



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA APLICADA

FELIPE FERNANDES DE LIMA MELO

**ANÁLISE DE INFRAESTRUTURAS ELÉTRICAS DE
DATA CENTERS: UMA ESTRATÉGIA BASEADA
EM MODELOS, POLÍTICAS DE MANUTENÇÃO E
PLANEJAMENTO DE EXPERIMENTOS**

RECIFE – PE

2020

FELIPE FERNANDES DE LIMA MELO

**ANÁLISE DE INFRAESTRUTURAS ELÉTRICAS DE
DATA CENTERS: UMA ESTRATÉGIA BASEADA
EM MODELOS, POLÍTICAS DE MANUTENÇÃO E
PLANEJAMENTO DE EXPERIMENTOS**

Dissertação submetida à Coordenação do
Programa de Pós-Graduação em Informática
Aplicada da Universidade Federal Rural
de Pernambuco, como parte dos requisitos
necessários para obtenção do grau de Mestre.

ORIENTADOR: Prof. Gustavo Rau de Almeida Callou

RECIFE – PE

2020

FICHA CATALOGRÁFICA

FELIPE FERNANDES DE LIMA MELO

**ANÁLISE DE INFRAESTRUTURAS ELÉTRICAS DE
DATA CENTERS: UMA ESTRATÉGIA BASEADA
EM MODELOS, POLÍTICAS DE MANUTENÇÃO E
PLANEJAMENTO DE EXPERIMENTOS**

Dissertação submetida à Coordenação do
Programa de Pós-Graduação em Informática
Aplicada á Universidade Federal Rural de
Pernambuco, como parte dos requisitos
necessários para obtenção do grau de Mestre.

Aprovada em: 01 de Junho de 2020.

BANCA EXAMINADORA

Gustavo Rau de Almeida Callou (Orientador)
Universidade Federal Rural de Pernambuco
Departamento de Computação

Erica Teixeira Gomes de Sousa
Universidade Federal Rural de Pernambuco
Departamento de Computação

Jean Carlos Teixeira de Araujo
Universidade Federal do Agreste de Pernambuco

Aos meus pais, irmã e avós.

Agradecimentos

Em primeiro lugar, a Deus pela oportunidade de alçar mais este degrau.

Aos meus pais, irmã e avó pelo apoio.

Em especial ao esforço e sacrifício dos meus pais que nunca deixaram faltar nada e sempre me apoiaram durante toda a minha vida.

Ao nosso professor Gustavo Callou, por sempre estar presente nos momentos que mais precisei, pela ajuda e apoio neste trabalho, por aceitar ser meu orientador, pela paciência e por fazer parte dessa trajetória, guiando minhas pesquisas e compartilhando seus conhecimentos comigo.

Agradeço também a todos os meus amigos que sempre acreditaram na realização deste trabalho.

"A persistência é o caminho do êxito."

(Charles Chaplin)

Resumo

Nos últimos anos, a demanda e o uso dos *data centers* aumentaram consideravelmente devido aos serviços oferecidos pela computação em nuvem. Os *data centers* são responsáveis por diversas aplicações de valor crítico e, assim, a análise da disponibilidade e confiabilidade de tais sistemas se torna de fundamental importância. Este trabalho adota estratégias de otimização baseadas nas técnicas de planejamento de experimentos (DoE). A estratégia de otimização adotada faz uso do índice de importância para a disponibilidade que é capaz de identificar os equipamentos mais sensíveis para tal métrica. Além disso, o trabalho proposto faz uso de uma modelagem híbrida que considera as vantagens das redes de Petri estocásticas (SPN) e diagramas de bloco de confiabilidade (RBD) para avaliar a disponibilidade. Políticas de manutenção também fazem parte do escopo deste trabalho a fim de se evitar possíveis interrupções no funcionamento do sistema. De forma mais específica, foram utilizadas duas técnicas para realização das manutenções: a manutenção corretiva e a manutenção preventiva.

Para ilustrar a aplicabilidade da abordagem proposta, foram realizados três estudos de casos, cada um deles aplicado a cenários distintos, com o objetivo de informar uma estratégia de otimização que leva em consideração técnicas de planejamento de experimentos, para otimizar a disponibilidade da infraestrutura elétrica de um *data center*. O primeiro estudo de caso foi conduzido comparando os resultados obtidos através tanto do algoritmos da força bruta quanto da estratégia de otimização proposta aplicados a infraestruturas elétricas de um *data center*. O segundo estudo de caso foi conduzido no cenário de manutenção da infraestrutura elétrica de *data center*, apresentando, assim uma comparação entre as políticas de manutenção corretiva e preventiva. Por fim, o terceiro estudo de caso, é considera diferentes políticas de manutenção sendo estas vinculadas a contratos de *SLA*, avaliando assim o impacto no uso de diferentes contratos de *SLA* na disponibilidade do sistema. Os experimentos realizados neste trabalho corroboram com a afirmação de que a estratégia de otimização proposta pode ser aplicada a diferentes cenários. Por exemplo, em experimentos realizados foi possível reduzir o tempo de execução em 26 vezes em relação a força bruta e, mesmo assim, chegou-se a solução ótima.

Palavras-chave: Disponibilidade, Planejamento de Experimentos, Redes de Petri, *Data Center*, Políticas de Manutenção e Diagramas de Bloco de Confiabilidade.

Abstract

In the last years, the demand and use of *data centers* have increased considerably due to the services offered by cloud computing. *Data centers* are responsible for several critical applications and, thus, the availability and reliability analysis of such systems becomes of fundamental importance. This work adopts optimization strategies based on the techniques of design of experiments (DoE). The optimization strategy adopted makes use of the importance index for availability, which can identify the most sensitive equipment for this metric. In addition, the proposed work makes use of hybrid modeling that considers the advantages of stochastic Petri nets (SPN) and reliability block diagrams (RBD) to assess availability. Maintenance policies are also adopted in this work to avoid possible interruptions in the operation of the system. To accomplish this, two techniques were used to fix the environment under analysis: corrective and preventive maintenances.

To illustrate the applicability of the proposed approach, three case studies were carried out, each applied to different scenarios, with the objective optimizing the availability of the electrical infrastructures of *data centers*. The first case study was conducted comparing the results obtained through both the brute force algorithms and the proposed optimization strategy. The second case study was conducted in the maintenance scenario of the electrical infrastructure of the *data center*, thus presenting a comparison between the corrective and preventive maintenance policies. Finally, the third case study, different maintenance policies are applied and these are linked to *SLA* contracts, thus evaluating the impact on the use of different *SLA* contracts on the system's availability. The experiments carried out in this work corroborate with the statement that the proposed optimization strategy can be applied to different scenarios. For example, in experiments carried out, it was possible to reduce the execution time by 26 times concerning brute force and, even so, the optimal solution was reached.

Keywords: Availability, Design of Experiments, Petri nets, *Data Center*, Maintenance Policy and Reliability block diagrams.

Lista de Figuras

Figura 1 – Modelo SPN	27
Figura 2 – Modelo RBD	28
Figura 3 – Diagrama de bloco de confiabilidade – Composto por três equipamentos em série	29
Figura 4 – Ciclo de vida do <i>SLA</i>	34
Figura 5 – Infraestruturas do <i>Data center</i>	36
Figura 6 – Pontos P_1 e P_2 marcados no plano (Fonte (NINA, 2011))	38
Figura 7 – Metodologia adotada	46
Figura 8 – Modelo SPN Base de Disponibilidade	52
Figura 9 – Modelo RBD Base de Disponibilidade	53
Figura 10 – Modelo Modificado RBD de disponibilidade	53
Figura 11 – Modelo de Manutenção Corretiva	54
Figura 12 – <i>clock</i> do Modelo de Manutenção Corretiva	55
Figura 13 – Modelo de Manutenção Preventiva	56
Figura 14 – <i>clock</i> do Modelo de Manutenção Preventiva	56
Figura 15 – Modelo de Manutenção Preventiva Modificado	57
Figura 16 – Arquiteturas	60
Figura 17 – Modelo RBD da Arquitetura A2	61
Figura 18 – Modelo SPN da Arquitetura A5	62
Figura 19 – Modelo da Arquitetura A1 de Manutenção Corretiva	68
Figura 20 – Modelo da Arquitetura A2 de Manutenção Preventiva	69
Figura 21 – Comparação da Manutenção Corretiva vs Manutenção Preventiva - Arquitetura A5	74
Figura 22 – <i>Downtime</i> e Custo da Manutenção Corretiva - Considerando a Arquitetura A4	82
Figura 23 – <i>Downtime</i> e Custo da Manutenção Preventiva - Considerando a Arquitetura A4	82

Lista de tabelas

Tabela 1 – Importância para a Disponibilidade	30
Tabela 2 – Resumo comparativo dos trabalhos relacionados.	45
Tabela 3 – Fórmulas adotadas para o cálculo da Disponibilidade	62
Tabela 4 – Valores de intervalo de referência para os dispositivos	63
Tabela 5 – Equipamentos mais sensíveis para a disponibilidade nas Arquiteturas de A1 até A5	65
Tabela 6 – Avaliação Comparativa dos Resultados do Algoritmo de Força vs a Estratégia de Otimização Proposta.	66
Tabela 7 – MTTFs e MTTRs das Transições Temporizadas do Modelo da Arquitetura A2 com Manutenção Preventiva	70
Tabela 8 – Equação do cálculo da degradação para o modelo da Figura 19.	71
Tabela 9 – Expressões de guarda Modelo da Arquitetura A2 com Manutenção Preventiva (ex., A2 - Figura 20).	71
Tabela 10 – Resultados das métricas considerando a realização da Manutenção Corretiva	72
Tabela 11 – Resultados das métricas considerando a realização da Manutenção Preventiva	72
Tabela 12 – Contratos <i>SLA</i>	75
Tabela 13 – Contrato <i>SLA</i> - Manutenção Corretiva	77
Tabela 14 – Contrato <i>SLA</i> - Manutenção Preventiva	78
Tabela 15 – Disponibilidade (9's) da Manutenção Corretiva	80
Tabela 16 – Disponibilidade (9's) da Manutenção Preventiva	80
Tabela 17 – Distância Euclidiana - Manutenção Corretiva	81
Tabela 18 – Distância Euclidiana - Manutenção Preventiva	81

Lista de Siglas

AG	<i>Genetic Algorithm</i>
CFD	<i>Computational Fluid Dynamic</i>
CMCF	<i>Multicommodity Capacited Minimum Cost Flow</i>
CRAC	<i>Computer Room Air Conditioning</i>
DEA	<i>Data Envelopment Analysis</i>
DEDS	<i>Discrete Events Dynamic Systems</i>
DOE	<i>Design of Experiments</i>
FEM	<i>Método de Elementos Finitos</i>
GEE	<i>Greenhouse gas</i>
HERO	<i>Hierarchical Energy Optimization</i>
HPLC	<i>Cromatógrafo Líquido de Alto Desempenho</i>
HT	<i>Percentage of High Tension</i>
ID	<i>Importance for Availability</i>
LRD	<i>long-range dependence</i>
MECDC	<i>Maintenance and Electricity Costs Data Center</i>
MTTF	<i>Mean Time to Failure</i>
MTTR	<i>Mean Time to Repair</i>
NAS	<i>Network Attached Storage</i>
NBS	<i>Nash Bargaining Solution</i>
OE	<i>Specific objectives</i>
OG	<i>General objectives</i>
RBD	<i>Reliability Block Diagram</i>
SAN	<i>Storage Area Network</i>
SDT	<i>Step Down Transformers</i>
SA	<i>Sensitivity Analysis</i>
SLA	<i>Service Level Agreement</i>

SPN	<i>Stochastic Petri Net</i>
STS	<i>Static Transfer Switch</i>
TS	<i>Tabu Search</i>
UG	<i>Underground Cables</i>
UPS	<i>Uninterruptible Power Supplies</i>

Sumário

1	Introdução	15
1.1	Motivação	17
1.2	Objetivos	18
1.3	Estrutura do Trabalho	18
2	Fundamentação Teórica	20
2.1	Disponibilidade	20
2.2	Políticas de Manutenção	21
2.2.1	Manutenção Preventiva	22
2.2.2	Manutenção Corretiva	23
2.2.3	Outros tipos de Manutenção	24
2.3	Redes de Petri Estocásticas	25
2.4	Diagramas de Bloco de Confiabilidade	27
2.5	Importância para a Disponibilidade	29
2.6	Planejamento de Experimentos	30
2.6.1	Planejamento de experimento de um único fator	31
2.6.2	Planejamento de experimento de dois fatores	31
2.6.3	Planejamento de experimento de vários fatores	32
2.6.4	Planejamento de experimento 2k	32
2.7	<i>Service Level Agreement</i>	33
2.8	<i>Data center</i>	35
2.8.1	Infraestruturas do <i>Data center</i>	36
2.8.2	Distância Euclidiana	37
3	Trabalhos Relacionados	39
4	Metodologia	46
4.1	Ferramenta	50
5	Modelos	52
5.1	Modelo Base de Disponibilidade	52
5.2	Modelo Base em RBD	52
5.3	Modelo de Manutenção Corretiva	54
5.4	Modelo de Manutenção Preventiva	55

5.4.0.1	Modelo de Custo	56
6	Estudos de Caso	59
6.1	Arquiteturas	59
6.2	Estudo de caso I	59
6.2.1	Modelos	60
6.2.1.1	Modelo RBD	61
6.2.1.2	Modelo SPN	61
6.2.2	Resultados	63
6.3	Estudo de caso II	67
6.3.1	Modelos	67
6.3.1.1	Modelo para a Manutenção Corretiva	67
6.3.1.2	Modelo para a Manutenção Preventiva	67
6.3.2	Resultados	69
6.4	Estudo de Caso III	74
6.4.1	Resultado	75
7	Conclusão	83
7.1	Considerações	83
7.2	Contribuições	84
7.3	Limitações e Trabalhos Futuros	85
	Referências	87

1 Introdução

Ao longo dos últimos anos, o número, o tamanho e o poder computacional dos *data centers* vêm aumentando consideravelmente devido aos serviços ofertados pela computação em nuvem, redes sociais e comércio eletrônico (FOX et al., 2009) (CHANDEL et al., 2018). Tendo em vista a constante utilização dos serviços oferecidos pela computação em nuvem, torna-se fundamental estudos direcionados para a infraestrutura deste ambiente, a fim de garantir uma alta disponibilidade.

Os *data centers* são formados por equipamentos e sistemas responsáveis pelo processamento e armazenamento de informações cruciais e de grande importância para a continuidade e funcionamento de diferentes aplicações e serviços (SOUSA, 2017). Eles são recursos de missão crítica, pois abrigam muitos serviços que são cruciais para as empresas. Por exemplo, eles podem ser utilizados como *host* de *sites* para processar transações comerciais e proteger dados bancários, registros financeiros e histórico médico de pacientes. Tais sistemas demandam que seus serviços estejam em operação 24 horas por dia, durante os 7 dias da semana (ARREGOCES; PORTOLANI, 2003).

O *data center* tem de estar disponível 365 dias ao ano, sem interrupção, incluindo toda a manutenção programada, avarias não programadas e garantir que os resultados de negócios críticos sejam entregues (SOUSA, 2017). Por se tratar de ambientes de missão crítica, os *data centers* devem estar sempre aptos a operar ininterruptamente, logo é essencial manter a disponibilidade e confiabilidade para permitir que o ambiente funcione corretamente (MARIN, 2016).

Atividades relacionadas à manutenção de equipamentos industriais vêm sendo adotadas desde a revolução industrial, onde diferentes estratégias de manutenção podem ser empregadas com a finalidade de reduzir a indisponibilidade dos sistemas (BORKOWSKI; HANS, 2007). Tendo em vista o constante crescimento no nível de utilização dos ambientes que proveem suporte à computação em nuvem, torna-se fundamental estudo em relação à infraestrutura deste serviço para garantir a alta disponibilidade exigida. Dependendo da aplicação que esteja fazendo uso desse ambiente, uma falha do sistema pode gerar um prejuízo financeiro e, também, na imagem da empresa. Sendo assim, espera-se que este ambiente esteja sempre operacional e sem ocorrência de falhas ou anomalias na sua infraestrutura.

Duas abordagens de manutenção distintas são utilizadas a fim de se evitar possíveis interrupções no funcionamento do sistema. A primeira, a manutenção corretiva, é realizada quando algum dispositivo falhar; a segunda abordagem, a manutenção preventiva, visa evitar que um determinado dispositivo venha a falhar. Utiliza-se também acordos de nível de serviço - *Service Level Agreement (SLA)* - para definir um conjunto de compromissos oferecidos por uma empresa prestadora de serviço, possibilitando assim a escolha do melhor contrato levando em consideração o custo e a disponibilidade oferecido para quem vai contratar o serviço.

Esse trabalho também tem como foco a avaliação no custo, que consiste do gasto relacionado a um bem ou serviço para produzir outros bens e serviços (WERNKE, 2016). Neste trabalho, busca-se a redução dos custos por meio da adoção das melhores estratégias de manutenção, e da redução da infraestrutura elétrica do *data center*. Porém, mesmo com a redução da infraestrutura, espera-se manter a oferta de serviços e capacidade de atender às necessidades do negócio (VERDI et al., 2010). Esse trabalho leva em consideração o custo dos contratos de *SLA*, o custo de manutenção e o do time de manutenção.

Esse trabalho também fez uso do planejamento de experimentos (DoE), sendo uma técnica utilizada para direcionar os projetistas na escolha dos experimentos que realmente precisam ser realizados. O DoE permite também determinar os efeitos de cada fator de entrada individual e os efeitos de suas interações sobre o desempenho (JALBY et al., 2012). Por meio do planejamento de experimento é possível obter uma maior quantidade de informação com o número mínimo de testes. Isso economiza trabalho considerável que teria de ser gasto na coleta de dados, por exemplo. Uma análise adequada dos experimentos ajuda a identificar os efeitos de vários fatores no desempenho (JAIN, 1990).

A avaliação da infraestrutura elétrica de *data centers* leva em consideração dois tipos distintos de manutenção (preventiva e corretiva). No caso, ambas as políticas de manutenção são utilizadas junto com contratos de *SLA*, apresentando variações nos tipos de contratos que podem ser utilizados. Sendo assim, com a utilização destas políticas de manutenção espera-se que o sistema como um todo fique mais tempo operacional sem a ocorrência de possíveis falhas nos equipamentos. Além disso, espera-se que a confiabilidade do ambiente aumente e o custo com o reparo dos equipamentos tenha uma redução.

Dessa forma, esta pesquisa propõe a utilização de uma estratégia de otimização que leva em consideração técnicas de planejamento de experimentos, com o intuito de otimizar

a disponibilidade da infraestrutura elétrica de um *data center*. Essa estratégia é vinculada a técnicas de modelagem híbrida que considera as vantagens da utilização de redes de Petri estocásticas (SILVA et al., 2006) e diagramas de bloco de confiabilidade (TORELL; AVELAR, 2004). Com a finalidade de ilustrar a aplicabilidade da estratégia proposta, diferentes estudos de casos foram propostos. Por exemplo, o Estudo de caso I faz uso do DoE em cenários de infraestrutura elétrica de *data centers* realizando uma comparação entre a solução proposta e os resultados obtidos através do algoritmo de força bruta. Já os Estudos de caso II e III fazem uso de políticas de manutenção vinculadas a estratégia proposta através do uso do DoE para encontrar a melhor relação custo benefício.

1.1 Motivação

A utilização de políticas de manutenção em instalações de *data center* é uma questão de valor crítico, pois devido às rigorosas restrições de disponibilidade e confiabilidade, as estratégias de gerenciamento desempenham um papel muito importante na infraestrutura que oferece suporte ao serviço comercial através da Internet (ex., computação em nuvem, serviços eletrônicos, big data).

Devido à falta de meios adequados para avaliar e estimar a disponibilidade e confiabilidade dos equipamentos de *data center* e, ainda mais, levando em consideração a manutenção de tais dispositivos, existe uma demanda de estudos nessa área. Esses estudos devem focar na redução do custo e da quantidade de manutenções corretivas demandadas, pois tal estratégia acaba por reduzir a disponibilidade dos sistemas. Sendo assim, o foco é no aumento da disponibilidade dos serviços ofertados.

Dependendo da aplicação que esteja fazendo uso desse ambiente, uma falha do sistema pode gerar um prejuízo financeiro e, também, denegrir a imagem da empresa. Falhas acabam por acarretar na elevação dos custos, prejudicando clientes e provedores de serviços. Isso acontece em virtude da perda da produção, ocasionada pela indisponibilidade do sistema, e também devido ao custo de reparo do equipamento/sistema.

Por meio do uso da técnica de planejamento de experimentos, o projetista pode fazer uso de métodos que possibilitam a diminuição do esforço necessário para encontrar uma determinada solução. Sendo assim, com o intuito de reduzir os cenários a serem observados, o DoE é de fundamental importância, onde a quantidade de interações necessárias pode

ser reduzida para se obter um resultado.

Com visto acima, este ambiente necessita de políticas de manutenção para assegurar seu pleno funcionamento, a fim de evitar possíveis problemas de indisponibilidade e confiabilidade do ambiente como um todo, como também a fim de reduzir os custos com manutenção. Dessa forma, a presente pesquisa tem como questão principal: Quais estratégias devem ser adotadas para análise e avaliação do custo e da disponibilidade de infraestruturas elétricas de *data centers*?

1.2 Objetivos

A presente pesquisa tem como objetivo a proposição de uma estratégia baseada em modelos, contratos de *SLA*, políticas de manutenção e técnicas de otimização para avaliação do custo e da disponibilidade de infraestruturas elétricas de *data centers*. Para atingir esse objetivo geral foram propostos os seguintes objetivos específicos:

- Construir modelos formais (RBD e SPN ([MARSAN et al., 1986](#))) para representar os sistemas da infraestrutura elétrica de *data center* com o objetivo de quantificar a disponibilidade;
- Propor modelos em SPN para representar diferentes políticas de manutenção (ex., preventiva e corretiva);
- Propor modelos para representar o custo e/ou *SLA* das políticas de manutenção;
- Desenvolver um modelo de otimização (ex., planejamento de experimentos (DoE)) para reduzir os cenários a serem avaliados e, também, maximizar a disponibilidade levando em consideração o custo.
- Propor uma estratégia integrada que relacione os modelos de disponibilidade, de custo, e da técnica de otimização para avaliar infraestruturas elétricas de *data centers*.

1.3 Estrutura do Trabalho

Esta dissertação está dividida, além da introdução, em mais seis capítulos. No Capítulo 3 são apresentados os trabalhos relacionados. No Capítulo 2 são introduzidos os conceitos necessários sobre disponibilidade, manutenção, redes de Petri, diagramas de bloco de confiabilidade, planejamento de experimentos e importância para disponibilidade. No

Capítulo 4 é explicada a metodologia proposta. No Capítulo 5 são apresentados detalhes dos modelos básicos propostos. No Capítulo 6 são expostos os estudos de caso que foram realizados a partir das abordagens desenvolvidas. No Capítulo 7 são apresentadas as conclusões a respeito do trabalho, as suas limitações e os trabalhos futuros.

2 Fundamentação Teórica

Este capítulo apresenta alguns conceitos básicos para um melhor entendimento da pesquisa desenvolvida. Em primeiro lugar, são apresentados conceitos sobre a disponibilidade, seguido pelas políticas de manutenção comumente utilizadas, redes de Petri estocásticas, diagramas de bloco de confiabilidade e posteriormente, definições sobre planejamentos de experimentos serão introduzidos, assim como, uma breve explicação sobre *SLA*, importância para a disponibilidade de *data centers*.

2.1 Disponibilidade

A disponibilidade corresponde a probabilidade do sistema se manter em funcionamento levando em consideração a ocorrência de falhas e reparos de dispositivos que compõem tal sistema (MACIEL K. S. TRIVEDI; KIM, 2011). A disponibilidade é uma métrica crítica, devido ao fato de que o tempo de inatividade geralmente afeta a produção e o lucro das empresas (CALLOU, 2013). O tempo durante o qual o sistema não está disponível é chamado de *downtime*.

Para avaliar a disponibilidade do sistema são considerados o tempo médio para falhar (*Mean Time to Failure - MTTF*), que representa tempo médio para a ocorrência de uma falha; e o tempo médio para o reparo (*Mean Time to Repair - MTTR*), que significa tempo médio para se realizar um reparo, após a ocorrência de uma falha. A disponibilidade (*A*) de estado estacionário pode ser calculado pela Equação 2.1. O resultado obtido pelo cálculo realizado na Equação 2.1, será um valor entre zero e um.

$$A = \frac{MTTF}{MTTF + MTTR} \quad (2.1)$$

As aplicações e sistemas computacionais necessitam de distintos níveis de disponibilidade, sendo assim podem ser classificados de acordo com esses níveis. Por exemplo, em (BAUER; ADAMS, 2012) as aplicações e sistemas computacionais são classificados conforme os seus níveis de criticidade demandados. Aplicações rotineiras demandam uma disponibilidade de 99%, aplicações essenciais tem a sua disponibilidade um pouco acima em 99,9%. No entanto, existem outros tipos de sistemas que demandam

uma disponibilidade bem superior, os sistemas críticos, que fazem a exigência de uma disponibilidade acima de 99,9999%. Já os sistemas críticos e seguros demandam uma disponibilidade ainda superior, chegando a disponibilidade mínima de 99,99999%.

