



BIG DATA EM NUVEM: uma pesquisa bibliográfica e comparativa entre bancos de dados e servidores locais e cloud

BIG DATA IN CLOUD: a bibliographic and comparative research between databases and local servers and cloud

Luciano Bérghamo^I
Vitor Coimbra de Oliveira^{II}

RESUMO

O objetivo do presente trabalho é apresentar um resultado conclusivo acerca do gerenciamento de banco de dados, no qual pretende ampliar o entendimento coletivo sobre as formas de administrar bancos de dados locais e em nuvem, além de dissertar sobre as novas tecnologias que irão influenciar a forma como as empresas tratam os dados armazenados em seus bancos de dados. O volume de dados armazenados no mundo digital cresce a cada ano, conforme pôde comprovar o CEO da Google em 2010, quando afirmou que a cada dois dias a sociedade mundial gerava a mesma informação que gerou desde os primórdios de sua existência até 2003. Este aumento constante no volume de dados mudou a forma como armazenamos e processamos estes dados, despertando o interesse dos autores deste artigo pelo estudo da tecnologia de computação em nuvem. A Pesquisa Descritiva foi a metodologia escolhida para este trabalho. Serão coletados dados através do conhecimento coletivo disponível para apresentar os fatos através de uma análise detalhada do problema.

Palavras-chave: Banco de dados. Computação em Nuvem. Armazenamento. Segurança.

ABSTRACT

The present work is presented as a conclusive result about database management, in which it intends to broaden the collective understanding of the ways of administering local and cloud databases, as well as to discuss the new Technologies that will influence the way companies treat data stored in their databases. The volume of data stored in the digital world grows every year, as it was able to prove the CEO of Google in 2010, when he stated that every two days the world society generated the same information that generated from the beginnings of its existence until 2003. This constant increase in data volume has changed the way we store and process this data, arousing the interest of the authors of this article by the study of cloud computing technology. Descriptive research was the methodology chosen for this work. Data will be collected through the collective knowledge available to present the facts through a detailed analysis of the problem.

Keywords: Database. Cloud Computing. Storage. Security.

^I Prof. Me. do Centro Universitário Claretiano de Batatais – São Paulo – Brasil. E-mail: luciano@claretiano.edu.br

^{II} Estudante do Curso de Pós-Graduação em Banco de Dados do Centro Universitário Claretiano de Batatais – São Paulo – Brasil. E-mail: vitorcoimbra@outlook.com



Data de submissão do artigo: 23/06/2019.

Data de aprovação do artigo: 01/10/2019.

DOI:

1 INTRODUÇÃO

A computação em nuvem está deixando de ser uma tendência e se tornando uma realidade presente nas grandes empresas, que aplicam tecnologias em nuvem por exemplo em sua infraestrutura, no armazenamento de backup de dados e no desenvolvimento de *softwares*.

As organizações alteram suas estratégias com mudanças de ambiente e de processos e isto exige uma infraestrutura de tecnologia da informação (TI) flexível, que permita a otimização de recursos e que possibilite o crescimento da organização sem grandes interrupções (VERAS, 2015).

Outro fator importante é o aumento significativo no volume de dados armazenados pelas empresas a cada ano. O investimento no armazenamento destes dados de forma segura é cada vez mais necessário.

Há uma tendência entre as grandes empresas de analisar uma quantidade inédita de dados digitais, que terá um grande impacto nos processos de negócios e nas tomadas de decisões. Este fenômeno é conhecido como *Big Data* e tem sua importância comparada à popularização da internet (TAURION, 2013).

Mediante estes fatos, justifica-se a necessidade deste estudo para apresentar as diferenças entre a utilização de uma infraestrutura em nuvem, comparado à implementação de uma infraestrutura local, apresentando algumas vantagens econômicas e de segurança da Computação em nuvem.

A tecnologia em nuvem oferece vantagens econômicas significativas, pois limita o investimento inicial em infraestrutura de computação e do custo total da propriedade (ELSMARI, 2011).

2 COMPUTAÇÃO EM NUVEM: A TI COMO SERVIÇO

Assim como aconteceu com o setor de energia elétrica, a *TI* será fornecida como um serviço público em breve, com a computação em nuvem. Este novo método de fornecer *TI* permite separar o ambiente de software do hardware e criar uma infraestrutura mais flexível. Segundo Veras (2015), isto tornou os *datacenters* mais eficientes, aumentou a sua disponibilidade e permitiu criar serviços do tipo “pague-conforme-crescer”, com preços cada vez mais competitivos.

