

UNIVERSIDADE FEDERAL DO PARÁ

Campus Universitário de Tucuruí – Faculdade de Engenharia da Computação

Ronison Caldas Cruz

Ruan Mateus Conceição Pereira

Tauan Jorge Miranda

CLOUD DATABASES

**BANCO DE DADOS DISPONIBILIZADOS COMO SERVIÇO EM PLATAFORMAS
DE NUVEM**

TUCURUÍ

2019

Ronison Caldas Cruz

Ruan Mateus Conceição Pereira

Tauan Jorge Miranda

CLOUD DATABASES

BANCO DE DADOS DISPONIBILIZADOS COMO SERVIÇO EM PLATAFORMAS DE NUVEM

Monografia apresentada à disciplina de Banco de Dados da Universidade Federal do Pará, como requisito parcial para obtenção de nota na referida disciplina.

Orientador: Prof. Renato Luz Cavalcante

TUCURUÍ

2019

Ronison Caldas Cruz

Ruan Mateus Conceição Pereira

Tauan Jorge Miranda

CLOUD DATABASES

BANCO DE DADOS DISPONIBILIZADOS COMO SERVIÇO EM PLATAFORMAS DE NUVEM

Monografia apresentada à disciplina de Banco de Dados da Universidade Federal do Pará, como requisito parcial para obtenção de nota na referida disciplina.

Orientador: Prof. Renato Luz Cavalcante

Prof. Marcos – UFPA

Tucuruí – PA, 24 de junho de 2019

A todos colegas, amigos e companheiros de jornada, que de alguma forma contribuíram para a conclusão desta monografia.

AGRADECIMENTOS

Agradecemos a Deus, pela saúde e disposição que nos permitiram a realização deste trabalho.

As nossas companheiras e filhos pela compreensão na ausência nos finais de semana durante esse mês.

Ao querido professor Belleti, pelos conselhos e palavras de carinho e consolo durante esta jornada

Ao magnânimo professor Clenilson, pela transmissão de todo conhecimento e em especial por segurar na mão do Ronison durante esse difícil período e não só transmitir boas energias, mas por mostrar por onde entra e sai.

Agradecemos também ao professor Marcos pela paciência e dedicação durante o curso que nos conduziu até aqui.

**“Tentei fugir de mim, mas onde eu ia eu
tava lá!!!”**

Tiririca

RESUMO

Computação em Nuvem surge como um novo paradigma na implantação de aplicações em que os recursos computacionais são fornecidos como um serviço através de uma conexão de rede. Esses serviços são disponibilizados em forma de software, infraestruturas e plataformas. Com a utilização desses serviços usuários estão disponibilizando informações, muitas vezes confidenciais em servidores remotos desconhecidos geograficamente, ou seja, os bancos de dados em nuvem. Os maiores desafios da atualidade na área de computação em nuvem são a manipulação e a técnica de armazenagem de vultosa quantidade de informações no sistema chamado de Big Data. Com a competição acirrada das empresas, as informações necessitam ser tratadas de um modo representativo, para que, dessa forma, possam auxiliar as empresas a gerar informações, a reduzir custos, aumentar os lucros, verificar atividades da empresa no passado e identificar novas oportunidades de negócio. Porém, para implantar e atualizar o arquivo de informações, alguns problemas podem ser encontrados, pois, mesmo com hardware e com software atualizados, alguns obstáculos se levantam, como os altos custos, mau desempenho, resposta inflexível de conhecimentos desordenados. Esta monografia apresenta uma visão geral desse tipo de banco de dados, assim como características, um breve resumo do seu surgimento, vantagens, possíveis problemas e dificuldades, além de suas mais diversas aplicações, além de tratar dos desafios gerais e de segurança, encontrados no gerenciamento dos dados armazenados nesse ambiente de nuvem.

Palavras-chave: computação em nuvem. banco de dados. sistema de segurança SGBD.

