

VILNEI MARINS DE FREITAS DAS NEVES

Computação em Nuvem: Um Estudo sobre IaaS

Trabalho Individual apresentado ao Programa de Pós-Graduação em Computação da Universidade Federal de Pelotas, como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação

Orientador: Prof. Dr. Maurício Lima Pilla
Co-orientador: Prof. Dr. Adenauer Corrêa Yamin

Pelotas, 2014

RESUMO

NEVES, Vilnei Marins de Freitas das. **Computação em Nuvem: Um Estudo sobre IaaS**. 2014. 40 f. Trabalho Individual (Mestrado) – Programa de Pós-Graduação em Computação. Universidade Federal de Pelotas, Pelotas.

A Computação em Nuvem é um paradigma de computação distribuída cuja ideia central é promover a redução de custos operacionais através da otimização no uso dos recursos computacionais, concentrando o esforço de administração e gerência desses recursos e de sua infraestrutura em um mesmo provedor. Dessa forma, ela se propõe a fornecer infraestrutura computacional, plataforma de desenvolvimento e software aplicativo na forma de serviços altamente escaláveis, dinâmicos e entregues sob demanda aos usuários finais, criando um modelo de negócios onde provedores de serviços em nuvem podem cobrar dos consumidores para acessar seus serviços, cuja qualidade de um serviço (QoS) entregue a um consumidor por um prestador de serviço é regulada através dos Acordos de Nível de Serviço (SLA).

Para contemplar esse conjunto de características e atingir seu objetivo, a Computação em Nuvem apoia-se nas tecnologias provenientes de paradigmas da computação paralela e distribuída como a Computação em Grade e Computação em Cluster, somadas ao uso dos mecanismos de virtualização e as técnicas de consolidação de máquinas virtuais. Isso permite que os recursos gerenciados por ela possam ser rapidamente alocados e desalocados conforme a demanda, fornecendo assim a noção de elasticidade característica desse paradigma, além de contribuir para sua alta escalabilidade. Além disso, busca nas premissas Computação Utilitária seu horizonte conceitual, para oferecer poder computacional com um serviço mensurado e cobrado de acordo com seu uso.

Essas premissas, somadas ao amadurecimento de suas tecnologias, tem promovido expansão da Computação em Nuvem, e tem trazendo consigo diversas iniciativas de desenvolvimento de ferramentas e propostas de uso. Da mesma forma, essa expansão também traz um conjunto de desafios a Computação em Nuvem, que se ao mesmo tempo tornam-se barreiras para o avanço da mesma, criam oportunidade para várias frentes de pesquisa relacionadas. O presente trabalho busca fornecer uma visão geral sobre as principais questões acerca da Computação em Nuvem, discutindo sua proposta, suas características e origens, além de trazer a discussão alguns de seus desafios e uma sucinta comparação entre as principais ferramentas para implantação de um ambiente em nuvem.

Palavras-chave: Computação em Nuvem, Sistemas Distribuídos, Virtualização, Infraestrutura como Serviço.

ABSTRACT

NEVES, Vilnei Marins de Freitas das. **Cloud Computing: A IaaS Study**. 2014. 40 f. Trabalho Individual (Mestrado) – Programa de Pós-Graduação em Computação. Universidade Federal de Pelotas, Pelotas.

The Cloud Computing is a distributed computing paradigm whose focus is to promote the reduction of operating costs by optimizing the use of computational resources, concentrating the administration and management efforts of these resources and their infrastructure in a single provider. Thus, it is proposed to provide computing infrastructure, development platform and application software in the form of highly scalable, dynamic and delivered on demand services, creating a business model where providers of cloud services can charge for their services access, whose quality of service (QoS) delivered to a consumer by a service provider is regulated through the service Level Agreements (SLA).

To address these set of features, and achieve your goal, Cloud Computing builds on technologies from paradigms of parallel and distributed computing such as Grid Computing and Cluster Computing, coupled with the use of mechanisms of virtualization and consolidation techniques virtual machines. This allows resources managed by it can be rapidly allocated and deallocated on demand, thereby providing the notion of elasticity characteristic of this paradigm, and contribute to their high scalability. Also, the Cloud Computing relies on Utility Computing assumptions, to provide computing power measured with a service and charged according to their usage.

These assumptions, coupled with the development of their technologies, has promoted the expansion of Cloud Computing, and brings many initiatives to develop tools and use proposes. Likewise, this expansion also brings several challenges Cloud Computing, which is at the same time become barriers to the advancement of same, to create opportunity several lines of research related. This paper aims to provide an overview of the main issues about cloud computing, discussing its proposal, its characteristics and origins, and bring the discussion some of their challenges and a brief comparison of the main tools for the implementation of an environment cloud.

Keywords: Cloud Computing, Distributed Systems, Virtualization, Infrastructure as a Service.

LISTA DE FIGURAS

Figura 1	Visão Geral da Computação em Nuvem	11
Figura 2	Computação em Nuvem (FOSTER et al., 2008).	15
Figura 3	Classificação de uma Nuvem (WIND, 2011).	18
Figura 4	Visão geral da Arquitetura do Eucalyptus. (EUCALYPTUS, 2012).	25
Figura 5	Visão geral da Arquitetura do OpenNebula. (OPENNEBULA, 2012).	28
Figura 6	Visão geral da Arquitetura do Nimbus. (NIMBUS, 2012).	29
Figura 7	Visão geral da Arquitetura do <i>OpenStack</i> . (OPENSTACK, 2012).	33

LISTA DE TABELAS

Tabela 1	Paradigmas da Computação Distribuída (BUYYYA et al., 2009a)	16
Tabela 2	Comparação das Plataformas para Implantação e Gerenciamento de Infraestruturas de Ambientes em Nuvem	34

LISTA DE ABREVIATURAS E SIGLAS

AWS	Amazon Web Services
EC2	Elastic Compute Cloud
IaaS	Infrastructure as a Service
LXC	Linux Containers
PaaS	Platform as a Service
QoS	Quality of Services
SaaS	Software as a Service
SLA	Services Level Agreement

SUMÁRIO

1	INTRODUÇÃO	8
1.1	Estrutura do Texto	9
2	COMPUTAÇÃO EM NUVEM: PREMISSAS E CARACTERÍSTICAS	10
2.1	Introdução	10
2.1.1	Premissas	12
2.1.2	Características	15
2.2	Considerações sobre o Capítulo	20
3	COMPUTAÇÃO EM NUVEM: DESAFIOS	21
3.1	Provisionamento Automático de Recursos	21
3.2	Migração de Máquinas Virtuais	22
3.3	Consolidação de Servidores	22
3.4	Gerenciamento de Energia	23
3.5	Considerações sobre o Capítulo	23
4	COMPUTAÇÃO EM NUVEM: PROJETOS PARA IMPLANTAÇÃO DE INFRAESTRUTURA COMO SERVIÇO	24
4.1	Eucalyptus	25
4.2	OpenNebula	26
4.3	Nimbus	28
4.4	OpenStack	30
4.5	Considerações sobre o Capítulo	32
5	CONSIDERAÇÕES FINAIS	36
	REFERÊNCIAS	38

1 INTRODUÇÃO

A Computação em Nuvem é um paradigma de computação distribuída que tem conquistado um significativo espaço e destaque dentre as tecnologias que dão suporte e promovem o atual cenário da tecnologia da informação. E isso ocorre principalmente em virtude de duas principais necessidades desse cenário: redução de custos e diminuição da complexidade na manutenção e administração das infraestruturas que o suportam. A Computação em Nuvem busca atender essas necessidades através da otimização no uso dos recursos computacionais (FOSTER et al., 2008), compartilhando-os entre os diversos usuários e da concentração dos esforços na administração das infraestruturas relacionadas, através da figura de um provedor que fornece sua capacidade computacional na forma de serviço.

A Computação em Nuvem oferece um modelo de negócio, que utilizando-se das premissas e tecnologias da Computação em Grade e da Computação Utilitária, propõe-se a fornecer infraestrutura computacional, plataforma de desenvolvimento e *software* aplicativo na forma de serviços altamente escaláveis, dinâmicos e entregues sob demanda aos usuários finais (FOX; GRIFFITH et al., 2009). O acesso a esses serviços pode, então, ser cobrados de seus consumidores, estabelecendo assim relação de consumo regulada através dos Acordos de Nível de Serviço (*Services Level Agreements* - SLA), o qual estabelece os níveis aceitáveis de qualidade de um serviço (QoS) entregue a um consumidor por um prestador de serviço (BUYA et al., 2009b).

Para garantir que esse modelo consiga atender as expectativas de seus usuários quanto a escalabilidade e a disponibilidades de seus serviços, a Computação em Nuvem apoia-se em mecanismos que contribuam para a flexibilidade de operação e manutenção do ambiente que dão suporte aos seus serviços. Um exemplo disso é o uso da virtualização como estratégia de provimento de capacidade computacional e armazenamento de maneira. Com ela, os recursos podem ser rapidamente alocados e desalocados conforme a demanda, fornecendo assim a noção de elasticidade e possibilitando expandir os serviços

para grandes escalas, mesmo com o rápido aumento na demanda.

Com essas premissas e com amadurecimento de suas tecnologias, a Computação em Nuvem tem se expandido e atraído a atenção de usuários e empresas de várias áreas, buscando nela reduzir os custos associados a operação e manutenção das infraestruturas que dão suporte a atual tecnologia de informação. Além disso, a comunidade científica também tem mostrado interesse na Computação em Nuvem, não só pela redução de custos, mas também pela capacidade de expansão do poder computacional disponível, em um cenário que precisa lidar com um grande volume de informações.

No entanto, sua natureza compartilhada e dinâmica, bem como a necessidade de respeitar o acordo de nível de serviço (SLA) e atender os requisitos de qualidade de serviço (QoS), potencializados por seus crescimento, tem contribuído para gerar diversos desafios quanto ao uso e gerenciamento de um ambiente típico de nuvem, principalmente aqueles relacionados ao uso mais eficiente de seus recursos.

