

**CENTRO UNIVERSITÁRIO DE ANÁPOLIS – UniEVANGÉLICA**  
**BACHARELADO EM ENGENHARIA DE COMPUTAÇÃO**

**VINICIUS FAVORITO RODRIGUES**

***CLOUD COMPUTING - DAAS: DATA AS A SERVICE***

**ANÁPOLIS**

**2020-12**

**VINICIUS FAVORITO RODRIGUES**

***CLOUD COMPUTING - DAAS: DATA AS A SERVICE***

Trabalho de Conclusão de Curso II apresentado como requisito parcial para a conclusão da disciplina de Trabalho de Conclusão de Curso II do curso de Bacharelado em Engenharia de Computação do Centro Universitário de Anápolis – UniEVANGÉLICA.

Orientador(a): Prof. Ma. Aline Dayany de Lemos.

**Anápolis  
2020-12**

**VINICIUS FAVORITO RODRIGUES**

***CLOUD COMPUTING - DAAS: DATA AS A SERVICE***

Trabalho de Conclusão de Curso II apresentado como requisito parcial para a obtenção de grau do curso de Bacharelado em Engenharia de Computação do Centro Universitário de Anápolis – UniEVANGÉLICA.

Aprovado(a) pela banca examinadora em 10 de dezembro de 2020, composta por:

---

Prof. Ma. Aline Dayany de Lemos  
Orientadora

---

Profa. Ma. Luciana Nishi

---

Prof. Me. Millys Fabrielle Araujo Carvalhaes

## Resumo

É notório que nesta década vivenciamos a era da conectividade através da *internet*, portanto, uma expressiva quantidade de dados é gerada diariamente por meio de compras *online*, serviços de streaming, tais como a *Netflix* e o *Spotify*, as redes sociais em geral e dentre outros serviços. Em consequência, geram bases de dados de vários formatos, assim, surgindo o desafio de administrá-los de maneira a reduzir os custos de gerenciamento com os bancos de dados durante o processamento e armazenamento desses dados. Contudo, uma possível maneira de solucionar as dificuldades do gerenciamento desses dados é a utilização do serviço *Data as a Service*. Esse serviço consegue identificar e disponibilizar dados precisos de acordo com a necessidade do usuário, descartando o processamento de dados irrelevantes e assim economizando recursos. Visando a utilização do *DaaS* no gerenciamento das bases de dados, será apresentado um *checklist* englobando três provedores de serviços *DaaS*, sendo eles: *Oracle*, *Amazon* e *Snowflake*. Que serão classificados de acordo com seus aspectos para possibilitar a identificação para os melhores casos a serem implementados. E será apresentado um estudo comparativo dos três provedores do serviço *DaaS*, destacando suas características *DaaS* e seus respectivos bancos de dados.

**Palavras-chave:** *Data as a Service*, *DaaS*, serviço, gerenciamento de dados.

## **Abstract**

*It is well known that in this decade we have experienced the era of connectivity through the internet, therefore, a significant amount of data is generated daily through online purchases, streaming services, such as Netflix and Spotify, social networks in general and other services . As a consequence, they generate databases of various formats, thus, the challenge of managing them in a way that reduces the management costs with the databases during the processing and storage of this data. However, a possible way to solve data management difficulties is to use the data service as a service. This service is able to identify and provide accurate data according to the user's need, discarding the processing of irrelevant data and thus saving resources. In order to use the DaaS in the management of the databases, a checklist will be presented covering three DaaS service providers, namely: Oracle, Amazon and Snowflake. Which will be classified according to their aspects to enable identification for the best cases to be implemented. And a comparative study of the three providers of the DaaS service will be presented, highlighting their DaaS characteristics and their respective databases.*

**Keyword:** *Data as a Service, DaaS, service, data management.*

## Lista de Ilustrações

<b>Figura 1.</b> Modelo de <i>DaaS</i> .....	8
<b>Figura 2.</b> <i>Autonomous DataBase</i> .....	11
<b>Figura 3.</b> Informações do Banco de Dados .....	11
<b>Figura 4.</b> Gráficos de gerenciamento.....	12
<b>Figura 5.</b> Ciclo do OCPUs .....	13
<b>Figura 6.</b> Ambiente de desenvolvimento: <i>SQL Developer Web</i> .....	13
<b>Figura 7.</b> Ambiente para <i>uploads</i> de arquivo .....	14
<b>Figura 8.</b> Gerenciar estrutura .....	14
<b>Figura 9.</b> Pré-visualização da estrutura .....	15
<b>Figura 10.</b> Interface de SGBD da <i>Oracle</i> .....	15
<b>Figura 11.</b> <i>Oracle Cloud Infrastructure</i> .....	16
<b>Figura 12.</b> <i>API's</i> .....	17
<b>Figura 13.</b> <i>DashBoard</i> da <i>Amazon RDS</i> .....	18
<b>Figura 14.</b> <i>Engine Type</i> .....	18
<b>Figura 15.</b> Configurações do Banco de Dados .....	19
<b>Figura 16.</b> Delimitadores de Capacidade .....	20
<b>Figura 17.</b> Ferramentas gráficas de monitoramento .....	20
<b>Figura 18.</b> <i>Dremio API</i> .....	21
<b>Figura 19.</b> Criação das Tabelas .....	23
<b>Figura 20.</b> <i>Stages</i> .....	23
<b>Figura 21.</b> Credenciais dos servidores .....	24
<b>Figura 22.</b> Exportação dos dados .....	24
<b>Figura 23.</b> Exportação dos dados em arquivo JSON .....	25
<b>Figura 24.</b> Criando uma tabela .....	25
<b>Figura 25.</b> Compartilhamento de dados .....	26
<b>Figura 26.</b> <i>SnowFlake Data Marketplace</i> .....	26
<b>Figura 27.</b> Breve descrição sobre a empresa que está disponibilizando os dados .....	27
<b>Figura 28.</b> Permissões para ter os dados .....	27

## **LISTA DE TABELAS**

<b>Tabela 1.</b> <i>Checklist</i> Proposto .....	34
--	----

## Sumário

Resumo	4
Abstract	5
Keyword: <i>Data as a Service, DaaS, service, data management.</i> Lista de Ilustrações	5
LISTA DE TABELAS	7
1. INTRODUÇÃO	1
2. FUNDAMENTAÇÃO TEÓRICA	3
2.1 Cloud Computing	3
2.2 O Big Data	5
2.3 O Data as a Service (DaaS)	7
2.4 Benefícios da utilização do DaaS	8
3. DESENVOLVIMENTO	1
3.1 Oracle DaaS	1
3.2 Amazon RDS	8
3.3 Snowflake	13
3.5. Comparação entre os Provedores DaaS	19
4. Limitações	1
5. Considerações Finais	1
6. Referências Bibliográficas	2

## 1. INTRODUÇÃO

A *internet* tornou-se uma ferramenta importante na vida moderna, a primeira transação de dados ocorreu há aproximadamente 50 anos, e com seu sucesso, o compartilhamento de dados é utilizado até os dias atuais este novo universo, que ainda está em expansão, vem facilitando os meios de comunicação, trabalho e até mesmo para o lazer (EQUIPE DE REDAÇÃO, 2019).