A disponibilidade na grande maioria das vezes é representada por meio de número de noves. Este corresponde a quantidade algarismos consecutivos iguais a 9, localizados após a vírgula (MARWAH et al., 2010). A Equação 2.2 representa o cálculo para obter o resultado em número de noves.

$$N = \log_{10}(1 - A) \quad (2.2)$$

onde A corresponde a disponibilidade computada, por exemplo, pela Equação 3.1.

Ao contrário da disponibilidade, a indisponibilidade pode ser calculada por meio da Equação 2.3. A indisponibilidade trata da probabilidade do sistema não está disponível.

$$UA = 1 - A \quad (2.3)$$

Por intermédio dos resultados obtidos pela indisponibilidade é possível realizar o cálculo do *downtime*. Por exemplo, o *downtime* anual, representa o número de horas no qual o sistema se encontra em um estado não disponível, no período de um ano, este pode ser calculado por meio da Equação 2.4.

$$D = UA \times T \quad (2.4)$$

onde T é o período considerado, por exemplo, 8760 (horas) para um ano.

2.2 Políticas de Manutenção

Define-se manutenção como sendo um conjunto de ações desenvolvidas com o intuito de assegurar o bom funcionamento das máquinas e instalações. Os tipos de manutenção são as formas de realizar as intervenções nos instrumentos de produção, ou seja, nos equipamentos que compõem um determinado sistema (VIANA, 2002). A manutenção deve ser realizada sobre determinadas circunstâncias normalmente adversas, e seu principal objetivo é restaurar rapidamente o equipamento para seu estado de prontidão

operacional usando os recursos disponíveis, como também evitar possíveis quebras e falhas dos equipamentos (DHILLON, 2002).

As atividades relacionadas a manutenção visam evitar a degradação dos dispositivos e instalações causadas pelo seu uso ou desgaste natural. Esta degradação pode aparecer de diferentes formas desde a perda de desempenho do equipamento, fabricação de produtos com má qualidade, entre outros. A manutenção de equipamentos pode incluir algumas atividades relacionadas com o tratamento de falhas (reparo, detecção, investigação da causa da falha e correção da possível falha). Sendo assim, as atividades relacionadas a manutenção estão limitadas ao retorno de um equipamento ao seu estado original (sem ocorrência de falhas). Contudo, essas atividades também devem envolver as modificações das condições originais dos equipamentos, por meio do uso de melhorias para prevenir a ocorrência de falhas, reduzir os custos de manutenção e proporcionar um aumento de produtividade (XENOS, 1998).

Os custos relacionados a manutenção são uma parte importante dos custos totais gastos em operações, fabricação ou produção. Dependendo do setor específico, as despesas de manutenção podem representar entre 15 e 60% dos custos dos produtos produzidos (MOBLEY, 2002). Por exemplo, nas indústrias relacionadas a alimentos, as despesas médias de manutenção representam cerca de 15% dos custos dos produtos produzidos, enquanto os custos de manutenção de ferro e aço, celulose e papel e outras indústrias pesadas representam até 60% dos custos totais de produção (MOBLEY, 2002). Talvez o mais importante seja o fato que o gerenciamento ineficaz da manutenção afeta significativamente a capacidade de fabricação dos produtos e a qualidade competitiva no mercado mundial (MOBLEY, 2002).

Existem alguns diferentes tipos de classificação relacionadas às técnicas na realização da manutenção. Os principais tipos de manutenção são apresentados nas subseções a seguir.

2.2.1 Manutenção Preventiva

Manutenção preventiva representa ações realizadas na tentativa de prevenir e reduzir a ocorrência de falhas ou queda no desempenho. Essa manutenção realiza a substituição de partes do sistema, obedecendo a um planejamento baseado em períodos de

tempo pré-estabelecidos (ALMEIDA, 2005). Vale ressaltar que este tipo de manutenção é realizada em equipamentos operacionais, sendo assim, os equipamentos podem estar em estado ativo e sem apresentar falhas (GUIMARÃES et al., 2012).

A manutenção preventiva fornece um mecanismo para armazenar e obter um histórico completo de desempenho e reparo de equipamentos, incluindo custos, para equipamentos dentro de um processo de instalação (GEITNER; BLOCH, 2019). A manutenção preventiva é projetada mediante os registros dos equipamentos, sendo realizada em torno dos dados de descrição dos equipamentos. Nesse caso, os intervalos de inspeção do equipamento são atribuídos de acordo com a gravidade ou de acordo com as leis ou códigos de segurança, ou de acordo com as recomendações dos fabricantes e outras com base na vida útil do equipamento (GEITNER; BLOCH, 2019).

Em comparação com a manutenção corretiva (ver Seção 2.2.2) a manutenção preventiva apresenta um maior custo, pois as peças devem ser trocadas e os equipamentos devem ser substituídos antes da ocorrência de falhas ou avarias. Porém, a ocorrência de falhas e paradas inesperadas são reduzidas, sendo assim há o aumento da disponibilidade. Desta forma, analisando o custo, a manutenção preventiva é mais vantajosa que a manutenção corretiva. Isso ocorre devido ao fato de que as paradas são programadas, diminuindo assim a ocorrência de paradas inesperadas dos equipamentos (XENOS, 1998).

2.2.2 Manutenção Corretiva

A manutenção corretiva realiza a correção de uma falha que ocorreu de maneira aleatória, ou seja, trata-se de uma correção da falha ou quando o sistema está apresentando um desempenho menor que o esperado. A manutenção corretiva implica em altos custos, pois causa perda da produção, além do custo de reparo do equipamento (OTANI; MACHADO, 2008). Normalmente envolve maior tempo de reparo, pois a atividade não foi planejada e maior custo. O custo está associado às horas mobilizadas para correção das falhas, parada de linha de produção, perdas de produtividade e perda da qualidade do serviço.

O objetivo da manutenção corretiva é trazer o equipamento de volta ao estado operante ou disponível no menor tempo possível (FOGLIATO; RIBEIRO, 2009). A manutenção corretiva é dividida em dois tipos: manutenção corretiva planejada e

manutenção corretiva não planejada.

- A manutenção corretiva planejada corresponde a um tipo de manutenção onde o equipamento apresenta um desempenho menor do que o esperado ou uma falha, isto é, pela atuação em função de acompanhamento preditivo ou pela decisão de operar até a quebra por decisão gerencial (PINTO; XAVIER, 2009).
- Manutenção corretiva não planejada é um tipo de manutenção que acontece após a ocorrência de uma falha ou perda de desempenho de um determinado equipamento, sem que haja tempo para a preparação dos serviços (OTANI; MACHADO, 2008). A manutenção corretiva não planejada pode resultar em altos custos e aspectos, tais como, a redução da vida útil das máquinas e/ou equipamentos e paradas de máquinas no processo produtivo (PINTO; XAVIER, 2012).

2.2.3 Outros tipos de Manutenção

Essa seção tem como objetivo apresentar, brevemente, outros tipos de manutenção existentes. Por exemplo, a manutenção preditiva é baseada nas condições do equipamento, permitindo o acompanhamento do dispositivo por meio de medições, acompanhamento de parâmetros ou condições de desempenho do equipamento (RIBEIRO; GOMES, 2016). A principal diferença entre a manutenção preventiva e a manutenção preditiva é que a manutenção preventiva é realizada assim que um intervalo de tempo predeterminado tiver decorrido, enquanto a manutenção preditiva requer verificação em intervalos predeterminados com a ação de manutenção realizada apenas se a inspeção mostrar que é necessário.

Existe também a manutenção detectiva que surgiu na década de 1990, e sua denominação está ligada à palavra Detectar. A identificação de falhas ocultas é de grande valia para garantir a confiabilidade (GARCIA, 2017). Manutenção detectiva atua de forma a detectar falhas ocultas nos equipamentos que não podem ser percebidas pela equipe de manutenção e operação. A manutenção detectiva é aplicada em sistemas de proteção, comando e controle em que a ocorrência de falha tem que ser a mínima possível (BOTELHO; CABRAL, 2017). Uma falha oculta é aquela em que a perda da sua função não é evidente para o operador. As falhas ocultas geralmente não são encontradas facilmente pelos operadores e não se trata de uma falha funcional (SEIXAS; BRASIL, 2011).

2.3 Redes de Petri Estocásticas

A rede de Petri (REISIG, 2012) é uma ferramenta para modelar e analisar sistemas de eventos discretos. O tempo e as escolhas probabilísticas são aspectos essenciais para um modelo de avaliação de desempenho. A rede de Petri estocástica (SPN) é uma extensão de rede de Petri (PN) (MURATA, 1989). As SPN adicionam tempo ao formalismo de redes de Petri, contudo os tempos que são associados às transições temporizadas são distribuídos exponencialmente, enquanto o tempo associado às transições imediatas é zero. As transições estocásticas modelam atividades através do tempo associado as transições, sendo assim o tempo da transição estocástica corresponde ao tempo de execução da atividade, e o disparo da transição estocástica corresponde ao término da atividade.

As redes de Petri estocásticas (SPN) surgiram em 1980 como um formalismo para a descrição de Sistemas Dinâmicos de Eventos Discretos (DEDS), cujo comportamento dinâmico poderia ser representado por meio de cadeias de Markov em tempo contínuo (BALBO, 2000) (MARSAN et al., 1986).

Uma SPN é definida pela 9-tupla $SPN = \{P, T, I, O, H, \Pi, G, M_0, Atts\}$ (MARSAN et al., 1986) (GERMAN, 2000), onde:

- $P = \{p_1, p_2, \dots, p_n\}$ é o conjunto de lugares;
- $T = \{t_1, t_2, \dots, t_m\}$ é o conjunto de transições imediatas e temporizadas, $P \cap T = \emptyset$
- $I \in (\mathbb{N}^n \rightarrow \mathbb{N})^{n \times m}$ é a matriz que representa os arcos de entrada (que podem ser dependentes de marcações). Uma transição é ativada apenas se houver tokens suficientes em todos os locais de entrada;
- $O \in (\mathbb{N}^n \rightarrow \mathbb{N})^{n \times m}$ é a matriz que representa os arcos de saída (que podem ser dependentes de marcações). Quando uma transição é acionada, remove o número de tokens especificados pelos arcos de entrada dos lugares de entrada e adiciona a quantidade de tokens fornecidos pelos arcos de saída a todos os lugares de saída;
- $H \in (\mathbb{N}^n \rightarrow \mathbb{N})^{n \times m}$ é a matriz que representa os arcos inibidores (que podem ser dependentes de marcações). Na presença de um arco inibidor, uma transição é ativada para disparar apenas se todos os lugares conectados por um arco inibidor contiverem menos tokens que a multiplicidade do arco;
- $\Pi \in \mathbb{N}^m$ é um vetor que associa o nível de prioridade a cada transição. Sempre que houver várias transições acionadas ao mesmo tempo aquela com a maior prioridade é acionada primeiro e leva a uma mudança de estado;

- $G \in (\mathbb{N}^m \rightarrow \{true, false\})^m$ é o vetor que associa uma condição de guarda relacionada a marcação do lugar à cada transição;
- $M_0 \in \mathbb{N}^n$ é o vetor que associa uma marcação inicial de cada lugar (estado inicial);
- $Atss: (Dist, W, G, Policy, Concurrency)^m$ compreende um conjunto de atributos para as transições m , em que
 - $Dist \in \mathbb{N}^m \rightarrow f$ é uma função de distribuição de probabilidade, com o disparo dependente da marcação. Em uma rede de Petri estocástica temporalizada, o tempo deve decorrer entre a ativação e o disparo de uma transição. O tempo real de disparo é uma variável aleatória, para a qual a distribuição é especificada por f . Diferimos entre transições imediatas ($f = 0$) e transições temporizadas, para as quais o domínio de f é $(0, \infty)$
 - $W \in \mathbb{R}^+$ corresponde a função de peso, que representa um peso para o acionamento de determinadas transições imediatas ou uma taxa λ_t para transições temporizadas. O último só é significativo para o caso padrão de transições temporizadas com atrasos de disparo distribuídos exponencialmente. Para transições imediatas, o valor especifica uma probabilidade relativa de acionamento da transição quando há várias transições imediatas ativadas em uma marcação e todas têm a mesma probabilidade. Uma escolha aleatória é então aplicada usando as probabilidades em peso w_t .
 - $G \in \mathbb{N}^n \rightarrow \{false, true\}$ corresponde a uma função que atribui uma condição de guarda relacionada às marcações para cada transição. Dependendo da marcação atual, as transições podem não ser acionadas (elas estão desativadas) quando a função de proteção retorna falso. Esta é uma extensão dos arcos inibidores.
 - $Policy \in \{prd, prs\}$ é a política de preempção. Onde prd corresponde a repetição preemptiva, significa que quando uma transição preemptiva é ativada novamente, o tempo de disparo decorrido anteriormente é perdido. Já o prs corresponde ao resumo preemptivo, no qual o tempo de disparo está relacionado a uma transição preemptiva que é retomada quando a transição é ativada novamente.
 - $Concurrency \in \{ss, is\}$ corresponde ao grau de simultaneidade das transições, em que ss representa a semântica de servidor único. Já o is representa a semântica de servidores infinitos no mesmo sentido que nos modelos de filas. Transições com política podem ser entendidas como sendo uma transição

individual para cada conjunto de tokens de entrada, todos executando em paralelo.

A Figura 1 representa uma rede de Petri estocástica através de dois tipos de nós: um nó representa os lugares e o outro nó representa as transições. Em uma SPN, os lugares são representados por círculos (LU, LU2, LU3 e LU4) e as transições por retângulos (T0, T1 e T2), um lugar só pode ser ligado a uma transição, através de um arco. E uma transição será ligada a um lugar através de um arco, ou seja, nunca haverá um lugar apontando para outro lugar ou transição apontando para uma outra transição (BRIS, 2013). Uma SPN pode apresentar dois tipos de transições, a imediata e estocástica. A transição imediata (T0) é disparada sempre que a condição de acionamento for satisfeita. Já as transições estocásticas (T1 e T2) são habilitadas somente se estiverem com as condições de acionamento satisfeitas e ter decorrido um período de tempo exponencialmente distribuído associado a cada transição. A primeira transição a ser acionada é T0, pois existe um token no lugar LU, habilitando assim o disparo desta transição. O disparo da transição T0 consome um token de LU e gera um token em LU2. Sendo assim, o disparo de T0 acaba por habilitar a transição T1, visto que existe um token no lugar LU2, porém esta transição só será disparada após ser decorrido o tempo de ativação desta transição. Quando este tempo é alcançado, esta transição é disparada consumindo o token em LU2 e gerando uma saída em LU3. Após o disparo de T1, a transição T2 pode ser disparada, visto que existe um token em LU3, porém este só é disparado após o tempo de ativação para a transição T2 for satisfeito. O disparo de T2 consome um token em LU3 e tem como saída um token em LU4.

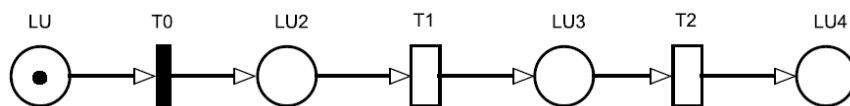


Figura 1 – Modelo SPN

2.4 Diagramas de Bloco de Confiabilidade

Os diagramas de bloco de confiabilidade (RBD) fornecem um método para avaliação de disponibilidade e/ou confiabilidade através do mapeamento em blocos, que se configuram em série ou paralelo. Cada bloco que compõe o modelo pode representar falhas em um

componente individual ou em um subsistema (TORELL; AVELAR, 2004) (MARSAN et al., 1986).

O RBD pode representar um sistema inteiro ou qualquer subconjunto ou combinação do sistema para o qual seja necessário realizar uma determinada análise. O RBD permite avaliar o impacto da disponibilidade e confiabilidade de um equipamento específico do sistema, na disponibilidade e confiabilidade de todo o sistema (TORELL; AVELAR, 2004).

A Figura 2(a) representa um modelo RBD, em série. Nesse caso, a falha de qualquer componente resulta em uma falha do sistema. Assumindo um sistema com n componentes independentes, a confiabilidade ou disponibilidade são obtidas por:

$$P_s = \prod_{i=1}^n P_i \quad (2.5)$$

onde P_i é a confiabilidade $R_i(t)$, disponibilidade instantânea ($A_i(t)$) ou disponibilidade estacionária (A_i) do bloco i .

Para uma disposição paralela (Figura 2(b)), o sistema funciona se pelo menos um componente funcionar. Assumindo um sistema com n componentes independentes, a confiabilidade, disponibilidade instantânea ou disponibilidade estacionária é obtida por: onde

$$P_s = 1 - \prod_{i=1}^n (1 - P_i) \quad (2.6)$$

onde P_i é a confiabilidade $R_i(t)$, disponibilidade instantânea ($A_i(t)$) ou disponibilidade estacionária (A_i) do bloco i .

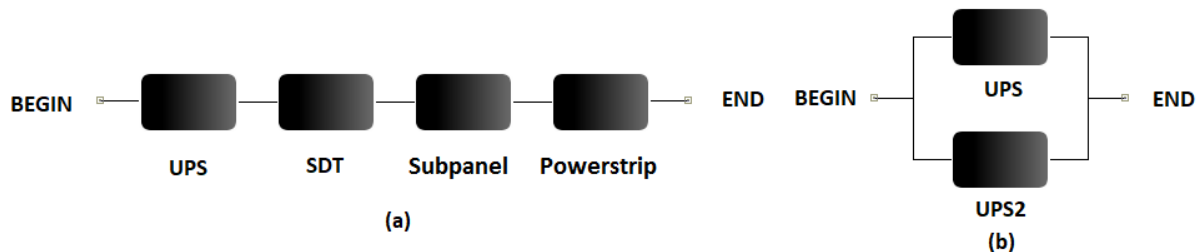


Figura 2 – Modelo RBD

2.5 Importância para a Disponibilidade

O índice de importância para a disponibilidade (I_i^A) é baseado no índice de importância para a confiabilidade. A importância para a confiabilidade considera o tempo médio para falha dos componentes e o tempo médio de recuperação dos componentes. Dessa forma, ao invés de utilizar a confiabilidade, a métrica utilizada é a disponibilidade. Este índice inclui além do tempo médio para falha, o tempo médio de recuperação dos componentes (FIGUEIRÊDO; MACIEL, 2011). O cálculo utilizado é mostrado na Equação 2.7, e os resultados podem ser normalizados pelo maior valor conforme mostrado na Equação 2.8.

$$I_i^A = A_s(1_i, P^i) - A_s(0_i, P^i), \quad (2.7)$$

onde P^i representa o vetor de disponibilidades dos componentes com o i ésimo componente removido; 0_i representa a condição quando o componente i é falho e 1_i a condição quando o componente i está no modo operacional. A Tabela 4 ilustra o exemplo do cálculo de importância para a disponibilidade.

$$I_{n_i} = \frac{I_i}{I_x} \quad (2.8)$$

onde I_{n_i} é o índice normalizado para o componente i ; I_i é o valor do índice não normalizado para o componente i e, I_x é o valor do maior índice não normalizado entre os componentes.

A Figura 3 apresenta três equipamentos ligados em séries. Deseja-se saber qual destes equipamentos é o dispositivo mais sensível para a disponibilidade do sistema e, assim, precisa de maior atenção.



Figura 3 – Diagrama de bloco de confiabilidade – Composto por três equipamentos em série

Os resultados do índice de importância para a disponibilidade podem ser obtidos na Tabela 4 que ilustra os valores para a disponibilidade e o resultado da importância para a disponibilidade para cada equipamento presente no modelo. Através do cálculo da importância para a disponibilidade, foi possível identificar que o equipamento 2 é o mais importante e precisa de maior atenção.

Tabela 1 – Importância para a Disponibilidade

	Equipamento 1	Equipamento 2	Equipamento 3
Valores da Disponibilidade	0,90	0,80	0,85
Importância para a Disponibilidade	0,68	0,76	0,72

2.6 Planejamento de Experimentos

Planejamento de experimentos - *Design of Experiments (DoE)* - é um método estruturado e organizado para determinar a relação entre os fatores e a saída de um processo (GU; JIN, 2006). Os fatores são as diferentes variáveis que afetam o desempenho de saída do sistema. Já os níveis de um fator são os valores específicos que eles podem ter. O DoE permite a minimização do número de experimentos que devem ser realizados reduzindo o tempo necessário para se obter a informação desejada. Ao se utilizar o DoE é possível identificar os efeitos de cada parâmetro de entrada no resultado obtido.

Os métodos DoE podem ser uma parte importante da otimização de um sistema, resultando em recomendações definidas de planejamento ou reprojetado. Todos esses métodos envolvem as atividades de planejamento experimental, realização de experimentos e adaptação dos modelos ao resultado das saídas. Um elemento essencial na aplicação dos métodos de DoE é o uso do procedimento chamado randomização. A randomização envolve tomar muitas decisões de planejamento experimental usando uma abordagem aleatória ou não padronizada (ALLEN, 2006).

O planejamento de experimentos ajuda a separar os efeitos de vários fatores que podem alterar o desempenho. Além de determinar se um fator tem um efeito significativo ou se a diferença observada é devido as variações aleatórias causadas por parâmetros que não foram controlados e/ou erros de medição.

O DoE projeta testes estratégicos e planejados para garantir que metas e objetivos

específicos sejam realizados por meio da análise dos dados gerados. A premissa básica do DoE é que um ou mais fatores são variados sistematicamente entre um conjunto de entradas (ao invés de apenas uma única condição) (CHUNG, 2003). Existe alguns tipos de projetos experimentais tais como: planejamento de experimento de um único fator, planejamento de experimento de dois fatores, planejamento de experimento de vários fatores e planejamento de experimento 2k (JALBY et al., 2012). Cada um desses tipos de projetos experimentais, são apresentados e detalhados nas subseções abaixo.

2.6.1 Planejamento de experimento de um único fator

Ocorre apenas a variação de um fator por vez, os fatores são as diferentes variáveis que afetam o desempenho de saída do sistema. Essas variáveis são controláveis, pois o projetista pode variar os níveis no sistema real nos modelos de simulação (JALBY et al., 2012).

Essa técnica permite ao projetista analisar um fator de entrada, variando esse valor, enquanto as outras entradas são constantes. Está técnica é usada para comparar várias alternativas de uma única variável. Não havendo limites para o número de níveis que o fator pode assumir, este número de níveis deve ser maior do que 2 (JAIN, 1990).

Contudo essa abordagem ignora possíveis interações entre duas variáveis de entrada. Sendo assim, se os fatores apresentarem interação entre si pode afetar os resultados obtidos, podendo gerar resultados incoerentes (JAIN, 1990).

2.6.2 Planejamento de experimento de dois fatores

Ocorre ao selecionar dois elementos particulares e estes são examinados em diferentes níveis. Um planejamento de dois fatores é usado quando existem dois parâmetros cuidadosamente controlados e variados para estudar seu impacto no desempenho. Por exemplo, supondo que existam os elementos A e B, e que o número de cenários a serem avaliados correspondem ao número de níveis do elemento A multiplicado pelo número de níveis do elemento B. Nesse caso, os elementos de A são um fator e os elementos de B são outro fator (JAIN, 1990) (JALBY et al., 2012).

Supõe-se que os fatores são categóricos ou estão sendo tratados como tal pelo

analista. Se os fatores forem quantitativos um modelo de regressão pode ser usado (JAIN, 1990).

2.6.3 Planejamento de experimento de vários fatores

Essa técnica é utilizada quando se têm vários fatores e vários níveis. Sendo assim, dependendo do número de fatores e níveis pode acarretar em um grande número de alternativas a serem analisadas (JALBY et al., 2012). A vantagem de utilizar essa técnica é que todas as combinações possíveis de configuração e carga de trabalho são analisadas. Dessa forma, pode-se encontrar o efeito de todos os fatores, incluindo os fatores secundários e suas interações (JAIN, 1990).

O problema dessa abordagem seria o custo do experimento. Podendo demandar muito tempo e recursos para realizar os experimentos, principalmente quando se leva em consideração a possibilidade de que cada um desses experimentos precisa ser repetido várias vezes (JAIN, 1990). Por exemplo, uma estação de trabalho, o número de experimentos seria $n = (4\text{CPUs}) \times (4\text{ memórias}) \times (5\text{ unidades de disco}) \times (4\text{ estações de trabalho})$ sendo assim, o valor de n é igual a 320 experimentos.

2.6.4 Planejamento de experimento 2k

Nesta técnica ao invés de utilizar todos os níveis de cada elemento, restringe-se esse número ao maior e ao menor nível de cada. Os resultados correspondem a dois níveis para cada elemento (JALBY et al., 2012).

Essa técnica facilita a análise e ajuda a classificar os fatores na ordem de impacto. No início de uma análise de desempenho, o número de fatores e seus níveis geralmente são elevados. Um planejamento fatorial completo (com vários níveis) com um número muito grande de níveis e fatores pode comprometer o esforço para obtenção dos resultados. O passo inicial seria reduzir o número de fatores e escolher apenas os que possuem um impacto significativo no desempenho (JAIN, 1990).

