2018 3rd International Conference on System Reliability and Safety (ICSRS)

Electrical and Thermal System Impact on the Availability of a Data Center's System

Walid Mokhtar Bennaceur
Laboratoire DAVID
Université de Versailles France
walid.mkh.ben@gmail.com

Leïla Kloul
Laboratoire DAVID
Université de Versailles France
leila.kloul@uvsq.fr

***Abstract*—In this article, we analyze the availability and reliability of a data center's system using Production Trees, a new modeling methodology for dealing with availability issues of production systems. Several factors affect the data center availability, among them the power energy production and temperature variations within the data center room. This paper proposes an approach to analyze the different interactions between data center's electrical and thermal systems.**

***Index Terms*—Production Trees, Reliability, Availability, Data Center.**

I. Introduction

In Information Technology (IT), the new technologies such as social networking, e-commerce and cloud computing are constantly growing. The companies offering the online services become competitive and this has led to a rapid increase of computing and communication capabilities provided by Data Centers (DC) [11]. Thus, ensuring a continuous service by avoiding downtime is becoming a competitive factor among companies. There are many reasons for a DC downtime. The power interruption and hardware failures are the major causes. Power interruption is caused by the inability of the electrical power system to provide sufficient energy to cooling and IT devices due to a failure. Moreover, in a DC, when the temperature rises above the recommended 20 to 22 *Celsius degrees* range, the hardware inside IT components may fail more frequently [2]. Therefore the cooling system of a DC impacts on the availability of the services provided by the IT system.

In this paper, we analyze the availability and reliability of a DC system through the analysis of its three main sub-systems: power sub-system, cooling sub-system and network or IT sub-system. In particular, we analyze the interactions between the electrical and the thermal sub-systems and their impact on the servers, the main network entities.

In order to perform availability analysis of a DC's system, several techniques are used. They are classified in two categories: Boolean formalisms and States/Transitions formalisms [4]. Boolean techniques look at the system components, critical events, and system characteristics. They have convenient graphical representations which is important for industrial models. However, they take into account neither the functional dependencies, nor temporal dependencies between events occurrences. States/Transitions techniques are able to represent dependencies between failures. They are very used to represent dynamic models and have a convenient graphical representation but this representation becomes unreadable for large scale models.

One of the DC system characteristics is dependency. Each component's failure in a sub-system can affect other components of the sub-system or the whole sub-system. Moreover, to satisfy load demands in the system, it is necessary to generate sufficient power energy and cooled air, and transport it to the load points (servers), taking into account the maximum capacity of each component in the system. Therefore the availability analysis of a DC's system depends on both the electrical sub-system and the cooling sub-system. The former is responsible for providing power while the latter is responsible for extracting the heat from the DC's room. Moreover, the availability analysis of each sub-system depends not only on the internal state of each component, but

978-1-7281-0238-2/18/$31.00 ©2018 IEEE
DOI 10.1109/ICSRS.2018.00031

also on its maximum production capacity.

A new modeling methodology called Production Trees (PT) was proposed in [10]. This formalism allows modeling the relationship between the components of a system with a particular attention to the flows circulating between these components. PTs look like Fault Trees (FT) with nodes that represent components and gates that represent behaviors. Therefore, this technique is very suitable for production availability analysis of production systems such as electrical and thermal systems.

In this paper, we investigate the availability and reliability of a DC's system using production trees. Firstly, we analyze the different sub-systems (Thermal and electrical) using PT modeling technique. Then we analyze the different interactions in the system , that is the impact of the electrical system on the cooling system and the impact of the cooling system on the IT system in term of temperature variations within the DC room.

Structure of the paper: in Section 2, we discuss related work. Section 3 presents both the electrical power and the cooling systems we are interested in. Section 4 presents an overview of PT modeling technique. Section 5 is dedicated to modeling the system using PT methodology. In Section 6 we analyze the model. Section 7 summarizes the obtained results. Finally, Section 8 concludes this work.

II. Related work

In the last few years, some works have been developed to perform reliability and availability analysis of DC systems. Wei [5] combines the advances of both RBD and GSPN for quantifying availability of Virtual Data Center (VDC). DC cooling architectures are not the focus of this work and the proposed models are specific for modeling VDC. A tooled approach to estimate availability of DC's power system called Mercury is presented in [7]. This tool supports RBD, SPN, CTMC and Energy Flow Model (EFM). The EFM verifies the energy flow model on the electrical power system, taking into account the power capacity that each component can provide. However, all techniques based on Petri nets in general may partly be categorized as simulation, since simulation methods often are necessary to solve the model. In [6], an analysis is carried out to ensure availability by providing adequate cooling resources to match the heat load. This work did not cover the impact of cooling component failures on the availability of the IT room. Finally, some works have been carried out in the context of both electrical and thermal systems. In [14], authors present an approach to calculate the reliability of different topologies and to compare them using SPN. However, the authors did not focus on the dependency between thermal and electrical systems.