1.1 Estrutura do Texto

O presente trabalho está organizado em cinco capítulos. No capítulo um, esta introdução, é realizada uma introdução apresentando as principais ideias que caracterizam a Computação da Computação. No capítulo dois, são tratados os fundamentos da Computação em Nuvem, tratando dos suas principais características, sua origem e premissas. No capítulo três, são abordados os principais desafios enfrentados pela Computação em Nuvem. No capítulo quatro, é realizada uma revisão dos principais projetos que se propõem a colaborar com a implantação de ambientes em nuvem responsáveis por fornecer Infraestrutura como Serviço. E finalizando, no capítulo cinco, são apresentadas as considerações finais, as quais compreendem as principais conclusões sobre o tema estudado.

2 COMPUTAÇÃO EM NUVEM: PREMISSAS E CARACTERÍSTICAS

2.1 Introdução

A premissa central da Computação em Nuvem é fornecer recursos computacionais complexos na forma de serviços que podem ser acessados através de uma rede. Isso permite que se crie um modelo de negócios onde provedores de serviços em nuvem podem cobrar dos consumidores para acessar seus serviços (BUYA et al., 2009b).

Segundo Mell (MELL; GRANCE, 2011), a Computação em Nuvem pode ser entendida como um modelo que visa permitir o acesso de forma conveniente, ubíqua e sob demanda, a um conjunto de recursos computacionais configuráveis, tais como: redes, servidores, armazenamento, aplicativos e serviços, que podem ser rapidamente provisionados e liberados com o mínimo esforço de gestão ou de interação com o provedor de serviços. Nesse sentido, a Figura 1 mostra uma visão geral da Computação em Nuvem, onde é possível notar brevemente como os aspectos citados se organizam para compor esse tipo de ambiente. Nessa figura, a abstração de um nuvem é caracterizada pelos principais tipos de serviço que a mesma propõem-se a fornecer: (i) Aplicação, (ii) Plataforma, e (iii) Infraestrutura e com seus respectivos exemplos de recursos que podem ser fornecidos. Ainda, a figura busca demonstrar a forma ubíqua de como a nuvem pode ser acessada, caracterizada através dos dispositivos de natureza diversa quanto sua capacidade computacional e finalidade que podem fazer uso desse modelo de sistema distribuído.

Do ponto de vista do uso de recursos de hardware, (ARMBRUST et al., 2010; FOX; GRIFFITH et al., 2009) destacam três pontos centrais que colaboraram com a adoção da Computação em Nuvem:

1. A ilusão de recursos de computação infinitos e disponíveis sob demanda, eliminando assim a necessidade dos usuários realizarem planejamentos de longo prazo quanto ao provisionamento de recursos.

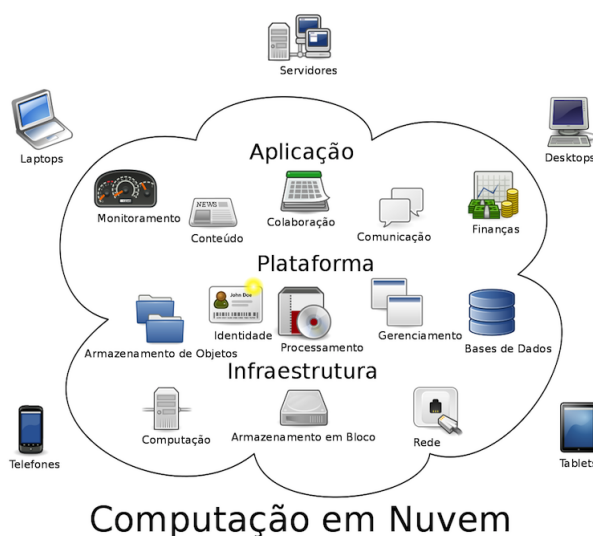


Figura 1: Visão Geral da Computação em Nuvem

2. A eliminação de um compromisso antecipado quanto ao uso de recursos e consequente alocação dos mesmos por parte dos usuário, permitindo assim que as empresas possam começar com um investimento menor e aumentar os recursos de hardware apenas quando há um aumento nas suas necessidades.
3. A capacidade de pagar pelo uso de recursos de computação com base em períodos de curto prazo (horas/dias), utilizando-os conforme necessário e liberá-los quando os mesmo não forem mais úteis.

A redução do custo operacional é um dos aspectos centrais para adoção de ambientes em nuvem. Essa redução é notada principalmente nos custos relacionados a energia elétrica, largura de banda de rede, na operação e administração de serviços ofertados por grandes *datacenters* (ARMBRUST et al., 2010; FOX; GRIFFITH et al., 2009). Como os recursos podem ser rapidamente alocados e desalocados conforme a demanda, não é necessário dimensioná-los para atender a maior carga de trabalho a longo prazo, fornecendo assim a noção de elasticidade característica desse paradigma. Da mesma forma, essa flexibilidade pode garantir uma alta escalabilidade, pois os provedores de infraestrutura podem assim disponibilizar uma grande quantidade de recursos a partir do seu conjunto de servidores e torná-los acessíveis, possibilitando expandir os serviços para grandes escalas, mesmo com o rápido aumento na demanda. No entanto para que se atinja esse objetivo e amplie ainda mais a capacidade de redução de custos, é necessário um gerenciamento eficiente dos recursos que fazem parte da infraestrutura da nuvem (BUYA et al., 2009b).

Do ponto de vista econômico, a Computação em Nuvem tem causado um impacto na área de TI semelhante ao impacto causado pela *terceirização* (*outsourcing*) na indústria de *hardware*. Nesse processo, diversas indústrias delegaram suas plantas produtivas para empresas que concentravam seus esforços e investimentos na pesquisa e melhorias na produção de semicondutores. Criou-se assim conceito de companhias *fab-less*, empresas que se concentram em buscar a inovação dos projetos de semicondutores, repassando a fabricação dos mesmos para outras indústrias, que focavam na otimização de suas linhas produtivas, e arrendavam para diversas companhias "fab-less". Com essa mesma visão - diminuição de riscos e custo de operação - a Computação em Nuvem busca criar um ambiente que permita a criação de empresas *datacenter-less*, empresas que forneçam produtos e serviços envolvendo tecnologia da informação, sem que as mesmas tenham que lidar com custos associados a administração e operação de *datacenters* (ARMBRUST et al., 2010; FOX; GRIFFITH et al., 2009) e com isso focando somente na prestação de seu serviço ou desenvolvimento de seu produto.

Outros aspectos também tem contribuído para a adoção da Computação em Nuvem e sua popularização quanto paradigma computacional, é possível destacar pelo menos dois aspectos principais: (i) aumento na capacidade de processamento e armazenamento de *hardware*, bem como sua rápida diminuição do custo, principalmente do *hardware commoditie*, incentivando sua adoção massiva na construção de *datacenters*; (ii) a adoção generalizada de serviços de computação e aplicações da Web, provocando crescimento exponencial da quantidade de dados disponíveis, especialmente os relacionados com o conteúdo da *Internet*, envolvendo rede sociais, comércio eletrônico, computação móvel (ARMBRUST et al., 2010; FOX; GRIFFITH et al., 2009; FOSTER et al., 2008).

Neste cenário, Computação em Nuvem tem gerado um considerável interesse e demonstrado crescimento em sua adoção nos últimos anos. Isso pode ser demonstrado no aumento de respectivamente, 9% e 14% no período que compreende os anos de 2011 e 2012, motivado por sua capacidade de redução de custos e por sua facilidade de acesso a recursos computacionais que a mesma pode proporcionar. Nas seções seguintes, serão discutidas sua origem e seu posicionamento quanto aos demais paradigmas de computação distribuída, bem como as características que a definem.

2.1.1 Premissas

A concepção de entregar poder computacional como serviço não é necessariamente uma concepção nova para a área de tecnologia da informação. Em

1961, John McCarthy (PARKHILL, 1966) sugeriu que “a computação pode algum dia ser organizada como uma utilidade pública”. Essa ideia também era compartilhada por Leonard Kleinrock, um dos principais cientistas da ARPANET, que idealizava em 1969, isso poderia acontecer no futuro através do crescimento das redes de computadores e sua evolução.

Essa ideia traduzia uma visão de futuro onde o acesso aos recursos computacionais estariam disponíveis ao público geral de forma transparente, possibilitando o uso da capacidade computacional disponível - tais como processamento e armazenamento, com a mesma facilidade com que são utilizados os serviços de fornecimento de água e energia elétrica, sendo assim, mensurados e cobrados. Essa forma de pensamento deu origem então ao modelo de *Utility Computing*, onde os recursos de computação, são vistos como serviços disponibilizados, mensurados e cobrados (FOSTER et al., 2008).

O modelo de *Utility Computing* inspira tanto a Computação em Grade, quanto a Computação em Nuvem em suas concepções. Ambas possuem como premissa a entrega de poder de processamento ao usuário final, encapsulado em uma estrutura mais abstrata e que pode ser mensurado e cobrado por seu uso.

A Computação em Nuvem, da mesma forma que apoia-se na *Utility Computing* como horizonte conceitual, usa como base, diversas premissas que foram herdadas de modelos de computação distribuída e de alto desempenho mais consolidados, possuindo diversos pontos em comum com a Computação em Grade e a Computação em Cluster.

Por exemplo, um *cluster* pode ser entendido como um tipo de sistema paralelo e distribuído que consiste de um conjunto de computadores interconectados, e que executam uma aplicação de forma conjunta, como se fosse um único recurso computacional. De forma geral, em um ambiente como esse, os nodos são interligados por uma rede de alta velocidade, sendo compostos por computadores ou servidores com características semelhantes, sendo seus recursos localizados concentrados sobre o mesmo domínio administrativo e gerenciado por uma única entidade. Desta forma, aspectos como a centralização da administração e uma maior homogeneidade quanto aos recursos físicos são mais evidenciados nesse tipo de ambiente.

Já na Computação em Grade, é um modelo de computação distribuída focado em resolver problemas de forma coordenada e dinâmica. Esse modelo busca fornecer mecanismos que permitam a diferentes instituições compartilharem recursos entre si, potencializando assim sua capacidade na resolução desses problemas. Em um ambiente em grade, compartilhamento de recursos e informações é feito através de uma abstração na forma de uma organização virtual, que realiza a integração dos recursos como processamento, armazenamento e rede de

diferentes instituições e os entrega como um conjunto mais abstrato de recursos.