Apesar de fazer parte do cotidiano da maioria das pessoas ter acesso a *internet*, as estatísticas mostram que ainda há lugares que não tem acesso à *internet*. Em uma pesquisa feita pelo *International Communications Union* (2019) indica que o ambiente *online* cresceu, e cerca de 56% da população possui acesso à *internet*, então, dos 7.7 bilhões de habitantes do planeta, em média 4.3 bilhões utiliza a rede mundial de computadores (NOGUEIRA, 2019).

Diante do aglomerado de dados, sua maioria é salva em bancos de dados locais que também são utilizados para processar suas informações (SOARES, 2014).

Os dados são armazenados por duas vertentes: dados estruturados e dados não estruturados. O armazenamento dos dados não estruturados não segue um padrão, não são armazenados em tabelas, linhas e colunas como os dados estruturados. E considerando que seu acesso é um tanto complexo, os meios para identificação dessas informações se tornam menos acessíveis (DEVMEDIA, 2014).

Tais dados dependem de um lugar para serem armazenados e comportar o processamento de sua estrutura, com a finalidade de fornecer informações da melhor maneira possível. Com a imposição progressiva em armazenar, manipular e analisar de uma forma célere e eficiente, capaz de organizar os dados em grandes volumes não estruturados, surgiu a inovação do *Big Data* (DEVMEDIA, 2014).

O *Big Data* é identificado por uma grande quantidade de dados, que atualmente seu processamento não cabe mais às técnicas comuns, como processar uma alta quantidade de dados em um banco de dados local (REDHAT, 2020). O conceito de *Big Data* quando trabalhado juntamente com o *Cloud Computing*, pode transformar a maneira de negócios realizarem suas funções (OPTIDATA, 2018).

O *Cloud Computing* gerencia e organiza os dados armazenados na nuvem, simplificando e tornando acessível o tratamento dos dados *Big Data* (OPTIDATA, 2018).

O *Cloud Computing* dispõe de serviços como por exemplo o *DaaS (Data as a Service)*. Um modelo focado para distribuição de serviços que possuem um banco de dados com um alto volume de informações e dependem de flexibilidade e velocidade no compartilhamento de informações (ALECRIM, 2020).

A realização de um estudo sobre a aplicabilidade desse serviço comparando três *softwares* provedores de serviços *DaaS*, *Oracle*, *Amazon* e *Snowflake* e construindo um *check-list* para identificar os critérios para utilização dos *softwares DaaS* é justificada nos custos e desafios apresentadas ao gerir grandes quantidades de dados, que de acordo com uma pesquisa feita pelo site *Data Science Academy* (2017), esclarece que cada etapa desde a coleta até disponibilização dos dados em geral para o usuário, depende de diferentes abordagens, para garantir que o resultado final seja entregue ao consumidor com velocidade e integridade.

Ante o exposto, como as características dos provedores de serviços *Data as a Service (DaaS)* podem auxiliar na escolha do melhor provedor diante de uma situação específica para tratar desafios de gerenciamento de dados?

Dessa forma, toma-se como objetivo geral deste trabalho realizar um estudo comparativo entre os *softwares* disponíveis para *Data as a Service*. E para que esse objetivo se torne possível, definiu-se as seguintes metas:

- Comparar as características do serviço *DaaS* e seus bancos de dados nos provedores *Oracle*, *Amazon* e *Snowflake*.
- Construir um *check-list* a fim de demonstrar as características dos provedores *DaaS*.

Nas próximas seções serão abordados tópicos descrevendo brevemente características de tecnologias que atuam juntamente aos serviços *DaaS*.

## 2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão discutidos e descritos conceitos de aumento de dados e envio destes dados para a nuvem, desafios ao tratar *Big Data*, aplicações de *Software* como serviço e Banco de Dados como serviço.

### 2.1 *Cloud Computing*

*Cloud computing* é uma estratégia de armazenamento de dados que utiliza a conexão com a Internet, facilitando desta forma, o acesso do usuário por meio de celulares, tablets e de forma padrão, o computador. Este tipo de tecnologia é fornecido de uma infraestrutura *online*, que não é necessário download, instalação de *software* específico ou servidor local, podendo ser utilizado apenas por um navegador de qualidade e uma conexão de *internet* estável (GIRALDO, 2018).

Segundo Giraldo (2018), *cloud computing* é um tipo de tecnologia que trouxe uma nova perspectiva de armazenamento de dados, ou como popularmente é conhecida, computação na nuvem. A computação em nuvem surgiu para ampliar e melhorar os recursos, trazendo em seu teor inúmeras vantagens para quem necessita que seus dados sejam distribuídos de forma mais simples e com menores custos, essas vantagens são:

- Redução de custos diante da infraestrutura: Como o *cloud computing* é uma tecnologia em nuvem e não há necessidade de investimento com compra de hardwares e *softwares* e também não haverá custos com instalação e a manutenção dos mesmos.
- Economia de espaço físico: Os dados são armazenados em uma nuvem, ou seja, essa tecnologia permite que um computador processe mais de um projeto simultaneamente, em vez de conter mais computadores para rodar projetos.
- Centralização das informações: É a concentração dos dados em um único servidor em nuvem, em vez de serem mantidos em diferentes servidores.
- Controle de recursos disponibilizado para o cliente sob a sua necessidade. Como a tecnologia do *Cloud Computing* tem a facilidade de ser flexível em mostrar a quantidade necessária de recursos, armazenamento e processamento.

- Trabalho remoto: Empresas que adotaram a *cloud computing* abrem a possibilidade de trabalho remoto, bastando apenas o funcionário ter acesso a *internet* e algum aparelho para o trabalho.

A ideia principal é o acesso à nuvem em qualquer tempo, lugar e meio, desde que seja feita com uma conexão com *internet*. Desta forma, as empresas que trabalham com *Cloud Computing* que também são denominados *Cloud Providers*, promovem uma grande mudança na forma padrão na infraestrutura, plataformas e nos *softwares*, pois além do acesso mais prático, sendo sua forma de pecúnia decorrente do uso regular da nuvem, da mesma forma que é cobrada água e luz de uma casa (DATA SCIENCE ACADEMY, 2017).