Currently, rare are the techniques that provide an approach to estimate the DC availability taking into account the interactions between the different subsystems. Production Trees are a simple technique that describes the interactions between system components taking into account the maximum capacity flow. It is useful to represent dynamic models and compute the production availability.

III. Electrical and Thermal Topologies

In general, the servers of a DC have two main requirements: provide them with a sufficient power energy (power system) and keep them in a constant and acceptable temperature (cooling system). Thus the servers depend on both the power system and the cooling system, which itself depends on the power system to operate properly.

In this paper, we consider the topology illustrated in Figure 1 which combines a power topology and a thermal topology. These are typical thermal and power systems of a real DC.

In a normal operating mode, servers are powered by two paths A and B. Each path is supplied by two Power Sources PS_1 and PS_2. However, if one of these power supplies fails, the power is supplied by a backup Power Generator (PG). The energy from the power sources is distributed between two Transformers Tr_A and Tr_B, one on each path. Then, the power enters the building with low-voltage going to FDP_A and FDP_B, the Front Distribution Panels, to supply two Uninterruptible Power Systems (UPS) per path noted UPS_{iA} and UPS_{iB}, $i = 1, 2$.

An UPS contains a battery to store energy and bridge the time between the utility failure and the availability of PG. The output flow from each UPS is finally routed to a Back low-voltage master Distribution Panel. We note BDP_{iX} the i^{th} distribution panel on flow path X, $X = A, B$ and $i = 1, 2$. Then, both $BDPs$ on path $X = A$ (respectively $X = B$) are connected to four (respectively two)

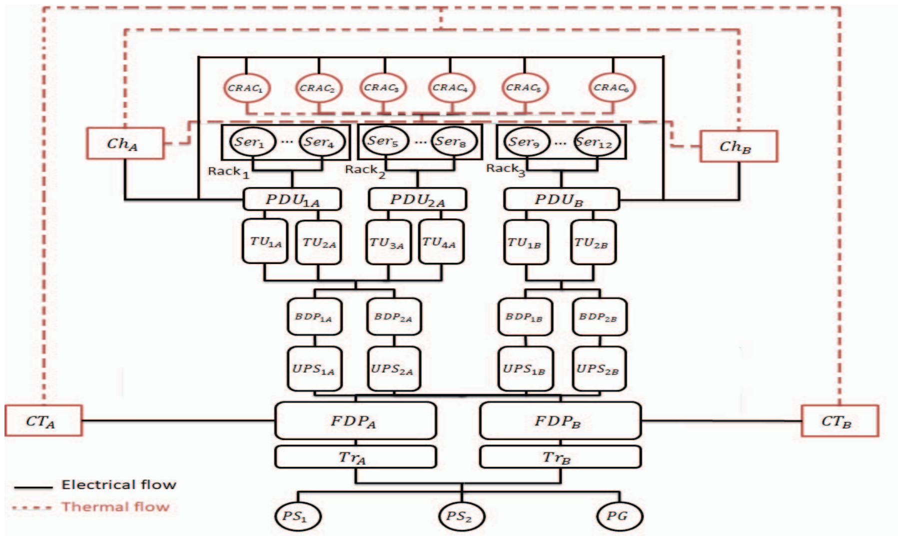

Figure 1. A typical thermal and power system of a DC

Transfer Units (TU). Finally, each two transfer units are connected to a Power Distribution Unit (PDU). Each PDU provides the electrical flow to four servers. The servers are noted Ser_j, $j = 1, \ldots, 12$, and are grouped in four racks (four servers per rack).

The second part of the system is the cooling system. All servers produce heat (known in the literature as *Joule heating*). Heat is energy and the common measure of heat output rate for equipment is $Joule/second$ ($1 Joule/sec$ equals $1 Watt$) [15].

