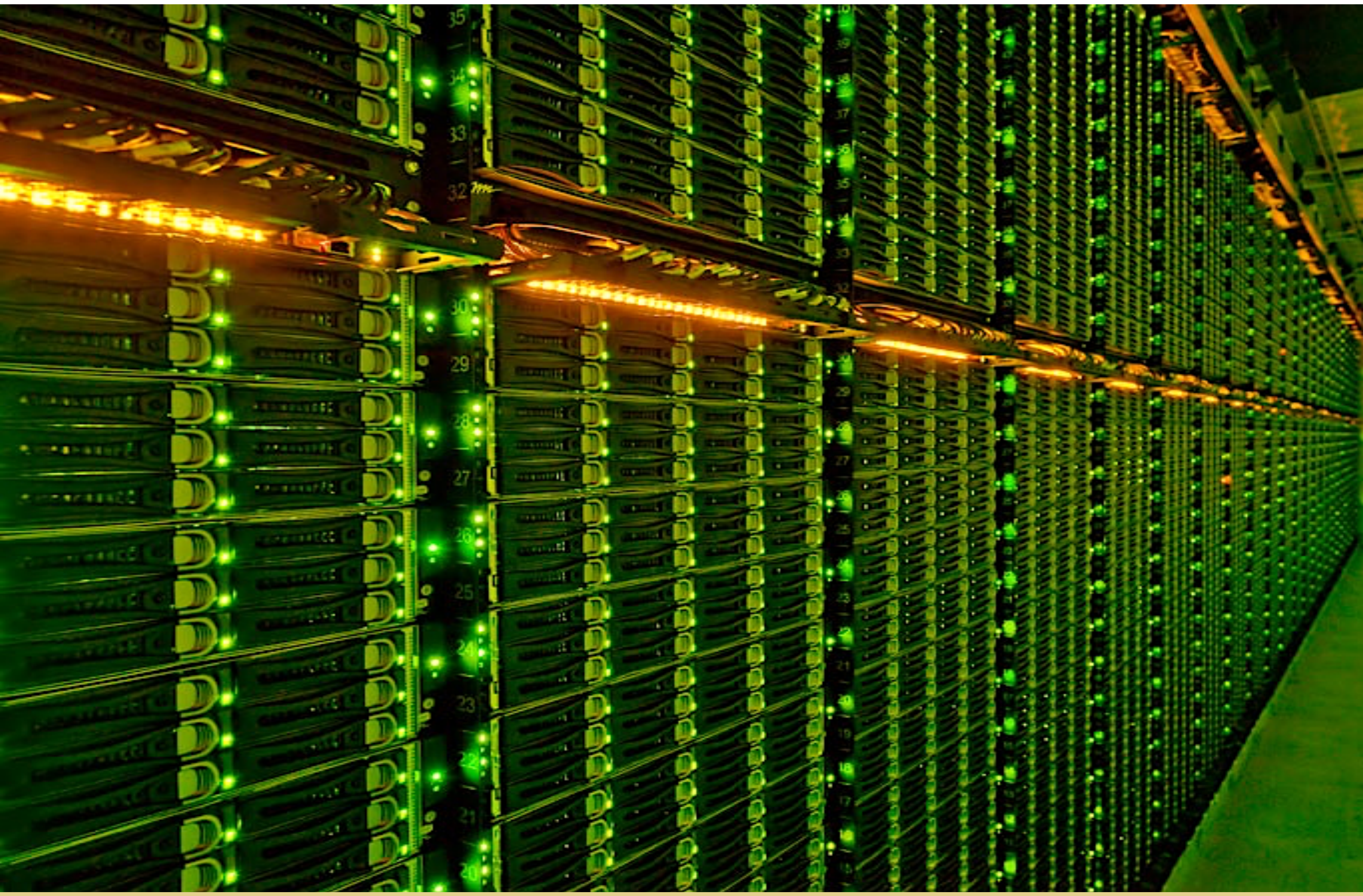


Thiago Lara Vasques

Load Management and Demand Response in Small and Medium Data Centers

PhD Thesis in Sustainable Energy Systems, supervised by Professor Pedro Manuel Soares Moura, submitted to the Department of Mechanical Engineering, Faculty of Sciences and Technology of the University of Coimbra

May 2018

· U C ·

UNIVERSIDADE DE COIMBRA

· U · C ·

UNIVERSIDADE DE COIMBRA

Load Management and Demand Response in Small and Medium Data Centers

by

Thiago Lara Vasques

PhD Thesis in Sustainable Energy Systems

in the framework of the Energy for Sustainability Initiative of the University of Coimbra and MIT Portugal Program, submitted to the Department of Mechanical Engineering, Faculty of Sciences and Technology of the University of Coimbra

Thesis Supervisor

Professor Pedro Manuel Soares Moura

Department of Electrical and Computers Engineering, University of Coimbra

May 2018

This thesis has been developed under the Energy for Sustainability Initiative of the University of Coimbra and been supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil).

Acknowledgements

First and foremost, I would like to thank God for coming to the conclusion of this work with health, courage, perseverance and above all, with a miraculous amount of love that surrounds and graces me.

The work of this thesis has also the direct and indirect contribution of many people, who I feel honored to thank.

I would like to express my gratitude to my supervisor, Professor Pedro Manuel Soares Moura, for his generosity in opening the doors of the University of Coimbra by giving me the possibility of his orientation when I was a stranger. Subsequently, by his teachings, guidance and support in difficult times. You, Professor, inspire me with your humbleness given the knowledge you possess.

A formal acknowledgement is also due to Institute of Systems and Robotics of the University of Coimbra (ISR-UC), Energy for Sustainability Initiative, MIT Portugal Program and specially to Federal University of Goiás and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES), for their support, which allowed me pursuing my Ph.D.

I am deeply thankful to my wife, who besides loving me all the time, has been my Ph.D. companion in this adventure that was to leave our country and live abroad changing our lives completely. Her encouragement in difficult periods, fruitful discussions and invaluable advices, allowed me to move forward in some stages of the Ph.D study. I love you!

I am very grateful to my family for their love, support and sacrifices by the time away and full of homesickness; to my parents, who always believed in me and invested all their energies in my education; to my brothers, who are my pride for being exactly who they are; to my niece, who renewed my days with a love that I did not know; to my sisters-in-law for the love and support; to the family of my wife, who is now my family and has been so precious; and to all my relatives, who allowed me to be away, but always near.

I owe a great debt of gratitude to the family that I built in Portugal, this country that is in my heart now. To the countless and unforgettable moments of love spent in paradise, at Quinta do Outeiro, beside my Mindinha, who is my Portuguese mother, my Portuguese and inspiring brother Miguel, my dear Juliana and my real squiggly Lana; to my Ph.D colleagues, who became my beloved friends in this journey, Lurian and Zuleika. Every moment we uniquely shared together will be deeply kept in my heart; to Joanna, Merielly and Kika Maria, my dear friends from Porto, who have given me so many moments of joy and happiness; to Francisco and Ruben, our dear friends from Lisbon, who have always treated me with the best of Portugal; to special meetings, with special people, who crossed my path in Portugal.

I would like to also thank all colleagues and professors from DEEC and DEM of FCTUC, and I want to express a special gratitude for the valuable help offered by dear Thiago Luz.

Finally, I would like to thank Coimbra and Portugal for the unforgettable days I spent knowing, living and feeling the best of this wonderful country.

ABSTRACT

Data centers are the backbone of a growing number of activities in modern economies. However, the large increase of digital content, big data, e-commerce, and Internet traffic is also making data centers one of the fastest-growing users of electricity. The total energy consumption of data centers corresponded to almost 1.5% of the global electricity consumption and has an approximated annual growth rate of 4.3%. Therefore, it is very important to increase the energy efficiency in data centers with actions such as power usage management, server consolidation, energy efficient components and systems, as well as demand response programs and renewable energy sources.

Small and medium data centers account for more than 50% of the total electricity consumption in this sector. In fact, surveys indicate that this data center profile waste more energy than larger facilities. Nevertheless, existing studies tend to be focused on the energy related issues for large data centers rather than small and medium data centers. Thus, this work aims to present how an intensive energy consumer, such as small and medium data centers, can become more efficient from the energy point of view and how they can take advantage of demand response programs to decrease costs and to cooperate with the grid to ensure higher reliability and sustainable development goals.

For this purpose, a set of actions have been taken in pursuit of this objective. Firstly, a meticulous state-of-the-art literature review of data centers energy efficiency and demand response perspectives was performed, providing fundamental information about the complex and technological world of small and medium data centers, in order to support the knowledge of their flexible load.

An energy efficiency survey directed at Brazil, Portugal and the United States was also conducted, showing an alarming reality regarding energy efficiency practices within small and medium data centers. In the same context, an impact and potential assessment on energy efficiency towards small and medium data centers was performed allowing to observe the main advantages, disadvantages and applications in three important methodologies.

A design of load management strategies and opportunities for demand response was developed considering a framework with two layers encompassing an energy efficiency methodology, the management of flexibility and the deployment of demand response scenarios with dynamic tariffs or an incentive-based contract. Alongside, the small and medium data centers load mathematical modeling and specification of optimization processes were provided within the same scope with two algorithms implemented with different approaches; one focused on small and medium data centers operators, and other dedicated to distribution system operators. At the data center level, a mixed-integer linear programming optimization technique is used to control flexible loads (considering ICT workload, cooling and UPS), in order to reduce the cost function in a price-based outline or to match the load reduction requirements in an incentive-based outline. Concerning the distribution system operator, a random-rotation and fairness algorithm selects the data centers to be used in each demand response event. Such algorithm deals with small and medium data centers in an aggregated and equitable way by drawing on their joint and flexible loads in contractual terms, as if they were one large data center.

Finally, the optimization results were firstly demonstrated in a small and medium data centers perspective, as well as their respective scenarios, in which it has been proven the algorithm operation and reliability. The potential for cost savings in demand response was also proved, being achieved with the considered incentives in the simulation process, savings of 1.33% for small and 5.15% for medium data centers in the incentive approach, and 0.21% for small and 0.68% medium data centers in the dynamic tariff approach. Subsequently, the same premises were utilized with focus on the distribution system operator point of view, enabling to predict specific contractual policies that can be adopted in this type of relationship through the best and the worst scenarios simulations. The scenario with a preponderance of incentives stimulates the adoption of demand response programs applied to small and medium data centers, nevertheless the unchanged and penalty cases enable to forecast an unfavorable scenario for this sort of demand response program. In this context, on one hand, the one single day analysis showed a variation in the reduction potential between 18.02% and 91.16%, while the highest value in the penalty profile was 356 € and in the incentive profile 10,580 €. The 22 business day analyses presented a fluctuation in the reduction potential in the range of 13.12% and 71.38%, whilst the highest value in the penalty profile was 1,200 € and in the incentive profile 84,700 €.

Keywords: Data Centers, Information and Communication Technologies, Energy Efficiency, Load Management, Demand Response

RESUMO

Os centros de dados são a base de um número crescente de atividades na economia moderna. No entanto, o grande aumento de conteúdo digital, *big data,* comércio eletrónico e tráfego da Internet está a tornar os centros de dados um dos utilizadores de eletricidade com um crescimento mais acentuado, com um consumo de energia correspondente a quase 1,5% do consumo global e com uma taxa de crescimento anual de aproximadamente 4,3%. Portanto, é de extrema importância aumentar as iniciativas de eficiência energética em centros de dados, direcionadas para a gestão do uso de energia, a consolidação de servidores, utilização de componentes e sistemas energeticamente eficientes, bem como a adoção de programas de resposta da procura e a utilização de fontes de energia renováveis.

Os pequenos e médios centros de dados representam mais de 50% do consumo total de energia neste setor. De fato, pesquisas indicam que esse perfil de centros de dados desperdiça mais energia do que as instalações de maior dimensão. No entanto, os estudos existentes tendem a estar focados nos aspetos do consumo de energia em grandes centros de dados, em vez dos pequenos e médios. Desse modo, este trabalho visa avaliar como um consumidor intensivo de energia, como os pequenos e médios centros de dados, se pode tornar mais eficiente do ponto de vista energético e ao mesmo tempo tirar proveito dos programas de resposta da procura para diminuir custos e cooperar com a rede de energia elétrica, garantindo maior fiabilidade e objetivos de desenvolvimento sustentável.

Para cumprir tal propósito, foi desenvolvido um conjunto de ações para alcançar esse objetivo. Em primeiro lugar, foi realizada uma meticulosa revisão bibliográfica abordando a eficiência energética e a resposta da procura no contexto dos centros de dados, fornecendo informações fundamentais sobre este complexo mundo tecnológico, de modo a permitir uma melhor compreensão da sua carga flexível.

Além disso, foi realizado um inquérito acerca de eficiência energética direcionada ao Brasil, a Portugal e aos Estados Unidos, mostrando uma realidade alarmante em relação às práticas de eficiência energética nos centros de dados de pequena e média dimensão. No mesmo contexto, foi desenvolvida uma avaliação do impacto e do potencial de eficiência energética nos pequenos e médios centros de dados, permitindo observar as principais vantagens, desvantagens e aplicações considerando três importantes metodologias.

Foi proposto um modelo, com duas camadas, para o planeamento de estratégias de gestão de cargas, que engloba oportunidades direcionadas à resposta da procura. Este abrange uma metodologia de eficiência energética, a gestão de cenários de flexibilidade e a implementação por meio de tarifas dinâmicas, ou por um contrato baseado em incentivos. Paralelamente, foram modeladas matematicamente as principais cargas dos pequenos e médio centros de dados e foi feita a especificação dos processos de otimização com dois algoritmos implementados com diferentes abordagens; um focado em operadores de pequenos e médios centros de dados e outro dedicado aos operadores de sistemas de distribuição. Em relação aos centros de dados, foi utilizada uma técnica de otimização de programação linear inteira mista para gestão das cargas flexíveis (considerando cargas de tecnologias da informação e comunicação, climatização e UPS) a fim de minimizar os custos de energia com uma abordagem direcionada à tarifa ou com um contrato baseado em incentivos. Já ao analisar a perspetiva do operador do sistema de distribuição, foi implementado um algoritmo de rotação aleatória e seleção justa dos centros de dados a serem usados em cada evento de resposta da procura, de forma agregada e equitativa, recorrendo às suas cargas flexíveis de forma conjunta em termos contratuais, tal como se tratasse de um grande centro de dados.

Por fim, os resultados da otimização foram em primeiro lugar demonstrados na perspetiva dos pequenos e médios centros de dados, bem como os cenários respetivos, em que ficou demonstrado o funcionamento e a fiabilidade do algoritmo. O potencial de poupança na abordagem de resposta da procura baseada em incentivos foi de 1,33% para pequenos centros de dados e 5,15% para médios. Já na abordagem por tarifárias dinâmicas os pequenos centros de dados alcançaram 0,21% e os médios, 0,68%. Posteriormente, as mesmas premissas foram utilizadas com foco no ponto de vista do operador do sistema de distribuição, possibilitando simular políticas contratuais específicas que podem ser adotadas nesse tipo de relação de consumo. Neste contexto, por um lado, a análise de um único dia mostrou uma variação no potencial de redução de potência entre 18,02% e 91,16%, enquanto o valor mais alto no perfil de penalização foi de 356 € e no perfil de incentivo 10,580 €. As análises que consideraram 22 dias úteis apresentaram uma variação no potencial de redução de potência no intervalo de 13,12% e 71,38%, enquanto que o valor mais elevado no perfil de penalizações foi de 1.200 € e no perfil de incentivo 84.700 €.

Palavras-chave: Centros de Dados, Tecnologia da Informação e Comunicação, Eficiência Energética, Gestão de Cargas, Resposta da Procura

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AC	Air Conditioning
ACK	Acknowledgment
APC	Adaptability Power Curve
APCren	Adaptability Power Curve at Renewable Energies
ASM	Ancillary Service Market
BEAMs	Building Environmental Assessment Methods
CHP	Combined Heat and Power
CM	Capacity Market
CPP	Critical Peaking Price
CPU	Central Processing Unit
CRAC	Computer Room Air Conditioners
CREW	CPU and RAM Energy Aware
CS	Curtailable Service
CUE	Carbon Usage Effectiveness
DB	Demand Bidding/Buy Back
DC	Direct Current
DC-EEP	Data Center Energy Efficiency and Productivity
DCA	DCAdapt
DCE	Power Usage Data Center Efficiency
DCEP	Data Center Energy Productivity
DCiE	Data Center infrastructure Efficiency
DCIM	Data Center Infrastructure Management
DDR3	Double Data Rate 3
DDR4	Double Data Rate 4
DDR5	Double Data Rate 5
DLC	Direct Load Control
DoE	Department of Energy
DPPE	Data Center Performance Per Energy
DR	Demand Response
DRAM	Dynamic Random-Access Memory
DSM	Demand Side Management
DSO	Distribution System Operators
DVFS	Dynamic Voltage and Frequency Scaling
EDE	Electronics Disposal Efficiency
EDR	Emergency Demand Response
EES	Energy ExpenseS
ENTSO-E	European Network of Transmission System Operators for Electricity
EP	Energy Provider
EPC	Energy Proportionality Coefficient
ERE	Energy Reuse Effectiveness
ERE	Energy Reuse Effectiveness
ERF	Energy Reuse Factor
ESDs	Energy Storage Devices ESDs

FERC	Federal Energy Regulatory Commission
FVER	Fixed-to-Variable Energy Ratio
FVER	Fixed to Variable Energy Ratio
GEC	Green Energy Coefficient
GHG	Greenhouse Gases
GUF	Grid Utilization Factor
HVAC	Heating, Ventilation, and Air Conditioning
I/O	Input and Output
ICT	Information Communication Technologies
IoT	Internet of Things
ISO	Independent System Operators
ITC	IT Customer
LBNL	Lawrence Berkeley National Laboratory
LCA	Life Cycle Assessment
LP	Liner Programming
MDC	Medium Data Center(s)
MIBEL	Iberian Electricity Market
MILP	Mixed Integer Linear Programming
NACK	Negative Acknowledgement
NIST	National Institute of Standards and Technology
NRDC	Natural Resources Defense Council
OP	Operations Research
PDE	Power Density Efficiency
PE Savings	Primary Energy Savings
PMSM	Power Monitor System and Management
PUE	Power Usage Effectiveness
QoS	Quality of Service
RCI	Rack Cooling Index
RDRAM	Rambus DRAM
RE	Rebound Effect
REF	Renewable Energy Factor
RHI	Return Heat Indexes
ROI	Return of Investment
RTI	Return Temperature Index
RTP	Real-Time Pricing
SCE	Server Compute Efficiency
SDA	Supply Demand Agreements
SDC	Small Data Center(s)
SDRAM	Synchronous DRAM
SHI	Supply Heat Indexes
SI-EER	Site Infrastructure Energy Efficiency Ratio
SLA	Service Level Agreement
SMDC	Small and Medium Data Centers
sPUE	System Power Usage Effectiveness
SSD	Solid State Drives
TCI	Thermal Correlation Index

TE	Thermodynamic Efficiency
TOU	Time-of-Use
TSO	Transmission System Operators
TSP	Travelling Salesman Problem
U.S.	The United States of America
UPS	Uninterruptible Power Supply
VMs	Virtual Machines
WSOA	Services Outsourcing Agreements
WUE	Water Usage Effectiveness
DEEC	Department of Electrical and Computer Engineering
UC	University of Coimbra
CoP	Coefficient of Performance

LIST OF MATHEMATICAL PARAMETERS

Sets	
$\mathcal{J}$	Set of DR time slots
$\mathcal{N}$	Set of SMDC, or ICT workload, or crack, or UPS units
$\mathcal{T}$	Set of time slots
Index	
i	Index of SMDC, or ICT workload, or crack, or UPS units, or server class, or device, array
j	Index of DR time slots, or space type, or device, or array
t	Index of time slots
Parameters	
a	Percentage of energy efficiency improvement of an ENERGY STAR server relatively to a "standard" unit
BW_{rw}	Read or write bandwidth
b	Percentage of energy efficiency improvement of an ENERGY STAR storage relatively to a "standard" unit
b_{ij}^{OFF}	Y-intercept of power-utilization function (DFVS disabled) for server class i in space type j
b_{ij}^{ON}	Y-intercept of power-utilization function (DFVS enabled) for server class i in space type j
C	Capacitance of the circuit
c_{min}	Minimum capacity level of the battery
c^t	UPS capacity level
c_{total}	Total capacity of the battery
CoP	Coefficient of Performance
D	Total DRAM channels
d	Constant factor
d_n	Total number of memory fetched in the storage disk
DoD	Depth-of-discharge
E_{AP}	Energy required to activate and pre-charge
E_{cost^j}	Energy cost during the DR event in time slot j
$E_{cost^{j+1}}$	Energy cost in a rebound effect situation in time slot j+1
E_{cost^t}	Baseline energy cost in a normal time slot t
$E_{dis^j}(DoD)$	Amount of UPS energy that can be discharged at the depth-of-discharge level in a DR event j
$E_{rec^{j+1}}(DoD)$	Amount of UPS energy that can be recharged at the depth-of-discharge level in a rebound effect process j+1
E^{DC}	Data center electricity demand
E_{disk}	Disk energy
E_{DRAM}	Dram energy
E_{cost^j}	Reduction of energy cost in a DR event in time slot j
$E_{cost^{j+1}}$	Reduction of energy cost in a DR event in time slot j+1
E_{icost^j}	Incentive in a normal time slot t

$E_{totalcost^{dt}}$	Total energy cost in a dynamic tariff approach
$E_{totalcost^{ic}}$	Total energy cost in an incentive-based contractual approach
$EM_{baseline}$	Efficiency metric for baseline server
$EM_{baseSB(j)}$	Efficiency metric for baseline storage equipment in array j
EM_{baseSE}	Efficiency metric for baseline storage equipment
EM_{EE}	Efficiency metric for efficient server
$EM_{EESB(j)}$	Efficiency metric for baseline storage equipment in array j
EM_{EESE}	Efficiency metric for energy-efficient storage equipment
E_{maxdis}	Maximum energy that can be discharged
E_{maxrec}	Maximum energy that can be charged
E_{net}	Networking energy
E_j^N	Electricity used by network devices in space type j
E_{ij}^S	Electricity used by servers of class i in space type j
E_j^{ST}	Electricity used by external storage devices in space type j
E_{rw}	Energy per read or write
$ě_{ij}^S$	Baseline annual electricity use per server of class i in space type j
$ě_j^{ST}$	Baseline annual electricity use per external storage device in space type j
e_{jk}^I	Ratio of electricity use by infrastructure system component k in space type j to ICT device electricity use in space type j
e_{ij}^S	Annual electricity use per server of class i in space type j
ES	ENERGY STAR server number
EUL	Expected useful life based on ICT upgrade cycle of data center
f	Frequency
$f(dt)$	Objective function based on dynamic tariffs
$f(ic)$	Objective function based on incentive-based contracts
f_{AP}	Frequency required to activate and pre-charge
$f_{baseSE(j)}$	Fraction of total TB stored on a baseline device/array j
$f_{EESB(i)}$	Fraction of total TB stored on energy-efficient device/array i
f_{max}	Maximum frequency
H	Amount of heat removed
lb	Lower boundary constraint
m_{ij}^{OFF}	Slope of power-utilization function (DFVS disabled) for server class i in space type j
m_{ij}^{ON}	Slope of power-utilization function (DFVS enabled) for server class i in space type j
$Ň_{ij}^S$	Baseline number of servers of class i installed in space type j
$Ň_j^{ST}$	Baseline number of external storage devices installed in space type j
N_{sw}	Average number of circuit switches per clock cycle
n_S	Switching number
P	Power consumption
P_{active}	Active state power
$P_{alloc_i^j}$	Allocated computational capacity to attend a DR event in a SMDC i in time slot j
$P_{baseline}$	Power draw of baseline servers
P_{crac_i}	Power consumption of CRAC units
P_{cool^j}	Potential of power reduction in a cooling DR event in time slot j

$P_{cool^{j+1}}$	Potential of power reduction in a cooling RE situation in time slot j
P_{EE}	Power draw of new efficient server equipment
P_{EESE}	Power draw of new energy-efficient storage equipment
$P_{ENERGY\ STAR}$	Power draw of ENERGY STAR server
$P_{ES\ STOR}$	Power draw of ENERGY STAR storage
$P_{ES,full\ load}$	Power draw of ENERGY STAR server at full load
$P_{ES,idle}$	Power draw of ENERGY STAR server at idle
P_{fan}	Power of CRAC fan
P_{idle}	Servers idle power
P_{ict}	Power consumption of ICT with their heat flow directed towards the CRAC unit
P_{max}	Maximum CPU power consumption
P_{peak}	Servers peak power
$P_{Post\ Ds\ Man}$	Total power draw of data storage after data storage management tools are implemented and after efficient data storage equipment is installed
$P_{Pre\ Ds\ Man}$	Total power draw of data storage before data storage management tool measures implemented (or with tool turned off) and after efficient data storage equipment is installed
$P_{sa,full\ load}$	Power draw of a single-application server at full load
$P_{sa,idle}$	Power draw of a single-application server at idle
$P_{smdc_i^t}$	Total SMDC power consumption i in time slot t
$P_{smdcdec_i^j}$	Decreased SMDC power consumption in a DR event j
$P_{smdcinc_i^{j+1}}$	Increased SMDC power consumption in a RE situation j+1
$P_{standby}$	Low power state power
$P_{total_i^t}$	Total computational power of SMDC i in time slot t
P_{twflex^j}	Total flexible workload in all SMDC's within time slot j
P_{ups^j}	Power reduction in a UPS DR event in time slot j
$P_{ups^{j+1}}$	Power increase in a UPS RE situation in time slot j
$P_{vh,full\ load}$	Power draw of a virtual host server at full load
$P_{vh,idle}$	Power draw of a virtual host server at idle
$P_{wflex_i^j}$	Flexible workload assigned to SMDC i in time slot j
$P_{wflex_i^{j+1}}$	Flexible workload assigned to SMDC i in time slot j+1
$P_{w\ Virt}$	Total power draw of all virtual hosts
PUE_i^t	Power usage effectiveness (PUE) in a SMDC i in time slot t
PUE_{Summer}	Average PUE over the summer peak demand period
PUE_{Winter}	Average PUE over the winter peak demand period
QoS	Quality of service
S_{max_i}	Total number of servers in data center i
$S_{select_i^j}$	Number of selected servers in data center i to attend a DR event in time slot j
S_{total^j}	Total number of selected servers in all SMDC data center in time slot j
sa	Single application servers, numbered 1 to n
t_{access}	Access time
t_{active}	Active state time
T_{adj}	Temperature of adjustment

t_{disk}	Disk time
t_{DRAM}	DRAM time
T_{max}	Maximum temperature of the server inlets
t_{miss}	Miss time
t_{net}	Network time
t_{RL}	Rotational latency
T_{safe}	Maximum permitted temperature at the server inlets
t_{seek}	Seek time
$t_{standby}$	Low power state time
T_{sup}	Temperature of the air supplied by CRAC units
$t_{switching}$	Switching time
$t_{transfer}$	Transfer time
t_{tt}	Transfer time from disk to higher level cache
ts	Time slot per hour
U_{ES}	Utilization of ENERGY STAR server
u_{ij}	Post-reduction processor utilization per server of class i in space type j
$\check{u}_{ij}$	Baseline processor utilization for active servers of class i in space type j
U_{sa}	Average utilization of a single-application server over the year
U_{vh}	Average virtual host server utilization over the year
ub	Upper boundary constraint
V_{dd}^2	Supply voltage of CPU
vh	Virtual host servers number
W	Amount of work necessary to remove the heat
α_{ij}^S	Fraction of servers of class i in space type j with energy efficient hardware
α_j^{ST}	Fraction of energy efficient external storage devices in space type j
β_{ij}^S	Fraction of servers of class i in space type j with dynamic voltage scaling enabled
γ_{ij}^S	Ratio of efficient server to baseline server electricity use for servers of class i in space type j
γ_j^{ST}	Ratio of efficient external storage device to baseline external storage device electricity use in space type j
$\delta_{ij}', \delta_{ij}''$	DFVS and utilization factors
ε_j^N	Ratio of network device to total ICT device electricity use in space type j
θ_{ij}^S	Baseline fraction of servers of class i in space type j that are legacy servers
α	Use factor percentage to a DR event
η	Efficiency
λ^t	Incentive given to participate in a DR event in the j time instant
ρ_{ij}^S	Device reduction ratio for servers of class i in space type j
ρ_j^{ST}	Device reduction ratio for external storage in space type j
σ_i^t	Average servers' utilization in a SMDC i in time slot t
φ^t	Electricity price in a time slot t

CHAPTER 1

INTRODUCTION

The availability of Internet and the relevance that Information and Communication Technology (ICT) has in modern society is changing the way in which computing resources are typically provisioned and allocated, where the computing infrastructure itself is provided as a service to its users. In a not very distant past, data were generated and communicated primarily among ICT systems – albeit of diminishing size. In the future, data-producing systems will increasingly involve small, low-power sensors and actuators embedded in the physical world – a network of cyber-physical systems, also referred to as the Internet of Things (IoT) (SIA 2015). However, the increasing demand of computing resources has brought an inevitable growth in the energy consumption associated with this infrastructure, fostering a set of ICT to reduce the environmental impacts called Green ICT (Craig-wood *et al.* 2010; Uddin; Rahman 2012).

Jiang et al. (2015) globally conceptualize a data center as a facility used to house enterprise's ICT equipment, such as servers, telecommunications, and storage systems, including also supporting infrastructures of high quality power delivery and cooling systems. More specifically, Pierson (2015) address data centers as being structures, or group of structures, dedicated to the centralized accommodation, interconnection and operation of ICT and network telecommunications equipment, providing data storage, processing and transport services, along with all the support facilities for high quality power supply and environmental control with the levels of resilience and security required to provide the desired service availability, as shown by Figure 1.1 in a more detailed and disaggregated flowchart.

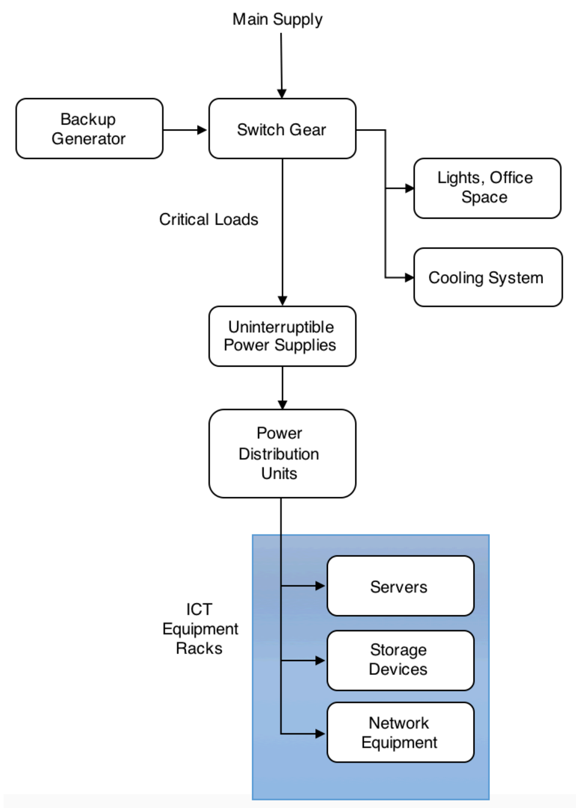
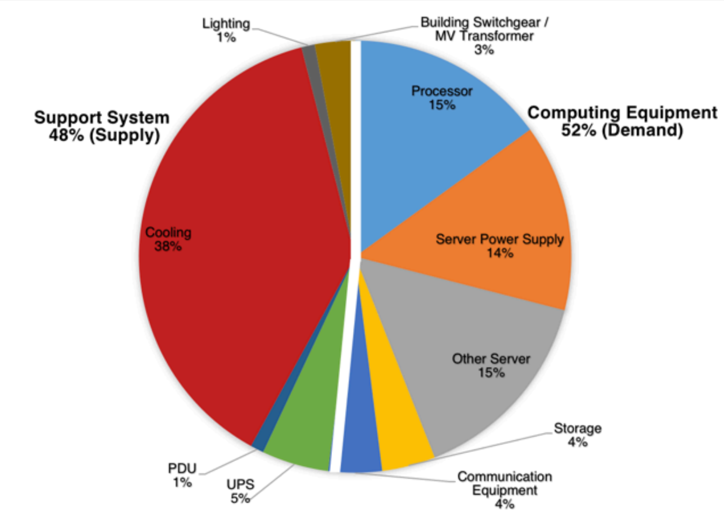


Figure 1.1 – Typical electrical and ICT components in a data center (Ghatikar et al. 2010)

This particular infrastructure is divided in three spaces: ICT room, data center support area and ancillary spaces (Rong *et al.* 2016). The ICT room is an environmentally controlled space that houses equipment and cabling directly related to computer and telecommunications systems which generate considerable amounts of heat. Moreover, the ICT equipment is highly sensitive to temperature and humidity fluctuations, so a data center must keep restricted environmental conditions for assuring the integrity and functionality of its hosted equipment. Data centers support areas are all those where different systems, such as the Uninterruptible Power Supply (UPS) systems, cooling control system and switch boards are located. Finally, the ancillary spaces include mainly offices, lobby and restrooms (Grice *et al.* 2013), (Oró *et al.* 2015).

Nowadays, data centers are the backbone of contemporary economies, having different profiles, such as server rooms that power small-to medium-sized organizations, enterprise data centers that support large corporations and server farms that run cloud computing services hosted by major market players.

Based on Ghatikar *et al.* (2010) data centers can include more than 100,000 hardware devices and the electrical load can range from about 1 kW to about 100 MW with different sizes and profiles. A typical example in terms of demand and supply, as well as the power draw unbundling of a medium data center is categorized in Figure 1.2.



Category	Power Draw
Computing	588 kW
Lighting	10 kW
UPS and distribution losses	72 kW
Cooling power draw for computing and UPS losses	429 kW
Building switchgear/ MV transformer/ other losses	28 kW
Total	**1,127 kW**

Figure 1.2 – Analysis of a typical 465 m² data center (Emerson 2015)

In the context of this huge load spectrum, according to Sheppy *et al.* (2011) data centers load profile is usually almost steady and up to 76% of existing facilities are oversized and therefore inefficient. In addition, it is estimated that up to 58% of energy is wasted in unnecessary and inefficient components, such as chips, slots, fans, voltage regulators, power supplies and many servers (International Energy Agency 2014).

1.1 MOTIVATION

The increase of digital content, big data, e-commerce, and Internet traffic is also making data centers one of the fastest-growing users of electricity (Josh and Delforge 2014). The total energy consumption of data centers in 2012 was about 270 TWh, which corresponds to almost 1.5% of the global electricity consumption, and has an approximated annual growth rate of 4.3% (Van Heddeghem *et al.* 2014). Just in 2014, U.S. data centers consumed 70 TWh of electricity and such consumption is projected to increase to roughly 73 TWh by 2020 (Shehabi *et al.* 2016), costing $13 billion annually in electricity bills and emitting nearly 150 million metric tons of Greenhouse Gases (GHG) emissions per year. If worldwide data centers were a country, they would be the globe's 12th-largest consumer of electricity, ranking somewhere between Spain and Italy (Josh and Delforge 2014). Thus, an understanding of data center energy use, disaggregated energy efficiency options, as well as the metrics used to characterize data center energy performance are fundamental to address this large load in the most sustainable way.

Connected with the above reality, conventional power systems have been facing a noticeable transition from a centralized supply side management to a decentralized supply and Demand Side Management (DSM), as a result of the inclusion of distributed renewable generation, among other factors (Wang *et al.* 2011). The electric power grid infrastructure, the so called smart grid (Aghaei and Alizadeh 2013), should ensure a higher efficiency and reliability through automated control, high-power converters, novel communication infrastructures, sensing and metering technologies, sophisticated energy management techniques, renewable energy, and network availability (Wiboonrat 2012; Cecati *et al.* 2010; Panajotovic *et al.* 2011; Fang *et al.* 2012; Güngör *et al.* 2011). Data centers are completely immersed in this grid context as a player and a typical architecture of them under smart grid environment is presented in Figure 1.3, where the flow of relationships between services requests and power supply can be identified. On the one hand, data centers can take advantage of the flexibility of their loads to implement load management strategies aiming at reducing the operation costs and, on the other hand, play an important role to ensure the efficient and reliable operation of electrical grids by providing which is named Demand Response (DR) services.

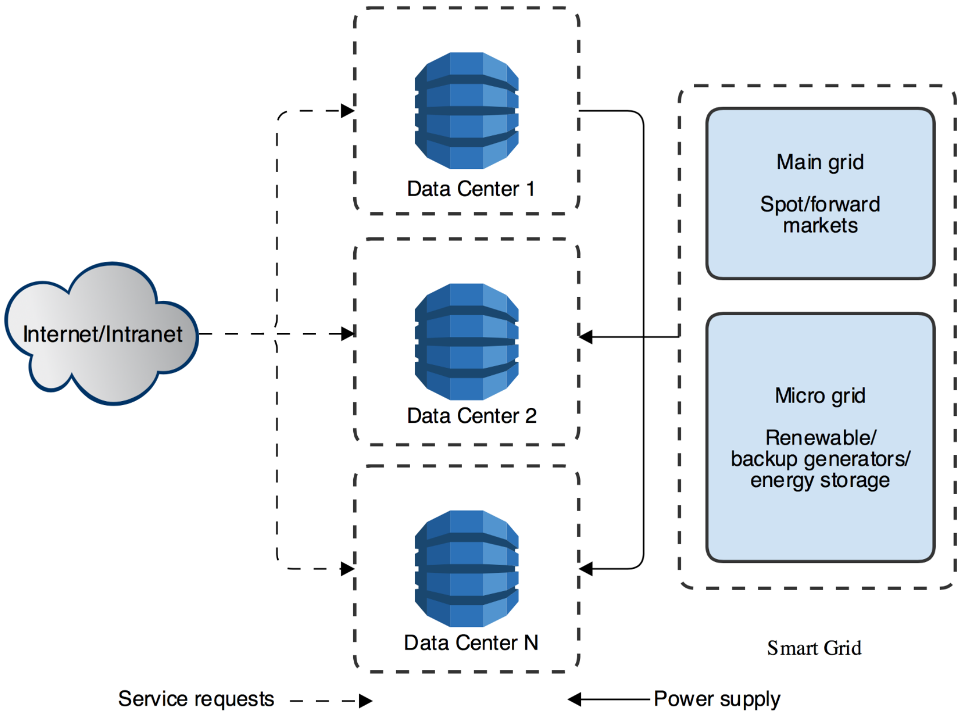


Figure 1.3 – A typical architecture of multiple data centers

This possibility is corroborated by The National Institute of Standards and Technology (NIST) and the Department of Energy (DoE) in the U.S., which have both identified DR as one of the precedence areas for the future smart grid and furthermore the National Assessment of Demand Response Potential report has identified that DR as the potential to reduce up to 20% of the total peak electricity demand in the U.S (Liu et al. 2013). The European Network of Transmission System Operators for Electricity (ENTSO-E) has also quantified a load reduction potential of about 11 GW available throughout continental Europe (Gils 2014).

Another reality is that existing studies tend to be focused on the energy related issues, such as power usage management, server consolidation, load management and DR programs for large data centers rather than small and medium data centers (SMDC). However, SMDC account for more than 50% of the total electricity consumption (Josh and Delforge 2014) and many organizations, such as laboratories, research institutes, universities, industries and enterprises have multiple SMDC scattered around their facilities. In fact, surveys indicate that this data center profile waste more energy than larger facilities, whereby the power consumption is often overlooked, because the energy cost of an individual data center usually accounts for just a portion of total spending. In this context, just as there is a neglected energy efficiency potential for SMDC, there is also a lack of DR programs aimed at these consumers that can contribute to the grid in aggregate form (Josh and Delforge 2014; Delforge 2014; Bennett and Delforge 2012).

The impact and the relevance of this research proposal directly affects data center owners and operators, the electric sector in general, i.e. Distribution System Operators (DSO), Transmission System Operators (TSO), Independent System Operators (ISO), aggregators, universities, researchers, research organizations or even the society in general.

1.2 RESEARCH FRAMEWORK

The research framework was structured to contemplate the research questions to be answered by this work, as well as the study plan to cover the main tasks to achieve the goals. Thus, the four research questions which arose were:

- How can SMDC take advantage of energy efficiency through centralized management of loads?
- Which loads should be defined by data center operators to respond to DR events?
- Can measures of energy efficiency, or participation in DR programs affect the quality and availability of computer services?

- How can data centers owners and utilities balance the maximization of benefits and sustainable goals?

In respect of the thesis research process, it was developed a study plan which was structured in several tasks over time, as follows:

- State of the art.
- Survey of technologies and best practices to reduce and optimize the energy consumption in SMDC.
- Characterization of equipment, associated energy consumption and consumption profile.
- Assessment of the impact and potential of energy efficiency.
- Design of strategies of load management and opportunities for DR.
- Modeling and simulation of the strategies.
- Assessment of the impact of such strategies to the user and to the grid.

1.3 MAIN GOALS

The general goal of this work is to understand how intensive energy consumers, as SMDC, can become more efficient from the energy point of view and how they can take advantage of DR programs to decrease costs and to cooperate with the grid to ensure higher reliability and sustainable development goals. Nonetheless, for each one of the research questions raised in this work there are also specific and consecutives objectives, namely:

- Analyzing the main technologies and best practices to reduce and optimize the energy consumption in SMDC, providing detailed information related to different type of data center dimensions, equipment characteristics, associated electricity consumption and power profile.
- Develop and implement a survey on the current reality regarding energy efficiency in SMDC.
- Definition of the most appropriated energy efficiency methodology for SMDC.
- Definition of load management strategies in the context of SMDC.
- Development of an approach to set specific loads for participation in DR programs.
- Development of an approach to create different DR scenarios based on metrics, thresholds and parameters applied in data centers.

- Development of a conceptual framework that balances sustainable development goals with benefits to the grid and SMDC.
- Development of algorithms that enable the implementation of DR scenarios simulations.

1.4 DISSEMINATION

The following publications resulted from the research reported in this thesis:

Journal papers:

- T. L. Vasques, P. S. Moura, A. T. de Almeida, A Review on Energy Efficiency and Demand Response in Small and Medium Data Centers, In Energy Efficiency (Springer 2018) – Accepted for Publication.
- T. L. Vasques, P. S. Moura, Demand Response Modeling and Optimization Applied to Small and Medium Data Centers, In Applied Energy (Elsevier 2018) – In Review.

Conference papers:

- T. Vasques, S. Araújo, F. Vieira, G. Júnior, M. Castro, G. Souza, P. Moura, Building the Brazilian Smart Grid: Implementation of Smart Grid Technologies in Goiás, In Energy for Sustainability 2015, Sustainable Cities: Designing for People and the Planet (EfS 2015).
- T. L. Vasques, P. S. Moura, An Energy Efficiency Perspective into Small and Medium Data Centers: Progress and Reality Based on Surveys, In Energy for Sustainability International Conference: Designing Cities & Communities for the Future (EfS 2017), 2017.
- T. L. Vasques, P. S. Moura, A. T. de Almeida, Energy Efficiency Insight into Small and Medium Data Centers: A Comparative Analysis Based on a Survey, In Summer Study on Energy Efficiency (ECEEE 2017), 2017.

Presentations:

- T. L. Vasques, P. S. Moura, Load Management and Demand Response in Small and Medium Data Centers, In Energy for Sustainability Research Day (EfS – Research Day 2018), 2018.

1.5 MAIN CONTRIBUTIONS

In this thesis, the main contribution is the participation analysis of a specific and currently neglected type of consumer in the energy market, the SMDC, in DR programs through the management of their loads.

For this purpose, the main technologies and best practices to decrease and optimize the energy consumption in SMDC were selected along with the mathematical models and information such as power, energy consumption, cost and CO_2 emissions.

A current energy efficiency survey directed at Brazil, Portugal and the United States was conducted, showing an alarming reality regarding energy efficiency practices within the SMDC. A comparative analysis with two other surveys with complementary profiles is also specified in order to provide a more holistic view of this scenario.

In the same context, an impact and potential assessment on energy efficiency towards SMDC was performed allowing to observe the main advantages, disadvantages and applications in three important methodologies.

A design of load management strategies and opportunities for DR were proposed in a framework with two layers encompassing an energy efficiency methodology, the management of flexibility and DR scenarios by dynamic tariffs and an incentive-based contract. The first layer has a responsibility-oriented approach, where energy efficiency actions were taken over by SMDC operators through the use of the most appropriate energy efficiency. Flexibility management was ensured by a SMDC algorithm that optimized the best time window to decrease load, whereby DR scenarios were defined by a DSO random-rotating and fairness algorithm in an aggregated and equitable way as if they were large data centers. The second layer is oriented towards the goals, constraints and deployment of DR strategies applied to SMDC and DSO. Alongside, the SMDC load mathematical modeling and specification of optimization processes were provided within the same scope.

After the evolution and conclusion of each of the contributions mentioned so far, simulations were carried out in different scenarios in order to support and provide a cost analysis in different contexts, but aimed at addressing the SMDC and DSO perspective.

1.6 Thesis Outline

This document is composed of 6 chapters that address the work conducted within this thesis.

Chapter 2 provides a comprehensive review in energy efficiency, DR and renewable energy integration, providing the state-of-the-art, perspectives and interconnections to SMDC.

An energy efficiency framework is discussed in chapter 3, whereas three current surveys are presented — one of which was carried out within the framework of this thesis —, underlining their premises and conclusions. Furthermore, on a proposal basis, three consolidated energy efficiency methodologies are analyzed, compared and assessed, allowing a more appropriate use in the context of the SMDC.

Chapter 4 presents a methodology framework addressing the DR phenomenon in different angles, providing an overview on the main approaches to optimize workloads in data centers with examples applied to SMDC. A framework proposal aligned with the goals of this thesis is presented and detailed. The mathematical models denoting the main workloads in a SMDC environment during demand response and rebound events are defined. From this ground, the two problems established in the context of this thesis, one from SMDC point of view and the other from DSO perspective are discussed alongside their resolution hypotheses through an algorithm optimization process.

The simulation results along with the arguments that describe the adopted case studies, input parameters, running, output data and a comparative analysis are presented in chapter 5. Firstly, the optimization outcomes in SMDC perspective and their respective scenarios are demonstrated. Subsequently, the same steps are applied with emphasis on DSO point of view, allowing to simulate the impact of specific actions present in this consumer relationship.

Chapter 6 summarizes the conclusions drawn in the course of this work, making a critical analysis of the obtained results, as well as answering the research questions of this thesis. Some suggestions of future work are also indicated.

CHAPTER 2

STATE-OF-THE-ART

This chapter starts by presenting some introductory concepts and definitions referred along this thesis. Based on the wide and currently heterogeneous scenario described in this thesis introduction, the main purpose of this chapter is bringing light to the above-mentioned issues through a comprehensive review in energy efficiency, DR and renewable integration providing the state-of-the-art, perspectives and interconnections to SMDC.

Several studies have conducted extensive literature reviews whose focus was specifically on data centers, energy efficiency, DR, or related issues. Ebrahimi *et al.* (2014), Fulpagare and Bhargav (2015) and Zhang *et al.* (2014) reviewed aspects related to the main cooling solutions used in data centers, as well as aspects related with waste heat recovery and advances in thermal management. Energy efficiency in networks, telecom systems, power efficient algorithms and server consolidation were addressed respectively in Hammadi and Mhamdi (2014), Garimella *et al.* (2013), Uddin *et al.* (2015) and Ahmad *et al.* (2015), nevertheless in a cloud data center context. Green metrics and renewable energy integration are the main aspects addressed by Uddin and Rahman (2012) and Oró *et* *al.* (2015), regardless of the dimension of data centers. Concerning DR, there are few broad reviews such as Oconnell *et al.* (2014) and (Paterakis *et al.* 2017), however not always focused on data centers specifically, mainly dealing with specific strategies.

Therefore, the present literature review arises from the initiative to relevantly increment and interconnect all the assumptions discussed hitherto separately, understanding that the panorama of this work, as structured in Figure 2.1, should be aligned with the future of data centers, where energy efficiency and DR should go hand in hand taking advantage of all joint potential from a technical (grid services and reliability) and economic (costs minimization) point of view. Thus, such role is ensured with a framework that prioritizes decreasing energy consumption, ensuring the remaining consumption with renewable sources and simultaneously providing DR services to the grid.

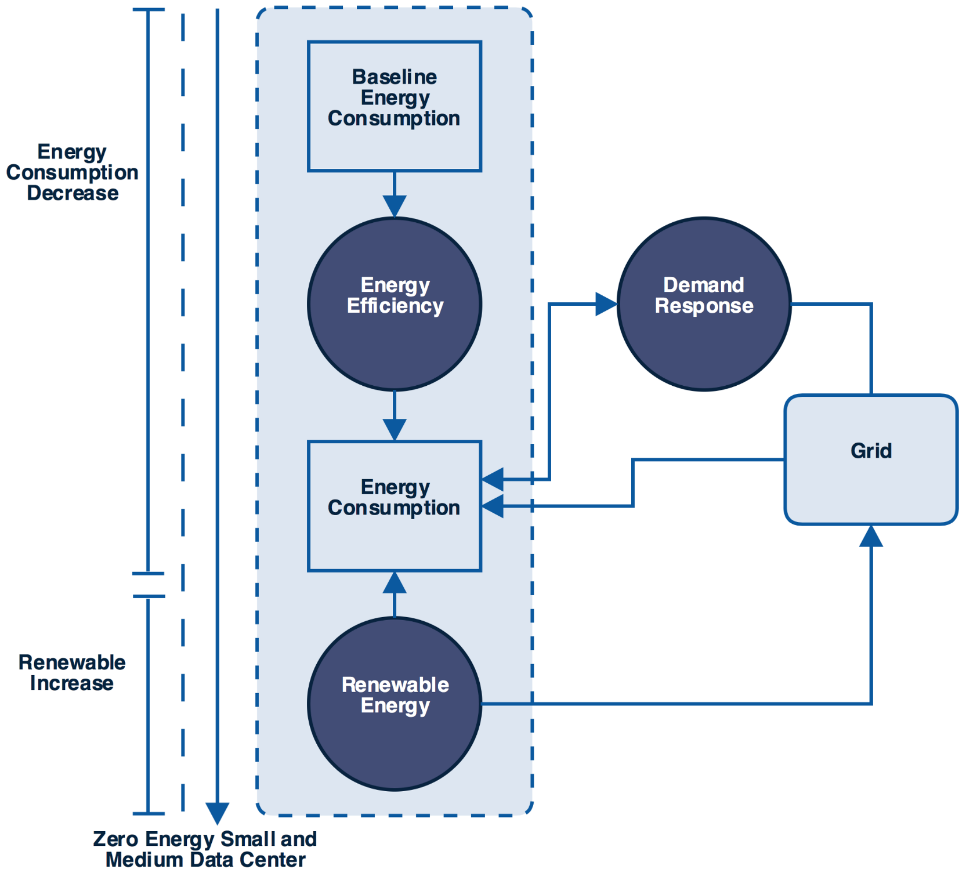


Figure 2.1 – Interconnected panorama of this work

The approach used in this chapter to achieve this goal is dismembering server components generically and analyzing their energy consumption profile and the energy efficiency strategies applied nowadays. By the same criterion, virtualization, cooling, UPS and energy management will be analyzed, as well as their DR strategies and related renewable sources integration. Firstly, this systematic analysis will be performed globally, considering data centers as a whole and finally a SMDC perspective will be given based on recent studies, implementations and respective adaptions.

2.1 METRICS, ENERGY EFFICIENCY AND ENERGY LOAD MANAGEMENT IN DATA CENTERS

The computational management with emphasis on energy efficiency was initially applied in the context of mobile devices powered by batteries. In such devices, energy consumption should be minimized in order to increase the battery lifetime (Beloglazov and Buyya 2013). Although servers with their different generalizations (e.g. blade, tower, and rackable) and data centers can use techniques developed for mobile devices, these systems require specific methods. Thus, the power consumption in data centers is affected by two main factors. The first one, from the hardware point of view, can also be divided in two points: one is caused by the amount of active computational resources and the other is the energy efficiency of the physical components. A way of dealing with the energy efficiency of physical components is the use of a power management system to keep their operation in proportion to the demand for use by applications. This has been done through solutions in hardware and firmware, which will be analyzed in the following sections.

Another factor in the problem of energy consumption, handled from a software point of view, is the inefficiency in the use of computational resources. A study including more than 5,000 production servers over a six-month period showed that even when they are not idle, most of the time, the server utilization is between 10% and 50% of its work capacity, generating heat and unnecessarily consuming energy (Barroso and Hölzle 2007). The existence of a set of computational resources much higher than the average use is justified by the need to deal with peak loads. Although this peak scenario occurs with a low frequency, it is necessary to ensure that performance is not adversely affected, which would happen if an application was executed on an overloaded server.

Judge *et al.* (2008) found that, even when standing idle, servers consume about 70% of the energy consumed during peak working hours. However, according to the data obtained from the SPEC power benchmark (SPEC 2017), the server configurations designed in the end of 2015 consume about 15% to 34% of the energy when idle. Despite the significant reduction of consumption, primarily due to the development of more efficient architectures, maintaining a server connected with a low level of usage is still highly inefficient from the energy consumption point of view.

Based on the highlighted context, it is fundamental to analyze the role of the servers and every supportive technological environment that surrounds it under the energy efficiency panorama in a detailed way, dismembering each component and pointing out the contemporary energy efficiency strategies used in each of them, as well as potential savings divided by CPU, memory, disk, network interface card and the impact of virtualization, cooling and UPS.

2.1.1 Metrics

Before addressing the specifics of a data center in terms of computational components, cooling technologies and power delivery, it is important to understand the metrics currently used for measuring energy efficiency variables in this type of heterogeneous environment. The power consumption of a data center consists of the power consumed by the computing equipment, cooling systems and other sources of power consumption such as lighting, and distribution and conversion losses.

Thereby, The Green Grid consortium has approached four features of efficient data center metrics: intuitive name, scalability to techno-economical changes, scientific accuracy, and the granularity to provide data-driven decisions (Uddin *et al.* 2014). Based on those trends Wahlroos *et al.* (2017) highlight that gauging and assessing the performance of an energy efficiency metric are essential actions in energy efficiency enhancement, allowing tracking the improvements, changes, comparisons between technologies, and benchmarking against average industry performance.

In practical terms, Power Usage Effectiveness (PUE) is a widely-used metric to measure the energy efficiency of non-computing equipment in data centers. It is calculated as the ratio between the total power consumed by a data center facility and the power consumed by the computing equipment (Brady *et al.* 2013), as given by Equation 2.1.

$$PUE = \frac{Total\ Power\ into\ the\ data\ center}{ICT\ equipment\ power} \tag{2.1}$$

A PUE of 1.0 is the best theoretical value, nevertheless not achievable, since in that case no power can be spend on cooling and other facilities. Nowadays, the average data center has a PUE between 1.5 and 2.0, and a highly optimized data center can reach a PUE of 1.1 (Beitelmal and Fabris 2014). Thus, the industry should gradually enhance measurement potentialities over time so that measuring of ICT energy consumption straight at the ICT load (e.g., servers, storage, network, etc.) turns into a widespread practice.

Whereas PUE has become the industry standard for reporting data center energy performance, nonetheless it is not unanimity. Horner and Azevedo (2016) claim that PUE remains an incomplete metric, failing to address hardware efficiency, energy productivity, and environmental performance. They propose a framework under which data center operators report general characteristics and performance metrics to provide a new scenario beyond the PUE inconsistences. In the same aspect, the study conducted by Chinnici *et al.* (2016) provides a general methodology that can be used to measure the energy efficiency of data centers through a holistic approach in which the advantages and the disadvantages of existing and emerging metrics are considered.

However, there are also a wide number of others metrics used in data centers (The Green Grid 2012) described by application, as follow.

In terms of utilization of renewables and reduction of CO_2 emissions Whitehead *et al.* (2014) highlights an important summary with some of the most ordinary metrics applied in data center industry, such as Green Energy Coefficient (GEC) that quantifies the portion of a facility's energy that comes from green sources, Energy Reuse Factor (ERF) and Carbon Usage Effectiveness (CUE). Jeong and Kim (2014) include in this category Water Usage Effectiveness (WUE), Renewable Energy Factor (REF) and Energy Proportionality Coefficient (EPC).

Additionally, eight European research projects have joined forces to introduce new metrics for the evaluation of data centers flexibility, as well as the effects of optimization to their general operational efficiency, such as: Adaptability Power Curve (APC), Adaptability Power Curve at Renewable Energies (APCren), DCAdapt (DCA), Grid Utilization Factor (GUF), Energy Reuse Effectiveness (ERE), Primary Energy Savings (PE Savings), CO_2 avoided emissions (CO_2Savings) and Energy ExpenseS (EES) (Aravanis *et al.* 2015).

Comprising the domain of heating, ventilation, and air conditioning in a data center Wahlroos *et al.* (2017) state the important energy efficiency metrics are: Thermal Correlation Index (TCI), Rack Cooling Index (RCI), Return Temperature Index (RTI), Supply and Return Heat Indexes (SHI) and (RHI), Power Density Efficiency (PDE), Thermodynamic Efficiency (TE) and Energy Reuse Effectiveness (ERE).

According to Nada and Elgelany (2014) generic data center energy efficiency metrics include Power Usage Data Center Efficiency (DCE), Data Center infrastructure Efficiency (DCiE), System Power Usage Effectiveness (sPUE), Fixed-to-Variable Energy Ratio (FVER), Data Center Energy Productivity (DCEP), Server Compute Efficiency (SCE), Data Center Performance Per Energy (DPPE), Fixed to Variable Energy Ratio metric (FVER), Data Center Energy Efficiency and Productivity (DC-EEP), Site Infrastructure Energy Efficiency Ratio (SI-EER) and Electronics Disposal Efficiency (EDE).

Finally, in comparison with single-issue metrics, Whitehead *et al.* (2015) emphasize the importance of Building Environmental Assessment Methods (BEAMs) and Life Cycle Assessment (LCA) to assess and decrease the impact of data centers on environment in the process to reduce their power infrastructures demand. Meanwhile BEAMs analyze the performance of buildings against benchmarks using a set of categories, their respective environment impact weight and a final rating awarded, LCA compiles an inventory to assess iteratively the impact of products, process and services on environment.

2.1.2 CPU

Processor is the central part of servers and as claimed by Wang *et al.* (2017), typically, CPU has been the largest, yet not prevailing, contributing to the power consumption, as characterized in Figure 2.2. In order to provide extra performance when necessary, traditional CPUs are equipped with additional procedures, whose purpose is to minimize the active and static power consumption in an energy efficiency procedure (Varrette *et al.* 2015).

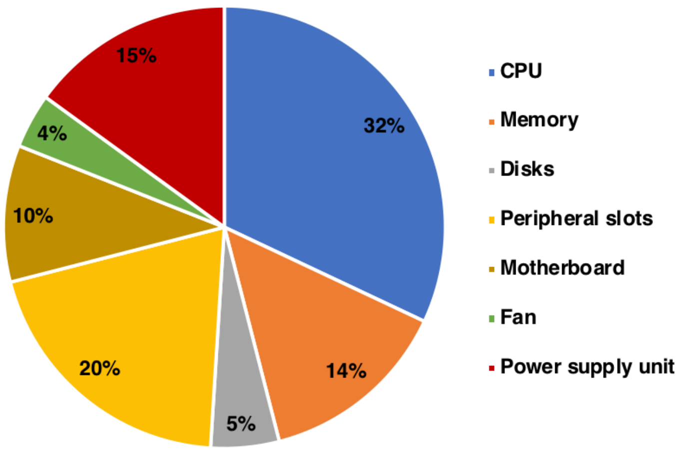


Figure 2.2 – Breakdown of power consumption in servers (Emerson 2015)

Dynamic Voltage and Frequency Scaling (DVFS) is conceptualized by Sueur and Heiser (2010) as being a commonly-used power-management technique where the clock frequency of a processor is reduced to allow a related decreasing in the supply voltage intending to establish an energy efficiency relationship, where the reduction of energy consumption is befitting with the processed workload. The reduction of power consumption leads to a meaningful decreasing in the energy requested for computation, specifically for memory-bound workloads. However, Lu *et al.* (2016) emphasizes that the lowest operating frequency is constrained by the stable voltage conditions of the circuit. The power consumption by DVFS at a frequency f is given by Equation 2.2.

$$P(f) = CN_{sw}V_{dd}^2 f \tag{2.2}$$

where C is the capacitance of the circuit, a significant percentage of which is wire-associated, N_{sw} is the average number of circuit switches per clock cycle and V_{dd} is the supply voltage of the CPU (Zhuravlev *et al.* 2013). As the maximum frequency is linearly subjected to the supply voltage, DVFS has a cubic effect on the power savings, as given by Equation 2.3.

$$P(f) = P_{peak}\left(1 - \left(\frac{f}{f_{max}}\right)^3\right) \tag{2.3}$$

DVFS minimizes the power requirements, but maximizes the application execution time or requests response time, whereas the DVFS total energy efficiency scheme concerns the fact that power reduction is a cubic effect of frequency at the cost of rise in the execution time, which is inversely proportional to the frequency. When considering the delay of sensitive applications (e.g., Internet services) it is fundamental to maintain the response time within certain thresholds. Furthermore, DVFS increases the overall execution time of the task in delay tolerant applications. As an outcome, there is a penalty of using the growth in the execution time even when instantaneous power savings are achieved. Nevertheless, to increase the energy savings without exceeding the execution deadline the energy delay trade-off can be used (Arianyan *et al.* 2017).

Other CPU procedures to provide energy efficiency are the Low-Power Sleep Modes and Dynamic Power Switching. The former, also known by Core Power Gating (Johannah *et al.* 2017) is utilized when the CPU is idle. By cutting the clock signal and power from idle units, the CPU might be commanded to enter in a low-power mode providing energy savings. Dynamic Power Switching technique enables the power management processing in all the domains and constantly monitoring to switch its state to a lower power mode when required.

A different approach was conducted by Krzywda *et al.* (2017) in which several actuators were analyzed jointly to optimize data center servers: DVFS and CPU pinning, which defines the set of CPU cores that each thread can run, were tested and results show that DVFS rarely reduces the power consumption of underloaded servers by more than 5%, but it can be used to limit the maximal power consumption of a saturated server by up to 20% (at a cost of performance degradation). CPU pinning reduces the power consumption of underloaded server (by up to 7%) at the cost of performance degradation, which can be limited by choosing an appropriate CPU pinning scheme.

2.1.3 Memory

Another important component is the local memory, which exists at different levels and it is interconnected to CPU and disk drives, as illustrated in Figure 2.3. The main memory, the Dynamic Random-Access Memory (DRAM), is responsible for a significant fraction of a server's power consumption. However, memory with various power states have been developed. Therefore, any memory power management should assure the performance of memory if DRAM low power states are present.

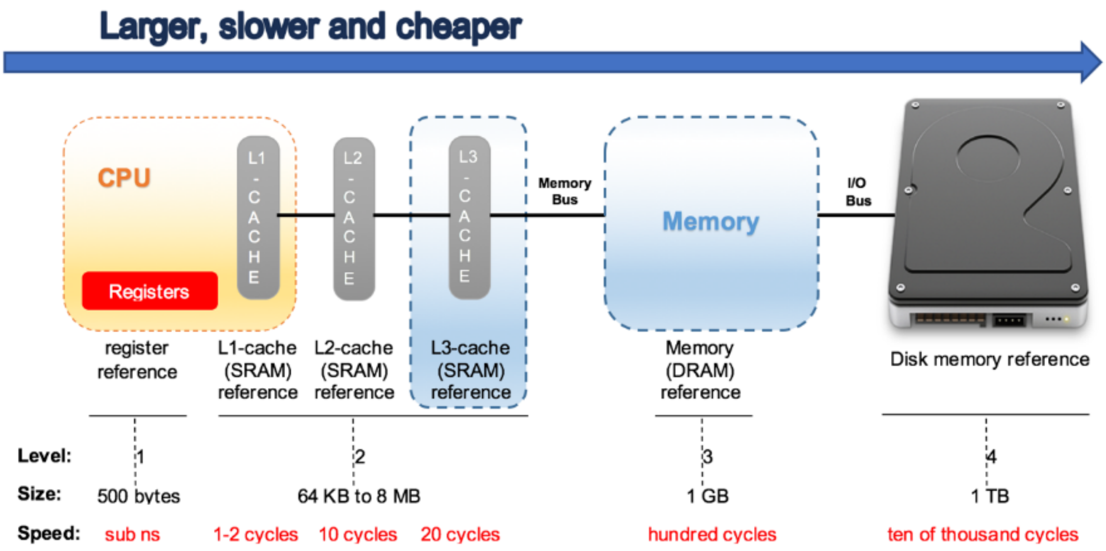


Figure 2.3 – Memory hierarchy key characteristic's in a computing system (Varrette et al. 2015)

The main memory contains *static* and *dynamic* energy consumption. In the decoding of addresses and fetching the data from the memory, dynamic energy is consumed. The static energy is consumed during the active period repaid over the number of data transfers. If E_{rw} is the energy per read or write, BW_{rw} is the read or write bandwidth, D are the total DRAM channels, E_{AP} is the energy required to activate and pre-charge, and f_{AP} their frequency then the energy consumption for each DRAM channel is given by Equation 2.4 (Ahn and Jouppi 2009).

$$E_{DRAM} = StaticEnergy + E_{rw}BW_{rw} + DE_{AP}f_{AP} \tag{2.4}$$

The time needed to regain data from the main memory impacts on the memory-based applications performance. To collect data from the main memory, the probability of hit and miss of the previous level of memory, i.e. caches, can be perceived. If the hit probability is p_{hit}, the miss probability is given by Equation 2.5. The time needed to obtain data from main memory with one level of cache, taking into account the access time t_{access} and miss time t_{miss}, is given by Equation 2.6.

$$p_{miss} = 1 - p_{hit} \tag{2.5}$$

$$t_{DRAM} = p_{hit}t_{\ access} + p_{miss}t_{miss} \tag{2.6}$$

Based on several studies, presented by Pore et al. (2015), the main technologies used in memory to promote a more efficient use of energy are:

- *Memory Architecture Modifications:* Dividing the memory into ranks and making use of smaller subsets of memory rather than the entire memory, results into activation savings and pre-charge energy associated with the rank subsets that are not accessed. Nevertheless, the prompt impact of this technique is the data path for each access becoming longer. The design of memory schemes incorporates different factors such as load balancing across memory ranks, number of memory ranks impacting the effective bandwidth, as well as the application features. Several other methods of power savings comprise managing the refresh rates of memory, use of memory buffer, etc.
- *Memory Low Power Modes*: Currently, new memory sorts have more power states, i.e., Rambus DRAM (RDRAM) establishes four different power states: active, standby, nap and shutdown. Power management patterns for the memory utilize these states to minimize the energy consumption.
  - *Static power management*: The memory is assigned to a low power state and when the memory access takes place, the chip has to resume to the active power state.
  - *Dynamic power management*: The low power state time interval is varied according to the access pattern. The limit time interval after which the memory is in low power state is a crucial design aspect of the power management. This limit is given for improving energy savings, nonetheless the delay is within the time limitation of the application.

Taking into account the Synchronous DRAM Memory (SDRAM), meaning numerous types of DRAM synchronized with the clock speed optimized by the microprocessor and increasing the number of instructions that the processor can perform in a given time, the current generation, Double Data Rate 4 (DDR4) differs due to a 20% decrease in energy consumption from its predecessor, DDR3 (Kim 2016). Whereas Double Data Rate 3 (DDR3) normally demands 1.5 V of electrical voltage, DDR4 demands as little as 1.2 V. In data centers implementing servers running as much as a terabyte of memory on a 24/7 profile, associated with onboard fans and external ventilation systems as cooling solutions, upgrading to DDR4 means big Return of Investment (ROI) in the form of energy savings (Islam *et al.* 2015). Furthermore, development of specifications for the new next generation, Double Data Rate 5 (DDR5) SDRAM, has started, which will be two times faster than DDR4, having double the density, twice the gigabyte capacity and also more power efficient (JEDEC 2017).

2.1.4 Disk

According to Dayarathna *et al.* (2016) and Tang *et al.* (2017) the majority of the storage disks have transition to on-off power states capability. They are either in idle state, standby state or off state when the disks are not in use. Taking into account that d_n is the total number of memory fetched in the storage disk, the active state power, P_{active} is proportional to d_n by a constant factor *d*, $P_{standby}$ is the power during disk Input and Output (I/O) in the low power state, t_{active} is time spent in active state and $t_{standby}$ is time spent in the low power state, then the energy consumption is given by Equation 2.7.

$$E_{disk} = dP_{active}t_{active} + P_{standby}t_{standby} \tag{2.7}$$

The time requested to obtain data from disk is essential in the disk design of power management patterns. If t_{seek} is seek time, t_{RL} is the rotational latency and t_{tt} is the transfer time from disk to higher level cache, then the time requested to obtain data from disk is given by Equation 2.8:

$$t_{disk} = d_n(t_{seek} + t_{RL} + t_{tt}) \tag{2.8}$$

Based on several studies assessed by Pore et al. (2015) the main technologies used in disk to promote more efficient use of energy are:

- *Disk Spinning Down*: Spinning down is the best-known procedure of power management in disks (i.e., switching the power off) when not in utilization. Nevertheless, the time to re-establish the disk to the active state takes few seconds and there are sudden fluctuations in the data center workload, being able to strongly degrade the delay sensitive applications performance. However, using prediction-based techniques to schedule the disk spinning down in the idle timeframe in the workloads it is possible the performance breakdown of the applications due to power state transitions.
- *Managing Data Storage and Replication*: There is often a large data set stored in multiple storage disks in data center applications, involving popular data, which are more frequently accessed than the common data remaining. Identifying these most requested data, storing on fewer disks and replicating them for performance, whereas the rest of data is stored on remaining disks is the role of this technique. The disks with most requested data are always in active state while more power management patterns are used to the remaining disks. Other power management techniques involve the usage of hybrid disk types such as a combination of Solid State Drives (SSD), Flash Storage Devices and DRAM to control the data storage based on the combination of their power, performance features and costs. The most requested data is migrated to more energy efficient devices, but moving the data frequently might exceed the savings obtained by a spinning down of disks.

