



Avaliação da Disponibilidade de Infraestrutura de Sincronização de Dados

Aluno: Carlos Alexandre Silva de Melo
Orientador: Paulo Romero Martins Maciel
e-mail: {casm3,prmm}@cin.ufpe.br



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO



WMoDCS 2016.2

CIn.ufpe.br



Agenda

- Introdução;
- Objetivos;
- Metodologia para Modelagem e avaliação de disponibilidade;
- Estudos de Caso;
- Conclusão e Trabalhos Futuros.



INTRODUÇÃO



Introdução

- Sincronização de dados trata-se da manutenção da consistência de um conjunto de dados distribuídos entre vários dispositivos;
- Com a popularidade adquirida por aplicações como Google Drive, Dropbox e One Drive, a sincronização de dados passou a ser popular também entre os usuários de sistemas computacionais;
- Graças a disseminação dos dispositivos móveis o número de usuários de sistemas computacionais apenas cresce;
- Computação em Nuvem é um dos paradigmas responsáveis pela evolução dos serviços disponibilizados a usuários.



Objetivo Geral

Proposição de modelos baseados em estados para avaliação da disponibilidade de infraestruturas de sincronização de dados hospedadas em plataforma privada de computação em nuvem.



Objetivos Específicos

- Implantação de serviço de sincronização de dados em uma infraestrutura privada de computação em nuvem;
- Elaboração de modelos que representem os modos operacionais da infraestrutura e do serviço de sincronização de dados;
- Validação de componente através da realização de experimentos;
- Aplicação de técnica de análise de sensibilidade para a identificação dos componentes de maior impacto na disponibilidade;

