

MARINHA DO BRASIL
DIRETORIA DE ENSINO DA MARINHA
CENTRO DE INSTRUÇÃO ALMIRANTE WANDENKOLK

CURSO DE APERFEIÇOAMENTO AVANÇADO EM
SEGURANÇA DAS INFORMAÇÕES E COMUNICAÇÕES

TRABALHO DE CONCLUSÃO DE CURSO

**ESTUDO DA VIABILIDADE DO USO DE SERVIÇOS DE NUVENS PARA BANCO
DE DADOS ADMINISTRATIVOS DA MARINHA DO BRASIL**



1º Ten (QC-CA) RAFAEL GOMES DOS SANTOS

Rio de Janeiro
2020

1º Ten (QC-CA) Rafael Gomes dos Santos

ESTUDO DA VIABILIDADE DO USO DE SERVIÇOS DE NUVENS PARA BANCO DE
DADOS ADMINISTRATIVOS DA MARINHA DO BRASIL

Monografia apresentada ao Centro de Instrução
Almirante Wandenkolk como requisito parcial à
conclusão do Curso de Aperfeiçoamento Avançado em
Segurança das Informações Digitais

Orientadores:

CF (RM1-EN) William Augusto Rodrigues de Souza

1ºTen(RM2-T) Rafael de Souza Nogueira

CIAW
Rio de Janeiro
2020

SANTOS, Rafael Gomes dos.

Estudo Da Viabilidade Do Uso De Serviços De Nuvens Para Banco De Dados Administrativos Da Marinha Do Brasil / Rafael Gomes dos Santos. - Rio de Janeiro, 2020.

56 f

Orientadores: CF (RM1-EN) William Augusto Rodrigues de Souza; 1ºTen(RM2-T) Rafael de Souza Nogueira.

Monografia (Curso de Aperfeiçoamento Avançado em Segurança da Informação e Comunicações). - Centro de Instrução Almirante Wandenkolk, Centro de Coordenação de Pós-Graduação, 2020.

1. Banco de dados 2. Armazenamento em nuvem 3. Dados administrativos I. Centro de Instrução Almirante Wandenkolk. Centro de Pós-Graduação Avançada. II. Título.

FOLHA DE APROVAÇÃO

1ºTen (QC-CA) Rafael Gomes dos Santos

ESTUDO DA VIABILIDADE DO USO DE SERVIÇOS DE NUVENS PARA BANCO DE
DADOS ADMINISTRATIVOS DA MARINHA DO BRASIL

Monografia apresentada ao Centro de Instrução Almirante Wandenkolk como requisito parcial à conclusão do Curso de Aperfeiçoamento Avançado em Segurança das Informações e Comunicações

Aprovada em 24 de março de 2020

Banca Examinadora:

Gian Karlo Huback Macedo de Almeida, Capitão de Mar e Guerra (RM1-EN) - Marinha do
Brasil

William Augusto Rodrigues de Souza, Capitão de Fragata (RM1-EN) - Marinha do Brasil

Rafael De Souza Nogueira, Primeiro-Tenente (RM2-T) - Marinha do Brasil

*Dedico este trabalho a todos os dinossauros
que sobreviveram ao impacto*

AGRADECIMENTOS

Agradeço a minha família pelo apoio emocional durante o curso, em especial a minha irmã, Sílvia, pelo apoio especial na produção deste trabalho, trabalhando com a organização dos capítulos, ideias e com a revisão dos textos.

Agradeço aos meus orientadores, CF William e ao 1º Ten Rafael Nogueira, pelo apoio no desenvolvimento do projeto, oferecendo materiais adicionais e direcionamentos na execução do trabalho.

Gostaria agradecer especialmente ao coordenador do curso CMG Huback pelo tratamento com a turma e no atendimento das demandas individuais dos alunos. O trabalho do senhor foi essencial para garantir o sucesso de todos os alunos no curso.

*“Things that I promised myself
fell apart
but I found my heart.”
(Burt Bacharach; Promises, Promises)*

RESUMO

A computação em nuvem é uma tendência entre grandes organizações. As organizações se aproveitam das vantagens desse modelo de computação para reduzir custos de serviços de armazenamento e processamento de dados. Os Bancos de Dados são desafios de processamento conhecidos da computação e sua utilização é ampla em ambientes corporativos. Os Bancos de Dados relacionais permitem a execução das ferramentas de gerência, e portanto, são usadas como ferramentas de auxílio na administração de recursos. Visando a melhor oferta desses recursos no ambiente corporativo da Marinha do Brasil, propõe-se a migração dos modelos de gestão de dados administrativos para uma solução em nuvem. Dessa forma, é de extrema importância, avaliar os custos e características das nuvens comerciais e verificar sua aplicabilidade em relação as demandas dessa organização. Nesse estudo, foram avaliadas as ferramentas de segurança e os custos de implementação de três soluções comerciais: Amazon Web Services, Microsoft Azure e Google Cloud. Concomitantemente, foram avaliados os custos de manutenção de um sistema físico local com as mesmas características, chamado de Custo Total de Propriedade. Ao final do trabalho conclui-se que soluções em nuvem podem ser utilizadas para o armazenamento de informações cujo nível de sigilo seja baixo, pois estes apresentam ferramentas de segurança, tais como criptografia em múltiplos estágios e protocolos de conexão segura com nível de segurança comprovado. Os sistemas em nuvem, quando aplicados a sistemas de bancos de dados únicos, com características que atendem ao modelo de negócio da Marinha do Brasil, apresentaram menor custo se comparados a estimativas de Custo Total de Propriedade.

Palavras-chaves: Banco de Dados. Zona Desmilitarizada. Segurança de Dados. Custo Total de Propriedade.

LISTA DE ILUSTRAÇÕES

Figura 1 – Tipos de Bancos de Dados	12
Figura 2 – Servidores Lenovo ThinkSystem	18
Figura 3 – Dispositivo De <i>Storage</i> Western Digital	18
Figura 4 – Esquema de uma rede segregada	19
Figura 5 – Distinção dos segmentos de rede monitorados	20
Figura 6 – Esquema geral de funcionamento de uma <i>screened subnet</i>	21
Figura 7 – Firewall Cisco	22
Figura 8 – Esquema de um <i>firewall</i>	22
Figura 9 – Esquema de funcionamento de um proxy	23
Figura 10 – Soluções de nuvens comerciais	28
Figura 11 – Logotipo do sistema de apoio à gestorias Quaestor/Marinha do Brasil	33
Figura 12 – Tendência de acessos ao sistema	35
Figura 13 – Calculadora para determinação de preços	37
Figura 14 – Modelos de Contração dos serviços de Bancos de Dados do Microsoft Azure	39
Figura 15 – Padrão para cobrança dinâmica baseada no uso de recursos de processamento	39
Figura 16 – Calculadora para determinação de preços	41
Figura 17 – Calculadora para determinação de preços	41
Figura 18 – Modelo de custos envolvidos na implementação de serviços de bancos de dados	43
Figura 19 – Aplicativo Web - Calculadora de TCO da Amazon Web Services	44
Figura 20 – Aplicativo Web - Calculadora de TCO da Microsoft Azure	45
Figura 21 – Comparação entre os custos de manutenção dos serviços em nuvem e físico local	47

LISTA DE QUADROS

Quadro 1 – Soluções de Segurança implementadas pelos serviços de armazenamento em nuvem	29
Quadro 2 – Soluções para migração de bancos de dados implementados pelos servidores de armazenamento em nuvem	31
Quadro 3 – Soluções para Escalabilidade e disponibilidade implementados nos servidores de armazenamento em nuvem	32
Quadro 4 – Características das soluções em nuvem mínimas desejadas	36
Quadro 5 – AWS - Custos por unidade de contratação	38
Quadro 6 – Especificações computacionais mínimas dos planos premium da Microsoft Azure	40
Quadro 7 – Microsoft Azure - Custos de produtos cobrados mensalmente	40
Quadro 8 – Google CloudSQL - Custos por unidade de contratação	42
Quadro 9 – AWS - Resumo da Discriminação do Custo Total de Propriedade	44
Quadro 10 – Microsoft Azure - Resumo da Discriminação do Custo Total de Propriedade	46
Quadro 11 – Comparativo entre custos das soluções	47

LISTA DE ABREVIATURAS E SIGLAS

DMZ	Desmilitarized Zone - Zona Desmilitarizada
TCO	Total cost of ownership - Custo total de propriedade
AWS	Amazon Web Services
OM	Organização Militar
SQL	Structured Query Language
NoSQL	Not Only SQL
TI	Tecnologia da Informação
VPN	Virtual Private Network

SUMÁRIO

1	INTRODUÇÃO	10
1.1	Apresentação	11
1.2	Justificativa	12
1.3	Relevância	13
1.4	Objetivos	13
1.4.1	Objetivo Geral	13
1.4.2	Objetivos Específicos	13
1.5	Metodologia	13
1.5.1	Limites e Escopo	14
1.5.2	Visão Geral do Trabalho	14
2	FUNDAMENTOS TEÓRICOS	16
2.1	Introdução	16
2.2	Definições e Fundamentos Teóricos	16
2.2.1	Banco de Dados	16
2.2.1.1	Estrutura de dados	17
2.2.1.2	Solução de armazenamento	17
2.2.2	Redes de Comunicação	19
2.2.2.1	Redes segregadas	19
2.2.2.2	DMZ - Zona desmilitarizada	20
2.2.2.3	Fronteiras das redes	21
2.2.2.4	Proxy	22
2.2.3	Segurança da Informação	23
2.2.3.1	Criptografia	24
2.2.4	Computação em nuvem	26
2.3	Estado da Arte	27
2.3.1	Bancos de Dados	28
2.3.1.1	Clientes	28
2.3.1.2	Segurança de dados	29
2.3.1.3	Migração	31
2.3.1.4	Escalabilidade e Disponibilidade	32
3	DESENVOLVIMENTO PRÁTICO DO TRABALHO	33
3.1	Escopo do Problema	33
3.1.1	Quantidade de acessos	34
3.1.2	Capacidade de Armazenamento	34

3.1.3	Transito de informações	35
3.1.4	Disponibilidade e Segurança	36
3.2	Soluções em nuvem	36
3.2.1	Amazon Web Services	36
3.2.2	Microsoft Azure	38
3.2.3	Google Cloud	41
3.3	Solução física	42
3.3.1	TCO da Amazon Web Services	43
3.3.2	TCO da Microsoft Azure	44
3.4	Comparativo entre soluções	46
4	CONCLUSÕES	49
	REFERÊNCIAS	50

1 INTRODUÇÃO

A evolução humana, da ciência e da tecnologia é marcada pelo registro de informações. Classificam-se como informações todas as observações do ambiente em que estamos inseridos (VOURAKIS, 2017).

O esforço para registro de eventos e de informações se justifica pela capacidade de apresentar esses dados aos demais. Esse comportamento, observado por Laignier e Fortes (2009), resulta na universalização da informação e torna-se um meio eficiente para a transferência e conservação do conhecimento.

Como descrito por Queiroz (2004), o desenvolvimento da escrita e do papel facilitou a codificação e interpretação das informações, porém está limitada àqueles detentores do conhecimento. Neste contexto o conhecimento é a informação mais valiosa e, conseqüentemente, mais difícil de gerenciar no mundo corporativo. Segundo Mendes (2015), atualmente é necessário que as empresas estejam preparadas para gerenciar o conhecimento, pois este fomenta a inovação, gera vantagens competitivas e sustentáveis.

Para que uma empresa possa aprender com o seu passado é necessário que ela esteja estruturada de forma para tal. Apenas desta maneira o aprendizado ocorrerá de maneira natural, o que agregará valor aos seus produtos e serviços. Assim, o armazenamento da informação reflete o estado do desenvolvimento tecnológico.

As empresas são responsáveis em criar, organizar e gerenciar o conhecimento que necessitam. Essa criação consiste em obter o conhecimento de fontes internas e externas à organização e armazená-lo para que seja possível sua utilização, pois a utilização desse conhecimento na empresa promove um processo de conversão do conhecimento por parte dos envolvidos que, por sua vez, permite uma inovação contínua do conhecimento (MENDES, 2015). Como observado por Paul (2009), a evolução tecnológica altera a forma como as informações são acessadas, suas dependências e custos, portanto, com a evolução dos circuitos integrados há um aumento da dependência da energia elétrica.

Além dos custos operacionais envolvidos no armazenamento de informações, deve-se observar que, muitas vezes, a maior parte dos conhecimentos que uma organização necessita para se manter competitiva, ela já possui, no entanto, estão, por diversos motivos, inacessíveis. Queiroz (2004) reitera que o desenvolvimento do armazenamento da informação é intimamente ligado com a necessidade da sua difusão. Nota-se ainda que, a partir do século XX, com o início da era da informação digital, a difusão da informação é facilitada com a introdução da internet e das redes de computadores.

A criação de um ambiente propício para identificar, criar, compartilhar e disseminar o conhecimento agrega valor à empresa e a coloca no rumo de atingir suas metas (MENDES, 2015).

1.1 Apresentação

A gestão de recursos em ambientes corporativos demanda um grande esforço coletivo de administradores empenhados a obter e manipular informações, indicativos sobre os recursos, projeções futuras e dados históricos. A manipulação dessas informações representa um reflexo da eficiência do sistema de gestão e permite avaliar a qualidade do trabalho executado. É importante que as informações sejam armazenadas de forma acessível, segura, rápida, íntegra e de alta disponibilidade para consultas futuras. (SILBERSCHATZ; STONEBRAKER; ULLMAN, 1991)

No passado, o armazenamento de informações era realizado de forma não ordenada, sendo carente de uma norma organizacional (VOURAKIS, 2017). A manipulação de grandes massas de informações é um dos problemas clássicos da Tecnologia da Informação e vem sendo estudado há mais de 50 anos (GREENSPOON, 2000), reconhecido pelo nome "*Database Management Systems*" ou "Sistemas de Bancos de Dados".