Several studies related with these technologies have been done (Zakarya and Gillam 2017) concluding that an ideally proportional system has an energy reduction potential of 40 – 75%. Examples of power management schemes proposed for disk power proportionality are:

- *Consolidation*: Moving data to a specific number of storage devices (Verma *et al.* 2010) or using a scalable architecture (Tsirogiannis *et al.* 2010).
- *Migration*: Storage of data in different energy efficient devices, e.g., SSD (Tsirogiannis *et al.* 2010), (Härder *et al.* 2011), hybrid disks, such as NAND flash storage and DRAM (Deng 2011) or in accordance with data popularity (Amur *et al.* 2010).
- *Aggregation*: In order to maximize the idle times among the operations to provide more opportunities for energy savings, read or writes are postponed (Gupta and Singh 2007).
- *Disk spin down*: During idle periods spinning down the disk is used regardless or in combination with other management policies (Verma *et al.* 2010).
- *Compression*: Utilization of Data Compression in some workload cases (Gupta and Singh 2007).

2.1.5 Networking Interface

Chen *et al.* (2016) highlights that the average use in the data centers is very low and idle networks devices such as ports, line cards, switches, are one of the considerable consumers of energy in the low utilization periods (Gupta and Singh 2007). Therefore, a power management procedure that is frequently used for network devices is to turn off the network components during idle timeframes (Gupta and Singh 2003). When the network components are not in use, they are either in idle or standby states. If P_{active} is the active state power and $P_{standby}$ is the power in the standby state, t_{active} is the time spent in the active state and $t_{standby}$ is the idle period, E_{net} is the energy needed for switching between the power states and n_s is the number of switching that occurred, then the energy consumption is given by Equation 2.9.

$$E_{net} = P_{active}t_{active} + P_{standby}t_{standby} + n_S \qquad (2.9)$$

The time, in seconds, needed to transfer the data through a network component in terms of the transfer time, $t_{transfer}$, and switching time, $t_{switching}$, is given by Equation 2.10.

$$t_{net} = t_{transfer} + t_{switching} \qquad (2.10)$$

Network cards are generally off-board in the main data center servers and therefore are included in the estimation of 20% of the peripheral slots presented in Figure 2.2. The main techniques to promote an efficiency use of energy associated with networking interfaces are (Pore *et al.* 2015):

- *Switching off the Network Component*: Use of a reactive scheme that switches-off the network component for a certain time after observing that there is no workload for few seconds. Some procedures imply proactive patterns where network interfaces are continually monitored for learning the inter-arrival time between packets in a window-based method.
- *Managing the Workload*: Data can be aggregated, stored in buffers for some period and sent if the application deadlines are not stringent, enabling the network components to be turned off during the idle period.
- *Sleep*: The network components, such as switches and routers are in sleeping mode or turned off in the idle timeframe between the workload arrivals, reaching energy savings between 10 and 20% (Gupta and Singh 2007; Nedevschi *et al.* 2008).
- *Aggregation*: The network topology is modified to consolidate the network flow on fewest possible routes, such that the data is sent on minimum active series of network devices. Bonetto et al. (2014) present that even simple policies allows to save from 30% to 50%.
- *Rate adaptation*: With this technique the workload rate is adjusted such that traffic is serviced within the required time constraints, achieving energy savings between 10 and 90% (Gupta and Singh 2007).
- *Traffic shaping*: The traffic is divided into bursts, in the Elastic Tree procedure. This traffic to same destinations is buffered before it is routed. This scheme increases the idle periods between the traffic bursts applied to transition the network devices into low power states and can save up to 50% of network energy, while maintaining the ability to handle traffic surges (Heller *et al.* 2010).

2.1.6 Virtualization Framework

Virtualization is an increasing and leading technology to mutualize the energy required by a single server operating multiple Virtual Machines (VMs) instances. Often virtualization is confused with colocation, which is the practice of housing privately-owned servers and networking equipment in a third-party data center. Nevertheless, short consensus has been produced about the capacity overhead in energy consumption and the throughput minimization for virtualized servers and/or computing components. Other way to address this topic is conceptualized by Mazumdar and Pranzo (2017), where virtualization helps to reduce the power consumption within a cloud infrastructure by enabling consolidation of heterogeneous applications in few active servers.

The main structure of any virtualization platform and thereby Cloud middleware, remains the hypervisor or Virtual Machine Manager. Thereafter, a VM operating managed by a hypervisor is called a guest machine. The main two categories of hypervisors are: native and hosted, but only the former (also called bare-metal) contains an interest in data center context. This category of hypervisor operates straight on the host's hardware to take over the hardware and to supervise guest operating systems. Nevertheless, from the energy efficiency point of view, to implement efficient virtualization mechanisms, it is required to define the allocation of a VM to a physical machine and live VM migration during overburdening situations.

It is important to take into account that a virtualized data center with servers hosts a sub-assembly of applications by supplying a virtual machine for each application hosted upon it. Hence, an application might have numerous tiers, multiple instances, operating across different VMs and being managed by a global VM controller in the data center, which is responsible by the admission control, load balance, assignment and migration of VMs, as it can be seen in Figure 2.4.

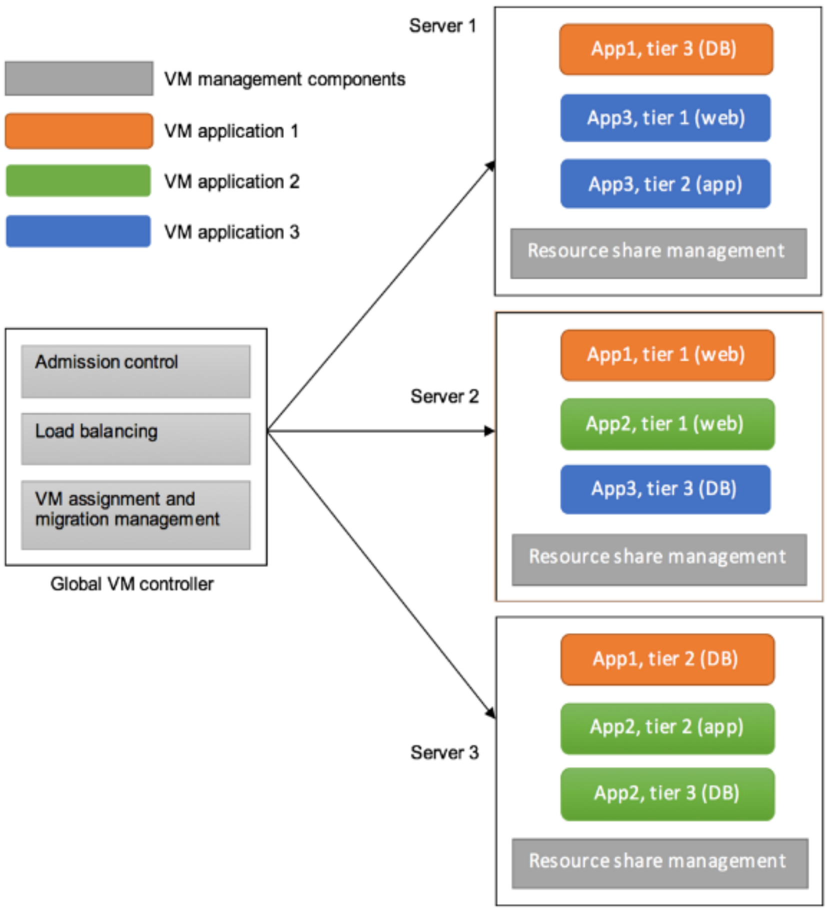


Figure 2.4 – A pictorial example of VM management components (Pore et al. 2015)

The VM management determines: how many VMs are necessary for each tier of an application; how the workload should be propagated among multiple instances of the application for each tier; how the resources should be held in common among VMs that are collocated in a VM; when and how VM migration should be conducted; and how the resources should be held in common among collocated VMs.

By analyzing the standard of VM Assignment and Migration, in practical terms, VM assignment is harder owing to the VM migration overhead and the uncertainty related to the resource constraints of the VMs. The overhead of VM migration produces a cost of VM assignment which is not time-independent. Thereby, the optimal solution for the VM assignment issue in non-negligible migration cost, is found offline where the information of resource demand of VMs are known over all times in advance. Hermenier et al. (2009) claim that consolidation based on the entropy definition, providing cluster resource management and a group of sensors, can achieve savings of more than 50% if compared to the static solution.

Several works have been addressing the problem of VM assignment and migration. Dhiman et al. (2010) highlights the variations in the power consumption for CPU-intensive and I/O-intensive applications through trials, where different applications have different characteristic specifications at different times and thereby VM allocation in physical server can be reached when power peak does not occur in parallel. It is stated that placing a hybrid of CPU-intensive and I/O-intensive application in a physical server produces less power consumption compared to any other arrangement. This occurs because it breaks up with the hotspot, which is an intense operation in CPU-intensive or I/O-intensive applications on one physical machine and splits the global system resources in a much more efficient form. An implementation of the system on a state-of-the-art testbed of server machines called vGreen, an open control loop to manage the application assignments (VMs) to the physical server, is presented maximizing average performance and system-level energy savings by around 40%, using benchmarks with different specificities (Dhiman *et al.* 2010).

The condition of holding fully used machines is the requirement of VMs being dynamically managed, as the resource demands of interactions-based applications (e.g. web-based application) and batch jobs are distinct. In the part of batch jobs, when a job is over, its resource requirement might not reach the resource requirement for the new job, thus, VMs should be repositioned to the physical machines to hold fully used machines. Hermenier et al. (2009), implements a dynamic restriction programming based on VM manager named Entropy, which operates VMs such that whenever an unallocated VM is accessible, it is applied to a physical machine that is capable to meet its resource demand. Every new allocation might cause some live virtual machine migration satisfying the objective.

From the VM Dynamic Resource Allocation point of view, to make the services in the form of VM applications available, the cloud presents autonomous management of the available physical resources. Two techniques are available for autonomic resource allocation to VMs: Static resource allocation (Hermenier *et al.* 2009), and Dynamic resource allocation (Wang and Wang 2011; Padala *et al.* 2009; Nathuji and Schwan 2007).

The static resource allocation model assumes that the resource demand of VMs are known in advance and the VMs' resource demands does not modify significantly during their life and the VM allocation is carried out in accordance with the peak resource requirements of the application. The pattern discussed in the literature for dynamic resource allocation interval analyzes how to optimize a utility model that catch Service Level Agreement (SLA) revenue cost as well as energy cost (Ardagna *et al.* 2012; Urgaonkar *et al.* 2008; Padala *et al.* 2009), machine learning techniques for learning resource requirement of applications (Tesauro *et al.* 2006), and control theory methodologies (Wang and Wang 2011).

An adaptive control scheme to define on VM resources share of multi-tier web based applications is used in Padala et al. (2009). An optimal control technique is used in Urgaonkar et al. (2008) to decide resource allocation and power management for time-varying workloads and heterogeneous applications. Wang and Wang (2011) studied power capping in a virtualized cluster to create a closed control loop by using Model Predictive Control and PI controller to supervise both the power consumption and performance goals of applications in a coordinated way. VM assignment and resource allocation were studied in Ardagna et al. (2012) in a combined way, considering a multi-tier virtualized system with the objective of increasing the SLAs revenue while decreasing energy costs, reaching substantial revenue earnings for the provider in comparison with alternative methods (up to 45%). A dynamic provisioning technique for multi-tier Internet applications was proposed by Urgaonkar et al. (2008) employing a queuing model to define how many resources have to be allocated to each tier of application and a hybrid of predictive and reactive mechanisms, determining when to provision these resources.

Castro et al. (2013) propose three new approaches for dynamic consolidation of VMs that take into account both CPU and RAM usage. A heuristic called CPU and RAM Energy Aware (CREW), which uses an energy model that jointly considers the consumption of CPU usage and RAM was proposed to define the allocation of VMs, ensuring the lowest possible power consumption. The implementation and evaluation of tenders done in the CloudSim (Calheiros *et al.* 2009) simulator used real workload VMs from the PlanetLab (Spring *et al.* 2006) and an enterprise cluster Google (Reiss *et al.* 2012). The results showed a reduction on the energy consumption in up to 33% and an increase on the Quality of Service (QoS) guarantee.

Based on this context, there are still demands in the VM management, given that VM migration overhead deteriorate not only the performance of the migrated VM, nevertheless also the performance of the VMs collocated in the source and destination physical machine (Lim *et al.* 2011). Constant cost per each migration is the base for many studies (Sanders *et al.* 2004), using migration time, which affects many elements as follow: (1) the memory update rate and the memory content of each virtual machine, (2) the virtual machines migrated total number, (3) network bandwidth availability for migration, and (4) the destination servers at the time of migration and the workload of the source (Dargie 2014). Thereafter, most of the presented solutions are greedy or heuristic, whose approximation ratio in comparison with optimal solutions are not derived, nevertheless VM assignment and resource management are given as a NP-hard. Lastly, virtualization approach simplifies dynamic power management and minimize power consumption, however the applicability of the virtualization under various situations such as real-time is not well researched. This is essential, considering that the VM overhead, the delay demands of some applications in an underused VM might not be satisfied.

2.1.7 Uninterrupted Power Supply (UPS)

UPS are key components of ICT systems, ensuring reliability by maintaining the continuity and quality of the systems' power supply. An UPS is understood to be a short duration (minutes to hours) power supply system that maintains the functions of the connected load when the main continuous power source has failed or has significant disturbances (IEC 2013). Therefore, the primary purpose of a UPS is to bridge an unexpected power gap and/or to provide the amount of power needed to safely power down the connected load. A UPS may also be used to continuously maintain the quality (e.g. harmonic content) and stability (voltage and frequency) of the power to the connected load. In data centers, the UPS systems are used to ensure the service continuity of ICT, to protect it from risk of halts in data processing, contributing to 7% of the total energy consumption (Moura *et al.* 2016).

The energy consumption of UPS should be an important consideration due to its high impact on the lifecycle costs, nevertheless in most applications of UPS, energy efficiency is not the most important issue, since the operational reliability of the ICT systems and the related security of data processing and storage are the major concerns. However, the conversion efficiency of UPS systems has been improving in recent years and large energy savings can be achieved with the adoption of new technologies without a reduction of the reliability levels.

Regarding UPS life cycle, Khan and Khan (2015) present the amount of power that can be stored or retrieved from the batteries, at a given time t, of the UPS, which are limited by their maximum amounts. The lifetime of the UPS is constrained by the number of cycles of UPS charging and discharging

(Wang *et al.* 2012) and therefore the operating cost of the UPS also depends upon UPS charging and discharging cycles.

Moura *et al.* (2016) analyzes the UPS efficiency clamming that product performance reliability and system configurations with high redundancy often conflict with optimized life-cycle costs. At a given level of supply security, these costs are an important consideration for the user.

In UPSs, a decrease in the consumed energy of the product and their installation system architecture leads not only to a direct decrease in the UPS energy costs, nonetheless also produces cooler operating conditions within the installation environment. This leads to a reduction in ventilating and air conditioning energy and infrastructure installation costs, an extension of the service life of UPS key components (e.g. such as energy storage batteries and capacitors) and an increase in the overall lifetime reliability of the UPS system.

The key factors that must be considered regarding energy efficiency are the size of the UPS, load type and load level. Larger UPS modules typically have higher energy efficiency than smaller ones because the power required for control electronics and auxiliary components becomes a smaller portion of the total capacity of the UPS system. The efficiency of an UPS depends on the load level, achieving the highest efficiency with a 100% load, as depicted Figure 2.5. However, the curve is relatively flat with load levels higher than 50%. An UPS operating with a low load level will have significant losses when compared with the same UPS operating at full load. In a realistic scenario, the load level is typically between 10 and 30%, which leads to a 4 – 17% reduction of efficiency (Moura *et al.* 2016).

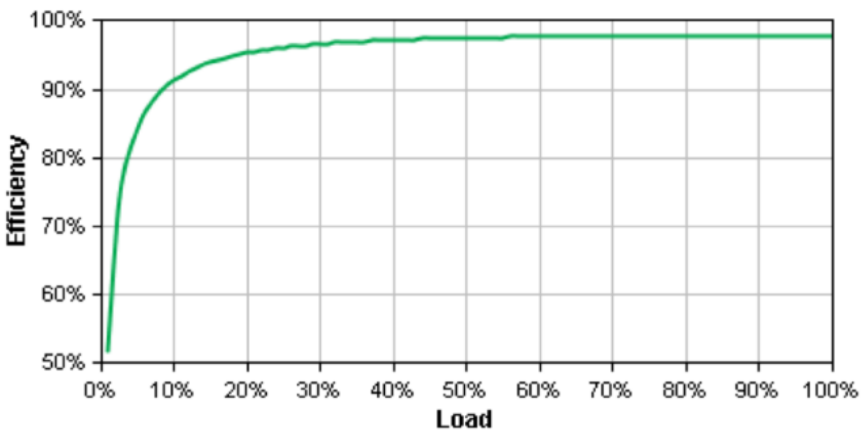


Figure 2.5 – Typical UPS efficiency curve (Moura et al. 2016)

The load type also has a strong influence on the achieved efficiency. UPS efficiency is usually tested with resistive or linear loads, but several UPSs are used with non-linear loads, with low power factor and high total harmonic distortion (Pier 2008). The low power factor will require a higher peak current from the UPS, decreasing its efficiency (Figure 2.6).

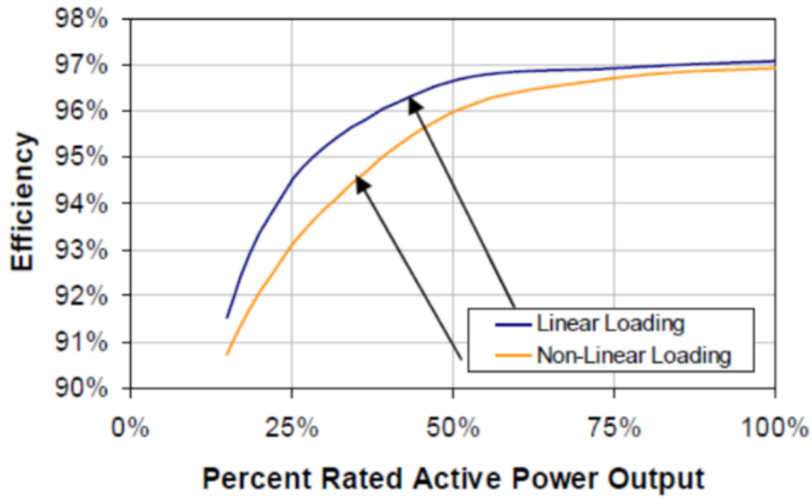


Figure 2.6 – UPS efficiency with linear and non-linear loads (Pier 2008)

Moura *et al.* (2016) assess the potential savings considering several technologies at UPS's component and product level, modeling the main design options and considering policy options focused on minimum efficiency performance standards and energy labelling. The results show a potential for energy savings in the European Union in 2025 of 11.4 TWh (65% energy saving relative to predicted energy requirement of EU ICT system UPS based on current practice).

Other alternative that can be considered concerning energy efficiency related to UPS's in accordance with Boulos et al. (2014) is the Direct Current (DC) distribution, since all data centers ICT equipment such as servers and storage devices are essentially DC-based loads. The backup supply system commonly required for critical facilities consists of batteries, which are also based on DC. Thus, by deploying a DC distribution rather than conventional AC, several conversion steps in the power delivery system can be eliminated, reducing distribution losses. By using DC entirely throughout a data center will save 10% to 20% in power costs and improve reliability (Sithimolada and Sauer 2010).

As seen in Table 2.1, DC-DC converters can reach an efficiency of 90 – 96% as compared to AC-DC power supplies which provide an efficiency of 65 – 75%. Even comparing best-in-class AC-DC to DC-DC a 2 – 5% advantage to the DC-DC solution can be reached. The efficiency findings show that 380 V DC provides the highest efficiency DC option, particularly when compared with the 48 V DC system, however, it requires a critical mass of 380 V DC commercial equipment to exist in the building before any user could decide for this option (Moreno-Munoz *et al.* 2011).

Table 2.1 – Power distribution efficiency comparing AC and different DC distribution methods (Moreno-Munoz et al. 2011)

	UPS	Distribution wiring + PDU (Power Distribution Unit)	PSU (Power Supply Unit)	Load Converter 12 V–1V	Total Efficiency
Facility AC-UPS	92.00%	99.00%	75.00%	88.00%	60.00%
Facility DC-UPS 48V/24V	92.86%	99.00%	91.54%	88.00%	74.00%
Facility DC-UPS 380V	96.00%	99.00%	91.75%	88.00%	76.73%
Distributed DC-UPS	92.00%	99.00%	94.00%	88.00%	75.34%

2.1.8 Cooling

One of the important elements to manage a data center reaching the expected performance is the cooling system. Depending on their power consumption, server systems need a specific amount of cool air at the intake and exhausts the same volume of heated air at the outtake. In such situation, the room is not able of supplying this amount of air, the server will draw in its own exhaust air, overheating the device. Thus, a proper cooling approach is inevitable for an uninterruptable server functioning, however Ni and Bai (2017) claim that more than half of the data centers' air conditioning systems are inefficient. Thereby, in order to increase the energy efficiency, it is fundamental to take into account many factors before adopting a cooling solution, such as energy use, facility location, density per rack, power density and other user specificities (Varrette *et al.* 2015). One practical example of this heterogeneity it that the power used by single racks might vary dramatically, with an average of around 1.7 kW up to 20 kW in high density servers (Varrette *et al.* 2015), directly affecting the adopted cooling solution.

There are three main cooling architectures used in Data Centers, such as room-, row- and rack-oriented. In the first scenario Computer Room Air Conditioners (CRAC) units are linked with the room (Fulpagare and Bhargav 2015) and cool air might be unrestricted or partly limited by ducts or vents when provided by the conditioners. Because the air supply uniformity is poor due to specific room designs, such as their shape or obstructions, the full rated capacity of the CRAC unit cannot be used in most situations. In row-oriented solutions, CRAC units are linked with a row, being their performance higher, as the airflow paths are shorter. Consequently, the requested CRAC fan power supply is smaller, decreasing the energy cost. In the last scenario, rack-oriented, CRAC units are linked with the rack allowing the cooling to be accurately adapted to the constraints of servers. On the other hand, the disadvantage is requiring a large amount of air conditioning equipment (Sohel Murshed and Nieto de Castro 2017).

Cooling systems mainly utilize air or a refrigerant and depending on the kind of cooling systems different operation temperatures will be required. In addition, depending of the different amount of heat exchanger processes, different elements will compose the cooling system, with different efficiency or losses (M. Vordem Berge et al. 2015). In order to reach energy-efficient data center cooling, numerous solutions have been presented, being the three most broadly adopted and effective: hot- and cold-aisle isolation, closed-coupled cooling and free cooling.

Considering the hot- and cold-aisle isolation, through a raised floor, a steel grid resting on stanchions installed 60 – 120 cm on the concrete floor, the cold air is supplied by CRAC units. By perforated tiles the cold air will flow into racks and then, hot air will be exhausted through a rear side of rack after absorbing heat produced by servers in the rack. One strategy to maximize the cooling efficiency is to avoid mixing the cold air supplied by CRAC and hot air exhausted by servers. This is guaranteed by a solution named hot- and cold-aisle isolation, arranging server racks such that the intakes of cold air in server racks are faced each other. The hot air is eventually drawn by the CRAC and the cold air is once more supplied to cold aisles by exchanging the heat with cold air (or water) delivered from chillers. This solution can achieve up to 40% energy savings according to Kim et al. (2015).

On the other hand, according to Alkharabsheh *et al.* (2015) and Capozzoli *et al.* (2015) there are inefficiencies associated with the cooling scheme in data centers, being air mixing one of the most important. The cold aisle–hot aisle implementation does not completely isolate the cold air streams in the cold aisle and the hot air streams in the rack exhaust for two causes: hot air recirculation and cold air bypass. The former, indicates that the hot air enters into the cold aisle from the top of the racks and the front end of the cold aisle closest to the cooling units. The recirculating warm air mixes with the cold air and increases the inlet temperature in a difficult way to foresee. Thereby, air mixing affects the reliability of the ICT equipment. Avoiding this would require the cooling system to have a complicated control system. The later, cold air bypass, happens by the time the cold air from the perforated tiles overshoots the racks and returns to the cooling unit at a lower temperature. This can also happen due to floor leakage between the cold aisle and the cooling unit. The resulting low cooling unit extract temperature decrease the efficiency of the cooling system by narrowing the temperature difference between the extract and the supply.

With the aim of decreasing the losses incurred throughout the supply of the cooling medium and quickly react to spatial temperature distribution, closed-couple cooling solutions place cooling units more nearly to computing units. There are mainly two categories according to the granularity of computing cluster covered by single cooling unit i.e., in-row and in-rack cooling's. An in-row cooling adapts the cooling requirements at every row in accordance with the corresponding conditions, while an in-row cooling adjusts its cooling settings in accordance with operating condition at each rack, achieving energy savings up to 40% (Kim *et al.* 2015). Nevertheless, the capital expenses for the installation is very high.

Free cooling is an approach to lowering the air temperature in a building or data center by using naturally cool air or water instead of mechanical refrigeration (Oró *et al.* 2015). The adoption of free cooling schemes is currently one of the most utilized techniques to turn data center more efficient. Zhang et al. (2014) and Ebrahimi et al. (2014) reviewed the improvements of data center free cooling schemes mainly with attention on setting characteristics and performances which can be roughly separated as follow:

- *Airside free cooling:* Use outside air for cooling data centers.
  - *Direct airside free cooling:* Drawing the cold outside air straight, after filtering into the data center.
  - *Indirect airside free cooling:* Running through air to air heat exchangers.
- *Waterside free cooling:* Use natural cold source by cooling water infrastructure.
  - *Direct water cooled system*: Natural cold water is used directly to cool the infrastructure through a heat exchange between the warm air and sea, river, ground water.
  - *Air cooled system:* Air cooler is utilized to cool the water circulating to CRACs when wet-bulb temperature of the outside air is low enough.
  - *Cooling tower system:* A cooling tower is adopted to cool the water circulating in CRACs and heat exchangers. Two water loops are required; a cooling (external) water loop and a chilled (internal) water loop.

Siriwadrana et al. (2013) researched the inclusion of outside air with desired supply air for data center cooling in the Australian climate conditions. It was found a significant potential for using this scheme in some states that could lead to significant savings on cooling costs. Subsequently, Lee and Chen (2013) using a dynamic building energy simulation program (eQUEST) have found some energy savings potential of airside free cooling for data centers in worldwide climate zones, highlighting that sizable direct air free cooling potential was reached in data centers positioned in humid climate zones. In addition, in dry climate conditions substantial humidification is needed using techniques such as evaporative cooling, where raising the humidity of air lowers the temperature and thereby the water consumption must be considered, since 1 MW in a data center can consume 68 m^3/day of water for cooling, as presented by Ristic *et al.* (2015).

However, there is a potential risk to damage ICT equipment using direct airside free cooling due to the risk of particulate contaminants entering data centers. In this context, Dai et al. (2012a); (2012b); (2013) studied different strategies to minimize the risks for ICT and telecommunication devices under this cooling policy. Nevertheless, even though Shehabi et al. (2007); (2008) claimed that with an appropriated filtration, the ASHRAE (2011) suggests particulate contaminant concentration for data centers is accomplished and significant economic saving can be achieved.

Some data center operators are leveraging to utilize seawater and geothermal energy to produce cooling for their infrastructure and pursuing green practices and minimize energy costs. After implementing a unique seawater cooling system, an 1600 m^2 data center located in Stockholm reduced energy costs by 80% (Oró *et al.* 2015). Furthermore, it was reused the sweater to heat local offices and residential buildings before returning it to the sea. Consequently, the data center has lowered its PUE value to 1.09 minimizing its ICT load enough to enable additional customers to distribute in the facility, reinforcing its economic benefit. Similarly, data centers in Iowa and Nebraska are cooled by a geothermal bore field utilizing the cool temperatures underground to cool down the servers (Oró *et al.* 2015).

2.1.9 Energy Load Management

Data center energy load management is an area of growing interest as it is supported by real preoccupations on energy usage and cost by modern computing systems. It has evolved to a category called Data Center Infrastructure Management (DCIM), i.e., software that reports granularly from the data center facility to the server and device level.

The DCIM concept involves the ICT integration along with facility management, aiming at centralized monitoring, management and intelligent capacity planning of data center systems. Capacity planning focuses firstly on energy, power, space, ICT equipment, cabling, network, cooling and environmental factors, such as temperature and relative humidity are covered (Cappuccio 2010). DCIM systems can provide more energy efficiency mapping and managing the entire power chain and therefore the energy capacity of the data center.

Even though DCIM systems are typically adapted for large data centers, the needs of small to medium data centers are not adequately contemplated nowadays, since such systems are generally complex, pricy, difficult to utilize and not modular enough. Additionally, currently solutions offered on the market are normally proprietary (Kim *et al.* 2015).

Besides enterprise solutions, relevant research effort was conducted on energy efficiency modeling and optimization to provide energy management. For example, models of servers' power usage were introduced in Basmadjian *et al.* (2011) whereas application of these models to energy-aware scheduling in Mämmelä *et al.* (2012). Moreover, Witkowski *et al.* (2013) address power modeling and estimation methodologies through specific application classes and Mukherjee *et al.* (2010) utilized thermodynamic information in scheduling. However, the previous studies are focused on research and optimization issues instead of offering models to simulate real data centers.

In Raghavendra *et al.* (2008), a power management solution that coordinates different individual approaches was proposed using simulations from 180 server traces from nine different real-world companies. Shah and Krishnan (2008) present the possibility of globally compute workloads to take advantage of local climatic conditions to reduce cooling energy costs, by undertaking an in-depth analysis of the environmental and economic burden of managing the thermal infrastructure of a globally connected data center network. SimWare (Yeo and Lee 2012) is a data warehouse simulator which computes energy efficiency analyzing the power consumption of servers, cooling units, fans as well as the effects of heat recirculation and air supply timing. A platform and application agnostic methodology for full-system power modeling in heterogeneous data centers was proposed (Canuto *et al.* 2016). It is based on collecting power and resource usage measurements while running a special training workload and fitting them through machine learning.

Innovative DCIM support systems for datacenter management are therefore needed. Power Monitor System and Management (PMSM) (Kim *et al.* 2015) is an example of such an innovation, as well as CoolemAll project (Cupertino *et al.* 2015), which proposes rethinking data center efficiency based on the interaction of all the factors involved rather just one set of technologies. The expected results included a data center monitoring, simulation and visualization software, design of energy efficient ICT hardware, contribution to existing and new energy efficiency metrics. In the same aspect, RenewIT project (Salom *et al.* 2017) developed a simulation tool to evaluate the energy performance of different technical solution integrating RES in several European climate regions. The public RenewIT tool helps actors from both the energy and ICT sectors to reduce the carbon footprint of planned data centers in the horizon of 2030, being based on selected meta-models extracted from advanced dynamic simulation models of challenging energy concepts for renewable energy supply of data centers.

Therefore, DCIM progressive adoption will produce technology updates, which in turn will enhance DCIM-compliant equipment and sensors. This is a fruitful and growing research area which is being updated as data centers improve.

2.2 DATA CENTER DEMAND RESPONSE

Nowadays, conventional power systems have been facing a noticeable transition from a centralized supply side management to a decentralized supply and demand side management, as a result of the inclusion of distributed renewable generation, among other factors (Wang *et al.* 2011). Traditionally, power systems were managed by scheduling the generation resources since loads were not controllable or even measurable with the required time resolution (Du and Lu 2011). However, with recent technological advances, consumers can be motivated to actively participate in the balancing between demand and supply by controlling their electricity consumption and using energy storage systems (Eu Commission Tf For Smart Grids Expert Group 2010). This is ensured by the smart grid, which is an electricity network that can intelligently integrate the actions of all users connected to it—generators, consumers and those that do both—in order to efficiently deliver sustainable, economic and secure electricity supplies. (Gellings 2011). Supported by ICT, the intelligent control of loads, generation and storage resources would increase the overall sustainability and reliability with potential benefits to the entire value-chain of generation companies, transmission system operators, distribution system operators, energy suppliers and end-users.

A DR event is when end-use customers reduce their use of electricity in response to power grid needs, economic signals from a competitive wholesale market or special retail rates. DR is another important competitive resource that can be used to maintain demand and supply in balance for grid operations and the associated wholesale markets (PJM 2013).

Therefore, large energy consumers are ideal candidates for participation in DR programs for having a significant impact on the load diagram. A feasible example of this case are data centers, which are intensive energy consumers.

A data center perspective on DR programs relies on the intersection of two important social issues. First, as ICT becomes increasingly pivotal to society, the associated energy demand are way up, being the growth in electricity demand for ICT ten times larger than the overall growth of electricity demand (Dreibholz *et al.* 2007; Koomey 2011; Ghatikar *et al.* 2012). Second, the integration of renewable energy into the power grid is essential for enhancing sustainability, however causes significant challenges for management of the grid (Zhu et al. 2012). The focus behind DR and energy efficiency in data centers is that these two challenges are in fact cooperative to reduce carbon emissions from electric power generation and to combat the effects of global climate change. Therefore, it is important not only to assess how an intensive energy consumer, such as data centers, can decrease costs by increasing their energy efficiency, nevertheless also how they can take advantage of DR programs to decrease costs and cooperate with the grid.

Data centers represent very large loads that can reach up to 100 MW and are particularly well-suited for participation in DR programs (Wierman *et al.* 2014), since they have flexible loads (Ghatikar et al. 2010) and are greatly automated and monitored, since the power load, the state of ICT equipment and cooling facilities are usually monitored and adjusted.

The survey developed by Liu *et al.* (2014) indicates that data centers can use 40 times more energy than conventional office facilities, and 5% of the load can typically be shed in 5 minutes and 10% in 15 minutes, with no impact on the ICT procedures. Moreover, if workload management approaches are exploited, the flexibility level can be even larger, without additional time needed to shed the load.

Ghatikar et al. (2010; 2012) concluded that data centers present significant load-reduction potential, nonetheless not all data centers can take advantage of all approaches because of different operational profile. Some strategies are appropriated for energy efficiency, however additional incremental benefits can be reached by temporarily decreasing service levels for a few hours a day and a few days a year for the implementation of DR. Consequently, data center DR strategies generally are divided between load-shedding (reduction or interruption of the load) and load-shifting (moving load from peak to off- peak periods).

A U.S. Federal Energy Regulatory Commission (FERC) assessment lists DR programs presented by Tang et al. (2012; 2014) as: dynamic pricing without enabling technology, dynamic pricing with enabling technology, Direct Load Control (DLC), interruptible tariffs, and other programs, such as capacity/demand bidding and wholesale programs. These programs are categorized into:

- *Price-based, market-led or stability-based programs* offer participants time-varying rates that reflect the value and cost of electricity in different time periods, as presented in Table 2.2.
- *Incentive-based, system-led, reliability-based, or economic-based programs* offer participants discount rates or rebates for their participation or load reduction performance on DR signals, as presented in Table 2.3.

The progress of advanced metering technologies will enable all types of customers to participate in automated DR programs and, in the data center case, taking advantage of the flexible features highlighted by (Irwin *et al.* 2011), as follows:

- Servers are equipped with programmable power management procedures, settling their power consumption by commands from selected interfaces.

- Many workloads are tolerant to delays or performance degrading, allowing data centers to suit the power consumption in response to price fluctuations.
- Data centers consume a massive amount of energy with a substantial impact on grid requirement.

Table 2.2 – Price-based programs (Tang et al. 2014)

Price-Based Programs	
Types	Description
Time-of-Use (TOU)	TOU rates differ in different blocks of time. The rate reflects the average cost of electricity during different periods.
Critical Peaking Price (CPP)	CPP benefits the participants by reducing their energy usage during CPP events.
Real-Time Pricing (RTP)	The price signal of RTP is released a day or an hour or even shorter ahead of the time for which it applies.

Table 2.3 – Incentive-based programs (Tang et al. 2014)

Incentive-Based Programs	
Types	Description
Direct Load Control (DLC)	DLC program operators offer a participant an incentive, usually financial, in the form of credits on the utility bill.
Interruptible/Curtailable Service (CS) Programs	Participants of these programs receive a rate discount or bill credit in return for agreeing to reduce load during certain time periods.
Demand Bidding/Buy Back (DB)	Participants offer their most cost-beneficial bids, price and reducible load, to an electricity market when the price has its highest value. The consumer benefits from cost savings and gaining rewards.
Emergency Demand Response (EDR)	Participants receive incentives for measured load reductions during emergency conditions, however curtailment is voluntary.
Capacity Market (CM)	Participants who commit to providing contracted load reductions when necessary receive incentives.
Ancillary Service Market (ASM)	Participants must adjust huge amount of load quickly when an event occurs. The response duration is typically in minutes rather than hours.

Tang et al. (2012; 2014) argue the achievement of demand control approaches relies on several factors, including: frequency, duration, local weather patterns, or electric grid conditions. However, to undertake a proper DR control strategy it is needed to assemble enough power consumption information of the participating facilities.

Even though simulation results of many studies have been demonstrating that it is possible to improve power grid reliability and provide an important source of economic benefits in DR market, nowadays data centers are largely non-participants due to four main challenges to overcome (Wierman et al. 2014):

1. *Regulation and Market Maturity:* Many of these DR programs are not yet available to data centers in several markets due to the need for adjustments in the regulatory aspects. As a result, the opportunities for their participation may be limited to simple and traditional programs, such as smart and coincident peak pricing, as demonstrated by Liu et al. (2013) and Brocanelli et al. (2014). The first step to overcome challenges in this aspect is transitioning to deregulated market with an independent energy regulator, providing more specific roles and competition in the whole value chain through, for example, aggregators, which are companies that pool the generation or flexible demand capacities of a number of smaller consumers (Flanagan 2013).
2. *Risk Management:* Camacho et al. (2014) claim that data centers prefer to negotiate long-term energy contracts with fixed usage prices because their main business is the maximizing uptime and performance, and energy issues are certainly secondary when compared with the need to maintain strong guarantees about these primary measures. However, the electric sector needs to provide information about this old mindset, using scientific works such as Basmadjian *et al.* (2015), in which new and appropriated data centers contracts are proposed and validated ensuring performance and reliability to data center operations.
3. *Control:* An active debate within the demand response field is about the entity that should have the control of DR actions. Grid operators would like to have a guaranteed response when they ask for it, which leads to “direct load control” programs for which the grid sends a signal to a controller of the program participant. However, this is not always acceptable to participants. The other extreme alternative is “prices-to-devices” where real-time prices are conveyed to participants; nevertheless, such programs typically require huge price variation in order to extract desired responses. This volatility is not acceptable given the risk tolerance of data centers, thus other programs must be developed in order to facilitate data center participation. (Wang and De Groot 2013). This volatility is not acceptable given the risk tolerance of data centers, thus other programs such as the pricing and operation strategy optimized used in (Jin *et al.* 2017) must be developed to facilitate their participation.
4. *Market Complexity:* The complexity to automate and incorporate the bidding process into a data center management system, as well as the high regulation have prevented data centers from entering these markets despite the financial opportunities (Wang *et al.* 2013).

Nevertheless, the ICT industry continues to refine technical capabilities in relation to power-capping, load management and virtualization of workload that will help manage any perceived risk and along with alternatives as microgrid optimal dispatch in Feng *et al.* (2017), changing some complexities.

Therefore, it is fundamental to promote the necessary adjustments with very specific polices in order that the progress made so far can widely also contemplate the emerging reality in the market of small and medium profile data centers.

2.3 RENEWABLE INTEGRATION IN DATA CENTERS

Data centers obtain their primary power from the electrical grid, however many data centers use on-site renewable energy sources, such as solar and wind power. Although progress in rising renewable-energy-powered data centers, with no large-scale Energy Storage Devices (ESDs) the use of available on-site renewable energy resources remains challenging due to the fluctuation in the power demand, as well as the intermittent characteristic of the renewable energy sources (Parra *et al.* 2017).

Gavald *et al.* (2014) claims that new renewable generation in data centers capacity is constructed to serve the infrastructure and typically requires an anticipated investment. In this context, the generated energy generation can be classified in according to Salom *et al.* (2017):

- *On-site generation from on-site renewables:* the renewable energy resources are directly supplied in the site.
- *On-site generation from off-site renewables:* the renewable energy source needs to be provided from outside the building site, instead of the generation of workable forms of energy occurring on the project site, i.e. energy carriers must be transported, such as biomass or biogas.

Other approaches are currently being adopted by data center operators, such as buying renewable energy generated from other organizations, where in this context the data center is no longer an active actor in the provisioning and operation of the energy source. Other important mechanisms in this framework are electricity tracking certificates and renewable electricity products (Depoorter *et al.* 2015).

Several studies address the renewable integration in data center on a wide spectrum. Malkamäki and Ovaska (2012) researched the free cooling potential and solar energy in European data centers, as well as existing interactions between solar energy, air temperature and ensuing data center cooling requirements, concluding that spots with high solar energy generation potential are slightly less ideal for free cooling, due to their higher ambient temperature. Arlit et al. (2012) presented an approach to manage data centers with renewable energy minimizing their dependence on grid power while decreasing capital cost. The load was designed and managed to utilize on-site renewables, mostly photovoltaic, and fully offset the utilization of non-renewable energy from the grid. For this purpose, it was combined the use of RES with dynamic ICT workload scheduling and incorporated management techniques to enhance overall data center utilization while enabling demand to be ‘shaped’ in accordance with the resource availability.

Stewart and Shen (2009) addressed a renewable energy management methodology in data centers stating that the specificities of their workloads reduce the dependence on non-renewable energy. Shuja *et al.* (2016) discuss the case of modular data centers based on shipping containers with capacity to be allocated to optimal sites with on-site availability of renewable energy, free cooling resources and waste heat recovery opportunities. Depoorter *et al.* (2015) suggests that forthcoming data centers rollout could consider site selection as a new strategy to restrict the environmental impact and their energy demand. To suggest so, a dynamic energy model incorporating free cooling and photovoltaic energy was developed assessing indicators of energy usage and the data center behavior located at different representative emplacements in Europe.

Mäsker *et al.* (2016) present the dispersed, intermittent and dissociated profile from energy demand nature in the process of expansion and establishment of renewable energies. By using a low average utilization workload profile, data centers providers can adapt the computational workload to be an energy price dependent through scheduling. The study compared two scheduling strategies for decreasing energy costs. The former using present values from smart meters to run the workloads and the later estimating the future energy price in the energy market based on weather forecasts, having a satisfactory effect on the use of renewable energy and on the mitigation of energy costs.

Combined Heat and Power (CHP) technologies in data centers aim at reducing consumption, by recycling wasted thermal energy, making them more cost-effective and energetically efficient assisting primary power. This technology uses absorption units for recovering heat unloaded by a thermal engine or a fuel cell, providing a decrease of electricity demand from large power plants and reduce congestion in electric transmission and distribution infrastructures as a main application (Darrow and Hedman 2009).

Guizzi et al. (2009) proposed a comparative analysis between a traditional data center and one utilizing CHP to generate electricity linked with an absorption machine to produce cold. Later on, Guizzi and Manno (2012) assessed a CHP system for a 100 kW ICT load data center, using a natural gas membrane steam reformer yielding a pure hydrogen flow for electric power generation in a polymer electrolyte membrane fuel cell. Heat was recouped from both the reforming unit and the fuel cell suppling the requirements of an office building located in the vicinity. The simulations demonstrated that a 47% cost reduction could be achieved when thermal energy from the CHP system is usefully recovered. Thereby, Salom *et al.* (2017) state that the concept is not subject to any geographical restrictions and depends on biogas availability. It might be deployed in very small, as well as large data centers (50 kW to 10 MW). However, the CHP plant demands a certain amount of annual operating hours to be cost-effective. Hence, it is necessary to have an appropriate heat demand available close to data center to absorb the heat especially during winter, by the time it is cooled by means of indirect air free cooling.

Little and Garimella (2012) presented the deployment of geothermal heat pumps to feed a district heating system. Since absorption units demand a heat medium temperature spectrum of 70 – 95 °C for chiller operation, the CPU must operate exceeding those values. Nevertheless, this operation might minimize the reliability of the CPU computing at this temperature interval, requiring further research for characterizing maximum operational CPU temperature will not impact its reliability and efficiency.

A data center in Utah is totally powered by 6 MW fuel cells, making the infrastructure more reliable to grid blackouts and representing an important environmental step. A data center research in Cheyenne, Wyoming was powered by a fuel cell supplied biogas generated by a wastewater treatment facility (Microsoft 2012). A data center in Maiden, North Carolina, doubled the size of the solid oxide fuel cell installation with a total of 10 MW installed capacity (Apple 2013).

Solar power has not been broadly used in data centers, due to the necessity of a large area of photovoltaic panels to generate even a portion of the energy needed by these high-energy density infrastructures. However, there are some successfully implementations, such as a 100 kW solar panel set occupying 730 m² situated on the data center roof in Missouri (Oró *et al.* 2015). For testing the capacity of using photovoltaic solar energy for data centers, a 10 kW of photovoltaic power system was installed in a data center in New Mexico (Oró *et al.* 2015). Goiri et al. (2013) presented a green data center prototype which covers a small container, an array of solar panels, an electrical battery and a grid-tie. In Portugal, a Data Center is 100% maintained by renewable energy; 40% of the energy is supplied by a 400 kW photovoltaic plant comprised by 1610 panels and 60% by a wind farm with 28 towers built near the data center (BCSD Portugal 2017).

In a similar way, a cluster of laptop motherboards supplied by two micro wind turbines and two solar panels was built (Sharma *et al.* 2011). On the other hand, by switching its daily operations energy requirements over to a 500-kW wind turbine, a small data center in Illinois became the first 100% on-site wind power data center in the U.S. (Oró *et al.* 2015) and a data center built a wind farm which is operational since late 2016 and generate 40% of its electrical usage (Nadjaran Toosi *et al.* 2017). Moreover, an Oklahoma's data centers purchased 20 years' wind energy from Iowa wind farm other operators are using the self-generated solar energy or wind energy in order to power their data centers (Gu *et al.* 2016).

As can be observed, the trend of renewable energy implementations in data centers is to grow, given the current requirements of international eco-committees for this purpose and the savings that can be obtained. However, the largest proportion should be in large data centers because they have higher investment capacity and larger available area to install generation systems. Therefore, new policies should be implemented to contemplate small and medium data centers as a singular and energy-intensive market. Projects such as DC4Cities (Klingert *et al.* 2015), which offer a technical and business related solution for optimizing the share of local renewable power sources when operating data centers in smart cities, are essential for the acceleration and integration in this scenario.

2.4 SMALL AND MEDIUM DATA CENTERS

Data centers fall into two general categories: internal and external. Internal data centers are dedicated to the needs of the organization that operates them and typically serve one of two main functions: production or research and development. External data centers provide services to companies that have outsourced some or all of their ICT functions. Both can still be subdivided into small, medium and large size profiles, depending on their purpose, mission and financial resources. Salom *et al.* (2017) present a way to express the dimension of data centers by using ICT power capacity, with the following breakdown:

- Server room: < 50 kW
- Very small data center: 50 – 250 kW
- Small data center: 250 – 1000 kW
- Medium size data center: 1 – 2 MW
- Large data center: 2 – 10 MW
- Very large data center: > 10 MW

However, Whitney and Josh (2014) also claim that the Small and Medium-Sized Organization category comprises four data center types: server closet, server room, localized and mid-tier as summarized in Table 2.4. In this work, a combination of these two approaches will be considered, where server closet, server room, very small, small and localized are examples of small data centers. The medium data center is common in the two classification methodologies.

Table 2.4 – Typical characteristics of data center space types (Masanet et al. 2011)

Space Type	Typical Size (m^2)	Typical ICT Features	Typical Infrastructure System Characteristics
Server closet	< 19	1-2 servers No external storage	Typically conditioned through an office Heating, Ventilation, and Air Conditioning (HVAC) system. Environmental conditions are not as tightly maintained as for other data center types. HVAC energy efficiency associated with server closets is probably similar to the efficiency of office HVAC systems.
Server room	< 47	A few to dozens of servers No external storage	Typically conditioned through an office HVAC system, with additional cooling capacity, probably in the form of a split system specifically designed to condition the room. The cooling system and power backup equipment are typically of average or low efficiency because there is no economy of scale to make efficient systems more first-cost competitive.
Localized data center	< 93	Dozens to hundreds of servers Moderate external storage	Typically use under-floor or overhead air distribution systems and a few CRAC units. Air conditioning units in localized data centers are more likely to be air cooled and have constant-speed fans and relatively low efficiency. Operational staff is likely to be minimal, which makes it likely that equipment orientation and airflow management are not optimized. Air temperature and humidity are tightly monitored. However, power and cooling redundancy may reduce overall efficiency.
Mid-tier data center	< 465	Hundreds of servers Extensive external storage	Typically use under-floor air distribution and CRAC units. The larger size of the center increases the probability that efficient cooling, e.g., a central chilled water plant and central air handling units with variable speed fans, is used. Staff at this size data center may be aware of equipment orientation and airflow management best practices. However, power and cooling redundancy may reduce overall efficiency.
Enterprise-class data center	> 465	Hundreds to thousands of servers	The most efficient equipment is expected to be found in these large data centers. Along with efficient cooling, these data centers may have energy management systems. Equipment orientation and airflow management best practices are most likely implemented. However,

		Extensive external storage	enterprise-class data centers are designed with maximum redundancy, which can reduce the benefits gained from the operational and technological efficiency measures.

Energy efficiency measures, demand response and integration of renewable sources discussed so far were analyzed without considering size, or proportion of data centers. However, while most technical sectors and even scientific community attention are focused on the largest data centers, these hyper-scale cloud computing profile represent only a small portion of overall energy consumption in this market. The extensive majority of data center energy is consumed in small, medium and multi-tenant environments. These profiles have generally made much less advancement than their hyper-scale cloud equivalent due to market barriers, lack of more specific data centers energy efficiency metrics and also misalignment of inducements according to Delforge (2014). The fact that the cost center is generally disassociated from the data center itself, since it is not the main business, is another key factor that influences the investment capacity in adopting new and more efficient technologies.

However, the above-mentioned realities are also applicable to SMDC, dimensions respected. Specifically, what can change is the amount of financial resources available to implement such measures, because in terms of technological level what will be changed is the proportion of a data center resources available.

Regarding data center energy efficiency, there has been significant advance in the last decade, with server farms operated by large companies leading the way. Nevertheless, these hyper-scale cloud computing enterprises account for approximately only 5% of all data center, as depicted by Figure 2.7. The corporate-owned enterprises, small and medium-sized organizations and multi-tenant data centers are far behind in terms of efficiency, requiring focused actions, such as utility incentive programs to reduce waste in the huge amount of electricity used by data centers of all sizes (Delforge 2014).

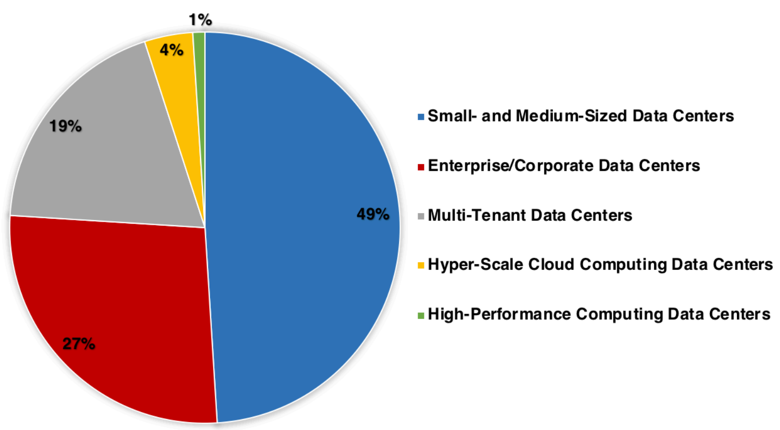


Figure 2.7 – Estimated U.S. data center electricity consumption by market segment (2011) (Josh and Delforge 2014)

Natural Resources Defense Council (NRDC) report states that large, mid-tier, and enterprise-class data centers represent only half of all U.S. servers, as can be seen in Table 2.5. The other half constitutes small server rooms and closets typically available in small and medium businesses and organizations, as well as in research institutes, departments and subsidiaries of larger organizations.

Even in the face of such statistics, in terms of energy efficiency an estimated 20 to 30% of servers (Delforge 2014) in even these large data centers are obsolete or unused because the completion or changes in project profiles, however they are still in operation and consuming electricity every day, for lack of awareness of their real need.

Table 2.5 – Estimated U.S. data center electricity consumption by market segment (2011) (Whitney and Delforge 2014)

Segment	% of stock (based on # of servers)	Average PUE	Average server utilization	Average server age (years)	2011 Electricity Use (GWh)	Server power at average utilization level (W)	Data center market segmentation by electricity consumption
Small and Medium Server Rooms	40	2.0	10%	3	37.5	149	49%
Enterprise/ Corporate Data Centers	30	1.8	20%	2	20.5	120	27%
Multi-Tenant Data Centers	22	1.8	15%	2	14.1	113	19%
Hyper-Scale Cloud Computing	7	1.5	40%	1	3.3	101	4%
High-Performance Computing	1	1.8	50%	2	1	169	1%
	100				**76.4**		**100%**

As the main focus of this present thesis is the small and medium data centers case, the aggregation of these two types of size profile is justified by the fact that the substantial difference between one and the other subsists in the proportion and quantity of equipment, and there are no major technological changes, unlike large data centers, where the complexity, technological profile and energy demand scenario are very different.

In this context, the characterization of such data centers addressing details in terms of thermal loads, temperature limits, electricity consumption, costs and carbon foot print, is primordially important.

From one part the raising demand for the ICT services and from the other part the straight proportionality between data center costs and floor area, have resulted in the design and development of more dense and higher power modules. Nowadays, the design and manufacture of thermal management systems is one of the most challenging aspects of data center design. The thermal management system must deal with the growth thermal loads while remaining the temperature of electronic components at an insured operating parameter. For designing such a system, it is needed to get specific, accurate and reliable information about the maximum thermal loads and temperature limits in each component of a data center as a whole. Therefore, Table 2.6 presents typical and modular data center thermal loads and temperature limits as a way of characterizing this aspect in this type of environment.

Table 2.6 – Summary of "typical" data center thermal loads and temperature limits. (Ebrahimi, et al. 2014)

Power loads	
Component	**Values**
Processors	60 – 75 W each (2 per server)
DIMM (Dual Inline Memory Module)	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	**Value (°C)**
Processor	85
DIMM	85
Disk drive	45

The ICT equipment are the main responsible by energy consumption and heat production in any data center. Typically, these singular infrastructures have had very contained environments due to its particularity. Suitable environmental conditions for electronic equipment divided by temperatures and relative humidity (RH) recommended by the American Society of Heating, Refrigerating and Air-

Conditioning Engineers (ASHRAE) (ASHRAE Technical Committee 9.9 2011) for all equipment classes can be seen in Table 2.7. These thresholds indicate the air inlet conditions in the ICT equipment and hence into the room or the cold aisles in cold/hot aisles settings. A wrong operation of humidity ranges will put at risk the reliability of the computing equipment, since a high humidity will cause water vapor to condensate on the equipment, while very low humidity can cause electrostatic discharges. Thereby the recommendation is a humidity envelope between 20 to 80%.

Table 2.7 – Summary of 2011 ASHRAE thermal guidelines for data centers

Data Centers Classes		Dry-bulb temperature	Humidity range	Maximum dew point
Class	**Definition**	**Recommended**		
A1 - A4		18 to 27 °C	5.5 °C DP to 60% RH and 15 °C DP	-
		Allowable		
A1	Enterprise servers and storage products tightly controlled	15 to 32 °C	20% to 80%	17 °C
A2	Volume servers, storage products, personal computers and workstations with some control	10 to 35 °C	20% to 80%	21 °C
A3	Volume servers, storage products, personal computers and workstations with some control along with use of free cooling techniques when allowable	5 to 40 °C	8% to 85%	24 °C
A4	Volume servers, storage products, personal computers and workstations with some control along with near full-time usage of free-cooling techniques	5 to 45 °C	8% to 90%	24 °C

In relation to the energy consumption, Table 2.8 presents the case of a medium data center whereby some ICT features are demonstrated, such as the amount of racks, servers, purchase year of servers and the total data center energy use per year. This type of information is essential to compare and analyze different data center profiles due to specificities and detail level.

Regarding a cost analyses, the increase in demand for computer resources has been leading to a raise of servers' numbers in data centers and correspondingly a considerable growth in power and capital costs for upgrading of capacity and development of new data centers. As with up-to-date computers, cloud environments provide multicore CPUs in the high-performance tires, which in fact demand parallelism for a cost-effective exploitation of the assets (Uddin *et al.* 2015). Consequently, the cost per process reduces since the number of processes performs in parallel in the instance rises up to the number of available cores, considering the dimension of data center standard costs related to components and sub component as presented in Table 2.9.

Table 2.8 – Hypothetical data center as a case study. (Fulpagare and Bhargav 2015)

Area (m²)	185
Number of racks	500
Servers per rack	20
Total servers	1000
Server vintage (year)	2006
Average server power (W)	273
Average data center connected load (total of all ICT) (kW)	273
Peak server power	427
Annual ICT energy use (kWh)	2,391,480
Annual energy loss in UPS to ICT (kWh)	144,000
Annual energy loss in PDU (kWh)	96,000
Annual energy use in lighting/security (kWh)	60,000
Annual energy use in chiller plant (kWh)	450,000
Annual energy use in CRAC units (kWh)	700,000
Annual energy loss in UPS to cooling plant (kWh)	50,000
Total energy use (crossing the data center boundary) (kWh)	3,891,480
Total ICT energy use (kWh)	2,391,480

Table 2.9 – Data center costs, component, and sub-components. (Uddin et al. 2015)

Cost	**Component**	**Sub-components**
~45%	Servers	CPU, memory, storage system
~25%	Infrastructure	Power distribution and cooling
~15%	Power draw	Electrical utility costs
~15%	Network	Links, transit, equipment
Standard costs of data center components		
Cost per CPU core	$0.040/h	
Cost per 1GB RAM	$0.025/h	
Cost per 1 GB storage	$0.0003/h	
Cost per server	$3.243/h	

In the context of carbon foot print it is important to highlight that greenhouse gases (GHG) emissions from aviation's, shipping, transportations, telecommunications and manufacturing industry are rising significantly. However the emissions from ICT are growing faster (Añón Higón *et al.* 2017). Reductions reached using Green ICT in strategic economic sectors would be five times superior than the increase in emissions from the ICT sector itself. Additionally, emissions from ICT is projected to rise from 3% of total global emissions in 2007 to a huge 6% by 2020 as can be seen in Table 2.10 *(Uddin and Rahman 2012*.

Table 2.10 – CO_2 emissions (carbon foot print) climate group and the global e sustainability initiative SMART 2020 (Uddin and Rahman 2012).

World	Emission 2007 $MtCO_2e$	Percentage 2007	Emission 2020 $MtCO_2e$	Percentage 2020
Server farms/data centers	116	14%	257	18%
Telecom infrastructure and devices	307	37%	358	25%
PCs and peripherals	407	49%	815	57%
Total	830	100%	1430	100%

Finally, this chapter has focused in researching the several available technologies to promote energy efficiency, the most prominent and recently discussed by various authors summarizing the most relevant aspects focusing on dismemberment strategy of computational components of servers, software solutions such as virtualization, cooling technologies, energy storage and management, addressing also the integration of renewable energy sources. All these aspects have been discussed on the framework and the perspective of SMDC, discussing demand response programs as a prominent solution to reduce costs and collaborate with the grid to pursue a higher joint reliability and sustainability in this process.

CHAPTER 3

ENERGY EFFICIENCY ASSESSMENT FRAMEWORK

In this chapter, based on the previous assumptions, an energy efficiency framework will be discussed, where three current surveys will be presented underlining their premises as well as their conclusions, one of which was conducted within the framework of this thesis. Furthermore, on a proposal basis, three consolidated energy efficiency methodologies are analyzed, compared and assessed to allow a more appropriate use in the context of the SMDC.

3.1 ENERGY EFFICIENCY SURVEYS

In order to verify the data centers energy efficiency reality, specifically in SMDC, three surveys with different approaches are presented. The first two surveys provide a comparative and argumentative basis from the literature and the third one was carried out in the context of this thesis.

The Green Grid (2016) questioned 150 key European ICT decision makers with data center responsibilities in the UK, France and Germany. It was found that while most organizations are facing growing pressures to improve the efficiency of their data centers, 43% of those surveyed have no energy efficiency objectives in place, as presented in Figure 3.1. Key findings addressed include:

- Energy efficiency and operating costs are the most common areas of the data center reported as requiring improvement, as can be seen in Figure 3.2.
- Two in five respondents reported that their data centers are expensive to run (48%) or upgrade (41%), demonstrating that cost is the most commonly reported impact of data center operations.
- The difficulty in predicting future costs (43%) and the cost of refreshing hardware (37%) are cited as the top challenges to developing resource efficient data centers, along with the difficulty in meeting environmental targets (33%).

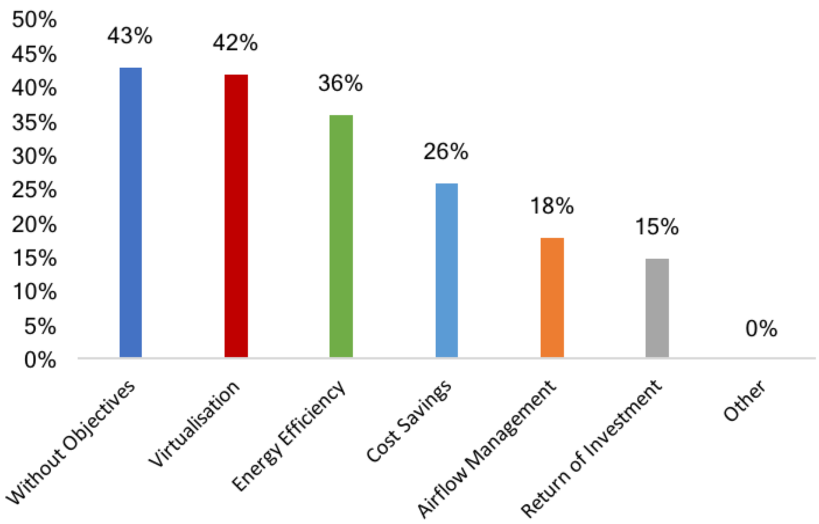


Figure 3.1 – Organizations' energy efficiency objectives (The Green Grid 2016)

Furthermore, 97% felt that they could improve their monitoring capabilities. Some of the findings presented a positive outlook for future innovations in data center resource efficiency, with nearly all those surveyed clearly seeing areas for improvement and 55% stating that energy efficiency was their highest priority when making changes.

However, to match the European Commission's expectation for data centers to be at least 80% powered by renewable energy by 2020, ICT leaders will need to commit to renovate their resource efficiency policies, since the share of renewable generation in data centers is still very low (The Green Grid 2016).

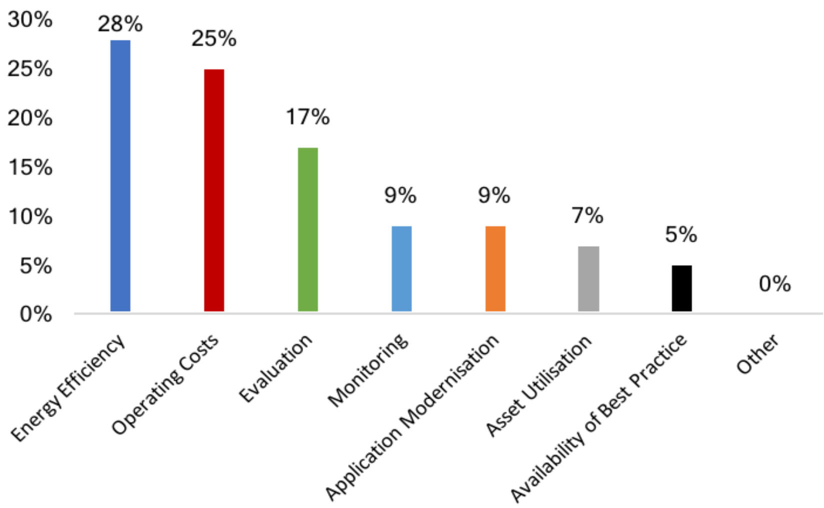


Figure 3.2 – Areas requiring most improvement (The Green Grid 2016)

The other survey, carried out by NRDC (Bennett and Delforge 2012), surveyed 30 U.S. ICT managers. The focus was on virtualization and server purchasing and replacement, since they are the largest and most cost-effective savings opportunities profile in SMDC. From the energy efficiency point of view, it is possible to reduce operating costs by replacing and using greener equipment, as well as to decrease the amount of hardware in use and idle processing time, improving the utilization of resources and providing energy savings through virtualization. Regarding DR, one of the used techniques to exploit the flexibility potential is the use of virtualization on servers to perform workloads shifting, or shedding, taking advantage of a time window where the energy cost is more profitable. In this context, the questions addressed various issues related to their current server fleet, virtualization practices, cloud computing, obstacles to implementing efficiency improvements, energy use, and billing. The results from this survey compared with results from other survey of large companies' virtualization practices conducted by the market research firm (Vanson-Bourne 2011), as presented in Figure 3.3 have shown that:

- Small and medium-sized businesses operate servers in a variety of ownership configurations that may make efficiency upgrades challenging. For example, if the company decides which hardware to lease or purchase, but the host pays the electricity bills at the off-site location, the server operator does not have a financial incentive to optimize the energy efficiency of their servers. This split-incentive situation is similar to the owner-tenant problem so common in commercial buildings.
- Indicators of virtualization adoption by smaller businesses lag behind larges ones. Most companies have tried virtualization and plan to do more in the future, but still have progress to make before achieving deep transitions to a mostly virtual server fleet. Their surveys' results revealed starkly different results, especially in the percentage of companies that have tried

virtualization and the percentage that plan to use or increase virtualization in the future. Whereas 90% of large enterprises have virtualized at least one server, only 37% of small companies have done so. This shows that there is still a lot of education and marketing to be done in this sector.

- When virtualization is used, small organizations tend to implement it more broadly than large ones, which narrows the penetration gap to 11%. However, there remains a large untapped virtualization opportunity in both markets and particularly in small and medium-sized organizations, where only 26% of all server stock has been virtualized.
- Almost all large companies have used virtualization, and many say they plan to increase its use in their operations, increasing the virtualization penetration rate. In contrast, only 23% of small companies said they plan to increase their virtualization in the next 12 months.
- Small companies have not adopted virtualization because of unaligned incentives and lack of information. Currently, 60% of the staff that make server purchasing decisions do not have access to their company's energy bill. This is critical, as server rooms can account for anywhere from 30 to 70% of an organization's electricity consumption (particularly in office-based organizations). Because over 90% of organizations do not have a way of monitoring server room electrical use, this opportunity is being overlooked as a strategic way to seriously reduce overhead costs and environmental impact.
- Half of all organizations surveyed plan on a server room upgrade in the subsequent year to the survey. Replacement of servers coming to the end of their warranty is the most cost-effective way to implement efficiency best-practices, since project benefits include the cost avoidance of investing in 1-for-1 server hardware and software replacement.

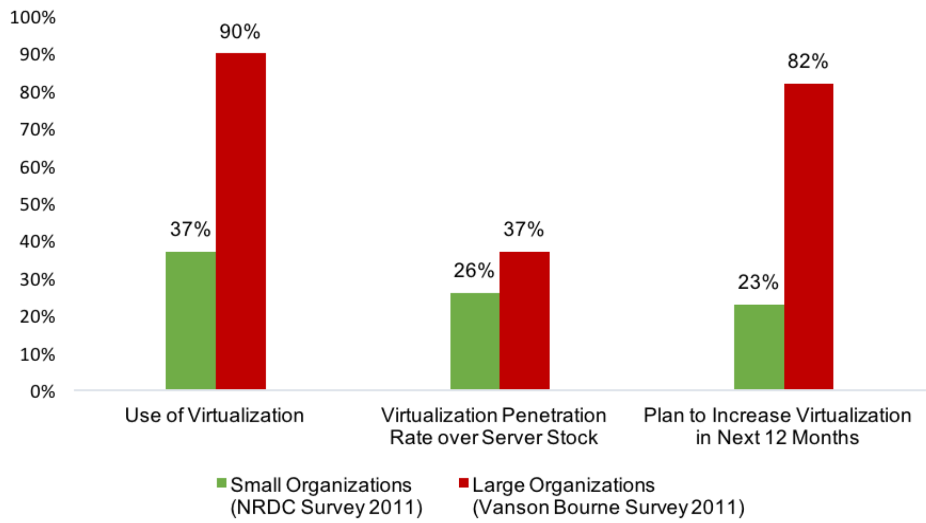


Figure 3.3 – Comparative survey results (Bennett and Delforge 2012)

Based on information from the main surveys available in the literature, their characteristics, particularities and focus, one of the tasks of this research was to develop its own survey aimed at better understanding the SMDC environment. Hence, it would be possible to analyze specific contexts in different countries, with several diversity of energy efficiency patterns and actions taken by SMDC actors.

