How to improve cloud services availability? Investigating the impact of power and IT subsystems failures

Daniel Rosendo*, Guto Leoni*, Demis Gomes*, André Moreira*, Glauco Gonçalves†,
Patricia Takako Endo‡, Judith Kelner*, Djamel Sadok*, and Mozhgan Mahloo§

*Federal University of Pernambuco, Recife, Brazil
Email: {daniel.rosendo, guto.leoni, demis.gomes, andre, jk, jamel}@gprt.ufpe.br
†Federal Rural University of Pernambuco, Recife, Brazil
Email: glauco.goncalves@ufrpe.br
‡University of Pernambuco, Caruaru, Brazil
Email: patricia.endo@upe.br
§Ericsson Research, Sweden
Email: mozhgan.mahloo@ericsson.com

Abstract

The cloud data center is a complex system composed of power, cooling, and IT subsystems. The power subsystem is crucial to feed the IT equipment. Power disruptions may result in service unavailability. This paper analyzes the impact of the power subsystem failures on IT services regarding different architecture configurations based on TIA-942 standard such as non-redundant, redundant, concurrently maintainable, and fault tolerant. We model both subsystems, power and IT, through Stochastic Petri Net (SPN). The availability results show that a fault tolerant power and IT configuration reduces the downtime from 54.1 to 34.5 hours/year when compared to a non-redundant architecture. The sensibility analysis results show that the failure and repair rates of the server component in a fault tolerant system present the highest impact on overall data center availability.

1. Introduction

According to Gartner, Inc[1], a company that does not employ the cloud will be as obsolete as the companies out of the Internet today; the migration to a cloud-based solution is practically unavoidable. Therefore, the cloud is essential to companies to provide their services with higher availability. Due to a growing number of enterprises using services on cloud infrastructure, many challenges begin to emerge. Cloud providers ought to estimate how physical and logical failures occurring on a data center can impact their users' perception, in order to reduce the service downtime and improve their service availability.

Service downtime affects user experience and is directly translated into revenue loss to the cloud provider [10]. The Ponemon Institute study [12] showed that among the data center interruptions origins, Uninterruptible Power Supply (UPS) failure appears at the top of the list causing around 25% of unplanned interruption.

Recently, an IT (Information Technology) failure in the British Airways data center resulted in 600 flights canceled affecting about 75,000 passengers with a cost of $112 million. The IT failure occurred due to a problem in the power subsystem. According to data center specialists, after IT problems, power disruptions are the second most common cause of data center unavailability[2]. Therefore, a cloud provider needs to understand the failures at power subsystem and their impacts on IT subsystem in order to avoid or mitigate service downtime. Several components may impact overall data center availability. Hence, identifying most critical components is useful to improve availability by increasing their redundancy or replacing them by more reliable components. Due to the difficulties on carrying out studies in an operational and real data center, we adopt modeling techniques to estimate the data center availability, and use real data from the literature.

This paper presents a set of stochastic models to represent the data center infrastructure based on Telecommunications Industry Association standard (TIA-942), focusing on two main subsystems, power and IT. We also present a model that integrates both power and IT subsystems. The main goal is to analyze the availability level of the cloud services hosted in a data center and how component redundancy can improve the overall data center availability.

---

[1]http://www.gartner.com/newsroom/id/3354117

[2]http://www.datacenterknowledge.com/archives/2017/05/30/british-air-data-center-outage-feeds-outrage-at-airline-cost-cuts/

HICSS

Moreover, we also evaluate, through sensitivity analysis, how components' failures and repair rates impact on cloud services' availability.

2. Background

In this section, we describe data center standards for availability evaluation, focusing on power and IT subsystems. Next, we present a brief description about Petri Nets.

2.1. Data Center Standardization

Data center standards define best practices and recommendations regarding data center design and infrastructure. Such standards denote the three subsystems that composes a data center: power, cooling and IT.

According to TIA-942 standard, a tier is different from another one regarding the number of redundant components (i.e., N means no redundancy; and $N + 1$ means redundancy with one component) and distribution paths (i.e. single or multiple paths that may be active or passive). TIA-942 defines four tiers, from I to IV, where higher tiers provide greater availability, resulting in higher costs and operational complexities. Therefore, the tier selection depends on the business requirements, such as required service availability, deployment costs, and downtime financial consequences.

2.2. Power subsystem components

The data center power subsystem is responsible for feeding non-critical, critical, and mechanical loads (Figure 1(a)). The non-critical loads may be interrupted without impacting the data center availability and consists of lighting, work stations, and supplementary equipment. On the other hand, the mechanical and critical loads impact the data center availability. The mechanical loads refer to cooling units. Mechanical loads affect indirectly the data center availability, i.e, once the power to the cooling subsystem is interrupted, the IT equipment will be operating for a while, until it gets overheated, and shut downs. 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 power system infrastructure includes an utility substation, an alternate power source, a transfer switch gear or an Automatic Transfer Switch (ATS), an UPS system, and a Power Distribution Unit (PDU). The main power supply of a data center is the utility substation. Data centers may also contain an alternative power feed (such

as solar, wind, bioenergy, hydroelectric and wave) [13]. Both primary and secondary power sources are connected to an ATS. The ATS provides input for the non-critical, critical, and mechanical 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 power to the IT equipment.
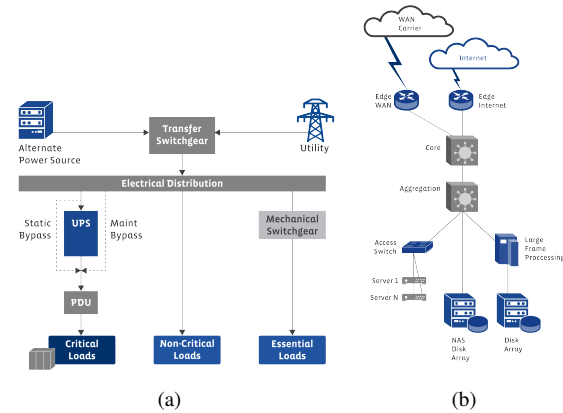

Fig. 1.   Tier I of (a) Power and (b) IT Subsystems (adapted from [1]).