Atualmente, a implementação e manutenção dos servidores de armazenamento de bancos de dados é uma tarefa dispendiosa. Os custos da compra de servidores, gasto energético, manutenção, escalabilidade, resfriamento, pessoal, obsolescência e renovação devem ser considerados quando se avalia as soluções disponíveis no mercado. Para que o acesso aos bancos de dados, armazenados em sistemas computacionais, sejam eficientes, sua operação tem sido tratada tanto pela visão de *software*, quanto pela visão de *hardware* e instalações (IBM, 2019).

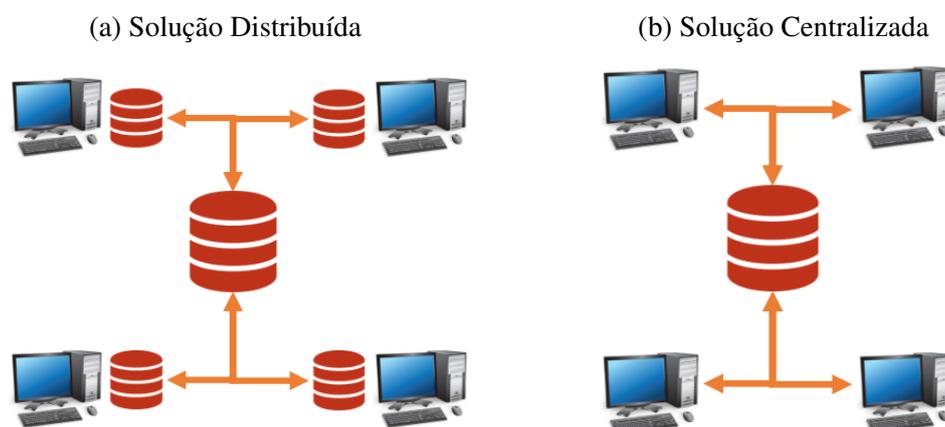
Em ambientes corporativos, os bancos de dados são acessados, modificados e auditados por diferentes usuários de forma simultânea, ainda que estes estejam distantes geograficamente. Entretanto, há a necessidade de garantir a disponibilidade dos bancos de dados aos usuários de forma transparente e diferentes soluções podem ser implementadas para cumprir tal objetivo (ORACLE, 2001). Entre as soluções, podemos citar:

- **Solução Distribuída:** Cada usuário possui uma cópia do banco de dados em sua Estação de Trabalho e os dados são sincronizados entre os usuários por meio de uma rotina periódica. Essa solução é representada na Figura 1a.
- **Solução Centralizada:** Todos os usuários acessam um servidor central que contém o banco de dados. As alterações são realizadas em tempo real. Essa solução é representada na Figura 1b.

Para ambas configurações apresentadas acima, há necessidade de uma rede de computadores para realizar a conexão entre os usuários. As redes de computadores, criadas desde 1969 (LEINER et al., 1997), permitem a transferência de informações em alta velocidade, mas estão susceptíveis à problemas de segurança e disponibilidade (CURTIN, 1997).

Recentemente, houve uma mudança no paradigma da gestão e manutenção de bancos de dados. Essa mudança aparece concomitante ao desenvolvimento de soluções de Computação em Nuvem (BUYA; BROBERG; GOSCINSKI, 2011). A Computação em Nuvem é uma

Figura 1 – Tipos de Bancos de Dados



Fonte – Elaborada pelo autor

ferramenta que permite a servidores espalhados geograficamente oferecer serviços de armazenamento e processamento de dados aos usuários, através de uma conexão com a internet, por meio de arrendamento. Esse tipo de solução atende a uma tendência mundial de virtualização de aplicações transparentes, ou seja, disponibilizar recursos complexos de uma maneira simples, para atendimento de uma heterogeneidade de usuários.

Na Marinha do Brasil, a gestão de municiação e caixa de economias é realizada através de uma solução centralizada, sendo o acesso aos bancos de dados realizado através do sistema Quaestor (MARINHA DO BRASIL, 2015). O sistema foi implementado em 2014 e possui um ciclo de vida estimado de aproximadamente 5 anos, dado o ciclo de vida dos servidores (NELSON; SIMEK; MASCHKE, 2009). Sendo assim, encontramos-nos em um ponto de guinada para a avaliação de custos e segurança em armazenamento de informações desses sistemas.

1.2 Justificativa

Em tempos de paz, as organizações militares podem aproveitar-se para desempenhar da melhor forma possível suas faculdades essenciais, apresentando exímio controle na gestão de recursos.

Quando tratamos de dados administrativos, é importante garantir a sua auditabilidade por longos períodos de tempo, assegurando-se de sua integridade e disponibilidade. A atualização do mecanismo para o armazenamento das informações dos sistemas de gestão da Marinha do Brasil acontecerá em um futuro próximo, portanto, é interessante discutir as possíveis soluções.

A gestão eficiente de recursos tem como foco principal a disponibilização de uma ferramenta de menor custo que atenda todos os requisitos de operação. Escolher uma ferramenta

para realizar a gestão de recursos poderá proporcionar uma redução de custos, aumento da produtividade e confiabilidade.

1.3 Relevância

O acesso aos dados administrativos é importante para o funcionamento das Organizações Militares. É responsabilidade dos órgãos envolvidos garantir organização, confiabilidade, segurança, confidencialidade, disponibilidade e acessibilidade a esses dados.

As soluções para o armazenamento de grandes bancos de dados podem ser comparadas pelos seus custos de implantação, manutenção, escalabilidade, ferramentas de segurança disponíveis, confiabilidade dos dados e disponibilidade.

A adoção de uma solução baseada em armazenamento em nuvem é uma tendência, a exemplo de grandes companhias, como o Facebook (MILLER, 2018).

1.4 Objetivos

Propõe-se com a execução desse trabalho, subsidiar dados que permitam atingir os seguintes objetivos:

1.4.1 Objetivo Geral

Como objetivo geral do trabalho procura-se verificar a viabilidade da substituição dos sistemas de armazenamento de bancos de dados administrativos (Gestorias de Municípios e Caixa de Economias) da Marinha do Brasil por sistemas baseados em clientes comerciais em nuvem.

1.4.2 Objetivos Específicos

São objetivos específicos do trabalho:

- Avaliar os custos operacionais de sistemas de armazenamento local de bancos de dados;
- Avaliar as ferramentas de segurança disponíveis para armazenamento de banco de dados;
- Identificar vantagens e desvantagens das soluções para armazenamento de bancos de dados.

1.5 Metodologia

A metodologia utilizada nesse trabalho divide-se em três etapas:

- a) **Pesquisa bibliográfica:** uma avaliação de diferentes fontes bibliográficas para identificar conceitos primordiais e os atributos comparativos para análise de custos e qualidade das soluções de armazenamento de bancos de dados;
- b) **Descrição das soluções que serão avaliadas:** uma pesquisa bibliográfica para levantamento das características comparativas atinentes às alternativas para o armazenamento de bancos de dados. Sendo estas: ferramentas de segurança disponíveis, ferramentas de migração, disponibilidade, capacidade de armazenamento, fluxo de informações e capacidade de atender múltiplos acessos. Nessa etapa, serão utilizadas as informações encontradas nas documentações disponíveis nas plataformas das soluções Microsoft Azure, Amazon Web Services e Google CloudSQL.
- c) **Análise quantitativa dos custos e vantagens de cada uma das soluções estudadas:** uma avaliação comparativa dos custos de manutenção de um sistema de bancos de dados que atenda os pré-requisitos dos sistemas administrativos de municiamento e caixa de economias da Marinha do Brasil. A avaliação dos custos será realizada para um período de tempo equivalente à vida útil dos sistemas físicos de armazenamento de bancos de dados. A aquisição dos custos das soluções em nuvem poderá ser realizada por ferramentas capazes de gerar orçamentos das soluções, tais como, as calculadoras de custos das soluções Microsoft Azure, Amazon Web Services e Google CloudSQL. Para a solução de armazenamento físico, os custos serão obtidos das ferramentas de avaliação dos custos de propriedade das mesmas soluções.

1.5.1 Limites e Escopo

A migração dos dados administrativos das plataformas de gestão da Marinha do Brasil, para o meio digital, concorre com os desafios resultantes do armazenamento de grandes massas de dados: segurança, custos de implantação, manutenção e escalabilidade.

A proposta desse trabalho consiste em realizar uma avaliação comparativa quantitativa dos custos e qualitativa em termos de segurança entre a implementação de um sistema de servidor de arquivos local e as soluções baseadas em Computação em Nuvem comerciais: *Google CloudSQL*, *Amazon Web Services* e *Microsoft Azure* para armazenamento de dados administrativos das plataformas de gestão de municiamento e caixa de economias da Marinha do Brasil.

1.5.2 Visão Geral do Trabalho

O trabalho será dividido em quatro partes:

- **Fundamentos Teóricos:** Exposição dos principais conceitos necessários para a discussão do trabalho executado, através da Pesquisa Bibliográfica;

- **Desenvolvimento Prático do Trabalho:** Definição das premissas do problema e características desejáveis e avaliação dos custos para implementação das soluções e medidas adicionais necessárias para a garantia do sigilo;
- **Conclusões:** Discussão final dos resultados obtidos no estudo de caso avaliado.

2 FUNDAMENTOS TEÓRICOS

2.1 Introdução

O seguinte capítulo tem como objetivo apresentar os conceitos necessários para o desenvolvimento do seguinte trabalho. Nele serão expostos definições, fundamentos teóricos, referências bibliográficas e estado da arte das tecnologias envolvidas no armazenamento em bancos de dados.

Como descrito por Constantini (2017), esse tipo de pesquisa preliminar deverá ser realizado de maneira descritiva onde serão avaliadas fontes confiáveis, tais como, livros, monografias, dissertações e artigos para tecer luz aos conceitos.

Esse capítulo será dividido em duas seções: "Definições e Fundamentos Teóricos", onde serão conceituados as principais definições elegidas importantes para a compreensão do trabalho; "Estado da Arte", onde serão avaliadas as características dos serviços para armazenamento em nuvem atuais.

2.2 Definições e Fundamentos Teóricos

Nessa seção serão apresentados, de forma reduzida, os conceitos necessários para a escolha de uma solução para armazenamento de dados:

2.2.1 Banco de Dados

Segundo Silberchatz, Korth e Sudarshan (2010), um banco de dados é uma coleção de dados inter-relacionados, representando informações sobre um domínio específico. São exemplos de banco de dados a lista telefônica e o sistema de controle de RH de uma empresa.

No caso das empresas, os bancos de dados geralmente são ferramentas computacionais organizadas para armazenar dados computacionais correlatos, conhecidos como objetos, registradores ou valores.

A organização dos dados é essencial para a garantia da velocidade de pesquisa e recuperação de informações mediante uma vasta massa de dados semelhantes. Em um caso ideal, os dados devem ser disponibilizados instantaneamente às operações que as requisitam.

Exemplificando, pode-se citar que as informações armazenadas pelos servidores de mídias sociais como o Facebook e Instagram estão organizadas em bancos de dados. Esses servidores armazenam informações de todos seus usuários como "*posts*", fotos, informações, dados de posicionamento, interesses, características de consumo entre outras (MONNAPPA, 2010). Entretanto, além do armazenamento de informações é necessário que se garanta que os dados possam ser disponibilizados de forma rápida ao usuário.

A gestão dos bancos de dados é realizada através de uma ferramenta computacional, o *Database Management System* (DBMS). Essa ferramenta faz uso de uma sequência de protocolos para identificar e organizar os dados. Esses protocolos identificam uma linguagem de programação que implementa funções de busca, inserção, deleção e modificação de dados. (TAKAI; ITALIANO; FERREIRA, 2005)

Por último, temos que conceituar um sistema de banco de dados como o conjunto de quatro componentes básicos: dados, *hardware*, *software* e usuários. Nas seções a seguir serão discutidos as características que diferem os bancos de dados e podem influenciar sua eficiência.

2.2.1.1 Estrutura de dados

A estrutura de dados de um Banco de Dados determinará como os dados estão organizados, sua coerência e seu relacionamento com os demais. Os fatores de eficiência e custo computacional de consultas e modificações são fortemente influenciados pela estrutura de dados adotada e pelo padrão das requisições realizadas. (GOMES, 2019)

Com relação à estrutura de dados, pode-se classificar os bancos de dados (GOMES, 2019) em:

- a) **Bancos de Dados Relacionais:** Estrutura de dados que permite a organização em grupos com características semelhantes. Nessa estrutura, os dados são organizados em tabelas ou outras estruturas hierárquicas, onde cada dado é identificado por atributos que indicam seu parentesco ou localização. Esse tipo de organização promoverá uma redução no tempo médio de recuperação e inserção de dados. Essas estruturas são comumente criadas utilizando linguagens computacionais SQL (*Structured Query Language*) tais como Oracle, MySQL, SQL Server e PostgreSQL.
- b) **Banco de Dados não relacionais:** Estrutura de dados utilizada em aplicações específicas, onde os dados não possuem relação perceptível entre si, ou para o armazenamento de informações de diferentes naturezas (fotos, audio, texto, tabelas e outros). Essas estruturas são comumente criadas utilizando linguagens computacionais NoSQL (*Not Only SQL*), tais como MongoDB, Redis e Cassandra.

A escolha de uma estrutura de dados deve ser realizada de forma a atender as requisições do projeto. Quando os bancos de dados são compartilhados por vários usuários ou em soluções em nuvem nota-se a adoção de soluções não relacionais mistas ou auto-contidas.

