

Doris Muzemba Mbayi

**Classificação dos Métodos de Tomada de Decisão para
Gerenciamento Autônomo da Infraestrutura de Computação em
Nuvem**

Trabalho de Conclusão de Curso
submetido ao Curso de Ciências da
Computação da Universidade Federal
de Santa Catarina para a obtenção do
Grau de Bacharel em Ciências da
Computação

Orientador: Prof. Dr. Carlos Becker
Westphall

Florianópolis
2014

Ficha de identificação da obra elaborada pelo autor
através do Programa de Geração Automática da Biblioteca Universitária
da UFSC.

A ficha de identificação é elaborada pelo próprio autor
Maiores informações em:
<http://portalbu.ufsc.br/ficha>

Doris Muzemba Mbayi

**Classificação dos Métodos de Tomada de Decisão para
Gerenciamento Autônomo de Infraestrutura de Computação em
Nuvem**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de “Bacharel em Ciências da Computação”, e aprovado em sua forma final pelo Curso de Bacharelado em Ciências da Computação.

Local, de de .

Prof., Dr. Renato Cislghi
Coordenador

Banca Examinadora:

Prof., Dr. Carlos Becker Westphall
Orientador

Prof.^a, Dr.^a Carla Merkle Westphall

M.Sr, Rafael de Souza Mendes

À Deus por tudo que me proporciona na vida.

À minha mãe e meu pai, os quais amo muito, pelo exemplo de vida e família.

À meus irmãos por tudo que me ajudaram até hoje.

À minha namorada, pelo carinho, compreensão e companheirismo.

À meus verdadeiros amigos pela alegria e companheirismo.

AGRADECIMENTOS

Agradecer primeiramente a Deus, por me iluminar e abençoar minha trajetória.

Ao meu pai Jean Mbayi Katungulu, e a minha mãe Rose Ibondo Tshitenga, pelo apoio e por tudo que sempre fizeram por mim, pela simplicidade, exemplo e carinho, fundamentais na construção do meu caráter.

Aos meus irmãos Marth Mbayi, Julie Mbayi, Simon Mbayi, Taty Mbayi, Aimée Mbayi, Patience Mbayi, Rosie Mbayi, Michoux Tshiela e meu falecido irmão Edo Mbayi pelo conhecimento e dicas importantes que contribuíram para minha formação.

A minha namorada Madalena Chaves, pacientemente, sempre me dando conselhos, forças e carinho.

Aos meus familiares por me ajudarem com conselhos, coragem e incentivos.

Agradeço o coordenador do Projeto Fortalezas Multimídia, o Arquiteto Roberto Toner e meus amigos do estagio Jéfté, Sanders, Pedro e Ludmilla por todo ensinamento e pela amizade de todos esses anos.

Ao meu orientador, professor Carlos Becker Westphall, pela grande colaboração, que muito contribui para conclusão deste trabalho.

Aos membros da banca, Rafael de Souza Mendes e a professora Carla Merkle Westphall, pelo tempo dedicado para revisão rigorosa e pelas sugestões que contribuíram com este trabalho.

Saber ligar um computador não significa saber informática como a maioria dos usuários fala, pois a informática por simples que as pessoas pensem guarda segredos que não imaginamos.

Irenaldo Alves

RESUMO

Os serviços de Computação em Nuvem são oferecidos de forma abstrata. Uma das grandes dificuldades com a qual os provedores de serviços da Nuvem se deparam é o gerenciamento da infraestrutura de Computação em Nuvem, pois várias aplicações diferentes tipos devem acessar concorrentemente aos recursos computacionais. Para atender a essas dificuldades, vários modelos de gerenciamento autônomo de recursos foram propostos por diferentes autores em artigos e trabalhos científicos. Devido a esta multiplicação de modelos, faz-se necessário criar uma taxonomia para classificar os trabalhos que abordam este tema. Para tanto o presente trabalho lança mão de estudos na área de Computação em Nuvem e da tomada de decisão para compor uma classificação se aplique aos trabalhos estudados, bem como trabalhos futuros.

Palavras-chave: Computação em Nuvem, Gerenciamento Autônomo de Recursos, Tomada de decisão, Taxonomia para Computação em Nuvem.

ABSTRACT

The Cloud Computing services are offered in abstract form. One of the great difficulties that the service providers of Cloud facing is the management of Cloud Computing infrastructure, because several different applications types must access concurrently to computing resources. To meet these difficulties, several models of autonomous management of resources have been proposed by different authors in articles and scientific papers. Due to the proliferation of models, it is necessary to create taxonomy to classify the jobs that deal with this theme. For both the present work launches hand of studies in the area of Cloud Computing and the decision-making process to compose a classification applies to work studied, as well as future work.

Keywords: Cloud Computing, Autonomous Management of Resources, Decision-making, and Taxonomy for Cloud Computing.

LISTA DE FIGURAS

Figura 1 – Arquitetura Orientada a Negocio de CN (MARY, 2013)	34
Figura 2 – Serviços de CN (WIKIPÉDIA, 2013)	35
Figura 3 – The NIST Cloud definition (JOYNER, 2011)	36
Figura 4 – Características do modelo de Computação em Nuvem (OPUS SOFTWARE, 2013)	37
Figura 5 – Infraestrutura de Computação em Nuvem (JESUS, 2012)	39
Figura 6 – Arquitetura da Computação em Nuvem (ZANG; CHENG; BOUTABA, 2010)	41
Figura 7 – Diagrama do Processo de Analise de Decisão (ADRIANE, 2011)	49
Figura 8 – Estrutura de Sistema de Ajuda a Tomada de Decisão (SPRAGUE, 1982)	50
Figura 9 – Estrutura de Sistema de Ajuda a Tomada de Decisão (MARAKAS, 2003)	51
Figura 10 – Estrutura de Sistema de Ajuda a Tomada de Decisão (TURBAN; ARONSON, 1998)	51
Figura 11 – Arquitetura de Automatic Computing (IBM Researches, 2014)	52
Figura 12 – Migração de Maquinas Virtual (DELGRADO, 2009)	63

LISTA DE TABELAS

Tabela 1 – Etapas do Processo de Tomada de Decisão e seus autores ..	47
Tabela 2 – Definição em que os modelos se baseiam	69
Tabela 3 – Necessidades de gerenciamento apontado nos modelos.....	71
Tabela 4 – Necessidades de cada categoria.....	74
Tabela 5 – Combinação de modelos e necessidades	74
Tabela 6 – Critério de Catisfação ou Otimização	76
Tabela 7 – Tomada de Decisão.....	79
Tabela 8 – Mudança no Tempo	81
Tabela 9 – Formalismo em que os modelos se baseam.....	83
Tabela 10 – Saída dos modelos	86
Tabela 11 – Aspecto da Nuvem abordada	89
Tabela 12 – Classificações dos modelos	91

LISTA DE ABREVIATURAS E SIGLAS

ABNT – Associação Brasileira de Normas Técnicas
AMC – Assosiation for Computing Machinery
API – Application Programming Interface
CE – Computação Evolucionaria
CN – Computação em Nuvem
DNS – Domain Name System
IA – Inteligência Artificial
IaaS – Infrastructure as a Service
ICA – Inteligência Computacional Aplicada
IEEE – Instituto de Engenheiros Eletricistas e Eletrônicos
IP – Internet Protocol
LAN – Local Area Network
LN – Logica Nebulosa
MB – MegaByte
MBPS – MegaByte Por Segundo
NIST – National Institute of Standards and Technology
PaaS – Plataform as a Service
PL – Programação Linear
QoS – Quality of Service
REST– Representational State Transfer
SaaS – Software as a Service
SGBC – Sistema de Gerenciamento de Banco de Conhecimento
SGBD – Sistema de Gerenciamento de Banco de Dados
SGBM – Sistema de Gerenciamento de Banco de Modelos
SLA – Service-Level Agreement
SO – Sistema Operacional
TF – Teoria de Fila
TJ – Teoria de Jogo
VPN – Virtual Private Network
WAN – Wide Area Network
WebDAV – Web-based Distributed Authoring and Versioning

SUMÁRIO

1	INTRODUÇÃO	31
1.1	Contexto	25
1.2	Objetivos do Trabalho	25
1.2.1	Objetivos gerais	28
1.2.2	Objetivos específicos	28
1.3	Organização do Trabalho	28
2	REVISÃO DA LITERATURA.....	27
2.1	Computação em Nuvem	28
2.1.1	Definição	28
2.1.2	Características de Computação em Nuvem	28
2.1.3	Modelos de Infraestrutura de Computação em Nuvem.....	28
2.1.3.1	Nuvem Publica	28
2.1.3.2	Nuvem Privada	28
2.1.3.3	Nuvem Comunitária	28
2.1.3.4	Nuvem Híbrida	28
2.1.4	Modelos de serviço de Cloud Computing	28
2.1.4.1	Software como Serviço (SaaS)	28
2.1.4.2	Plataforma como Serviço (PaaS)	28
2.1.4.3	Infraestrutura como Serviço (IaaS)	28
2.1.5	Segurança e Aspectos Legais do Computação em Nuvem	28
2.1.6	Benefícios	28
2.1.7	Desafios	28
2.2	Teoria de Decisão	28
2.2.1	Introdução.....	28
2.2.2	Questão da Tomada de Decisão	28
2.2.3	Sistemas de Ajuda a Tomada de Decisão	28
2.2.4	Tomada de Decisão em Computação em Nuvem	28
2.3	Autonomic Computing	28
2.4	Programação Linear.....	28
2.5	Teoria do Jogo	28
2.6	Inteligência Computacional	28
2.6.1	Redes Neurais	28
2.6.2	Logica Nebulosa	28
2.6.3	Computação Evolucionaria.....	28
2.6.3.1	Algoritmo genética	28
2.6.3.2	Programação Genética	28
2.6.3.3	Programação Evolucionaria	28
2.6.3.4	Algoritmo Evolucionário	28
2.6.4	Inteligência Artificial.....	28
2.6.5	Sistemas Especialistas	28
2.6.6	Data Mining	28
2.6.7	Neuro Genética	28
2.6.8	Neuro Fuzzy	28

2.6.9	Fuzzy Database	28
2.6.10	Soft Computing	28
2.7	Gerenciamento da Infraestrutura de Nuvem	28
2.7.1	Gerenciamento em Nível da Empresa Provedora.....	28
2.7.1.1	Alocação de Recursos	28
2.7.1.2	Configuração e Inicialização.....	28
2.7.1.3	Reconfiguração	28
2.7.1.4	Monitoração	28
2.7.2	Gerenciamento em Nível da Empresa Cliente	28
2.7.2.1	Criação e Implantação de VMs	28
2.7.2.2	Alocação de Recursos	28
2.7.2.3	Configuração e Inicialização.....	28
2.7.2.3	Reconfiguração	28
2.7.2.4	Monitoração	28
2.7.3	Alocação Dinamica de Recursos.....	28
2.7.3.1	Migração de Maquinas Virutais	28
2.7.3.2	Elasticidade.....	28
2.7.4	Sistemas de Gerenciamento Autonomo para Nuvem.....	28
3.	UMA TAXONOMIA PARA CLASSIFICAÇÃO	
	MÉTODOS PARA GERENCIAMENTO AUTONOMO DE	
	RECURSOS EM CLOUD.....	28
3.1	Metodologias de pesquisa para chegar à amostra piloto	28
3.2	Taxonomias propostas	28
3.3	Classificações dos modelos propostas nos trabalhos	28
4	CONCLUSÃO E TRABALHOS FUTUROS	28
5	REFERÊNCIAS	33

1. INTRODUÇÃO

1.1 Contexto

Computação em Nuvem (CN) ou *Cloud Computing* em inglês oferece uma solução interessante para o desenvolvimento de software e acesso de conteúdo com a transparência da localização da infraestrutura utilizada. A infraestrutura de *CN* é geralmente composta por vários datacenters em que os consumidores têm acesso a apenas uma parte do poder computacional. A alocação desses recursos computacionais é controlada por um provedor, e os recursos são alocados em uma maneira elástica, de acordo com as necessidades dos consumidores (OLIVEIRA; TARCISIO; SANTOS, 2011). A maior vantagem deste modelo é utilizar, em qualquer lugar e independentemente de plataforma, as mais variadas aplicações por meio da Internet com a mesma facilidade como se eles estivessem instalados em nossos próprios computadores.

Para acomodar demandas de recursos de forma escalável e elástica, o processo de alocação e realocação em *Nuvem* deve ser dinâmico. Outra característica essencial dos mecanismos de alocação de recursos em *Nuvem* é garantir que os requisitos de todas as aplicações sejam devidamente atendidos.

CN permite que os consumidores reduzam ou eliminem custos associados com a infraestrutura interna para a prestação dos seus serviços. Esta oportunidade de redução de custos faz dela uma alternativa atraente, especialmente para iniciativas de negócios. O gerenciamento de recursos computacionais continua um grande desafio para os provedores de *Nuvem*, visto a complexidade da infraestrutura de *Nuvem*, este gerenciamento deve ser feito de forma autônoma.

Vários pesquisadores e autores propuseram modelos diferentes para gerenciamento autônomo de recursos em *Nuvem*, cada um resolvendo problema de gerencia de recursos baseado em formalismo como programação linear, teoria de jogo, inteligência computacional, algoritmo genético, etc.

Propor uma taxonomia para estes modelos é o objetivo deste trabalho.

1.2 Objetivos do Trabalho

1.2.1 Objetivos gerais

O objetivo deste trabalho é classificar os modelos utilizados na tomada de decisão para gerenciamento autônomo da *Nuvem* levando em conta suas principais características e particularidade. Antes de chegar à classificação, será feito uma avaliação e comparação desses modelos. Assim o presente trabalho será de grande utilidade, pois ajudará aos Administradores da *Nuvem* a escolher um modelo que combina com as características da sua *Nuvem* para um bom gerenciamento de recursos.

1.2.2 Objetivos específicos

1. Propor uma classificação (taxonomia) para os modelos utilizados na tomada de decisão para gerenciamento autônomo da infraestrutura de *CN*,
2. Classificar um grande número de modelos da *Nuvem* segundo a taxonomia apresentada,
3. Classificar uma ferramenta de gerenciamento da *Nuvem* segundo a taxonomia.

1.3 Organização do Trabalho

A continuidade deste trabalho se articula em torno de quatro capítulos, além da Introdução. No capítulo dois, apresentaremos a revisão literária sobre os conceitos como *CN*, Teoria de Decisão, Automatic Computing, Programação Linear, Teoria de Jogo, Inteligência Computacional e Gerenciamento da Infraestrutura de Nuvem, fundamentais para entendimento do nosso trabalho. No terceiro capítulo, será proposta uma taxonomia para modelos de tomada de decisão para gerenciamento da infraestrutura de Nuvem. No quarto Capítulo serão feitas as considerações finais.

2 REVISÃO DA LITERATURA

2.1 Computação em Nuvem

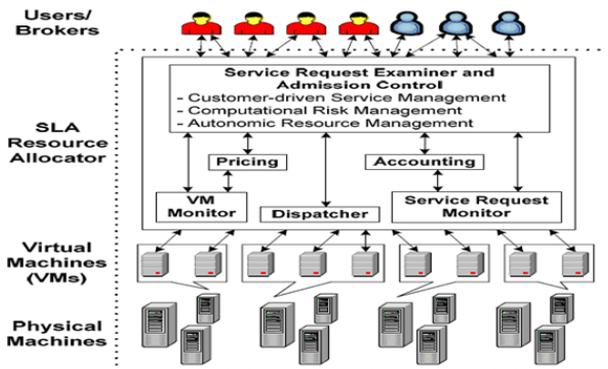
2.1.1 Definição

CN representa atualmente uma revolução para indústria da Informática. Ela pode ser definida como um conceito que faz referência à utilização de recursos tecnológicos e capacidade de processamento e armazenamento dos Servidores espalhados no mundo inteiro, fornecido como serviço que pode ser acessado via internet a partir de qualquer computador (JOSET, 2011).

Com *CN*, o poder de calculo e a capacidade de armazenamento de informações são disponibilizados sobre demanda e cobrado em função da utilização, gerando uma grande diminuição de custos. Comparando com quem instala a sua própria infraestrutura, os usuários de *CN* podem economizar muito a suas despesas e tirar proveito da confiabilidade e escalabilidade do sistema. Além disso, os usuários podem tirar proveito dos recursos computacionais (Rede, Servidores, Aplicações e Serviços) potencialmente ilimitados para realizar suas tarefas de calculo intensivo e acessar aos sistemas de armazenamento em massa (ALBURQUEQUE; RABELO; SOARES, 2012).

Para **Buyya et al. (2008)**, *CN* é um sistema distribuído e paralelo, constituído por uma coleção de computadores virtuais e físicos interconectados que proveem um ou mais recursos computacionais, baseados em Acordos de Níveis de Serviços (SLAs) preestabelecidos através de uma negociação entre consumidor e provedor.

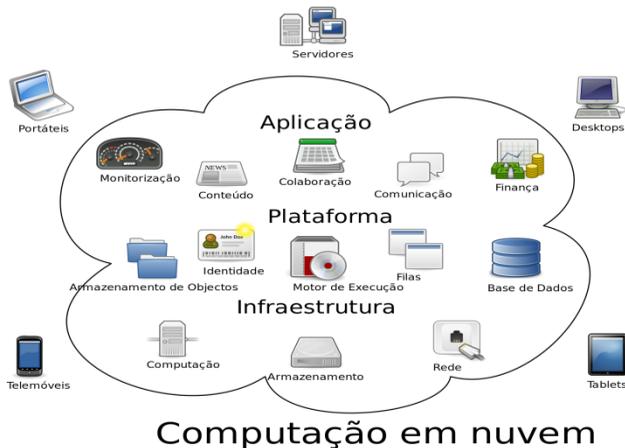
Figura 1 Arquitetura Orientada a Negocio de CN



Fontes: (MARY, 2013)

Já para (ARMBRUST, et al., 2009), *CN* é um conjunto de serviços de rede ativados, proporcionando escalabilidade, Qualidade de Serviço (QoS), infraestrutura barata de computação sob demanda e acessada de forma simples e continua.