Basically, the cooling system consists of three layers: the *production layer* which contains Cooling Towers (CT), the *chilling layer* which contains Chillers (Ch) and the *distribution layer* containing Computer Room Air Conditioning (CRAC) units [15]. The system considered in this paper consists of ten components: six CRAC units, two Chillers and two CTs. The two redundant cooling towers CT_A and CT_B in the *production layer* pump the water to other components and need to be powered to operate properly. The power energy is provided by FDP_A and FDP_B, respectively. The pumped water is routed to the *chilling layer* (Ch_A and Ch_B). The main role of a chiller is to cool the water as long as it is powered by the electrical system. Chillers Ch_A and Ch_B are powered by PDU_{1A} and PDU_B, respectively. Once the water chilled, it is delivered to the *distribution layer* containing CRAC units. Each CRAC unit extracts the air from the chilled water on condition that it is powered by at least one PDU. Finally, the CRAC units provide the cooled air to the servers. In this scenario each two redundant CRAC units provide air to a rack containing four servers and the cooling system is considered to be operational if at least one of the two CRAC units associated with each rack is working, the other one being in a standby mode or failure state (see Figure 1).

Electronic components are usually designed to function only within specified temperature ranges. Once this range is exceeded, their failure rates increase significantly. In fact, the failure rate of a semiconductor device doubles for every 10° *Celsius* increase [15]. Therefore, we consider the Arrhenius model in order to analyze the temperature variation effects on the the availability of the servers in the IT room. Arrhenius model relates the lifetime of an electronic component to the operating temperature [9]. The following equation estimates the relationship between the temperature and the Mean Time To Failure (MTTF) of the device:

$$r = A * e^{(\frac{-E_a}{K*TP})} \tag{1}$$

Where:

- r is the reaction rate. A measure to quantify the speed of a chemical reaction.

- TP is the temperature (in $^\circ$ *Kelvin*) at which component breaks down.
- K and A are, respectively, the Boltzmann constant and a pre-exponential constant.
- E_a is the activation energy usually within the range $[0.3eV - 0.7eV]$ [13].

In this paper, to assess the temperature impact TP on MTTF of the servers, we consider the initial temperature TP_0, such that $TP > TP_0$. This leads to the following MTTF expression:

$$MTTF_{TP} = MTTF_{TP_0} * e^{(\frac{E_a}{k} * (\frac{1}{TP} - \frac{1}{TP_0}))} \quad (2)$$

Equation 2 allows us to compute a new value of MTTF at elevated temperature [8].

IV. Production Trees

Production Trees (PT) are a new modeling methodology for production availability analysis [10]. They allow modeling the relationship between basic components of a system with a particular attention to the flows circulating between these components.

Production Trees provide two types of components to model a production system: basic components and gates. Basic components represent the production or treatment units of the system whereas the gates model the interactions between these units and thus the behavior of the whole system. Basic components are similar to basic events in a Fault Tree (FT). However, unlike the gates of FT, the gates of PT are not logical. They allow dealing with production flows upstream and downstream a production line, according to the type of these flows. Three types of flows circulate in a PT:

- *Capacity flow* moving forward (source to target).
- *Demand flow* moving backward (target to source).
- *Production flow* moving forward (source to target).

The production depends on the demand which itself depends on the capacity. First, each component (production unit) exports its actual production capacity, noted $outCapacity$. This capacity is null if the component is failed and equal to its intrinsic capacity ($intraCapacity$) otherwise. Then, the component receives a demand, noted $inDemand$, which, in stabilized situations, should not exceed the component capacity. Finally, the component exports a production ($outProduction$), which is the minimum of its actual capacity and the input demand. If the demand is null, the component is considered in standby mode.

In [10], three types of gates are defined: the *PLUS-gate*, the *MIN-gate* and the *SPLITTER-gate*.

1) The *MIN-gate*: it has one parent and two or more children. Its output capacity is the minimum of the output capacities of its children and of its intrinsic capacity. Its input demand (coming from its parent) is propagated unchanged to its children. Finally, its output production is the minimum of the output production of its children.

2) The *PLUS-gate*: it has one parent and several children. Its output capacity is the minimum of its intrinsic capacity and the sum of its children output capacities. Its input demand is propagated unchanged to its children. Finally, its output production is the sum of its children output productions. In the case where the output capacity of the gate is not equal to the output capacity of its children, the input demand of the gate is propagated to its children according to an allocation strategy. For example, the demand can be allocated according to a pro-rata of their capacities. Another strategy consists to allocate the maximum production to the first child, the maximum of the rest to the second child, etc (priority).

