

AN INNOVATIVE APPROACH FOR HARMONISING AVAILABILITY AND ENERGY EFFICIENCY IN DATA CENTRES: A CASE STUDY

MAHMOUD KUNBAZ*, JÜRGEN BIESER#

Apleona HSG GmbH, Germany

*E-mail: mahmoud.kunbaz@apleona.com

#E-mail: juergen.bieser@apleona.com

A Data Centre (DC) as critical, energy-hungry infrastructure is dominated by two main driving forces: i) Availability and ii) Energy requirements. As a result of increasing energy costs, algorithms for optimising energy efficiency have been devised. However, these algorithms do not take availability into consideration.

This paper aims to present a combination of Failure Mode, Effects and Criticality Analysis (FMECA)/Reliability, Availability, Maintainability (RAM)/Energy Analysis as an innovative approach for harmonising availability and energy efficiency in DC. Based on various measures defined by FMECA/RAM, corresponding availability and reliability are modelled and calculated. In parallel, potential energy saving measures are included in RAM simulation to quantify their influence on the availability and reliability of DC infrastructure. As a result, a set of the most promising optimisation measures is selected.

Results show that some energy saving measures are highly correlated with availability. However, required data centre availability can be achieved with improved energy efficiency if the right set of optimisation measures is implemented. This approach guides DC managers to identify improvement potentials in terms of availability and energy efficiency, providing a reliable decision basis for future investments.

Keywords: data centre, energy performance, reliability, availability, criticality analysis

Introduction

A data centre (DC) can be defined according to DIN EN 50600 [1] as a structure or a group of structures intended for the interconnection as well as the operation of IT and network facilities which provide data storage, data processing and data transmission services. Such IT infrastructure is supported with a complex technical infrastructure including critical facilities like power supply, cooling, and monitoring systems, in order to ensure service availability along with high levels of reliability, security and environmental control. The key role of Facility Management (FM) in DC is mainly related to the operation and maintenance of such critical facilities [2].

The global market size of data centres reached 152 billion US dollars in 2016 [3], which allowed them to be touted as one of the key enabling technologies that play a vital role in the exponential growth of the IT industry. Nevertheless, the world's fast growing energy consumption as well as DC high availability requirements stand as the main challenges associated

with such growth in the DC market. It is estimated that the annual electricity cost of cloud data centres in the U.S. will reach up to 13 billion US dollars by 2020, while the equipment in these data centres will require cooling that costs between two and five million US dollars per year [4]. On the other hand, the cost of an unplanned outage in a DC could far exceed \$17,000 per minute [5]. The average cost of an unplanned outage in data centres is nearly \$9,000 per minute [5].

The above mentioned factors have motivated many researchers to devise algorithms for optimising energy efficiency, nonetheless, the impact of such optimisation algorithms on the availability and reliability of the infrastructure have not been taken into account.

The FMECA/RAM/Energy Analysis approach focuses on finding the best harmony among availability, reliability and energy efficiency in data centres.

Applying such approach on the mechanical and electrical system of a large data centre in the UK reveals that some optimisation measures resulted from the FMECA have negative impact on the availability and the reliability of the DC. Simultaneously, RAM

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purposes, provided the original author and source are credited, a link to the CC License is provided, and changes – if any – are indicated.

simulations of energy efficiency improvement measures show that their impact on the infrastructure severely differs. The best results can be only achieved when the most promising set of both measures is selected as a result of performing a number of simulations for different combinations of the most promising improvements.

The paper is structured as follows: In the following section, FMECA approach is described and the outcomes of this stage are presented. Next, RAM procedure is illustrated. Following that, energy analysis stage is demonstrated. Finally, we discuss the results of the case study and the conclusions.

2. Failure Modes Effects and Criticality Analysis (FMECA)

FMECA is just an evolution of the Failure Modes and Effects Analysis (FMEA) methodology which was developed in 1949 by the American Army [6]. FMEA aims to evaluate and analyse potential reliability risks as well as the impact of various system and equipment failure modes on the operation process, mission successes and safety of personnel and systems [7]. It can be also defined according to DIN EN 60812 [8] as a systematic approach of analysing and identifying potential failures together with their causes and effects. The FMEA method is a qualitative assessment of risk that facilitates the procedure of prioritising actions [9].