Geralmente o desempenho diminui ou aumenta consideravelmente conforme o fator é aumentado do mínimo para o máximo. Por exemplo, espera-se que o desempenho melhore à medida que o tamanho da memória aumenta ou o número de unidades de disco

amente. Nesses casos, pode-se começar o experimento com o nível mínimo e máximo do fator (JAIN, 1990).

2.7 *Service Level Agreement*

Service Level Agreement (SLA) é um acordo de nível de serviço, e resulta em um documento que define um conjunto de compromissos entre uma empresa prestadora de serviços e um cliente (ALPENDRE, 2006). Sendo assim, por meio deste documento o cliente tem acesso aos serviços oferecidos pela empresa. Este documento descreve os serviços a serem obtidos e as taxas a serem alcançadas para o cumprimento de todos os compromissos acordados (STURM et al., 2000).

Um contrato de *SLA* pode ser definido como sendo um acordo entre o serviço de suporte e o usuário que quantifica o serviço mínimo aceitável (HILES, 2004). Os contratos de *SLA* são particularmente úteis no processamento crítico, podendo ser documentos bem complexos e longos ou podem ser simples. De toda forma, *SLAs* são indispensáveis para proporcionar bons relacionamentos de serviço entre o provedor de serviço e o cliente (HILES, 2004).

Um *SLA* também especifica as medidas a serem tomadas em caso de desvio e falta de atendimento ao serviço contratado. As afirmações do provedor de serviços são baseadas em uma definição detalhada dos parâmetros de serviço (LUDWIG et al., 2003). O uso de contratos de *SLA* continuará a crescer e, eventualmente, os *SLAs* se tornarão o modelo de negócios predominante para a entrega de um grande número de produtos e serviços. O ciclo de vida do *SLA* consiste nas seguintes fases: desenvolvimento do *SLA*, negociação e vendas, implementação, execução e avaliação conforme apresentado na Figura 4 (BLOKDIJK, 2008).

Na fase de desenvolvimento do *SLA*, os provedores de serviços devem entender e levar em consideração a importância de uma boa definição e caracterização do produto. Esse processo de desenvolvimento leva em consideração as características, definições e busca encontrar todos os aspectos necessários para a oferta do produto ou serviço. Contudo, mesmo que os contratos de *SLAs* sejam frequentemente tratados de forma similar a um produto, eles são recursos com valor agregado ao produto ou serviço (BLOKDIJK, 2008).

Após a fase de desenvolvimento do *SLA*, este é colocado no mercado ou associado

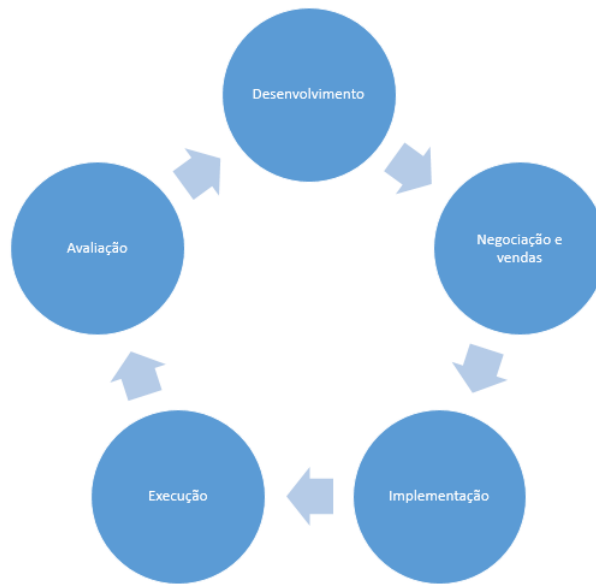


Figura 4 – Ciclo de vida do *SLA*

a determinado produto. Em alguns cenários, o contrato de *SLA* pode estar vinculado a um modelo, geralmente isso ocorre nas ofertas de serviços mais genéricas e tecnicamente rotineiras ou quando o provedor de serviços está desenvolvendo o contrato de *SLA* de nível mais baixo e/ou padrão em uma oferta de *SLA* de várias camadas (BLOKDIJK, 2008). O objetivo esperado nesta fase é a venda e execução dos contratos. Em alguns casos, os requisitos do cliente podem ser tão rigorosos ou únicos que o *SLA* deve ser desenvolvido através de negociações preestabelecidas entre ambas as partes (BLOKDIJK, 2008).

Na fase de implementação, os serviços são solicitados, ativados e configurados em conformidade com o contrato de *SLA*. Com isso algumas medidas são tomadas, novos recursos de monitoramento são instalados, limites são definidos, relatórios adicionais são estabelecidos (BLOKDIJK, 2008).

Na fase de execução correspondente às atividades relacionadas ao serviço que está sendo entregue, isso inclui a avaliação direta dos contratos de *SLA*. Eventos extraordinários, como por exemplo, interrupções, tempo de inatividade de manutenção e até falha da capacidade de medir o desempenho, devem ser registrados e medidos, e o impacto na realização das demandas do negócio deve ser avaliado e relatado nesta fase (BLOKDIJK, 2008).

O *SLA* deve ser avaliado constantemente, existem dois tipos de avaliações: avaliação focada no cliente e avaliação focada no provedor. As avaliações focadas no cliente concentram-se no desempenho do provedor de serviços do ponto de vista do cliente.

A principal métrica desse tipo de avaliação é a conformidade com o contrato de *SLA*, principalmente a disponibilidade e a satisfação do cliente. Já as avaliações focadas no provedor concentram-se na execução do contrato de *SLA* como um caso de negócios dentro da estratégia do *SLA*. A intenção desse tipo de revisão é otimizar o uso do *SLA* pelo provedor de serviços, com o objetivo de melhorar a lucratividade ou reduzir as possíveis penalidades ou multas recebidas, devido ao descumprimento de determinadas cláusulas do contrato ofertado. As principais métricas que envolve esse tipo de avaliação são a qualidade do serviço ofertado, lucratividade do contrato e melhorias recomendadas (BLOKDIJK, 2008).

2.8 *Data center*

Os *data centers* são responsáveis pela hospedagem de recursos essenciais para a computação, pois trata-se de ambientes controlados que possuem um gerenciamento centralizado, permitindo assim que a empresa opere de acordo com as necessidades do negócio. Um *data center* é uma instalação composta por servidores que apresentam dispositivos de armazenamento, de rede (switches, roteadores e cabos), e necessitam de sistemas de distribuição de energia, sistemas de resfriamento, servidores web e de aplicativos, servidores de arquivo e impressão (ARREGOCES; PORTOLANI, 2003). Normalmente, os *data centers* são adotados como host de sites para processar transações comerciais proteger dados (ex., registros financeiros e históricos médicos de pacientes).

Portanto, os *data centers* devem possuir altos níveis de disponibilidade para suportar novos requisitos de tecnologia. Para construir um *data center* algumas infraestruturas devem ser analisadas, por exemplo, elétrica, resfriamento, conectividade, espaço, proteções contra incêndio e monitoramento de temperatura (CALLOU, 2013). Os *data centers* precisam ser confiáveis. Para obter altos valores de confiabilidade, as infraestruturas de *data centers* devem ser compostas de, por exemplo: (i) rede redundante para manter o sistema funcionando se um dispositivo de rede falhar, e (ii) múltiplas fontes de energia elétrica, responsáveis pelo suporte da carga elétrica do *data center* (CALLOU, 2013).

Os *data centers* proporcionam vantagens na utilização de tal serviço, por exemplo: suporte das operações do negócio 24 horas, redução do custo total, gasto com manutenção e operação do ambiente, facilidade na implantação de aplicativos, flexibilidade,

disponibilidade, escalabilidade, segurança e gerenciabilidade. Os *data centers* por serem responsáveis por armazenar recursos importantes para as empresas, este deve funcionar com uma operação de 24X7. Sendo assim, é necessário que este ambiente seja provido de alguns recursos para poder atender as demandas, tais como: capacidade de energia, capacidade de refrigeração, cabeamento, controles de temperatura e umidade, sistemas de prevenção contra incêndio, segurança física, espaço em rack e pisos elevados (ARREGOCES; PORTOLANI, 2003).

2.8.1 Infraestruturas do *Data center*

Esta subseção apresenta os componentes básicos que compõem a infraestrutura de um *data center*. Os *data centers* são compostos por três tipos básicos de infraestrutura: infraestrutura de TI, infraestrutura de refrigeração e infraestrutura elétrica, conforme pode ser observado na Figura 5.

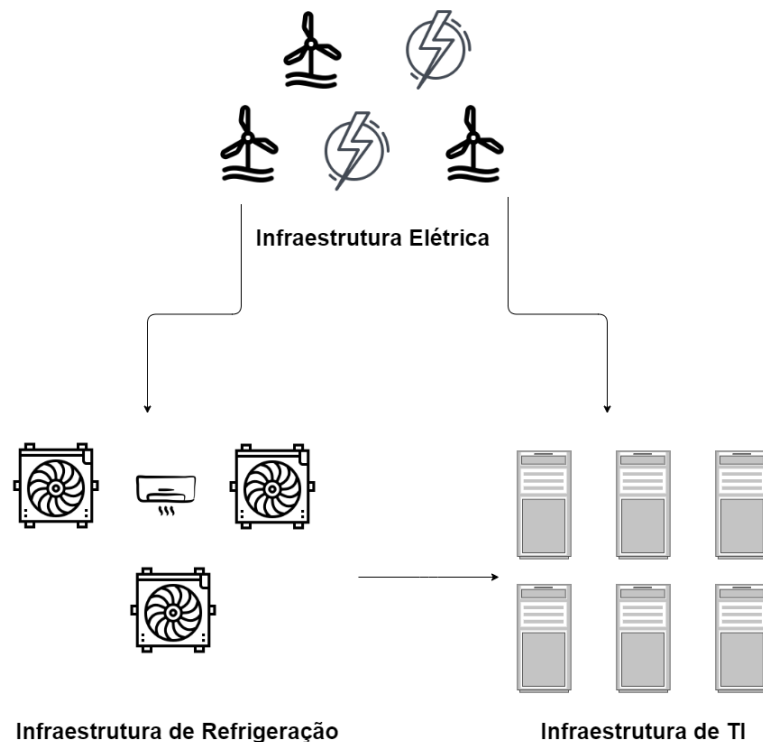


Figura 5 – Infraestruturas do *Data center*

A infraestrutura de TI é responsável pelo processamento e armazenamento de informações utilizadas por determinadas empresas. Este ambiente geralmente é composto por três componentes principais, equipamentos de armazenamento, equipamento de rede e

servidores (MACIEL et al., 2010). Os dispositivos de armazenamento normalmente são conectados por meio de uma rede SAN (*Storage Area Network*), porém os dispositivos de armazenamento também podem ser conectados a redes NAS (*Network Attached Storage*) através da Ethernet. Os equipamentos de rede são compostos geralmente por switches, hubs, roteadores, pontes e gateways, esses são responsáveis por toda comunicação interna ou externa e por último os servidores, este é organizado em uma arquitetura composta por multicamada, ou seja, ocorre a separação das camadas (BARROSO et al., 2003).

A infraestrutura de refrigeração é responsável por controlar e regular a temperatura da sala onde está a infraestrutura de TI. Geralmente, o CRAC (*Computer Room Air Conditioning*), é responsável pela retirada do calor liberado pelos dispositivos de TI através de um sistema de distribuição de água refrigerada. O calor é removido da água através de trabalho termodinâmico, a fim de melhorar o ciclo de refrigeração do chiller. A água refrigerada é devolvida ao CRAC, já o calor absorvido pelo refrigerador é levado para uma corrente de água secundária. O circuito secundário transfere o calor para o ambiente externo em uma torre de resfriamento. A infraestrutura de refrigeração poderá necessitar de cerca de 38% do consumo total de energia do *data center* (NI; BAI, 2017).

A infraestrutura elétrica é responsável pelo fornecimento ininterrupto de energia elétrica condicionada à frequência corretas e a tensão para a infraestruturas de refrigeração e de TI. Os *data centers* são tipicamente energizados através da rede elétrica. No entanto, também existem *data center* que usam diesel, energia solar, eólica, hidrogênio (células de combustível) entre outras fontes de energia. A energia elétrica proveniente de fontes externas é dividida entre o equipamento de TI, as instalações de infraestrutura e os sistemas de suporte pelos equipamentos de comutação (DAYARATHNA et al., 2015).

A energia principal que alimenta o *data center* corresponde à energia elétrica que pode vir de diferentes subestações elétricas. Uma subestação elétrica corresponde a uma estação secundária de um sistema que proporciona a geração, transmissão e distribuição de eletricidade. A energia elétrica pode fluir através de diferentes subestações.

2.8.2 Distância Euclidiana

A distância Euclidiana é a matriz mais usada para calcular a distância entre dois vetores (DATTORRO, 2010). A distância euclidiana computa a distância entre os pontos

$P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$, que correspondem a dois pontos conforme pode ser observado na Figura 6. Então, a distância Euclidiana entre x e y é obtida através da fórmula:

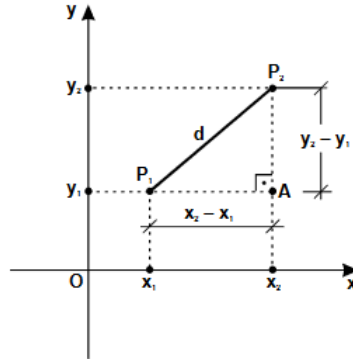


Figura 6 – Pontos P_1 e P_2 marcados no plano (Fonte (NINA, 2011))

$$D = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (2.9)$$

onde: X_1 ou Y_1 correspondem a distância inicial, e X_2 ou Y_2 representam a distância final. Sendo assim, a distância Euclidiana é mede a distância entre esses dois pontos.

3 Trabalhos Relacionados

Nesse capítulo são apresentados alguns trabalhos desenvolvidos com o objetivo de aumentar a disponibilidade e reduzir a ocorrência de falhas através da utilização de políticas de manutenção. Além disso, outro fator que deve ser levado em consideração é a necessidade de se possuir ambientes cada vez mais disponíveis. Sendo assim, pesquisas são realizadas com o intuito de permitir que tais ambientes fiquem sempre disponíveis. Outro fator que também deve ser levado em consideração é a utilização de contratos de *SLA*, realizando assim, uma análise referente ao custo de tais contratos e o impacto de cada contrato na disponibilidade e no custo do sistema. Os trabalhos relacionados foram encontrados na literatura ao se utilizar a seguinte string de busca: (*"All Metadata":maintenance policies*) OR (*"All Metadata":data center*) OR (*"All Metadata":Design of Experiments*) OR (*"All Metadata":Service Level Agreement*) OR (*"All Metadata":electrical infrastructure*) OR (*"All Metadata":optimization*) OR (*"All Metadata":availability*).

Uma proposta de otimização da emissão de gases de efeito estufa (*GEE - Greenhouse gas*) em ambiente de *data center* tolerante a falhas foi apresentada em (BOSSE et al., 2016). A modelagem adotada pelos autores fez uso de cadeia de Markov, com o objetivo de minimizar os custos e as emissões de GEE sem reduzir a disponibilidade do sistema. Os autores fizeram uso de dois algoritmos distintos, um algoritmo genético (AG), é um algoritmo *tabu search* (*TS*). Estes algoritmos foram adaptados para fornecer a tolerância as falhas, reduzir custos e emissões de GEE. Os autores adotaram técnicas de otimização com foco na redução da emissão de gases de efeito estufa, porém nada foi mencionado em relação a DoE e infraestrutura elétrica de *data center*.

Os autores em (Callou et al., 2010) apresentaram uma avaliação referente às políticas de manutenção com diferentes acordos de nível de serviço (*SLA - Service Level Agreement*). Uma modelagem híbrida foi adotada para considerar as vantagens de diferentes redes de Petri estocásticas (*SPN - Stochastic Petri Net*) e diagramas de bloco de confiabilidade (*RBD - Reliability Block Diagram*) para avaliar infraestruturas de energia do *data center*, levando em consideração políticas de manutenção. Questões de confiabilidade, custo e sustentabilidade também foram computadas na avaliação de infraestruturas do *data center*. No entanto, este trabalho focou somente na manutenção corretiva.

Os autores em (ADDABBO et al., 2015) propuseram modelos de disponibilidade

e confiabilidade que levam em consideração *SLA*. Os autores analisaram um caso específico em um ambiente farmacêutico de diagnóstico. O foco está na classe de instrumentação científica, Cromatógrafo Líquido de Alto Desempenho (*HPLC - Cromatógrafo Líquido de Alto Desempenho*). O objetivo deste trabalho é mostrar como essa análise pode ajudar no agendamento de manutenção preventiva, reduzindo os custos, os contratos de *SLA* para fornecedores de serviços, aumentando simultaneamente a disponibilidade e a confiabilidade dos instrumentos. Os autores analisaram diferentes cenários, aplicando diferentes políticas de manutenção vinculadas a contrato de serviço. O foco dos autores foi somente na manutenção preventiva, cenários para a manutenção corretiva não foram avaliados.

Em (SAID; TAGHIPOUR, 2016) é apresentada uma análise para quantificar os efeitos de múltiplas políticas de manutenção preventivas sobre a confiabilidade e disponibilidade de equipamentos móveis. Nesse estudo foi considerado um sistema complexo e reparável, como um caminhão de transporte subterrâneo submetido a manutenção corretiva e vários tipos de manutenção preventiva. Os autores desenvolveram uma função de probabilidade para estimar os parâmetros do modelo, que incluem os parâmetros do processo de falha e os efeitos de manutenção preventiva. Este trabalho propôs modelos de disponibilidade e confiabilidade e manutenção, mas o foco não foi em *data centers*.

Os autores em (WANG et al., 2017) fizeram uma análise de desempenho disponibilidade e confiabilidade de um sistema degradado multi-estados, através de cadeias de Markov. Este trabalho analisa sistemas multiestados com diferentes níveis de desempenho para cada estado. Os autores levam em consideração possíveis degradações do ambiente analisado que podem afetar o sistema. Os autores consideram a disponibilidade do sistema, quando a disponibilidade do sistema está em um de seus últimos estados degradados, é realizada uma manutenção preventiva com o intuito de restaurar o sistema para um dos estados anteriores com maior desempenho. Este trabalho não mencionou manutenção corretiva.

Em (LIU; TRIVEDI, 2006) é proposta uma abordagem geral de quantificação de sobrevivência que é aplicável a uma ampla gama de arquiteturas de sistemas e aplicativos. Essa abordagem fornece informações úteis para melhor compreensão do estado estacionário do sistema e comportamento transitório sob várias condições adversas. Os autores fazem uso de cadeia de Markov para modelar manutenção preventiva, porém nada foi mencionado em relação a manutenção corretiva.

Silva et al. ([SILVA et al., 2013](#)) apresentaram uma metodologia para estimar métricas de impacto na sustentabilidade e confiabilidade, apoiada por um ambiente integrado conhecido como ASTRO. Essa ferramenta considera as vantagens dos diagramas de blocos de confiabilidade e das redes de Petri estocásticas, para estimar o impacto da sustentabilidade com base na avaliação do ciclo de vida. A ferramenta suporta uma modelagem hierárquica para permitir a criação de modelos mais simples, a fim de diminuir problemas relacionados à representação de sistemas maiores. Sendo assim, no artigo foi proposto um ambiente integrado para dar suporte à avaliação de confiabilidade e sustentabilidade de infraestruturas do *data center*. No entanto, manutenção não foi o foco desta pesquisa.

Os autores em ([SILVA et al., 2014](#)) propuseram modelos de desempenho com a finalidade de avaliar os sistemas de computação em nuvem distribuídos em vários *data centers*, levando em consideração a ocorrência de desastres. Este trabalho apresenta uma abordagem para avaliar a capacidade de desempenho nos sistemas IaaS implantados em *data centers* geograficamente distribuídos. O ambiente proposto é conhecido como GeoClouds Modcs, permite a avaliação do impacto da ocorrência de desastres, o tempo de transmissão da máquina virtual e as diferentes cargas do cliente no desempenho do sistema. Os modelos de disponibilidade e confiabilidade, mas não levaram em consideração a manutenção preventiva e nem a corretiva.

Um modelo de benchmarking para índices de confiabilidade é apresentado usando a *Data Envelopment Analysis (DEA)* em ([BICHPURIYA et al., 2011](#)). O modelo *DEA* é apresentado para comparar os utilitários com base em seus índices de confiabilidade. O modelo encontra o componente de melhor desempenho entre o conjunto de componentes semelhantes com base nos conjuntos de entradas (porcentagem de alta-tensão (*HT - Percentage of High Tension*), alimentadores, porcentagem de cabos subterrâneos (*UG - Underground Cables*), densidade de carga e vida útil restante dos transformadores) e saídas (*SAIDI, SAIFI, ASAI*). Nada foi mencionado em relação a políticas de manutenção.

Os autores em ([BALANICI; PACHNICKE, 2018](#)) apresentaram redes eletrostáticas híbridas em um *data center* sob medida para diferentes classes de tráfego. Uma modelagem matemática foi adotada para possibilitar a otimização pela aplicação de modelos precisos e pelo uso de hipóteses sobre o hardware do *data center*. Propriedades de dependência de longo alcance (*LRD - long-range dependence*) foram aplicados na modelagem do tráfego

de rede do *data center*. Com base nos modelos criados, os autores introduziram um novo conceito de classes de tráfego, definindo cinco diferentes aplicativos, normalmente hospedados em *data centers*. Sendo assim, não foi o foco dessa pesquisa a infraestrutura elétrica de *data center*.

Os autores em (DUMITRESCU et al., 2018) propõem métodos para reduzir o consumo de energia em uma instalação de *data center*. A modelagem adotada é matemática e foi criada a partir do software *CFD* (*Computational Fluid Dynamic*). O modelo simulou o ambiente dinâmico no *data center* com a variação de carga e mudança na infraestrutura de resfriamento. Os autores fazem uso de otimização com foco na infraestrutura de resfriamento, porém nada foi mencionado em relação a DoE e infraestrutura elétrica de *data center*.

Huang et al. (HUANG et al., 2018b) propuseram uma otimização dos custos de eletricidade para serviços interativos de *data center* com foco no equipamento UPS. Os autores apresentaram um algoritmo de dois estágios para otimizar o custo de eletricidade dos serviços interativos do *data center*, incluindo tanto a carga de energia quanto a cobrança de demanda. A execução do algoritmo de dois estágios é utilizada para reduzir as cargas de trabalho de acordo com a restrição de qualidade do serviço, o UPS é usado para reduzir a potência de pico sem reduzir o desempenho. Este trabalho se restringiu a estudar o dispositivo UPS.

Uma proposta de otimização energética hierárquica das redes de *data center* foi proposta em (ZHANG; ANSARI, 2015). Os autores demonstram um modelo de otimização de energia, chamado *HERO* (*Hierarchical Energy Optimization*) para reduzir o consumo de energia de equipamentos de rede, desligando os *switches* e links de rede, garantindo a conectividade total e maximizando a utilização do link. O modelo HERO proposto para topologias da rede hierárquica de *data center* se enquadra na classe do problema do *CMCF* (*Multicommodity Capacited Minimum Cost Flow*), ou seja, é um problema da classe NP difícil. Problemas *CMCF* são problemas da classe NP difícil, portanto métodos exatos só podem ser usados para resolver casos triviais. O modelo proposto é avaliado por várias experiências com diferentes padrões de tráfego. Porém, nada foi mencionado em relação a DoE e políticas de manutenção.

Jung et al. (JUNG et al., 2016) propuseram uma técnica de otimização para implantação de sensor no *data center*. Os autores criaram um sistema de monitoramento

de temperatura sem fio para otimizar o número de sensores de temperatura implantados em um *data center*. Os autores realizaram a otimização através da análise dos dados históricos, coletados por meio dos sensores em uma rede de sensores sem fio. Os dados de temperatura obtidos em tempo real, também são usados para prever o próximo estado da temperatura do *data center* para detectar anomalias em potencial na geração de calor. Sendo assim, a infraestrutura elétrica não era, diretamente, o foco dos autores.

Os autores em (SANKARAN; SIVALINGAM, 2017) apresentaram a análise de um algoritmo de agendamento para redes de *data center* preparadas opticamente. O objetivo desse trabalho é analisar o problema do agendamento de transmissão em uma rede híbrida de *data center* por meio de pacotes ópticos. Os autores fazem uso de modelos matemáticos para simulação de dados e atraso de pacotes de controle. Fazendo uso de um algoritmo guloso para obter os resultados referentes ao desempenho do agendamento. Técnicas de otimização foram adotadas, porém nada é mencionado em relação a políticas de manutenção ou ao uso de planejamento de experimentos.

Em (LI et al., 2015) os autores fizeram uma otimização conjunta da largura de banda do provedor e atraso do usuário em *data centers* através do software Wenxin. O objetivo desse trabalho é a alocação de solicitações em *data centers* distribuídos geograficamente. Um modelo de otimização integrado foi proposto para considerar uma alta utilização da largura de banda para o provedor associado a um baixo atraso para os usuários. Isso foi obtido através do método baseado na solução de barganha de Nash (*NBS - Nash Bargaining Solution*), usado para modelar os requisitos de alta largura de banda do provedor e o baixo atraso para os usuários finais. O foco do trabalho não foi relacionado com infraestruturas elétricas de *data centers*.