2.3. IT subsystem components

A data center IT subsystem is basically composed of servers, storage, and network components. The storage is illustrated as Network Attached Storage (NAS) Disk Array. Network is represented by Edge, Core, Aggregation routers, and Access switch (Figure 1(b)). The Storage Area Network (SAN) is a network component used in Tier IV to connect array disks to servers.

Servers host applications and is composed of CPUs, NICs, and RAMs. The storage system stores server's application data and consist of disk drives or flash devices connected through a SAN. Technologies such as NAS and Redundant Array of Independent Drives (RAID) connects servers to remote storage systems and provides high-availability, respectively. Lastly, networking devices connect servers and storage systems and manage all data flow from/to the data center. The network design is typically based on the core, distribution, and edge levels and relies on switches, hubs, routers, among others.

2.4. Petri Nets

A Petri Net is a well-known model to model system with respect to evaluate performance and dependability [6]. White circles represent places, receiving tokens, where a

set of tokens (markup) represents the state of a system, such as number of processors, tasks, and clients. The rectangles comprise transitions, executing actions such as consume and/or produce tokens. There are three types of transitions: timed exponential, which fires after a time parameter related to a exponential distribution; immediate, that fires instantly; and timed non-exponential. Arcs establish a connection between places and transitions. There are two arc types: directed and inhibitor.

Transitions are only enabled to fire if all preconditions are fulfilled as, for example, if there are enough tokens in the input places. When a transition fires and it is connected with input and output places via a directed arc, the transition consumes an amount of tokens from its input places and produces tokens at all of its output places. Otherwise, if the arc is an inhibitor one, the precondition on firing is that there is no token in the input place. Transitions also can be activated with a guard function, a boolean expression composed of places, transitions, and tokens. For instance, consider the guard function $GF_{systemDown} = (\#machine_{on} = 0) AND(\#routers_{on} = 0)$. It means that the transition $GF_{systemDown}$ will be enabled only if there are no tokens in places $\#machine_{on}$ and $\#routers_{on}$.

The stochastic transitions can have two different firing semantics: single server and infinite server. In case of single server semantics, when a transition fires, a new delay is associated with this transition, and the fires are sequential. In case of infinite server transition, a set of tokens in determined place are processed in parallel and one fire does not interfere in others [5].

Petri Nets 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 Chain, where all transitions follows exponential distributions; or *(b)* simulations using theory of discrete event simulation. In this work, we use the term **building block** to refer to a set of places and transitions representing a system component.

2.5. Reliability Block Diagram

Reliability Diagram Block (RBD) is a mathematical tool that can be used to calculate reliability, availability, and maintainability of systems. The components of systems is represented by a set of blocks, that can be arranged logically in series or in parallel. In series configurations, if one component fails, the overall system will fail too. However, in parallel configuration, if only one component is working, the system will be operational [4].

3. Power Subsystem Model

This Section describes a data center power infrastructure regarding Tiers I to IV. The following models are based on the TIA-942 standard.

3.1. Power Subsystem Tier I

A data center power subsystem is mainly composed of utility, generator, ATS, UPS, and PDU. In our models, we represent those components as a building block containing two places and two transitions (may be timed or immediate transitions). The places refer to the state of the component, that may be *UP* (available) or *DOWN* (unavailable). The transitions indicate the repair and failure rates of the component and may have guard functions. All the guard functions used in our models (Tiers I to IV) are listed in Table I. In the next subsections, we explain in detail the power components and how they are interconnected.