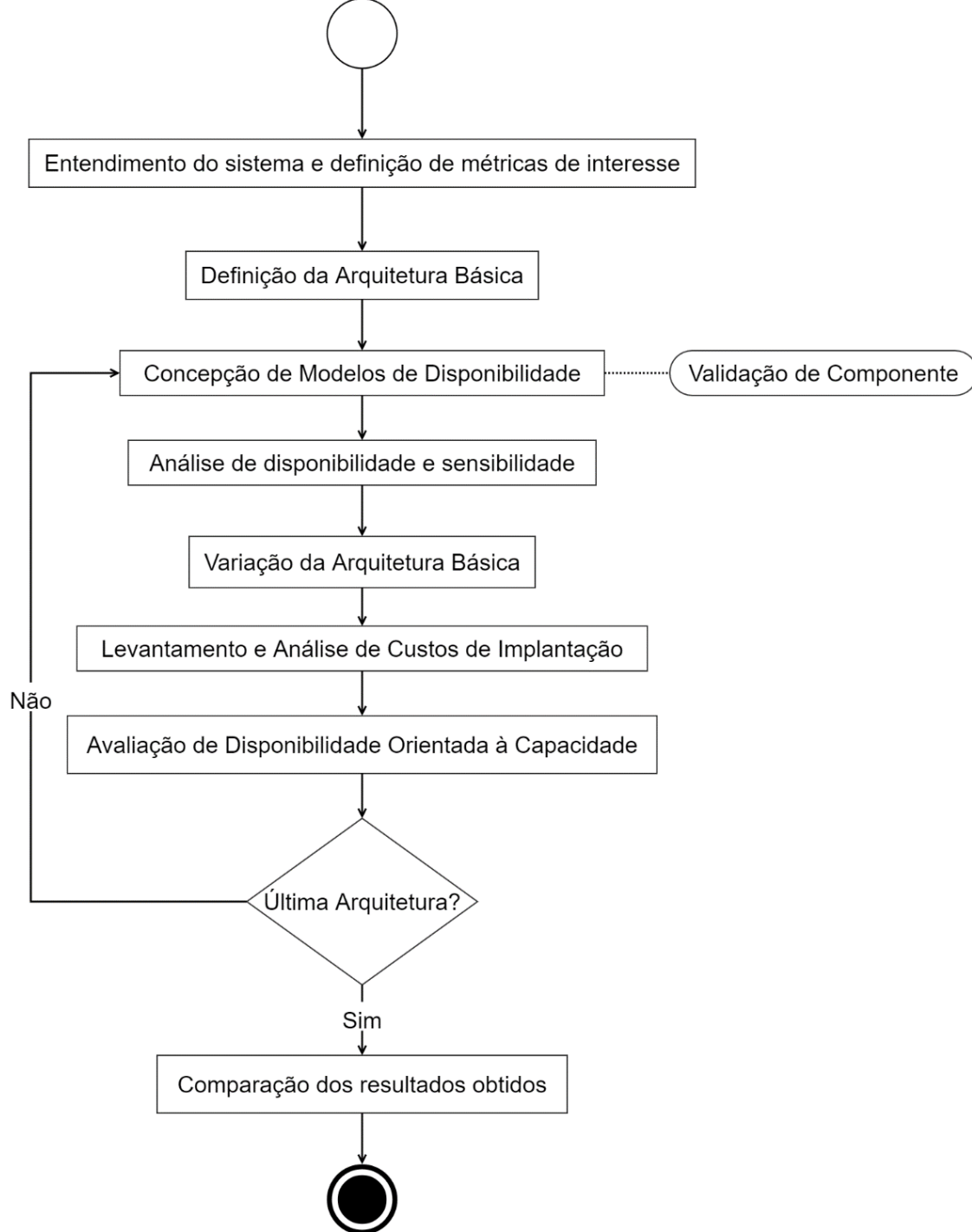


METODOLOGIA PARA MODELAGEM E AVALIAÇÃO DE DISPONIBILIDADE



Metodologia

Figura 3.1: Metodologia Adotada





Arquitetura Básica

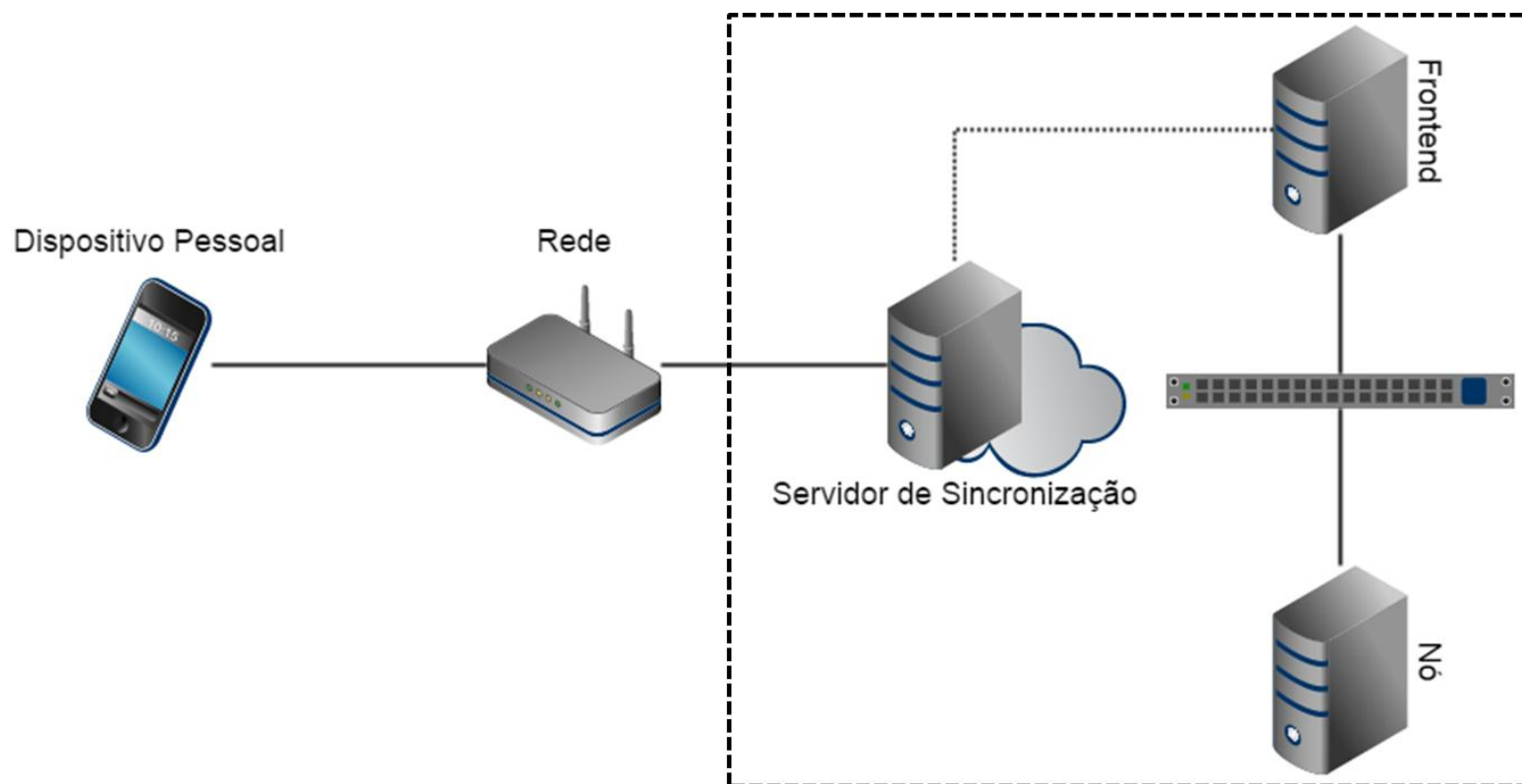
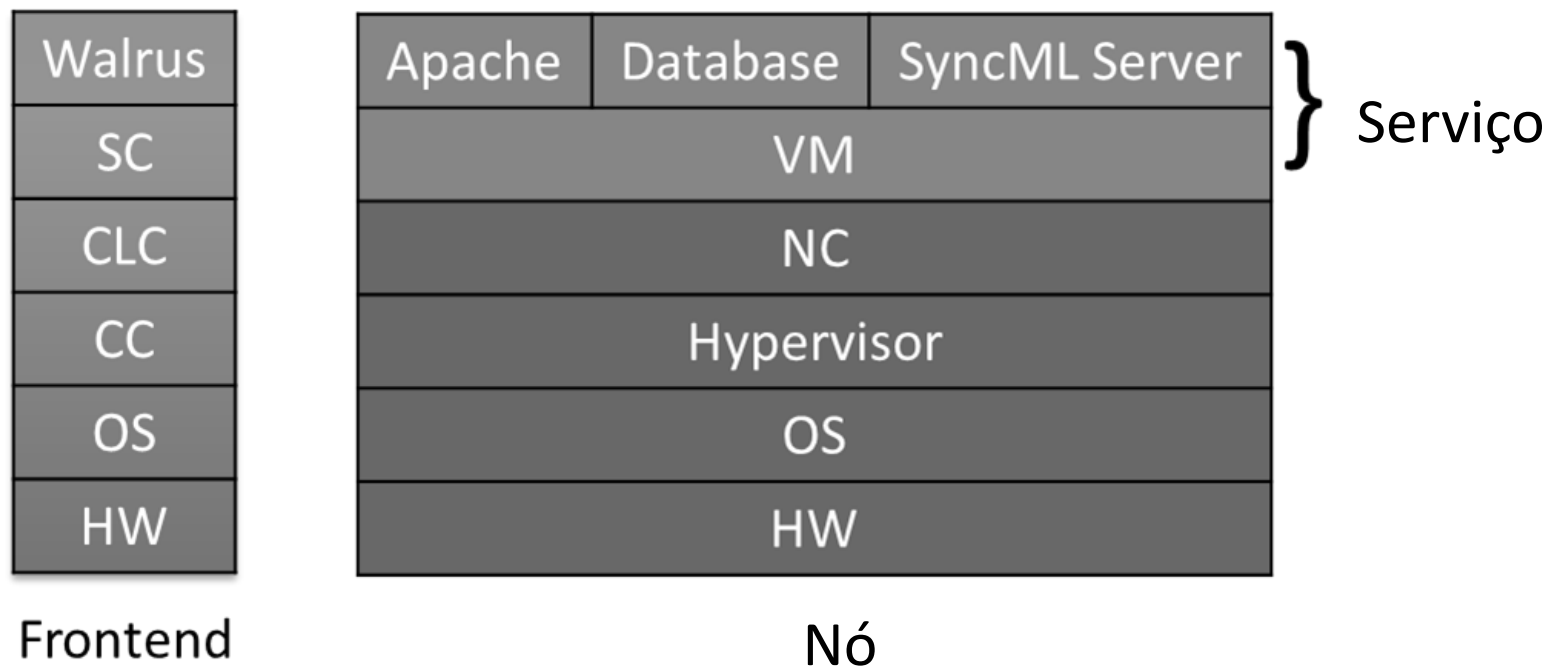