ABSTRACT

Cloud Computing emerges as a new paradigm in deploying applications in which computing resources are delivered as a service over a network connection. These services are provided in the form of software, infrastructures and platforms. With the use of these services users are providing information, often confidential on geographically unknown remote servers, that is, the cloud databases. The biggest challenges of today in the area of cloud computing are the manipulation and technique of storing a large amount of information in the system called Big Data. With the fierce competition of companies, the information needs to be treated in a representative way, so they can help companies generate information, reduce costs, increase profits, verify company activities in the past and identify new opportunities for business. However, in order to deploy and update the information file, some problems can be encountered because even with updated hardware and software, some obstacles arise, such as high costs, poor performance, and inflexible response to disordered knowledge. This monograph presents an overview of this type of database, as well as features, advantages, possible problems and difficulties besides its most diverse applications, as well as dealing with the general and security challenges found in managing the data stored in this cloud environment.

Keywords: cloud computing. database. SGBD security system.

LISTA DE FIGURAS

FIGURA 1- ILUSTRAÇÃO BANCO DE DADOS EMPRESARIAL.....	15
FIGURA 2- JHON MCCARTHY	16
FIGURA 3- REPRESENTAÇÃO DO RESCIMENTO DO B.D EM NUVEM	17
FIGURA 4- EMPRESAS QUE UTILIZAM O BDN	20
FIGURA 5- SEGURANÇA DE UM BDN	21

LISTA DE SIGLAS

AWS	- Amazon Web Services
ARPA	- Advanced Research and Project Agency
BD	- Banco de Dados
BDN	- Banco de Dados em Nuvem
GAE	- Google App Engine
SGBD	- Sistema de Gerenciamento de Banco de Dados
SGDN	-Sistemas de Gerenciamento de Dados em Nuvem
SGBD	- Sistema de Gerenciamento de dados

SUMÁRIO

1 INTRODUÇÃO	12
1.1 CONTEXTO.....	12
1.2 OBJETIVOS	13
1.3 JUSTIFICATIVA.....	14
2 REVISÃO TEÓRICO-IMPÍRICA.....	15
2.1 BANCO DE DADOS EM NUVEM	15
2.2	16
CONTEXTO HISTÓRICO	16
2.2.1 ACENSSÃO DO CLOUD COMPUTING	17
2.3 VANTAGENS E DESVANTAGENS.....	18
3 APLICAÇÕES.....	20
4 DESAFIOS E SEGURANÇA.....	21
5 CONCLUSÃO	23

1 INTRODUÇÃO

1.1 CONTEXTO

A computação em nuvem é um sistema de serviços computacionais acessados via internet e sob demanda, onde usuários pagam apenas pelos serviços que utilizam. Esse modelo de serviço é semelhante a alguns modelos já conhecidos, como serviço de água, luz, telefone, onde o usuário só paga pelo que efetivamente ele consumiu.

A computação em nuvem trabalha com recursos dinamicamente escaláveis e normalmente virtualizados através de serviços sobre a Internet. Nesse modelo, desenvolvedores de serviços para Internet não precisam mais investir muito capital em hardware e em recursos humanos para manter o serviço em operação. Além disso, as empresas podem usar a infraestrutura fornecida por provedores de serviços em nuvem, caso precisem de um poder de processamento alto, ao invés de investir recursos em máquinas físicas e equipe técnica especializada para fazer reparos.

Do ponto de vista do usuário, infraestruturas de computação em nuvem disponibilizam sistemas operacionais que possibilitam que os usuários, utilizando qualquer dispositivo conectado à Internet, possam ter acesso aos serviços, arquivos, informações e programas situados “na nuvem”. Assim, os dispositivos utilizados pelos usuários não necessitam de grandes recursos computacionais, aproximando-se de simples terminais.

Hoje em dia há vários ambientes de nuvem disponíveis no mercado que prestam suporte ao desenvolvimento de aplicações, como a Amazon Web Services (AWS), Google App Engine (GAE), Microsoft Azure Platform³ e OpenStack⁴ são alguns exemplos de ambientes de desenvolvimento de aplicações em nuvem. Esses ambientes seguem diferentes tipos de modelos de provisão de serviços e cada um fornece uma gama de serviços que se diferenciam em diversos aspectos tais como preço, facilidade de uso e desempenho.