TABLE I
GUARD FUNCTIONS OF POWER SUBSYSTEM - TIERS I, II, III AND IV

Transition	Guard Function
PW_IT_1	$(\#PW\_GENERATOR\_1\_UP = 0)$
PW_IT_9	$(\#PW\_GENERATOR\_2\_UP = 0)$
PW_IT_3	$((\#PW\_UTILITY\_1\_UP{>}0) OR (\#PW\_GENERATOR\_1\_UP{>}0))$ $AND (\#PW\_ATS\_1\_UP{>}0)$
PW_IT_11	$((\#PW\_UTILITY\_2\_UP{>}0) OR (\#PW\_GENERATOR\_2\_UP{>}0))$ $AND (\#PW\_ATS\_2\_UP{>}0)$
PW_IT_4	$((\#PW\_UTILITY\_1\_UP = 0) AND (\#PW\_GENERATOR\_1\_UP = 0))$ $OR (\#PW\_ATS\_1\_UP = 0)$
PW_IT_12	$((\#PW\_UTILITY\_2\_UP = 0) AND (\#PW\_GENERATOR\_2\_UP = 0))$ $OR (\#PW\_ATS\_2\_UP = 0)$
PW_IT_5	$(\#PW\_UPSMODULE\_1\_UP{>}0) OR (\#PW\_MAINSWB\_1\_UP = 1)$
PW_IT_13	$(\#PW\_UPSMODULE\_2\_UP{>}0) OR (\#PW\_MAINSWB\_2\_UP = 1)$
PW_IT_6	$(\#PW\_UPSMODULE\_1\_UP = 0) AND (\#PW\_MAINSWB\_1\_UP = 0)$
PW_IT_14	$(\#PW\_UPSMODULE\_2\_UP = 0) AND (\#PW\_MAINSWB\_2\_UP = 0)$
PW_IT_7 *(Tier I and II)*	$(\#PW\_SECONDSWB\_1\_UP = 1) AND (\#PW\_PDU\_1\_UP{>}0)$
PW_IT_8 *(Tier I and II)*	$(\#PW\_SECONDSWB\_1\_UP = 0) OR (\#PW\_PDU\_1\_UP = 0)$
PW_IT_7 *(Tier III)*	$((\#PW\_SECONDSWB\_1\_UP = 1) AND (\#PW\_PDU\_1\_UP{>}0))$ $OR ((\#PW\_MAINSWB\_2\_UP = 1) AND (\#PW\_PDU\_2\_UP{>}0))$
PW_IT_8 *(Tier III)*	$((\#PW\_SECONDSWB\_1\_UP = 0) OR (\#PW\_PDU\_1\_UP = 0))$ $AND ((\#PW\_MAINSWB\_2\_UP = 0) OR (\#PW\_PDU\_2\_UP = 0))$
PW_IT_7 *(Tier IV)*	$((\#PW\_SECONDSWB\_1\_UP = 1) AND (\#PW\_PDU\_1\_UP{>}0))$ $OR ((\#PW\_SECONDSWB\_2\_UP = 1) AND (\#PW\_PDU\_2\_UP{>}0))$
PW_IT_8 *(Tier IV)*	$((\#PW\_SECONDSWB\_1\_UP = 0) OR (\#PW\_PDU\_1\_UP = 0))$ $AND ((\#PW\_SECONDSWB\_2\_UP = 0) OR (\#PW\_PDU\_2\_UP = 0))$

3.1.1 Utility, generator and ATS components

In a tier I data center, the utility (*PW_UTILITY_1_UP/DOWN*) is the main power supply. A generator (*PW_GENERATOR_1_UP/DOWN*) if available (*PW_AVGENERATOR_1_UP*) is used when the utility fails (*PW_IT_1*). We highlight that in order to simplify our models and to obtain the simulation results of tiers III and IV in a feasible time, we do not consider the required time to start the generator. The

guard function of this immediate transition ensures that a single generator is turned on at a time. On the other hand, the transition (*PW_IT_2*) turn off the generator after the utility recovery.
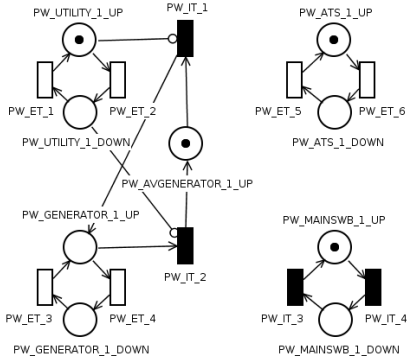

Fig. 2.   Utility, generator, and ATS components

The ATS (*PW_ATS_1_UP/DOWN*) component switches between utility and generator power sources. The ATS supplies power to the inside of the data center. Lastly, we have the building block (*PW_MAINSWB_1_UP/DOWN*) that put together the utility, generator, and ATS components, indicating if they are able or not to provide energy to the data center. Therefore, it will be able if (*PW_IT_3*) the utility or the generator is available and the ATS is working. On the other hand, it won't be able if (*PW_IT_4*) both utility and generator fails or the ATS fails. Note that the ATS is a single point of failure in the tier I power subsystem. The Figure 2 shows these power components.

3.1.2   UPS and PDU components

Following the description of the components of the power distribution path in a data center, next we present the UPS and PDU components. The UPS (*PW_UPSMODULE_1_UP/DOWN*) aims to provide energy to critical loads (IT equipment) when the utility and the generator are unavailable. The PDU (*PW_PDU_1_UP/DOWN*) distributes energy to the racks to feed the IT equipment.

We also have two more building blocks. The first one (*PW_SECONDSWB_1_UP/DOWN*) has the same idea of the *PW_MAINSWB_1_UP/DOWN* but including different components, to simplify, in our model it includes only the UPS module, but it could also include batteries or a bypass, for instance. The secondary switchboard (*PW_SECONDSWB_1_UP/DOWN*) will be able to forward energy to the PDU component if (*PW_IT_5*) the building block *PW_MAINSWB_1_UP/DOWN* or the UPS component is available. Otherwise (*PW_IT_6*) it will not be able to feed the IT equipment. The second one

(*PW_SSPOWERTOIT_UP/DOWN*) represents the overall (all power components) power subsystem, indicating if they are able to provide (*PW_IT_7*) energy to the IT equipment or not (*PW_IT_8*). As described on these guard functions the IT equipment will be powered if both *PW_SECONDSWB_1_UP* and PDU components are available. Otherwise, the data center will be unavailable. The Figure 2 shows these power components.
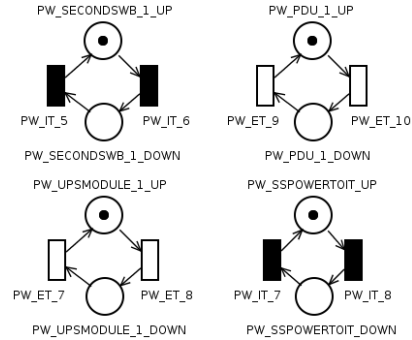

Fig. 3.   The UPS and PDU components

3.2. Power Subsystem Tier II

Tier II data center power system has redundant generator and UPS in addition to all components of tier I. Therefore, we have $N + 1$ generator (*PW_AVGENERATOR_1_UP*) and UPS (*PW_UPSMODULE_1_UP/DOWN*). This redundant configuration eliminates the single point of failure in the UPS component, providing a higher availability than the tier I. Both UPS modules work simultaneously and synchronizes the output voltage to the IT equipment. The guard functions are described in Table I.