A Computação em Grade possui um modelo de negócio orientado a projeto, onde as instituições ligadas ao mesmo e que podem estar distribuídas geograficamente, disponibilizam seus recursos computacionais de forma a atingir um objetivo comum. Cada instituição é responsável da sua estrutura computacional e a fornece aos demais "parceiros" conforme sua capacidade, disponibilidade e interesse, criando assim um ambiente heterogêneo e dinâmico.

Assim, a Computação em Grade busca concentrar toda capacidade de computação fornecida através dos diversos dispositivos disponíveis na resolução de uma grande tarefa, buscando com isso, extrair o máximo de computação para resolução da mesma.

Por sua vez, a Computação em Nuvem vem aproveitando os diversos conceitos e estratégias desenvolvidos ao longo do tempo para atender as demandas e enfrentar os desafios encontrados por estes paradigmas. No entanto, apesar das similaridades quanto sua origem e quanto alguns conceitos, a Computação em Nuvem se diferencia dos demais paradigmas principalmente pelos problemas que buscam solucionar e na forma de gerenciar e disponibilizar seus recursos. Um ambiente em nuvem tende a concentrar seus recursos em um único fornecedor, que administra e mantém toda infraestrutura física e disponibiliza seus recursos na forma de um serviço mensurado e cobrado conforme seu uso.

Além disso, a Computação em Nuvem vem tentando consolidar um modelo de negócio onde é possível contratar recursos computacionais de grande porte, como capacidade de armazenamento, processamento e transferências de dados, através de uma operação de compra eletrônica, de forma semelhante a qualquer outro item comprado em empresas de *e-commerce*, tendo acesso a esses recursos de forma praticamente imediata e na forma de um serviço fornecido.

Nessa perspectiva, a Computação em Nuvem busca disponibilizar essa capacidade computacional para diversos conjuntos de problemas, que de forma geral, pertencem a um escopo mais direcionado como as relações comerciais que emergem junto com o crescimento da Internet. Esses problemas se caracterizam por possuírem necessidade de poder computacional que oscilam muito durante o tempo. Isso exige que o ambiente em nuvem seja capaz de gerenciar e escalar seus recursos conforme a demanda dos diversos usuários que compartilham e competem pelos mesmos.

Uma visão sobre o posicionamento da Computação em Nuvem quando aos demais paradigmas da Computação Distribuída pode ser visto na figura 2. Nela, a Computação em Nuvem situa-se em uma região que demonstra sua flexibilidade quanto a escalabilidade, podendo a mesma ser tão escalável quanto aos

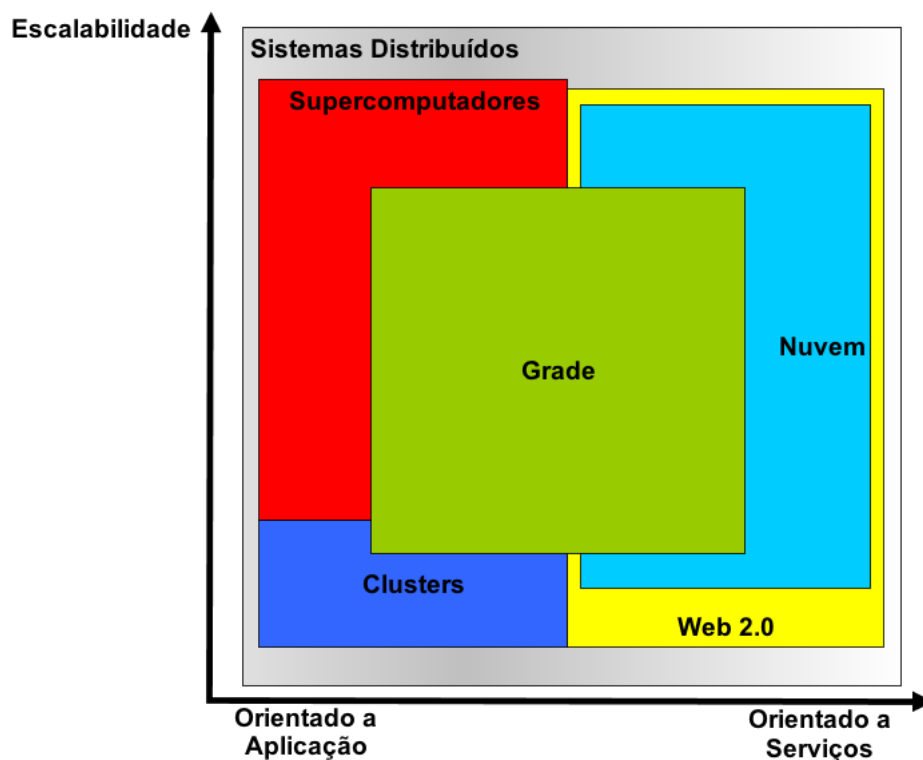


Figura 2: Computação em Nuvem (FOSTER et al., 2008).

demaís paradigmas de Computação Distribuída. Além disso, demonstra a sua vocação para atender aplicações orientadas a serviço, especialmente aquelas cuja natureza estejam vinculadas à *Internet* e à *Web 2.0*, da mesma forma que tenta mostrar sua forte ligação com a Computação em Grade.

Um resumo das diferenças pode ser visto na tabela 1, de forma a sumarizar e permitir uma comparação mais objetiva.

Além disso, uma discussão mais aprofundada sobre Computação em Grade, bem como uma comparação mais detalhada com a Computação em Nuvem, podem ser vistas em (FOSTER; KESSELMAN; TUECKE, 2001) e (FOSTER et al., 2008), respectivamente.

2.1.2 Características

A Computação em Nuvem pode ser caracterizada como um tipo de sistema paralelo e distribuído que consiste em um conjunto de computadores interligados e virtualizados, que são dinamicamente provisionados e apresentados como um ou mais recursos de computação unificada, os quais são regidos com base em acordos de nível de serviço estabelecidos através de negociação entre o

Característica	Cluster	Grids	Nuvem
Hardware	De Prateleira	De Alto Desempenho	Ambos
SOs instalados no Nodos	Um S.O padrão (Win/Linux)	Diversos	Hipervisor com SOs executados sobre VMs
Proprietário	Único	Múltiplos	Único
Interconexão	Dedicada Alto desempenho Baixa latência Alta largura de banda	Maioria através da Internet Alta latência Baixa largura de banda	Dedicada Alto desempenho Baixa latência Alta largura de banda
Negociação de Serviço	Limitada	Sim, baseada em SLA	Sim, baseada em SLA
Gerenciamento de Usuários	Centralizado	Descentralizado	Centralizado Possibilidade de delegação
Gerenciamento de Recursos	Centralizado	Distribuído	Centralizado/ Distribuído
Interoperabilidade	VIA (Virtual Interface Architecture)	Padrões Abertos promovidos pelo OGF <i>Open Grid Forum</i>)	<i>Web Services</i> (REST/SOAP)
Capacidade	Estável e Garantida	Variável	Provisionada sob Demanda
Tipos de Aplicações	Científicas Corporativas Comerciais	Colaboração Científica e de Alto Desempenho	Legadas/Web dinamicamente provisionadas Entrega de Conteúdo

Tabela 1: Paradigmas da Computação Distribuída (BUYA et al., 2009a)

prestador de serviço e os consumidores (BUYYA et al., 2009b). De acordo com Mell (MELL; GRANCE, 2011), esse modelo apresenta cinco características chave:

- *Auto-Atendimento sob demanda*: um consumidor pode obter imediatamente um recurso por um determinado tempo, sem que para isso exista qualquer tipo de intervenção humana.
- *Ampla Acesso à Rede*: os recursos fornecidos pela nuvem são entregues através da rede - Internet, por exemplo - sendo usado por vários aplicativos clientes de diferentes plataformas.
- *Agrupamento de Recursos*: os recursos de computacionais de um fornecedor são agrupados, com a intenção de servir a vários consumidores utilizando um modelo de *multi-tenancy*, onde os diferentes recursos físicos e/ou virtuais são atribuídos e reatribuídos conforme a demanda do consumidor.
- *Elasticidade rápida*: essa característica dá ao consumidor a sensação de que os recursos computacionais podem ser fornecido de maneira quase que infinita. A premissa é que os recursos podem ser fornecidos de forma rápida e elástica, de acordo com a demanda do usuário, e assim que não forem mais necessário, são liberados. Isso permite que o ambiente consiga tratar os momentos de pico das cargas de trabalho, escalando rapidamente.
- *Serviços Mensurados*: os sistemas em nuvem são capazes de controlar seus recursos de forma automática, com o objetivo de otimizar seus recursos, permitindo compartilhar os mesmos entre vários usuários, em diferentes configurações. Da mesma forma, esses sistemas possuem a capacidade de mensurar o uso desses recursos, podendo dessa forma cobrar pelo uso de processamento, armazenamento, largura de banda, oferecendo assim transparência tanto para o provedor e consumidor do serviço utilizado.

A premissa de entregar capacidade computacional como um serviço tem sua origem no modelo de *Utility Computing*. Os serviços são regulados através dos Acordos de Nível de Serviço (*Services Level Agreements* - SLA), os quais estabelecem os níveis aceitáveis de qualidade de um serviço que devem ser fornecidos. Dessa forma, o surgimento da Computação em Nuvem pode ser entendida como uma mudança de foco de uma infraestrutura que oferece armazenamento e recursos de computação – como é o caso da Computação em Grade – para

outra cuja a economia está focada em oferecer recursos mais abstratos na forma de serviços (FOSTER et al., 2008).

Revisar
e Avaliar
transformar em
Seção

Outra característica chave da Computação em Nuvem é o uso da virtualização. A virtualização oferece a abstração necessária para que os recursos computacionais físicos (processamento, armazenamento, recursos de rede) possam ser unificados como um conjunto de recursos mais abstratos, permitindo que outros serviços possam ser construídos sobre estes (FOSTER et al., 2008). Através dela é possível compartilhar a mesma máquina física com múltiplos sistemas operacionais e/ou aplicações de usuários finais, mantendo garantias quanto ao desempenho (ADDIS et al., 2010) e isolamento entre as mesmas. Essa propriedade garante um nível de encapsulamento e flexibilidade que permite que tais máquinas possam ser manipuladas e tenham suas aplicações migradas entre *servidores* com maior facilidade do que se isto ocorresse usando-se somente máquinas físicas. Essas características fornecem melhor segurança, gerenciamento e isolamento.