No cenário que temos um dos principais desafios em termos de desenvolvimento de aplicações e com a implementação, implantação e migração de aplicações na nuvem, que são tarefas desafiadoras no sentido da maior complexidade inerente à heterogeneidade desses ambientes (Galán et al. 2009). Esses ambientes não são implementados utilizando padrões comuns; cada um possui sua própria API, ferramentas de desenvolvimento, mecanismos de virtualização, características de

governança, tecnologias de implantação e gerenciamento de recursos, etc. Consequentemente, os usuários tendem a optar por implementar suas aplicações em um ambiente específico, tornando-se altamente dependentes de um único provedor de nuvem, problema conhecido como cloud lock-in (Armbrust et al. 2009), assim impossibilitando o uso de recursos de ambientes diferentes.

1.2 OBJETIVOS

- **OBJETIVO GERAL**

Identificar os principais aspectos que cercam e corroboram para o contexto referente a segurança de um B.D (banco de dados) em nuvem, demonstrando como são afetados através de falhas de segurança e as contramedidas adotadas por todo o sistema que asseguram a confiabilidade e a integridades dos dados nesse contexto.

1.3 JUSTIFICATIVA

O gerenciamento de dados é um fator muito importante dentro do contexto de computação em nuvem, uma vez que a segurança desses dados é fator crucial em ambientes de Cloud Computing e deve ser tratado com certa atenção. Outro fator relevante é que os Sistemas de Gerenciamento de Banco de Dados (SGBDs) relacionais não possuem escalabilidade quando existe uma grande quantidade de dados armazenados por eles. Assim, aspectos de armazenamento de dados, processamento de consultas e controle transacional têm 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). A escalabilidade do sistema deve ser transparente para os usuários, podendo estes armazenar seus dados na nuvem sem a necessidade de saber a localização dos dados ou a forma de acesso (Sousa et al., 2010). A escalabilidade do sistema ocorre quando o mesmo possui a capacidade de crescer conforme a demanda de uso. Já a disponibilidade do serviço possibilita aos usuários acesso aos dados em qualquer momento, onde quiserem ou precisarem, os Sistemas de Gerenciamento de Dados em Nuvem (SGDN) devem dispor de uma alta disponibilidade, visto que o meio de comunicação entre o usuário e o sistema é a internet, e que pode ocorrer atrasos e indisponibilidade do sistema. Outro item relevante no contexto de gerenciamento de dados em nuvem é a consistência de dados, onde todos os nós devem ter a mesma visão dos dados ao mesmo tempo, fazendo assim, com que todos os dados armazenados sejam iguais para todos os nós do sistema, por exemplo, se um determinado dado for atualizado, então ele deverá ser atualizado para todos os nós do sistema, mantendo assim os dados íntegros, coerentes, consistentes.

2 REVISÃO TEÓRICO-IMPÍRICA

2.1 BANCO DE DADOS EM NUVEM

Banco de Dados em Nuvem (BDN) estão começando a ser utilizados e têm o potencial de atrair clientes de diversos setores do mercado, desde pequenas empresas com o objetivo de reduzir o custo total, por meio da utilização de infraestrutura e sistemas de terceiros, até grandes empresas que buscam soluções para gerenciar milhares de máquinas e permitir o atendimento de um aumento inesperado de tráfego (Abadi, 2009). Segundo Cottman (2011), há uma palavra que resume bem a origem dos BDN: volume. A evolução da Internet tem gerado grandes volumes de dados (terabytes, petabytes, e assim por diante) que surpreendem até mesmo a mente de um profissional de TI, em contrapartida os bancos de dados relacionais convencionais não se mostraram suficientemente preparados para lidarem da melhor forma possível com um quantitativo de dados dessa magnitude. Não é por acaso que quando se pergunta como portais bem estabelecidos como o Google, o Facebook e o Twitter poderiam armazenar, processar e arquivar essa vasta quantidade de dados, a resposta é uma só: através de BDN.

FIGURA 1- ILUSTRAÇÃO BANCO DE DADOS EMPRESARIAL.