The 27 questions were addressed to key decision makers with technical and management responsibilities for these data centers. The survey obtained 22 responses. Since the total population (number of data centers) is not high (at least comparing it with other types of buildings and facilities) it is normal to have the assessment based on a relatively small number, of relevant, responses. Such number is not too different from the 33 responses obtained in the survey conducted in the U.S. by NRDC. The 22 responses are distributed between the three different countries, as follows: Brazil with 13 responses, the U.S. with 5 and Portugal with 4. The response rate was 65% in Brazil, 50% in the U.S. and 40% in Portugal.

The main objective was to analyze the energy efficiency of small and medium data centers in Brazil. The U.S. was chosen as a North America representative and Portugal as the European Union representative, of regions where surveys had already been published, to provide a good comparison.

Two questions were asked to map the data center respondent's professional role, and their institutional profile. As depicted in Figure 3.4, the major professional role in the responses was technical or encompassing both management and technical responsibilities. The planning/design and just the management role had 2 and 1 responses, respectively. In addition, most responses were from the educational sector, such as: universities, laboratories, or companies related to this type of activity in the different participants' countries. There were two responses from business, electricity and electoral justice government sector, contrasting with only one from healthcare sector.

Respondents were asked about the size of their data centers based on floor area. Floor area was chosen because many servers are expected to be located in server closets and server rooms, which have different technology characteristics — and, hence, different efficiency opportunities — than larger data centers. It also facilitates the characterization of electricity costs and potential cost savings, since small and medium data centers are the objects of this work.

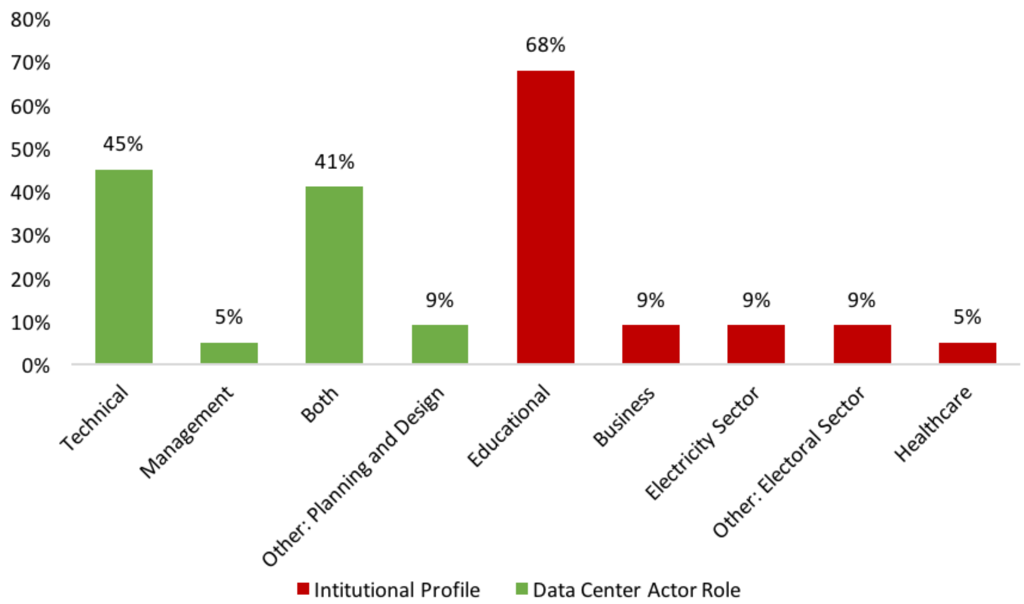


Figure 3.4 – Institutional profile and data center actor role

The distribution of size of data center collected by this survey is presented in Figure 3.5. This shows that 65% of those surveyed operated a small data center, and 32% a medium. Just one response came from the operator of an enterprise class data center. Howsoever, it is important to highlight that this survey was directed to just the small and medium-sized centers operators and managers.

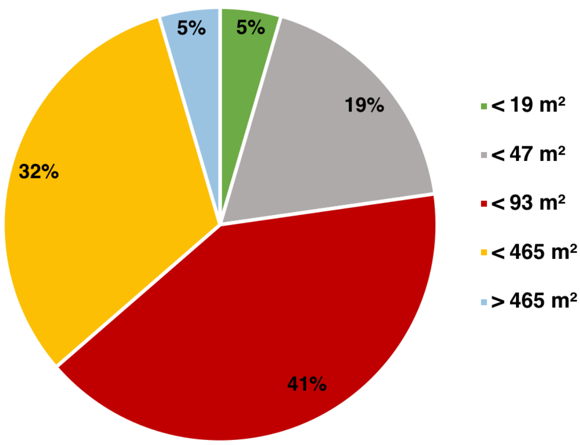


Figure 3.5 – Data center size distribution

Thereafter, a set of questions was asked in order to find out the level of energy efficiency of the servers, such as:

- How many servers does the data center have?
- How many legacy servers are there?
- What is the number of servers with some energy efficiency label?

- How many Dynamic Voltage and Frequency Scaling (DVFS)-enabled servers does the data center have?

The average obtained proportion, considering the order of questions, was 10-3-8-2, i.e., for each set of 10 servers, 3 are legacy, 8 have some energy efficiency label and only two have DVFS enabled.

In order to verify energy efficiency actions at the storage devices level the following questions were asked:

- How many storages appliances does the data center have?
- What is the maximum storage capacity?
- What is the current raw storage?
- How many storage appliances with energy efficiency label does the data center have?

The average proportion was 1-1, i.e., all the actors who have storage appliances in their facilities have already acquired them with some energy efficiency label. However, the average proportion to those data centers which do not have storages appliances in their facilities was 10-3, i.e., for each 10 data centers, nearly 3 of them do not have storage appliances.

A set of questions was asked on cooling solutions: which are the most popular solutions in this data center profile, the amount of equipment in the facility, and their joint power consumption. 15 out of 22 (69%) of participants have answered these questions with the following outcome:

- Split: 55%
- Hot and Cold Aisle: 45%
- Chillers: 27%
- High Precision: 18%
- In Row: 18%
- Free cooling: 10%
- Rack Air Distribution: 0%
- Perimetral: 0%
- Other: 0%

The participants were able to choose more than one solution and split cooling is the most used solution in small and medium-sized data center profiles, followed by hot and cold aisle, whereas solutions such as free cooling tends to be adopted in larger spaces.

The rest of the questions sought to assess, through a yes or no methodology, if servers, storage appliances and network devices have energy monitoring. Respondents were also asked about the adoption of simple measures, such as the use of metrics like PUE, the annual calculation of data center energy consumption and the knowledge of the electric load diagram. Figure 3.6 presents the responses and the results are alarming from the energy efficiency point of view as the energy consumption of most of the servers, storage appliances and network devices are not monitored. A higher proportion did not calculate annual energy consumption. Moreover, 14% of respondents were unaware of metrics such as PUE and 18% did not know if there was an electrical load diagram in the facility.

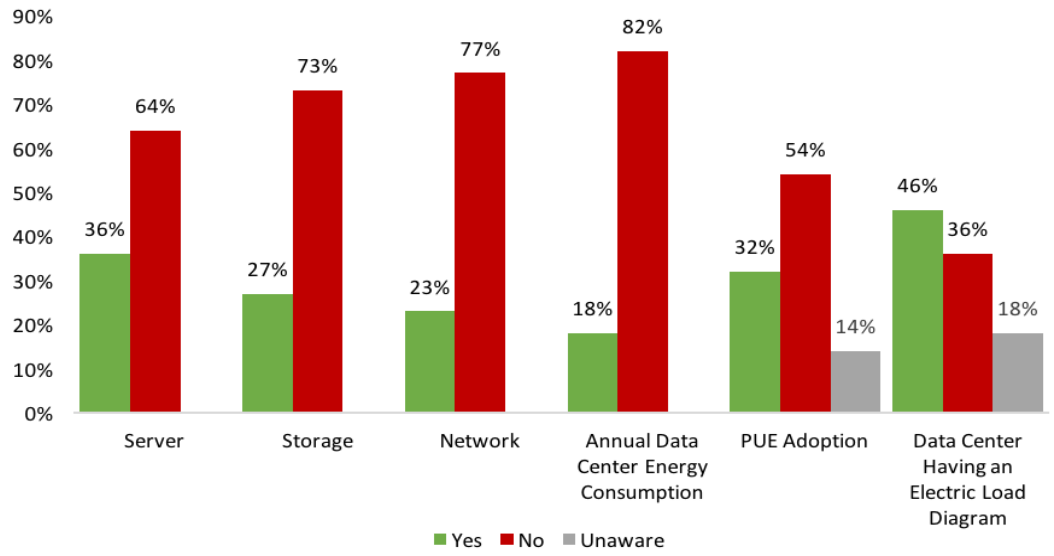


Figure 3.6 – Energy efficiency monitoring actions

The answers to the questions on monitoring actions can be used to calculate a propensity for energy efficiency actions by country, as shown in Figure 3.7. For the overall number of answers, the average data center actors most effectively participating in energy efficiency actions are 4 out of 5 for the U.S., 2 out of 4 for Portugal and 3 out of 13 for Brazil.

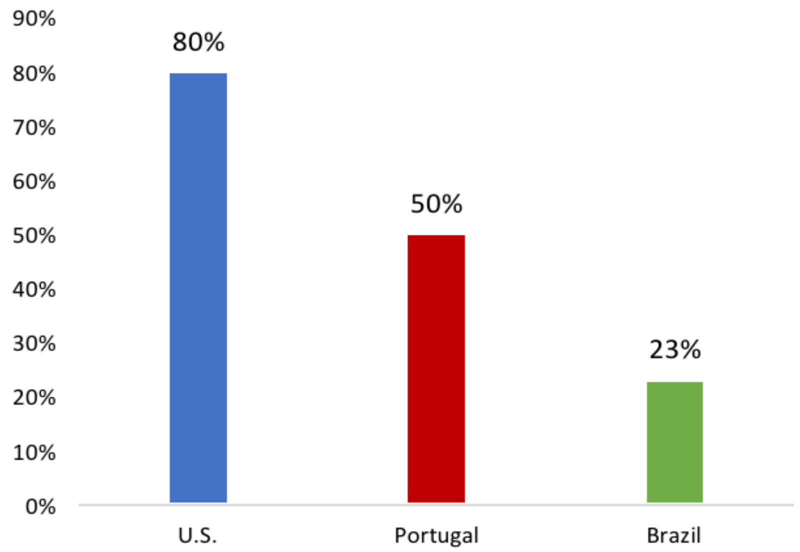


Figure 3.7 – Energy efficiency action tendency

The answer to the last two survey questions can be used as indicators to establish a correlation with the low index of actions in energy efficiency by the majority of those surveyed, where nearly 73% answered that another sector is responsible for managing the overall data center energy consumption and nearly 91% answered that payment of energy bill is the responsibility of the institution that owns the data center the and not the data center itself.

3.2 IMPACT ASSESSMENT OF ENERGY EFFICIENCY METHODOLOGY

Based on the worrisome reality in terms of energy efficiency demonstrated by the specificities of the three presented surveys, especially in SMDC, it is important to analyze methodologies already consolidated in the literature that, when implemented in the most appropriate context, allow the change of this scenario.

Thus, the main purpose of this section is to discuss three different energy efficiency methodologies and compare them highlighting the best scenario for implementations, as well as their advantages and disadvantages in an impact assessment. The chosen methodologies were respectively proposed by “Energy Logic”, “Lawrence Berkeley National Laboratory” and “National Renewable Energy Laboratory” and the selection criterion prioritized the relevance of the research institutes, how updated they are and the number of mentions during the literature review stage.

3.2.1 Energy Logic Methodology (EL)

The methodology presented in Emerson 2015 highlights the best energy efficiency practices applied to a 465 m^2 data center based on real-world technologies and operating parameters, as depicted in Figure 1.2. Through this model, it was possible to quantify the savings in five years due to each action at the system level, as well as to assess how energy reduction in some systems affects consumption in supporting systems. The results have shown nearly 50% reduction in data center energy consumption without compromising performance or availability, as presented in Table 3.1.

The model indicates that decreases in energy consumption at the ICT equipment level have the greatest impact on total consumption because they cascade across all supporting systems, as can be demonstrated in Figure 3.8. However, to ensure these savings a sequential approach to reduce energy costs is needed, applying the 10 technologies and best practices that exhibited the most potential in the order in which they have the greatest impact. Even though the sequence is not fundamental because it is not a step-by-step methodology, the energy-saving measures contained should be considered a guide with flexibility in implementing this sequence depending on each context. The first measurement in this approach is establishing an ICT equipment procurement policy that exploits the energy efficiency benefits of low-power processors and high-efficiency power supplies. Thus, inefficient servers will be removed and replaced with higher-efficiency units, providing a more energy-optimized data center.

Table 3.1 – Benefits from efficiency improvement actions (Emerson 2015)

Efficiency Improvement Area		Power Reduction (kW)	Estimated Cumulative Yearly Reduction				
			Year 1	Year 2	Year 3	Year 4	Year 5
ICT Polices	Low-power processors	111	6	22	45	78	111
	High-efficiency power supplies	124	12	43	68	99	124
	Server power management	86	9	26	43	65	86
ICT Projects	Blade servers	7	1	7	7	7	7
	Virtualization	86	9	65	69	86	86
Best Practices	Higher voltage AC power distribution	20	0	0	20	20	20
	Cooling best practices	15	15	15	15	15	15
	Variable-capacity cooling	49	49	49	49	49	49
Infrastructure Projects	High-density supplemental cooling	72	0	57	72	72	72
	Monitoring and optimization	15	0	15	15	15	15
Total		**585**	**100**	**299**	**402**	**505**	**585**

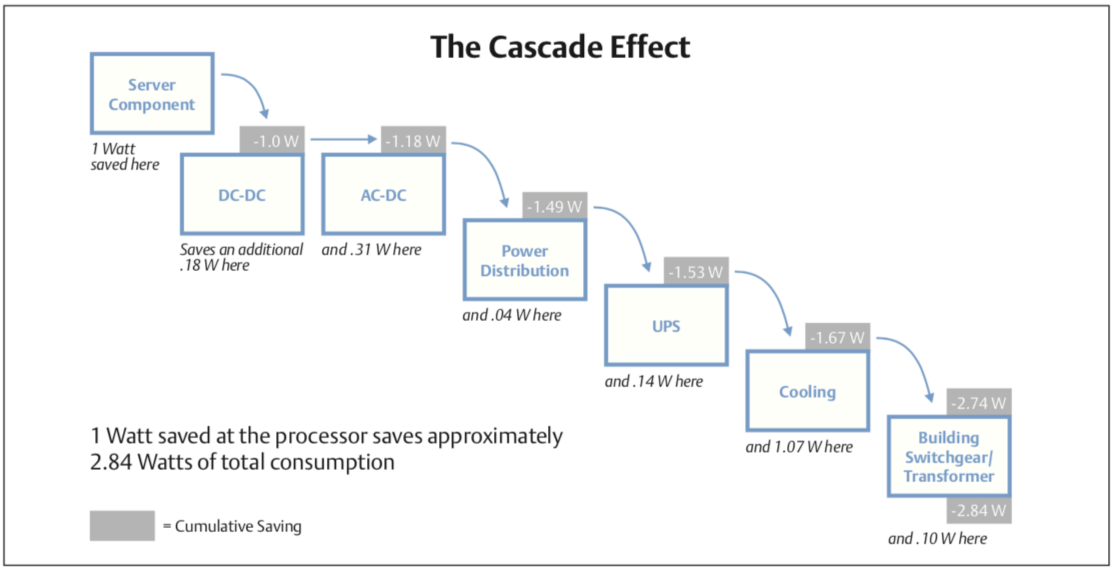


Figure 3.8 – Cascade effect for cumulative savings (Emerson 2015)

Power management software has great capacity to decrease energy costs and should be considered as part of an energy optimization strategy, mainly for data centers with large differences between peak and average utilization rates.

The next initiative involves ICT projects with impact on energy consumption such as the utilization of blade servers — which is a server architecture that houses multiple server modules ("blades") in a single chassis and it is widely used in data centers to save space and improve system management — and server virtualization techniques. These technologies have arisen as “best practice” procedures to data center management and are fundamental in the process of optimizing a data center for efficiency, performance and manageability.

After applying policies and plans to optimize ICT systems, the focus changes to supply-side systems. The most effective procedures to infrastructure optimization involve cooling best practices, higher voltage AC power distribution, variable-capacity cooling, supplemental cooling, monitoring and optimization. Employing these measures on the model reduced the energy use by 52% without compromising performance or availability.

In its unoptimized state, the 465 m² data center example used to develop the methodology had a total computational load of 588 kW and total facility load of 1,127 kW. Through the presented optimization strategies, this facility has been changed to allow the same level of performance using eventfully less power and space. Overall compute load was decreased to 367 kW, while rack density was raised from 2.8 kW per rack to 6.1 kW per rack. The number of racks required to support the compute load as reduced from 210 to 60 and eliminated power, cooling and space limitations constraining growth. Total energy consumption was reduced to 542 kW and the total floor space required for ICT equipment was reduced by 65%.

This methodology intends to be adaptable for every type of data center, however the sequence might be affected by facility profile. 24/7 facilities will want to focus initial efforts on sourcing ICT equipment with low-power processors and high-efficiency power supplies. Facilities that experience predictable peaks in activity might reach the greatest results from power management technology.

3.2.2 Lawrence Berkeley National Laboratory Methodology (LBNL)

The methodology proposed by Masanet et al. 2011 presents a general bottom-up approach, which enables to model an estimation of energy demand in different data center space types, considering energy use of servers, storage and network devices in order to estimate the potential electricity savings associated with a set of broad data center efficiency improvements. The study further estimated that

2011 electricity demand could be reduced by as much as 70% through adoption of energy efficient technologies and operating practices.

The approach facilitates analysis of energy demand in five data center space types already presented in Table 2.4: server closets, server rooms, localized data centers, mid-tier data centers, and enterprise-class data centers.

The data center energy model to estimate the energy use and efficiency potential in data centers is described in general form by Equation 3.1.

$$E^{DC} = \sum_{j} \left[\sum_{i} E_{ij}^{S} + E_{j}^{ST} + E_{j}^{N} \right] PUE_j \tag{3.1}$$

where,

E^{DC} = data center electricity demand (kWh/year)

E_{ij}^{S} = electricity used by servers of class i in space type j (kWh/year)

E_{j}^{ST} = electricity used by external storage devices in space type j (kWh/year)

E_{j}^{N} = electricity used by network devices in space type j (kWh/year)

PUE_j = power utilization effectiveness of infrastructure equipment in space type j (pu)

In Equation 3.1, the total electricity use of ICT devices within a given space type is established through summation of the electricity utilization of servers, external storage, and network equipment. The total electricity use of ICT devices is then multiplied by an assumed power utilization effectiveness (PUE) for that space type. The variables in Equation 3.1 rely on several parameters correlated with to the adoption of energy efficiency measures. This feature enables the model to estimate current electricity demand, as well as potential electricity savings in different measure implementation scenarios. The measures inserted in the model capture the major classes of data center devices and operations efficiency strategies listed in the EPA study Dreibholz et al. 2007, which extensively reviewed such strategies.

An important remark is that a number of calculations in the model are made relatively to static baseline values that reflect current data center features. This enables the estimation of electricity savings potentials between scenarios in a consistent mode.

Equation 3.2 is applied to estimate server electricity utilization by space type based on server class, the number of servers in each space type, and the annual electricity utilization per server in each class. The model employs three server class definitions from market research firm International Data Corporation (IDC) (Koomey 2011), measuring power data by server class, and estimating infrastructure device energy use based on unit sales prices: volume servers (<$25,000), mid-range servers ($25,000 to $500,000), and high-end servers (>$500,000).

$$E_{ij}^{S} = {\check{N}_{ij}^{S}}\Big/{\rho_{ij}^{S}}\, e_{ij}^{S} \tag{3.2}$$

where,

E_{ij}^{S} = electricity used by servers of class i in space type j (kWh/year)

$\check{N}_{ij}^{S}$ = baseline number of servers of class i installed in space type j

ρ_{ij}^{S} = device reduction ratio for servers of class i in space type j

e_{ij}^{S} = annual electricity use per server of class i in space type j (kWh/year)

Equation 2 estimates the number of installed servers in each class using a baseline value—defined as the actual number of installed servers—divided by a "device reduction ratio." The device reduction ratio accounts for the relative reduction in servers that can occur through efficiency strategies that diminishes server counts, such as virtualization, consolidation of applications, and legacy server removal. For example, a device decreasing ratio of 3 specifies that three servers have been replaced by one server (i.e., a 3:1 reduction ratio). Annual electricity use per server is estimated using Equations 3, which demonstrated the relationships between server electricity utilization and the enactment of key efficiency measures.

Namely, the potentials for three major efficiency strategies are characterized: (1) use of efficient server hardware; (2) use of dynamic frequency and voltage scaling (DFVS); and (3) reducing the number of physical servers. Efficient server hardware relates itself broadly to hardware measures such as high efficiency power supplies, multiple-core processors, more efficient memory, and variable speed fans. Equation 3.3 presents the net effect of such measures relatively to baseline server electricity use for each server class. DFVS is a usual energy saving characteristic that ensures a processor's clock speed to ramp down during intervals of low utilization, hence reducing power use. The fractions of a server population with efficient hardware and DFVS enabled can be differed in Equation 3.3 to estimate server electricity utilization at different levels of measure adoption.

$$e_{ij}^{S} = \check{e}_{ij}^{S}(\alpha_{ij}^{S}(\gamma_{ij}^{S} - 1) + 1)(\beta_{ij}^{S}\delta'_{ij} + (1 - \beta_{ij}^{S})\delta''_{ij}) \tag{3.3}$$

where,

e_{ij}^{S} = annual electricity use per server of class i in space type j (kWh/year)

$\check{e}_{ij}^{S}$ = baseline annual electricity use per server of class i in space type j (kWh/year)

α_{ij}^{S} = fraction of servers of class i in space type j with energy efficient hardware

γ_{ij}^{S} = ratio of efficient server to baseline server electricity use for servers of class i in space type j

β_{ij}^{S} = fraction of servers of class i in space type j with dynamic voltage scaling enabled

$\delta'_{ij}, \delta''_{ij}$ = DFVS and utilization factors

The net effect of decreasing the number of physical servers is assigned in Equation 3.3 through two "DFVS and utilization factors." These two factors represent the dynamic relationship between the number of installed servers that exist after device reduction initiatives, the average processor use of these remaining servers, and the utilization of DFVS.

In a virtualization scenario, several physical servers are replaced by "virtual" servers that reside on a single physical "host" server. An important implication is that the processor use of the remaining host servers will rise due to the increased computational demand necessary to support the virtual servers. In spite of the increase in server electricity utilization that accompanies virtualization, data centers can notice substantial electricity savings through large decreasing in the number of servers.

Equations 3.4 and 3.5 measure the DFVS and usage factors based on server power-utilization functions. For simplification, these functions are assumed to be linear and are thus described using slopes and y-axis intercepts in the model.

$$\delta'_{ij} = {(m_{ij}^{ON} u_{ij} + b_{ij}^{ON})}\Big/{(m_{ij}^{OFF} \check{u}_{ij} + b_{ij}^{OFF})} \tag{3.4}$$

where,

m_{ij}^{ON} = slope of power-utilization function (DFVS enabled) for server class i in space type j

u_{ij} = post-reduction processor utilization per server of class i in space type j (%)

b_{ij}^{ON} = Y-intercept of power-utilization function (DFVS enabled) for server class i in space type j

m_{ij}^{OFF} = slope of power-utilization function (DFVS disabled) for server class i in space type j

$\check{u}_{ij}$ = baseline processor utilization for active servers of class i in space type j (%)

b_{ij}^{OFF} = Y-intercept of power-utilization function (DFVS disabled) for server class i in space type j

$$\delta''_{ij} = {(m_{ij}^{OFF} u_{ij} + b_{ij}^{OFF})}\Big/{(m_{ij}^{OFF} \check{u}_{ij} + b_{ij}^{OFF})} \tag{3.5}$$

where,

m_{ij}^{OFF} = slope of power-utilization function (DFVS disabled) for server class i in space type j

u_{ij} = post-reduction processor utilization per server of class i in space type j (%)

b_{ij}^{OFF} = Y-intercept of power-utilization function (DFVS disabled) for server class i in space type j

$\check{u}_{ij}$ = baseline processor utilization for active servers of class i in space type j (%)

The average utilization per server, after device reduction, is estimated through Equation 3.6. The post-reduction use is a function of four variables: (1) the device reduction ratio for servers (defined as the baseline number of installed servers divided by the number that remain after server reduction); (2) the baseline utilization of active servers prior to reduction; (3) the fraction of removed servers that are legacy servers; and (4) the average utilization "overhead" of virtualization software. Legacy servers are those that are functionally obsolete (e.g., hosting applications that are no longer used), but still draw power. Although the presence of legacy servers varies greatly by data center, some industry analysts indicate that they can comprise up to 10% (or more) of the server population at a typical large data center. For simplifying, it is assumed that legacy servers have negligible usage and will be completely removed in server reduction efforts; thus, they have no impact on post-reduction processor utilization. The utilization overhead variable accounts for the processor utilization growth needed to run virtualization software on the remaining host servers. This software overhead is in addition to utilization increases related to the computational demands of virtual servers.

$$u_{ij} = \check{u}_{ij}\rho_{ij}^{S}\left(1-\theta_{ij}^{S}\right) + \acute{u}_{ij} \tag{3.6}$$

where,

u_{ij} = post-reduction processor utilization per server of class i in space type j (%)

$\check{u}_{ij}$ = baseline processor utilization for active servers of class i in space type j (%)

ρ_{ij}^{S} = device reduction ratio for servers of class i in space type j

θ_{ij}^{S} = baseline fraction of servers of class i in space type j that are legacy servers

$\acute{u}_{ij}$ = Post-reduction processor utilization overhead per server of class i in space type j

Equation 3.7 is utilized to calculate the electricity use of external storage equipment by space type. The electricity use of external storage is expressed as a function of the baseline (i.e., current) number of installed devices, the device reduction ratio, baseline storage device electricity use, and assumed adoption levels of key efficiency measures. Equation 3.7 highlights the savings potentials associated with two broad efficiency strategies: (1) efficient storage devices and management; and (2) reducing the number of external storage devices. Efficient storage devices and management refers to actions aimed at improving the efficiency of both the physical device (e.g., a switch to high efficiency hard disk drives (HDDs) and solid state devices (SSD)) and data management (e.g., tiered storage and/or spinning down HDDs). Device decreasing strategies for external storage include actions, such as data de-duplication, virtualization, and growing capacity utilization.

$$E_j^{ST} = \frac{\check{N}_j^{ST}}{\rho_j^{ST}} \check{e}_j^{ST} \left(1 + \alpha_j^{ST} \left(\gamma_j^{ST} - 1\right)\right) \tag{3.7}$$

where,

E_j^{ST} = electricity used by external storage devices in space type j (kWh/year)

$\check{N}_j^{ST}$ = baseline number of external storage devices installed in space type j

ρ_j^{ST} = device reduction ratio for external storage in space type j

$\check{e}_j^{ST}$ = baseline annual electricity use per external storage device in space type j (kWh/year)

α_j^{ST} = fraction of energy efficient external storage devices in space type j

γ_j^{ST} = ratio of efficient external storage device to baseline external storage device electricity use in space type j

The model estimates the electricity usage of network devices as a fraction of total ICT electricity demand for each space type using Equation 3.8 (rather than in the bottom-up fashion used for servers and storage devices). Thus, the model allows the use of available (albeit limited) data on network devices in a mode that is consistent with the manner those data are reported. Still, Equation 3.8 could be used to coarsely calculate the effects of network efficiency improvements by adjusting downward the network device scaling term (i.e., the second term within the brackets).

$$E_j^N = \sum_j \left[\left(\sum_i E_{ij}^S + E_j^{ST} \right) \left(\varepsilon_j^N \Big/ (1 - \varepsilon_j^N) \right) \right] \tag{3.8}$$

where,

E_j^N = electricity used by network devices in space type j (kWh/year)

E_{ij}^S = electricity used by servers of class i in space type j (kWh/year)

E_j^{ST} = electricity used by external storage devices in space type j (kWh/year)

ε_j^N = ratio of network device to total ICT device electricity use in space type j (kWh/kWh)

The electricity use of infrastructure devices is estimated by an assumed PUE for each space type. Equation 3.9 is utilized to calculate each PUE, based on assumptions for the electricity use of four major infrastructure system components: power transformers, UPS, cooling systems, and lighting. The cooling systems component represents the broadest class of infrastructure devices in the model, including primary refrigeration units (e.g., air conditioners and water chillers), coolant pumps, fans and air handlers, cooling towers, and similar devices. Since the types and configurations of such equipment vary greatly across data centers, cooling system electricity utilization is represented in aggregate by space type. The effects of efficiency measurements are estimated through modifications to the ratio of component to ICT device energy demand in Equation 3.9.

$$PUE_j = 1 + \sum_j e_{jk}^I \tag{3.9}$$

where,

PUE_j = PUE of infrastructure equipment in space type j (kWh/kWh)

e_{jk}^I = ratio of electricity use by infrastructure system component k in space type j to ICT device electricity use in space type j (kWh/kWh)

3.2.3 National Renewable Energy Laboratory Methodology (NREL)

The methodology addressed in Masanet and Robert (2014) examines the techniques and analysis methods utilized to verify savings from improving the efficiency of two specific parts of ICT equipment: servers and data storage. The discussion analyzes the premise using more efficient server and data storage equipment, as well as managing them to work more efficiently through measures such as:

- **Server virtualization:** where instead of operating many servers at low utilization rates, virtualized environments will combine the processing power onto fewer servers, operating at higher total utilization rates.
- **More efficient servers:** using ENERGY STAR servers the energy consumption can be reduced in 30% when compared with standard servers. The servers run efficiently at low loads, since the processor power management demands lower power consumption by the time servers are idle and provide the follow characteristics:
    - Efficient power supplies to limit power conversion losses.
    - Improved power quality.
    - Idle power draw limits for rack-mounted or pedestal servers with one or two processors.
    - Results of the Server Efficiency Rating Tool (SERT) tests to accommodate comparisons of server efficiency under various usage scenarios.

- Ability to measure real-time power use, processor utilization, and air inlet temperatures.
- Advanced power management features and efficient components that save energy across various operating states (including idle).
- A power and performance data sheet for purchasers; this standardizes key information on energy performance, features, and other capabilities.

- **Data storage management:** through tools such as:
  - Automated storage provisioning, which improves storage efficiency by right-sizing, identifying and reallocating unused storage, and increases server capacity by improving existing storage use.
  - Deduplication software, which condenses the data stored at many organizations by more than 95% by finding and eliminating unnecessary copies.
  - Thin provisioning, which allocates just enough storage just in time by centrally controlling capacity and allocating space only as applications require it.
  - Redundant Array of Independent Disks (RAID), combining multiple disk drive components into a single logical unit.
  - Tiering storage, which stores automatically low-priority data, that are rarely accessed, on higher-latency equipment that uses less energy.
- **More efficient data storage equipment:** which uses less energy by adoptions technologies such as:
  - Lower speed drives.
  - Massive array of idle disks (MAID).
  - Solid-state drives (SSDs).
  - ENERGY STAR-certified data storage.

The simple algorithm is proposed in order to estimate annual savings for data center ICT measures, using Equation 3.10.

$$\begin{aligned} Annual\ Energy\ Savings \\ &= 8760 * (Power\ Draw_{Pre-existing\ measure} \\ &- \ Power\ Draw_{Efficient\ Measure}) \end{aligned} \tag{3.10}$$

Equation 3.11 estimates the savings when server efficient metrics increase due to units with higher efficiency (e.g., operations/Watt).

$$Annual\ Energy\ Savings_{Efficient\ Servers} = P_{EE} * \left({}^{1 - EM_{EE}}/_{EM_{baseline}} \right) * 8760 \tag{3.11}$$

where,

P_{EE} = power draw of new efficient server equipment (kW)

EM_{EE} = efficiency metric for efficient server

$EM_{baseline}$ = efficiency metric for baseline server

8760 = number of hours in a year as servers run 24/7 in a data center

Another form to compute savings for servers is considering ENERGY STAR-certified servers as "efficient servers." Using EPA estimates of percentage savings compared to standard or typical servers, savings can be calculated using Equations 3.12, 3.13 and 3.14.

$$Annual\ Energy\ Savings_{ES\ Servers} = (P_{baseline} - P_{ENERGY\ STAR}) * 8760 \tag{3.12}$$

$$P_{ENERGY\ STAR} = \sum_{ES=1}^{n} (P_{ES,idle} + U_{ES} * (P_{ES,full\ load} - P_{ES,idle}) \tag{3.13}$$

$$P_{baseline} = P_{ENERGY\ STAR} / (1 - a) \tag{3.14}$$

This approach conducts to the simplified expression represented in Equation 3.15.

$$Annual\ Energy\ Savings_{ES\ Servers} = \left({}^{1}/_{(1-a)} - 1 \right) P_{ENERGY\ STAR} * 8760 \tag{3.15}$$

where,

$P_{ENERGY\ STAR}$ = power draw of ENERGY STAR server (kW)

ES = ENERGY STAR server, numbered 1 to *n*

$P_{ES,idle}$ = power draw of ENERGY STAR server at idle (kW)

$P_{ES,full\ load}$ = power draw of ENERGY STAR server at full load (kW)

U_{ES} = utilization of ENERGY STAR server

$P_{baseline}$ = power draw of baseline servers

a = percentage ENERGY STAR server is more efficient than baseline "standard" or "typical" unit

8760 = number of hours in a year as servers run 24/7 in a data center

Server virtualization savings compare baseline energy use of a large combination of single application servers that would have been purchased normally during a server upgrade, without virtualization to a smaller set of virtual host servers, as can be seen in Equations 3.16, 3.17 and 3.18.

$$P_{baseline} = \sum_1^n (P_{sa,idle} + U_{sa} * (P_{sa,full\ load} - P_{sa,idle})) \tag{3.16}$$

$$P_{w\ Virt} = \sum_1^m (P_{vh,idle} + U_{vh} * (P_{vh,full\ load} - P_{vh,idle})) \tag{3.17}$$

$$Annual\ Energy\ Savings_{ES\ Servers} = (P_{baseline} - P_{w\ Virt}) * 8760 \tag{3.18}$$

where,

$P_{baseline}$= total power draw of all single-application servers without virtualization during server refresh (kW)

sa = single application servers, numbered 1 to *n*

$P_{sa,idle}$= power draw of a single-application server at idle (kW)

$P_{sa,full\ load}$ = power draw of a single-application server at full load (kW)

U_{sa}= average utilization of a single-application server over the year

$P_{w\ Virt}$ = total power draw in kilowatts of all virtual hosts

vh = virtual host servers, numbered 1 to *m*

$P_{vh,idle}$= power draw of a virtual host server at idle (kW)

$P_{vh,full\ load}$ = power draw of a virtual host server at full load (kW)

U_{vh}= average virtual host server utilization over the year

Savings from upgrading to more efficient storage devices can be calculated using Equations from 3.19 to 3.22. Equations 3.19, 3.29 and 3.21 utilize efficiency metrics of the efficient and baseline unit to estimate savings. Equation 3.22 makes use of the percentage savings for an ENERGY STAR-certified data storage to estimate savings. To calculate savings from software management tools Equation 3.23 relies on measuring power draws before and after storage management tools are implemented.

$$EM_{baseSE} = (\sum_{j=1}^{m} f_{baseSE(j)} EM_{baseSE}(j) \tag{3.19}$$

$$EM_{EESE} = (\sum_{i=1}^{n} f_{EESE(j)} EM_{EESE}(j) \tag{3.20}$$

$$Annual\ Energy\ Savings_{Efficient\ Storage} = P_{EESE} * \left({EM_{baseSE}}/{(EM_{EESE})} - 1\right) * 8760 \tag{3.21}$$

where,

P_{EESE} = power draw of new energy-efficient storage equipment

EM_{EESE} = efficiency metric for energy-efficient storage equipment

EM_{baseSE} = efficiency metric for baseline storage equipment

$EM_{EESB(j)}$= watts per terabyte (TB) of energy-efficient storage device/array j (this value can come from product specifications for devices and/or arrays)

$EM_{baseSB(j)}$ = watts per terabyte (TB) of energy-efficient storage device/array j (this value can come from product specifications for devices and/or arrays)

$f_{EESB(i)}$= fraction of total TB stored on energy-efficient device/array i

j = baseline devices/arrays, numbered 1 to m

$f_{baseSE(j)}$ = fraction of total TB stored on a baseline device/array j

i = energy-efficient devices/arrays, numbered 1 to n

8760 = number of hours in a year as servers run 24/7 in a data center

$$Annual\ Energy\ Savings_{ES\ Storage} = \left({1}/{(1-b)} - 1\right) P_{ES\ STOR}\ * 8760 \tag{3.22}$$

where,

$P_{ES\ STOR}$ = power draw in kilowatts of ENERGY STAR storage

b = percentage of ENERGY STAR storage more efficient than typical or standard storage

8760 = number of hours in a year as servers run 24/7 in a data center

$$Annual\ Energy\ Savings_{DS\ Man} = 8760 * (kW_{Pre\ Ds\ Man} - kW_{Post\ Ds\ Man}) \tag{3.23}$$

where,

$P_{Pre\ Ds\ Man}$ = total power draw of data storage before data storage management tool measures implemented (or with tool turned off) and after efficient data storage equipment is installed, if that was part of the measure (the savings from the efficient storage equipment can be calculated using Equations from 3.19 to 3.22) (kW)

$P_{Post\ Ds\ Man}$ = total power draw of data storage after data storage management tools are implemented and after efficient data storage equipment is installed, if that was part of the measure (the savings from the efficient storage equipment can be calculated using Equations from 3.19 to 3.22) (kW)

8760 = number of hours in a year as servers run 24/7 in a data center

Total energy savings, which contain additional cooling and power infrastructure savings, is calculated by multiplying energy savings from an ICT upgrade by the data center's power usage effectiveness (PUE). As a data center becomes more efficient, PUE moves toward 1. Equation 3.24 estimates total energy and demand savings.

$$Annual\ Energy\ Savings_{Total} = PUE * \ Annual\ Energy\ Savings_{IT} \tag{3.24}$$

where,

PUE= average PUE determined over the entire year

Equation 3.25 describes ICT lifetime savings for server virtualization, efficient server upgrades, or efficient storage.

$$Lifetime\ Energy\ Savings_{IT} = \ Annual\ Energy\ Savings_{Total} * EUL \tag{3.25}$$

where,

EUL= expected useful life based on IT upgrade cycle of data center

Equations 3.26 and 3.27 estimate seasonal peak demand savings, based on server and storage 24/7 operations.

$$Peak\ Demand\ Savings_{Winter} = PUE_{Winter} * Annual\ Energy\ Savings_{IT}/8760 \tag{3.26}$$

$$Peak\ Demand\ Savings_{Summer} = PUE_{Summer} * Annual\ Energy\ Savings_{IT}/8760 \tag{3.27}$$

where,

PUE_{Winter}= average PUE over the winter peak demand period, which can be tracked over an entire year. PUE_{Winter} may be smaller in winter due to free cooling)

PUE_{Summer} = average PUE over the summer peak demand period. PUE_{Summer} may be much higher during the summer as free cooling options may not be available as often.

3.2.4 Applications

For each of the presented methodologies there is a more appropriate implementation scenario, after all, each approach has particular characteristics that can make it more appropriate for a given situation.

The EL approach is indicated to policies-based situations, where the use of efficient ICT technologies to be introduced into the data center is a part of the normal equipment replacement cycle. It is advisable to previously use methodologies with consistent mathematical models in order to estimate the current energy scenario and savings perspectives.

The LBNL model is more appropriated to the assessment of electricity use and efficiency due to calculations complexity. Given the bottom up nature of the model, improved data on installed device numbers, classes, and equipment electricity use in different space types would particularly improve its accuracy.

The NREL method is more suitable where information about ENERGY STAR, SERT and particularities on virtualization and storage metrics and devices lifetime are present to calculate data center ICT savings.

3.2.5 Advantages and Disadvantages

In order to analyze each approach in its specificities Table 3.2 presents their main advantages and disadvantages.

Table 3.2 – Methodologies advantages and disadvantages

Methodologies	Advantages	Disadvantages
EL	Demand and supply power consumption-based	Model based on a reduced number of technologies and best practices
	Vendor-neutral approach	Not mathematically clear
	Provide cascade effect of savings	Do not contemplate network devices
	Flexibility regarding the sequence of implementation measures	Do not contemplate storage devices
LBNL	Bottom-up approach based on reliable studies	Equations not fully based on ENERGY STAR
	Concise mathematical modelling framework	More comprehensive and complex
	Estimates data center energy use and efficiency potentials at different geographic scales and can be replicated or refined by others	Restricted data coming from paid consulting firms
	Contemplate network and storage devices	Do not contemplate the virtualization impact
NREL	Bottom-up approach based on servers and data storage-utilities incentives oriented	Do not contemplate network devices
	Concise mathematical modelling framework with equations based on ENERGY STAR and SERT requirements	Difficulty in determining useful life
	Contemplate storage devices and virtualization impact	Do not contemplate different geographic scales, just different equipment's capacity
	Easier data obtaining and updated in terms of technologies	Regulators do not define "typical" or "standard" efficiencies for ICT equipment

3.2.6 Comparative Analysis

The analysis of the presented methodologies clarifies that their objectives are not the same, as can be seen in Table 3.3, although they present some similarities. Therefore, it is important to highlight the main features found in the respective approaches in order to compare them enabling future decision-making in the choice of the most appropriate approach.

Table 3.3 – Methodologies comparative analysis.

Features	Methodologies		
	EL	LBNL	NREL
Guide and list-based	X		
Police recommendations	X		
Bottom-up approach	X	X	X
Vendor-neutral evaluation	X	X	X
Quantitative analysis	X	X	X
Cascade effect of savings	X		
Mathematical modelling framework		X	X
Different data center size profiles		X	
Storage devices-base		X	X
Different type of disks			X
Network devices-based		X	
Virtualization impact	X		X
Cooling solutions	X	X	X
UPS's solutions	X		
Lightning solutions	X	X	
Different classes of servers	X	X	X
Lifetime-based	X		X
Designing for high density	X		
Power management	X		
DVFS-based		X	
Airflow management	X		
ENERGY STAR-based	X	X	X
SERT-based			X
Utilities incentive-based			X
Upgrading	X		X
Replacement perspective	X		X
Peak demand savings			X
PUE-based		X	X
Weather considerations			X
Difficult to obtain data		X	X

Considering the goals of this work, as well as the described characteristics, the more appropriated methodology for energy efficiency adoption is from EL. The reasons that have supported this choice were being more suitable for the SMDC case, taking into account the devices replacement window, the capacity to estimate the savings in five years with cascade effect and using equipment's suggested by ENERGY STARS. However, it is important to complement the model using network, backup and storage solutions to improve the comprehensiveness of the approach in terms of energy efficiency.

Finally, this chapter has presented an insight on the energy efficiency perspective based on three surveys which assessed the importance of energy use to operators of SMDC. This shows that, even though there have been major advances in energy efficiency for large data centers over the last decade, with nearly 52% savings, in SMDC the reality is greatly different – 43% have no energy efficiency objectives in place. Furthermore, the survey conducted in this work highlighted alarming statistics, where 64%, 73% and 77% of surveyed participants do not monitor servers', storages appliances' and network devices' energy use, respectively.

Complementally, three energy efficiency methodologies have been discussed emphasizing their main characteristics as well as their applications. Subsequently, they have been assessed concluding that EL approach is more suitable and appropriated to SMDC profile.

The results of the presented surveys demonstrated a paradoxical reality in relation to the technological possibilities in the field of energy efficiency in the data center environment as a whole, including those of small and medium profile. Although best practices using technologies that enable more efficient use of energy are increasing, it has proven that the absorption of this reality in the data center market does not occur in the same pace. Thus, it was important to carry out the evaluation of energy efficiency methodologies in SMDC already applied in real scenarios in order to mitigate this mismatch. Combining the information provided by recent surveys with approaches that alter certain scenarios, enables the creation of new policies and frameworks for this neglected market. Thus, the results of the survey carried out in this thesis will be taken into account in the elaboration of scenarios for the future simulations in this work. Likewise, the methodology assessed will be considered in the development of the framework that will be presented in the next section.

Therefore, it is important to note that lack of accurate information and corporate misalignment are the main causes of slow deployment of energy efficiency in this sector. Thus, the policy recommendation is that it is essential to treat small and medium data centers as an individual market case, which requires attention, specific incentives and policies to address their particularities. Innovative possibilities include making joint energy analysis of multiples data centers at the same time and proposing an aggregator to mediate operations between data centers and utilities.

It is a desirable goal to achieve a trade-off between the maximum performance of data centers and the minimum environmental impact by considering various aspects such as cost, and energy consumption as constraints. Thus, DR programs emerge in this context as an alternative to the minimization of energy-related costs, since an inversely proportional relationship is established, i.e., the less energy efficient a data center is, the more opportunities there are to explore DR alternatives and the next sections will cover this alternative.

CHAPTER 4

DEMAND RESPONSE METHODOLOGY FRAMEWORK

This chapter presents a methodology framework addressing demand response in different perspectives. Firstly, it is provided an overview on the main approaches to optimize workloads in data centers, with examples applied to SMDC. Then, a framework proposal in accordance with the goals of this thesis is presented and discussed in detail based on other examples from the literature. The mathematical models denoting the key workloads in a SMDC environment during power reduction and rebound events are also defined. From these premises, the two problems established in the context of this work, one from SMDC point of view and the other from DSO perspective are discussed. Their resolution hypotheses through algorithm optimization processes are presented with detail in order to further introduce the simulations and data collection.

4.1 APPROACHES TO OPTIMIZE DATA CENTERS ENERGY LOADS

For the purpose of reducing the energy cost and perform a gradual inclusion of data centers in DR programs, recent and relevant works have been conducted proposing theoretical frameworks of robust optimization and low computational complexity to obtain close-to-optimal solutions.

Cioara *et al.* (2016) propose an electronic marketplace designed for trading energy flexibility and ancillary services, enacting data centers to shape their energy demand to buy additional energy when prices are low and sell energy surplus when prices are high.

A new pricing mechanism, which extracts load reductions from tenants in colocation data centers during emergence demand response events, was proposed by Chen *et al.* (2015). Their results present benefits to the environment and data center operators by decreasing the need for backup diesel generation and providing payments for load reductions.

The dynamic interactions between smart grid and data centers, as a two-stage price optimization problem, were presented in Wang *et al.* (2016), using an heuristic algorithm and simulations to achieve a win-win solution for both the utility and data centers.

Fridgen *et al.* (2017) present an economic analysis of spatial load migration using geographically distributed data centers as an alternative form of demand side flexibility compared to load shifting and load shedding, finding that spatially migrating load provides an interesting alternative to economically balancing a grid, as well as realistic opportunities to virtually transfer balancing power between different market areas worldwide.

A multi-objective energy-efficient task scheduling problem on a green data center partially powered by renewable energy, where the computing nodes are DVFS-enabled is highlighted in Lei *et al.* (2016). The solution is provided by a multi-objective co-evolutionary algorithm that searches the suitable computing node, supply voltage and clock frequency for the task computation, and the smart time scheduling strategy is employed to determine the start and finish time of the task on the chosen node.

Wang *et al.* (2017) propose scheduling algorithms to adjust the scheduling policies for the incoming jobs according to the performance target and the behavior of other competitors based on the game theory. The results show the capacity to reduce the conflict in scheduling decisions made by different schedulers and hence improve the scheduling performance in the data centers deployed with clusters and distributed schedulers.

A novel method based on demand response was proposed to control the cooling supply related to ICT dynamic load in Zhu *et al.* (2017) and the assessment pointed out reductions of 7.9%, 14.2%, 15.6% and 17.9% of energy consumption in a cooling demand response at room, row, rack and server levels, respectively during the demand response period. Determining the cooling demand according to the ICT load at server level, the reduction of the electricity consumption of cooling systems during the demand response period was by 0.9% and considering the dynamic energy efficiency of cooling units, ICT load shifting could be optimized in 1.2%.

As well as there are energy efficiency strategies exploiting workloads with direct impact on CPU, memory, disk and network, virtualization, cooling, and UPS, there are data centers demand response approaches in terms of flexibility. The technique addressed in Cioara *et al.* (2016) provides flexibility mechanisms defined for hardware components, such as load time shifting, alternative usage of non-electrical cooling devices and charging/discharging UPS, evidencing potential to shape and modify the data center baseline energy profile to meet energy network levels goals and to provide several types of energy and balancing services.

Tran *et al.* (2014) and Li *et al.* (2015) define optimization problems, where processing, virtualization, quality of service and cooling solutions are analyzed together in a demand response scenario, resulting in a significant energy consumption and electricity cost reduction.

In Ghatikar *et al.* (2012), an initial set of control and load migration strategies and economic feasibility for four data centers was evaluated. The findings show that with minimal or no impact to data center operations demand savings of 25% at the data center level, or 10% to 12% at the whole building level can be achieved with demand response strategies, such as server and CRAC unit shutdown, load shifting or queuing ICT jobs while server are idle, temperature set point adjustment and load migration between homogeneous and heterogeneous cluster systems.

In the same context, it is important to consider the adaptation to data centers of measures already used in other sectors with a higher tradition of participation in DR programs. In this way, Paterakis *et al.* (2017) perform an extensive overview on the theme covering different sectors and international experiences.

Regarding industrial customers, the demand can usually be decreased by on-site generation, energy storage, consumption shifting, non-critical load curtailment and temporary shut-down of several processes. Temporarily interrupting one or more processes may result in significant load reductions. Nevertheless, several constraints such as the criticality of a process, the number of available production lines, the required production target, inventory restrictions, etc., may have longer term impacts on the process line, in a very similar way that occurs in the case of data centers, respecting the differences of equipment and business profile.

Analyzing commercial and other non-residential customers, the main DR strategy is load reduction, where air conditioner is the most significant load that can be controlled. However, energy intelligent buildings with energy consumption monitoring and management of locally available resources, as well as the energy procurement from the grid, has been introduced (Christantoni *et al.* 2016) and can also be used in SMDC.

Residential customers are suitable for DLC and price-based DR programs and can invest on an automated system, which monitors and controls the consumption of several appliances (Paul *et al.* 2017). Following this same premise, data centers can use algorithms, software, or even appliances to provide this level of supervision in their loads (Paul *et al.* 2017), taking advantage of specific demand response programs. Nevertheless, the main similarity of the residential sector compared to the SMDC market is the fact that their loads have a low potential to participate in DR events acting individually, and therefore there is a need for aggregation, as suggested the review conducted by Carreiro *et al.* (2017), covering several cases of aggregators in the context of end users, namely in the residential case.

However, although the contribution of each of the above-mentioned works is undeniable in the DR field, there are researches that are closer to the purpose of this thesis and that will be used as the guiding and argumentative elements in this process.

From the SMDC point of view, studies such as Tran *et al.* (2014) and Wang *et al.* (2016) prioritize a detailed mathematical description of the load systems present in data centers, as well as the problem formulation oriented to an objective function and its respective constraints. Thereby, this work will follow the same criterion in sections 4.2, 4.3 and 4.4.

Regarding the Distribution System Operator (DSO) point of view, with the aim of taking advantage from the massively automated infrastructure and expressive energy usage, the definition of power adaptation collaboration was researched by Basmadjian, Lovasz, *et al.* (2013) and shortly thereafter a supply demand exploring this concept between energy provider and data centers was presented in Basmadjian *et al.* (2013). Those two works were utilized as a comparative basis to the algorithms and polices proposed by (Basmadjian *et al.* 2015) and they will be used jointly as references to this present thesis, specifically in sections 4.2 and 4.5.

4.2 FRAMEWORK PROPOSAL

The framework proposed in Basmadjian *et al.* (2013) take into account the following entities: Energy Provider (EP), Data Center and Data Center IT Customer (ITC) sub-ecosystems. Aiming at contemplating them, as Figure 4.1 displays, a three-tier architecture has been established.

The Level I, called Connection, comprises all the particularities of the EPs and data centers involved infrastructure. Accordingly, the monitoring and control infrastructure is carried out by the Connection level. The Level II, known as Negotiation, represents the decision-making logic deployed in the form of agents to allow power adaptation collaboration. Hence, this level must interact with Level I in order to read the current status (e.g. shortage situation) and adopt certain power adaptation requests to the involved infrastructure, as well as with Level III, which includes the contracts to stimulate power adaptation collaboration between EP – Data Centers – ITC.

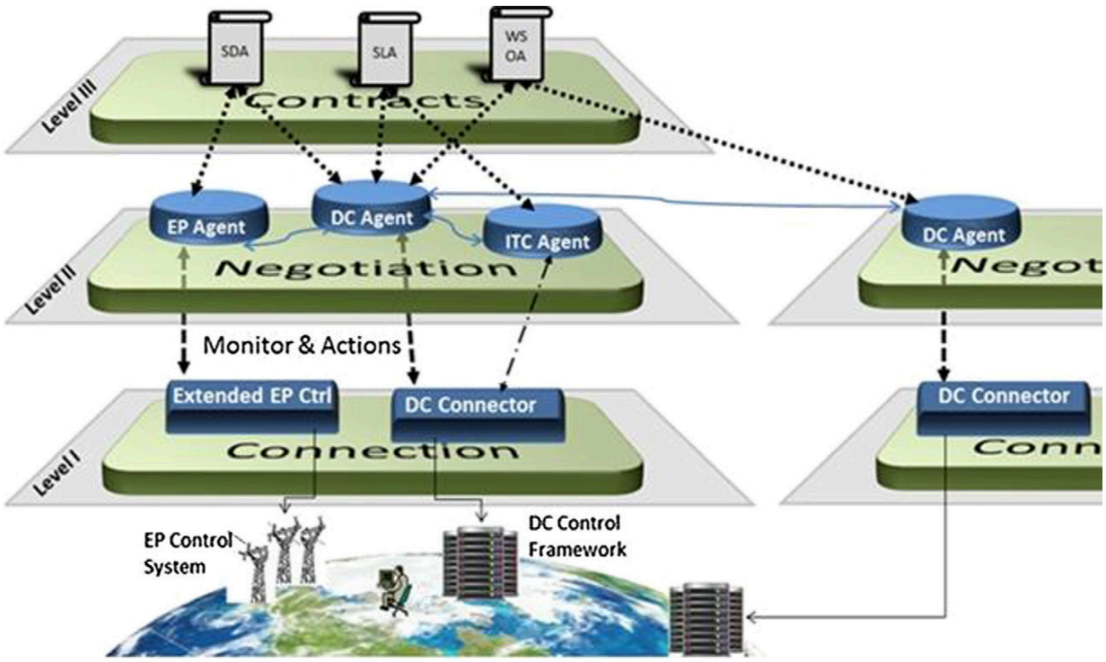


Figure 4.1 – Framework overview of Basmadjian et al. (2013)

However, the framework proposed by this work intends to simplify the process by reducing the number of actors, interactions and restricting the structure. In the same way it is intended to present more detail in the processes belonging to each of the strategies described in the aforementioned layers, as represented in more descriptive details in Figure 4.2.

The main actors in this proposed framework are restrict just to SMDC and DSO, as well as their relationships and interactions. One of its premises is that the connections between data centers and their consumers are not part of the DSO responsibilities; in other words, they are transparent to this fact, as opposed to the three actors in Basmadjian *et al.* (2013) and the strong relationship between them.

This framework is initially SMDC-oriented, however it can be adapted to cover different data centers profiles, since each layer is properly contextualized. Nevertheless, in the context of this work, SMDC are analyzed jointly and the intermediation between their operators and the energy provider is performed by a DSO, or an Aggregator, which sees each SMDC as a single component that has flexible loads and along with other SMDC contain a significant load to be used in DR programs.

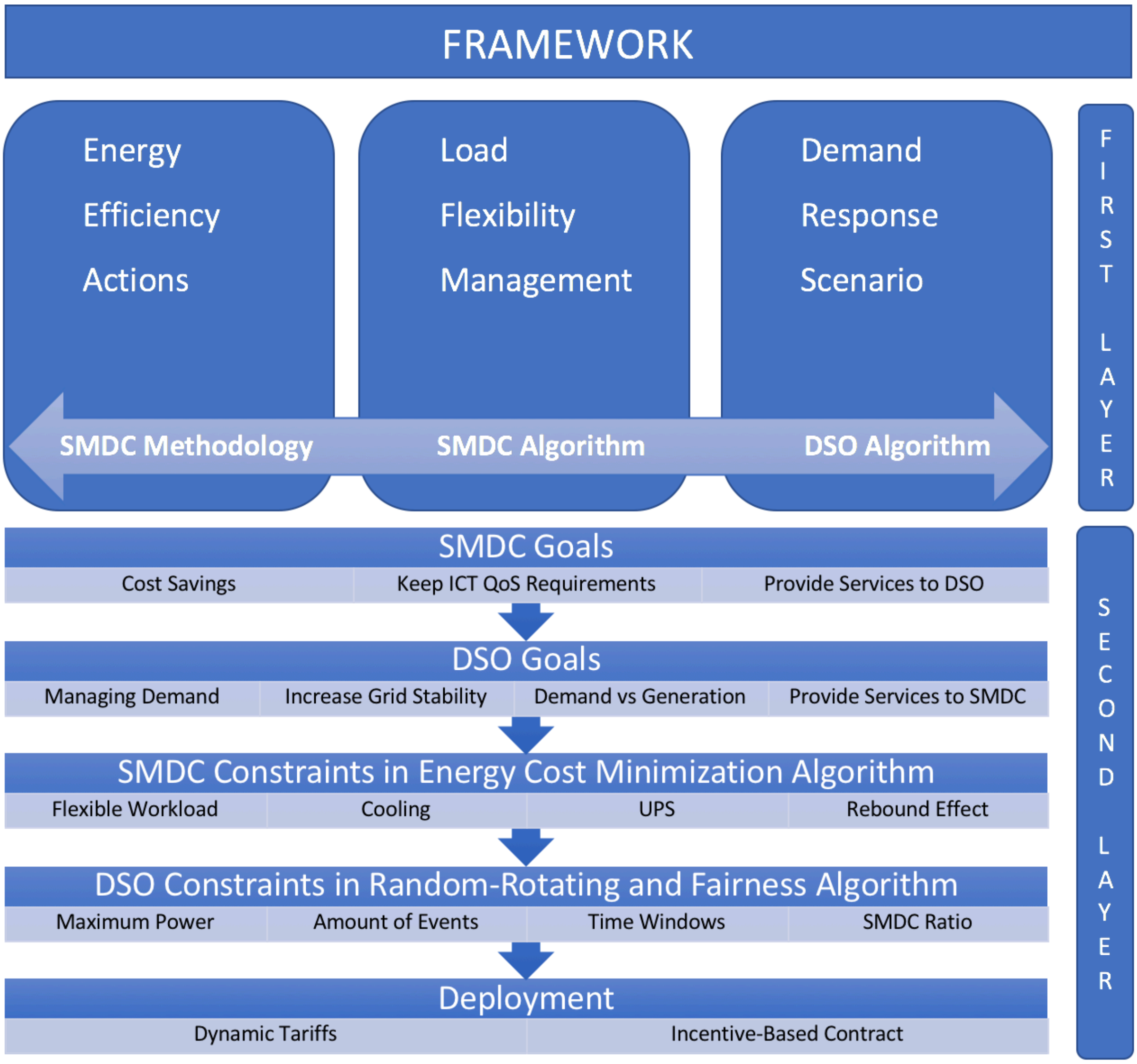


Figure 4.2 – Framework proposal

Thus, the main objective is to combine energy efficiency initiatives, which in turn provide a better knowledge and management of the flexible data center load, as well as the mutual possibility of taking financial advantages in DR programs established between SMDC and DSO. The framework is divided into two layers, as can be seen in Figure 4.2.

The first layer has a responsibility-oriented approach, where energy efficiency actions are taken over by SMDC operators through the consistent use of the most appropriate energy efficiency methodology, as analyzed in Chapter 3. Flexibility management will be the responsibility of a SMDC algorithm that optimizes the best time window to decrease load, whereas DR scenarios will be defined by a DSO random-rotating and fairness algorithm.

Besides the advantage of reducing the number of layers, actors and interactions, another important feature of this framework is to detail the specificities of the lower layer. The second layer is oriented towards the goals, constraints and deployment of DR strategies applied to SMDC and DSO. On one hand, the main goal of SMDC is to decrease expenses while providing DR services to DSOs, and simultaneously maintain their requirements and QoS thresholds. On the other hand, the DSO aims at providing services to SMDC, enabling the management of demand, reduction of costs (demand vs. generation and implementation of DR program) and increasing the grid stability.

In the SMDC constraints of energy cost minimization algorithm sublayer, it is possible to note that the subdivision is given by ICT flexible workload, where a set of delay-tolerant loads will be run in a given time. Non-ICT workloads comprise cooling solutions and UPS strategy. Rebound Effect (RE), or "payback effect" as it is sometimes referred to, has been defined as the tendency of electrical loads to produce a demand spike while "catching-up" to normal (Wrinch *et al.* 2012), is the last constraint in this sublayer. An example of such behavior is that once a given data center has reduced power during a DR event, the same reduced power must be taken over shortly thereafter, to only recover its baseline values.

In other words, SMDC will make use of these three profiles of flexible loads to reduce power during a DR event, first prioritizing Non-ICT loads and only using ICT loads after it, as presented by Table 4.1.

In the next sublayer, a DSO random-rotating and fairness algorithm will define which SMDC will be used in DR events, ensuring that they are utilized in a balanced and fair manner over a given contractual period, until a new usage cycle begins. However, there are several constraints, such as: maximum reduction of power that cannot be exceeded, number participations in events by cycle, time windows per event and the correct ratio between SMDC, which must be maintained in each selection process.

In the deployment sublayer, the structure to be used can be based on dynamic tariffs, in which the energy price changes over time or an incentive-based contract established between the parties, where SMDC might receive a financial compensation to reduce power in a given period, or in the case of not accomplishing with a specific contractual clause, to pay a penalty.

Table 4.1 – Workloads hierarchy

Resources	Elements	Component	Flexibility Technologies	DR Action	Constraints	Priority
ICT Load	Server	CPU	DVFS	Shedding Shifting	Load Profile Delay Time	3
			Dynamic Shutdown	Shedding Shifting	Load Profile Off Time	5
		Memories and Disks	Consolidation Live Migration	Shifting	Load Profile Delay Time	4
			Dynamic Shutdown	Shedding Shifting	Load Profile Off Time	3
	Networking	CPU	DVFS	Shedding Shifting	Load Profile Delay Time	3
			Active Connections	Shedding Shifting	Load Profile Delay Time	4
			Dynamic Shutdown	Shedding Shifting	Load Profile Off Time	5
Non-ICT Load	Cooling	HVAC	Set Point Adjustment	Shedding Shifting	Delay Time	2
			Dynamic Shutdown	Shedding Shifting	Off Time	-
	Lightning	Sensors	Dynamic Shutdown	Shedding	Off Time	-
	Storage	UPS	Discharge	Shifting	Load Profile Delay Time	1
		Generator	Use	Shifting	Load Profile Delay Time	-
		Renewable Generation	Use	On-Site Generation	What How Much	-

In the upcoming sections the contract premises proposed by this work will be discussed in detail within the DSO optimization particularities. However, it is important to highlight that the background to this type of agreement has been established by Basmadjian *et al.* (2013) with three different strands:

- **GreenSLA** (Green Service Level Agreements) contracts are agreements between data centers and ITCs, which reflect the agreed scope for the data center to operate in an energy-aware manner and at the same time guarantee a certain level of QoS for the IT customers.
- **GreenSDA** (Green Supply Demand Agreements) contracts are agreements between EPs and data centers, which define the flexibilities and energy-related contractual terms that these parties grant to each other.
- **GreenWSOA** (Workload Services Outsourcing Agreements) contracts are agreements among federated data centers that set rules for the geographical shifting of workload.

This thesis is focused just in the GreenSDA profile as prescribed by the proposed framework, however the contractual terms discussed in Basmadjian, Niedermeier, *et al.* (2013), will be adopted as comparative parameter of analysis.

While the terms in Basmadjian, Niedermeier, *et al.* (2013) are based on power adaptation profiles during a DR event, the proposed terms of the contract in this thesis stem from the premise that these profiles must be previously established, in the act of signing the agreement between the parties, i.e., before a data center participates in the contract it is necessary that it can guarantee the reduction profile stipulated by DSO. For such purposes, this work is going to consider the profile values determined in the trials of Basmadjian *et al.* (2015), which are: 105 kW for small data centers and 550 kW for medium data centers.

4.3 SMDC MATHEMATICAL MODELS

SMDC can support a broad range of workloads in a normal operation during a time window t. These loads are typically divided in computational (e.g., CPU, memory, network and storage) and non-computational (e.g., cooling, lighting, and power facility) and then subdivided in delay-sensitive, with no flexibility and delay-tolerant, with flexibility of scheduling.

Assuming that, for the purpose of DR strategies, the loads should be primarily flexible, all data centers will be managed at DR time j responsible for shifting or shedding the flexible workload of geo-dispersed data centers, in order to attend a DR event decreasing their loads in response to the DSO needs established in a contract, or to dynamic tariffs.

However, after the DR period the reduced power tends to return to its baseline (t), but before it, in a time window $j+1$ occurs a process of RE, wherein there is an energy consumption increase equivalent to 100% of the energy consumption reduction during the DR period aggravated by the losses, to ensure the service replenishment that was decreased during DR. For example, as the temperature of cooling devices increases in a DR event, the power decreases. Subsequently, to recover the temperature after the DR event, a higher power is needed, before resuming the usual value.

It is considered a discrete time model $t \in \mathcal{T} = \{1, \dots, T\}$ representing a temporal baseline (e.g., typically a day, or a month), where the normal operation of SMDC and their critical mission occurs. Let $j \in \mathcal{J} = \{1, \dots, J\}$ denote a DR period and $j+1 \in \mathcal{J} = \{1, \dots, J\}$ denote a RE situation. Let $i \in \mathcal{N} = \{1, \dots, I\}$ represent the set of SMDC, ICT workloads, or CRAC units.

In the same context, it is important to highlight the general constraints of this problem as follows:

- ICT workload.
- Cooling.
- Uninterrupted Power Supplies.

4.3.1 ICT Workload

Each data center i, has hundreds of servers providing services to meet users' requests. S_{max_i} represent the maximum number of servers per data center. The data center operator can switch on or off servers to adjust the service time, as well as shift or shed certain workloads in selected servers $S_{select_i^j}$ in order to attend a DR event, as formulated by Equation 4.1. Since the number of servers is typically large, it can relax the integer constraint on the number of selected servers without significantly affecting the optimal result.

$$\sum_{i=1}^{N} S_{select_i^j} = S_{total^j}, S_{select_i^j} \geq 0\ , \forall i \in \mathcal{N}, j \in \mathcal{J} \tag{4.1}$$

where Equation 4.2 S_{total^j} is the total of selected server in a set of SMDC.

$$0 \leq S_{select_i^j} \leq S_{total^j} \leq S_{max_i}\ , \forall i \in \mathcal{N}, j \in \mathcal{J} \tag{4.2}$$

In a specific time slot j, flexible workloads can be ready to run in each data center during a demand response signal. There are different types of workloads, and each one may correspond to a specific application. Then, a total of $\mathcal{N}$ data centers might work jointly from the DSO point of view to complete the total flexible workload of SMDC P_{twflex^j} with the attribution of a single data center i running $P_{wflex_i^t}$ or a set of them, $i+1$, working together, as can be seen in Equation 4.3.

$$\sum_{i=1}^{N} P_{wflex_i^j} = P_{twflex^j}\ , P_{wflex_i^j} \geq 0, \forall i \in \mathcal{N}, j \in \mathcal{J} \tag{4.3}$$

$P_{wflex_i^j}$ represents the computation demand that can be decreased in terms of power in a DR event and executed in an appropriated cost-effectiveness time window in a load shed DR strategy. In this approach, the workload service demand, considering memory, storage and network, is neglected for simplicity. In order to consider it, the service demand have to be changed from a scalar to a vector, in which each element corresponds to one type of demand. Equation 4.4 models the workload.

$$\sum\nolimits_{i=1}^{N} P_{alloc_i^j} = P_{wflex_i^j}\ , 0 \leq P_{alloc_i^j} \leq P_{resid_i^{t+j}}\ , \forall i \in \mathcal{N}, t \in \mathcal{T}, j \in \mathcal{J} \tag{4.4}$$

where, j is the maximum number of time slots allowed for finishing the flexible workload i from its arrival time t, in a DR event i.e., the work i must be fully served before the beginning of time slot $t+j$.

$P_{alloc_i^j}$ is the SMDC allocated capacity in slot t, where $(0 \leq j \leq t)$. $P_{resid_i^{t+j}}$ represents the SMDC baseline residual capacity to run the workload after a DR event in time slot $t + j$ of a load shifting DR strategy, where $(0 \leq P_{resid_i^{t+j}} \leq P_{total_i^t})$, with $P_{total_i^t}$ denoting the SMDC total capacity. On the other hand, $P_{wflex_i^{j+1}}$ denotes the flexible power assuming the residual and baseline values.

