

# Identificação de gargalos de disponibilidade e desempenho em sistemas de computação em nuvem

Rubens de Souza Matos Júnior  
Orientador: Prof. Paulo Maciel

# Agenda

---

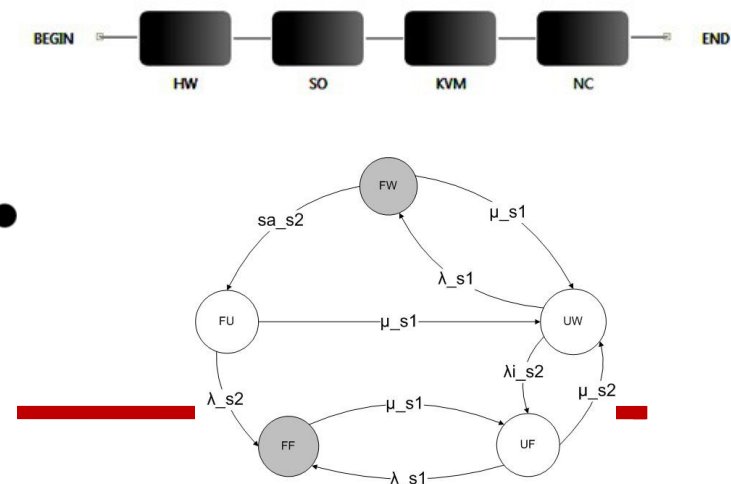
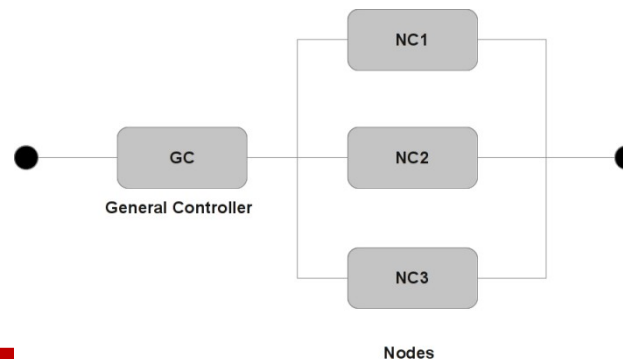
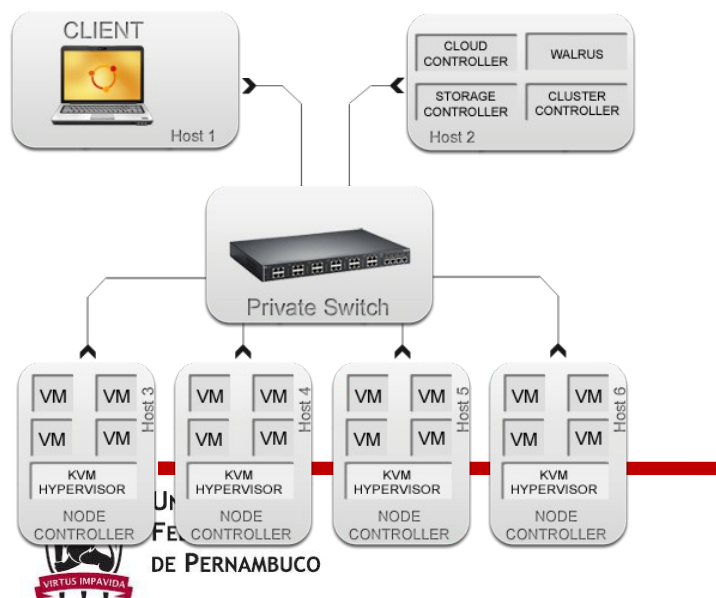


- Motivação
- Objetivos
- Metodologia proposta
- Estudos de caso
- Próximos passos



# Motivação

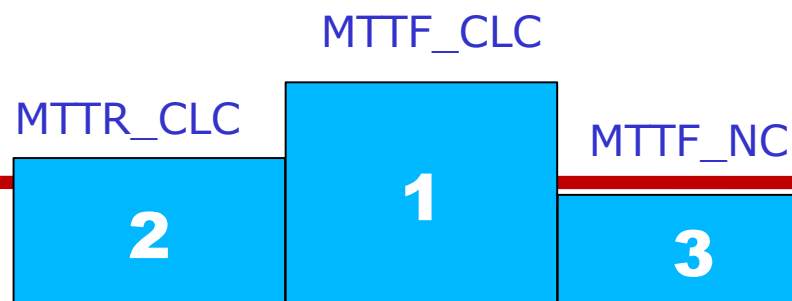
- Computação em nuvem é usada/almejada por várias empresas
- Como propor melhorias de desempenho/dependabilidade para sistemas complexos tais como as nuvens computacionais?
  - Muitos componentes de hardware e software
  - Modelos hierárquicos facilitam a descrição desses sistemas e o tratamento de largeness e stiffness.