Uma solução baseada na Computação em Nuvem pode, de forma geral, ser classificada de acordo com o modelo de implantação e de acordo com o tipo de serviço prestado. Na Figura 3 é possível ter uma visão geral dessa classificação.

De acordo com o modelo de implantação, as "nuvens" podem ser classificadas como: (i) Nuvens Privadas, quando a nuvem só pode ser usada por uma instituição ou empresa, que mantém dentro de sua própria estrutura computacional as aplicações e recursos que usados por ela; (ii) Nuvens Públicas, quando os serviços são fornecidos e comercializados por uma organização a outras instituições/empresas ou ao público; (iii) Nuvens Comunitárias são aquelas cuja sua infraestrutura é composta por vários fornecedores com a finalidade de prestar serviços a uma comunidade específica; e (iv) Nuvens Híbridas são aquelas cuja infraestrutura de nuvem é uma composição de dois ou mais tipos de nuvem, que permanecem como entidades únicas, mas são unidas por uma tecnologia padronizada ou proprietária que permite a portabilidade de dados e de aplicação, permitindo por exemplo, um balanceamento de carga entre nuvens distintas (WIND, 2011).

A classificação quanto ao tipo de serviço ofertado pode ser feito de acordo com o nível de abstração do serviço fornecido. Em geral, a Computação em Nuvem fornece três níveis de serviço:

- **Infraestrutura como Serviço (IaaS):** nesse nível de serviço, o provedor oferece uma infraestrutura virtualizada de servidores, armazenamento e rede, que pode ser fornecida sob demanda e cobrada conforme seu uso. Nesse tipo de serviço, o cliente pode criar uma máquina virtual e instalar sobre a mesma, uma pilha completa de software, incluindo nisso, o sistema

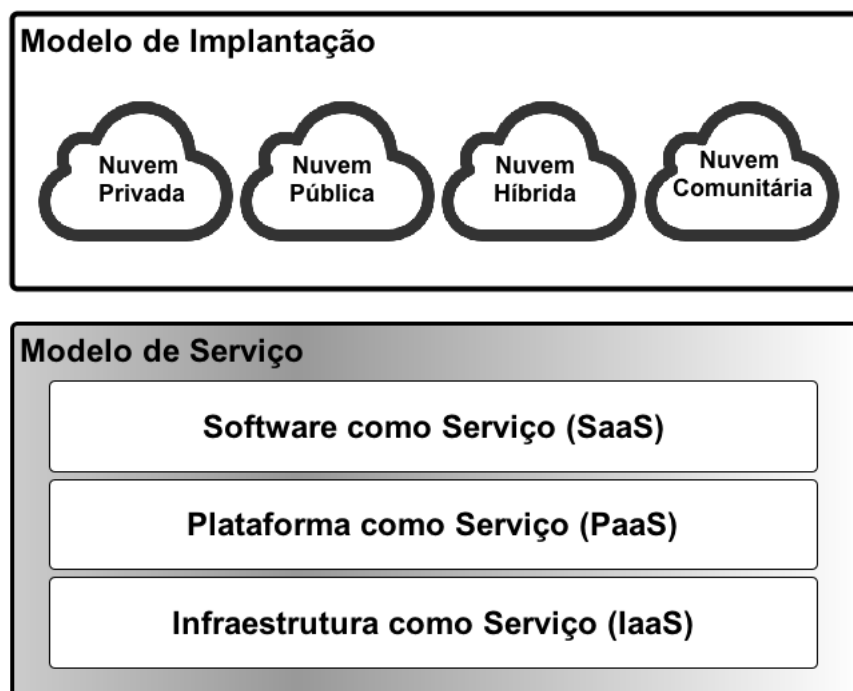


Figura 3: Classificação de uma Nuvem (WIND, 2011).

operacional. O acesso ao servidor ou servidores, e aos serviços fornecidos por esse, acontece normalmente através da rede. Esse formato de serviço garante uma maior flexibilidade, no entanto, exige do cliente maior conhecimento para que esse possa usá-lo. Um exemplo para este tipo de serviço é o fornecido pela Amazon através do *Amazon's Elastic Compute Cloud* (EC2) (AMAZON, 2013a).

- **Plataforma como Serviço (PaaS):** oferece um ambiente de alto nível integrado para construir, testar e implementar aplicativos personalizados. A vantagem imediata desse tipo de serviço é a facilidade de uso e a possibilidade dos clientes em manter o foco na criação e manutenção de seus próprios serviços. A manutenção do *hardware* (virtual) e software é responsabilidade do provedor de nuvem. Uma restrição nesse nível, deve-se ao fato de que os serviços dos clientes precisam ser compatíveis com a nuvem. Como exemplo, pode-se citar o *Google Apps* (GOOGLE, 2013).
- **Software como Serviço (SaaS):** fornece software de propósito especial que é acessível remotamente pelos consumidores através da Internet com

um modelo de cobrança baseada em uso. A principal característica desse tipo de serviço é o fato dos fornecedores ficarem responsáveis por gerenciar toda a plataforma virtual e as aplicações sobre esta. Como toda a manutenção é feita pelo fornecedor, os clientes só precisam se concentrar na avaliação de serviços de *software* e usá-los. Nesse nível de serviço não é possível negociar atualizações do software. Serviços como *Dropbox* (DROPBOX, 2012) , *Facebook* (FACEBOOK, 2012), *Gmail* (GMAIL, 2012), *Salesforce* (SALESFORCE, 2012) são exemplos desse nível de serviço.

2.2 Considerações sobre o Capítulo

Esse capítulo buscou realizar uma introdução sobre os principais aspectos relacionado com a Computação em Nuvem, trazendo suas premissas, origens e características que a definem. O próximo capítulo traz à discussão os principais desafios que a Computação em Nuvem precisa enfrentar durante seu processo de desenvolvimento como paradigma de Computação Distribuída.

3 COMPUTAÇÃO EM NUVEM: DESAFIOS

Como discutido anteriormente, a Computação em Nuvem busca otimizar o uso de recursos, compartilhando-os entre os diversos usuários, tentando criar um ambiente altamente escalável e flexível quanto as necessidades e ofertas de processamento e armazenamento para seus usuários. E para isso, apoia-se nos mecanismos de virtualização para garantir um conjunto de características que permitam facilitar o gerenciamento desse ambiente. Esse cenário gera uma série de desafios ligados à natureza compartilhada e dinâmica do ambiente em nuvem, bem como a necessidade de respeitar o acordo de nível de serviço (SLA) e atender os requisitos de qualidade de serviço (QoS). (ZHANG; CHENG; BOU-TABA, 2010).

3.1 Provisionamento Automático de Recursos

Uma das principais características da Computação em Nuvem é sua capacidade de aquisição e de liberação de recursos conforme a demanda à qual o ambiente é exposto. Nesse sentido, é esperado que o ambiente seja capaz de realizar a alocação de recursos e a sua subsequente liberação conforme a necessidade da aplicação, visando sempre atender os requisitos de QoS, estabelecidos no SLA, e minimizar os custos relacionados a sua operação. No entanto, não é uma tarefa trivial alcançar esse tipo de objetivo. Em ambientes dinâmicos como os em nuvem, onde as demandas por recursos oscilam drasticamente em um curto período de tempo, prever o quanto de recurso é necessário em um determinado momento torna-se um desafio, especialmente quando é necessário atender requisitos de QoS. É preciso agilidade para atender rapidamente a flutuações de demanda, em especial aquelas que são caracterizadas por um crescimento súbito, como em picos de acesso e utilização.

3.2 Migração de Máquinas Virtuais

A virtualização tornou a principal tecnologia usada em infraestruturas em nuvem. Através do uso da virtualização, é possível compartilhar a mesma máquina física com múltiplos sistemas operacionais e/ou aplicações de usuários finais, promovendo o isolamento entre os mesmos. Da mesma forma, o seu uso torna possível realizar balanceamento de carga entre estruturas físicas, através de ações como a migração de máquinas virtuais entre servidores.

A técnica de migração de máquinas virtuais evoluiu das técnicas de migração de processos. As ferramentas de implementação de infraestruturas como *OpenStack* (OPENSTACK, 2012), *OpenNebula* (MORENO-VOZMEDIANO; MONTERO; LLORENTE, 2012a), *Eucalyptus* (EUCALYPTUS, 2012), dentre outros, implementam o conceito de *live-migration*, onde as máquinas virtuais (VM) são movidas em um processo que envolve paradas extremamente curtas que variam de dezenas de milissegundos a um segundo.

Além disso, uma migração de um sistema operacional inteiro e todas as suas aplicações como uma unidade permite evitar dificuldades enfrentadas por abordagens em nível de processo. Os maiores benefícios da migração de VM é evitar sobrecarga em pontos específicos da infraestrutura de servidores, no entanto, esta não é uma tarefa simples. Atualmente, a detecção de pontos de sobrecarga e o disparo de uma migração não tem a agilidade suficiente para responder às mudanças de carga de trabalho repentinas. Além disso, é necessário transferir o estado dos dados em memória de forma consistente e eficiente, considerando os recursos para as aplicações e para os servidores físicos.

3.3 Consolidação de Servidores

A consolidação de servidores é uma abordagem que visa maximizar a utilização de recursos, buscando como contrapartida a minimização do consumo de energia em um ambiente em nuvem.

A idéia central dessa abordagem é deslocar as VMs que residem em múltiplos servidores que estejam subutilizados para um único servidor usando *live-migration*, de modo que os demais servidores podem ser colocados em um estado de economia de energia. Do ponto de vista tecnológico, a consolidação de várias cargas de trabalho provenientes de diferentes usuários, em uma mesma máquina física, permite reduzir custos operacionais relacionados a sua manutenção, promovendo com isso, uma maior utilização dos recursos físicos disponíveis. No entanto, as flutuações de carga de trabalho imprevistas ou falhas de *hardware* podem causar um impacto considerável entre vários aplicativos que estejam em execução nesse ambiente (ADDIS et al., 2010).

Além disso, a alocação de vários servidores em uma única máquina física, pode gerar um problema de otimização, formulado como uma variante do problema da *mochila booleana* (?), que é classificado como *NP-HARD*, o que pode dificultar o uso dessa abordagem (ZHANG; CHENG; BOUTABA, 2010).