2.2.1.2 Solução de armazenamento

As informações armazenadas em bancos de dados são dependentes de uma solução de armazenamento. A escolha de uma solução dependerá da natureza dos dados que deverão ser armazenados e o modo de acesso. Podemos descrever três soluções para o armazenamento de dados (MARTINI, 2016):

- a) **Servidores** : Equipamentos que permitem acesso remoto às informações armazenadas à clientes conectados ao servidor. Os clientes são autorizados a realizar a busca, inclusão e modificação dos dados nele armazenados. São exemplos dessa solução: serviços de e-mail corporativo, serviços de compartilhamento de dados e sistemas de gerência organizacionais.

A escolha de um servidor dependerá das necessidades de: sistema operacional, capacidade de armazenamento, velocidade de acesso, dimensões e necessidades energéticas. Na Figura 2 estão apresentados os encapsulamentos físicos disponíveis: torres, *racks* e *blades*.

Figura 2 – Servidores Lenovo ThinkSystem



Fonte – Retirada da página na internet da Lenovo

- b) **Dispositivos de storage**: Equipamentos que permitem o acesso remoto à informações armazenadas com baixa frequência de acesso. Essa solução, embora de custo menor, é utilizada para o armazenamento de arquivos que não precisam ser acessados por muitos usuários ou cuja demanda seja maior que sua capacidade de transferência. Um desses dispositivos pode ser observado na Figura 3.

Essa solução, para ser adotada em grandes ambientes corporativos, é comumente combinada com o uso de servidores para redução dos tempos de acesso e expansão da capacidade de compartilhamento.

Figura 3 – Dispositivo De *Storage* Western Digital



Fonte – Retirado da página da internet da Western Digital

- c) **Nuvem** : Tecnologia, dependente da internet, para disponibilização de capacidade de armazenamento de dados conforme demanda e tarifação. Essa tecnologia garante maior

agilidade, escalabilidade, elasticidade e redução de custos de manutenção. Em contrapartida, o acesso de dados armazenados em nuvem é dependente da conexão com a internet, da contratação do serviço e está sujeita à problemas de segurança relacionados à privacidade.

Essa solução consiste na terceirização do armazenamento a grandes centros de dados. Essa transferência de responsabilidade, pode reduzir a demanda por infraestrutura para a instalação das outras soluções apresentadas, tais como espaço físico, estrutura energética, necessidade de resfriamento constante e dispositivos de segurança contra incêndios. Essas características devem ser levadas em conta ao se avaliar o investimento realizado para a adoção dessa solução.

A escolha de uma das soluções deve respeitar as necessidades organizacionais bem como os investimentos necessários para sua instalação.

2.2.2 Redes de Comunicação

A conexão entre computadores possibilita a troca de informações entre usuários de uma corporação. Essa conexão dá-se-a através de uma rede de comunicação.

Os padrões de redes de comunicação determinam o meio de conexão, os protocolos de comunicação, tecnologias para segurança e aplicações (FRANCISCATTO; CRISTO; PERLIN, 2014).

A seguir, serão discutidos as principais características e demandas das redes corporativas.

2.2.2.1 Redes segregadas

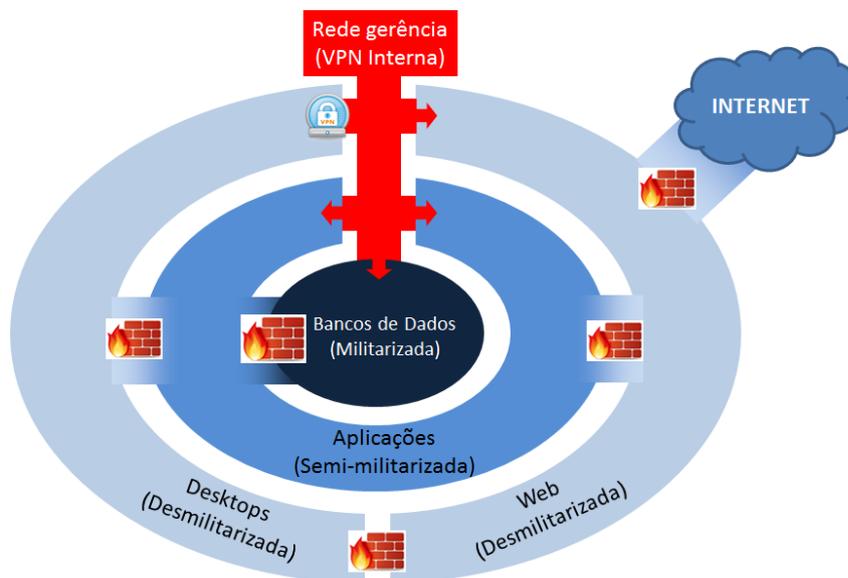
O gerenciamento de segurança dos computadores de uma empresa é uma tarefa complexa, pois em um ambiente conectado indiscriminadamente à internet há inúmeras portas de entrada para ameaças. Sendo assim, uma prática comum dos gerentes de TI é segregar a rede corporativa das redes para acesso externo. A Figura 4 representa uma rede segregada.

Figura 4 – Esquema de uma rede segregada



Fonte – (ALECRIN, 2013)

Figura 5 – Distinção dos segmentos de rede monitorados



Fonte – (BENETTI, 2016)

Uma rede segregada permite o monitoramento do tráfego de entrada e saída de forma centralizada, através de regras gerais que independem da quantidade de usuários instalados à rede. Procura-se com esse monitoramento eliminar vulnerabilidades e garantir a segurança dos dados que circulam na rede corporativa. (NAKAMURA; GEUS, 2007)

O nível de segurança atribuído define três camadas de rede ou perímetros de segurança: Uma zona desmilitarizada, uma zona semi-militarizada e uma zona militarizada. Aqui militarização corresponde ao monitoramento e controle do tráfego de entrada e saída, bem como à negação de acesso ao conteúdo por agentes externos. (NAKAMURA; GEUS, 2007). A Figura 5 apresenta um esquema dos perímetros de segurança de uma rede.

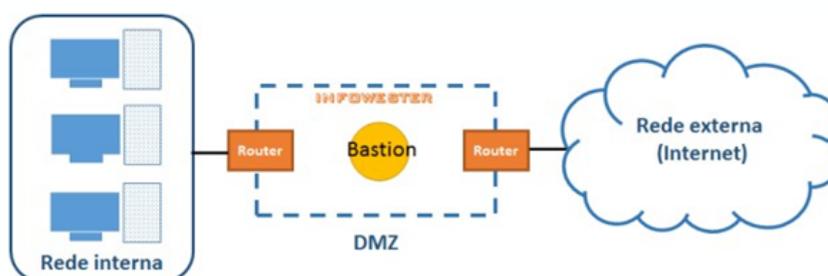
Ainda que, o tráfego em perímetros de segurança militarizados seja monitorado e isolado do acesso à ameaças externas, sua segurança não é garantida, pois deve-se levar em consideração que os usuários introduzem elementos não previsíveis e que podem ser alvo de ataques.

2.2.2.2 DMZ - Zona desmilitarizada

As zonas desmilitarizadas (*Demilitarized Zone - DMZ*) são porções de uma rede cujos perímetros de segurança não são confiáveis e onde os ativos deverão ser configurados pela equipe de TI individualmente para redução das vulnerabilidades (PEREIRA; SENA, 2018). Nesse ambiente, o monitoramento dos ativos computacionais deve ser realizada individualmente e o cuidado com a segurança dos dados devem ser extremos.

Nessa porção da rede estão localizados os proxies de saída para a internet, ou seja, um sistema de computador que age como um intermediário para requisições de clientes solicitando recursos de outros servidores com suporte às conexões oriundas da internet. Logo esse

Figura 6 – Esquema geral de funcionamento de uma *screened subnet*



Fonte – (ALECRIN, 2013)

ambiente deve ser caracterizado como hostil e seu acesso pelos computadores da rede corporativa deve ser limitado.

A DMZ, por sua vez, fica entre a rede interna e a rede externa, havendo um roteador que normalmente trabalha com filtros de pacotes entre a rede interna e a DMZ e entre a DMZ e a rede externa (ALECRIN, 2013), essa arquitetura é conhecida como *Screened Subnet*, conforme a Figuras 6.

O nível segurança e a flexibilidade de configuração fazem da *Screened Subnet*, que utiliza a DMZ uma arquitetura normalmente mais complexa e, conseqüentemente, de custo mais elevado.

2.2.2.3 Fronteiras das redes

As fronteiras entre as porções das camadas de rede são monitoradas por ferramentas responsáveis pela filtragem das conexões e do bloqueio dos ataques. Esses dispositivos são chamados de *firewalls*, podendo ser implementados por *software* ou *hardware*, sendo mais comum a utilização de *software*, já que o *hardware* nada mais é do que um equipamento com um *software* de *firewall* instalado. (ALECRIN, 2013).

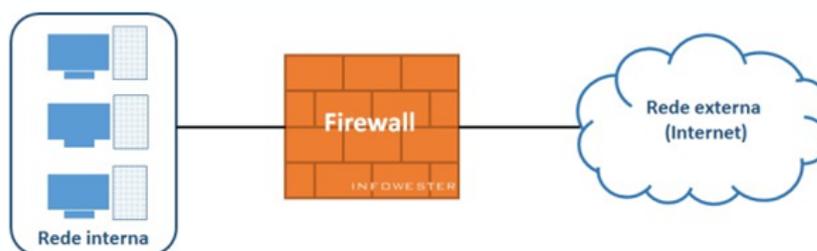
Os *firewalls* executam suas funções através de um conjunto de regras de filtragem de conexões que se baseia na análise do tráfego da rede. Como resultado dessa análise, os *firewalls* realizam uma atividade de bloqueio de segurança sendo capazes de determinar quais operações são permitidas e quais representam riscos aos dados que trafegam na rede. (MJR; RANUM, 1994) "Parede de fogo", a tradução literal do nome, evidencia que o *firewall* é uma ferramenta de proteção. A sua missão, por assim dizer, consiste basicamente em bloquear tráfego de dados indesejado e liberar acessos bem-vindos. (ALECRIN, 2013)

A partir de um conjunto de condições (regras), o *firewall* analisa o tráfego de rede para determinar quais operações de transmissão ou recepção de dados podem ser executadas. Além disso, recursos adicionais podem ser oferecidos por equipamentos *firewall*, tais como o redirecionamento de tráfego para outros sistemas de avaliação de segurança e para reforço dos serviço de autenticação dos clientes. *Firewalls* implementados em *hardware* específico podem apresentar ainda capacidades mais apuradas para lidar com grandes volumes de dados. (MJR;

Figura 7 – Firewall Cisco



Fonte – Cisco

Figura 8 – Esquema de um *firewall*

Fonte – (ALECRIN, 2013)

RANUM, 1994)

Nas Figuras 7 e 8 ilustram-se os dispositivos *Firewall*.

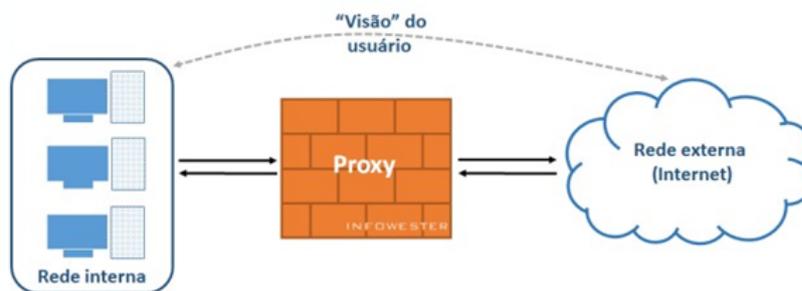
É possível encontrar, por exemplo, roteadores ou equipamentos semelhantes a estes que exercem a função de um *firewall*. Neste caso, o objetivo normalmente é o de proteger uma rede com tráfego considerável ou com dados muito importantes. (ALECRIN, 2013)

Firewalls são soluções importantes de segurança são amplamente utilizados até os dias de hoje, mas um *firewall* não é capaz de proteger totalmente uma rede ou um computador, razão pela deve ser utilizado em conjunto com outros recursos, como antivírus, sistemas de detecção de intrusos, VPN (*Virtual Private Network*) e assim por diante. (ALECRIN, 2013)

2.2.2.4 Proxy

Proxy de serviços são *firewalls* de aplicação específica para segurança das comunicações externas que tem como função atuar como um filtro de solicitações que intermediará as conexões entre a rede interna e uma rede externa (FRANCISCATTO; CRISTO; PERLIN, 2014), ou seja, é uma solução de segurança que atua como intermediário entre um computador ou uma rede interna e outra rede, externa - normalmente, a internet.

Figura 9 – Esquema de funcionamento de um proxy



Fonte – (ALECRIN, 2013)

São geralmente instalados em servidores potentes por precisarem lidar com um grande número de solicitações, *firewalls* deste tipo são opções interessantes de segurança porque não permitem a comunicação direta entre origem e destino (ALECRIN, 2013). Isso pode ser percebido na Figura 9.

Os *proxies* são implementados nos servidores para garantia da segregação entre a rede interna e externa. Esse elemento de ligação estabelece as regras para o controle do fluxo de entrada e saída, interpretando as requisições e devolvendo ao usuário os resultados.

A implementação de um *proxy* envolve complicações, haja visto a enorme quantidade de serviços e protocolos existentes na internet, fazendo com que, dependendo das circunstâncias, este tipo de *firewall* não consiga ou exija muito trabalho de configuração para bloquear ou autorizar determinados acessos. (ALECRIN, 2013)

Em alguns casos o *proxy* pode armazenar informações sobre o conteúdo das requisições anteriores para formulação de *logs* ou para eliminar a necessidade do reenvio das informações entre sítios da rede externa e clientes da rede interna.