Neste contexto, existem três tipos de nuvem no mercado: as públicas, privadas e híbridas. Para as empresas, é necessário melhor discernimento sobre os tipos de nuvem, pois apesar de elas terem o mesmo propósito, é preciso identificar o propósito da nuvem e o objetivo da empresa para selecionar a que melhor se enquadra no ambiente (GIRALDO, 2018).

A nuvem pública é a mais utilizada entre as empresas por serem de baixo custo, ou seja, não é necessário um alto investimento em TI, a empresa somente irá custear a capacidade que será utilizada. O segundo tipo de nuvem é a privada neste modelo mantém um domínio interno tendo o acesso restrito apenas para usuários e parceiros, além de atender a um único negócio. Por fim, o terceiro tipo de nuvem é a híbrida que caracteriza a união dos dois tipos de modelos anteriores, podendo ser utilizada para extrair o melhor desempenho dependendo da finalidade do uso (RUSSO, 2017).

Ante o exposto é válido ressaltar que os serviços disponibilizados pela nuvem de forma flexível são obtidos por uma rede *web* de *data centers* de forma segura, além de serem atualizados de forma efetiva de acordo com a atual geração de *hardware*. Desta forma há múltiplos benefícios em serviços terceirizados de *internet* e *data centers* corporativos. O *Cloud Provider* são empresas que oferecem algum tipo de serviço de *Cloud Computing* para outras empresas (DATA SCIENCE ACADEMY, 2017; EQUIPE DE REDAÇÃO, 2019).

É incorporado no *Cloud Computing* o *Software as a Service (SaaS)* que são *softwares* que utilizam a computação em nuvem para oferecer serviços *online* como os serviços de *streaming* (VALIN, 2009).

Para deixar mais claro um exemplo clássico de *SaaS* é o *Salesforce* onde oferecem um serviço que fica hospedado em um servidor na nuvem (SANTANNA, 2018).

Antes que uma companhia passe a utilizar o *SaaS* é importante adequar o seu servidor para receber um grande volume de conexões em seu site e assim reforçar a segurança para evitar invasões (VALIN, 2009).

## 2.2 O Big Data

O *Big Data* tem a competência de registrar, juntar e processar um enorme número de dados em uma pesquisa de relações. E no *Big Data* não tem uma pesquisa exata até encontrarmos o que queremos. Para melhor entendimento do conceito, seria como procurar uma agulha em um palheiro, a agulha seria a pesquisa exata do que queremos e o palheiro seria as buscas até encontrar o que realmente queremos (LOMBARDO, 2013).

O *Big Data* é a forma que caracteriza um grande volume de dados estruturados ou não estruturados, que submerge uma empresa no cotidiano. No entanto, não se trata da escala de dados que são notáveis, e sim da manipulação dos dados feita pela organização que se mostra relevante. O *big data* pode ser observado pela retenção de informações, para que as empresas possam criar e escolher as melhores estratégias para o negócio (SAS INSIGHTS, 2019).

Observando o contexto do alto volume de dados armazenados que a maioria das empresas se encontra, juntamente com a obtenção de conhecimento sobre os *softwares* e configurações que compõem o processamento dos dados para o *Big data*, podem mostrar como um conjunto dados são capazes de identificar tendências sobre a conduta de consumo, e padrões de vendas para os meses futuros. O processamento *Big Data* é apto para avaliar mudanças que ocorrem através da transformação digital, que pode trazer benefícios para o negócio. O resultado seria uma atuação mais precisa que foca no bom uso dos dados (SANTOS, 2018).

Segundo Guimarães (2018), a ideia de gestão de dados maciços e de uma larga escala de informações obtidas, as maneiras onde os mesmo podem retornar de uma forma ordenada e tornar a admissão com um fácil entendimento. Com os conceitos dos V's dentro do *Big Data* que fomentam a modernização e a competitividade de uma empresa.

Cada um dos V's será descrito abaixo (NORMANDEAU , 2013; TAKE, 2017; GUIMARÃES, 2018):

- Volume: Para entender melhor o V mais conhecimentos dentro do *Big Data* que é o Volume, ao longo dos anos foram gerados megadados que se instituiu uma gigantesca quantidade de dados conforme o acesso com a *internet*. Até o próximo ano de 2021, teremos um tráfego de dados móveis que alcançará em torno de 587 exabytes anuais. Com um volume que tende a crescer, é necessário fazer investimentos em ferramentas e na expansão dos servidores para suportar o volume de dados.
- Velocidade: Os dados têm uma rapidez para serem criados, produzidos e acessados, com isso a análise dessa quantidade de dados deverá ser feita com velocidade.
- Veracidade: A veracidade dos dados do *Big data* tem como verificar a qualidades dos dados obtidos. Tendo a compreensão da veracidade dos dados irá ajudar as companhias analisarem os dados limpos e tomarem decisões de riscos.
- Variedade: Os dados são estruturados, não estruturados e semiestruturados. E os dados aparecem de vários tipos como fotos, vídeos, imagens, cookies entre outros tipos.
- Variabilidade: A variabilidade é a oscilação de dados nos níveis que são colhidas e manipuladas. Portanto, a variabilidade é mostrar que os dados vêm de várias fontes.

O *Big Data* apresenta dificuldades em sua implantação, como a definição da arquitetura que os dados serão estruturados, armazenados e analisados. Sua análise deve ser feita antes que as informações fiquem desprezíveis e ultrapassadas. E também deve-se levar em consideração o custo gerado para realizar essas ações. E como consequências de segurança de acesso a esses dados, quando armazenados por bancos de dados locais, se tornam suscetíveis a ataques, principalmente roubo de informações. (DATA SCIENCE ACADEMY, 2017).

Em uma época em que a conectividade é de alta velocidade e que tem a movimentação de grandes conjuntos de dados, e suprir os detalhes necessários para acessá-los é um problema. O grande volume de dados em que o *Big Data* é englobado, na maioria das vezes possuem dados confidenciais, de alta importância para empresas e pessoas. Tal

situação se torna uma séria preocupação para provedores *DaaS* e empresas que consomem os serviços oferecidos pelos *DaaS*. Uma solução para problemas citados no decorrer desta Seção, seria a junção entre *Cloud Computing* e *Big Data*, a união delas apresenta uma possível solução escalável, moldável para grandes volumes de dados. Tornando os recursos facilmente acessados e com custo menor para a empresa (DATA SCIENCE ACADEMY, 2017).