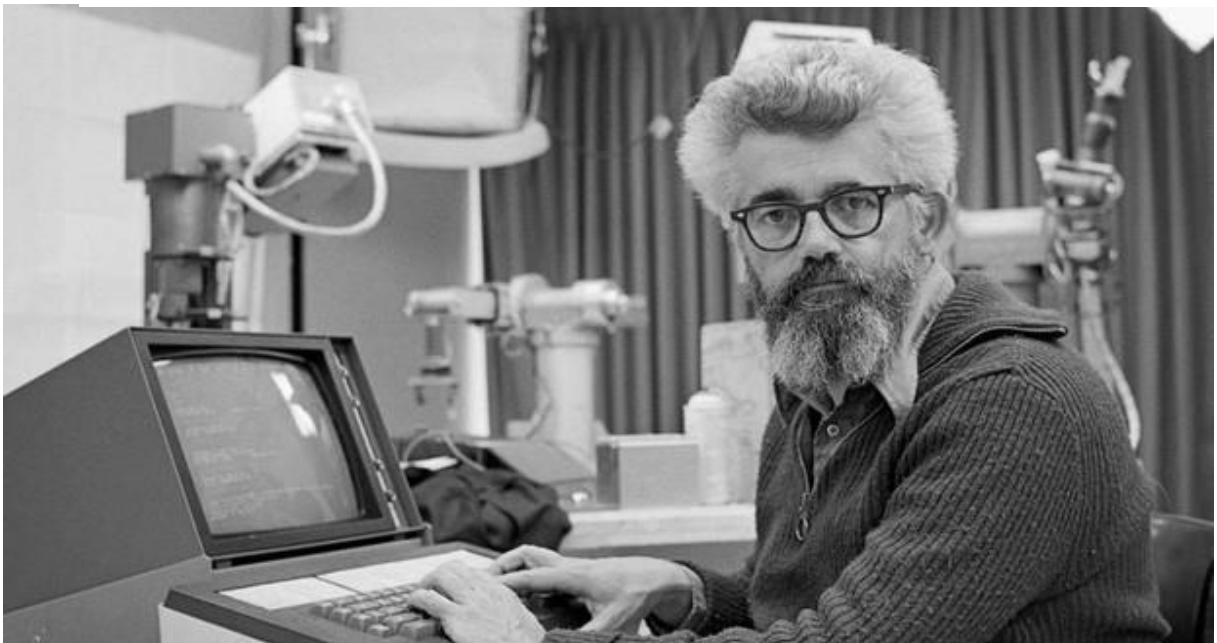
Em linhas gerais pode-se defini-lo como uma coleção de dados inter-relacionados que estão armazenados na web, e que podem ser gerenciados e manipulados através de Sistemas de Gerenciamento de Dados em Nuvem (SGDN), softwares estes que segundo Cottman (2011) são especialmente concebidos para

atingir o processamento massivamente paralelo – na ordem das centenas de milhares de tarefas simultâneas -, utilizando os recursos distribuídos de computação em grid.

2.2 CONTEXTO HISTÓRICO

A história da computação em nuvem é bem mais antiga do que se imagina. Pode-se dizer que o conceito de computação em nuvem surgiu com dois grandes especialistas na área. O primeiro deles, o americano John McCarthy, sugeriu uma ideia que se assemelha ao conceito de internet atual: ele queria que a computação

FIGURA 2- JHON MCCARTHY



fosse oferecida como um serviço público, a exemplo dos serviços de água e luz, em que só se paga o que se usa. O pesquisador também discutiu o uso compartilhado do computador, de forma simultânea, por dois ou mais usuários. Essas sugestões foram apresentadas em um discurso no Massachusetts Institute of Technology (MIT), nos EUA, em 1961. O conceito foi chamado por ele de "*Utility Computing*". Em 1962 foi a vez de Joseph Carl Robnett Licklider estudar novas formas de usar o computador. Foi aí que o físico encontrou um jeito de as pessoas compartilharem dados de forma global. Com essa descoberta ele criou a rede Arpanet, conhecida como a "mãe" da internet, para a agência americana Advanced Research and Projects Agency (Arpa). O objetivo da rede era unir as bases militares com os

departamentos de pesquisa do governo americano. Tanto McCarthy quanto Licklider são vistos como os grandes pioneiros da computação em nuvem.

As ideias dos pioneiros da nuvem ficaram em *standby* por um bom tempo. Foi só quando a população começou a ter mais acesso à tecnologia, em 1997, que o termo "computação em nuvem" foi utilizado pela primeira vez pelo professor de sistemas de informação, Ramnath Chellappa, em uma palestra acadêmica. O termo foi inspirado no símbolo da internet, a "nuvem", que representa algo que está "no ar". Já em 1999 a americana Salesforce.com ficou conhecida como a primeira empresa a disponibilizar aplicações empresariais pela web.