# Motivação



- Análise de sensibilidade:
  - Métodos para verificar o quanto as mudanças nos parâmetros de entrada irão afetar as saídas (resultados) de um sistema/modelo.
  - Técnica essencial para **detecção de “gargalos”** de desempenho/dependabilidade
    - Variação dos parâmetros, um por vez
    - Análise diferencial
    - Análise de correlação/regressão
    - Design of Experiments: Full-factorial, 2k-Factorial, ...



# Objetivos específicos



- Principais objetivos:
  - Criação de **modelos** abrangendo **nível de aplicação e de infraestrutura** (Software executando em IaaS)
  - Prover **métodos e ferramentas** para **análise de sensibilidade** automatizada dos modelos hierárquicos
  - Incorporação de métodos de **A.S. em algoritmos de otimização** de infraestruturas/serviços de nuvem



# Focos do trabalho



- Disponibilidade de infraestruturas de nuvem (IaaS)
  - Nuvens privadas (Eucalyptus)
  - Mobile cloud
- Desempenho/disponibilidade de aplicações executando na nuvem
  - Servidores web com mecanismos de recuperação de desastres
  - Composições de web services
- Desempenho de alocação de VMs (auto-scaling)



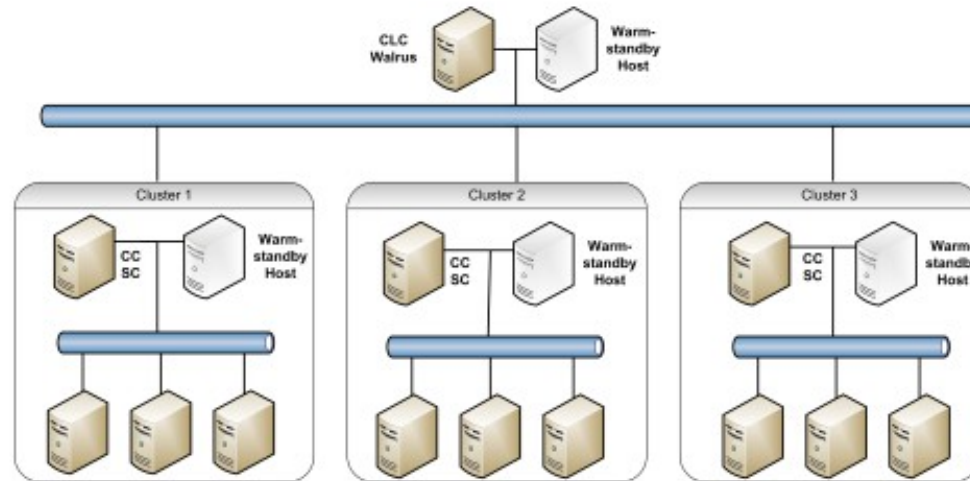
# Metodologia proposta



1. Criação de **modelo de alto nível**
2. Criação de **sub-modelos** para componentes com comportamento **detalhado** mais conhecido (hosts, VMs, aplicações)
3. **Identificação das métricas** de interesse
4. **Resolução** do modelo hierárquico
5. Caso métricas **não** sejam **satisfatórias**:
6. Análise de **sensibilidade** dos modelos de mais **baixo nível**
7. **Composição** dos índices de sensibilidade **de baixo para cima**.
8. Construção e **análise** do **ranking geral**
9. Efetua **ações prioritárias** e reavalia métricas

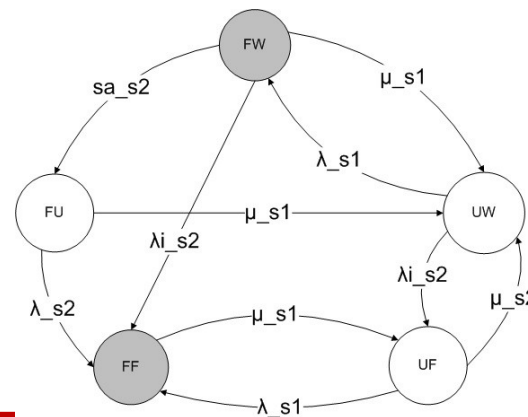
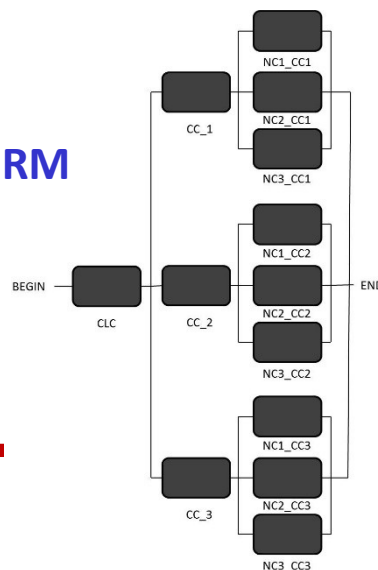


