Modeling and Analyzing Power System Failures on Cloud Services

Daniel Rosendo*, Patricia Takako Endo†, Glauco Gonçalves‡, André Moreira*,
Guto Leoni Santos *, Demis Moacir Gomes*, Judith Kelner*, Djamel Sadok* and Mozhgan Mahloo§
*Federal University of Pernambuco, Recife, Brazil
Email: {daniel.rosendo, andre, guto.leoni, demis.gomes, jk, jamel}@gprt.ufpe.br
†University of Pernambuco, Caruaru, Brazil
Email: patricia.endo@upe.br
‡Federal Rural University of Pernambuco, Recife, Brazil
E-mail: glauco.goncalves@ufrpe.br
§Ericsson Research, Stockholm, Sweden
E-mail: mozhgan.mahloo@ericsson.com

***Abstract*—Many enterprises rely on cloud infrastructure to host their critical applications (such as trading, bank's transaction, airline reservation system, and credit card authorization). The unavailability of these applications may lead to several consequences that go beyond the financial losses, reaching the cloud provider reputation too. However, maintain high availability in a cloud data center is a difficult task due to its complexity. The power subsystem is crucial for the entire operation of the data center because it supplies power for all other subsystems, including IT components and cooling equipment. Some studies have already proposed models to evaluate the availability of the power subsystem, but none of them are based on standards. Standards are a guide to cloud providers regarding availability, points of failure, and watts per square foot based on components' redundancy. This paper proposes Petri Net models based on the TIA-942 standard to estimate the availability of the data center power subsystem, it also provides a case study to analyze how failures on power subsystem impact on the critical applications availability. These models are important to resource planning and decision making by the cloud providers, because we can identify which components they can invest in order to improve the availability level. For instance, our results show that adding two redundant components can decrease the downtime in 3.32min/year.**

I. Introduction

A recent study analyzed the cost behavior of unplanned data center outages in the past 6 years, and found that the average cost has increased from $505,502 in 2010 to $740,357 in 2016 [1]. UPS (uninterruptible power supply) failure remains at the top of the list causing around 25% of unplanned failures; followed by DDoS attacks, human error, and cooling system failure. The UPS is one of many components that are part of the power subsystem. This subsystem is one of the largest and the most complex systems of a data center. It consists of many facilities and structures, components and equipment, with complex interactions among all those [2].

From the cloud provider, these failures are still hard to manage and prevent. Note also that the larger a data center is the greater is its downtime cost. A strategy is to apply component redundancy to ensure reaching higher availability at minimum costs.

Some standardization bodies work on defining a classification that allows comparing data centers according to their availability. Such classification may be based on Tiers (ITU and TIA-942 standards) or Classes of availability (BICSI-002 and EN-50600 standards) and describes how redundant components can be added to increase the availability level.

In this work, we propose a set of power subsystem models based on the TIA-942 standard; for that, we use Petri Net in order to understand how their failures impact on the overall data center availability. We also provide a sensitivity analysis to detect which power components are more sensible to failures and how they bias the availability level.

In summary, our contributions can be defined as: *(a)* to propose stochastic models based on TIA-942 standard to represent the power subsystem of a cloud data center; *(b)* to evaluate those models regarding their availability, reliability, and sensitivity; and *(c)* to model a critical application and propose a model to integrate the critical application with the power subsystem.

This paper is organized as follows: section II presents some basic concepts about data center infrastructure; section III describes some modeling techniques, such as RBD (Reliability Block Diagram) and Petri Net; section IV presents our stochastic models of power subsystem; section V has results regarding availability and sensitivity analysis; while section VI presents a case study with critical application; section VII describes some related work; and finally section VIII concludes our paper and delineates some future works.

II. Data Center Infrastructure

Data center standards define fundamental aspects, best practices, and recommendations regarding data center design and infrastructure. According to them, a generic data center system is basically composed of three subsystems: i) power infrastructure, ii) cooling infrastructure, and iii) IT infrastructure,

Those standards also define a classification that allows comparing data centers according to their availability level. Such classification may be based on Tiers (ITU and TIA-942 standards) or Classes of availability (BICSI-002 and EN-50600 standards). In this work, we focus on the power subsystem and our proposal models are based on the TIA-942 Tier classification.

Tier classification goes from I to IV and higher tiers inherit requirements of lower ones. For instance, in Tier I, there is no redundancy, Tier II must be less susceptible to system disruptions, Tier III may avoid system disruptions, and Tier IV should be fault tolerant. Higher tiers provide greater availability, which understandably result in higher costs and operational complexities. Therefore, the tier selection depends on the business requirements, such as minimum service availability, employment costs, and downtime financial consequences.

Basically, a tier is different from another tier due to the number of redundant components and distribution paths. The redundant components refer to the number of IT equipment, cooling, and power components that comprises the data center infrastructure. In Tier I, N means no redundancy indicating that system failures will result in outages. While in Tiers II, III, and IV $N+1$ means that there is some level of component redundancy. The number of delivery paths refers to the number of distribution paths of the power and cooling systems serving the IT equipment [3].

A. Power Subsystem

The data center power subsystem is responsible for feeding non-essential, essential, and critical loads. The non-essential (lighting, work stations, and supplementary equipment) loads may be interrupted without impacting the data center availability. The essential and critical loads impact in the data center availability. The essential loads refer to mechanical and cooling units. Essential loads affect indirectly the data center availability, that means, once the power to the cooling subsystem is interrupted, the IT equipment will continue operating for a while, until it gets warm, and turns off. Lastly, we have the critical loads composed of the IT equipment. Faults in the power components that feed the critical load directly affect the overall data center availability.