A nível mundial, existem exemplos de sucesso de migração para tecnologia em nuvem desde 2008. Segundo a Netflix (2016), provedora mundial de filmes e séries que são transmitidos para os assinantes do serviço através da internet, a sua jornada de migração para sistemas em nuvem começou em agosto de 2008, quando sua base de dados se corrompeu de forma grave, impossibilitando a empresa de enviar os DVDs para os clientes por 3 dias.

Este incidente fez a Netflix perceber que precisava se distanciar do armazenamento local em seu datacenter e seguir em direção ao armazenamento em nuvem, para suportar o crescimento acelerado da plataforma com a disponibilidade, integridade dos dados e confiabilidade necessárias para a continuidade dos serviços fornecidos. Foi definido o provedor de nuvem e, após sete anos de trabalho, a empresa completou a migração de todos seus serviços para nuvem e fechou seu último datacenter.



No Brasil, o “Reclame Aqui” (conhecido serviço de defesa do consumidor e pesquisa online de reputação de marcas) é outro caso de sucesso. Segundo a empresa Rivendel Tecnologia, o Reclame Aqui migrou sua estrutura tecnológica para a nuvem em apenas quatro meses, com o objetivo de garantir a disponibilidade no acesso que o site não fique indisponível, mesmo com uma grande quantidade de acessos simultâneos. (MANDIC, 2019).

O crescimento na utilização de bancos de dados em nuvem reforça mais ainda a importância da Segurança da Informação. Para que uma empresa tenha bancos de dados seguros, é necessário se preocupar não somente com ataques e roubos realizados por pessoas de fora da organização. É preciso garantir também que crimes virtuais não sejam cometidos por membros da própria organização (GALVÃO, 2015).

2.1 Computação empresarial baseada em cliente/servidor

Com o surgimento das redes de computadores locais, a arquitetura padrão de *TI* das organizações passou a ter aplicações utilizando parcialmente os recursos dos servidores locais e uma parte dos recursos dos computadores dos usuários, distribuindo assim o processamento dos dados.

As organizações começaram a adquirir computadores servidores e de usuários (*desktops*), implantando uma infraestrutura de *TI* local para possibilitar armazenar os dados e utilizar as aplicações dentro da própria organização.

Estas redes locais conectavam servidores e *desktops* utilizando protocolos não padronizados, fornecendo recursos baratos para a sua configuração e atraindo as organizações, que passaram a priorizar a flexibilidade ao invés da utilização adequada dos recursos de *TI* (VERAS, 2015).

Interligar redes locais entre empresas distintas era uma tarefa cara e complexa, onde era quase sempre necessário existir recursos intermediários que gerenciavam a comunicação entre as redes internas de duas ou mais organizações.

Com o surgimento da *internet*, as organizações encontraram um meio barato para interligar suas diferentes redes, através de provedores de acesso que permitiram interligar sistemas de organizações distintas através da grande rede mundial (a *internet*).

Tratando-se de banco de dados, as organizações utilizavam *SGBD* (Sistemas Gerenciadores de Banco de Dados) centralizados, onde todo o processamento dos dados era realizado no servidor do Banco, o resultado era fornecido à uma aplicação e então apresentado ao usuário em seu computador, através desta aplicação. Com o avanço da tecnologia, os computadores aumentaram seu poder de processamento e o custo de aquisição caiu consideravelmente. Os *SGBD* passaram a utilizar também o poder de processamento disponível nos computadores clientes (GALVÃO, 2015).

2.2 A computação em nuvem e a mudança na infraestrutura de *TI*

O conceito de computação em nuvem ainda se aperfeiçoa na medida em que é disseminado entre as organizações.

A tendência atual em gerenciamento de dados é totalmente voltada para a *internet*, onde é possível gerenciar uma quantidade muito grande de dados de forma dinâmica e confiável. A computação em nuvem oferece uma infraestrutura tecnológica, além de plataforma e softwares como serviços, utilizando a *internet* (ELSMARI, 2011).

