

Soluções para criar a sua própria nuvem

Maicon de Vargas Pereira

¹Faculdade de Tecnologia Senac Pelotas
Rua Gonçalves Chaves, 602 – Pelotas – RS – Brasil – Caixa Postal – 96.015-560
Curso Superior de Tecnologia em Redes de Computadores

maicon.tec.inf@hotmail.com

Resumo. *Este artigo tem como objetivo demonstrar soluções de um servidor de arquivos em nuvem analisando seus recursos, desempenho e suas vantagens e desvantagens.*

Abstract. *This article aims to demonstrate solutions to a file server in the cloud by analyzing its features, performance, and their advantages and disadvantages.*

1. Introdução

Com o avanço da sociedade humana moderna, serviços básicos e essenciais são quase todos entregues de uma forma completamente transparente. Serviços de utilidade pública como água, gás, eletricidade e telefone tornaram-se fundamentais a vida diária e são explorados através de um modelo de pagamento baseado no uso [Vecchiola et al. 2009]. As infraestruturas existentes permitem entregar tais serviços em qualquer lugar e a qualquer hora, de forma que se pode simplesmente acender a luz, abrir a torneira ou usar o fogão. O uso desses serviços é, então, cobrado de acordo com as diferentes políticas para o usuário final. Recentemente, a mesma ideia de utilidade tem sido aplicada no contexto da informática e uma mudança consistente neste sentido tem sido feita com a disseminação de *Cloud Computing* ou Computação em Nuvem. A Computação em Nuvem é uma tendência recente de tecnologia cujo objetivo é proporcionar serviços de Tecnologia da Informação (TI). Tendências anteriores à computação em nuvem foram limitadas a uma determinada classe de usuários ou focadas em tornar disponível uma demanda específica de recursos de TI, principalmente de informática [Buyya et al. 2009b]. Computação em nuvem pretende ser global e prover serviços para as massas que vão desde o usuário final que hospeda seus documentos pessoais na Internet até empresas que terceirizarão toda a parte de TI para outras empresas. Nunca uma abordagem para a utilização real foi tão global e completa: não apenas recursos de computação e armazenamento são entregues sob demanda, mas toda a pilha de computação pode ser aproveitada na nuvem. Os usuários estão movendo seus dados e aplicações para a nuvem e assim acessá-los de forma simples e de qualquer local. Isso é novamente um caso de utilização de processamento centralizado. Um cenário semelhante ocorreu há aproximadamente 50 anos: um servidor de tempo compartilhado acessado por vários usuários [mc7 2014]. Contudo, nas últimas décadas, quando os computadores pessoais surgiram, os dados e as aplicações começaram a serem utilizados localmente. A figura 1 mostra uma visão geral de uma nuvem.

2. Conceitos da Computação em Nuvem

A Computação em Nuvem está se tornando uma das palavras chaves da indústria de TI. A nuvem é uma metáfora para a Internet ou infraestrutura de comunicação entre os compo-



Figura 1. Visão geral de uma nuvem [itaraujo 2014]