3.4 Gerenciamento de Energia

O consumo de energia é um aspecto importante na Computação em Nuvem, decorrente do crescimento da infraestrutura responsável por seu funcionamento e manutenção, como equipamento de rede, armazenamento, processamento, redundância e etc.

Segundo Koomey (KOOMEY, 2011), no período de 2005 a 2010 a energia consumida na manutenção dos centro de dados aumentou em 56%. No final desse período, a energia utilizada para manter essas estruturas, correspondia a aproximadamente 1,5% do total da energia consumida globalmente. Esse consumo não contribui somente para causar impacto na economia, aumentando os custos operacionais, mas também influenciam muito no impacto ambiental, principalmente no aumento da emissão de dióxido carbono resultante das fontes de geração de energia. Dessa forma, existe uma preocupação em otimizar o uso de recursos, observando a eficiência energética e sem comprometer o desempenho dos serviços fornecidos, visto em trabalhos como (BELOGLAZOV; ABAWAJY; BUYYA, 2012).

Assim, através do uso mais eficiente da infraestrutura computacional, a Computação em Nuvem pode contribuir com a redução nas emissões de dióxido carbono, colaborando com a diminuição no impacto causado ao meio ambiente e evitando o esgotamento de fontes de energia não-renováveis (THEGREENGRID, 2012).

3.5 Considerações sobre o Capítulo

Esse capítulo trouxe a discussão os principais desafios que a Computação em Nuvem precisa enfrentar durante seu desenvolvimento e amadurecimento. O próximo capítulo apresentar os principais projetos *opensource* com foco na implantação de estruturas em nuvem que forneçam infraestruturas computacionais como serviço (*IaaS*).

4 COMPUTAÇÃO EM NUVEM: PROJETOS PARA IMPLANTAÇÃO DE INFRAESTRUTURA COMO SERVIÇO

A Computação em Nuvem permite a elaboração de diversas estratégias e tecnologias na construção de serviços que atendam as demandas de diversos públicos, mesclando níveis de abstração distintos. Como comentado no capítulo anterior, a Computação em Nuvem fornece diferentes modelos de implantação, com abordagens que englobam desde a manipulação de estruturas no nível de sistema operacionais ao uso de recursos no nível de aplicativos.

Neste texto, optou-se por apresentar os principais projetos que se propõem a colaborar com a implantação de estruturas em nuvem que forneçam infraestruturas computacionais como serviço (*IaaS*). No entanto, antes de explorarmos as características de cada projeto, é oportuno observarmos as características arquiteturais que projetos como esses costumam compartilhar.

Um típico ambiente de *IaaS* é caracterizado pela existência de um grande número de componentes, fazendo uso de diversas tecnologias e estratégias para alcançar seus objetivos. De forma geral, em ambientes como esses é possível identificar seis componentes (SEMPOLINSKI; THAIN, 2010) que formam suas bases: (i) *hardware* e sistemas operacionais, em geral, diversos; (ii) rede, com mecanismos que permitam criar e gerenciar redes virtuais que interligam a VMs; (iii) gerenciadores de virtualização ou hipervisores, que permitam criar e gerenciar diversas VMs; (iv) um repositório com imagens de máquinas virtuais e discos; (v) interface de administração que permita ao usuário gerenciar os recursos que lhe foram fornecidos; e por último (vi) os projetos que se a propõem a implementar e gerenciar ambientes em nuvem, unificando e coordenando diversas tecnologias com propósito de fornecer capacidade computacional como um serviço homogêneo, flexível e sob demanda.

Nas seções seguintes, serão apresentados os principais projetos de código abertos que se propõem a colaborar com a implantação de estruturas em nuvem no formato (*IaaS*), mostrando suas principais características e estruturas.

4.1 Eucalyptus

O projeto Eucalyptus (NURMI et al., 2009) (EUCALYPTUS, 2012) é uma solução de código aberto que permite a instalação de infraestruturas de nuvens privadas e híbridas, sendo capaz de gerenciar e controlar grandes coleções de recursos distribuídos. Ele foi projetado para atender as características centrais da computação em nuvem, possuindo mecanismos que facilitem o crescimento e comunicação de sua infraestrutura com outras estruturas em nuvem. Para isso, o projeto Eucalyptus permitem integrar recursos computacionais e de armazenamento que possam ser facilmente adquiridos, através do uso de *hardware commodity* e que se comuniquem através de interfaces padronizadas, sendo por exemplo, compatível com as plataformas da *Amazon EC2* e *S3*.

A arquitetura do Eucalyptus - vista na Figura 4 - é uma arquitetura modular e flexível, desenvolvida usando combinação das linguagens Java e C, e construída sobre 5 componentes que se comunicam através do uso de interfaces bem definidas e com papéis claros quanto a suas atribuições:

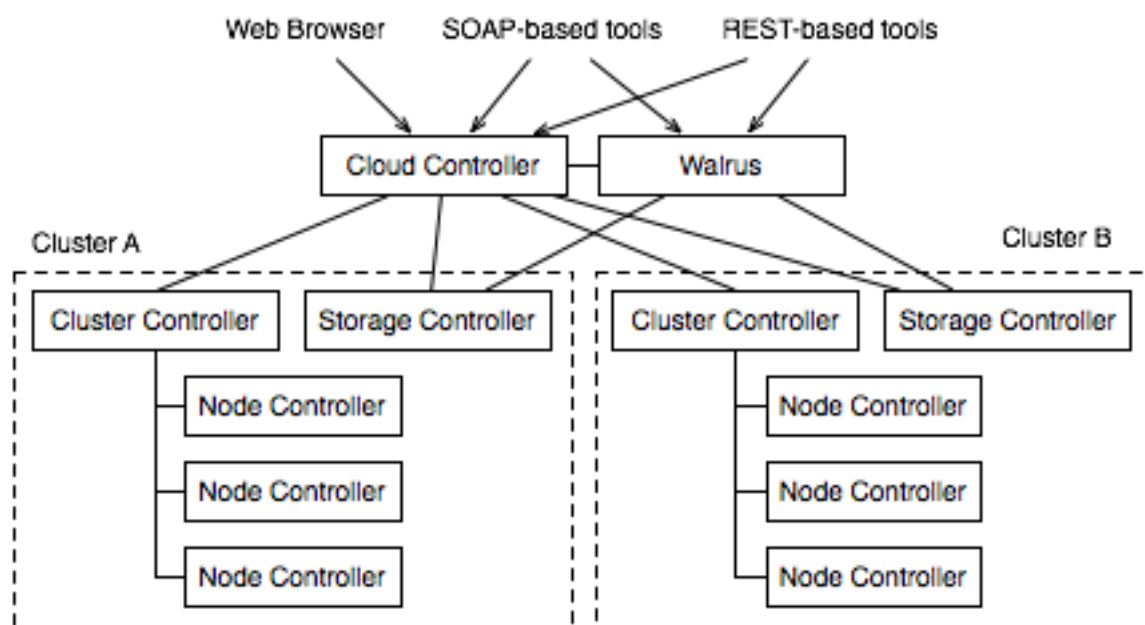


Figura 4: Visão geral da Arquitetura do Eucalyptus. (EUCALYPTUS, 2012).

- **Controlador de Nodo (Node Controller - NC):** é componente básico de cada nodo, desenvolvido em C, sendo responsável por gerenciar o ciclo de vida das instâncias em execução em cada um deles, interagindo com o sistema operacional, hipervisores suportados pelo projeto - Xen, KVM, VMWare - e com o Controlador de Cluster;
- **Controlador de Cluster (Cluster Controller - CC):** É responsável por gerenciar os Controladores de Nodo, possuindo a responsabilidade de

implantação e gerenciamento das instâncias em execução nesses. Para isso, ele se comunica com o Controlador de Nós e com o Controlador de Nuvem de forma simultânea, sendo às vezes, responsável por gerenciar também o acesso às redes de algumas instâncias em execução. Da mesma forma que o *Node Controller* é desenvolvido em C;

- **Controlador de Nuvem (Cloud Controller - CLC):** Desenvolvido em Java, é responsável por fornecer um front-end para acesso ao ecossistema do Eucalyptus. Ele fornece um conjunto de webservices que interage tanto com ferramentas clientes quanto com os demais componentes da infraestrutura do Eucalyptus;
- **Walrus Storage Controller (WS3):** componente responsável por armazenar as imagens utilizadas pelas VMs, dados de usuários e os *snapshots* de volumes e pode ser acessado através do uso da API do Amazon S3. Desenvolvido em Java;
- **Storage Controller (SC):** Fornece armazenamento em bloco persistente - sobre AoE ou iSCSI - para as instâncias, permitindo a criação de *snapshots* de volume. Também desenvolvido em Java.

Até o desenvolvimento do *Walrus*, o Eucalyptus não possuía um mecanismo próprio que suportasse *live-migration*, dependendo assim das funcionalidades do hipervisor usado. Com o *Walrus*, as imagens de discos virtuais estão distribuídas, e quando existe a necessidade de migrar um servidor, esse disco é replicado em nível de blocos, para então possam realizar a migração (GRAUBNER; SCHMIDT; FREISLEBEN, 2011).

Ficou bom?

Na mais recente versão, o projeto *Eucalyptus* aumentou a compatibilidade com as estruturas de nuvem fornecidas pela Amazon, através da implementação de mecanismos que utilizam APIs e ferramentas compatíveis com Amazon EC2, como: (i) *Auto Scaling*, que permite manipular a escalabilidade dos recursos; (ii) *Elastic Load Balancing*, que distribui o tráfego de várias aplicações entre as diversas instâncias do *Eucalyptus*, de forma a aumentar a tolerância a falhas das aplicações; e o (iii) *CloudWatch*, que provê o monitoramento de recursos e aplicações que estão sendo executadas nas nuvens que usam o *Eucalyptus*.

4.2 OpenNebula

O projeto OpenNebula (MORENO-VOZMEDIANO; MONTERO; LLORENTE, 2012b) reúne um conjunto de ferramentas de código aberto e visa auxiliar na

implantação de um ambiente de infraestrutura em nuvem. Ele tem como objetivo fornecer uma camada de gerenciamento desse ambiente de forma aberta, flexível, extensível e abrangente para automatizar e orquestrar o funcionamento de vários servidores virtualizados. Para isso, se propõe a aproveitar e integrar soluções já existentes implantadas para a rede, monitoramento, virtualização, armazenamento e gerenciamento de usuários.

