3 are represented the PUE and DCiE efficiency parameters and electrical energy variation with IT load during 2006-2015. The PUE vary from most inefficient value of 2.3 (10% load in 2007) to designed value of 1.56 (90%load in 2015).

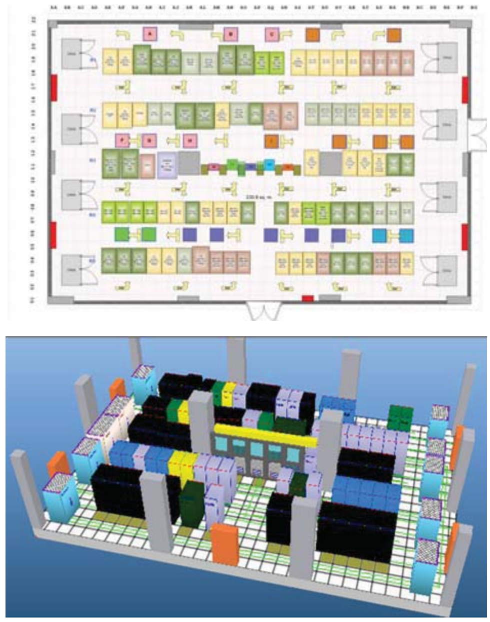

Fig. 2. Server room – real situation. (heterogenous IT cabinets).

The efficiency metrics variation of the in time due to the computer room load, demonstrated that the optimum values are obtained when the data center works close to 100% capacity in terms of IT load (space) and power consumption [17].

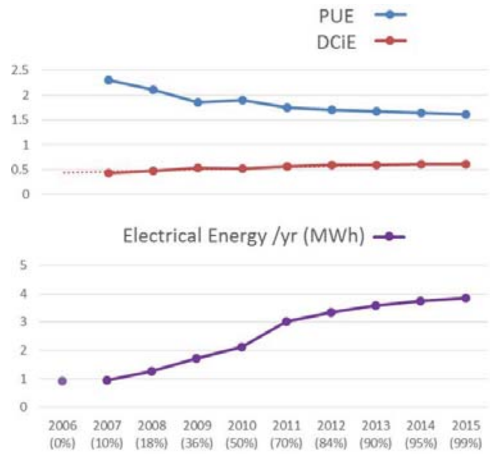

Fig. 3. PUE and DCiE, electrical energy graphs graphs due to load increase yearly (space load by IT equipment - percentage) [17].

During 2015-2017, due to IT demand for virtualization of 30% from installed infrastructure, the data center facility was upgraded in terms of power capacity (increase with 25%). Due to the server consolidation, in some zone power consumption increased that should produce hot spot areas. Air management and optimization methods have to be applied, with proper containment to avoid air recirculation and providing standard temperature parameters for ICT equipment.

To analyze changes in the computer room, a mathematical model was created using CFD (Computational Fluid Dynamic) software (Fig.4). In the computer room a DCIM (Data Center infrastructure Management) software was implemented, each rack have power and temperatures metered, collected by sensors with 1 min. interval. The graphs and reports were created with Cacti software. The level of details included in the model is high, after comparison of computing data with metered parameters the error is less than 5% which validate the simulations.

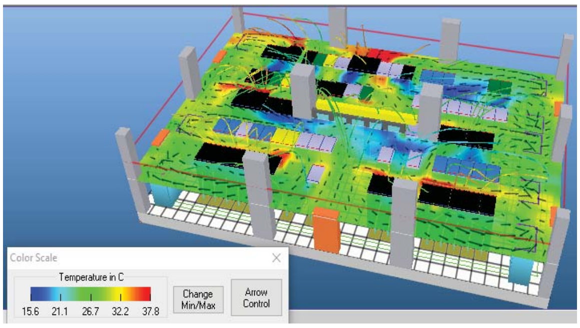

Fig. 4. Computer room CFD simulation.

As results of airflow optimization, the efficiency metrics in the data center show a considerable improvement (Fig. 5). In 2016 the PUE=1.72 (increase) due to inefficiency of cooling, hot spot occurred in row 5 as showed in CFD (Fig. 4). In 2017 with airflow management (partial containment), the PUE rich value of 1.33 and total power consumption also decreases.

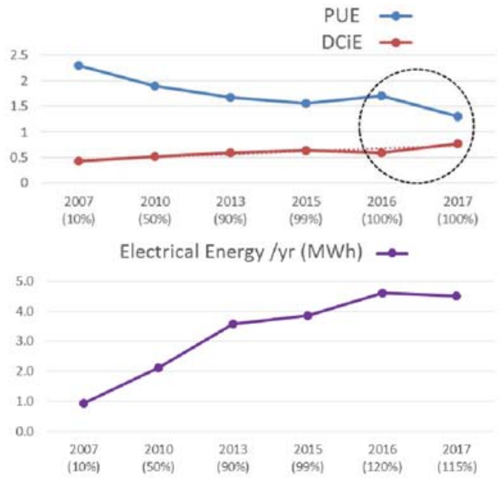

Fig. 5. PUE & DCiE metrics variation with optimization airflow.

