**UNIVERSIDADE VEIGA DE ALMEIDA – UVA**

**BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**ARQUITETURAS ESCALÁVEIS E TOLERANTES À FALHAS**

**Thiago Soares da Cruz**

**RIO DE JANEIRO**

**2016**

**UNIVERSIDADE VEIGA DE ALMEIDA - UVA**

**THIAGO SOARES DA CRUZ**

Monografia apresentada ao curso de Ciência da Computação da Universidade Veiga de Almeida, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Carlos Lemos

**ARQUITETURAS ESCALÁVEIS E TOLERANTES À FALHAS**

**RIO DE JANEIRO**

**2016**

image003

**UNIVERSIDADE VEIGA DE ALMEIDA - UVA**

**BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO**

**THIAGO SOARES DA CRUZ**

**ARQUITETURAS ESCALÁVEIS E TOLERANTES À FALHAS**

Monografia apresentada como requisito parcial à conclusão do curso em Bacharel em Ciência da Computação.

APROVADA EM:

CONCEITO: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

BANCA EXAMINADORA:

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**PROF.**

**ORIENTADOR**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**PROF.**

**CO-ORIENTADOR**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**PROF.**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**PROF .**

**Coordenação de Ciência da Computação**

Rio de Janeiro, 10 de junho de 2017

*Dedico este trabalho à minha família, em especial a minha mãe, que sempre acreditou no meu potencial, me impulsionando sempre a ir em frente, fazendo desta caminhada, um grande aprendizado e tornando-a possível.*

*E todos aqueles que sempre estiveram comigo diante dos problemas e das adversidades me incentivando sempre a levantar e a lutar, para concluir mais essa etapa da minha caminhada profissional e pessoal.*

“Citação”

Fonte: Fonte da Citação

**RESUMO**

Palavras-Chave:

**ABSTRACT**

**LISTA DE ILUSTRAÇÕES**

Figure 1: Clound 15

Figure 2: Joseph Carl 16

Figure 3: John McCarthy 16

Figure 4: Ramnath Chellappa 17

Figure 5: Modelos de Implementação 18

Figure 6: Nuvem pública 19

Figure 7: Nuvem Privada 20

Figure 8: Nuvem Comunitária 21

Figure 9: Modelos de Implementação 22

Figure 10: Modelos de Serviço 25

**LISTA DE ABREVIATURAS E SIGLAS**

Paas – Plataforma as a service – Plataforma como serviço

**SUMÁRIO**

[INTRODUÇÃO 12](#_Toc467757049)

[2 COMPUTAÇÃO EM NUVEM 15](#_Toc467757050)

[2.1 História 16](#_Toc467757051)

[2.2 Modelos de implantação 18](#_Toc467757052)

[2.2.1 Nuvem Pública 19](#_Toc467757053)

[2.2.2 Nuvem Privada 20](#_Toc467757054)

[2.2.3 Nuvem Comunitária 21](#_Toc467757055)

[2.2.3 Nuvem Híbrida 22](#_Toc467757056)

[2.3 Princípios da computação em nuvem 23](#_Toc467757057)

[2.4 Modelos de Serviços 25](#_Toc467757058)

[3 Paas 26](#_Toc467757059)

[3.1 PaaS: Plataforma como serviço 27](#_Toc467757060)

[3.2 Programação descentralizada: GRID e Cluster 28](#_Toc467757061)

[3.3 Programação em micro-serviços 29](#_Toc467757062)

[3.4 Docker 30](#_Toc467757063)

[3.4.1 Conceito 31](#_Toc467757064)

[4 SOFTWARES DE DEPLOY 32](#_Toc467757065)

[4.1 Tsuru 33](#_Toc467757066)

[4.1.1 Conceito 34](#_Toc467757067)

[4.1.2 Implementação em Cluster próprio 35](#_Toc467757068)

[4.2 Vagran 36](#_Toc467757069)