Além disso, *OpenNebula* possui suporte à infraestruturas em nuvens federadas através de interfaces com nuvens externas, o que proporciona isolamento, escalabilidade e suporte a múltiplos sítios. Isso permite que as organizações possam complementar a sua infraestrutura local com a capacidade de computação de outras nuvens, de forma a atender às demandas durante picos de uso, ou implementar estratégias de alta disponibilidade. Assim, um único ponto de acesso e um sistema de gestão centralizada pode ser utilizado para controlar várias implementações de *OpenNebula* (LASZEWSKI et al., 2012).

O *OpenNebula*, além de expor uma interface XML-RPC para diversas ações sobre o ambiente - e.g. invocar VMs -, e para comunicação entre os módulos do conjunto, o *OpenNebula* suporta outras diferentes interfaces de acesso: interfaces baseadas em REST, baseadas em serviços OGF OCCI, e a mais recente API padrão para Computação em Nuvem, a API fornecida pelo AWS EC2.

A autorização é baseada em senhas, pares de chaves SSH RSA, certificados *X.509* ou *LDAP*. Ele também suporta ACLs mais detalhadas para gerenciar usuários com múltiplas funções, como parte de suas capacidades de autorização.

Quanto ao armazenamento, o *OpenNebula* suporta várias configurações de servidor: de sistemas de arquivos não-compartilhados com transferências de imagens via SSH, a sistemas de arquivos compartilhados (NFS, GlusterFS, Lustre) ou LVM com CoW (*copy-on-write*). Além disso, possui suporte a qualquer tipo de servidor de armazenamento, desde soluções baseadas em hardware padrão a soluções de nível empresarial (*storages*).

A arquitetura interna do *OpenNebula* pode ser dividida em três camadas: (i) *Tools*, que possui as ferramentas de gerenciamento providas pela camada *Core*; (ii) *Core*, responsável pelo gerenciamento da máquinas virtuais, rede virtual e pelos mecanismo de hospedagem; e a camada de (iii) *Driver*, responsável por gerenciar mecanismo de virtualização, armazenamento e monitoramento na forma de módulos conectáveis para permitir a interação da camada de *Core* com *middleware* específico. A organização de sua arquitetura pode ser vista na Figura 5

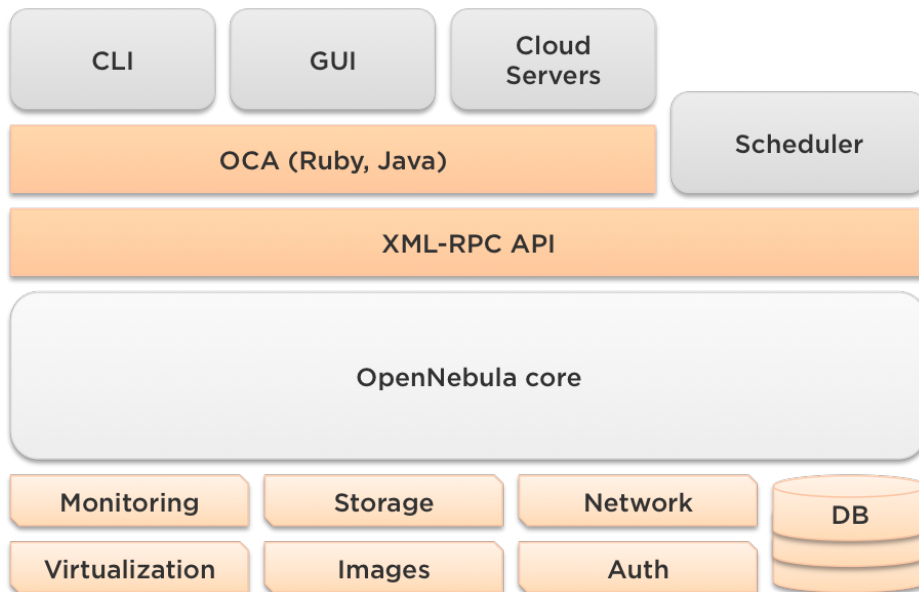


Figura 5: Visão geral da Arquitetura do OpenNebula. (OPENNEBULA, 2012).

4.3 Nimbus

Nimbus (NIMBUS, 2012) é um conjunto de ferramentas de código aberto, que tem como missão fornecer de um serviço no formato *IaaS* para a comunidade científica através do acesso e utilização de ambientes em *cluster* e/ou grade.

De forma a alcançar o objetivo proposto, os esforços de trabalho do projeto *Nimbus* estão direcionados em duas frentes de desenvolvimento: (i) a *Infraestrutura Nimbus* e a (ii) *Plataforma Nimbus*.

A *Infraestrutura Nimbus* é uma implementação de código aberto no formato *IaaS* compatível com os serviços EC2/S3 da *Amazon* (AMAZON, 2013b), e tem como intenção atender demandas da comunidade científica, suportando funcionalidades inerentes de ambientes em grade, como *batch-schedulers*, acesso via interface WSRF (*Web Services Resource Framework*) e etc, além de suportar mecanismos de virtualização como Xen e KVM.

A *Plataforma Nimbus*, por sua vez, concentra o foco em fornecer um conjunto de ferramentas adicionais para simplificar a gestão dos serviços de infraestrutura e facilitar a integração com outras nuvens existentes, como *OpenStack* (OPENS-TACK, 2012) e *Amazon* (AMAZON, 2013b).

O *Nimbus*, por possuir em sua proposta grande afinidade com os projetos e ambientes em grade, é capaz de se integrar a mecanismos como o *Globus*, utilizando diversos componentes do mesmo, como os mecanismos que gerem as credenciais e autenticação de usuários, além de mecanismos de armazenamento como *GridFTP*, também pertencente ao projeto *Globus*, sendo a mesma, substituída por uma nova abordagem originária do próprio projeto *Nimbus*: o

Nimbus Cumulus.

O projeto *Cumulus* é ferramenta importante dentro do projeto *Nimbus*, apresentado-se como uma implementação de código aberto da API REST do *Amazon S3* integrado ao projeto. Ele permite fornecer um *frontend* para o repositório de imagens de VMs usadas pelo *Nimbus*, funcionando como interface primária para transferir imagens dentro e fora da nuvem. Ao inicializar uma imagem em uma determinada nuvem *Nimbus*, essa imagem deve ser primeiro colocado no repositório *Cumulus* mesmo da nuvem. Um visão geral de sua arquitetura pode ser observada na Figura 6.

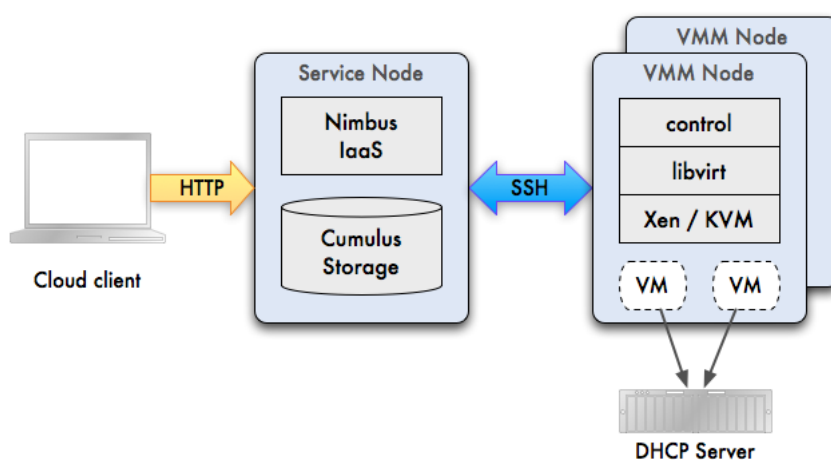


Figura 6: Visão geral da Arquitetura do Nimbus. (NIMBUS, 2012).

Dada sua natureza mais próxima de ambientes em grade - em especial aqueles direcionados a comunidade científica, o *Nimbus* foca em fornecer uma maior flexibilidade quanto a configuração e administração do ambiente, sendo por outro lado, constante e menos flexível quanto sua infraestrutura. No entanto, o projeto dá uma significativa atenção quanto a gestão da capacidade de alocação de recursos, permitindo diferentes direitos de uso, de acordo como perfil uso do usuário. Seguindo esse raciocínio, o suporte à API do EC2 e outras implementações baseadas em REST. Isso qualifica seu gerenciamento, oferecendo capacidade do *Nimbus* utilizar recursos federados de outras nuvens, como estratégia para lidar com sobrecargas e aumentar sua capacidade de alocação de recursos.

Combinando essas ferramentas e capacidades de diferentes maneiras, permite que os usuários possam rapidamente desenvolver soluções personalizadas para comunidade específicas. Por exemplo, uma comunidade pode querer usar *Nimbus* para configurar uma nuvem privada ou comunitária, enquanto outro pode querer se concentrar em aumentar recursos de uma nuvem já existentes com recursos da comunidade e várias nuvens públicas.

4.4 OpenStack

Dentre as implementações de *software* para Computação em Nuvem, o *OpenStack* (OPENSTACK, 2012) vem se destacando com uma opção promissora. O *OpenStack* é um conjunto de tecnologias capaz de controlar múltiplas infraestruturas de computação, armazenamento e recursos de rede, permitindo que administradores e pesquisadores possam fornecer recursos de computação *on-demand*, através do provisionamento e gerenciamento de grandes redes de máquinas virtuais. Seu objetivo é entregar uma infraestrutura em nuvem, para a implantação de nuvens públicas e privadas, com bases nas premissas de simplicidade e escalabilidade massiva, permitindo o gerenciamento dos recursos disponíveis, tanto em ambientes de infraestrutura pequena, a ambientes compostos por diversos centros de dados distribuídos geograficamente, propondo-se a ser um "sistema operacional em nuvem" de código aberto.

Tais recursos são acessíveis através de APIs e interfaces *Web* para desenvolvedores, administradores e usuários. Além disso, o *OpenStack* é projetado para escalar horizontalmente usando *hardware* padrão, permitindo assim redução de custos.