FMECA composes a Criticality Analysis (CA) in addition to FMEA, providing an added benefit by ranking system/subsystem failure modes quantitatively. Such added value allows system analysts to determine the severity related to specific systems or components. FMECA focuses mainly on, i) identifying the importance and urgency of defusing any type of failure considering the functionality and the performance of a specific system as well as the effects on the operation process, ii) determining potential failure mechanisms and their sequential effects, iii) classifying identified failure types, including their detailed characteristics such as recognisability, diagnostic capability, testability along with considering the arrangements related to replacement parts and maintenance activities, iv) and evaluating the severity of the failure/outage [10].

The first step of carrying out FMECA is to undertake specific preparatory steps including general analysis of the system(s) being studied, describing their operation and understanding the overall process as well as different environmental conditions, since these vary heavily depending on the complexity of the system and the nature of business. Participants and milestones are also defined at this stage. In the next step, a decomposition of all components of the DC and their functions is made and a hierarchy structure of systems/

subsystems/components as well as their relationships is established using a prepared FMECA worksheet. The potential failure modes of each component are then determined where single component can have more than one failure mode. A failure mode is simply defined as the manner in which a component can no longer perform its intended function [11]. A practical example of a failure mode is a mechanical defect in a switching element of the low-voltage distribution system in DC. If this switching element fails and no redundant component is present, then this might lead to an IT outage in the DC. Afterwards, criticality is evaluated by producing a Risk Priority Number (RPN) for the components by multiplying three factors, which are the probability that a failure occurs, the severity of such failure and the likelihood of detecting the failure [12, 13]. Such factors take many aspects into account such as information related to failure, redundancy of systems, spare parts availability, maintenance details, failure costs, etc.

Based on the assessed criticality of each component, optimisation measures are suggested in order to improve the availability and reliability of system(s) by reducing the risk of an identified potential failure mode(s). However, each single optimisation measure can have either positive or negative impact on the availability of the DC. It becomes even more complex to decide upon the impact on the availability of the DC, especially in case of implementing more than one single measure. Moreover, it is of utmost importance to consider the relationship between different systems/components and how they together interact. Therefore, in order to evaluate the impact of the optimisation measure(s) on DC availability, a RAM simulation of the whole DC must be carried out, illustrating how different systems are interacting with each other when implementing one or more changes, thus producing better understanding of the impact of the considered improvements on DC availability and reliability. The findings of the FMECA will serve as the basis for the System modelling in the RAM software.

3. Reliability, Availability, and Maintainability (RAM)

RAM modelling is a methodology used to predict asset performance over a defined time span. This method provides a statistical foundation for evaluating the impact of proposed changes obtained from the FMECA in terms of DC availability and reliability, offering a strategic view of how various systems perform over longer time spans.

At the first step of RAM modelling, all DC components are modelled as blocks forming a block model called Reliability Block Diagram (RBD) (see Fig. 1).