nentes arquiteturais, baseada em uma abstração que oculta a complexidade de infraestrutura. Cada parte desta infraestrutura é provida como um serviço e, estes serviços são normalmente alocados em *data centers*, utilizando *hardware* compartilhado para computação e armazenamento. Para utilizarem os serviços, os usuários necessitam apenas ter em suas máquinas um sistema operacional, um navegador e acesso a Internet. Todos os recursos e processamentos computacionais estão disponíveis na Internet. Assim, as máquinas dos usuários não necessitam ter altos recursos computacionais, diminuindo assim o custo na aquisição de máquinas por parte destes usuários. Todo hardware pode ser utilizado para realizar alguma tarefa que seja adequada ao seu poder de processamento. Novos recursos de hardware podem ser adicionados a fim de aumentar o poder de processamento e cooperar com os recursos existentes. O modelo de Computação em Nuvem foi desenvolvido com o objetivo de fornecer serviços de fácil acesso e de baixo custo e garantir características tais como disponibilidade e escalabilidade. Este modelo visa fornecer, basicamente, três benefícios. O primeiro benefício é reduzir o custo na aquisição e composição de toda infraestrutura requerida para atender as necessidades das empresas, podendo essa infraestrutura ser composta sob demanda e com recursos heterogêneos e de menor custo. O segundo é a flexibilidade que esse modelo oferece no que diz respeito à adição e troca de recursos computacionais, podendo assim, escalar tanto em nível de recursos de *hardware* quanto *software* para atender as necessidades das empresas e usuários. Já [Armbrust et al. 2009] propõem a seguinte definição: 'A Computação em Nuvem é um conjunto de serviços de rede ativados, proporcionando escalabilidade, qualidade de serviço, infraestrutura barata de computação sob demanda e que pode ser acessada de uma forma simples e pervasivo'. Um levantamento com algumas propostas de definição para computação em nuvem pode ser encontrado em [Vaquero et al. 2009].

2.1. Arquitetura de uma Nuvem

A arquitetura é baseada em camadas, sendo que cada uma destas trata de uma particularidade na disponibilização de recursos para as aplicações [Buyya et al. 2009b]. Uma camada é uma divisão lógica de componentes de *hardware* e *software*. Alguns destes recursos computacionais podem ser agrupados e organizados para realizar uma determinada tarefa do sistema como um todo. Cada camada pode ter seu gerenciamento ou monitoramento de forma independente das outras camadas, melhorando a flexibilidade, reusabilidade e escalabilidade no tocante a substituição ou adição de recursos computacionais sem

afetar as outras camadas.



Figura 2. arquitetura de uma nuvem [Teleco 2014]

A figura 2 ilustra as camadas referentes à arquitetura da computação em nuvem e suas respectivas associações. É possível distinguir na figura quatro diferentes camadas que se deslocam progressivamente da perspectiva do sistema para a do usuário final.

[Teleco 2014] A camada de mais baixo nível é a de infraestrutura física, que pode conter dispositivos de armazenamento, *clusters*, *desktops* e outros recursos de *hardware*. Esta camada fornece flexibilidade e facilidade de agregação de novos recursos à medida que se tornem necessários. Uma camada de *middleware* é responsável por gerenciar a infraestrutura física e tem por objetivos prover um ambiente de execução apropriado para as aplicações e explorar de maneira eficaz os recursos físicos. Esta camada pode ser dividida em duas subcamadas: uma responsável por garantir o isolamento de processos e aplicações, qualidade de serviço, podendo utilizar tecnologias de virtualização e outra camada responsável por prover um conjunto de serviços que auxiliam os provedores de serviços comerciais e profissionais para os usuários finais? dentre os serviços dessa camada podem ser encontrados negociação de Qualidade de Serviço (*QoS*), serviços de cobrança, gerenciamento de requisições, entre outros. No nível acima da camada de *middleware*, encontra-se a camada responsável por prover suporte para a construção de aplicações e que contém ferramentas ou ambientes de desenvolvimento. Estes ambientes possuem interfaces *Web 2.0*, *mashups*, componentes, recursos de programação concorrente e distribuída, suporte a *workflows*, bibliotecas de programação e linguagens de programação. Esta camada de desenvolvimento não é utilizada pelos usuários finais, e sim, pelos usuários mais experientes, aqueles que desenvolvem as soluções para computação em nuvem. Esta camada *middleware* no nível de usuário constitui o ponto de acesso das aplicações à infraestrutura da nuvem. Por fim, encontra-se a camada das aplicações de computação em nuvem. Esta camada é de interesse do usuário, pois é por meio dela que eles utilizam os aplicativos. As camadas abaixo desta são responsáveis

pelas características de escalabilidade, disponibilidade, ilusão de recursos infinitos e alto desempenho.