Figura 3.2: Visão de Alto nível da Arquitetura Básica



Arquitetura Básica



$$MO_{servico} = frontend \wedge nó \wedge servico$$



Modelos

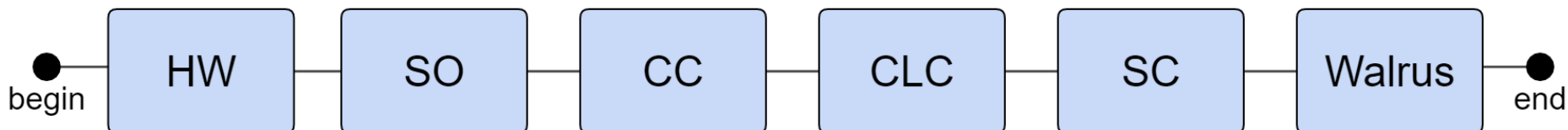


Figura 3.4: RBD: Subsistema do frontend

$$A_{Frontend} = A_{HW} \times A_{OS} \times A_{CC} \times A_{CLC} \times A_{SC} \times A_{Walrus}$$

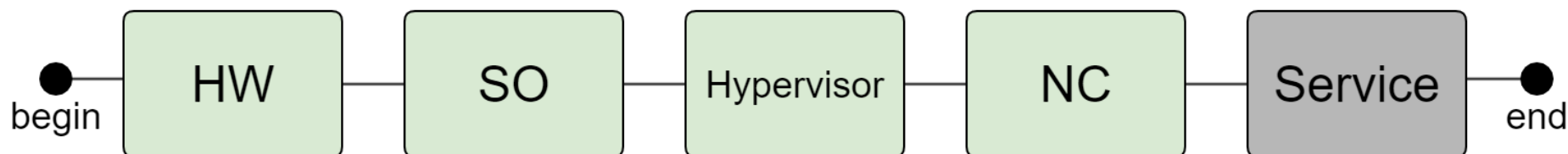


Figura 3.5: RBD: Subsistema do nó

$$A_{Node} = A_{HW} \times A_{SO} \times A_{Hypervisor} \times A_{NC} \times A_{\text{serviço}}$$