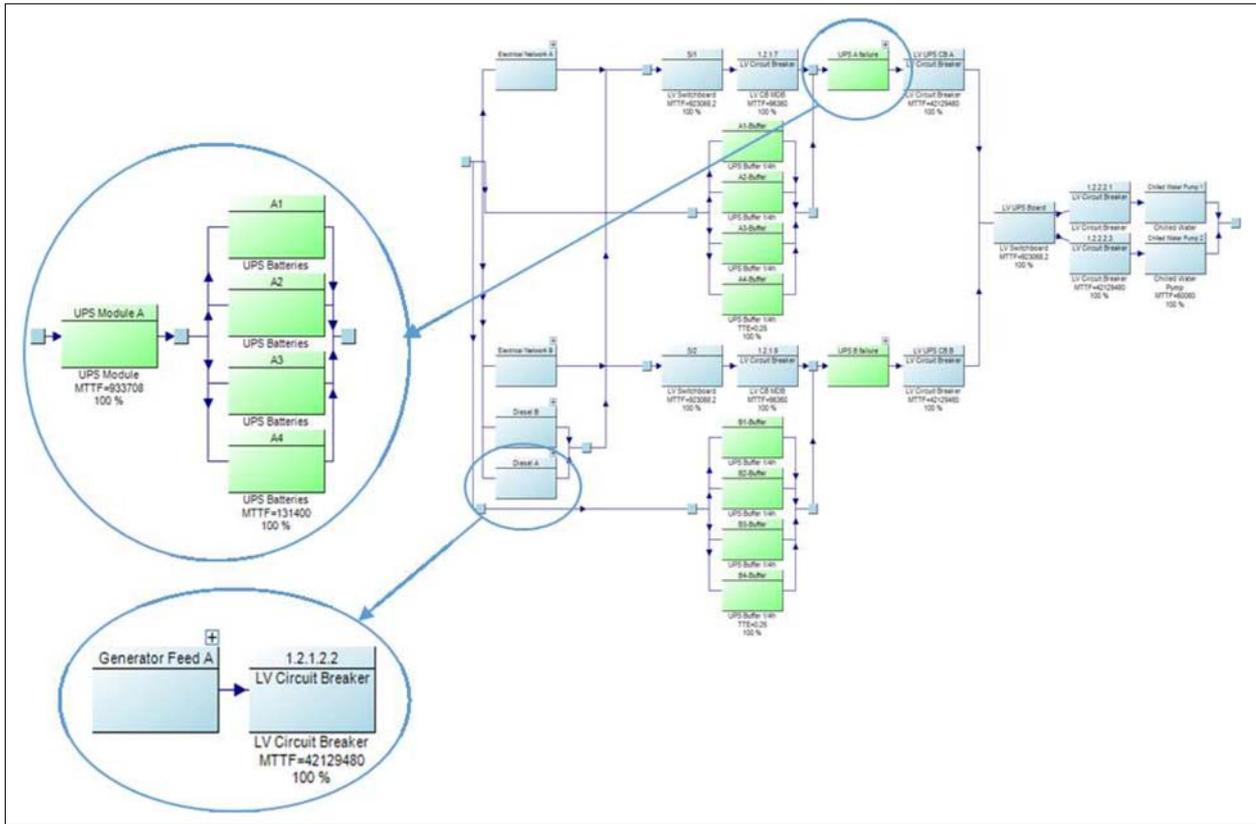


Fig. 1. Part of RBD Diagram

Each RBD consists of a combination of serial and parallel block formations, depending on the design (e.g. redundancy concept) and functionality of different DC systems, representing various functional interdependencies of different systems and how their components interact to produce system failures. A failure mode is defined for each block in the RBD using two parameters: Mean Time Between Failure (MTBF) and the Mean Time To Repair (MTTR). The failure mode of each block also includes parameters related to maintenance program (e.g. corrective maintenance), inspection frequencies, etc. [14]. Such information are already gathered in the FMECA sheet in the previous stage using various sources such as collected historical failure records, IEEE Gold Book, OREDA manual, suppliers' reliability reports, etc. The aim here is to produce forecasts that can be reliably utilised in practice. Single Line Diagrams (SLDs) of the DC stand as the basis for building the RAM model.

The software used for RAM modelling and simulation in our case study is Availability Workbench software from Isograph. Using Monte Carlo simulations, quantitative predictions of the reliability and availability of the whole DC are made. Examples of the simulation output parameters are Mean Total Down Time (TDT) in hours, Standard deviation of TDT in hours, Mean Number of Outages, etc. Deciding upon the duration as well as the required number of simula-

tions is very crucial in order to obtain accurate results. This heavily depends on the size and complexity of the model, the reference period in which the DC is evaluated, the installed redundancy and the components failure behaviour.

The as-is-state of the DC is firstly simulated in order to benchmark current availability and reliability. Afterwards, improvement measures identified by FMECA and those obtained from the energy efficiency analysis are incorporated into the RBD in order to test their influence on the availability and reliability of the data centre.

4. Energy analysis

The main aim of this stage is to identify measures that can contribute to energy savings then quantify measures out of them, which can have impact on the availability and reliability of the DC. These quantified measures will be included in the RAM analysis along with the optimisation measures obtained from the FMECA in order to find the optimal combination that ensures reliable DC with enhanced availability and reduced energy consumption.