Data centers should provide to their users guarantees through a QoS policy, ensured by a previous established service level agreement (SLA), where average performance for the data center operation is guaranteed in a time window. Small and medium data centers usually do not have SLA in terms of general data center operation. However, it is common that several applications and services have a certain QoS level in their operation.

Hence, in a DR event the prior assumption is that the workload will be reduced, i.e., it will have quality constraints in terms of processing, memory and disk allocation. Therefore, the first criterion to meet the DR QoS requirements, as shown Equation 4.5, is being a flexible workload.

$$QoS = {P_{wflex_i^j}}/{P_{alloc_i^j}}, \forall i \in \mathcal{N}, t \in \mathcal{T}, j \in \mathcal{J} \tag{4.5}$$

The second criterion lies in the fact of strictly observing the flexible workloads reduction thresholds and subsequently returning to a previous baseline where the prices are cheaper. Thus, these thresholds are mathematically described in Equation 4.6.

$$0 \leq P_{alloc_i^j} \leq P_{resid_i^{t+j}} \leq P_{total_i^t},\ \forall i \in \mathcal{N}, t \in \mathcal{T}, j \in \mathcal{J} \tag{4.6}$$

where the workload to be reduced by a previous cut of resources should fulfil the allocated capacity, or remain below, if there are no requirements of computational elements. These QoS thresholds are lower than the baseline residual capacity, which in turn is lower than the total capacity.

4.3.2 Cooling Solutions

Besides server power consumption and UPS's, cooling solutions also greatly contribute to the energy consumption of a data center. One of the simplest power models for computer room air conditioner (CRAC), used in the majority of SMDC, was demonstrated by Zhan and Reda (2013), in which the total power budget is split among computational environment and cooling solutions. The power consumption was modeled by Equation 4.7 with P_{crac} of a CRAC unit i.

$$P_{crac_i} = {\sum_i P_{ict}}/{CoP}, \forall i \in \mathcal{N} \tag{4.7}$$

where P_{ict} is the power consumption of the ICT load with their heat flow directed towards the CRAC unit, being the CoP, given by Equation 4.8, the Coefficient of Performance, which is the ratio between the removed heat H and the amount of work necessary W to remove that heat and varying in function of the temperature T_{sup} and T_{adj}. T_{sup} is the temperature of the air supplied by CRAC units, $T_{adj} = T_{safe} - T_{max}$, with T_{safe} denoting the maximum permitted temperature at the server inlets in order to prevent device damage and T_{max} the maximum temperature of the server inlets in SMDC; if T_{adj} is negative, it indicates that a server inlet exceeds the maximum safe temperature. In response, it necessary to decrease T_{sup} to bring the servers back below the system redline level

$$CoP = {}^{H}\!/_{W} \tag{4.8}$$

The temperature set-point of the computer room air conditioner (CRAC) recommended by ASHRAE is between 10 to 35 °C for A2 class. In a DR event, in a time window j, the environment temperature can increase to the maximum set-point, in order to reduce the consumption. Based on Doyle *et al.* (2013), the potential of power reduction in a cooling DR event can be obtained by Equation 4.9.

$$P_{cool_i^j} = {}^{P_{ict}}\!/_{CoP} + P_{fan}\ , \forall j \in \mathcal{J} \tag{4.9}$$

where P_{ict} is the total power of ICT load consumption, P_{fan} is the power required by the fans of the CRAC units and $P_{cool_i^{j+1}}$ it the RE situation where the temperature decreases and assumes the previous baseline values.

4.3.3 Uninterrupted Power Supplies

In a DR event, SMDC have to make a decision on whether they use UPS's as a resource by discharging energy from the available batteries, being $E_{dis^j}(DoD)$ the discharged energy and E_{maxdis} the amount of UPS energy that can be discharged at the depth-of-discharge level (DoD) where $\mathrm{DoD} \in [0,1]$, as given by Equation 4.10.

$$E_{maxdis} = (-DoD + DoD_{max})c_{total} \tag{4.10}$$

The recharge function at the depth-of-discharge level is represented by $E_{rec^{j+1}}(DoD)$ and its maximum value E_{maxrec} is formulated by Equation 4.11.

$$E_{maxrec} = (DoD)c_{total} \tag{4.11}$$

In order to preserve the useful lifetime, the UPS battery must reserve a minimum energy level of its capacity c_{min}. Additionally, the UPS should not be totally discharged during the DR event, in order to preserve energy for the data center operation. Therefore, the maximum amount of energy for discharge during a DR event should be limited, being α is use factor percentage to a DR event. Assuming that an UPS has a capacity of c_{total} and the efficiency of UPS (including the battery) charging and discharging are the same, represented by $\eta \in [0,1]$, where $\eta = 0.9$. The UPS capacity level c^t is expressed by the set of Equations from 4.12 to 4.15:

$$0 \leq E_{dis^j}(DoD) \leq E_{maxdis}\, \alpha, \forall j \in \mathcal{J} \tag{4.12}$$

$$0 \leq E_{rec^{j+1}}(DoD) \leq E_{maxrec}, \forall j \in \mathcal{J} \tag{4.13}$$

$$c^{t+1} = c^t + \eta E_{rec^{j+1}}(DoD) - {E_{dis^j}(DoD)}/{\eta} \tag{4.14}$$

$$c_{min} \leq c^t \leq c_{total}, \forall t \in \mathcal{T} \tag{4.15}$$

Using a different mathematical notation in function of the power during the DR event, the formulation of Equation 4.14 can be rewritten in Equation 4.16

$$c^t = c^{t-1} + \eta P_{ups^{j+1}}(j+1) - P_{ups_i^j}\, j/\eta, \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.16}$$

where, $P_{ups_i^j}$ is the power to be discharged during the event in a time window j, c^{t-1} is the capacity in the previous time instant and $P_{ups_i^{j+1}}$ is the recharging process in a RE situation.

4.3.4 Power Consumption

The data centers power consumption can be divided between ICT power consumption and non-ICT power consumption. The quantitative relation between them is measured by the PUE, which is represented as the ratio between the total power consumption and the ICT energy consumption. One on hand, the power used by computing devices is considered productive. On the other hand, the power for support infrastructure is auxiliary. Thus, PUE helps to understand the total power consumption based on the ICT power consumption. Therefore, in Equation 4.17 the total data center energy consumption $P_{smdc_i^t}$ can be calculated in a data center i in the time slot t.

$$P_{smdc_i^t} = S_{max_i}[P_{idle} + (P_{peak} - P_{idle})\,\sigma_i^t] + S_{max_i}[(P_{idle} + (PUE_i^t - 1)P_{peak}], \quad \forall i \in \mathcal{N}, t \in \mathcal{T} \tag{4.17}$$

where, S_{max_i} is the maximum number of servers in the data center, P_{peak} and P_{idle} are the servers peak and idle power, respectively and σ_i^t is the average server utilization (between 0 and 1) at time t, assuming that servers and their utilization rate are homogeneous. The expression in (4.17) has two key terms. The first term, i.e., $S_{max_i}[P_{idle} + (P_{peak} - P_{idle})\, \sigma_i^t]$ represents the power dependent on ICT load and the second term, i.e., $S_{max_i}[(P_{idle} + (PUE\,_i^t - 1)P_{peak}]$represents the power incident on non-ICT load.

4.3.5 Energy Cost

In a normal operation, Equation 4.18 presents the energy cost in SMDC considering all time slots t.

$$E_{cost^t} = \sum_{i=1}^{I} \left[\sum_{t=1}^{T} \left(P_{smdc_i^t}\, \varphi^t \right) \right] \tag{4.18}$$

$$\forall i \in \mathcal{N}, t \in \mathcal{T}$$

where,

E_{cost^t} = is the baseline energy cost in a time normal time slot t

$P_{smdc_i^t}$ = is the total SMDC power consumption

φ^t= is the electricity price in a time slot t

When dynamic tariffs are considered, it is possible to create variations in the tariff, in order to motivate DR events, since a high cost will encourage a consumption reduction. However, immediately afterwards the DR event, the rebound effect occurs, in which the previously reduced power will be fully taken over until it reaches the baseline value. Thereby, the total energy cost, taking into account the normal operation, the demand response and rebound effect windows is given by Equation 4.19.

$$\begin{aligned} E_{totalcost^{dt}} = \sum_{i=1}^{I} \frac{1}{ts} \Bigg\{ \left[\sum_{t=1}^{T} P_{smdc_i^t}\, \varphi^t \right] - \left[\sum_{j=1}^{J} \left(P_{smdcdec_i^j}\, \varphi^j \right) \right] \\ + \left[\sum_{j=1}^{J} \left(P_{smdcinc_i^{j+1}}\, \varphi^{j+1} \right) \right] \Bigg\} \end{aligned} \tag{4.19}$$

$$\forall i \in \mathcal{N}, t \in \mathcal{T}, j \in \mathcal{J}$$

where,

$E_{totalcost^{dt}}$ = is the total energy cost in a dynamic tariff approach

$P_{smdcdec_i^j}$ = is the decreased SMDC power consumption in a DR event j

$P_{smdcinc_i^{j+1}}$ = is the increased SMDC power consumption in a RE situation $j+1$

φ^t = is the electricity price in a time slot t

φ^j = is the electricity price in a time slot j

φ^{j+1} = is the electricity price in a time slot $j+1$

ts = time slots per hour

or in other terms as formulated by Equation 4.20.

$$E_{totalcost^{dt}} = \sum_{i=1}^{I} \left(E_{cost^t} - E_{cost^j} + E_{cost^{j+1}} \right) \qquad (4.20)$$

$$\forall i \in \mathcal{N}, t \in \mathcal{T}, j \in \mathcal{J}$$

where,

E_{cost^j} = is the reduction of energy cost in a DR event in time slot j

$E_{cost^{j+1}}$ = is the increase of energy cost in a RE situation in time slot $j+1$

On the other hand, in the case where the approach to be considered is an incentive-based contract, the formulation presented in Equation 4.21 is depicted as:

$$\begin{aligned} E_{totalcost^{ic}} = \sum_{i=1}^{I} \frac{1}{ts} \Bigg\{ & \left[\sum_{t=1}^{T} \left(P_{smdc_i^t}\, \varphi^t \right) \right] - \left[\sum_{t=1}^{T} \left(P_{smdcdec_i^j}\, \varphi^t \right) \right] \\ & + \left[\sum_{t=1}^{T} \left(P_{smdcinc_i^{j+1}}\, \varphi^t \right) \right] - \left[\sum_{j=1}^{J} \left(P_{smdcdec_i^j}\, \lambda^j \right) \right] \Bigg\} \end{aligned} \qquad (4.21)$$

$$\forall i \in \mathcal{N}, t \in \mathcal{T}, j \in \mathcal{J}$$

where,

$E_{totalcost^{ic}}$ = is the total energy cost in an incentive-based contract approach

λ^j = is the incentive given to participate in a DR event in the j time instant or in other terms as formulated by Equation 4.22.

$$E_{totalcost^{ic}} = \sum_{i=1}^{I} \left(E_{cost^t} - E_{cost^j} + E_{cost^{j+1}} - E_{icost^j} \right), \forall i \in \mathcal{N}, t \in \mathcal{T}, j \in \mathcal{J} \tag{4.22}$$

where,

E_{icost^j} = is the incentive in a normal time slot t

Upon accepting a DR signal, in order to decrease power, a SMDC can fall into three situations based on the power reduction percentage in the defined time window:

- Reducing between 80% and 100% of the power target and therefore receiving the incentive;
- Reducing between 20% and 79% of the power target and thus, not receiving the incentive and not paying the penalty;
- Reducing below 20% of the power target and then paying the penalty.

With the above-mentioned constraints, it is possible to formulate the SMDC energy costs minimization problem. Thus, the objective is to minimize the SMDC total energy cost in DR events, both in the dynamic tariff scenario in Equation 4.23 and in the incentive-based contract in Equation 4.24, as follows:

- *Objective function 1: Dynamic tariff SMDC energy cost minimization*

$$\min f(dt) = \sum_{i=1}^{I} \frac{1}{ts} \left\{ \left[\sum_{t=1}^{T} P_{smdc_i^t}\, \varphi^t \right] - \left[\sum_{j=1}^{J} \left(P_{smdcdec_i^j}\, \varphi^j \right) \right] + \left[\sum_{j=1}^{J} \left(P_{smdcinc_i^{j+1}}\, \varphi^{j+1} \right) \right] \right\} \tag{4.23}$$

$$subject\ to\ constraints$$

$$(4.3) - (4.6)$$

$$(4.7) - (4.9)$$

$$(4.10) - (4.16)$$

- *Objective function 2: Incentive-based contract SMDC energy cost minimization*

$$\min f(ic) = \sum_{i=1}^{I} \frac{1}{ts} \left\{ \left[\sum_{t=1}^{T} \left(P_{smdc_i^t}\, \varphi^t \right) \right] - \left[\sum_{t=1}^{T} \left(P_{smdcdec_i^j}\, \varphi^t \right) \right] + \left[\sum_{t=1}^{T} \left(P_{smdcinc_i^{j+1}}\, \varphi^t \right) \right] - \left[\sum_{j=1}^{J} \left(P_{smdcdec_i^j}\, \lambda^j \right) \right] \right\} \quad (4.24)$$

$$subject\ to\ constraints$$

$$(4.3) - (4.6)$$

$$(4.7) - (4.9)$$

$$(4.10) - (4.16)$$

4.4 PROBLEM I: SMALL AND MEDIUM DATA CENTERS OPTIMIZATION

The first optimization problem, established by the second layer of the proposed framework, is specifically focused on SMDC and successively on their relationship with DSOs. Through the optimization techniques and strategies that will be detailed, it will be possible to carry out an implementation process capable of extracting the best advantages in terms of energy costs.

4.4.1 Linear Optimization Programming Techniques

Operations Research (OP) techniques have been used to solve many problems since 1950s (Clímaco *et al.* 2003). With each passing year, new forms of solving different sort of problems through OR techniques are proposed. Nonetheless, Linear Programming (LP), which was introduced in 1827 by Joseph Fourier (Sierksma 2001), developed in 1939, during World War II by Leonid Kantorovich (Schrijver 1986), and was framed in the literature by the invention of the Simplex Method in 1947 by George Dantzig (Schrijver 1986), has never lost its acceptance among all those used techniques.

In LP, all functions and constraints are correlated to a variable x with the form $a^t x + b$. These linear optimization techniques can be divided into three main categories: continuous, integer and mixed-integer. The continuous linear method optimizes variables that are commonly real numbers. It is resolved using algorithms, which produce iterated values of the variables until a solution is found (Wright 2010). The integer programming method is comparable to the continuous, however includes an additional constraint that claims that some or all the optimization variables need to be integers. The programming technique used in this thesis is the Mixed Integer Linear Programming (MILP), that was composed on the base of Dantzig's Simplex Method. According to Bixby (2012) the first commercially used mixed integer linear programming code dates back to 1960s. It is characterized by the fact that it associates continuous and discrete variables. Mathematically, the mixed integer technique search for a vector x that maximize or minimize an objective function under a set of constraints (Iqbal *et al.* 2014). The formal mathematical expression is given in the set of Equations and Inequations 4.25, 4.26 and 4.27.

$$min\ c^T \tag{4.25}$$

$$s.t\ Ax \leq b \tag{4.26}$$

$$and\ x \geq 0 \tag{4.27}$$

In this classical demonstration of LP model, min or max depicts the objective of creating the model. For example, the objective function of a cost problem normally aims to minimize the cost, while a profit problem usually tends to maximize the objective function. c is a row vector that denotes coefficients of unknowns in the objective function. b is a column vector that demonstrates right hand side coefficients of constraints. T is the transpose symbol that converts the row vector to a column vector. The abbreviation $s.t$ corresponds to 'subject to' in optimization context. A is a matrix that comprises coefficients of unknowns in constraints. x is a column vector that symbolizes unknown variables which can assume continuous or discrete values. In this context, MILP is applicable; if the problem adopts x as a continuous variable, then the problem is a LP problem, otherwise (x is discrete or integer), it is a MILP problem. Therefore, adding one more constraint, such as $x \in Z^n$, becomes the LP problem in a MILP problem.

MILP is a notorious utilized optimization method for solving many sorts of engineering and business problems, such as travelling salesman problem (a very well-known problem as TSP), transportation, optimal scheduling, optimal dispatch of power generator, etc. There are many optimization package programs, such as MATLAB's Optimization Toolbox, LINDO, Gurobi, Mathcad, MOSEK, GAMS, OptimJ, AMPL. Many other packages that uses C, C++, CPLEX, Java, FORTRAN, Visual Basic, .net, MATLAB interfaces to solve optimization problems are available as a possibility. Moreover, the Microsoft Excel has an Excel Solver add-on that is applicable to deal with optimization problems. However, this study uses MATLAB's 'intlinprog' algorithm that enables the user to choose the preferred solving method among these existing techniques. This algorithm and its corresponding options will be explained with more details in the next section.

4.4.2 MATLAB's 'intlinprog' Algorithm

MATLAB contains many algorithms for a broad diversity of optimization problems, such as linear, nonlinear, and quadratic programming problems in its optimization toolbox. MATLAB's 'intlinprog' algorithm is one of the tools present in this toolbox for solving MILP problems. The word 'intlinprog' is the abbreviation of integer linear programming in the perspective. This algorithm was added into the optimization toolbox in 2014. Considering that MILP has become a known mathematical language in the literature since 1960s, the inclusion date of this algorithm might be comprehended as a late release. However, MATLAB is a modular software with continuous updates. Meanwhile, there has already existed other codes for other programming languages to solve MILP problems among optimization software. The difference that makes this algorithm more compelling is its user-friendliness. The referred algorithm contains a considerably large variety of settings that can be customized by the user in order to address the problems solutions from different aspects. There are essentially three methods used in this algorithm to find optimal solutions for MILP problems: B&B method, cutting plane method, and MATLAB's heuristics method.

The algorithm is developed by default as a minimizer and requires an objective function aiming at minimizing the solution. Hence, a maximization problem has to be transformed in a minimization structure by multiplying its objective function by '-1' in order to allow this solver's solution approach. Furthermore, all inequality constraints of minimization or maximization problem type, which contains greater than or equal to (≥) signs should be also converted into a less than or equal to (≤) by multiplying both sides of constraints by '-1'. Summarizing, the 'intlinprog' algorithm solves problems in the following form the set of Equations and Inequations from 4.28 to 4.33.

$$min\ f^T x \tag{4.28}$$

$$s.t\ A \cdot x \leq b \tag{4.29}$$

$$Aeq \cdot x = beq \tag{4.30}$$

$$lb \leq x \tag{4.31}$$

$$x \leq ub \tag{4.32}$$

$$x(intcon)\ \text{values are integers} \tag{4.33}$$

The 'intlinprog' algorithm fundamentally requires some matrices and vectors, such as A matrix for coefficients of inequality constraints, b column vector for inequality constraints corresponding right hand sides, Aeq matrix for coefficients of equality constraints, beq column vector for equality constraints corresponding right hand sides, f vector for coefficients of variables in the objective function, ub and lb column vectors to set upper bounds and lower bounds for values of variables given in the objective function, and $intcon$ row vector to differentiate discrete variables from continuous variables. After setting these matrices and vectors for a specific MILP problem, the 'intlinprog' algorithm applies the above-mentioned three MILP solving procedures and identifies the optimal solution.

4.4.3 Electricity Price Fluctuation

Before defining the mathematical model based on a MILP algorithm to implement the SMDC optimization, different electricity price fluctuation scenarios are considered and the electricity market is regulated depending on the economic conjuncture and the energy policy of a given country. The electricity price methodology that will be adopted is from Portugal.

According to *Energy Services Regulatory Authority* (2018) in Portugal "in the context of the markets liberalization process, where the activities of the network operators are considered natural monopolies and are therefore subject to economic regulation, the production and selling of electricity are open to competition.

The Iberian Electricity Market (MIBEL) is a platform where electricity is traded for delivery the day following that of the negotiation. This market forms the price for each one of the 24 hours of each day and for each one of the 365 or 366 days of each year. The market price at each hour is found through a process where the price of the production offers is placed in ascending order (supply curve) and the price of the electricity buying offers are placed in descending order (demand curve) for the same time. The market price (corresponding to the crossing of the supply and demand curves on a graph) is the lowest price which guarantees that supply satisfies demand. The operating rules of this organized market are specific to the market operator (OMEL).

The process development for the electricity sector liberalization dictated the trading market opening, and, in the present framework, any consumer can freely choose their electricity supplier.

The evolution of the retail market conditions, namely pertaining in what concerns the price of electricity, is clearly restricted by the evolution of the wholesale market, since the latter determines a substantial part (energy costs) of the total costs of the supply of electricity."

In this context, three electricity price fluctuation scenarios are taken into account:

- The first one is based on an hourly electricity price fluctuation, whereby the tariffs are dynamic and fluctuate hourly depending on the energy generation, consumption and the peak hourly loads, as shows in Figure 4.3.

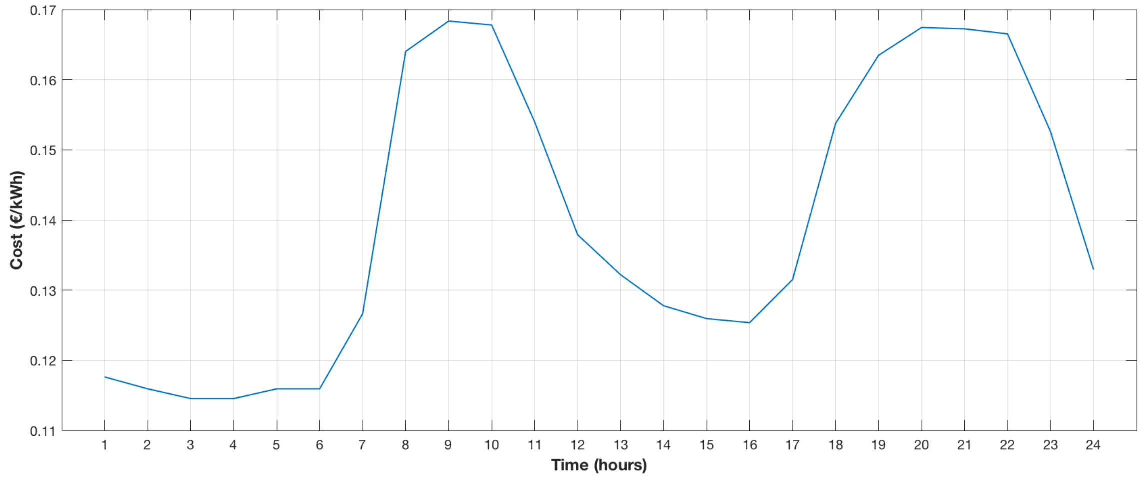


Figure 4.3 – Hourly electricity price fluctuation

This tariff was firstly composed using as sample the prices of energy in the wholesale market on February 22, 2018, which is a representative day in terms of the price variation in Winter. The next step was to use the average energy price of a service building, namely the Department of Electrical and Computer Engineering (DEEC) of the University of Coimbra (UC) in Portugal. The ratio between the average prices of the service building and the wholesale market was calculated to have a tariff with the hourly variation of the wholesale market and the average price of the building.

Following the same premises, in order to present the same prices in 20 minute intervals, each hour is divided into three periods, totalizing 72 different markings, as depicted in Figure 4.4. Although it is more common to consider periods of a quarter of an hour, i.e., 4 periods of 15 minutes, for the purpose of simplification this work will consider 20 minutes to coincide with the number of constraints described in optimization problems in SMDC.

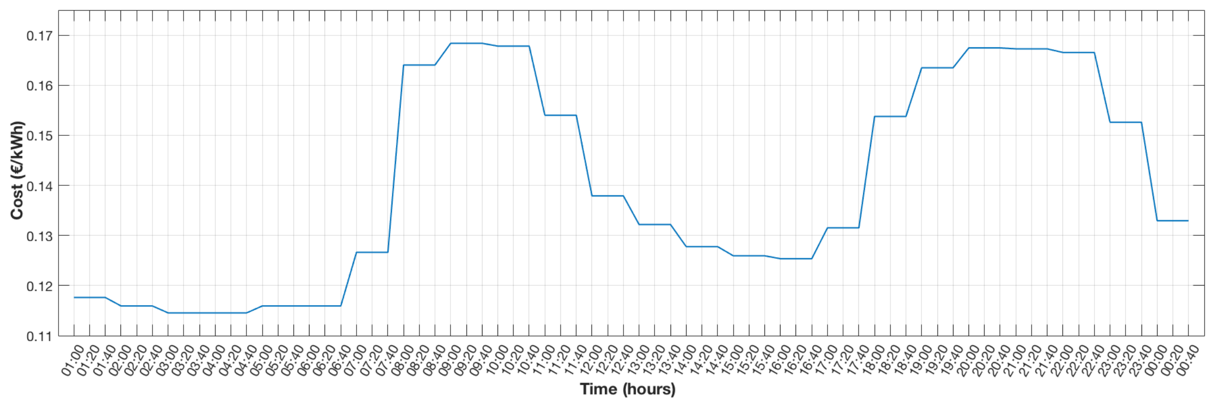


Figure 4.4 – Hourly electricity price fluctuation in 20 minute intervals

- In the second scenario, the same hourly electricity price fluctuation is taken into account, however in a specific 20-minutes period the price is induced to stimulate a DR action (price increase) and shortly thereafter contemplate a RE situation (price decrease), as shown in Figure 4.5.

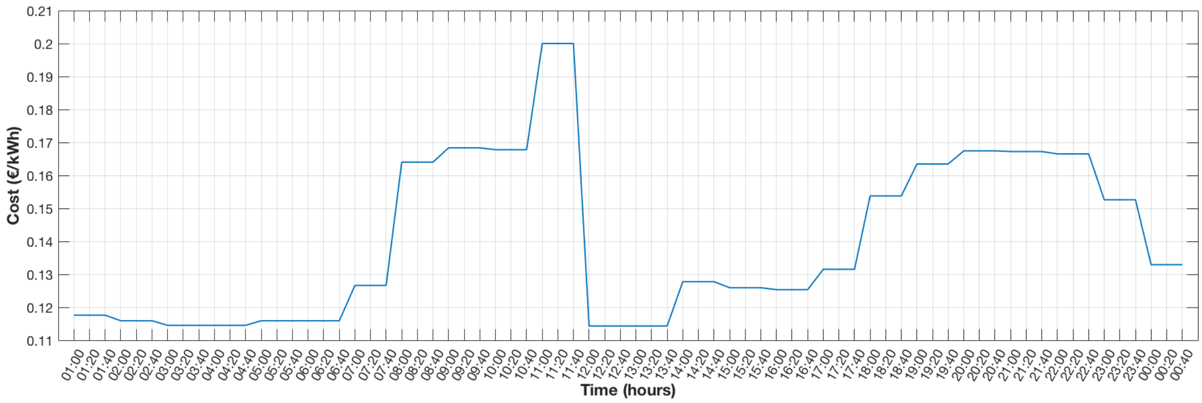


Figure 4.5 – Hourly electricity price fluctuation applied to a DR action

- The third case also considers three periods of 20 minutes in each hour, but using a Time-of Use tariff with four periods, as can be seen in Figure 4.6, namely: peak, half-peak, normal off-peak and super off-peak hours. The prices established in this composition are originated from those practiced by the DEEC service building at the UC. Additionally, in the DR period, it is considered the application of a financial incentive.

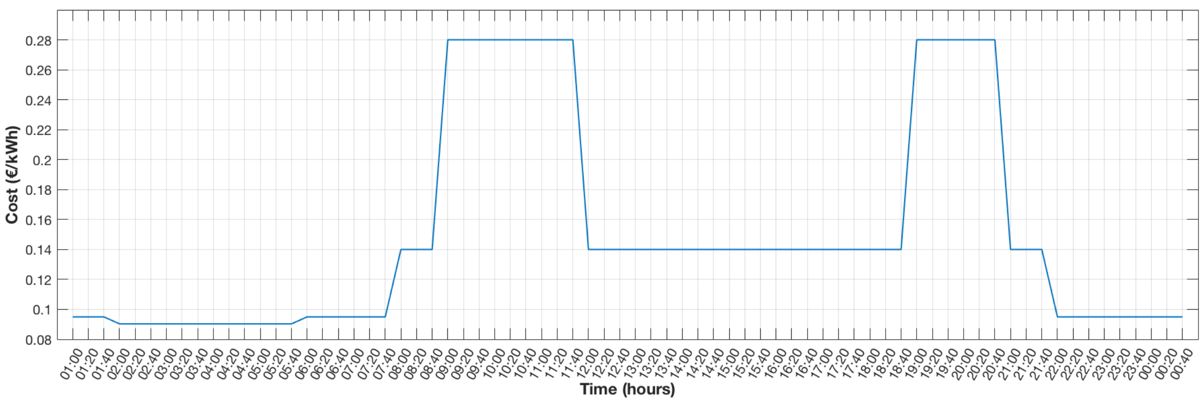


Figure 4.6 – Tariff periods electricity price fluctuation

4.4.4 Optimization Scenarios

Taking into account the electricity price fluctuations already defined, three different optimization scenarios will be established for SMDC using the MATLAB 'intlinprog' algorithm in MILP simulations with their respective constraints.

- Considering that SMDC can take advantage of dynamic tariffs in an hourly electricity price fluctuation to reduce power in a DR event, the first scenario aims to find an optimal price time for this reduction, as well as setting the best tariff windows to smooth the rebound effect process, as presented in the example code of Annex 1.1.1. Thus, the optimization function is given by the defined electricity cost value φ^{24} and the following $I \times 24$ matrix in Equation 4.34, where I is the number of SMDC.

$$\sum_{i=1}^{I}\left[\sum_{t=1}^{T}\left(P_{smdc_i^t}\,\varphi^t\right)\right] = \begin{bmatrix} P_{smdc_1^1} & P_{smdc_1^2} & P_{smdc_1^3} & . & . & . & P_{smdc_1^{24}} \\ P_{smdc_2^1} & P_{smdc_2^2} & P_{smdc_2^3} & . & . & . & P_{smdc_2^{24}} \\ P_{smdc_3^1} & P_{smdc_3^2} & P_{smdc_3^3} & . & . & . & P_{smdc_3^{24}} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ P_{smdc_I^1} & P_{smdc_I^2} & P_{smdc_I^3} & . & . & . & P_{smdc_I^{24}} \end{bmatrix} \times \begin{matrix} \varphi^1 \\ \varphi^2 \\ \varphi^3 \\ . \\ . \\ . \\ \varphi^{24} \end{matrix}$$

$$= \begin{bmatrix} \varphi^1 \times P_{smdc_1^1} & \varphi^2 \times P_{smdc_1^2} & \varphi^3 \times P_{smdc_1^3} & . & . & . & \varphi^{24} \times P_{smdc_1^{24}} \\ \varphi^1 \times P_{smdc_2^1} & \varphi^2 \times P_{smdc_2^2} & \varphi^3 \times P_{smdc_2^3} & . & . & . & \varphi^{24} \times P_{smdc_2^{24}} \\ \varphi^1 \times P_{smdc_3^1} & \varphi^2 * P_{smdc_3^2} & \varphi^3 \times P_{smdc_3^3} & . & . & . & \varphi^{24} \times P_{smdc_3^{24}} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ \varphi^1 \times P_{smdc_I^1} & \varphi^2 \times P_{smdc_I^2} & \varphi^3 \times P_{smdc_I^3} & . & . & . & \varphi^{24} \times P_{smdc_I^{24}} \end{bmatrix} \tag{4.34}$$

With the same price variation, but a different aggregation of data defined by periods of 20 minutes every hour, the optimization function is given by the defined electricity cost value φ^{72} along with $I \times 72$ normal operation matrix, $I \times 1$ DR matrix and $I \times 1$ RE matrix, compounding Equation 4.35, where I is the number of SMDC.

$$\sum_{i=1}^{I}\left\{\left[\sum_{t=1}^{T} P_{smdc_i^t}\,\varphi^t\right] - \left[\sum_{j=1}^{J}\left(P_{smdcdec_i^j}\,\varphi^j\right)\right] + \left[\sum_{j=1}^{J}\left(P_{smdcinc_i^{j+1}}\,\varphi^{j+1}\right)\right]\right\}$$

$$= \sum_{i=1}^{I}\left\{\left[\sum_{t=1}^{T}\left(\begin{bmatrix} P_{smdc_1^1} & P_{smdc_1^2} & P_{smdc_1^3} & . & . & . & P_{smdc_1^{72}} \\ P_{smdc_2^1} & P_{smdc_2^2} & P_{smdc_2^3} & . & . & . & P_{smdc_2^{72}} \\ P_{smdc_3^1} & P_{smdc_3^2} & P_{smdc_3^3} & . & . & . & P_{smdc_3^{72}} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ P_{smdc_I^1} & P_{smdc_I^2} & P_{smdc_I^3} & . & . & . & P_{smdc_I^{72}} \end{bmatrix} \times \begin{bmatrix} \varphi^1 \\ \varphi^2 \\ \varphi^3 \\ . \\ . \\ . \\ \varphi^{72} \end{bmatrix}\right)\right] - \left[\sum_{j=1}^{J}\left(\left[P_{smdcdec_i^j}\right] \times \left[\varphi^j\right]\right)\right] + \left[\sum_{j=1}^{J}\left(\left[P_{smdcdec_i^{j+1}}\right] \times \left[\varphi^{j+1}\right]\right)\right]\right\} \tag{4.35}$$

$$= \sum_{i=1}^{I}\left\{\left[\sum_{t=1}^{T}\left(\begin{bmatrix} \varphi^1 \times P_{smdc_1^1} & \varphi^2 \times P_{smdc_1^2} & \varphi^3 \times P_{smdc_1^3} & . & . & . & \varphi^{72} \times P_{smdc_1^{72}} \\ \varphi^1 \times P_{smdc_2^1} & \varphi^2 \times P_{smdc_2^2} & \varphi^3 \times P_{smdc_2^3} & . & . & . & \varphi^{72} \times P_{smdc_2^{72}} \\ \varphi^1 \times P_{smdc_3^1} & \varphi^2 \times P_{smdc_3^2} & \varphi^3 \times P_{smdc_3^3} & . & . & . & \varphi^{72} \times P_{smdc_3^{72}} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ \varphi^1 \times P_{smdc_I^1} & \varphi^2 \times P_{smdc_I^2} & \varphi^3 \times P_{smdc_I^3} & . & . & . & \varphi^{72} \times P_{smdc_I^{72}} \end{bmatrix}\right)\right] - \left[\sum_{j=1}^{J}\left(\left[P_{smdcdec_i^j} \times \varphi^j\right]\right)\right] + \left[\sum_{j=1}^{J}\left(\left[P_{smdcdec_i^{j+1}} \times \varphi^{j+1}\right]\right)\right]\right\}$$

- The second scenario used the same matrix arrangement of Equation 4.35, but the change of prices in DR e RE circumstances vary sharply as previous defined e to stimulate a DR event and contemplate a RE, as presented in the example codes of Annexes 1.1.2 and 1.1.3.

- In the third scenario, the objective is that SMDC can receive incentives for each reduction in the power percentage agreed in the contract and for this purpose it is considered that DR and RE occurs in the same tariff period, as presented in the example code of Annex 1.1.4. Thereby, the optimization function is given by the defined electricity cost value φ^{72} along with $I \times 72$ normal operation matrix, $I \times 1$ DR matrix, $I \times 1$ RE matrix and the addition of $I \times 1$ incentive matrix, compounding Equation 4.36, where I is the number of SMDC.

$$
\begin{aligned}
&\sum_{i=1}^{I}\left\{\left[\sum_{t=1}^{T}\left(P_{smdc_i^t}\,\varphi^t\right)\right]-\left[\sum_{j=1}^{J}\left(P_{smdcdec_i^j}\,\varphi^t\right)\right]+\left[\sum_{j=1}^{J}\left(P_{smdcinc_i^{j+1}}\,\varphi^t\right)\right]-\left[\sum_{j=1}^{J}\left(P_{smdcdec_i^j}\,\lambda^j\right)\right]\right\}\\
&=\sum_{i=1}^{I}\left\{\left[\sum_{t=1}^{T}\left(\begin{bmatrix} P_{smdc_1^1} & P_{smdc_1^2} & P_{smdc_1^3} & \cdots & P_{smdc_1^{72}}\\ P_{smdc_2^1} & P_{smdc_2^2} & P_{smdc_2^3} & \cdots & P_{smdc_2^{72}}\\ P_{smdc_3^1} & P_{smdc_3^2} & P_{smdc_3^3} & \cdots & P_{smdc_3^{72}}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ P_{smdc_I^1} & P_{smdc_I^2} & P_{smdc_I^3} & \cdots & P_{smdc_I^{72}}\end{bmatrix}\times\begin{bmatrix}\varphi^1\\ \varphi^2\\ \varphi^3\\ \vdots\\ \varphi^{72}\end{bmatrix}\right)\right]-\left[\sum_{t=1}^{T}\left([P_{smdcdec_i^t}]\times[\varphi^t]\right)\right]+\left[\sum_{t=1}^{T}\left([P_{smdcinc_i^t}]\times[\varphi^t]\right)\right]-\left[\sum_{j=1}^{J}\left([P_{smdcdec_i^j}]\times[\lambda^j]\right)\right]\right\}\\
&=\sum_{i=1}^{I}\left\{\left[\sum_{t=1}^{T}\left(\begin{bmatrix} \varphi^1\times P_{smdc_1^1} & \varphi^2\times P_{smdc_1^2} & \varphi^3\times P_{smdc_1^3} & \cdots & \varphi^{72}\times P_{smdc_1^{72}}\\ \varphi^1\times P_{smdc_2^1} & \varphi^2\times P_{smdc_2^2} & \varphi^3\times P_{smdc_2^3} & \cdots & \varphi^{72}\times P_{smdc_2^{72}}\\ \varphi^1\times P_{smdc_3^1} & \varphi^2\times P_{smdc_3^2} & \varphi^3\times P_{smdc_3^3} & \cdots & \varphi^{72}\times P_{smdc_3^{72}}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \varphi^1\times P_{smdc_I^1} & \varphi^2\times P_{smdc_I^2} & \varphi^3\times P_{smdc_I^3} & \cdots & \varphi^{72}\times P_{smdc_I^{72}}\end{bmatrix}\right)\right]-\left[\sum_{t=1}^{T}\left([P_{smdcdec_i^t}\times\varphi^t]\right)\right]+\left[\sum_{t=1}^{T}\left([P_{smdcinc_i^t}\times\varphi^t]\right)\right]-\left[\sum_{j=1}^{J}\left([P_{smdcdec_i^j}\times\lambda^j]\right)\right]\right\}
\end{aligned}
\tag{4.36}
$$

The three assumed scenarios are subjected to three different constraints. The SMDC total power is expressed by $P_{smdc_i^t}$, however either $P_{smdcdec_i^j}$ in a DR event, or $P_{smdcinc_i^{j+1}}$ in a RE situation are respectively composed in Equation 4.37 and 4.38 of the sum of the ICT and non-ICT loads, the latter being the aggregation of the cooling and the UPS workloads.

$$P_{smdcdec_i^j} = P_{wflex_i^j} + P_{ups_i^j} + P_{cool_i^j} \tag{4.37}$$

$$P_{smdcinc_i^{j+1}} = P_{wflex_i^{j+1}} + P_{ups_i^{j+1}} + P_{cool_i^{j+1}} \tag{4.38}$$

Thereby, in the first constraint the sum of the decreased power $P_{wflex_i^j}$, $P_{drups_i^j}$, $P_{drcool_i^j}$ in a DR event cannot exceed the contractual power reduction target of -105 kW for small data centers, as formulated by Equation 4.39 and -550 kW for medium data centers, as set by Equation 4.40.

$$P_{wflex_i^j} + P_{ups_i^j} + P_{cool_i^j} \leq -105 \tag{4.39}$$

$$P_{wflex_i^{j+1}} + P_{ups_i^{j+1}} + P_{cool_i^{j+1}} \leq -550 \tag{4.40}$$

In the second constraint, the sum of the increased power $P_{wflex_i^j}, P_{drups_i^j}, P_{drcool_i^j}$ in a RE situation cannot exceed 105 kW for small data centers, as formulated by Equation 4.41 and 550 kW for medium data centers, as set by Equation 4.42.

$$P_{wflex_i^j} + P_{ups_i^j} + P_{cool_i^j} \leq 105 \tag{4.41}$$

$$P_{wflex_i^{j+1}} + P_{ups_i^{j+1}} + P_{cool_i^{j+1}} \leq 550 \tag{4.42}$$

The third constraint establishes the SMDC lower and upper boundaries for each strategy of power reduction and raising in a DR event and RE situation respectively, as can be seen in Equations 4.43, 4.44 and 4.45.

$$lb \leq P_{wflex_i} \leq ub \tag{4.43}$$

$$lb \leq P_{ups_i} \leq ub \tag{4.44}$$

$$lb \leq P_{cool_i} \leq ub \tag{4.45}$$

4.5 PROBLEM 2: DISTRIBUTION SYSTEM OPERATOR OPTIMIZATION

In the course of this work it has been highlighted the importance to implement new policies that contemplate the relationships established between the DSO and SMDCs. The best way to build these bases will be through the growth of optimization, simulation and analysis of results process. Specifically, in the case of DSOs, it is extremely important to focus on optimization processes that allow predicting and anticipating the key relationships and interactions that can occur contractually between the parties. In this way, it is possible to strengthen existing policies and propose new ones from such observations.

In the first problem, already described, the interaction proposed by the framework between the actors established in the scope of this work, occurs fundamentally from SMDC responding to a direct, or indirect request of DSO. The implementation of this process can happen through dynamic tariffs, or by contracts. In other words, SMDC can act individually to take advantage of the best tariff window by reducing power in a DR event and consequently minimizing their energy costs. In another form, they can sign a DSO contract where incentives are received if a given percentage of power is reduced, or a penalty payment will be due otherwise.

However, the interaction flow between the actors in the second problem is inverse, i.e., it fundamentally happens with DSOs direct, or indirectly requiring a response from SMDC. In this case, the mechanisms of implementation can also occur by tariffs, or contractually, whereby the operationalization terms are established by DSO and if it is advantageous, signed by any SMDC that meet the prerequisites.

4.5.1 Contractual Terms

In this context, the DSO contract is optimized by a random-rotating and fairness algorithm capable to define, after manifest interest, which SMDC will be chosen in each DR event. In order to promote fairness among SMDC participations, the algorithm also establishes a justice criterion balancing the various selections during a given contractual period.

On one hand, the contract proposed in Basmadjian, Niedermeier, *et al.* (2013) and also adopted in Basmadjian *et al.* (2015) advocates eight terms.

- The first and second contractual terms correspond respectively to a minimum and maximum power reduction expressed in kW. For each of the above-mentioned terms, two sub-terms are defined that specify the minimum and maximum duration (in minutes) of the corresponding power adaptation.
- The third term represents the maximum amount of time (in minutes) DC needs to send back a reply to EP.
- The fourth term specifies the maximum number of rejections by DC to EP's power adaptation requests on a monthly basis.
- The fifth term defines the maximum number of successive rejections allowed by DC.
- The sixth term states the maximum number of requests EP can send to each DC on a monthly basis.
- Since recovering from a power adaptation needs to be performed within a reasonable amount of time, the seventh term specifies the minimum period of time (in minutes) between two successive EP's power adaptation requests to DC.
- The eighth term guarantees that the DC has enough time (in minutes) to receive a notice from EP regarding a potential power adaptation request. Every time one of the parties' breaches one or more contractual terms of this contract, a penalty is applied. Also, incentives are created based on the signed terms.

Therefore, as can be noticed, this contractual profile in addition to establishing multiple terms adding greater complexity, increases the probability of incurring in penalty cases, making this approach more restrictive, penalty-oriented, and hindering adoption by data centers data centers, whose the primary mission is to provide ICT services rather than energy services.

On the other hand, the proposed contract by this work establishes only four terms, in order to reduce the complexity, making the contractual model more attractive to SMDC. Nonetheless, as already mentioned, one of the added values of this contract profile is that SMDC need, before signing the contract, to ensure that they have conditions to reduce the power objective established by the DSO: 105 kW for small data centers and 550 kW for medium data centers. This premise, in addition to avoid the power adaptation profile phenomenon adopted by Basmadjian *et al.* (2015) and explained in the scope of the Section 4.2, also avoids the surprise and complexity of dealing with various load profiles not long before the DR event began. Hence, the four terms are:

- In the first term, each SMDC can establish along with the DSO the sum of power (accumulated during a period of time) that can be reduced during a cycle, for example, a monthly cycle. This liberation alternates the penalty-oriented flow and causes the goal of reducing a critical mission power to gain financial advantage, becoming more attractive, anyway if a data center has stipulated a maximum power to be accomplished during a cycle, it will do its utmost to achieve that value to receive the proposed incentive.
- Regarding the number of DR events to participate during a cycle, the second term states that to data center shall be granted the right not to participate in a given event, claiming operational or maintenance reasons. Non-participation will not count for incentive, or penalty effects. Therefore, the number of participations of a data center is conditioned to the maximum power agreed per cycle and the criterion of randomness and fairness of the algorithm. This policy reinforces the DSO interest of contemplating and creating mechanisms to make the aggregate participation of consumers as small and medium-sized data centers feasible and operational.
- The third term specifies that a DR event can occur in one single time window, or more and the power required cannot exceed the total power considering the sum of all data centers present in the contract. This policy assures SMDCs that the DSO will not take advantage of this type of contract in addition to what was previously established.
- The fourth term defines that the correct ratio between small and medium data centers present in the contract must be obeyed at each DR event in order that DSO does not take advantage of larger loads or penalize a group of data centers over others, ensuring equal conditions among participants.

4.5.2 Random-Rotating and Fairness Algorithm Overview

The algorithm behavior must meet the following conditions:

- The DSO will send a demand response signal to all SMDC included in the contract. This DR signal is characterized by two vectors announcing periods of time and the associated power to be reduced. Thus, from the time vector it can extract the total duration of the DR event and the power vector, the maximum and average values.
- Before sending a DR signal, the DSO can predict the amount of SMDC that will be required to attend the DR event based on the premise that ensures data centers with different dimensions will participate on equal terms of choice, following the ratio criterion of participating data centers established in the contract.
- After receiving such a signal, one or more SMDC, according to their contractual premises, will give an answer expressing interest in participating in the DR event (ACK), making themselves available to try reducing their loads during the time set by DSO.
- A certain data center can deny the participation in a DR event (NACK) claiming operational reasons, e.g. maintenance, change or devices upgrade.
- If after sending the DR signal there are more SMDC interested in participating of the DR event than necessary to achieve the reduction, a first random selection with a uniform distribution is done, creating a list with all the respondents. Then, the SMDC needed for reduction over the defined time periods are going to be used. In a second selection process until the end of cycle, make a new random selection, but now contemplating the criterion of fairness, which will prioritize those data centers who have not yet participated. In case of all data centers have already participated, the selection process will be set randomly.
- Then, reduce certain power in each data center within the time window defined in the agreement (e.g. minutes, or hours per day) and in the signal of DR. This reduction must occur consecutively using the random and fair list of choice with a uniform distribution within an interval time window of 10s, in order to minimize the impact of the rebound effect.
- The DR event is analyzed in small time fractions enacting the above reduction criterion, rather than the entire DR time window. This mechanism will provide more reliability in terms of grid stability and security since the distribution of power reduction by time will be steadier and smoother.
- This power is going to be decrease through priority workloads defined by SMDC operators and mentioned in the contract.

- The workloads should follow the criteria of initially prioritizing the non-ICT loads and only after the ICT loads.
- During the time window where the DR event will happen, the power to be reduced is monitored in real time to assess if the defined objectives are achieved. If not, new remaining SMDC from the previous list are called to meet the requirements.
- The reduced power from each chosen SMDC is decreased from the previously established accumulated power reduction present in the contract.
- At the next DR event, the remaining accumulated power is considered to the reduction and no more the total power.
- The above-mentioned procedures at each DR event are repeated.
- After the contractual cycle closes, perform calculations to grant incentives, apply penalties, or present data centers that will not have to receive or pay any amount penalties in an unchanged scenario.
- Print the needed outputs based on a balance sheet on function of events, acceptance rate, reduced power and financial information. The parameters to do these calculations should be the amount of reduced power, the associated reduction percentage and the incentive and penalty rates considered: 1 € for small data centers and 1.5 € for medium ones. The adoption criterion for these values was the cost of power during peak demand hours, i.e., 10 and 15 times the cost of the additional that is paid during such hours.
- Note that at the beginning of every cycle, the contract reduction parameters need to be reset;
- The code development of this algorithm was also implemented in MATLAB and its main functions were separated into different files, as represented by Figure 4.7.

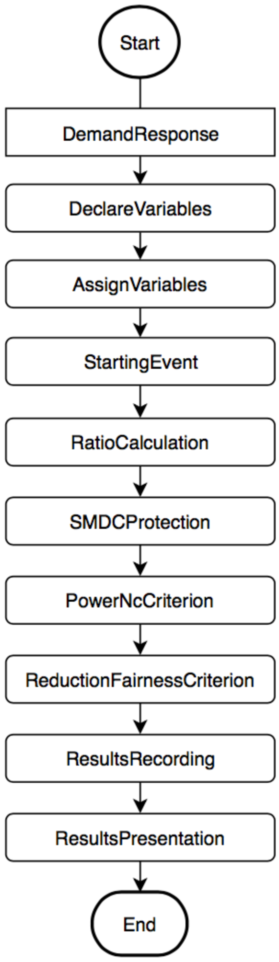


Figure 4.7 – Main algorithm files

It is important to highlight the assignment of each file, as follows:

- **Start:** it runs the "Demand Response" file and at the same time imposes a constraint preventing that, at each event, DSO requests a larger amount of power to reduce than the total power offered by the data centers present in the contract.
- **DemandResponse:** it is the main file where the other instructions are called. It is presented in Annex 1.2.1 an example of this code.
- **DeclareVariables:** it stores the main table inputs in variables format, setting the amount of DR events and the percentage of data centers participation by event. In practice, the instructions in this file simulate the sending of the DR signal and all data centers that might reject, or accept to participate in each event in order to take advantage of power reduction.
- **AssignVariables:** the flow continues and the average power to be decreased in multiple time windows in the DR events is calculated calling the "Starting Event" file. Thereafter, in this stage yet, the process of random selection of all the data centers that previously accepted the demand response signal occurs.

- **StartingEvent:** this file is responsible to calculate the average power that will be reduced in each time window required in the DR event.
- **RatioCalculation:** three types of different ratio calculations are performed by this file. Firstly, it is calculated the ratio of SMDC present in the contract by dividing the number of small participants by the total of data centers and doing the same process for the medium profile. Thereafter, it is needed to calculate the amount of SMDC to participate on the DR event, as well as their respective time windows by dividing the amount of small data centers that accepted the demand response signal by the amount of time windows and doing the same process for the medium profile. Finally, multiplying separately the average power to be decreased by the respective SMDC ratio previously calculated, it is assessed how much power should be decreased by each small and medium profile, as presented in the example code in Annex 1.2.2.
- **PowerNcCriterion:** this file, presented as an example in Annex 1.2.3, uses an approach based on five different power profiles to set how much power each data center will decrease by event. In the contract, data centers commit themselves to a defined amount of flexible load to be reduced, once they accept to participate in a DR event. However, during the DR event this situation can be occur as agreed, or in a different way. Thereby, the power values considered in Table 4.2 to simulate a real scenario are:
    - Accomplish: 105 kW for small data centers and 550 kW for medium data centers.
    - Slightly above: 20% above of the set power.
    - Slightly below: 20% below of the set power.
    - Below: between 20% and 79% of the set power.
    - Quite bellow: between 0.1% and 19% of the set power.

The values used to fulfill the input table are derived from the result of equations that demonstrate the sum of the three flexible load profiles used in this work: flexible ICT workload, set point adjustment of cooling devices and UPS discharge.

Once the values have been defined, percentages will be assigned to each of the 5 profiles, which in turn will be distributed randomly among all participants data centers in the contract, for simulation purposes. For example, if a contract has 16 small and 16 medium data centers, at each demand response event these 5 power profiles will be randomly distributed among them. It is important to emphasize the possibility of assigning different weights between these profiles in order to simulate scenarios where, for example, the predominance is data centers that meet the agreed, or that are a little above, significantly below and so forth.

Table 4.2 – Power SMDC values

Small Data Centers					Medium Data Centers			
Pdrwork	Pdrups	Pdrcool	Total	Profile	Pdrwork	Pdrups	Pdrcool	Total
40	30	35	**105**	accomplished	**220**	150	180	**550**
23	**55**	27	**105**	accomplished	120	**230**	200	**550**
30	23	**52**	**105**	accomplished	150	100	**300**	**550**
52	33	22	**107**	slightly above	**215**	185	155	**555**
23	**59**	27	**109**	slightly above	140	**230**	200	**570**
28	35	**47**	**110**	slightly above	166	190	**195**	**551**
37	28	30	**95**	slightly below	**219**	150	180	**549**
29	**41**	24	**94**	slightly below	170	**200**	160	**530**
20	29	**37**	**86**	slightly below	100	190	**195**	**485**
24	18	23	**65**	below	**200**	100	130	**430**
23	**31**	10	**64**	below	70	**100**	60	**230**
14	10	**16**	**40**	below	119	50	**186**	**355**
9	6	4	**19**	quite below	**53**	33	22	**108**
4	**10**	4	**18**	quite below	32	**48**	25	**105**
1	2	**5**	**8**	quite below	20	29	**37**	**86**

- **ReductionFairnessCriterion:** the sending of the demand response signal is, in practice, to broadcast to all data centers of the contract the average power to be reduced during the event, as it occurs in the “Star Event” file. Subsequently, the present file is responsible to check if the number of SMDC that accepted to participate in the event is less than or equal to the available amount of data centers. If this condition is met, the algorithm will choose the data center with the least amount of participations. If it is the first event of the cycle, where there was no participation, or all of them have the same amount of participations in other events, then the choice will be randomly selected. After this fairness process, the number of requests and the amount of power reduced by the chosen data centers in the event are stored, as presented in the example code of Annex 1.2.4.
- **ResultsRecording:** Considering that a cycle is a quantity of demand response events that have occurred, regardless of whether it is weekly or monthly, at the end of this process all the results are stored in their respective variables, being ready to be presented in charts, or tables.
- **ResultsPresentation:** This file neatly displays all charts and tables generated after the end of the DSO demand response cycle, such as:
    - Acceptance percentage by data center dimension.
    - Denial percentage by data center dimension.
    - Joint acceptance and denial percentage by data center dimension.
    - Total accepted calls.

- Total rejected calls.
- Average power and percentage of non-accomplishment.
- Participation by events.
- Power reduction percentage.
- Data centers power reduction.
- Minimum, average and maximum reduction per each data center.
- DSO minimum, average and maximum reduction.
- Required and reduced power by event.
- Small and medium data center average power reduction by event.
- Required and achieved power.
- Total reduction achievement.
- Incentive, unchanged and penalty numbers by data center dimension.
- Incentive, unchanged and penalty percentage.
- Financial return.

4.5.3 Random-Rotating and Fairness Algorithm Flow

It is fundamental to have a real perspective of the algorithm flow, mainly of the loop and conditional structures. Through the flowchart presented in Figure 4.8 it is possible reinforce the core elements of the code, such as: the specific ratio calculations, how the Nc power profiles are assigned, the main contractual thresholds, in addition to random, fairness and reduction criteria, which are the basis where the algorithm was designed.

Before the initialization process it is important to point out that the main input element of the algorithm is a table that makes use of Excel software and internally has three structured spreadsheets in the form of matrices.

- Ssmdc: which represents the contractual information matrix assigned to data centers of small and medium profile and their respective ratio.
- Sdso: defines the representative matrix of the DR signal sent by the DSO in each DR event.
- Snc: is the matrix that has the responsibility for randomly establishing the different power profiles that will be part of the simulation process.

When the algorithm starts each of these variables is assigned, and the total amount of DR events is stored, being part of the loop that defines a complete cycle and the random and fair simulation of SMDC participation and choice begins.

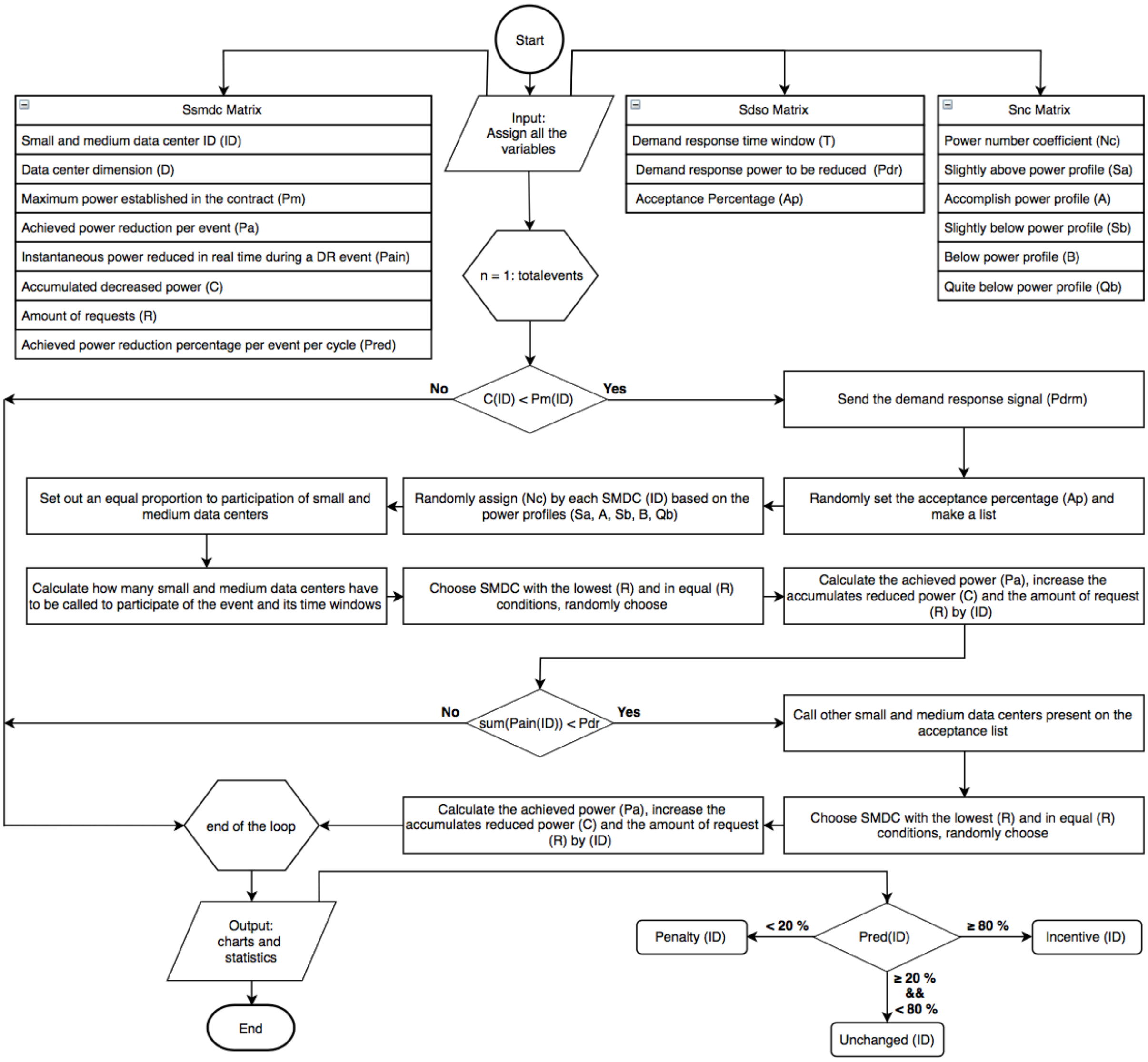


Figure 4.8 – Random-rotating and fairness algorithm flowchart

The first condition structure checks whether the data center has its maximum power limit per cycle reached. If yes, it leaves the loop, finishes the cycle, prints the results and finds out if it received an incentive, has to pay a penalty, or is in an unchanged situation. If not, it is able to participate and eventually be chosen in a DR event following all the participation and power reduction process until it goes to the next condition structure. In this case, the DSO checks through the instantaneous power whether the required power broadcasted in DR signal was achieved. If yes, it leaves the loop, finishes the cycle, prints the results and follow the same process of the other condition. But if not, other data centers are called to try reaching the required power and follow the previous condition steps.

In order to better comprehend the flow and specificities of this algorithm a sequential line by line description is provided in Algorithm 4.1.

Algorithm 4.1: Random-rotating and fairness in small and medium data centers

Require: ***Scon:*** **Set of all SMDC present in the DR contract**

Ensure: ***Ssmdc:*** **Matrix (*ID; D; Pm; Pa; Pain; C; R; Pred*), where**

ID: is the unique data center recognized by one single identification

D: data center dimension (small-size (SDC) or medium size (MDC))

Pm: is the maximum power reduction allowed per cycle

Pa: is the achieved power reduction per event

Pain: is the instantaneous power reduced in real time during a DR event

C: is the accumulated of power reduced during each DR event

R: is the amount of request

Pred: is the achieved power reduction percentage per event per cycle

Sdso: **Matrix (*T; Pdr; Ap*), where:**

T: DR time window vector

Pdr: Vector of power to be reduced in a DR event by all chosen SMDC

Pdrm: Average power to be reduced in a DR event and its time windows by all chosen SMDC

Ap: Acceptance percentage by DR event

Snc: **Matrix (*Nc; Sa; A; Sb; B; Qb*), where:**

Nc: Power number coefficient

Sa: Slightly above power profile

A: Accomplish power profile

Sb: Slightly below power profile

B: Below power profile

Qb: Quite below power profile

1: **for** *Ssmdc* $\in$ *Scon && n = 1: totalevents* **do**

2: **if** $C \leq Pm \in Ssmdc$ **then**

3: send the DR signal *Sdso* to all SMDC included in the contract with

Pdrm = Pdr(1) +Pdr(2)…+Pdr(N) / T(1) +T(2)…+T(N)

4: randomly set *Ap* (ACK) and make a list

5: randomly assign Nc by each SMDC ID based on the power profiles (*Sa, A, Sb, B, Qb)* $\in$ *Snc*

6: set out an equal ratio to participation of SMDC that have accepted the DR signal

pSDC = TotalSDCAcceptance / TotalIDList && pMDC = TotalMDCAcceptance/ TotalIDList

7: calculate how many SDMC have to be called to participate of the event and its time windows

chSDC = TotalSDCAcceptance / TNumber && chMDC = totalMDCAcceptance/ TNumber

8: choose SMDC with the lowest *R* and in equal *R* conditions, randomly choose

9: calculate *Pa by each ID* in order to decrease power

*PaSDC=Pdrm * pSDC && PaMDC = Pdrm * pMDC*

10: increase *C by each* participant *ID*

```
        C(ID) = C(ID) + PaSDC && C(ID) = C(ID) + PaMDC
        increase R by each participant ID
        R(ID) = R(ID) + 1
          if  Pain(ID1) + Pain(ID2) + Pain(IDN) ∈ Ssmdc ≤ Pdr(T) then
            choose one or more SMDC from the previous list according to D
            choose SMDC with the lowest R and with equal R conditions, randomly choose.
            calculate Pa by each ID in order to decrease power
           PaSDC=Pdrm * pSDC && PaMDC = Pdrm * pMDC
           increase C by each participant ID
           C(ID) = C(ID) + PaSDC && C(ID) = C(ID) + PaMDC
           increase R by each participant ID
           R(ID) = R(ID) + 1
         else
              go to the end of the loop
         end if
      else
        randomly send (NACK) rejecting the DR signal Sdso
        keep the previous values of power
        C(ID) = C(ID) + 0
         keep the previous values of request
         R(ID) = R(ID) + 0
         go to the end of the loop
        end if
end for
assign the reduction percentage per cycle
Pred(ID) = (C / (D * R)) * 100
if at the end of the cycle for each SMDC Pred(ID)  ≥ 80 % then
   SMDC will receive the incentive
Pred(ID) / 100 * C  * 1 && Pred(ID) / 100 * C  * 1.5
elseif at the end of the cycle for each SMDC Pred(ID)  ≥ 20 % && Pred(ID) < 80 % then
         SMDC will not pay the penalty and will not receive the incentive
else at the end of the cycle for each SMDC Pred(ID)  < 20 % then
        SMDC will pay the penalty
(20% - Pred(ID) / 100) *(21 kW * R - C) * 1 && (20% - Pred(ID) / 100) *(110 kW * R - C) * 1.5
end if
print the charts and tables related to the output variables
all parameters will be reset at the end of the cycle
```

Finally, this chapter proposed a new framework according to the previously set goals and study approaches. Using two layers it was possible to define the main relations and interactions between SMDC and DSO in order to pursue the best practices in terms of energy efficiency and at the same time to take advantage from DR programs. Specific domains such as the mathematical modeling and optimization strategies and techniques were detailed and discussed aiming to highlight their characteristics, added value and means of implementation. The next chapter will be focused on simulating scenarios based on real power data coming from devices specifications.

CHAPTER 5

SIMULATION RESULTS AND DISCUSSION

This chapter aims to present the simulation results, along with the arguments that describe the adopted case studies, input parameters, running, output data and a comparative analysis. Firstly, the optimization outcomes in SMDC perspective and their respective scenarios are presented. Thereafter, the same steps are applied with focus on DSO point of view, allowing to simulate the impact of specific actions present in this type of relationship, such as a contract with different amounts of SMDC and with different DR cycles.

5.1 SMALL AND MEDIUM DATA CENTERS OPTIMIZATION RESULTS

The optimization process described in this section contemplates different goals and consequently different scenarios. Thus, the following case studies will be established:

- In the first case, the objective is to find the optimal tariff windows in order that SMDC can individually decide which are the best hours to reduce power when participating in a DR event, as well as considering a successive power increase in a RE situation. Thus, knowing that the RE should occur immediately after the DR, it is necessary to manage the flexible loads in order that the DR occurs at the very end of a high price window for the RE to start at a low price window.
- In a different context, a DSO action to induce the prices in specific time windows in order to stimulate a SMDC participation in DR events is presented in the second case also aiming to present different scenarios, with diverse power reduction possibilities, where it is possible to notice the DR and RE occurrence by one small and one medium data center.
- The third case presents respectively the DR and RE impact on the daily load diagram scope along with a subsequent cost analysis involving different amounts of SMDC, i.e., while in the second case the analysis happens individually in a small and a medium data center, in this specific case it occurs in aggregate form for different quantities of SMDC.
- The impact on costs and savings prospects based on an incentive-based approach is the objective of the fourth case study.

It is important to emphasize that during these case studies the different strategies directed to flexible loads, such as ICT workload, set point adjustment in cooling devices and batteries discharge in UPS equipment will be demonstrated. The power values utilized in all case studies have different source reference, but are real-based data and applied according to the data center size. For the ICT workload the values are based on Table 2.6 and Table 2.8, as well as in the size range profile proposed by Salom *et al.* (2017). The set point adjustment implemented on CRAC devices as cooling strategies has its premises in the ASHRAE thermal guidelines from Table 2.7 and the specification of the used fan device is presented in Table 5.1.

Table 5.1 – Cooling fan specification (Vertiv 2017)

Speed	MODEL FH 600C, 72F°/50% RH, 45E WT, 10° WATER TD 0.3" EXTERNAL STATIC PRESSURE	NET SENSIBLE COOLING CAPACITY (kJ)	MOTOR kW
100%	Centrifugal blowers with VSD	299	11.0
	EC motorized impeller under floor	312	7.6
90%	Centrifugal blowers with VSD	274	8.0
	EC motorized impeller under floor	282	5.5
80%	Centrifugal blowers with VSD	246	5.6
	EC motorized impeller under floor	252	3.9
70%	Centrifugal blowers with VSD	203	3.8
	EC motorized impeller under floor	205	2.6
60%	Centrifugal blowers with VSD	206	2.5
	EC motorized impeller under floor	208	1.7

Regarding the UPS equipment utilized, the specification, arranged by diverse power levels, is presented in Table 5.2.

Table 5.2 – UPS specification (Delta 2018)

Manufacture Model	Delta HPH-20K	Delta HPH-30K	Delta HPH-40K	Delta HPH-60K	Delta HPH-70K	Delta HPH-80K
Power	20 kW	30 kW	40 kW	60 kW	70 kW	80 kW
Application (Data Center Profile)	Small	Small	Small	Medium	Medium	Medium
Input/Output Voltage	380/220 Vac, 400/230 Vac, 415/240 Vac					
Input Frequency	40~70 Hz					
Output Frequency	50/60 Hz +/- 0,05 Hz					
Power Factor	> 0,99					
Efficiency	Up To 96%					

Through these equipment' and the mentioned references, Table 5.3 shows how it is possible to obtain data to feed the input parameters that compose the equations formulated in the section of mathematical modeling and therefore obtain results. The complete table can be seen in Annex 2.

For the sake of simplicity, the data presented in this table also consider the five different power reduction profiles and their respective variations from the threshold of 105 kW for small data centers and 550 kW for medium data centers, as already mentioned in the contractual context, i.e., based on the contractually agreed objective, data centers can accomplish the power reduction, being slightly above, slightly below, below, or quite below.

For cooling data, the input parameters are:

- Power consumption of ICT load.
- Coefficient of performance.
- Supplied temperature.
- Adjustment temperature.
- Fans power.

For UPS data, the input parameters are:

- Discharge energy.
- Capacity in the previous stage.
- Total capacity.
- Efficiency of the UPS, including the battery.
- DR time.