A typical data center power system infrastructure includes an utility substation, an alternate power source, a transfer switchgear or an Automatic Transfer Switch (ATS), an Uninterruptible Power Supply (UPS) system, and a Power Distribution Unit (PDU).

A data center is powered by an utility substation and may also contain an alternative power feed (such as solar, wind, bioenergy, hydroelectric and wave) [4]. Both primary and secondary power sources are connected to an ATS. The ATS provides input for the non-essential, essential, and critical loads. Following the critical load distribution path, the ATS feeds the UPS system (batteries). Then, the UPS system routes power to the PDU (rack socket for cabinets). Lastly, the PDU distributes electrical energy to the IT equipment.

According to the TIA-942 standard, in a Tier I data center power subsystem there is no redundancy. A single distribution path (non-redundant) feeds the IT equipment.The standard recommends $N+1$ redundant generator and UPS to the Tier II. As in Tier I, the Tier II has a single distribution path serving the computer equipment. Figure 1 shows a Tier II power subsystem with redundant generators and UPS.

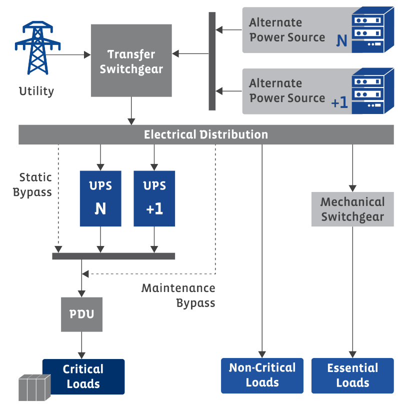

Fig. 1. Example of a Tier II power subsystem (adapted from [5])

III. Modeling Techniques

This section describes the modeling techniques used in our models, including RBD and Petri Net. We also present the Percentage Difference method, a method used to perform parametric sensitivity analysis.

A. Reliability Block Diagram

RBD is a graphical representation, based in block structures, that can be used to calculate dependability metrics about systems, such as reliability, availability, and maintainability [6] [7].

The availability can be expressed as the ratio of the service uptime and total service time. Total time is the sum of service uptime and service downtime. For the availability calculation, these concepts can be associated with average behavior of a system.

Considering a series configuration with N components, the availability of each component, A_x, is given by the division between MTTF (Mean Time To Failure) and the MTBF (Mean Time Between Failures) of each component (Eq. 1). In addition, MTBF is the sum of MTTF and MTTR (Mean Time to Repair), and indicates the time between detection of a failure and the detection of the next failure.

$$A_x = \frac{MTTF_x}{MTBF_x} = \frac{MTTF_x}{MTTF_x + MTTR_x} \quad (1)$$

Thus, the availability of the overall system, A_s, is calculated as shown in Eq.2.

$$A_s = \prod_{x=0}^{N} A_x \quad (2)$$

B. Petri Nets

A Petri Net comprises a state-space model, which allows to represent the system behavior through states and transitions.A Petri Net C can be defined as $C = (P, T, I, O)$, where $P = \{p_1, p_2, ..., p_n\}$ is a finite set of places, $n \geq 0$; $T = \{t_1, t_2, ..., t_m\}$ is a finite set of transitions, $m \geq 0$; $I : T \rightarrow P^{\infty}$ is the input function (that represents a mapping from transitions to a set of places); and $O : T \rightarrow P^{\infty}$ is the output function (that represents a mapping from transitions to a set of places).

Figure 2 illustrates a Petri Net components. The white circles (Figure 2.a) represent places, the passive components of a system. The rectangles comprise transitions, the active components. There are three types of transitions: timed (Figure 2.b), which fire after a time parameter related to a certain probability distribution; immediate (Figure 2.c), that fire instantly; and timed non-exponential (Figure 2.d). Arcs establish a connection between places and transitions. There are two arc types: directed (Figure 2.e) and inhibitor (Figure 2.f) a token (Figure 2.g) represents resources of a system, such as number of processors, tasks, and clients, for example. A place contains tokens, and a transition processes them.

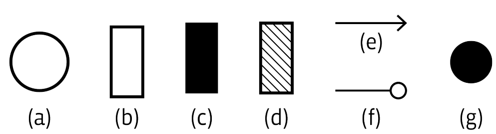

Fig. 2. Petri Net components

Petri Net is a specification method, that we can use to estimate the system availability (or any other metric). To solve it, one can use two options: *(a)* analytic solution by using Markov Chains, or *(b)* simulations using theory of discrete event simulation. In this work, a set of places and transitions representing a system component is called a **building block**, and we are using simulations to solve our Petri Nets.

C. Parametric Sensitivity Analysis

Parametric sensitivity analysis addresses the identification of factors for which the smallest variation implies the highest impact in models output measure [8], [9]. It predicts the effect on outputs with respect to variations in inputs, assisting to find bottlenecks, and conducting an optimization process [8].