[3.2.1 Conceito 37](#_Toc467757070)

[4.2.2 Implementação em Cluster próprio 38](#_Toc467757071)

[4.3 Kubernet 39](#_Toc467757072)

[4.3.1 Conceito 40](#_Toc467757073)

[4.3.2 Implementação em Cluster próprio 41](#_Toc467757074)

[5 SISTEMAS ESCALÁVEIS 42](#_Toc467757075)

[5.1 Sistemas Operacionais escaláveis 43](#_Toc467757076)

[5.2 Máquinas Virtuais 44](#_Toc467757077)

[5.3 Comparações 45](#_Toc467757078)

[6 COMPARAÇÃO ENTRE PLATAFORMAS 46](#_Toc467757079)

[6.1 Plataformas 47](#_Toc467757080)

[6.2 Tolerâncias à falhas 48](#_Toc467757081)

[6.2 Segurança das plataformas 49](#_Toc467757082)

[7 PROGRAMAÇÃO EM MICROSERVIÇOS 50](#_Toc467757083)

[7.1 Boas práticas de programação distribuída (Doze fatores) 51](#_Toc467757084)

[7.2 Empregabilidade da técnica 52](#_Toc467757085)

[7.3 Casos de Sucesso 53](#_Toc467757086)

[8 ESTUDO DE CASO 54](#_Toc467757087)

[8.1 Infraestrutura 55](#_Toc467757088)

[8.2 Métricas de Escalabilidade 56](#_Toc467757089)

[9 CONCLUSÃO 57](#_Toc467757090)

[10 Referências 58](#_Toc467757091)

INTRODUÇÃO

A computação em nuvem tem como enfoco proporcionar soluções com baixo custo de forma eficiente para o processamento, armazenamento e distribuições de montantes de dados. Atualmente, existem diversas definições e conceitos para a computação em nuvem. Neste estudo, iremos utilizar a definição de [Mell and Grance 2009], onde de acordo com o mesmo, se pode definir computação em nuvem como sendo um modelo que provê acesso sob demanda a um conjunto de recursos computacionais, onde estes podem ser configurado de acordo com as necessidades, como CPU, armazenamento, memória e outros.

Estes recursos podem ser fornecidos e liberados de forma rápida, utilizado o mínimo de esforço de gerenciamento ou assistência do provedor da nuvem.

Com a revolução de dispositivos móveis e com a demanda crescente de aplicativos, dados, informações (esta que abalou os diversos fatores culturais e áreas da sociedade), processamento e o armazenamento. Os dados foram se tornando cada vez mais importante, onde ao passar das épocas foi-se tornando visível o avanço quanto a necessidade de dados e consequentemente a esta visualização, se foi percebendo um acumulo cada vez maior de informações.

A muito invisível para os usuários como seus aplicativos processam os seus recursos (meta-dados), dados e informações relevantes para a entrega de serviços; para engenheiros da computação e analistas é uma crescente preocupação de como poder crescente cada vez mais provendo o máximo de recursos e dados frente a constante e crescente demanda dos clientes, sem ter o mínimo de downtime (tempo de falha) para novas implementações em produção (deploy).

Um novo conceito está surgindo e revolucionando a maneira de programar e suprir recursos para o ambiente das aplicações: programação em micro-serviços.

Este tipo de programação separa o a aplicação em serviços, da qual a aplicação vai se encaixando como se fosse um bloco de caixas, sendo cada caixa um serviço específico. Estes serviços podem ser providos por módulos ligados a própria aplicação, ou módulos extensores de outras aplicações, que fazem interfaceamento por meio de midlewares (APIs).

Estes serviços não precisam estar no mesma servidor que a aplicação, podem estar em outro servidor, em outro Data center, em outro Cluster (de servidores) e até em outro continente.