3.3. Power Subsystem Tier III

In a tier III power subsystem we have redundant components and two independent distribution paths serving the computer equipment (N+1 generator and UPS in the primary path). A tier III configuration is concurrently maintainable, allowing the maintenance of components in a distribution path without system disruptions. Besides, all IT equipment is dual powered (each one linked to independent distribution paths).

The guard functions are described in Table I. We highlight that, differently from tier I and II, in a tier III configuration the power subsystem will not be able to feed the IT equipment if (*PW_IT_8*) the *PW_SECONDSWB_1_UP/DOWN* or PDU component of the primary distribution path fails along with a failure in the *PW_MAINSWB_2_UP/DOWN* or PDU component in

the secondary path. Otherwise (*PW_IT_7*), the subsystem is able to provide energy to the critical load.

3.4. Power Subsystem Tier IV

The tier IV power subsystem configuration is composed of two complete (all power components) distributions paths serving IT equipment (N+1 generator and UPS in both paths). This configuration is fault tolerant and the main power feed are the UPS modules. Besides, it eliminates single point of failure. The tier IV model was modeled as show in Figure 4. Differently from the tier III, tier IV power subsystem has a fully redundant distribution path.

The guard functions are described in Table I. The building block (*PW_MAINSWB_2_UP/DOWN*) puts together the utility, generator, and ATS components of the secondary distribution path and indicates if they are able (*PW_IT_11*) or not (*PW_IT_12*) to provide energy.

We also point out that, differently from tier III, the critical load in the second distribution path of tier IV may be powered by the utility, generator, or UPS modules, providing a higher availability. The building block (*PW_SECONDSWB_2_UP/DOWN*) of the secondary path will be able to forward energy to the PDU component if (*PW_IT_13*) the *PW_MAINSWB_2_UP/DOWN* or the UPS component is available. Otherwise (*PW_IT_14*) it will not be able. The tier IV subsystem will be able to feed the IT equipment if (*PW_IT_7*) both *PW_SECONDSWB_1_UP* and PDU component of the first distribution path are available or both *PW_SECONDSWB_2_UP* and PDU component of the second path are available.

4. IT Subsystem Model

In our models, we are considering an application running on top of the IT infrastructure in the data center. To model service, we used a Reliability Block Diagram (RBD) to represent dependency between service components. The IT infrastructure is comprised by network components, storage, and servers. However, the behavior of the IT infrastructure components are modeled using SPN. These models are described next.

4.1. RBD model of service

Figure 5 shows the RBD that represent service component. We considered that a server is composed of: hardware (HW), operating system (OS), virtual machine (VM) and the application (APP) instance that is running on this server. It is worth mentioning that any application can
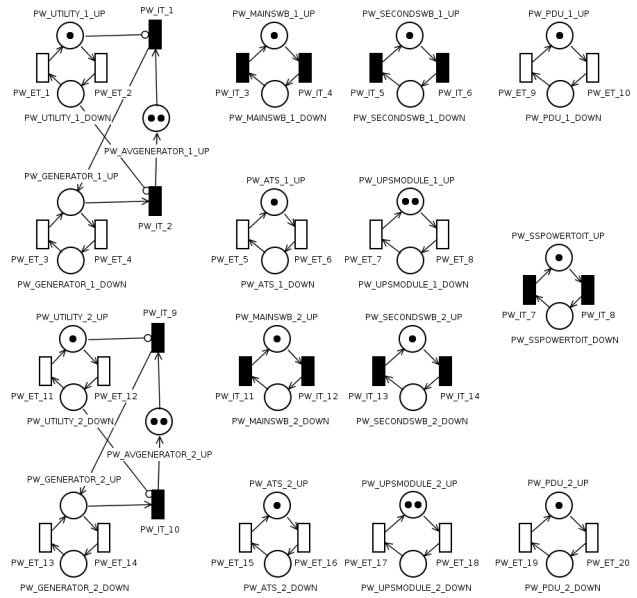

Fig. 4.   Tier IV power subsystem

be modeled, once the Mean Time To Failure (MTTF) and Mean Time To Repair (MTTR) values are known.

Solving the RBD ([19]), we obtain the MTTF and MTTR values of the entire service component. These values will be added in transitions referring to the server in our SPN model (Figure 7), described next.
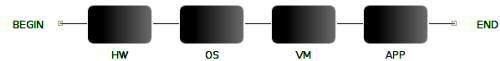

Fig. 5.   RBD model of service, based from [4]

4.2. IT Subsystem Tier I

In this work, we disregarded some components of the infrastructure illustrated in the Figure 1(b). The Edge router WAN was disregarded because is used for corporate networks, and does not affect the availability of service. We also disregard the large frame processing and disk array attached to it, because is used for backup, and does not impact the availability service.