2.2.1 ACENSSÃO DO CLOUD COMPUTING

A partir dos anos 2000 a cloud computing começou a ganhar mais força e passou a ser oferecida comercialmente. Em 2002 a Amazon lançou um conjunto de serviços que incluía armazenamento, computação e inteligência humana, tudo baseado em nuvem. Anos depois lançou a Amazon EC2/S3, primeiro serviço de infraestrutura de computação. Já em 2008 a Google App Engine lançou seu serviço no mercado de nuvem com custos baixos e inovações. No ano seguinte a Microsoft Azure criou sua

FIGURA 3- REPRESENTAÇÃO DO RESCIMENTO DO B.D EM NUVEM



plataforma especial para execução de aplicativos e serviços, baseada também nos conceitos da computação em nuvem. Após isso empresas de vários tamanhos começaram a adotar o armazenamento e de lá pra cá o investimento tem sido cada vez maior.

2.3 VANTAGENS E DESVANTAGENS

Com a disseminação de dispositivos móveis e computadores, produziu-se cada vez mais arquivos: são textos para a escola e o trabalho, vídeos e, em especial, fotos. Estima-se, por exemplo, que a cada dois minutos sejam tiradas mais fotos no mundo que o total de fotos tiradas em todo o século XIX. E onde guardar tudo isso? Nesse exato contexto, com a evolução da tecnologia, surgiu o armazenamento em nuvem.

Segurança: os sites do também chamado cloud computing fracionam os documentos e espalham esses pedaços em diferentes servidores ao redor do mundo. Muitos utilizam a duplicidade (cópia) de arquivos. Além disso, os dados passam por criptografia, o que impede possíveis invasores de terem acesso ao conteúdo completo. Ou seja, dificilmente uma pessoa perderá seus diretórios e tudo que houver dentro deles, mesmo em caso de ataques à plataforma.

Mobilidade: um arquivo armazenado na nuvem pode ser acessado em qualquer lugar com conexão à internet, por meio dos mais diversos dispositivos. E a atualização é feita automaticamente. Se o usuário altera um dado, por exemplo, a modificação ocorre em todos os aparelhos simultaneamente – tablet, celular, notebook ou computador de mesa.

Compartilhamento: Nesse tipo de tecnologia, basta enviar o link do arquivo ou criar uma pasta compartilhada. Assim, todos os colegas têm acesso ao mesmo conteúdo e podem editá-lo. Não precisando mais de dezenas Pen drives ou CD's.

Esse recurso é útil para trabalhos realizados em grupo, ou então para acompanhar planilhas com fluxos de produção de uma empresa. Agilizando assim todo os serviços como um todo.

Backup das alterações: o armazenamento cloud anula o risco de sobrescrever algum arquivo acidentalmente, já que o backup contém um histórico de alterações. É possível resgatar versões anteriores! Sem o perigo do retrabalho, a produtividade aumenta.

Espaço livre: ao passar diretórios à nuvem, o usuário libera espaço no disco rígido local para usar da melhor forma que desejar.

Editável: No armazenamento cloud o salvamento ocorre logo após as modificações, e o usuário não precisa ficar se lembrando de clicar no botão salvar a cada poucos minutos. Aumentando assim a produtividade do trabalho e erradicando a possibilidade de o usuário perder um arquivo por não o ter salvo em tempo hábil.

Preço: a maioria desses serviços é gratuita, mas possui uma capacidade de armazenamento limitada. Escritórios, ou mesmo pessoas físicas com muitos anos de vida digital, geralmente precisam de mais espaço para o backup de todos os arquivos.

Dropbox, One Drive e similares são serviços oferecem soluções pagas que ampliam essa margem. Contudo, exige-se uma assinatura mensal. Ou seja, para uma empresa ou pessoa solicitar esses serviços, ela precisa encaixar esse investimento nas despesas fixas do mês.