Table 5.3 – Cooling and UPS data per profile

Small Data Center												
Cooling							UPS					
Unit	P_{ict}	CoP	T_{sup}	T_{adj}	P_{fan}	Profile	Unit	E_{dis^j}	c^{t-1}	c_{total}	η	j
2	190	2	10	25	8	accomplished	2	35.5	40	40	0.9	0.3
2	60	2	10	25	8	slightly above	2	32.8	40	40	0.9	0.3
2	96	2	12	27	11	slightly below	2	25.8	30	30	0.9	0.3
2	99	2	15	30	1.7	below	2	18.5	20	20	0.9	0.3
2	39	2	15	30	1.7	quite below	2	18.5	20	20	0.9	0.3
Medium Data Center												
Cooling							UPS					
Unit	P_{ict}	CoP	T_{sup}	T_{adj}	P_{fan}	Profile	Unit	E_{dis^j}	c^{t-1}	c_{total}	η	j
6	440	2	10	25	8	accomplished	2	50.3	80	80	0.9	0.3
6	880	2	12	27	11	slightly above	2	57.7	80	80	0.9	0.3
6	376	2	12	27	11	slightly below	2	47.7	70	70	0.9	0.3
6	249	2	15	30	1.7	below	2	40.6	60	60	0.9	0.3
6	125	2	15	30	1.7	quite below	2	51.2	60	60	0.9	0.3

5.1.1 Case Study 1

Based on an hourly electricity price fluctuation, whereby the tariffs are dynamic in a 1 x 24 vector with the prices of Figure 4.3, the optimization process chooses higher prices to decrease power in a DR event and lower prices to increase power in a RE. As this time window has 24 positions, representing the number of hours in a day, the 12 highest prices are assigned to periods with potential for DR, while the others 12 will contemplate a RE situation, as demonstrated in Figure 5.1.

The main objective of this case is to demonstrate how a small or medium data center can differently manage their flexible loads, graphically showing how they act in face of the different prices.

For threshold effects, the maximum power values adopted in the DR reduction process are: 105 kW for small and 550 kW for medium data centers. It is also possible to notice that among the selected flexible loads to be reduced in this case, the greatest predominance occurs in ICT workload with 40 kW in the small data center and 220 kW in the medium data center. Whereas for cooling the power is 35 kW in the small and 180 kW in the medium and for UPS the reduction was 30 kW in the small and 150 kW in the medium data center. However, the small difference among the loads prevails.

The same situation can be seen in Figure 5.2, where the predominance is clear in the cooling devices set point adjustment, in which the achieved decreased power was 42 kW in the small and 300 kW in the medium profile. The ICT workload reached 28 kW in the small and 150 kW in the medium. The UPS battery was used to discharge in this case 35 kW in the small and 100 kW in the medium data center. An interesting aspect to be observed in this case is the difference between the loads of the small and medium data center. While in the former the difference is small, in the latter it is possible to perceive a greater discrepancy in relation to the notoriety of the reduced load by the cooling strategy.

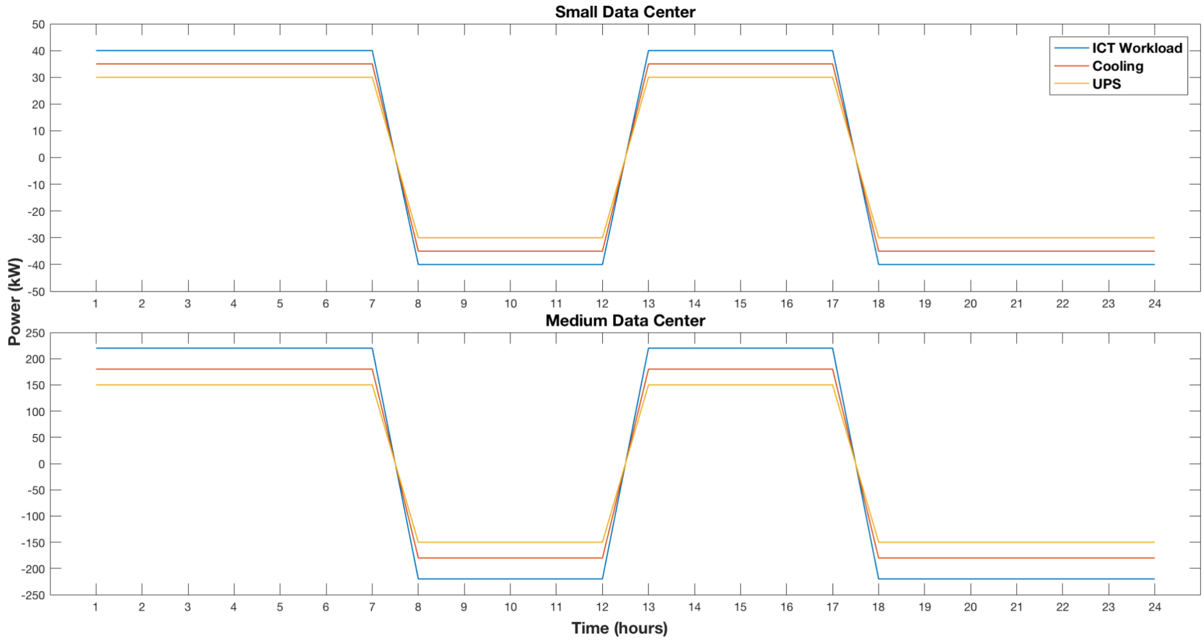


Figure 5.1 – ICT workload predominance in DR and RE tariff window

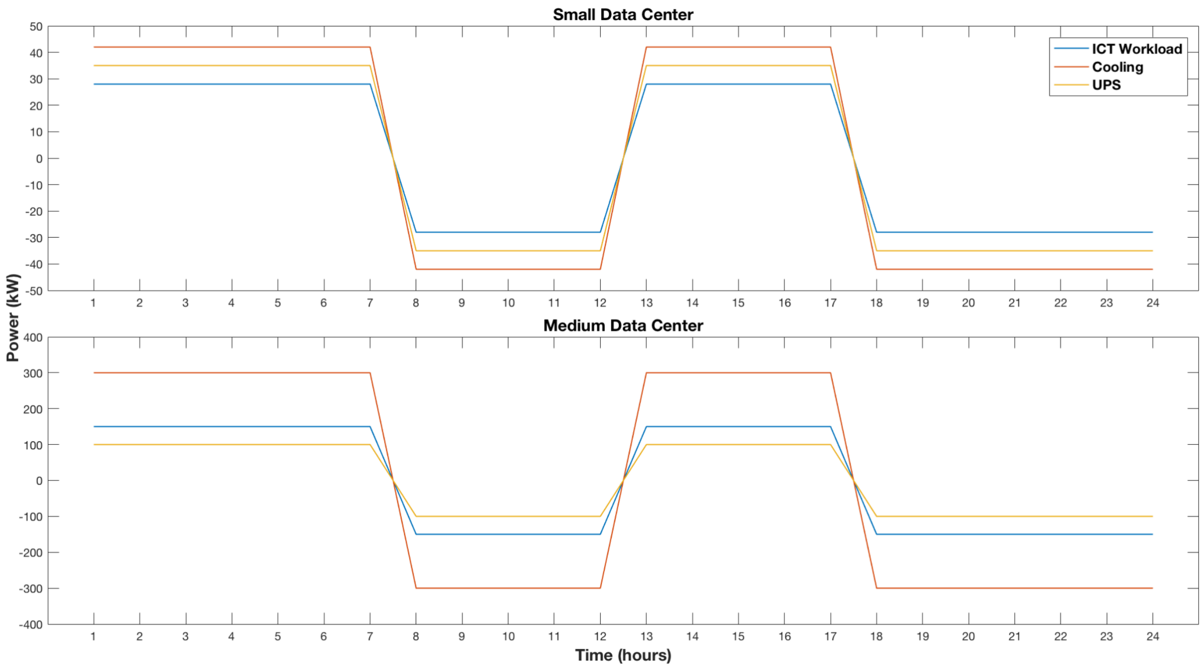


Figure 5.2 – Cooling predominance in DR and RE tariff window

In this last case, the dominant strategy is the discharge of battery in UPS equipment shown in Figure 5.3. The respective power reduction was 45 kW in the small and 230 kW in the medium data center. On second place, in terms of higher power is the cooling solution, in which the reduced power was 25 kW in the small and 200 kW in the medium profile. The ICT workload achieved 35 kW in the small and 120 kW in the medium data center on the third position. Regarding the uniformity between the difference of loads, while in the small data center is clear, in the medium data center the discrepancy can be realized in the ICT workload, since the difference noticed in the UPS and cooling solutions is uniform.

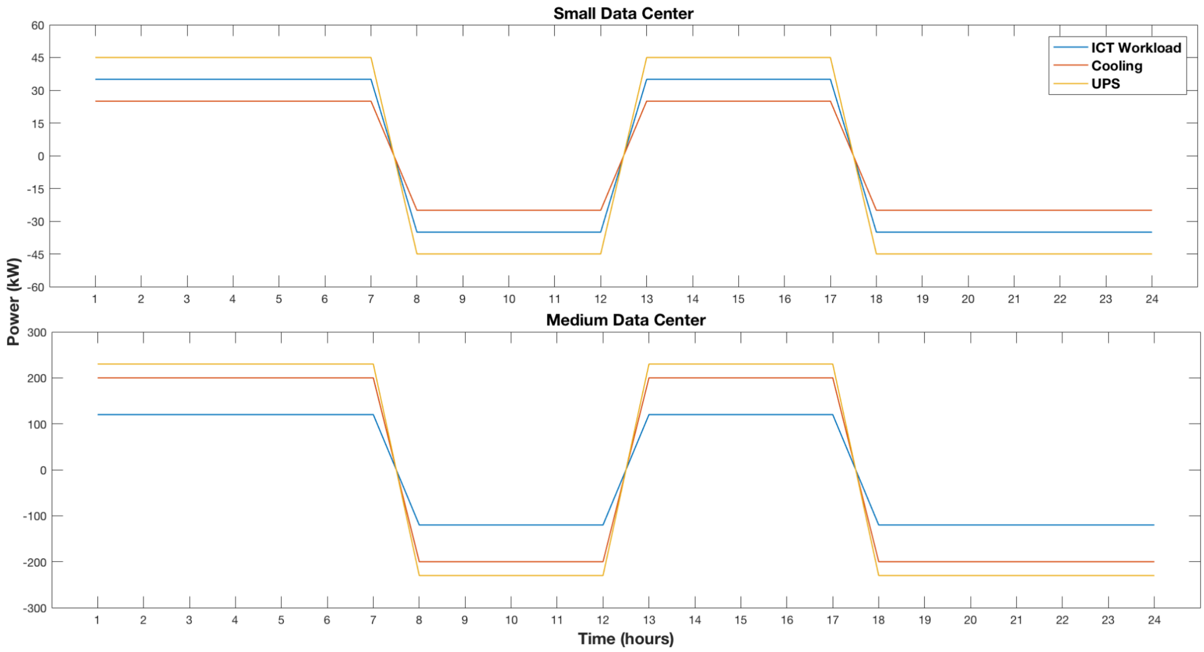


Figure 5.3 – UPS predominance in DR and RE tariff window

5.1.2 Case Study 2

Regarding this price scenario, a more detailed hourly electricity price fluctuation with 20 minute intervals per hour is used and the second study case is structured in a 1 x 72 vector with the prices of Figure 4.5. As can be noticed, the stimulus to a power reduction occurs at 11:00 h, where the energy price rises to 0.20 €/kWh and oppositely, the power increase stimulus happens between 12:00 and 14:00 h, whereby the energy cost decreases to 0.114 €/kWh. The DR event starts at 11:40 h and ends at 12:00 h. Immediately, the RE gradually begins at 12:00 h and ends at 14:00 h in order to avoid a peak situation during the power rebound which was decreased previously in the DR action. Thus, this pattern will be adopted in this case study. The phenomenon of the DR event happening at 11:40 am is due to this period being the exactly prior to the price decline for the rebound. If it happened at 11am, for example, the rebound would happen at 11:20 and 11:40, where prices would still be high.

In Figure 5.4 an accomplished scenario is represented, i.e., the DSO had a reduction expectation of 105 kW from a small data center and 550 kW from a medium one and this reduction has happened exactly as expected. As can be noticed, during the DR event the small data center decreases a power of 40 kW from ICT workload, 35 kW from cooling and 30 kW from UPS. Nonetheless, the highest decreased power with the medium data center occurs through UPS devices with 200 kW, ICT slightly below with 190 kW and afterwards 160 kW by cooling equipment'. Moreover, it is possible to observe the RE smoothly happening in both data centers profiles, but with 2 periods of 20 kW, 17 kW, and 15 kW by ICT, cooling and UPS respectively in the small data center. This is due to the strategy, which gradually resumes power in a non-simultaneous way in order to avoid a peak power increase. For the medium one the RE occurs with the same period division, but with the power increasing 100 kW in UPS, 95 kW in ICT and 80 kW in cooling strategies. The predominant yellow line before and after de DR e RE actions is as a matter of fact the overlapping of strategies in a baseline normal operation circumstance, because the main objective of this optimization process is to just highlight the DR and RE in the induced time windows.

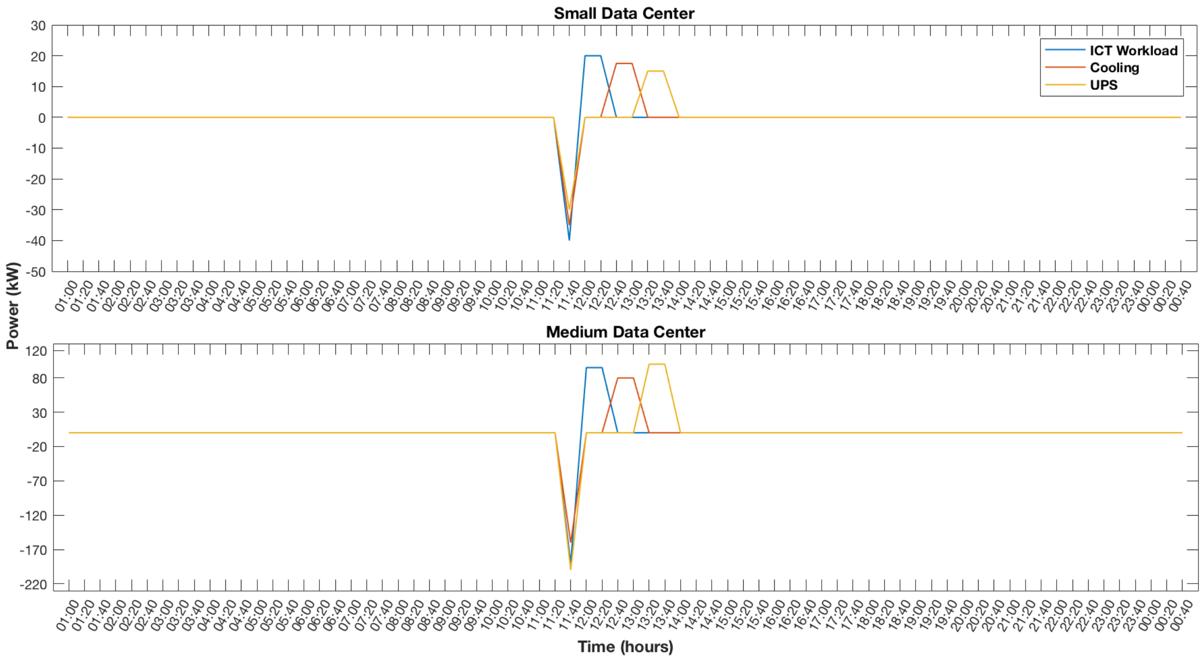


Figure 5.4 – SMDC accomplished DR

On the other hand, data centers can ensure a certain reduction and stay slightly above the desired value. Figure 5.5 shows this scenario, emphasizing a UPS power predominance in the small data center and a cooling power predominance in the medium one. The values of DR in the small data center are 59 kW for UPS, 27 kW for cooling and 23 kW for ICT, totalizing 109 kW. In the medium data center, the values are distributed by 286 kW for cooling, 176 kW for ICT and 100 kW for UPS, totalizing 562 kW. Looking at RE time windows is evident the heterogeneity among the powers of the loads.

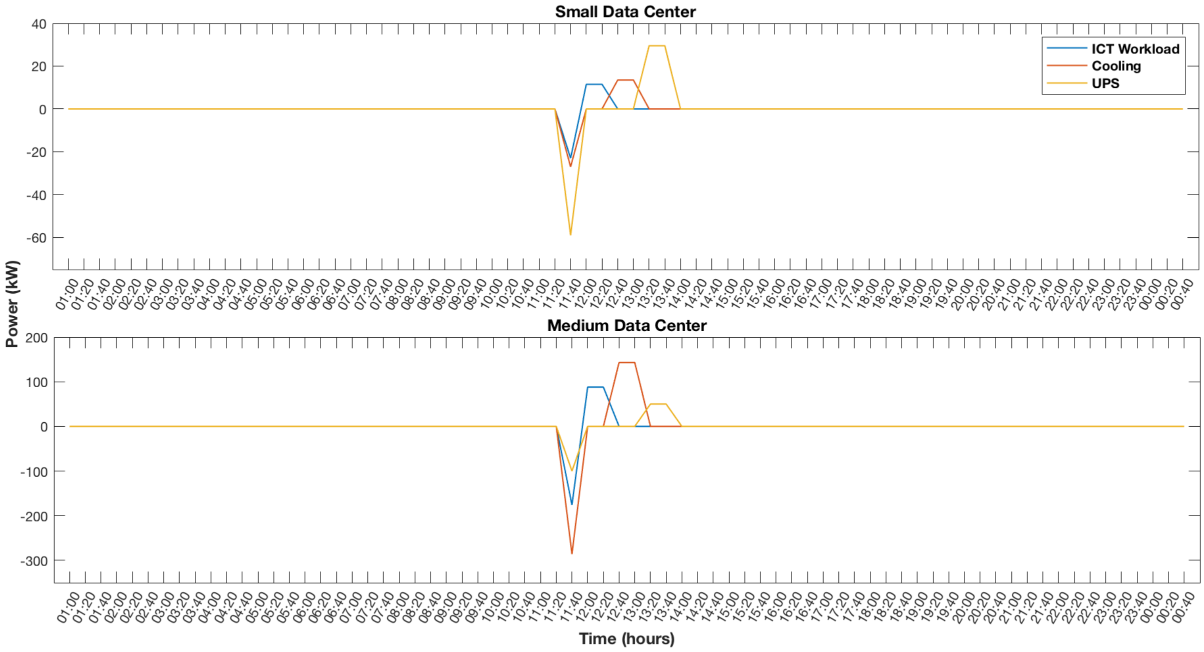


Figure 5.5 – SMDC slightly above DR

The opposite scenario might also happen, where a data center operator wants to reduce the pre-set value and stays slightly below of that goal. In Figure 5.6 the DR event happening in the small data center has a cooling predominance with 37 kW, afterwards 23 kW with UPS and 20 kW with ICT workload. In the medium data center, the ICT workload has the highest power with 219 kW, followed by cooling with 180 kW and UPS with 150 kW. The rebound pattern remains the same, with the power going up every period of 20 minutes over a 2-hour interval.

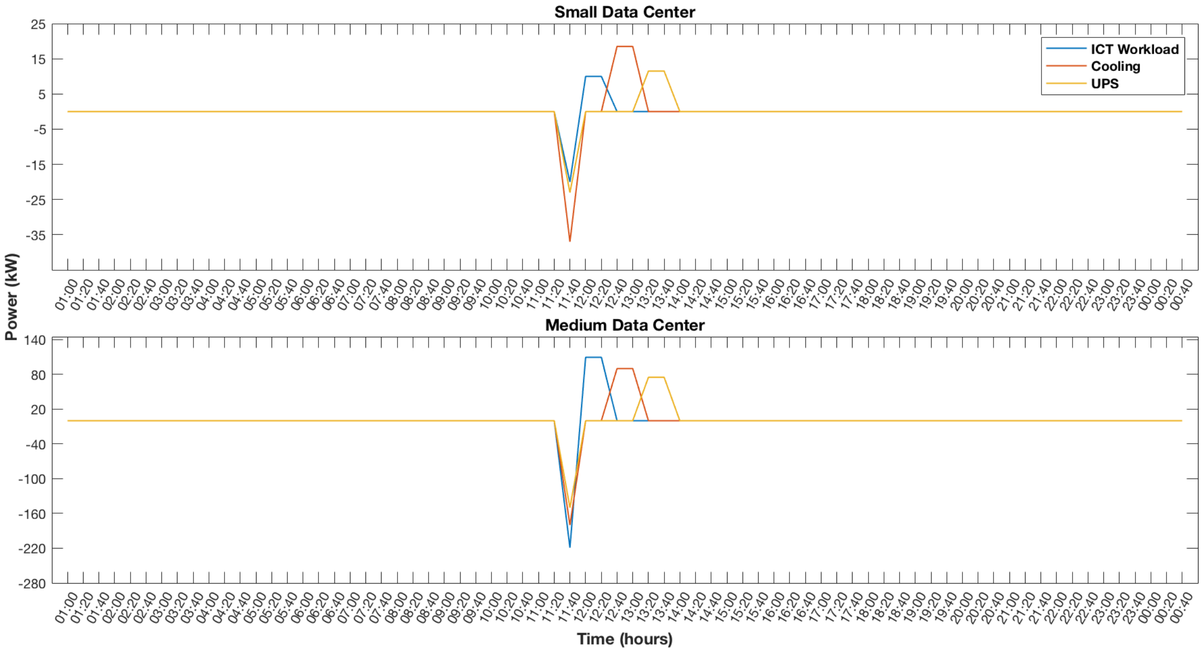


Figure 5.6 – SMDC slightly below DR

Another type of scenario is when the power to be reduced only reaches an intermediate value getting below of the intended, as it is the case depicted in Figure 5.7. The highest power in both data centers belongs to the UPS strategy in this DR event, nevertheless the main difference is in the fact that in the small data center the second place come from the ICT strategy and finally the cooling. In the medium data center is exactly the opposite, the second place belongs to the cooling reduction and the last one is from the ICT workload. The power values in the small data center are 31 kW, 23 kW and 10 kW for UPS, ICT and cooling, respectively. For the medium data center, the reached values are 130 kW, 90 kW and 80 kW for UPS, cooling and ICT, consecutively.

The worst case that can occur is if a data center committed to reducing a certain value and falls quite below, as presented in Figure 5.8. This might denote some unexpected problem, characterizing some failure in a determined process, or in a set of them, such as inability to reschedule a certain ICT load, intense external temperature increase, or a low availability for discharging UPS batteries.

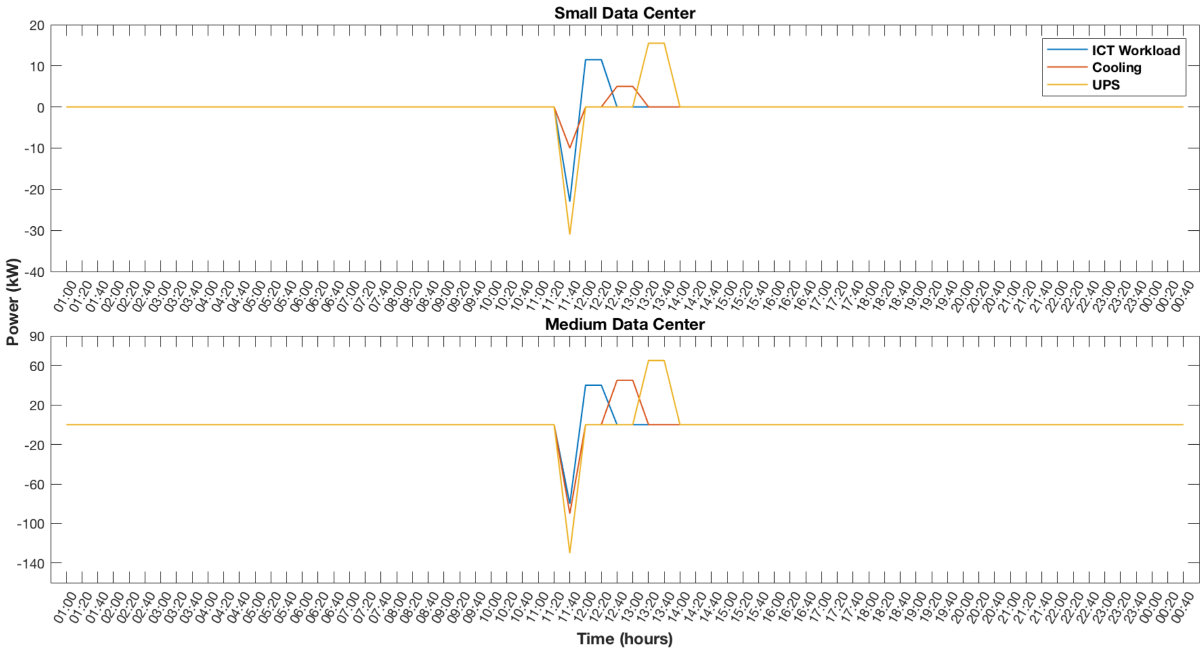


Figure 5.7 – SMDC bellow DR

It is clear in Figure 5.8 the predominance of the set point adjustment of cooling devices strategy in both data centers with exactly the same distribution regarding the other power loads, just with different values. In the small data center, the power reduction was 5 kW with cooling, 2 kW in UPS and 1 kW from ICT workload, whereas 37 kW with cooling, 29 kW in UPS and 20 kW from ICT workload were obtained in the medium data center.

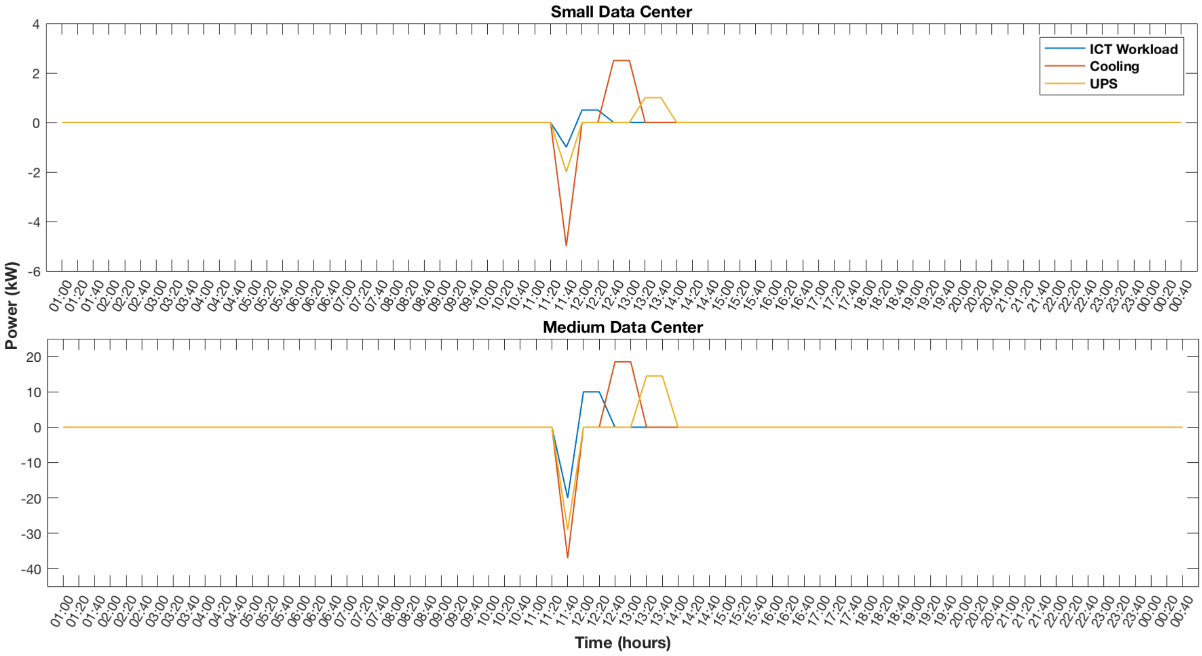


Figure 5.8 – SMDC quite bellow DR

Finally, differently from Figure 5.4 where the flexile loads are reduced without any criterion, the ideal scenario from SMDC point of view is that the intended power reduction is exactly accomplished and for this purpose, that the use of flexible loads follows the hierarchy proposed in Table 4.1 considering the critical mission state of a data center: firstly the UPS load, subsequently cooling devices and lastly the ICT workload, which is the core business of a data center. In this regard, Figure 5.9 shows both the small data centers and the medium one achieving 105 kW and 550 kW of power reduction respectively in this DR event and the strategies distribution being arranged as follows: 50 kW for UPS, 35 kW with cooling and 20 kW from ICT in the small data center and 220 kW for UPS, 180 kW with cooling and 150 kW from ICT in the medium data center. The RE smoothly occurs following the same process in a 2 hours' time window.

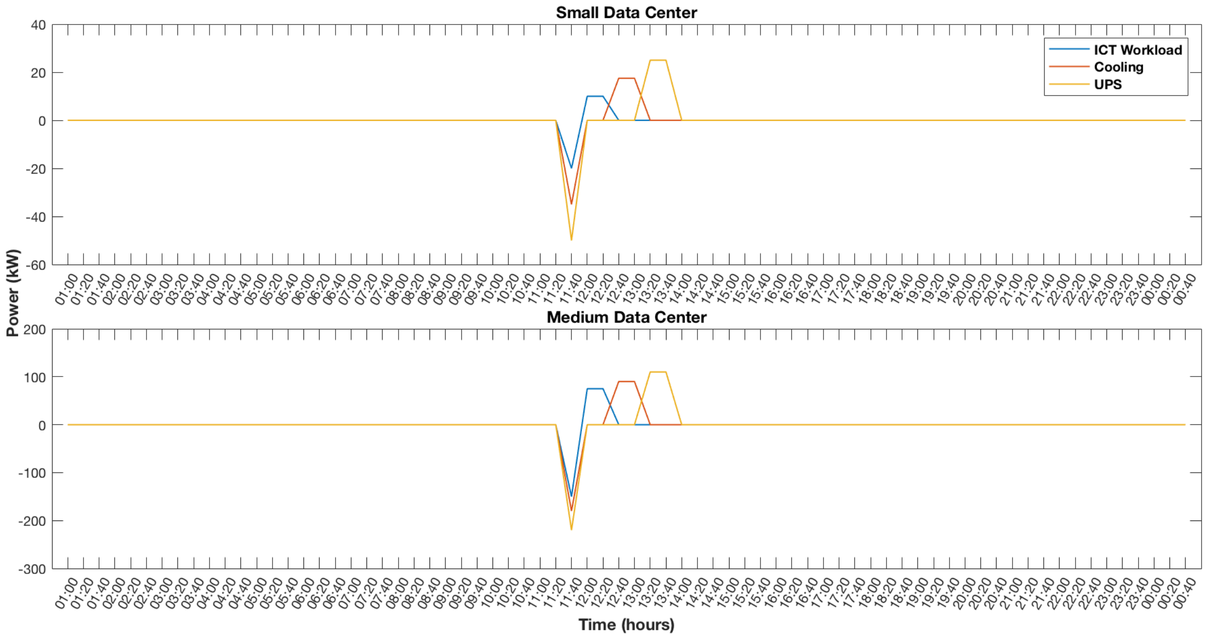


Figure 5.9 – SMDC ideal DR

5.1.3 Case Study 3

In this third case study the main goal is to observe how DR and RE happens taking into account the daily load diagram of SMDC, as shown in Figure 5.10. The adopted tariff strategy is the same as in the last case study, i.e., an hourly electricity price fluctuation with 20 minutes intervals, whereby the DR event starts at 11:40 h and ends at 12:00h and shortly after the RE process begins, finishing at 14:00 h. The others time periods are framed in a baseline power values.

In this context, the average energy price is 0.1402 €/kWh, but during DR the price is increased to 0.2 €/kWh. Thereafter, a cost analysis will be performed considering three different scenarios. The first one in equal conditions with 16 small and 16 medium data centers, the second one with a preponderance of 21 small and 15 medium and the third one with 10 small and a predominance of 17 medium data centers. It is important to point out that in this case it will be considered that all data centers that respond to the tariff stimulus at the same time will reduce the intended power by following the accomplished profile and jointly totalizing an approximate 10 MW of power reduction impact in this type of DR.

In the cost analysis described by Table 5.4 it is possible to notice the exact power decreased by small and medium data center, as well as by their summation during the DR event. As can be seen, the amount of reduced power is different considering the size of data center and the respective scenario; and the determinant factor for that is the diversity of SMDC combination established in each situation.

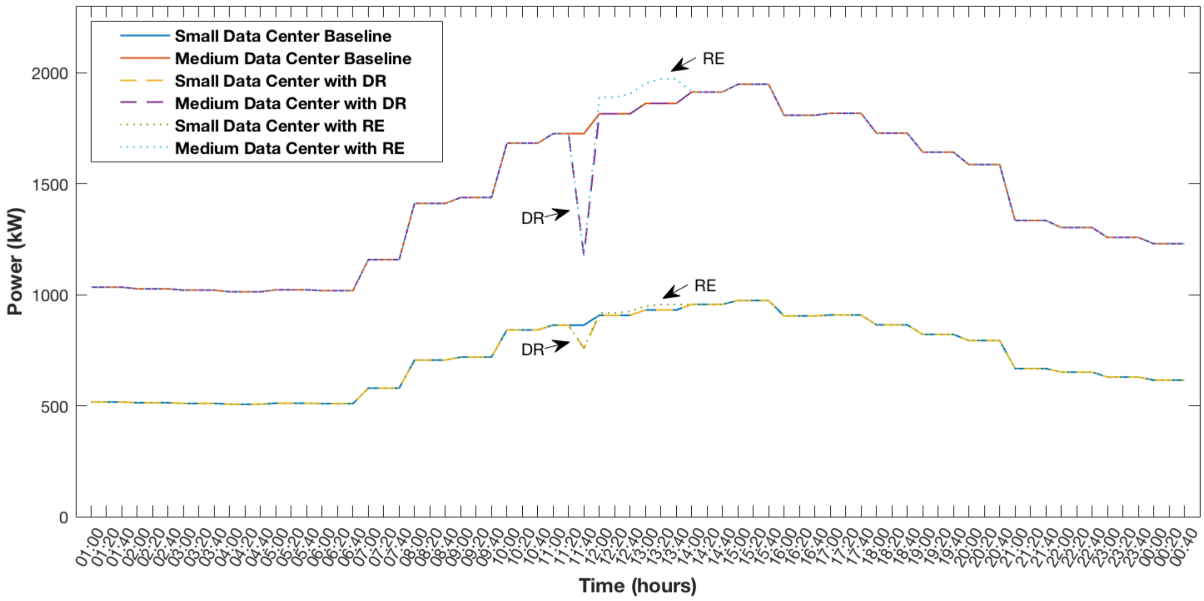


Figure 5.10 – DR and RE in daily load diagram

Table 5.4 – DR cost analysis with different scenarios

Analysis Parameters	1st Scenario	2nd Scenario	3rd Scenario
Average price of energy	0.14 €/kWh	0.14 €/kWh	0.14 €/kWh
Average price of energy during DR	0.2 €/kWh	0.2 €/kWh	0.2 €/kWh
Small Data Centers DR participation	16	21	10
Medium Data Centers DR participation	16	15	17
Decreased power by Small Data Center	1,680 kW	2,205 kW	1,050 kW
Decreased power by Medium Data Center	8,800 kW	8,250 kW	9,350 kW
Accumulated decreased power during the DR event	10,480 kW	10,455 kW	10,400 kW
DR cost considering RE	898 €	896 €	891 €
Operational cost by Small Data Center	7,386 €	7,386 €	7,386 €
Operational cost by Medium Data Center	14,770 €	14,770 €	14,770 €
Operational cost with all Small Data Centers	118,200 €	155,100 €	73,860 €
Operational cost with all Medium Data Centers	236,400 €	221,600 €	251,100 €
Total operational cost	354,500 €	376,700 €	325,000 €
Operational cost with DR by Small Data Center	7,370 €	7,370 €	7,370 €
Operational cost with DR by Medium Data Center	14,670 €	14,670 €	14,670 €
Operational cost with DR with all Small Data Centers	117,900 €	154,800 €	73,700 €
Operational cost with DR with all Medium Data Centers	234,700 €	220,100 €	249,400 €
Total operational cost with DR	352,600 €	374,900 €	323,100 €

For comparative purposes, the total reduced power is approximately 10 MW with slight variations among the scenarios. In this context, the DR cost takes into account the RE process, i.e., the achieved savings in the reduction process is directly affected by the power increase after the DR period. What will determine the amount of the savings is exactly the tariff stimulus allocated for each time window in the respective actions. Thus, as the power reduction variation is derisory, the difference in the DR cost by scenario will also be.

Moreover, this analysis also includes the data center operational cost using the hourly price fluctuation, in order to compare it with the operational cost influenced by the price stimulus induced by DSO. Both calculations are addressed individually and encompassing the diversity of data centers in each scenario.

Verifying the savings potential in a DR event, the small data center reached 15.91 €, which represents 0.22% of its daily operational cost. On the other side, the medium one achieved 99.82 €, i.e., 0.68% or its daily expense. Regarding the three scenarios, the second one with predominance of small data centers was the most profitable given their total power, followed by the first one with equivalence of quantity and finally the third one, with preponderance of medium data centers.

5.1.4 Case Study 4

The fourth case study also considers three periods of 20 minutes in each hour, however there are four different tariff periods: peak, half-peak, normal off-peak and super off-peak. This type of tariff is used in a situation in which a given data center enter into a contract with a DSO in order to receive a certain financial incentive to reduce power in a DR event receiving a direct incentive for the power reduction.

Specifically, for this scenario presented in Figure 5.11, the diagram shows the DR starting at 14:40 h and ending at 15:00 h, whereas the RE occurs from that point onto 17:00 h. The others time periods are framed in a baseline power values. The DR event occurrence in this time window happened because the DSO requested, through the contract, a power reduction in that period, paying an incentive for that purpose.

The DR cost analysis in an incentive-based contract depicted in Table 5.5 presents an energy average price of 0.1496 €/kWh (considering the entire period of 24h). Regarding the DR and RE occurrence, the energy price is 0.1400 €/kWh in a half-peak time window. In this case the operational cost with tariff price is calculated by multiplying the data center power by the energy price in each hour taking into account the four different time windows and prices, consequently. The operational cost is presented and arranged by data center profile, where it is possible to realize that the medium data center is exactly double in terms of power in comparison with the small one, and therefore both values are doubled.

The incentives are estimated taking into consideration values that provide a relevant stimulus to motivate energy-intensive consumers such as SMDC. The adopted values are 1 € for small and 1.5 € for medium data centers. Thereby, the incentive application will be granted in accordance with the amount of reduced power, namely 105 kW and 550 kW for the small and medium data center in an accomplished profile respectively. Therefore, the cost difference between them is due to the greater reduction potential found in mid-sized data centers.

In terms of savings, the small data center achieved 110 € by event, or in other words, 1.33% of its daily operational cost. Regarding the medium one, it received from the DSO a rebate of 850 €, i.e., 5.15% of its daily energy expense.

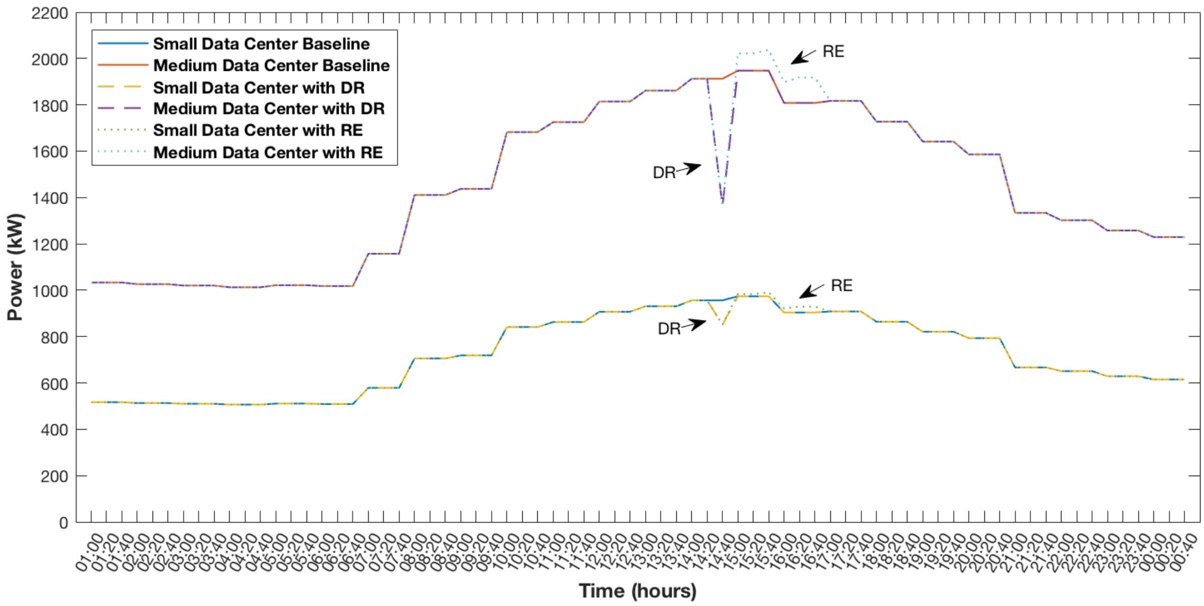


Figure 5.11 – DR and RE in an incentive-based daily load diagram

Table 5.5 – DR cost analysis in an incentive-based contract

Analysis Parameters	Costs
Average price	0.1496 €/kWh
Average price during the DR and RE period	0.14 €/kWh
Operational cost with tariff price by Small Data Center	8,220 €
Operational cost with tariff price by Medium Data Center	16,450 €
Small Data Center incentive	1 €/kW
Medium Data Center incentive	1.5 €/kW
Operational cost with DR and RE with incentive by Small Data Center	8,110 €
Operational cost with DR and RE with incentive by Medium Data Center	15,600 €

5.2 Distribution System Operator Optimization Results

The results presented in the optimization process of this section are intended to demonstrate how a particular contractual policy oriented towards DR can be advantageous in order to include energy consumers, previously excluded from the DR market because their core business has a mission-critical characteristic, such as the SMDC. Accordingly, the developed algorithm allows to optimize different contractual scenarios in egalitarian situations or with predominance of small, or medium data centers in the context of DR. Another advantage is to enable the simulation of a contract with diversified time cycles, such as day or month, evaluating the potential reactions of this type of consumer in response to a contractual agreement that provides incentives for the successfully implementation of DR, but also imposes penalties when the contractual commitments are not achieved.

Hence, three case studies based on different quantities of data centers will be addressed with the aim of standardizing and simplifying the approach. The first case will include an egalitarian contract with 16 small and 16 medium data centers. The second one with a predominance of small data centers with 21 participants and 15 medium ones. The third case study has 10 small and a preponderance of medium data centers, with 17 participants.

For each case study it will be possible to predict a monthly contract considering 22 working days, as well as to analyze a single DR event per day, to financially analyze different scenarios, with prevalence of incentives, penalties, or an unchanged state. In this context, it was considered that small data centers signed the contract with a DR power objective of 105 kW and medium data centers with 550 kW.

Some characteristics of the optimization process are common to all case studies. Overall, at each demand response event the DSO might request a reduction corresponding to the sum of the power of all data centers participating in the contract. However, knowing the possibility of declination, the DSO will consider as its objective an adhesion rate of 80% of the total power of the contract when sending the demand response signal. Hereinafter, as already mentioned, it is possible to set as input the percentage of data centers that will accept the DR signal. Hence, this optimization process considered an acceptance rate between 60 and 100% as default values, with the final value being randomly set.

Specifically, in the scenario where only one demand response event will be addressed, it will occur always during one hour, simulating four time windows of 15 minutes. Nevertheless, in the 22-days scenario, the DR time window will have other terms: in 50% of the days the signal broadcasted will have 4 temporary periods, occurring in 1h and in the other 50% there will be intercalated periods of 45, 30 and 15 minutes with the same 80% of adhesion rate by the participants SMDC.

Specifically, for a monthly contract with 22 working days, in the first case study the set of time windows will be:

- Four periods of 15 minutes with 2.09 MW of power during one hour.
- Three periods with 2.79 MW in 45 minutes.
- Two periods of 4.17 MW in 30 minutes.
- One period of 8.38 MW in 15 minutes, which corresponds to the DSO objective, i.e., 80% of the total power capacity of the contract considering the adhesion rate.

In the second case study:

- Four periods of 15 minutes with 2.09 MW of power.
- Three periods with 2.78 MW, two periods of 4.18 MW.
- One period of 8.36 MW.

In the third case study:

- Four periods of 15 minutes with 2.08 MW of power,
- Three periods with 2.77 MW,
- Two periods of 4.16 MW.
- One period of 8.32 MW.

Finally, it should be noted that the present simulation process was conducted in order to consider all the potential results that are possible to achieve with the algorithm. Thus, for each of the mentioned quantitative case studies, a simulation of 1 and 22 days respectively was performed. For each of these simulations, 3 scenarios of predominance with incentive, penalty and unchanged profile were generated. Finally, each of these scenarios provides 18 different charts, totaling 324 charts. Thus, in order to simplify this results presentation, the first case study will address the incentive scenario, the second one the unchanged case and the third case will be responsible for presenting the penalty scenario. However, the other important charts related to the case studies and scenarios that were not covered in this section are presented in Annex 3.

5.2.1 Case Study 1

Analyzing a single event in a specific perspective, the DSO broadcasts the DR signal to the 32 participants, namely 16 small and 16 medium data centers, totalizing 8.38 MW of power, with an expected constant reduction of 2.09 MW during four periods of 15 minutes.

As described in Table 5.6, in this scenario all medium-sized data centers present in the contract randomly accepted the signal, while 93.75% of the small ones accepted to participate in a total and non-imposed adhesion rate of 98%. Thereby, in order to ensure the power reduction required by the DSO, all data centers that accepted to participate in the event were randomly sorted in a list and selected immediately afterwards based on such list. For this reduction process, 16 calls were made to small data centers with 1 rejection and 14 calls to medium data centers without any rejection.

It is possible to note the criterion of fair choice among data centers through Figure 5.12, where the ID’s from one to 16 are small data centers and from 17 to 32 are medium data centers. However, in a single event it is not possible to have a complete overview of this algorithm feature, but it can still perceive that even 31 data centers accepting to participate in the DR event, two medium data centers (17 and 18) were not selected by the DSO algorithm for power reduction, in order to maintain the balance between small and medium data centers.

Table 5.6 – Acceptance and denial statistics for one DR event in case study 1

Parameters	Statistics
DR Event Acceptance Percentage (%)	98
DR Event Denial Percentage (%)	2
Small Data Center Acceptance Percentage (%)	93.75
Small Data Center Accepted Calls	15
Small Data Center Denial Percentage (%)	6.25
Small Data Center Rejected Calls	1
Medium Data Center Acceptance Percentage (%)	100
Medium Data Center Accepted Calls	14
Medium Data Center Denial Percentage (%)	0
Medium Data Center Rejected Calls	0

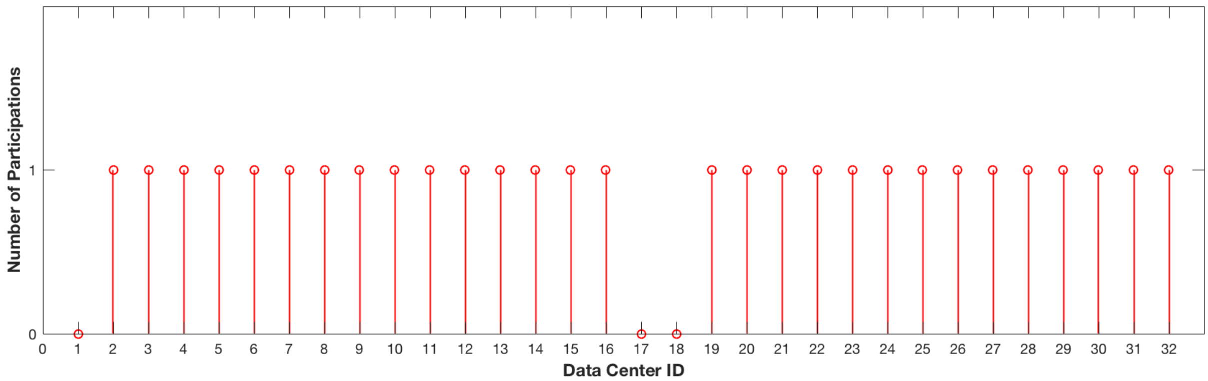


Figure 5.12 – Fair choice criterion for one DR event in case study 1

Analyzing the power reduction potential in Figure 5.13, on one hand a few data centers do not participate of the DR event and obviously do not reduce any amount of power. On the other hand, several others reach a reduction of 100% of the contractual power reduction target and even exceed. Other data centers have an intermediate reduction percentage of 20%, while others have a minimum reduction of 1%.

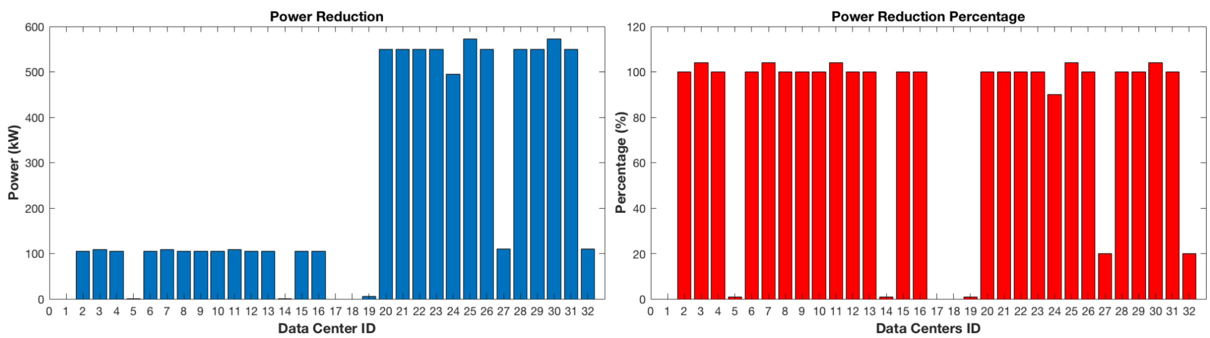


Figure 5.13 – SMDC power reduction potential for one DR event in case study 1

Specifically, as shown in Figure 5.14, the maximum power reduced in this event is 572 kW, the average 238.8 kW and the minimum 1.05 kW, whereas the total power reduced considering all participations is 7.64 MW, i.e., 91.16% of the required power reduction. What justifies such a percentage is that although the algorithm might call other data centers present on the DR event acceptance list, there is no guarantee that those called will reduce the agreed power. In this case what happened was exactly that, i.e., the DSO called other data centers, but these did not meet the reduction percentage previously established. Concerning the contractual premises, 72% of the data centers achieved their objective.

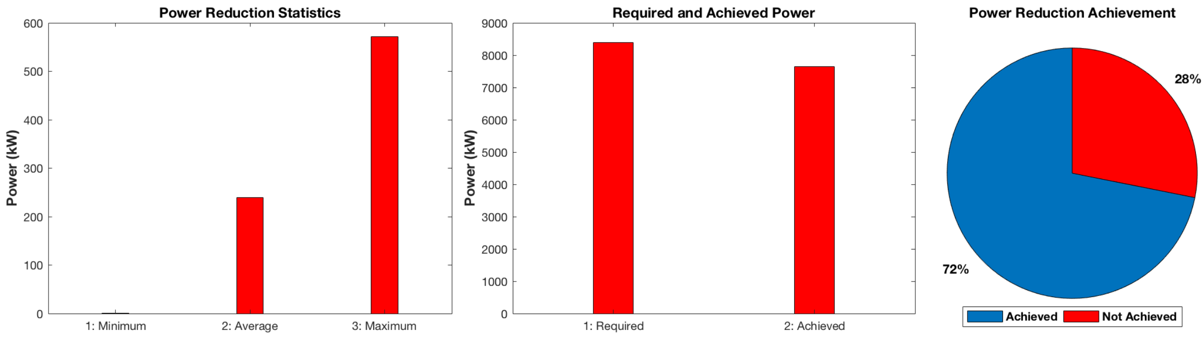


Figure 5.14 – DR power reduction outcomes for one DR event in case study 1

At the end of the DR event and the conclusion of the contractual cycle, it is possible to extract information that determines how many data centers received incentive, how many were penalized and how many were left unchanged, as can be seen in Figure 5.15. As in this scenario the predominance is of incentive, 75% of the data centers received this financial payment, or credit, of which 13 are of small profile and 11 of medium one. Of those that remained unchanged, one is small and four are medium and of those who will pay penalty, two are small data centers and only one is medium.

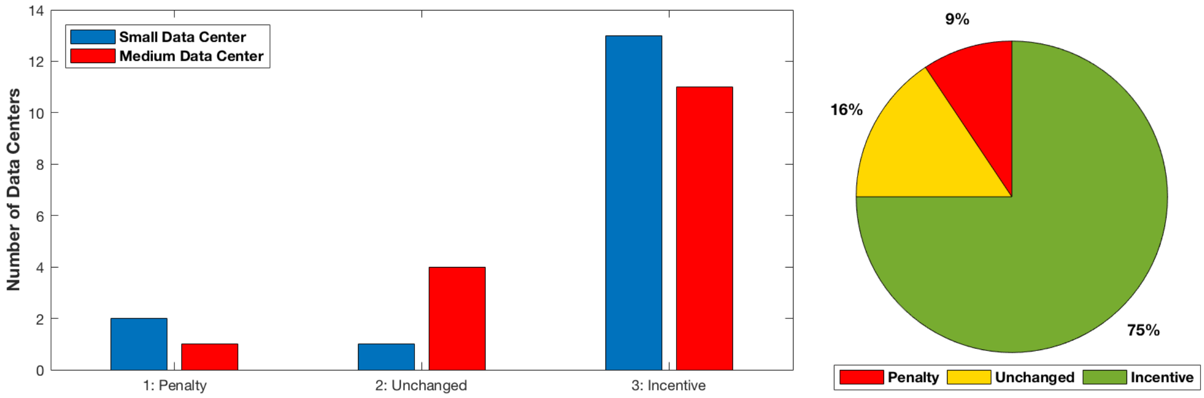


Figure 5.15 – SMDC financial profile for one DR event in case study 1

Therefore, the highest value received from the DSO to a small data center is 109.2 € and to a medium one is 847 € for having reduced more than 80% of the agreed power. Five data centers will not receive, or pay any amount, for having reduced between 20 and 79%. Finally, the highest amount to be paid to the DSO will be 3.94 € to a small and 30.57 € to a medium data center, for having reduced less than 20%, as presented in Figure 5.16.

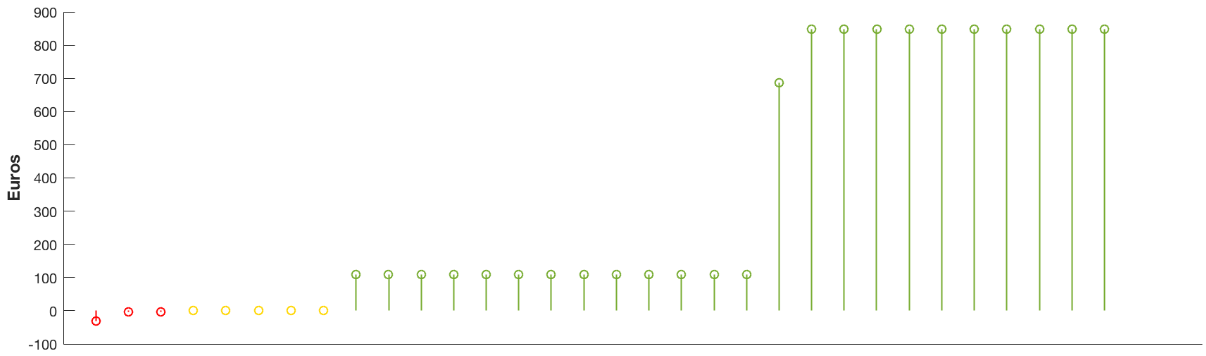


Figure 5.16 – SMDC financial index for one DR event in case study 1

Going toward a more comprehensive analysis, looking into a realistic scenario, with DR occurring in 22 workdays, it is understood that the DSO will perform one DR event per workday and will send the DR signal to the same number of data centers covered in this case study. Thus, in Figure 5.17 the acceptance and denial percentages are characterized by small and medium data centers. It is possible to note in this cycle that small data centers denied DR events more frequently than the medium ones and the highest acceptance rate was 99% and the lowest 62%.

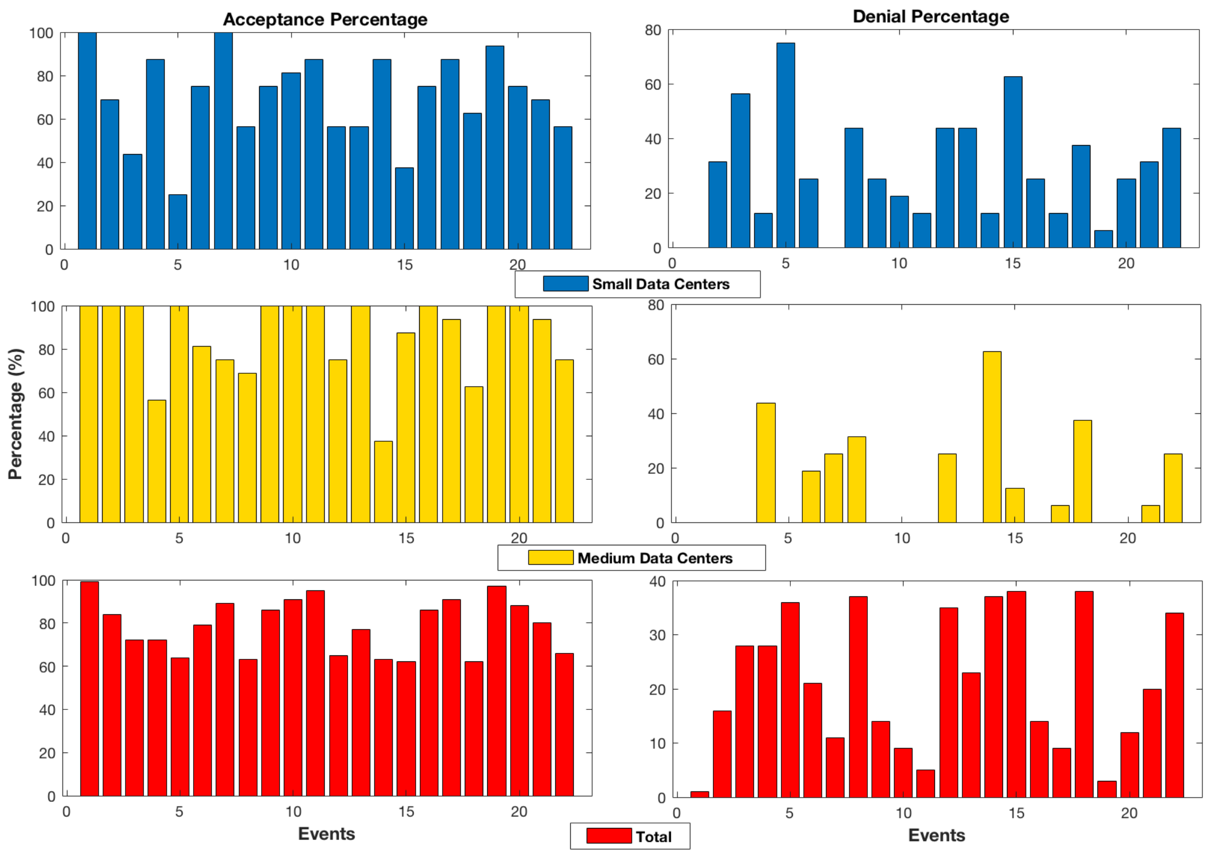


Figure 5.17 – Acceptance and denial statistics for 22-DR events in case study 1

On the other side, in Figure 5.18 all calls by event are specified and can be visualized in terms of acceptance, or rejection with an aggregation defined by size, where the number of calls made to small data centers was 249, to medium ones 255 and the total number of rejections was 103 in small data centers and 47 in medium ones.

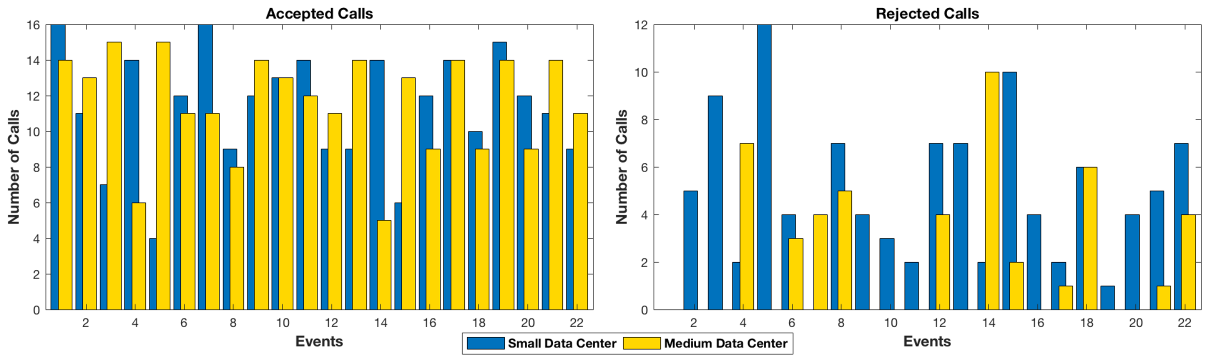


Figure 5.18 – Accepted and rejected calls for 22-DR events in case study 1

Another substantial difference that can be perceived when analyzing a monthly cycle is the functioning of the fairness criterion, as shown in Figure 5.19, where the ID's from one to 16 are small data centers and from 17 to 32 are medium data centers. As there are several events occurring, it is evident a nearly flat feature of the chart, with slight variations of 15 and 16 participations.

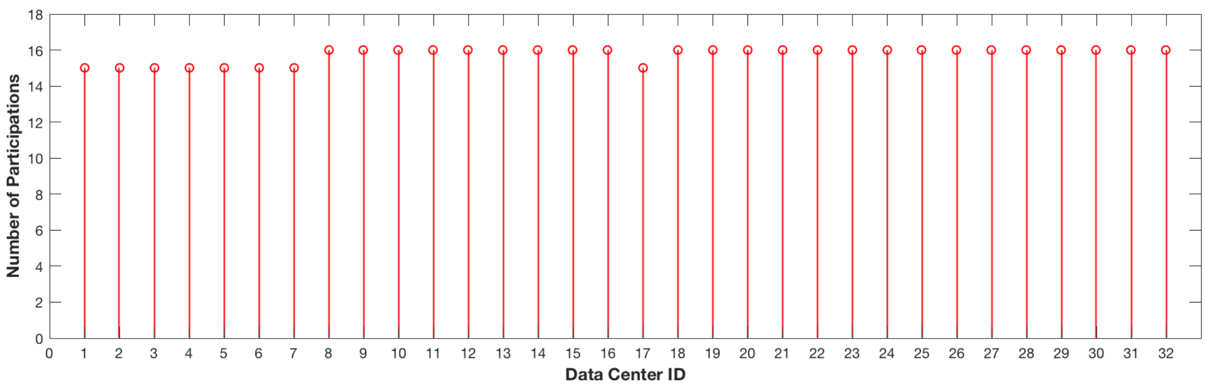


Figure 5.19 – Fair choice criterion for 22-DR events in case study 1

Observing the SMDC power reduction in Figure 5.20, there are fluctuations in the range of 48.37% and 89.87% per SMDC, proving that data centers are not always able to reduce the totality of the power agreed in contract.

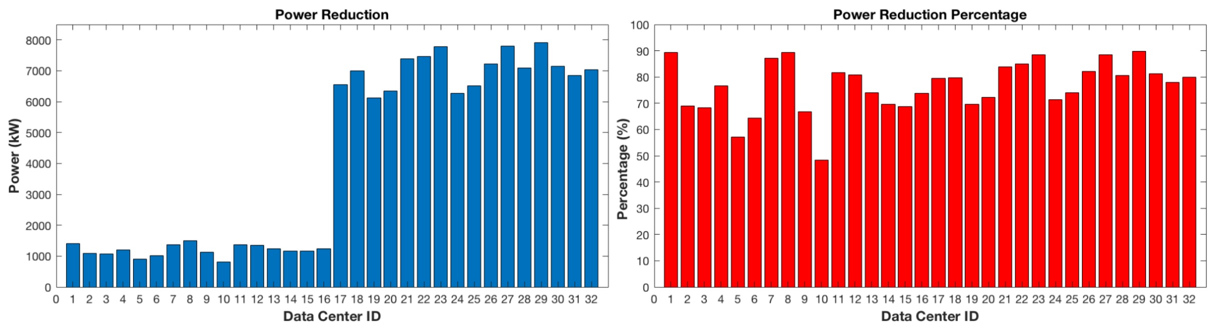


Figure 5.20 – SMDC power reduction for 22-DR events in case study 1

Figure 5.21, on the other hand, revels that no data center has been able to reach 100% of the maximum power agreed in contract. Analyzing all 22 DR events, the maximum power reduced was 7,909 kW, the average 4,112 kW and the minimum 812.7 kW. The total accumulated power required by the DSO was 184.33 MW and the accumulated reduced power by SMDC was 131.57 MW, i.e., a DR cycle with 71.38% of power reduction.

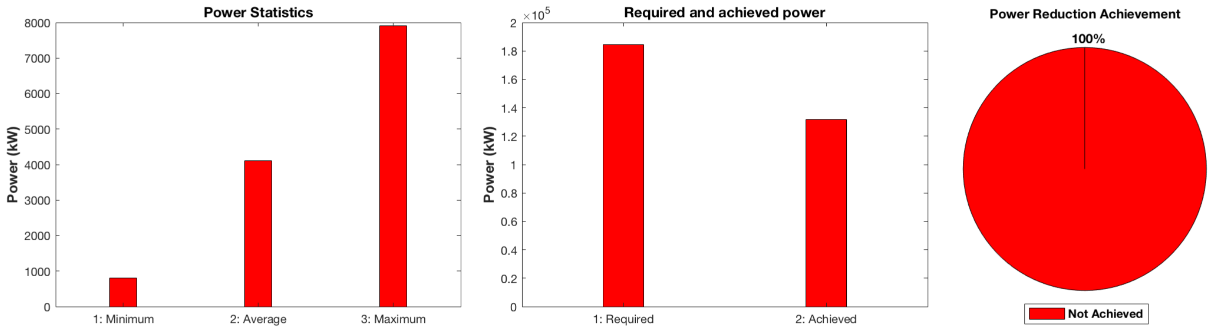


Figure 5.21 – DR power reduction outcomes for 22-DR events in case study 1

Regarding the financial aspects, it is interesting to note that due to the randomness degree implemented in the algorithm, even with the input data contemplating a scenario of predominance of incentives, unchanged and penalty in that order, the participation of data centers in various events have changed the final results for a majority of unchanged, as shown by Figure 5.22.

The same input data used for a DR event was adopted for the 22-day scenario, however the reason for the chart changing, where can be perceived the lack of penalized data centers, was the fact that a data center that reduced a certain power below, or quite below that agreed upon in a particular event, will not necessarily repeat that same action in another event. On the contrary, since this type of contract is already part of a SMDC reduction expectation, data center operators want to take financial advantage and reduce the agreed power. Thus, the final report of this cycle presents 11 small and 8 medium data centers in an unchanged condition, whereas 5 small and 8 medium data centers will receive financial incentive.

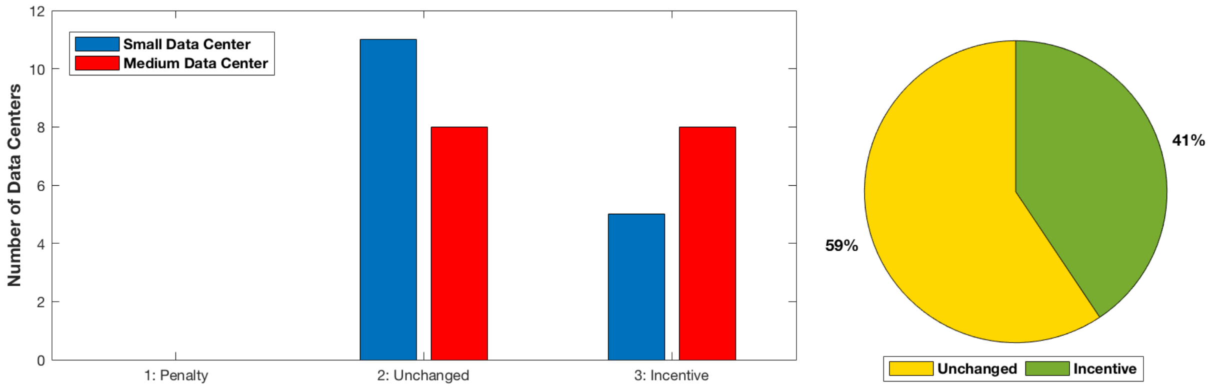


Figure 5.22 – SMDC financial profile for 22-DR events in case study 1

Consequently, based on this final report the highest value received from the DSO by a small data center is 1,394 € and to the medium one is 10,950 € for having reduced the agreed power in the contract. All other data centers did not receive or pay any amount to the DSO, as presented by Figure 5.23.