O *OpenStack* possui suporte a hipervisores como LXC (*Linux Containers*), Xen/XCP, Microsoft Hyper-V, QEMU, IBM PowerVM, e o mais largamente usado nos projetos com *OpenStack*, o KVM. Quanto ao armazenamento, o *OpenStack* é capaz de suportar três modos de armazenamento: de objetos, acessando-os através de uma API REST ; em bloco, disponibilizando volumes de armazenamento; ou em nível de sistema de arquivo, usando a interface de manipulação de arquivos fornecida pelo sistema operacional. Atualmente, o projeto possui suporte a tecnologias de armazenamento distribuído como: NFS, LVM, ZFS, GlusterFS, Ceph, MooseFS e outros.

A arquitetura do *OpenStack* é constituída por sete componentes que formam o núcleo do projeto. Eles são responsáveis por gerenciar diferentes mecanismos como rede, armazenamento, gerenciamento das imagens de VMs, gerenciamento de usuários e identidades e outros que estejam dentro do espectro de um ambiente dessa natureza. Esses componentes de núcleo podem ser assim descritos:

1. ***OpenStack Compute (Nova)***, responsável por fornecer e gerenciar conjuntos de máquinas virtuais distribuídas entre diversas máquinas físicas interligadas - independente de *hardware* ou mecanismo de virtualização. Ele busca criar um ambiente em nuvem altamente escalável e redundante, se ajustando de acordo com a demanda necessária. Por essas características, pode ser considerado o componente mais complexo e distribuído do pro-

jeto. É possível compará-lo ao serviço EC2 (*Elastic Compute Cloud*) da Amazon;

2. **OpenStack Object Storage (Swift)**, responsável por criar um espaço de armazenamento distribuído, redundante e escalável dentro da infraestrutura provida através do *OpenStack*, permitindo armazenar e recuperar arquivos ao longo dessa estrutura;
3. **OpenStack Imaging Service (Glance)**, responsável por fornecer um catálogo e um repositório de imagens de disco virtual - usadas normalmente pelo *Nova*, permitindo a descoberta, registro e serviços de entrega das mesmas;
4. **OpenStack Identity (Keystone)** gerencia a autenticação e autorização de todos os serviços pertencentes ao *OpenStack*, criando um ponto comum para a integração de políticas de segurança e gerenciamento de usuário dentro da infraestrutura em nuvem. Ele fornece serviços de catálogo, políticas, *token* de segurança e identidades, possuindo suporte ao uso de tecnologias como LDAP, SQL e outras que sejam capazes de armazenar esse tipo de informação;
5. **OpenStack Block Storage (Cinder)** - diferentemente do *Swift*, que manipula objetos - permite a manipulação de volumes de armazenamento, expondo um dispositivo de bloco para o usuário. Isso permite que seja possível anexar volumes as instâncias de máquinas virtuais em execução gerenciadas pelo *OpenStack Compute*;
6. **OpenStack Network (Quantum)** é responsável por permitir que os usuários que estejam usando os outros serviços do *OpenStack* (e.g *Nova*) possam criar suas próprias redes, através do fornecimento de um serviço de conectividade, onde é possível anexar interfaces de rede e interligá-los, criando assim uma rede virtual;
7. **OpenStack Dashboard (Horizon)** provê uma interface gráfica - baseada em Web - para os usuários acessarem os demais serviços do *OpenStack*. Através dela, os usuários podem realizar diversas ações, como iniciar a instância de uma máquina virtual, criar novas VMs, configurar rede e controle de acesso desse ambiente e etc.

No entanto, sua arquitetura orientada a serviços, permite o desenvolvimento de diversos outros componentes que se comunicam através da API fornecida. Além disso, a comunicação através de uma API é uma característica que influencia diretamente na dinâmica de funcionamento do *OpenStack*, pois permite

que cada componente possa fornecer e consumir serviços dos demais, de uma forma mais independente. Isso possibilita que mudanças e extensões nas características destes componentes possam ser feitas - desde que as definições estabelecidas pela API sejam respeitadas, contribuindo para natureza mais flexível e massivamente escalável do *OpenStack*.

De forma resumida, a relação entre os principais componentes do OpenStack pode ser descrita da seguinte forma:

- *OpenStack Compute* ("Nova") armazena e recupera discos virtuais ("imagens") e metadados associados no *Image* ("Glance") instanciando e executando máquinas virtuais para atender as demanda de computação;
- *OpenStack Network* ("Quantum") fornece uma rede virtual para o Nova;
- *OpenStack Block Storage* ("Cinder") fornece volumes de armazenamento para o Nova;
- *OpenStack Image* ("Glance") armazena os arquivos dos discos virtuais no *OpenStack Object Store* ("Swift");
- *OpenStack Identity* ("Keystone") fornece um serviço de autenticação para os demais componentes do *OpenStack*
- *OpenStack Dashboard* ("Horizon") fornece para esses, um *frontend web*.

Uma visão geral da arquitetura do *OpenStack* e do relacionamento dos seus principais componentes pode ser vista na figura 7.

Avaliar: <http://www.ibm.com/developerworks/br/cloud/library/cl-openstack-cloud/>

4.5 Considerações sobre o Capítulo

Nesse capítulo, foram apresentados os projetos que mais se destacam para implantação de ambientes em nuvem. Eles compartilham características comuns quanto sua proposta de fornecer um ambiente em nuvem redundante e escalável, diferindo-se nas estratégias, tecnologias e em alguns momentos, no foco do problema que se propõem a enfrentar.

Projetos mais maduros como o *Eucalyptus*, começam a caminhar com maior velocidade na direção de aumentar sua capacidade de integração com outras tecnologias em nuvem mais consolidadas no mercado, como as da Amazon, permitindo construir soluções em nuvem com características federadas. De certa forma, esse foco na compatibilidade de tecnologias consolidadas no mercado,

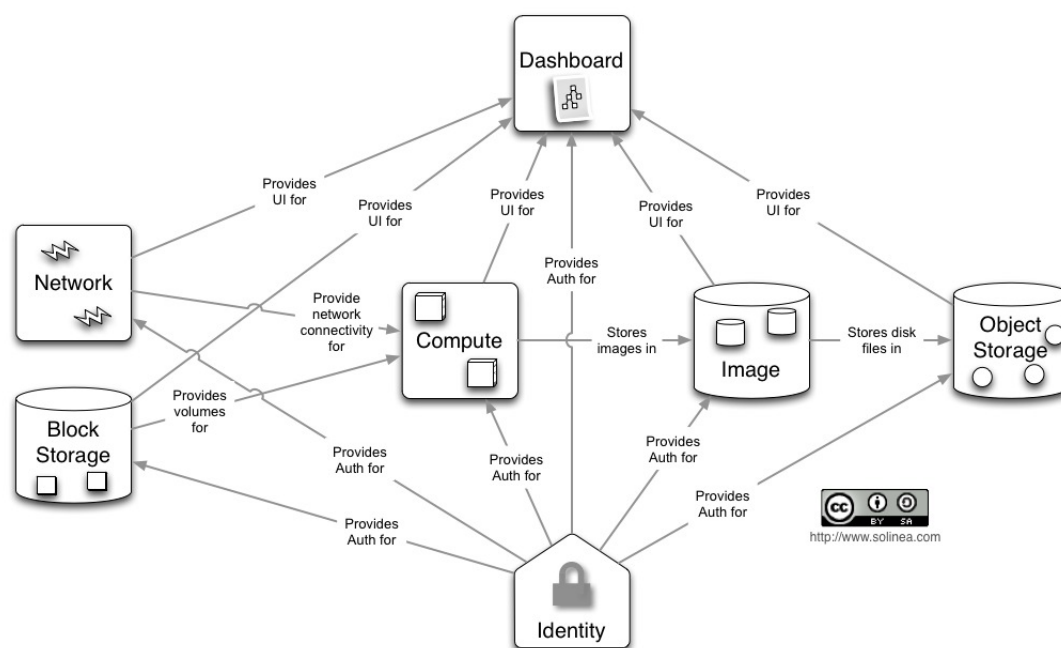


Figura 7: Visão geral da Arquitetura do *OpenStack*. (OPENSTACK, 2012).

colabora para que o *Eucalyptus* se fortaleça como solução na implantação de infraestruturas em nuvem.

Outros projetos como o *OpenNebula* e o *Nimbus*, se propõem a oferecer uma estrutura de nuvem através do uso de tecnologias já existentes, permitindo compor soluções em nuvem com diversos fins. Isso, por exemplo, colabora para que o *OpenNebula* possa ser usado tanto para a criação de ambientes em nuvem de uso corporativo/comercial, quanto para uso científico, onde é possível aproveitar todo o conhecimento e técnicas já amadurecidos provenientes da administração infraestruturas de servidores, apoiando-se nessa base para construir um ambiente em nuvem. Seguindo o mesmo raciocínio, o projeto *Nimbus* busca ampliar o uso da Computação em Nuvem com fins científicos, propondo integrar ambientes de computação em grade e em *cluster* com soluções tipicamente de nuvem, buscando potencializar as capacidades de ambas.

Por sua vez, projetos com *OpenStack* possuem uma abordagem mais ampla, e como sua própria apresentação do projeto demonstra, se propõem a fornecer um "sistema operacional em nuvem", caracterizado por sua flexibilidade na composição de ambiente e comunicação, escalabilidade e administração. Como forma de auxiliar na comparação entre esses diferentes projetos, é apresentada uma tabela que sintetiza suas principais características na Tabela 2.

No próximo capítulo, será realizado um fechamento do presente trabalho trazendo principais entendimentos surgidos com o estudo realizado. Também será feita uma breve análise dos projetos que se propõem a implantar uma ambiente

Projeto	OpenStack	OpenNebula	Nimbus	Eucalyptus
Foco	Científico Comercial	Científico Comercial	Científico	Científico Comercial
Linguagens	Python	Java/Ruby	Python/Java	C/Java
Tipos de Nuvens	Privadas Públicas Híbridas	Privadas Públicas Híbridas	Privadas	Privadas Públicas Híbridas
Hipervisores	XEN KVM VMWare LXC	XEN KVM VMWare	XEN KVM	XEN KVM VMWare
Armazenamento	NFS GlusterFS MooseFS Lustre	NFS GlusterFS Lustre LVM	<i>GridFTP</i> <i>Cumulus</i>	<i>Walrus</i>
Interoperabilidade	REST	XML-RPC REST OGF OCCI EC2 WS API	WSRF REST EC2 WS API	<i>EC2 WS API</i>
Licença	Apache v2	Apache v2	Apache v2	GNUv3/BSD

Tabela 2: Comparação das Plataformas para Implantação e Gerenciamento de Infraestruturas de Ambientes em Nuvem

em nuvem, destacando suas principais características.