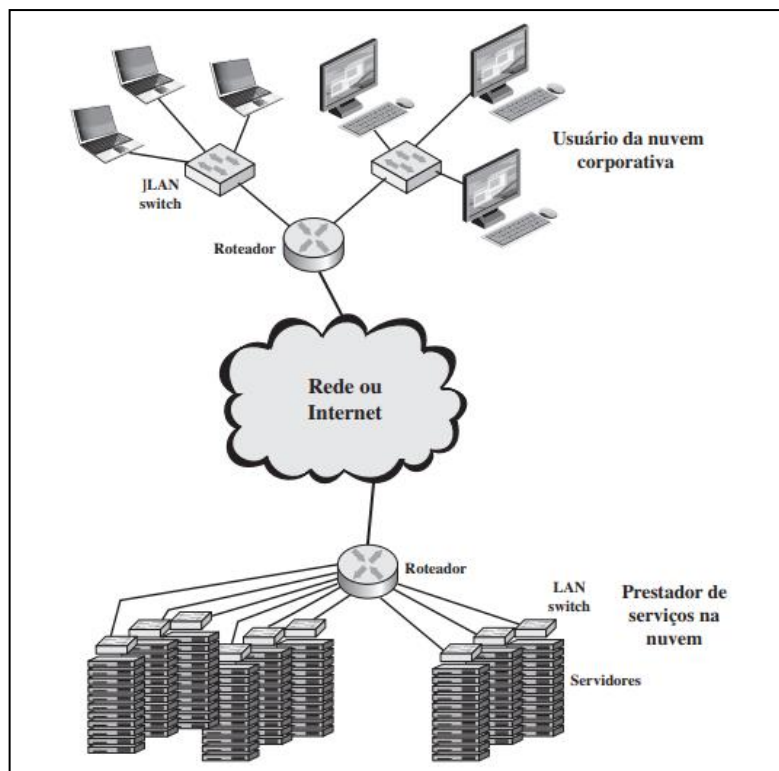


Utilizar a computação em nuvem exige uma quebra de paradigma na forma de utilizar a *TI*, onde a organização deixa de adquirir equipamentos para o seu Datacenter e passa a adquirir serviços de um provedor. Isto permite utilizar medidores de nível de serviço (como disponibilidade e desempenho) para dar ao cliente um maior controle sobre o serviço de nuvem contratado. (VERAS, 2015).

Na utilização de serviços em nuvem, a organização mantém uma rede local de computadores interligados (estações de trabalho), conectados à equipamentos roteadores de tráfego que fazem a ponte entre a rede interna da empresa e o provedor de serviços em nuvem. (STALLINGS, 2015).

A figura 1 demonstra a estrutura típica de uma organização que utiliza serviços em nuvem:

Figura 1 - Contexto da computação em nuvem



Fonte: Stallings (2015, p. 399)

Atualmente, a computação em nuvem é fornecida como serviço em três tipos: pública, privada e híbrida. Na contratação de uma nuvem pública, a infraestrutura do provedor de nuvem é compartilhada entre dois ou mais clientes, o que possibilita por exemplo que uma mesma aplicação seja utilizada coletivamente, reduzindo os custos.

Ao contratar um serviço de nuvem privada, o provedor irá fornecer uma infraestrutura não compartilhada, que traz a garantia de que todos os recursos alocados em nuvem pela organização serão de acesso e utilização exclusivos dela.

Em se tratando de nuvem híbrida, o provedor fornece uma infraestrutura mista entre pública e privada, que deixa de possuir algumas características da nuvem pública. Este tipo



permite que o cliente utilize a nuvem pública para executar aplicações e armazenar dados menos sensíveis, e utilize a nuvem privada para armazenar dados confidenciais e aplicações da organização que não podem ser compartilhadas. O quadro abaixo apresenta os três modelos de serviços para fornecer computação em nuvem.

Quadro 1 - Modelos de Serviço de computação em nuvem

Modelo	Descrição
Infraestrutura como serviço (IaaS)	Neste modelo, O provedor oferece toda a infraestrutura necessária para a organização, para armazenamento de dados e processamento de informações. A organização não se torna responsável por administrar a infraestrutura física, ficando responsável apenas por gerenciar recursos como: armazenamento, aplicações e máquinas virtuais.
Plataforma como serviço (PaaS)	O provedor fornece recursos para os desenvolvedores executarem suas aplicações e disponibilizá-las em um ambiente em nuvem para utilização de seus clientes.
Software como um serviço (SaaS)	As aplicações são hospedadas em nuvem e fornecidas pelo provedor como serviços para seus clientes. O gerenciamento de toda a infraestrutura e dos recursos (como controle de armazenamento) é de responsabilidade do provedor, de forma transparente para o usuário dos serviços contratados.

Fonte: Veras (2015)

Estes três modelos podem interagir entre si, onde a *IaaS* fornece recursos (como hardware e software) para a *PaaS*, que por sua vez fornece tecnologias e ferramentas para que o desenvolvedor possa implementar e executar seus serviços, que serão posteriormente disponibilizados para seus clientes através da *SaaS*. Apesar desta interação, o provedor não precisa obrigatoriamente oferecer estes três modelos para os seus consumidores (VERAS, 2015).