# Nuvem privada Eucalyptus com componentes redundantes



Colaboração com Jamilson e Jean

Combinação de **RBD + MRM**





# Nuvem privada Eucalyptus com componentes redundantes



Measure	Architectures		
	A1	A2	A3
Steady-state Availab.	0.999938749	<u>0.999969376</u>	<u>0.999969377</u>
Number of 9's	4.21288	4.51394	4.51395
Annual downtime	32.194 min	16.096 min	16.095 min

Praticamente nenhum ganho com 2 para 3 clusters

A1		A2		A3	
Param.	$SS_{\theta_i}(A_{Sys})$	Param.	$SS_{\theta_i}(A_{Sys})$	Param.	$SS_{\theta_i}(A_{Sys})$
$\lambda_{CLC}$	$-3.82 \times 10^{-5}$	$\lambda_{CLC}$	$-3.82 \times 10^{-5}$	$\lambda_{CLC}$	$-3.82 \times 10^{-5}$
$\lambda_{CC}$	$-3.82 \times 10^{-5}$	$\mu_{CLC}$	$2.83 \times 10^{-5}$	$\mu_{CLC}$	$2.83 \times 10^{-5}$
$sa$	$3.27 \times 10^{-5}$	$sa$	$1.64 \times 10^{-5}$	$sa$	$1.63 \times 10^{-5}$
$\mu_{CLC}$	$2.83 \times 10^{-5}$	$\lambda_{CLC_i}$	$-6.41 \times 10^{-6}$	$\lambda_{CLC_i}$	$-6.41 \times 10^{-6}$
$\mu_{CC}$	$2.82 \times 10^{-5}$	$\lambda_{CC}$	$-2.34 \times 10^{-9}$	$\lambda_{CC}$	$-1.08 \times 10^{-13}$
$\lambda_{CLC_i}$	$-6.41 \times 10^{-6}$	$\mu_{CC}$	$1.73 \times 10^{-9}$	$\mu_{CC}$	$7.96 \times 10^{-14}$
$\lambda_{CC_i}$	$-6.41 \times 10^{-6}$	$\lambda_{CC_i}$	$-3.92 \times 10^{-10}$	$\lambda_{CC_i}$	$-1.80 \times 10^{-14}$
$\lambda_{Node}$	$-2.01 \times 10^{-8}$	$\lambda_{Node}$	$-1.23 \times 10^{-12}$	$\lambda_{Node}$	$-5.66 \times 10^{-17}$
$\mu_{Node}$	$2.01 \times 10^{-8}$	$\mu_{Node}$	$1.23 \times 10^{-12}$	$\mu_{Node}$	$5.66 \times 10^{-17}$

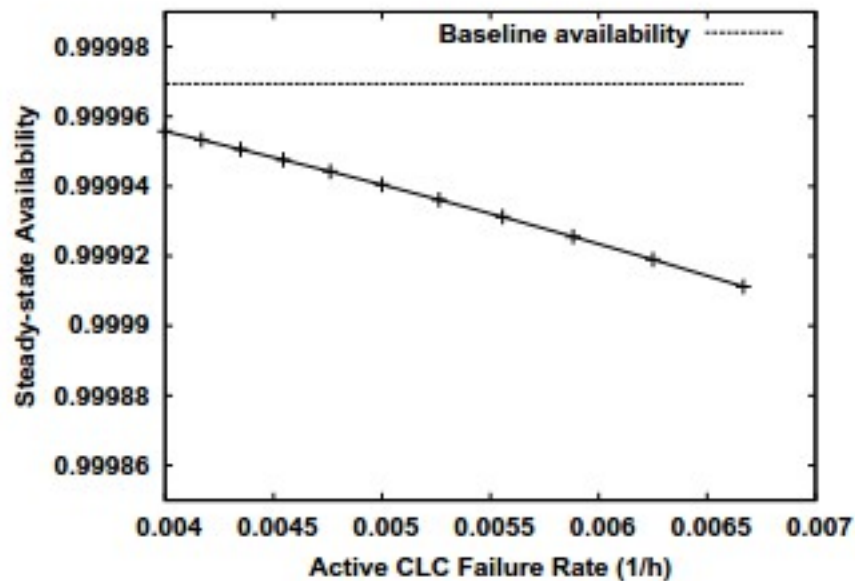
Cloud controller é o ponto de maior impacto na disponibilidade