Modelos

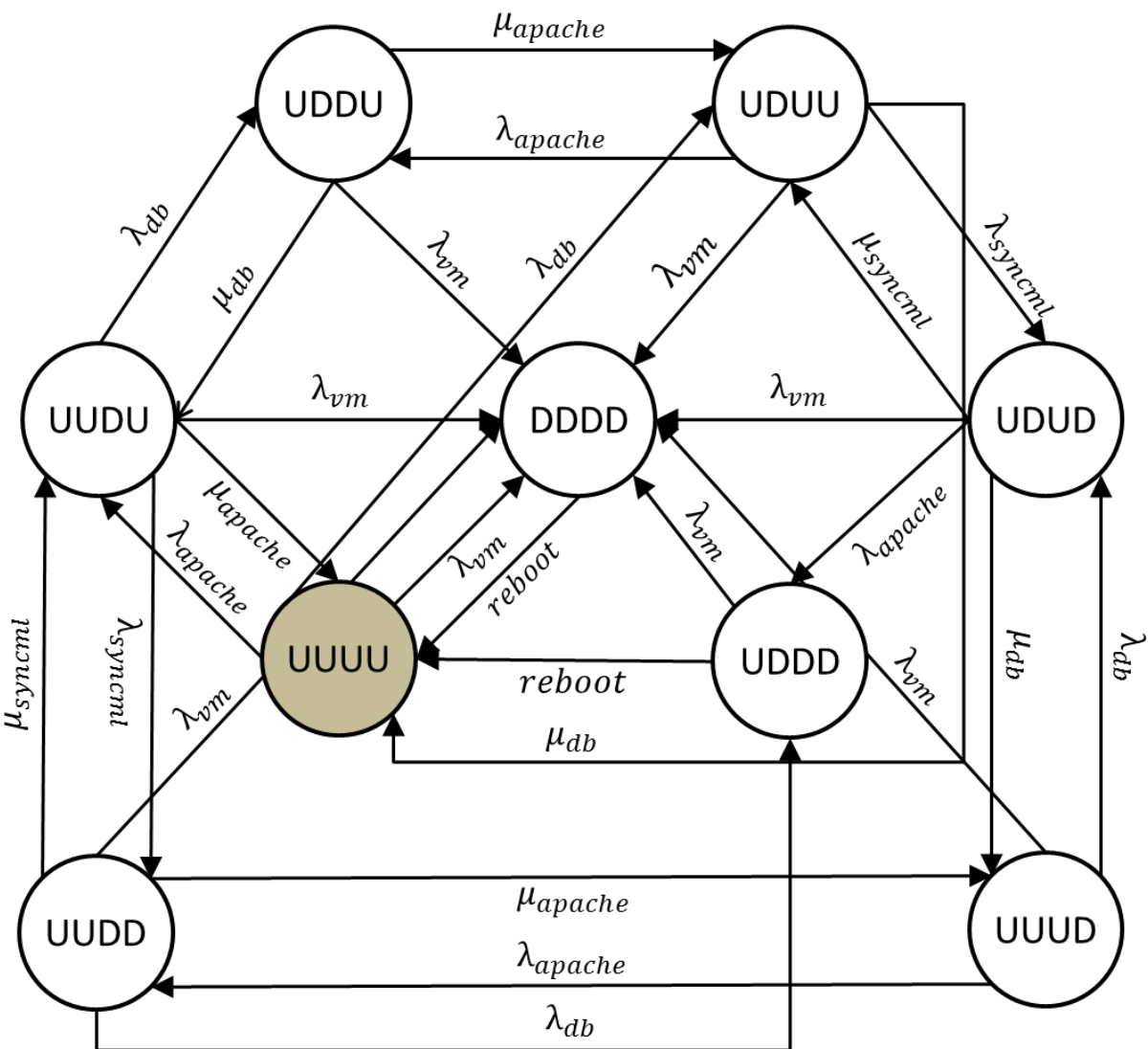


Tabela 3.2: Taxas de Falha e Reparo do Serviço - CTMC

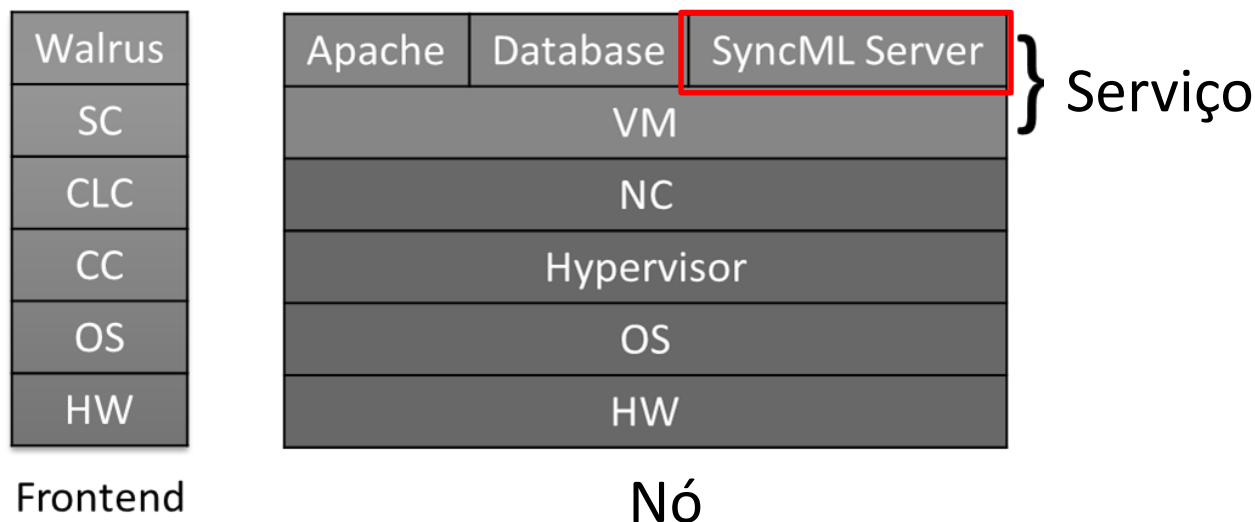
Taxa	Significado
λ_{syncml}	Falha do SyncML
μ_{syncml}	Reparo do SyncML
λ_{apache}	Falha do Apache
μ_{apache}	Reparo do Apache
λ_{db}	Falha do Banco de Dados
μ_{db}	Reparo do Banco de Dados
λ_{vm}	Falha da VM
reboot	Taxa para reparo da VM com os serviços

Figura 3.6: CTMC: Subsistema do Serviço



Validação do servidor de sincronização

- *Uma infraestrutura de testes foi montada;*
- *Utilizamos do método proposto em EJLALI et al. (2003) para provocar falhas e promover reparos de forma intencional desse componente em nossa infraestrutura de testes.*



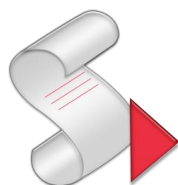


Validação do servidor de sincronização

- *Em virtude da quantidade de tempo necessária para a ocorrência de falhas, técnicas conhecidas como "fatores de redução" ARAUJO et al. (2011) foram utilizadas em nossos scripts.*

Tabela 4.3: Fator de redução de MTTF

Componente	Falha Real	Falha com redução	Reparo
SyncML	788,4 h	7,88 h	1 h



- *Scripts de Falha e Reparo*
- *Scripts de Monitoramento*



Validação do servidor de sincronização

- *Um script de monitoramento armazenou em log o número de ocorrências de falha e reparo.*

Tabela 4.4: Número de Ocorrências

Descrição	Quantidade
SyncML Falha	17
SyncML Reparos	17
Total	34



Validação do servidor de sincronização

- *Logo após, os cálculos estatísticos propostos em KEESE (1965) foram aplicados, com base no número de ocorrências e no tempo para cada uma delas fomos capazes de encontrar os valores mínimos e máximos para distribuição F.*

Tabela 4.5: Intervalos para distribuição-F

Descrição	Valor
Valor Mínimo	0,5948
Valor Máximo	1,981



Validação do servidor de sincronização

- O valor encontrado (0,917664329) está com uma certeza de 95% dentro do intervalo de confiança delimitado pelo A_U (maior valor de disponibilidade) e A_L (menor valor de disponibilidade).*

Tabela 4.6: Intervalo de confiança para A e ρ

Intervalo de Confiança de 95%		
ρ	ρ_U	0,045292
	ρ_L	0,17774
A	A_U	0,956671
	A_L	0,849084