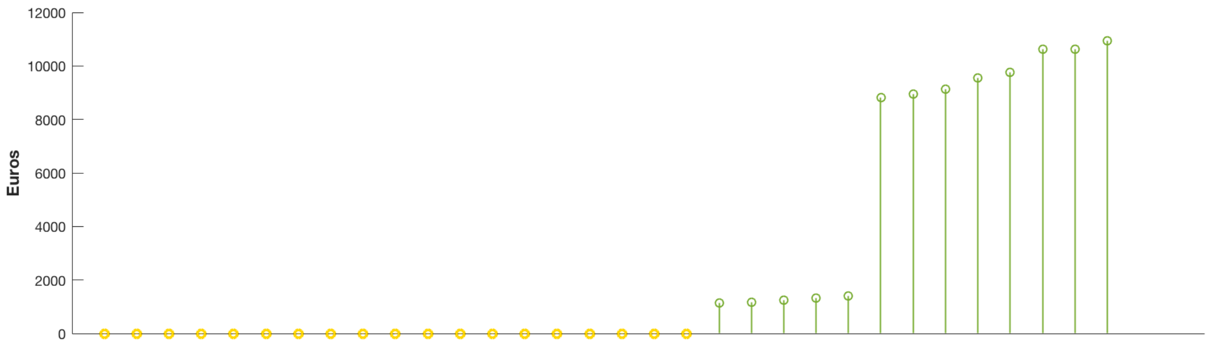


Figure 5.23 – SMDC financial index for 22-DR events in case study 1

5.2.2 Case Study 2

In this second case study, the DSO broadcasts the DR signal to 36 participants, being 21 small and 15 medium data centers in one single DR event. This event follows exactly the same time determinations as the previous case study, but with a difference in the amount of the 80% of power to be reduced by the given adhesion rate, totalizing an accumulated power of 8.36 MW in four periods of 15 minutes each, with 2.09 MW of constant power. It was randomly stipulated for this scenario simulation an acceptance rate of 63%, as described by Table 5.7, where all medium-sized data centers present in the contract accepted the signal, while only 38.09% of the small ones accepted to participate in the DR event, which represents a value below that previously considered by the DSO on sending the signal, being possible to test extreme and non-representative cases in comparison with an ordinary situation. For this decrease process, 8 calls were made to small data centers with 13 rejections and 14 calls to medium data centers without any rejection.

Table 5.7 – Acceptance and denial statistics for one DR event in case study 2

Parameters	Statistics
DR Event Acceptance Percentage (%)	63
DR Event Denial Percentage (%)	37
Small Data Center Acceptance Percentage (%)	38.09
Small Data Center Accepted Calls	8
Small Data Center Denial Percentage (%)	61.91
Small Data Center Rejected Calls	13
Medium Data Center Acceptance Percentage (%)	100
Medium Data Center Accepted Calls	14
Medium Data Center Denial Percentage (%)	0
Medium Data Center Rejected Calls	0

The criterion of fair choice among data centers showed in Figure 5.24 presents almost 40% of non-acceptance rate and in addition one medium data center (with ID 22) which has not been chosen after the random selection process. The ID's from one to 21 are small data centers and from 22 to 36 are medium data centers.

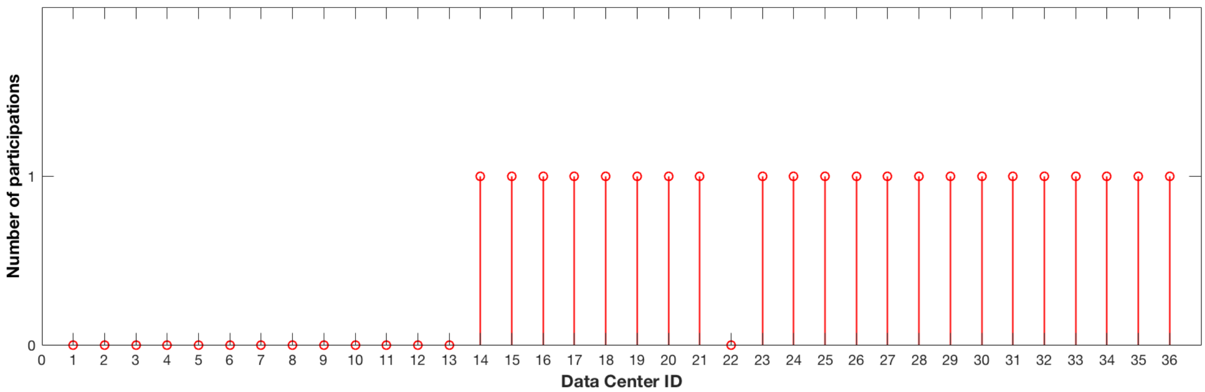


Figure 5.24 – Fair choice criterion for one DR event in case study 2

Observing the power reduction potential in Figure 5.25, it can be noticed that 14 data centers do not reduce any amount of power, four others reached a reduction of 100% and three reached 4% beyond agreed. Additionality, eight data centers have an intermediate reduction percentage of 20%, while five others have just 1% of reduction.

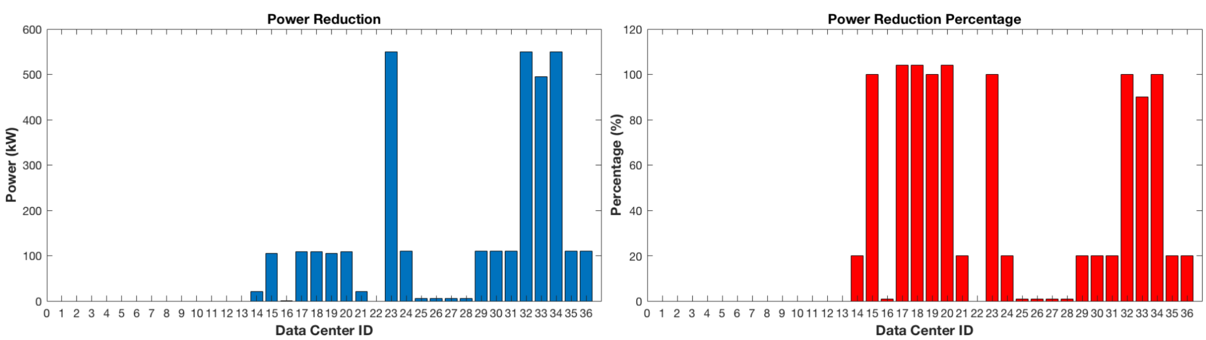


Figure 5.25 – SMDC power reduction for one DR event in case study 2

Particularly, Figure 5.26 depicts the maximum power reduced in this event, 550 kW, the average, 94.6 kW, and the minimum, 1.05 kW. The total power reduced considering all data center contributions is 3.4 MW, which corresponds to just 40.66% of the required power reduction, a value that is justified by the low acceptance percentage of data centers participating in the contract. About the contractual terms, 22% of the data centers accomplished their power reduction objectives, at the same time as 78% did not accomplish.

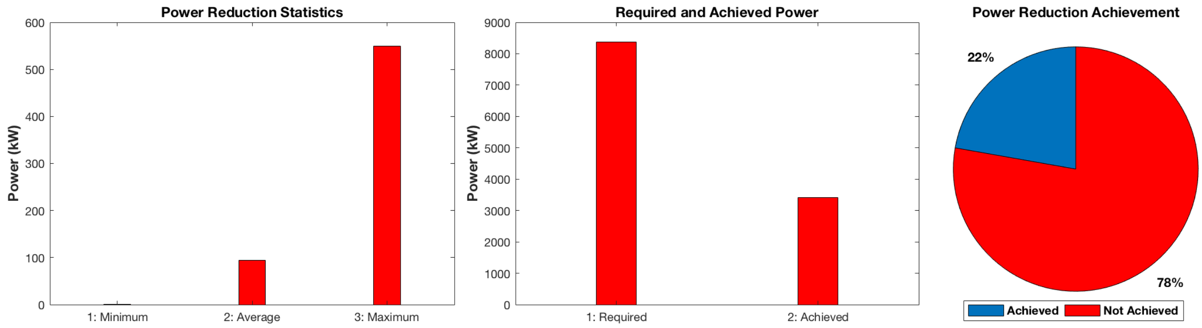


Figure 5.26 – DR power reduction outcomes for one DR event in case study 2

In the results highlighted by Figure 5.27 the majority of data centers falls under the unchanged profile (totalizing 61%), with 15 small data centers and 7 medium ones. 25% of the data centers received a financial incentive of which 5 are of small profile and 4 of medium one. The 14% remnant will pay a penalty, where 1 is small-sized and 4 are medium data centers.

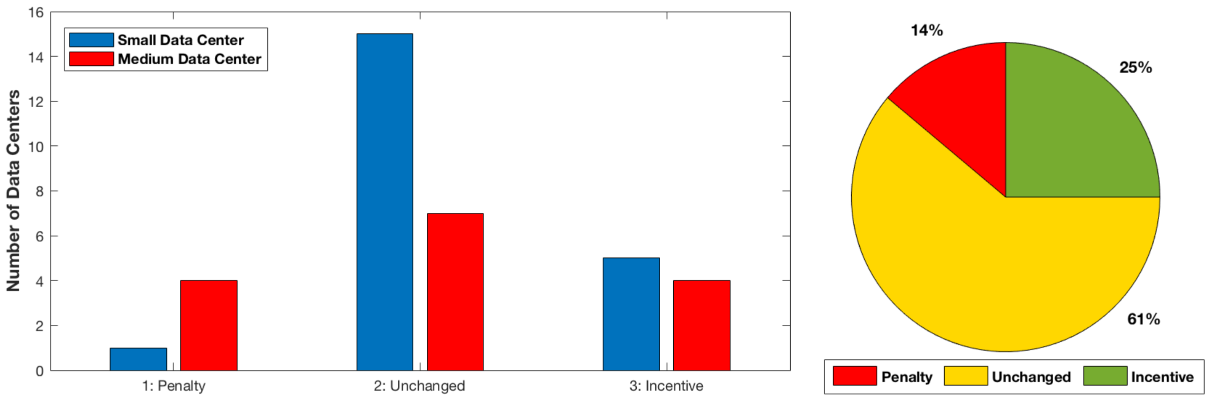


Figure 5.27 – SMDC financial profile for one DR event in case study 2

Therefore, based on this final information the highest value received from the DSO by a small data center is 109.2 € and by a medium one is 847 €, respectively. 22 data centers will not receive, or pay any amount, whilst the highest amount to be paid to the DSO due to a penalty will be 3.94 € to a small and 30.57 € to a medium data center, as shown in Figure 5.28.

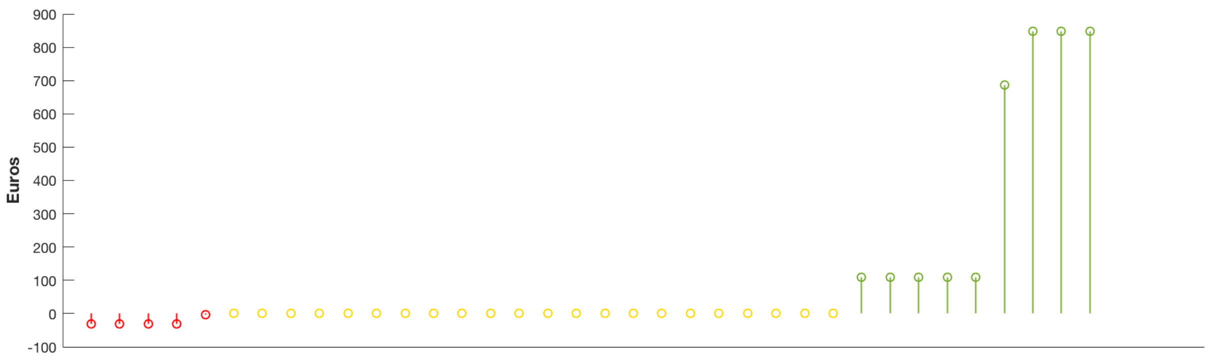


Figure 5.28 – SMDC financial index for one DR event in case study 2

Figure 5.29 indicates the acceptance and denial statistics considering 22 days of cycle. As in this cycle the contractual majority belongs to small data centers, similarly the evident denial actions are more frequent in small data centers than in the medium ones. The highest acceptance rate found was 99% and the lowest was 61%.

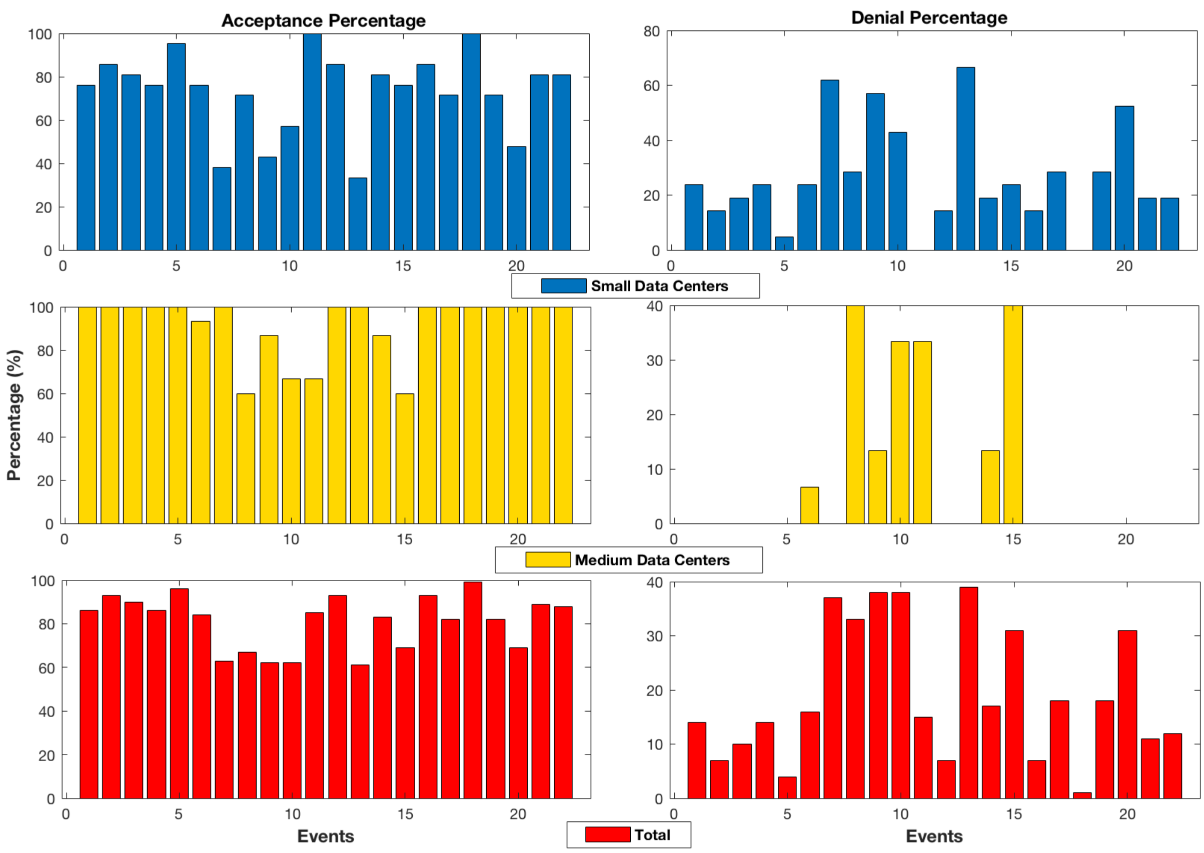


Figure 5.29 – Acceptance and denial statistics for 22-DR events in case study 2

In Figure 5.30, on one hand the number of calls made by small data centers was 339 and by medium ones 244. On the other hand, the total number of rejections was 123 in small data centers and 27 in medium ones.

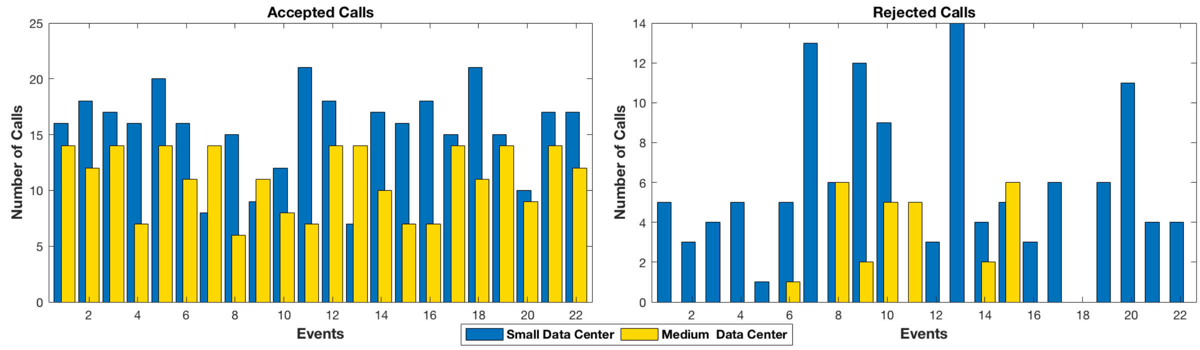


Figure 5.30 – Accepted and rejected calls for 22-DR events in case study 2

The fair choice criterion in Figure 5.31 presents 80.6% of data centers with 16 participations and 19.4% with 17 participations, demonstrating that even considering the possibility of denying participation in a certain event, the fairness criterion of the algorithm works in a balanced way. The ID's from one to 21 are small data centers and from 22 to 36 are medium data centers.

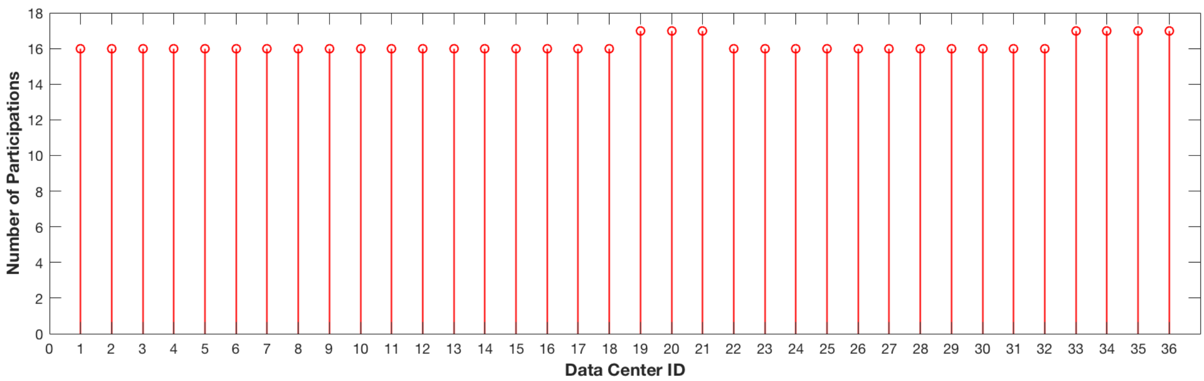


Figure 5.31 – Fair choice criterion for 22-DR events in case study 2

The SMDC achieved power reduction presented in Figure 5.32 shows fluctuations in the range of 12.87% and 67.87%, whilst Figure 5.33 indicates that no data center has been able to achieve 100% of the maximum power agreed in contract. Analyzing all 22 DR events as whole the maximum power reduced was 5,863 kW, the average 2,062 kW and the minimum 216.3 kW. The total accumulated power required by DSO was 184 MW and the accumulated reduced power by SMDC was 74.22 MW, i.e., a DR cycle with just 40.33% of power reduction, i.e., a scenario to test the performance of the algorithm in an extreme and non-representative case of an ordinary contractual situation.

Similarly to the previous case study, the randomness degree implemented in the algorithm made the same input data that initially included the three financial profiles in a single event was changed to a 2-profile chart with an unchanged profile dominance of 97%. Hence, only one small data center was penalized, whereas 15 medium ones and 20 small data centers will not have to receive, or pay any amount to DSO, as depicted by Figure 5.34.

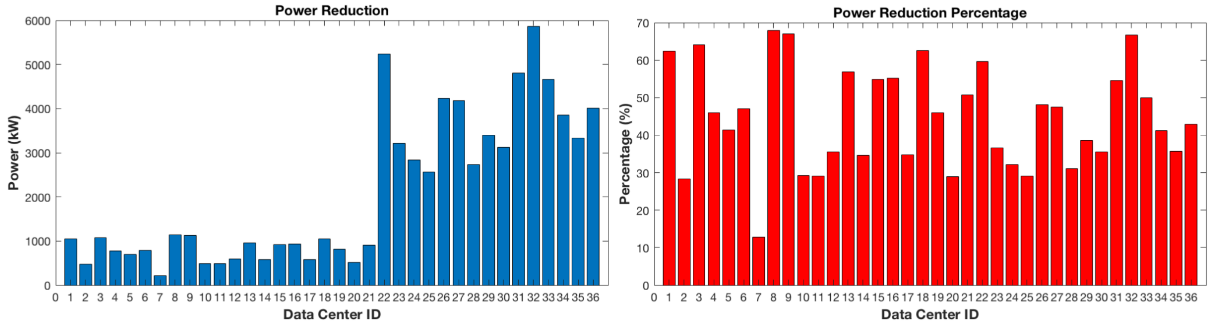


Figure 5.32 – SMDC power reduction potential for 22-DR events in case study 2

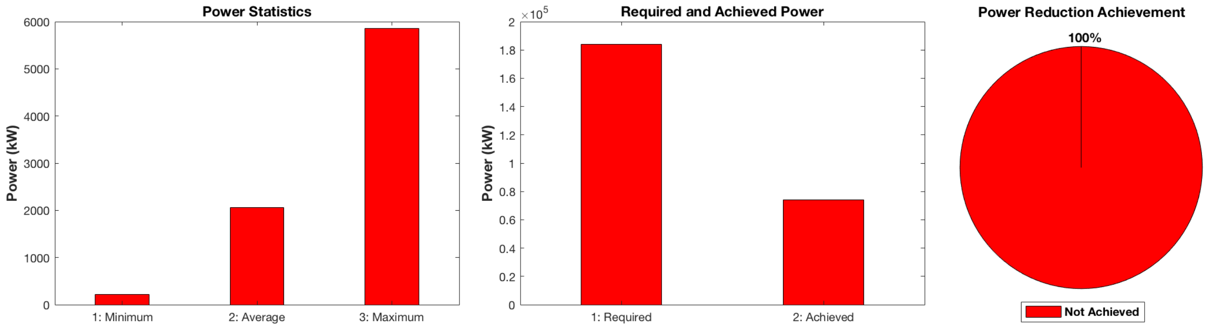


Figure 5.33 – DR power reduction outcomes for 22-DR events in case study 2

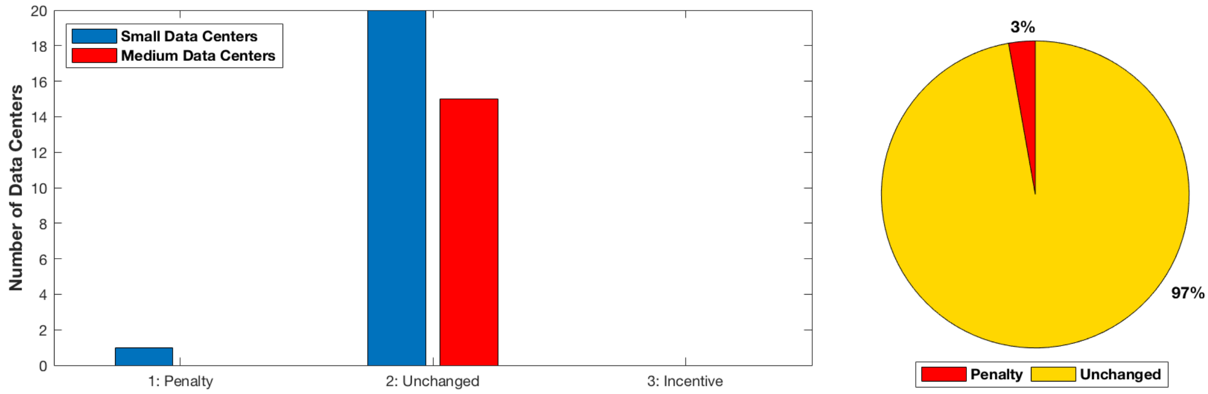


Figure 5.34 – SMDC financial profile for 22-DR events in case study 2

Therefore, according to the final results presented in Figure 5.35, of the 36 data centers present in the contract, 35 had no loss nor profit, and only one small data center should pay to the DSO the amount of 14.47 € for having reduced less than 20% of the previous agreed power. In other words, the algorithm enables to predict such scenario in order to avoid this type of contractual situation.

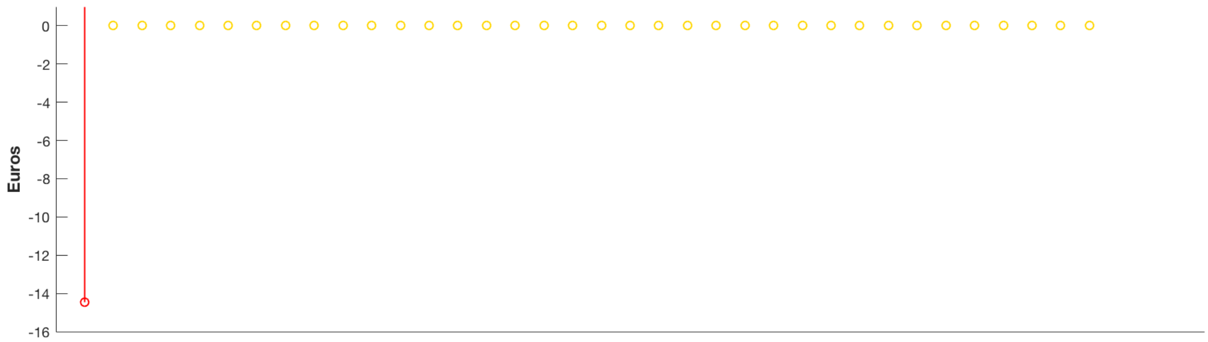


Figure 5.35 – SMDC financial index for 22-DR events in case study 2

5.2.3 Case Study 3

In the third case study the DSO broadcasts the DR signal to 27 participants, in which 10 are small and 17 are medium data centers. This event will follow exactly the same time assumption as the previous case studies, totalizing 8.32 MW in four periods of 15 minutes each, with 2.08 MW of constant power referring to the 80% adhesion. It was randomly stipulated for this scenario simulation an acceptance rate of 85%, as presented in Table 5.8, where all medium data centers existing in the contract accepted the signal, while 60% of the small ones accepted to participate in the DR event. For this reduction activity, six calls were made to small data centers with four rejections and 15 calls to medium data centers without any rejection.

Table 5.8 – Acceptance and denial statistics for one DR event in case study 3

Parameters	Statistics
DR Event Acceptance Percentage (%)	85
DR Event Denial Percentage (%)	15
Small Data Center Acceptance Percentage (%)	60
Small Data Center Accepted Calls	6
Small Data Center Denial Percentage (%)	40
Small Data Center Rejected Calls	4
Medium Data Center Acceptance Percentage (%)	100
Medium Data Center Accepted Calls	15
Medium Data Center Denial Percentage (%)	0
Medium Data Center Rejected Calls	0

The criterion of fair choice among data centers presented in Figure 5.36 shows that although denials to the DR signal occurred only from small data centers, two medium-sized data centers (with ID 11 and 12) were not called upon to reduce the stipulated power after defining the list of participants, to ensure balance between SMDC. The ID's from one to 10 are small data centers and from 11 to 27 are medium data centers.

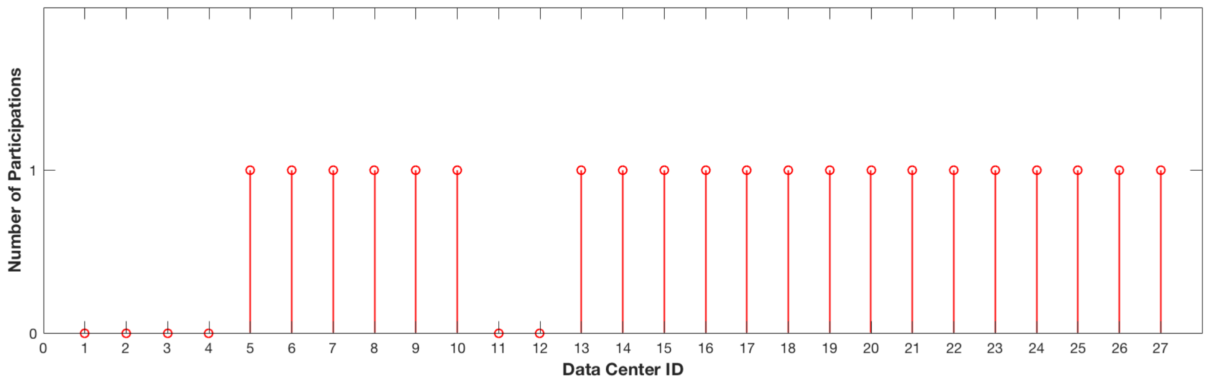


Figure 5.36 – Fair choice criterion for one DR event in case study 3

Analyzing the power reduction in Figure 5.37, on one hand several data centers do not participate of the DR event and obviously do not reduce any amount of power. On the other hand, the majority of the others reach a minimum reduction of 1%, characterizing a very low reduction scenario. In the group that decreases the agreed power, one data center accomplished 100% of the power reduction, other slightly exceeds this value with 104% and other is slightly below with 90%. Of those who reduced an intermediate amount of power, only two data centers reached 20%.

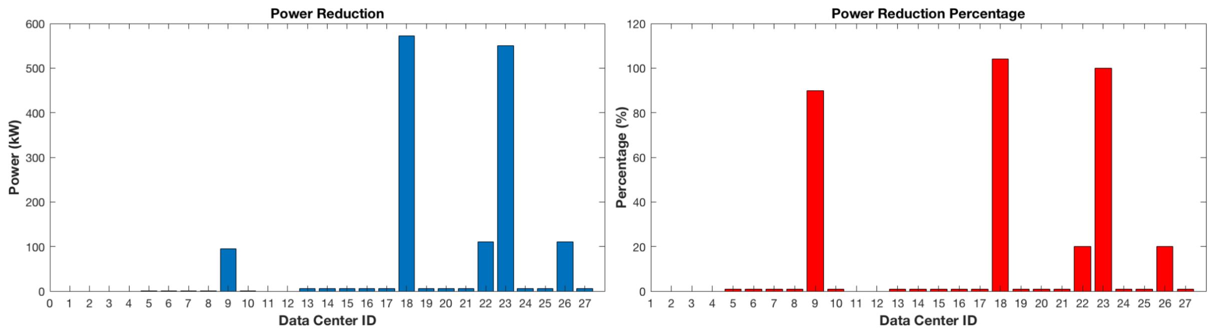


Figure 5.37 – SMDC power reduction for one DR event in case study 3

In a detailed way, as present Figure 5.38, the maximum power reduced in this event is 572 kW, the average 55.63 kW and the minimum 1.05 kW, whereas the total accumulated power reduced considering all participations is 1.5 MW, i.e., 18.02% of power reduction. Regarding the contractual premises, only 7% of the data centers achieved their objective, whilst 93% did not achieve.

In this scenario shown by Figure 5.39, the predominance is of penalty state characterized by 59% of the data centers, where five are small-sized and 11 are medium profile. The unchanged state has 30% of data centers with both small and medium ones represented with four and with only 11% belonging to the incentive state, where there is only one small data center and two medium ones.

Therefore, these final results highlight that the highest value received from the DSO by a small data center is € 88.45 and to a medium one is 847 €. A set of eight data centers will not receive, or pay any amount and the highest amount to be paid to DSO will be 3.94 € to a small and 30.7 € to a medium data center, as can be seen in Figure 5.40. In other words, this is the worst case analyzing a contractual perspective, and this is the main function of the algorithm, allowing to predict this sort of unwanted contractual situations.

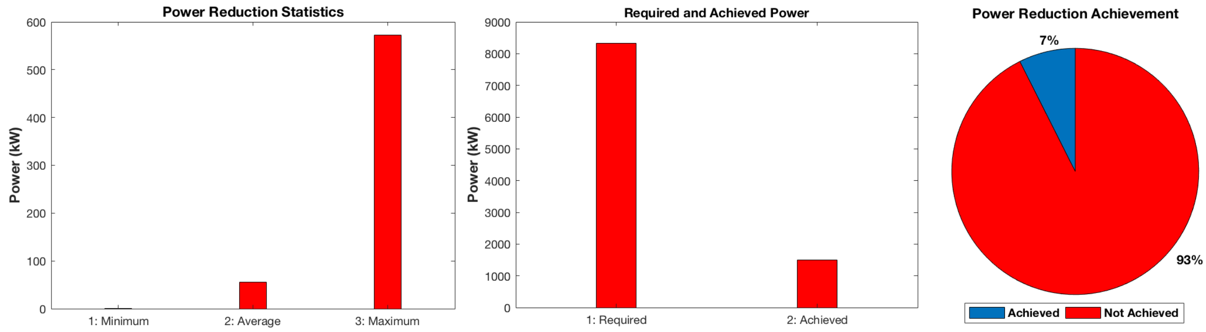


Figure 5.38 – DR power reduction outcomes for one DR event in case study 3

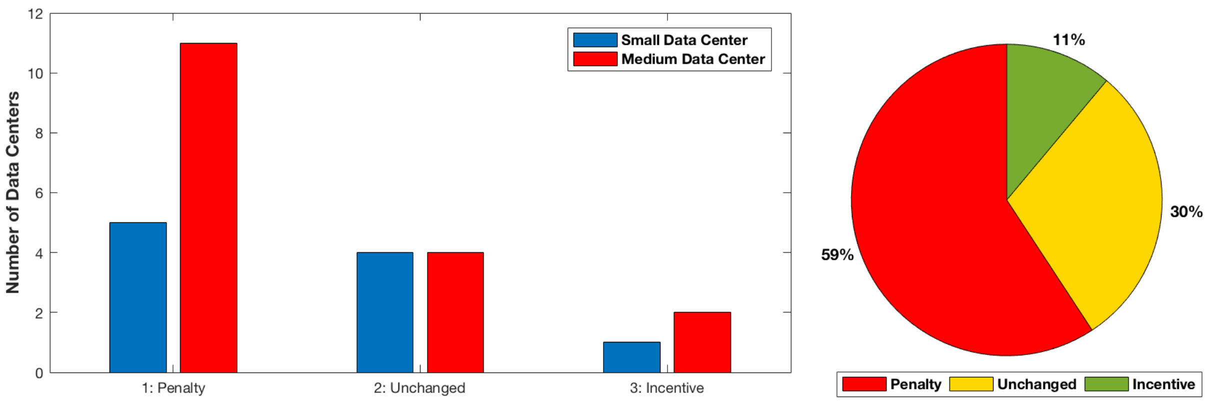


Figure 5.39 – SMDC financial profile for one DR event in case study 3

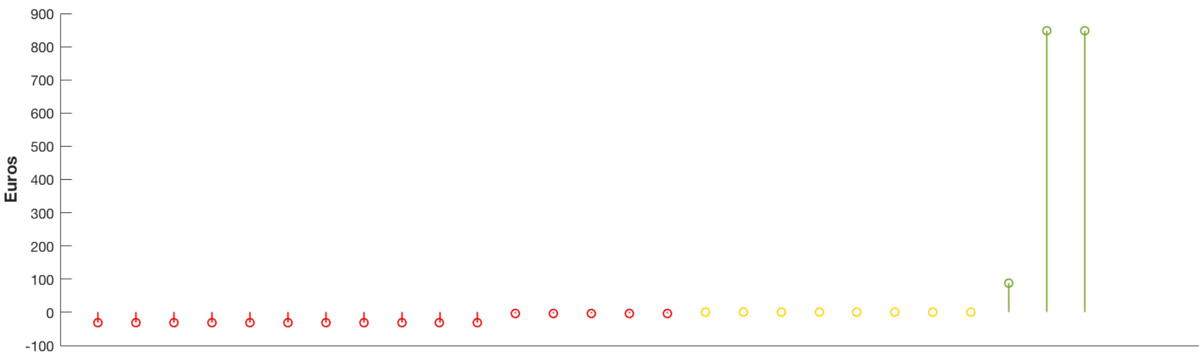


Figure 5.40 – SMDC financial index for one DR event in case study 3

Following the same pattern and evaluating the statistics of acceptance and denial for 22-DR events, this contractual cycle has mostly medium data centers, but the same feature is repeated in this scenario, where the highest percentage of denial occurs through small data centers according to Figure 5.41. In this context, the highest acceptance rate was 99% and the lowest 62%.

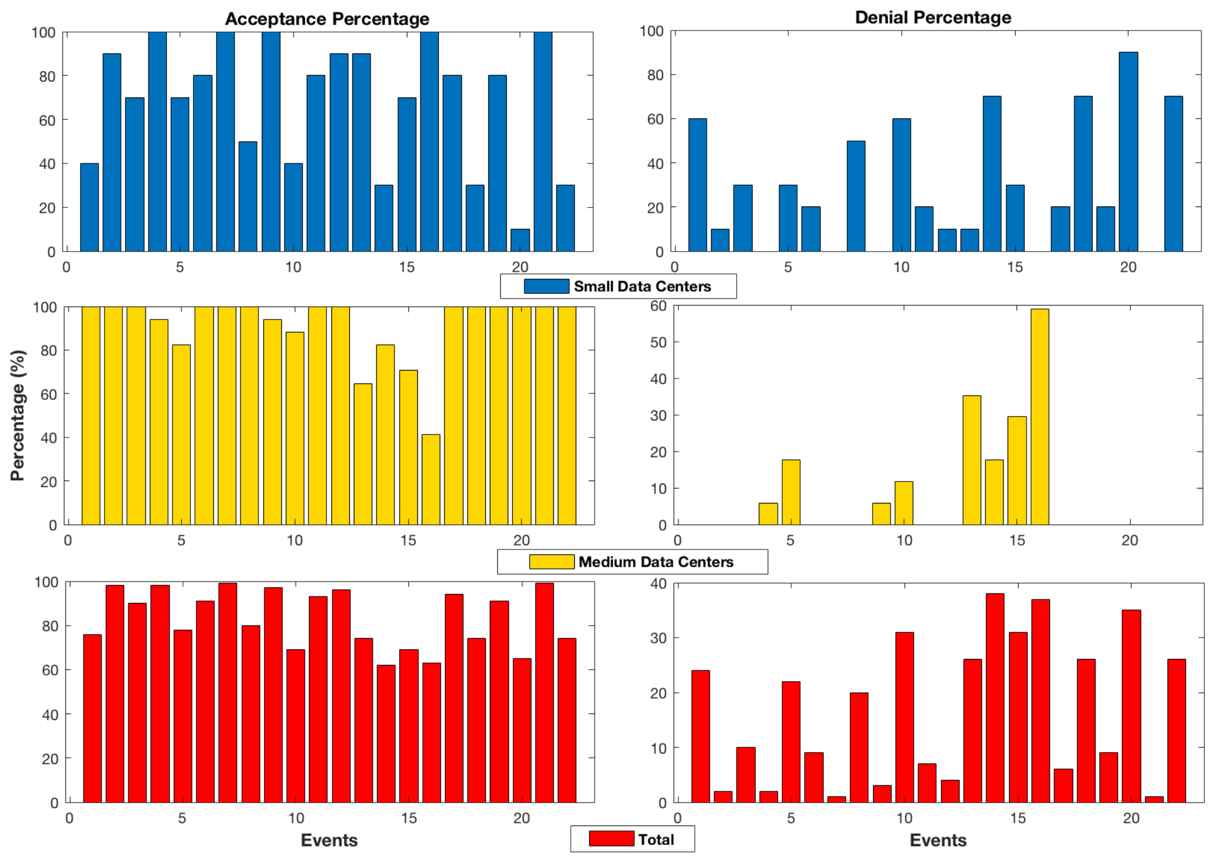


Figure 5.41 – Acceptance and denial statistics for 22-DR events in case study 3

Analyzing in terms of call, Figure 5.42 demonstrates that the number of calls made to small data centers was 153 and to medium ones 202. On the other side, the total number of rejections was 67 in small data centers and 31 in medium ones.

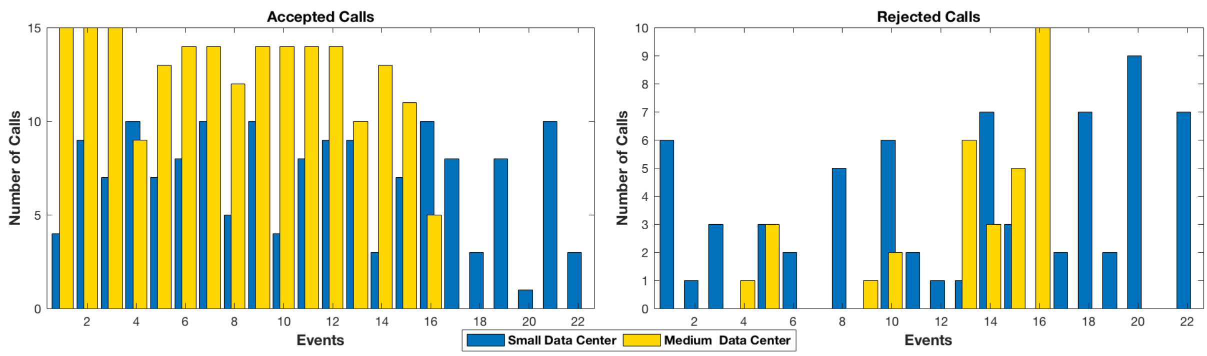


Figure 5.42 – Accepted and rejected calls for 22-DR events in case study 3

In relation to the fair choice criterion for 22-DR events, the chart presented in Figure 5.43 is the one with the most distinct characteristic, not presenting a flatter performance. This phenomenon can be explained because the flat character is observed in the right division between data centers, in which the ID's from one to 10 are small data centers and from 11 to 27 are medium data centers. Thus, on one hand, from the 10 small data centers three ones, i.e., 30% had 16 participations, while seven ones, or 70% had 15 attendances. On the other hand, regarding the medium data centers, two, i.e., 12% had 11 participations, unlike the others 15 (88%), which had 12 attendances.

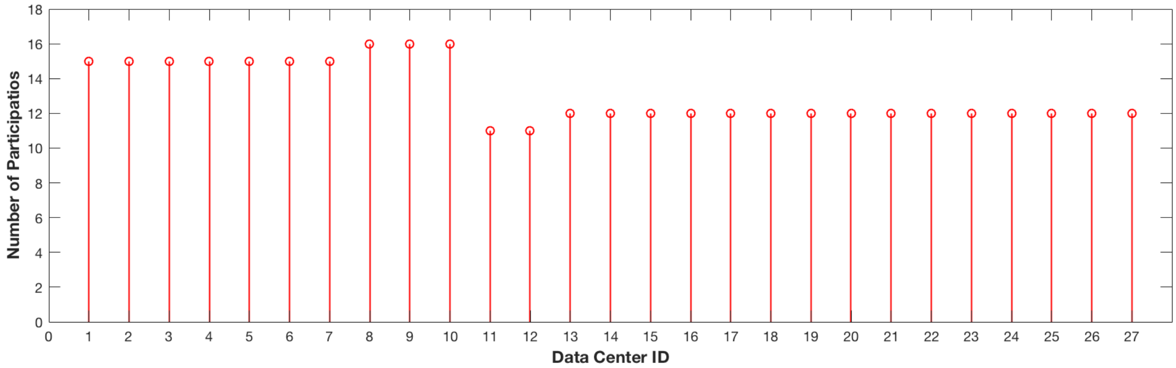


Figure 5.43 – Fair choice criterion for 22-DR events in case study 3

Regarding, the SMDC power reduction, on one hand Figure 5.44 depicts the fluctuations in the range of 2.58% and 33.5%. On the other hand, Figure 5.45 specifies, similarly to the previous case studies, that no data center has been able to achieve 100% of the maximum power agreed in contract. Evaluating all 22 DR events as whole, the maximum power reduced was 2,211 kW, the average 889.7 kW and the minimum 119.7 kW. The total power required by the DSO was 183.03 MW and the reduced power by SMDC was 24.02 MW, i.e., a DR cycle with just 13.12% of power reduction, repeating the statement of how the algorithm is able to predict this type of undesired contractual situation in order to avoid the worst case.

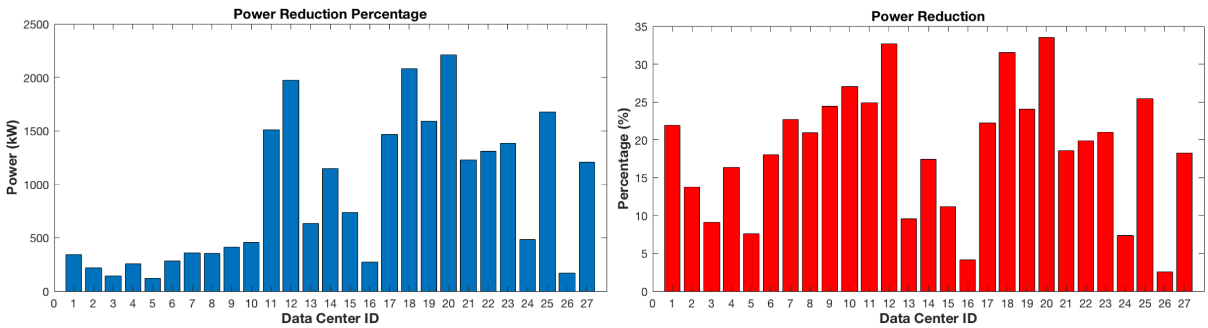


Figure 5.44 – SMDC power reduction potential for 22-DR events in case study 3

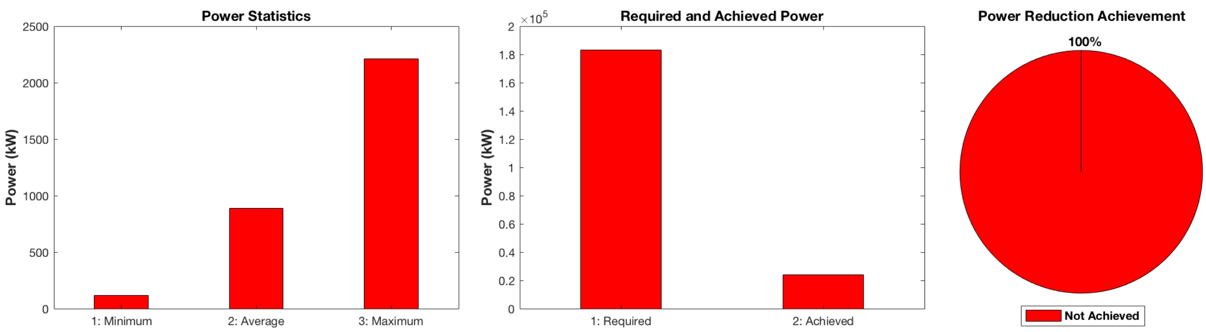


Figure 5.45 – DR power reduction outcomes for 22-DR events in case study 2

In the same way as in previous cases, the randomness degree applied in the algorithm made the same input data that originally contained the three financial profiles in a single event was changed to a 2-profile chart with a penalty profile dominance of 52%. Henceforth, five small data centers and nine medium ones were penalized, whereas five small data centers and eight medium ones will not have to receive, or pay any amount to DSO, as represented by Figure 5.46.

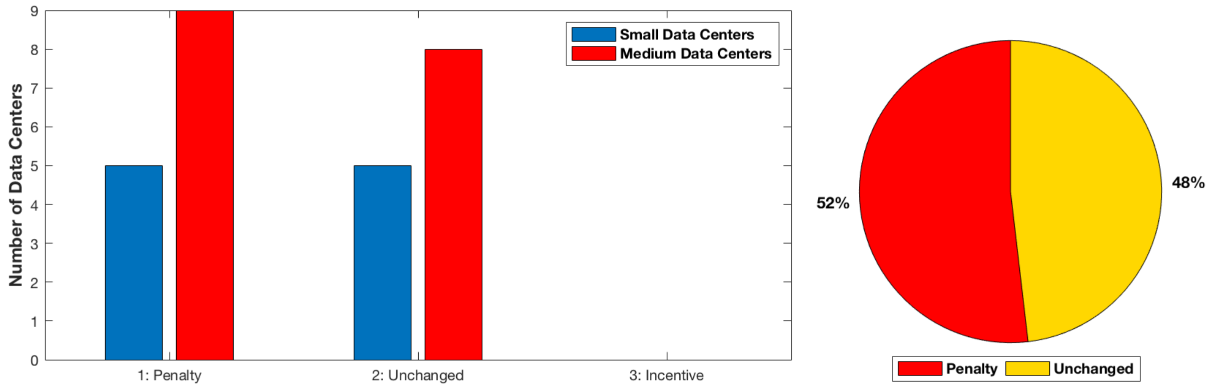


Figure 5.46 – SMDC financial profile for 22-DR events in case study 3

Therefore, in accordance with the final report specified by Figure 5.47, of the 27 data centers present in the contract, 13 of them had no loss nor profit and 14 should pay a penalty to the DSO. Accordingly, the highest value due from a small data center to the DSO was 41.63 €, while from a medium data center was 209.4 € for not having reduced the agreed power in the contract.

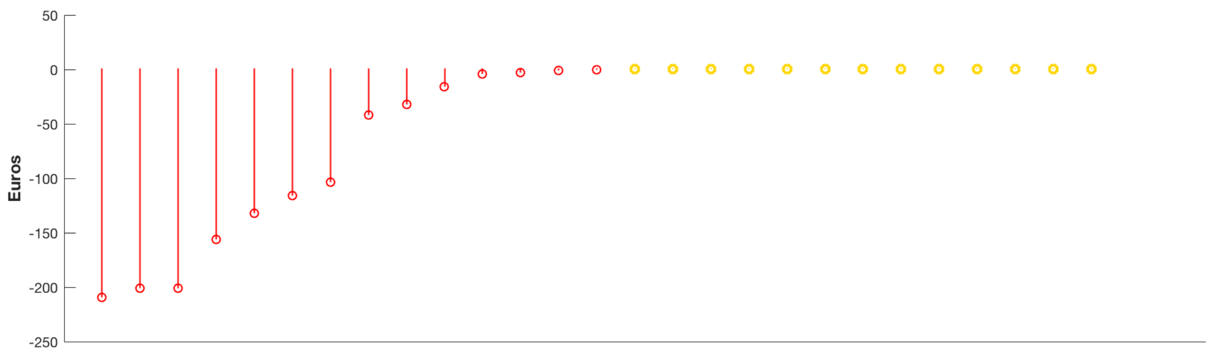


Figure 5.47 – SMDC financial index for 22-DR events in case study 3

5.2.4 Case Study Analysis

Finally, in order to compose a more refined analysis, Table 5.9 presents the compilation of the 3 cases of studies discussed so far in this subsection, specified by the number of data centers used in the contract, as well as whether the determining profile was incentive, unchanged, or penalty. Each case was analyzed taking into account a single event, or a monthly cycle of 22 business days.

Table 5.9 – DSO power and financial balance sheet

1 Event						
Case Study 1 (16/16) Incentive	**DSO Power Balance**					
	Minimum (kW)	Average (kW)	Maximum (kW)	Required (MW)	Achieved (MW)	**Reduction Potential (%)**
	1.05	238.8	572	8.38	7.64	91.16
	DSO Financial Balance					
	Penalty			Incentive		
	Small	Medium	**Total (€)**	Small	Medium	**Total (€)**
	2	1	38.46	13	11	10.580
Case Study 2 (21/17) Unchanged	**DSO Power Balance**					
	Minimum (kW)	Average (kW)	Maximum (kW)	Required (MW)	Achieved (MW)	**Reduction Potential (%)**
	1.05	94.6	550	8.36	3.40	40.66
	DSO Financial Balance					
	Penalty			Incentive		
	Small	Medium	**Total (€)**	Small	Medium	**Total (€)**
	1	4	126.3	5	4	3,336
Case Study 3 (10/17) Penalty	**DSO Power Balance**					
	Minimum (kW)	Average (kW)	Maximum (kW)	Required (MW)	Achieved (MW)	**Reduction Potential (%)**
	1.05	55.63	572	8.32	1.50	18.02
	DSO Financial Balance					
	Penalty			Incentive		
	Small	Medium	**Total (€)**	Small	Medium	**Total (€)**
	5	11	356	1	2	1,782
22 Events						
Case Study 1 (16/16) Incentive	**DSO Power Balance**					
	Minimum (kW)	Average (kW)	Maximum (kW)	Required (MW)	Achieved (MW)	**Reduction Potential (%)**
	812.7	4,112	7,909	184.3	131.6	71.38
	DSO Financial Balance					
	Penalty			Incentive		
	Small	Medium	**Total (€)**	Small	Medium	**Total (€)**
	0	0	0	5	8	84,680
Case Study 2 (21/17) Unchanged	**DSO Power Balance**					
	Minimum (kW)	Average (kW)	Maximum (kW)	Required (MW)	Achieved (MW)	**Reduction Potential (%)**
	216.3	2,062	5,863	184	74.22	40.33
	DSO Financial Balance					
	Penalty			Incentive		
	Small	Medium	**Total (€)**	Small	Medium	**Total (€)**
	1	0	14.47	0	0	0
Case Study 3 (10/17) Penalty	**DSO Power Balance**					
	Minimum (kW)	Average (kW)	Maximum (kW)	Required (MW)	Achieved (MW)	**Reduction Potential (%)**
	119.7	889.7	2,211	183.03	24.02	13.12
	DSO Financial Balance					
	Penalty			Incentive		
	Small	Medium	**Total (€)**	Small	Medium	**Total (€)**
	5	9	1,218	0	0	0

The purpose of this analysis is to provide a comparation from the DSO perspective and a balance sheet as a function of the achieved power reduction and the financial aspect of how much the DSO will have to receive or pay in respect of the SMDC.

The case where the predominance was incentive represents the ideal scenario to justify DR programs applied to consumers such as SMDC, since both the DSO and the data center operators achieved their goals. The former adjusting the load diagram to its needs instead of resorting to power generation and the later to obtain financial advantages in order to reduce energy costs. In other words, this is the case that is expected in the vast majority of situations.

However, the unchanged and penalty cases represent the opposite, i.e., the worst scenarios in which programs do not present benefits for any of the contractually involved actors; data centers would have to pay penalties, or not receive any amount of money per power reduction, and the DSO would be far from its main objectives, which are the power reduction geared to demand management and increase of the grid stability. Hence, these cases are extreme and non-representative, but that were demonstrated in order to reinforce the predictive character of the algorithm to avoid this scenario.

Therefore, the way to minimize bad results is to ensure through the value of incentive and more restricted contractual clauses, a higher DR signal acceptance rate by contract participants and that SMDC operators are able to keep the power reduction rate as close as possible to the values agreed in the contract. Another alternative is to have a larger SMDC universe. If the total number of SMDC is greater, the percentages required to achieve the objectives are lower. In conclusion, this chapter presented the simulation process results and a discussion process towards the adopted case studies, input parameters, running, output data and a comparative analysis addressing all the involved particularities.

The optimization outcomes were firstly demonstrated in SMDC perspective, as well as their respective scenarios. It has been proven the algorithm operation and reliability. The potential for cost savings in DR was also proved, being achieved with the considered incentives in the simulation process, savings of 1.33% for small and 5.15% for medium data centers in the incentive approach, and 0.21% for small and 0.68% for medium data centers in the dynamic tariff approach.

Subsequently, the same premises were utilized with focus on DSO point of view, enabling to predict specific contractual policies that can be adopted in this type of relationship through the best and the worst scenarios simulations. In this context, the scenario with a preponderance of incentives strengthens and proves the need for DR programs applied to SMDCs, however the unchanged and penalty cases allow to foresee a discouraging scenario for this type of DR program and avoid it through the value of incentive, more restricted contractual clauses and increasing the universe of SMDC to reach the objectives with more flexibility. Therefore, in terms of the random values used in the simulation process, the one single day analysis showed a variation in the power reduction potential between 18.02% and 91.16%, while the highest value in the penalty profile was 356 € and in the incentive profile 10,580 €. On the other hand, the 22 business day analyses presented a fluctuation in the power reduction potential in the range of 13.12% and 71.38%, whilst the highest value in the penalty profile was 1,219 € and in the incentive profile 84,680 €.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

In this chapter, a summary of the conclusions drawn in the course of this work is presented, making a critical analysis of the obtained results and answering the research questions raised at the beginning of this thesis. Some suggestions of future work are also indicated.

6.1 CONCLUSIONS

The growth of data centers demand in recent years has led to an increase in their requirements, power, and therefore in their electricity consumption, being the impact of these unique infrastructures on the worldwide energy map increasingly relevant. Modern data centers are filled by computing equipment, which in a critical mission operate uninterruptedly to maintain on-demand network requests. They vary in size from individual servers located in literal closets to expansive warehouses populated with thousands of servers, networking, data storage equipment, as well as a support infrastructure with cooling and UPS solutions.

Data centers designs incorporate oversized and redundant systems, in most cases, running at partial load, and the power and cooling requirements are greatly overstated, leading to extra investment and operational cost expenses. Overall, the data concerning to this type of energy consumer are alarming, since the total energy consumption of data centers is about 1.5% of the global electricity consumption and presents an approximated annual growth rate of 4.3%.

Due to the large savings potential, many researchers have been focused on the development of knowledge, tools and systematic standards to use efficient technologies and methodologies to reduce the data center energy consumption and integrate renewables in their energy portfolio within smart grid context, whereas the power grid evolves together with the integration of smart embedded systems combining instrumentation, analytics and control contributing to deliver electricity in a more reliable and efficient way.

In this perspective, efficient measures in load management and DR increasingly emerge as proposals to mitigate this reality as a whole, being confirmed by important and backed organizations in the U.S. context, such as NIST and DoE and ENTSO-E in the European context.

Thus, several studies have analyzed the evolving landscape of data centers mostly focusing on largest data centers, which are able to obtain large operational cost savings from efficiency improvements and DR programs already in place for industry and great consumers. Companies running large data centers benefit from economies of scale and have the resources to invest in efficiency measures, as well as being an influential player in the energy market. Given high operational costs of large data centers, upfront capital expenditures on efficiency measures are quickly recovered by operating cost savings, in addition to having very expressive individual loads able to take advantage from DR energy programs.

Nevertheless, considerably less attention has been given to small and medium-scale data centers, which account for more than 50% of the total electricity consumption and where the small and medium physical footprint is characteristically embedded within a larger building, becoming difficult an identification and classification process focused on efficiency measures. These spaces are commonly operated by institutions unfamiliar with ICT systems and best practices for data center management, hence the often improvisation, the ad hoc nature and an overlooked power consumption, making SMDC considerably less efficient relative to their large counterparts. Simultaneously, several studies have been addressing DR programs, but almost always applied to large data centers, being evidenced how DR is still insipid in this sort of small and medium profile environment, as well as the policies for this type of energy consumer has often been neglected, as this thesis has demonstrated.

Based on this big picture, this work aimed at alerting about this reality, but not remaining inert to it, rather, asking questions and seeking propositional responses that would provide changes in this current scenario. Thereby, it was intended to know through the research questions how SMDC could take advantage of energy efficiency through centralized management of loads, which loads should be defined by data center operators to respond to DR events, whether measures of energy efficiency, or participation in DR programs could affect the quality and availability of computer services and lastly how data centers owners and utilities could balance the maximization of benefits and sustainable goals.

Mapped to these questions, general and specific goals were drawn in order to find such answers. The general goal of this work was to understand how intensive energy consumers, as SMDC, can become more efficient from the energy point of view and how they can take advantage of DR programs to decrease costs and to cooperate with the grid to ensure higher reliability and sustainable development goals. The specific goals were to define a group of centralized load management strategies, the development of an approach to set specific loads for participation in DR programs, to build an approach to create different DR scenarios based on metrics, thresholds and parameters applied in data centers and lastly, the development of a conceptual framework that balances sustainable development goals with benefits to the grid and SMDC. Thus, the development of this work sought to answer such questions and rigorously follow these goals through the time-distributed work packages that were discussed in sections form in this thesis.

In order to shine a light to the above-mentioned issues, an extensive review on energy efficiency and DR, specifying the state-of-the-art, outlook and connection to SMDC was performed. Several available and prominent technologies recently examined by various authors to stimulate energy efficiency actions were presented. In addition, contemporary studies related to the DR field were addressed highlighting the main advances in the optimization area. All this effort has allowed to have a broad vision of the investigated problem, as well as to penetrate the specificities and peculiarities of this complex technological mix that is a data center.

Through the literature review process, it was possible to collect the main mathematical models related to the energy consumption of the main components of a server, which is the central element of a data center, regardless of size. It was also possible to realize how virtualization, cooling and UPS solutions could contribute in the future DR actions. And mainly, this process allowed to extract valuable data about power, energy consumption, costs and CO_2 emissions regarding SMDC in a characterization mapping, which enabled to use some of these values in the simulation process.

Thereafter, based on an energy efficiency background, three current surveys were presented underlining their premises, one of which was carried out within the framework of this thesis. The main conclusions demonstrated that even though there have been major advances in energy efficiency for large data centers over the last decade with nearly 52% savings, in SMDC the reality is greatly different, whereas 43% of the surveyed have no energy efficiency objectives in place. Furthermore, the survey conducted in this work highlighted alarming statistics, where 64%, 73% and 77% of surveyed participants do not monitor servers', storages appliances' and network devices' energy use, respectively.

Additionally, three energy efficiency methodologies were discussed and assessed emphasizing their main features, as well as their applications. Taking into account the goals of this work the more appropriated methodology was from EL, by the fact of being more suitable to the SMDC case, considering the devices replacement window, the capacity to estimate the savings in five years with cascade effect and using equipment's suggested by ENERGY STARS. Nonetheless, it was pointed out the importance to complement the model utilizing network, backup and storage solutions to improve the extensiveness of the approach in terms of energy efficiency.

According to the previously set goals and study approaches a new framework was proposed. Using two layers was possible to define the main relations and interactions between SMDC and DSO, in order to pursue to combine energy efficiency initiatives, which in turn provide a better knowledge and management of the flexible data center load, as well as the mutual possibility of taking financial advantages in DR programs. To implement it, the mathematical models denoting the key workloads (ICT, cooling and UPS) in a SMDC environment during power decreasing and rebound power conditions were set in an approach aimed at elaborating the respective objective functions and their constraints.

From that point, the two problems established in the context of this work, one from SMDC point of view and the other from DSO perspective were thoroughly specified along with their resolution hypotheses through algorithm optimization process and two strategies: one founded on dynamic tariffs and the other in an incentive-based approach. In this context, three electricity price fluctuation scenarios were adopted to conduct the simulations. The first one considered an hourly electricity price fluctuation, whereby the tariffs are dynamic and fluctuate hourly. The second scenario utilized the same hourly electricity price fluctuation, however in a specific 20-minutes period, where the price was induced to stimulate a DR action and hereinafter contemplated a RE situation. The third one took into account three periods of 20 minutes in each hour with four different tariff periods: peak, half-peak, normal off-peak and super off-peak hours with variations in summer and winter seasons, and contemplating the application of a financial incentive to promote DR actions.

The first problem from SMDC point of view was solved by the use of linear optimization programming techniques, more specifically applying Mixed Integer Linear Programming. The second problem was optimized implementing new polices in form of contractual terms by the development of a random-rotating and fairness algorithm qualified to define, after manifest interest, which SMDC will be chosen in DR events.

The optimization results were firstly validated in SMDC perspective, as well as their corresponding scenarios, in which it has been demonstrated the algorithm operation and reliability. According to the adopted values in the simulation process, the potential for savings in DR approach by incentive showed that the small data center achieved 110 € of saving by event, or in other words, 1.33% of its daily operational cost. Concerning the medium one, it received from the DSO a rebate of 847 €, i.e., 5.15% of its daily energy expense. In the DR approach by dynamic tariffs the small data center reached 15.9 € of savings, which represents 0.22% of its daily operational cost. On the other side, the medium one achieved 99.82 € of savings, i.e., 0.68% or its daily expense. In relation to the three simulated scenarios, in terms of operation the one that was most profitable for DSO, considering or not a DR situation, it was the second scenario with dominance of small data centers, followed by the first one with equivalence of quantity and the third one, with the majority of medium data centers. However, it is important to point out that this did not only depend on the predominance of small or medium data centers, but on the random conditions that varied among scenarios.

Successively, the same properties were employed with focus on DSO point of view, allowing to simulate certain contractual policies that can be implemented in this sort of association through the best and the worst scenarios simulations. The scenario with a preponderance of incentives stimulates the adoption of DR programs applied to SMDCs, nevertheless the unchanged and penalty cases enable to test an unfavorable scenario for this sort of DR program. Therefore, in terms of the random values used in the simulation process, the one single day analysis showed a fluctuation in the power reduction potential between 18.02% and 91.16%, while the highest value in the penalty profile was 356 € and in the incentive profile 10,580 €. The 22 business day analyses presented a variation in the power reduction potential in the range of 13.12% and 71.38%, whilst the highest value in the penalty profile was 1,220 € and in the incentive profile 84,700 €.

Therefore, answering the research questions made by this work consecutively, SMDC can take advantage of energy efficiency through centralized management of loads using specific methodologies such as EL assessed by this work, in addition to keeping up to date with the best practices pointed out by the main surveys. It was also concluded that the loads that should be defined by data center operators to respond to DR events are UPS, cooling and finally ICT workloads, for being inserted more properly in the main business of data centers. The suggested framework along with the optimization process responded to the last two questions, demonstrating that it is necessary to establish specific energy policies appropriate to the particularity of the SMDC market, jointly dealing with it and not isolated as in the case of large data centers. Thus, it is possible not to affect the quality and availability of computer services, extract financial benefits and be in line with sustainable goals.

6.2 FUTURE WORK

Related to the developed work there are several alternatives that could be used to make this research either more comprehensive on one hand, or more specific on the other hand, such as:

- The profile of server room and very small data centers can be inserted in the domain of the analyzed data center size.
- Data from real data centers with different characteristics can be used in simulations with their consumption profiles.
- The same research can be applied to other countries, or specific energy markets.
- This work addressed load management with energy efficiency measures and participation in DR programs directed to the SMDC, but it would be important to analyze the participation of this type of energy consumer in other energy markets, such as Ancillary Services. Similar research questions, specific objectives, applied literature review, inclusion on the specter of this framework, modeling and optimization could bring answers to SMDC, DSOs, TSOs and aggregators in this direction.
- Different tariffs models can be considered, analyzing their particularities and impact.
- Other incentive values can be used to analyze the optimal balance for data center operators and DSO.
- Only three scenarios were analyzed in terms of SMDC quantity and two scenarios in relation to the number of events. This quantitative could be extrapolated in order to study if such a change would directly affect the policies adopted and the suggestion of new ones.
- The algorithm from the DSO point of view can include the impact assessment of the RE in the contractual terms, as well as the costs minimization through tariff variations, or incentives.
- A multi-objective analysis with other objective functions such as reliability and QoS could be considered in the future, as well as the use of other techniques and algorithms.
- A stronger integration between the DR methodology and energy efficiency can be implemented, for example, by adding the ability to simulate the impact of energy efficiency measures.

REFERENCES

Aghaei, J. and Alizadeh, M. I. (2013) ‘Demand response in smart electricity grids equipped with renewable energy sources: A review’, *Renewable and Sustainable Energy Reviews*, pp. 64–72. doi: 10.1016/j.rser.2012.09.019.

Ahmad, R. W., Gani, A., Hamid, S. H. A., Shiraz, M., Yousafzai, A. and Xia, F. (2015) ‘A survey on virtual machine migration and server consolidation frameworks for cloud data centers’, *Journal of Network and Computer Applications*, pp. 11–25. doi: 10.1016/j.jnca.2015.02.002.

Ahn, J. and Jouppi, N. (2009) ‘Future scaling of processor-memory interfaces’, *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis - SC ’09.* New York, New York, USA: ACM Press, p. 1. doi: 10.1145/1654059.1654102.

Alkharabsheh, S., Fernandes, J., Gebrehiwot, B., Agonafer, D., Ghose, K., Ortega, A., Joshi, Y. and Sammakia, B. (2015) ‘A Brief Overview of Recent Developments in Thermal Management in Data Centers’, *Journal of Electronic Packaging*, 137(4), p. 40801. doi: 10.1115/1.4031326.

Amur, H., Cipar, J. and Gupta, V. (2010) ‘Robust and Flexible Power-Proportional Storage’, *Proceedings of the 1st ACM symposium on Cloud computing - SoCC ’10*, pp. 217--228. doi: 10.1145/1807128.1807164.

Añón Higón, D., Gholami, R. and Shirazi, F. (2017) ‘ICT and environmental sustainability: A global perspective’, *Telematics and Informatics*, 34(4), pp. 85–95. doi: 10.1016/j.tele.2017.01.001.

‘Apple doubles fuel cell capacity at NC data centre with Bloom’ (2013) *Fuel Cells Bulletin*, 2013(1), p. 5. doi: 10.1016/S1464-2859(13)70011-7.

Aravanis, A. I., Voulkidis, A., Salom, J., Townley, J., Georgiadou, V., Oleksiak, A., Porto, M. R., Roudet, F. and Zahariadis, T. (2015) ‘Metrics for assessing flexibility and sustainability of next generation data centers’, *2015 IEEE Globecom Workshops, GC Wkshps 2015 - Proceedings*. doi: 10.1109/GLOCOMW.2015.7414182.

Ardagna, D., Panicucci, B., Trubian, M. and Zhang, L. (2012) ‘Energy-aware autonomic resource allocation in multitier virtualized environments’, *IEEE Transactions on Services Computing*, 5(1), pp. 2–19. doi: 10.1109/TSC.2010.42.