Este conceito de serviços para servir aplicações está diretamente ligado ao novo paradigma programação e processamento em nuvem, da qual cada serviço possa estar fora do namespace (ambiente) da aplicação e ser chamado conforme a sua demanda.

Devido a grande demanda de determinados serviços, essenciais a determinadas aplicações, estes precisam estar disponíveis quase que o tempo todo; porém grandes demandas pode ser um problema para os engenheiros de software, administradores de sistemas e a equipe de infraestrutura, sendo necessário aumentar a disponibilidade desse(s) serviço(s) subindo uma nova instância do mesmo (escalonamento);

O conceito de container visa resolver estes e outros problemas; se encaixa perfeitamente para a programação em microseviço, pois cada serviço fica isolado em um único container e este pode ser replicado (escalonado) conforme a demanda.

A minha linha de pesquisa não se fundamenta na aplicação em si, mais no conceito de escalabilidade que se pode ter para cada serviço, recurso, metadados, banco de dados e outros que possam estar servindo à aplicação. Irei fazer uma aplicação para fundamentar as teorias aplicadas nesse trabalho, porém a mesma será meramente aplicável ao trabalho.

Este trabalho se dividirá além da introdução em outros capítulos: 2 Computação em nuvem da qual irei fazer uma fundamentação histórica e teórica sobre a computação em nuvem, modelos de implantação, suas vantagens e desvantagens, seus modelos de serviços, a elasticidade para os seus serviços e pagamento pela utilização de recursos; além de argumentar sobre esses modelos e falar sobre falhas e suas tolerâncias.

No capítulo 3, irei apresentar o conceito de PaaS, plataforma como serviço e alguns outros conceitos relativos a computação em nuvem, falarei sobre modelos de programação descentralizada e os conceitos de grid, cluster e docker implementados para esse tipo de plataforma e modelo de programação.

No capítulo 4, irei falar sobre alguns softwares de deploy (implementação) e os conceitos relativos a essa nova arquitetura de Paas, me fundamentarei em alguns como: o Tsuru, Vagran, Kubernet; irei fazer fundamentação teórica sobre cada plataforma, além de fazer uma breve comparação (no capítulo 4).

No capítulo 5, irei falar sobre sistemas escalares, ou seja sistemas operacionais que já fazem a implementação de escalabilidade de forma nativa como se fosse um verdadeiro Paas, esses sistemas foram criados para este tipo de arquitetura;

No capítulo 6, irei fazer uma comparação teórica com pontos positivos e negativos sobre as PaaS abordadas; Irei fazer esta comparação somente no capítulo 6 e não no capítulo 4 pois faço necessário informar sobre sistemas escalares e irei fazer comparação desses com os PaaS.

No capítlulo 7, irei abordar e fundamentar a teoria sobre programação em micro-serviço, boas práticas desse modelo de programação para o novo modelo de arquitetura descentralizado e escalável.

No capítulo 8, irei falar sobre o estudo de casos abordado no trabalho, farei fundamentação teórica para o motivo de meu estudo e trabalho, além de falar sobre infraestrutura relativa à este tipo de arquitetura e aplicabilidade deste modelo de computação.

No capítulo 9, irei fazer a conclusão do meu trabalho; Abordando meu ponto de vista e críticas e soluções relativas à este novo modelo e paradigma de programação.

2 COMPUTAÇÃO EM NUVEM

O termo de computação em nuvem se refere à entrega de recursos computacionais através da Web ou de uma rede própria.

De acordo com o Instituto Nacional de Padrões e Tecnologia do Departamento de Comércio Norte-Americano:

Computação em nuvem é um modelo para permitir acesso ubíquo, conveniente e

sob demanda via rede a um agrupamento compartilhado e configurável de recursos computacionais (por exemplo, redes, servidores, equipamentos de armazenamento,

aplicações e serviços), que pode ser rapidamente fornecido e liberado com esforços

mínimos de gerenciamento ou interação com o provedor de serviços. (NIST, 2015).