2.2.3 Segurança da Informação

A segurança das informações é uma das vertentes de TI que estuda medidas para preservar a confidencialidade, integridade, disponibilidade e não repúdio das informações. Essas características podem ser descritas como (DIAS, 2000):

- a) **Confidencialidade:** Garantia do acesso à informação unicamente por agentes autorizados;
- b) **Disponibilidade:** Garantia do acesso à informação sempre que esta for solicitada;
- c) **Integridade:** Garantia de que o conteúdo da informação não será modificado por terceiros não autorizados;
- d) **Não Repúdio:** Garantia de que a autoria da informação não possa ser negada;

Uma das soluções para a garantia da segurança das informações é o uso de criptografia.

2.2.3.1 Criptografia

Em 2005 a editora Novatec publicou o livro “Criptografia em *software e hardware*”, o qual foi utilizado como base na construção deste texto.

Neste livro são abordadas várias questões relativas ao uso da criptografia como proteção para impedir o acesso a informações privadas. Os autores afirmam que:

A proteção por criptografia é uma ferramenta prática para proteger informações confidenciais. Independentemente do algoritmo criptográfico utilizado, sempre ocorrerá transformação de um texto legível em um ilegível. Mesmo que o invasor obtenha o conteúdo de um arquivo, este será ilegível. Para ter acesso à informação original, o invasor terá que resolver um problema matemático de difícil solução. A criptografia pode adicionar também maior segurança ao processo de identificação de pessoas, criando identidades digitais fortes. (MORENO; PEREIRA; CHIARAMONTE, 2005)

Com o conceito de criptografia, têm-se alguns termos oficiais utilizados: encriptar e decriptar. O ato de transformar um texto legível (texto claro, texto original, texto simples) em algo ilegível (cifra, texto cifrado, texto código) é chamado de “encriptar” (codificar, criptografar, cifrar). A transformação inversa é chamada de “decriptar” (decodificar, descriptografar, decifrar). (MORENO; PEREIRA; CHIARAMONTE, 2005)

Além destes termos, fala-se bastante em algoritmos:

O algoritmo de criptografia é uma sequência de procedimentos que envolvem uma matemática capaz de cifrar e decifrar dados sigilosos. O algoritmo pode ser executado por um computador, por um *hardware* dedicado e por um humano. Em todas as situações, o que diferencia um de outro é a velocidade de execução e a probabilidade de erros. Existem vários algoritmos de criptografia. (MORENO; PEREIRA; CHIARAMONTE, 2005)

Além do algoritmo, utiliza-se uma chave, a qual foi definida pelos autores como:

A chave na criptografia computadorizada é um número ou um conjunto de números. A chave protege a informação cifrada. Para decifrar o texto cifrado, o algoritmo deve ser alimentado com a chave correta, que é única. (MORENO; PEREIRA; CHIARAMONTE, 2005)

Os algoritmos criptográficos fazem uso de chaves para encriptar ou decriptar as informações. É pelo tipo de chave que podemos caracterizar a criptografia em dois tipos: simétrica e assimétrica. (MICROSOFT, 2010)

A diferença entre os dois tipos de criptografia é que algoritmos de encriptação simétrica usam a mesma chave para executar as funções de encriptação e decriptamento, já o algoritmo de encriptação assimétrica usa uma chave para encriptar os dados e outra chave para

decriptá-los. Outra diferença funcional entre encriptação simétrica e assimétrica está relacionada ao comprimento das chaves, que é medido em bits e está diretamente relacionado ao nível de segurança fornecido por cada algoritmo de encriptação. (STALLINGS, 2017)

Em sistemas assimétricos, a chave usada para encriptação é conhecida como Chave Pública e pode ser livremente compartilhada com outras pessoas. Por outro lado, a chave usada para decriptar é a Chave Privada e deve ser mantida em segredo. (STALLINGS, 2017)

Em sistemas simétricos, as chaves são selecionadas aleatoriamente e seus comprimentos são normalmente definidos com 128 ou 256 bits, dependendo do nível de segurança exigido. Na encriptação assimétrica, no entanto, deve haver uma relação matemática entre as chaves pública e privada, ou seja, há um padrão matemático entre as duas. Devido ao fato de que esse padrão pode ser potencialmente explorado por invasores para quebrar a encriptação, as chaves assimétricas precisam ter muito um comprimento muito maior para apresentar um nível equivalente de segurança. A diferença relacionada ao comprimento da chave é tão grande que uma chave simétrica de 128 bits e uma chave assimétrica de 2.048 bits oferecem níveis de segurança aproximadamente iguais. (STALLINGS, 2017)

Ambos os tipos de encriptação têm vantagens e desvantagens. Algoritmos de encriptação simétrica são muito mais rápidos e exigem menos poder de processamento computacional mas sua principal fraqueza é a distribuição de chaves. Como a mesma chave é usada para encriptar e decriptar informações, essa chave deve ser distribuída para qualquer pessoa que precise acessar os dados, o que naturalmente gera um aumento dos riscos relacionados a segurança. (STALLINGS, 2017)

Por outro lado, a encriptação assimétrica resolve o problema de distribuição de chaves usando chaves públicas para encriptação e chaves privadas para decriptação. A desvantagem, no entanto, é que os sistemas de encriptação assimétrica são muito lentos em comparação aos sistemas simétricos e exigem muito mais poder de processamento devido aos seus comprimentos de chave mais longos. (STALLINGS, 2017)

O estudo sobre a quebra de sistemas criptográficos é conhecido como análise criptográfica. Semelhantemente ao invasor, o criptoanalista procura as fraquezas dos algoritmos. O criptógrafo desenvolve sistemas de criptografia. É importante que a comunidade de criptografia conheça as fraquezas, pois os invasores também as estão procurando. (MORENO; PEREIRA; CHIARAMONTE, 2005)

De modo algum a criptografia é a única ferramenta necessária para assegurar a segurança de dados, nem resolverá todos os problemas de segurança. Segundo autores:

A criptografia é um instrumento entre vários outros, não sendo à prova de falhas. Toda criptografia pode ser quebrada e, sobretudo, se for implementada incorretamente, não agregará nenhuma segurança real (MORENO; PEREIRA; CHIARAMONTE, 2005)

2.2.4 Computação em nuvem

Com uma breve busca pelo termo na internet, é notório que a Computação em Nuvem está se tornando uma das palavras chaves da indústria de TI. Segundo várias publicações, a nuvem é uma metáfora para a Internet, sendo baseada em uma abstração que oculta à complexidade de infraestrutura. Cada parte desta infraestrutura é provida como um serviço e, estes são normalmente alocados em centros de dados, utilizando *hardware* compartilhado para computação e armazenamento. (BUYYA; RANJAN; CALHEIROS, 2009) (SOUSA; MOREIRA; MACHADO, 2009)

O termo computação em nuvem, segundo TAURION (2009), surgiu em 2006 em uma palestra de Eric Schmidt, da Google, sobre como sua empresa gerenciava seus data centers. Segundo TAURION (2009) hoje, computação em nuvem, se apresenta como o cerne de um movimento de profundas transformações do mundo da tecnologia. (SOUSA; MOREIRA; MACHADO, 2009)

Segundo Borges et al. (2011) a computação em nuvem pode ser definida como:

um paradigma de infra-estrutura que permite o estabelecimento do SaaS (*software* como serviço), sendo um grande conjunto de serviços baseados na web com o objetivo de fornecer funcionalidades, que até então, necessitavam de grandes investimentos em *hardware* e *software*, e que funciona através de um modelo de pagamento pelo uso. (BORGES et al., 2011)

Explorando a definição de Borges et al. (2011), os estudiosos elencaram três objetivos principais que são atribuídos a utilização de Computação em Nuvem: fornecer serviços de fácil acesso, baixo custo e com garantias de disponibilidade e escalabilidade:

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 substituição de recursos computacionais, podendo escalar tanto em nível de recursos de *hardware* quanto *software* para atender as necessidades das empresas e usuários. O último benefício é prover uma abstração e facilidade de acesso aos usuários destes serviços.

A computação em nuvem é uma evolução dos serviços e produtos de tecnologia da informação sob demanda, também chamada de *Utility Computing* (BRANTNER et al., 2008). O objetivo da *Utility Computing* é fornecer componentes básicos como armazenamento, processamento e largura de banda de uma rede como uma “mercadoria” através de provedores especializados com um baixo custo por unidade utilizada. (BORGES et al., 2011)

Uma definição disponibilizada pelo grupo Gartner é que se trata de modelo de computação onde as capacidades relacionadas a tecnologias da informação são escaláveis e elásticas, sendo que as mesmas são providas como serviços para os usuários finais através da internet. O grupo ainda ressalta que poucos investimentos iniciais em infra-estrutura são necessários.

Apresenta-se a definição do NIST (*National Institute of Standards and Technology* - USA):

Computação em nuvem representa um conveniente modelo de acesso, sempre que for necessário, a um conjunto compartilhado de recursos computacionais configuráveis, tais como, redes, servidores, armazenamento, aplicações e serviços, que podem ser disponibilizados rapidamente, e para isto o esforço de gerenciamento e interação com o provedor dos serviços é mínimo ou nenhum. (BORGES et al., 2011)

De forma geral, a computação em nuvem é o fornecimento de serviços de computação, incluindo servidores, armazenamento, bancos de dados, rede, *software*, análise e inteligência, pela Internet (“a nuvem”) para oferecer inovações mais rápidas, recursos flexíveis e economias de escala.

Para utilizarem os serviços, os usuários necessitam apenas ter em suas máquinas um sistema operacional, um navegador e acesso à Internet. Segundo pesquisadores:

todos os recursos computacionais estão disponíveis na nuvem e as máquinas dos usuários não necessitam ter altos recursos computacionais, diminuindo o custo na aquisição de máquinas. 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. (SOUSA; MOREIRA; MACHADO, 2009)

Segundo a Microsoft (2010), há três maneiras diferentes de implantar serviços de nuvem: em uma nuvem pública, nuvem privada ou nuvem híbrida. As definições de cada uma seguem:

As nuvens públicas pertencem a um provedor de serviço de nuvem terceirizado e são administradas por ele, que fornece recursos de computação (tais como servidores e armazenamento) pela Internet. Com uma nuvem pública, todo o *hardware*, *software* e outras infraestruturas de suporte são de propriedade e gerenciadas pelo provedor de nuvem. Uma nuvem privada se refere aos recursos de computação em nuvem usados exclusivamente por uma única empresa ou organização. Uma nuvem privada pode estar localizada fisicamente no *datacenter* local da empresa. Já as Nuvens híbridas combinam nuvens públicas e privadas ligadas por uma tecnologia que permite que dados e aplicativos sejam compartilhados entre elas. Permitindo que os dados e os aplicativos se movam entre nuvens privadas e públicas, uma nuvem híbrida oferece à sua empresa maior flexibilidade, mais opções de implantação e ajuda a otimizar sua infraestrutura, segurança e conformidade existentes. (MICROSOFT, 2010)

A escolha do tipo de implantação de uma computação em nuvem depende da demanda corporativa, pois apresentam características distintas.

2.3 Estado da Arte

A computação em nuvem envolve uma grande quantidade de conceitos e tecnologias. Empresas e corporações como a Amazon, Google e Microsoft estão publicando serviços computacionais seguindo a lógica da infraestrutura de computação em nuvem, sendo a Amazon a pioneira em disponibilizar e comercializar infraestrutura deste tipo. (SOUSA; MOREIRA; MACHADO, 2009)

Figura 10 – Soluções de nuvens comerciais



Nessa seção será avaliado o estado da arte, uma avaliação comparativa entre as características das soluções comerciais de maior popularidade, apresentados na Figura 10: Amazon Web Services (AMAZON, 2006), Microsoft Azure (MICROSOFT, 2010) e Google Cloud (GOOGLE, 2008).

2.3.1 Bancos de Dados

Todas as plataformas citadas possibilitam o armazenamento e a gerencia de bancos de dados relacionais e não relacionais.

Para avaliação comparativa entre as soluções, supõe-se um banco de dados relacional, pois o banco de dados relacional oferece maior consistência e confiabilidade, mas exige o relacionamento entre várias tabelas para o acesso à informação.

As soluções para o tratamento desse tipo de bancos de dados recebem nomes fantasia, sendo estes: Amazon Aurora, da Amazon Web Services; Google CloudSQL, da Google Cloud; e o Banco de Dados SQL do Azure, da Microsoft Azure.

2.3.1.1 Clientes

Como estratégia de marketing é comum que as soluções de computação em nuvem apresentem quais são os clientes que utilizam seus recursos. Uma boa cartela de clientes é um indicativo da qualidade do serviço prestado, do futuro organizacional da plataforma em nuvem e da tendência do mercado a adotar a solução, levando em conta fatores como facilidades, usabilidade, disponibilidade e custos.

São clientes do Amazon Aurora as empresas: Dow Jones, Arizona State University, Verizon, Capital One, ONU, Nielsen, Autodesk e outras. O Google CloudSQL possui como clientes as empresas: Evernote, Optiva, Pantheon, Natura, Hurb(Hotel Urbano), Findep, Magazine Luiza, Mercado Livre dentre outros. Alguns clientes do Microsoft Azure são Paychek, Komatsu, Allscripts, KMD.ABB, AccuWeather dentre outras.

Pela cartela de clientes descrita acima é possível notar que empresas que precisam de grande disponibilidade e órgãos governamentais preferem a solução Amazon Aurora; empresas brasileiras e lojistas preferem utilizar a solução Google Cloud e grandes empresas que possuem setores de confecção de produtos optam pela solução Microsoft Azure.