Arianyan, E., Taheri, H. and Khoshdel, V. (2017) ‘Novel fuzzy multi objective DVFS-aware consolidation heuristics for energy and SLA efficient resource management in cloud data centers’, *Journal of Network and Computer Applications*, 78, pp. 43–61. doi: 10.1016/j.jnca.2016.09.016.

Arlitt, M., Bash, C., Blagodurov, S., Chen, Y., Christian, T., Gmach, D., Hyser, C., Kumari, N., Liu, Z., Marwah, M., McReynolds, A., Patel, C., Shah, A., Wang, Z. and Zhou, R. (2012) 'Towards the design and operation of net-zero energy data centers', *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2012 13th IEEE Intersociety Conference on*, pp. 552–561. doi: 10.1109/itherm.2012.6231479.

ASHRAE (2011) '2011 Gaseous and Particulate Contamination Guidelines For Data Centers', *American Society of Heating, Refrigerating and Air-Conditioning Engineers*, pp. 1–22.

ASHRAE Technical Committee 9.9 (2011) 'Thermal Guidelines for Data Processing Environments', *American Society of Heating, Refrigerating, and Air-Conditioning Engineers Inc*.

Barroso, L. A. and Hölzle, U. (2007) 'The case for energy-proportional computing', *Computer*, 40(12), pp. 33–37. doi: 10.1109/MC.2007.443.

Basmadjian, R., Ali, N., Niedermeier, F., de Meer, H. and Giuliani, G. (2011) 'A methodology to predict the power consumption of servers in data centres', *Proceedings of the 2nd International Conference on Energy-Efficient Computing and Networking - e-Energy '11*. New York, New York, USA: ACM Press, p. 1. doi: 10.1145/2318716.2318718.

Basmadjian, R., Lovasz, G., Beck, M., Meer, H. De, Hesselbach-Serra, X., Botero, J. F., Klingert, S., Ortega, M. P., Lopez, J. C., Stam, A., Krevelen, R. Van and Girolamo, M. Di (2013) 'A Generic Architecture for Demand Response: The ALL4Green Approach', *2013 International Conference on Cloud and Green Computing*. IEEE, pp. 464–471. doi: 10.1109/CGC.2013.79.

Basmadjian, R., Müller, L. and De Meer, H. (2015) 'Data centres' power profile selecting policies for Demand Response: Insights of Green Supply Demand Agreement', *Ad Hoc Networks*, 25(PB), pp. 581–594. doi: 10.1016/j.adhoc.2014.11.007.

Basmadjian, R., Niedermeier, F., Lovasz, G., De Meer, H. and Klingert, S. (2013) 'GreenSDAs leveraging power adaption collaboration between energy provider and data centres', *2013 Sustainable Internet and ICT for Sustainability, SustainIT 2013*. doi: 10.1109/SustainIT.2013.6685195.

BCSDPortugal (2017) 'Data Center da PT Portugal', pp. 1–9. *Accessed: 14-01-2017*.

Beitelmal, A. H. and Fabris, D. (2014) 'Servers and data centers energy performance metrics', *Energy and Buildings*, 80, pp. 562–569. doi: 10.1016/j.enbuild.2014.04.036.

Beloglazov, A. and Buyya, R. (2013) 'Managing overloaded hosts for dynamic consolidation of virtual machines in cloud data centers under quality of service constraints', *IEEE Transactions on Parallel and Distributed Systems*, 24(7), pp. 1366–1379. doi: 10.1109/TPDS.2012.240.

Bennett, D. and Delforge, P. (2012) 'Small Server Rooms , Big Energy Savings Opportunities and Barriers to Energy Efficiency on the Small Server Room Market', *NRDC*, (February).

Bixby, R. E. (2012) 'A Brief History of Linear and Mixed-Integer Programming Computation', *Documenta Mathematica · Extra*, ISMP ISMP, pp. 107–121.

Bonetto, E., Finamore, A., Mellia, M. and Fiandra, R. (2014) 'Energy efficiency in access and aggregation networks: From current traffic to potential savings', *Computer Networks*, 65, pp. 151–166. doi: 10.1016/j.comnet.2014.03.008.

Boulos, S., Nuttall, C., Harrison, B., Moura, P. and Jehle, C. (2014) 'ErP Lot 27 – Uninterruptible Power Supplies: Preparatory Study - Final Report'. *Ricardo-AEA, Intertek, ISR – University of Coimbra*.

Brady, G. A., Kapur, N., Summers, J. L. and Thompson, H. M. (2013) 'A case study and critical assessment in calculating power usage effectiveness for a data centre', *Energy Conversion and Management*, 76, pp. 155–161. doi: 10.1016/j.enconman.2013.07.035.

Brocanelli, M., Li, S., Wang, X. and Zhang, W. (2014) 'Maximizing the revenues of data centers in regulation market by coordinating with electric vehicles', *Sustainable Computing: Informatics and Systems*. Elsevier Inc., pp. 1–13. doi: 10.1016/j.suscom.2014.03.004.

Calheiros, R. N., Ranjan, R., De Rose, C. A. F. and Buyya, R. (2009) 'CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services', *arXiv preprint arXiv:0903.2525*, p. 9.

Camacho, J., Zhang, Y., Chen, M. and Chiu, D. M. (2014) 'Balance your Bids before your Bits : The Economics of Geographic Load-Balancing', *Proc. of the 5th Int. Conf. on Future Energy Systems (ACM e-Energy)*, pp. 75–85. doi: 10.1145/2602044.2602068.

Canuto, M., Bosch, R., Macias, M. and Guitart, J. (2016) 'A methodology for full-system power modeling in heterogeneous data centers', *Proceedings of the 9th International Conference on Utility and Cloud Computing - UCC '16*. New York, New York, USA: ACM Press, pp. 20–29. doi: 10.1145/2996890.2996899.

Capozzoli, A., Chinnici, M., Perino, M. and Serale, G. (2015) 'Review on Performance Metrics for Energy Efficiency in Data Center: The Role of Thermal Management', *Energy Efficient Data Centers*, pp. 135–151. doi: 10.1007/978-3-319-15786-3_9.

Cappuccio, D. J. (2010) 'DCIM : Going Beyond IT', *Research ID Number: G00174769*, (March).

Carreiro, A. M., Jorge, H. M. and Antunes, C. H. (2017) 'Energy management systems aggregators: A literature survey', *Renewable and Sustainable Energy Reviews*, pp. 1160–1172. doi: 10.1016/j.rser.2017.01.179.

Castro, P. H. P., Cardoso, K. V and Corrêa, S. (2013) 'Uma Abordagem Baseada no Consumo de CPU e RAM para a Eficiencia Energetica em Centros de Dados para Computação em Nuvem', *Wscad-Ssc 2013*, (February), p. 8. doi: 10.13140/2.1.4681.0242.

Cecati, C., Mokryani, G., Piccolo, A. and Siano, P. (2010) 'An overview on the Smart Grid concept', *IECON Proceedings (Industrial Electronics Conference)*, pp. 3322–3327. doi: 10.1109/IECON.2010.5675310.

Chen, N., Ren, X., Ren, S. and Wierman, A. (2015) 'Greening multi-tenant data center demand response', *Performance Evaluation*. Elsevier B.V., 91, pp. 229–254. doi: 10.1016/j.peva.2015.06.014.

Chen, T., Gao, X. and Chen, G. (2016) 'The features, hardware, and architectures of data center networks: A survey', *Journal of Parallel and Distributed Computing*, 96, pp. 45–74. doi: 10.1016/j.jpdc.2016.05.009.

Chinnici, M., Capozzoli, A. and Serale, G. (2016) 'Measuring energy efficiency in data centers', *Pervasive Computing*. Elsevier, pp. 299–351. doi: 10.1016/B978-0-12-803663-1.00010-3.

Christantoni, D., Oxizidis, S., Flynn, D. and Finn, D. P. (2016) 'Implementation of demand response strategies in a multi-purpose commercial building using a whole-building simulation model approach', *Energy and Buildings*, 131, pp. 76–86. doi: 10.1016/j.enbuild.2016.09.017.

Cioara, T., Anghel, I., Bertoncini, M., Salomie, I., Arnone, D., Mammina, M., Velivassaki, T.-H. and Antal, M. (2016) 'Optimized flexibility management enacting Data Centres participation in Smart Demand Response programs', *Future Generation Computer Systems*. doi: 10.1016/j.future.2016.05.010.

Clímaco, J. N., Antunes, C. H. and Alves, M. J. G. (2003) *Programação linear multiobjectivo: do modelo de programação linear clássico à consideração explícita de várias funções objectivo*. Imprensa da Universidade de Coimbra. doi: 10.14195/978-989-26-0479-4.

Craig-wood, K., Krause, P. and Mason, A. (2010) 'Green ICT : Oxymoron or call to innovation?', *Transport*, (Ict), pp. 978–981. doi: 10.5176/978-981-08-7240-3.

Cupertino, L., Da Costa, G., Oleksiak, A., Pia¸tek, W., Pierson, J.-M., Salom, J., Sisó, L., Stolf, P., Sun, H. and Zilio, T. (2015) 'Energy-efficient, thermal-aware modeling and simulation of data centers: The CoolEmAll approach and evaluation results', *Ad Hoc Networks*, 25, pp. 535–553. doi: 10.1016/j.adhoc.2014.11.002.

Dai, J., Das, D., Ohadi, M. and Pecht, M. (2013) 'Reliability risk mitigation of free air cooling through prognostics and health management', *Applied Energy*, 111, pp. 104–112. doi: 10.1016/j.apenergy.2013.04.047.

Dai, J., Das, D. and Pecht, M. (2012a) 'A multiple stage approach to mitigate the risks of telecommunication equipment under free air cooling conditions', *Energy Conversion and Management*, 64, pp. 424–432. doi: 10.1016/j.enconman.2012.06.018.

Dai, J., Das, D. and Pecht, M. (2012b) 'Prognostics-based risk mitigation for telecom equipment under free air cooling conditions', *Applied Energy*, 99, pp. 423–429. doi: 10.1016/j.apenergy.2012.05.055.

Dargie, W. (2014) 'Estimation of the cost of VM migration', *Proceedings - International Conference on Computer Communications and Networks, ICCCN*. doi: 10.1109/ICCCN.2014.6911756.

Darrow, K. and Hedman, B. (2009) 'Opportunities for Combined Heat and Power in Data Centers', *Cogeneration and Distributed Generation*.

Dayarathna, M., Wen, Y. and Fan, R. (2016) 'Data Center Energy Consumption Modeling: A Survey', *IEEE Communications Surveys & Tutorials*, 18(1), pp. 732–794. doi: 10.1109/COMST.2015.2481183.

Delforge, P. (2014) 'America's Data Centers Are Wasting Huge Amounts of Energy', *Natural Resources Defense Council (NRDC)*, IB:14-08-A(August), pp. 1–5.

Delta (2018) *Delta Power Solutions*. *Available at: http://www.deltapowersolutions.com. Accessed: 22-02-2018.*

Deng, Y. (2011) 'What is the future of disk drives, death or rebirth?', *ACM Computing Surveys*, 43(3), pp. 1–27. doi: 10.1145/1922649.1922660.

Depoorter, V., Oró, E. and Salom, J. (2015) 'The location as an energy efficiency and renewable energy supply measure for data centres in Europe', *Applied Energy*, 140, pp. 338–349. doi: 10.1016/j.apenergy.2014.11.067.

Dhiman, G., Marchetti, G. and Rosing, T. S. (2010) 'VGreen: A system for energy-efficient management of virtual machines', *ACM Transactions on Design Automation of Electronic Systems*, 16(1), pp. 1–27. doi: 10.1145/1870109.1870115.

Doyle, J., Shorten, R. and O'mahony, D. (2013) 'Stratus: Load balancing the cloud for carbon emissions control', *IEEE Transactions on Cloud Computing*, 1(1), pp. 116–128. doi: 10.1109/TCC.2013.4.

Dreibholz, T., Becke, M. and Adhari, H. (2007) 'Report to Congress on Server and Data Center Energy Efficiency Public Law 109-431', *tdr.wiwi.uni-due.de*, 109, p. 431.

Du, P. and Lu, N. (2011) 'Appliance commitment for household load scheduling', *IEEE Transactions on Smart Grid*, 2(2), pp. 411–419. doi: 10.1109/TSG.2011.2140344.

Ebrahimi, K., Jones, G. F. and Fleischer, A. S. (2014) 'A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities', *Renewable and Sustainable Energy Reviews*, 31, pp. 622–638. doi: 10.1016/j.rser.2013.12.007.

Emerson (2015) ‘Energy Logic : Reducing Data Center Energy Consumption by Creating Savings that Cascade Across Systems’, *White Paper, Emerson Network Power*, pp. 1–21.

Energy Services Regulatory Authority (2018) *Electricity Market*. *Available at: http://www.erse.pt/eng/mktsupervision/electricitymkt/Paginas/default.aspx. Accessed: 21-04-2018.*

Eu Commission Tf For Smart Grids Expert Group (2010) ‘Functionalities of smart grids and smart meters’, *Group*, 23(December), pp. 1–69.

Fang, X., Misra, S., Xue, G. and Yang, D. (2012) ‘Smart Grid — The New and Improved Power Grid: A Survey’, *IEEE Communications Surveys & Tutorials*, 14(4), pp. 944–980. doi: 10.1109/SURV.2011.101911.00087.

Flanagan, C. (2013) ‘A Data Center Perspective on Demand Response’, *Data Centers Dynamics*.

Fridgen, G., Keller, R., Thimmel, M. and Wederhake, L. (2017) ‘Shifting load through space–The economics of spatial demand side management using distributed data centers’, *Energy Policy*, 109, pp. 400–413. doi: 10.1016/j.enpol.2017.07.018.

Fulpagare, Y. and Bhargav, A. (2015) ‘Advances in data center thermal management’, *Renewable and Sustainable Energy Reviews*, pp. 981–996. doi: 10.1016/j.rser.2014.11.056.

Garimella, S. V., Persoons, T., Weibel, J. and Yeh, L. T. (2013) ‘Technological drivers in data centers and telecom systems: Multiscale thermal, electrical, and energy management’, *Applied Energy*, pp. 66–80. doi: 10.1016/j.apenergy.2013.02.047.

Gavald, O., Depoorter, V. and Oppelt, T. (2014) ‘Report of different options for renewable energy supply in Data Centres in Europe’, *RenewIT*.

Gellings, C. (EPRI) (2011) 'Estimating the Costs and Benefits of the Smart Grid', *Power*.

Ghatikar, G., Ganti, V. and Matson, N. (2012) 'Demand Response Opportunities and Enabling Technologies for Data Centers : Findings from Field Studie's, *Lawrence Berkeley National Laboratory*.

Ghatikar, G., Piette, M. A., Fujita, S., McKane, A., J.H., D., Radspieler, A., Mares, K. C. and Shroyer, D. (2010) 'Demand Response and Open Automated Demand Response Opportunities for Data Centers', *Lawrence Berkeley National Laboratory*.

Goiri, Í., Katsak, W., Le, K., Nguyen, T. D. and Bianchini, R. (2013) ‘Parasol and GreenSwitch’, *Proceedings of the eighteenth international conference on Architectural support for programming languages and operating systems - ASPLOS ’13*. New York, New York, USA: ACM Press, p. 51. doi: 10.1145/2451116.2451123.

Grice, J. W., Dean, N. and Eddie, S. (2013) ‘Sustainable Site Selection : The Convergence of Data Center

Site', *The Green Grid. Research Report*, (Sustainable Site Selection), p. 20.

Gu, C., Fan, L., Wu, W., Huang, H. and Jia, X. (2016) 'Greening cloud data centers in an economical way by energy trading with power grid', *Future Generation Computer Systems*. doi: 10.1016/j.future.2016.12.029.

Guizzi, G. L. and Manno, M. (2012) 'Fuel cell-based cogeneration system covering data centers' energy needs', *Energy*, 41(1), pp. 56–64. doi: 10.1016/j.energy.2011.07.030.

Guizzi, G., Manno, M. and Zaccagnini, A. (2009) 'Comparative analysis of combined cooling, heating and power systems (CCHP) covering data centers energy needs', *ECOS 2009 - 22nd International Conference on Efficiency, Cost, Optimization Simulation and Environmental Impact of Energy Systems*.

Güngör, V. C., Sahin, D., Kocak, T., Ergüt, S., Buccella, C., Cecati, C. and Hancke, G. P. (2011) 'Smart grid technologies: Communication technologies and standards', *IEEE Transactions on Industrial Informatics*, 7(4), pp. 529–539. doi: 10.1109/TII.2011.2166794.

Gupta, M. and Singh, S. (2003) 'Greening of the internet', *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications - SIGCOMM '03*, pp. 19–26. doi: 10.1145/863956.863959.

Gupta, M. and Singh, S. (2007) 'Using Low-Power Modes for Energy Conservation in Ethernet LANs', *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, pp. 2451–2455. doi: 10.1109/INFCOM.2007.299.

Hammadi, A. and Mhamdi, L. (2014) 'A survey on architectures and energy efficiency in Data Center Networks', *Computer Communications*, 40, pp. 1–21. doi: 10.1016/j.comcom.2013.11.005.

Härder, T., Hudlet, V., Ou, Y. and Schall, D. (2011) 'Energy Efficiency Is Not Enough, Energy Proportionality Is Needed!', *DASFAA Workshops*, pp. 226–239. doi: 10.1007/978-3-642-20244-5.

Van Heddeghem, W., Lambert, S., Lannoo, B., Colle, D., Pickavet, M. and Demeester, P. (2014) 'Trends in worldwide ICT electricity consumption from 2007 to 2012', *Computer Communications*. Elsevier B.V., 50, pp. 64–76. doi: 10.1016/j.comcom.2014.02.008.

Heller, B., Seetharaman, S., Mahadevan, P., Yiakoumis, Y., Sharma, P., Banerjee, S. and McKeown, N. (2010) 'ElasticTree : Saving Energy in Data Center Networks', *Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation*, pp. 17–17. doi: 10.1021/ci2004779.

Hermenier, F., Lorca, X., Menaud, J.-M., Muller, G. and Lawall, J. (2009) 'Entropy: a Consolidation Manager for Clusters', *Proceedings of the 2009 ACM SIGPLAN/SIGOPS international conference on Virtual execution environments - VEE '09*, p. 41. doi: 10.1145/1508293.1508300.

Horner, N. and Azevedo, I. (2016) ‘Power usage effectiveness in data centers: Overloaded and underachieving’, *Electricity Journal*, 29(4), pp. 61–69. doi: 10.1016/j.tej.2016.04.011.

Iec (2013) *Uninterruptible power systems (UPS) – Part 1 : General and safety requirements for UPS. International Electrotechnical Commission. Reference number. IEC 62040-1*.

International Energy Agency (2014) ‘ENERGY SUPPLY SECURITY 2014 Part 3’, *Energy Supply Security: The Emergency Response of IEA Countries - 2014 Edition*, pp. 1–105.

Iqbal, M., Azam, M., Naeem, M., Khwaja, A. S. and Anpalagan, A. (2014) ‘Optimization classification, algorithms and tools for renewable energy: A review’, *Renewable and Sustainable Energy Reviews*, pp. 640–654. doi: 10.1016/j.rser.2014.07.120.

Irwin, D., Sharma, N. and Shenoy, P. (2011) ‘Towards continuous policy-driven demand response in data centers’, *Proceedings of the 2nd ACM SIGCOMM workshop on Green networking - GreenNets ’11*, p. 19. doi: 10.1145/2018536.2018541.

Islam, M. A., Arafath, M. Y. and Hasan, M. J. (2015) ‘Design of DDR4 SDRAM controller’, *8th International Conference on Electrical and Computer Engineering: Advancing Technology for a Better Tomorrow, ICECE 2014*, pp. 148–151. doi: 10.1109/ICECE.2014.7026950.

JEDEC (2017) ‘JEDEC DDR5 & NVDIMM-P Standards Under Development’, *JEDEC’s Server Forum*.

Jeong, S. and Kim, Y. W. (2014) ‘A holistic investigation method for data center resource efficiency’, *International Conference on ICT Convergence*, pp. 548–551. doi: 10.1109/ICTC.2014.6983207.

Jiang, T., Yu, L. and Cao, Y. (2015) 'Energy Management of Internet Data Centers in Smart Grid'. Springer. doi: 10.1007/978-3-662-45676-7.

Jin, M., Feng, W., Liu, P., Marnay, C. and Spanos, C. (2017) ‘MOD-DR: Microgrid optimal dispatch with demand response’, *Applied Energy*, 187, pp. 758–776. doi: 10.1016/j.apenergy.2016.11.093.

Jin, M., Feng, W., Marnay, C. and Spanos, C. (2017) ‘Microgrid to enable optimal distributed energy retail and end-user demand response’, *Applied Energy*. doi: 10.1016/j.apenergy.2017.05.103.

Johannah, J. J., Korah, R., Kalavathy, M. and Sivanandham (2017) ‘Standby and dynamic power minimization using enhanced hybrid power gating structure for deep-submicron CMOS VLSI’, *Microelectronics Journal*, 62, pp. 137–145. doi: 10.1016/j.mejo.2017.02.003.

Josh, W. and Delforge, P. (2014) 'Data Center Efficiency Assessment: Scaling Up Energy Efficiency Across the Data Center Industry: Evaluating Key Drivers and Barriers'. *NRDC and Anthesis*.

Judge, J., Pouchet, J., Ekbote, A. and Dixit, S. (2008) ‘Reducing data center energy consumption’, *ASHRAE Journal*, 50(November).

Khan, M. U. S. and Khan, S. U. (2015) 'Smart Data Center', *Handbook on Data Centers*. New York, NY: Springer New York, pp. 247–262. doi: 10.1007/978-1-4939-2092-1_7.

Kim, J. (2016) *Strong,* 'Thorough, and Efficient Memory Protection against Existing and Emerging DRAM Errors'. University of Texas.

Kim, J., Sabry, M. M., Ruggiero, M. and Atienza, D. (2015) 'Power-Thermal Modeling and Control of Energy-Efficient Servers and Datacenters', *Handbook on Data Centers*. New York, NY: Springer New York, pp. 857–913. doi: 10.1007/978-1-4939-2092-1_29.

Klingert, S., Niedermeier, F., Dupont, C., Giuliani, G., Schulze, T. and de Meer, H. (2015) 'Renewable Energy-aware Data Centre Operations for Smart Cities - The DC4Cities Approach', *Proceedings of the 4th International Conference on Smart Cities and Green ICT Systems*. SCITEPRESS - Science and and Technology Publications, pp. 26–34. doi: 10.5220/0005430600260034.

Koomey (2011) 'Growth in Data Center Electricity use 2005 to 2010', *Oakland, CA: Analytics Press. August*. doi: 10.1088/1748-9326/3/3/034008.

Krzywda, J., Ali-Eldin, A., Carlson, T. E., Östberg, P.-O. and Elmroth, E. (2017) 'Power-performance tradeoffs in data center servers: DVFS, CPU pinning, horizontal, and vertical scaling', *Future Generation Computer Systems*. doi: 10.1016/j.future.2017.10.044.

Lee, K. P. and Chen, H. L. (2013) 'Analysis of energy saving potential of air-side free cooling for data centers in worldwide climate zones', *Energy and Buildings*, 64, pp. 103–112. doi: 10.1016/j.enbuild.2013.04.013.

Lei, H., Wang, R., Zhang, T., Liu, Y. and Zha, Y. (2016) 'A multi-objective co-evolutionary algorithm for energy-efficient scheduling on a green data center', *Computers and Operations Research*, 75, pp. 103–117. doi: 10.1016/j.cor.2016.05.014.

Li, J., Bao, Z. and Li, Z. (2015) 'Modeling demand response capability by internet data centers processing batch computing jobs', *IEEE Transactions on Smart Grid*, 6(2), pp. 737–747. doi: 10.1109/TSG.2014.2363583.

Lim, S.-H., Huh, J.-S., Kim, Y. and Das, C. R. (2011) 'Migration, Assignment, and Scheduling of Jobs in Virtualized Environment', *HotCloud 2011*, pp. 1–5.

Little, A. B. and Garimella, S. (2012) 'Waste Heat Recovery in Data Centers Using Sorption Systems', *Journal of Thermal Science and Engineering Applications*, 4(2), p. 21007. doi: 10.1115/1.4005813.

Liu, Z., Liu, I., Low, S. and Wierman, A. (2014) 'Pricing Data Center Demand Response', *Acm Sigmetrics*, pp. 111–123. doi: 10.1145/2591971.2592004.

Liu, Z., Wierman, A., Chen, Y., Razon, B. and Chen, N. (2013) 'Data center demand response: Avoiding the coincident peak via workload shifting and local generation', *Performance Evaluation*. Elsevier B.V., 70(10), pp. 770–791. doi: 10.1016/j.peva.2013.08.014.

Lu, T., Pande, P. P. and Shirazi, B. (2016) 'A dynamic, compiler guided DVFS mechanism to achieve energy-efficiency in multi-core processors', *Sustainable Computing: Informatics and Systems*, 12, pp. 1–9. doi: 10.1016/j.suscom.2016.04.003.

Malkamäki, T. and Ovaska, S. J. (2012) 'Solar energy and free cooling potential in European data centers', in *Procedia Computer Science*, pp. 1004–1009. doi: 10.1016/j.procs.2012.06.138.

Mämmelä, O., Majanen, M., Basmadjian, R., De Meer, H., Giesler, A. and Homberg, W. (2012) 'Energy-aware job scheduler for high-performance computing', *Computer Science - Research and Development*, 27(4), pp. 265–275. doi: 10.1007/s00450-011-0189-6.

Masanet, E. R., Brown, R. E., Shehabi, A., Koomey, J. G. and Nordman, B. (2011) 'Estimating the energy use and efficiency potential of U.S. data centers', *Proceedings of the IEEE*, pp. 1440–1453. doi: 10.1109/JPROC.2011.2155610.

Masanet, E. and Robert, H. (2014) 'Data Center IT Efficiency Measures', *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*, (January 2015).

Mäsker, M., Nagel, L., Brinkmann, A., Lotfifar, F. and Johnson, M. (2016) 'Smart Grid-aware scheduling in data centres', *Computer Communications*, 96, pp. 73–85. doi: 10.1016/j.comcom.2016.04.021.

Mazumdar, S. and Pranzo, M. (2017) 'Power efficient server consolidation for Cloud data center', *Future Generation Computer Systems*, 70, pp. 4–16. doi: 10.1016/j.future.2016.12.022.

'Microsoft data centre in Wyoming to use FuelCell Energy unit' (2012) *Fuel Cells Bulletin*, 2012(12), p. 5. doi: 10.1016/S1464-2859(12)70348-6.

Moreno-Munoz, A., De La Rosa, J. J. G., Pallarés-Lopez, V., Real-Calvo, R. J. and Gil-De-Castro, A. (2011) 'Distributed DC-UPS for energy smart buildings', *Energy and Buildings*, 43(1), pp. 93–100. doi: 10.1016/j.enbuild.2010.08.018.

Moura, P., Nuttall, C., Harrison, B., Jehle, C. and de Almeida, A. (2016) 'Energy savings potential of uninterruptible power supplies in European Union', *Energy Efficiency*, 9(5), pp. 993–1013. doi: 10.1007/s12053-015-9406-7.

Mukherjee, T., Banerjee, A., Varsamopoulos, G. and Gupta, S. K. S. (2010) 'Model-driven coordinated management of data centers', *Computer Networks*, 54(16), pp. 2869–2886. doi: 10.1016/j.comnet.2010.08.011.

Nada, N. and Elgelany, A. (2014) ‘Green Technology, Cloud Computing and Data Centers: the Need for Integrated Energy Efficiency Framework and Effective Metric’, *International Journal of Advanced Computer Science and Applications*, 5(5), pp. 89–93. doi: 10.14569/IJACSA.2014.050513.

Nadjaran Toosi, A., Qu, C., de Assunção, M. D. and Buyya, R. (2017) ‘Renewable-aware geographical load balancing of web applications for sustainable data centers’, *Journal of Network and Computer Applications*, 83, pp. 155–168. doi: 10.1016/j.jnca.2017.01.036.

Nathuji, R. and Schwan, K. (2007) ‘VirtualPower: coordinated power management in virtualized enterprise systems’, *Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles - SOSP ’07*, p. 265. doi: 10.1145/1294261.1294287.

Nedevschi, S., Popa, L., Iannaccone, G., Ratnasamy, S. and Wetherall, D. (2008) ‘Reducing network energy consumption via sleeping and rate-adaptation’, *Symposium A Quarterly Journal In Modern Foreign Literatures*, 21(3), pp. 323–336. doi: 10.1.1.143.3471.

Ni, J. and Bai, X. (2017) ‘A review of air conditioning energy performance in data centers’, *Renewable and Sustainable Energy Reviews*, pp. 625–640. doi: 10.1016/j.rser.2016.09.050.

Oconnell, N., Pinson, P., Madsen, H. and Omalley, M. (2014) ‘Benefits and challenges of electrical demand response: A critical review’, *Renewable and Sustainable Energy Reviews*, pp. 686–699. doi: 10.1016/j.rser.2014.07.098.

Oró, E., Depoorter, V., Garcia, A. and Salom, J. (2015) ‘Energy efficiency and renewable energy integration in data centres. Strategies and modelling review’, *Renewable and Sustainable Energy Reviews*, pp. 429–445. doi: 10.1016/j.rser.2014.10.035.

Padala, P., Hou, K.-Y., Shin, K. G., Zhu, X., Uysal, M., Wang, Z., Singhal, S. and Merchant, A. (2009) ‘Automated control of multiple virtualized resources’, *Proceedings of the 4th ACM European conference on Computer systems*, pp. 13–26. doi: 10.1145/1519065.1519068.

Panajotovic, B., Jankovic, M. and Odadzic, B. (2011) ‘ICT and smart grid’, *2011 10th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services, TELSIKS 2011 - Proceedings of Papers*, pp. 118–121. doi: 10.1109/TELSKS.2011.6112018.

Parra, D., Swierczynski, M., Stroe, D. I., Norman, S. A., Abdon, A., Worlitschek, J., O?Doherty, T., Rodrigues, L., Gillott, M., Zhang, X., Bauer, C. and Patel, M. K. (2017) ‘An interdisciplinary review of energy storage for communities: Challenges and perspectives’, *Renewable and Sustainable Energy Reviews*, 79, pp. 730–749. doi: 10.1016/j.rser.2017.05.003.

Paterakis, N. G., Erdinç, O. and Catalão, J. P. S. (2017) ‘An overview of Demand Response: Key-elements and international experience’, *Renewable and Sustainable Energy Reviews*, pp. 871–891. doi:

10.1016/j.rser.2016.11.167.

Paul, D., Zhong, W. De and Bose, S. K. (2017) 'Demand Response in Data Centers Through Energy-Efficient Scheduling and Simple Incentivization', *IEEE Systems Journal*, 11(2), pp. 613–624. doi: 10.1109/JSYST.2015.2476357.

Pier (2008) 'Uninterruptible Power Supplies : A Data Center Efficiency Opportunity. Technical brief. California Energy Commission's Public Interest Energy Research (PIER) Program'.

Pierson, J.-M. (ed.) (2015) *Large-Scale Distributed Systems and Energy Efficiency*. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/9781118981122.

PJM (2013) 'Retail Electricity Consumer Opportunities for Demand Response in PJM's Wholesale Markets', pp. 1–3.

Pore, M., Abbasi, Z., Gupta, S. K. S. and Varsamopoulos, G. (2015) 'Techniques to Achieve Energy Proportionality in Data Centers: A Survey', *Handbook on Data Centers*. New York, NY: Springer New York, pp. 109–162. doi: 10.1007/978-1-4939-2092-1_4.

Raghavendra, R., Ranganathan, P., Talwar, V., Wang, Z. and Zhu, X. (2008) 'No " Power " Struggles : Coordinated Multi-level Power Management for the Data Center', *Solutions*, 36, pp. 48–59. doi: http://doi.acm.org/10.1145/1346281.1346289.

Reiss, C., Tumanov, A. and Ganger, G. (2012) 'Towards understanding heterogeneous clouds at scale: Google trace analysis', *Intel science & technology center for cloud computing*.

Ristic, B., Madani, K. and Makuch, Z. (2015) 'The Water Footprint of Data Centers', *Sustainability*, 7(8), pp. 11260–11284. doi: 10.3390/su70811260.

Rong, H., Zhang, H., Xiao, S., Li, C. and Hu, C. (2016) 'Optimizing energy consumption for data centers', *Renewable and Sustainable Energy Reviews*. Elsevier, 58, pp. 674–691. doi: 10.1016/j.rser.2015.12.283.

Salom, J., Urbaneck, T. and Oró, E. (2017) *Advanced Concepts for Renewable Energy Supply of Data Centres*. River Publishers. doi: 10.13052/rp-9788793519411.

Sanders, P., Sivadasan, N. and Skutella, M. (2004) 'Online scheduling with bounded migration', *Automata , Languages and Programming, Proceedings*, pp. 1111–1122. doi: 10.1287/moor.1090.0381.

Schrijver, A. (1986) 'Theory of linear and integer programming', *Wiley-Interscience series in discrete mathematics and optimization*, p. xi, 471 . doi: 10.1016/0378-4754(87)90121-2.

Shah, A. J. and Krishnan, N. (2008) 'Optimization of Global Data Center Thermal Management Workload for Minimal Environmental and Economic Burden', *IEEE Transactions on Components and Packaging Technologies*, 31(1), pp. 39–45. doi: 10.1109/TCAPT.2007.906721.

Sharma, N., Barker, S., Irwin, D. and Shenoy, P. (2011) ‘Blink: managing server clusters on intermittent power’, *ACM SIGPLAN Notices*, 46(3), p. 185. doi: 10.1145/1961296.1950389.

Shehabi, A., Horvath, A., Tschudi, W., Gadgil, A. J. and Nazaroff, W. W. (2008) ‘Particle concentrations in data centers’, *Atmospheric Environment*, 42(24), pp. 5978–5990. doi: 10.1016/j.atmosenv.2008.03.049.

Shehabi, A., Smith, S. J., Sartor, D. A., Brown, R. E., Herrlin, M., Koomey, J. G., Masanet, E. R., Horner, N., Azevedo, I. L. and Lintner, W. (2016) ‘United States Data Center Energy Usage Report’, *Ernest Orlando Lawrence Barkeley National Laboratory*, (June). doi: LBNL-1005775.

Shehabi, A., Tschudi, W. and Gagdil, A. (2007) 'Data Center Economizer Contamination and Humidity Study'*. Lawrence Berkeley National Laboratory*.

Sheppy, M., Lobato, C., Geet, O. Van, Pless, S., Donovan, K. and Powers Chuck (2011) 'Reducing Data Center Loads for a Large- scale , Low-energy Office Building : NREL ’ s Research Support Facility'. *Golden, Colorado.*

Shuja, J., Gani, A., Shamshirband, S., Ahmad, R. W. and Bilal, K. (2016) ‘Sustainable Cloud Data Centers: A survey of enabling techniques and technologies’, *Renewable and Sustainable Energy Reviews*, 62, pp. 195–214. doi: 10.1016/j.rser.2016.04.034.

SIA (2015) ‘Rebooting the IT Revolution: A Call to Action. Semiconductor Industry Association and the Semiconductor Research Corporation’, p. 40.

Sierksma, G. (2001) *Linear and Integer Programming: Theory and Practice*. Boca Raton: CRC Press.

Siriwardana, J., Jayasekara, S. and Halgamuge, S. K. (2013) ‘Potential of air-side economizers for data center cooling: A case study for key Australian cities’, *Applied Energy*, 104, pp. 207–219. doi: 10.1016/j.apenergy.2012.10.046.

Sithimolada, V. and Sauer, P. W. (2010) ‘Facility-level DC vs. typical AC distribution for data centers: A comparative reliability study’, *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, pp. 2102–2107. doi: 10.1109/TENCON.2010.5686625.

Sohel Murshed, S. M. and Nieto de Castro, C. A. (2017) ‘A critical review of traditional and emerging techniques and fluids for electronics cooling’, *Renewable and Sustainable Energy Reviews*, 78, pp. 821–833. doi: 10.1016/j.rser.2017.04.112.

Spec (2018) *SPECpower_ssj2008 benchmark. http://www.spec.org/power_ssj2008/results/. Accessed: 14-01-2016*.

Spring, N., Peterson, L., Bavier, A. and Pai, V. (2006) ‘Using PlanetLab for Network Research: Myths, Realities, and Best Practices’, *ACM SIGOPS Operating Systems Review*, 40, pp. 17–24. doi: http://doi.acm.org/10.1145/1113361.1113368.

Stewart, C. and Shen, K. (2009) ‘Some joules are more precious than others: Managing renewable energy in the datacenter’, *Proceedings of the Workshop on Power Aware Computing and Systems*.

Sueur, E. L. and Heiser, G. (2010) ‘Dynamic voltage and frequency scaling: the laws of diminishing returns’, *Proceedings of the 2010 international conference on Power aware computing and systems*, pp. 1–8.

Tang, C.-J., Dai, M.-R., Chuang, C.-C., Chiu, Y.-S. and Lin, W. S. (2014) ‘A load control method for small data centers participating in demand response programs’, *Future Generation Computer Systems*. Elsevier B.V., 32, pp. 232–245. doi: 10.1016/j.future.2013.07.020.

Tang, C., Dai, M. and Chuang, C.-C. (2012) ‘Demand Response Control Strategies for On-campus Small Data Centers’, *Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), 2012 9th International Conference on*, pp. 217–224. doi: 10.1109/UIC-ATC.2012.97.

Tang, Y., Sun, H., Wang, X. and Liu, X. (2017) ‘Achieving convergent causal consistency and high availability for cloud storage’, *Future Generation Computer Systems*, 74, pp. 20–31. doi: 10.1016/j.future.2017.04.016.

Tesauro, G., Jong, N. K., Das, R. and Bennani, M. N. (2006) ‘A hybrid reinforcement learning approach to autonomic resource allocation’, *Proceedings of the IEEE International Conference on Autonomic Computing, 2006. ICAC’06.*, pp. 65–73. doi: 10.1109/ICAC.2006.1662383.

The Green Grid (2012) *Harmonzing Global Metrics for Data Centers Energy Efficiency*, *Journal of Chemical Information and Modeling*. doi: 10.1017/CBO9781107415324.004.

The Green Grid (2016) *‘The Green Grid’. http://www.thegreengrid.org. Accessed: 14-01-2016*.

Tran, N., Ren, S., Han, Z., Man Jang, S., Il Moon, S. and Seon Hong, C. (2014) ‘Demand Response of Data Centers: A Real-time Pricing Game between Utilities in Smart Grid’, *9th International Workshop on Feedback Computing*.

Tsirogiannis, D., Harizopoulos, S. and Shah, M. a (2010) ‘Analyzing the Energy Efficiency of a Database Server’, *the 2010 International Conference*, p. 231. doi: 10.1145/1807167.1807194.

Uddin, M., Alsaqour, R., Shah, A. and Saba, T. (2014) ‘Power usage effectiveness metrics to measure efficiency and performance of data centers’, *Applied Mathematics and Information Sciences*, 8(5), pp. 2207–2216. doi: 10.12785/amis/080514.

Uddin, M., Darabidarabkhani, Y., Shah, A. and Memon, J. (2015) ‘Evaluating power efficient algorithms for efficiency and carbon emissions in cloud data centers: A review’, *Renewable and Sustainable Energy Reviews*, pp. 1553–1563. doi: 10.1016/j.rser.2015.07.061.

Uddin, M. and Rahman, A. A. (2012) ‘Energy efficiency and low carbon enabler green IT framework for data centers considering green metrics’, *Renewable and Sustainable Energy Reviews*, pp. 4078–4094. doi: 10.1016/j.rser.2012.03.014.

Urgaonkar, B., Shenoy, P., Chandra, A., Goyal, P. and Wood, T. (2008) ‘Agile dynamic provisioning of multi-tier Internet applications’, *ACM Transactions on Autonomous and Adaptive Systems*, 3(1), pp. 1–39. doi: 10.1145/1342171.1342172.

Vanson-Bourne (2011) 'V-Index: Virtualization Penetration Rate in the Enterprise'. Available at: http://www.v-index.com/full-report.html.

Varrette, S., Bouvry, P., Jarus, M. and Oleksiak, A. (2015) ‘Energy Efficiency in HPC Data Centers: Latest Advances to Build the Path to Exascale’, *Handbook on Data Centers*. New York, NY: Springer New York, pp. 81–107. doi: 10.1007/978-1-4939-2092-1_3.

Verma, A., Koller, R., Useche, L. and Rangaswami, R. (2010) ‘SRCMap: energy proportional storage using dynamic consolidation’, *FAST’10 Proceedings of the 8th USENIX conference on File and storage technologies*, (VM), p. 20.

Vertiv (2017) 'Energy Efficiency Solutions for Your Data Center'.

vor dem Berge, M., Buchholz, J., Cupertino, L., Da Costa, G., Donoghue, A., Gallizo, G., Jarus, M., Lopez, L., Oleksiak, A., Pages, E., Piątek, W., Pierson, J.-M., Piontek, T., Rathgeb, D., Salom, J., Sisó, L., Volk, E., Wössner, U. and Zilio, T. (2015) ‘CoolEmAll: Models and Tools for Planning and Operating Energy Efficient Data Centres’, *Handbook on Data Centers*. New York, NY: Springer New York, pp. 191–245. doi: 10.1007/978-1-4939-2092-1_6.

Wahlroos, M., Pärssinen, M., Manner, J. and Syri, S. (2017) ‘Utilizing data center waste heat in district heating – Impacts on energy efficiency and prospects for low-temperature district heating networks’, *Energy*, 140, pp. 1228–1238. doi: 10.1016/j.energy.2017.08.078.

Wang, B., Chen, C., He, L., Gao, B., Ren, J., Fu, Z., Fu, S., Hu, Y. and Li, C. T. (2017) ‘Modelling and developing conflict-aware scheduling on large-scale data centres’, *Future Generation Computer Systems*. doi: 10.1016/j.future.2017.07.043.

Wang, C. and De Groot, M. (2013) ‘Enabling demand response in a computer cluster’, *2013 IEEE International Conference on Smart Grid Communications, SmartGridComm 2013*, pp. 181–186. doi: 10.1109/SmartGridComm.2013.6687954.

Wang, C., Urgaonkar, B., Wang, Q. and Kesidis, G. (2013) 'A hierarchical demand response framework for data center power cost optimization under real-world electricity pricing', pp. 1–10.

Wang, D., Ren, C., Sivasubramaniam, A., Urgaonkar, B. and Fathy, H. (2012) 'Energy storage in datacenters: what, where and how much?', *ACM SIGMETRICS Performance Evaluation Review*, 40(1), p. 187. doi: 10.1145/2318857.2254780.

Wang, H., Huang, J., Lin, X. and Mohsenian-Rad, H. (2016) 'Proactive Demand Response for Data Centers: A Win-Win Solution', *IEEE Transactions on Smart Grid*, 7(3), pp. 1584–1596. doi: 10.1109/TSG.2015.2501808.

Wang, P., Huang, J. Y., Ding, Y., Loh, P. and Goel, L. (2011) 'Demand side load management of smart grids using intelligent trading/metering/ billing system', *2011 IEEE Trondheim PowerTech*, pp. 1–6. doi: 10.1109/PTC.2011.6019420.

Wang, X. and Wang, Y. (2011) 'Coordinating power control and performance management for virtualized server clusters', *IEEE Transactions on Parallel and Distributed Systems*, 22(2), pp. 245–259. doi: 10.1109/TPDS.2010.91.

Wang, Y., Brun, E., Malvagi, F. and Calvin, C. (2017) 'Competing energy lookup algorithms in Monte Carlo neutron transport calculations and their optimization on CPU and Intel MIC architectures', *Journal of Computational Science*. doi: 10.1016/j.jocs.2017.01.006.

Whitehead, B., Andrews, D., Shah, A. and Maidment, G. (2014) 'Assessing the environmental impact of data centres part 1: Background, energy use and metrics', *Building and Environment*, 82, pp. 151–159. doi: 10.1016/j.buildenv.2014.08.021.

Whitehead, B., Andrews, D., Shah, A. and Maidment, G. (2015) 'Assessing the environmental impact of data centres part 2: Building environmental assessment methods and life cycle assessment', *Building and Environment*, pp. 395–405. doi: 10.1016/j.buildenv.2014.08.015.

Whitney, J. and Delforge, P. (2014) 'Data Center Efficiency Assessment Scaling Up Energy Efficiency Across the Data Center Industry : Evaluating Key Drivers and Barriers', (August).

Wiboonrat, M. (2012) 'Next Generation Data Center design under Smart Grid', *Ubiquitous and Future Networks (ICUFN), 2012 Fourth International Conference on*, pp. 103–108. doi: 10.1109/ICUFN.2012.6261673.

Wierman, A., Liu, Z., Liu, I. and Mohsenian-Rad, H. (2014) 'Opportunities and challenges for data center demand response', *International Green Computing Conference*. IEEE, pp. 1–10. doi: 10.1109/IGCC.2014.7039172.

Witkowski, M., Oleksiak, A., Piontek, T. and Węglarz, J. (2013) 'Practical power consumption estimation for real life HPC applications', *Future Generation Computer Systems*, 29(1), pp. 208–217. doi: 10.1016/j.future.2012.06.003.

Wright, S. (2010) 'Continuous Optimization (Nonlinear and Linear Programming)', *Foundations of Computer-Aided Process Design*, pp. 1–14.

Wrinch, M., Dennis, G., El-fouly, T. H. M. and Wong, S. (2012) 'Demand Response Implementation for Improved System Efficiency in Remote Communities Pilot Results from the Village of Hartley Bay', *IEEE annual Electrical Power and Energy Conference (EPEC 2012)*, pp. 105–110.

Yeo, S. and Lee, H.-H. S. (2012) 'SimWare: A Holistic Warehouse-Scale Computer Simulator', *Computer*, 45(9), pp. 48–55. doi: 10.1109/MC.2012.251.

Zakarya, M. and Gillam, L. (2017) 'Energy efficient computing, clusters, grids and clouds: A taxonomy and survey', *Sustainable Computing: Informatics and Systems*, 14, pp. 13–33. doi: 10.1016/j.suscom.2017.03.002.

Zhan, X. and Reda, S. (2013) 'Techniques for energy-efficient power budgeting in data centers', *Proceedings of the 50th Annual Design Automation Conference on - DAC '13*, p. 1. doi: 10.1145/2463209.2488951.

Zhang, H., Shao, S., Xu, H., Zou, H. and Tian, C. (2014) 'Free cooling of data centers: A review', *Renewable and Sustainable Energy Reviews*, 35, pp. 171–182. doi: 10.1016/j.rser.2014.04.017.

Zhu, K., Cui, Z., Wang, Y., Li, H., Zhang, X. and Franke, C. (2017) 'Estimating the maximum energy-saving potential based on IT load and IT load shifting', *Energy*, 138, pp. 902–909. doi: 10.1016/j.energy.2017.07.092.

Zhu, W., Garrett, D., Butkowski, J. and Wang, Y. (2012) 'Overview of distributive energy storage systems for residential communities', *2012 IEEE Energytech, Energytech 2012*, 1, pp. 1–6. doi: 10.1109/EnergyTech.2012.6304671.

Zhuravlev, S., Saez, J. C., Blagodurov, S., Fedorova, A. and Prieto, M. (2013) 'Survey of Energy-Cognizant Scheduling Techniques', *Parallel and Distributed Systems, IEEE Transactions on*, 24(7), pp. 1447–1464. doi: 10.1109/TPDS.2012.20.

ANNEX I: ALGORITHMS

I.1 SMDC

I.1.1 The Best Time Window

```
%% Data capture from Excel table dcdata.xlsx
filename = 'dcdata.xlsx';
sheetDC1 = 1;
sheetDC2 = 2;
smdc_num1 = xlsread(filename,sheetDC1);
smdc_num2 = xlsread(filename,sheetDC2);
%% Variables declaration with the values present in the table
price24 = smdc_num1(:,4); % daily tariff variable by hour
price72 = smdc_num2(:,1); % daily tariff variable by a third of an hour
%% Formulation of objective function of minimization to small data center
% f = (x1 + x2 +x3)*p
f = [price24;price24;price24];
intcon = 1:24;
%% Formulation of constraints
% x1 + x2 + x3 ≤ -150 (Demand Response)
% -x1 -x2 -x3 ≤ 150 (Rebound Effect)
vet3 = ones(1,24);
vet4 = zeros(1,24);
Aeq1 = [vet3 vet4 vet4];
Aeq2 = [vet4 vet3 vet4];
Aeq3 = [vet4 vet4 vet3];
Aeq = [Aeq1; Aeq2; Aeq3];
beq = [0; 0; 0];
%% Boundaries and division of power by demand response strategies
% lb = -105
% ub = 105
lb1(1:24,1) = -28; % workload demand response
lb2(1:24,1) = -42; % cooling demand response
lb3(1:24,1) = -35; % ups demand response
ub1(1:24,1) = 28; % workload rebound effect
ub2(1:24,1) = 42; % cooling rebound effect
ub3(1:24,1) = 35; % ups rebound effect
lb = [lb1;lb2;lb3];
ub = [ub1;ub2;ub3];
%% Optimization running
[x,fval] = intlinprog(f, intcon,[],[],Aeq,beq,lb,ub);
%% Charts
x1 = x(1:24);
x2 = x(25:48);
x3 = x(49:72);
figure('Name','Small Data Center');
subplot (2,1,1);
plot(x1,'LineWidth',1);
hold on;
plot (x2,'LineWidth',1);
plot (x3,'LineWidth',1);
grid on;
title('Daily Diagram of Demand Response and Rebound Effect');
xlabel('Time (hours)');
ylabel('Power (kW)');
```

```
xticks([1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24]);
legend('ICT Workload', 'Cooling', 'UPS');
hold off
subplot(2,1,2);
plot(f(1:24),'LineWidth',1);
grid on;
title('Daily Tariff');
xlabel('Time (hours)');
ylabel('Cost (€/kWh)');
xticks([1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24]);
```

I.1.2 Cost Minimization

```
%% Data capture from Excel table dcdata.xlsx
filename = 'dcdata.xlsx';
sheetDC1 = 1;
sheetDC2 = 2;
%% Data capture from Excel table dcdata.xlsx
filename = 'dcdata.xlsx';
sheetDC1 = 1;
sheetDC2 = 2;
indataM = 4;
smdc_num1 = xlsread(filename,sheetDC1);
smdc_num2 = xlsread(filename,sheetDC2);
smdc_num4 = xlsread(filename,indataM);
%% Input
in = input('Which line do you want to record your data on? ');
my_cell1 = sprintf( 'B%s',num2str(in) );
my_cell2 = sprintf( 'J%s',num2str(in) );
on = in-1;
%% Variables declaration with the values present in the table
price24 = smdc_num1(:,4); % daily tariff variable by hour
price72 = smdc_num2(:,1); % daily tariff variable by a third of an hour
% input equations data
pdrwork = smdc_num4(on,1); % workload demand response equation
pdrcool = round
(smdc_num4(on,3)*(smdc_num4(on,4)/(smdc_num4(on,5)*(smdc_num4(on,6)+smdc_nu
m4(on,7)-smdc_num4(on,8)))+smdc_num4(on,9))); % cooling demand response
equation
% ups demand response equation
if (smdc_num4(on,12) <= smdc_num4(on,14))&&(smdc_num4(on,12) >=
smdc_num4(on,14)*0.5)
   pdrups = round (smdc_num4(on,11)*(-(smdc_num4(on,12)-
smdc_num4(on,13))/(smdc_num4(on,15)*smdc_num4(on,16))));
elseif smdc_num4(on,12) > smdc_num4(on,14)
       smdc_num4(on,12) = smdc_num4(on,14)
     pdrups = round (smdc_num4(on,11)*(-(smdc_num4(on,12)-
smdc_num4(on,13))/(smdc_num4(on,15)*smdc_num4(on,16))));
elseif smdc_num4(on,12) < smdc_num4(on,14)*0.5
       smdc_num4(on,12) = smdc_num4(on,14)*0.5
      pdrups = round (smdc_num4(on,11)*(-(smdc_num4(on,12)-
smdc_num4(on,13))/(smdc_num4(on,15)*smdc_num4(on,16))));
end
xlswrite('dcdata.xlsx',pdrcool,'indataM',my_cell1); % writing pdrcool data
on Excel table
xlswrite('dcdata.xlsx',pdrups,'indataM',my_cell2); % writing pdrups data on
Excel table

%% Formulation of objective function of minimization to medium data center
% f = (x4 + x5 +x6)*p
f = [price72;price72;price72];
intcon = 1:72;
```

```
%% Formulation of constraints
% x1 + x2 + x3 ≤ -550 (demand response 10:40h)
% -x1 -x2 -x3 ≤ 550 (rebound effect 11h-13h)
vet3 = ones(1,72);
vet4 = zeros(1,72);
DRtime = 33; % DR
Aeq1 = [vet3 vet4 vet4];
Aeq2 = [vet4 vet3 vet4];
Aeq3 = [vet4 vet4 vet3];
Aeq = [Aeq1; Aeq2; Aeq3];
beq = [0; 0; 0];
%% Boundaries and division of power by demand response strategies
lb1(1:72,1) = 0; % workload demand response
lb2(1:72,1) = 0; % cooling demand response
lb3(1:72,1) = 0; % ups demand response
lb1(DRtime,1) = -pdrwork; % workload demand response
lb2(DRtime,1) = -pdrcool; % cooling demand response
lb3(DRtime,1) = -pdrups; % ups demand response
ub1(1:72,1) = 0; % workload demand response
ub2(1:72,1) = 0; % cooling demand response
ub3(1:72,1) = 0; % ups demand response
ub1(34:35,1) = pdrwork/2; % workload demand response
ub2(36:37,1) = pdrcool/2; % cooling demand response
ub3(38:39,1) = pdrups/2; % ups demand response
lb = [lb1;lb2;lb3];
ub = [ub1;ub2;ub3];
%% Optimization running
[x,fval] = intlinprog(f,intcon,[],[],Aeq,beq,lb,ub);
%% Charts
x1 = x(1:72);
x2 = x(73:144);
x3 = x(145:216);
figure('Name','Medium Data Center');
subplot (2,1,1);
plot(x1,'LineWidth',1);
hold on;
plot (x2,'LineWidth',1);
plot (x3,'LineWidth',1);
grid on;
title('Daily Diagram of Demand Response and Rebound Effect');
xlabel('Time (hours)');
ylabel('Power (kW)');
xtickangle(60);
legend('ICT Workload', 'Cooling', 'UPS');
hold off
subplot(2,1,2);
plot(f(1:72),'LineWidth',1);
grid on;
title('Daily Tariff');
xlabel('Time (hours)');
ylabel('Cost (€/kWh)');
xtickangle(60);
```

1.1.3 DR and RE in Daily Load Diagram

```
%% Data capture from Excel table dcdata.xlsx
filename = 'dcdata.xlsx';
% sheetDC24 = 1;
sheetDC72 = 2;
% smdc_num24 = xlsread(filename,sheetDC24);
smdc_num72 = xlsread(filename,sheetDC72);
```

```
%% Variables declaration with the values present in the table
pt_s = smdc_num72(:,2); % SDC power
pt_m = smdc_num72(:,3); % MDC power
% preco24 = smdc_num24(:,4); % 24h tariff variable
price72 = smdc_num72(:,4); % 24h*3 tariff variable
price_avg = mean(price72); %  average price
price_avg1(1:32,:) = price_avg;
price_avg2 = price72(33:39);
price_avg3(1:33,:) = price_avg;
price_dr = [price_avg1;price_avg2;price_avg3]; % DR price
dr_s = smdc_num72(:,5); % DRE in SDC
dr_m = smdc_num72(:,6); % DRE in MDC
rb_s = smdc_num72(:,7); % RE in SDC
rb_m = smdc_num72(:,8); % RE in MDC
pdr_s = pt_s-dr_s; % decreased power in SDC DRE
pdr_m = pt_m-dr_m; % decreased power in MDC DRE
prb_s = pdr_s+rb_s; % increased power in SDC RE
prb_m = pdr_m+rb_m; % increased power in MDC RE
t = 1:1:72; % 24h/3 time variable
smalldc; % call smalldc function
mediumdc; % call mediumdc function
%% Interactions and Results
disp('----- The Demand Response event will occur between 11:40 p.m. and
12:00 p.m. -----');
disp('----- O Rebound Effect will occur between 12h e 14h -----');
fprintf('The kWh average price is: %.4f €\n',price_avg);
fprintf('The kWh price during the Demand Response event is: %.4f
€\n',price_dr(33));
N_s = input('How many Small Data Centers will participate in the event: ');
N_m = input('How many Medium Data Centers will participate in the event:
');
c1 = sum(pt_s.*price_avg);
c2 = sum(pt_m.*price_avg);
c3 = N_s.*sum(pt_s.*price_avg);
c4 = N_m.*sum(pt_m.*price_avg);
c5 = c3+c4;
c6 = sum(pt_s.*price_dr);
c7 = sum(pt_m.*price_dr);
c8 = N_s.*sum(pt_s.*price_dr);
c9 = N_m.*sum(pt_m.*price_dr);
c10 = c8+c9;
fprintf('The decreased power by Small Data Centers was: %.f kW\n',N_s*(-
b2_s));
fprintf('The decreased power by Medium Data Centers was: %.f kW\n',N_m*(-
b2_m));
fprintf('The total decreased power during the event was: %.f kW\n',N_s*(-
b2_s)+N_m*(-b2_m));
fprintf('The cost to the DSO of Demand Response considering Rebound Effect
is: %.2f ,Ç¨\n',(-fval_s*N_s)+(-fval_m*N_m));
fprintf('The average operational cost by Small Data Centers is: %.2f
,Ç¨\n',c1);
fprintf('The average operational cost by Medium Data Centers is: %.2f
,Ç¨\n',c2);
fprintf('The average operational cost by all Small Data Centers is: %.2f
,Ç¨\n',c3);
fprintf('The average operational cost by all Medium Data Centers is: %.2f
,Ç¨\n',c4);
fprintf('The total average operational cost is: %.2f €\n',c5);
fprintf('The operational cost with Demand Response by Small Data Centers
is: %.2f €\n',c6);
fprintf('The operational cost with Demand Response by Medium Data Centers
is: %.2f €\n',c7);
```

```
fprintf('The operational cost with Demand Response by all Small Data
Centers is: %.2f €\n',c8);
fprintf('The operational cost with Demand Response by all Medium Data
Centers is: %.2f €\n',c9);
fprintf('The total average operational cost with Demand Response is: %.2f
€\n',c10);
%% Charts
x1 = x_s(1:72);
x2 = x_s(73:144);
x3 = x_s(145:216);
x4 = x_m(1:72);
x5 = x_m(73:144);
x6 = x_m(145:216);
figure('Name','Demand Response e Rebound Effect');
subplot (3,1,1);
plot(f(1:72),'LineWidth',1);
grid on;
title('Daily Tariff');
xlabel('Time (hours)');
ylabel('Cost (€/kWh)');
xtickangle(60);
subplot (3,1,2);
plot(x1,'LineWidth',1);
hold on;
plot (x2,'LineWidth',1);
plot (x3,'LineWidth',1);
grid on;
title('Daily Diagram of Demand Response (DR) and Rebound Effect (RE) in
Small Data Center');
xlabel('Time (hours)');
ylabel('Power (kW)');
ylim([-80 80]);
xtickangle(60);
legend('ICT Workload', 'Cooling', 'UPS');
subplot(3,1,3);
plot(x4,'LineWidth',1);
hold on;
plot (x5,'LineWidth',1);
plot (x6,'LineWidth',1);
grid on;
title('Daily Diagram of Demand Response (DR) and Rebound Effect (RE) in
Medium Data Center');
xlabel('Time (hours)');
ylabel('Power (kW)');
xtickangle(60);
legend('ICT Workload', 'Cooling', 'UPS');
hold off
figure('Name','Load Diagram and Related Costs');
subplot(2,1,1);
plot(t,pt_s,t,pt_m,t,pdr_s,'--',t,pdr_m,'--
',t,prb_s,':',t,prb_m,':','LineWidth',1);
grid on
title('Daily Load Diagram');
xlabel('Time (hours)');
ylabel('Power (kW)');
legend('Small Data Centers', 'Medium Data Centers', 'Small Data Centers
with DR', 'Medium Data Centers with DR', 'Small Data Centers com RE',
'Medium Data Centers with RE','Location','northeastoutside');
subplot(2,1,2);
CT = [c1 c2 c3 c4 c5; c6 c7 c8 c9 c10];
c = categorical({'Average Operational Cost','Demand Response Operational
Cost'});
bar(c,CT);
```

```
title('Related Costs');
ylabel('Euros(€)');
grid on;
legend('By Small Data Center', 'By Medium Data Center', 'All Small Data
Centers', 'All Medium Data Centers','Total','Location','northeastoutside');
```

1.1.4 DR and RE in an Incentive-based Daily Load Diagram

```
%% Data capture from Excel table dcdata.xlsx
filename = 'dcdata.xlsx';
% sheetDC24 = 1;
sheetDC72 = 2;
% smdc_num24 = xlsread(filename,sheetDC24);
smdc_num72 = xlsread(filename,sheetDC72);
%% Variables declaration with the values present in the table
pt_s = smdc_num72(:,2); % SDC power
pt_m = smdc_num72(:,3); % MDC power
price72 = smdc_num72(:,4); % 24h*3 tariff variable
price_avg = mean(price72); % average price
price_avg1(1:42,:) = price_avg;
price_avg2 = price72(43:49);
price_avg3(1:23,:) = price_avg;
price_dr = [price_avg1;price_avg2;price_avg3]; % DR price
dr_s = smdc_num72(:,5); % DRE in SDC
dr_m = smdc_num72(:,6); % DRE in MDC
rb_s = smdc_num72(:,7); % RE in SDC
rb_m = smdc_num72(:,8); % RE in MDC
pdr_s = pt_s-dr_s; % decreased power in SDC DRE
pdr_m = pt_m-dr_m; % decreased power in MDC DRE
prb_s = pdr_s+rb_s; % increased power in SDC RE
prb_m = pdr_m+rb_m; % increased power in MDC RE
t = 1:1:72; % 24h/3 time variable
smalldc; % call smalldc function
mediumdc; % call mediumdc function
 %% Interactions and Results
disp('----- The Demand Response event will occur between 2:40 p.m and 3 p.m
-----');
disp('----- O Rebound Effect will occur between 3 p.m and 5 p.m -----');
fprintf('The kWh average price is: %.4f €\n',price_avg);
fprintf('The kWh price during the Demand Response event is: %.4f
€\n',price_dr(43));
c1 = sum(pt_s.*price_avg);
c2 = sum(pt_m.*price_avg);
c11 = sum(pt_s.*price72); % normal daily cost
c12 = sum(pt_m.*price72);
c13 = c11-price72(43)*105; % DR cost
c14 = c12-price72(43)*550;
c15 = c11-c13; % difference between normal cost and demand response
c16 = c12-c14;
c17 = 7.1*c15/100; % incentive considering the above difference
c18 = 2*c16/100;
c19 = c11-c17*105;
c20 = c12-c18*550;
fprintf('The daily energy cost by Small Data Center is: %.2f €\n',c11);
fprintf('The daily energy cost by Medium Data Center is: %.2f €\n',c12);
fprintf('The average operational cost by Small Data Center is: %.2f
€\n',c1);
fprintf('The average operational cost by Medium Data Center is: %.2f
€\n',c2);
fprintf('The incentive to Small Data Center is: %.2f €\n',c17);
fprintf('The incentive to Medium Center √©: %.2f €\n',c18);
```

```
fprintf('The operational cost with Demand Response with no incentive by
Small Data Center is: %.2f €\n',c13);
fprintf('The operational cost with Demand Response with no incentive by
Medium Data Center is: %.2f €\n',c14);
fprintf('The operational cost with Demand Response with incentive by Small
Data Center is: %.2f €\n',c19);
fprintf('The operational cost with Demand Response with incentive by Medium
Data Center is: %.2f €\n',c20);
%% Chart
x1 = x_s(1:72);
x2 = x_s(73:144);
x3 = x_s(145:216);
x4 = x_m(1:72);
x5 = x_m(73:144);
x6 = x_m(145:216);
figure('Name','Demand Response e Rebound Effect by Contract');
subplot (4,1,1);
plot(f(1:72),'LineWidth',1);
grid on;
title('Daily Tariff');
xlabel('Time (hours)');
ylabel('Cost (€/kWh)');
xtickangle(60);
subplot (4,1,2);
plot(x1,'LineWidth',1);
hold on;
plot (x2,'LineWidth',1);
plot (x3,'LineWidth',1);
grid on;
title('Daily Diagram of Demand Response in Small Data Center');
xlabel('Time (hours)');
ylabel('Power (kW)');
ylim([-80 80]);
xtickangle(60);
legend('ICT Workload', 'Cooling', 'UPS','Location','northeastoutside');
subplot(4,1,3);
plot(x4,'LineWidth',1);
hold on;
plot (x5,'LineWidth',1);
plot (x6,'LineWidth',1);
grid on;
title('Daily Diagram of Demand Response in Medium Data Center');
xlabel('Time (hours)');
ylabel('Power (kW)');
xtickangle(60);
legend('ICT Workload', 'Cooling', 'UPS','Location','northeastoutside');
hold off
% figure('Name','Diagrama de Carga dos Data Centers');
subplot(4,1,4);
plot(t,pt_s,t,pt_m,t,pdr_s,'--',t,pdr_m,'--
',t,prb_s,':',t,prb_m,':','LineWidth',1);
grid on
title('Daily Diagram');
xlabel('Time (hours)');
ylabel('Power (kW)');
:00','16:20','16:40','17:00','17:20','17:40','18:00','18:20','18:40','19:00
','19:20','19:40','20:00','20:20','20:40','21:00','21:20','21:40','22:00','
22:20','22:40','23:00','23:20','23:40','00:00','00:20','00:40'});
xtickangle(60);
legend('Small Data Centers', 'Mediuns Data Centers', 'Small Data Centers
with DR', 'Medium Data Centers with DR', 'Small Data Centers with RE',
'Medium Data Centers with RE','Location','northeastoutside');
```