2.2. Segurança

A Computação em Nuvem é um modelo que utiliza a Internet para disponibilizar seus serviços. Assim, deve-se ter formas para impedir o acesso não autorizado a informações e que os dados sensíveis permaneçam privados, pois estes serão processados fora da empresa. Questões de segurança devem ser consideradas para prover a autenticidade, confidencialidade e integridade. No que diz respeito à confiabilidade e responsabilidade, o provedor deve fornecer recursos confiáveis, especialmente se a computação a ser realizada é crítica e existindo uma clara delimitação de responsabilidade.

2.3. Gerenciamento de Dados

O gerenciamento de dados é considerado um ponto crítico no contexto de computação em nuvem. Os Sistemas de Gerenciamento de Banco de Dados (*SGBDs*) relacionais não possuem escalabilidade quando milhares de sites são considerados [Wei et al. 2009]. Assim, aspectos de armazenamento de dados, processamento de consultas e controle transacional tem sido flexibilizados por algumas abordagens para garantir a escalabilidade, mas ainda não existem soluções que combinem estes aspectos de forma a melhorar o desempenho sem comprometer a consistência dos dados [Abadi 2009].

2.4. Autonomia

A Computação em Nuvem é um sistema autônomo gerenciado de forma transparente para os usuários. *Hardware* e *software* dentro de nuvens podem ser automaticamente reconfigurados, orquestrados e estas modificações são apresentadas ao usuário como uma imagem única. Essa autonomia é importante, pois reduz o custo de equipe de monitoramento do sistema tanto no âmbito centralizado quanto distribuído [Birman et al.2009].

2.5. Disponibilidade de Serviços

A disponibilidade de serviços permite aos usuários acessar e utilizar a nuvem onde e quando desejarem. Como se trata da Internet podem ocorrer atrasos e sistemas indisponíveis. Os ambientes de computação em nuvem devem prover alta disponibilidade.

2.6. Escalabilidade e Desempenho

A escalabilidade foi uma das características fundamentais que conduziram ao surgimento da Computação em Nuvem. As nuvens de serviços e as plataformas oferecidas podem ser dimensionadas por vários fatores, tais como localizações geográficas, desempenho ou configurações. Apesar das limitações de rede e segurança, as soluções de computação em nuvem devem fornecer elevado desempenho, além de ser flexível para se adaptar diante de uma determinada quantidade de requisições. Como os ambientes que possuem acesso público, é imprevisível e variável a quantidade de requisições realizadas, tornando mais complexo fazer estimativas e garantias de QoS.

2.7. Aplicações de Computação em Nuvem

Diversos tipos de aplicações baseadas em computação em nuvem estão disponíveis. Serviços de *webmail*, sites, *softwares* como serviços em geral são alguns exemplos. Outras aplicações estão relacionadas à bioinformática [Deelman et al. 2008] e processamento de imagens. Neste artigo será apresentada uma aplicação executada em nuvem para compartilhamento de arquivos.

3. Tipos de soluções para armazenamento de arquivos em nuvens privadas

Nesse artigo serão apresentados alguns *softwares* para armazenamento de arquivos na nuvem semelhante a solução *Dropbox* [Dropbox 2014], com o diferencial da utilização estar focada em nuvens privadas, permitindo, assim, uma maior segurança e controle do acesso a dados mais sensíveis das organizações.

3.1. Owncloud

A ferramenta *Owncloud* [Owncloud 2014] é uma plataforma *opensource* criada pelos responsáveis pelo *K Desktop Environment (KDE)* que permite ter a própria nuvem e, sendo assim, ter dados sob controle. É uma solução que roda em várias distribuições Linux. A instalação é simples e rápida, não precisando de permissões especiais. Sem dúvida, uma boa solução para implantar dentro da Intranet das empresas, oferecendo assim um lugar onde todos os funcionários possam compartilhar e produzir sem necessidade de depender de serviços de terceiros. O acesso aos arquivos pode ser feito via *browser* ou utilizando o protocolo *webdav* [WebDAV 2014].