IV. Conclusion

The overall data center facility energy efficiency assessment has to take into account all factors that influence the energy consumption.

The ICT technology evolution, changes and consolidation of IT infrastructure in an operational data center impacted the data center efficiency, overall PUE will increase to very inefficient value (+2). Specific measures have to be applied in existing facilities, starting from the energy chain (grid, transformer, transport, back-up, UPSs and cooling system) correlated with IT infrastructure evolution.

Due to the fast evolution of IT, in modern data centers changes are most often in computer room, so the dynamic data center model has to be considered during the facility life cycle.

The IT operational performance is increasing with energy efficient systems, but the overall data center efficiency parameters have to be corrected with special measures applied in the computer room to optimize the cooling capacity and performance. Cooling efficiency is a key factor to increase energy efficiency in a data center, air flow optimization is the most efficient way. The geometry in the computer room, load variation and changes can be simulated with special computer program. Using mathematical models for data center is the most effective method to reduce power consumption and improve the efficiency, to adopt the proper solution in computer room for cooling and air flow optimization.

The mathematical model simulating the dynamic environment in the data center, load variation and changes is the best method to analyze and compare different scenarios and solutions and should be the objective for future study.

References

[1] Telecommunications Industry Association (TIA), ANSI/TIA-942-2005 “Telecommunications Infrastructure Standard for Data Centers”.

[2] Y. Joshi, P. Kumar (ed.), “Energy Efficient Thermal Management of Data Centers. Introduction to Data Center Energy Flow and Thermal Management”, Springer, 2012.

[3] KW. Richey, “The Eniac”, Virginia Tech., 1997, http://ei.cs.vt.edu/~history/ENIAC.Richey.HTML.

[4] M. Owen et al., “Thermal guidelines for Data processing environments”, ASHRAE TC 9.9, Ashrae Datacom Series, Vol.1, Second Edition, 2012.

[5] E. Oró, V. Depoorter, A. Garcia, J. Salom, “Energy efficiency and renewable energy integration in datacentres. Strategies and modelling review”, Renewable and Sustainable Energy Reviews, 2015.

[6] W. Van Heddeghem, S. Lambert, B. Lanoo, D. Colle, M. Pickavet, P. Demmester, “Trends in worldwide ITC electricity consumption from 2007 to 2012”, Comput. commun., vol. 50, 64-76, Sept. 2014.

[7] R. Brown, “Report to Congress on server and data center energy efficiency”, Public Law 109-431, 2008.

[8] European Commission, JRC Technical Reports, “2018 Best Practice Guidelines for the EU Code of Conduct on Data Centre Energy Efficiency”, Ver.9.1.0., 2018.

[9] PW. Turner, “Dollars per kW plus Dollars per Square Foot are a Better Data Center Cost Model”, The Uptime Institute, 2006.

[10] A.H..Khalaj, T. Scherer, S.K. Halgamuge, “Energy, environmental and economical saving potential of data centers with various economizers across Australia”, Appl. Energy, vo. 183, 1528-1549, 2016.

[11] J. Ni, X. Bai, “A review of air conditioning energy performance in data centers”, Renew. Sustain. Energy Rev., Vol. 67, 625-640, 2017.

[12] UNFCCC, “ICT Sector Helping to Tackle Climate Change”, https://unfccc.int/news/ict-sector-helping-to-tackle-climate-change,2016.

[13] B. Whitehead, D. Andrews, A. Shah, G. Maidment, “Assessing the environmental impact of data centers part 1: Background, energy use and metrics”, Build. Environ. 82:151-159, 2014;.

[14] R. Sadler, “Video demand drives up global CO_2 emissions”, Climate News Network, Jan. 2017.

[15] P. Bertoldi, “A market transformation programme for improving energy efficiency in data centers”, ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA, USA, 2014.

[16] J. Koomey, “Growth in Data Center Electricity Use 2005 to 2010”, Analytics Press, Oakland, CA, USA, Aug. 2011.

[17] C. Dumitrescu, A. Plesca, “Overview on Energy Efficiency Parameters in Data Centers”, Proc. of EPE 2016, 153-156, Iasi, 2016.

[18] D. Sartor, “Data Center Energy Efficiency Best Practices”, training course, U.S. Department of Energy's Federal Energy Management Program, 2018