3) The *SPLITTER-gate*: unlike the other gates, this gate has only one child and several parents. Its output capacity is the minimum of its intrinsic capacity and the output capacity of its unique child. It is transmitted unchanged to its parents. The output demand of the gate is the sum of its parents demands. Finally, the output production of the gate is split among its parents following an allocation strategy (priority, pro-rata, ...), as for *PLUS-gate*. Figure 2 gives a graphical representation for *MIN-gate*, *PLUS-gate* and *SPLITTER-gate*. □

Currently, the gates defined in the PT modeling technique allow dealing with one kind of production flow at once. Therefore, in this paper, we extend this modeling technique by introducing a new gate, namely the *COND-gate* to be able to deal with different kinds of flows.

The *COND-gate:* it has one parent and several children. Each child represents a specific kind of flow and f is the dependency function associated to the gate. in our case f is the dependency function between the thermal system and the electrical system.

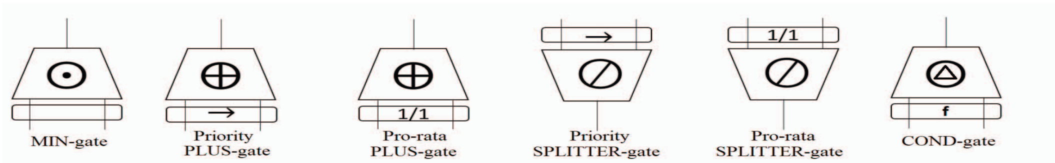

Figure 2. Graphical representation of gates of Production Trees

V. Modeling the system using PT

The electrical system is responsible for providing power to cooling system. To model the global system (thermal and electrical), we model first the thermal system, then the electrical one, taking into account the functional dependencies between them.

The production tree modeling the thermal system has to catch the different interactions between the production and the treatment units. Moreover, it has to take into account dependencies between this system and the electrical system. Indeed, thermal system components become operational only if they are powered by the electrical system. Thus, the production tree modeling the thermal system has to take into account two different kinds of flows: air flow and electrical flow. For that, we use the *COND-gate*, the new gate introduced in Section IV, which allows modeling dependency behaviors.

Building the PT model goes through 3 steps: the capacity exportation, the demand exportation and the production exportation.

Step 1: the first step is the transmission of production capacity of cooled air to the racks (servers). let start by the *production layer*. Both cooling towers CT_A and CT_B will export their production capacities $outCapacity_{CT_A}$ and $outCapacity_{CT_B}$, respectively, under the condition that they are powered by the electrical system. This is modeled using two *COND-gates*, one for each cooling tower. The power to CT_A and CT_B is transmitted by FDP_A and FDP_B, respectively. In order to simplify the graphical representation and prevent duplicating sub-branches at multiple tree locations, transfer functions represented by triangles 1 and 2 are used in Figure 3 to refer to the power path from FDP_A and FDP_B, respectively. Then, since the water production is the sum of CT_A and CT_B water production, the outputs of both *COND-gates* are combined using a *PLUS-gate*. Finally, the output capacity of this gate is propagated unchanged to the *chilling layer* using a *SPLITTER-gate*. As in the *production layer*, the same process is applied in the *chilling layer* and the *chilling layer*.

Step 2: once a server Ser_i, $1 \leq i \leq 12$ has been informed about the production capacity of cooled air, it sends its demand $outDemand_{ser_i}$. This demand is propagated unchanged through the PT and continues its traversal until it reaches the thermal production source, that is the cooling towers. Note that, when a demand is routed through a *COND-gate*, required power demand is sent to electrical system through corresponding power path.

Step 3: the third step is the cooled air transmission to servers according to the demand received by each CT. As in the first step, the production is distributed between components through PT gates.

The complete model is presented in Figure 3 in which, for the sake of readability, only the production flows are represented. □

Unlike the thermal components, the electrical components will not export their capacities. The servers export directly their demand, in term of energy, to the other system components. Then the power sources produce the energy taking into account their maximum capacities. This energy is transmitted to servers, through the other system components. Therefore building the PT model goes through 2 steps only: the exportation of the demand by the servers and the transmission of the energy production according to the demand received. Moreover as the electrical system does not depend on any other system, no *COND-gate* is required in the PT model. Because of the paper size constraint, the PT model of the electrical system is not provided.