### 2.3 O Data as a Service (DaaS)

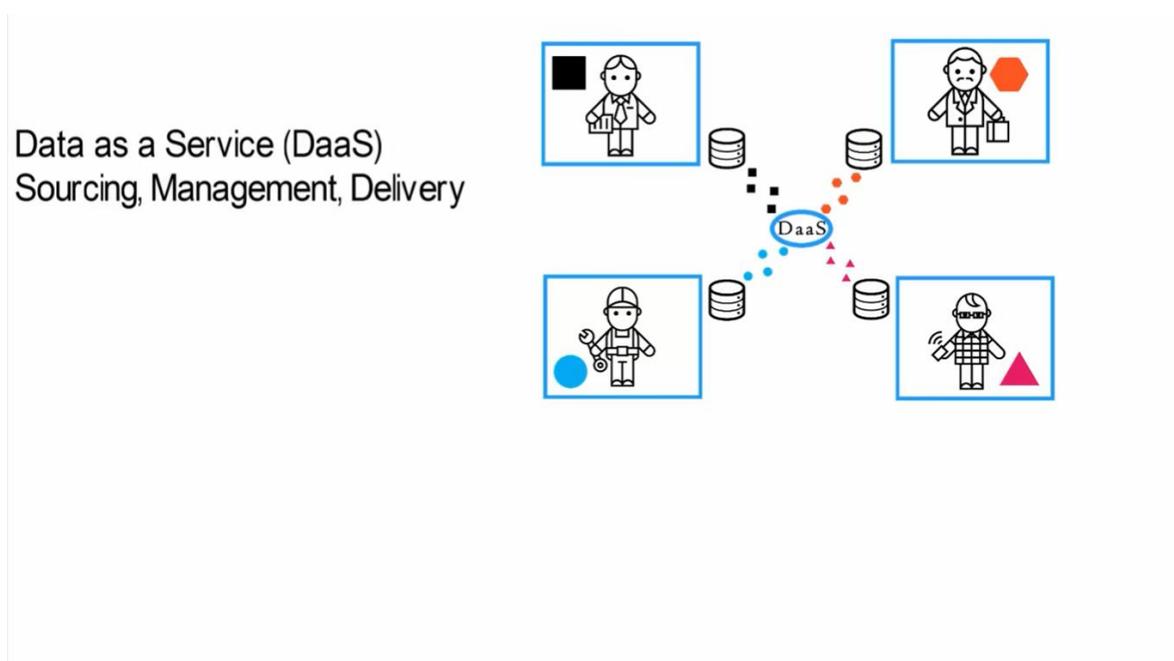
O *DaaS* é um conceito relativamente novo e que só a partir de 2008 começou a chegar às instituições, que passaram a perceber os benefícios que a adoção deste modelo pode levar. Esta oportunidade surgiu diante dos serviços já existentes não terem sido estruturados para trabalhar com grandes quantidades de dados e alta velocidade de processamento. De modo que alguns provedores que oferecem o serviço de armazenamento de dados disponibilizam somente serviços básicos para armazenamento e hospedagens de aplicativos, diferentemente do *DaaS* que dispõe de serviços de integração, análise e processamento de dados (MCDANIEL, 2019).

O *Data as a Service* é um serviço de gestão de dados podendo ser acessado apenas com o acesso à *internet*, pois usa a nuvem para disponibilizar seus serviços de armazenamento, inclusão, tratamento e diagnóstico de dados (MCDANIEL, 2019).

O *DaaS* é uma das novas tecnologias que possui um comportamento “*as a Service*”, ou seja, para propiciar uma maneira simples a gestão de dados e economia de recursos, dispensa tais como tarefas de *softwares* complexos e que podem gerar uns altos custos de processamento e armazenamento e conseqüentemente gastos financeiros (FELDMAN, 2020).

O *SaaS* é um modelo de computação em nuvem que dispõe de suas aplicações aos usuários somente tendo acesso à *internet*, sem a necessidade de instalar programas em suas máquinas. O *SaaS* dispensa a necessidade da instalação e gerenciamento de programas instalados nas máquinas locais, o *DaaS* possui funcionamento semelhante tentando evitar o aumento dos custos (MCDANIEL, 2019).

Observe a Figura 1 de um modelo de *DaaS*, que mostra várias informações vindas de fontes separadas sendo reunidas para serem processados pelo serviço *DaaS*.

Figura 1 - Modelo de *DaaS*

Fonte: Data as a Service (*DaaS*), (2015).

Como uma tecnologia usada internamente em uma empresa, a plataforma de dados como serviço é uma solução *end-to-end* e pode ser considerada como um facilitador entre várias fontes de dados e ferramentas, como relatórios de autoatendimento, BI, microsserviços e formulários. Depois que a plataforma é implantada, os usuários finais podem acessar os dados sempre que quiserem, usando *SQL* padrão sobre *ODBC*, *JDBC* ou *REST*. As empresas também podem usar serviços *DaaS* externos para acessar dados. Várias empresas fornecem serviços *DaaS* por meio de *APIs* simples. Por exemplo, estes são alguns dos principais fornecedores que fornecem dados sobre empresas: *Clearbit*, *Crunchbase* (KANTARCI, 2020).

Como uma *DaaS* é um modelo alternativo de serviço de computação em nuvem, diferente dos modelos tradicionais *IaaS*, *PaaS* e *SaaS*, onde os dados são disponibilizados aos usuários como um serviço por meio de rede. Uma vez que os dados são o valor deste modelo, é fundamental ser capaz de gerenciar e processar a maior quantidade de dados heterogêneos, a fim de permitir acesso oportuno a informações críticas de conhecimento (TERZO et al., 2013).

#### 2.4 Benefícios da utilização do *DaaS*

Como na atualidade a acessibilidade à *internet*, o armazenamento de dados na nuvem está se tornando comuns, quando utilizados juntamente aos serviços fornecidos pela nuvem desenvolvidos exclusivamente para gerir e tratar com agilidade grandes quantidade

de dados, surge um ambiente propício onde o serviço *DaaS* se destaca com suas praticidades e benefícios (MCDANIEL, 2019).

O *DaaS* combina os dados de CRM (gerenciamento de relacionamento com o cliente) de uma empresa com gatilhos em tempo real e fontes de dados difíceis de encontrar para fornecer melhor segmentação e um fluxo de consumidores no mercado (BEDGOOD, 2015).