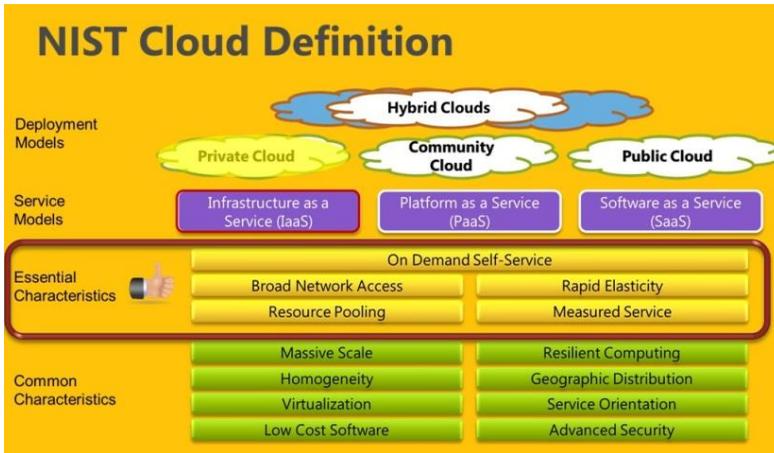
Figura 2 Serviços de CN



Fonte: (WIKIPÉDIA, 2013)

A definição do **NIST** lista cinco características essenciais para *CN*: autosserviço sob demanda, acesso a rede de banda larga, *pool* de recursos, rápida elasticidade ou expansão e serviço de mensuração. A definição também cita três modelos de serviços (SaaS, PaaS e IaaS) e quatro modelos de desenvolvimento (privado, comunitário, público e híbrido) que, juntos, categorizam modos de entrega de serviços de *CN* (MELL; GRANCE, 2011).

Figura 3 The NIST cloud definition



Fonte: (JOYNER, 2011)

2.1.2 Características de Computação em Nuvem

Segundo Mell e Grance (2011), o serviço de *CN* deve responder a certo número de critérios como, por exemplo:

Autoatendimento sob demanda: o usuário aloca e libera recursos computacionais (espaço de armazenamento, banda passante, tempo de servidor) segundo suas necessidades, automaticamente, sem requerer uma intervenção humana com fornecedor de serviço, via API, por exemplo.

Ampla acesso em rede: os recursos são disponibilizados em rede e acessíveis de qualquer plataforma. São utilizados mecanismos padrão que promovem o uso de plataformas heterogêneas. Como por exemplo: celular e PC, ou qualquer outra plataforma.

Pool de recursos: os recursos computacionais (armazenamento, tratamento, memória, banda passante) do fornecedor são reunidos para servir vários Usuários, com alocação e liberação dinâmica de diferentes recursos físicos e virtuais em função da demanda.

Elasticidade Rápida: os recursos são alocados e liberados no momento que for necessário, com agilidade. Para o cliente esta alocação de recursos parece ilimitada e pode ser feito a qualquer momento e na quantidade que precisar.

Serviços Mensuráveis: os sistemas de *CN* ajustam-se automaticamente e aperfeiçoam o uso de recursos de maneira que fique transparente para o consumidor e para o fornecedor. O usuário poderá gerenciar e aperfeiçoar a sua utilização de recursos, ajudando fornecedor a cobrar do cliente pelos recursos utilizados.

Figura 4 Características do modelo de Computação em Nuvem



Fonte: (OPUS SOFTWARE, 2013)

2.1.3 Modelos de Infraestrutura de Computação em Nuvem

Existem vários fatores a levar em conta antes de migrar as aplicações de uma empresa para um ambiente de *CN*. Por exemplo, certos fornecedores de serviços tem às vezes interesse de baixar o custo de funcionamento, mas os consumidores podem preferir quem oferece um nível elevado de segurança, a fiabilidade e desempenho a preços variados. Como consequência existe diferentes tipos de Nuvem, cada um com suas vantagens e desvantagens (DILLON; WU; CHANG, 2010).

Cada organização devera analisar de forma independente cada situação, sendo da sua responsabilidade a escolha do modelo de infraestrutura a adotar. A escolha entre estes modelos citado abaixo poderá ser importante para o seus usuários em termos de disponibilização do serviço, tipos do contrato e preço. Existe um conjunto de fatores a levar em conta para esta decisão (STRUKHOFF et al., 2008):

Investimento inicial: a *Nuvem Privada* tem associado um investimento inicial que depende da complexidade e exigência da solução, mas tipicamente bem mais elevado que o custo inicial de uma *Nuvem Publica*.

Volume de dados: o armazenamento em *Nuvem* é conhecido por sua grande escalabilidade, na maioria das empresas o volume de informação começa pequeno. Na *Nuvem Privada* fornece uma fácil escalabilidade de recursos computacionais quando este volume aumenta.

Longevidades de dados: na *Nuvem Publica* quanto mais informação acumulada, mais elevados serão os custos de utilização do serviço.

Desempenho exigido: *Nuvens Privadas* são implantadas dentro de firewalls e tipicamente com uma largura de banda de no mínimo 100MBPS por nó. *Nuvens Publicas* são acessadas via Internet e enfrentam os limites da largura da sua banda. Isso varia em torno de 10MB.

Padrão de acesso e questões de localização: *Nuvens Publicas* têm mecanismo de replicação de dados para vários locais geograficamente dispersos, às vezes por um custo extra. Isso poderá ser de grande vantagem para aplicações com acesso remoto. As *Nuvens Privadas* são normalmente implantadas em um único local para acesso baseado em LAN. Os usuários remotos trabalham se ligando na WAN via Internet.

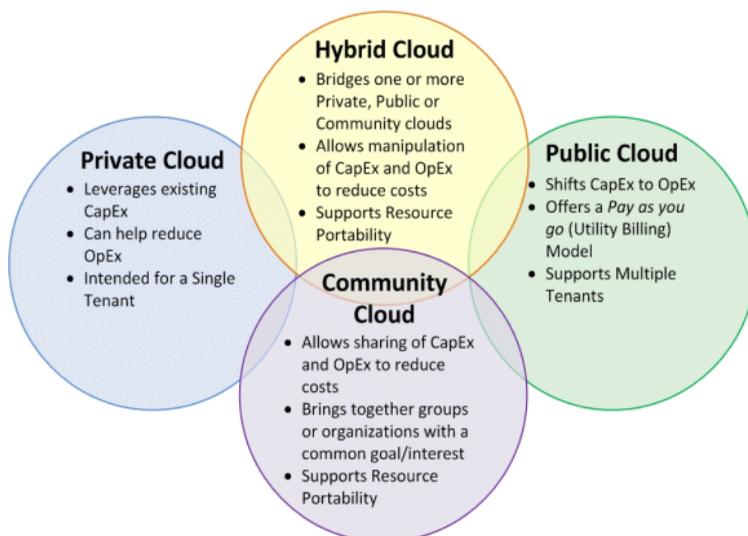
Segurança e isolamento de dados: as responsabilidades pela segurança de dados em *Nuvem* mudam de acordo com tipo de serviço e tipo de infraestrutura. Se existe uma preocupação com os dados a serem armazenados fora da empresa, *Nuvem Publica* não é uma escolha adequada. A solução seria uma *Nuvem Privada* onde os dados são isolados com base em suas necessidades de segurança.

Confidencialidade e destruição de dados: assim como a segurança, confidencialidade dos dados é um fator a considerar ao escolher a solução de armazenamento em *Nuvem*. Com *Nuvens Privadas*, se tem controle sobre o acesso a dados e o conhecimento sobre as atividades legais. Quando chegar a hora destruir os dados, esta no poder do cliente para fazer isso.

Acordo de Nível de Serviço (SLA): *Nuvens Privadas* tem mecanismos de disponibilidade e serviço de acesso a dados. Falhas individuais de um servidor não derrubam a *Nuvem* ou geram perda de dados de modo a quebrar o *SLA*. Em *Nuvens Publicas* *SLAs* são publicados pelo provedor e são da sua responsabilidade. Em caso de perda de dados, o fornecedor faz o seu melhor para recupera-los, cobrando um custo extra sem honrar nenhuma garantia de disponibilidade de serviço.

Recursos técnicos próprios: *Nuvens Publicas* eliminam necessidades de alocação de recursos da organização para gerenciamento e manutenção da infraestrutura. Devido a utilização de protocolos como o WebDAV e REST, as aplicações terão que ser adaptadas para comunicar com outros serviços via esses mesmos protocolos necessitando de know-how para desenvolvimento e migração.

Figura 5 Infraestrutura de Computação em Nuvem



Fonte: (JESUS, 2012)

2.1.3.1 Nuvem Publica

A infraestrutura de *Nuvem* é colocada à disposição do público em geral ou de um grande número de empresas via rede. O fornecedor gerencia a infraestrutura e pode alocar recursos entre vários clientes. Este tipo de *Nuvem* oferece várias vantagens aos clientes, como exemplo, a redução de investimento inicial, otimização de custos operacionais e a transferência de riscos tecnológicos para o fornecedor da infraestrutura. A *Nuvem Publica* não oferece uma disponibilidade permanente ou controle preciso sobre os dados, à rede e parâmetros de segurança, isso faz com que a maioria das empresas optem por *Nuvem Privada*.

2.1.2.2 Nuvem Privada

A infraestrutura é concebida para ser usada exclusivamente por uma organização. A *Nuvem Privada* pode ser hospedada e gerenciada por uma organização ou por um terceiro. No caso de fornecedor terceiro, a organização pode utilizar redes seguras do tipo *VPN* para acessar a infraestrutura. Este modelo de *Nuvem* oferece o mais alto grau de controle sobre o desempenho, a fiabilidade e a segurança.

2.1.3.3 Nuvem Comunitária

Consiste em dividir a infraestrutura entre organismos de uma mesma comunidade que tem interesses comuns. Deve-se ressaltar, ainda, que este consenso entre os usuários integrantes da *Nuvem Comunitária* deveram estabelecer políticas de uso e de segurança das informações. Além disso, os custos podem ser divididos entre o numero de usuários, gerando uma economia mais significativa do que a *Nuvem Privada*.

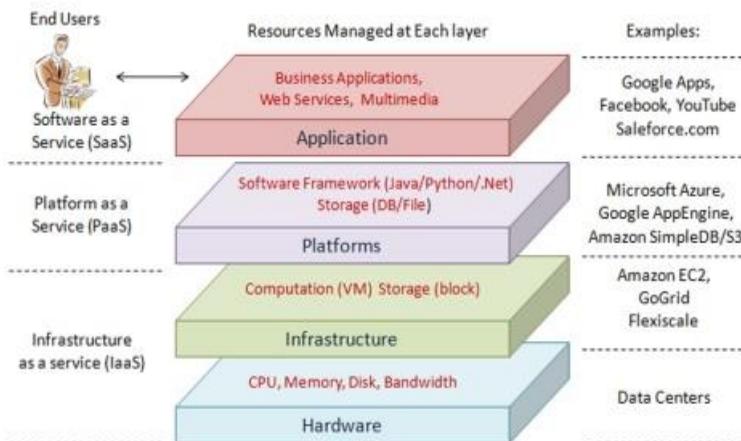
2.1.3.4 Nuvem Híbrida

Nuvem Híbrida é uma combinação de dois ou mais modelos de *Nuvem* (Privada, Pública ou Comunitária) que tende a remediar as limitações de cada um desses modelos e permitir a portabilidade de dados e aplicações entre as *Nuvens*. Na *Nuvem Híbrida*, uma parte da infraestrutura de serviço pode se tornar uma *Nuvem Privada* e a outra parte pode ser executada como *Nuvem Pública*. A *Nuvem Híbrida* oferece mais flexibilidade do que a *Nuvem Pública* e a *Nuvem Privada*. Especialmente ela fornece um melhor controle e segurança sobre os dados de aplicações do que a *Nuvem Pública*, oferecendo uma flexibilidade e uma escalabilidade.

2.1.4 Modelos de serviço da Computação em Nuvem

CN fornece recursos computacionais físicos e virtuais como serviço de um jeito flexível e em função da demanda. A figura 6 ilustra a arquitetura deste conceito que pode ser dividido em quatro camadas: camada física, camada de infraestrutura, camada de plataforma e camada de aplicação. Conceitualmente, cada camada da arquitetura ilustrada na figura 6 pode ser implementada como serviço da camada superior. Inversamente, cada camada pode ser concebida como cliente da camada inferior. Assim na prática, *CN* oferece serviços que podem ser reagrupados em três categorias: o Software as a Service (SaaS), o Platform as a Service (PaaS) e o Infrastructure as a Service (IaaS) (Zhang, et al.,2010).

Figura 6 Arquitetura da Computação em Nuvem



Fontes: (ZHANG; CHENG; BOUTABA, 2010).

2.1.4.1 Software as a Service (SaaS)

Este modelo conhecido como Software como Serviço em português faz referência a fornecer softwares completos em *Nuvem*. Em sua essência, trata-se de uma forma de trabalho onde o software é oferecido como serviço, assim, o usuário não precisa adquirir licenças de uso para instalação ou mesmo comprar computadores ou servidores para executá-lo. Nesta modalidade, no máximo, paga-se um valor periódico, como se fosse uma assinatura, somente pelos recursos utilizados ou pelo tempo de uso.

2.1.4.2 Platform as a Service (PaaS)

Este modelo, conhecido em português como Plataforma como Serviço, consiste em oferecer um ambiente de hospedagem e desenvolvimento em *Nuvem* permitindo criar aplicações com ajuda de linguagens de programação e outras ferramentas oferecidas pelo fornecedor. O PaaS permite também fornecer aos usuários mais flexibilidade reduzindo a complexidade de desenvolvimento e evitando a perda de tempo no desenvolvimento de aplicações. O usuário não gerencia nem controla a infraestrutura, mas ele pode controlar as aplicações desenvolvidas e eventualmente configurar o ambiente de hospedagem de aplicações.

2.1.4.3 Infraestrutura as a Service (IaaS)

Este modelo conhecido em português como Infraestrutura como Serviço consiste em fornecer recursos físicos como unidade de processamento, memória e capacidades de armazenamento, frequentemente como VMs. O usuário pode implantar e controlar os sistemas operacionais, armazenar dados, rodar suas aplicações e eventualmente configurar certos componentes de rede.

2.1.5 Segurança e aspectos legais da Computação em Nuvem

Apesar da diversidade de benefícios oferecida pela *CN*, existe algumas situações que precisam ser melhor analisadas, para que a empresa possa usufruir desta tecnologia sem maiores problemas. É o caso da segurança da informação. Como ocorre a segurança da informação na computação nas nuvens? Ela é realmente segura e eficaz?

A norma **NBR-ISO-IEC 17999** faz três recomendações que devem ser seguidos pelos responsáveis sobre introdução, implementação ou manutenção da segurança de informação (**LOPES, 2010**).

Confiabilidade: a garantia que o acesso de informação seja obtido somente por pessoas autorizadas.

Disponibilidade: a informação ou o sistema computacional devem estar disponíveis no momento que é necessário.

Integridade: a informação deve ser retornada em sua forma original no momento que foi armazenada.

Muitas organizações questionam como as três recomendações desta norma deve se aplicar para *CN*. A questão de segurança em Nuvem já evoluiu bastante, os mesmo mecanismos de proteção de dados em ambientes tradicionais também são aplicados nas estruturas baseadas em *Nuvens*.

O **Gartner (2008)** sugeriu alguns cuidados para tomar antes da aquisição de serviço de provedor de *CN*. A organização deve se fazer as seguintes perguntas:

- Como é feito o acesso dos usuários?
- Como o provedor obedece às normas de regulação?
- Onde os dados são hospedados?
- Como os dados são segregados?
- Como os dados são recuperados?
- Como é feito o suporte?
- Qual a validade do provedor no longo prazo?

Além das questões citada acima **Durkee (2010)** sugerem mais três questões a serem levados em conta:

- Qual é SLA?
- Qual a conectividade do provedor?

- O conhecimento dos responsáveis pela política de segurança de informação está de acordo com a segurança do provedor?

A adoção e implementação de *CN* apresenta uma série de questões legais que deve ser tomadas pelas organizações ao contratar esses serviços, *SLA* serve para definir pontos de acordo sobre a qualidade de serviço prestado pelo fornecedor, definido no contrato. Ele especifica os níveis de serviços que o cliente pode esperar do fornecedor, incluído metas de desempenho (**BLOMBERG, 2008**).

2.1.6 Benefícios

Os possíveis benefícios do uso da *CN* são os seguintes (**VERAS, 2011**):

Menores custos de infraestrutura: cobrança sobre demanda, o usuário paga pelo uso.

Aumento da utilização da infraestrutura: a infraestrutura compartilhada passa a ser mais utilizada e permite a redução de custo por parte do provedor.

Aumento de segurança: uma infraestrutura centralizada pode melhorar a segurança incluído rotinas de backup otimizadas.

Acesso a aplicações sofisticadas: aplicações caras podem ser utilizadas no modelo sobre demanda.

Economia de energia: redução de custos de energia e refrigeração é um grande benefício de estrutura da *Nuvem*.

Aumento de produtividade por usuário: o usuário pode trabalhar em qualquer lugar, e isso pode ter impacto positivo na sua produtividade.

Aumento de confiabilidade: a existência de estrutura de contingência obrigatória em *Nuvem* pode melhorar a confiabilidade de aplicações.

Escalabilidade sob demanda: a organização não precisa projetar a infraestrutura pelo pico. *CN* permite alocar recursos sob demanda, tornando a escalabilidade uma realidade.

2.1.7 Desafios

Assim como a implantação de qualquer tecnologia, a *CN* possui desafios. A interoperabilidade, juntamente com a segurança são dois fatores chaves na adoção da *CN* (**Rochwerger, et al., 2009**). Os principais desafios para adoção da *CN* são:

Falta de Interoperabilidade: este problema esta relacionada com a falta de padrões unificados entre os diferentes fornecedores, dificultando a utilização de múltiplos serviços de *CN* de forma unificada.

Compatibilidade entre aplicações: as aplicações construídas para a *Nuvem* são muitas vezes incompatíveis com as outras plataformas que não são baseados em *Nuvem*.

Dificuldade de obedecer a normas regulatórias: varias organização trabalham na definição de padrões abertos para lidar com problemas normais em ambiente de *CN*.

Segurança inadequada: os consumidores de *CN* podem realizar tarefas como processamento, transferências e armazenamento de informações. E estas tarefas devem ser realizadas de forma segura. Os provedores suportam a arquitetura na qual todos os usuários e aplicativos compartilham a mesma infraestrutura. Desta se estabelece um grande desafio para fazer o gerenciamento de segurança no ambiente.