Each network component is modeled using a building block with two places and two transitions. Places means the state of components (*UP* or *DOWN*) and transitions represent actions of these components (fail and repair). For example, the Core router is modeled as shown in figure 6. The place $IT\_CORE\_1\_UP$ represents when the Core router is UP, while the place $IT\_CORE\_1\_DOWN$ represents when this component is down. The failure

of this component is modeled by transition IT_ET_3, that consumes a token in place $IT\_CORE\_1\_UP$ and produces a token in place $IT\_CORE\_1\_DOWN$. Core router repair is modeled by transition IT_ET_4, and follows the inverse path of the failure transition. The other network components are similarly modeled, each one with own building block, as shown in Figure 6.
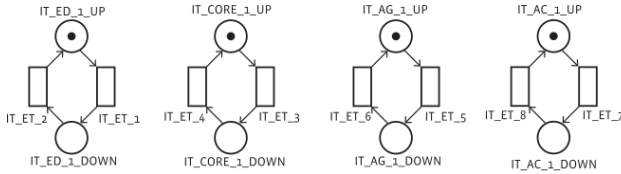

Fig. 6.   SPN network components tier I

Storage and service differ from network components due to the addition of more places and transitions, as shown in Figure 7. This is due to the fact that when one network component fails, server and storage also become unavailable. So, we used an additional place and transitions to model this behavior. For example, service are modeled using a building block with places up/down and transitions failure/repair, like network components. The immediate transition IT_IT_1 fires when a network component connected to the servers fails. When this component is repaired, the immediate transition IT_IT_2 fires, and the server become available again. Storage is modeled in a similar way.
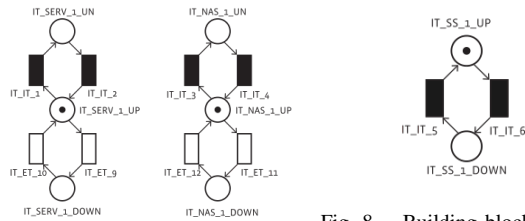

Fig. 7.   Storage and server tier I

Fig. 8.   Building block regarding IT infrastructure status

The behavior mentioned above is assured by guard functions, described in Table II. As can be seen in Figure 1(b), only Edge router, Core router, Aggregation router, and Access Switch are connected to servers. So, if one these components fails, servers will become unavailable. This behavior is modeled by guard function associated with transition IT_IT_1. On the other hand, if all components are running, the servers are available, and this behavior is ensured by guard function associated with transition IT_IT_2. Guard functions present in immediate transition associated with transition IT_IT_3 and associated with transition IT_IT_4 are similar, but only Edge router, Core router, Aggregation router are connected to storage,

therefore the failure of one of these components will make the storage unavailable. This behavior is modeled by the guard function associated with transition IT_IT_3. The guard function present in transition IT_IT_4 models these components repair, allowing storage to become available again.

TABLE II
GUARD FUNCTIONS OF IT SUBSYSTEM - TIERS I TO IV

Transition	Guard Function
IT_IT_1	$((\#IT\_ED\_1\_UP = 0)OR(\#IT\_CORE\_1\_UP = 0)$ $OR(\#IT\_AG\_1\_UP = 0)OR(\#IT\_AC\_1\_UP = 0))$
IT_IT_2	$((\#IT\_ED\_1\_UP > 0)AND(\#IT\_CORE\_1\_UP > 0)$ $AND(\#IT\_AG\_1\_UP > 0)AND(\#IT\_AC\_1\_UP > 0))$
IT_IT_3	$((\#IT\_ED\_1\_UP = 0)OR(\#IT\_CORE\_1\_UP = 0)$ $OR(\#IT\_AG\_1\_UP = 0))$
IT_IT_4	$((\#IT\_ED\_1\_UP > 0)AND(\#IT\_CORE\_1\_UP > 0)$ $AND(\#IT\_AG\_1\_UP > 0))$
IT_IT_5 $(tier\,I\,and\,II)$	$((\#IT\_NAS\_1\_UP > 0)AND(\#IT\_SERV\_1\_UP > 0))$
IT_IT_6 $(tier\,I\,and\,II)$	$((\#IT\_NAS\_1\_UP = 0)OR(\#IT\_SERV\_1\_UP = 0))$

The Figure 8 shows the building block that represents the status of the IT infrastructure. When the IT infrastructure is working (token in place $IT\_SS\_1\_UP$), if servers or storage fails, the immediate transition IT_IT_6 fires, making the system unavailable (token in place $IT\_SS\_1\_DOWN$). When server and storage become running again, the immediate transition IT_IT_5 fires, making IT infrastructure available again. The guard functions of these transitions are described in Table II.

4.3. IT Subsystem Tier II

The main difference between tier I and II is the dual Internet access link of Edge router. If one link fails, the second one will still keep the data center Internet connection. To represent this redundancy in our model, we just added a token in place $IT\_IT\_ED\_1\_UP$. So, if one link fails, one token will go to place $IT\_IT\_ED\_1\_DOWN$ but there will still be another token in the UP place.

4.4. IT Subsystem Tier III and IV

According TIA-942 specification, the differences between tiers III and IV are about low-level components such as cabling redundancy, component location, etc. However, since we are modeling components at a higher level, we disregard these differences. Therefore, tiers III and IV are modeled in a similar way.