Ao invés de manter servidores e infraestrutura especializada para processamento de dados, storage de arquivos e outros serviços, utiliza-se serviços online (pode ser de outras empresas) para tal fim. Porém ao se expor a esse tipo de serviço, torna-se necessário ter algum tipo de questionamento e expertise para se manter a segurança, a confiabilidade e a privacidade desses dados.



Figure 1: Clound Fonte: http://www.synergixtech.com/wp-content/uploads/2016/09/Cloud-Computing-Benefits.png

2.1 História

A principal funcionalidade da computação em nuvem é a utilização de recursos computacionais por meio da web, e a idéia não é tão recente, a ideia já existia em 1960, com Joseph Carl Robnett Licklider.

Joseph Carl foi um dos desenvolvedores da ARPANET (*Advanced Research Projects Agency Network*), o antecessor direto da internet, que tinha o objetivo de interligar as bases militares e os departamentos de pesquisa do governo americano.

Nesse período, Joseph já imaginava uma rede de computadores intergaláctica em que todos estariam conectados acessando programas e dados de qualquer lugar.



Figure 2: Joseph Carl Fonte: http://www.psynergie.com/psychologie-internet/photo-joseph-licklider.jpg

Na década de 1960, John McCarthy, um importante americano pesquisador da área da informática e também um dos pioneiros da inteligência artificial, propôs a ideia de que a computação deveria ser organizada na forma de um serviço de utilidade pública, assim como os serviços de água e energia, em que os usuários só pagam pelo que usam, sendo precursor da idéia de PaaS.

Figure 3: John McCarthy Fonte: http://www-formal.stanford.edu/jmc/jmccolor.jpg

Mesmo com a existência dessas ideias há tanto tempo, o termo computação em nuvem só veio a ser mencionado em 1997, numa palestra acadêmica do professor de sistemas da informação Ramnath Chellappa, e só foi desenvolvida no ano de 1999 com o surgimento da Salesforce.com, primeira empresa a disponibilizar aplicações pela internet.



Figure 4: Ramnath Chellappa Fonte: http://goizueta.emory.edu/profiles/images/portrait/chellapa.jpg

A partir do sucesso dessa empresa, outras grandes começaram a investir na área, como à Amazon, a Google, a IBM e a Microsoft.

O que conhecemos hoje como computação em nuvem, nasceu com os sistemas distribuídos, caracterizados por serem um conjunto de unidades de processamento independentes, que, por meio da troca de comunicação e gerenciamento de sincronização, pode processar uma aplicação em diferentes localidades, de forma transparente para o usuário, ou seja, o usuário da aplicação vê apenas o todo. A computação em nuvem vai além disso, trata-se de um formato de computação a partir do qual aplicativos, serviços, dados e recursos de TI são disponibilizados aos usuários como serviço, por meio da internet.

Não é mais necessário, para algumas empresas, ter supercomputadores, pois o poder de processamento e os dados ficam nas nuvens. Só precisamos de dispositivos que nos dêem acesso a esses recursos. Dispositivos estes que consequentemente são mais baratos e possuem uma maior portabilidade e flexibilidade, como smartphones, *tablets* e *netbooks*. Só precisamos agora de dispositivos de entrada e saída e acesso à internet.

2.2 Modelos de implantação

Figure 5: Modelos de Implementação fonte: http://www.vividdynamics.com/wp-content/uploads/2013/12/cloud-hosting.jpg

É possível implantar soluções utilizando computação em nuvem de maneiras diferentes, dependendo de fatores, como:

* Requisitos de segurança;
* Hospedagem dos serviços;
* Capacidade de customização;
* Nível de acesso;
* Gerenciamento de serviços;

Existem modelos principais que são determinados pelo nível de acesso: Nuvem pública, Nuvem privada, Nuvem Comunitária e Nuvem híbrida.