$$A_L = \frac{1}{1 + \rho_L} \quad (4.1)$$

$$A_U = \frac{1}{1 + \rho_U} \quad (4.2)$$



Modelo para COA

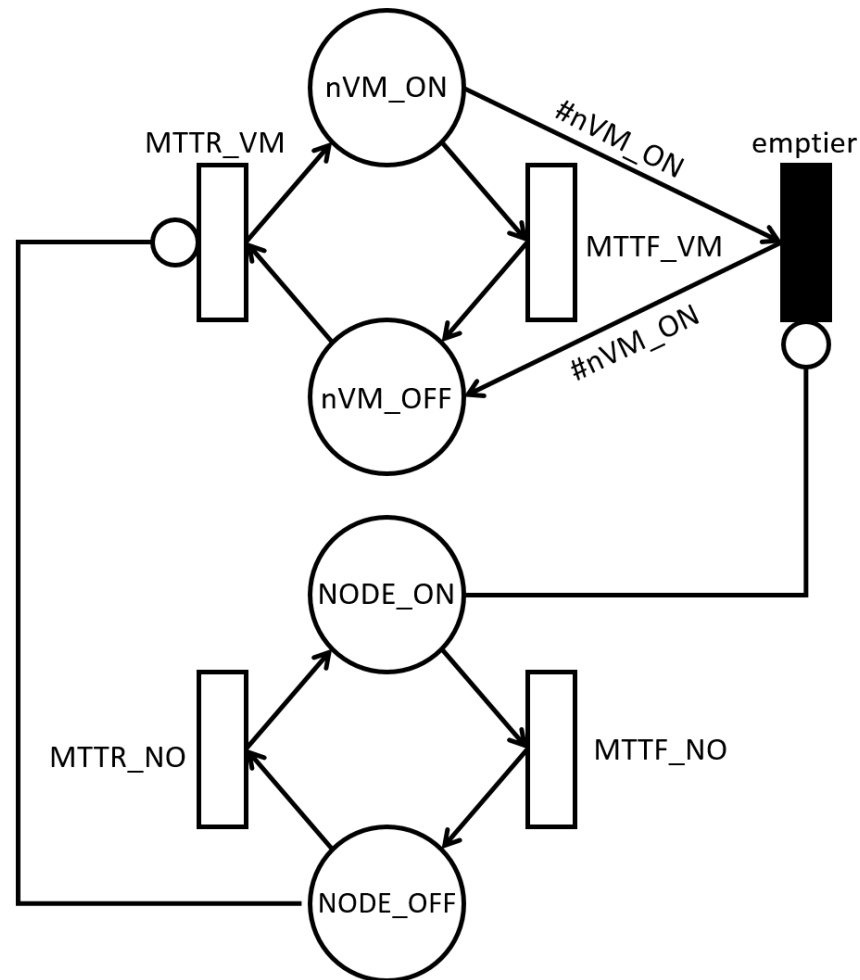


Figura 3.12: Modelo de Disponibilidade Orientada à Capacidade



ESTUDOS DE CASO



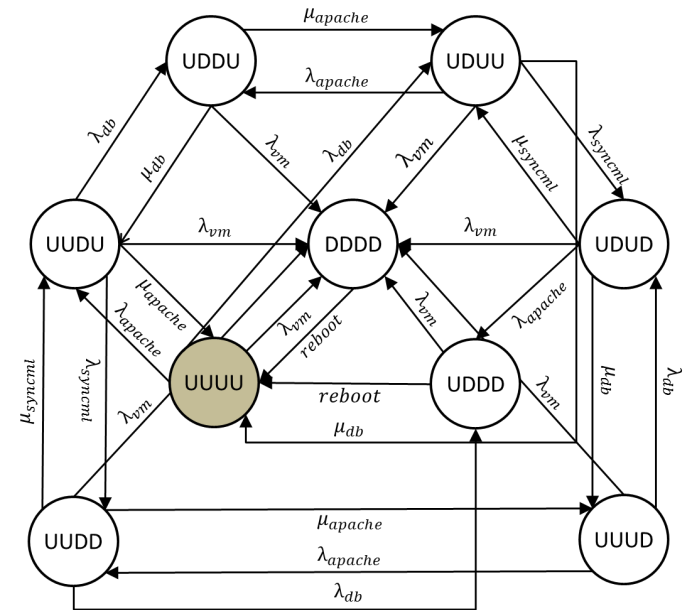
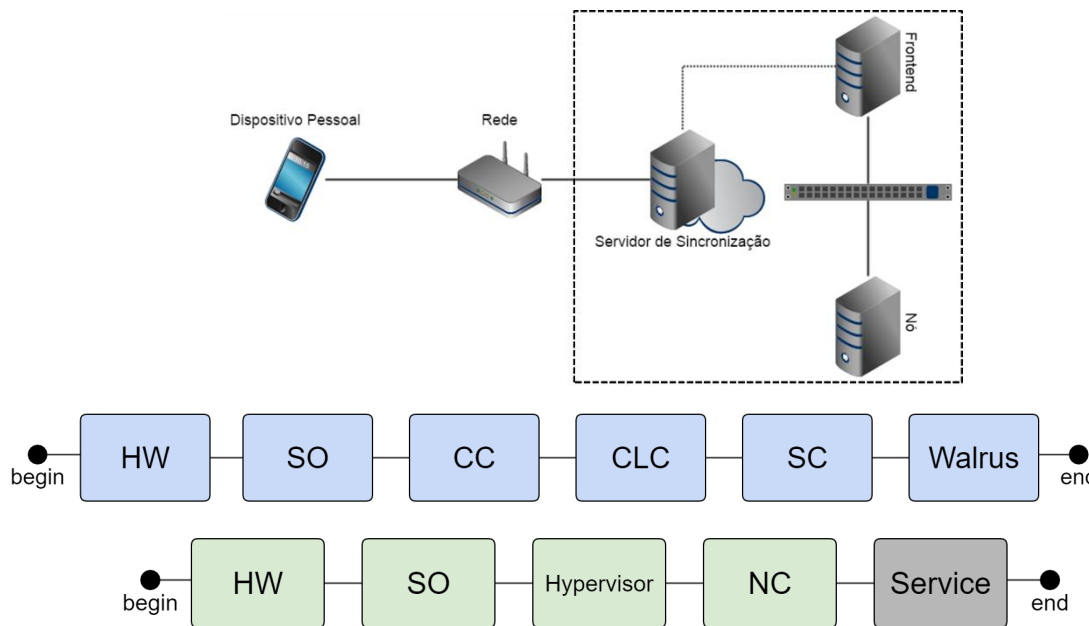
Estudos de caso

- *Estudo de Caso I – Avaliação dos modelos de disponibilidade da arquitetura básica*
 - *Análise de sensibilidade;*
- *Estudo de Caso II – Avaliação de disponibilidade orientada a capacidade*



Estudo de Caso I – Avaliação da Disponibilidade da arquitetura básica

- Para a avaliação dos modelos baseados em estado apresentados, e que representam o funcionamento da arquitetura básica, alguns valores de entrada tornaram-se necessários.





Estudo de Caso I – Valores de Entrada para avaliação

Tabela 4.1: Parâmetros de Entrada para os modelos

Componentes	MTTF	MTTR
HW	8760 h	100 min
SO	2893 h	15 min
CLC	788,4 h	1 h
CC	788,4 h	1 h
SC	788,4 h	1 h
Walrus	788,4 h	1 h
Hypervisor	2990 h	1 h
NC	788,4 h	1 h
SyncML	788,4 h	1 h
Apache	788,4 h	1 h
Database	1440 h	3 h
VM	2880 h	69 s



Estudo de Caso I – Avaliação da Disponibilidade

- Com base nos parâmetros de entrada, pudemos extrair como saída valores relacionados à disponibilidade do RBD que representa frontend e nó, e da CTMC do serviço.*

Tabela 4.2: Resultados da Avaliação de Disponibilidade para os modelos da arquitetura básica

Modelo Métrica	Frontend	Nó	Serviço
Disponibilidade (%)	99,46	99,81	99,53
Disponibilidade (Número de Noves)	2,27	2,72	2,33
<i>Uptime</i> anual (h)	8713,22	8743,53	8719,63
<i>Downtime</i> anual (h)	46,78	16,47	40,37



Estudo de Caso I – Análise de sensibilidade da arquitetura básica

- *A técnica de diferença percentual foi utilizada para diagnosticar qual/quais componentes possuem maior impacto na métrica de interesse.*

$$SI = \left(\frac{D_{max} - D_{min}}{D_{max}} \right), \quad (2.17)$$



Estudo de Caso I – Parâmetros de entrada

Tabela 4.7: Parâmetros de entrada para análise de sensibilidade das arquiteturas propostas

Parâmetros	Valores
$MTTF_{NC}, MTTF_{walrus}, MTTF_{SC}, MTTF_{CC}, MTTF_{CLC}$ (h)	788,4
$MTTF_{hypervisor}$ (h)	2990
$\lambda_{syncml}, \lambda_{apache}$ (h^{-1})	1/788.4
λ_{db} (h^{-1})	1/1440
μ_{db} (h^{-1})	1/3
$reboot$ (h^{-1})	1/0.019166
λ_{VM} (h^{-1})	1/2880
$MTTR_{NC}, MTTR_{walrus}, MTTR_{SC}, MTTR_{CC}, \mu_{syncml}, \mu_{apache}$ (h)	1
$MTTR_{SO}$ (h)	0,25
$MTTF_{SO}$ (h)	2893
$MTTR_{HW}$ (h)	1,66
$MTTF_{HW}$ (h)	8760