2.1.8 Conclusão

CN é bem posicionada atualmente na lista das novas tecnologias. Ela se caracteriza pela extensibilidade, elasticidade e operabilidade que seu grande número de usuários se beneficia ao redor do mundo. Os recursos computacionais são disponibilizados por demanda e cobrados em função da utilização. Este conceito se baseia principalmente em Grid Computing, Utility Computing e evidentemente em Virtualização. Assim, como toda inovação tecnológica, *CN* não só ajuda a economizar dinheiro, mas também permite minimizar o impacto ambiental com o uso eficiente de recursos computacionais economizando energia e diminuindo a produção de lixos eletrônicos.

2.2 Teoria de decisão

2.2.1 Introdução

A Teoria de Decisão é uma área interdisciplinar de estudo, com definições que relaciona filosofia, matemática e estatística, aplicável a quase todos os ramos de ciências (NUNES, 2006). Nesta sessão será abordado os Conceitos Gerais da Teoria de Decisão, Decisão em Tecnologia de Informação, a Tomada de Decisão em *CN*.

2.2.2 Questão da Tomada de Decisão

A decisão é um termo que se refere ao processo cognitivo pelo qual uma pessoa pode escolher entre duas ou mais alternativas, mesmo a não escolha de nenhuma das opções é uma tomada de decisão (NUNES, 2006).

No ponto de vista de (LONGUEVILLE, 2003), existem cinco tipos de problemas no qual pode se aplicar a teoria de decisão:

- Problemas de Descrição: associado à caracterização real do estado corrente de uma organização;
- Problemas de Investigação: associados às relações entre dois ou mais elementos ou fenômenos;
- Problemas de Explicação: associados a estabelecimento de uma relação causa efeito;

- Problemas de Predição: associados à projeção futuro baseado sobre dados históricos;
- Problemas de Prescrição: associado à projeção normativa baseada sobre dados históricos;

A questão da tomada de decisão sempre esteve presente no dia-a-dia do ser humano. As decisões são difíceis de serem tomadas por envolver na maioria das vezes vários fatores que devem ser levados em conta.

Existem quatro fatores que dificultam a tomada de decisão: a complexidade de decisão, as incertezas inerentes à situação de decisão, os objetivos do tomador de decisão podem ser múltiplos e as diferentes perspectivas de um problema de decisão (NUNES, 2006).

Já **Robert Thrall (1984)**, no seu artigo sobre a Taxonomia de Modelo de Decisão afirmou que a teoria de decisão pode se aplicar sobre problema do domínio público ou privado e um modelo de tomada de decisão deve levar em conta três características: a primeira característica se refere ao número de decisões a tomar, se preocupando principalmente se houve ou não conflito de interesses na decisão; a segunda característica é referente ao número de critério da decisão, quando o problema envolve vários critérios como a maioria dos problemas da sociedade, esta característica deve ser levada em conta no modelo; a terceira característica é o número da decisão, importante para definir a política de tomada de decisões.

Assim a tomada de decisão pode ser interpretada como uma função f definida da seguinte forma:

$$\textit{Tomada de decisão} = f(x1, x2, x3);$$

Onde,

$x1$: representa o número de decisão a tomar;

$x2$: representa o número de critério da decisão;

$x3$: representa o número de decisão.

Antes do surgimento de métodos de Decisão com Múltiplos Critérios, os problemas de decisão nos levava na maioria dos casos a otimização de uma função econômica, constituído somente de um critério de seleção. No período da segunda guerra mundial surgiu um campo de conhecimento chamado Pesquisa Operacional aplicada à estruturação e análise de tomada de decisão. Com o passar do tempo, os pesquisadores se depararam com a dificuldade de tomar decisão baseado em um único critério. Assim surgiram as metodologias de múltiplos critérios para suprir a esta dificuldade (NUNES, 2006).

Patrícia Marcon (2006) no seu trabalho de mestrado estudou várias modelos de processos de tomada de decisão segundo os autores Simon (1963), Lacombe e Heilborn (2003), Bateman e Snel (1998), Robbins (2002),

Megginson, Mosley e Pietri (1998), Montana (1999), e Stoner e Freeman (1999) e elaborou uma tabela seguinte com Etapas do Processo de Tomada de Decisão:

Tabela 1: Etapas do Processo de Tomada de Decisão e seus autores

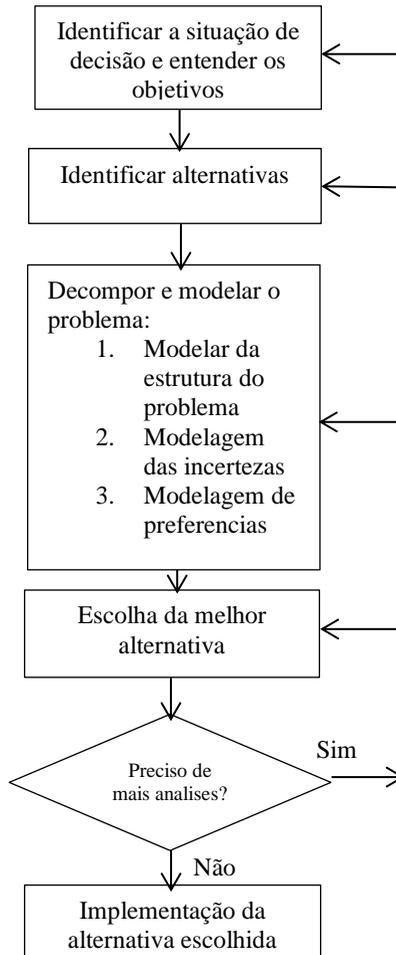
	SIMON (1963)	LACOMBE e HEILBORN (2003)	BATEMAN e SNEL (1998)	ROBBINS (2002)	MEGGISON, MOSLEY e PIETRI (1998)	MONTANA (1999)	STONER e FREEMAN (1999)
1	Coleta de informações	Identificação de sinais e sintomas	Identificação e diagnóstico do problema	Definição do problema	Compreensão e definição da situação	Análise situacional	Definição do problema
2	Atividade de estruturação	Análise do problema ou da oportunidade existente	Elaboração de soluções alternativas	Identificação dos critérios de decisão	Elaboração das alternativas de ação	Estabelecimento do padrão de desempenho	Criação das alternativas
3	Atividade de escolha	Identificação de soluções alternativas	Avaliação das alternativas	Peso dos critérios	Avaliação das vantagens e desvantagens de cada alternativa	Geração de alternativas	Avaliação e seleção da melhor alternativa
4		Análise das soluções alternativas	Escolha	Criação das alternativas	Escolha e implementação da alternativa	Avaliação de alternativas	Implementação e monitoramento da decisão

5		Avaliação das alternativas	Implementação da decisão	Classificação das alternativas segundo os critérios	Avaliação dos resultados da decisão	Teste piloto	
6		Comunicação da decisão escolhida	Avaliação da decisão	Cálculo da decisão ótima			
7		Implementação da decisão					

Fonte: (MARCON, 2006).

Adriane Araújo de Oliveira (2011), na sua dissertação de Mestrado sobre Aplicação de Método de Análise Hierárquica na Tomada de decisão para Adoção de Computação em Nuvem apresenta um fluxograma conforme demonstrado na Figura 7, proposta por **Clemen (1995)** que resume todo o processo de análise que envolve tomada de decisão.

Figura 7 Diagrama do Processo de Análise de Decisão



Fonte: (ADRIANE, 2011)

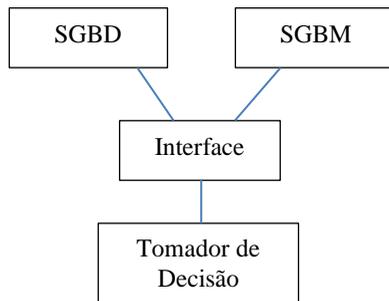
2.2.3 Sistemas de Ajuda a Tomada de Decisão

Um Sistema de Ajuda a Tomada de Decisão é um sistema concebido para resolver problemas de decisão, estes problemas podem ter características seguintes (KLEIN, 1997) os critérios de decisão são múltiplos e conflitantes; a procura de uma solução envolve calculo e manipulação de dados; a sequencia de operações para achar a solução não é conhecido à priori; a solução deve ser obtida em um tempo limitado.

De uma maneira precisa podemos definir um Sistema de Ajuda a Tomada de Decisão como um sistema computacional automatizado, interativo, flexível, adaptável e especialmente desenvolvido para auxiliar na resolução de problemas de tomada de decisão (KLEIN, 1997).

Sprague e Carlson (1982) propuseram a estrutura mostrada na Figura 8 e identificaram três componentes fundamentais na Estrutura de um Sistema de Ajuda a Tomada de Decisão: Interface usuário-maquina que permite uma interatividade entre o usuário e o sistema, Sistema de Gestão de Base de Dados (SGBD) para armazenamento e gerenciamento de dados e um Sistema de Gestão de Base de Modelos (SGBM) que contem um conjunto de modelos e processos de cálculos utilizados em diferente tratamento de dados disponível para o usuário.

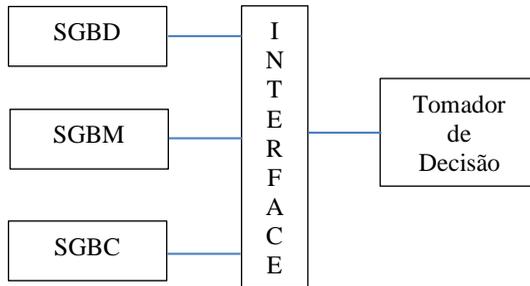
Figura 8 Estrutura de Sistema de Ajuda a Tomada de Decisão



Fonte: (SPRAGUE; CARLSON, 1982)

Já Maracas propõe uma estrutura feita de cinco partes distintas: um SGBD, um SGBM, um Sistema de Gestão de Base de Conhecimento (SGBC) que a tarefa de reconhecimento do problema e geração de soluções finais ou intermediárias e também faz a função de gerenciamento do processo de resolução de problema e a Interface de usuário, ilustrado na Figura 9.

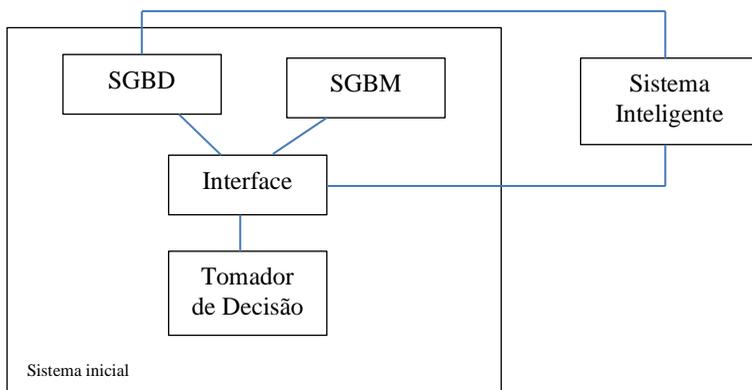
Figura 9 Estrutura de Sistema de Ajuda a Tomada de Decisão



Fonte: (MARAKAS, 2003)

É mais comum a busca de técnicas de Inteligência Artificial para melhorar a desempenho de Sistemas de Ajuda a Tomada de Decisão. **Turban e Aronson (1998)** sugeriram dois modos fundamentais para integração de Sistemas Inteligentes com Sistemas de Ajuda a Tomada de Decisão, um deles está apresentada na Figura 10.

Figura 10 Sistemas Inteligentes com Sistemas de Ajuda a Tomada de Decisão



Fonte: (MARAKAS, 2003)

Sistema de Ajuda a Tomada de Decisão inteligente ou Sistema de Ajuda a Tomada de Decisão a base de conhecimento é baseado sobre a arquitetura do modelo proposta por Sprague e Carlson integrando a ele um Sistema Inteligente, permitindo a capturar conhecimentos do domínio e fornecer uma ajuda inteligente durante o processo de tomada de decisão.

2.2.4 Tomada de Decisão em Nuvem

O problema de alocação dinâmica de recursos aparece em todos os cenários onde a quantidade de tarefas é sempre maior que a quantidade de recursos para resolvê-las. Neste caso é necessário tomar decisão sobre qual recurso será alocado a uma determinada tarefa (**FRANSISCO, 2012**).

Atualmente Sistemas Computacionais necessitam cada vez mais de recursos de hardware para serem executados, em *Nuvem* são mais comuns à utilização de técnicas de Inteligência Artificial, Redes Neurais, Programação Linear, Teoria de Fila, Teoria de Jogo, modelos baseados em Algoritmos Genéticos para auxiliar nos problemas de alocação automatizada de recursos.

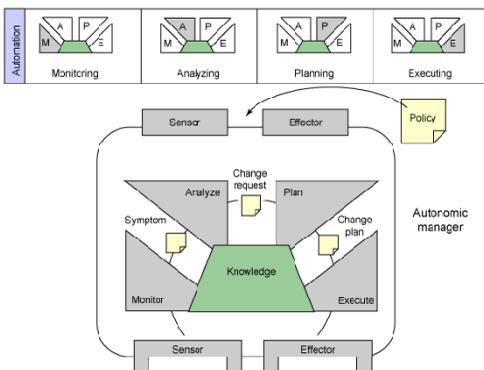
Então a Teoria de Tomada de Decisão é importante também em *CN*, pois ela fundamenta todos os modelos de alocação autônoma de recursos para gerenciamento da infraestrutura de *Nuvem*.

2.3 Autonomic Computing

Autonomic Computing ou computação autônoma em português é um modelo de autogerenciamento que deriva de conceito da biologia. O controle de corpo humano funciona de forma tão autônomo que geralmente não precisa da interferência dos seres humanos. Da mesma forma, o objetivo da Automatic Computing criar sistema que sejam capazes de executar de forma autônoma, mantendo a complexidade do sistema invisível para o usuário (**BUYYA; RAHMAN; RANJAN, 2011**). Sistemas baseados em Automatic computing tomam decisões sobre sua própria política de funcionamento para atingir suas metas, vivificando constantemente, monitora e otimiza o seu status.

Russel e Norvig (2003) propuseram um modelo no qual, uma agente monitora seu ambiente através de sensores, assim ele determina as ações para executar no ambiente.

Figura 11 Arquitetura de Automatic Computing



Fonte: IBM Researches

Segundo **Buyya, Rahman e Ranjan (2011)** um sistema autônomo deve possuir quatro principais características seguintes:

- Autoconfiguração: habilidade de se adaptar e mudar o sistema, por exemplo, instalar um *plugin* quando ele detecta que é pré-requisito para funcionamento de um software do sistema;
- Auto otimização: habilidade de melhorar o desempenho do sistema, por exemplo, ajustar a carga de trabalho atual quando se observa aumento ou diminuição de capacidade;
- Auto cura: habilidade de detectar, diagnosticar e corrigir falha no sistema, por exemplo, reenviar um arquivo quando detecta que foi corrompido durante o primeiro envio;
- Autoproteção: habilidade de antecipar, detectar, identificar e proteger contra intrusos no sistema, tornar um recurso inacessível quando detecta um intruso.

Além desses quatro características, **Buyya, Rahman e Ranjan (2011)** citam também mais outros quatro características de sistemas autônomos seguintes: Autoconhecimento, capacidade de se reconhecer si próprio; O conhecimento do contexto, capacidade de reconhecer em qual contexto ele esta executando; Abertura, capacidade de se comunicar com outros sistemas através de protocolos estabelecidos; Antecipatória, capacidade de antecipar as necessidades do usuário.

2.4 Programação linear

A *Programação Linear* é uma das técnicas da Pesquisa Operacional das mais utilizadas em se tratando de problemas de otimização. Os problemas de Programação Linear buscam a distribuição eficiente de recursos limitados para atender um determinado objetivo, em geral, maximizar lucros ou minimizar custos. Este objetivo é expresso através de uma função linear, denominada de “Função Objetivo” (**CORRAR; GARCIA, 2001**).

Quando se análise um problema, para tentar enquadrá-lo em um modelo de *programação linear* é fundamental que se consiga distinguir, de um lado, quais são as variáveis fora do controle do analista, ou parâmetros, cujos valores já estão fixados, e, de outro, quais são as variáveis de decisão, ou seja, aquelas cujo valor se quer conhecer. A solução de um modelo dará exatamente o valor dessas variáveis de decisão. As variáveis de decisão compõem tanto a função objetivo como as restrições, a função objetivo é uma expressão onde cada variável de decisão é ponderada por algum parâmetro (**CORRAR; GARCIA, 2001**).

As empresas podem fazer uso da *programação linear* para resolver seus problemas de otimização de recursos, seja com o objetivo de redução de custos, otimização de resultados, ou qualquer outro processo de tomada de decisão

gerencial.

2.5 Teoria dos jogos

A *teoria dos jogos* é uma teoria matemática criada para se modelar fenômenos que podem ser observados quando dois ou mais agentes de decisão interagem entre si. Ela fornece a linguagem para a descrição de processos de decisão conscientes e objetivos envolvendo mais do que um indivíduo (SARTINI et al., 2004).

Para **Fábio Portela Lopes de Almeida (2006)** em seu artigo sobre “*A teoria dos jogos: uma fundamentação teórica dos métodos de resolução de disputa*”, o desenvolvimento da *teoria de jogo* iniciou-se no Século XX, em especial após a Primeira Guerra Mundial. Seu objeto de estudo é o conflito, o qual “ocorre quando atividades incompatíveis acontecem. Estas atividades podem ser originadas em uma pessoa, grupo ou nação” Na teoria dos jogos, o conflito pode ser entendido como a situação na qual duas pessoas têm que desenvolver estratégias para maximizar seus ganhos, de acordo com certas regras pré-estabelecidas.

A teoria dos jogos, desde a década de 1940, tem sido de grande utilidade estratégica. No início, como já disposto, a teoria tinha finalidades eminentemente militares, tendo sido utilizada com grande sucesso na II Guerra Mundial e, mais tarde, na Guerra Fria e na Guerra da Coreia. Computadores quânticos, criptografia, a formulação de algoritmos e mesmo a definição de estratégias de investimento na bolsa de valores têm sido consideradas aplicações da teoria dos jogos.

2.6 Inteligência Computacional

Segundo a **ICA (2014)**, a Inteligência Computacional busca, através de técnicas inspiradas na natureza (Redes Neurais, Lógica Nebulosa, Computação Evolucionária, Inteligência Artificial, Sistema Especialista, Data Mining, Neuro Genética, Neuro Fuzzy, Fuzzy Database, Soft Computing), o desenvolvimento de sistemas inteligentes que imitem aspectos do comportamento humano, tais como: aprendizado, percepção, raciocínio, evolução e adaptação.