1.2 DSO

1.2.1 DemandResponse

```
%call variables
callvariables;
for evt=1:numEvent
  ApplyVariables
  if Acceptedlist(1) > 0
        startingevent
        %ratio criterion
        RatioCalculation;
        %  Power NC criterion
        PowerNcCriterion;
        instantPotDC  = zeros(1,length(SMDC(:,1))).';
                format shortG
                if rate_SDC > 0 && rate_SDC < maxSMDCsmall
                    for aj=1:rate_SDC
                        aux_randi =
returnPositionPositionSmaller(smdc_small(:,5));
                        if smdc_small(aux_randi,3) <
smdc_small(aux_randi,4)
                            smdc_small(aux_randi,5) =
smdc_small(aux_randi,5) +1;
                            power_reduced = power_reduced +
smdc_small(aux_randi,2)*SMDC(smdc_small(aux_randi,6),7);
                            smdc_small(aux_randi,3) =
smdc_small(aux_randi,2) +
smdc_small(aux_randi,3)*SMDC(smdc_small(aux_randi,6),7);
                            reducSDCperEvent(evt) = reducSDCperEvent(evt) +
smdc_small(aux_randi,2)*SMDC(smdc_small(aux_randi,6),7);
                            SMDC(smdc_small(aux_randi,6),3) =
SMDC(smdc_small(aux_randi,6),3) +
SMDC(smdc_small(aux_randi,6),2)*SMDC(smdc_small(aux_randi,6),7);
                            instantPotDC(smdc_small(aux_randi,6)) =
SMDC(smdc_small(aux_randi,6),2)*SMDC(smdc_small(aux_randi,6),7);
                            SMDC(smdc_small(aux_randi,6),5) =
SMDC(smdc_small(aux_randi,6),5) + 1;
                            selectedSDC =[selectedSDC
smdc_small(aux_randi,1)]; % SDC called list
                            indexSDCCham(evt) = indexSDCCham(evt) + 1;
                        end
                    end
                end
                if rate_MDC > 0 && rate_MDC < maxSMDCmedium
                    for aj=1:rate_MDC
                         aux_randi =
returnPositionPositionSmaller(smdc_medium(:,5));
                         if smdc_medium(aux_randi,3) <=
smdc_medium(aux_randi,4)
                             smdc_medium(aux_randi,5) =
smdc_medium(aux_randi,5) +1;
                             power_reduced =power_reduced +
smdc_medium(aux_randi,2)*SMDC(smdc_medium(aux_randi,6),7);
                             smdc_medium(aux_randi,3) =
smdc_medium(aux_randi,3) +
smdc_medium(aux_randi,2)*SMDC(smdc_medium(aux_randi,6),7);
                             reducMDCperEvent(evt) = reducMDCperEvent(evt) +
smdc_medium(aux_randi,2)*SMDC(smdc_medium(aux_randi,6),7);
```

```
                            SMDC(smdc_medium(aux_randi,6),3) =
SMDC(smdc_medium(aux_randi,6),3) +
SMDC(smdc_medium(aux_randi,6),2)*SMDC(smdc_medium(aux_randi,6),7);
                            instantPotDC(smdc_medium(aux_randi,6)) =
SMDC(smdc_medium(aux_randi,6),2)*SMDC(smdc_medium(aux_randi,6),7);
                            SMDC(smdc_medium(aux_randi,6),5) =
SMDC(smdc_medium(aux_randi,6),5) + 1;
                            selectedMDC =[selectedMDC
smdc_medium(aux_randi,1)]; %MDC called list
                            indexMDCCham(evt) = indexMDCCham(evt) + 1;
                         end
                    end
                end

                checkPowerReduced = zeros(1,length(DRsignal(:,1)));
                checkPowerReduced(1) = power_reduced;
                selectedSDC = selectedSDC(2:end);
                selectedMDC = selectedMDC(2:end);

                format shortG
        disp("-------------------------------------------------------------
-----------------------------------------------------------");
                disp("Small data centers (SDC) chosen: ");
                disp([['Dc(s): '],[num2str(selectedSDC)]]);
                disp("Medium data centers (MDC) chosen: ");
                disp([['Dc(s): '],[num2str(selectedMDC)]]);
                disp(" ");
        disp("-------------------------------------------------------------
-----------------------------------------------------------");
                                                                for i=1:length(Pdr(:))

powerTotalrequired(evt) = powerTotalrequired(evt) + Pdr(i);
                                                                end
        for c=1:length(DRsignal(:,1))
                disp("***** - Time window: "+DRsignal(c,1)+" *****");
                disp(" ");



                if power_reduced < pdrInstante(c,Pdr)
                 if length(smdc_small) > 0
                  for ev=1:chamadaSDC
                      if power_reduced >= pdrInstante(c,Pdr) break; end
                     aux_randi =
returnPositionPositionSmaller(smdc_small(:,5));
                      if smdc_small(aux_randi,3) <= smdc_small(aux_randi,4)
                        if (length(selectedSDC)+length(neededCallSDC)) <
(length(smdc_small(:,1))+1)
                           power_reduced = power_reduced+
smdc_small(aux_randi,2)*SMDC(smdc_small(aux_randi,6),7);
                           checkPowerReduced(c) = checkPowerReduced(c) +
smdc_small(aux_randi,2)*SMDC(smdc_small(aux_randi,6),7);
                           SMDC(smdc_small(aux_randi,6),5) =
SMDC(smdc_small(aux_randi,6),5)+1;
                           SMDC(smdc_small(aux_randi,6),3) =
SMDC(smdc_small(aux_randi,6),2)*SMDC(smdc_small(aux_randi,6),7)+SMDC(smdc_s
mall(aux_randi,6),3);
                           instantPotDC(smdc_small(aux_randi,6)) =
SMDC(smdc_small(aux_randi,6),2)*SMDC(smdc_small(aux_randi,6),7);
                           smdc_small(aux_randi,5) =
smdc_small(aux_randi,5)+ 1;
```

```
                           smdc_small(aux_randi,3) =
smdc_small(aux_randi,2)*SMDC(smdc_small(aux_randi,6),7)+
smdc_small(aux_randi,3);
                           reducSDCperEvent(evt) = reducSDCperEvent(evt) +
smdc_small(aux_randi,2)*SMDC(smdc_small(aux_randi,6),7);
                           neededCallSDC = [neededCallSDC
smdc_small(aux_randi,1)];
                           indexSDCCham(evt) = indexSDCCham(evt) + 1;
                           disp(["SDC that had to be called! Dc:
",smdc_small(aux_randi,1)]);
                        end
                      end
                  end
                 end
                if length(smdc_medium) > 0
                  for ev=1:chamadaMDC
                      if power_reduced >= pdrInstante(c,Pdr) break; end
                      aux_randi =
returnPositionDoSmaller(smdc_medium(:,5));
                      if smdc_medium(aux_randi,3) <=
smdc_medium(aux_randi,4)
                        if (length(selectedMDC)+length(neededCallMDC)) <
length(smdc_medium(:,1)+1)
                           power_reduced = power_reduced+
smdc_medium(aux_randi,2)*SMDC(smdc_medium(aux_randi,6),7);
                           checkPowerReduced(c) = checkPowerReduced(c) +
smdc_medium(aux_randi,2)*SMDC(smdc_medium(aux_randi,6),7);
                           SMDC(smdc_medium(aux_randi,6),5) =
SMDC(smdc_medium(aux_randi,6),5) + 1;
                           SMDC(smdc_medium(aux_randi,6),3) =
SMDC(smdc_medium(aux_randi,6),3) +
SMDC(smdc_medium(aux_randi,6),2)*SMDC(smdc_medium(aux_randi,6),7);
                           instantPotDC(smdc_medium(aux_randi,6)) =
SMDC(smdc_medium(aux_randi,6),2)*SMDC(smdc_medium(aux_randi,6),7);
                           smdc_medium(aux_randi,5) =
smdc_medium(aux_randi,5) + 1;
                           smdc_medium(aux_randi,3) =
smdc_medium(aux_randi,2)*SMDC(smdc_medium(aux_randi,6),7) +
smdc_medium(aux_randi,3);
                           reducMDCperEvent(evt) = reducMDCperEvent(evt) +
smdc_medium(aux_randi,2)*SMDC(smdc_medium(aux_randi,6),7);
                           neededCallMDC = [neededCallMDC
smdc_medium(aux_randi,1)];
                           indexMDCCham(evt) = indexMDCCham(evt) + 1;
                           disp(["MDC that had to be called! Dc:
",smdc_medium(aux_randi,1)]);
                        end
                      end
                  end
                end
                end
                disp("Power to be decreased: "+Pdr(c)+" kW");
                disp("Reduction achieved: "+abs(checkPowerReduced(c))+"
kW");
                disp("Reduction of each data center called in this time
window");
                disp(" ")
                aux_Dc = [0];
                aux_D = [0];
                aux_instPot = [0];
                aux_per = [0];
                for i=1:length(SMDC(:,1))
                    if instantPotDC(i) ~= 0
```

```
                        aux_Dc = [aux_Dc SMDC(i,1)];
                        aux_D = [aux_D SMDC(i,2)];
                        aux_instPot = [aux_instPot instantPotDC(i)];
                        aux_per = [aux_per SMDC(i,7)*100];
                    end
                end
                    aux_Dc  = aux_Dc(2:end).';
                    aux_D  = aux_D(2:end).';
                    aux_instPot = aux_instPot(2:end).';
                    aux_per = aux_per(2:end).';


disp(table(aux_Dc,aux_D,aux_instPot,aux_per,'VariableNames',{'Dc' 'D'
'Decreased' 'DecreasedPercentage'}))
                disp(" ");
                instantPotDC = instantPotDC*0;
                clear aux_D;
                clear aux_instPot;
                clear aux_per;
              end
              format shortG

        gravarDados
        apresentacaoResultados

        limparVariaveis
  else
      disp("Event number: "+evt);
      disp("There was 0% of acceptance of data centers!");
  end
end %end of loop
disp(" ");
disp("Total number of events: "+numEvent);
disp(" ");
percentChartQMA_small = percentChartQMA_small(2:end,:);
percentChartQMA_medium = percentChartQMA_medium(2:end,:);
percentChartQDM_small = percentChartQDM_small(2:end,:);
percentChartQDM_medium = percentChartQDM_medium(2:end,:);
filename = 'input.xlsx';
sheetSMDC = 1;
smdc_num = xlsread(filename,sheetSMDC);

for iv=1:length(smdc_num(:,1))
    if (smdc_num(iv,5)*smdc_num(iv,2)) == 0
        smdc_num(iv,6) = 0;
    else
        smdc_num(iv,6) = (smdc_num(iv,3) /
(smdc_num(iv,5)*smdc_num(iv,2) ))*100;
    end
    if smdc_num(iv,3) == 0 && smdc_num(iv,5) == 1
        smdc_num(iv,5) = 0;
    end
end
chartsDR
```

1.2.2 RatioCalculation

```
disp("--------------------------------------------------------------------
-------------------------------------------------------------");
disp("Ratio of SMDC list:");
if ratio_SDC == 0
    if round(ratio_MDC*100) < 10
        tab = [[num2str(round(ratio_SDC*100)) ,    '% -  Small Data Centers
']];
        tab = [tab;[num2str(round(ratio_MDC*100)) ,'% -  Medium Data
Centers']];
    else
        tab = [[num2str(round(ratio_SDC*100)) ,    '% -  Small Data Centers
']];
        tab = [tab;[num2str(round(ratio_MDC*100)) ,'% -  Medium Data
Centers']];
    end
elseif  ratio_MDC == 0
    if round(ratio_SDC*100) < 10
        tab = [[num2str(round(ratio_SDC*100)) ,    '% -  Small Data Centers
']];
        tab = [tab;[num2str(round(ratio_MDC*100)) ,'% -  Medium Data
Centers']];
    else
        tab = [[num2str(round(ratio_SDC*100)) ,    '% -  Small Data Centers
']];
        tab = [tab;[num2str(round(ratio_MDC*100)) ,'% -  Medium Data
Centers  ']];
    end
else
    if round(ratio_SDC*100) < 10
        tab = [[num2str(round(ratio_SDC*100)) ,    '% -  Small Data Centers
']];
        tab = [tab;[num2str(round(ratio_MDC*100)) ,'% -  Medium Data
Centers']];
    elseif round(ratio_MDC*100) < 10
        tab = [[num2str(round(ratio_SDC*100)) ,    '% -  Small Data Centers
']];
        tab = [tab;[num2str(round(ratio_MDC*100)) ,'% -  Medium Data
Centers ']];
    else
        tab = [[num2str(round(ratio_SDC*100)) ,    '% -  Small Data Centers
']];
        tab = [tab;[num2str(round(ratio_MDC*100)) ,'% -  Medium Data
Centers']];
    end
end
disp(tab);
disp("List of SMDC that accepted the invitation:");
tab = [["Dc:";SMDC(:,1)],["D: ";SMDC(:,2)]];
disp(tab);
disp("List of SMDCs that denied the invitation");
lisDeny = listDeny(smdc_num,SMDC);
tab = [["Dc:";lisDeny(:,1)],["D: ";lisDeny(:,2)]];
disp(tab);
disp("--------------------------------------------------------------------
-------------------------------------------------------------");
 % Acceptance percentage in SMDC
 if length(lisDeny(:,1)) > 0
     listGeneralSmall    = geteDC(smdc_num,"SMALL");    listGeneralSmall    =
length(listGeneralSmall(:,1));
```

```
     totalDeniedSmall  = geteDC(lisDeny,"SMALL");  totalDeniedSmall  =
length(totalDeniedSmall(:,1));
     listGeneralMedium  = geteDC(smdc_num,"MEDIUM");   listGeneralMedium  =
length(listGeneralMedium(:,1));
     totalDeniedMedium = geteDC(lisDeny,"MEDIUM"); totalDeniedMedium =
length(totalDeniedMedium(:,1));

     percentChartQMA_small  = [percentChartQMA_small;  (1-(totalDeniedSmall
/ listGeneralSmall)) *100];
     percentChartQMA_medium = [percentChartQMA_medium; (1-
(totalDeniedMedium / listGeneralMedium ))*100];
     percentChartQDM_small  = [percentChartQDM_small;  (totalDeniedSmall
/ listGeneralSmall)  *100];
     percentChartQDM_medium = [percentChartQDM_medium; (totalDeniedMedium
/ listGeneralMedium) *100];

 else
     percentChartQMA_small  = [percentChartQMA_small;[100]];
     percentChartQMA_medium = [percentChartQMA_medium;[100]];
     percentChartQDM_small  = [percentChartQDM_small;[0]];
     percentChartQDM_medium = [percentChartQDM_medium;[0]];
     totalDeniedSmall = 0;
     totalDeniedMedium = 0;
 end
        indexSDCRejCham(evt) = totalDeniedSmall;
        indexMDCRejCham(evt) = totalDeniedMedium;

if ratio_SDC > maxSMDCsmall && ratio_MDC > maxSMDCmedium
   disp("The number of data centers that accepted the invitation is not
enough to reduce the desired power of:");DRsignal
   disp("Mega Watts");
end
ratio_SDC_aux = ratio_SDC;
ratio_MDC_aux = ratio_MDC;

ratio_SDC = round( (Pdrm*ratio_SDC) /sizeSMALL_DC ); ratio_MDC =
round( (Pdrm*ratio_MDC) /sizeMEDIUM_DC );
disp('Amount of SDC needed to achieve the power reduction in comparison
with the average Pdrm:');
        disp(ratio_SDC);
disp('Amount of MDC needed to achieve the power reduction in comparison
with the average Pdrm:');
        disp(ratio_MDC);

% Achieve maximum reduction by event
  crtSDC = abs(ratio_SDC - ((Pdrm*ratio_SDC_aux) /sizeSMALL_DC ));
  crtMDC = abs(ratio_MDC - ((Pdrm*ratio_MDC_aux) /sizeMEDIUM_DC));
  clear ratio_SDC_aux;
  clear ratio_MDC_aux;
  if crtMDC < 0.0001
      crtMDC = 0.05;
  end
  if crtSDC < 0.0001
      crtSDC = 0.05;
  end
```

1.2.3 PowerNCCriterion

```
Nc_num = xlsread(filename,sheetNc);

    [ accomplished, slightly_above, slightly_below, quite_below, below ] =
homogeneousDistribution(Nc_num(3,3:7),SMDC);

    % Reduction percentage
    Nc_num = Nc_num(7,3:7);

    % Slightly above
    percent_slightly_above = Nc_num(1);
    % Accomplished
    percent_accomplished     = Nc_num(2);
    % Slightly bellow
    percent_slightly_below = Nc_num(3);
    % Below
    percent_below         = Nc_num(4);
    % Quite below
    percent_quite_below       = Nc_num(5);

    % accomplished column
    SMDC = [SMDC,zeros(1,length(SMDC(:,1))).'];

%% accomplished group
    cont = length(SMDC(:,7));
    sorting = 0;
    if accomplished > 0
        for i=1: accomplished

           if cont > 0
                sorting =  randi(length(SMDC(:,1)));
                while SMDC(sorting,7) ~= 0
                      sorting =  randi(length(SMDC(:,1)));
                end

                SMDC(sorting,7) = percent_accomplished;
                cont = cont -1;
           else
               break;
           end
        end
    else
        bool_accomplished = false;
    end
%% slightly above group
    if slightly_above > 0
        for i=1: slightly_above
            if cont > 0
                sorting =  randi(length(SMDC(:,1)));
                while SMDC(sorting,7) ~= 0
                      sorting =  randi(length(SMDC(:,1)));
                end

                SMDC(sorting,7) = percent_slightly_above;
                cont = cont -1;
           else
               break;
           end
        end
```

```
    else
        bool_slightly_above = false;
    end
%% slightly below
    if slightly_below > 0
        for i=1: slightly_below
             if cont > 0
                sorting =  randi(length(SMDC(:,1)));
                while SMDC(sorting,7) ~= 0
                      sorting =  randi(length(SMDC(:,1)));
                end

                SMDC(sorting,7) = percent_slightly_below;
                cont = cont -1;
           else
               break;
           end
        end
    else
        bool_slightly_below = false;
    end
%% quite below group
    if quite_below > 0
        for i=1: quite_below
            if cont > 0
                sorting =  randi(length(SMDC(:,1)));
                while SMDC(sorting,7) ~= 0
                      sorting =  randi(length(SMDC(:,1)));
                end

                SMDC(sorting,7) = percent_quite_below;
                cont = cont -1;
           else
               break;
           end
        end
    else
        bool_quite_below = false;
    end
%% below group
if  > 0
    for i=1:
        if cont > 0
            sorting =  randi(length(SMDC(:,1)));
            while SMDC(sorting,7) ~= 0
                  sorting =  randi(length(SMDC(:,1)));
            end

            SMDC(sorting,7) = percent_;
            cont = cont -1;
       else
           break;
       end
    end
else
    bool_quite_below = false;
end
```

I.2.4 ReductionFairnessCriterion

```
%  reduction fairness criterion
if ratio_SDC >= maxSMDCsmall
    ratio_SDC = 0;
    for i=1:length(smdc_small(:,1))% small data center check
        if SMDC(i,3) < SMDC(i,4)
            selectedSDC =[selectedSDC smdc_small(i,1)];
            power_reduced = power_reduced +
smdc_small(i,2)*SMDC(smdc_small(i,6),7);            smdc_small(i,5) =
smdc_small(i,5)+1;            smdc_small(i,3) =
smdc_small(i,3)+smdc_small(i,2)*SMDC(smdc_small(i,6),7);
            reducSDCperEvent(evt) = reducSDCperEvent(evt) +
smdc_small(i,2)*SMDC(smdc_small(i,6),7);
            instantPotDC(smdc_small(i,6)) =
SMDC(smdc_small(i,6),2)*SMDC(smdc_small(i,6),7);
            indexSDCCall(evt) = indexSDCCall(evt) + 1;
            SMDC(i,5) = SMDC(i,5) + 1;
            SMDC(i,3) = SMDC(i,3) + SMDC(i,2)*SMDC(smdc_small(i,6),7);
        end
    end
end
% medium data center check
if ratio_MDC >= maxSMDCmedium
    ratio_MDC = 0    for i=1:length(smdc_medium(:,3))
        if SMDC(i+maxSMDCsmall,3) < SMDC(maxSMDCsmall+i,4)
            selectedMDC =[selectedMDC smdc_medium(i,1)];
            power_reduced = power_reduced +
smdc_medium(i,2)*SMDC(smdc_medium(i,6),7            smdc_medium(i,5) =
smdc_medium(i,5)+1;            smdc_medium(i,3) =
smdc_medium(i,3)+smdc_medium(i,2)*SMDC(smdc_medium(i,6),7);
            reducMDCperEvent(evt) = reducMDCperEvent(evt) +
smdc_medium(i,2)*SMDC(smdc_medium(i,6),7);
            indexMDCCall(evt) = indexMDCCall(evt) + 1;
            instantPotDC(smdc_medium(i,6)) =
SMDC(smdc_medium(i,6),2)*SMDC(smdc_medium(i,6),7);
            SMDC(i+maxSMDCsmall,5) = SMDC(i+maxSMDCsmall,5) + 1;
            SMDC(i+maxSMDCsmall,3) = SMDC(i+maxSMDCsmall,3) +
SMDC(i+maxSMDCsmall,2)*SMDC(smdc_medium(i,6),7);
        end
    end
end
```

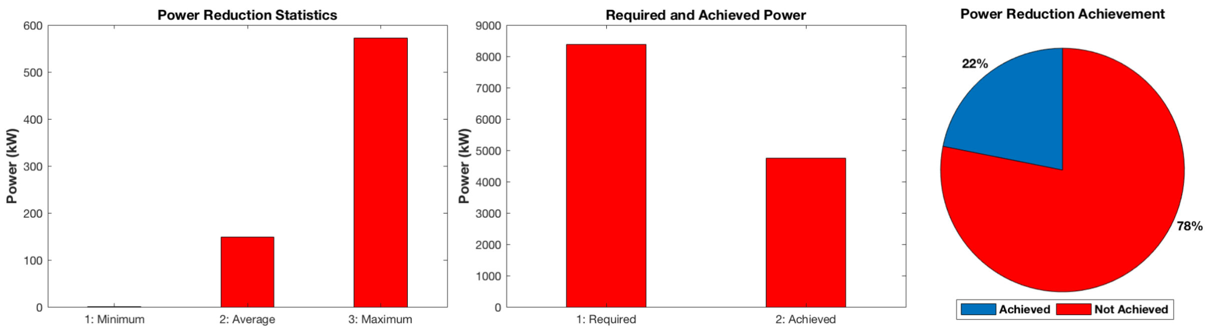
ANNEX 2: DR SIMULATION DATA

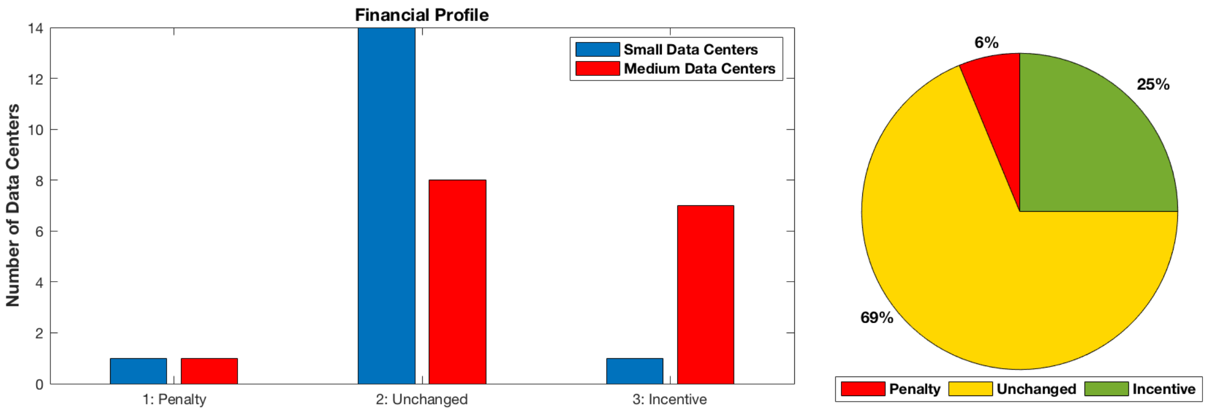
pdrwork	pdrcool	cracam	pict	cop	tsup	tsafe	tmax	pfan	pdrups	upsam	esto	estopre	emax	eff	drtime	total	profile
Small Data Centers																	
40	35	2	190	2	10	25	25	8	30	2	35,5	40	40	0,9	0,33	105	accomplished
50	22	2	60	2	10	25	25	8	33	2	35,09	40	40	0,9	0,33	105	accomplished
35	25	2	90	2	10	25	25	8	45	2	33,3	40	40	0,9	0,33	105	accomplished
23	27	2	110	2	10	25	25	8	55	2	32,57	40	40	0,9	0,33	105	accomplished
28	42	2	260	2	10	25	25	8	35	2	34,8	40	40	0,9	0,33	105	accomplished
30	52	2	360	2	10	25	25	8	23	2	36,58	40	40	0,9	0,33	105	accomplished
41	35	2	190	2	10	25	25	8	30	2	35,5	40	40	0,9	0,33	106	slightly above
52	22	2	60	2	10	25	25	8	33	2	35,09	40	40	0,9	0,33	107	slightly above
35	25	2	90	2	10	25	25	8	48	2	32,87	40	40	0,9	0,33	108	slightly above
23	27	2	110	2	10	25	25	8	59	2	31,23	40	40	0,9	0,33	109	slightly above
28	47	2	300	2	12	27	27	11	35	2	24,8	30	30	0,9	0,33	110	slightly above
30	58	2	432	2	12	27	27	11	23	2	26,57	30	30	0,9	0,33	111	slightly above
37	30	2	96	2	12	27	27	11	28	2	25,83	30	30	0,9	0,33	95	slightly below
36	24	2	24	2	12	27	27	11	27	2	25,98	30	30	0,9	0,33	87	slightly below
29	24	2	24	2	12	27	27	11	41	2	23,9	30	30	0,9	0,33	94	slightly below
17	23	2	12	2	12	27	27	11	53	2	22,12	30	30	0,9	0,33	93	slightly below
20	37	2	180	2	12	27	27	11	29	2	26,57	30	30	0,9	0,33	86	slightly below
29	51	2	348	2	12	27	27	11	22	2	26,73	30	30	0,9	0,33	102	slightly below
24	23	2	12	2	12	27	27	11	18	2	27,32	30	30	0,9	0,33	65	below
27	23	2	12	2	12	27	27	11	23	2	26,57	30	30	0,9	0,33	73	below
23	10	2	99	2	15	30	30	1,7	31	2	15,39	20	20	0,9	0,33	64	below
10	12	2	129	2	15	30	30	1,7	23	2	16,57	20	20	0,9	0,33	45	below
20	32	2	32	2	15	30	30	1,7	29	2	15,69	20	20	0,9	0,33	81	below
14	16	2	189	2	15	30	30	1,7	10	2	18,51	20	20	0,9	0,33	40	below
11	5	2	24	2	15	30	30	1,7	4	2	19,4	20	20	0,9	0,33	20	quite below
9	4	2	9	2	15	30	30	1,7	6	2	19,1	20	20	0,9	0,33	19	quite below
4	4	2	9	2	15	30	30	1,7	10	2	18,51	20	20	0,9	0,33	18	quite below
4	6	2	39	2	15	30	30	1,7	7	2	18,8	20	20	0,9	0,33	17	quite below
1	5	2	24	2	15	30	30	1,7	2	2	19,7	20	20	0,9	0,33	8	quite below
4	5	2	24	2	15	30	30	1,7	1	2	19,84	20	20	0,9	0,33	10	quite below
Medium Data Centers																	
220	180	6	440	2	10	25	25	8	150	2	57,72	80	80	0,9	0,33	550	accomplished
210	155	6	357	2	10	25	25	8	185	2	52,52	80	80	0,9	0,33	550	accomplished
190	160	6	374	2	10	25	25	8	200	2	50,3	80	80	0,9	0,33	550	accomplished
120	200	6	507	2	10	25	25	8	230	2	45,84	80	80	0,9	0,33	550	accomplished
165	195	6	490	2	10	25	25	8	190	2	51,78	80	80	0,9	0,33	550	accomplished
150	300	6	840	2	10	25	25	8	100	2	65,15	80	80	0,9	0,33	550	accomplished
230	180	6	440	2	10	25	25	8	150	2	57,72	80	80	0,9	0,33	560	slightly above
215	155	6	357	2	10	25	25	8	185	2	52,52	80	80	0,9	0,33	555	slightly above
197	160	6	374	2	10	25	25	8	200	2	50,3	80	80	0,9	0,33	557	slightly above
140	200	6	507	2	10	25	25	8	230	2	45,84	80	80	0,9	0,33	570	slightly above
166	195	6	516	2	12	27	27	11	190	2	41,78	70	70	0,9	0,33	551	slightly above
176	286	6	880	2	12	27	27	11	100	2	55,15	70	70	0,9	0,33	562	slightly above
219	180	6	456	2	12	27	27	11	150	2	47,72	70	70	0,9	0,33	549	slightly below
205	155	6	356	2	12	27	27	11	185	2	42,52	70	70	0,9	0,33	545	slightly below
170	160	6	376	2	12	27	27	11	200	2	40,3	70	70	0,9	0,33	530	slightly below
100	190	6	496	2	12	27	27	11	230	2	35,84	70	70	0,9	0,33	520	slightly below
100	195	6	516	2	12	27	27	11	190	2	41,78	70	70	0,9	0,33	485	slightly below
115	250	6	736	2	12	27	27	11	100	2	55,15	70	70	0,9	0,33	465	slightly below
200	130	6	256	2	12	27	27	11	100	2	55,15	70	70	0,9	0,33	430	below
195	90	6	90	2	12	27	27	11	105	2	54,44	70	70	0,9	0,33	390	below
70	60	6	249	2	15	30	30	1,7	100	2	45,15	60	60	0,9	0,33	230	below
80	90	6	399	2	15	30	30	1,7	130	2	40,69	60	60	0,9	0,33	300	below
20	95	6	424	2	15	30	30	1,7	60	2	51,08	60	60	0,9	0,33	175	below
119	186	6	880	2	15	30	30	1,7	50	2	52,57	60	60	0,9	0,33	355	below
40	35	6	125	2	15	30	30	1,7	30	2	55,54	60	60	0,9	0,33	105	quite below
53	22	6	59	2	15	30	30	1,7	33	2	55,09	60	60	0,9	0,33	108	quite below
32	25	6	74	2	15	30	30	1,7	48	2	52,87	60	60	0,9	0,33	105	quite below
21	27	6	84	2	15	30	30	1,7	59	2	51,23	60	60	0,9	0,33	107	quite below
20	37	6	134	2	15	30	30	1,7	29	2	55,68	60	60	0,9	0,33	86	quite below
20	32	6	109	2	15	30	30	1,7	29	2	55,68	60	60	0,9	0,33	81	quite below

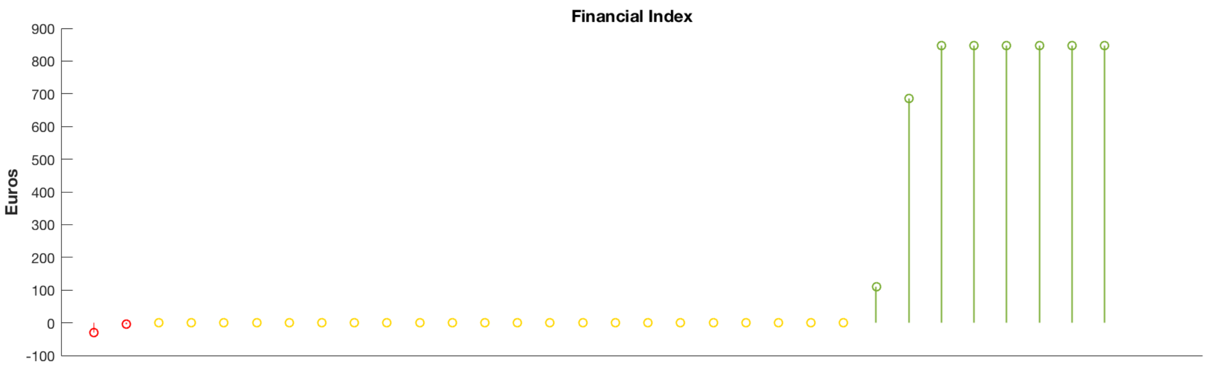
ANNEX 3: DR SIMULATION CHARTS

3.1 ONE DAY SCENARIO

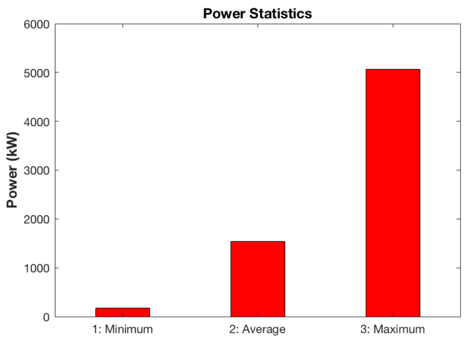
3.1.1 16 Small Data Centers 16 Medium Data Centers (Unchanged)

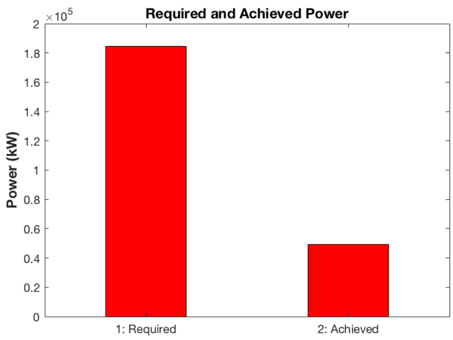


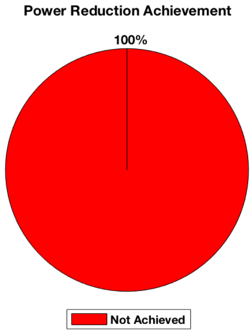


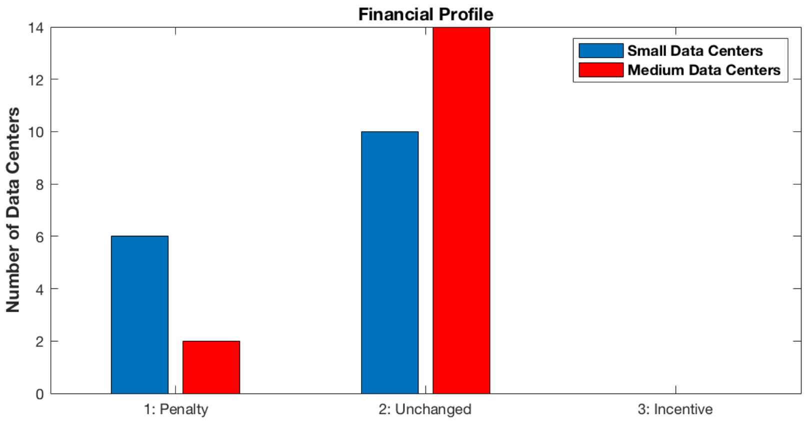


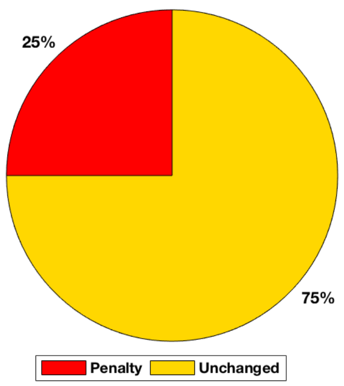
3.1.2 16 Small Data Centers 16 Medium Data Centers (Penalty)



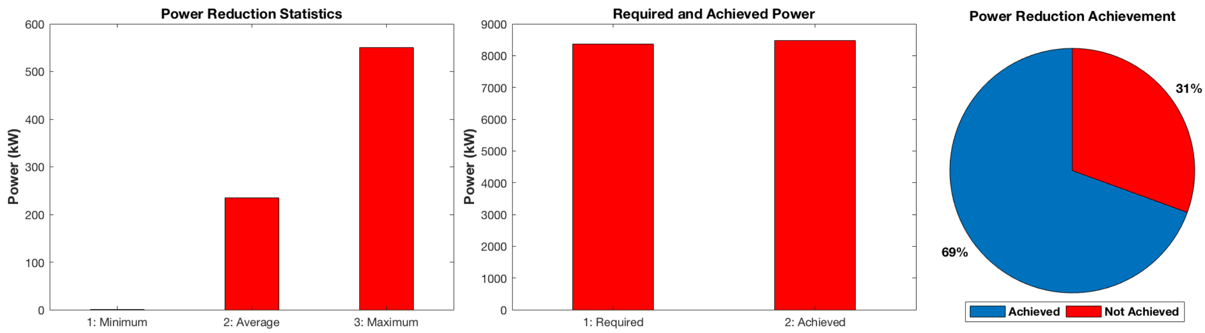


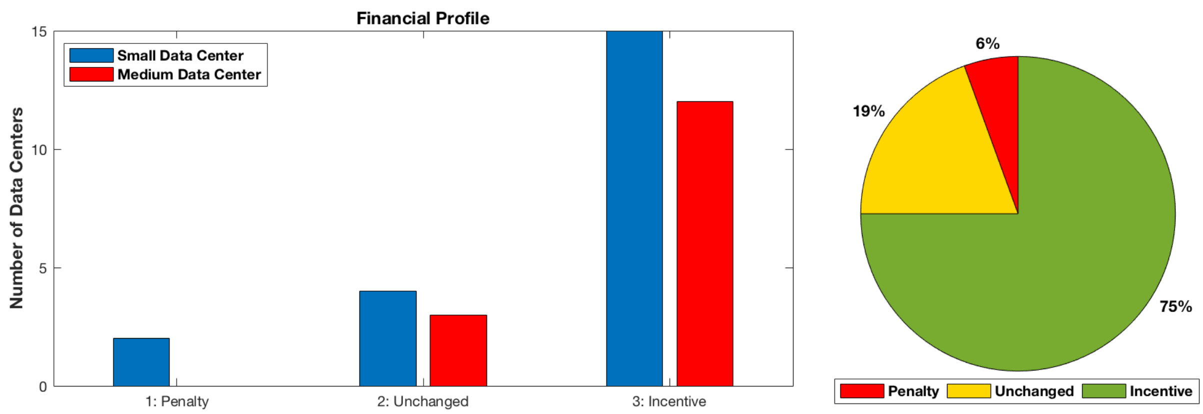


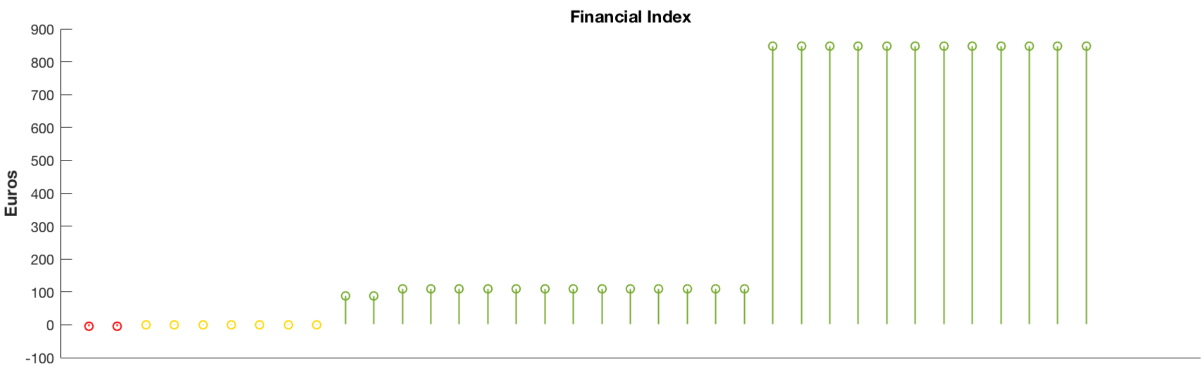




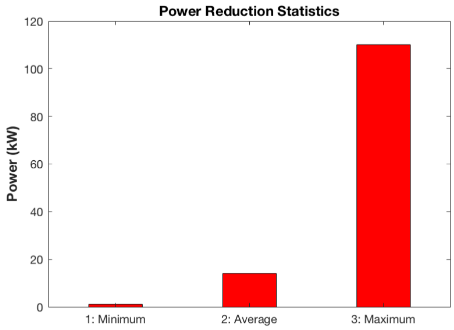
3.1.3 21 Small Data Centers 15 Medium Data Centers (Incentive)

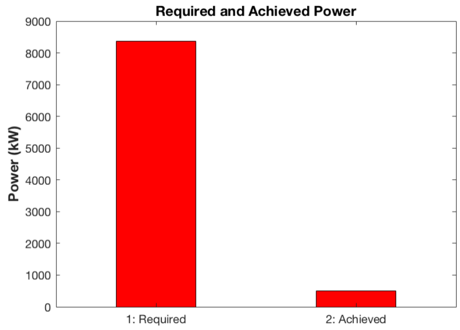


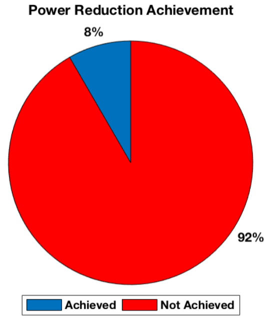


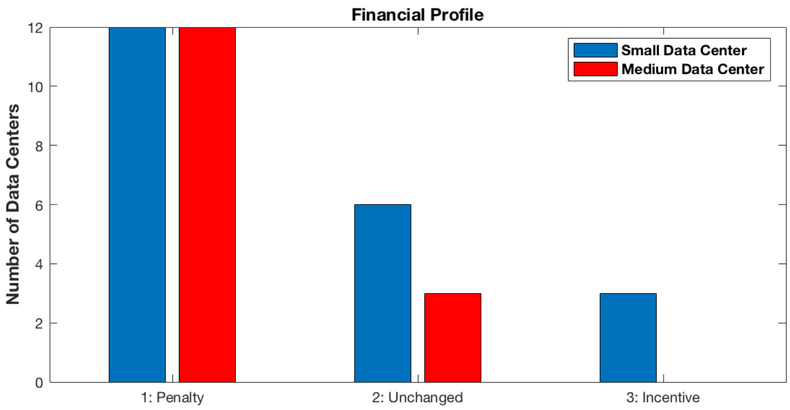


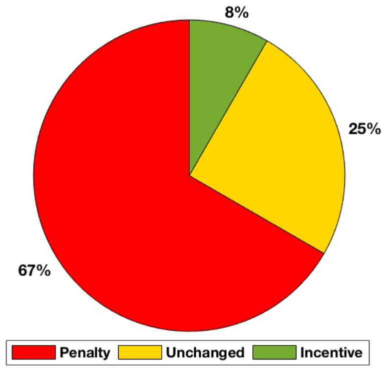
3.1.4 21 Small Data Centers 15 Medium Data Centers (Penalty)

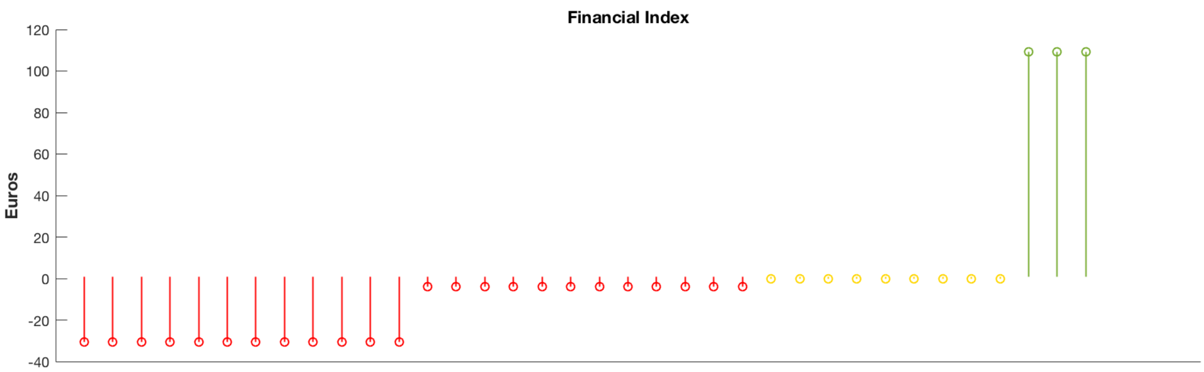




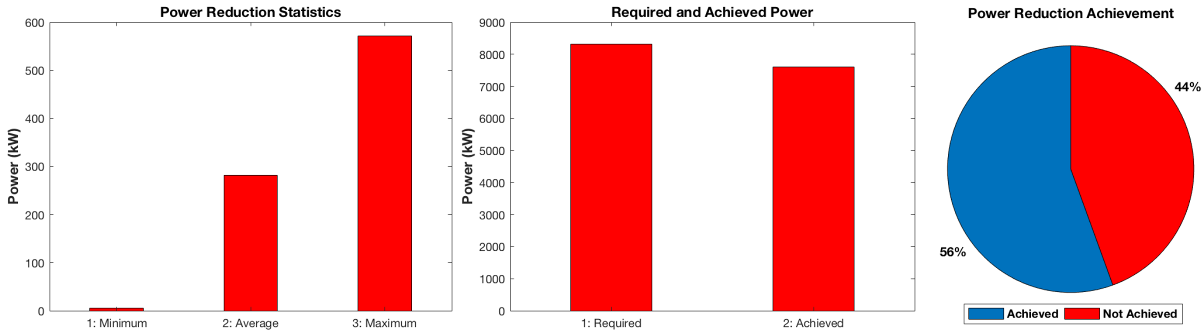


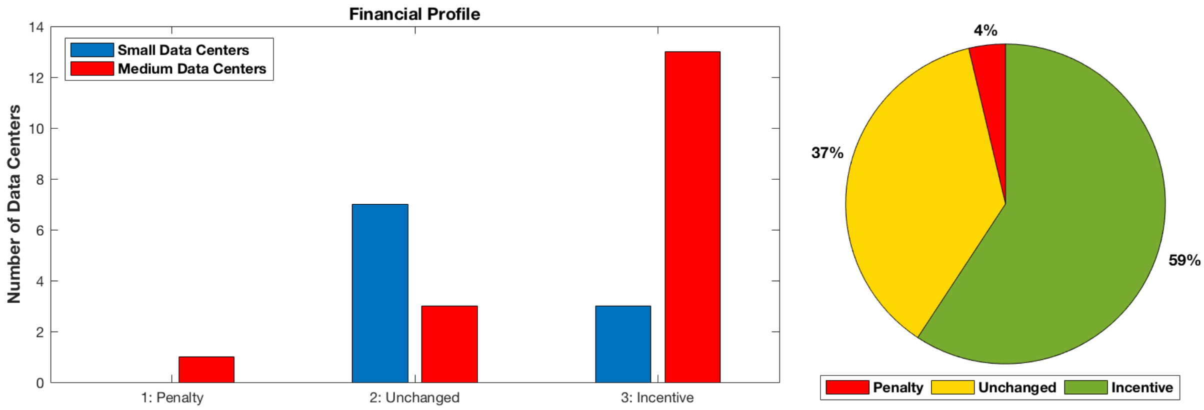


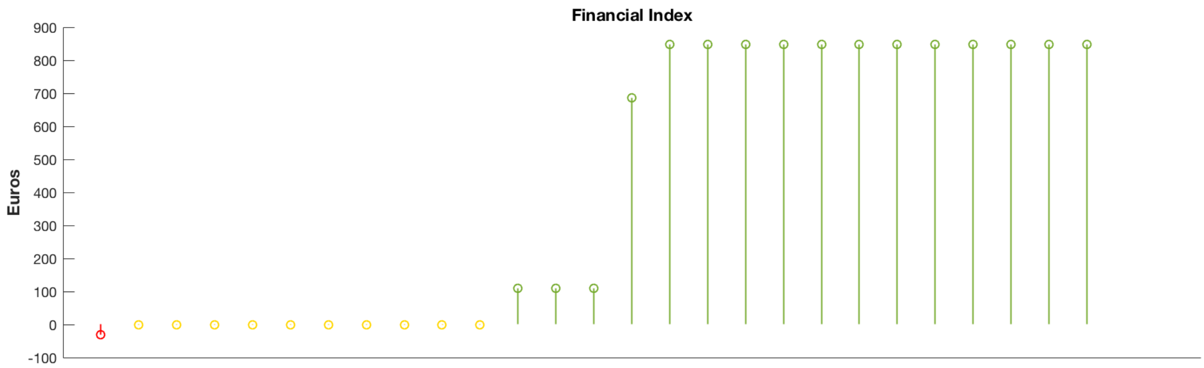




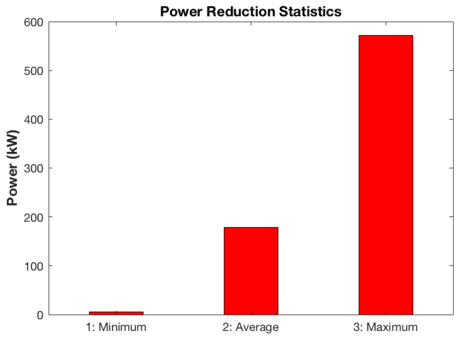
3.1.5 10 Small Data Centers 17 Medium Data Centers (Incentive)

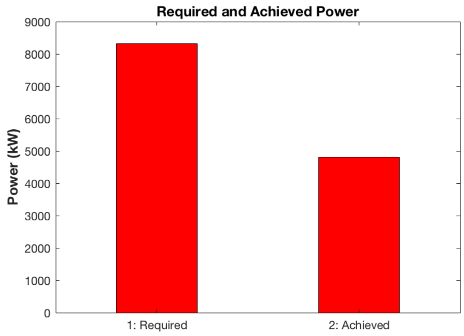


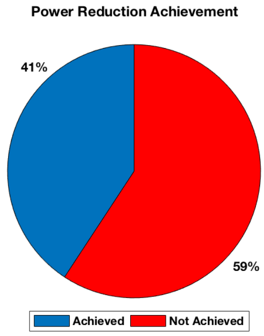


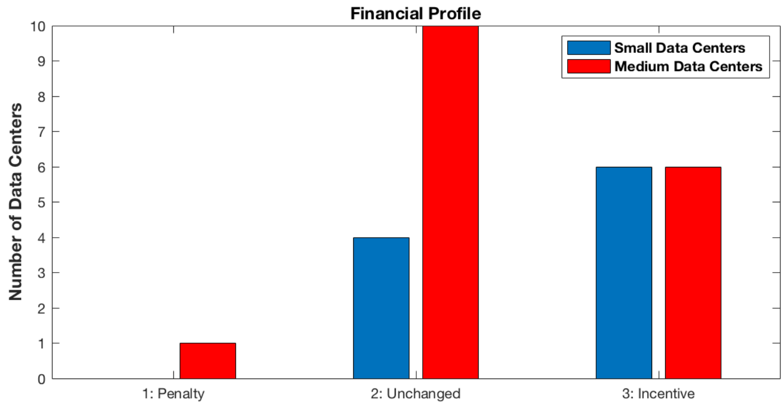


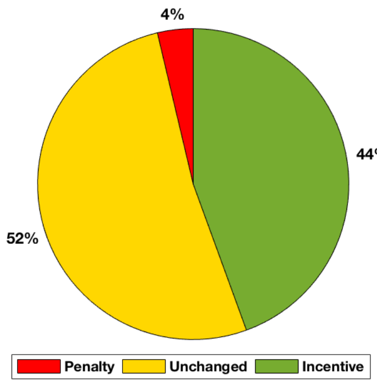
3.1.6 10 Small Data Centers 17 Medium Data Centers (Unchanged)

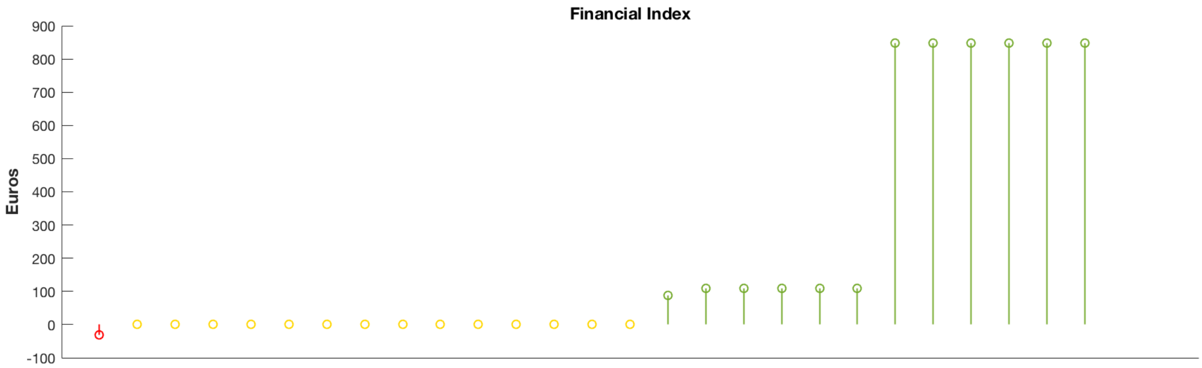






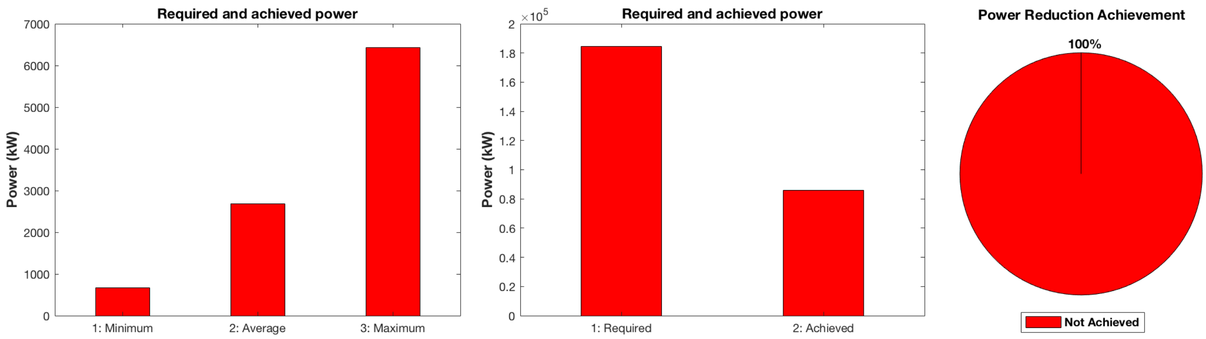


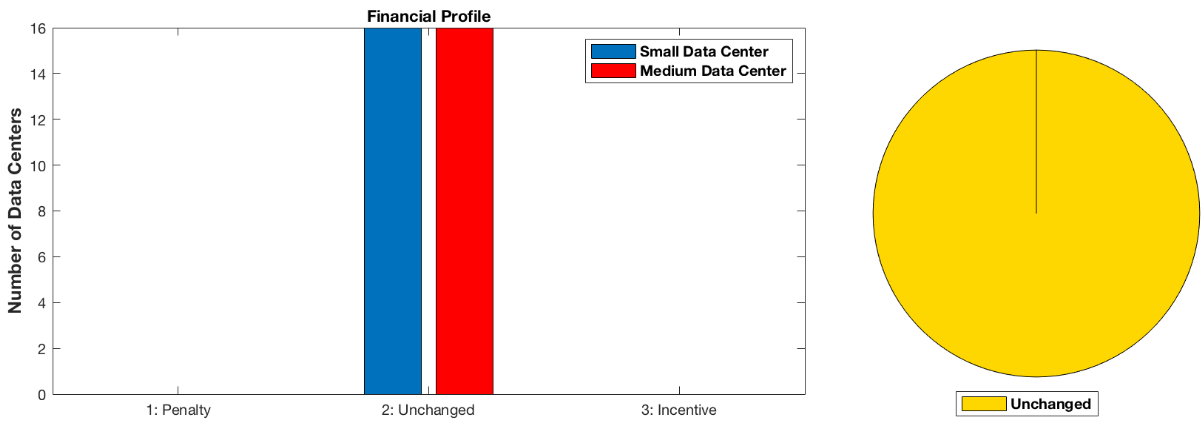




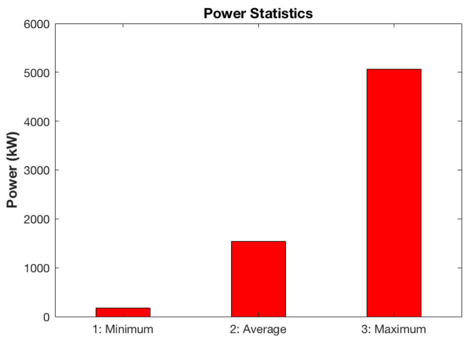
3.2 22 DAYS SCENARIO

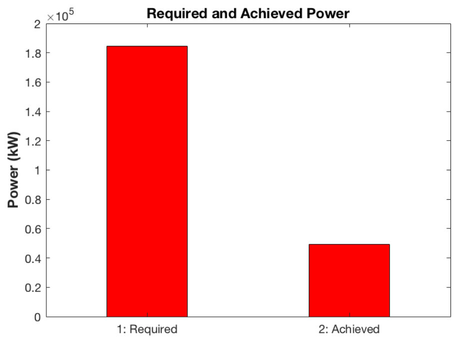
3.2.1 16 Small Data Centers 16 Medium Data Centers (Unchanged)

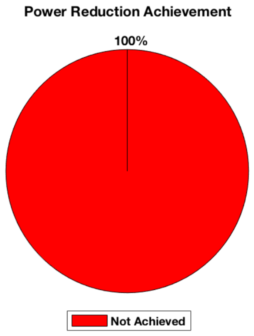


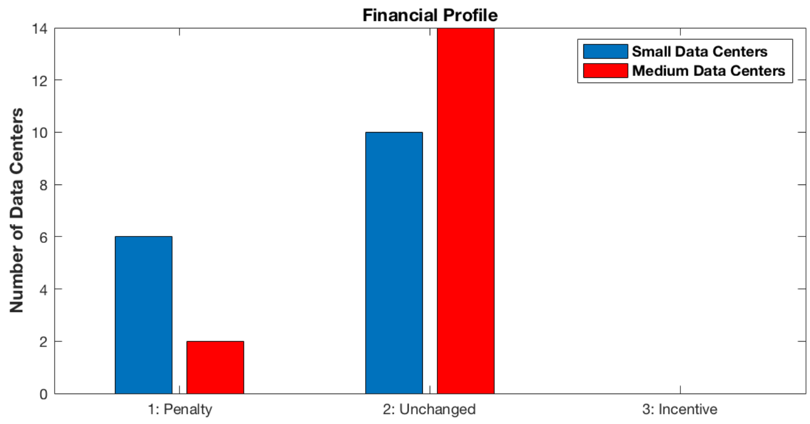


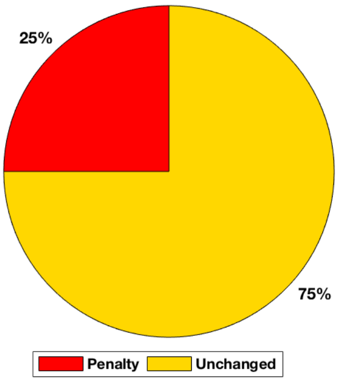
3.2.2 16 Small Data Centers 16 Medium Data Centers (Penalty)



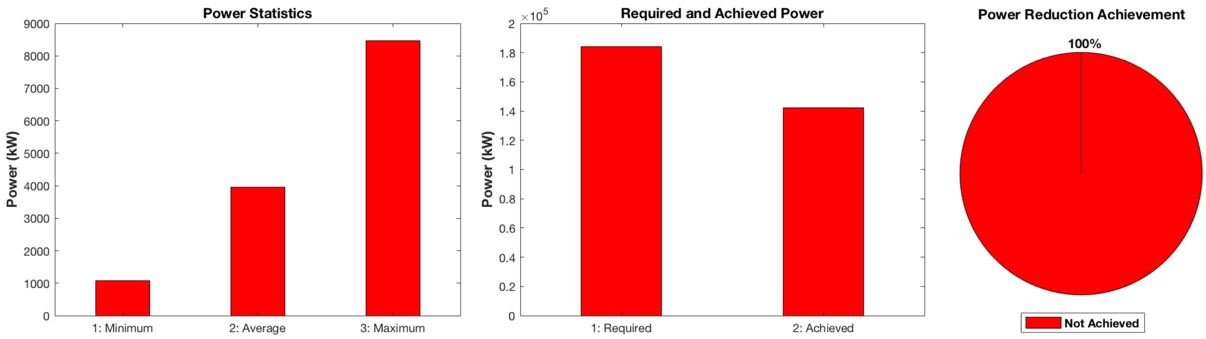


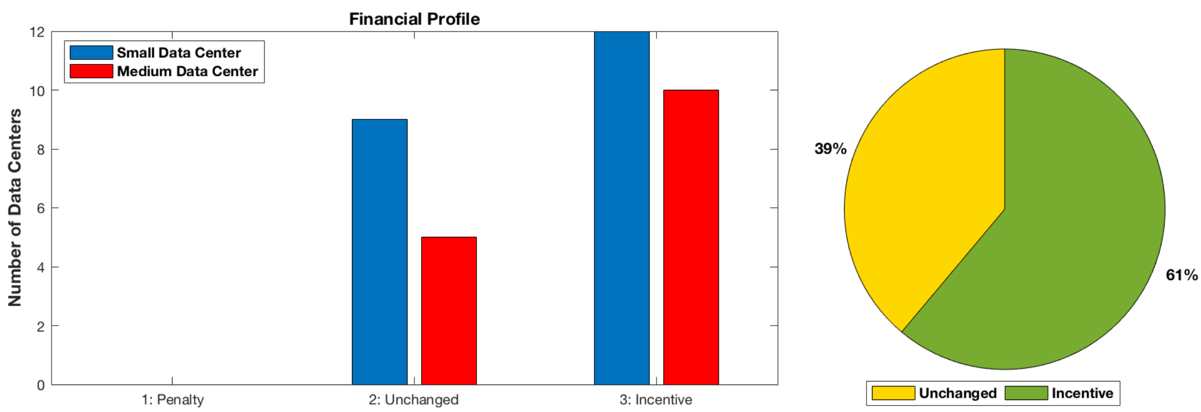


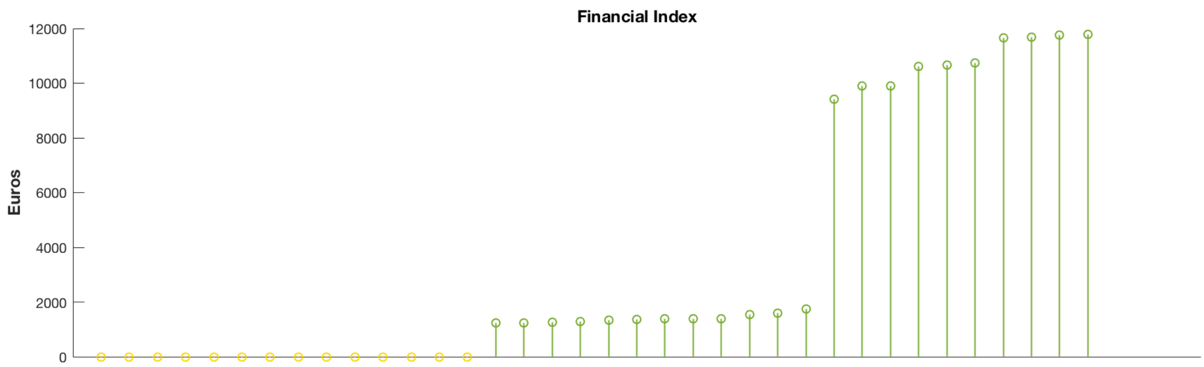




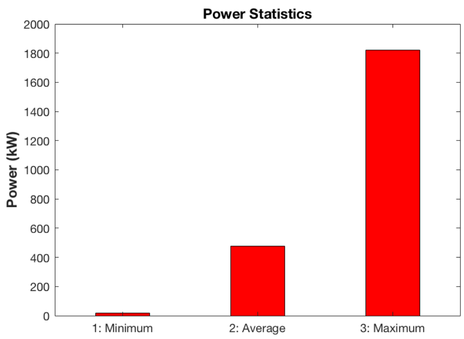
3.2.3 21 Small Data Centers 15 Medium Data Centers (Incentive)

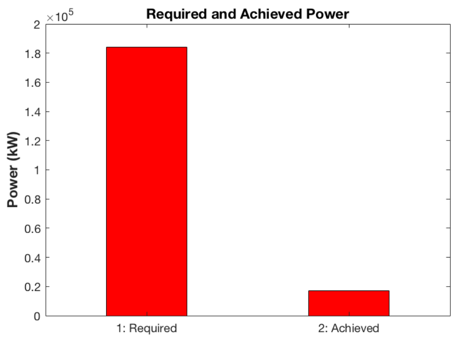


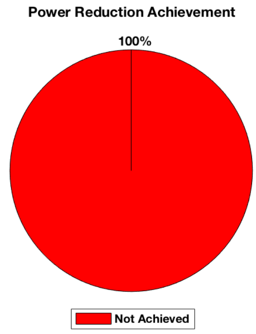


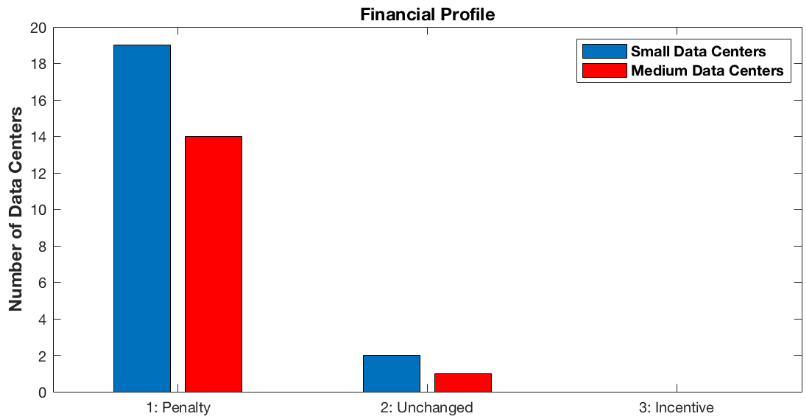


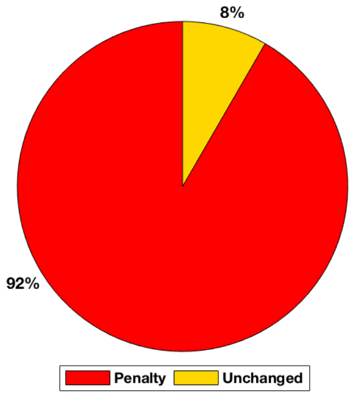
3.2.4 21 Small Data Centers 15 Medium Data Centers (Penalty)

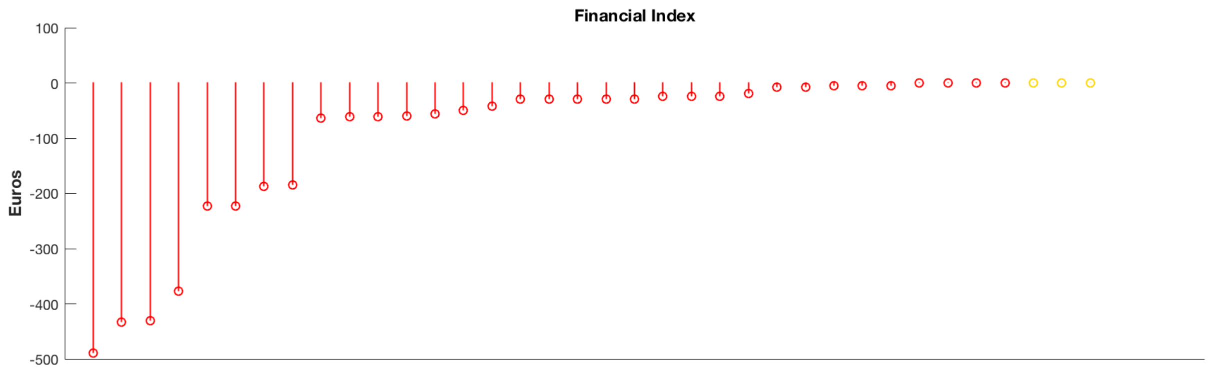




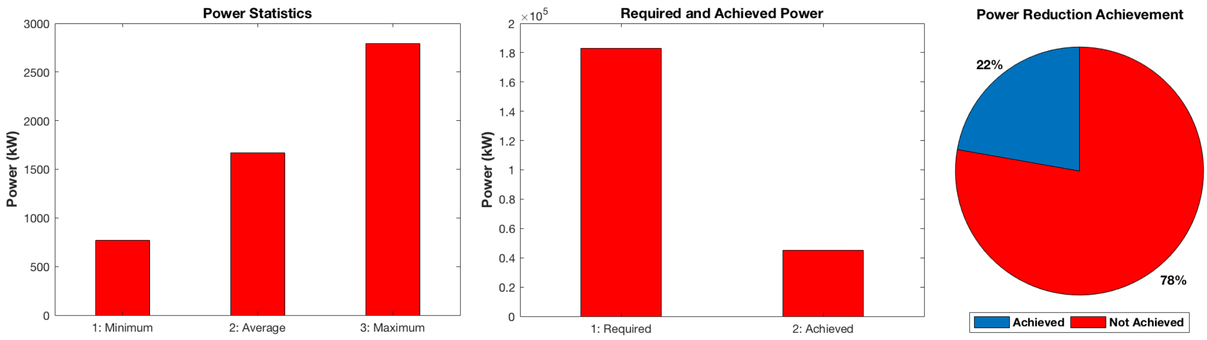


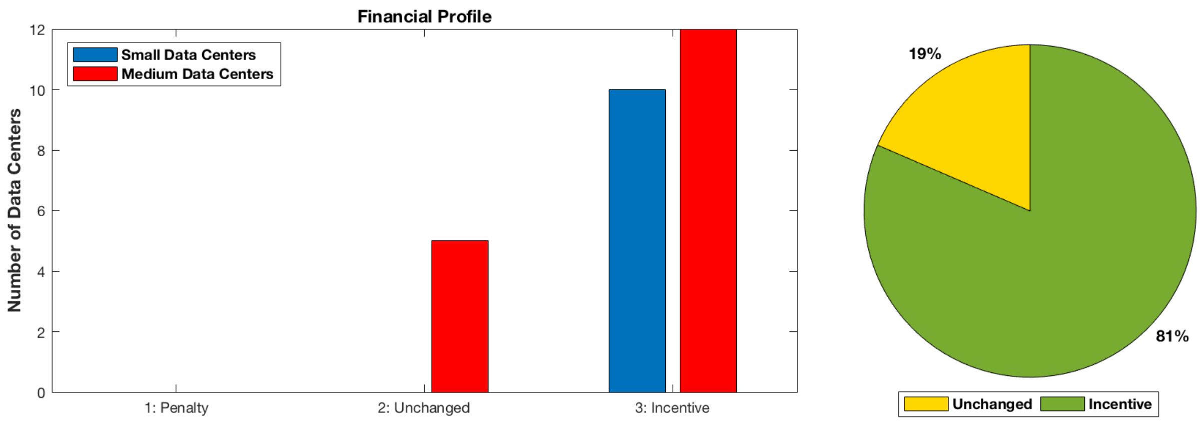


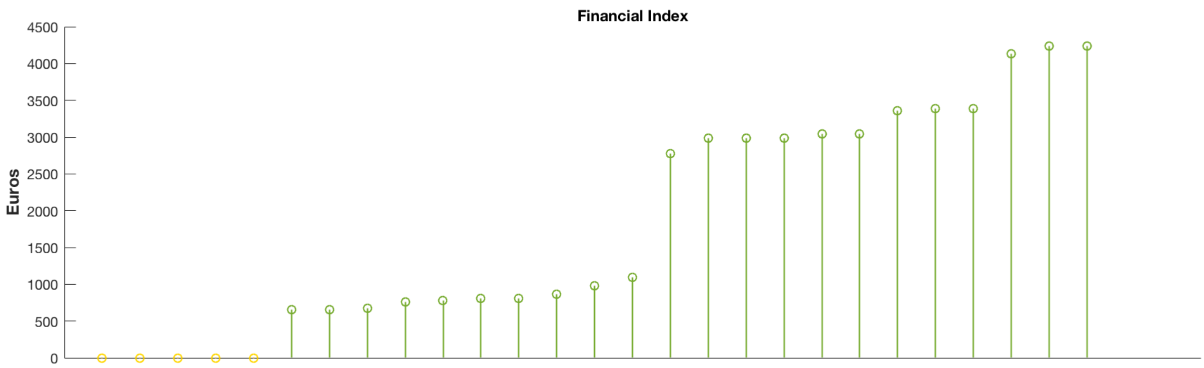




3.2.5 10 Small Data Centers 17 Medium Data Centers (Incentive)







3.2.6 10 Small Data Centers 17 Medium Data Centers (Unchanged)