There are many ways to perform a sensitivity analyses, such as the percentage difference, factorial experimental design, correlation analysis and regression analysis. In this work, we are using the Mercury tool [10] configured to perform the sensitivity analysis with percentage difference technique. Hence, it requires a minimum and maximum value for each parameter in order to compute the corresponding percentage variation on the chosen metric (see Eq. 3 from [8]).

$$S_{\theta}(Y) = \frac{max\{Y(\theta)\} - min\{Y(\theta)\}}{max\{Y(\theta)\}} \quad (3)$$

TABLE I
GUARD FUNCTIONS OF POWER SUBSYSTEM - TIER I AND II

Transition	Guard Function
PW_NET_6	#PW_GENERATOR_1_UP=0
PW_IT_17	((#PW_UTILITY_1_UP=0)AND(#PW_GENERATOR_1_UP=0)) OR(#PW_ATS_1_UP=0)
PW_IT_16	((#PW_UTILITY_1_UP>0)OR(#PW_GENERATOR_1_UP>0)) AND(#PW_ATS_1_UP>0)
PW_IT_4	(#PW_MAINSWB_1_UP=0)AND(#PW_UPSMODULE_1_UP>0) AND(#PW_BATTERY_1_UP=0)
PW_IT_3 PW_IT_5	#PW_MAINSWB_1_UP=1
PW_IT_2	((#PW_MAINSWB_1_UP=1)AND((#PW_UPSBYPASS_1_UP=0) AND(#PW_UPSMODULE_1_UP=0)))OR((#PW_MAINSWB_1_DOWN=1) AND(#PW_UPSMODULE_1_UP=0))OR((#PW_BATTERY_1_DOWN>0) AND((#PW_AVBATTERY=0)AND(#PW_BATTERY_1_UP=0)))
PW_IT_1	((#PW_MAINSWB_1_UP=1)AND((#PW_UPSBYPASS_1_UP>0) OR(#PW_UPSMODULE_1_UP>0)))OR((#PW_UPSMODULE_1_UP>0) AND(#PW_BATTERY_1_UP=1))
PW_IT_9	#PW_MAINSWB_1_UP=0
PW_IT_8	#PW_MAINSWB_1_UP=1
PW_IT_11	(#PW_SECONDSWB_1_UP=0)OR(#PW_PDU_1_UP=0)
PW_IT_10	(#PW_SECONDSWB_1_UP=1)AND(#PW_PDU_1_UP>0)

where $max\{Y(\theta)\} = max\{Y(\theta_1), Y(\theta_1), ..., Y(\theta_n)\}$; $min\{Y(\theta)\} = min\{Y(\theta_1), Y(\theta_1), ..., Y(\theta_n)\}$; and the parameter θ varies over the range of its n possible values of interest.

IV. POWER SUBSYSTEM MODELS

This section presents our models regarding power subsystem Tier I (subsection IV-A) and II (subsection IV-B). To better understand our proposal, we present the model divided in a set of operational building blocks.

A. Tier I

Each electrical component (utility, generator, ATS, UPS, and PDU) is modeled as a building block composed of two places (indicating its state, UP or $DOWN$) and two timed transitions (indicating its repair and failure rates). The list of all guard functions used in the Petri Net are described in Table I. Table II presents complementary variables of our models, that have constant values, such as the battery discharge time and the required time to start the generator.

TABLE II
ADDITIONAL VARIABLES

Variable	Value (h)	Comments	Transition
PW_DELAY_GENERATOR_1 [11]	0.05278	Time to start the generator	*PW_NET_6*
PW_DISCHARGE_BAT_1 [11]	0.117	Battery runtime	*PW_NET_15*

Next, we explain in detail these building blocks, how they are interconnected and their guard functions.

1) Utility, generator and ATS components: Figure 3 shows the Petri Net regarding utility (*PW_UTILITY_1_UP/DOWN*), generator (*PW_GENERATOR_1_UP/DOWN*), and ATS (*PW_ATS_1_UP/DOWN*) components, connected through the (*PW_MAINSWB_1_UP/DOWN*) building block.

The utility is the main power supply to the Data center and when it fails, the generator, if available

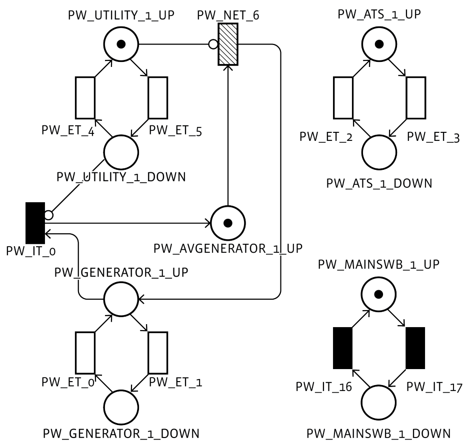

Fig. 3. Utility, generator, and ATS components

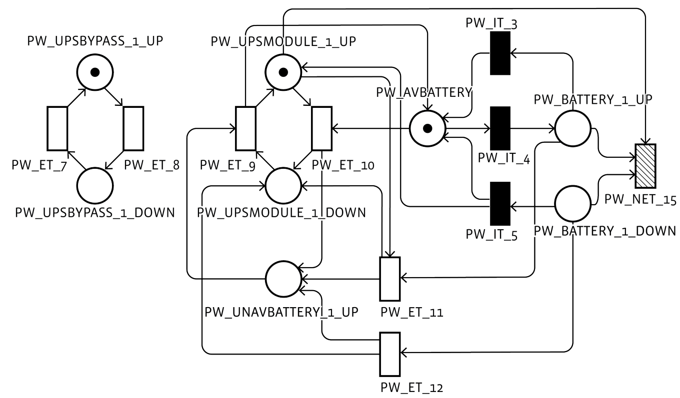

Fig. 4. The UPS components: static bypass, UPS module and batteries

(*PW_AVGENERATOR_1_UP*), will feed the data center. The time needed to start the generator is represented by the non-exponential transition *PW_NET_6*. Once the utility is repaired, the generator is turned off and it will be available to future needs (see the immediate transition *PW_IT_0* with its guard function).