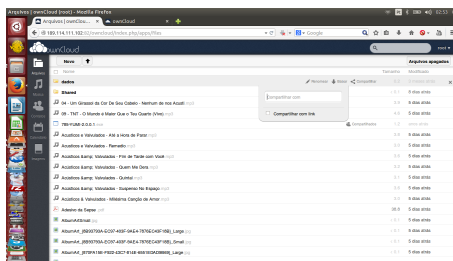


Figura 3. OwnCloud

3.2. iFolder

O *iFolder* [Novell 2014] Inicialmente desenvolvido pela *Novell* e que faz parte do produto *Novell Open Enterprise Server*, teve o seu código aberto e agora está disponível de forma gratuita para livre utilização. Ele usa o *Mono*, que é uma implementação em software livre do *.Net Framework* da *Microsoft*, e há pacotes prontos do *software* servidor para *Suse* e clientes para *Windows*, *Suse* e *MacOS X*. O *iFolder* é uma solução simples de armazenamento que pode aumentar a produtividade permitindo que o usuário faça *backups*, acesse e gerencie seus arquivos pessoais de qualquer lugar e a qualquer hora. Uma vez instalado o *iFolder*, simplesmente os arquivos serão salvos localmente como sempre e o *iFolder* automaticamente atualiza os arquivos em um servidor da rede, deixando os dados sempre disponíveis em outras máquinas utilizadas. O funcionamento é semelhante a serviços como *Dropbox*, com a diferença que qualquer diretório local pode ser sincronizado com o servidor, de forma segura usando criptografia. O acesso pode também ser feito via *browser*, possibilitando assim a utilização de sistemas operacionais não suportados.