2.6.1 Redes Neurais

Redes Neurais são modelos computacionais não lineares, inspirados na estrutura e operação do cérebro humano, que procuram reproduzir características humanas, tais como: *aprendizado, associação, generalização e abstração*. Devido à sua estrutura, as Redes Neurais são bastante efetivas no aprendizado de padrões a partir de dados não lineares, incompletos, com ruído e até compostos de exemplos contraditórios.

2.6.2 Lógica Nebulosa

Lógica Nebulosa (Fuzzy Logic) tem por objetivo modelar o modo aproximado de raciocínio humano, visando desenvolver sistemas computacionais capazes de tomar decisões racionais em um ambiente de incerteza e imprecisão. A **Lógica Nebulosa** fornece um mecanismo para manipular informações imprecisas, tais como os conceitos de muito, pouco, pequeno, alto, bom, quente, frio, etc, fornecendo uma resposta aproximada para uma questão baseada em um conhecimento que é inexato, incompleto ou não totalmente confiável.

2.6.3 Computação Evolucionária

É uma área de pesquisa interdisciplinar que compreende diversos paradigmas inspirados no princípio Darwiniano da evolução das espécies. O estágio de pesquisa considera, entre outros, os seguintes paradigmas: *Algoritmos Genéticos*, *Programação Genética*, *Programação Evolucionária*, *Algoritmos Evolucionário*.

2.6.3.1 Algoritmos Genéticos

Algoritmos Genéticos são algoritmos matemáticos inspirados nos mecanismos de evolução natural e recombinação genética. A técnica de *Algoritmos Genéticos* fornece um mecanismo de busca adaptativa que se baseia no princípio Darwiniano de reprodução e sobrevivência dos mais aptos. Isto é obtido a partir de uma população de indivíduos, representados por cromossomas, cada um associado a uma aptidão, que são submetidos a um processo de evolução por vários ciclos.

2.6.3.2 Programação Genética

Programação Genética é uma técnica automática de programação que propicia a evolução de programas de computadores que resolvem (ou aproximadamente resolvem) problemas.

2.6.3.3 Programação Evolucionária

Programação Evolucionária é uma estratégia estocástica de otimização, similar a *Algoritmos Genéticos*, que dispensa a representação genômica de soluções e envolve outros tipos de operadores de mutação. Com *Algoritmos Genéticos*, a representação dos indivíduos é livre e segue a natureza do problema. Indivíduos são submetidos a diferentes tipos de mutação que

simplesmente alteram aspectos da solução de acordo com uma distribuição estatística que pondera variações menores ou maiores conforme os indivíduos se encontrem mais provavelmente próximos ou afastados do ótimo global.

2.6.3.4 Algoritmos Evolucionários

Algoritmos Evolucionários usam modelos computacionais de processos evolucionários como elementos chave no projeto e implementação de sistemas para a solução de problemas. *Algoritmos Evolucionários* compartilham o conceito básico de simulação da evolução de estruturas individuais via o processo de seleção, mutação e reprodução, com base no desempenho das estruturas no ambiente.

2.6.4 Inteligência Artificial

Inteligência Artificial (IA) é uma área da computação que primeiramente trabalha com o processamento simbólico de conhecimento, criando programas que fazem os computadores parecer inteligentes. As soluções dos problemas são heurísticas e respostas satisfatórias são aceitas. *IA* é constituída de técnicas próprias para a solução de problemas.

2.6.5 Sistemas Especialistas

Sistemas Especialistas são programas computacionais destinados a solucionar problemas em um campo especializado do conhecimento humano. Usa técnicas de IA, base de conhecimento e raciocínio inferencial.

2.6.6 Data Mining

Data Mining refere-se ao processo de se “descobrir conhecimento” embutido em um determinado banco de dados. *Data Mining* é a área onde se aplica determinados algoritmos para se extrair padrões dos dados de um banco. Paralelamente, *Knowledge Discovery Databases* refere-se aos processos de se descobrir a utilidade do conhecimento nos dados do banco.

2.6.7 Neuro Genética

Neuro Genética estuda o emprego de *Algoritmos Genéticos* no projeto de redes neurais. Isto envolve dois casos distintos: a otimização da arquitetura da rede e seus respectivos parâmetros (topologia da rede, parâmetros de aprendizado, etc.); e a otimização dos pesos sinápticos, isto é, a utilização de *Algoritmos Genéticos* como algoritmo de atualização dos pesos.

2.6.8 Neuro Fuzzy

Neuro Fuzzy são sistemas híbridos que combinam *Redes Neurais* e *Lógicas Nebulosas* para incorporar num só sistema o conhecimento explícito de especialistas e o conhecimento implícito inerente a um conjunto de dados.

2.6.9 Fuzzy Database

Fuzzy Database são sistemas que visam tornar mais maleável o modelo de banco de dados relacionais. Estes sistemas buscam formas de introduzir características "fuzzy" no banco de dados tais como, pesquisa nebulosa (fuzzy queries) e armazenamento de informações imprecisas.

2.6.10 Soft Computing

Soft Computing distingue a computação convencional baseada na lógica binária, da computação denominada "soft", baseada na *lógica nebulosa*, *redes neurais*, *raciocínio probabilístico*, *algoritmos genéticos*, *teoria do caos* etc. Através da Computação Soft, o sistema computacional torna-se tolerante à imprecisão, incertezas e verdades parciais.

2.7 Gerenciamento da Infraestrutura de Nuvem

O gerenciamento da infraestrutura de Nuvem inclui organização das atividades de servidor, rede e armazenamento, gerenciamento de configuração, monitoramento de desempenho, gerenciamento de recursos de armazenamento e medição de uso. É importante a utilização de um gerenciador de infraestrutura de Nuvem, responsável pelo ciclo de vida das VMs e pelo monitoramento dos recursos computacionais disponíveis.

O gerenciador disponibiliza aos usuários um ponto de controle central, através do quais novos recursos são adicionados, VMs são criadas, redes virtuais são configuradas e novas imagens são disponibilizadas. Este gerenciamento pode ser feito pela empresa provedora ou pelo usuário, dependendo do modelo da Nuvem adotado.

2.7.1 Gerenciamento em nível da empresa provedora

O Gerenciamento em nível da empresa provedora se resume a gestão de VMs, e Servidores físicos para uso eficiente de recursos. Entre as operações de gerenciamento em nível da empresa provedora, as mais comuns são: alocação

de recursos, implantação de servidores e de VMs, configuração e inicialização de VMs e servidores e monitoração.

2.7.1.1 Alocação de Recursos

É a primeira operação realizada em Nuvem. A provedora atribui a Empresa Cliente sua porção de recursos. Este recurso é um conjunto de memória, processador, espaço de armazenamento, banda passante e hardware. A política de alocação de recurso em Nuvem depende da estratégia comercial do fornecedor. Como os recursos concorridos entre vários usuários o fornecedor pode propor várias maneiras de reservação de recursos (TCHANA; BROTO; HAGIMONT, 2012).

Reserva para duração indeterminado: deve ser estabelecido no SLA que a empresa Cliente tem recurso disponível por uma duração indeterminada e de forma contínua.

Reserva para usuário futuro e com duração limitada: neste caso, a dificuldade do gerenciamento das reservas.

Reserva para usuário atual com duração limitado: neste caso, o recurso deve ser disponibilizado imediatamente.

2.7.1.2 Configuração e inicialização

A operação de iniciar de VMs necessita da configuração das mesmas. A configuração depende em parte da demanda da Empresa Cliente e em outra parte da política de alocação de recursos em Nuvem. Durante a fase de reserva de recursos, a Empresa cliente fornece as características de VMs que ela vai querer. Estas características são ligadas a: quantidade de memória, o número de processadores, o lugar onde vão executar as VMs e o Sistema de Arquivos, SO de VMs. Outras configurações como de rede, por exemplo, virão do IaaS. De fato, o IaaS pode implementar várias configurações de acesso a rede, por exemplo a atribuição de um rede virtual (VLAN) aos VMs pertencendo a mesma Empresa. O IaaS deve também configurar a rede em função do pedido da Empresa Cliente.

2.7.1.3 Reconfiguração

A maioria das tarefas de gerenciamento em nível de IaaS podem ser interpretados como operação de reconfiguração. Nesta sessão apresentaremos algumas tarefas de reconfiguração ligadas à gestão de VMs: a Consolidação e Reparação.

Consolidação de VM: permite ao fornecedor implementar diferentes políticas de alocação e redistribuição de recursos com objetivo de economizar ou respeitar o SLA. A consolidação pode ser também feita fora das operações de alocação de recursos. A IaaS pode reorganizar a disposição de suas VMs em

maquinas físicas para liberar algumas, para redução da taxa de consumo de energia.

Reparação: levando em conta a complexidade da infraestrutura de Nuvem e o numero de usuários, o risco de aparição de colapso é importante. A aparição de um erro de funcionamento deve ser rapidamente detectada e corrigida para não penalizar a Empresa.

2.7.1.4 Monitoração

Todas as tarefas de gerenciamento da infraestrutura de Nuvem dependem das informações obtidas na monitoração nos ambientes físicos, virtuais e softwares. Entre as informações de monitoração que interessa nos podemos citar a taxa de uso de processadores, de discos, da rede, da memória, etc.

CN é um sistema distribuído e heterogêneo. Na maioria das vezes, o administrador se interessa as informações agregadas e obtidas por cálculos de observações locais. Por exemplo, o calculo pode ser feito combinando informações vindas de um conjunto de maquinas pertencentes à mesma zona geográfica ou mesma Empresa.

2.7.2 Gerenciamento em Nível de Empresa Cliente

Mesmo com as vantagens que ele oferece CM tem ainda dificuldade de seduzir as Empresas desde a sua popularização. Este receio de ordem ideológica: a ideia de confiar seus dados a um terceiro ainda não é aceito facilmente nas Empresas. Frente a esta ideia, CN responde com mais desenvolvimento em segurança e confiabilidade. O gerenciamento de aplicações em Nuvem é uma tarefa da empresa fornecedora, ela inclui as tarefas seguintes: Criação de VMs; Reserva e Alocação se recursos; Instalação de inicialização de Softwares; Acompanhamento da execução de Softwares e VMs; Reconfiguração e Otimização.

2.7.2.1 Criação de VMs e Implantação

A execução de todas as aplicações em Nuvem e feito em VMs. Esta execução requer que as aplicações estejam hospedadas em Nuvem. Para Empresa cliente a implantação pode ser feito em duas ocasiões: durante o carregamento de sistema de arquivos dos VMs e durante a gravação dos binários de suas aplicações. A construção de sistema de arquivos contem as seguintes fases: instalação do SO, configuração do SO e a instalação dos pacotes ou binários dos futuros softwares.

2.7.2.2 Alocação de Recursos

A Alocação de Recursos para empresa cliente consiste em reservar local para armazenamento ou reservar VMs. A primeira opção permite a armazenar a Imagem do SO criado ou de dados a serem usados pelas VMs. As VMs são visto pela Empresa cliente como maquina físicas disponibilizadas pela Nuvem de sua propriedade. Esta operação é feito em forma de contrato detalhando cada recurso solicitado (processador, memória, banda passante etc.) e sua disponibilidade. O fim do contrato gera a parada de VMs e liberação de recursos que foi alocado para Empresa.

2.7.2.3 Configuração e inicialização

Uma vez que as VMs iniciadas e os binários de softwares instalados, o administrador da Empresa deve planejar a configuração e a inicialização dos softwares. Esta operação necessita o acesso as VMs via IP ou nome DNS das VMs. Em alguns casos, a inicialização ode necessitar de configuração de meio de comunicação entre softwares situados em VMs diferentes. Por causa de múltiplos softwares, estas operações são às vezes sujeitas a erros. Além disso, certas aplicações impõem uma ordem de inicialização entre softwares. É o caso dos servidores web Apache em aplicações J2EE.

2.7.2.4 Reconfiguração

As operações de reconfiguração realizada na Empresa cliente são infinitas. Nessa sessão, apresentaremos duas operação que são mais comuns no gerenciamento de aplicações distribuídas:

Escalabilidade: uma vez as aplicações inicializadas, a Empresa deve ser capaz de seguir o estado de diferentes softwares enfim de intervir caso necessita.

Reparação: dois tipos de problema podem ocorrer durante a execução da aplicação: um erro na VM ou um erro no software. Sem ter nenhuma ação reparadora sobre as suas VMs, a Empresa cliente deve solicitar ação de reparo para o fornecedor. Neste caso a Empresa cliente pode ser levada a fazer a reimplantação e inicialização dos seus softwares.

2.7.2.5 Monitoração

A maioria das informações de monitoração recolhidos vem diretamente dos API da IaaS. Assim, em certos casos esta informações são insuficientes. Por exemplo, determinação do estado do servidor de base de dados é mais significativo quando se observa o seu tempo de resposta ao em vez da carga de CPU do servidor que o hospeda. Neste caso em partícula o administrador coloca em suas aplicações uns espece de agentes pra recolher informações que vão ser usadas para tomada de decisão para Reconfiguração.

2.7.3 Alocação Dinâmica de Recursos

Entender a migração de processo é mais importante para abordagem da técnica de alocação dinâmica de recursos. Um processo é uma abstração do SO que representa uma instância de um programa de computador em execução. A migração de processos é o ato de transferir um processo de uma máquina para outra durante a sua execução. A técnica de migração de processos permite distribuir cargas dinamicamente, migrando processos de máquinas sobrecarregados para máquinas com menos carga (**MILOJICIC; LORENTE; MONTERO, 2011**).

Os objetivos de se utilizar a migração de processos estão sempre vinculados ao tipo de aplicação que utilizará a migração. Segundo **Milojicic, Lorente e Montero (2011)**, os objetivos principais da migração de processos são:

- Acesso a maior capacidade de processamento: A utilização da capacidade de processamento das máquinas de uma *data center* é variável no tempo. Desta forma, pode-se utilizar a migração de processos para aproveitar de forma mais eficiente os recursos disponíveis, migrando processos para máquinas que possuam capacidade de processamento ociosa;
- Exploração da localidade de recursos: Em determinadas aplicações, a localidade dos dados manipulados exerce papel fundamental no desempenho das aplicações. Desta forma, através da migração de processos, pode-se mover um processo para uma máquina próxima aos dados que estão sendo operados, reduzindo a latência de acesso e a utilização de recursos de rede;
- Resiliência a falhas: A resiliência a falhas pode ser melhorada através da migração de processos de nós que apresentam falhas parciais, ou em situações onde aplicações que executam por longos períodos de tempo, quando diferentes tipos de falha como falhas em algum disco ou memória ou falhas na rede podem ocorrer;
- Administração do sistema: A migração de processos facilita o processo de administração dos sistemas computacionais. Podem-se transferir temporariamente processos de uma máquina que precisará ser desligada para a manutenção e após a manutenção o processo é migrado de volta para a máquina de origem;
- Computação móvel: A computação móvel pode ser beneficiada com a utilização da migração, pois as aplicações podem migrar dinamicamente de forma a acompanhar o usuário da aplicação independente de sua localização física.
- Aplicações de longa duração: As aplicações que executam por longos períodos de tempo como dias e semanas podem sofrer diferentes interrupções como a

falha parcial de máquinas ou desligamentos de origem administrativa. A migração de processos pode realocar estas aplicações de forma transparente para evitar as interrupções;

- Múltiplas cargas de trabalho genéricas: Em alguns ambientes como os ambientes das universidades, um conjunto de cargas de trabalhos distintas e não previsíveis pode ser executado ao mesmo tempo. Desta forma, conforme os usuários submetem tarefas e executam aplicações, a carga em cada um dos nós individuais dos sistemas varia muito. A migração de processo permite a distribuição destas cargas entre todos os nós;

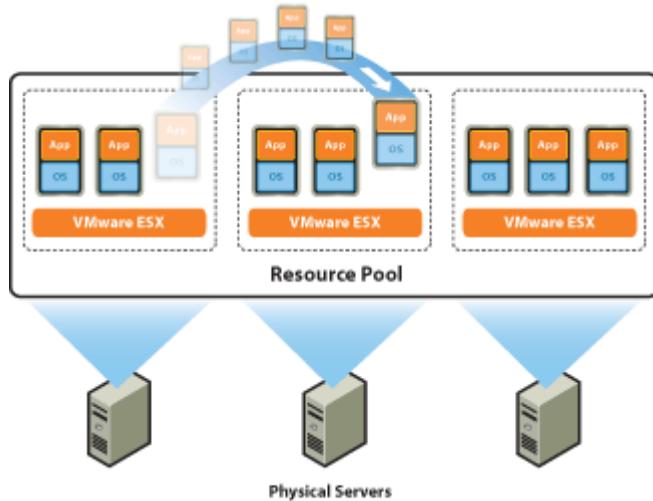
- Aplicações de rede: As aplicações de rede são umas das aplicações que mais podem se beneficiar da migração de processos. Agentes móveis voltados para a mobilidade podem utilizar a migração para aproximar os dados dos usuários ou para acompanhar os usuários conforme eles se movimentam na rede.

2.7.3.1 Migração de Máquina Virtual

A migração de VMs é uma técnica bastante semelhante à migração de processo. A principal diferença é que a migração de VMs envolve migração de SO completo com todas as aplicações em execução por uma máquina física. Então podemos dizer que a migração de VMs é uma técnica que move VMs de uma máquina física para outra (**CLARK; FRASER; HAND, 2005**).

Em geral deve se considerar três estado para que a migração seja realizada: o estado dos dispositivos virtuais, as conexões externas com os dispositivos e a memória física da VM. Todos estes estados devem permanecer o mesmo após a migração. A consistência de sistema de arquivo é outra característica importante na migração de VMs. Ao termino da migração é necessário que o sistema de arquivos no computador de destino idêntico ao sistema de arquivo no computador de origem (**NELSON, LIM, & HUTCHINS, 2005**).

Figura 12 Migração de Maquinas Virtual



Fonte: (DELGRADO, 2009)

A migração de VMs em ambientes virtualizados de duas formas principais: a migração ao vivo baseada em pré-cópia e a migração ao vivo baseado em pós-cópia (Hines & Gopalan, 2009).