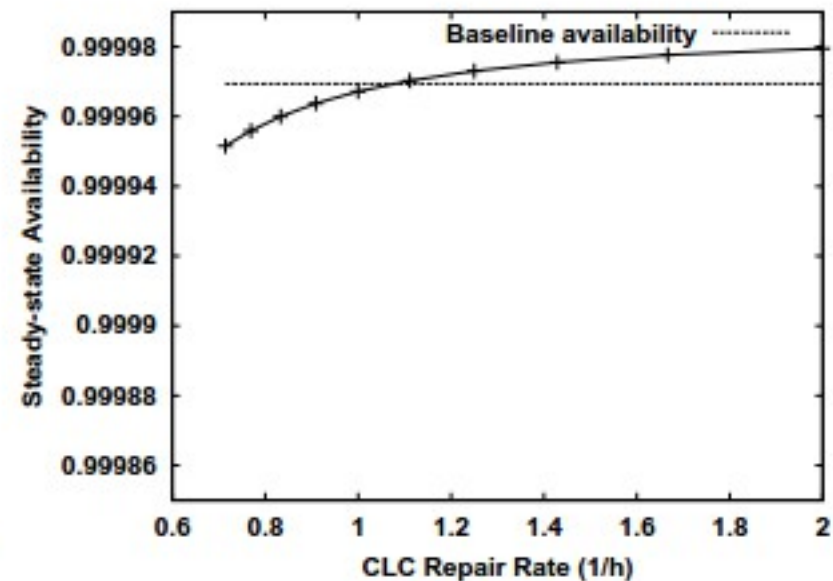
# Nuvem privada Eucalyptus com componentes redundantes



- Identificação do **CLC como ponto crítico** mesmo já possuindo redundância warm-standby



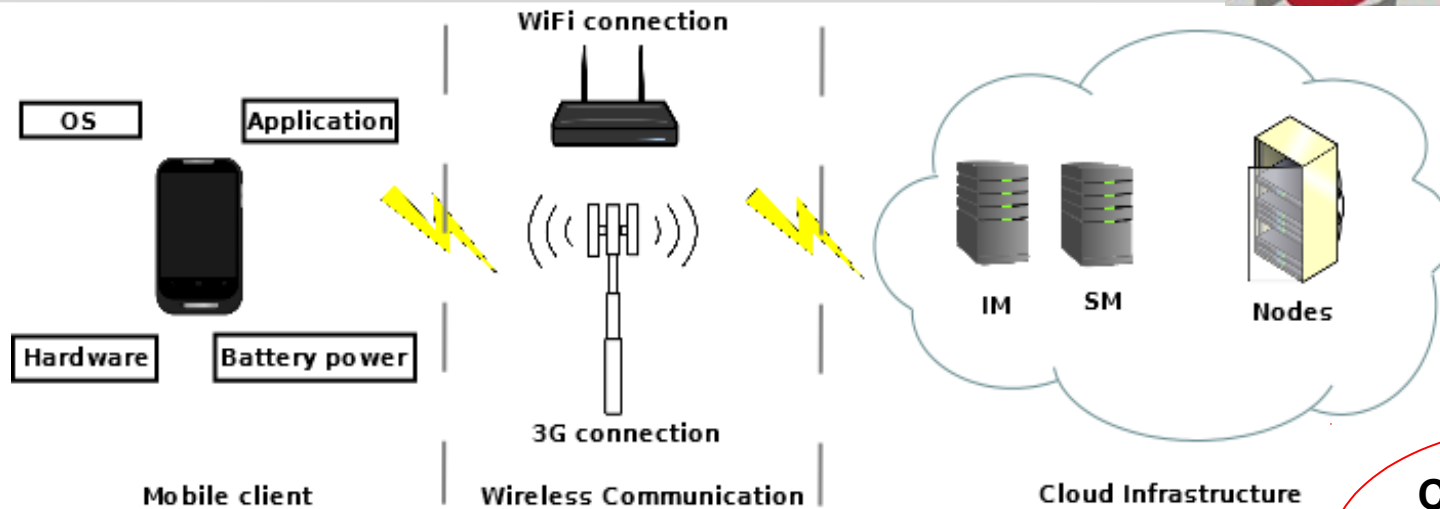
(a) Active CLC failure rate ( $\lambda_{clc}$ )



(b) CLC repair rate ( $\mu_{clc}$ )