VI. System Analysis

To analyze the system availability depicted in Figure 1, a new analytical approach we have recently developed based on the Probability Distributions of Capacity (PDC) is applied. This approach calculates

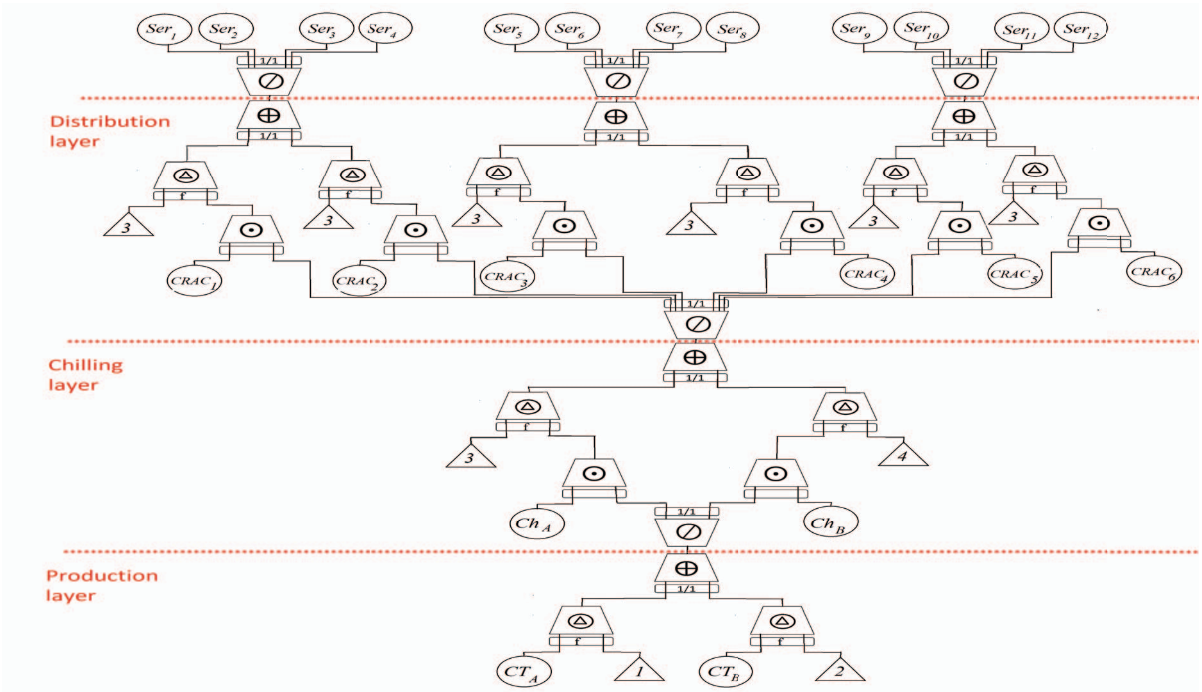

Figure 3. PT of the thermal System

the system availability using a set of predefined formulas. It is more concise and provides more accuracy in terms of system availability than simulation methods. The basic idea of this assessment algorithm of the PTs is inspired from [1].

We consider the components reliability data in [3]. The system capacity is 200 kW, provided by both PS_1 and PS_2 (100 kW each). The default temperature is 22° C which allows the servers to operate properly. The revised MTTF provided by Arrhenius model is inserted into our model.

To validate the results of our approach, we implement the PT models of both systems using the AltaRica 3.0 modeling language [12]. We use the stochastic simulator as the assessment tool.

To identify the dependencies between components and sub-systems, in the following, we study component's failure rates variation and its impact on the system availability.

Figure 4 provides the variation of failure rates of the power source PS_1 (the same for PS_2) and its impact on the system availability. According to these results, it is clear that the availability of the power source decreases when the failure rate increases. However, this does not affect the system availability since the power production is ensured by both PS_2 and PG (PS_1 and PG if PS_2 is considered). The simulation results of the AltaRica model match those obtained using our approach.

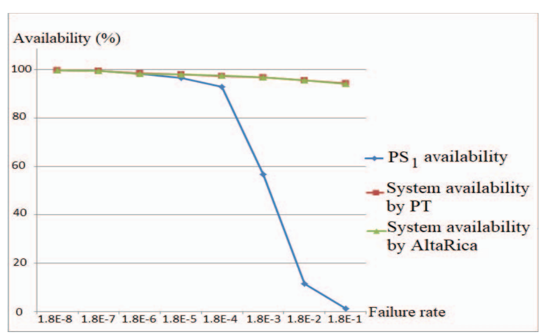

Figure 4. PS1 and system availability

Figure 5 shows the impact of the servers demand on the system temperature. The servers demand varies from $30kW$ to $240kW$. When the demand is less than $60kW$, the CRAC units produce air to extract the heat from the IT room and the temperature remains the same. When the demand exceeds $60kW$, the temperature within the IT starts increasing progressively because the electrical system is not able to produce a sufficient energy to satisfy the CRAC (thermal) demands. In this case the cooled air production decreases and leads to an increase in temperature.