Com relação à arquitetura, as soluções tradicionais de bancos de dados podem gerar gargalos de desempenho em se tratando de computação em nuvem, pois podem apresentar falhas parciais e problemas de sincronismo global. Com isto surgem novas arquiteturas, como por exemplo os sistemas de bancos de dados *P2P* (ponto-a-ponto), que garantem maior escalabilidade, alta segurança contra-ataques e gerenciamento de dados descentralizado (ELSMARI, 2011).

3 PROCEDIMENTOS METODOLÓGICOS

A Pesquisa Descritiva foi a metodologia escolhida para este trabalho. Foram coletados dados através do conhecimento coletivo disponível para apresentar os fatos através de uma análise detalhada do problema.

3.1 Identificação do crescimento da computação em nuvem

O uso pessoal da computação em nuvem se popularizou nos últimos anos, com a utilização de aplicações que armazenam fotos, dados e documentos pessoais em nuvem,



trazendo uma maior segurança para o usuário de que seus dados estarão menos propícios à perda, quando comparado ao armazenamento em mídias físicas (como HD, DVD e pen drive). Porém o forte crescimento desta tecnologia de armazenamento nas grandes organizações despertou o interesse dos autores na realização de uma pesquisa descritiva para definir como a computação em nuvem e o Big Data transformam o modelo de negócios das organizações.

3.2 Investimento atual na computação em nuvem

Foi realizada uma revisão bibliográfica para indicar as empresas que estão investindo nesta tecnologia nos últimos anos, e que já migraram ou estão migrando sua infraestrutura para a nuvem. Com isto é possível mensurar os impactos e os benefícios que a utilização de serviços em nuvem traz para uma organização.

4 DISCUSSÃO SOBRE A SEGURANÇA DA COMPUTAÇÃO EM NUVEM

Em termos de segurança da infraestrutura e das informações, a computação em nuvem possui características muito similares aos ambientes locais de TI. Apesar deste fato, o ambiente em nuvem requer uma maior preocupação, pois a organização que utiliza o ambiente perde consideravelmente o controle sobre todos os recursos e serviços fornecidos pelo provedor de nuvem, mas precisa manter a responsabilidade sobre a privacidade e as políticas de segurança do ambiente (STALLINGS, 2015).

4.1 Os riscos de um ambiente em nuvem

Alguns riscos como: vazamento ou perda de dados, roubo da conta empresarial de nuvem e uso abusivo e malicioso da computação em nuvem são preocupantes tanto para o provedor, quanto para a organização que está contratando os serviços.

Existem diversas maneiras de comprometer os dados de uma organização. Armazenar os dados em mídias não confiáveis e a perda de uma chave criptográfica podem resultar em uma destruição definitiva dos dados.

Desta forma, partes não autorizadas devem ser impedidas de acessar os dados sensíveis da organização. Um meio simples para aumentar a segurança e solucionar este problema é encriptar e proteger a integridade dos dados que são trafegados e não fornecer as chaves de autenticação ao provedor de nuvem (STALLINGS, 2015).

Nos Estados unidos foi criado o *National Institute of Standards and Technology* (NIST), uma agência governamental para prover padrões de tecnologia que ampliem a segurança econômica virtual. No âmbito da computação em nuvem, a norma NIST SP-800-14 estabelece diversas orientações quanto à segurança e privacidade (STALLINGS, 2015):

Quadro 2 - Orientações NIST SP-800-14

Orientação	Detalhamento
Governança	Cria práticas organizacionais que abordam políticas de segurança, procedimentos e padrões utilizados no desenvolvimento de aplicações, criando meios de auditoria e ferramentas para garantir as boas práticas, que devem ser seguidas pela organização durante todo o período em que utilizar o ambiente em nuvem.



Gerenciamento de acesso	Garante que os métodos de autenticação, acesso e autorizações sejam adequados à necessidade da organização que utiliza o ambiente em nuvem.
Proteção dos dados	Considera riscos do ambiente em nuvem compartilhado, como cruzamento dos dados com outros clientes da nuvem que possam oferecer alto risco à organização.
Disponibilidade	O Provedor de nuvem deve oferecer disponibilidade do ambiente, backups e planos de recuperação de desastres baseados nas cláusulas contratuais, com o objetivo de atender aos requisitos de continuidade da organização e para garantir que em caso de indisponibilidade dos serviços, as operações críticas da organização possam ser reestabelecidas imediatamente.