Brechas na segurança: como já dito, os dados são criptografados e fracionados entre diversos servidores. Ainda assim, qualquer sistema on-line esta suscetível a hackers. A maior preocupação é com o vazamento de informações pessoais, como nome, e-mail e número do cartão de crédito dos clientes. Essa é uma responsabilidade tanto das companhias quanto dos próprios usuários, que devem escolher senhas fortes e adotar boas práticas na rede.

Legislação: as leis que penalizam crimes cibernéticos são diferentes em cada nação. Como os servidores de cloud computing ficam espalhados pelo globo, isso pode dificultar muito as investigações e o julgamento de piratas envolvidos em roubos de dados.

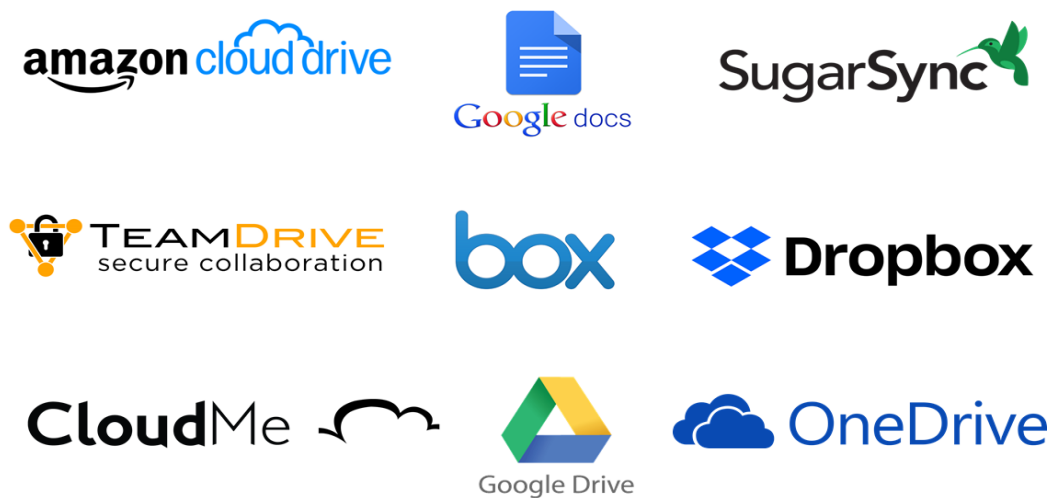
Há dúvidas, por exemplo, sobre qual país deveria dar seguimento ao processo: o local onde o servidor está situado? A região de onde partiu o ataque? Ou seria o país em que o usuário estava quando fez upload dos documentos?

Conexão: infelizmente, para o usuário ter um bom desempenho com banco de dados em nuvem, ele precisa dispor de uma conexão de internet razoável, por exemplo, usuário com conexões de telefonia móvel (3G), possuem dificuldades para acessar arquivos mais pesados, como vídeos e outros, dependendo da região onde se encontram.

3 APLICAÇÕES

Muitas empresas de grande porte já mudaram para a nuvem. Suas necessidades de segurança são resolvidas por provedores, o que permite que eles se concentrem mais nos negócios. Até mesmo alguns dos departamentos governamentais com os dados mais confidenciais têm dados na nuvem. A nuvem permite que eles mantenham seus requisitos de conformidade com facilidade. Empresas como a Amazon, VMware, Microsoft e Google, Netflix, Nasa e o Youtube, são empresas que já utilizam desses serviços. Podemos citar uma aplicação bastante interessante disponibilizado pela Google, por exemplo, que é o Google docs, o mesmo permite a edição de arquivos de texto, elaboração de slides e planilhas eletrônicas, contribuindo assim com a elaboração de documentos que podem ser acessados em qualquer lugar do mundo e a qualquer hora.

FIGURA 4- EMPRESAS QUE UTILIZAM O BDN



4 DESAFIOS E SEGURANÇA

FIGURA 5- SEGURANÇA DE UM BDN



a computação em nuvens traz inúmeros desafios quanto à segurança dos dados incorporados a esses sistemas. As informações trafegadas nas redes, principalmente corporativas, possuem algum tipo de valor e devem ser resguardadas, exigindo assim um nível de segurança eficiente nos sistemas que gerenciam a rede. Essas informações podem ser alvos de diversos problemas de infraestrutura física ou de ataques virtuais. De modo geral, a segurança de recursos de informação possui três componentes primordiais:

Integridade de dados - que significa proteger contra alteração indevida ou danos.