In order to power and cool the complex IT infrastructure of data centres which include servers, storage systems, communications networking, etc., extensive mechanical and electrical infrastructures are required

[15]. Since DC must be available and running all times at all year round, they can consume as much energy as a city [16]. Typical power densities in data centres are about 538–2153 W/m² and can reach up to 10 kW/m² [17]. Data centres in the U.S. consumed approximately 70 billion kW/h in 2014 [18]. In addition to these facts, it has been reported that 2% of the global CO₂ emissions are generated by the information and communication technology sector including data centres. This high amount of power consumption attributes primarily to IT equipment requirements, cooling demands, power distribution, lightning, etc.

Energy efficiency in data centres is usually measured using Power usage effectiveness (PUE). PUE represents the ratio of energy consumed by the DC to the total amount of energy delivered to IT. In order to keep the PUE as low as possible, aspects such as cooling, uninterruptible power supply (UPS) systems, energy distribution, lightning, etc. have to be adjusted.

The energy analysis approach starts by dividing the DC into three main zones starting from inside to outside aiming to find optimisation potentials related to each considered zone. Zone one includes the area inside the building and inside the server rooms. Potentials in this zone include temperature increase, optimisation of the cold transfer, rack, etc. Zone two covers the area inside the building and outside the server rooms. Optimisation potentials taken into account here are air conditioning systems, heat generation (distribution and transfer), UPS systems, adjustment of the Genetic Algorithm (GA) parameters, lighting, etc. The last considered area is the zone three which is the area outside of the building, where aspects like constructive measures, refrigeration optimisation, co-generation or combined heat and power (CHP) are optimised.

As a summary of this stage, energy efficiency analysis focuses on the typical areas where energy can

be saved along the energy flow within the DC. These include Emergency power systems, UPS, air flow optimisation, computer room air conditioning (CRAC) units, chillers, dry coolers/cooling towers, etc. The potential improvement measures are also examined with respect to their effect on the availability of the DC in RAM simulations.

5. Case study results

In this section, we present the results of applying the FMECA/RAM/Energy Analysis approach on the mechanical and electrical system of a giant data centre located in the UK. Figure 2 provides an overview of FMECA/RAM/Energy Analysis approach.

As a first step in this study, two site visits were accomplished in order to get insights into the DC and how different systems are operating together. Basic information for the FMECA session were also collected including site and building plans, schematics, descriptions, screenshots, asset lists and energy consumption data. Afterwards, FMECA session was carried out by the project team in cooperation with the DC operation team in order to assess current risks and systematically identify potential failure mechanisms of all components. After generating a risk priority number to quantify the overall risk of each failure mode, components were categorised into three groups, non-critical, medium critical and high critical based on criticality assessment and in consultation with the data centre operator. As an output of the FMECA, high-risk components/failure modes were determined and six risk reducing measures were identified and discussed within Apleona's team. An example of such measures is to replace generator control units as this will reduce MTTR in case of a failure. Another optimisation measure is to install additional valves. This ensures that sufficient CRAC units are



Fig. 2. FMECA/RAM/Energy Analysis Approach

available to support the cooling of IT load, in case of valve failure or related maintenance activities.

In parallel with FMECA analysis, twelve improvement measures were identified as a result of the energy analysis, aiming to reduce energy consumption and energy costs in the DC. As previously explained, the DC was divided into three different zones, identifying optimisation potentials related to each considered zone. When considering zone one for instance, sealing the raised floor in the computer room saves 91,000 kW/h, resulting in about 1.8% savings through air loss. Also retrofitting non-return flaps and reducing the amount of CRAC-units in operation saves 199,700 kW/h, generating about 1.1% of savings in CRAC-units consumption. It is important to notice that only energy optimisation measures that could have an impact on the availability and reliability of the DC are considered for the RAM simulation. For instance, measures like light optimisation in different areas are therefore excluded from RAM simulation. In our case study, three energy efficiency improvement measures were quantified for the simulation.