Estudo de Caso I - Ranking de sensibilidade para a arquitetura básica

Tabela 4.8: Ranking de sensibilidade para a Arquitetura Básica

Componente	Posição	Valor
λ_{DB}	1°	$1,05 \times 10^{-2}$
μ_{DB}	2°	$8,47 \times 10^{-3}$
$MTTF_{NC}, MTTF_{walrus}, MTTF_{CC}, MTTF_{SC}, MTTF_{CLC}$	3°	$1,69 \times 10^{-3}$
$\lambda_{apache}, \lambda_{syncml}$	7°	$1,50 \times 10^{-3}$
$\mu_{apache}, \mu_{syncml}$	9°	$1,26 \times 10^{-3}$
$MTTR_{NC}, MTTR_{walrus}, MTTR_{CLC}, MTTR_{SC}, MTTR_{CC}$	11°	$1,13 \times 10^{-3}$
$MTTF_{hypervisor}$	16°	$4,46 \times 10^{-4}$
$MTTR_{hypervisor}$	17°	$2,97 \times 10^{-4}$
$MTTF_{HW_{frontend}}, MTTF_{HW_{no}}$	18°	$2,42 \times 10^{-4}$
$MTTF_{SO_{frontend}}, MTTF_{SO_{no}}$	20°	$1,15 \times 10^{-4}$
$MTTR_{HW_{frontend}}, MTTR_{HW_{no}}$	22°	$1,01 \times 10^{-4}$
$MTTR_{SO_{frontend}}, MTTR_{SO_{no}}$	24°	$7,68 \times 10^{-5}$
λ_{vm}	25°	$4,05 \times 10^{-5}$
reboot	26°	$6,66 \times 10^{-6}$



Estudo de Caso I - Resultados da análise de sensibilidade para a arquitetura básica

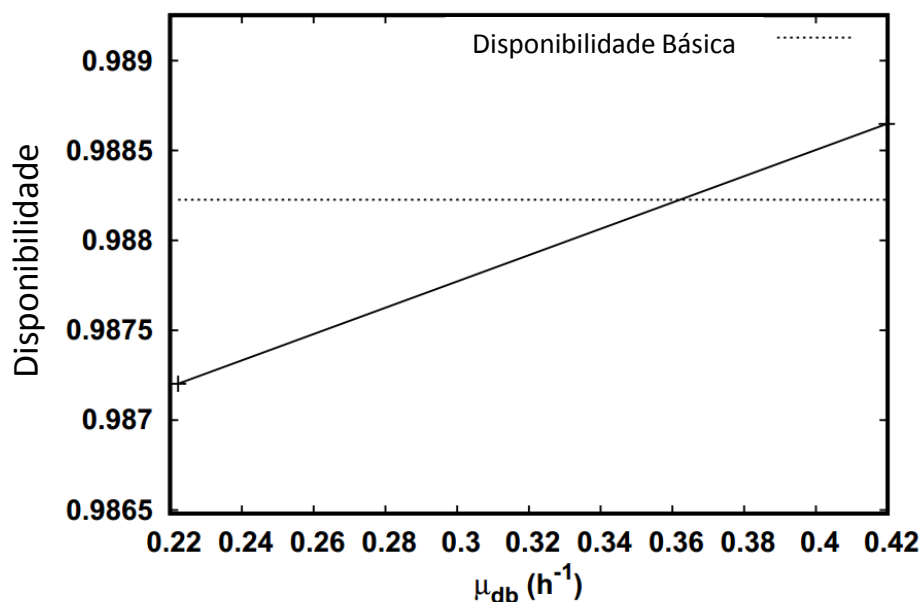


Figura 4.2: Variação da taxa de reparo μ_{DB}

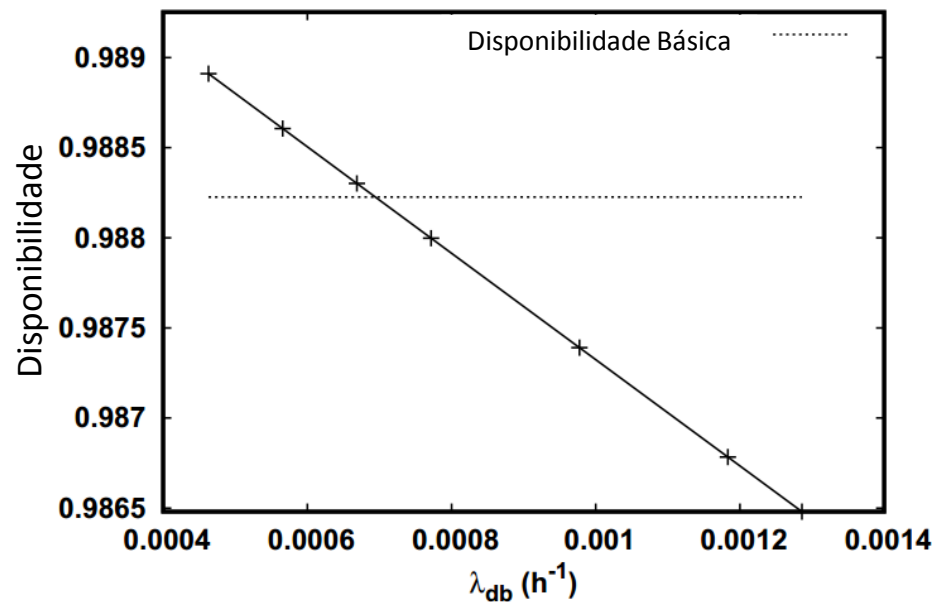
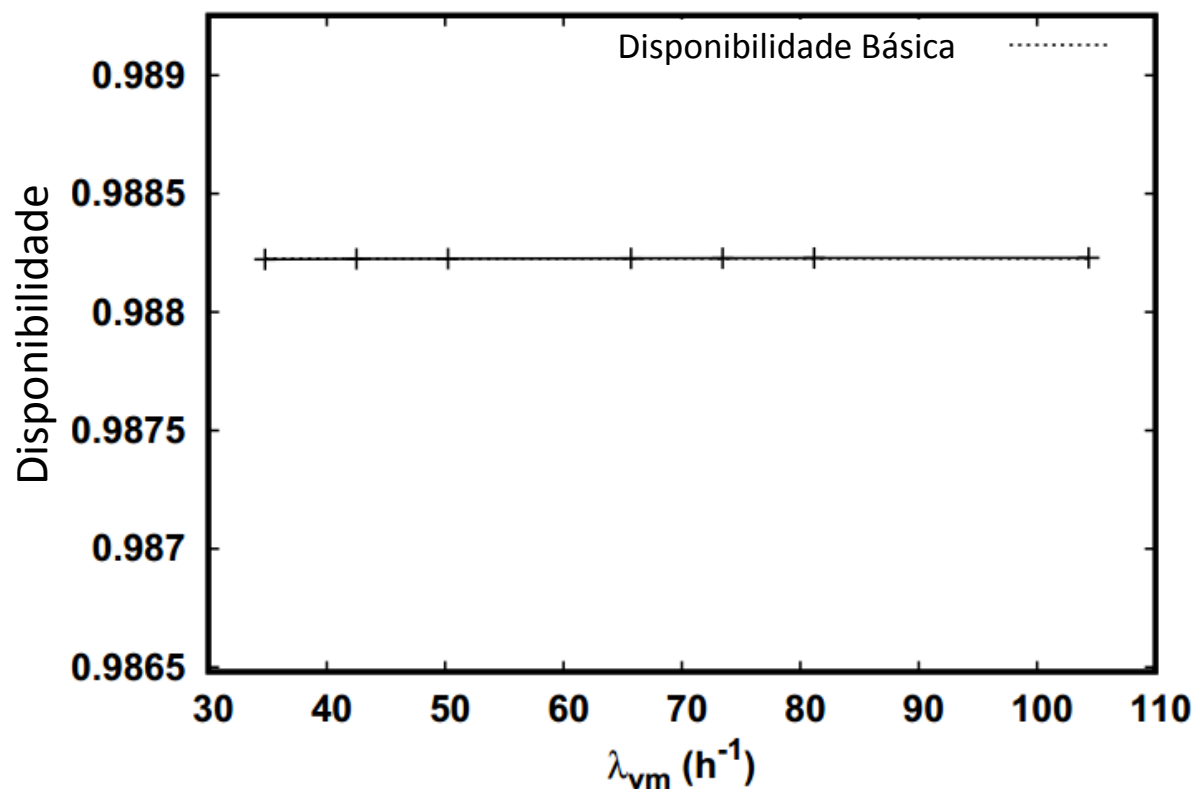