Diante dos desafios de processamento, armazenamento e gestão de dados em bancos locais, demonstra que o *DaaS* oferece várias vantagens relevantes relacionadas à velocidade, confiabilidade e desempenho em relação aos bancos de dados locais. O *DaaS* possui uma maior flexibilidade e escalabilidade em comparação aos bancos de dados que estão em uma máquina local, um exemplo típico seria o *software PostgreSQL*, um banco de dados relacional que pode ter seus recursos controlados e disponibilizados instantaneamente de acordo com a necessidade atual do cliente (MCDANIEL, 2019).

Os custos de gerenciamento e processamento de dados em bancos locais são calculados de maneira fixa, onde processar certa quantidade de dados sempre gera um custo base pré-determinado, com o pré-requisito de ter o aumento de dados processados pode haver um aumento no valor cobrado. As plataformas *DaaS* também geram seus custos de acordo com o processamento, porém seus provedores podem alocar a quantidade suficiente de recursos para suas cargas de trabalho, podendo controlar a distribuição conforme a necessidade, economizando gastos desnecessários. E após os serviços são adquiridos, os provedores permitem quase que imediatamente que as organizações já possam começar a processar e armazenar os dados (MCDANIEL, 2019)

Como o *DaaS* é um serviço em nuvem, suas funcionalidades são aprimoradas a um nível onde a probabilidade de falha no funcionamento dos servidores é mínima, ou seja, os servidores têm pouca chance de cair. Deixando assim os usuários mais seguros com menores chances de terem prejuízos. Além disso, seus usuários não precisam se preocupar em buscar novas atualizações a serem instaladas, já que são atualizados automaticamente pelos provedores do serviço (MCDANIEL, 2019).

A utilização dos serviços *DaaS* reduz o acúmulo e a repetição de dados, em relação aos modelos de banco de dados tradicionais como exemplo bancos relacionais, permitindo que os usuários os enviem o mais rápido possível até ao consumidor, dependendo somente da velocidade de conexão com a *internet* (DREMIO'S DATA LAKE ENGINE, 2020).

### 3. DESENVOLVIMENTO

Com base nos aspectos dos provedores *Oracle*, *Amazon* e *Snowflake* será apresentado um estudo comparativo identificando suas características e posteriormente um *checklist* exibindo uma comparação entre os atributos destacados para que possibilite um entendimento melhor de onde cada um desses provedores pode ser mais aproveitado.

#### 3.1 Oracle DaaS

Utilizando os serviços da *Oracle*, um cadastro grátis foi feito. Utilizando o *Oracle Cloud*, criamos um banco de dados relacional denominado *Autonomous DataBase*. A versão grátis permite que seja usado um núcleo de processamento para leitura dos dados e 20GB de armazenamento conforme exibido na Figura 2.

Figura 2 - *Autonomous DataBase*

The image shows a screenshot of the Oracle Autonomous Database configuration interface. At the top, it says "Configure o banco de dados". Below that, there is a section for "Always Free" with a toggle switch that is turned on, labeled "Mostrar apenas as opções de configuração Always Free". A yellow warning box contains the text: "Se o seu Always Free Autonomous Database ficar sem atividade por 7 dias consecutivos, o banco de dados será interrompido automaticamente. Seus dados serão preservados e você poderá reiniciar o banco de dados para continuar utilizando-o. Se o banco de dados permanecer parado por 3 meses, ele será reivindicado. Saiba mais". Below the warning, there is a dropdown menu for "Escolher versão do banco de dados" with "19c" selected. At the bottom, there are two input fields: "Contagem de CPUs" (with "1" entered) and "Armazenamento (TB)" (with "0.02" entered). Both fields are labeled "SOMENTE PARA LEITURA". Small text below the CPU field states: "Os Autonomous Databases Always Free podem utilizar até 1 núcleo. A contagem de núcleos de CPU não pode ser ajustada." Small text below the storage field states: "Os bancos de dados Autônomos Always Free podem utilizar até 0.02 TB (20 GB) de armazenamento. O tamanho do armazenamento não pode ser ajustado."

Fonte: *Oracle* (2020)

Com o banco de dados criado, antes de acessá-lo é possível visualizar por uma interface, as informações gerais, como demonstrado na Figura 3.

Figura 3 - Informações do Banco de Dados

A screenshot of a web interface showing general information for a database. The title is "Informações Gerais". Below it, several key-value pairs are listed: "Nome do Banco de Dados: DB202009231851", "Compartimento: tchucubira2 (raiz)", "OCID: ...2vtdsq" with links "Mostrar" and "Copiar", "Criado: qua., 23 de set. de 2020 21:54:43 UTC", "Contagem de OCPUs: 1", "Armazenamento: 20 GB", "Tipo de Licença: Licença incluída", "Versão do Banco de Dados: 19c", "Dimensionamento Automático: Desativado" with an information icon, "Estado do ciclo de vida: Disponível", "Tipo de Instância: Gratuito" with a link "Upgrade para Pago", and "Modo: Leitura/Gravação" with a link "Editar".