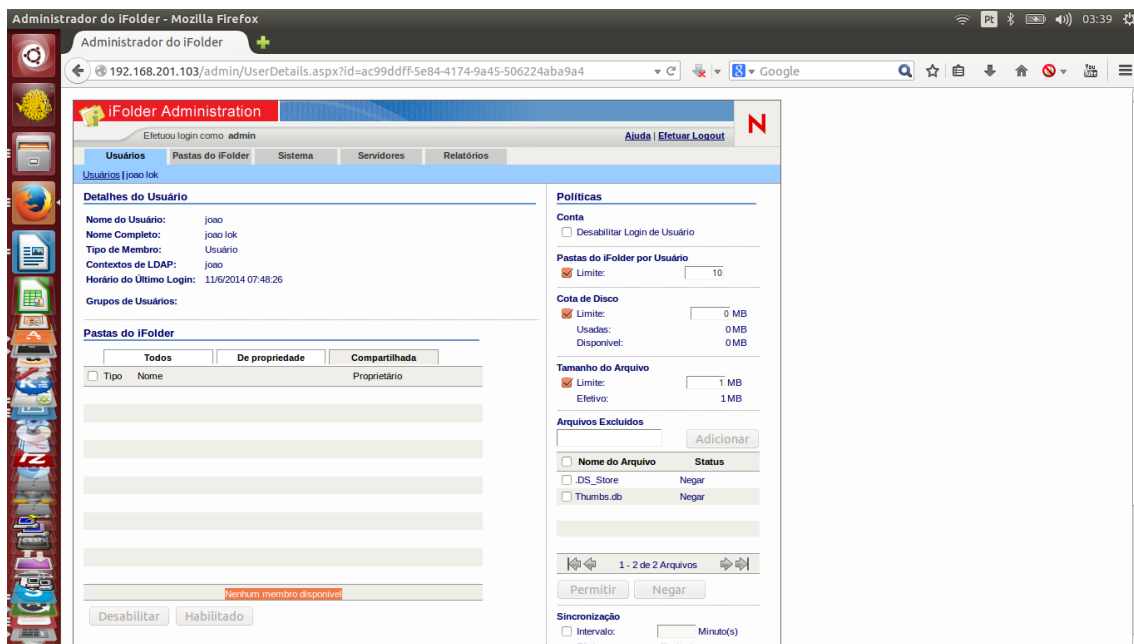


Figura 4. ifolder

3.3. SparkleShare

O SparkleShare [SparkleShare 2014] é uma ferramenta de colaboração e compartilhamento de arquivos que permite manter, guardar dados e realizar a sincronização automática destes dados em um repositório *Git* [Git 2014] privado ou até mesmo em servidores públicos como o Github [Git 2014]. Atualmente é possível utilizar clientes *Linux*, *Mac OSX* e *Android* como clientes. Há projetos para suporte à *Windows* e à *IOS*.

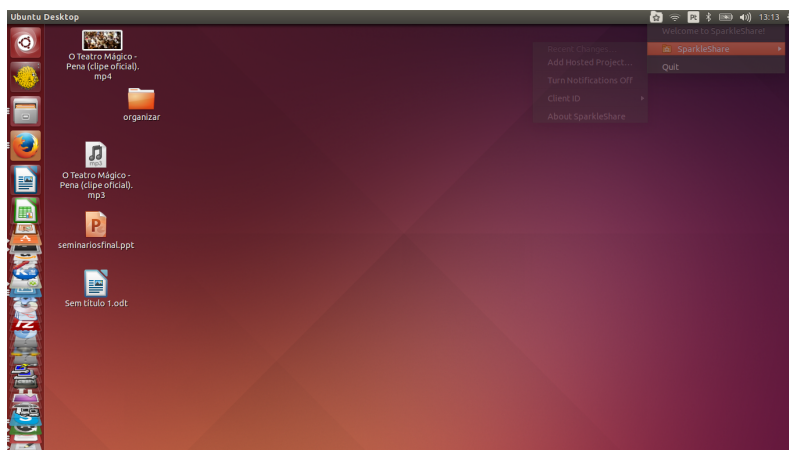


Figura 5. sparkleshare

4. Cenário atual

O cenário de testes consiste em um servidor *Ubuntu Server* e *hosts* cliente com *Windows XP*, onde serão realizados os teste, tais como desempenho. Serão analisados os tempos para realização das tarefas, visto que todos os *hosts* servidores e clientes possuem a mesma configuração, sendo assim, todos *hosts* clientes são virtualizados.

Tabela 1. Configuração de *hardware* das máquinas servidores.

Softwares	CPU	Memória RAM	HD	Sistema Operacional
Owncloud	1,73 GHz	1525 MB	80 GB	Ubuntu Server 12.04
iFolder	1,73 GHz	1525 MB	80 GB	Ubuntu Server 12.04
SpakleShare	1,73 GHz	1525 MB	80 GB	Ubuntu 12.04

Tabela 2. Configuração de *hardware* das máquinas virtuais.

Softwares	CPU	Memória RAM	HD	Sistema Operacional
OwncloudClient	1,73 GHz	512 MB	20 GB	XP
iFolderClient	1,73 GHz	512 MB	20 GB	XP
SparkleShareClient	1,73 GHz	512 MB	20 GB	Ubuntu

5. Testes realizados

O cenário utilizado para os testes foi um servidor de arquivos em nuvem e oito VMs clientes, conforme mostra a figura 6, foi realizado desempenho de consumo de banda para cada um dos casos, que onde obteve o seguinte resultado que esta demonstrado na figura 7. As soluções *iFolder* e *Owncloud* se comportaram muito bem em seus desempenho por outro lado a ferramenta *SparkleShare* teve alguns desafios uma, foi seu não funcionamento no *Windows XP*.