Finally, a RAM study is conducted using a software as earlier explained considering a review period of 10 years. The data centre is modelled as blocks

based on site and building plans, schematics, descriptions and incorporating the data acquired from the FMECA, representing how various components interact to produce system failure. The current situation of the DC is then simulated in order to benchmark the current availability and reliability, creating a reference for future comparisons considering arising availability and reliability variations after simulating the DC with various optimisation measures. Table 1 shows the results of simulating each single optimisation measure in comparison with the results obtained from simulating the as-is-state of the DC in terms of availability and reliability. Availability always refers to the expected DC uptime percentage, while reliability refers to the expected number of outages that could occur in the considered time span. The values obtained from the simulation are calculated with a 95% confidence interval, in order to ensure high reliable results.

Table 1 shows clearly that each single optimisation measure has different impact on the availability and reliability of the DC. For instance, performing maintenance in accordance with the manufacturer requirements as well as operating ATS in automatic mode in order to ensure sufficient power supply to the CRAC units in case of a failure appears to be a prom-

Table 1. Results of simulations

Measure number	System	Availability	Availability	Availability	Reliability (no. of failures)	Reliability (no. of failures)
		Current situation (%)	With measure (%)	Increase/Decrease (%)	Current situation	With measures
Infrastructure						
1	Low voltage (LV) Circuit breaker	99.767	99.770	0.003	5.10	5
2	Automatic Transfer Switch (ATS)	99.767	99.765	-0.002	5.10	5.12
3	Generator Panel	99.767	99.767	0.000	5.10	5.03
4	Buffer Tank	99.767	99.771	0.004	5.10	5.07
5	Additional valves	99.767	99.810	0.043	5.10	5.05
6	Building Management System (BMS)	99.767	99.825	0.058	5.10	4.33
Energy						
7	CRAC (Shutdown of 5 CRACs in Computer Suite)	99.767	99.764	-0.003	5.10	5.26
8	CRAC (Set up hot aisle containment 1 CRAC per cabinet)	99.767	99.620	-0.147	5.10	35.92
9	CRAC (Set up hot aisle containment 2 CRACs per containment)	99.767	99.766	-0.001	5.10	5.10

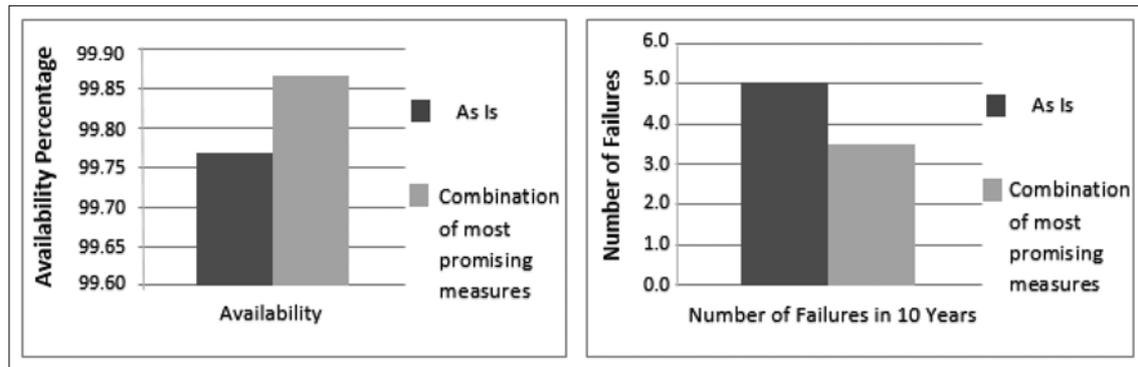


Fig. 3. Achieved Improvements with the Most Promising Set of Optimisation Measures

ising improvement measure. However, RAM simulation results show that such improvement will have a negative impact on the DC availability. The best set of improvement measures can be only determined after a number of simulations of different combinations.

Figure 3 demonstrates the results achieved after selecting the most promising set of improvement measures. Availability can be raised from 99.76% to 99.87%, while also reducing the number of outages within 10 years by approx. 31%. In addition to that, an energy saving potential of 40.6% can be achieved.