Os autores em (YEGANEH et al., 2019) propuseram um método de otimização de custos para computação em nuvem pelo planejamento de capacidade do *data center* com preços dinâmicos. O objetivo desse trabalho é otimizar o custo do consumo de energia dos *data centers* ecológicos e maximizar o lucro a longo prazo do provedor de serviços de rede móvel. Os autores fizeram uso de modelos matemáticos para a proposição de um algoritmo que determina o número ideal de servidores e trata da relação entre o custo da operação e o atraso dos serviços. Nada foi mencionado em relação as políticas de manutenção ou ao uso de planejamento de experimentos.

Os autores em (CHIARAVIGLIO et al., 2018) apresentam uma abordagem para

equilibrar o custos de manutenção e consumo de energia em *data center*. Este trabalho visa o gerenciamento da energia nos servidores em um *data center* em nuvem para minimizar conjuntamente o consumo de eletricidade e os custos de manutenção derivados da variação de energia (e conseqüentemente da temperatura) na CPU dos servidores. Os autores fazem uso de um algoritmo chamado de *MECDC* (*Maintenance and Electricity Costs Data Center*), com o objetivo de realizar o equilíbrio entre o consumo de energia e o custo de manutenção. Eles consideram o custo de manutenção, porém nada foi mencionado em relação a manutenção corretiva ou manutenção preventiva.

Os autores em (HUANG et al., 2018a) propuseram uma política de manutenção em tempo real para sistemas de fabricação de vários estágios, considerando os efeitos imperfeitos de manutenção. O objetivo desse trabalho é reduzir os custos relacionados à manutenção. O custo de manutenção inclui custo de recursos e perda de produção devido à parada da máquina. Os autores fazem uso de modelos matemáticos orientado por dados do sistema de fabricação, obtendo assim, os riscos de perda permanente de produção e perda de produção incorridos pela ação de manutenção. Políticas de manutenção foram adotadas, porém nada é mencionado em relação a otimização ou ao uso de planejamento de experimentos.

Um modelo dinâmico de manutenção prescritiva considerando o envelhecimento e a degradação do sistema de locomotivas foi proposto em (LIU et al., 2019). No modelo proposto de manutenção considera-se a influência da degradação e do envelhecimento na falha do sistema e determina dinamicamente o tempo ideal de inspeção. A influência do nível de degradação e do envelhecimento na taxa de falha do sistema é modelada de maneira aditiva. Com base no nível de degradação observado na inspeção, o reparo ou a substituição são realizados no sistema. Os autores fazem uso de manutenção prescritiva, porém nada foi mencionado em relação a políticas de manutenção corretiva e/ou preventiva.

A Tabela 2 ilustra os trabalhos relacionados, ressaltando as diferenças entre o foco de cada pesquisa. Conforme pode ser observado, poucos trabalhos abordam de forma integrada a modelagem híbrida por meio do uso de SPN e RBD para otimização das infraestruturas elétricas de *data centers*, fazendo uso de técnicas de planejamento de experimentos, vinculadas a diferentes políticas de manutenção e contratos de *SLA*. Além disso, a estratégia de otimização adotada, os modelos elaborados para a manutenção preventiva e a degradação do tempo médio para ocorrência de uma falha, podem ser

destacadas como contribuições desse trabalho.

Tabela 2 – Resumo comparativo dos trabalhos relacionados.

Trabalhos Relacionados	Manutenção	SLA	Disponibilidade	Infraestrutura elétrica	Otimização	DoE	Modelagem
[Bosse et al. 2016]	Corretiva e Preventiva		X				Cadeia de Markov
[Callou et al. 2010]	Corretiva	X	X	X			RBD e SPN
[Addabbo et al. 2015]	Preventiva	X	X				
[Said and Taghipour 2016]	Corretiva e Preventiva		X				
[Wang et al. 2017]	Preventiva		X				
[Liu and Trivedi 2006]	Preventiva						
[Silva et al. 2013]			X	X			
[Silva et al. 2014]			X	X			
[Bichpuriya et al. 2011]			X	X			
[Balanici et al. 2018]			X			X	Equação Matemática
[Dumitrescu et al. 2018]			X			X	Matemática (software CFD)
[Huang et al. 2018]				X		X	
[Zhang et al. 2015]				X		X	Cadeia de Markov
[Jung et al. 2016]						X	
[SANKARAN; SIVALINGAM, 2017]						X	Equação Matemática
[LI et al., 2015]						X	
[YEGANEH et al., 2019]						X	Equação Matemática
[Chiaraviglio et al. 2018]				X			
] [Huang et al. 2018]							Equação Matemática
[Liu et al. 2019]	Prescritiva						
[Trabalho Proposto]	Corretiva e Preventiva	X	X	X	X	X	RBD e SPN

4 Metodologia

Esse capítulo apresenta a metodologia utilizada para avaliar a disponibilidade da infraestrutura elétrica de *data centers*. Essa metodologia leva em consideração a estratégia de otimização adotada, por meio do planejamento de experimentos, e das políticas de manutenção vinculadas a contratos de *SLA*. A Figura 7 mostra uma visão geral da metodologia.

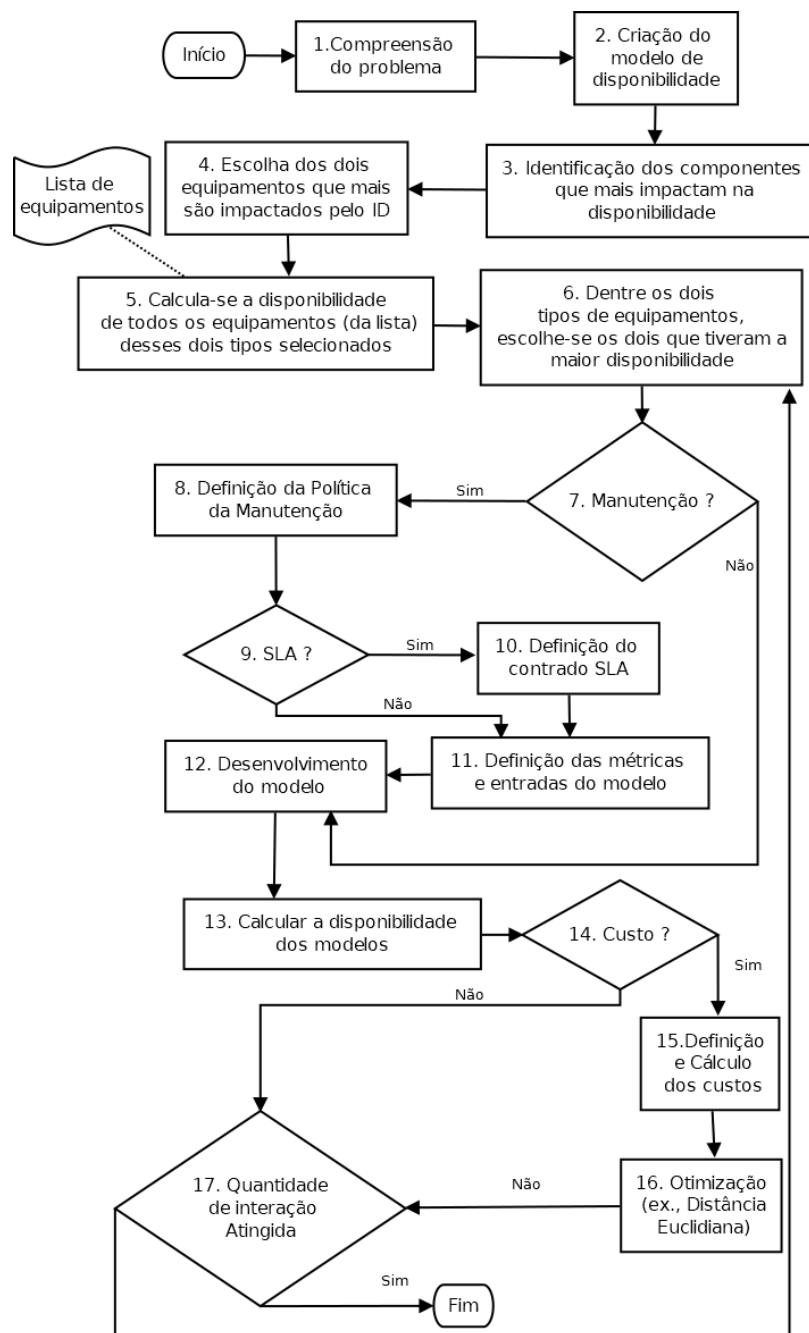


Figura 7 – Metodologia adotada

A primeira fase corresponde a “compreensão do problema”, onde é primordial entender o funcionamento do sistema real e, a partir deste, ser capaz de representá-lo por meio de modelos. A segunda etapa, “criação dos modelos de disponibilidade” realiza a criação dos modelos de disponibilidade para cada arquitetura utilizada. Nessa fase, optou-se por realizar a criação de modelos em RBD com o objetivo de quantificar o índice da importância para a disponibilidade de cada equipamento. Após a criação desses modelos, avança-se para a fase de “identificação dos componentes que mais impactam na disponibilidade”, onde é possível verificar quais componentes de determinada arquitetura em análise são mais sensíveis para a disponibilidade. Essa identificação ocorre através do cálculo de importância da disponibilidade que foi apresentado no Capítulo 2. Vale ressaltar que esta análise é realizada para cada cenário que será analisado. Sendo assim, para cada cenário analisado é criado um modelo em RBD para poder verificar qual equipamento mais impacta na disponibilidade do sistema analisado.

Em seguida, segue-se para a fase de escolha dos dois equipamentos que mais impactam na disponibilidade. A identificação acontece através do cálculo do índice de importância da disponibilidade (ID), que aponta quais equipamento são mais sensíveis. Com a relação dos equipamentos mais sensíveis, seleciona-se os dois dispositivos que mais impactam na disponibilidade. Esses dispositivos são utilizados nas fases posteriores.

Na fase de “cálculos de disponibilidade” é realizado o cálculo da disponibilidade dos dois dispositivos (selecionados a partir da fase de “identificação dos equipamentos que mais impactam a disponibilidade”), levando em consideração uma lista de equipamentos com diferentes parâmetros de entrada como, por exemplo, diferentes MTTFs. Na lista de equipamentos existe uma quantidade determinada de dispositivos para cada equipamento, cada um desses equipamentos possui informações distintas de MTTF, custo e eficiência. Vale ressaltar que nesta fase são criados modelos em SPN para representar cada dispositivo, e a partir desses modelo são obtidos os resultados da disponibilidade de cada dispositivo.

Na fase de "escolha dos dois dispositivos com maior disponibilidade", os 2 dispositivos que mais impactam a disponibilidade são avaliados para todas as opções de equipamentos disponíveis na lista. Após ter obtido a disponibilidade individual de cada um desses 2 dispositivos que compõem o sistema em análise, são escolhidos os 2 dispositivos do conjunto de N equipamentos (lista de equipamentos) que tiveram a melhor disponibilidade. Para os demais equipamentos que compõem a arquitetura em análise,

seleciona-se apenas um dispositivo da lista de equipamentos de forma aleatória.

Na fase seguinte, “escolha das políticas de manutenção”, o projetista realiza uma análise referente ao uso das políticas de manutenção. Dependendo da escolha do projetista, pode-se definir pela utilização de políticas de manutenção. Caso haja a necessidade da realização de manutenções, avança-se para a fase de “definição da política de manutenção”, são definidas as políticas de manutenção que serão adotadas. Assumindo que não serão realizadas manutenções, a metodologia avança para a fase de “desenvolvimento do modelo”, onde são realizados o desenvolvimento dos modelos.

A fase de “definição das políticas de manutenção” corresponde a definição das políticas de manutenção que serão adotadas. Por exemplo, pode-se adotar somente manutenção corretiva ou manutenção preventiva em conjunto com a corretiva, visto que esses dois tipos de política de manutenção são as mais utilizadas. Na manutenção corretiva são levantados todos os parâmetros e características deste tipo de manutenção, tais como: MTTF e MTTR de cada equipamento. Assumindo a manutenção preventiva e corretiva, considera-se todos os aspectos citados para a manutenção corretiva com o acréscimo de parâmetros como a frequência em que serão realizadas as manutenções preventivas e o tempo que levam para serem realizadas.

Na fase de “escolha de *SLA*”, o projetista verifica a necessidade da utilização de contratos de *SLA*. Caso seja necessário a utilização de contratos *SLA*, avança-se para a fase de “definição dos contratos de *SLA*”. Caso contrário, avança-se para a fase de “definição de métricas”, onde ocorre a definição das métricas de entrada para o modelo de manutenção.

Na fase de “definição dos contratos de *SLA*”, onde são definidos os contratos de *SLA* para cada tipo de manutenção (corretiva e preventiva/corretiva) são estabelecidas as condições necessárias para cada nível do contrato. Os contratos *SLAs* para a manutenção corretiva considera as métricas: (i) custo de equipamento; (ii) custo da manutenção corretiva; (iii) custo com o time de manutenção. Para a manutenção preventiva, além dos custos citados, deve-se considerar também o custo extra da realização de uma manutenção periódica. Tendo em vista as métricas apresentadas, podem ser criados os seguintes contratos: *SLA 1*, *SLA 2*, *SLA 3* e *SLA 4* para a manutenção corretiva com o tempo de reparo de 120h, 48h, 24h e 8h, respectivamente. Já para a manutenção preventiva, além dos tempos utilizados, faz-se uso de um tempo extra, que corresponde a frequência da manutenção preventiva. Essa frequência pode ser realizada a cada dois anos (17520 horas),

um ano e meio (13140 horas), um ano (8760 horas) e seis meses (4380 horas), por exemplo. Vale ressaltar que no Capítulo 6 são definidos os custos e os contratos estabelecidos de forma mais detalhada.

A próxima fase corresponde a “definição de métricas”. Essa etapa está relacionada com a definição das métricas e dados de entrada para o modelo. Os dados de entrada podem ser obtidos tanto experimentalmente como a partir dos dados fornecidos por manuais de fabricantes. Os dados de entrada correspondem ao MTTF, MTTR, quantidade de equipes de manutenção. Vale ressaltar que caso tenha sido considerada a utilização de diferentes contratos de *SLA*, uma nova métrica de entrada correspondente ao custo e a frequência da realização das manutenções que devem ser analisadas para cada contrato *SLA*.

Uma vez definida a política de manutenção, segue-se com a criação do modelo de disponibilidade levando em consideração a política escolhida (ex., preventiva, corretiva). Nessa fase de “desenvolvimento do modelo”, é necessário escolher a ferramenta de modelagem (ex. Mercury). Além disso, também é definido o tipo de modelo que será utilizado para representar o sistema sendo modelado. Por exemplo, pode-se adotar o formalismo de redes de Petri estocástica para representar o sistema modelado que leva em consideração as políticas de manutenção estabelecidas. Os modelos em SPN são criados com intuito de serem aplicados em modelos que tenham dependência, como, por exemplo, os modelos de manutenção.

Uma vez criado os modelos, deve-se partir para a fase de cálculo da disponibilidade para cada modelo criado na fase anterior. Esses cálculos ocorrem através da interação da aplicação desenvolvida em Java com a ferramenta Mercury. Essa ferramenta foi utilizada para realizar os cálculos da disponibilidade dos equipamentos. A geração dos resultados é realizada de forma automática, obtendo a disponibilidade dos equipamentos selecionados previamente. Os resultados são armazenados em arquivos, apresentando o resultado da disponibilidade de cada conjunto de equipamentos.

Na fase de “escolha da métrica de custo”, nessa etapa existe a opção de se levar em consideração a métrica de custo. Caso o custo venha a ser analisado, avança-se para a fase de “definição e cálculo do custo”. É nesse momento que são definidos os custos que serão levados em consideração, por exemplo, custo de aquisição e custo operacional. Caso o custo não seja necessário, o foco da análise que será realizada, segue-se a metodologia para o passo onde se verifica se o número de iterações desejadas foram executadas. Em

caso afirmativo, termina-se a avaliação.

Na fase de “definição e cálculo do custo”, é feita a definição para os cálculos dos custos. Esta fase leva em consideração todos os parâmetros necessários para a realização dos cálculos dos custos. Essa etapa define também as equações utilizadas para os cálculos dos custos, por exemplo, da manutenção corretiva e o custo da manutenção preventiva.

A fase “otimização” adota técnicas de otimização. Por exemplo, pode-se adotar a distância euclidiana para se detectar a melhor relação custo-benefício em que se deseja reduzir tanto o *downtime* quanto o custo. Esta técnica é utilizada para relacionar as métricas conflitantes, disponibilidade e custo, encontrando assim a solução que mais se aproxima da solução que apresente a menor distância Euclidiana. Nesse exemplo, busca-se um resultado que possibilite o menor custo e downtime.

A fase seguinte da metodologia testa se a quantidade de iterações definida pelo projetista foi atingida. Um maior número de iterações pode fazer com que a estratégia proposta se aproxime mais do resultado ótimo. No entanto, isso irá acarretar em uma maior demanda por recursos computacionais. Caso a quantidade de iterações não tenha sido atingida, o sistema voltará para a fase “escolha dos dois dispositivos com maior disponibilidade”. Em seguida, todo o processo será repetido até que seja atingido o número de iterações.

Os modelos utilizados para o cálculo da importância para a disponibilidade são desenvolvidos em RBD, pois não há dependências entre os componentes. Os demais modelos e cálculos desenvolvidos são em SPN, devido ao fato de existir dependência no caso da utilização de manutenção. Sendo assim é utilizado o RBD para identificar os componentes que são mais impactados pela disponibilidade. Já o SPN é utilizado para realizar o cálculo da disponibilidade de cada equipamento. Como também, para a construção do modelo onde ocorre a combinação de todos os componentes previamente escolhidos.

4.1 Ferramenta

Foi realizado o desenvolvimento de uma ferramenta em Java ([BITBUCKET, 2008](#)) que possibilitou a automação de partes dos experimentos realizados. A ferramenta possibilita a contabilização da disponibilidade dos modelos analisados.

É importante destacar que a ferramenta desenvolvida é responsável pelo cálculo da disponibilidade a partir de uma lista de equipamentos pré-estabelecidos (quinta fase da metodologia). Sendo responsável também pela seleção dos dois equipamentos que apresentam a maior disponibilidade dentre os dispositivos analisados (sexta fase da metodologia). A aplicação também realiza o cálculo da disponibilidade dos modelos (décima terceira fase da metodologia), além de realizar as iterações solicitadas pelo projetista, e da geração dos resultados em arquivos.

Os passos de cálculo do ID não estão presentes na ferramenta, como também os cálculos dos custos dos contratos e distância Euclidiana. Esses passos estão sendo realizados de forma não automatizada. Deve-se ficar claro que todos os cálculos obtidos na ferramenta desenvolvida são realizados por meio da integração dos scripts desenvolvidos em Java com a ferramenta Mercury. Através dessa integração, é possível realizar os cálculos da disponibilidade dos modelos.

A geração dos resultados é realizada de forma automática, obtendo a disponibilidade dos equipamentos selecionados previamente. Os resultados são armazenados em arquivos, apresentando o resultado da disponibilidade de cada conjunto de equipamentos.

5 Modelos

Neste capítulo serão apresentados os modelos propostos (modelos para avaliação da disponibilidade, modelos para a avaliação do índice importância para a disponibilidade e por fim, modelos de manutenção).

5.1 Modelo Base de Disponibilidade

A Figura 8 ilustra o modelo base de disponibilidade em SPN para representar um sistema com um único componente. Neste modelo, o equipamento pode adotar dois possíveis estados: em funcionamento ou falho. Como pode ser observado, existem duas transições que representam respectivamente o tempo médio para a falha e reparo deste equipamento. Neste caso, a transição MTTF representa o tempo médio para o equipamento UPS falhar, já a transição MTTR representa o tempo médio para reparo do equipamento. Um token no lugar UPS_ON ou UPS_OFF representam, respectivamente, o funcionamento ou falha de um equipamento, no caso, de um UPS. A disponibilidade pode ser computada ao calcular a probabilidade do sistema se encontrar em funcionamento, ou seja, a probabilidade de haver token no lugar UPS_ON ($P\{\#UPS_ON > 0\}$) (ZIMMERMANN et al., 2006).

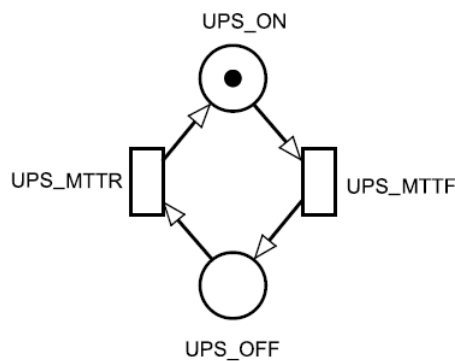


Figura 8 – Modelo SPN Base de Disponibilidade

5.2 Modelo Base em RBD

A Figura 9 apresenta um modelo RBD básico adotado para computar a disponibilidade da infraestrutura elétrica do *data center*. Essa infraestrutura é composta

pelos seguintes equipamentos: UPS, SDT, Subpanel e *Powerstrip*. Conforme pode ser observado, o modelo encontra-se em série, ou seja, a falha de qualquer componente resulta em uma falha do sistema. Sendo assim, a disponibilidade é computada pela multiplicação da disponibilidade individual de cada equipamento que compõem a arquitetura, ou seja, a disponibilidade do $UPS \times SDT \times Subpanel \times Powerstrip$.



Figura 9 – Modelo RBD Base de Disponibilidade

Já a Figura 10 apresenta uma variação, onde essa nova infraestrutura é composta dos seguintes equipamentos: UPS, UPS2, SDT, Subpanel e *Powerstrip*. Conforme pode ser observado, o modelo agora apresenta um componente, no caso o UPS, com redundância. Logo, o sistema funciona se pelo menos um componente dos que se encontram redundantes funcionar e todos os demais componentes do modelo também estiverem em funcionamento. Em outras palavras, para ocorrer a falha do sistema é necessário que os componentes UPS e UPS2 falhem, ou ocorra a falha de qualquer um dos demais componentes (ex., STS, SDT, Subpanel, *powerstrip*) do modelo que se encontram conectados em série. Sendo assim, a disponibilidade pode ser computada através da equação $RS = (1 - ((1 - disp\ UPS) \times (1 - disp\ UPS2))) \times disp\ SDT \times disp\ Subpanel \times disp\ Powerstrip$, onde disp corresponde a disponibilidade dos dispositivos.

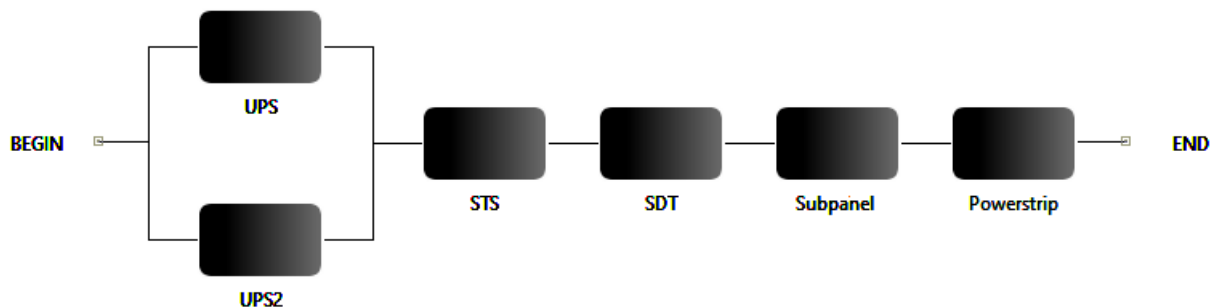


Figura 10 – Modelo Modificado RBD de disponibilidade

5.3 Modelo de Manutenção Corretiva

A Figura 11 ilustra o modelo base, onde um time de manutenção corretiva é alocado para realizar a manutenção. Um token no lugar `repararUPS` representa o momento em que o equipamento se encontra em falha e foi submetido a uma manutenção corretiva. Entretanto, para que este lugar seja alcançado, faz-se necessário que uma equipe de manutenção esteja disponível (um token no lugar `equipeManutencao`). Pode-se observar que existem três transições no modelo: `MTTR`, `MTTF` e `acionarManutencao`. As duas primeiras foram explicadas anteriormente no modelo base. A transição imediata `acionarManutencaoUPS` é disparada assim que o dispositivo falha (disparo da transição `MTTF`) e a equipe de manutenção está disponível para iniciar o reparo. Vale ressaltar que o número de tokens no lugar `equipeManutencao` representa a quantidade de times de manutenção disponíveis para serem alocados. Assumindo uma única equipe de manutenção e se estiverem presentes vários equipamentos no sistema, a equipe de manutenção só vai poder ser acionada para fazer um único reparo por vez.