2.2.1 Nuvem Pública



Figure 6: Nuvem pública Fonte: http://www.ximedica.info/images/uploads/the\_cloud-resized-600.jpg

Uma infraestrutura de nuvem pública é disponibilizada para o público geral e é de propriedade de um provedor de serviços de web (terceirizado). Em uma nuvem pública, os recursos de computação são disponibilizados dinamicamente através do provedor e são fornecidos para seus clientes (rede própria utilizando VPNs tendo como base a infraestrutura da web); A cobrança feita pelo provedor é feita pela quantidade de recursos que o cliente utiliza (isto pode variar em contrato).

Esse modelo possui boa relação custo/benefício para o cliente, uma vez que oferece a flexibilidade de disponibilizar apenas os recursos necessários e entregar todos os serviços com certa consistência de disponibilidade, resiliência, segurança e facilidade de gerenciamento. Como este modelo é baseado na web e possui infraestrutura e meios de acesso administrados pelo provedor de acesso o cliente precisa aceitar o controle reduzido e o monitoramento feito do provedor, além de confiar na governança e segurança da infraestrutura.

2.2.2 Nuvem Privada

Figure 7: Nuvem Privada Fonte: https://web-material3.yokogawa.com/image\_8952.jpg

Uma infraestrutura de nuvem privada é operada apenas para um único cliente, o provedor dedica serviços específicos para aquele cliente. O cliente especifica, arquiteta e controla toda a gama de recursos computacionais que o provedor fornece. Um motivo comum que leva os clientes a procurarem um serviço de nuvem privado é a capacidade de controlar e garantir seus próprios padrões de segurança dos dados.

Um cliente normalmente fará uso de uma nuvem pública utilizando conexões através de links privados e esses recursos apenas serão compartilhados internamente. Como os recursos não são compartilhados entre várias organizações, o cliente, que contratou o serviço, paga o valor total pelos recursos da nuvem, independente da quantidade que foi utilizada.

Sendo assim, a organização contratante pode realocar os recursos para subáreas da mesma empresa conforme suas necessidades.

2.2.3 Nuvem Comunitária

Figure 8: Nuvem Comunitária Fonte: https://lh4.googleusercontent.com/NWf67CzmfbXLVsj60ZRyC-eX-UcdAt3ITRZcgedyN4dBbGU0BOWdCSNdtuqz9DxZ4fHNC6GLnUlreeoRX\_\_8c07l61YMMoY3zxKJvMkbFfx92vjjDYLm1ai2STm0h4XQfA

Uma infraestrutura de nuvem comunitária é contratada por um grupo de organizações em conjunto ou programas que compartilham necessidades especificas, como segurança e aspectos legais. O controle da nuvem pode ser feito pelo cliente ou pelo provedor, de acordo com o que foi combinado no contrato.

Quando organizações possuem o mesmo conjunto de requisitos e clientes, a nuvem comunitária permite a eles combinarem ferramentas e compartilharem recursos computacionais, dados e capacidades. Ao eliminar a duplicidade de sistemas similares, as organizações podem economizar dinheiro e alocar seus recursos de maneira mais eficiente.

2.2.3 Nuvem Híbrida



Figure 9: Modelos de Implementação Fonte: https://puserscontentstorage.blob.core.windows.net/userimages/de1cc483-bb71-4170-bd25-0c04f167acf5/c9851e30-da98-4765-92bb-d33ca089ff49image32.png

Uma infraestrutura de nuvem híbrida abrange duas ou mais nuvens, podendo estas serem nuvens públicas, comunitárias ou privadas, com o conjunto dos serviços que são hospedados internamente ou externamente.

Os clientes geralmente não se limitam a um único modelo de implantação, mas sim preferem incorporar serviços de nuvem diferentes e sobrepostos para atingir seus requisitos específicos. Modelos de implantação híbridos são complexos e requerem um planejamento

cuidadoso para serem executados e gerenciados especialmente quando é necessária a comunicação entre dois tipos diferentes de implantações em nuvem.