A migração ao vivo baseado em pré-cópia é a migração que a maioria das plataformas de virtualização como Xen, KVM e o VMWare adotaram. No processo de pré-cópia, antes da VM a ser migrado começar a executar na máquina de destino, a sua memória é totalmente replicado na máquina de destino (Hines & Gopalan, 2009). A estimacão de tempo de migração do mecanismo pré-cópia é definido por

$$TempoTotal = \frac{TamanhoDaMemória}{Banda} + \beta$$

Onde β é um parâmetro que reflete a taxa de atualização da memória que pode variara muito, dependendo caso pode inviabilizar a conclusão da migração.

O mecanismo de migração ao vivo baseado em pós-cópia parte do pressuposto que é mais eficiente transferir as paginas de memória após o resumo da VM na máquina de destino. A VM é suspensa na máquina de origem ao iniciar o processo de migração e o conteúdo dos registradores dos processadores virtuais e os estados dos dispositivos são copiados na máquina de destino. Assim a VM é copiada na máquina de destino sem nenhuma pagina de memória. Caso a VM necessita de uma pagina que foi copiada ainda, ela é

pausada até a página seja copiada (HINES; GOPALAN, 2009). O tempo total de migração é dado por

$$TempoTotal = \frac{\text{TamanhoDaMemória}}{\text{Banda}}$$

Dependendo somente do tamanho da memória e da banda disponível.

2.7.3.2 Elasticidade

A Elasticidade é uma das maiores qualidades da Nuvem. Ela pode ser assimilada como a capacidade da Nuvem de prover recursos dinamicamente, de acordo com a demanda da Empresa cliente.

A Elasticidade é vista de forma diferente pelo consumidor do serviço de Nuvem e pela empresa provedora de Nuvem. O usuário de serviço de Nuvem não precisa entender as tecnologias que fazem a Elasticidade acontecer. Ele interage apenas pelo portal de acesso no qual solicita provisionamento e alocação de recursos computacionais. A Elasticidade aparece em Nuvem Pública, quando o usuário não precisa dispor de nenhuma capacidade por si. Usa apenas serviços provisionados pela Nuvem, de acordo com sua demanda. Já na Nuvem privada, a empresa precisa adquirir e implementar os recursos computacionais que serão alocados. A elasticidade a Elasticidade é sentida apenas no nível de usuários internos, mas não no Data Center.

2.7.4 Sistemas de Gerenciamento autônomo para Nuvem

O gerenciamento em Nuvem é essencialmente ligado a sua utilização em nível da empresa provedora e em nível da Empresa cliente. Nestes dois níveis, encontramos operações de gerenciamento do mesmo tipo. Em nível da empresa provedora o gerenciamento consiste em fazer gerenciamento de recursos. Quanto a Empresa cliente, suas tarefas administrativas são de origem das configurações em nível da empresa provedora.

Frente a este número de tarefas, oferecer uma solução estática para gerenciamento será inapropriado em certas situações. Além disso, nem a provedora, nem a Empresa cliente podem disponibilizar mão de obra humana para assumir de forma permanente e contínua todas essas tarefas. Assim o mais adequado é usar Sistemas de Gerenciamento Autônomo para melhorar o gerenciamento da infraestrutura. Este sistema deve geralmente preencher as seguintes condições (HINES; GOPALAN, 2009):

Interoperabilidade: capacidade de se comunicar com outros sistemas de gestão da Nuvem e sistemas de gerenciamento de aplicações que ele hospeda.

Auto reparável: capacidade de corrigir erros no ambiente que ele administra.

Extensível: capacidade de gerenciar novos recursos (Máquinas físicas, VMs, Softwares) no seu ambiente.

Programável: capacidade de assimilar novas políticas de gerenciamento (novo método de consolidação, de escalonamento, etc.).

Adaptável: capacidade de assimilar as substituições e modificações de um de seus módulos.

Linguagem dedicada: proporcionar facilidade de utilização através de uma linguagem dominada por partes interessadas.

3. UMA TAXONOMIA PARA CLASSIFICAÇÃO DE MÉTODOS PARA GERENCIAMENTO AUTÔNOMO DE RECURSOS EM CLOUD

4.1 Metodologias de pesquisa para chegar à amostra piloto

Fizemos um levantamento de trabalhos que abordam o tema gerenciamento de recursos em Nuvem. Este levantamento foi organizado da seguinte forma:

Criamos uma amostra inicial com sessenta artigos pesquisados na Internet com a palavra *Cloud Computing*, com esta amostra, para os artigos que abordam gerenciamento de recursos, selecionamos as palavras chaves para elaborar a nossa expressão de busca.

Assim, as palavras chaves selecionados Segundo o interesse da nossa pesquisa foram: *decision making* (tomada de decisão), *decision-theoretic* (decisão teórica), *decision theory* (teoria de decisão), *decision-theoretic planning* (planejamento decisão teórico), *managmente* (gerenciamento), *automatic* (automático), *autonomic* (autônômica), *self-configuration* (autoconfiguração), *self-managmente* (autogerenciamento), *control theory* (teoria de controle), *adaptative* (adaptative), *recursos alocation* (alocação de recursos), *automatic policy* (política automática), *cloud computing* (computação em nuvem) e *grid computing* (computação em grade).

Com isso a nossa string de busca ficou da seguinte forma: (“*decision making*” OR “*decision-theoretic*” OR “*decision theory*” OR “*decision-theoretic planning*” OR “*managmente*” OR “*automatic*” OR “*autonomic*” OR “*self-configuration*” OR “*self-managmente*” OR “*control theory*” OR “*adaptative*” OR “*recursos allocation*” OR “*automatic policy*”) AND (“*cloud computing*” OR “*grid computing*”).

Com nossa expressão de busca elabora, usamos ela para fazer uma busca nos repositórios de trabalhos científicos seguintes:

Google Acadêmico, que fornece de maneiras simples, diversas livros, artigos, teses, resumos e disciplinas de forma bastante abrangente;

IEEE Xplore Digital Library, que é um banco de dados de pesquisas acadêmico que fornece artigos e trabalhos sobre a ciência da computação, engenharia elétrica e eletrônica;

ScienceDirect, que é um site operado pela editora *Elsevier* que fornece trabalhos científicos agrupados em quatro sessões principais: ciências físicas e engenharia, ciências biológica, ciências de saúde e ciências sociais e humanas;

ACM Digital Library, que é um bando de dados de trabalhos científicos criado pela *Assosiation for Computing Machinery* (AMC) que contem todos os artigos publicados pela AMC;

Springer, que é uma editora baseada na Alemanha que publica livros de livros textos, livros referencia acadêmica e artigos com foco em ciências, tecnologia, matemática e medicina.

Em cada um desses repositórios citado acima, fizemos uma busca no dia 10 de abril de 2014 dividida em dois períodos, primeiramente buscamos os trabalhos produzidos somente em 2014 e em seguida buscamos trabalhos produzidos entre 2010 e 2014 usando o nosso expressão de busca e selecionamos os primeiros duzentos artigos para cada repositório de busca citado acima, por ordem de relevância pela classificação do repositório, com total de cem trabalhos para cada período. Isso nos leva a uma amostra piloto com total de mil trabalhos, incluído artigos, capítulos de livros e citações.

Coma a amostra piloto pronta, o primeiro passo, aplicamos um filtro para eliminando os artigos duplicados gerando uma nova amostra de novecentos cinquenta sets trabalhos. Em seguida os trabalhos que não tinham “cloud” ou “grid” no titulo foi excluído, gerando um novo conjunto de quinhentos oitenta e seis trabalhos.

Aplicamos mais filtro analisando os títulos dos trabalhos e os trabalhos considerados com características seguintes:

- *Generic cloud definitions*: os trabalhos que abordam o funcionamento da Nuvem como, por exemplo, “*A view of Cloud Computing*” (**ARMBRUST, et al., 2009**), serão considerados amplos de mais, por isso serão descartados da amostra;
- *Security*: a segurança foi considerada como uma subárea do gerenciamento, então os trabalhos que apresentam tomadas de decisão e segurança serão descartados;
- *Cloud and Grid application in knowledge areas*: os trabalhos que contem “utilizing cloud” no seu titulo como, por exemplo, “*Mobile healthcare information management utilizing Cloud Computing and Android OS*” , (**DOUKAS; PLIAKAS; MAGLOGIANNIS, 2010**) serão considerados como aplicação do CN e Computação em Grade e devem ser descartados;
- *Cloud managers*: trabalhos como “*OpenNebula: A Cloud Managment Tool*” (**MILOJICIC; LIORANTE; MONTERO, 2011**), serão considerados específicos de mais, mesmo abordando algoritmos de decisão que pode ser considerado importante, a implementação utilizando ferramentas de gestão da Cloud estão fora do foco deste trabalho;
- *Trust management*: mesmo abordando tema relacionado diretamente com tomada de decisão, o gerenciamento de confidencialidade foi considerado muito específico e aplicado à segurança da informação e controle de acesso, assim os trabalhos sobre gerenciamento de confidencialidade serão descartados;
- *Content management*: os trabalhos onde o tema central é gestão de conteúdo como, por exemplo, “*Media cloud: an open cloud computing middleware for*

content management” (DIAZ-SANCHEZ et. al, 2011), serão também descartados;

- *General cloud surveys*: trabalhos sobre pesquisas gerais sobre Cloud Computing como, por exemplo, “*Cloud computing: state-of-the-art and research challenges*” (ZHANG; CHENG; BOUTABA, 2010) serão considerados fora do foco do nosso trabalho;
- *Simulators*: trabalhos apresentando simulador foram excluídos.
- *Data management*: os trabalhos como, por exemplo, “*Challenges for Mobile Data Management in the Era of Cloud and Social Computing*” (BISDIKIAN et. al, 2011) serão considerados como abordar tema de gerenciamento de dados em Cloud;
- *Smart grid*: trabalhos sobre funcionamento de grid foram excluídos.

Com os critérios de exclusão estabelecidos, aplicamos mais um filtro nos quinhentos oitenta e seis trabalhos, fazendo a leitura dos resumos dos mesmos. Com isso conseguimos reduzir a amostra para trezentos oitenta e quatro trabalhos. Depois aplicamos novamente o mesmo filtro, mas desta vez fazendo a leitura dos trabalhos por inteiro e conseguimos reduzir a nossa amostra por duzentos noventa e set trabalhos.

4.3 Taxonomias propostas

Por fim nosso trabalho, vendo tempo disponível para realizar o nosso trabalho, decidimos classificar dez por centos do total dos modelos propostas nos duzentos noventa e set modelos, isso nos dá um total de trinta modelos a ser classificados, escolhendo os mais atuais que apresentam experimentos, resultados e análise de resultados e que são melhores formados segundo a nossa classificação. Também tivemos que estabelecer critérios para nos servir de base na classificação desses modelos. Assim levamos em conta as características seguintes:

- Definição de Nuvem computing em que o modelo se baseia;
- As necessidades de gerenciamento apontadas no modelo;
- Se o modelo resolve problema de satisfação ou de otimização;
- Critérios visada para satisfazer ou otimizar;
- Parâmetros de entrada na tomada de decisão;
- Camadas que vão influenciar a tomada de decisão;
- Numero de decisores;
- Numero de critérios;
- Numero de decisões;
- Previsão de mudança no tempo;
- Formalismo em que se o modelo baseia;
- Métricas medidas no experimento;
- Resultado final dos algoritmos de tomada de decisão;
- O aspecto da Nuvem abordado no modelo.

a) Definição de CN em que o trabalho se baseia

A primeira classificação é com relação a qual definição de CN o modelo se baseia. Para nosso trabalho, dividimos esta classificação quatro categorias: i) modelos definindo CN baseado se na definição do **Notional Institute of Standards and Technology (NIST)**; ii) baseado na definição do **Rajkumar Buyya**; iii) baseado na definição do **Michael Armbrust**; iv) baseado em outras definições de CN.

A definição do **NIST** lista cinco característica essencial para CN: autosserviço sob demanda, acesso a rede de banda larga, pool de recursos, rápida elasticidade ou expansão e serviço de mensuração. A definição também cita três modelos de serviços (SaaS, PaaS e IaaS) e quatro modelos de desenvolvimento (privado, comunitário, publico e híbrido) que junto, categorizam modos de entrega de serviços em Nuvem (**MELL, 2011**).

Para **Buyya et al. (2008)**, CN é um sistema distribuído e paralelo, constituído por uma coleção de computadores virtuais e físicos interconectados que proveem uma ou mais recursos computacionais, baseados em acordos de SLAs preestabelecidos através de uma negociação entre consumidor e o provedor.

Já para **Armbrust (2009)** CN é um conjunto de serviços de rede ativados, proporcionando escalabilidade, qualidade de serviço, infraestrutura barata de computação sob demanda e que acessada de forma simples.

Tabela 2 Definição em que os modelos se baseiam

Nº	Modelos	i	ii	iii	iv
1	(RAVINDRAN, 2013), (KURIBAYASHI, 2011), (YOU, 2011), (WUHIB; STADLER; LINDGREN, 2012), (WEI; BLAKE; SALEH, 2013), (FITÓ et al., 2012)	X			
2	(WU; KUMAR; BUYYA, 2011), (WEI; VASILAKOS; ZHENG, 2010), (GOUDARZI; PEDRAM, 2011), (FENG et al., 2012), (ERGU et al., 2011)		X		
3	(BELOGLAZOV; ABAWAJY; BUYYA, 2011), (ESPADA et al., 2013), (LIN et al., 2011), (NGUYEN et al., 2013), (CHUNLIN; LAYUAN, 2013)			X	

4	(WU et al., 2012), (ZHANG; ZHU; BOUTABA, 2011), (RAMEZANI; LU; HUSSAIN, 2013), (GARG; BUYYA; CALHERIROS, 2011), (BILGAIYAN; SANGNIKA; DAS, 2014), (DI; KONDO; CIRNE, 2012), (AN et al., 2010), (KUMAR et al., 2011), (RAO et al., 2011), (GUAN; FU; ZHANG, 2012), (GE et al., 2012), (DHINGRA; PAUL, 2014), (ENDO; SADOK; KELNER, 2011), (TAI et al., 2011)				X
---	--	--	--	--	---

Analisando os modelos propostas na nossa amostra piloto estabelecidas acima, percebemos que a maioria dos autores, quase cinquenta por centos, não se baseiam em uma definição de CN. Chegamos a uma conclusão que isso pode ser devido à popularização de CN, o autor consideram que os leitores estão cientes de o que é a CN. Relembrando que esta conclusão não pode ser generalizada, já que examinamos somente dez por centos da amostra inicial.

b) Necessidade de gerenciamento de recursos

Para fornecer serviço de qualidade aos clientes, os provedores de serviços de Nuvem precisam adotar estratégia de gerenciamento de suas infraestruturas cada um segundo duas necessidades.

Pesquisando sobre gerenciamento de recurso em Nuvem, nos no deparamos com varias abordagens que os trabalhos científicos apontam como necessidade para gerenciar a infraestrutura da Nuvem, como por exemplo, minimizar a violação do SLA, melhorar o QoS, otimizar o uso de recursos (alcançar o melhor desempenho, com sobrecarga aceitável), otimizar os lucros, gerenciar a infraestrutura de forma eficiente com menor custo possível, economizar o consumo de energia para minimizar impactos ecológicos, maximizar satisfação do usuário.

A nossa proposta é de dividir eles em quatro categorias seguintes:

- i) Resolve problema satisfação de usuário:** onde se enquadra trabalhos que propõem modelos que visa satisfazer problemas relacionados à SLA e QoS para satisfação do usuário;
- ii) Resolve problemas da infraestrutura:** onde se enquadra trabalhos que propõem modelos melhorar o gerenciamento da infraestrutura;
- iii) Resolve problemas de negócio:** onde se enquadra todos os trabalhos que propõem modelos que visam à maximização de lucro;

iv) **Resolve problema ecológico:** onde se enquadra os trabalhos que propõem modelos para minimizar o consumo de energia e diminuir o impacto ambiental.