Figura 4.1: Variação da taxa de falha λ_{db}



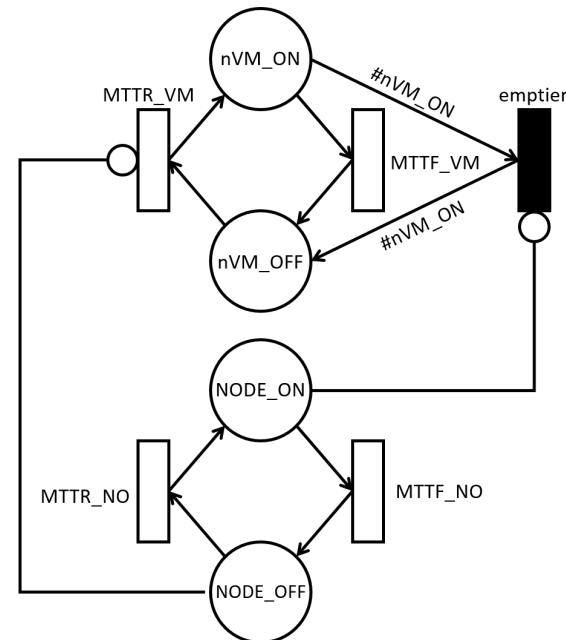
Estudo de Caso I - Resultados da análise de sensibilidade para a arquitetura básica





Estudo de Caso II – Avaliação da COA

- *A primeira etapa, antes de se definir qual a disponibilidade orientada à capacidade do sistema é definir o que vem a ser nossa capacidade;*
- *Este estudo de caso avalia o número de máquinas virtuais realmente disponível aos usuários*





Estudo de Caso II – Tipos de VM's criáveis

Tabela 4.12: Tipos e Configurações de VM criáveis com Eucalyptus

Tipos de VM	CPU's Utilizáveis	Disco (GB)	RAM (MB)
m1.small	1	5	256
m1.medium	1	10	512
m1.large	2	10	512
m1.xlarge	2	10	1024
m2.xlarge	4	15	2048
m3.2xlarge	4	30	4096
c1.medium	2	10	512
c1.xlarge	2	10	2048
cc1.4xlarge	8	60	3072
cc2.8xlarge	16	120	6144
m2.xlarge	2	10	2048
m2.2xlarge	2	30	4096
m2.4xlarge	8	60	4096
cr1.8xlarge	16	240	16384
t1.micro	1	5	256



Estudo de Caso II – Tipos de VM's criáveis

- De posse de cada tipo de VM criável, pudemos calcular o número de máquinas virtuais criáveis por cada nó, tendo como computador base um Dell PowerEdge T320 com Intel Xeon E5-2420, com seis núcleos físicos e doze threads, 1TB de disco, e 24GB de RAM.



Estudo de Caso II – Número de Usuários

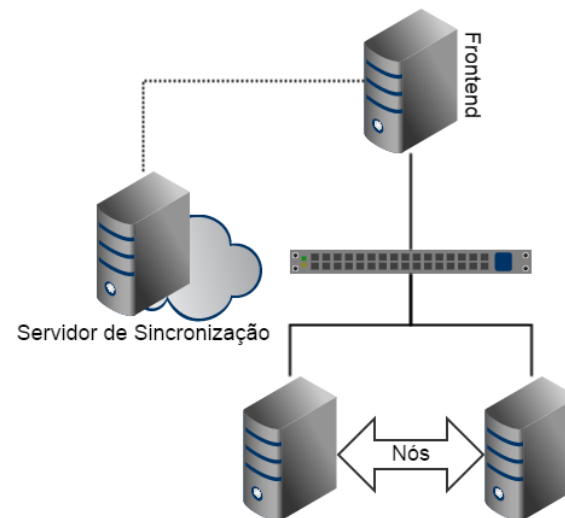
Tabela 4.13: Número de VMs e Usuários por Nó

Tipos de VM	Nº de VM's Criáveis por Nó	Nº de Usuários por VM	Nº de Usuários por Nó
m1.small	12	0	0
m1.medium	12	2,5	30
m1.large	6	2,5	15
m1.xlarge	6	2,5	15
m2.xlarge	3	5	15
m3.2xlarge	3	12,5	37,5
c1.medium	6	2,5	15
c1.xlarge	6	2,5	15
cc1.4xlarge	1,5	27,5	41,25
cc2.8xlarge	0,75	57,5	43,12
m2.xlarge	6	2,5	15
m2.2xlarge	6	12,5	75
m2.4xlarge	1,5	27,5	41,25
cr1.8xlarge	0,75	117,5	88,12
t1.micro	12	0	0