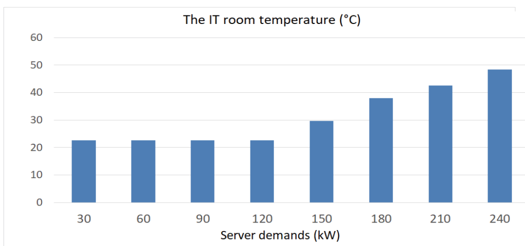

Figure 5. Electrical system impact on thermal system

Figure 6 shows the reliability of the system. According to the obtained results, the probability that the system will be fully operational is higher when the server demand is $10kW$. When the demand increases, the electrical system must provide a sufficient energy by activating components, initially in standby mode (PG for example), and this leads to decrease slowly the reliability of the system. When the demand is higher than $60kW$, the electrical system is able to generate the demanded energy but unable to generate sufficient energy for the thermal system which is responsible of extracting the heat from the IT room. This is consistent with the results obtained in Figure 5 which show the impact of the electrical system on the thermal system.

VII. Conclusion

The paper presents a formal model for estimating the system reliability in relation to the impact of temperature variation on IT system. For that we have used production tree modeling technique to model the electrical and thermal systems of the DC. To be able to model functional dependencies, we have extended PT by introducing a new gate. Using this technique, we were able to assess the impact of the cooling system on the IT system. As future work, we intend to consider the complete network system.

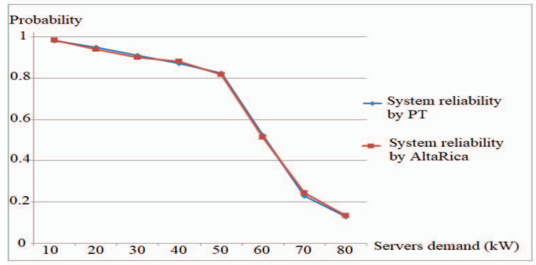

Figure 6. The system reliability

References

[1] User's guide for the UNIRAM version 4.1 for Windows—availability assessment methodology, Availability Systems . (1996)

[2] Ashrae, thermal guidleines for data processing environment. Ashrae 1st edition (2000)

[3] Component reliability data for use in probabilistic safety assessment. IEEE Gold Book 473 (2013)

[4] Andrews, J., Moss, T.: Reliability and risk assessment. John Wiley and Sons (1993)

[5] B, W., C, L., X, K.: Dependability modeling and analysis for the virtual data center of cloud computing. the IEEE 13th International Conference on High Performance Computing and Communications (HPCC) (2011)

[6] Beitelmal, A., Patel, C.: Thermo-fluids provisioning of a high performance high density. Springer Science (2006)

[7] Callou, G., Ferreira, J., Maciel, P., Tutsch, D., Souza, R.: An integrated modeling approach to evaluate and optimize data center sustainability, dependability and cost. Energies 2014 7(1), 238–277 (January 2014)

[8] G Arrheni, K.C., Wood., S.: A tribute to the memory of svante arrhenius . (2008)

[9] Jedec: Jedec global standards for the microelectronics industry. arrhenius equation for reliability (2005), http://www.jedec.org/standards-documents/dictionary/terms/arrhenius-equation-reliability

[10] Kloul, L., Rauzy, A.: Production trees: a new modeling methodology for production availability analyses. Reliability Engineering & System Safety 167, 561–571 (2017)

[11] Microsoft: Creating a greaner data center. (2013), http://www.microsoft.com/presspass/features/2009/apr09/04-02Greendatacenters.mspx

[12] Noble, J., Taivalsaari, A., Moore, I.: Prototype-based programming: Concepts, languages and applications. Springer-Verlag, Berlin and Heidelberg (1999)

[13] Onsemi: On semiconductor quality and reliability handbook. (2011), http://www.onsemi.com/publink/Collateral/HBD851-D.PDF

[14] R Souza, G Callou, K.C.J.F.P.M.: The effects of temperature variation on data center it systems. Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference. (2013)

[15] Rasmussen.l, N.: Calculating total tooling requirements for data centers. Schneider Electric White Pape 25r (2015)