2.3 Princípios da computação em nuvem

Segundo NIST (2011), um modelo de Computação em Nuvem deve apresentar algumas

características essenciais:

* Autoatendimento sob demanda: o usuário pode usufruir das funcionalidades computacionais sem a necessidade da interação humana com o provedor de serviço, ou seja, o provedor identifica as necessidades do usuário, podendo assim automaticamente reconfigurar todo hardware e software, e essas modificações devem ser apresentadas ao usuário de forma transparente.
* Amplo acesso a serviços de rede: os recursos computacionais são acessados através da internet, que são acessados por mecanismos padronizados, que pode ser um navegador simples que use poucos recursos computacionais, sem a necessidade do usuário modificar o ambiente de trabalho de seu dispositivo, como por exemplo, linguagem de programação e sistema operacional.
* Pool de recursos: os recursos computacionais (físicos ou virtuais) do provedor são divididos em pools para que possam atender a múltiplos usuários simultaneamente. Esses recursos são alocados e realocados dinamicamente, de acordo com a demanda dos usuários. Os usuários por sua vez não precisam saber a localização física dos recursos computacionais, essas informações podem ser proporcionadas de maneira de alta abstração podendo apenas ser informados o

país, estado ou centro de dados.

* Elasticidade rápida: as funcionalidades computacionais devem ser rápidas e elásticas, assim como rapidamente liberadas, podendo em alguns casos serem liberadas automaticamente caso haja necessidade devido a demanda. O usuário deve ter a impressão de ter recursos ilimitados que podem ser comprados ou adquiridos em qualquer quantidade e a qualquer momento. A elasticidade deve ter três componentes: escalabilidade linear, utilização on-demand e pagamento por unidades consumidas de um recurso. Outro recurso que pode auxiliar nesse processo é a virtualização que pode criar várias instâncias de recursos requisitados usando apenas um recurso físico. A virtualização também torna possível abstrair características físicas de uma plataforma computacional, emulando vários ambientes que podem ser independentes ou não.
* Serviços mensuráveis: os sistemas em nuvem automaticamente controlam e monitoram os recursos necessários para cada tipo de serviço, tais como armazenamento, processamento e largura de banda. Esse recurso deve ser monitorado e controlado de forma transparente tanto para o provedor de serviço quanto para o usuário.

Além dessas características, algumas outras não tão essenciais em um ambiente de computação em nuvem, mais que definem para o bom serviço às aplicações estão hospedadas e o bom relacionamento com o cliente, são essas:

* Tolerância a falhas: O provedor de serviço tem que ser totalmente redundante em sua infraestrutura e possibilitar alta disponibilidade de dados para os seus clientes, de modo que se houver alguma falha ou problema em sua estrutura de rede, servidor, processamento, armazenamento e outros, os dados e aplicações dos clientes possam ser realocados para outras máquinas (servidores), datacenters, cluster e até continentes sem que o cliente perceba que tal falha ocorreu.
* Níveis de Qualidade de Serviço de SLA: Esta questão está diretamente ligada com a tolerância à falhas, da qual em último caso de não poder ser tolerante à falhas e a falha de fato ocorrer que o provedor de serviço possa dar o mínimo de assistências aos seus usuários e que o mesmo possam ter alguma garantia de que os seus serviços, possam ser restabelecidos dentro do tempo referenciado em contrato.