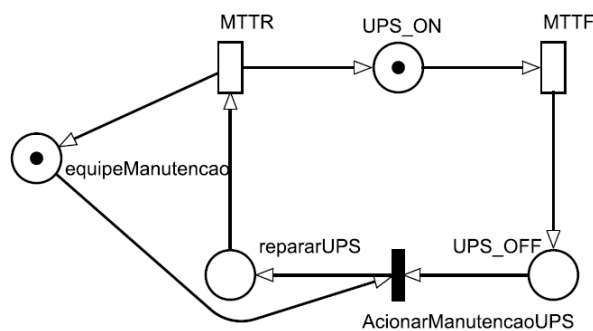


Figura 11 – Modelo de Manutenção Corretiva

O modelo de manutenção corretiva contempla ainda a modelagem de um *clock* (ver Figura 12) para representar o passar do tempo durante o período de avaliação, por exemplo 5 anos, e para poder representar a degradação dos equipamentos que passam a falhar com uma maior frequência. Um token no lugar `primeiroAno`, `segundoAno`, `terceiroAno`, `quartoAno` ou `quintoAno`, representa a simulação do tempo decorrido de um ano, dois, três, quatro ou cinco, respectivamente. As transições estocásticas apresentadas nesse modelo possuem um tempo de um ano (8760 horas). Este *clock* é utilizado para representar o desgaste que o equipamento terá, ao passar dos anos, onde se tem uma degradação de 10%, 20%, 30%, 40% e 50%, respectivamente no tempo entre falhas. Em outras palavras, com o passar do tempo, os equipamentos tendem a falhar com uma maior frequência.

Conforme ilustra a Figura 12, o token estará no lugar correspondente a simulação do tempo decorrido. Dependendo de onde o token esteja, a degradação pode aumentar até atingir o máximo de 50%. Esse comportamento é representado pela fórmula do *delay* que é associada às transições de falha do modelo. Por exemplo, a expressão associada a transição MTTF é $MTTF_UPS - (MTTF_UPS \times 0.1 \times \#primeiroAno) - (MTTF_UPS \times 0.2 \times \#segundoAno) - (MTTF_UPS \times 0.3 \times \#terceiroAno) - (MTTF_UPS \times 0.4 \times \#quartoAno) - (MTTF_UPS \times 0.5 \times \#quintoAno)$. Pode-se observar que se o *clock* estiver no lugar primeiroAno, o MTTF do UPS será reduzido em 10%.

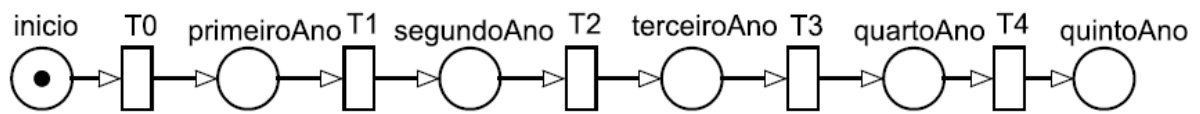


Figura 12 – *clock* do Modelo de Manutenção Corretiva

5.4 Modelo de Manutenção Preventiva

A Figura 13 representa o modelo base de manutenção corretiva acrescido da possibilidade da realização de manutenção preventiva. Neste modelo existe a adição de um novo lugar, REPARAR_UPS_PREV, para representar o momento da ocorrência da manutenção preventiva. A transição INICIAR_REPARO_UPS simula a periodicidade para a realização da manutenção preventiva, por exemplo, 8760 horas representa um ano. A transição TERMINAR_REPARO_UPS representa o tempo necessário para a realização da manutenção preventiva no dispositivo, 4 horas, por exemplo. A grande diferença representada aqui neste modelo da manutenção preventiva para a corretiva é que o MTTF do dispositivo ao passar por uma manutenção preventiva é reparado para o valor do MTTF do equipamento novo.

A Figura 14 representa o *clock* da manutenção preventiva. Observa-se a existência das transições imediatas T6, T7, T8, T9 e T10 que representam a volta ao estado inicial, ou seja, quando ocorrer uma manutenção preventiva a degradação passa a ser zero (0), indicando que o dispositivo volta ao estado de novo. As transições T6, T7, T8, T9 e T10 serão ativadas caso a condição de guarda seja satisfeita. Para o exemplo da Figura 15, todas as transições imediatas do modelo da Figura 14 têm a seguinte expressão de guarda $((\#REPARAR_UPS_PREV > 0) \text{OR} (\#REPARAR_UPS_PREV2 > 0))$.

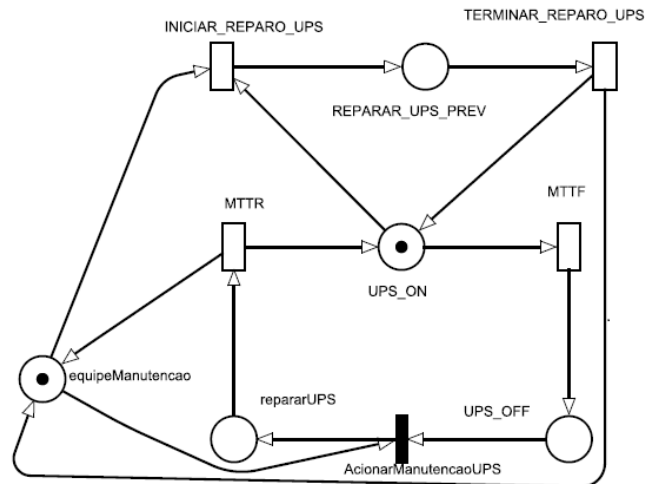


Figura 13 – Modelo de Manutenção Preventiva

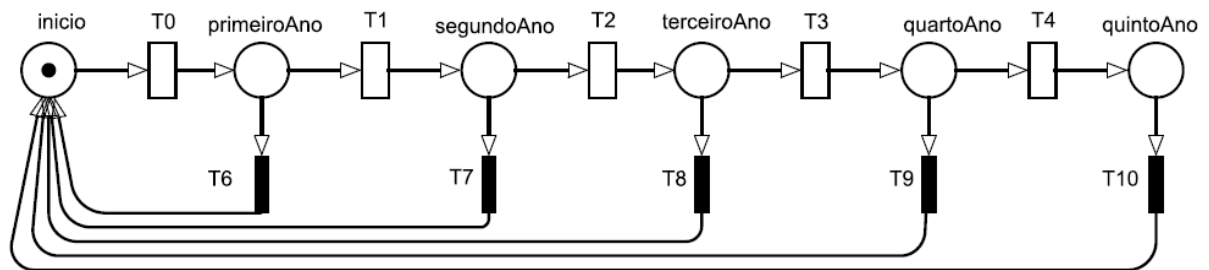


Figura 14 – *clock* do Modelo de Manutenção Preventiva

Vale salientar que se temos dois dispositivos redundantes, só faz sentido realizar a manutenção preventiva em cada um dos equipamentos por vez. Dessa forma, a Figura 15 mostra dois dispositivos UPS redundantes, com o acréscimo de um artifício para não permitir que os dispositivos UPS_ON e UPS_ON2 entrem em manutenção preventiva ao mesmo tempo. Esse artifício, trata-se de uma ligação entre UPS_ON e a transição INICIAR_REPARO_UPS2, e a ligação entre UPS_ON2 e a transição INICIAR_REPARO_UPS.

5.4.0.1 Modelo de Custo

Esta seção apresenta as equações utilizadas para obter o custo da manutenção corretiva e o custo da manutenção preventiva. A Equação 5.1 corresponde ao cálculo adotado para computar o custo da manutenção corretiva. Esse custo corresponde ao preço da hora da manutenção corretiva multiplicada pelo tempo que foi demandado para se realizar tal operação, no caso, esse tempo corresponde ao *downtime* do sistema modelado.

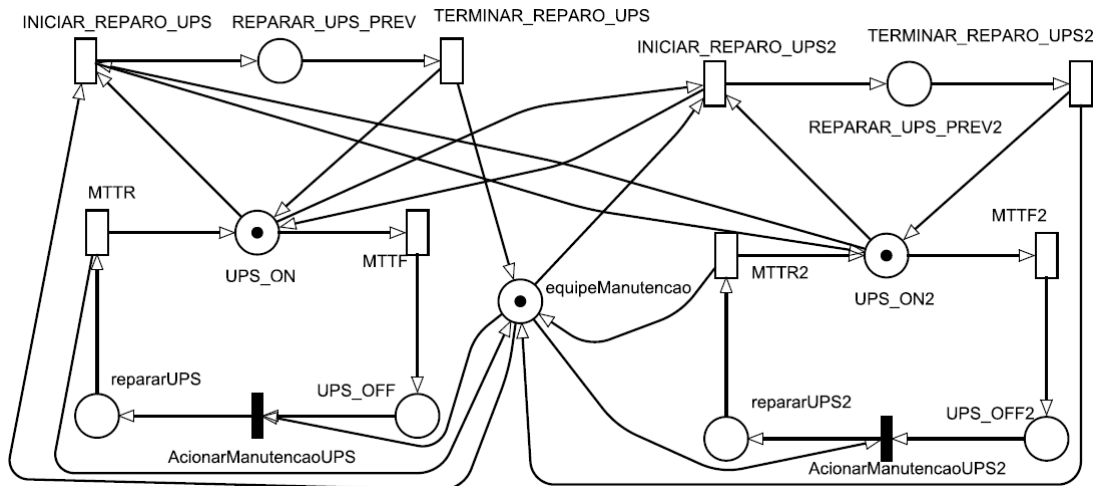


Figura 15 – Modelo de Manutenção Preventiva Modificado

Observe que cada equipamento pode ter um preço distinto para a manutenção corretiva.

$$CMC = \sum_{i=1}^n P_i \times T_i \quad (5.1)$$

onde P representa o preço da manutenção por hora i corresponde ao equipamento, n é o total de equipamentos no sistema em análise, e T representa o tempo. Esta equação representa o custo para a realização da manutenção corretiva, onde temos o somatório dos equipamentos que compõem cada infraestrutura.

O custo total da manutenção corretiva (Equação 5.2) corresponde ao custo demandado para se ter um time de manutenção somado ao custo do reparo (Equação 5.1).

$$CTMC = \sum_{i=1}^n TM_i + CMC \quad (5.2)$$

onde TM_i representa o custo do time de manutenção i e CMC representa o custo da manutenção corretiva. Esta equação representa o custo total para realização da manutenção corretiva.

O custo da manutenção preventiva, isoladamente, corresponde ao que é computado através da Equação 5.3. Esse custo corresponde ao preço da manutenção preventiva específica de cada equipamento por hora, multiplicado pelo tempo da realização do serviço e, também, pela frequência em que foi realizado no período de análise.

$$CMP = \sum_{i=1}^n P_i \times T_i \times F_i \quad (5.3)$$

onde P representa o preço da manutenção preventiva por hora de cada equipamento i , T é o tempo para a realização da manutenção e F a frequência da realização da manutenção.

Por fim, tem-se o custo total da manutenção preventiva, o qual corresponde ao custo do time de manutenção em conjunto com os custos da manutenção preventiva e da manutenção corretiva conforme pode ser observado na Equação 5.4.

$$CTMP = \sum_{i=1}^n TM_i + CMC + CMP \quad (5.4)$$

onde TM_i representa o custo do time de manutenção i ; CMC representa o custo da manutenção corretiva; e CMP representa o custo da manutenção preventiva.

6 Estudos de Caso

Este capítulo apresenta três estudos de caso. O primeiro estudo tem como objetivo apresentar a estratégia de otimização proposta para maximizar a disponibilidade da infraestrutura elétrica de um *data center*. O segundo estudo de caso tem como objetivo além de otimizar a disponibilidade da infraestrutura elétrica de um *data center* levando em consideração políticas de manutenção corretiva e preventiva. O terceiro estudo de caso faz uso da estratégia de otimização proposta, levando em consideração políticas de manutenção vinculadas a diferentes contratos de *SLA*.

6.1 Arquiteturas

A Figura 16 apresenta as cinco arquiteturas das infraestruturas elétricas de *data center* adotadas nesses estudos (AVELAR, 2003). Essas arquiteturas são compostas por: fontes de alimentação sem interrupção (UPS), transformador (SDT), chaveador de transferência estática (STS), quadro de energia (subpanel) e régua de energia elétrica (*powerstrip*).

A Figura 16(a) ilustra a arquitetura A1, esta não tem nenhuma redundância, sendo composta por UPS, SDT, SubPanel e Powerstrip. A Figura 16(b) mostra a arquitetura A2, que é semelhante a arquitetura A1 com a adição de uma replicação para o UPS. A Figura 16(c) ilustra a arquitetura A3, com os mesmos componentes presentes na arquitetura A1, com a diferença na replicação do UPS e SDT, como também a adição de uma chave de transferência estática (STS). A Figura 16(d) apresenta a arquitetura A4, onde temos os mesmos componentes presentes na arquitetura A1, com SubPanel adicional. A Figura 16(e) ilustra a arquitetura A5, que corresponde a arquitetura A1 com todos os equipamentos replicados. Vale ressaltar que essa arquitetura não tem a necessidade do equipamento STS.

6.2 Estudo de caso I

Esse estudo apresenta a aplicabilidade da estratégia de otimização proposta, que leva em consideração técnicas de planejamento de experimentos para otimizar a disponibilidade da infraestrutura elétrica de um *data center*. As soluções obtidas com a técnica proposta e

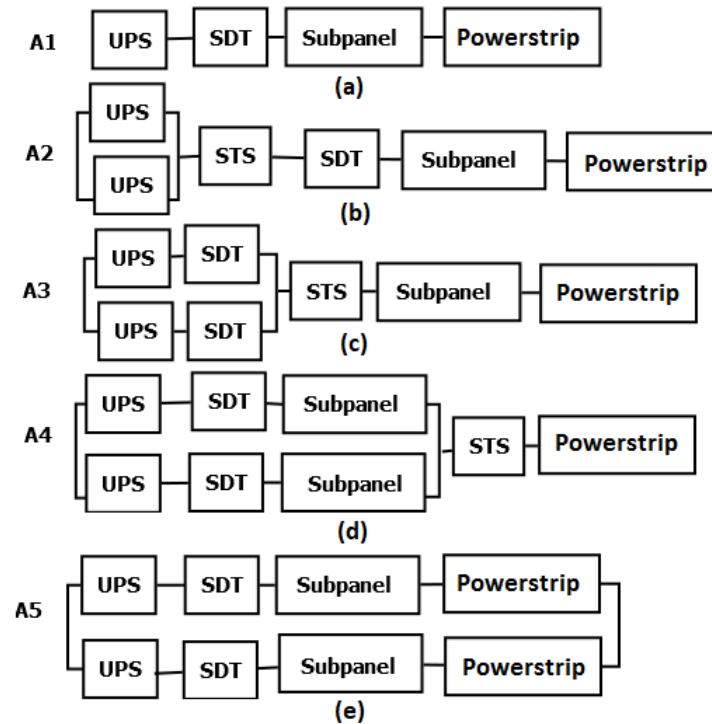


Figura 16 – Arquiteturas

o tempo de execução dessa estratégia são comparados com o algoritmo de força bruta, que obtém a solução ótima. As arquiteturas de *data center* A1 a A5 previamente detalhadas foram avaliadas nesse estudo. Sendo assim, esse estudo tem como objetivo apresentar a estratégia de otimização proposta com o intuito de maximizar a disponibilidade da infraestrutura elétrica de um *data center*. Além disso, uma comparação dos resultados do algoritmo de força bruta com a estratégia de otimização adotada é apresentada para mostrar a viabilidade da solução proposta.

6.2.1 Modelos

Nesta seção são apresentados os modelos propostos, por exemplo, modelos em RBD, usados no cálculo do índice para importância da disponibilidade, modelos em SPN utilizados para obtenção da disponibilidade. Para todas as arquiteturas (A1, ..., A5) foram analisadas a disponibilidade dos equipamentos e o índice da importância para a disponibilidade.

6.2.1.1 Modelo RBD

A Figura 17 representa o modelo RBD para a arquitetura A2, observa-se que existe redundância do equipamento UPS, e os demais equipamentos estão ligados em série. Através deste modelo é possível realizar o cálculo do índice da importância para a disponibilidade. Esse cálculo identifica quais equipamentos mais impactam na disponibilidade de todo o sistema. Como parâmetros de entrada para o modelo, em cada dispositivo, são informados os tempos médios de falha e de reparo (Ver Tabela 4).

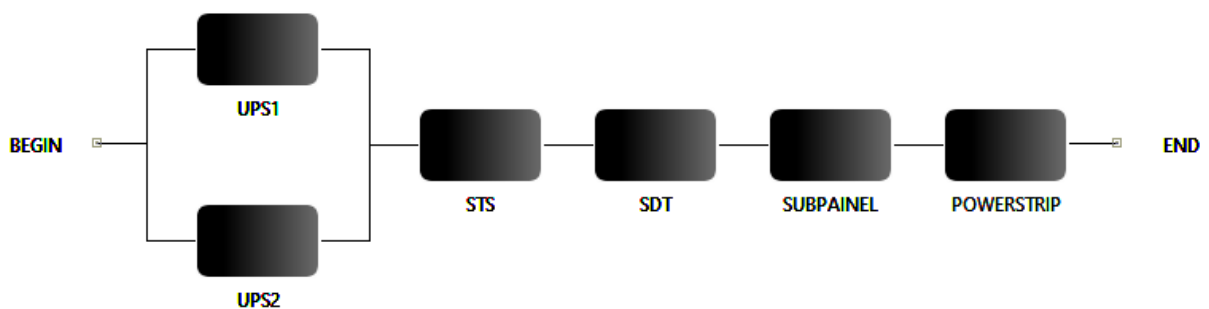


Figura 17 – Modelo RBD da Arquitetura A2

6.2.1.2 Modelo SPN

A Figura 18 apresenta o modelo SPN proposto para representar a disponibilidade da arquitetura A5. Esse modelo representa o comportamento dos equipamentos que podem estar em funcionamento ou em estado de falha. Para calcular sua disponibilidade, o tempo médio de falha e o tempo médio de reparo são os únicos parâmetros necessários (ver Tabela 4). A disponibilidade é computada através da probabilidade de $P\{((\#UPS_ON = 1) \text{ AND } (\#SDT_ON = 1) \text{ AND } (\#SubPainel_ON = 1) \text{ AND } (\#PowerStrip_ON = 1)) \text{ OR } ((\#UPS_ON2 = 1) \text{ AND } (\#SDT_ON2 = 1) \text{ AND } (\#SubPainel_ON2 = 1) \text{ AND } (\#PowerStrip_ON2 = 1))\}$. A Tabela 3 apresenta as fórmulas adotadas para o cálculo da disponibilidade.

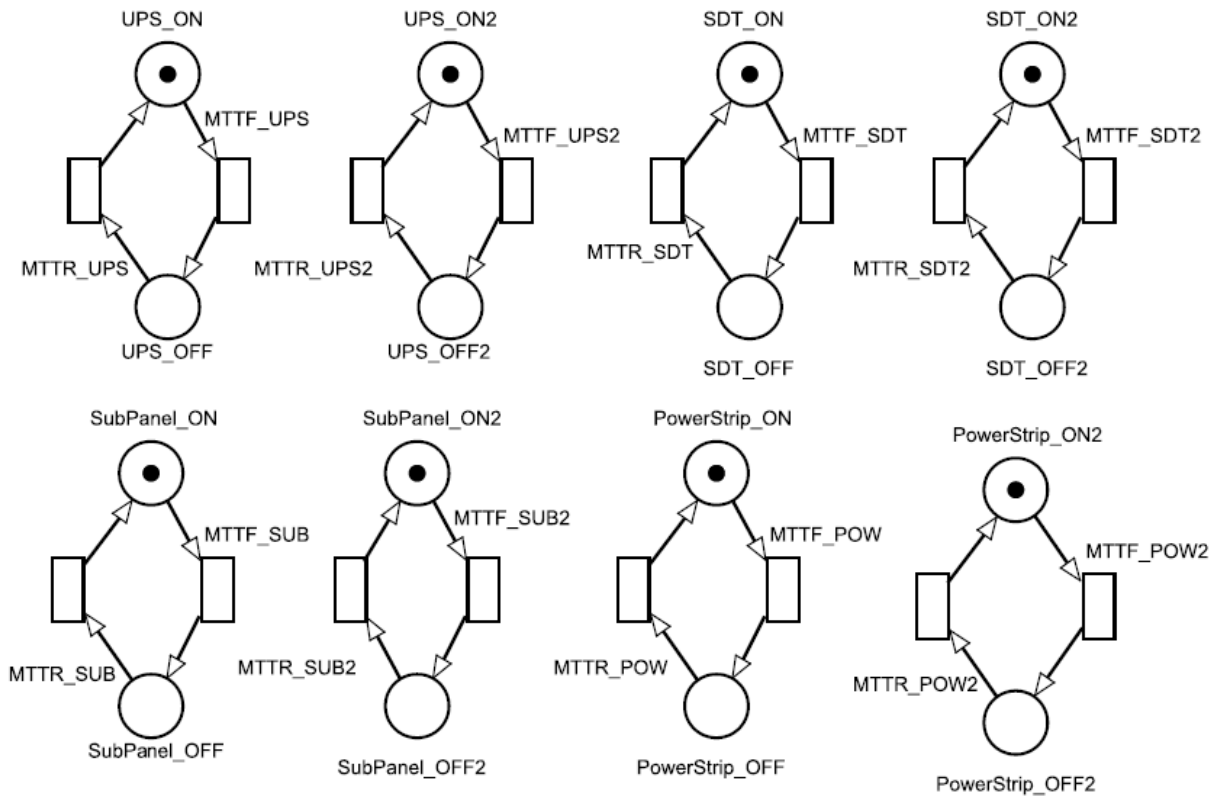


Figura 18 – Modelo SPN da Arquitetura A5

Tabela 3 – Fórmulas adotadas para o cálculo da Disponibilidade

Arquiteturas	Fórmula para o cálculo da Disponibilidade
A1	$(\#UPS_ON==1) \text{ AND } (\#SDT_ON==1) \text{ AND } (\#SubPainel_ON==1) \text{ AND } (\#PowerStrip_ON==1)$
A2	$((\#UPS_ON==1) \text{ OR } (\#UPS_ON2==1)) \text{ AND } (\#SDT_ON==1) \text{ AND } (\#SubPainel_ON==1) \text{ AND } (\#PowerStrip_ON==1) \text{ AND } (\#STS_ON==1)$
A3	$((\#UPS_ON==1) \text{ AND } (\#SDT_ON==1)) \text{ OR } ((\#UPS_ON2==1) \text{ AND } (\#SDT_ON2==1)) \text{ AND } (\#SubPainel_ON==1) \text{ AND } (\#PowerStrip_ON==1) \text{ AND } (\#STS_ON==1)$
A4	$((\#UPS_ON==1) \text{ AND } (\#SDT_ON==1) \text{ AND } (\#SubPainel_ON==1)) \text{ OR } ((\#UPS_ON2==1) \text{ AND } (\#SDT_ON2==1) \text{ AND } (\#SubPainel_ON2==1)) \text{ AND } (\#PowerStrip_ON==1) \text{ AND } (\#STS_ON==1)$
A5	$((\#UPS_ON==1) \text{ AND } (\#SDT_ON==1) \text{ AND } (\#SubPainel_ON==1) \text{ AND } (\#PowerStrip_ON==1)) \text{ OR } ((\#UPS_ON2==1) \text{ AND } (\#SDT_ON2==1) \text{ AND } (\#SubPainel_ON2==1) \text{ AND } (\#PowerStrip_ON2==1))$

6.2.2 Resultados

Para cada dispositivo presente nas arquiteturas analisadas foi adotada uma lista com 10 equipamentos, com MTTF, custo e eficiência distintos. Dessa forma, a estratégia de otimização adotada tem como objetivo de encontrar, entre as combinações destes equipamentos, a combinação ótima ou a que mais se aproxime da solução ótima, com menor tempo de execução. A Tabela 4 apresenta a faixa de valores adotada para o MTTF, custo de aquisição e eficiência para a criação da lista de componentes candidatos. A lista é criada baseada nos componentes da arquitetura e através de um sorteio randômico, para a definição dos parâmetros do dispositivo, a partir dos intervalos mostrados na Tabela 4. Por exemplo, no caso do dispositivo UPS, o seu MTTF terá um valor entre 25000 e 75000, o seu custo de aquisição entre 43558,31 e 72597,19 e, por fim, apresentará uma eficiência entre 90,535 e 99,9. Na Tabela 4, AC corresponde ao custo de aquisição, Eff é a eficiência do dispositivo da arquitetura.

Tabela 4 – Valores de intervalo de referência para os dispositivos

Equipamento	MTTF(hs)	AC(R\$)	Eff(%)
UPS	[25000, 75000]	[43558,31, 72597,19]	[90,535, 99,9]
STS	[24038, 72114]	[2323,11, 3871,85]	[94,525, 99,9]
Subpanel	[152000, 456000]	[580,78, 967,96]	[94,905, 99,9]
SDT	[141290,5, 423871,5]	[1597,14, 2661,90]	[93,575, 99,9]
<i>Powerstrip</i>	[199576,756, 313455,756]	[681,45, 898,27]	[86,575, 97,575]

A ferramenta utilizada para a modelagem foi o Mercury (SILVA et al., 2015), que possibilita a avaliação de modelos de confiabilidade, desempenho e fluxo de energia. Destaco que foi criada uma aplicação (BITBUCKET, 2008), na linguagem Java, para realizar a otimização. A aplicação acessa a lista de dispositivos, apresentada anteriormente, e se comunica com a ferramenta Mercury para realizar o cálculo da disponibilidade de cada equipamento.