Tabela 3 Necessidades de gerenciamento apontado nos modelos

N°	Modelos	Necessidades
1	(BELOGLAZOV; ABAWAJY; BUYYA, 2011)	Minimizar consumo de energia total, Minimizar migrações das VMs, Minimizar a violação do SLA
2	(WU; KUMAR; BUYYA, 2011)	Gerenciar mudança dinâmica dos requisitos dos clientes, Minimizar a violação do SLA, Maximizar lucros, Minimizar custos
3	(WU et al., 2012)	Reduzir o custo de provisionamento de recursos, Reduzir a violação do SLA
4	(ZHANG; ZHU; BOUTABA, 2011)	Maximizar lucros, Otimizar a alocação de recursos, Controle do preço, Minimizar o consumo de energia, Minimizar o tempo de espera
5	(RAMEZANI; LU; HUSSAIN, 2013)	Minimizar o tempo de execução das tarefas, Minimizar o consumo de energia e custo, Otimizar da carga de trabalho e alcançar um melhor equilíbrio de carga, Prevenção de sobrecarga na rede, Otimizar a utilização da Nuvem, Diminuir o tempo de inatividade das VMs
6	(WEI; VASILAKOS; ZHENG, 2010)	Maximizar lucro, Alocação eficiente de recursos
7	(GOUDARZI; PEDRAM, 2011)	Otimizar o lucro total esperado, Reduzir a violação do SLA
8	(RAVINDRAN, 2013)	Melhorar o tempo de resposta (QoS), Tolerância à falha, Avaliação do SLA, Melhorar a confiabilidade
9	(KURIBAYASHI, 2011)	Minimizar a probabilidade de perda de requisições, com isso reduziria a

		quantidade de recursos a ser alocado, Alocação justo de recursos para varias usuários sem comprometer a eficiência
10	(GARG; BUYYA; CALHERIROS, 2011)	Minimizar violação do SLA, Maximizar lucros e minimizar custos relacionados à infraestruturra, Melhorar o Tempo de resposta
11	(YOU et al, 2011)	Atingir o equilíbrio no preço cobrado pelo serviço
12	(BILGAIYAN; SANGNIKA; DAS, 2014)	Melhorar a taxa de utilização de recursos, Minimizar o tempo de resposta
13	(DI; KONDO; CIRNE, 2012)	Manter SLA, Minimizar custo, Prevendo carga no ambiente de Nuvem
14	(WUHIB; STADLER; LINDGREN, 2012)	Satisfazer o cliente conforme o SLA, Minimizar consumo de energia, Satisfazer diferentes objetivos do provedor
15	(WEI; BLAKE; SALEH, 2013)	Maximizar uso de recursos, Manter SLA cumprido
16	(FITÓ et al, 2012)	Minimizar impactos ecológicos, Minimizar consumo energético Maximizar satisfação do usuário, Maximizar lucro,
17	(AN et al., 2010)	Negociação de preço entre provedores e clientes para uma cobrança justa de serviços oferecida, Alocação eficiente de recursos
18	(ESPADA et al., 2013)	Utilização eficiente de recursos computacionais
19	(KUMAR et al., 2011)	Alocar recursos de forma eficiente, Minimizar custos econômicos, Cumprimento de prazos (SLA)
20	(LIN et al., 2011)	Otimizar a decisão de realocação dos recursos, Reduzir o custo de utilização do usuário,
21	(FENG et al., 2012)	Maximizar lucros, Otimizar a alocação de recursos no

		ambiente de nuvem, Melhorar o QoS, Preservar o SLA
22	(RAO et al., 2011)	Melhorar a desempenho do ambiente de nuvem, gerenciando melhor os recursos.
23	(GUAN; FU; ZHANG, 2012)	Gerenciamento proativo de falha em ambiente de nuvem
24	(GE et al., 2012)	Reduzir o consumo total de energia da infraestrutura de nuvem, Reduzir o custo de utilizadores (preço de serviço),
25	(DHINGRA; PAUL, 2014)	Otimizar a alocação de recursos, Reduzir o consumo de energia, Minimizar a violação do SLA
26	(NGUYEN et al., 2013)	Gerenciar a infraestrutura de forma mais eficiente, Eliminar falhas, Poupar energia, Reduzir custos
27	(ENDO; SADOK; KELNER, 2011)	Otimizar a tomada de decisão sobre uso de recursos computacionais
28	(ERGU et al, 2011)	Alocação de recursos de forma dinâmica e eficiente
29	(TAI et al., 2011)	Equilibrar as cargas de trabalho para melhorar o desempenho geral do sistema
30	(CHUNLIN; LAYUAN, 2013)	Otimização global de uso de recursos para o provedor nas camadas IaaS e SaaS, para o usuário na camada de aplicação, Maximizar lucros

Tabela 4 Necessidades de cada categoria

I	ii	iii	iv
<ul style="list-style-type: none"> - Comprimento do SLA - Qualidade de Serviço (QoS) - Uso eficiente de recursos - Diminuir custos de serviço 	<ul style="list-style-type: none"> - Aperfeiçoar o gerenciamento da infraestrutura - Melhorar o tempo de resposta das tarefas - Alocação eficiente de recursos - Diminuir custos da infraestrutura 	<ul style="list-style-type: none"> - Maximizar lucros - Negociação de preço entre provedor e cliente - Diminuir custos relacionados à infraestrutura 	<ul style="list-style-type: none"> - Diminuir o consumo de energia - Diminuir o impacto da infraestrutura sobre o meio ambiente

Tabela 5 Combinação de modelos e necessidades

Numero	Modelos	Categorias
1	(BELOGLAZOV; ABAWAJY; BUYYA, 2011)	i, ii, iv
2	(WU; KUMAR; BUYYA, 2011)	i, ii, iii
3	(WU et al., 2012)	i, ii
4	(ZHANG; ZHU; BOUTABA, 2011)	ii, iii, iv
5	(RAMEZANI; LU; HUSSAIN, 2013)	i, ii, iii, iv
6	(WEI; VASILAKOS; ZHENG, 2010)	iii
7	(GOUDARZI ;PEDRAM, 2011)	i, iii
8	(RAVINDRAN, 2013)	i, ii, iii, iv
9	(KURIBAYASHI, 2011)	ii
10	(GARG; BUYYA; CALHERIROS, 2011)	i, ii, iii
11	(YOU et al., 2011)	iii
12	(BILGAIYAN; SANGNIKA; DAS, 2014)	ii
13	(DI; KONDO; CIRNE, 2012)	i, ii, iii

14	(WUHIB; STADLER; LINDGREN, 2012)	i, ii, iv
15	(WEI; BLAKE; SALEH, 2013)	i, ii
16	(FITÓ et al., 2012)	i, ii, iii, iv
17	(AN et al., 2010)	ii, iii
18	(ESPADA et al., 2013)	ii
19	(KUMAR et al., 2011)	i, ii, iii
20	(LIN et al., 2011)	ii
21	(FENG et al., 2012)	i, ii, iii
22	(RAO et al., 2011)	ii
23	(GUAN; FU; ZHANG, 2012)	ii
24	(GE et al., 2012)	i, iv
25	(DHINGRA ; PAUL, 2014)	i, ii, iv
26	(NGUYEN et al., 2013)	ii, iii, iv
27	(ENDO; SADOK; KELNER, 2011)	ii
28	(ERGU et al., 2011)	ii
29	(TAI et al., 2011)	ii
30	(CHUNLIN ; LAYUAN, 2013)	i, ii

Analisando os modelos propostas, deu para perceber a correlação entre as categorias proposta no critério sobre necessidades de gerenciamento e as necessidades de gerenciamento apontado nos modelos.

Por exemplo, **Feng et. al (2012)** no artigo “*Revenue Maximization Using Adaptive Resource Provisioning in Cloud Computing Environments*” (Maximização de Lucro Usando Provisionamento Adaptativa de Recursos em Ambiente de Nuvem) que propõe um modelo baseado no SLA para maximizar os lucros da empresa provedora de serviço fazendo uma alocação otimizado de recurso, considerando o QoS, o mecanismo de fixação de preço, a taxa de chegada e disponibilidade de serviço. Com isso este modelo foi classificado como:

- i) Resolve problema satisfação de usuário: já que é o modelo baseado no SLA e implementa um mecanismo de fixação de preço de serviço;
- ii) Resolve problemas da infraestrutura: pois modelo propõe uma alocação otimizada de recursos;
- iii) Resolve problemas de negócio: pois o modelo, nos seus resultados aponta um aumento de até 65,9 % de lucro para a empresa provedora de serviço.

Podemos citar também o exemplo de **Nguyen et. al (2013)**, no seu artigo “*An Improvement of Resource Allocation for Migration Process in Cloud Environment*” (Uma Melhoria da Alocação de Recursos para Processos de

Migração em Ambiente de Nuvem) apresenta uma estratégia eficaz utilizado para processo de migração com objetivo de manter em nível alto a disponibilidade de recursos, reduzindo assim o tempo de espera pelo processo de migração. Assim o Data Center pode usar menos recursos e ainda garantindo a qualidade de serviço, gerando economia de energia e diminuição de custos relacionados a infraestrutura. Com isso este modelo foi classificado como:

- ii) Resolve problemas da infraestrutura: pois a redução de tempo pelo processo de imigração é um grande benefício para gerenciamento da infraestrutura;
- iii) Resolve problemas de negócio: pois o modelo proposta resulta também uma diminuição de custos de cerca de 20% relacionados à infraestrutura;
- iv) Resolve problema ecológico: pois o modelo proposta resulta também uma diminuição de consumo de energia.

Já **Tai et. al (2011)**, no artigo “*ArA: Adaptive resource allocation for cloud computing environments under bursty workloads*” (ArA: Alocação de Recursos Adaptável para Ambiente de Computação em Nuvem com Balanceamento de Carga) propõe um Algoritmo para fazer balanceamento de carga inteligente, o resultados do experimento mostram que este modelo pode se adaptar rapidamente as mudanças nas demandas dos usuários e assim o desempenho da infraestrutura de nuvem. Assim este modelo foi classificado como:

- ii) Resolve problemas da infraestrutura: pois ele propõe ele propõe uma ótima solução para problemas de gerenciamento da infraestrutura de nuvem utilizando balanceamento de carga.

Da mesma forma foram analisado todo o resto dos modelos da nossa amostra piloto e classificado segundo este critério.

c) Resolve problema de satisfação ou de otimização

Os modelos de gerenciamento de recursos em Nuvem estudados, proposta nos trabalhos resolvem problemas de satisfação (SLA, QoS, etc.) ou de otimização (uso de recurso, lucro, consumo de energia, etc.).

Por isso, propomos um critério de classificação que vai definir se o modelo resolve problema de satisfação ou de otimização e indicar quais critérios ele visa satisfazer ou otimizar.

Tabela 6 Critério de satisfação ou otimização

Cat.	Modelos	Satisfação	Otimização
1	(BELOGLAZOV; ABAWAJY; BUYYA, 2011), (WU; KUMAR; BUYYA, 2011), (WU et al., 2012), (GARG; BUYYA; CALHERIROS, 2011),	X	X

	(DI; KONDO; CIRNE, 2012), (KUMAR et al., 2011), (FENG et al., 2012), (GE et al., 2012), (GUAN; FU; ZHANG, 2012) (DHINGRA; PAUL, 2014)		
2	(ZHANG; ZHU; BOUTABA, 2011), (RAMEZANI; LU; HUSSAIN, 2013), (WEI; VASILAKOS; ZHENG, 2010), (GOUDARZI; PEDRAM, 2011), (KURIBAYASHI, 2011), (BILGAIYAN; SANGNIKA; DAS, 2014), (WUHIB; STADLER; LINDGREN, 2012), (WEI; BLAKE; SALEH, 2013), (FITÓ et al., 2012), (EESPADA et al., 2013), (LIN et al., 2011), (RAO et al., 2011), (NGUYEN et al., 2013), (ENDO; SADOK; KELNER, 2011), (ERGU et al., 2011), (TAI et al., 2011)		X
3	(RAVINDRAN, 2013), (YOU et al., 2011), (AN et al., 2010)	X	

Analisando os modelos da nossa amostra piloto, constatamos que cerca de sessenta por cento resolvem problemas de otimização (otimização de uso de recursos, minimização de custo e consumo de energia, maximização de lucro), cerca de trinta por cento resolvem problemas de otimização e satisfação junto (cumprimento de SLA, melhora do QoS, cobrança de preço justo pelo serviço, otimização de uso de recursos, minimização de custo e consumo de energia, maximização de lucro), somente dez por cento dos modelos resolvem problemas de satisfação (cumprimento de SLA, melhora do QoS, cobrança de preço justo pelo serviço).

Este desequilíbrio constatado entre a satisfação e a otimização mostra um cenário que mais comum no mundo de negócio, onde as empresas visam maximizar seus lucros e deixam a satisfação do cliente em segundo plano.

Feng et. al (2012) no artigo “*Revenue Maximization Using Adaptive Resource Provisioning in Cloud Computing Environments*” (Maximização de Lucro Usando Provisionamento Adaptativa de Recursos em Ambiente de

Nuvem) propõe um modelo baseado no SLA para maximizar os lucros da empresa provedora de serviço fazendo uma alocação otimizado de recurso, considerando o QoS, o mecanismo de fixação de preço, a taxa de chegada e disponibilidade de serviço, e este modelo foi classificado como resolvendo problema de otimização e satisfação ao mesmo, pois o modelo proposto visa satisfazer o cliente com QoS e cumprindo o SLA, além de otimizar o uso de recursos computacionais e maximizar os lucros da empresa provedora de serviços de nuvem, todos os modelos com as mesmas característica foram classificados também como resolvendo problemas de otimização e satisfação.

Karibayashi (2011) no artigo “*Optimal Joint Multiple Resource Allocation Method for Cloud Computing Environments*” (Método Ideal para Alocação de Múltiplos recursos em Ambiente de Computação em Nuvem) propõe um modelo que otimiza o uso de recursos computacionais fazendo uma distribuição justa de recursos entre os usuários, além de minimizar a probabilidade de perda de pacotes, este modelo foi classificado como resolvendo problemas de otimização, assim como todos os modelos com as mesma característica,

An et. al (2010) no artigo “*Automated Negotiation with Decommitment for Dynamic Resource Allocation in Cloud Computing*” (Negociação Automatizada com Anulação de Alocação de Recursos em Computação em Nuvem) propõe um modelo que usa um mecanismo onde agentes distribuídos negociam o preço de serviço e as penas a ser aplicadas em caso de violação de SLA na forma de um leilão, este modelo foi classificado como resolvendo problema de satisfação, assim como todos os modelos com as mesmas características.

d) Tomada de decisão

Em um ambiente de computação distribuída como a Nuvem, os componentes se comunicam ou interage uns com os outros para efeito de tomada de decisão individual ou de todo o sistema, por exemplo, para alocar, migrar e reconfigurar VMs, fazer agendamento de tarefas, balanceamento de carga, etc.

Para efeito de nosso trabalho, adotaremos a abordagem apresentado por **Robert Thrall (1984)** no se artigo sobre Taxonomia de Modelo de Decisão, para ele, um modelo de tomada de decisão deve levar em conta três características na sua tomada de decisão: a primeira característica se refere ao numero de *decisores*, a segunda característica é referente ao numero de *critério* da decisão e à terceira característica é o numero de *decisões* a ser tomado.

Tabela 7 Tomada de decisão

Nº	Modelos	Numero de decisores		Numero de critérios		Numero de decisões	
		Um	Muitos	Um	Muitos	Um	Muito
1	(BELOGLAZOV; ABAWAJY; BUYYA, 2011)	X			X	X	
2	(WU; KUMAR; BUYYA, 2011)	X			X		X
3	(WU et al., 2012)	X			X	X	
4	(ZHANG; ZHU; BOUTABA, 2011)	X			X		X
5	(RAMEZANI; LU; HUSSAIN, 2013)		X		X		X
6	(WEI; VASILAKOS; ZHENG, 2010)		X		X		X
7	(GOUDARZI; PEDRAM, 2011)	X			X		X
8	(RAVINDRAN, 2013)	X			X	X	
9	(KURIBAYASHI, 2011)		X		X	X	
10	(GARG; BUYYA; CALHERIROS, 2011)	X			X	X	
11	(YOU et al., 2011)	X			X		X
12	(BILGAIYAN; SANGNIKA; DAS, 2014)	X			X	X	
13	(DI; KONDO; CIRNE, 2012)	X			X	X	
14	(WUHB; STADLER; LINDGREN, 2012)		X		X		X
15	(WEI; BLAKE; SALEH, 2013)		X		X		X
16	(FITÓ et al., 2012)		X		X		X
17	(AN et al., 2010)		X		X	X	
18	(ESPADA et al., 2013)		X		X		X
19	(KUMAR et al., 2011)	X			X	X	
20	(LIN et al., 2011)		X		X	X	

21	(FENG et al., 2012)		X		X	X	
22	(RAO et al., 2011)		X	X	X		
23	(GUAN; FU; ZHANG, 2012)		X		X	X	
24	(GE et al., 2012)		X		X	X	
25	(DHINGRA; PAUL, 2014)		X		X		X
26	(NGUYEN et al., 2013)	X			X	X	
27	(ENDO; SADOK; KELNER, 2011)	X			X		X
28	(ERGU et al., 2011)		X		X		X
29	(TAI et al., 2011)		X		X	X	
30	(CHUNLIN; LAYUAN, 2013)		X		X	X	

O processo de decisão é vista como uma busca sequencial, onde *decisões* (balanceamento de carga, ligar ou desligar VMs, migrar VMs, diminuir ou aumentar o preço de serviço, etc.) são tomadas levando em conta *critérios* (SLA, QoS, consumo de energia, custos, tolerância a falha, lucros, recursos físicos e virtuais disponíveis, confiabilidade, etc.), todos nos limites do conhecimento dos *decisores* (agentes inteligentes, mecanismo utilizando um determinado algoritmo, ajuste ou adaptação nos *decisores* já existente nas ferramentas de gerenciamento do CN, balanceador de carga, alocador de tarefa, etc.). *Decisores* podem ser considerados como minimizadores das *decisões* buscam apenas resultados satisfatórios operando em ambientes de racionalidade limitada com a Nuvem por exemplo. O desafio dos *decisores* é o de identificar qual das alternativas no conjunto de solução é a melhor.

Chunlin et. al (2013) no seu artigo “*Multi-Layer Resource Management in Cloud Computing*” (Gerenciamento de Recursos Multicamada em Computação em Nuvem), propõe um modelo que utiliza três otimizadores diferentes, um para camada IaaS, um para camada SaaS e outro para camada de usuários. Este modelo tem como objetivo maximizar os lucros e aperfeiçoar o uso de recursos, utilizando a alocação e deslocação de tarefas. Assim este modelo foi classificado como, segundo o numero de *decisores* como *muito* (três otimizadores, uma para cada camada), segundo numero de *critérios* como *muito* (maximizar lucro e aperfeiçoar uso de recursos) e segundo as decisões a ser tomada como *um* (alocação e deslocação de tarefas).

Beloglazov et. al (2011) no seu artigo “*Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing*” (Heurística para Gerenciamento eficiente do Data Center para Computação em Nuvem) propõe um modelo para gerenciamento de recursos em data center orientado a computação verde utilizando um algoritmo baseado na

teoria de fila que visa minimizar consumo de energia, minimizar migrações das VMs e minimizar a violação do SLA fazendo migração de VMs. Este modelo foi classificado como, segundo o numero de *decisores* como *um* (um Maximizador de migração), segundo numero de *critérios* como *muito* (minimizar consumo de energia, minimizar migrações das VMs e minimizar a violação do SLA) e segundo as decisões a ser tomada como *um* (migração de VMs).

Da mesma forma fora classificado todo o resto dos modelos segundo o critério, tomada de decisão.

e) Previsão de mudança no tempo

O tempo é um parâmetro importante na tomada de decisão, um modelo de tomada de decisão pode ou não prever mudança com relação ao tempo, e este tempo pode ser continuo ou discreto. Há outros modelos que não consideram o tempo.

Por isso propomos um critério de classificação de modelos de tomada de decisão baseado na previsão de mudança de tempo.