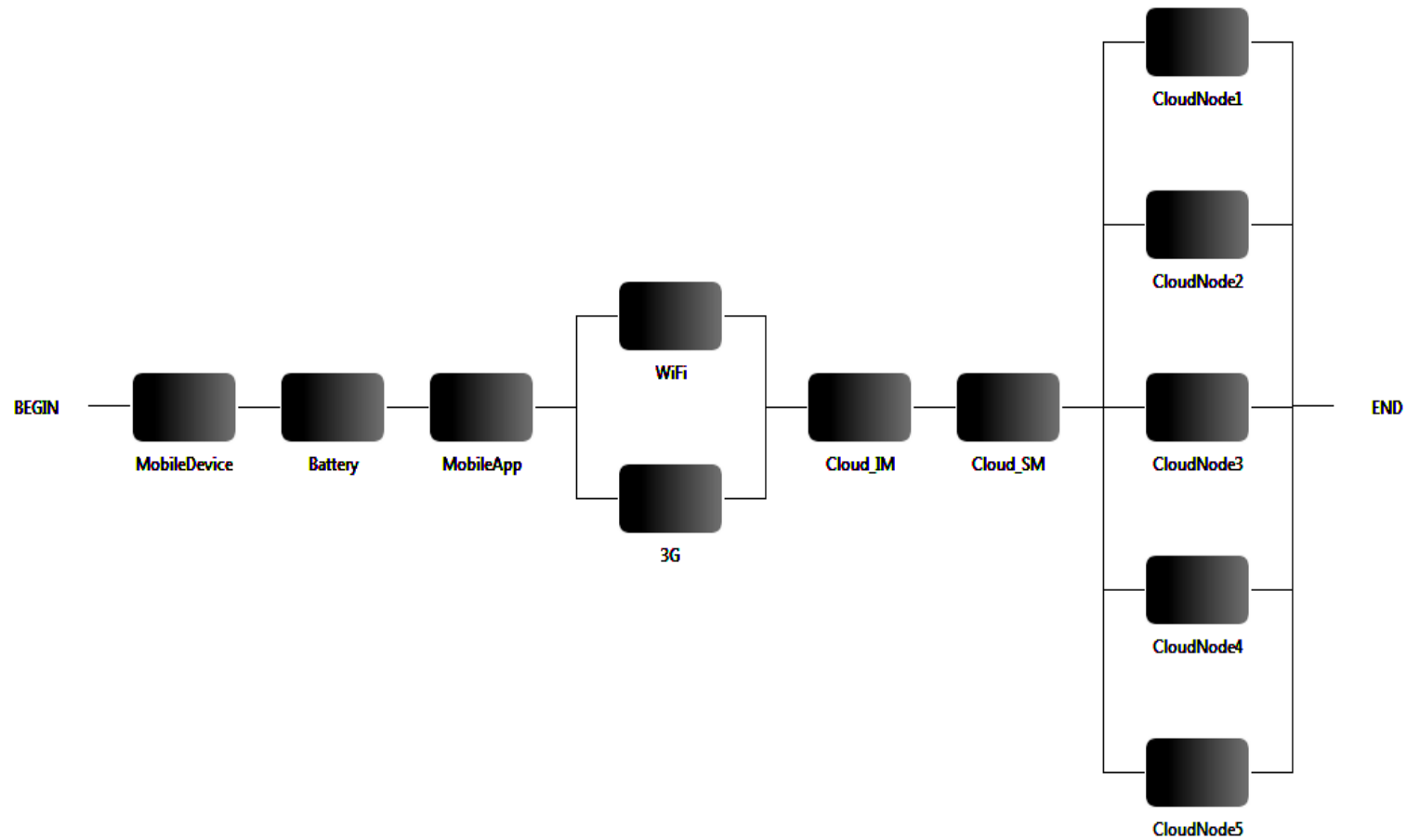
# Mobile cloud



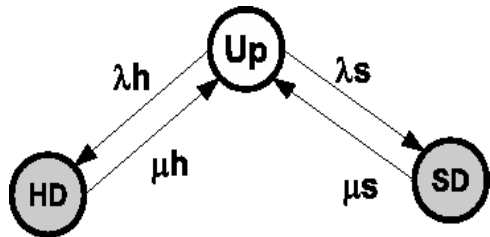
Colaboração  
com Danilo e  
Jean

- **Modelos de disponibilidade** para sistemas de mobile cloud
  - Falha/reparo de **hardware, rede e software**
  - **Descarga da bateria** do dispositivo móvel
  - Atualizações do **software**
  - **Fatores de cobertura** das falhas