2.3.1.2 Segurança de dados

A computação em nuvem utiliza a Internet para disponibilizar seus serviços. Isso se torna mais complexo visto que os recursos computacionais utilizam diferentes domínios de redes, sistemas operacionais, *software*, criptografia, políticas de segurança, entre outros. Questões de segurança devem ser consideradas para prover a autenticidade, confidencialidade e integridade. (SOUSA; MOREIRA; MACHADO, 2009)

Cada provedor do serviço em nuvem, deverá procurar meios para proteger a segurança dos dados. Serviços como: segurança física, criptografia, acesso seguro e proteção avançada contra ameaças. A seguir, no Quadro 1, vamos descrever as soluções apresentadas na documentação para os três serviços em forma de quadro: (AMAZON, 2006), (MICROSOFT, 2010), (GOOGLE, 2008)

Quadro 1 – Soluções de Segurança implementadas pelos serviços de armazenamento em nuvem

Solução	Características de cada solução de acordo com as empresas
Amazon Aurora	<ul style="list-style-type: none"> ●Cinco vezes mais rápido que bancos de dados MySQL padrão e três vezes mais rápido que bancos de dados PostgreSQL padrão. ●Oferece a segurança, a disponibilidade e a confiabilidade de bancos de dados comerciais por um décimo do custo. ●Oferece um sistema de armazenamento distribuído, tolerante a falhas e com recuperação automática que escala automaticamente para até 64 TB por instância de banco de dados. ●Altos níveis de performance e disponibilidade, com até 15 réplicas de leitura de baixa latência, recuperação <i>point-in-time</i>, backup contínuo para o Amazon S3 e replicação entre três zonas de disponibilidade (AZs). ●Oferece vários níveis de segurança para o banco de dados. ●Isolamento de redes usando o Amazon VPC ●A criptografia de dados ociosos usando chaves criadas e controladas por você no <i>AWS Key Management Service (KMS)</i>

	<ul style="list-style-type: none"> ●Criptografia de dados em trânsito usando o SSL. ●Em uma instância criptografada do Amazon Aurora, os dados do armazenamento subjacente são criptografados, bem como os <i>backups</i>, <i>snapshots</i> e réplicas automatizados no mesmo <i>cluster</i>.
Microsoft Azure	<ul style="list-style-type: none"> ●"Firewall que impede o acesso de rede ao servidor de banco de dados" ●Tecnologia <i>Always Encrypted</i> ●Conformidade com padrões para tecnologia de proteção de informações e auditoria interna ●Segurança multicamadas: Segurança em <i>datacenters</i> físicos, infraestrutura e operações. ●Regras de <i>Firewall</i> de IP: protege o acesso de IPs que não estejam cadastrados ●Suporte a dois tipos de autenticação: autenticação do SQL usando nome de usuário e senha; e autenticação indireta ●Segurança em nível de linha: permite controlar os acessos às linhas de uma tabela com base nas características do usuário executando uma consulta ●Proteção contra ameaças: Rastreamento de atividades e histórico para identificação de possíveis ameaças ou suspeitas de violações de segurança e de abuso ●Criptografia em trânsito: Faz uso de um protocolo TLS para criptografia dos dados em movimento ●Criptografia em repouso: Criptografia AES, as chaves de criptografia são armazenadas nos servidores da Microsoft ●Criptografia em uso: proteção de dados confidenciais
Google CloudSQL	<ul style="list-style-type: none"> ●Transparência nos recursos de segurança ●Adoção de padrões internacionais de segurança e privacidade, validada por auditores independentes. ●Inexistência de acesso "<i>backdoor</i>" dos dados às entidades governamentais ●Confecção de relatórios de transparência sobre solicitações governamentais. ●Manutenção e Evolução contínua dos recursos de segurança ●Cumprimento das leis de proteção de dados ●Defesa em profundidade ●Segurança Física dos <i>datacenters</i> aprimorada e conforme padrões internacionais ●<i>Hardware</i> personalizado específico para uso seguro

- Proteção dos dados em trânsito por tecnologias de segurança, como HTTPS e *Transport Layer Security*.
- Criptografia de e-mails em repouso e em trânsito, além de *cookies* de identidade.

Conforme descrito no Quadro 1 podemos identificar que todas as empresas listam ferramentas: para garantia da confidencialidade, tais como criptografia em todos os estágios do armazenamento e transporte da informação; planos de negócio para garantia da disponibilidade e integridade dos dados; e ferramentas para garantia da privacidade dos usuários.

2.3.1.3 Migração

A transição de um ambiente tradicional para provedores de nuvem permite ao negócio optar pela solução de custo mais efetivo. No entanto, o processo de migração não é simples e pode oferecer riscos, tanto aos sistemas que foram fundidos, ou migrados parcialmente para o ambiente de nuvem, causando, por exemplo, exposição de informações críticas do negócio. (MORAIS, 2015)

Devido a complexidade do processo de migração, as empresas especificam como deve ser realizado o processo de migração para a computação em nuvem, as quais estão descritas no Quadro 2. (AMAZON, 2006) (MICROSOFT, 2010) (GOOGLE, 2008).

Quadro 2 – Soluções para migração de bancos de dados implementados pelos servidores de armazenamento em nuvem

Solução	Processo de Migração de acordo com as empresas
Amazon Aurora	<ul style="list-style-type: none"> • Facilidade para migração a partir de outros sistemas da Amazon • Documentação com informações simples e detalhadas do guia do usuário para executar a migração. • Aplicativos MySQL e o PostgreSQL podem ser reconectados facilmente à nova instância, sem qualquer modificação. • Suporte para as ferramentas padrão de importação e exportação e para a replicação de <i>binlog</i> do MySQL. • Ferramenta para migração do Oracle e do Microsoft SQL Server
Microsoft Azure	<ul style="list-style-type: none"> • Migração sem atrito • Migração de SQL Server locais sem alterar o código do aplicativo com a Instância Gerenciada. • Atualização constante do SQL Server, reduz o esforço de migração.
Google Cloud	<ul style="list-style-type: none"> • Migração do tipo "<i>Lift and Shift</i>" de soluções de bancos de dados de código aberto • Ferramentas disponibilizadas por terceiros podem ser utilizadas para facilitar a migração • Independência do tipo dos bancos de dados migrados

Observa-se que o processo de migração é um diferencial importante na escolha de uma solução em computação de nuvem. A Amazon apresenta maior detalhamento do processo

de migração no seu material de marketing disponibilizado na internet.

2.3.1.4 Escalabilidade e Disponibilidade

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 de computação em nuvem possuem acesso público, é imprevisível e variável a quantidade de requisições realizadas, tornando mais complexo fazer estimativas e garantias. (SOUSA; MOREIRA; MACHADO, 2009)

O Quadro 3 apresenta argumentos relativos a escalabilidade e disponibilidade declaradas de cada plataforma pelos seus criadores. (AMAZON, 2006) (MICROSOFT, 2010) (GOOGLE, 2008)

Quadro 3 – Soluções para Escalabilidade e disponibilidade implementados nos servidores de armazenamento em nuvem

Solução	Escalabilidade e disponibilidade declarados pelos criadores
Amazon Aurora	<ul style="list-style-type: none"> ● Sistema de Armazenamento distribuído, tolerante a falhas e com recuperação automática ● Tolerância a falhas ● Recuperação automática ● Escalamento automático para até 64 TB por instância de bancos de dados ● Oferece altos níveis de performance e disponibilidade. ● Até 15 réplicas de leitura de baixa latência ● Recuperação <i>point-in-time</i> ● <i>Backup</i> contínuo para o Amazon S3 ● Replicação entre três zonas de disponibilidade (AZs) ● Disponibilidade de 99,999999999%.
Microsoft Azure	<ul style="list-style-type: none"> ● Disponibilidade de até 99,995%. ● Escalamento automático - Hiperescala - com suporte para até 100 TB ● Recuperação de bancos de dados enormes em poucos minutos. ● Escalamento de recursos de computação e de armazenamento de modo independente
Google Cloud	<ul style="list-style-type: none"> ● Segurança e confiabilidade são incorporadas em todas as camadas ● Otimizada para custo e agilidade ● Bancos de dados totalmente gerenciados e compatíveis com código aberto. ● Bancos de dados unicamente escalonáveis e altamente confiáveis ● Suporte aos serviços do Google ● Possibilidade de aplicar soluções de análise de dados e <i>machine learning</i>

3 DESENVOLVIMENTO PRÁTICO DO TRABALHO

O seguinte capítulo tem como objetivo descrever as características do problema a ser solucionado e das soluções que serão sugeridas. Como apresentado por Morais (2015), a definição clara das necessidades de projeto determinará o desempenho do sistema, bem como sua necessidade de migração.

Esse capítulo será dividido em três seções: "Escopo do Problema", onde serão descritas as características do problema a ser resolvido, quanto em respeito a quantidade de acessos, capacidade de armazenamento, volume de informações a ser transitadas por período de tempo, periodicidade e probabilidade de carga, necessidade de disponibilidade e segurança; "Solução em nuvem", onde são descritos os custos envolvidos para manutenção mensal dos serviços em nuvem comercial para o atendimento das necessidades do escopo e "Solução Física", onde são avaliados os custos de uma solução física para atendimento das demandas do escopo.

3.1 Escopo do Problema

Os sistemas de gestão de dados administrativos de uma empresa são dependentes das suas atividades de auditoria e controle.

Na Marinha do Brasil, a gestão de dados administrativos pode ser dividida em quatro controles: o Controle Diário, controle dos agentes gestores locais; Controle Mensal, apresentação e controle dos ordenadores de despesa; Controle Semestral, controle dos órgãos de auditoria; Controle Quinquenal, necessidade de armazenamento de dados por período auditável. (MARINHA DO BRASIL, 2015)

O volume de dados administrativos é dependente da natureza dos dados relacionados e da estrutura dos dados de bancos de dados (GOMES, 2019). A solução atual, o sistema Quaestor, cujo símbolo é apresentado na Figura 11, apresenta um banco de dados relacional de natureza textual, onde cada usuário só possui acesso a uma parcela dos dados que lhe faz jus.

A migração desse sistema que encontra-se no fim da sua vida útil, tal como constatado por Nelson, Simek e Maschke (2009), nos leva a avaliar as soluções para armazenamento.

Figura 11 – Logotipo do sistema de apoio à gestorias Quaestor/Marinha do Brasil



Fonte – Marinha do Brasil

Para tanto devemos conhecer as especificações de projeto, através de uma avaliação preliminar.

3.1.1 Quantidade de acessos

As Organizações Militares (OMs) são divididas em dois grupos: OMs Apoiadoras, com rancho e gerência de caixa de economias próprios e OMs Apoiadas, que são arranchadas em outras unidades e "terceirizam" a gerência de seus recursos.

Os dados sobre a quantidade de acessos não é encontrada disponível em literaturas por serem dados de comportamento dinâmico. Dessa forma, este trabalho abordará o tema com estimativas majoradas, com uma margem de segurança para possível expansão futura. Supõe-se que existam quinhentas (500) organizações militares que precisam gerenciar seus recursos e ter acesso ao sistema.

Em cada Organização Militar é formada uma equipe responsável pela gerência desses recursos, sendo esta composta pelos seguintes agentes: Digitadores, gestores, equipe de pessoal, agentes fiscais e ordenadores de despesas (MARINHA DO BRASIL, 2015). Supondo que a equipe média possua quinze (15) agentes e que cada um destes possa acessar simultaneamente o banco de dados teremos, no máximo, sete mil e quinhentos (7500) acessos simultâneos aos bancos de dados.

Entretanto, existe a improbabilidade que essa quantidade de acessos ocorra simultaneamente. Um modelo mais fidedigno, adaptado de a partir de uma curva estatística, prevê que a quantidade de requisições varie de forma sazonal (BASTOS, 2015). Podemos identificar algumas características da rotina que podem apresentar tendências de acesso em diferentes horas do dia, dias de semana e períodos do mês, sendo estes:

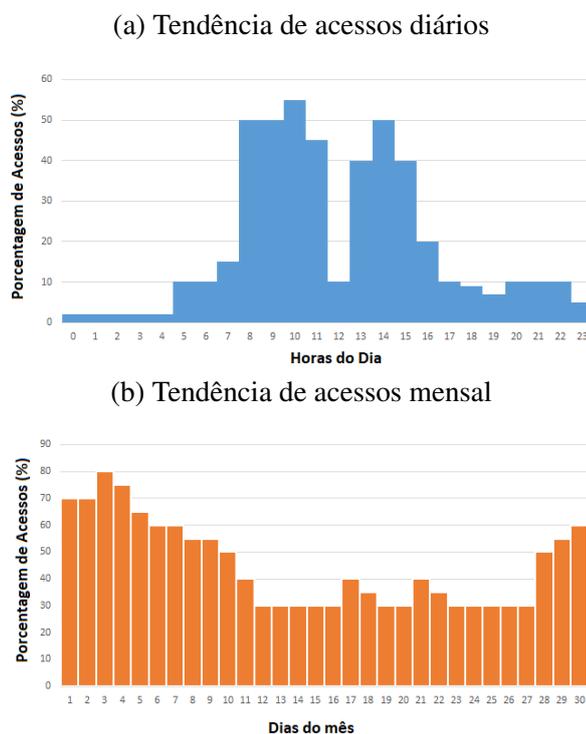
- A maior quantidade de acessos ocorrerá durante o expediente organizacional (das 8 horas as 16 horas)
- Maior quantidade de acessos entre os dias 5 dias finais de um mês e os 10 dias iniciais do mês seguinte (Realização de comprovações)
- Maior quantidade de acessos entre nos dias iniciais da semana
- O sistema sempre possuirá usuários conectados

Representa-se essa tendência em um gráfico de cargas mensal e diário. Essa tendência é apresentada na Figura 12.

3.1.2 Capacidade de Armazenamento

Segundo Milani (2006), o tamanho médio, em *bytes*, de uma entrada textual não comprimida mantida em um banco de dados é igual a sua quantidade de caracteres adicionado de uma unidade.

