

Análise de sensibilidade em modelos hierárquicos para sistemas de cloud computing

Proposta de tese de doutorado

Rubens de Souza Matos Júnior

Universidade Federal de Pernambuco
Centro de Informática - CIn

16 de Março de 2012

Sumário

- 1 Motivação para computação em nuvem
- 2 Modelagem hierárquica e Análise de sensibilidade
- 3 Estudos de caso

Questões sobre computação em nuvem

- Vários tipos de sistemas distribuídos migrando para/nascendo nesse modelo
 - Aplicações web clássicas
 - Redes sociais
 - Serviços web isolados e compostos
- Falhas no serviço Amazon EC2
 - Cascadeamento de falhas nos volumes EBS
 - Replicação negligenciada pelos projetistas das aplicações
- Interdependência entre aplicações
 - Serviços mashup
 - Aplicações e jogos dentro de redes sociais (Facebook, Orkut, Twitter)

Infraestrutura MoDCS Cloud

- Know-how adquirido
 - Implantação da nuvem Eucalytus
 - Experimentos de envelhecimento de software
- Perspectiva de trabalhos envolvendo aspectos em outros níveis da nuvem
 - Armazenamento distribuído
 - Desempenho de alocação de VMs
 - Desempenho/disponibilidade de aplicações executando na nuvem

Questões a serem respondidas

Perguntas

Considerando um dado cenário de aplicação, como descobrir o que produzirá maiores ganhos na disponibilidade/desempenho?

- Hardware mais potente/confiável
- Arquitetura mais sofisticada
- Software mais flexível/resiliente

Questões a serem respondidas

Perguntas

Considerando um dado cenário de aplicação, como descobrir o que produzirá maiores ganhos na disponibilidade/desempenho?

- Hardware mais potente/confiável
- Arquitetura mais sofisticada
- Software mais flexível/resiliente

Possíveis respostas

Modelos hierárquicos, considerando os componentes de diferentes níveis.

Análise de Sensibilidade para obter respostas diretas.

Tipos de modelos a serem considerados

1 Para dependabilidade

- Fault Tree
- RBD
- Markov chains

Tipos de modelos a serem considerados

1 Para dependabilidade

- Fault Tree
- RBD
- Markov chains

2 Para desempenho

- Markov chains
- GSPN
- Redes de Fila?

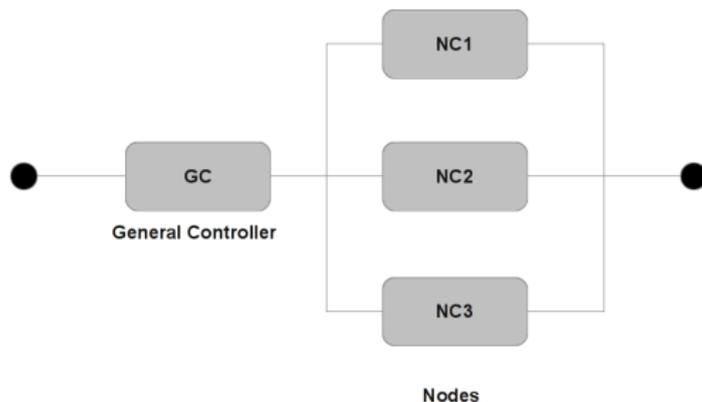
Tipos de modelos a serem considerados

- 1 Para dependabilidade
 - Fault Tree
 - RBD
 - Markov chains
- 2 Para desempenho
 - Markov chains
 - GSPN
 - Redes de Fila?
- 3 Combinações: $\binom{1}{1}$; $\binom{1}{2}$; $\binom{2}{1}$; $\binom{2}{2}$

Diferenciais almejados

- Junção de modelos a nível de aplicação com modelos de infraestrutura (Eucalyptus)
- Consideração de aspectos de dependabilidade/desempenho relacionados a software aging
- Análise de sensibilidade automatizada
- Foco em aspectos de cloud/elastic computing: dynamic provisioning

Eucalyptus: Modelo RBD (alto nível)



Eucalyptus: Modelo CTMC (nível GC)

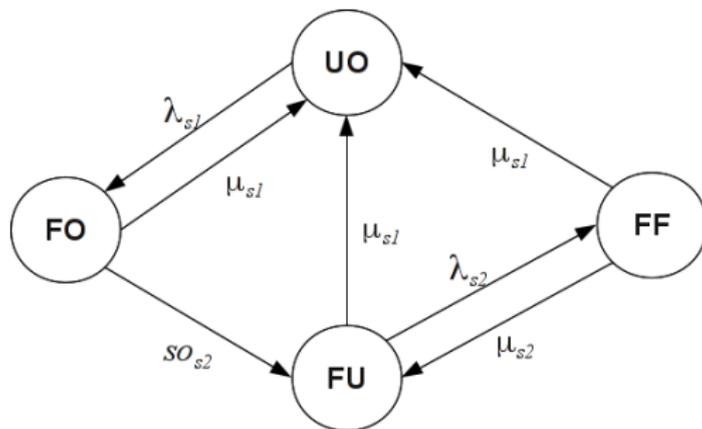
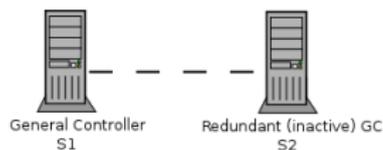
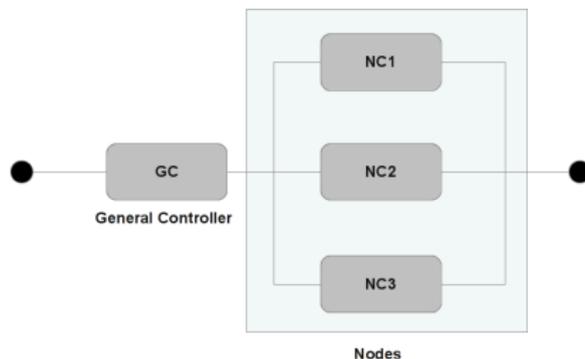