Tabela 8 Mudança no tempo

Numero	Modelos	Continuo	Discreto	Não considera o tempo
1	(WU; KUMAR; BUYYA, 2011), (WU et al., 2012), (GOUDARZI; PEDRAM, 2011), (KURIBAYASHI, 2011), (BILGAIYAN; SANGNIKA; DAS, 2014), (DI; KONDO; CIRNE, 2012), (WEI; BLAKE; SALEH, 2013), (FITÓ et al., 2012), (NA et al., 2010), (ESPADA et al., 2013), (KUMAR et al., 2011), (LIN et al., 2011), (FENG et al., 2012), (RAO et al., 2011), (NGUYEN et al., 2013), (TAI et al., 2011), (CHUNLIN; LAYUAN,	X		

	2013), (ENDO; SADOK; KELNER, 2011), (BELOGLAZOV; ABAWAJY; BUYYA, 2011),			
2	(ZHANG; ZHU; BOUTABA, 2011), (RAMEZANI; LU; HUSSAIN, 2013), (GUAN; FU; ZHANG, 2012), (GE et al., 2012), (DHINGRA; PAUL, 2014), (ERGU et al, 2011)		X	
3	(WEI; VASILAKOS; ZHENG, 2010), (RAVINDRAN, 2013), (GARG; BUYYA; CALHERIROS, 2011), (YOU et al., 2011), (WUHIB; STADLER; LINDGREN, 2012),			X

Tai et. al (2011) no artigo “ArA: Adaptive resource allocation for cloud computing environments under bursty workloads” (ArA: Alocação de Recursos Adaptativa para Ambientes de computação em Nuvem sob Carga de Trabalho) propõe um modelo de gerenciamento de recursos em Nuvem que nas especificações de uma tarefa é incluído os intervalos de tempo de chegada e o tempo de execução das tarefas, assim este modelo foi classificado segundo o critério previsão de mudança no tempo como *continua*.

Ge et. al (2012), no artigo “A game theoretic resource allocation for overall energy minimization in mobile cloud computing system” (Alocação de Recursos utilizando Teoria de Jogo para minimização de consumo de energia total no Sistema de Computação em Nuvem Move) propõe um modelo de gerenciamento de recursos que usa um algoritmo que atinge seu nível mais alto de eficiência em tempo polinomial, assim este modelo foi classificado segundo o critério previsão de mudança no tempo como *discreto*.

Os modelos que não mencionaram a mudança no tempo foram classificados com *não consideram o temp.*

Da mesma forma classificado os outros modelos segundo o critério previsão de mudança no tempo.

f) Formalismo em que os modelos se baseiam

Um modelo de tomada de decisão deve ter um formalismo em que se baseia. Existe enumeras formalismos em se baseia esses modelos, mas para fim de nosso trabalho, agruparmos eles em set categorias seguintes:

- i) Baseada em técnicas de programação linear (PL);
- ii) Baseada em técnicas da teoria de jogo (TJ);
- iii) Baseada em técnicas da lógica nebulosa (LN);
- iv) Baseado em técnicas de computação evolucionaria (CE);
- v) Baseado em técnicas de inteligência artificial (IA);
- vi) Baseado em Teoria de Fila (TF)

Assim propomos um critério de classificação de modelos de tomada de decisão em *Nuvem* baseado em formalismo em que se baseia o modelo.

Tabela 9 Formalismo em que os modelos se baseiam

Nº	Modelos	IA	PL	TJ	LN	CE	TF	Outros
1	(WU et al., 2012), (RAO et al., 2011), (ENDO; SADOK; KELNER, 2011)	X						
2	(ZHANG; ZHU; BOUTABA, 2011), (GOUDARZI ; PEDRAM, 2011), (WEI; BLAKE; SALEH, 2013), (FITÓ et al., 2012), (AN et al., 2010), (ESPADA et al., 2013), (KUMAR et al., 2011), (NGUYEN et al., 2013), (CHUNLIN ; LAYUAN, 2013), (ZHANG; ZHU; BOUTABA, 2011), (WU; KUMAR; BUYYA, 2011), (KURIBAYASHI, 2011)		X					
3	(WEI; VASILAKOS; ZHENG, 2010), (GE et al., 2012)			X				

4	(DI; KONDO; CIRNE, 2012), (GUAN; FU; ZHANG, 2012), (RAMEZANI; LU; HUSSAIN, 2013)				X			
5	(YOU et al., 2011), (BILGAIYAN; SANGNIKA; DAS, 2014)					X		
6	(BELOGLAZOV; ABAWAJY; BUYYA, 2011), (FENG et al, 2012)						X	
7	(GARG; BUYYA; CALHERIROS, 2011), (WUHIB; STADLER; LINDGREN, 2012), (LIN et al., 2011), (DHINGRA; PAUL, 2014), (ERGU et al., 2011), (TAI et al., 2011), (RAVINDRAN, 2013)							X

Nos modelos da nossa amostra examinada, constatamos que cerca de sessenta e set por centos dos modelos foram classificados em categorias propostas, segundo o critério de classificação *formalismo em que os modelos se baseiam*. Como por exemplo, o modelo proposta por **Endo et. al (2011)** no artigo “*Autonomic Cloud Computing: Giving Intelligence to Simpleton Nodes*” (Computação em Nuvem Autônomo: Oferecendo Inteligência para Nodos Simpleton), propões um modelo de gerenciamento de recursos em Nuvem utilizando Computação Autônoma que usa técnicas de IA para otimizar a tomada de decisão sobre utilização de recursos. Neste modelo são usadas as “*Automatic Cloud Agents*” (agentes Clouds autônomos) para enviar de forma assíncrona informações de status do ambiente para a entidade de gerenciamento. Estes agentes são responsáveis pelo monitoramento de recursos, compartilhamento de informações e tomar decisões sobre o uso de recursos. Este modelo foi classificado como baseado no formalismo de IA, pois ele usa técnicas de IA para aperfeiçoar a tomada de decisão sobre uso de recursos computacionais.

Goudarzi e Pedram (2011) no artigo “*Multi-dimensional SLA-based Resource Allocation for Multi-tier Cloud Computing Systems*” (Alocação de

Recursos baseado em SLA multidimensional para Sistemas de Computação Nuvem) apresenta um modelo de gerenciamento de recursos em SLA que gera uma otimização de lucros gerados devido à satisfação dos contratos de SLA dos clientes e a diminuição de consumo de energia. Neste modelo, a taxa de chegada das requisições do cliente é determinada por uma função linear de distribuição probabilística. E para resolver problema de maximização de lucro, o autor formula o problema como problema de PL, assim este modelo foi classificado como baseado no formalismo PL.

Ge et. al (2012) no artigo “*A game theoretic resource allocation for overall energy minimization in mobile cloud computing system*” (Teoria de Jogo na Alocação de Recursos para Minimização de consumo de energia no Sistema de Computação em Nuvem Móvel) propõe um abordagem da TG para otimizar consumo de energia em sistema de CN. Neste modelo o problema de minimização de consumo de energia é formulador como um jogo, onde cada dispositivo móvel é um jogador e sua estratégia de jogo é selecionar um dos servidores e equilibrar a sua carga de trabalho, assim minimizando o consumo global de energia. Os resultados experimentais mostram que esta abordagem é capaz de reduzir até 28,9 % de consumo de energia em alguns casos. Este modelo foi classificado como baseado no formalismo TJ.

Di et. al (2012), no artigo “*Host Load Prediction in a Google Compute Cloud with a Bayesian Model*” (Previsão de Carga de hospedagem na Google Compute Cloud com Modelo Bayesian) propõe um modelo de previsão de carga do host baseado no modelo de Bayes. Para fazer a previsão de carga são identificados a previsibilidade, a tendência e padrões de carga de host. Os experimentos mostraram que este modelo baseado no modelo de Bayes melhora a precisão da previsão de carga por 5,6 até 50 % em relação a outros modelos com base em medias moveis, auto regressão, etc. Este modelo foi classificado como baseado no formalismo LN, pois o método de Bayes faz parte de LN.

You et. al (2011), no artigo “*ARAS-M: Automatic Resource Allocation Strategy based on Market Mechanism in Cloud Computing*” (ARAS-M: Estratégia para Alocação Automática de Recursos com base no Mecanismo de Mercado em Computação em Nuvem) propõe um modelo para alocação automática de recursos usando algoritmo genética baseado em algoritmo de ajuste de preço de serviço, para lidar com problema da demanda e da procura. Este modelo foi classificado como baseado no formalismo CE, assim como todos os modelos que seguem as mesmas características.

Feng et. al (2012) no artigo “*Revenue Maximization Using Adaptive Resource Provisioning in Cloud Computing Environments*” (Maximização de Lucros Utilizando Provisionamento de Recursos em ambiente de Computação em Nuvem) propõe um modelo de gerenciamento de recursos baseado em teoria de fila para maximizar lucros da empresa provedora, melhorar o QoS, simulando vários políticas de escalonamento de fila e comparando os resultados. Este modelo foi classificado como baseado em TF, assim como todos os outros modelos que seguiram as mesmas características que este.

Alguns modelos da nossa amostra piloto não se encaixaram nos formalismos estabelecidos acima, eles foram classificados como baseado em outros formalismos. Como por exemplo, o modelo proposta por **Garg et. al (2011)** orientado a SLA, utilizando algoritmo próprio baseado em alocação de tarefas, o modelo proposta por **Ravindran (2013)**, que faz gerenciamento autônomo de recursos baseado em QoS e SLA Auditing oferecendo meios para avaliar serviço prestado pela empresa provedora, reforçando a importância do respeito do SLA, o modelo proposta por **Tai et. al (2011)**, utilizando um algoritmo adaptado para alocação de recurso baseado em balanceamento de carga, etc.

g) Saídas dos modelos de tomada de decisão

Todo modelo de tomada de decisão tem que apresentar um resultado final, que pode um conjunto de configurações a ser feito no sistema, um conjunto de ações a ser realizado no sistema, conjunto de planos a ser aplicado no sistema, isso na forma de uma ideia, de uma implementação, de experimento que pode apresentar somente a análise dos resultados durante o experimento, ou além de apresentar resultados pode comparar com outros trabalhos.

Assim criamos uma classificação dos modelos, baseado em resultados finais apresentado nos trabalhos (configurações, ações, planos ou outro).

Tabela 10 Saída dos modelos

N°	Modelos	Configurações	Ações	Planos	Outros
1	(GOUDARZI; PEDRAM, 2011), (YOU et al., 2011), (GE et al., 2012), (ENDO; SADOK; KELNER, 2011)	X			
2	(WU; KUMAR; BUYYA, 2011), (NA et al., 2010), (LIN et al., 2011),	X		X	
3	(WU et al., 2012), (BILGAIYAN; SANGNIKA; DAS, 2014), (DI; KONDO; CIRNE, 2012), (ESPADA et al., 2013),	X	X	X	

	(KUMAR et al., 2011), (RAO et al., 2011), (NGUYEN et al., 2013), (CHUNLIN; LAYUAN, 2013). (BELOGLAZOV; ABAWAJY; BUYYA, 2011),				
4	(RAMEZANI; LU; HUSSAIN, 2013), (ZHANG; ZHU; BOUTABA, 2011), (RAVINDRAN, 2013), (WUHB; STADLER; LINDGREN, 2012), (GUAN; FU; ZHANG, 2012), (DHINGRA ; PAUL, 2014), (ENDO; SADOK; KELNER, 2011), (TAI et al., 2011), (WEI; VASILAKOS; ZHENG, 2010), (ERGU et al., 2011)	X	X		
5	(FITÓ et al., 2012)	X		X	
6	(KURIBAYASHI, 2011), (GARG; BUYYA; CALHERIROS, 2011), (WEI; BLAKE; SALEH, 2013),		X		
7	(FENG et al., 2012)			X	

Segundo o nosso critério estabelecido o produto final de modelos avaliados foi agrupado em quatro classes. Então o produto final pode ser um conjunto de configurações (atualização de softwares, configurações de segurança para corrigir falhas no ambiente, atualizações de VMs, etc.) a ser feita no ambiente, um conjunto de ações (migrar de VMs, ligar VMS, Desligar VMs,

alocar VMs, balancear carga de trabalho, etc.) a ser executado no ambiente, um plano de utilização de recursos, ou outros.

Beloglazov et. al no artigo “*Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing*” (Energia Consciente heurísticas de alocação de recursos para gerenciamento de data center para computação em nuvem) propõe um modelo de gerenciamento autônomo de recursos que minimiza o uso de recurso e a violação do SLA. Este modelo é composto de vários componentes, um deles chamado *Comsumer Profiler* reúne características específicas dos consumidores para que os consumidores mais importantes possam ser concedidos privilégios especiais e prioridade sobre os de mais, fazendo configurações no ambiente (conjunto de configurações). No outro componente chamado *VM Manager* que matem o controle da disponibilidade de VMs e seu uso de recursos. Este componente é carregado de provisionamento de novos VMs, bem como realocação de VMs (um conjunto de ações) e o componente *Accounting* monitora o uso real de recursos e os dados históricos de uso de recursos são usados para melhorar as decisões planejamento de alocação futura de recursos. A saída deste modelo foi classificada como sendo um conjunto *Configurações*, um conjunto de *Ações* e um *Plano* de alocações de recursos. Da mesma forma foram classificados todos os modelos que seguem o mesmo padrão.

Wubib et. al (2012) no artigo “*Dynamic Resource Allocation with Management Objectives-Implementation for an OpenStack Cloud*” (Alocação Dinâmica de Recursos para Gerenciamento Orientada a Implementação para OpenStack Cloud) propõe um modelo de gerenciamento dinâmica de recurso que tem como objetivo equilibrar a carga de trabalho e a eficiência energética, fazendo um conjunto de configurações no ambiente OpenStack e o balanceamento de carga. A saída deste modelo foi classificada como sendo um conjunto *Configurações*, um conjunto de *Ações*. Da mesma forma foram classificados todos os modelos que seguem o mesmo padrão.

h) O aspecto da Nuvem abordada no modelo

CN é uma área bem complexa, quando alguém faz um estudo sobre gerenciamento de recursos em Nuvem, por exemplo, ele tem que especificar quais os aspectos de Nuvem que vai abordar. Aspecto técnico se o estudo refere-se a configurações e ações, aspecto segurança, foram excluídos da amostra todos os artigos que trata de segurança, aqui consideraremos como aspecto de segurança se modelo traz o estudo é sobre recuperação e prevenção de falha, aspecto de negocio quando o estudo visa maximizar lucros e minimizar custos e aspecto político quando o objetivo é satisfazer as partes interessadas.

Assim propomos um critério de classificação baseado no aspecto de Nuvem abordado no trabalho.

Tabela 11 Aspecto da Nuvem abordada

Numero	Modelos	Técnico	Segurança	Negocio	Política
1	(BELOGLAZOV; ABAWAJY; BUYYA, 2011), (WU et al., 2012)	X			X
2	(WU; KUMAR; BUYYA, 2011), (GOUDARZI; PEDRAM, 2011), (WEI; VASILAKOS; ZHENG, 2010), (GARG; BUYYA; CALHERIROS, 2011), (FITÓ et al., 2012), (KUMAR et al., 2011), (GE et al., 2012), (DHINGRA; PAUL, 2014), (FENG et al., 2012)	X		X	X
3	(ZHANG; ZHU; BOUTABA, 2011), (RAMEZANI; LU; HUSSAIN, 2013), (DI; KONDO; CIRNE, 2012), (AN et al., 2010), (YOU et al., 2011)	X		X	
4	(RAVINDRAN, 2013), (KURIBAYASHI, 2011), (GUAN; FU; ZHANG, 2012)	X	X		
5	(WUHIB; STADLER; LINDGREN, 2012), (WEI; BLAKE; SALEH, 2013), (ESPADA et al., 2013), (LIN et al., 2011), (RAO et al., 2011), (BILGAIYAN; SANGNIKA; DAS, 2014), (ENDO; SADOK; KELNER, 2011),	X			

	(ERGU et al., 2011), (TAI et al., 2011), (CHUNLIN; LAYUAN, 2013)				
6	(NGUYEN et al., 2013)	X	X	X	X

Analisando os modelos da nossa amostra, constatamos que todos eles abordam o aspecto técnico (conjunto de configurações e ações) associado outros aspectos estabelecidos acima (segurança, negocio e politica). Por exemplo, **Nguyen et. al (2013)**, propõe um modelo de gerenciamento de recursos que contem vários estratégia para diminuir a solicitação de processo de migração de VMs como objetivo de manter o data center com alta disponibilidade de recursos, o modelo também propõe um método para reduzir o tempo de espera pelo processo de migração, com isso o data center pode usar menos recursos e ainda garantir o QoS e preservar o SLA. Segundo os resultados apresentado, este modelo este modelo ajuda a gerenciar de forma eficiente e segura o ambiente de CN, economizando o consumo de energia, minimizando custos e maximizando lucros. Este modelo foi classificado como abordando os aspectos *técnico, segurança, negocio e politica*. Da mesma forma foram analisados e classificados os outros modelos.

4.4 Classificações dos modelos propostas nos trabalhos

A classificação dos trabalhos da nossa amostra será feita segundo a taxonomia proposta acima. Assim teremos oito critérios para classificar os modelos de tomada de decisão para gerenciamento da infraestrutura de Nuvem.