**Informações Gerais**

**Nome do Banco de Dados:** DB202009231851

**Compartimento:** tchucubira2 (raiz)

**OCID:** ...2vtdsq [Mostrar](#) [Copiar](#)

**Criado:** qua., 23 de set. de 2020 21:54:43 UTC

**Contagem de OCPUs:** 1

**Armazenamento:** 20 GB

**Tipo de Licença:** Licença incluída

**Versão do Banco de Dados:** 19c

**Dimensionamento Automático:** Desativado ⓘ

**Estado do ciclo de vida:** Disponível

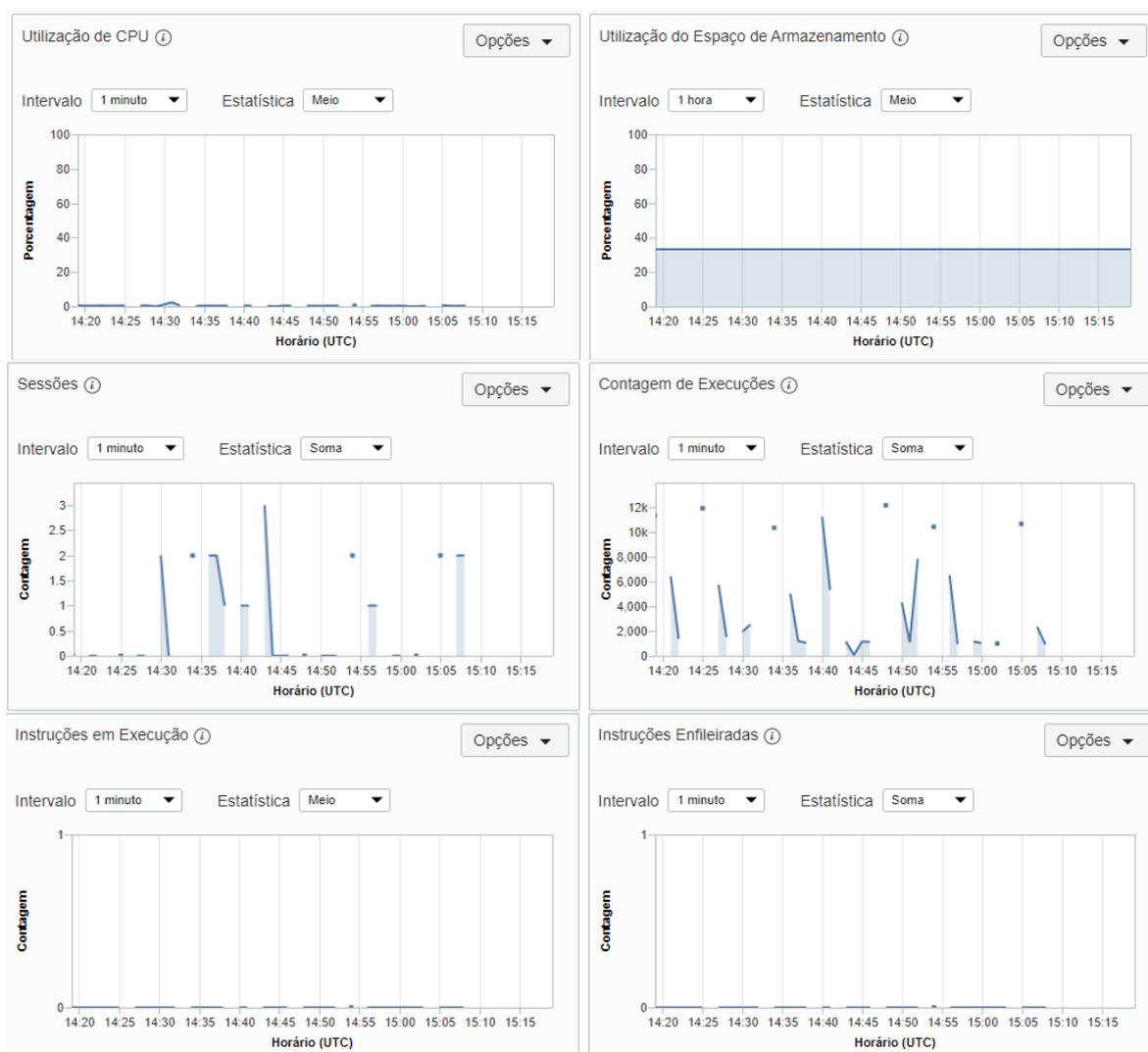
**Tipo de Instância:** Gratuito [Upgrade para Pago](#)

**Modo:** Leitura/Gravação [Editar](#)

Fonte: *Oracle* (2020)

Nessa mesma interface de visão geral, pudemos acompanhar métricas fornecidas por meio de gráficos que demonstram a Utilização do CPU, utilização do espaço de armazenamento, número de sessões, contagem de execuções, Instruções de execução e instruções enfileiradas como mostrado na Figura 4.

Figura 4 - Gráficos de gerenciamento



Fonte: Oracle (2020)

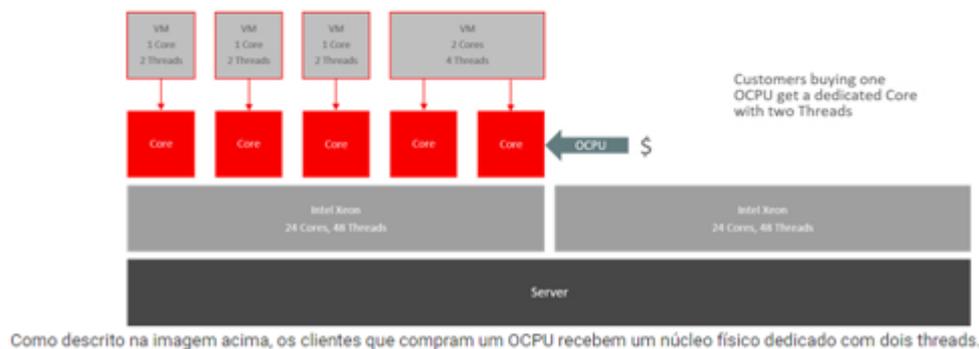
Mesmo após a criação do banco de dados, a *Oracle* permite que sejam redimensionados parâmetros como o espaço de armazenamento e a contagem de OCPUs, que é um medidor de processamento próprio da *Oracle*, como descrito na Figura 5.

Figura 5 - Ciclo do OCPUs

## 2. OCPU

Medida oracle cloud infrastructure para computação. Uma OCPU é definida como a capacidade de CPU equivalente a um núcleo físico de um processador Intel Xeon com hiper threading ativado, ou um núcleo físico de um processador Oracle SPARC.

Para processador Intel Xeon, cada OCPU corresponde a dois threads de execução de hardware, conhecidos como vCPUs. Para processadores Oracle SPARC, um OCPU corresponde a oito threads de execução de hardware, também conhecidos como vCPUs.



Fonte: Nohejl (2020)

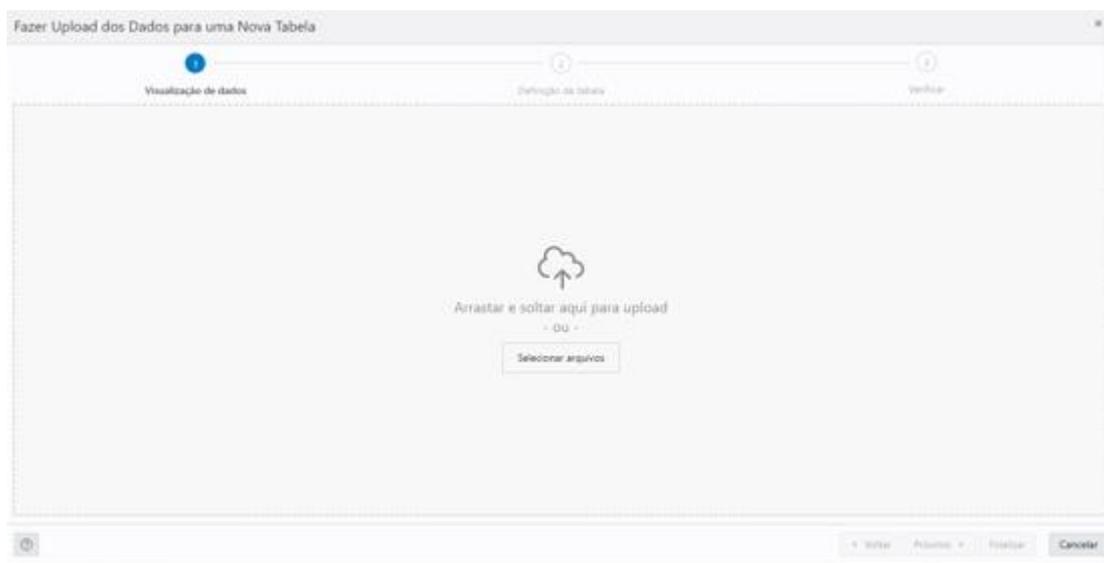
O banco de dados inclui para a versão gratuita, uma ferramenta que gerencia o sistema de dados e o integra a um ambiente de desenvolvimento denominado *SQL Developer Web*. Ou seja, a *Oracle* não exige que seja utilizado um Sistema de Gerenciamento de Banco Dados à parte para gerenciar os dados, como apresentado na Figura 6.