Figura 12 – Tendência de acessos ao sistema



Fonte – Elaborado pelo autor

Suponhamos que o tamanho médio da sequência de dados inseridos pelas OMs seja de 50 caracteres e que em média cada organização militar insira no banco de dados uma quantidade de três mil (3000) dados mensalmente, resultando num total de setenta e seis mil e quinhentos (76500) *bytes*.

No total, para se manter um histórico de dez(10) anos de dados das quinhentas(500) organizações militares devemos garantir o armazenamento mínimo de 4,3 GB.

3.1.3 Transito de informações

O trânsito de informações será dado pela média mensal do volume de dados inseridos ou requisitados pelas organizações militares.

Supondo que diariamente cada organização militar faça a requisição de uma quantidade de dados equivalente ao volume de dados inseridos e supondo a distribuição dos dados seja uniforme ao longo do horário de expediente, de 8 horas, temos um volume de carga por hora de 18,67 kB/h por OM.

Extrapolando o valor para a quantidade de clientes, o um volume total de dados requisitados mensal será de aproximadamente 2,13GB/mês.

3.1.4 Disponibilidade e Segurança

A disponibilidade do sistema desejada deve ser superior a 95%. Garantindo no mínimo 4 instâncias do servidor alocados em localidades geograficamente distantes.

Os dados devem ter sua confidencialidade e integridade garantidos. Mecanismos para garantir essas ações devem ser implementados.

3.2 Soluções em nuvem

Para fins de comparação de preço, serão avaliadas soluções que atendam as demandas do escopo adicionadas de algumas características relacionadas à desempenho listadas no Quadro 4.

Quadro 4 – Características das soluções em nuvem mínimas desejadas

Característica	Quantidade Mínima Solicitada
Número de Servidores virtuais	10
Número de Bancos de Dados armazenados	1
Capacidade de armazenamento	10 GB
Capacidade de armazenamento de <i>backup</i>	10 GB
Capacidade de armazenamento para registro de alterações	30 GB
Suporte à Bancos de Dados MySQL	Sim
Volume máximo de tráfego mensal	3000 dados, 3 GB
Mínimo de acessos simultâneos	7500
Tempo Mínimo de Contrato	3 anos

Fonte – Elaborado pelo autor

Serão avaliadas as opções mais simples e que resultam em menor custo, não sendo adverso ao compromisso de adoção de um serviço de processamento sem servidor (*Serverless*).

Especificações adicionais, utilizadas para cobrança, particulares de cada uma das soluções serão avaliadas tendo em vista o melhor atendimento das necessidades levantadas no escopo

3.2.1 Amazon Web Services

O serviço Amazon Web Services disponibiliza uma calculadora para determinação de custos totais e mensais aproximados em cada serviço. Essa calculadora é apresentada na Figura 13. Ademais, a política de cobrança é explicada em detalhes no artigo "*How AWS Pricing Works*" (AMAZON WEB SERVICES, 2018).

Conforme descrito por Amazon Web Services (2018) e apresentado na Figura 13, a determinação de custos baseia-se em três características: Computação, armazenamento e taxa de dados transferidos. Além disso, são oferecidos três formas de cobrança: sobre demanda, onde a cobrança é realizada sobre o processamento de dados e capacidade dos bancos de dados,

Figura 13 – Calculadora para determinação de preços

Fonte – (AMAZON WEB SERVICES, 2018)

pagamento temporal, onde o equipamento é reservado a um usuário por determinado tempo e reserva de recursos, onde os recursos são reservados com antecedência sem a opção de escalar-se.

Para o armazenamento de bancos de dados relacionais, o modelo de cobrança apresentado é chamado "Amazon RDS"(AMAZON WEB SERVICES, 2018). Nele, a cobrança é determinada pelos seguintes fatores:

- Horas de funcionamento : Quantidade de tempo que o Banco de Dados está disponível;
- Tipo de Bancos de Dados: Características dos dados armazenados e de como são acessados;
- Tipo de cobrança escolhida: sobre demanda, temporal ou reserva;
- Número de conexões;
- Armazenamento disponível;
- Armazenamento adicional sobre demanda;
- Número de requisições de acesso e gravação de dados;
- Número de nuvens redundantes solicitadas, e
- Transferência de dados.

A adoção de um serviço de armazenamento de bancos de dados sem servidor(*serverless*) representa o melhor custo-benefício. São características desse sistema: a cobrança é realizada mensalmente, conforme o tempo de uso, e os recursos são disponibilizados de forma linear e de maneira elástica.

No Quadro 5 estão relacionados os parâmetros utilizados para cobrança em pagamentos mensais, listam-se também suas unidades de cobrança e o custo mensal para a solução que atenda as características do banco de dados descrito no escopo.

Quadro 5 – AWS - Custos por unidade de contratação

Requisito	Custo por unidade de contratação	Custo para a solução
Instância (AGU)	0,06 por hora por instância	691,20 USD por mês por 16 instâncias
Armazenamento	0,10 USD por GB/mês/instância	32 USD por 20GB/mês para 16 instâncias
Taxa de Requisições	0,20 USD por 1 milhão de solicitações	0,20 USD por mês
Armazenamento de <i>Backup</i>	0,021 USD por GB/mês	0,21 USD por mês
Registros de alterações	0,012 USD por milhão de registros de alterações	0,012 USD por 300 dias
Transferência de dados de leitura	0,09 USD por GB	0,18 USD por mês
Transferência de dados de escrita	0,00 USD por GB	0,00 USD por mês

Fonte – Amazon Web Services

Nessa tabela, pode-se observar que a maior parcela dos custos é atribuída à contratação de instâncias e que os custos para o armazenamento de *backups* e registros de alterações são menores se comparados ao armazenamento convencional, pois o armazenamento de *backup* apresenta menor quantidade de gravações e requisições de dados.

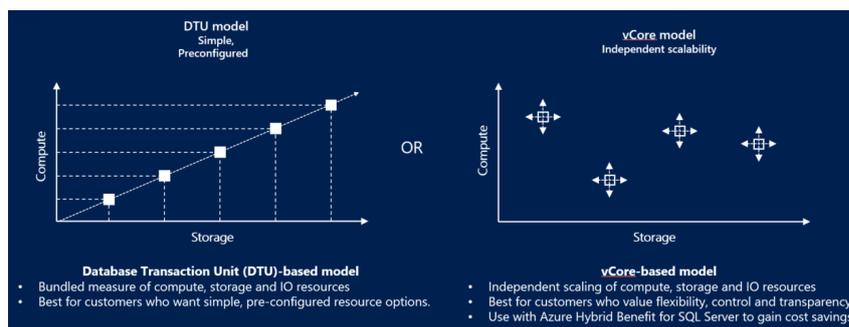
O custo total da solução estimado, para o período de cinco (5) anos, será de: \$43.428,12 USD.

3.2.2 Microsoft Azure

Conforme Microsoft (2010), estão disponíveis dois modelos de implementação de uma plataforma de bancos de dados: uma virtualizada, onde o usuário escolhe a quantidade de recursos computacionais que sempre serão disponibilizados, sendo possível a pausa automática conforme consumo; e o modelo DTU, onde serviços são vendidos em pacotes proporcionais, que mimizam consumos convencionais. A diferença entre os modelos pode ser observada na Figura 14.

Um sistema virtualizado, cujo preço é maior, torna-se uma solução viável para cargas intermitentes, pois nessa solução as cobranças são realizadas conforme quantidade de acessos e requisições. Um exemplo desse tipo de cobrança é apresentado na Figura 15. Entretanto, como a carga esperada para o sistema, apresentada na Figura 12, é sempre não nula, o modelo

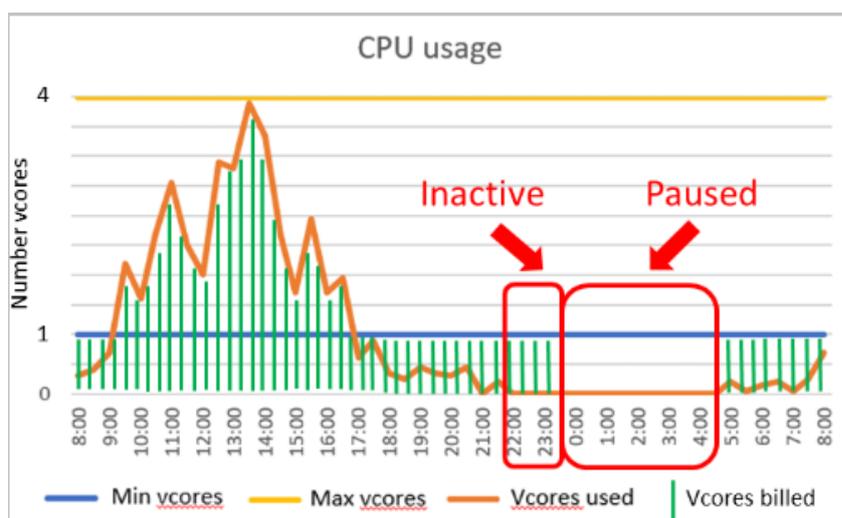
Figura 14 – Modelos de Contração dos serviços de Bancos de Dados do Microsoft Azure



Fonte – (MICROSOFT, 2010)

não apresentará muitas vantagens, sendo mais interessante avaliar um pacote de ferramentas fechado no formato DTU.

Figura 15 – Padrão para cobrança dinâmica baseada no uso de recursos de processamento



Fonte – Microsoft Azure

Os parâmetros para cobrança do uso de serviços de armazenamento de bancos de dados DTU são:

- Plano de Serviços: Planos com conjuntos de especificações de processamento e armazenamento;
- Quantidade de bancos de dados: Quantidade de bancos de dados relacionais que estarão disponíveis;
- Retenção de Longo Prazo: Tamanho médio do banco de dados durante o período de retenção;
- Criptografia de Dados Transparente;

- Períodos de retenção de *backup*, semanal, mensal e anual;
- Plano para Suporte Técnico;

Os planos de serviços, com as suas especificações garantidas, são apresentados no Quadro 6. Nela podemos notar que as premissas do escopo são atendidas no plano P1 e que a quantidade de DTUs em cada plano define as demais características.

Quadro 6 – Especificações computacionais mínimas dos planos premium da Microsoft Azure

Tamanho da computação	P1	P2	P4	P6	P11	P15
Número máximo de DTUs	125	250	500	1.000	1750	4000
Armazenamento incluído(GB)	500	500	500	500	4096	4096
Opções de espaço de armazenamento máximo(GB)	500, 750, 1024	500, 750, 1024	500, 750, 1024	500, 750, 1024	4096	4096
Armazenamento máximo OLTP na memória(GB)	1	2	4	8	14	32
Máximo de trabalhos simultâneos(solicitações)	200	400	800	1600	2400	6400
Máximo de sessões simultâneas	30000	30000	30000	30000	30000	30000

Fonte – Microsoft Azure

Existe uma calculadora, apresentada na Figura 16, para o cálculo do preço conforme as ferramentas solicitadas. No Quadro 7 estão relacionados os parâmetros utilizados para cobrança, juntamente com suas unidades de cobrança e o custo mensal para a solução que atenda as características do banco de dados descrito no escopo.

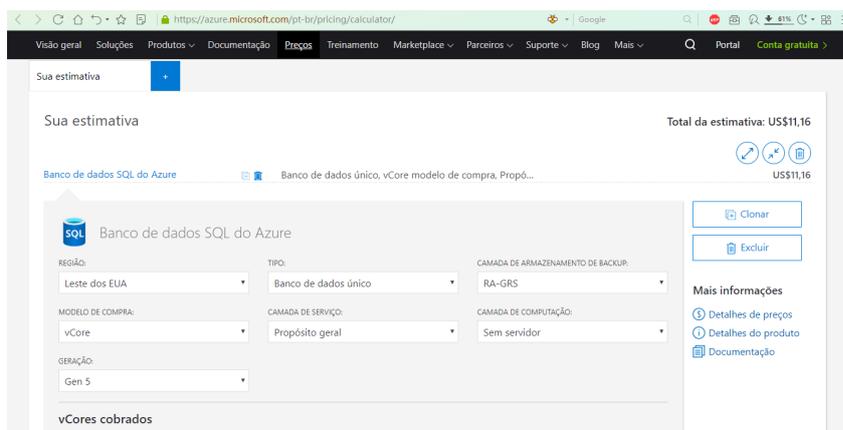
Quadro 7 – Microsoft Azure - Custos de produtos cobrados mensalmente

Requisito	Custo por unidade de contratação	Custo para a solução
Plano de Serviços Premium P1	456,25 USD por mês	456,25 USD por mês
Retenção de Longo Prazo	0,05 USD por GB/mês/ <i>backup</i>	30,00 USD por mês

Fonte – Microsoft Azure

No Quadro 7, nota-se que o custo total da solução estimado, para o período de cinco (5) anos, será de: \$29.175,00 USD.

Figura 16 – Calculadora para determinação de preços

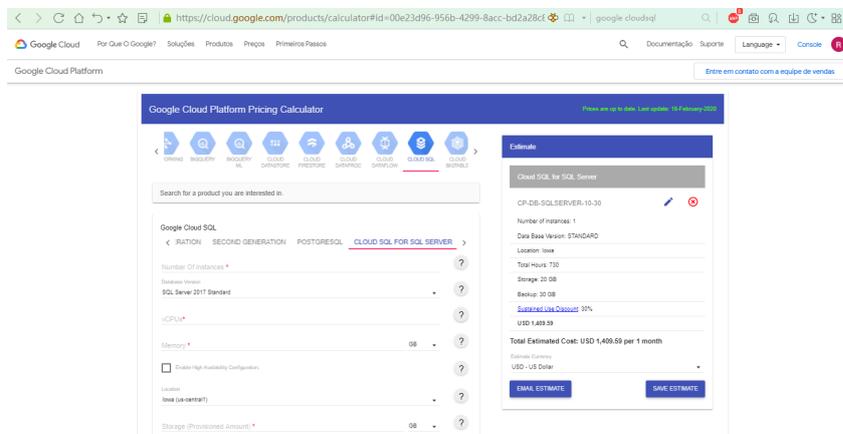


Fonte – Microsoft Azure

3.2.3 Google Cloud

O serviço Google Cloud disponibiliza uma aplicação calculadora para determinação de custos totais e mensais aproximados em cada serviço. Essa calculadora é apresentada na Figura 17.