The ATS component switches between utility and generator power sources and forwards the energy to the entire data center. There is a building block composed of places (*PW_MAINSWB_1_UP/DOWN*) that represents if the utility, generator, and ATS are feeding the data center (non-essential, essential, and critical loads).

In that case, they will be able to determine if the utility or the generator is working and the ATS is operating (immediate transition *PW_IT_16*). On the other hand, they will not be able to feed the data center if both utility and generator have failed or if only the ATS has failed (see the immediate transition *PW_IT_17*).

2) UPS component: The UPS component Petri Net is shown in Figure 4. The UPS is composed of a static bypass and a UPS module integrated with the batteries. We point out that the UPS only feeds the data center critical loads, such as the IT equipment. Normally, the energy from the ATS passes through the UPS and the UPS module, charging the batteries. Therefore, the static bypass is only used when the UPS module fails; so it acts as an alternative path inside the UPS component. In that configuration, the data center critical load will not be protected by the batteries. Hence, it will be more susceptible to disruptions.

As the UPS module and the batteries are strictly related to each other, we modeled them together. There are a UPS module and a battery available for use (see places *PW_UPSMODULE_1_UP/DOWN* and *PW_AVBATTERY*).

We modeled the failure of the UPS module while it is in the following three different states (1) not using the batteries (see the exponential transition *PW_ET_10*), (2) while using the batteries (see *PW_ET_11*), and (3) when the battery is discharged (see *PW_ET_12*). All of these failure states will put the UPS module in the place named *PW_UPSMODULE_1_DOWN* and its battery in place *PW_UNAVBATTERY_1_UP*.

The immediate transitions denote when the use of batteries is required or not. The batteries must be used while the generator is starting after a utility failure or if the utility, generator, and ATS are not able to feed the data center, in other words, if the state of the building block is *PW_MAINSWB_1_DOWN*.

Besides, to use the batteries, the UPS module must be working and we do not have another UPS module (batteries) feeding the IT equipment (in the case of redundant UPS components) (see the immediate transition *PW_IT_4* and its guard function). On the other hand, the use of batteries is not required anymore if the data center main power feed is working (*PW_MAINSWB_1_UP*). When the main feed become, the batteries may be in use (see the immediate transition *PW_IT_3*) or may be discharged (see *PW_IT_5*).

In order to simplify our model, we do not consider the time it takes to charge the batteries. But, we consider its discharging time while in use (see the non-exponential transition *PW_NET_15*). We also modeled the UPS static bypass (*PW_UPSBYPASS_1_UP/DOWN*), considering its two states (UP or DOWN) and failure and repair rates.

3) PDU component: The PDU Petri Net is shown in Figure 6, and the building block (*PW_SECONDSWB_1_UP*) represents when all the previous components are able to feed the PDU, connecting the building blocks described previously.

The building block *PW_SECONDSWB_1_UP/DOWN* has two immediate transitions. The first one defines the conditions in which we have energy to the PDU (see immediate transition *PW_IT_1* and its guard function). They are: (1) if *PW_MAINSWB_1_UP* is working and the UPS component (static bypass or UPS module) is able to forward the energy, or (2) if the UPS module is working and is using the battery power. On the other hand, second one determines the circumstances in which we do not have energy to the PDU (see immediate transition *PW_IT_2* and its guard function). They are: (1) if *PW_SECONDSWB_1_UP* works and both static bypass and UPS module of the UPS component are down, or (2) if *PW_SECONDSWB_1_DOWN* is not working and the

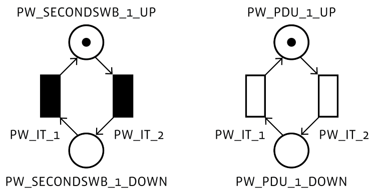

Fig. 6. The PDU component

UPS module (batteries) inside the UPS component is down, or lastly (3) if the battery is discharged (which signifies that the main power feed is down) and, in case of redundant UPS components, there is neither one available nor currently being used.

4) Essential and critical loads availability: Lastly, there are two building blocks regarding the power feed to the essential (mechanical and cooling) and critical loads (IT equipment), shown in Figure 7.

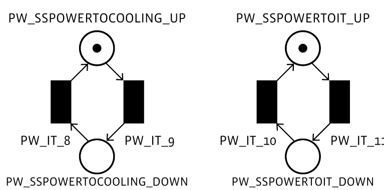

Fig. 7. Building blocks regarding the power feed to the essential and critical loads

These building blocks represent whether the electrical components are able or not to provide energy to the cooling and IT equipment (`PW_SSPOWERTOCOOLING_UP/DOWN` and `PW_SSPOWERTOIT_UP/DOWN`).

The building block of the essential load (utility, generator, and ATS) has two immediate transitions `PW_IT_8` and `PW_IT_9` and its state depends of the building block `PW_MAINSWB_1_UP` once the cooling feed comes from it. While, the building block of the critical load has two immediate transitions `PW_IT_10` and `PW_IT_11` and its state depends of the building block `PW_SECONDSWB_1_UP/DOWN` and the `PW_PDU_1_UP/DOWN` component, which implies that, the feed to the IT equipment depends of all electrical components in the power architecture.