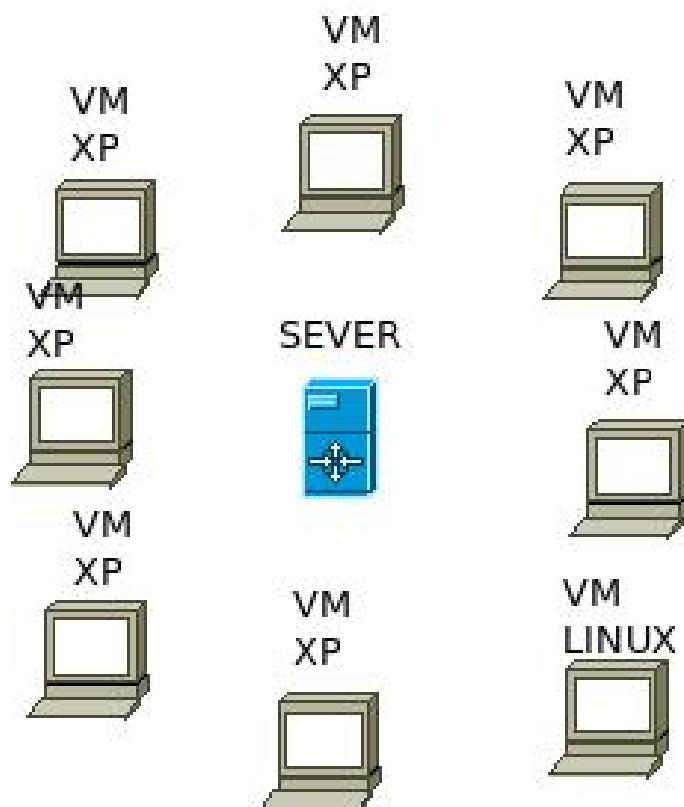
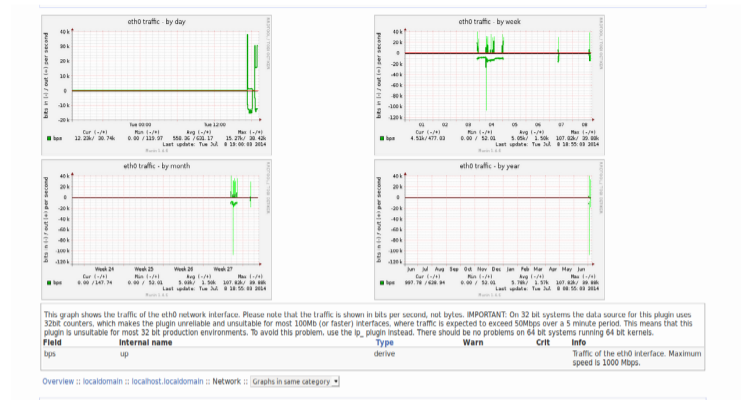