Figura 17 – Calculadora para determinação de preços



Fonte – Google Cloud

Os preços do Google CloudSQL dependem da licença do MySQL contratada, para a nossa solução devemos considerar os preços para uma instância SQL Server. Para essa instância a cobrança dá-se através dos seguintes fatores:

- Agente computacional: Mede o processamento necessário para execução, considerando quantidade CPUs e memória necessários para execução de tarefas.
- Armazenamento e rede: Valores que dependem da localização dos servidores
- Período de operação: Quantidade de horas e dias que o sistema deve estar disponível por semana

Não há opção de utilização do sistema sem servidor(*serverless*), assim adota-se para fins de cálculo uma estrutura semelhante à encontrada na Tabela 5. Os custos dessa solução são encontrados no Quadro 8.

Quadro 8 – Google CloudSQL - Custos por unidade de contratação

Requisito	Custo por unidade de contratação	Custo para a solução
Agente Computacional CPU	30,15 USD por mês/unidade	482,40 USD por mês para 16 instâncias
Agente Computacional Memória	5,11 USD por GB/mês	163,52 USD por mês para 32GB
Armazenamento	0,17 USD por GB/mês	3,40 USD por mês para 20GB
Armazenamento de <i>Backup</i>	0,08 USD por GB/mês	2,40 USD por mês para 30GB
Saída de Internet	0,05 USD por GB/mês	0,15 USD por mês

Fonte – Google Cloud

Pode-se notar no Quadro 8 que a maior parcela do custo é resultado da contratação do agente computacional. O custo total da solução, para o período de cinco (5) anos, será de: \$39.112,20 USD.

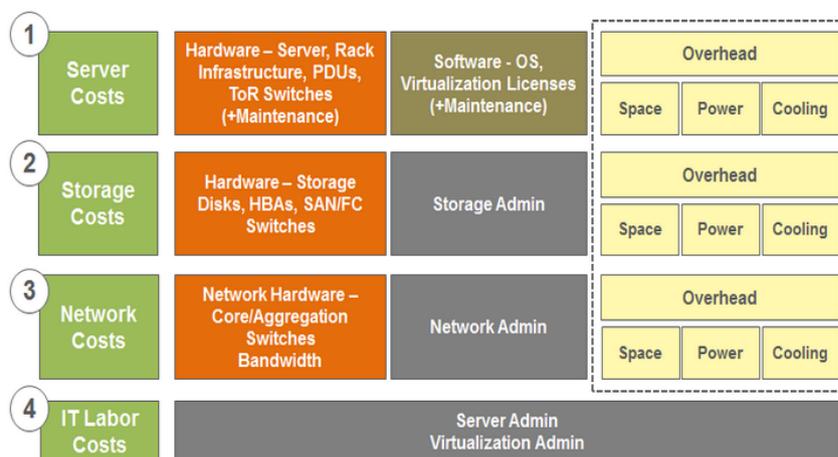
3.3 Solução física

A criação de um sistema físico que apresente as mesmas capacidades e disponibilidade dos sistemas em nuvem é uma tarefa complexa. Pode-se identificar alguns aspectos com diferentes características de custos sendo eles:

- **Custo de Implementação:** envolve os custos para a instalação (obras civis, instalação elétrica, proteção contra incêndios, entre outros) e os custos de aquisição de *hardware*(Compra das unidades de processamento, *racks*, ar condicionado, controle de segurança para as portas, entre outros), estes custos são singulares e realizados previamente à operação do sistema.
- **Custos de Manutenção :** envolve os custos de insumos(energia elétrica, etc), pessoal (equipe de suporte, limpeza, segurança) e manutenção dos equipamentos. Estes custos tem duas características: constante no tempo, como por exemplo o pagamento de salários; e estocásticos, que dependem do eventos aleatórios, como por exemplo o clima que pode afetar a durabilidade dos equipamentos e o preço da energia. Esses pagamentos são realizados por demanda ou mensais.

Outra divisão destes custos é apresentada por Barr (2014), nesse modelo os custos de implementação são divididos em: Custos de Servidor; Custos de Armazenamento; Custos de Rede; e Custos de mão-de-obra. Esse modelo é apresentado na Figura 18.

Figura 18 – Modelo de custos envolvidos na implementação de serviços de bancos de dados



Fonte – (BARR, 2014)

A título de curiosidade e comparação, as empresas fornecedoras de serviços em nuvem disponibilizam uma estimativa dos valores dessa instalação. Essa ferramenta é conhecida como Calculadora de Custos Totais de Propriedade (*TCO - Total Cost of Ownership*). O objetivo dessa seção é comparar os valores apresentados nessas calculadoras com os valores encontrados anteriormente levando em consideração o prazo de vida apresentado por (NELSON; SIMEK; MASCHKE, 2009).

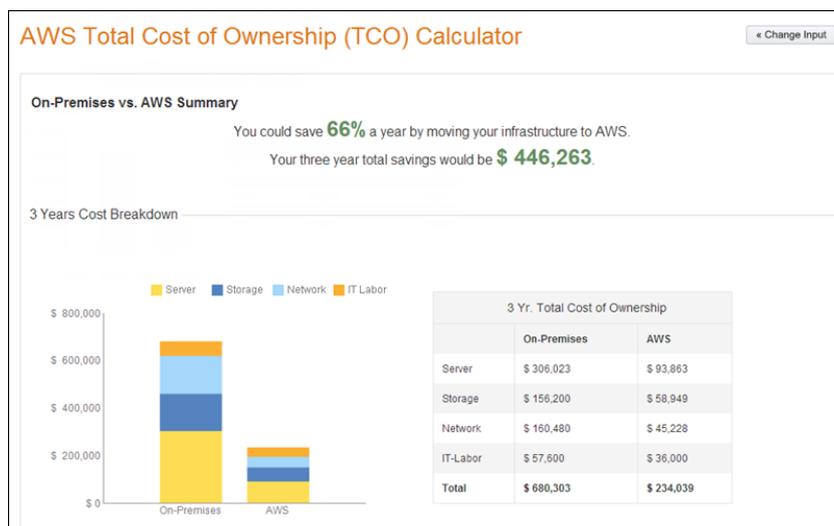
3.3.1 TCO da Amazon Web Services

A AWS fornece uma aplicação de calculadora de TCO, apresentada na Figura 19.

Nesse sistema os custos são avaliados em função das seguintes características:

- **Configuração de *Hardware*** : Quantidade máquinas virtuais, núcleos de processamento, memória RAM, etc.
- **Configuração de *Software***: Licença de *Softwares*
- **Tipo de configuração**: *On-premise* vs *Colocation*
- **Localização Geográfica**: Identifica o *datacenter* da Amazon mais próximo do local de instalação para aquisição de tarifas
- **Virtualização de servidores**: Capacidade de usar servidores virtuais
- **Armazenamento**: Capacidade de armazenamento e de *backup*

Figura 19 – Aplicativo Web - Calculadora de TCO da Amazon Web Services



Fonte – Amazon Web Services

Simulando uma plataforma para bancos de dados com as mesmas características da solução avaliada previamente encontramos que para um período de três anos o custo total de operação do servidor físico seria de R\$4.653.406,00.

O custo obtido pode ser expandido conforme o Quadro 9.

Quadro 9 – AWS - Resumo da Discriminação do Custo Total de Propriedade

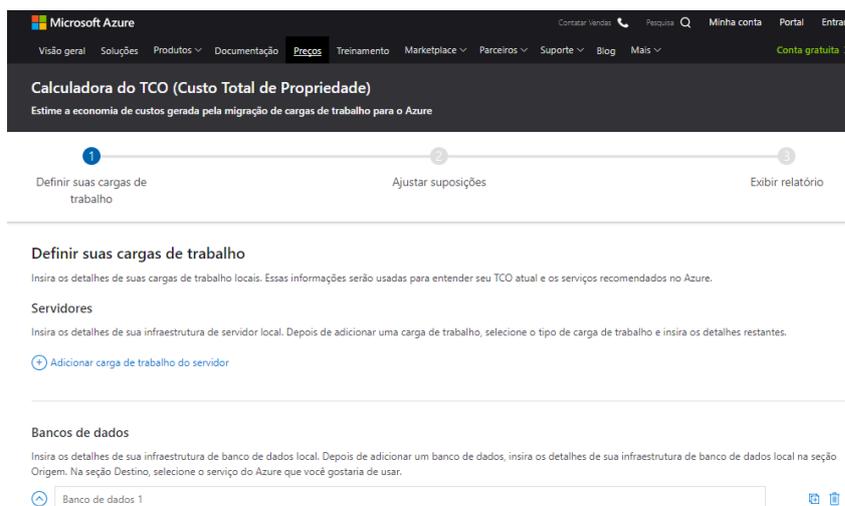
Categoria	Descrição	Custo total para o período de 3 anos(R\$)
Servidor	<i>Hardware</i>	193.210,00
	Manutenção	86.944,00
	Instalação dos <i>Racks</i>	363.577,00
	Virtualização (licenças de <i>software</i>)	141.744,00
	<i>Software</i> (Licenças)	3.926.761,00
	Energia e Refrigeração	315.032,00
Armazenamento	<i>Hardware</i>	525,00
	<i>Backup</i>	7.107,00
	Energia e Refrigeração	155.736,00
Rede	Disponibilização	238.119,00
Mão de obra	Total	10.125,00

Fonte – Amazon Web Services

3.3.2 TCO da Microsoft Azure

A plataforma Microsoft Azure fornece uma aplicação de calculadora de TCO, apresentada na Figura 20.

Figura 20 – Aplicativo Web - Calculadora de TCO da Microsoft Azure



Fonte – Microsoft Azure

Nesse sistema os custos são avaliados em função das seguintes características:

- **Configuração do *Hardware*:** Quantidade de processadores, memória, armazenamento, armazenamento para *backup* e possibilidade de virtualização;
- **Configuração de *Software* disponível:** Licenças de programas e sistemas operacionais;
- **Capacidade de Rede:** Largura de banda de saída;
- **Localização Geográfica:** Custo com energia elétrica e demais serviços
- **Custos de Suporte:** Suporte e seguro dos componentes e *softwares* comprados
- **Custos de espalhamento geográfico:** Redundância do sistema em outra localidade geograficamente distante.
- **Custos com Insumos :** Custos de Eletricidade e mão-de-obra de TI.
- **Período de Tempo:** Tempo estimado para substituição dos equipamentos

Essa estimativa é realizada levando em consideração a seguinte configuração:

- Um servidor único com 16 núcleos e 30 GB RAM
- Custo de manutenção do servidor igual a 20% do seu valor
- Licença para sistema operacional do servidor Windows
- Manutenção do sistema operacional
- Consumo energético mensal : 324 kWh

- Licença do serviço de bancos de dados SQL Server Enterprise
- Custos de construção civil para instalação dos *racks* e outras estruturas
- Custos de instalação dos *racks* de armazenamento
- Custos de manutenção dos sistemas de armazenamento igual a 10% do seu valor

O resultado da simulação nos mostra que em 5 anos os custos totais de propriedade do sistema totalizarão R\$4.190.628,73 ou \$ 958.953,94 USD. Esse valor, muito superior aos dos sistemas em nuvem simulados anteriormente, está distribuído conforme o Quadro 10. Nesse quadro, nota-se que a maior porção do custo é dado pela aquisição de *software* para gerência de bancos de dados.

Quadro 10 – Microsoft Azure - Resumo da Discriminação do Custo Total de Propriedade

Categoria	Custo (R\$/5 anos)
<i>Hardware</i>	51.739,73
<i>Software</i>	14.292,70
Consumo de Energia	6.853,81
Virtualização	5.911,00
Banco de Dados	4.075.682,40
Centro de Dados	16.843,09
Custos de Rede	11.557,57
Armazenamento	626,13
Mão de Obra	7.121,46

Fonte – Microsoft Azure

Alguns dos custos apresentados acima não se aplicam às características do negócio, principalmente os custos com mão de obra e custos de rede. O resulta em um custo total após 5 anos de R\$4.171.948,86.

Outro dado obtido nessa avaliação é a relação entre os custos de manutenção de ambas as soluções. Essa relação é apresentada na Figura 21, nela observa-se que a estimativa do custo de manutenção é diretamente proporcional ao tempo, mas que a taxa de variação, representativa dos custos mensais, é maior no sistema físico que nos sistemas em nuvem.