Figura 6 - Ambiente de desenvolvimento: *SQL Developer Web*



Fonte: Oracle (2020)

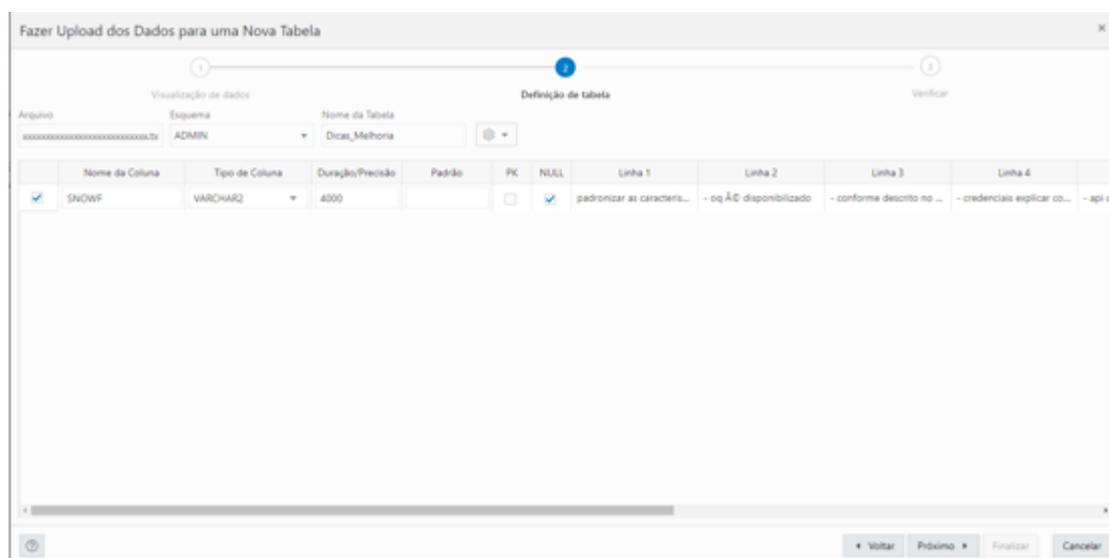
O banco de dados da *Oracle*, quando criado, não dispõe de base de dados predefinida. Porém é possível importar dados por meio de *uploads* de arquivos, como exibido nas Figuras 7,8 e 9.

Figura 7 - Ambiente para *uploads* de arquivos

Fonte: *Oracle* (2020)

Durante o upload a *Oracle* não exige formato específico de arquivo pois após a leitura do arquivo a estrutura da base pode ser editada antes de os dados serem copiados.

Figura 8 - Gerenciar estrutura



Fonte: *Oracle* (2020)

E por fim é possível verificar uma pré-visualização da estrutura que será formada, conforme pode ser observado na Figura 8.



a uma base de dados já refinados para serem adicionadas ao banco de dados criado anteriormente.

Figura 11- Oracle Cloud Infrastructure

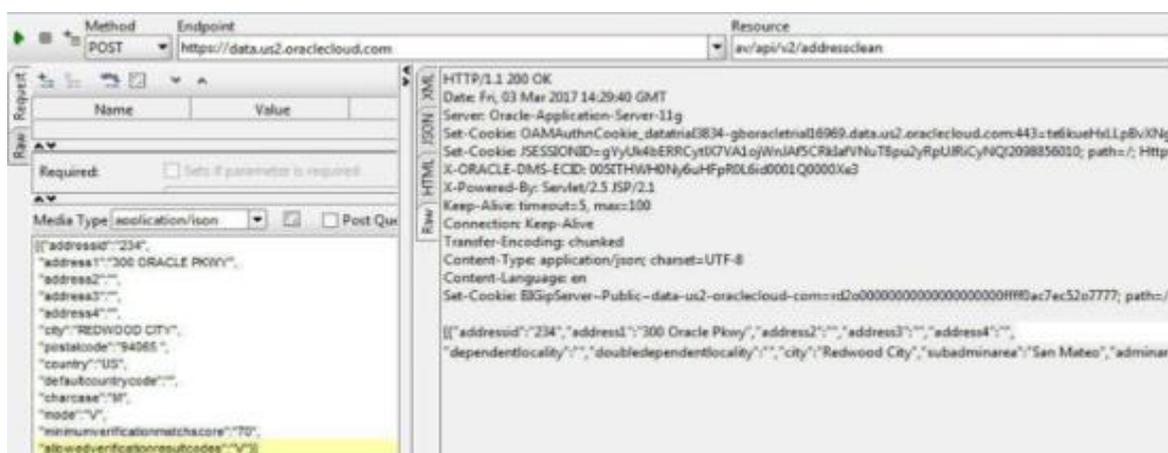
The screenshot shows a search interface for Oracle Cloud Infrastructure. The search term 'oracle' is entered in the search bar, and the results are displayed in a table. The table has columns for Name, Industry Category, Location, Revenue, Employees, and Contacts. The results are filtered by 'Companies in Business' and 'Enterprise Companies'. The table shows 40 results, with the first 15 visible. The results include Oracle Corporation, Oracle Flexible Packagi..., Oracle America, Inc., Oracle Systems Corpora..., Oc Acquisition LLC, Oracle Financial Service..., NetScite Inc., Bea Systems, Inc., Responsys, Inc., Opower, Inc., Micror Systems, Inc., ORACLE POLSKA SP Z..., ORACLE ARGENTINA..., and ORACLE ITALIA SRL.