In tiers III and IV, all components are redundant in order to keep the data center available, in case of unplanned outages. In addition, in tier III and IV there are

more components compared to tier I. Two SAN switches connect servers to disk arrays to provide storage. Similarly to previous tiers, in tiers III and IV large frame processing and edge router are disregarded. To model redundant components, there are two tokens in *UP* places and infinite server policy is used in transitions. The SPN model of tiers III and IV is presented in Figure 9.
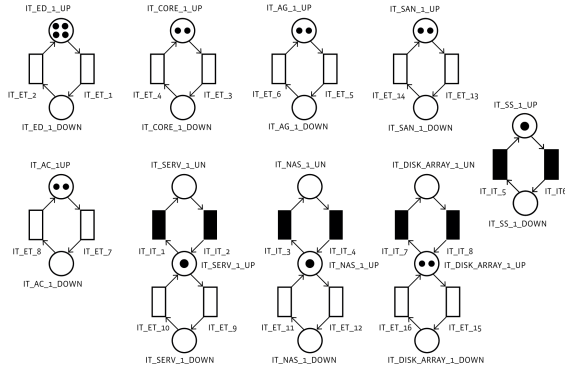

Fig. 9.   SPN model of IT infrastructure - tiers III and IV

As the model is very similar to tier I, all stochastic and immediate transitions of tier I are also used in tier III and IV models. SAN switch is represented by a building block with two places $IT\_SAN\_1\_UP$ and $IT\_SAN\_1\_DOWN$, and repair/failure transitions IT_ET_14 and IT_ET_13, like other components. The component that represents new array disk is similar to server and NAS components. The place $IT\_DISK\_ARRAY\_1\_UP$ represents when new disk array is *UP*, while place $IT\_DISK\_ARRAY\_1\_DOWN$ represents when it is *DOWN*.

Stochastic transitions IT_ET_15 and IT_ET_16 model repair and failure of new disk array, respectively. However, when both SAN switches fail, the second array disks will be unavailable ($IT\_DISK\_ARRAY\_1\_UN$), and this behavior is assured by the guard function present in immediate transition IT_IT_7. This transition consumes tokens of place $IT\_DISK\_ARRAY\_1\_UP$ and produces a token in place $IT\_DISK\_ARRAY\_1\_UN$. When a single SAN switch is repaired, the second array disk become available again, as modeled in guard function present in immediate transition IT_IT_8. This transition consumes all tokens present in place $IT\_DISK\_ARRAY\_1\_UN$ and produces a token in place $IT\_DISK\_ARRAY\_1\_UP$. These guard functions and respective transitions are presented in Table III.

As tier III and IV have an addition array disks to storage, guard functions presents in immediate transitions IT_IT_5 and IT_IT_6 change. Now, when at least one

of the storage way is working and servers are working, system is available. On the other hand, when servers are not working or both storage way are not working, system are unavailable. These guard functions are described in Table III.

TABLE III
GUARD FUNCTIONS OF IMMEDIATE TRANSITIONS - TIER III AND IV

Transition	Guard Function
IT_IT_5	$(\#IT\_SERV\_1\_UP>0)AND$ $((\#IT\_NAS\_1\_UP>0)\ OR(\#IT\_DISK\_ARRAY\_1\_UP>0))$
IT_IT_6	$(\#IT\_SERV\_1\_UP=0)OR$ $((\#IT\_NAS\_1\_UP=0)\ AND(\#IT\_DISK\_ARRAY\_1\_UP=0))$
IT_IT_7	$(\#IT\_SAN\_1\_UP=0)$
IT_IT_8	$(\#IT\_SAN\_1\_UP>0)$

5. Integrating Power and IT Subsystems

In order to integrate iers I to IV, we added the building block shown in Figure 10. This building block represents the overall (power and IT) data center availability, which may be available (DC_1_UP) or not (DC_1_DOWN).
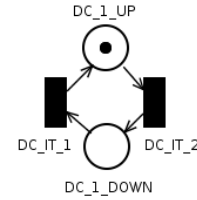

Fig. 10.   Building block regarding the power and IT integration.

The immediate transitions DC_IT_1 and DC_IT_2 have guard functions representing the repair and failure conditions of the data center, see Table IV. As described in Table IV, the data center will be available only if the power subsystem is able to feed the IT equipment ($\#PW\_SSPOWERTOIT\_UP=1$) and the IT subsystem is working ($\#IT\_SS\_1\_UP=1$), (see immediate transition DC_IT_1). On the other hand, a power subsystem failure will result in IT system disruptions. Besides, a failure in the IT subsystem will also result in the data center unavailability, (see transition DC_IT_2).

The overall data center availability (integration of the power and IT subsystems) is given by Equation 1 bellow.

$$DC_{availability} = P\{(\#DC\_1\_UP=1)\} \qquad (1)$$

From this equation, the data center availability is the probability of having a token in place DC_1_UP. This way, we may analyze the impact of power failures on the availability of the IT subsystem.

Transition	Guard Function
DC_IT_1	$(\#PW\_SSPOWERTOIT\_UP = 1)\,AND$ $(\#IT\_SS\_1\_UP = 1)$
DC_IT_2	$(\#PW\_SSPOWERTOIT\_UP = 0)\,OR$ $(\#IT\_SS\_1\_UP = 0)$

6. Evaluation

The MTTF and MTTR values of the video stream application (considering VLC to Encode/decode of videos) are shown in Table V and parameters related to the power and IT subsystems are represented in Table VI. We used the Mercury tool [18] for all analysis.