Estudo de Caso II – Resultado COA

- 99,78% com uma degradação de 0,22% no fornecimento de máquinas virtuais em cada nó da arquitetura;
- Em uma arquitetura com dois nós em redundância do tipo *ativo-ativo* onde o serviço é provido em ambos os nós ao mesmo tempo a relação entre COA e VM é dada por:





Estudo de Caso II – Resultados da COA

Tabela 4.14: Eucalyptus - COA (Disponibilidade $\times$ VM)

Tipo de VM	Número de VM por nó	Número de VMs perdidas	COA
m1.small	12	0,05	23,95
m1.medium	12	0,05	23,95
m1.large	6	0,03	11,97
m1.xlarge	6	0,03	11,97
m2.xlarge	3	0,02	5,98
m3.2xlarge	3	0,02	5,98
c1.medium	6	0,03	11,97
c1.xlarge	6	0,03	11,97
cc1.4xlarge	1,5	0,01	1,99
cc2.8xlarge	0,75	-	-
m2.xlarge	6	0,03	11,97
m2.2xlarge	6	0,03	11,97
m2.4xlarge	1,5	0,01	1,99
cr1.8xlarge	0,75	-	-
t1.micro	12	0,05	23,95



Estudo de Caso II - Resultados para análise de sensibilidade do modelo COA

Tabela 4.15: Análise de Sensibilidade para modelo de COA

Parâmetro	S(COA)
$MTTF_{NODE}$	$2,83 \times 10^{-3}$
$MTTR_{NODE}$	$1,66 \times 10^{-3}$
$MTTR_{VM}$	$2,32 \times 10^{-4}$
$MTTF_{VM}$	$8,60 \times 10^{-6}$



Estudo de Caso II - Resultados para análise de sensibilidade do modelo COA

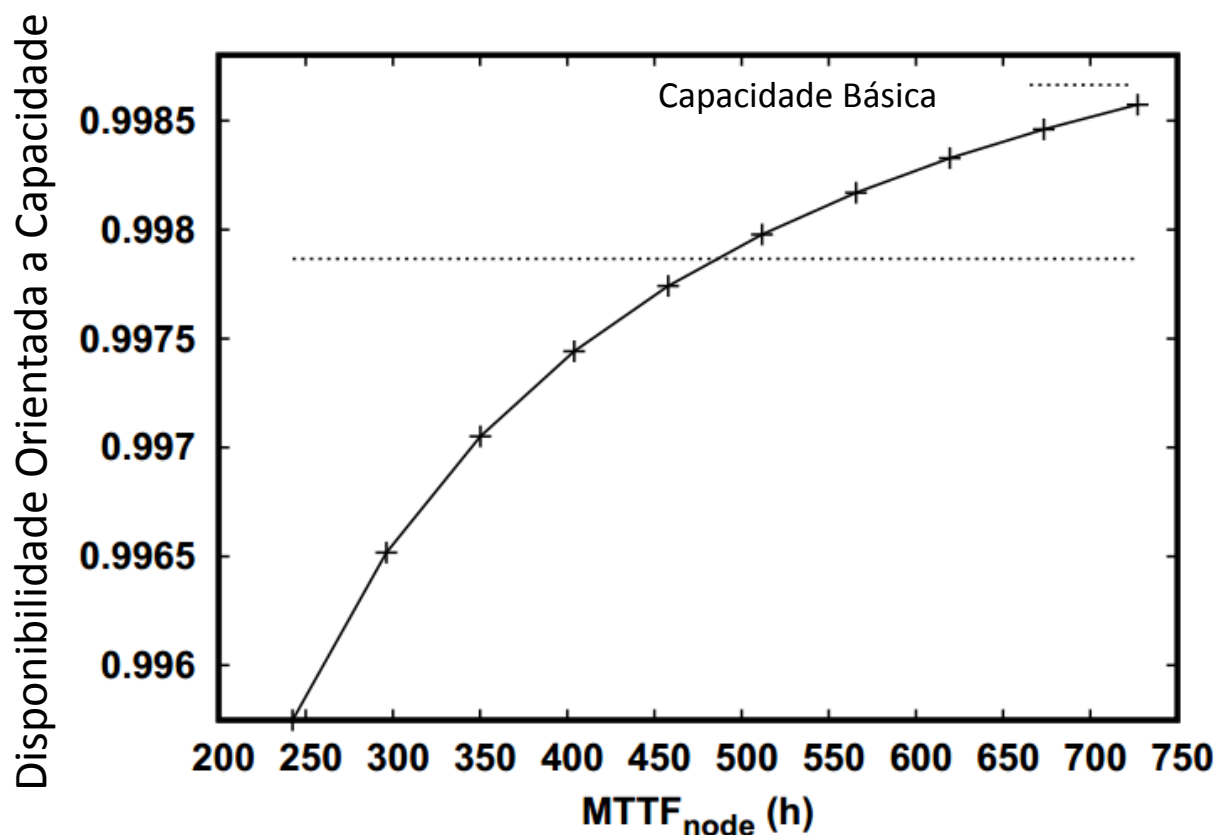


Figura 4.5: Variação do Nó



CONCLUSÃO E TRABALHOS FUTUROS



Conclusões

- Este trabalho propôs uma série de modelos que representam infraestruturas de sincronização de dados hospedadas em uma plataforma de computação em nuvem.



Contribuições

- Modelos de avaliação de disponibilidade e disponibilidade orientada à capacidade;
- Validação da aplicação responsável pela sincronização de dados;
- Aplicação de técnicas de análise de sensibilidade para estabelecer componentes de maior impacto na disponibilidade básica;



Publicações

Aceitas

Carlos Melo, Jamilson Dantas, Jean Araújo, and Paulo Maciel. **Availability Models for Synchronization Server Infrastructure**. 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC). October 9-12, 2016 – Budapest, Hungary.

Carlos Melo, Jamilson Dantas, Jean Araújo, and Paulo Maciel. **Availability Models for Synchronization Server Infrastructure**. 2016 IEEE Pacific Rim International Symposium on Dependable Computing (PRDC). January 22-25, 2017 – Christchurch, New Zealand.

Em Revisão

Carlos Melo, Rubens Matos, Jamilson Dantas and Paulo Maciel. **The Impact of Redundancy Techniques on the Availability of a Synchronization Server Infrastructure**. Elsevier Reliability Engineering & System Safety.



Trabalhos Futuros

- Avaliação de desempenho de servidor de sincronização;
- Avaliação de métricas de dependabilidade e desempenho de aplicação de sincronização do lado cliente.



OBRIGADO!