Todo o histórico de equipamentos escolhidos e combinações realizadas é salvo em arquivos .TXT. Vale ressaltar que o MTTF utilizado para realizar o cálculo da disponibilidade, faz uso dos valores obtidos na lista de equipamentos, gerado a partir da faixa de valores apresentados na Tabela 3. Já para o MTTR foi adotado um valor

constante para todos os equipamentos, neste caso foi adotado o MTTR de oito horas. O experimento foi realizado no notebook com as seguintes configurações: processador intel core I7-7500U sétima geração, HD de 1 Terabyte, 8 GB de memória RAM, placa de vídeo de 2 GB.

Com a estratégia adotada, ao invés de verificar todas as possíveis combinações, foram verificados um número reduzido de cenários. Vale ressaltar que as arquiteturas A2, A3 e A4 têm a necessidade de um equipamento STS. Isso não ocorre nas arquiteturas A1 e A5, visto que não necessitam fazer nenhuma interrupção de transferência estática. Levando em consideração a arquitetura A1, e a quantidade de equipamentos disponíveis na lista para cada componente da arquitetura, tem-se 10000 combinações (10 UPS x 10 STS x 10 Subpanel x 10 PS) possíveis a serem avaliadas. Esse número de combinações para cada cenário a ser avaliado cresce de acordo com a quantidade de equipamentos na arquitetura analisada. No entanto, com a estratégia adotada temos a verificação de apenas 4 cenários, os quais serão mais detalhados a seguir. Caso seja preciso melhorar os resultados obtidos a partir desses 4 cenários, existe a possibilidade de repetir este processo quantas vezes forem necessárias, visto que a ferramenta desenvolvida permite que o projetista indique a quantidade de replicações que devem ser realizadas. Para cada arquitetura, o processo foi repetido cinco vezes e se escolheu dentre os vinte resultados que apresentou a melhor disponibilidade.

Para cada arquitetura apresentada anteriormente, foi realizado o cálculo da importância para a disponibilidade com o objetivo de identificar os dois equipamentos mais sensíveis. A Tabela 5 apresenta os valores obtidos para o índice de importância para a disponibilidade onde, em negrito, são destacados os equipamentos que mais impactam na disponibilidade da arquitetura. O passo seguinte da metodologia é o de computar a disponibilidade, individualmente, destes dois equipamentos selecionados pelo ID, levando em consideração a lista de equipamentos com N componentes (no caso, o N foi 10). Desta lista são escolhidos os dois dispositivos que apresentam maior disponibilidade, os demais equipamentos da arquitetura são selecionados um a um, de forma aleatória, a partir da lista de equipamentos e de acordo com a arquitetura analisada. Dessa forma, teremos 4 combinações de cenários a serem analisados, sendo 2 opções de equipamentos para cada um dos 2 dispositivos com maior índice ID, e o restante da arquitetura será composta por uma única opção de dispositivo para cada equipamento através de uma escolha aleatória.

Dessa forma, temos um número bem reduzido de combinações de cenários, bem diferente da alternativa sem aplicar a técnica de otimização proposta, onde teríamos 10000 cenários a serem analisados.

Tabela 5 – Equipamentos mais sensíveis para a disponibilidade nas Arquiteturas de A1 até A5

Arquitetura	UPS	SDT	Subpainel	<i>Powerstrip</i>	STS
A1	1	0,9998865297	0,9998851571	0,9998767245	–
A2	1,57E-04	1	0,9999986273	0,9999901937	0,9999570246
A3	2,00E-04	2,00E-04	1	0,9999915665	0,9999583973
A4	2,42E-04	2,42E-04	2,42E-04	1	0,9999668305
A5	1	0,9998865297	0,9998851571	0,9998767245	–

Analisando o cenário da arquitetura A1, o ID indicou que os dois equipamentos mais importantes foram o UPS e SDT. Dessa forma, foi realizado o cálculo da disponibilidade de cada um dos dez equipamentos (da lista) e escolheu-se os dois dispositivos que apresentaram a maior disponibilidade para cada equipamento analisado. Sendo assim, escolhem-se dois dispositivo para o UPS e dois dispositivos para o SDT. Após isso, foi realizada a combinação destes equipamentos com um subpainel e um *powerstrip*, estes sendo escolhidos de forma aleatória da lista de 10 componentes fornecida. Este mesmo processo foi replicado para as demais arquiteturas.

Com a estratégia adotada ao invés de verificar todas as possíveis combinações verificam-se poucos cenários, ou seja, se fossem avaliados todos os possíveis cenários para as arquitetura A1 e A5 haveriam as seguintes combinações 10x10x10x10, logo, existiriam 10000 possibilidades possíveis a serem verificadas. Já para as arquiteturas A2, A3 e A4 tem-se a seguinte combinação 10x10x10x10x10, sendo assim existem 100000 possibilidades a serem analisadas. Com a estratégia adotada para A1 teríamos 2x2x1x1, apenas 4 cenários com a possibilidade de repetir este processo quantas vezes fossem necessárias, visto que a aplicação permite que o projetista indique a quantidade de replicações da estratégia de otimização que devem ser realizadas para cada arquitetura.

A Tabela 6 apresenta os resultados obtidos ao realizar a otimização das cinco arquiteturas, sendo possível verificar a disponibilidade, tempo de execução e, também a comparação em relação ao algoritmo que testa todas as combinações possíveis (força bruta).

Tabela 6 – Avaliação Comparativa dos Resultados do Algoritmo de Força vs a Estratégia de Otimização Proposta.

Arq	Disponibilidade FB (9's) [IC]	Disponibilidade Otimizada (9's) [IC]	Erro	Tempo Exec FB	Tempo Exec Otimizada
A1	0,9998256821 (3,75) [0,9997761; 0,999797]	0,9998256821 (3,75) [0,9997942; 0,9998066]	0,00%	2 minutos 12 segundos	5 segundos
A2	0,9998093175 (3,71) [0,9996802; 0,999726]	0,9997816687 (3,66) [0,9997020; 0,9997353]	1,58%	47 minutos 49 segundos	5 segundos
A3	0,9998284835 (3,76) [0,9997157; 0,999759]	0,9997830982 (3,66) [0,9997541; 0,9997755]	0,004%	60 minutos 42 segundos	6 segundos
A4	0,9998485231 (3,82) [0,9997454; 0,999788]	0,9998484833 (3,82) [0,9998392; 0,9998440]	0,000004%	112 minutos 51 segundos	6 segundos
A5	0,999999696 (7,51) [0,9999999; 0,9999995]	0,999999682 (7,49) [0,9999996; 0,9999996]	0,000001%	9 minutos 30 segundos	7 segundos

Legenda: Arq - Arquiteturas; Disponibilidade FB (9's) [IC] - Disponibilidade por Força bruta (Em Números de noes) [Intervalo de Confiança com nível de significância de 95%]; Tempo Exec FB - Tempo de Execução por Força Bruta; Tempo Exec Otimizada - Tempo de Execução da Estratégia de Otimização.

Analisando os resultados obtidos pode ser observado que a disponibilidade da solução por meio de força bruta e da solução otimizada foram bem próximas, apresentando uma diferença inferior a 2%, no pior caso. Contudo devemos observar que o tempo de execução entre as duas abordagens apresentadas são bem distintos, e na abordagem utilizando força bruta o tempo de execução tende a crescer de acordo com a arquitetura analisada.

Conforme pode ser observado na Tabela 6, para a arquitetura A1, a estratégia proposta foi capaz de encontrar a solução ótima. Vale salientar que se analisarmos o tempo de execução da solução pelo algoritmo de força bruta e a solução otimizada para a mesma arquitetura A1, percebe-se que o tempo de execução da solução pelo algoritmo de força bruta (132s) foi superior ao tempo de execução da solução otimizada (5s), ou seja, houve uma redução acima de vinte e seis vezes no tempo de execução do experimento.

Ainda em relação aos resultados presentes na Tabela 6, observou-se que houve um aumento considerável no tempo de execução dos experimentos ao utilizar o algoritmo de força bruta, porém com a estratégia de otimização adotada, esse aumento no tempo de execução é desprezível. Por exemplo: a variação do tempo de execução entre os experimentos das arquiteturas A1 e A2 utilizando força bruta varia em cerca de 45 minutos. Já na estratégia proposta o tempo de execução de ambas as arquiteturas foi de 5 segundos.

Pode-se destacar que há uma redução no tempo de execução, por meio do algoritmo de força bruta, entre as arquiteturas A4 e A5. Isso ocorre devido ao fato que na arquitetura A5 existe a remoção do dispositivo STS diminuindo assim a quantidade de cenários a serem analisados, ou seja, na arquitetura A4 temos cinco dispositivos e conforme mencionado anteriormente para cada dispositivo temos dez equipamentos distintos a serem combinados, resultando em $10 \times 10 \times 10 \times 10 \times 10 = 100.000$ possibilidades. Já na arquitetura A5 temos

apenas quatro dispositivos, ou seja, 10.000 possibilidades. Devido a redução da quantidade de possibilidades a serem analisadas, o tempo de execução por meio da força bruta diminui.

6.3 Estudo de caso II

Essa seção apresenta um segundo estudo de caso com o objetivo de ilustrar a estratégia de otimização que leva em consideração técnicas de planejamento de experimentos e diferentes políticas de manutenção, para otimizar a disponibilidade da infraestrutura elétrica de um *data center*. Apresentando assim, uma comparação entre as políticas de manutenção corretiva e preventiva. Sendo assim, este estudo de caso demonstra que a estratégia proposta pode ser aplicada a diferentes cenários. As arquiteturas A1 a A5 de *data centers*, previamente apresentadas, foram avaliadas nesse estudo de caso, e cada uma dessas arquiteturas é formada por conjuntos específicos de equipamentos com redundâncias.

6.3.1 Modelos

Nesta seção são apresentados os modelos propostos em SPN para o cálculo da disponibilidade levando em consideração a manutenção corretiva e a manutenção preventiva.

6.3.1.1 Modelo para a Manutenção Corretiva

A Figura 19 apresenta o modelo SPN proposto para computar a disponibilidade da arquitetura A1 considerando a realização da manutenção corretiva. Neste modelo, o equipamento só irá ser reparado caso ao menos uma equipe de manutenção esteja disponível para ser alocada para realizar o reparo. A disponibilidade da arquitetura A1 é computada através da equação: $A = (\#UPS_ON = 1)AND(\#SDT_ON = 1)AND(\#Subpainel_ON = 1)AND(\#powerstrip_ON = 1)$.

6.3.1.2 Modelo para a Manutenção Preventiva

A Figura 20 apresenta o modelo SPN proposto para computar o impacto da manutenção preventiva na disponibilidade da arquitetura A2. Nesta arquitetura, observa-se

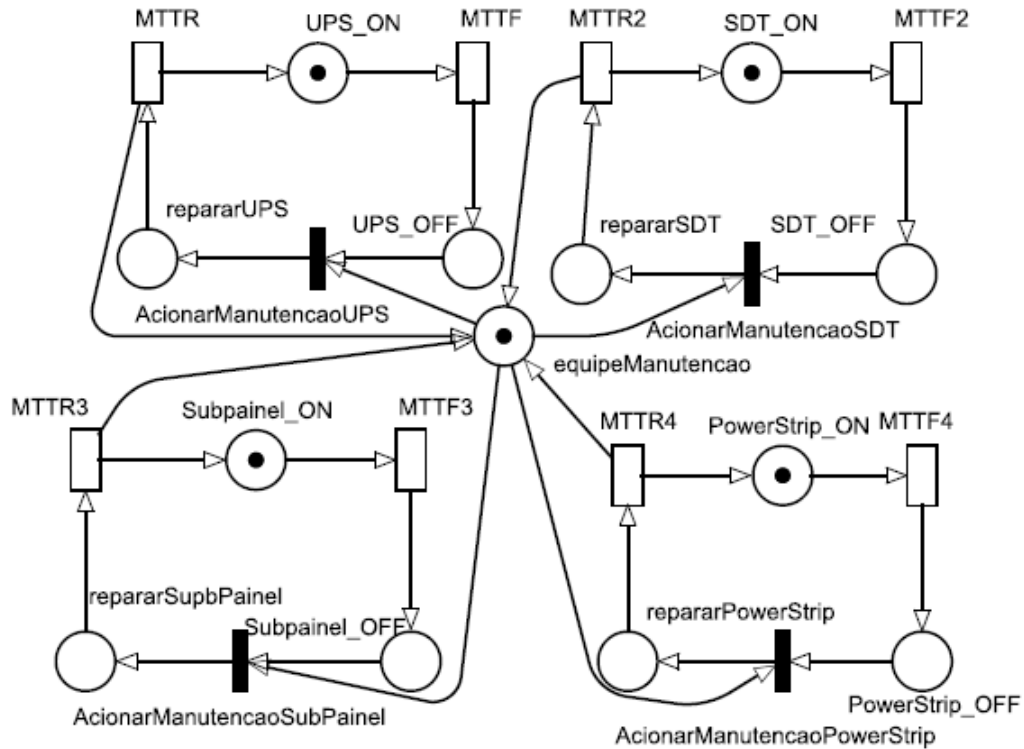


Figura 19 – Modelo da Arquitetura A1 de Manutenção Corretiva

que existe uma redundância do equipamento UPS. Neste modelo, o equipamento é reparado em caso de falha (manutenção corretiva), ou o equipamento pode ser submetido à manutenção preventiva (somente o UPS). A disponibilidade é computada através da equação: $((\#UPS_ON = 1)OR(\#UPS_ON2 = 1))AND(\#SDT_ON = 1)AND(\#Subpainel_ON = 1)AND(\#PowerStrip_ON = 1)AND(\#STS_ON = 1)$. Para a manutenção preventiva, a arquitetura A1 não foi analisada, visto que para melhorar a disponibilidade do serviço só faz sentido realizar a manutenção preventiva em equipamentos redundantes, ou seja, das arquiteturas A2 a A5.

Vale ressaltar que para efeitos de simplificação da representação, os modelos das Figuras 19 e 20 foram apresentados sem os seus respectivos *clocks*, previamente descritos (Seção 5.3 e 5.4). Destaca-se ainda, que todas as transições temporizadas que têm um tempo para falha associado, têm o tempo definido de acordo com o *clocks*, conforme mostra a expressão associada a cada transição $MTTF_X - (MTTF_X \times 0.1 \times \#primeiroAno) - (MTTF_X \times 0.2 \times \#segundoAno) - (MTTF_X \times 0.3 \times \#terceiroAno) - (MTTF_X \times 0.4 \times \#quartoAno) - (MTTF_X \times 0.5 \times \#quintoAno)$, onde X representa cada um dos equipamentos.

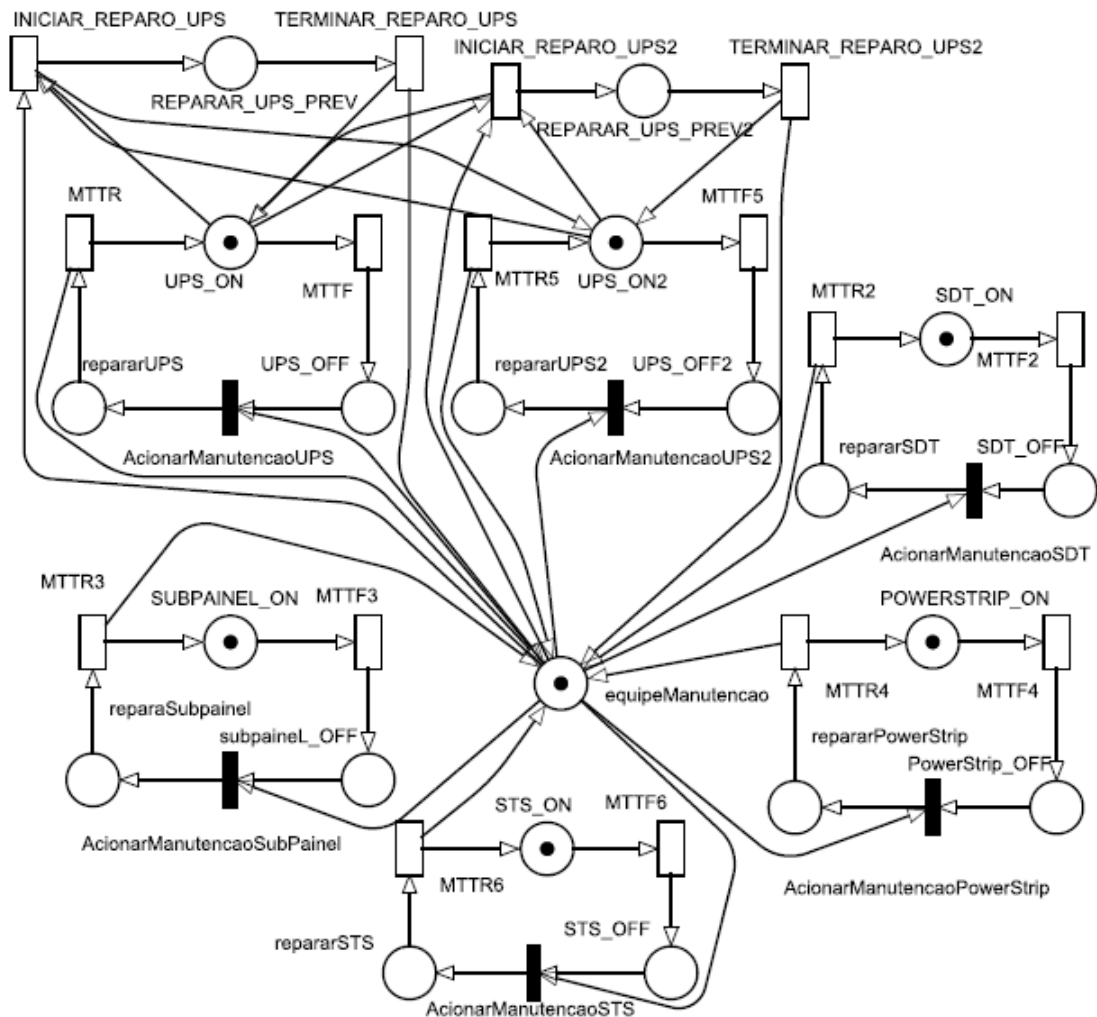


Figura 20 – Modelo da Arquitetura A2 de Manutenção Preventiva

6.3.2 Resultados

Vale ressaltar que a faixa de valores adotada para o MTTF, custo de aquisição e eficiência energética para a criação da lista de componentes candidatos são os mesmos que foram apresentados na Tabela 4. Além disso, as equações utilizadas para obter os resultados da disponibilidade dos modelos são as mesmas conforme pode ser visto na Tabela 3. Já os valores associados às transições MTTR, MTTR2, MTTR3 e MTTR4 presentes na Figura 19 são 8h. Já para a manutenção preventiva, a Tabela 7 apresenta o tempo de cada transição presente no modelo SPN da Figura 20.

Este estudo de caso faz uso das mesmas abordagens utilizadas no primeiro estudo de caso, com a diferença que este estudo leva em consideração duas políticas de manutenção distintas. Fazendo uso assim da política de manutenção corretiva e preventiva. Diferente do estudo de caso anterior, neste não foi possível realizar a verificação de todos os cenários

Tabela 7 – MTTFs e MTTRs das Transições Temporizadas do Modelo da Arquitetura A2 com Manutenção Preventiva

Transição Temporizada	Tempo (h)
MTTR, MTTR2, MTTR3, MTTR4, MTTR5 e MTTR6	8
INICIA_REPARO_UPS e INICIA_REPARO_UPS2	8760
TERMINAR_REPARO_UPS e TERMINAR_REPARO_UPS2	4
AcionarManutenção	Transição Imediata

(por força bruta), visto que se tratam de modelos com maiores dependências e sua execução acaba sendo mais custosa. Com isso, a verificação de todas as possibilidades através de um algoritmo de força bruta não é indicado, visto que a demanda de recurso computacional e o tempo para obtenção dos resultados inviabiliza tal abordagem.

A Tabela 8 apresenta as equações utilizadas para realizar o cálculo da degradação do tempo de falha utilizado para cada equipamento presente na infraestrutura analisada. Conforme apresentado no Capítulo 5, a degradação pode acontecer numa taxa de 10%, 20%, 30%, 40%, e por fim pode atingir uma degradação de até 50%. Esta degradação é utilizada para ambas as políticas de manutenção que estão sendo analisadas.

Os clocks utilizados para a manutenção corretiva e para a manutenção preventiva foram apresentados no Capítulo 5. A Tabela 9 ilustra as expressões de guarda utilizadas para ativação das transições imediatas no Modelo da arquitetura A2 com Manutenção Preventiva (Figura 20). As expressões de guarda foram utilizadas devido ao fato de que quando um equipamento é submetido a uma manutenção preventiva ele passa ao status de novo, ou seja, não existe nenhuma degradação no tempo de falha deste dispositivo. Já para a manutenção corretiva utiliza-se também um clock, porém neste caso não existe a volta para o estado inicial, sendo assim não existe a transição imediata voltando para o estado inicial.

As Tabelas 10 e 11 apresentam os resultados da disponibilidade para as arquiteturas (A1, . . . , A5). É possível perceber que houve um aumento considerável na disponibilidade do sistema para cada arquitetura. Nessas tabelas são apresentados os resultados da disponibilidade, disponibilidade em número de noves e do tempo de execução demandado para a obtenção dos resultados. O cálculo da disponibilidade em número de noves foi feito a partir da fórmula $(-\lg[1 - A/100])$, sendo A a disponibilidade.

Tabela 8 – Equação do cálculo da degradação para o modelo da Figura 19.

Equipamento	Transição	Fórmula da Degradação de cada Equipamento
UPS	MTTF	$MTTF_UPS - (MTTF_UPS * 0.1 * \#primeiroAno) - (MTTF_UPS * 0.2 * \#segundoAno) - (MTTF_UPS * 0.3 * \#terceiroAno) - (MTTF_UPS * 0.4 * \#quartoAno) - (MTTF_UPS * 0.5 * \#quintoAno)$
SDT	MTTF2	$MTTF_SDT - (MTTF_SDT * 0.1 * \#primeiroAno) - (MTTF_SDT * 0.2 * \#segundoAno) - (MTTF_SDT * 0.3 * \#terceiroAno) - (MTTF_SDT * 0.4 * \#quartoAno) - (MTTF_SDT * 0.5 * \#quintoAno)$
SUBPAINEL	MTTF3	$MTTF_SUBPAINEL - (MTTF_SUBPAINEL * 0.1 * \#primeiroAno) - (MTTF_SUBPAINEL * 0.2 * \#segundoAno) - (MTTF_SUBPAINEL * 0.3 * \#terceiroAno) - (MTTF_SUBPAINEL * 0.4 * \#quartoAno) - (MTTF_SUBPAINEL * 0.5 * \#quintoAno)$
<i>Powerstrip</i>	MTTF4	$MTTF_POWERSTRIP - (MTTF_POWERSTRIP * 0.1 * \#primeiroAno) - (MTTF_POWERSTRIP * 0.2 * \#segundoAno) - (MTTF_POWERSTRIP * 0.3 * \#terceiroAno) - (MTTF_POWERSTRIP * 0.4 * \#quartoAno) - (MTTF_POWERSTRIP * 0.5 * \#quintoAno)$

Tabela 9 – Expressões de guarda Modelo da Arquitetura A2 com Manutenção Preventiva (ex., A2 - Figura 20).

Arquitetura	Transição	Expressão de Guarda
A2	T6, T7, T8, T9, T10	$((\#UPS_OFF_PREV > 0) \text{ OR } (\#UPS_OFF_PREV2 > 0))$
A3	T6, T7, T8, T9, T10	$((\#UPS_OFF_PREV > 0) \text{ OR } (\#UPS_OFF_PREV2 > 0) \text{ OR } (\#SDT_OFF_PREV > 0) \text{ OR } (\#SDT_OFF_PREV2 > 0))$
A4	T6, T7, T8, T9, T10	$((\#UPS_OFF_PREV > 0) \text{ OR } (\#UPS_OFF_PREV2 > 0) \text{ OR } (\#SDT_OFF_PREV > 0) \text{ OR } (\#SDT_OFF_PREV2 > 0) \text{ OR } (\#SubPainel_OFF_PREV > 0) \text{ OR } (\#SubPainel_OFF_PREV2 > 0))$
A5	T6, T7, T8, T9, T10	$((\#UPS_OFF_PREV > 0) \text{ OR } (\#UPS_OFF_PREV2 > 0) \text{ OR } (\#SDT_OFF_PREV > 0) \text{ OR } (\#SDT_OFF_PREV2 > 0) \text{ OR } (\#SubPainel_OFF_PREV > 0) \text{ OR } (\#SubPainel_OFF_PREV2 > 0) \text{ OR } (\#PowerStrip_OFF_PREV > 0) \text{ OR } (\#PowerStrip_OFF_PREV2 > 0))$

Analisando as Tabelas 10 e 11, pode-se observar que todos os valores das métricas apresentadas tendem a aumentar. Além da disponibilidade, os tempos de execução dos resultados também tiveram um aumento considerável. Isso ocorre devido ao fato que cada arquitetura analisada tem uma composição distinta de dispositivos. A quantidade de equipamentos que compõem a arquitetura A1 é bem menor que a quantidade de equipamentos que compõem a arquitetura A5, por exemplo.