The availability of the essential, A_{EL}, and critical, A_{CL}, load of the Tier I is given by Eq. 4 and 5, respectively.

$$A_{EL} = P\{\#PW\_SSPOWERTOCOOLING\_UP = 1\} \quad (4)$$

$$A_{CL} = P\{\#PW\_SSPOWERTOIT\_UP = 1\} \quad (5)$$

Hence, the probability of having a token in place `PW_SSPOWERTOCOOLING_UP` (for essential loads) and `PW_SSPOWERTOIT_UP` (for critical loads).

B. Tier II

As stated in section II, according to the TIA-942 standard, a Tier II data center power subsystem is composed of redundant components. The standard recommends $N+1$ generator and UPS redundancy.

As Tier I, the Tier II also has a single distribution path serving the IT equipment. Figure 5 shows our Petri Net model regarding a Tier II power subsystem.

It is important to highlight that the Tier II Petri Net is very similar to the Tier I Petri Net model, meaning that they have the same components, places, transitions, and guard functions (see Table I). However, they differ in the number of tokens in places regarding the generator (`PW_AVGENERATOR_1_UP`) and the UPS component (`PW_UPSBYPASS_1_UP`, `PW_UPSMODULE_1_UP`, and `PW_AVBATTERY`). These places have two tokens representing the Tier II redundant generator and UPS components.

The availability of Tier II is obtained by the same equation presented for the Tier I, Eq. 4 and 5

V. Evaluation Results

We evaluated the availability and sensitivity of Tier I and II power subsystem models using the Mercury tool. Furthermore,

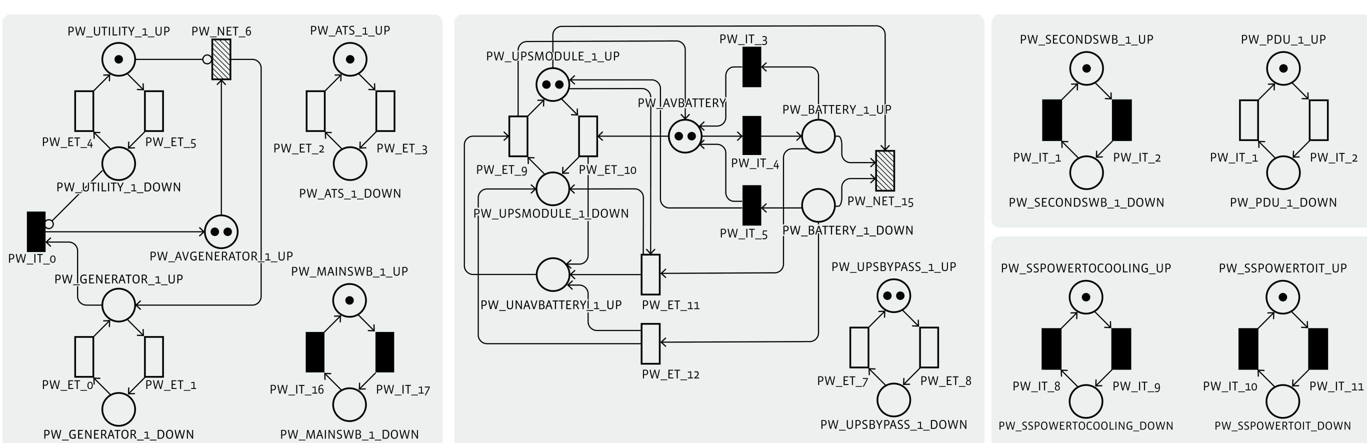

Fig. 5. Petri Net model regarding Tier II power infrastructure

after the sensitivity analysis we conducted a reliability analysis in the components with most impact on availability. We used stationary simulations with a confidence level of 98%, maximum relative error of 5%, and simulation time of 50000. Table III presents the MTTF and MTTR values of each power component we are considering in our models.

TABLE III
MTTF AND MTTR VALUES OF POWER SUBSYSTEM COMPONENTS

Component	MTTF (in hours)	MTTR (in hours)
Utility [11]	257.2678158	0.03280086594
Generator [11]	9,733.307378	3.9000039
ATS [11]	102,093.9468	5.739869131
UPS Bypass [12]	50,000.0	8
UPS Module [11]	27,472.52747	8
PDU [11]	282,581.0	156.0062402

A. Availability Analysis

The simulation results regarding availability level are shown in Table IV.

The Tier I provides an availability of 99.98414% for the mechanical and cooling equipment (essential load), and 99.93907% for the IT equipment (critical load). On the other hand, the Tier II provides a slightly higher 99.98421% and 99.93970% of availability level for the essential and critical loads, respectively.

TABLE IV
AVAILABILITY ANALYSIS REGARDING THE TIER I AND II POWER SUBSYSTEM

Tier	Load classification	Availability Level (%) Confidence Interval (98%)	Number of nines (9's)	Downtime (Hour/Year)
Tier I	Essential load	99.98414 [99.98414 - 99.98415]	3.79	1.38
	Critical load	99.93907 [99.93904 - 99.93910]	3.21	5.33
Tier II	Essential load	99.98421 [99.98421 - 99.98421]	3.80	1.38
	Critical load	99.93970 [99.93969 - 99.93972]	3.22	5.28

B. Sensitivity Analysis

The sensitivity analysis will allow us to find the power components with most impact in each subsystem availability, as well as the impact of design changes from Tier I to Tier II. We used the percentage difference method implemented in the Mercury tool (see subsection III-C).