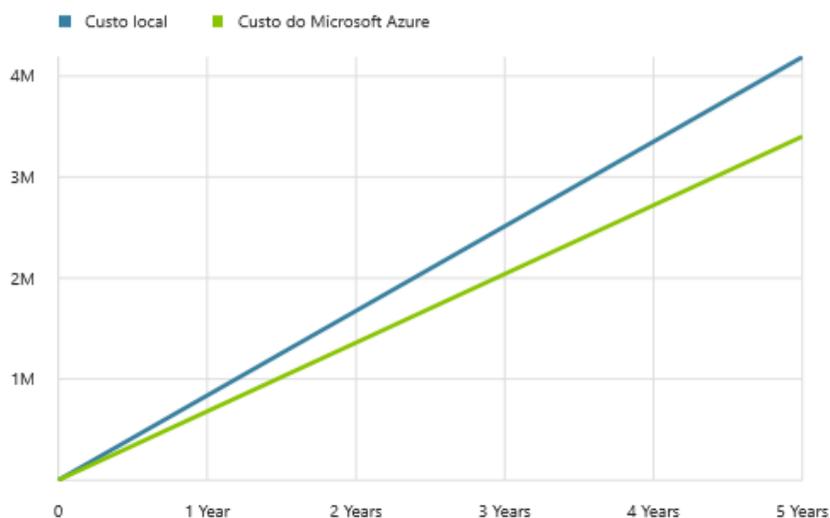
3.4 Comparativo entre soluções

No Quadro 11 encontram-se um resumo dos preços das soluções avaliados para um período de cinco anos, conforme apresentados anteriormente. Para a conversão Dólar x Real foi utilizada a seguinte relação:

$$\$1,00USD = R\$4,39 \quad (3.1)$$

De posse desses resultados podemos traçar algumas constatações, sendo estas:

Figura 21 – Comparação entre os custos de manutenção dos serviços em nuvem e físico local



Fonte – Microsoft Azure

Quadro 11 – Comparativo entre custos das soluções

Solução avaliada	Custo total para um período de 5 anos (R\$)
Nuvem - Amazon Web Services	190.649,45
Nuvem - Microsoft Azure	128.078,25
Nuvem - Google Cloud	171.702,56
Físico - TCO AWS	4.653.406,00
Físico - TCO Microsoft Azure	4.190.628,73

Fonte – Elaborado pelo autor

- Soluções em nuvem apresentam custos reduzidos se comparadas às soluções físicas.
- Para um escopo como o apresentado a solução em nuvem Microsoft Azure apresenta menor custo
- Soluções físicas são mais confiáveis com relação à confidencialidade dos dados, uma vez que estes estão armazenados somente na rede interna protegida por firewalls.
- Soluções em nuvem apresentam melhor disponibilidade, pois permitem a replicação de dados em data centers de excelência e distribuição geográfica de *backups*.
- Soluções Híbridas podem ser alvo de estudos futuros, onde os dados sejam cifrados localmente nos clientes de dados antes de ser enviados à nuvem, garantindo a confidencialidade dos dados.
- A avaliação do sigilo dos dados armazenados nos bancos de dados é decisiva para escolha da melhor solução.

- O maior custo encontrado nas soluções físicas correspondem à compra de licenças de *software*.
- Diferentes plataformas de soluções em nuvem listam produtos semelhantes com descrições distintas.
- Os custos resultantes das avaliações das calculadoras de TCO apresentam dados estimados e que nem sempre são críticos ao negócio.
- Os servidores de serviços em nuvem comerciais apresentam ferramentas para garantia da segurança dos dados
- Caso soluções de *software* livre sejam utilizadas, as soluções de armazenamento local tornam-se mais competitivas.
- A implementação de outros serviços nas soluções físicas e nuvens privadas permitem maior competitividade entre as soluções.

4 CONCLUSÕES

Conclui-se que soluções em nuvem podem ser utilizadas para o armazenamento de informações cujo nível de sigilo seja baixo e que não comprometam a organização, pois estes apresentam ferramentas de segurança, tais como criptografia em múltiplos estágios e protocolos de conexão segura, que são difundidas e de nível de segurança comprovado.

Os sistemas em nuvem, quando aplicados a sistemas de bancos de dados únicos, com características que atendem ao modelo de negócio da Marinha do Brasil, apresentaram menor custo se comparados a estimativas de Custo Total de Propriedade.

Entre as soluções avaliadas: Amazon Web Services, Microsoft Azure e Google Cloud, a Microsoft Azure apresentou menores custos em um sistema equivalente de manutenção de bancos de dados que atendesse às características estimadas de uso do sistema.

Os aspectos que afetam a avaliação de custos desse sistema são: licenças de *software*, custos de manutenção (energéticos, mão-de-obra, limpeza, manutenção predial, etc) e custos de aquisição. A escolha de um sistema baseado em *software* livre pode reduzir o custo de sistemas em até 90%, portanto, é importante difundir essa prática entre os desenvolvedores e gestores de TI com o intuito de reduzir custos e normatizar as soluções para interoperacionalidade entre sistemas operacionais.

Os modelos de cobrança das plataformas em nuvem são feitos de tal forma a permitir a cobrança da utilização dos recursos por demanda, o que pode ser uma vantagem em função das características da carga. Entretanto, o sistema em nuvem apresenta algumas desvantagens, entre elas, a dependência do acesso à internet e a terceirização da propriedade das informações.

REFERÊNCIAS

- ALECRIN, E. O que é firewall? conceito, tipos e arquiteturas. InfoWester, 2013. Disponível em: <<https://www.infowester.com/firewall.php>>. Citado 4 vezes nas páginas 19, 21, 22 e 23.
- AMAZON. *Amazon Web Services*. [S.l.], 2006. Disponível em: <<https://aws.amazon.com/pt/>>. Citado 4 vezes nas páginas 28, 29, 31 e 32.
- AMAZON WEB SERVICES. *How AWS Pricing Works*. [S.l.], 2018. Disponível em: <https://d0.awsstatic.com/whitepapers/aws_pricing_overview.pdf>. Citado 2 vezes nas páginas 36 e 37.
- BARR, J. The new aws tco (total cost of ownership) calculator. 2014. Disponível em: <<https://aws.amazon.com/pt/blogs/aws/the-new-aws-tco-calculator/>>. Citado na página 43.
- BASTOS, F. de A. A. *Estatística e Probabilidade*. Recife: Editora da Universidade Estadual do Ceará, 2015. Citado na página 34.
- BENETTI, T. Segurança da informação: Rede segregada em camadas. 2016. Disponível em: <<https://www.professionaisti.com.br/2016/11/seguranca-da-informacao-rede-segregada-em-camadas/>>. Citado na página 20.
- BORGES, H. P. et al. *Computação em nuvem*. [S.l.: s.n.], 2011. 48 p p. Citado 2 vezes nas páginas 26 e 27.
- BRANTNER, M. et al. Building a database on s3. In: *2008 ACM SIGMOD international conference on Management of data - SIGMOD '08*. New York: ACM Press, 2008. Citado na página 26.
- BUYYA, R.; BROBERG, J.; GOSCINSK, A. *Cloud Computing: Principles and Paradigms*. 1^a ed.. ed. [S.l.]: Wiley, 2011. Citado na página 11.
- BUYYA, R.; RANJAN, R.; CALHEIROS, R. N. Modeling and simulation of scalable cloud computing environments and the cloudsim toolkit: Challenges and opportunities. International Conference on High Performance Computing and Simulation, Leipzig, 2009. Citado na página 26.
- CONSTANTINI, A. Pesquisa científica: primeiros passos. 2017. Disponível em: <<https://augustobene.com/pesquisa-cientifica-primeiros-passos/>>. Citado na página 16.
- CURTIN, M. Introduction to network security. 1997. Disponível em: <<http://www.interhack.net/pubs/network-security.pdf>>. Citado na página 11.
- DIAS, C. *Segurança e Auditoria da Tecnologia da Informação*. [S.l.]: Axcel Books, 2000. Citado na página 23.
- FRANCISCATTO, R.; CRISTO, F. de; PERLIN, T. *Redes de Computadores*. [S.l.]: Universidade Federal de Santa Maria, Colégio Agrícola de Frederico Westphalen, 2014. Citado 2 vezes nas páginas 19 e 22.

- GOMES, P. C. T. Quais os principais bancos de dados e quais suas diferenças? 2019. Disponível em: <<https://www.opservices.com.br/banco-de-dados/e>>. Citado 2 vezes nas páginas 17 e 33.
- GOOGLE. *Google Cloud*. [S.l.], 2008. Disponível em: <<https://cloud.google.com/>>. Citado 4 vezes nas páginas 28, 29, 31 e 32.
- GREENSPOON, P. *SQL for Web Nerds*. 1a ed.. ed. [S.l.]: Morgan Kaufmann Pub, 2000. Citado na página 11.
- IBM. *Database Security*. [S.l.], 2019. Disponível em: <<https://www.ibm.com/cloud/learn/database-security>>. Citado na página 11.
- LAIGNIER, P.; FORTES, R. Introdução à história da comunicação. *E-papers*, 2009. Citado na página 10.
- LEINER, B. M. et al. *Brief History of the Internet*. Internet Society, 1997. Disponível em: <<https://www.internetsociety.org/wp-content/uploads/2017/09/ISOC-History-of-the-Internet-1997.pdf>>. Citado na página 11.
- MARINHA DO BRASIL. *SGM-305 Normas sobre Municiamto*. [S.l.], 2015. Citado 3 vezes nas páginas 12, 33 e 34.
- MARTINI, R. Principais soluções e tipos de armazenamento de dados para empresas. 2016. Disponível em: <<https://www.baguete.com.br/colunas/ramiro-martini/27/09/2016/principais-solucoes-e-tipos-de-armazenamento-de-dados-para-empresas>>. Citado na página 17.
- MENDES, C. Transferência do conhecimento nas organizações. março 2015. Disponível em: <<https://www.profissionaisti.com.br/2015/03/transferencia-do-conhecimento-nas-organizacaoes/>>. Citado na página 10.
- MICROSOFT. *Microsoft Azure*. [S.l.], 2010. Disponível em: <<http://www.microsoft.com/azure/>>. Citado 8 vezes nas páginas 24, 27, 28, 29, 31, 32, 38 e 39.
- MILANI, A. *MySQL - Guia do Programador*. 1a. ed. [S.l.]: Novatec, 2006. Citado na página 34.
- MILLER, R. Facebook accelerates its data center expansion. *Data Center Frontier*, 2018. Disponível em: <<https://datacenterfrontier.com/facebooks-accelerates-data-center-expansion/>>. Citado na página 13.
- MJR, M. R.; RANUM, M. J. Thinking about firewalls. In: *Proceedings of Second International Conference on Systems and Network Security and Management (SANS-II)*. [S.l.: s.n.], 1994. Citado 2 vezes nas páginas 21 e 22.
- MONNAPPA, A. How facebook is using big data - the good, the bad and the ugly. 2010. Disponível em: <<https://www.simplilearn.com/how-facebook-is-using-big-data-article>>. Citado na página 16.
- MORAIS, N. S. de. *Proposta de modelo de migração de sistemas de ambiente tradicional para nuvem privada para o Polo de Tecnologia da Informação do Exército brasileiro*. 88 p. Dissertação (Mestrado Profissional em Computação Aplicada) — Universidade de Brasília, 2015. Citado 2 vezes nas páginas 31 e 33.

MORENO, E.; PEREIRA, F. D.; CHIARAMONTE, R. B. *Criptografia em software e hardware*. São Paulo: Novatec, 2005. 21-42 p. Citado 2 vezes nas páginas 24 e 25.

NAKAMURA, E. T.; GEUS, P. L. de. *Segurança de redes em ambientes cooperativos*. São Paulo: Novatec Editora, 2007. Citado na página 20.

NELSON, S. D.; SIMEK, J. W.; MASCHKE, M. C. *The 2009 Solo and Small Firm Legal Technology Guide: Critical Decisions Made Simple*. [S.l.]: American Bar Association, 2009. 35-45 p. Citado 3 vezes nas páginas 12, 33 e 43.

ORACLE. *Oracle 9i Database Administrator's Guide*. [S.l.], 2001. Disponível em: <https://docs.oracle.com/cd/B19306_01/server.102/b14231/ds_concepts.htm>. Citado na página 11.

PAUL, M. S. Futuro ambiente estratégico em uma era de conflito persistente. *Military Review*, p. 78 – 91, 2009. Citado na página 10.

PEREIRA, V. da S.; SENA, Y. A. B. L. de. Dmzvisor: um firewall para a segurança de zona desmilitarizada corporativa em redes definidas por software. In: *Anais do XVIII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais*. Porto Alegre, RS, Brasil: SBC, 2018. p. 97–104. Disponível em: <<https://sol.sbc.org.br/index.php/sbseg/article/view/4275>>. Citado na página 20.

QUEIROZ, R. de C. R. de. *A Informação Escrita: Do Manuscrito ao Texto Virtual*. Tese — Universidade Estadual de Feira de Santana, 2004. Citado na página 10.

SILBERCHATZ, A.; KORTH, H.; SUDARSHAN, S. *Database Systems Concepts*. 6a ed.. ed. [S.l.]: McGraw-Hill, 2010. Citado na página 16.

SILBERSCHATZ, A.; STONEBRAKER, M.; ULLMAN, J. Database systems: Achievements and opportunities. *Communications of the ACM*, v. 34, n. 10, 1991. Citado na página 11.

SOUSA, F.; MOREIRA, L. O.; MACHADO, J. C. Computação em nuvem: Conceitos, tecnologias, aplicações e desafios. p. 150–175, 2009. Citado 4 vezes nas páginas 26, 27, 29 e 32.

STALLINGS, W. *Cryptography and Network Security - Principles and Practice*. 7ª ed.. ed. [S.l.]: Pearson Education Limited 2017, 2017. 767 p. Citado na página 25.

TAKAI, O.; ITALIANO, I.; FERREIRA, J. *Introdução a Banco de Dados*. [S.l.], 2005. Citado na página 17.

TAURION, C. Computação em nuvem: Transformando o mundo da tecnologia da informação. Brasport, Rio de Janeiro, 2009. Citado na página 26.

VOURAKIS, R. M. *A Evolução do Armazenamento da Informação*. Trabalho de Conclusão de Curso — Fundação Getúlio Vargas, 2017. Citado 2 vezes nas páginas 10 e 11.