Tabela 10 – Resultados das métricas considerando a realização da Manutenção Corretiva

1 Equipe de Manutenção			2 Equipes de Manutenção	
Arq	Disponibilidade (Nº de Noves)	Tempo Exec (s)	Disponibilidade (Nº de Noves)	Tempo Exec (s)
A1	0,9996184504 (3,41844901)	6	0,9996184943 (3,418498914)	6
A1	0,9996180635 (3,418498914)	21	0,9996185572 (3,418570553)	33
A3	0,9996562967 (3,463816356)	125	0,9996568399 (3,464503252)	529
A4	0,9996965223 (3,517873225)	1295	0,9996969583 (3,518497644)	7042
A5	0,9999988585 (5,94253639)	819	0,9999994296 (6,24385231)	2067

Legenda: Arq - Arquiteturas; Nº de Noves - Disponibilidade em Números de noves; Tempo Exec (s) - Tempo Execução em segundos.

Tabela 11 – Resultados das métricas considerando a realização da Manutenção Preventiva

1 Equipe de Manutenção			2 Equipes de Manutenção	
Arq	Disponibilidade (Nº de Noves)	Tempo Exec (s)	Disponibilidade (Nº de Noves)	Tempo Exec (s)
A2	0,9997940369 (3,686210569)	42	0,9997942974 (3,686760168)	83
A3	0,9998255071 (3,758222121)	375	0,999825796 (3,758941835)	1584
A4	0,9998461995 (3,813042363)	4693	0,9998465791 (3,814115369)	37126
A5	0,9999996099 (6,408794069)	5305	0,9999997493 (6,600883992)	47037

Legenda: Arq - Arquiteturas; Nº de Noves - Disponibilidade em Números de noves; Tempo Exec (s) - Tempo Execução em segundos.

Foi realizado também a verificação da utilização de duas equipes de manutenção, a fim de verificar o impacto que teria no sistema na utilização de duas equipes de manutenção. Conforme pode ser observado, houve um aumento da disponibilidade do sistema ao fazer uso de duas equipes de manutenção. Porém, o tempo de execução também teve um aumento considerável.

Observando os resultados da Tabela 10, pode-se perceber que para a manutenção

corretiva a arquitetura A5 apresenta a melhor disponibilidade em relação as demais arquiteturas, apresentando uma disponibilidade de 5,94 com uma equipe de manutenção e 6,24 com a utilização de duas equipes de manutenção. Houve um maior ganho de disponibilidade para a arquitetura A5 ao fazer uso de duas equipes de manutenção, em relação as demais arquiteturas a variação foi menor. Já com relação ao tempo de execução, observou-se também um aumento, tanto ao utilizar uma equipe ou duas equipes de manutenção.

Realizando a mesma análise para os resultados da disponibilidade apresentados na Tabela 11 para a manutenção preventiva, observou-se que a arquitetura A5 apresenta a melhor disponibilidade em relação as demais arquiteturas, apresentando uma disponibilidade de 6,4 com uma equipe de manutenção e 6,60 com a utilização de duas equipes de manutenção. Houve um maior ganho de disponibilidade para a arquitetura A5 ao fazer uso de duas equipes de manutenção, em relação as demais arquiteturas também existe um aumento na disponibilidade, porém menor. O tempo de execução também aumentou, tanto na utilização de uma equipe de manutenção, com a utilização de duas equipes de manutenção.

Vale ressaltar que para o caso da manutenção corretiva são avaliadas as arquiteturas (A1, ..., A5). Porém, para a manutenção preventiva foram avaliadas apenas as arquiteturas (A2, ..., A5). Isso ocorre pelo fato desta política de manutenção ser utilizada apenas em equipamentos redundantes, cenário observado a partir da arquitetura A2.

Realizando uma comparação entre os resultados das Tabelas 10 e 11 com o uso de uma equipe de manutenção, pode-se observar que na manutenção preventiva obteve os melhores resultados, quando comparada a manutenção corretiva.

A manutenção preventiva obteve uma maior disponibilidade quando comparada com a manutenção corretiva (ver Tabelas 10 e 11), ou seja, o sistema ficou mais disponível com a utilização da manutenção preventiva. A Figura 21 mostra um comparativo dos resultados obtidos, para a arquitetura A5, em relação aos cenários analisados para a disponibilidade. É possível observar que a manutenção preventiva com um time de manutenção (disponibilidade de 6,4 em números de noves) foi mais efetiva do que a manutenção corretiva com duas equipes de manutenção. Vale ressaltar que a manutenção preventiva com dois times de manutenção (disponibilidade de 6,6 em números de noves) foi a mais efetiva das analisadas.

A manutenção preventiva apresenta uma melhor disponibilidade, pois após o período de um ano (8760 horas) o equipamento é submetido a uma manutenção, permitindo assim que este equipamento volte a ficar como um novo. Já na manutenção corretiva é necessário que o equipamento entre em um estado de falha para ser submetido a uma manutenção. Os resultados obtidos mostram que o uso de uma manutenção preventiva é de grande valia, pois permite que o sistema tenha uma disponibilidade melhor quando comparada com a manutenção corretiva.

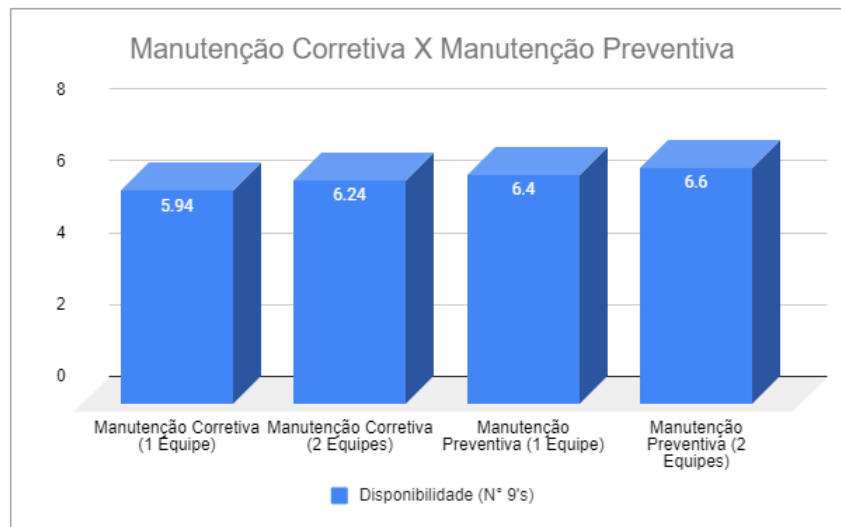


Figura 21 – Comparação da Manutenção Corretiva vs Manutenção Preventiva - Arquitetura A5

6.4 Estudo de Caso III

Esse estudo difere dos anteriores realizados por adicionar a métrica de custo. Sendo assim, pode-se analisar o impacto de diferentes políticas de manutenção, que levam em consideração contratos de *SLA*, em relação a disponibilidade e custo. Vale salientar que também analisamos o impacto da variação no número de equipes de manutenção. Os resultados da disponibilidade alcançada por cada contrato são apresentados percentualmente e em número de noves (9's). A distância euclidiana foi utilizada considerando ambas as métricas de custo e *downtime* para auxiliar na identificação do contrato com a melhor relação de custo-benefício. Vale ressaltar que os modelos em SPN para a manutenção corretiva e manutenção preventiva são os mesmos modelos apresentados no segundo estudo de caso. Sendo assim, busca-se através desse estudo demonstrar que a

estratégia de otimização proposta pode ser aplicada no cenário que leva em consideração políticas de manutenção vinculadas a diferentes contratos de *SLA*.

6.4.1 Resultado

Esse estudo de caso faz uso da mesma abordagem utilizada no primeiro. Contudo, não foi possível analisar todos os cenários através do algoritmo de força bruta devido à complexidade dos modelos analisados. O tempo para realizar a avaliação com força bruta acabou por inviabilizar a sua utilização. Por isso, foram avaliados apenas os resultados de disponibilidade da estratégia de otimização proposta. Sendo assim, são analisados apenas quatro cenários, de acordo com os equipamentos selecionados pelo índice de importância para a disponibilidade, juntamente com os dispositivos selecionados de forma aleatória.

A Tabela 12 apresenta as particularidades existentes para cada um dos contratos *SLAs* adotados. Nessa tabela, assumindo a manutenção corretiva, pode-se observar que TTM representa o tempo total de manutenção (TTM + TA + MTTR) e envolve o tempo de deslocamento (TTD), o tempo de administrativo (TA) e o MTTR. Assumindo a manutenção preventiva, outros tempos devem ser levados em consideração. Sendo assim, para a manutenção preventiva, o TTM é composto por TA, FreqMP (frequência da realização da manutenção preventiva) e MTTR. O TTD (tempo de deslocamento) considera o tempo médio gasto pelas equipes de manutenção para o trajeto de seu local de trabalho até a empresa que solicitou o reparo; TA (tempo administrativo) corresponde ao tempo para notar a falha e solicitar o reparo; MTTR e o tempo médio para reparo. No caso da manutenção preventiva/corretiva, também se utiliza a métrica FreqMP, que corresponde a frequência (periodicidade) da realização da manutenção preventiva.

Tabela 12 – Contratos *SLA*

<i>SLA</i>	TTM	TTD	TA	MTTR	FreqMP
I	120h	18h	98h	4h	17520h (2 anos)
II	48h	10h	34h	4h	13140h (1 ano e meio)
III	24h	6h	14h	4h	8760h (1 anos)
IV	8h	3h	1h	4h	4380h (6 meses)

O contrato de *SLA* para a manutenção corretiva considera as métricas: (i) custo

da manutenção corretiva, este é composto pelo percentual de 15% do custo de aquisição de cada componente que compõem a infraestrutura e o *downtime* deste equipamento; (ii) o custo com o time de manutenção, este custo pode variar em função da quantidade de times de manutenção disponíveis para fazer o reparo, sendo o custo mensal assumido para cada equipe de R\$3.000,00. Para a manutenção preventiva, deve-se considerar também o custo extra da realização de uma manutenção periódica, valor obtido considerando um percentual de 10% do valor do custo com a manutenção corretiva. Vale ressaltar que os custos variam de acordo com o contrato de *SLA* adotado. Lembrando que os custos de aquisição dos equipamentos são os mesmos apresentados anteriormente, conforme pode ser visto na Tabela 4.

A Tabela 13 apresenta os contratos de *SLA* para a manutenção corretiva. Conforme pode ser observado, foram criados os seguintes contratos: *SLA* 1, *SLA* 2, *SLA* 3 e *SLA* 4. O tempo de reparo total para cada contrato *SLA* foi de 120 horas, 48 horas, 24 horas e 8 horas, respectivamente (Callou et al., 2010).

A Tabela 14 apresenta os contratos de *SLA* para a manutenção preventiva. Existe, porém um dado extra que é a frequência da manutenção preventiva, podendo ser realizada a cada dois anos (17520 horas), um ano e meio (13140 horas), um ano (8760 horas) e seis meses (4380), dependendo do contrato escolhido (Callou et al., 2010).

Vale salientar que para ambas as políticas de manutenção alguns contratos de *SLA* apresentam os seus custos com valores elevados, por exemplo no contrato de *SLA* para a manutenção corretiva esperava-se que os custos do contrato da arquitetura A4 fossem mais elevados que os custos do contrato de *SLA* da arquitetura A3. Porém, acontece justamente o inverso, onde se tem o contrato da arquitetura A3 sendo mais elevado que o contrato da arquitetura A4, isso ocorre devido ao fato que os componentes que compõem estas arquiteturas são distintos uns dos outros, já que são escolhidos de acordo com o índice de importância para a disponibilidade e, também, de forma aleatória. Logo, os equipamentos que compõem cada arquitetura podem variar, conseqüentemente, o custo associado a manutenção também pode variar.

Vale ressaltar que para o caso da manutenção corretiva são avaliadas as arquiteturas (A1, ..., A5). Porém, para a manutenção preventiva, foram avaliadas apenas as arquiteturas (A2, ..., A5). Isso ocorreu pelo fato desta política de manutenção ser utilizada apenas em equipamentos redundantes, cenário observado a partir da arquitetura A2.

Tabela 13 – Contrato *SLA* - Manutenção Corretiva

Arquitetura A1			
<i>SLA</i>	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA</i> I	120h	R\$8.493,36	R\$12.993,36
<i>SLA</i> II	48h	R\$8.892,70	R\$13.392,70
<i>SLA</i> III	24	R\$9.331,97	R\$13.831,97
<i>SLA</i> IV	8h	R\$9.815,17	R\$14.315,17
Arquitetura A2			
<i>SLA</i>	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA</i> I	120h	R\$19.945,34	R\$24.445,34
<i>SLA</i> II	48h	R\$21.489,88	R\$25.989,88
<i>SLA</i> III	24	R\$23.188,86	R\$27.688,86
<i>SLA</i> IV	8h	R\$25.057,75	R\$29.557,75
Arquitetura A3			
<i>SLA</i>	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA</i> I	120h	R\$20.103,46	R\$24.603,46
<i>SLA</i> II	48h	R\$21.663,80	R\$26.163,80
<i>SLA</i> III	24	R\$23.380,18	R\$27.880,18
<i>SLA</i> IV	8h	R\$25.268,20	R\$29.768,20
Arquitetura A4			
<i>SLA</i>	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA</i> I	120h	R\$20.810,72	R\$24.580,62
<i>SLA</i> II	48h	R\$22.441,80	R\$26.138,69
<i>SLA</i> III	24	R\$24.235,98	R\$27.852,56
<i>SLA</i> IV	8h	R\$26.209,57	R\$29.737,81
Arquitetura A5			
<i>SLA</i>	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA</i> I	120h	R\$20.245,39	R\$24.745,39
<i>SLA</i> II	48h	R\$21.819,93	R\$26.319,93
<i>SLA</i> III	24	R\$23.551,93	R\$28.051,93
<i>SLA</i> IV	8h	R\$25.457,12	R\$29.957,12

Legenda: TTM - Tempo total de manutenção; EQ - Equipe de manutenção;

Tabela 14 – Contato *SLA* - Manutenção Preventiva

Arquitetura A2				
<i>SLA</i>	FreqMP	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA I</i>	(2 anos) = 26280	120h	R\$21.489,88	R\$25.989,88
<i>SLA II</i>	(1.5) = 17520	48h	R\$27.018,50	R\$29.518,50
<i>SLA III</i>	(1 anos) = 8760	24h	R\$25.057,75	R\$29.557,75
<i>SLA IV</i>	(6 meses) = 4380	8h	R\$27.113,53	R\$31.613,53
Arquitetura A3				
<i>SLA</i>	FreqMP	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA I</i>	(2 anos) = 26280	120h	R\$21.663,80	R\$26.163,80
<i>SLA II</i>	(1.5) = 17520	48h	R\$23.380,18	R\$27.880,18
<i>SLA III</i>	(1 anos) = 8760	24h	R\$25.268,20	R\$29.768,20
<i>SLA IV</i>	(6 meses) = 4380	8h	R\$27.345,02	R\$31.845,02
Arquitetura A4				
<i>SLA</i>	FreqMP	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA I</i>	(2 anos) = 26280	120h	R\$22.441,80	R\$26.941,80
<i>SLA II</i>	(1.5) = 17520	48h	R\$24.235,98	R\$28.735,98
<i>SLA III</i>	(1 anos) = 8760	24h	R\$26.209,57	R\$30.709,57
<i>SLA IV</i>	(6 meses) = 4380	8h	R\$28.380,53	R\$32.880,53
Arquitetura A5				
<i>SLA</i>	FreqMP	TTM(h)	Custo do Contrato 1 EQ	Custo do Contrato 2 EQ
<i>SLA I</i>	(2 anos) = 26280	120h	R\$21.819,93	R\$26.319,93
<i>SLA II</i>	(1.5) = 17520	48h	R\$23.551,93	R\$28.051,93
<i>SLA III</i>	(1 anos) = 8760	24h	R\$25.457,12	R\$29.957,12
<i>SLA IV</i>	(6 meses) = 4380	8h	R\$27.552,83	R\$32.052,83

Legenda: FreqMP - Frequência na realização da manutenção preventiva; TTM - Tempo total de manutenção; EQ - Equipe de manutenção;

As Tabelas 15 e 16 apresentam, respectivamente, os resultados da disponibilidade das manutenções corretiva e preventiva, obtidos para cada arquitetura levando em consideração os diferentes *SLAs* e considerando diferentes números de equipes de manutenção disponíveis.

A arquitetura A4 foi escolhida para ser analisada, pois entre as arquiteturas apresentadas nas Tabelas 15 e 16, ela é a que apresenta a melhor relação custo-benefício.

Conforme pode ser observado nas Tabelas 15 e 16, o contrato *SLA* do nível 4 oferece uma maior disponibilidade do que os demais contratos. No caso da manutenção corretiva, tem-se a disponibilidade (números de noves) para o contrato de nível 1 e 4 de 3,40 e 3,52 respectivamente. No caso da manutenção preventiva, tem-se, respectivamente, para o Contrato de nível 1 e 4 a disponibilidade de 3,75 e 3,81.

As Tabelas 15 e 16 mostram que a manutenção preventiva apresenta melhores resultados para a disponibilidade quando comparada a manutenção corretiva. Vale ressaltar que o comportamento dos resultados de disponibilidade obtidos para a arquitetura A4 se repete para as arquiteturas A2 e A3, onde os resultados obtidos pela manutenção preventiva apresentam melhores resultados. Porém, para a arquitetura A5, observou-se uma particularidade no contrato *SLA4*, onde ocorrem manutenções preventivas com intervalos de seis meses (4380h), a disponibilidade da manutenção preventiva apresenta valores menores que o cenário levando em consideração somente a manutenção corretiva. Nesse caso, foi verificado que a utilização da manutenção preventiva no período de seis meses acarretou em uma queda na disponibilidade do sistema. Sendo assim, neste cenário é mais indicado utilizar a manutenção corretiva com duas equipes de manutenção. Porém, para os *SLA1*, *SLA2* e *SLA3* com a manutenção ocorrendo em um intervalo de dois anos (17520h), um ano meio (13140h) e um ano (8760), obteve-se melhores resultados.

A manutenção preventiva permite que os equipamentos voltem a funcionar em estado de novo (sem ter o MTTF degradado pela ação do tempo). Já na manutenção corretiva, com o passar do tempo, os equipamentos passam a falhar com mais frequência (MTTF degradado). Os resultados obtidos mostram que o uso de uma manutenção preventiva é de grande valia, pois permite que o sistema tenha uma disponibilidade melhor quando comparada unicamente com a manutenção corretiva.

Foi avaliado também o impacto da utilização de contratos com mais equipes de manutenção no sistema, tendo assim análises referentes a uma e duas equipes de manutenção. É possível observar que a utilização de mais equipes de manutenção aumenta a disponibilidade do sistema. Porém para ambas políticas de manutenção, a utilização de duas equipes são suficientes para obter a maior disponibilidade do sistema.

Avaliando o custo dos contratos de *SLA* associado às políticas de manutenção (Tabelas 13 e 14), percebe-se que para cada nível do contrato os preços tendem a ficar maiores. Porém mesmo existindo o aumento no valor do contrato, analisando as Tabelas

Tabela 15 – Disponibilidade (9's) da Manutenção Corretiva

SLA	Arquitetura A1		Arquitetura A2		Arquitetura A3		Arquitetura A4		Arquitetura A5	
	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's
	1 equipe	2 equipes	1 equipe	2 equipes	1 equipe	2 equipes	1 equipe	2 equipes	1 equipe	2 equipes
SLA 1	2,21	2,21	3,33	3,40	3,31	3,40	3,40	3,49	4,26	4,57
SLA 2	2,61	2,61	3,40	3,42	3,41	3,43	3,50	3,51	5,06	5,36
SLA 3	2,91	2,91	3,41	3,42	3,43	3,43	3,51	3,52	5,66	5,96
SLA 4	3,39	3,39	3,42	3,42	3,43	3,43	3,52	3,52	6,61	6,92

Tabela 16 – Disponibilidade (9's) da Manutenção Preventiva

SLA	Arquitetura A2		Arquitetura A3		Arquitetura A4		Arquitetura A5	
	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's	Disp. 9's
	1 equipe	2 equipes	1 equipe	2 equipes	1 equipe	2 equipes	1 equipe	2 equipes
SLA 1	3,60	3,68	3,60	3,68	3,75	3,79	4,82	5,12
SLA 2	3,70	3,72	3,70	3,72	3,80	3,81	5,59	5,86
SLA 3	3,72	3,73	3,72	3,73	3,81	3,81	6,04	6,24
SLA 4	3,73	3,73	3,73	3,73	3,81	3,82	6,14	6,33

15 e 16, a disponibilidade para cada nível do contrato também tende a crescer.

Este estudo de caso utiliza a distância Euclidiana para detectar a melhor relação custo-benefício. As Tabelas 17 e 18 apresentam as distâncias euclidianas para as políticas de manutenção e associa a cada contrato de *SLA*, considerando o *downtime* e custo do contrato. Vale salientar que os resultados apresentados consideram o eixo “X” como sendo os valores obtidos para o *downtime*, vinculado a cada contrato e o eixo “Y” como sendo o custo de cada contrato de *SLA*. A distância euclidiana foi obtida através da fórmula: $D = \sqrt{(X_f - X_i)^2 + (Y_f - Y_i)^2}$; onde X_i ou Y_i é a distância inicial e X_f ou Y_f é a distância final. Como as métricas citadas apresentam ordem de grandezas distintas, os resultados são apresentados como valores absolutos e normalizados em relação ao componente de maior valor, através da equação: $ln_i = \frac{I_i}{I_x}$, sendo ln_i o índice normalizado para o componente i ; I_i o valor não normalizado para o componente i ; e I_x corresponde ao maior índice não normalizado entre os componentes.

Assumindo os resultados das distâncias euclidianas para ambas políticas de manutenção, pode-se identificar a de menor valor que é 1,1009 para o *SLA2* com um time de manutenção corretiva. Vale ressaltar que o contrato do nível 2 com apenas uma equipe

para a manutenção preventiva apresenta resultados bem próximos com uma distância de 1,1479.

Tabela 17 – Distância Euclidiana - Manutenção Corretiva

<i>SLA</i>	1 Equipe	2 Equipes
<i>SLA 1</i>	1,220544747	1,296553079
<i>SLA 2</i>	1,100930757	1,132042095
<i>SLA 3</i>	1,123345903	1,165943502
<i>SLA 4</i>	1,166828283	1,208587690

Tabela 18 – Distância Euclidiana - Manutenção Preventiva

<i>SLA</i>	1 Equipe	2 Equipes
<i>SLA 1</i>	1,210719229	1,214310516
<i>SLA 2</i>	1,147947465	1,228085576
<i>SLA 3</i>	1,173510903	1,266784679
<i>SLA 4</i>	1,214830148	1,313257084

A Figura 22 apresenta os resultados da manutenção corretiva da arquitetura A4, para todos os contratos *SLAs*, levando em consideração uma e duas equipes de manutenção. Similarmente, a Figura 23 apresenta os resultados para a manutenção preventiva. Observa-se que entre os contratos *SLAs* da manutenção corretiva, o contrato de *SLA* do nível dois com apenas uma equipe de manutenção apresenta a menor distância euclidiana, sendo o mais indicado para ser utilizado. Similar resultado foi detectado na manutenção preventiva, onde o *SLA 2* com uma equipe de manutenção tem a menor distância Euclidiana. Para ambos os casos, esses contratos são apresentados como sendo os melhores, pois apresentam a menor distância euclidiana.