1) Sensitivity Analysis - Tier I: Table V shows the sensitivity ranking of the top four power components with most impact on availability in tier I. Those components are the PDU failure rate (`PW_MTTF_PDU_1`), PDU repair rate (`PW_MTTR_PDU_1`), ATS failure rate (`PW_MTTF_ATS_1`), and lastly ATS repair rate (`PW_MTTR_ATS_1`).

As presented in Figure 8, from the baseline availability (point of interaction), an increase variation of 10% in the MTTF of the PDU resulted in an availability increase from 0.99939% to 0.99944% (or 0.47 hours). While, an increase

TABLE V
SENSITIVITY RANKING OF TIER I. TOP FOUR.

Parameter	Sensitivity Index
PW_MTTF_PDU_1	0.0001119
PW_MTTR_PDU_1	0.0001052
PW_MTTF_ATS_1	0.0000180
PW_MTTR_ATS_1	0.0000099

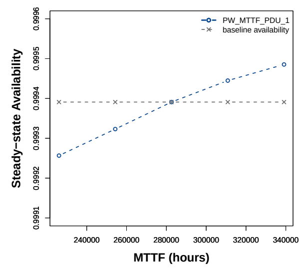

Fig. 8. Tier I: Sensitivity analysis of PW_MTTF_PDU_1.

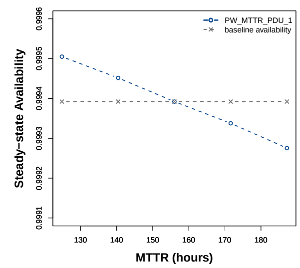

Fig. 9. Tier I: Sensitivity analysis of PW_MTTR_PDU_1.

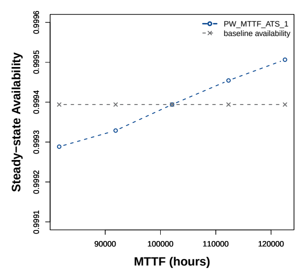

Fig. 10. Tier I: Sensitivity analysis of PW_MTTF_ATS_1.

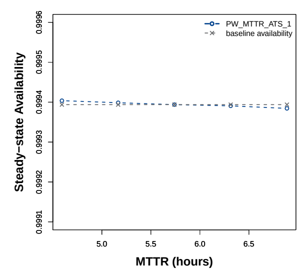

Fig. 11. Tier I: Sensitivity analysis of PW_MTTR_ATS_1.

variation of 20% resulted in an availability increase of 0.83 hours. Figure 10 represents the MTTR of the PDU. From the baseline availability, an increase of 15.6 hours in the repair time resulted in an availability decrease of 0.48% or 42.04 hours. On the other hand, decreasing its repair time in 10%, increased the up time to 0.52% or 45.5 hours.

Figure 10 and Figure 11 represent the availability variation of the ATS component. The ATS failure rate has a higher impact on availability than its repair rate. Therefore, investments in preventive maintenance in ATS may be a better choice.

2) Sensitivity Analysis - Tier II: Table VI shows the sensitivity ranking regarding the four power components with higher impact on availability.

TABLE VI
SENSITIVITY RANKING OF TIER II. TOP FOUR.

Parameter	Sensitivity Index
PW_MTTF_PDU_1	0.000116
PW_MTTR_PDU_1	0.000109
PW_MTTR_ATS_1	0.000014
PW_DELAY_GENERATOR_1	0.000013

The failure of the PDU component (*PW_MTTF_PDU_1*) has the greatest impact on the subsystem availability of Tier II. Following we have the PDU repair rate (*PW_MTTR_PDU_1*), ATS repair rate (*PW_MTTR_ATS_1*), and the delay to start the generator (*PW_DELAY_GENERATOR_1*).

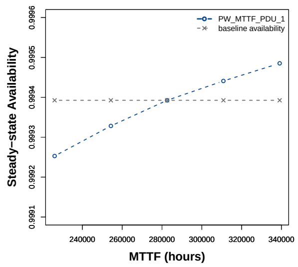

Fig. 12. Tier II: Sensitivity analysis of PW_MTTF_PDU_1.

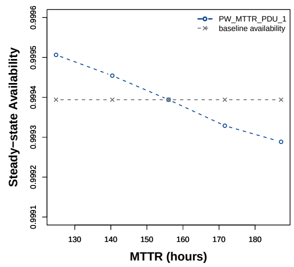

Fig. 13. Tier II: Sensitivity analysis of PW_MTTR_PDU_1.

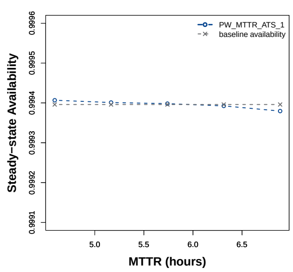

Fig. 14. Tier II: Sensitivity analysis of PW_MTTR_ATS_1.

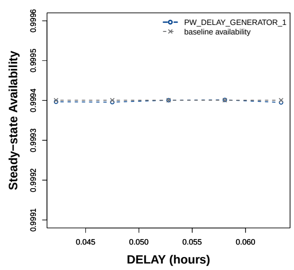

Fig. 15. Tier II: Sensitivity analysis of PW_DELAY_GENERATOR_1.

Figure 12 shows that an increase of 31.2 hours in the failure time of the PDU results in an availability increase of 0.81 hours. While, a decrease of 31.2 hours in its time to fail results in an decrease of 1.22 hours. Therefore, the decrease of the PDU failure time is more significant than its increase. In Figure 13 the impact of the repair time of the PDU in the overall availability varies from 0.98 hours (decreasing 31.2 hours) to 0.92 hours (increasing 31.2 hours).