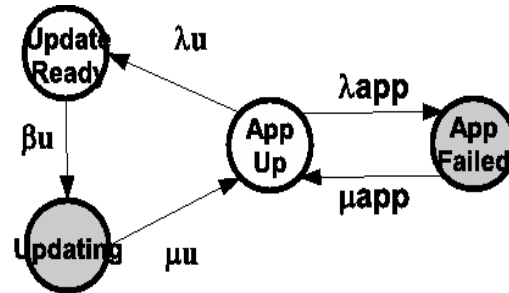
# Modelo RBD



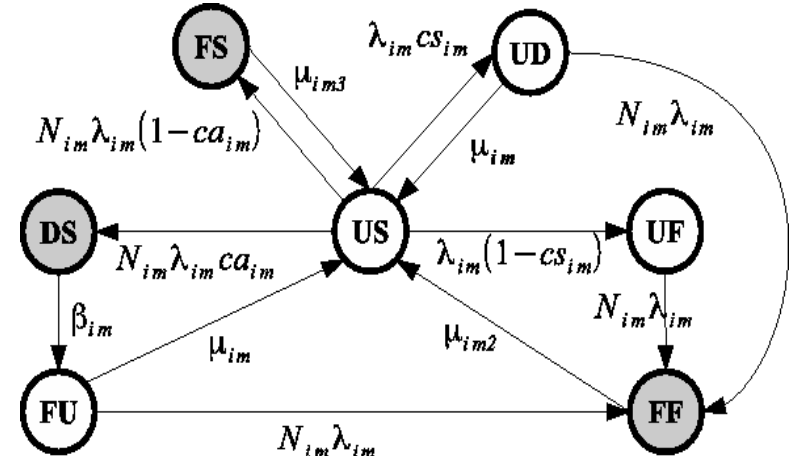
# Modelos CTMC



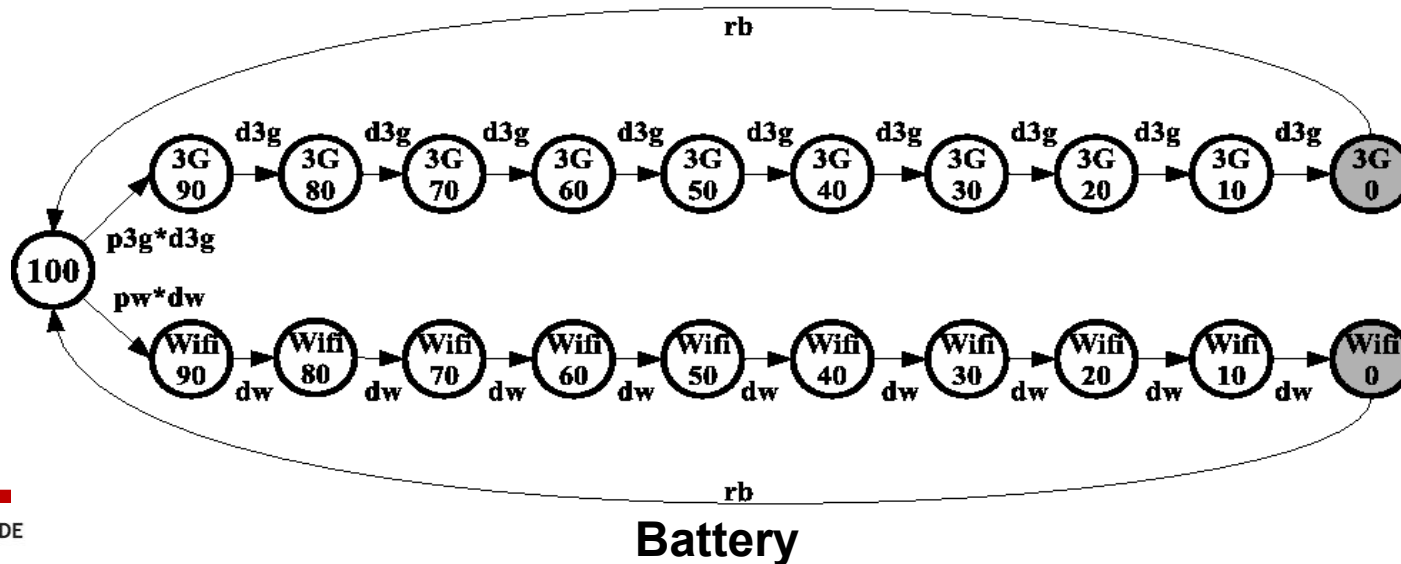
Mobile Device



Mobile Application



Infrastructure Manager



Battery

# Sumarização de resultados



## Derivadas parciais

Parameter	$ SS(A) $
$CS_{sm}$	0.006242686368
$ca_{sm}$	0.005679147263
$CS_{im}$	0.003946053118
$ca_{im}$	0.003579911742
$rb$	0.002173576439

## Diferença percentual

Parameter	$ S(A) $
$N_{sm}$	0.0068660546
$rb$	0.0058528304
$N_{im}$	0.003583039
$ca_{sm}$	0.0026962696
$CS_{sm}$	0.0026128065

## Análise de DoE

Parameter	Effect
$rb$	0.010521
$ca_{sm}$	0.007525
$N_{sm}$	-0.006857
$\mu_{sm2}$	0.004468
$\lambda_{sm}$	-0.004225

Os parâmetros que aparecem entre os **5 primeiros** do ranking em pelo menos **2 dos 3 métodos** são:

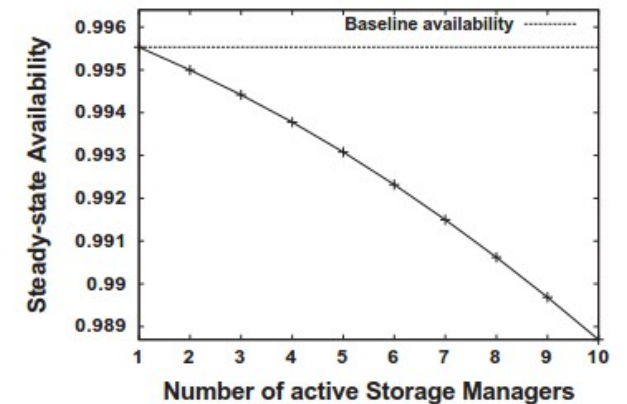
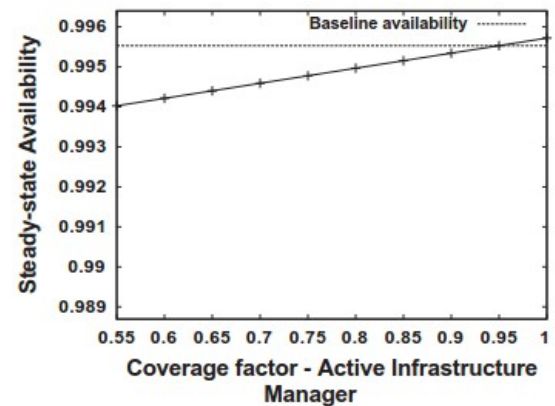
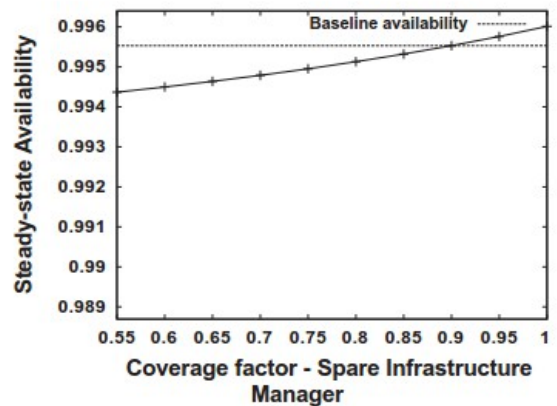
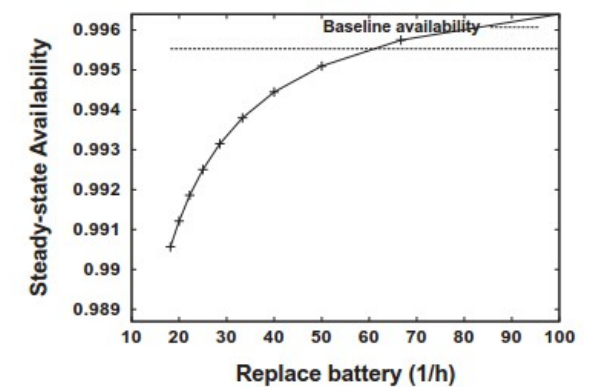
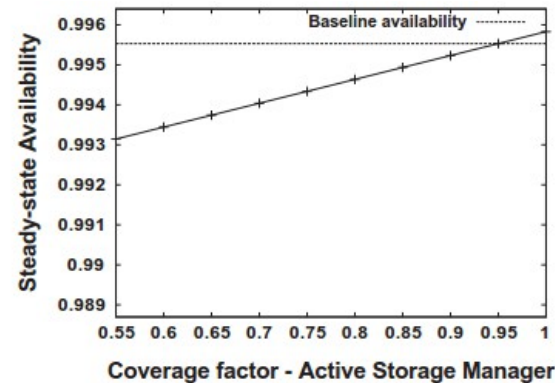
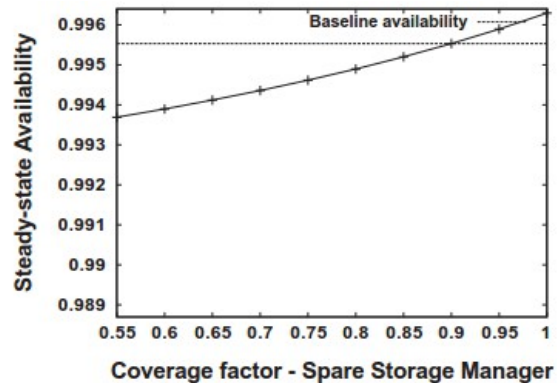
- $ca_{sm}$  • **Fator de cobertura** da falha do servidor primário no **Storage Manager**
- $CS_{sm}$  • **Fator de cobertura** da falha do servidor secundário no **Storage Manager**
- $N_{sm}$  • **Número** necessário de **Storage Managers** ativos
- ' • Taxa de **substituição da bateria** do dispositivo móvel





# Sumarização de resultados

Availability	Number of nines	Downtime (h/yr)
0.99553119	2.349808	39.147



# Web services em nuvem privada com autoscaling



- Composição de web services para recomendação de eventos musicais
- Criação e encerramento automático de VMs de acordo com a carga
- 3 modelos: 1 SPN + 2 CTMC:
  - Carga / autoscaling
  - Instanciação das VMs
  - Execução do web service (tempo de resp.)