Analisando os contratos destinados a arquitetura A4, o *SLA 1* com duas equipes de manutenção é considerado o contrato menos indicado para ser utilizado, pois apresenta a maior distância euclidiana para a manutenção corretiva. Realizando a mesma comparação para a manutenção preventiva verificou-se que o contrato de *SLA 4*, com duas equipes de manutenção, é o menos indicado para ser utilizado. Vale ressaltar que esses contratos

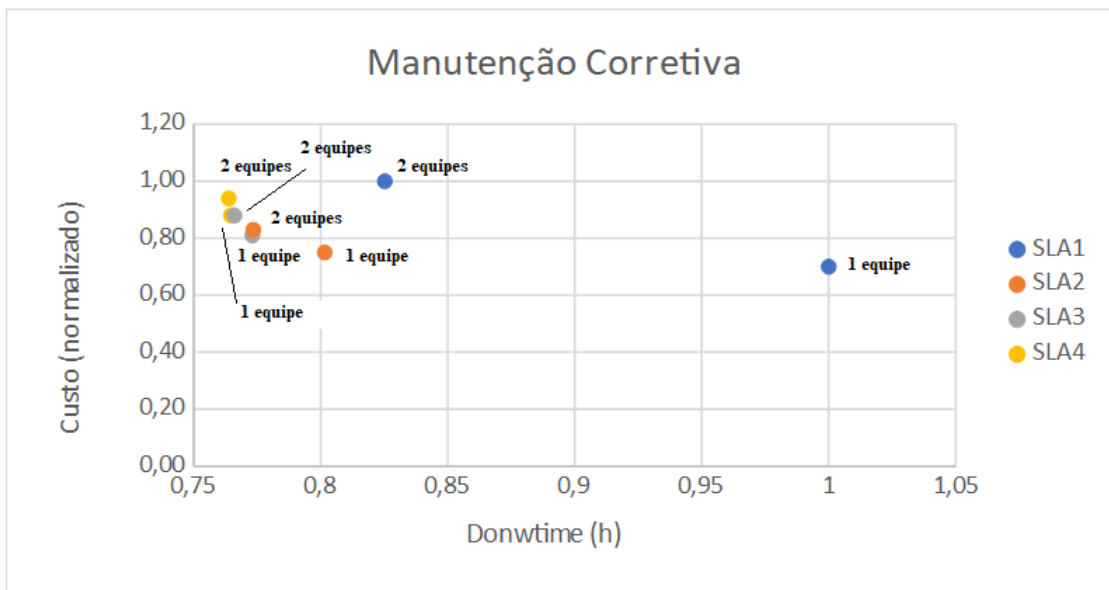


Figura 22 – *Downtime* e Custo da Manutenção Corretiva - Considerando a Arquitetura A4

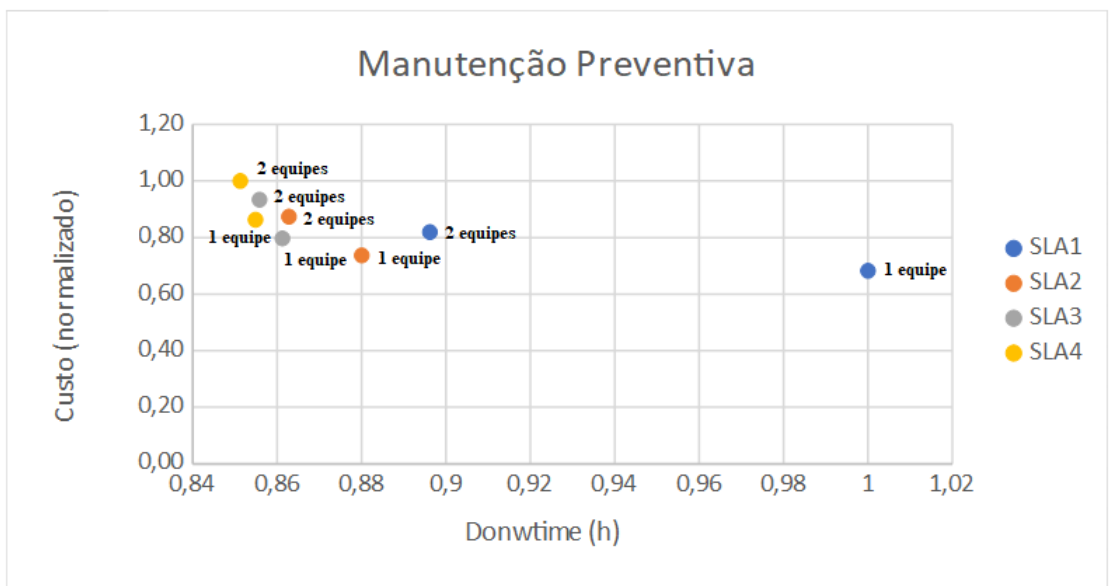


Figura 23 – *Downtime* e Custo da Manutenção Preventiva - Considerando a Arquitetura A4

apontados como os menos indicados a serem utilizados, pois apresentam a maior distância euclidiana, logo possuem o maior *downtime* e/ou custo entre os contratos analisados para a arquitetura A4.

7 Conclusão

Este trabalho de pesquisa compreende o desenvolvimento de uma estratégia de otimização que leva em consideração técnicas de planejamento de experimentos para a elaboração de um algoritmo de otimização. Neste capítulo são apresentadas as contribuições, as limitações encontradas e os trabalhos futuros.

7.1 Considerações

Este trabalho apresentou uma estratégia de otimização através do uso de planejamento de experimentos para infraestruturas elétricas de *data centers*. Fazendo uso de redes de Petri Estocásticas para analisar a disponibilidade de diferentes arquiteturas e diagrama de bloco de confiabilidade para encontrar os equipamentos mais sensíveis à disponibilidade das arquiteturas.

Experimentos evidenciaram a viabilidade da abordagem proposta, onde a estratégia de otimização fornece resultados importantes redução do número de cenários avaliados e conseqüentemente do tempo de execução. Conforme observado no Estudo de caso I, a utilização da otimização proposta foi de grande valia, pois os resultados obtidos evidenciaram que a disponibilidade do sistema ficou próximo da solução ótima. Além disso, o tempo de execução dos experimentos utilizando a força bruta foram superiores ao tempo de execução da estratégia proposta, logo o tempo de execução dos experimentos foram reduzidos consideravelmente ao utilizar a estratégia proposta em no mínimo 26 vezes.

Vale ressaltar que a estratégia de otimização proposta pode ser aplicada a diferentes cenários, conforme foi realizado no Estudo de caso II, onde tal estratégia foi aplicada ao contexto de manutenção. Sendo assim, foi possível realizar experimentos voltados a otimização de diferentes políticas de manutenções. Conforme observado no Estudo de caso II, a utilização da manutenção preventiva foi de grande valia, pois os resultados de disponibilidade do sistema foram superiores em relação aos resultados de disponibilidade obtidos através da manutenção corretiva. Além do fato de consolidar, que a estratégia de otimização desenvolvida pode ser adaptada a diferentes experimentos.

Conforme observado no Estudo de caso III, a utilização da manutenção preventiva apresentou resultados da disponibilidade do sistema superiores em relação aos resultados

de disponibilidade apresentados com a utilização da manutenção corretiva, porém os custos dos contratos também foram superiores. Estes custo superior ocorre de acordo com o contrato de *SLA* adotado. Visto que, cada contrato utilizado possui um custos associado. Fazendo uso da distância euclidiana considerando o downtime e custo observou-se que é possível se detectar os melhores contratos de *SLA* levando em consideração essas métricas.

7.2 Contribuições

Esta seção apresenta todas as principais contribuições deste trabalho. Abaixo seguem os resultados mais relevantes para esta pesquisa:

- Modelos SPN para as manutenções Corretiva e Preventiva
 - Os modelos SPN foram propostos para representar a manutenção corretiva e preventiva. Além disso, foi adotado o mecanismo de degradação para o MTTF, que variar de acordo com o tempo podendo ter perdas de 10%, 20%, 30%, 40% e 50%. A degradação do MTTF é computada através de um clock, que com o passar do tempo pode infligir uma perda no MTTF do equipamento.
- Desenvolvimento de uma estratégia de otimização através do uso da técnica de planejamento de experimentos:
 - Com a utilização da estratégia proposta de otimização é possível realizar experimentos mais direcionados, avaliando cenários mais reduzidos. Possibilitando também uma redução no tempo de execução dos experimentos.
 - Ao realizar a proposta de otimização foi possível obter os resultados em um menor tempo de execução. Se observar os tempos de execução dos experimentos no Estudo de caso I, pode-se perceber que houve uma diminuição considerável no tempo de execução do experimento. Essa redução apresentada neste estudo de caso chegou, no mínimo a, 26 vezes.
 - Com a estratégia proposta os cenários analisados reduziram consideravelmente. Conforme foi apresentado nos estudos de casos com a realização da proposta de otimização são analisados apenas 4 cenários. Sendo assim, no lugar de analisar todos os possíveis cenários há uma redução para apenas 4 cenários.
- Aplicabilidade em diferentes cenários:
 - Conforme pode ser apresentado através dos estudos de casos realizados, a

proposta de otimização realizada pode ser aplicada a diferentes cenários. Neste trabalho a otimização proposta foi empregada em três cenários e atendeu todos sem a necessidade de mudanças na estratégia proposta.

- Ferramenta desenvolvida:
 - Foi realizado o desenvolvimento de uma ferramenta que possibilitou a automação de partes dos experimentos realizados. A ferramenta desenvolvida em Java, possibilita o cálculo da disponibilidade dos modelos analisados. Esse cálculo ocorre através da interação da aplicação desenvolvida com a ferramenta Mercury (SILVA et al., 2013). A geração dos resultados é realizado de forma automática, obtendo a disponibilidade dos equipamentos selecionados previamente. Os resultados são armazenados em arquivos, apresentando o resultado da disponibilidade de cada conjunto de equipamentos.

7.3 Limitações e Trabalhos Futuros

Os modelos propostos para representar manutenção preventiva em conjunto com a manutenção corretiva tornaram-se complexos para a concepção manual deles. Nem todos os passos da metodologia foram automatizados, demandando assim, a intervenção do usuário.

Como trabalhos futuros podemos elencar:

- A metodologia proposta pode ser considerada para avaliar outros cenários:
 - Realizar a ampliação da estratégia de otimização proposta para outros tipos de infraestruturas de *data center*, TI e refrigeração.
- Avaliar outras infraestruturas de *data center*:
 - Um outro trabalho futuro seria aplicar a estratégia proposta a arquiteturas de *data center* mais complexas. Sendo assim, aplicá-la a cenários maiores para comprovar que esta proposta também pode ser aplicada a cenários com mais dependências e componentes, por exemplo, TIERs I, II, III e IV.
- Utilização de outros algoritmos de otimização:
 - Realizar experimentos com a utilização de outros algoritmos de otimização, (por exemplo: colônia de formigas e algoritmos genéticos) e comparar com a otimização proposta neste trabalho.

- Ferramenta de criação de modelos:
 - Desenvolvimento de uma ferramenta para automatizar o processo de criação dos modelos de manutenção.
- Levar em consideração a visão do ponto de vista do provedor do serviço:
 - Um estudo pode ser realizado sob outra perspectiva, a do provedor do serviço. Sendo assim, ajustes podem ser necessários para se modelar e analisar essa diferente ponto de vista do mesmo sistema.
- Previsão de falhas:
 - Um estudo verificando a possibilidade de se detectar quando cada dispositivo poderá vir a falhar é indicado como trabalho futuro. Por exemplo, pode-se vir a utilizar redes neurais, séries temporais ou inteligência artificial para se realizar tais previsões comportamentais.

Referências

- ADDABBO, T.; BIXIO, G.; FORT, A.; MUGNAINI, M.; VIGNOLI, V.; VIGNI, F. High performance liquid chromatography lcc analysis. In: IEEE. **Instrumentation and Measurement Technology Conference (I2MTC), 2015 IEEE International**. [S.l.], 2015. p. 945–950.
- ALLEN, T. T. **Introduction to engineering statistics and six sigma: statistical quality control and design of experiments and systems**. [S.l.]: Springer Science & Business Media, 2006.
- ALMEIDA, A. T. d. Modelagem multicritério para seleção de intervalos de manutenção preventiva baseada na teoria da utilidade multiatributo. **Pesquisa Operacional**, SciELO Brasil, v. 25, n. 1, p. 69–81, 2005.
- ALPENDRE, M. Service level agreement: Um conceito a saber usar. **Departamento de Engenharia Informtica, Universidade de Coimbra**, 2006.
- ARREGOCES, M.; PORTOLANI, M. **Data center fundamentals**. [S.l.]: Cisco Press, 2003.
- AVELAR, V. Comparing availability of various rack power redundancy configurations. **APC White Paper**, v. 48, p. 1–22, 2003.
- BALANICI, M.; PACHNICKE, S. Hybrid electro-optical intra-data center networks tailored for different traffic classes. **IEEE/OSA Journal of Optical Communications and Networking**, IEEE, v. 10, n. 11, p. 889–901, 2018.
- BALBO, G. Introduction to stochastic petri nets. In: SPRINGER. **School organized by the European Educational Forum**. [S.l.], 2000. p. 84–155.
- BARROSO, L. A.; DEAN, J.; HÖLZLE, U. Web search for a planet: The google cluster architecture. **IEEE micro**, IEEE, n. 2, p. 22–28, 2003.
- BAUER, E.; ADAMS, R. **Reliability and availability of cloud computing**. [S.l.]: John Wiley & Sons, 2012.
- BICHPURIYA, Y. K.; NAVALKAR, P. V.; SOMAN, S. Benchmarking of reliability indices for electricity distribution utilities: approach and discussion. IET, 2011.
- BITBUCKET. **bitbucket**. 2008. Disponível em: <https://bitbucket.org/felipe_lmelo/projconclusao>. Acesso em: 03 julho. 2019.
- BLOKDIJK, G. **The Service Level Agreement SLA Guide-SLA book, Templates for Service Level Management and Service Level Agreement Forms. Fast and Easy Way to Write your SLA**. [S.l.]: Lulu. com, 2008.
- BORKOWSKI, M.; HANS, P. Reliability centered maintenance (rcm) handbook. **Naval Sea Systems Command, US**, 2007.
- BOSSE, S.; JAMOUS, N.; KRAMER, F.; TUROWSKI, K. Introducing greenhouse emissions in cost optimization of fault-tolerant data center design. In: IEEE. **2016 IEEE 18th Conference on Business Informatics (CBI)**. [S.l.], 2016. v. 1, p. 163–172.

BOTELHO, D. C. d. C.; CABRAL, V. R. Estudo do impacto do desalinhamento de máquina rotativas no consumo de energia elétrica. Renato Couto de Almeida, 2017.

BRIŠ, R. Evaluation of the production availability of an offshore installation by stochastic petri nets modeling. In: IEEE. **The International Conference on Digital Technologies 2013**. [S.l.], 2013. p. 147–155.

Callou, G.; Sousa, E.; Maciel, P.; Tavares, E.; Araujo, C.; Silva, B.; Rosa, N.; Marwah, M.; Sharma, R.; Shah, A.; Christian, T.; Pires, J. P.; Magnani, F. S. Impact analysis of maintenance policies on data center power infrastructure. In: **2010 IEEE International Conference on Systems, Man and Cybernetics**. [S.l.: s.n.], 2010. p. 526–533.

CALLOU, G. R. d. A. Assessment to support the planning of sustainable data centers with high availability. Universidade Federal de Pernambuco, 2013.

CHANDEL, S.; NI, T.-Y.; YANG, G. Enterprise cloud: Its growth & security challenges in china. In: IEEE. **2018 5th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2018 4th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)**. [S.l.], 2018. p. 144–152.

CHIARAVIGLIO, L.; D'ANDREAGIOVANNI, F.; LANCELLOTTI, R.; SHOJAFAR, M.; BLEFARI-MELAZZI, N.; CANALI, C. An approach to balance maintenance costs and electricity consumption in cloud data centers. **IEEE Transactions on Sustainable Computing**, IEEE, v. 3, n. 4, p. 274–288, 2018.

CHUNG, C. A. Simulation modeling handbook: a practical approach. In: **eight edition**. [S.l.: s.n.], 2003. p. 192–pp.

DATTORRO, J. **Convex optimization & Euclidean distance geometry**. [S.l.]: Lulu.com, 2010.

DAYARATHNA, M.; WEN, Y.; FAN, R. Data center energy consumption modeling: A survey. **IEEE Communications Surveys & Tutorials**, IEEE, v. 18, n. 1, p. 732–794, 2015.

DHILLON, B. S. **Engineering maintenance: a modern approach**. [S.l.]: cRc press, 2002.

DUMITRESCU, C.; PLESCA, A.; ADAM, M.; NITUCA, C.; DRAGOMIR, A. Methods for reducing energy consumption, optimization in operational data centers. In: IEEE. **2018 International Conference and Exposition on Electrical And Power Engineering (EPE)**. [S.l.], 2018. p. 0483–0486.

FIGUEIRÊDO, J. J. C. de; MACIEL, P. R. M. Análise de dependabilidade de sistemas data center baseada em índices de importância. 2011.

FOGLIATO, F.; RIBEIRO, J. L. D. **Confiabilidade e manutenção industrial**. [S.l.]: Elsevier Brasil, 2009.

FOX, A.; GRIFFITH, R.; JOSEPH, A.; KATZ, R.; KONWINSKI, A.; LEE, G.; PATTERSON, D.; RABKIN, A.; STOICA, I. Above the clouds: A berkeley view of cloud computing. **Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS**, v. 28, n. 13, p. 2009, 2009.

- GARCIA, F. L. Proposta de implantação de manutenção preventiva em um centro de usinagem vertical: Um estudo de caso. **Revista Tecnologia e Tendências**, v. 9, n. 2, p. 88–115, 2017.
- GEITNER, F. K.; BLOCH, H. P. **Machinery component maintenance and repair**. [S.l.]: Gulf Professional Publishing, 2019. v. 3.
- GERMAN, R. **Performance analysis of communication systems with non-Markovian stochastic Petri nets**. [S.l.]: John Wiley & Sons, Inc., 2000.
- GU, Y.; JIN, D. Drop test simulation and doe analysis for design optimization of microelectronics packages. In: IEEE. **56th Electronic Components and Technology Conference 2006**. [S.l.], 2006. p. 6–pp.
- GUIMARÃES, L. M.; NOGUEIRA, C. F.; SILVA, M. D. B. da. Manutenção industrial: implementação da manutenção produtiva total (tpm). **e-xacta**, v. 5, n. 1, 2012.
- HILES, A. **Service-Level Agreements in Business Continuity Management**. 2004.
- HUANG, J.; CHANG, Q.; ZOU, J.; ARINEZ, J. A real-time maintenance policy for multi-stage manufacturing systems considering imperfect maintenance effects. **IEEE Access**, IEEE, v. 6, p. 62174–62183, 2018.
- HUANG, Y.; LI, G.; WANG, P.; CHANG, F.; LI, J. Electricity cost optimization of data center interactive services with ups. In: IEEE. **2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)**. [S.l.], 2018. p. 181–184.
- JAIN, R. **The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling**. [S.l.]: John Wiley & Sons, 1990.
- JALBY, W.; WONG, D. C.; KUCK, D. J.; ACQUAVIVA, J.-T.; BEYLER, J.-C. Measuring computer performance. In: **High-Performance Scientific Computing**. [S.l.]: Springer, 2012. p. 75–95.
- JUNG, L. T.; FOONG, O.-M.; WINATA, P. Optimization of sensor deployment in data center. In: IEEE. **2016 3rd International Conference on Computer and Information Sciences (ICCOINS)**. [S.l.], 2016. p. 645–649.
- LI, W.; QI, H.; LI, K.; STOJMENOVIC, I.; LAN, J. Joint optimization of bandwidth for provider and delay for user in software defined data centers. **IEEE Transactions on Cloud Computing**, IEEE, v. 5, n. 2, p. 331–343, 2015.
- LIU, B.; LIN, J.; ZHANG, L.; KUMAR, U. A dynamic prescriptive maintenance model considering system aging and degradation. **IEEE Access**, IEEE, v. 7, p. 94931–94943, 2019.
- LIU, Y.; TRIVEDI, K. S. Survivability quantification: The analytical modeling approach. **International Journal of Performability Engineering**, v. 2, n. 1, p. 29–44, 2006.
- LUDWIG, H.; KELLER, A.; DAN, A.; KING, R. P.; FRANCK, R. Web service level agreement (wsla) language specification. **Ibm corporation**, p. 815–824, 2003.

- MACIEL K. S. TRIVEDI, R. M. P.; KIM, D. S. **Performance and Dependability in Service Computing: Concepts, Techniques and Research Directions**, chapter **Dependability Modeling**. [S.l.]: Premier Reference Source, IGI Global, 2011.
- MACIEL, P.; TRIVEDI, K.; KIM, D. Dependability modeling in: Performance and dependability in service computing: Concepts, techniques and research directions. **Hershey: IGI Global, Pennsylvania, USA**, v. 13, 2010.
- MARIN, P. S. **Data Centers—Engenharia: Infraestrutura Física**. [S.l.]: São Paulo: PM Books, 2016.
- MARSAN, M. A.; BALBO, G.; CONTE, G. et al. **Performance models of multiprocessor systems**. [S.l.]: the MIT Press, 1986.
- MARWAH, M.; MACIEL, P.; SHAH, A.; SHARMA, R.; CHRISTIAN, T.; ALMEIDA, V.; ARAÚJO, C.; SOUZA, E.; CALLOU, G.; SILVA, B. et al. Quantifying the sustainability impact of data center availability. **ACM SIGMETRICS Performance Evaluation Review**, ACM, v. 37, n. 4, p. 64–68, 2010.
- MOBLEY, R. K. **An introduction to predictive maintenance**. [S.l.]: Elsevier, 2002.
- MURATA, T. Petri nets: Properties, analysis and applications. **Proceedings of the IEEE**, IEEE, v. 77, n. 4, p. 541–580, 1989.
- NI, J.; BAI, X. A review of air conditioning energy performance in data centers. **Renewable and sustainable energy reviews**, Elsevier, v. 67, p. 625–640, 2017.
- NINA, B. L. F. A menor distância entre dois pontos. 2011.
- OTANI, M.; MACHADO, W. V. A proposta de desenvolvimento de gestão da manutenção industrial na busca da excelência ou classe mundial. **Revista Gestão Industrial**, v. 4, n. 2, 2008.
- PINTO, A. K.; XAVIER, J. A. N. **Manutenção: função estratégica**. [S.l.]: Qualitymark, 2009.
- _____. **Manutenção: função estratégica**. [S.l.]: Qualitymark, 2012.
- REISIG, W. **Petri nets: an introduction**. [S.l.]: Springer Science & Business Media, 2012. v. 4.
- RIBEIRO, J. B.; GOMES, M. C. Manutenção preditiva em motores de combustão interna por análise do óleo lubrificante. Arruda, Alberto Luiz de Luna, 2016.
- SAID, U.; TAGHIPOUR, S. Modeling failure and maintenance effects of a system subject to multiple preventive maintenance types. In: IEEE. **Reliability and Maintainability Symposium (RAMS), 2016 Annual**. [S.l.], 2016. p. 1–7.
- SANKARAN, G. C.; SIVALINGAM, K. M. Design and analysis of scheduling algorithms for optically groomed data center networks. **IEEE/ACM Transactions on Networking**, IEEE, v. 25, n. 6, p. 3282–3293, 2017.
- SEIXAS, E. de S.; BRASIL, R. Determinação do intervalo ótimo para manutenção: Preventiva, preditiva e detectiva. 2011.

- SILVA, A. N. da; LINS, F. A.; JÚNIOR, J. C.; ROSA, N. S.; QUENTAL, N. C.; MACIEL, P. R. Avaliação de desempenho da composição de web services usando redes de petri. In: **Brazilian Symposium on Computer Networks. Curitiba, Paraná, Brazil.** [S.l.: s.n.], 2006.
- SILVA, B.; CALLOU, G.; TAVARES, E.; MACIEL, P.; FIGUEIREDO, J.; SOUSA, E.; ARAUJO, C.; MAGNANI, F.; NEVES, F. Astro: An integrated environment for dependability and sustainability evaluation. **Sustainable computing: informatics and systems**, Elsevier, v. 3, n. 1, p. 1–17, 2013.
- SILVA, B.; MACIEL, P.; BRILHANTE, J.; ZIMMERMANN, A. Geoclouds modes: A performance evaluation tool for disaster tolerant iaas clouds. In: **IEEE. Systems Conference (SysCon), 2014 8th Annual IEEE.** [S.l.], 2014. p. 116–122.
- SILVA, B.; MATOS, R.; CALLOU, G.; FIGUEIREDO, J.; OLIVEIRA, D.; FERREIRA, J.; DANTAS, J.; JUNIOR, A.; ALVES, V.; MACIEL, P. 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.** [S.l.: s.n.], 2015.
- SOUSA, E. d. C. Os benefícios do investimento na capacitação profissional da equipe de manutenção e operação de data center. **Datacenter: projeto, operação e serviços- Unisul Virtual**, 2017.
- STURM, R.; MORRIS, W.; JANDER, M. **Foundations of service level management.** [S.l.]: Sams Indianapolis, IN, 2000. v. 13.
- TORELL, W.; AVELAR, V. Tempo médio entre falhas: Explicação e padrões. **W. Kingston, Rhode Island: American Power Conversion**, 2004.
- VERDI, F. L.; ROTHENBERG, C. E.; PASQUINI, R.; MAGALHÃES, M. Novas arquiteturas de data center para cloud computing. **Minicursos do XXVIII SBRC**, p. 103–152, 2010.
- VIANA, H. R. G. **PCM-Planejamento e Controle da Manutenção.** [S.l.]: Qualitymark Editora Ltda, 2002.
- WANG, W.; XU, Y.; FAN, B.; XIONG, J. On multi-state system with interval-valued states under preventive maintenance and minimal repairs. In: **IEEE. Reliability Systems Engineering (ICRSE), 2017 Second International Conference on.** [S.l.], 2017. p. 1–7.
- WERNKE, R. Gestão de custos: uma abordagem prática. são paulo: Atlas, 2001. __. **Gestão de custos: uma abordagem prática**, v. 2, 2016.
- XENOS, H. G. Gerenciando a manutenção produtiva. **Belo Horizonte: Editora de desenvolvimento gerencial**, v. 171, 1998.
- YEGANEH, H.; SALAHI, A.; POURMINA, M. A. A novel cost optimization method for mobile cloud computing by capacity planning of green data center with dynamic pricing. **Canadian Journal of Electrical and Computer Engineering**, IEEE, v. 42, n. 1, p. 41–51, 2019.

ZHANG, Y.; ANSARI, N. Hero: Hierarchical energy optimization for data center networks. **IEEE Systems Journal**, IEEE, v. 9, n. 2, p. 406–415, 2015.

ZIMMERMANN, A.; KNOKE, M.; HUCK, A.; HOMMEL, G. Towards version 4.0 of timenet. In: VDE. **13th GI/ITG Conference-Measuring, Modelling and Evaluation of Computer and Communication Systems**. [S.l.], 2006. p. 1–4.