Fonte: Stallings (2015)

4.2 A Segurança como um serviço

Além dos três modelos de computação em nuvem citados anteriormente, o provedor ainda pode fornecer um pacote de serviços de segurança, através da Segurança como um serviço (SecaS). Este pacote pode ser fornecido pelo provedor em conjunto com a *SaaS* e pode incluir ferramentas de segurança como antivírus, autenticação em duas etapas, segurança de e-mail e detecção de invasões.

5 CONSIDERAÇÕES FINAIS

Este artigo apresentou as novas tecnologias que estão influenciando a forma de armazenamento de dados nas empresas, onde se observou uma tendência muito forte de utilização do ambiente em nuvem, seja privado ou compartilhado, para garantir a continuidade da operação empresarial e a segurança dos dados da organização.

Conclui-se que a computação em nuvem está alterando significativamente a estrutura da organização, pois ela deixa de possuir uma estrutura de *datacenter* própria e passa a se tornar cliente de provedores terceirizados, responsáveis por fornecer a infraestrutura que a organização precisa para armazenar seus dados.

Com isto, cria-se uma relação cliente e fornecedor, que possibilita o fornecimento de serviços do tipo “pague conforme a sua necessidade”. Isto traz uma economia para as organizações, pois não precisam mais preocupar-se em manter uma infraestrutura própria, e permite que elas contratem serviços em nuvem de acordo com a necessidade real da instituição.

O Big Data já não é mais um conceito, mas faz parte da nossa realidade: em março de 2013, a Apple registrou 25 bilhões de downloads de aplicativos em sua Apple Store. Entre 78% e 84% dos consumidores se baseiam nas mídias sociais para tomar decisão de compra de um produto. Ou seja, as informações fornecidas sobre um produto se tornaram mais importantes do que o próprio produto (TAURION, 2013).

Hoje não pensamos se vamos utilizar a computação em nuvem, mas pensamos em como e quando iremos migrar nossa infraestrutura empresarial para a nuvem. Toda a mobilidade fornecida por esta nova tecnologia está revolucionando o mundo atual: hoje a maioria dos consumidores opta por adquirir dispositivos móveis ao invés de computadores de mesa.



Atualmente as empresas ainda estão compreendendo o verdadeiro significado do *Big Data*, buscando definir as tecnologias envolvidas para então definir como irão aplicar este fenômeno em sua realidade. Dados são a matéria-prima da nova Revolução Industrial, matéria-prima esta que cria conhecimento sobre produtos e direcionam os consumidores, mudando o modelo de negócio das empresas.

Um exemplo do poder do Big Data é uma grande rede varejista americana, que analisou um volume imenso de dados de seus clientes e descobriu um padrão de compras e comentários em redes sociais que definiu uma alta possibilidade de uma cliente estar grávida. Com base nestes dados, enviou correspondências para a residência destes clientes (TAURION, 2013).

Através do *Big Data* e do *Business Intelligence* (BI) é possível identificar padrões de consumo, analisar hábitos dos consumidores e direcionar por exemplo campanhas de marketing para o público-alvo correto. Todo este novo modelo de negócio baseado no comportamento dos consumidores não seria possível sem a tecnologia de computação em nuvem.

REFERÊNCIAS

ELSMARI, Ramez. **Sistemas de Banco de Dados**. 6.ed. São Paulo: Pearson Addison Wesley, 2011.

GALVÃO, Michele da Costa. **Fundamentos em Segurança da Informação**. São Paulo: Pearson Education do Brasil, 2015.

MANDIC. **Para crescer, Reclame Aqui migra para a nuvem da Amazon AWS**. Disponível em: <https://www.mandic.com.br/cases/migracao-on-premises-para-a-nuvem-da-amazon-reclame-aqui/>. Acesso em: 26 mar. 2019.

NETFLIX. **Concluindo a migração para a nuvem**. 2016. Disponível em: https://media.netflix.com/pt_br/company-blog/completing-the-netflix-cloud-migration Acesso em: 25 mar. 2019.

STALLINGS, William. **Criptografia e Segurança de Redes: Princípios e práticas**. 6.ed. São Paulo: Pearson Education do Brasil, 2015.

TAURION, Cezar. **Big Data**. Rio de Janeiro: Brasport, 2013.

VERAS, Manoel. **Computação em Nuvem: Nova Arquitetura de TI**. Rio de Janeiro: Brasport, 2015.