Figura 6. cenário de testes

Owncloud



iFolder

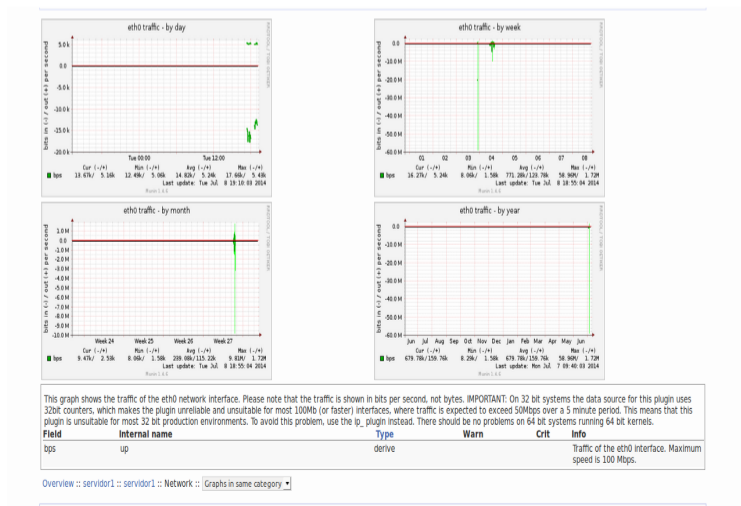


Figura 7. desempenho de consumo de banda

6. Conclusões

Este projeto teve como base um estudo de caso comparativo entre três soluções de servidores de arquivos em nuvem. As empresas podem prestar serviços diretamente aos usuários através da Internet de acordo com as necessidades destes usuários. O servidor certamente tem uma grande funcionalidade e papel em um cenário em uma empresa, por dois simples motivos, a integridade da informação do usuário e custo benefício. Foi criado um cenário real contendo um servidor com sistema Linux com a distribuição *ubuntu*, fazendo todas as configurações necessárias para o funcionamento das três soluções assim constado no artigo e diversas configurações e testes para realizar a melhor solução. Com tudo as três soluções são excelentes. O SparkShare aconteceu algumas falhas por ser uma ferramenta em processo de desenvolvimento mas foi resolvido e as outras duas soluções ocorreram perfeitamente.

Referências

Abadi et al. 2009. Data management in the cloud: Limitations and opportunities. Disponível em: <<http://cs-www.cs.yale.edu/homes/dna/talks/abadinedbday09.pdf>>. Acesso em: abril 2014.

Armbrust et al. 2009. Above the clouds: A berkeley view of cloud computing. Disponível em: <<http://d1smfj0g31qzek.cloudfront.net/abovetheclouds.pdf>>. Acesso em: abril 2014.

Buyya et al. 2009a. Modeling and simulation of scalable cloud computing environments and the cloudsim toolkit: Challenges and opportunities. Disponível em: <<http://arxiv.org/pdf/0907.4878.pdf>>. Acesso em: abril 2014.

Buyya et al. 2009b. Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167739X08001957>>. Acesso em: abril 2014.

Birman et al. 2009. Toward a cloud computing research agenda. Disponível em: <<http://www.cs.cornell.edu/projects/quicksilver/publicpdfs/sigact2.pdf>>. Acesso em: abril 2014.

Deelman et al. 2008. The cost of doing science on the cloud: the montage example. Disponível em: <<http://irsa.ipac.caltech.edu/Montage/publications/deelmansc08corrected.pdf>>. Acesso em: abril 2014.

Dropbox 2014. Dropbox. Disponível em: <<http://www.Dropbox.com.br>>. Acesso em: abril 2014.

Git 2014. Git e Github conceitos bsicos. Disponível em: <<http://pt.slideshare.net/HenriqueSilvestre1/git-e-github-conceitos-bsicos>>. Acesso em: abril 2014.

Ifolder. Ifolder. Disponível em: <<http://www.novell.com/products/openenterpriseserver/features/online-file-storage-ifolder.html/>>. Acesso em: abril 2014.

Itaraujo 2014. computacao em nuvem. Disponível em: <<http://www.itaraujo.com.br/computacao-em-nuvem/>>. Acesso em: abril 2014.

Owncloud. Owncloud. Disponível em: <<http://owncloud.org/>>. Acesso em: abril 2014.

Sparkleshare. Sparkleshare. Disponível em: <<http://sparkleshare.org/>>. Acesso em: abril 2014.

Teleco 2014. Teleco. Disponível em: <<http://www.teleco.com.br/tutoriais/tutorialservnuvopers1/pagina2.as>>. Acesso em: abril 2014.

Vecchiola et al. 2009. Aneka: A software platform for .net-based cloud computing. Disponível em: <<http://www.cloudbus.org/reports/AnekaCloudPlatform2009.pdf>>. Acesso em: abril 2014.

Vaquero et al. 2009. A break in the clouds: towards a cloud definition. Disponível em: <<http://www.sigcomm.org/sites/default/files/ccr/papers/2009/January/1496091-1496100.pdf>>. Acesso em: abril 2014.

Vivaolinux. Vivaolinux. Disponível em: <<http://www.vivaolinux.com.br/>> Acesso em :
abril 2014.

Wei et al. 2009. Scalable transactions for web applications in the cloud. Disponível em:
<<http://www.researchgate.net/publication/220767879ScalableTransactionsforWebApplicationsintheClo>>
Acesso em: abril 2014.