TABLE V
MTTF AND MTTR VALUES OF THE VLC APPLICATION (FROM [4])

Components	MTTF (hr/yr)	MTTR (hr/yr)
HW	8,760	1.667
OS	1,440	1
VM	1,880	0.167
APP	336	1

TABLE VI
MTTF AND MTTR VALUES OF POWER (FROM [15] AND [8]) AND IT
COMPONENTS (FROM [11], [17], AND [16])

Component	MTTF (hr/yr)	MTTR (hr/yr)
Utility	257.2678158	0.03280086594
Generator	9,733.307378	3.9000039
ATS	102,093.9468	5.739869131
UPS Module	27,472.52747	8
PDU	282,581.0	156.0062402
Edge Router	796	1
Core Router	16,243	0.78
Aggregation Router	8,247	0.63
Access Switch	13,043.48	0.35
Service (obtained from RBD)	231.6581	0.9157
NAS	1,200,000	12
SAN	255,358	7.66
Disk Array	1,200,000	12

6.1. Availability Analysis

The availability analysis results were obtained through stationary simulations, and are shown in Table VII. This table is composed of the availability level, the number of 9's, and the downtime related to tiers I to IV.

Tier I offers approximately 99.38% of availability with a downtime of 54.16 hours/year. The $N + 1$ redundancy

in the generator, UPS module, and access link of the edge router on tier II reduced the downtime to 13.3 hours when compared to tier I. In tier III, the use of a secondary distribution path (no UPS components) and redundant IT equipment reduced by downtime even more (to 6.3 hours) compared to tier II. Lastly, in tier IV a complete and secondary distribution path (with UPS components) reduced by the downtime in 9.72 seconds compared with tier III. A downtime reduction of 9.72 seconds may be expressive depending on the business needs. We also point out that from tier I (no redundant) to tier IV (fully redundant) reduced by the downtime in 19.65 hours.

TABLE VII
SERVICE AVAILABILITY

Tier	Availability (%) (Min - Average - Max)	Number of 9's	Downtime (hr/yr)
I	99.38164 - 99.381690 - 99.38173	2.2087	54.1639
II	99.53373 - 99.533775 - 99.53381	2.3314	40.8412
III	99.606034 - 99.606035 - 99.606036	2.4045	34.5112
IV	99.606065 - 99.606066 - 99.606067	2.4045	34.5085

6.2. Sensitivity Analysis

The sensitivity analysis results were obtained using the percentage difference method. From this analysis we obtained the sensitivity ranking of all power and IT components. Table VIII presents the top three most sensitive components from tier I to IV.

Then, we conducted a further sensitivity analysis in the top two component's rate. With this analysis we may verify the impact in overall data center availability varying (in 10% using five sample points) the MTTF and MTTR rates of a single component at a time. The results are shown in Figures 11 to 18. From the results, we highlight that increasing the mean time to failure of server in 20% (46.33 hours) has the highest impact in case of Tier IV, where it reduced by the downtime in 5.73 hours (from 34.50% to 28.77%). On the other hand, decreasing the repair time of server by 20% (10.98 minutes) has a higher availability impact on Tier IV (compared to other tiers), reducing the overall downtime in 6.87 hours (from 34.50% to 27.62%). Therefore, shortening the repair time of the server can considerably improve data center availability.

7. Related Work

There are some works that model data center infrastructure, such as [2; 8; 3; 7; 14]. For instance, a RBD model for evaluating availability and reliability of modular UPS
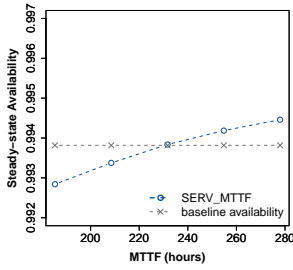
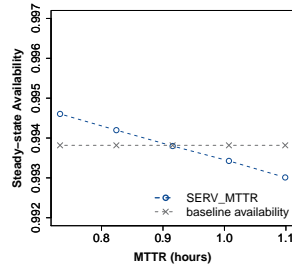Fig. 11. Tier I: Sensitivity analysis of the MTTF server.

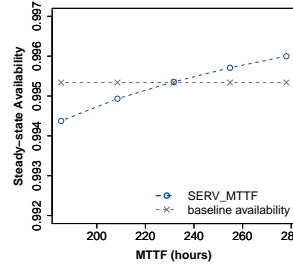Fig. 12. Tier I: Sensitivity analysis of the MTTR server.

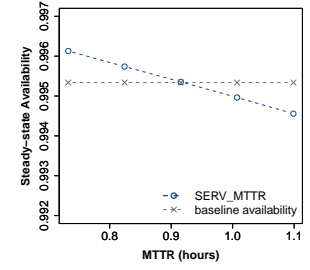Fig. 13. Tier II: Sensitivity analysis of the MTTF server.

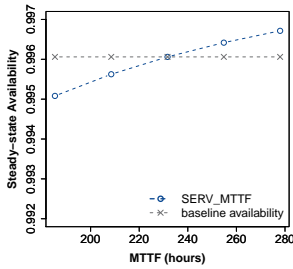Fig. 14. Tier II: Sensitivity analysis of the MTTR server.

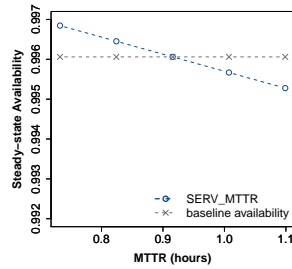Fig. 15. Tier III: Sensitivity analysis of the MTTF server.

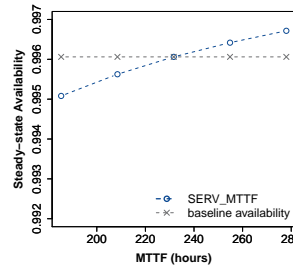Fig. 16. Tier III: Sensitivity analysis of the MTTR server.

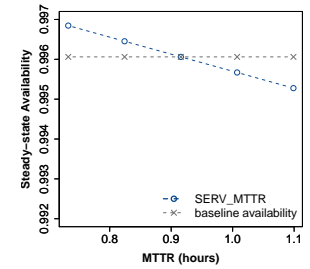Fig. 17. Tier IV: Sensitivity analysis of the MTTF server.