Figura: Controlador geral com redundância cold-standby

Sensibilidade da Disponibilidade: Modelo de alto nível



$$A_S = A_{GC} * A_{Nodes},$$
$$\frac{\partial A_S}{\partial \theta} = \frac{\partial A_{GC}}{\partial \theta} A_{Nodes} + A_{GC} \frac{\partial A_{Nodes}}{\partial \theta}$$

Sensibilidade da Disponibilidade: Nós

$$A_{Nodes} = (1 - \prod_{i=1}^3 (1 - A_{NC_i})),$$
$$\frac{\partial A_{Nodes}}{\partial \theta} = \frac{\partial (1 - \prod_i^3 (1 - A_{NC_i}))}{\partial \theta}$$
$$\frac{\partial A_{Nodes}}{\partial \theta} = \frac{\partial (1 - (1 - A_{NC_i})^3)}{\partial \theta}$$

Sensibilidade da Disponibilidade: Nós

$$A_{Nodes} = (1 - \prod_{i=1}^3 (1 - A_{NC_i})),$$
$$\frac{\partial A_{Nodes}}{\partial \theta} = \frac{\partial (1 - \prod_i^3 (1 - A_{NC_i}))}{\partial \theta}$$
$$\frac{\partial A_{Nodes}}{\partial \theta} = \frac{\partial (1 - (1 - A_{NC_i})^3)}{\partial \theta}$$

$$A_{NC_i} = \frac{\mu_{NC}}{\lambda_{NC} + \mu_{NC}}$$

Sensibilidade da Disponibilidade: General Controller

$$A_{GC} = \frac{\mu_{s1} + \frac{SO_{s2} * \lambda_{s1} * (\mu_{s1} + \mu_{s2})}{(SO_{s2} + \mu_{s1}) * (\lambda_{s2} + \mu_{s1} + \mu_{s2})}}{\lambda_{s1} + \mu_{s1}}$$

Sensibilidade da Disponibilidade: General Controller

$$A_{GC} = \frac{\mu_{s1} + \frac{SO_{s2} * \lambda_{s1} * (\mu_{s1} + \mu_{s2})}{(SO_{s2} + \mu_{s1}) * (\lambda_{s2} + \mu_{s1} + \mu_{s2})}}{\lambda_{s1} + \mu_{s1}}$$

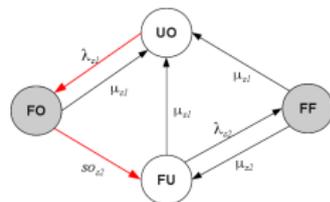
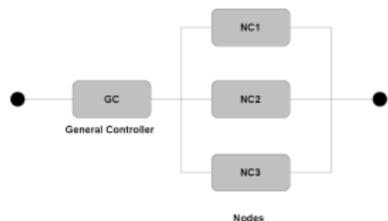
$$\frac{\partial A_{GC}}{\partial \lambda_{s1}} = - \left(\frac{\lambda_{s1} * SO_{s2} * (\mu_{s2} + \mu_{s1})}{(\mu_{s2} + \mu_{s1} + \lambda_{s2}) * (\mu_{s1} + SO_{s2}) * (\lambda_{s1} + \mu_{s1})^2} + \mu_{s1} \right) + \frac{(\mu_{s2} + \mu_{s1}) * SO_{s2}}{(\mu_{s2} + \mu_{s1} + \lambda_{s2}) * (\mu_{s1} + SO_{s2}) * (\lambda_{s1} + \mu_{s1})}$$

Resultados

Parâmetro	Valor
λ_{s1}	0,0001
λ_{s2}	0,0002
μ_{s1}	0,2
μ_{s2}	0,2
SO_{s2}	20
λ_{nc}	0,0005
μ_{nc}	0,2

Disponibilidade do sistema: 0,9999947892

Resultados



Parâmetro	$SS_k(A)$
λ_{s1}	$-5,192 \times 10^{-6}$
SO_{s2}	$4,896 \times 10^{-6}$
μ_{s1}	$4,196 \times 10^{-7}$
λ_{s2}	$-2,471 \times 10^{-7}$
μ_{s2}	$1,235 \times 10^{-7}$
μ_{nc}	$4,640 \times 10^{-8}$
λ_{nc}	$-4,640 \times 10^{-8}$

Próximos passos

- Comparar resultados da análise diferencial com resultados de DoE
- Integrar métricas como Birnbaum's Importance, Structural Importance, etc.
- Incorporar modelos de:
 - Envelhecimento/rejuvenescimento do Eucalyptus
 - Desempenho de alocação de máquinas virtuais
 - Desempenho/confiabilidade de aplicações Mashup (web services)