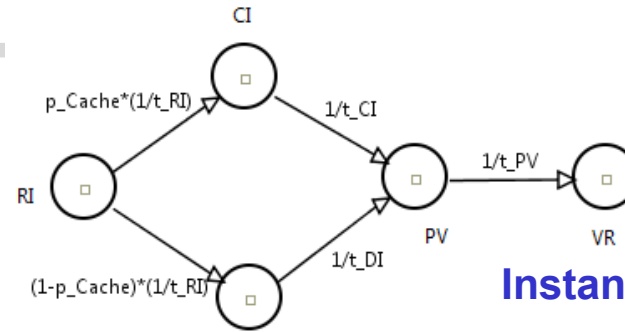




# Web services em nuvem privada com autoscaling

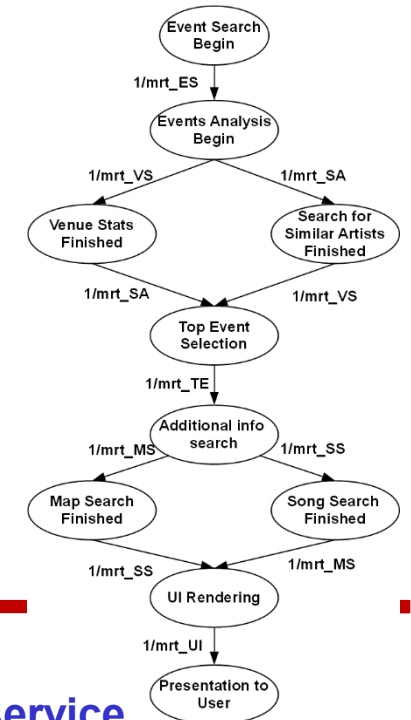
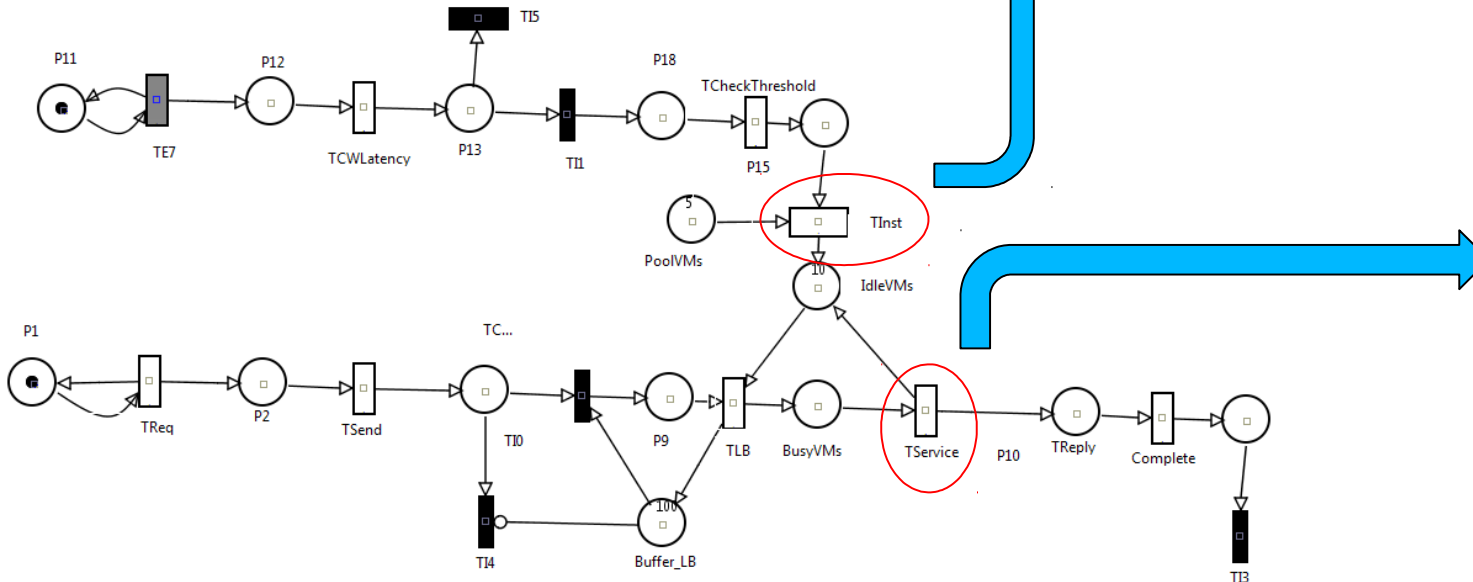


Colaboração com Eliomar e Artur



Instanciação de VMs

## Representação global do sistema



Execução do web service



# Próximos passos

---



- Incluir funções de A.S. em CTMC no Mercury para:
  - Probabilidades transientes
  - Tempo médio de absorção
- Testar o uso de A.S. para otimização baseada no GRASP (em andamento)
- Experimentos para validar os resultados de alguns estudos de caso (em andamento)