Confidencialidade - que significa proteger os dados contra a exposição a pessoas não autorizadas.

Disponibilidade - que significa proteger os dados contra a interferência com os meios de acesso aos recursos.

Naturais: As ameaças naturais são as providas pela própria natureza, como terremotos, enchentes, tempestades que acabam muitas das vezes por levar a perdas físicas e prejuízos financeiros. Sabemos que não existe segurança 100% ou imunidade a desastres, mas é possível se prevenir ao detectar as ameaças com a antecedência necessária para tomar providências que impeçam impactos na rotina de trabalho. Para isso, é necessário conhecer e diagnosticar os riscos que o data center pode sofrer. O local do data center é o primeiro ponto a ser avaliado. O entendimento dos riscos de interrupção por agentes externos como acidentes, fenômenos naturais e manifestações populares devem ser considerados nesta etapa.

Involuntárias: Já as ameaças involuntárias, são as que não tem o intuito de causar nenhum dano, por exemplo, falta de energia, porem muitas das vezes acabam desencadeando algum tipo de problema ao dado armazenado.

Voluntarias: as ameaças voluntárias são realizadas por pessoas mal-intencionadas, que querem tirar proveito das outras, como espiões e crackers. Esta ameaça é umas das mais perigosas pois os seus causadores muitas das vezes não estão interessados em causar danos aos dados em si, mas sim em roubar informações muitas das vezes confidenciais

5 CONCLUSÃO

O serviço de cloud data bases vem passando por desafios e consequentemente evoluções desde de que foi concebido, e a medida do seu desenvolvimento vem ganhando espaço no mercado principalmente com as grandes empresas, embora ainda existam barreiras a serem superadas referentes ao aperfeiçoamento principalmente no que diz respeito a manutenção e segurança e a deste serviço ele tem se mostrado de certo modo confiável

Entretanto importante que ocorra um amadurecimento de pontos chave, superando os atuais desafios que a computação em nuvem enfrenta, com atenção especial à segurança de dados, padronização e modelos de negócios que possam assegurar um comércio justo e seguro para todas as partes envolvidas.

Por fim, foram discutidos alguns desafios importantes no gerenciamento de dados, tais como escalabilidade, segurança, qualidade do serviço de dados, entre outros. É importante ressaltar que, várias soluções, existentes em outros modelos computacionais, que solucionem ou atenuem estes desafios, podem ser aplicadas em ambientes de computação em nuvem. Estes desafios geram oportunidades de pesquisa que devem ser superados, de forma que computação em nuvem seja amplamente aceita e utilizada por todos.

REFERENCIAS

SOUZA, Flávio R. C.; MOREIRA, L. O.; MACÊDO, J. F.; MACHADO, J. C. Gerenciamento de Dados em Nuvem: **Conceitos, Sistemas e Desafios 1 2**. Universidade Federal do Ceará (UFC).

ARRUDA, D. F.; MOURA JÚNIOR, J. A. F. Banco de Dados em Nuvem: **Conceitos, Características, Gerenciamento e Desafios**. Faculdade de Ciência e Tecnologia de Caruaru - Universidade de Pernambuco (UPE) -Caruaru –PE-Brasil Luís Alexandre Tavares Silva Faculdade de Ciência e Tecnologia de Caruaru -Universidade de Pernambuco (UPE) -Caruaru –PE-Brasil.

<https://www.tca.com.br/blog/vantagens-e-desvantagens-do-armazenamento-na-nuvem/>

<https://www.projuris.com.br/o-que-quer-dizer-banco-de-dados-na-nuvem/>

<https://www.ipm.com.br/blog/historia-da-computacao-em-nuvem-como-surgiu-a-cloud-computing/>

<<https://www.santodigital.com.br/como-escolher-entre-o-google-cloud-sql-ou-cloud-spanner/>>

<<https://cloud.google.com/sql/?hl=pt-br>>

<https://www.portal.ufpr.br/tutoriais/tutoriais_normaliza/modelo_trabalho_academico.pdf>

<<https://www.portal.ufpr.br/normalizacao.html>>

<https://www.portal.ufpr.br/tutoriais_normaliza/modelo_monografia.pdf>