Regarding the ATS component, Figure 14 shows that changes in its repair time do not significantly impact the overall power availability. Decreasing 1.14 hours from the baseline repair time (5.74 hours), we have an availability increase of 0.07 hours. In Figure 15, the impact of the increase of the generator start up time on availability varied from 99.9396% for a start delay of 2.53 minutes to 99.9395% for 3.8 minutes.

3) Reliability Analysis: Figure 16 and 17 show the estimated reliability curve regarding PDU and ATS power components. We chose them, once their parameters are the top ones with most impact on availability of both Tier I and II. We do not analyze the generator start up delay (*PW_DELAY_GENERATOR_1*) once its parameter does not refer to an MTTF value.

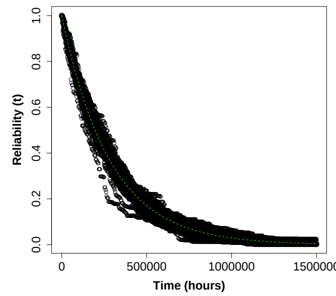

Fig. 16. PDU reliability curve.

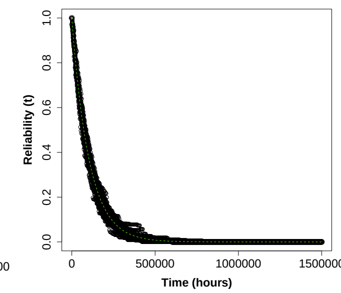

Fig. 17. ATS reliability curve.

We ran 30 transient simulations (black points) and performed a non-linear regression to plot the reliability curve (green line), $R(t)$. The results show that the ATS component is less reliable than the PDU component.

VI. Critical Application Availability

A critical application is essential to the survival of a business or enterprise. This way, when a critical application fails, business operations are significantly impacted. Credit card authorization is a good example of critical application.

We modeled a generic critical application using RBD. Next, we obtain its MTTF and MTTR reliability metrics in order to integrate the application to the Petri Net model that represents the power subsystem. With this, we can estimate the overall availability. We highlight that the availability evaluation considers the critical load (IT equipment) feeding.

Figure 18 shows the RBD model that represents the critical application. This system is composed of four serial components: hardware (HW), operating system (OS), virtual machine (VM) and the critical application (APP). The MTTF and MTTR values are described in Table VII.

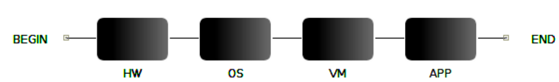

Fig. 18. RBD model of server components based from [7]

TABLE VII
RBD parameters obtained from [7], [13]

Components	MTTF (hours)	MTTR (hours)
HW	8760	1.667
OS	1440	1
VM	1880	0.167
APP	240	0.056

By using the Structure Function method available in Mercury Tool, we obtained the reliability results regarding the RBD model (see Table VIII). These values will be used in the Petri Net model described next.

Figure 19 shows the Petri Net model that integrates our power subsystem to the critical application. On the left side,

TABLE VIII
RELIABILITY ANALYSIS OF THE CRITICAL APPLICATION

Metric	Result	Metric	Result
MTTF	181.58	Number of Nines	2.91
MTTR	0.21	Uptime (hours)	8755.24
Availability	0.99879	Downtime (hours)	10.57

the building block represents the power feeding the IT subsystem. The guard functions (*PW_IT_10* and *PW_IT_11*) are the same for Tier I and II and were described in Table I. On the right side, the application building block is composed of two places *CRITICAL_APP_UP* and *CRITICAL_APP_DOWN*, representing the application status. It also has two exponential transitions *IT_ET_1* and *IT_ET_2*, which MTTR and MTTRF values were obtained from the RBD (see Table VIII).

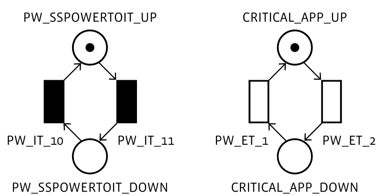

Fig. 19. Petri Net model of the integration between the power subsystem and the critical application

The application availability integrated to the power subsystem is given by the following Eq. 6.

$$A_{app} = P\{\{\#PW\_SSPOWERTOIT\_UP = 1\} \\ AND\{\#CRITICAL\_APP\_UP = 1\}\} \quad (6)$$

Therefore, the probability of both the power to IT subsystem and the critical application are working properly.

Table IX presents the availability results of the simulation. Tier I provides an availability of 99.8186%, and the Tier II 99.8190%, being 1.5 minutes more available than Tier I. To critical applications, such as PayPal, this little difference of 1.5 minutes can represent a meaningful lost of $ 5,625 (considering that PayPal has a downtime cost of 225,000 USD/hour [14]).

TABLE IX
AVAILABILITY OF THE CRITICAL APPLICATION INTEGRATED TO THE POWER SUBSYSTEM

Tier	Load classification	Availability Level (%) Confidence Interval (98%)	Number of nines (9's)	Downtime (Hour/Year)
Tier I	Critical load	99.81861 [99.81860 - 99.81862]	2.741	15.88
Tier II	Critical load	99.81900 [99.81897 - 99.81903]	2.742	15.85

VII. RELATED WORK

In [11], authors explain about different UPS configurations, showing benefits and limitations of each one. The architectures were based in TIA-942 standard. Five power architectures were modeled using Markov chain in order to evaluate their availability levels.