2.4 Modelos de Serviços

A idéia de Computação em Nuvem é composta por modelos de serviços, esses modelos são pagos conforme a necessidade e o uso dos mesmos (pay-per-use), dando ao cliente a possiblidade de usar mais ou menos recursos de acordo com sua necessidade. Os modelos de serviços são os seguintes:

* Software como Serviço (SaaS): um aplicativo pode ser utilizado por uma grande quantidade de usuários simultaneamente. Esse tipo de serviço é disponibilizado por provedores e acessado pelos usuários através de aplicações como o navegador. Todo o controle e gerenciamento da rede, sistemas operacionais, armazenamento e possíveis manutenções será de responsabilidade do provedor de serviço (Aulbach, 2009).
* Plataforma como Serviço (PaaS): é a disponibilização de plataformas de desenvolvimento que facilitam a implantação de aplicações assim como o gerenciamento do hardware subjacente e das camadas de software. O usuário não tem controle sobre a rede, sistemas operacionais ou armazenamento, mas poderá controlar a aplicação implementada na nuvem. A linguagem de programação bem como o ambiente de desenvolvimento é fornecida pelo provedor (NOGUEIRA, 2010).
* Infraestrutura como Serviço (IaaS): consiste no fornecimento de infraestrutura de processamento, armazenamento, redes, entre outros. Este serviço, assim como os demais, tem seus recursos – neste caso a infraestrutura – compartilhados com diversos usuários simultaneamente. Isso se torna possível através do processo de virtualização, no qual o usuário terá controle sobre máquinas virtuais, armazenamento, aplicativos instalados e possivelmente um controle limitado sobre os recursos de rede (VERAS, 2012).



Figure 10: Modelos de Serviço Fonte: http://ftp.unipar.br/~seinpar/2013/artigos/Rogerio%20Schueroff%20Vandresen.pdf

3 Paas

3.1 PaaS: Plataforma como serviço

3.2 Programação descentralizada: GRID e Cluster

3.3 Programação em micro-serviços

3.4 Docker

3.4.1 Conceito

4 SOFTWARES DE DEPLOY

4.1 Tsuru

4.1.1 Conceito

4.1.2 Implementação em Cluster próprio

4.2 Vagran

4.2.1 Conceito

4.2.2 Implementação em Cluster próprio

4.3 Kubernet

4.3.1 Conceito

4.3.2 Implementação em Cluster próprio

5 SISTEMAS ESCALÁVEIS

Os sistemas de virtualização passaram por algumas mudanças ao longo do tempo até chegarem aos sistemas atuais, em primeira fase eram servidores em grande escala que ocupavam grande parte de uma sala, sem virtualização e que rodavam somente uma aplicação, não usando nem metade do seu poder de processamento e com grandes gastos de energia; Se necessário ampliação havia custos de compra de hardware e de espaços, energia, gerenciamento centralizado das máquinas, segurança entre outros.

Com o surgimento da virtualização houve a inserção de um hypervisor no sistema operacional, que consiste em um gerenciador para a virtualização.

( imagem)

A evolução deste modelo consiste em não usar mais a camada do sistema operacional, e sim que o próprio hypervisor é que faz a gestão em cima da camada do hardware, se tornando um sistema operacional de gerenciamento.

(imagem)

O modelo de containers elimina a camada do sistema operacional que existia para cada máquina virtualizada e o hypervisor para gerenciar as instâncias virtualizadas. A virtualização pelo container utiliza o kernel do sistema operacional nativo da máquina, geralmente sendo o linux, mais hoje já é possível ser feito em qualquer sistema operacional.

(imagem)

O kernel do sistema é responsável por fornecer as bibliotecas e os binários necessários para que o container possa rodar a aplicação de forma necessária.

Basicamente há uma abstração do nível de sistema operacional, pois as configurações ficam no container e o mesmo pode ser tem um alto nível de provisionamento, podendo ser replicado sem menores problemas, simplificando muito a implementação em diferentes máquinas e ambientes de homologação (QA – Quality Qssurance, Staging, Produção); o containter pode ser baixado em qualquer lugar, e máquina hospedeira utilizando o dockerhub.

Esse dinamismo foi um dos alicerces da computação em nuvem, permitindo replicações de aplicações e escalabilidade da mesma a níveis nunca vistos antes.