Fig. 18. Tier IV: Sensitivity analysis of the MTTR server.

TABLE VIII
SENSITIVITY RANKING

Tier	Parameter	Sensitivity Index	Tier	Parameter	Sensitivity Index
I	MTTF Server	7.8595×10^{-4}	III	MTTF Server	7.9549×10^{-4}
	MTTR Server	7.7678×10^{-4}		MTTR Server	7.8704×10^{-4}
	MTTF Edge Router	2.5944×10^{-4}		MTTF Edge Router	1.2862×10^{-6}
II	MTTF Server	7.9386×10^{-4}	IV	MTTF Server	7.9528×10^{-4}
	MTTR Server	7.8786×10^{-4}		MTTR Server	7.8741×10^{-4}
	MTTR PDU	1.1233×10^{-4}		MTTR Edge Router	1.5268×10^{-6}

was proposed by [2]. Authors compared different UPS configurations under temperature variation. Their results showed that varying the UPS configuration increased the availability and reliability system. However, authors do not assess the impact of UPS system in service availability.

In [8], authors presented models to evaluate the impact, cost, and dependability of data center cooling and power infrastructures. RBD and SPN was used to model different architectures. They proposed an incremental architecture where components with most reliability and cost importance were replicated. Authors focused on sustainable impact and availability and reliability power system. While, we evaluate impact of power and IT in service availability.

Authors in [9] proposed models to evaluate an emergency call center, with focus on evaluating power system impacts in call center service downtime. The call center

architecture modeled is composed of energy, networking, voice, and customer service structures. They used RBD to represent logical dependencies of the components and SPN to represent the energy components' behavior. Differently from authors, our work considers power and IT subsystems and their impact on service availability.

Our work differs from the literature because we proposed models based on TIA-942 standard. At the same time, we evaluate the impact of power infrastructure in cloud services availability running an IT infrastructure of data center. We also performed sensitivity analysis to detect components that most impact on service availability.

8. Final Considerations

This paper presented a set of RBD and SPN models integrating power and IT subsystems of a data center. Based on the availability results of simulations, using redundant component in power and IT systems, can considerably decrease downtime of a service running in the data centers. Changing the architecture from a tier I data center to tier IV, lead to roughly 36.28 percent (19.65 hours/year) reduction in the service downtime. From sensitivity analysis, we observed that the server

failure and repair rates have the highest impact on the data center availability. Moreover, we analyzed the server component individually to see how much it affects the overall availability. Results showed that in Tier IV the server repair and failure rates have the highest impact on availability when compared to other tiers. As future work we aim to integrate the power, cooling, and IT subsystems by modeling their interconnections in order to better estimate the overall data center availability.

Acknowledgment

References

[1] "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.

[2] T. Addabbo, A. Fort, M. Mugnaini, V. Vignoli, E. Simoni, and M. Mancini, "Availability and reliability modeling of multicore controlled ups for datacenter applications," *Reliability Engineering & System Safety*, vol. 149, pp. 56–62, 2016.

[3] H. A. Alissa, K. Nemati, B. G. Sammakia, M. J. Seymour, R. Tipton, D. Mendo, D. W. Demetriou, and K. Schneebeli, "Chip to chiller experimental cooling failure analysis of data centers: The interaction between it and facility," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 6, no. 9, pp. 1361–1378, 2016.

[4] 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*. IEEE, 2014, pp. 666–671.

[5] S. Balsamo, A. Marin, and I. Stojic, "Deriving the performance indices in product-form stochastic petri nets: open problems and simulation," in *EUROSIS European Simulation Multiconference*, 2015.

[6] F. Bause and P. S. Kritzinger, *Stochastic Petri Nets*. Vieweg Wiesbaden, 2002, vol. 1.

[7] G. Callou, P. Maciel, D. Tutsch, and J. Araújo, "Models for dependability and sustainability analysis of data center cooling architectures," in *Dependable Systems and Networks Workshops (DSN-W), 2012 IEEE/IFIP 42nd International Conference on*. IEEE, 2012, pp. 1–6.

[8] 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.

[9] M. A. de QV Lima, P. R. Maciel, B. Silva, and A. P. Guimarães, "Performability evaluation of emergency call center," *Performance Evaluation*, vol. 80, pp. 27–42, 2014.

[10] P. T. Endo, M. Rodrigues, G. E. Gonçalves, J. Kelner, D. H. Sadok, and C. Curescu, "High availability in clouds: systematic review and research challenges," *Journal of Cloud Computing*, vol. 5, no. 1, p. 16, 2016.

[11] A. P. Guimaraes, P. Maciel, and R. MATLAs JR, "Design of it infrastructures of data centers: An approach based on business and technical metrics," *Quantitative Assessments of Distributed Systems: Methodologies and Techniques*, p. 265, 2015.

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

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

[14] Z. Liu, Y. Chen, C. Bash, A. Wierman, D. Gmach, Z. Wang, M. Marwah, and C. Hyser, "Renewable and cooling aware workload management for sustainable data centers," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, no. 1. ACM, 2012, pp. 175–186.

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

[16] B. Schroeder and G. A. Gibson, "Disk failures in the real world: What does an mttf of 1, 000, 000 hours mean to you?" in *FAST*, vol. 7, 2007, pp. 1–16.

[17] S. Shetty, "Determining the availability and reliability of storage configurations," *Dell Power Solutions*, 2002.

[18] 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 *DSN*, 2015.

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