6. Conclusions

Availability and energy requirements are the main driving forces in the field of data centres. The tier topology methodology is used widely in the market in order to assess the availability of DC. Such approach classifies DC between four different classes (TIER I to IV) [19], so if the topology meets a certain class or TIER standard, then its corresponding defined availability percentage is reached. Nevertheless, the FMECA/RAM/Energy Analysis methodology follows a different approach. Independently of the current state of the DC, whether it is a new DC in the design phase or an existing one in operation, the proposed methodology in this paper examines the topology as well as the systems and components in operation, so that the availability and reliability of a data centre or even any other complex system is evaluated. Such approach requires completely different analysis of the DC than the TIER methodology and delivers a comprehensive set of information that support decision making process. Furthermore, this methodology suggests optimisation measures considering risky components, combined with proposed improvements to enhance energy efficiency. Since such measures include modifications in the systems, the RAM analysis offers additionally the possibility to quantify the impacts of such measures before the actual implementation. This is extremely important in terms of cost reduction, since implementing the wrong bundle of im-

provement measures can have negative impacts on the availability and therefore could lead to unnecessary investments due to renewal of the technical installations up to loss of reputation in the event of critical infrastructure failures. The results of this approach are usually combined with an investment plan in order to provide a sound basis for decisions related to the implementation of action plans derived from the analysis.

Acknowledgements

This paper has been developed with the support of APLEONA HSG Facility Management GmbH.

References

- [1] DIN EN 50600 (2012), Information technology – Data centre facilities and infrastructures. Deutsches Institut für Normung.
- [2] Bieser J., Menzel K. (2018), Assessing a Facility Maintenance Model of Data Centers: A Methodology for Advanced Maintenance Management for Data Centers. In: ICCCB, Finland.
- [3] Intel Intelligent Power Technology. Canals (2012), [Online, cited 2018 January 25. Available on: <https://www.canals.com/newsroom/data-center-infrastructure-market-will-be-worth-152-billion-2016>]
- [4] Sitaram, Phalachandra H. L., Gautham S., Swathi H. V., Sagar T. (2015), Energy efficient data center management under availability constraints. In: Systems Conference (SysCon), 9th Annual IEEE International, 2015 June.
- [5] Ponemon (2016), Cost of Data Center Outages. Michigan: Ponemon Institute.
- [6] U.S. Department of Defense. Procedures for Performing a Failure Mode, Effects and Criticality Analysis. MIL-P-1629, 1949.
- [7] Teoh P. C., Case K. (2005), An evaluation of failure modes and effects analysis. International Journal Computer Integrated Manufacturing, 18(4), 279–293.

- [8] DIN EN 60812 (2006), Analysis techniques for system reliability – Procedure for failure mode and effects analysis (FMEA). Deutsches Institut für Normung.
- [9] BS5760 (2009), Reliability of systems, equipment and components. British Standards Institution.
- [10] Lipol L. S., Haq J. (2011), Risk analysis method: FMEA/FMECA in the organizations. *International Journal of Basic & Applied Sciences IJBAS-IJENS*, 11(5), 49–57.
- [11] Melani H. A., Murad A., Netto C., de Souza F. M., Nabeta I. (2018), Criticality-based maintenance of a coal-fired power plant. *Energy*, 147, 767–781.
- [12] Dieter G. E. (1983), *Engineering Design: A Materials and Processing Approach*: McGraw-Hill.
- [13] Stamatis D. H. (2003), *Failure Mode and Effect Analysis: FMEA from Theory to Execution*: ASQ Quality Press.
- [14] Isograph (2017), *Availability Workbench User Manual*.
- [15] Bieser J., Kunbaz M. (2018), The role of facility maintenance in data centres: A case study. In: *European Facility Management Conference*, Sofia.
- [16] Pruhs K. (2011), Green computing algorithmics. In: *IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 3–4.
- [17] Beaty D. L. (2013), Internal IT load profile variability. ASHRAE.
- [18] Wolfgang B. (2018), *Energy Saving in Data Centers*. MDPI. 2018 January.
- [19] Uptime Institute. *Data Center Site Infrastructure Tier Standard*: Topology, 2009.