5.1 LXC Containers

Conforme citação do Rogério dos Anjos (em [Linux](http://3way.com.br/category/linux/),[Novidades](http://3way.com.br/category/novidades/)) - “O LXC é uma sistema leve de virtualização que usa múltiplos containers de forma isolada no kernel linux. Ele cria um ambiente muito próximo de um sistema linux sem precisar instalar um kernel separado. Com o LXC é possível criar processos separados para usuários, espaço em disco, memória, CPU, rede e muito mais”

O container tem a característica de isolar os recursos do sistema dos recursos da aplicação, criando assim um ambiente isolado; o mesmo poder consumir recursos do sistema como: namespace, chroot, cgroups entre outros. Funciona com se cada container fosse uma máquina virtual completa, podendo ter inclusive os mesmos problemas de uma máquina virtual, como criação de usuários e permissões de escritas em diretórios e etc.

O resultado é uma máquina virtual sem a camada do hypervisor, isolada e com controle de recursos.

Conforme citação do Rogério dos Anjos (em [Linux](http://3way.com.br/category/linux/),[Novidades](http://3way.com.br/category/novidades/)) - “ Os containers fornecem um ambiente mais próximo possível de um sistema operacional do que você conseguiria de uma VM, mas sem a sobrecarga da execução separada do kernel e da simulação de hardware do sistema.“

(imagem)

Os benefícios da utilização do LXC container, viabilizou grandes benefícios para a virtualização e para a computação em nuvem, como:

\* Provisionamento:

\* Escalabilidade:

\* Custo:

5.2 Máquinas Virtuais

5.3 Comparações

6 COMPARAÇÃO ENTRE PLATAFORMAS

6.1 Plataformas

6.2 Tolerâncias à falhas

6.2 Segurança das plataformas

7 PROGRAMAÇÃO EM MICROSERVIÇOS

7.1 Boas práticas de programação distribuída (Doze fatores)

7.2 Empregabilidade da técnica

7.3 Casos de Sucesso

8 ESTUDO DE CASO

8.1 Infraestrutura

8.2 Métricas de Escalabilidade

9 CONCLUSÃO

10 Referências

História da computação em nuvens. Disponível em <http://www.dsc.ufcg.edu.br/~pet/jornal/agosto2012/materias/historia\_da\_computacao.html>. Acessado em 18/11/2016

Clound Computing. Disponível em <https://en.wikipedia.org/wiki/Cloud\_computing>. Acessado em 20/11/2016.

O que é clound computing. Disponível em <http://www.infowester.com/cloudcomputing.php>. Acessado em 20/11/2016.

Fundamentos de Computação Nuvem para Governos. Disponível em <http://www4.serpro.gov.br/wcge2010/artigos/Artigo-Fundamentos%20de%20Computacao%20Nuvem%20para%20Governos.pdf>. Acessado em 21/11/2016.

Clound computing estamos nas nuvens voando. Disponível em <http://ppgi.ci.ufpb.br/wp-content/uploads/palestra-cloud-computing.pdf>. Acessado em 21/11/2016

“The NIST Definition of Cloud Computing”. Disponível em <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf >. Acessado em 24/11/2016

Conceitos de computação em nuvem. Disponível em < http://ftp.unipar.br/~seinpar/2013/artigos/Rogerio%20Schueroff%20Vandresen.pdf>. Acessado em 24/11/2016

Nogueira, Matheus Cadori1; PEZZI, Daniel da Cunha (2010) “A Computação Agora é nas Nuvens” Universidade de Cruz Alta (UNICRUZ) – Cruz Alta, RS – Brasil.

Veras, Manoel.(2012)”Cloud Computing: Nova Arquitetura da TI”. Editora Brasport: Rio de Janeiro, Brasil.

Souza, Flávio R. C.; Moreira, Leonardo O.; Machado, Javam C. Computação em Nuvem: Conceitos, Tecnologias, Aplicações e Desafios. ERCEMAPI 2009.