In [15], authors modeled the power system of data center based on TIA-942 standard. The four tiers were modeled using a simplified model based in blocks interconnected to evaluate the availability.In order to improve the availability, they added an ATS mechanism that provides power all time, even if the primary bus fails or if it is shut down for maintenance. Results showed that this mechanism increased considerably the availability when compared with previous architectures.

Govindan, et al. [16] aimed to model a power infrastructure through CTMC (Continuous Time Markov Chains) and RBD to estimate the cost of the availability. The authors complement the work described in [11] varying UPS placement and redundancy, in order to study the best topology to decrease the costs. Results showed that the use of a hybrid topology, which utilizes UPS at rack-level and at server-level, provides a six nines of availability with a 34% cost reduction in relation to a centralized rack-level UPS with a five nines of availability.

Callou, et al. [12] estimated the availability, sustainability, and cost of power and cooling infrastructures using SPN, Energy Flow Model (EFM), and RBD approaches. Although the authors presented four architectures, they do not follow any standard.

This paper presented models that follow the TIA-942 standard in a higher granularity level in comparison with the existing works, modelling many components not covered by them. Additionally, we also carried out sensitivity analysis, in order to check which components most impact on the power subsystem availability.

VIII. CONCLUSIONS AND FUTURE WORKS

This paper presented a set of data center power subsystem models based on TIA-942 standard. We modeled a critical application and integrated it with the power subsystem model to analyze the impact of power failures in the application availability. This work is useful in answering data center planning decisions such as by investing into a subsystem how component redundancy reduces downtime. Furthermore, the results allow data center operators to make better design decisions considering critical components in order to increase the overall data center availability.

The availability results showed that using redundant generator and UPS (tier II) provided a higher availability than tier I (no redundancy). The sensitivity analysis results showed that PDU and ATS are the two components that most impact on data center availability.

As future work, we plan to propose models regarding tier III and IV of power subsystemWe also plan to integrate the three basic subsystem of a data center (power, cooling and IT) and evaluate how failures of those subsystems affect the overall data center availability.

ACKNOWLEDGMENT

This work was supported by the RLAM Innovation Center, Ericsson Telecomunicações S.A., Brazil.

References

[1] P. Institute, "Cost of data center outages: Data center performance benchmark series," 2016.

[2] M. Cepin, *Assessment of power system reliability: methods and applications*. Springer Science & Business Media, 2011.

[3] H. Geng, *Data center handbook*. John Wiley & Sons, 2014.

[4] W. Kao and H. Geng, "Renewable and clean energy for data centers," *Data Center Handbook*, pp. 559–576, 2015.

[5] "Ansi/bicsi 002, data center design and implementation best practices." https://www.bicsi.org/uploadedFiles/BICSI_Website/Global_Community/Presentations/CALA/Ciordia_002_Colombia_2016.pdf. Last access: November, 2016.

[6] A. K. Verma, A. Srividya, and D. R. Karanki, *Reliability and safety engineering*, vol. 43. Springer, 2010.

[7] J. Araujo, P. Maciel, M. Torquato, G. Callou, and E. Andrade, "Availability evaluation of digital library cloud services," in *Dependable Systems and Networks (DSN), 2014 44th Annual IEEE/IFIP International Conference on*, pp. 666–671, IEEE, 2014.

[8] R. d. S. M. Junior, A. P. Guimaraes, K. M. Camboim, P. R. Maciel, and K. S. Trivedi, "Sensitivity analysis of availability of redundancy in computer networks," *CTRQ 2011*, p. 122, 2011.

[9] D. Hamby, "A review of techniques for parameter sensitivity analysis of environmental models," *Environmental monitoring and assessment*, vol. 32, no. 2, pp. 135–154, 1994.

[10] B. Silva, R. Matos, G. Callou, J. Figueiredo, D. Oliveira, J. Ferreira, J. Dantas, A. Lobo, V. Alves, and P. Maciel, "Mercury: An integrated environment for performance and dependability evaluation of general systems," in *Proceedings of Industrial Track at 45th Dependable Systems and Networks Conference, DSN*, 2015.

[11] K. McCarthy and V. Avelar, "Comparing ups system design configurations," *APC white paper 75, Schneider electric Data center science center*, 2005.

[12] G. Callou, P. Maciel, D. Tutsch, J. Ferreira, J. Araújo, and R. Souza, "Estimating sustainability impact of high dependable data centers: a comparative study between brazilian and us energy mixes," *Computing*, vol. 95, no. 12, pp. 1137–1170, 2013.

[13] M. Martinello, M. Kaaniche, and K. Kanoun, "Web service availability-impact of error recovery and traffic model," *Reliability Engineering & System Safety*, vol. 89, no. 1, pp. 6–16, 2005.

[14] C. Cérin, C. Coti, P. Delort, F. Diaz, M. Gagnaire, Q. Gaumer, N. Guillaume, J. Lous, S. Lubiarz, J. Raffaelli, *et al.*, "Downtime statistics of current cloud solutions," *International Working Group on Cloud Computing Resiliency, Tech. Rep*, 2014.

[15] E. N. Power, "Using static transfer switches to enhance data center availability and maintainability," 2010.

[16] S. Govindan, D. Wang, L. Chen, A. Sivasubramaniam, and B. Urgaonkar, "Towards realizing a low cost and highly available datacenter power infrastructure," in *Proceedings of the 4th Workshop on Power-Aware Computing and Systems*, p. 7, ACM, 2011.