Tabela 12 Classificações dos modelos

Nº	Modelos Classes	A	b	c	d	e	f	g	h
1	(BELOGLAZOV; ABAWAJY; BUYYA, 2011)	iii	i, ii, iv	Satisfação e Otimização	Um, Muito, Um	Continua	TF	Configuração, Ações, Planos	Técnico, Político
2	(WU; KUMAR; BUYYA, 2011)	li	i, ii, iii	Satisfação e Otimização	Um, Muito, Muito	Continua	PL	Configuração, Planos	Técnico, Negocio, Político
3	(WU et al., 2012)	iv	i, ii	Satisfação e Otimização	Um, Muito, Um	Continua	IA	Configuração, Ações, Planos	Técnico, Político
4	(ZHANG; ZHU; BOUTABA, 2011)	iv	ii, iii, iv	Otimização	Um, Muito, Muito	Discreto	PL	Configuração, Ações	Técnico, Negocio
5	(RAMEZANI; LU; HUSSAIN, 2013)	iv	i, ii, iii, iv	Otimização	Muito, Muito, Muito	Discreto	LN	Configuração, Ações	Técnico, Negocio
6	(WEI; VASILAKOS; ZHENG, 2010)	li	iii	Otimização	Muito, Muito, Muito	Não considera tempo	TJ	Configurações, Ações	Técnico, Negocio, Política
7	(GOUDARZI; PEDRAM, 2011)	li	i, iii	Otimização	Um, Muito, Muito	Continua	PL	Configurações	Técnico, Negocio, Política
8	(RAVINDRAN, 2013)	I	i, ii, iii, iv	Satisfação	Um, Muito, Um	Não considera tempo	Outros	Configurações, Ações	Técnico, Segurança

9	(KURIBAYASHI, 2011)	I	ii	Otimização	Muito, Muito, Um	Continua	PL	Ações	Técnico, Segurança
10	(GARG; BUYYA; CALHERIROS, 2011)	iv	i, ii, iii	Otimização e Satisfação	Um, Muito, Um	Não considera tempo	Outros	Ações	Técnico, Negocio, Política
11	(YOU et al, 2011)	I	iii	Satisfação	Um, Muito, Muito	Não considera tempo	CE	Configurações	Técnico, Negocio
12	(BILGAIYAN; SANGNIKA; DAS, 2014)	iv	ii	Otimização	Um, Muito, Um	Continua	CE	Ações, Configurações, Plano	Técnico
13	(DI; KONDO; CIRNE, 2012)	iv	i, ii, iii	Otimização e Satisfação	Um, Muito, Um	Continua	LN	Ações, Configurações, Plano	Técnico, Negocio
14	(WUHIB; STADLER; LINDGREN, 2012)	I	i, ii, iv	Otimização	Muito, Muito, Muito	Não considera tempo	Outros	Ações, Configurações	Técnico
15	(WEI; BLAKE; SALEH, 2013)	i	i, ii	Otimização	Muito, Muito, Muito	Continua	PL	Ações	Técnico
16	(FITÓ et al, 2012)	i	i, ii, iii, iv	Otimização	Muito, Muito, Muito	Continua	PL	Configurações, Plano	Técnico, Negocio, Política
17	(AN et al., 2010)	iv	ii, iii	Satisfação	Muito, Muito, Um	Continua	PL	Configurações, Plano	Técnico, Negocio

18	(ESPADA et al., 2013)	iii	ii	Otimização	Muito, Muito, Muito	Continua	PL	Ações, Configurações, Plano	Técnico
19	(KUMAR et al., 2011)	iv	i, ii, iii	Satisfação e Otimização	Um, Muito, Um	Continua	PL	Ações, Configurações, Plano	Técnico, Negocio, Política
20	(LIN et al., 2011)	iii	ii	Otimização	Muito, Muito, Um	Continua	Outros	Configurações, Plano	Técnica
21	(FENG et al., 2012)	ii	i, ii, iii	Satisfação e Otimização	Muito, Muito, Um	Continua	TF	Configurações, Plano	Técnico, Negocio, Política
22	(RAO et al., 2011)	iv	ii	Otimização	Muito, Muito, Um	Continua	IA	Ações, Configurações, Plano	Técnico
23	(GUAN; FU; ZHANG, 2012)	iv	ii	Satisfação e Otimização	Muito, Muito, Um	Discreto	LN	Ações, Configurações	Técnico, Segurança
24	GE et al., 2012)	iv	i, iv	Satisfação e Otimização	Muito, Muito, Um	Discreto	TJ	Configurações	Técnico, Negocio, Política
25	(DHINGRA; PAUL, 2014)	iv	i, ii, iv	Satisfação e Otimização	Muito, Muito, Muito	Discreto	Outros	Ações, Configurações	Técnico, Negocio, Política

26	(NGUYEN et al., 2013)	iii	ii, iii, iv	Otimização	Um, Muito, Um	Continua	PL	Ações, Configurações, Plano	Técnico, Negocio, Politica, Segurança
27	(ENDO; SADOK; KELNER, 2011)	iv	ii	Otimização	Um, Muito, Muito	Continua	IA	Configurações	Técnico
28	(ERGU et al, 2011)	ii	ii	Otimização	Muito, Muito, Muito	Discreto	Outros	Ações, Configurações	Técnico
29	(TAI et al., 2011)	iv	ii	Otimização	Muito, Muito, Um	Continua	Outro s	Ações, Configurações	Técnico
30	(CHUNLIN; LAYUAN, 2013)	iii	i, ii	Otimização	Muito, Muito, Um	Continua	PL	Ações, Configurações, Plano	Técnico

Analisando os modelos de tomada de decisão para infraestrutura de Nuvem da nossa amostra piloto, conseguimos perceber uma correlação entre os critérios de classificação estabelecidos acima.

A maioria dos modelos que no critério de *previsão de mudança no tempo*, que foram classificados como *continua*, no critério de classificação *saída dos modelos de tomada de decisão*, além das *configurações e ações* a ser tomadas no ambiente, eles propõem um *plano* de utilização de recursos.

A maioria dos modelos baseados na definição de **Rajkumar Buyya** tem tendência de abordar os aspectos de *Negocio* e *Politica* da CN e a maioria dos modelos baseados na definição do **NIST** tem tendência a abordar o aspecto técnico do CN.

Todos os modelos baseados em formalismo de Programação Linear têm tendência a propor otimização de uso de recurso, maximização de lucros e minimização de consumo de energia e custos.

4. CONCLUSÃO E TRABALHOS FUTUROS

O *CN* é uma tendência que cresceu muito nos últimos anos, conforme vimos, esse paradigma oferece diversos benefícios proporcionando as empresas economia de recursos, com a possibilidade de expansão elástica da capacidade, e foco em seus negócios e aos usuários comuns comodidade de acessar de qualquer lugar e menos gastos na utilização de softwares do dia-a-dia, não mais pagando sobre a licença completa e sim um aluguel pela sua utilização, de forma semelhante ao que hoje ocorre com eletricidade. As atividades como instalação de servidores, manutenção da operação, armazenamento de dados, e a atualização dos aplicativos são realizadas fora do alcance das empresas usuárias dos serviços. E para essas empresas usuárias, bastaria à contratação do serviço desejado, que seria prestado via Internet.

Neste trabalho, propomos uma taxonomia para os modelos de tomada de decisão para gerenciamento da infraestrutura de Nuvem, como isso isto visamos ajudar os administradores de infraestruturas de Nuvens na escolha ou na implementação de um modelo para gerenciar seus ambientes de Nuvens, segundo os objetivos que cada um visa alcançar.

Depois de fazer todos os filtros estabelecidos, a nossa amostra ficou de duzentos noventa e set modelos, nos quais decidimos classificar somente trinta modelos, por causa dos prazos estabelecidos para realização do nosso trabalho. Assim como trabalhos futuras propomos entender esta taxonomia para todos os modelos da amostra, implementar alguns modelos da amostra escolhido segundo critérios que serão estabelecidos para comprar os resultados. Isto além de avaliar qualitativamente os modelos, serviria como auxiliaria os administradores de ambientes de Nuvens na escolha de modelo para implantar ou adaptar para gerenciar o ambiente de Nuvem.

REFERÊNCIAS

- Albuquerque, M. G., Rabelo, D. d., & Soares, J. M. (2012). Uma Abordagem para Simulação e Avaliação na Redução do Consumo de Energia em Ambientes Computacionais Virtualizados baseada em Redes de Petri Colorida . *Departamento de Teleinformática – Universidade Federal do Ceará (UFC), Campus do PICI, S/Nº, Bloco 725 – Fortaleza – CE – Brasil.*
- Almeida, F. P. (2006). A teoria dos jogos: uma fundamentação teórica dos métodos de resolução de disputa. *Universidade de Brasília.*
- An, B., Lesser, V., Irwin, D., & Zink, M. (2010). Automated Negotiation with Decommittment for Dynamic Resource Allocation in Cloud Computing. *ACM.*
- Armbrust, M., Fox, A., Griffith, R., Joseph, D. A., Katz, H. R., Konwinski, A., et al. (2009). Above the Clouds: A Berkeley View of Cloud.
- Beloglazov, A., Abawajy, J., & Buyya, R. (2011). Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing. *Cloud Computing and Distributed Systems (CLOUDS) Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Australia.*
- Bilgaiyan, S., Sangnika, S., & Das, M. (2014). An Analysis of Task Scheduling in Cloud Computing using Evolutionary and Swarm-based Algorithms. *Internationall Journal of Computer Applications .*
- Bisdikian, C., Mitschang, B., Pedreschi, D., & Tseng, V. S. (2011). Challenges for Mobile Data Management in the Era of Cloud and Social Computing. *IEEE.*
- Blomberg, J. (2008). Negotiating meaning of shared information in service system encounters. *European Management Journal, 213-222.*
- Buyya, R., Rahman, M., & Ranjan, R. (2011). A taxonomy and survey on autonomic management of applications in grid computing environments. *IEEE.*
- Buyya, R., Yeo, S. C., Venugopal, S., Broberg, J., & Brandic, I. (2008). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems.*

- Chunlin, L., & Layuan, L. (2013). Multi-Layer Resource Management in Cloud Computing. *Springer Link*, 100-120.
- Clark, C., Fraser, K., & Hand, S. (2005). Live migration of virtual machines. *2nd conference on Symposium on Networked Systems Design and Implementation*.
- Corrar, L. J., & Garcia, E. A. (2001). CORRAR, Luis João Programação linear: uma aplicação à contabilidade de custos no processo de tomada de decisão. *Congresso Internacional de Custos*.
- Dhingra, A., & Paul, S. (2014). Green Cloud Computing: Towards Optimizing Data Centre Resource Allocation. *International Journal of Engineering Research & Technology*.
- Di, S., Kondo, D., & Cirne, W. (2012). Host Load Prediction in a Google Compute Cloud with a Bayesian Model. *IEEE*.
- Diaz-Sanchez, D., Almenarez, F., Marin, A., & Preserpio, D. (2011). Media cloud: an open cloud computing middleware for content management. *IEEE*.
- Dillon, T., Wu, C., & Chang, E. (2010). Cloud Computing: Issues and Challenges. *IEEE*, 27-33.
- Doukas, C., Pliakas, T., & Maglogiannis, I. (2010). Mobile healthcare information management utilizing Cloud Computing and Android OS. *IEEE*.
- Durkee, D. (2010). Why Cloud Computing Will Never Be Free. In: D. Durkee. *Communications of the ACM*.
- Endo, P. T., Sadok, D., & Kelner, J. (2011). Autonomic Cloud Computing: giving intelligence to simpleton nodes. *IEEE*, 502-505.
- Ergu, F., Kou, G., Peng, Y., Shi, Y., & Shi, Y. (2011). The analytic hierarchy process: task scheduling and resource allocation in cloud computing environment. *Springer Link*, 835-848.
- Espada, J., Molina, A., Jiménez, G., Molina, M., Ramírez, R., & Concha, D. (2013). A tenant-based resource allocation model for scaling Software-as-a-Service applications over cloud computing infrastructures. *ELSEVIER*, 273-286.

- Feng, G., Garg, S., Buyya, R., & Li, W. (2012). Revenue Maximization Using Adaptive Resource Provisioning in Cloud Computing Environments. *IEEE*, 192-200.
- Fitó, O., Ferram, J., Macias, M., & Guitart, J. (2012). Business-Driven IT Management for Cloud Computing Providers. *IEEE*.
- Francisco, C. X. (2012). Cognare: Um Sistema para alocação dinâmica de recursos baseado em técnicas de Inteligência Artificial. *Dissertação de Mestrado, Universidade Federal de Goiás*.
- Garg, S. K., Buyya, R., & Calheriros, R. N. (2011). SLA-oriented resource provisioning for cloud computing: Challenges, architecture, and solutions. *Cloud Computing and Distributed Systems (CLOUDS) Laboratory, The University of Melbourne, Australia*.
- Gartner, G. (2008). *Gartner Highlights Five Attributes of Cloud Computing*. Acesso em 22 de 01 de 2014, disponível em Gartner Group: <http://www.gartner.com/newsroom/id/1035013>
- Ge, Y., Zhang, Y., Qiu, Q., & Lu, Y.-H. (2012). A game theoretic resource allocation for overall energy minimization in mobile cloud computing system. *ACM*, 279-284.
- Goudarzi, H., & Pedram, M. (2011). Multi-dimensional SLA-based Resource Allocation for Multi-tier Cloud Computing Systems. *University of Southern California, Los Angeles, CA 90089*.
- Guan, Q., Fu, S., & Zhang, Z. (2012). Ensemble of Bayesian Predictors and Decision Trees for Proactive Failure Management in Cloud Computing Systems. *Journal of Communications*, 52-61.
- Hines, M., & Gopalan, K. (2009). Post-copy based live virtual machine migration using adaptive pre-paging and dynamic self-ballooning. *ACM SIGPLAN/SIGOPS*, 51-60.
- ICA, I. C. (2014). Inteligencia Computacional. *Laboratorio de Inteligencia Computacional Aplicada, PUC Rio de Janeiro*.
- Jesus, J. D. (31 de 06 de 2012). *Navegando na Nuvem IBM, Parte 1: Um Manual sobre Tecnologias em Nuvem*. Acesso em 24 de 03 de 2014, disponível em IBM Developer Works: http://www.ibm.com/developerworks/br/websphere/techjournal/1206_dejesus/1206_dejesus.html

- Joset, P. (2011). *Cloud Computing, tentative de définition*. Acesso em 2014 de 03 de 23, disponível em ABISSA informatique: http://www.abissa.ch/data/fichiers/tec_cloud_computing.pdf
- Klein, G. (1997). Sources of Power – How People Make Decisions. *MIT Press*.
- Kumar, K., Feng, J., Nimmagadda, Y., & Lu, Y. H. (2011). Resource Allocation for Real-Time Tasks using Cloud Computing. *IEEE*, 1-7.
- Kuribayashi, S.-i. (2011). Optimal Joint Multiple Resource Allocation Method for Cloud Computing Environments. *International Journal of Research and Reviews in Computer Science (IJRRCS)*.
- Lin, W., Wanh, J. Z., Liang, C., & Qi, D. (2011). A Threshold-based Dynamic Resource Allocation Scheme for Cloud Computing. *ScienceDirect*, 695-703.
- Lopes, G. (2010). *NBR-ISO-IEC-17999*. Acesso em 2014 de 05 de 21, disponível em Scribd: <http://pt.scribd.com/doc/37346638/NBR-ISO-IEC-17999>
- Marcon, P. M. (2006). O Processo de Tomada de Decisão do Enfermeiro no Cenário Administrativo. *Universidade Federal do Paraná, Programa de Pós-Graduação em Enfermagem*.
- Mell, P., & Grance, T. (2011). The NIST Definition of Cloud Computing. *Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology (NIST)*.
- Milojicic, D., Liorente, I. M., & Montero, R. S. (2011). OpenNebula: A Cloud Management Tool. *IEEE*.
- Nelson, M., Lim, B., & Hutchins, G. (2005). Fast transparent migration for virtual machines. *USENIX Annual Technical Conference*, 391-394.
- Nguyen, T. D., Nguyen, A. T., Nguyen, M. D., Nguyen, V. M., & Huh, E. M. (2013). An Improvement of Resource Allocation for Migration Process in Cloud Environment. *The Computer Journal Advance*.
- Nunes, L. F. (2006). Tomada de Decisão com Múltiplos Critérios: Pesquisa Ação sobre o Método AHP em Pequenas Empresas. *Universidade de Taubaté*.
- Oliveira, A. A. (2011). Aplicação de Método de Análise Hierárquica na Tomada de decisão para Adoção de Computação em Nuvem. *Universidade Federal do Rio Grande do Norte, Programa de Pós-graduação em Engenharia de Produção*.

- Oliverira, D. F., Tarcisio, A. V., & Santos, N. d. (2011). Computação em Nuvem para Gerenciamento de Rede de Ensino. *Universidade Federal de Rio de Janeiro*.
- Ramezani, F., Lu, J., & Hussain, F. (2013). An Online Fuzzy Decision Support System for Resource Management in Cloud Environments. *IEEE*.
- Rao, J., Bu, X., Wang, K., & Xu, C.-Z. (2011). Self-adaptive provisioning of virtualized resources in cloud computing. *ACM*, 129-130.
- Ravindran, K. (2013). QoS Auditing for Evaluation of SLA in Cloud-based Distributed Services. *IEEE Ninth World Congress on Services*.
- Rochwerger, B., Breitbang, D., Levy, E., Galis, A., Nagin, K., Liorente, I. M., et al. (2009). The RESERVOIR Model and Architecture for Open Federal Cloud Computing. *IBM Journal of Research and Development*.
- Russel, S., & Norvig, P. (2003). Artificial Intelligence: A Modern Approach. *Prentice-Hall: Englewood Cliffs, NJ, U.S.A.*
- Sartini, B. A., Garbugio, G., Bortolossi, H. J., Santos, P. A., & Barreto, L. S. (2004). Uma Introdução a Teoria dos Jogos. *Universidade Federal da Bahia*.
- Strukhoff, R., Deniz, Y., White, E., Romanski, P., & McMillanLiz. (14 de 10 de 2008). *Cloud Computing Public or Private? How to Choose Cloud Storage*. Acesso em 23 de 02 de 2014, disponível em Cloud Computing Journal: <http://cloudcomputing.sys-con.com/node/707840?page=0,1>
- Tai, J., Zhang, J., Li, J., & Meleis, W. (2011). ArA: Adaptive resource allocation for cloud computing environments under bursty workloads. *IEEE*, 1-8.
- Tchana, A., Broto, L., & Hagimont, D. (2012). Fault Tolerant Approaches in Cloud Computing Infrastructures. *Think Mind*.
- Thrall, R. M. (1984). A taxonomy for decision models. *Springer Link*, 23-27.
- Veras, M. (2011). *Cloud Computing: Nova Arquitetura da Ti*. Rio de Janeiro: Editora Brasport.
- Wei, G., Vasilakos, A. V., & Zheng, Y. (2010). A game-theoretic method of fair resource allocation.

- Wei, Y., Blake, B., & Saleh, I. (2013). Adaptive Resource Management for Service Workflows in Cloud Environments. *IEEE*.
- Wu, H., Zhang, W., Zhang, J., Wei, J., & Huang, T. (2012). A benefit-aware on-demand provisioning approach for multi-tier applications in cloud computing. *Institute of Software, Chinese Academy of Sciences, Beijing 100190, China*.
- Wu, L., Kumar, S., & Buyya, R. (2011). SLA-based Resource Allocation for Software as a Service Provider (SaaS) in Cloud Computing Environments . *Cloud Computing and Distributed Systems (CLOUDS) Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Australia*.
- Wuhib, F., Stadler, R., & Lindgren, H. (2012). Dynamic Resource Allocation with Management Objectives-Implementation for an OpenStack Cloud. *IEEE*.
- You, X., Wan, J., Xu, X., Jiang, C., Zhang, W., & Zhang, J. (2011). ARAS-M: Automatic Resource Allocation Strategy based on Market Mechanism in Cloud Computing. *School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou , China*.
- Zhang, Q., Cheng, L., & Boutaba, R. (2010). Cloud computing: state-of-the-art and research challenges. *The Brazilian Computer Society* .
- Zhang, Q., Zhu, Q., & Boutaba, R. (2011). Dynamic Resource Allocation for Spot Markets in Cloud Computing Environments. *Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, USA*.