5 CONSIDERAÇÕES FINAIS

Com a ampliação do uso e oferta dos serviços e produtos relacionados a tecnologia de informação, em especial aquele que possuem estreita ligação com a Internet, criou-se também a necessidade de reduzir custos e complexidade na operação e administração das estruturas que garantem seus funcionamento. Nesse cenário, a Computação em Nuvem mostra ter um papel fundamental na busca atender essas necessidades, promovendo o uso mais eficiente dos recursos computacionais, através do compartilhamento desses recursos e da concentração dos esforços na administração das infraestruturas relacionadas.

A Computação em Nuvem, por possuir em seu cerne diversas premissas e tecnologias provenientes da Computação em Grade e da Computação Utilitária, conseguiu moldar um modelo de computação distribuída que provê capacidade computacional na forma de um serviço elástico, escalável e entregues sob demanda, sendo esse serviço cobrado de acordo com seu uso. Nesse modelo, um recurso pode ser obtido imediatamente - sem intervenção humana; acessado de forma ubíqua, cuja seu fornecimento se ajusta conforme a demanda, sendo cobrado de acordo com essa. Além disso, nesse modelo, os recursos disponíveis tendem a estar agrupados, sendo compartilhados em um modelo de *multi-tenancy*. Essas características passam aos usuários a ilusão de que esse modelo pode fornecer recursos de computação infinitos.

Dessa forma, ela tem avançado em estabelecer um modelo de negócio que é capaz de ofertar um conjunto de soluções de T.I, que mesclam infraestrutura computacional, plataforma de desenvolvimento e *software* aplicativo, buscando atender as necessidades tecnológicas de seus usuário em vários níveis de abstração. Isso tem atraído a atenção de diversas áreas - corporativas ou acadêmicas -, que enxergam a Computação em Nuvem com uma ferramenta capaz de colaborar com suas atividades.

No entanto, a Computação enfrenta diversos desafios decorrentes de sua crescente adoção. Dentre os presentes, neste trabalhos foram destacados principalmente os relacionados ao uso eficiente dos recursos, como os problemas re-

lacionados provisionamento automático de recursos, a migração e consolidação de máquinas virtuais. Esses desafios estão intrinsecamente ligados a natureza dinâmica das cargas de trabalhos ao qual um ambiente em nuvem está exposto. A capacidade de aquisição e liberação de recursos, a migração de máquinas virtuais de forma consistente como resposta à sobrecargas nos servidores físicos, e a otimização no uso de recurso através de um processo de consolidação de máquinas virtuais que promova o uso eficiente desses recursos, encontram diversas barreiras em ambientes dinâmicos como os em nuvem, onde as demandas por recursos oscilam drasticamente em um curto período de tempo. Da mesma forma, o gerenciamento de energia de forma eficiente nesses ambiente sofre com os mesmos desafios. E diante de um cenário de crescimento tanto da Computação em Nuvem, quanto das preocupações com o impacto ambiental causados pelas tecnologias de informação, buscar melhorias na forma em que esses ambiente consomem energia ganha um espaço de destaque nas pesquisas relacionadas com a área.

Da mesma forma, esse crescente interesse na Computação em Nuvem, tem motivado o desenvolvimento de diversas ferramentas que possam colaborar com a adoção e uso da Computação em Nuvem. Atualmente, existem várias frentes de trabalho - suportadas por empresas, comunidades de código aberto, ou ambas - que tem focado no desenvolvimento de um número significativo de projetos para Computação em Nuvem, principalmente aqueles direcionados a implantação de sua infraestrutura. Esses projetos buscam, cada um de sua forma, explorar o potencial uso da Computação em Nuvem em atender as necessidades de diversas áreas.

Assim, a Computação em Nuvem, com suas características, desafios e potencialidades, consegue constituir campo de pesquisa científica que possui um significativo espaço a ser explorado, e uma grande motivação para o desenvolvimento de trabalhos relacionada a sua área.

REFERÊNCIAS

ADDIS, B.; ARDAGNA, D.; PANICUCCI, B.; ZHANG, L. Autonomic Management of Cloud Service Centers with Availability Guarantees. In: CLOUD COMPUTING (CLOUD), 2010 IEEE 3RD INTERNATIONAL CONFERENCE ON, 2010. **Anais...** [S.l.: s.n.], 2010. p.220 –227.

AMAZON. **Amazon Elastic Compute Cloud (Amazon EC2)**. <http://aws.amazon.com/pt/ec2/>.

AMAZON. **Amazon Web Services**.

ARMBRUST, M.; FOX, A.; GRIFFITH, R.; JOSEPH, A. D.; KATZ, R.; KONWINSKI, A.; LEE, G.; PATTERSON, D.; RABKIN, A.; STOICA, I. et al. A view of cloud computing. **Communications of the ACM**, [S.l.], v.53, n.4, p.50–58, 2010.

BELOGLAZOV, A.; ABAWAJY, J.; BUYYA, R. Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing. **Future Generation Computer Systems**, [S.l.], v.28, n.5, p.755 – 768, 2012.

BUYYA, R.; YEO, C. S.; VENUGOPAL, S.; BROBERG, J.; BRANDIC, I. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. **Future Gener. Comput. Syst.**, Amsterdam, The Netherlands, The Netherlands, v.25, n.6, p.599–616, June 2009.

BUYYA, R.; YEO, C.; VENUGOPAL, S.; BROBERG, J.; BRANDIC, I. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. **Future Generation computer systems**, [S.l.], v.25, n.6, p.599–616, 2009.

DROPBOX. **Página Web**. <https://www.dropbox.com/>.

EUCALYPTUS. **Página do projeto**. <http://www.eucalyptus.com>.

FACEBOOK. **Página Web**. <http://www.facebook.com/>.

FOSTER, I.; KESSELMAN, C.; TUECKE, S. The Anatomy of the Grid - Enabling Scalable Virtual Organizations. **International Journal of Supercomputer Applications**, [S.l.], v.15, p.2001, 2001.

FOSTER, I.; ZHAO, Y.; RAICU, I.; LU, S. Cloud Computing and Grid Computing 360-Degree Compared. In: GRID COMPUTING ENVIRONMENTS WORKSHOP, 2008. GCE '08, 2008. **Anais...** [S.l.: s.n.], 2008. p.1 –10.

FOX, A.; GRIFFITH, R. et al. Above the clouds: A Berkeley view of cloud computing. **Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Tech. Rep. UCB/EECS**, [S.l.], v.28, 2009.

GMAIL. **Página Web**. <http://mail.google.com/>.

GOOGLE. **Google App Engine**. <https://developers.google.com/appengine/>.

GRAUBNER, P.; SCHMIDT, M.; FREISLEBEN, B. Energy-Efficient Management of Virtual Machines in Eucalyptus. In: CLOUD COMPUTING (CLOUD), 2011 IEEE INTERNATIONAL CONFERENCE ON, 2011. **Anais...** [S.l.: s.n.], 2011. p.243 –250.

KOOMEY, J. Growth in data center electricity use 2005 to 2010. **Oakland, CA: Analytics Press. August**, [S.l.], v.1, p.2010, 2011.

LASZEWSKI, G. von; DIAZ, J.; WANG, F.; FOX, G. Comparison of Multiple Cloud Frameworks. In: CLOUD COMPUTING (CLOUD), 2012 IEEE 5TH INTERNATIONAL CONFERENCE ON, 2012. **Anais...** [S.l.: s.n.], 2012. p.734 –741.

MELL, P.; GRANCE, T. The NIST definition of cloud computing (draft). **NIST special publication**, [S.l.], v.800, p.145, 2011.

MORENO-VOZMEDIANO, R.; MONTERO, R.; LLORENTE, I. IaaS Cloud Architecture: From Virtualized Datacenters to Federated Cloud Infrastructures. **Computer**, [S.l.], v.45, n.12, p.65 –72, dec. 2012.

MORENO-VOZMEDIANO, R.; MONTERO, R. S.; LLORENTE, I. M. IaaS Cloud Architecture: From Virtualized Datacenters to Federated Cloud Infrastructures. **Computer**, Los Alamitos, CA, USA, v.45, n.12, p.65–72, 2012.

NIMBUS. **Página do projeto**. <http://www.nimbusproject.org>.

NURMI, D.; WOLSKI, R.; GRZEGORCZYK, C.; OBERTELLI, G.; SOMAN, S.; YOUSEFF, L.; ZAGORODNOV, D. The eucalyptus open-source cloud-computing system. In: CLUSTER COMPUTING AND THE GRID, 2009. CCGRID'09. 9TH

IEEE/ACM INTERNATIONAL SYMPOSIUM ON, 2009. **Anais...** [S.l.: s.n.], 2009. p.124–131.

OPENNEBULA. **Página do projeto**. <http://www.opennebula.org>.

OPENSTACK. **Página do projeto**. <http://www.openstack.org>.

PARKHILL, D. F. **The challenge of the computer utility**. [S.l.]: Addison-Wesley Publishing Company Reading, 1966. v.2.

SALESFORCE. **Página Web**. <http://www.salesforce.com>.

SEMPOLINSKI, P.; THAIN, D. A Comparison and Critique of Eucalyptus, Open-Nebula and Nimbus. In: CLOUD COMPUTING TECHNOLOGY AND SCIENCE (CLOUDCOM), 2010 IEEE SECOND INTERNATIONAL CONFERENCE ON, 2010. **Anais...** [S.l.: s.n.], 2010. p.417 –426.

THEGREENGRID. **Página do projeto**. <http://www.thegreengrid.org>.

WIND, S. Open source cloud computing management platforms: Introduction, comparison, and recommendations for implementation. In: OPEN SYSTEMS (ICOS), 2011 IEEE CONFERENCE ON, 2011. **Anais...** [S.l.: s.n.], 2011. p.175 –179.

ZHANG, Q.; CHENG, L.; BOUTABA, R. Cloud computing: state-of-the-art and research challenges. **Journal of Internet Services and Applications**, [S.l.], v.1, p.7–18, 2010.