Name	Industry Category	Location	Revenue	Employees	Contacts
Oracle Corporation	Services	Redwood City, CA, U...	\$37,047,000,000	136,034	8876
Oracle Flexible Packagi...	Manufacturing	Winston Salem, NC...	\$107,488,110	285	15
Oracle America, Inc.	Manufacturing	Redwood City, CA, U...	\$11,478,710,298	29,000	635
Oracle Systems Corpora...	Services	Redwood City, CA, U...	\$4,293,618,293	56,134	158
Oc Acquisition LLC	Services	Redwood City, CA, U...	\$1,797,910,422	5,202	691
Oracle Financial Service...	Wholesale Trade	Edison, NJ, USA	\$106,454,213	693	5
NetScite Inc.	Services	San Mateo, CA, USA	\$741,149,000	4,603	448
Bea Systems, Inc.	Services	San Jose, CA, USA	\$100,804,314	4,278	14
Responsys, Inc.	Services	San Bruno, CA, USA	\$113,134,675	866	78
Opower, Inc.	Services	Arlington, VA, USA	\$148,691,000	599	0
Micror Systems, Inc.	Services	Columbia, MD, USA	\$794,935,747	6,506	231
ORACLE POLSKA SP Z...		Warszawa, POLAND	\$158,655,939	200	4
ORACLE ARGENTINA...		Ciudad de Buenos A...	\$138,782,924	439	5
ORACLE ITALIA SRL		CNISELLO BALSAM...	\$257,065,588	1,148	11

Fonte: Bingham (2017)

Outra característica *DaaS* que podemos identificar, é a utilização de API's para refinar os dados. A *Oracle* dispõe de API's para determinados serviços de refinamento de dados. Como por exemplo a API de verificação de endereços, *Oracle DaaS for Sales* e *Oracle DaaS for Marketing*. A Figura 12 apresenta uma interface da API de verificação de endereços.

Figura 12- API's



Fonte: Bingham (2017)

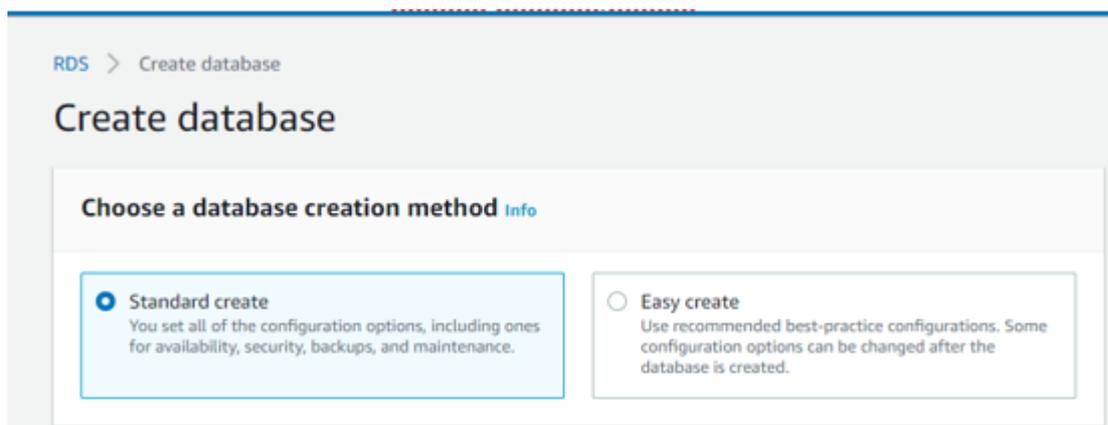
Após o refinamento dos dados pelas API's, eles podem ser encontrados no banco de dados vinculado à API em um repositório à parte. Ficando assim, disponíveis para serem utilizados ou compartilhados em um *Marketplace*.

### 3.2 Amazon RDS

Ao acessar o site da *Amazon*, logo na tela inicial é possível identificar que podemos utilizar as funcionalidades do banco de dados na Nuvem por um cadastro gratuito de 12 meses, com 20 GB de armazenamento disponíveis.

Quando o *login* é realizado, *DashBoard* da *Amazon RDS* é acessado inicialmente, e aparece um botão para criar um banco de dados ou restaurar algum outro banco do servidor S3 da *Amazon* como demonstrado na Figura 13. Na criação do banco de dados, podemos selecionar se desejamos configurar manualmente todo o banco de dados ou se preferimos usar configurações padrões que a própria *Amazon* dispõe.

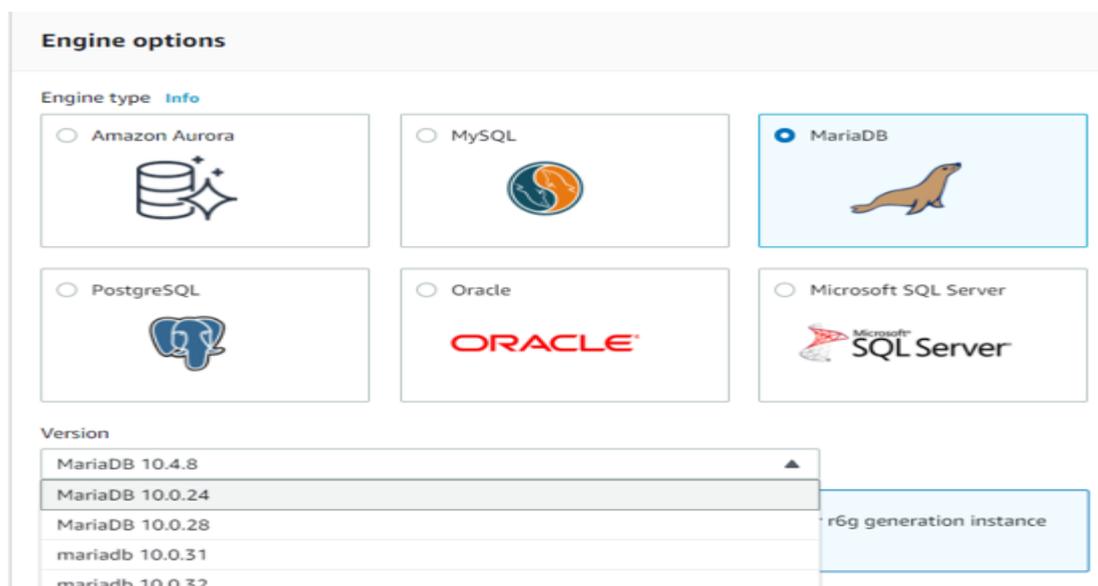
Figura 13- Dashboard da Amazon RDS



Fonte: Amazon (2020)

Logo em seguida podemos selecionar o *engine type*, que é o mecanismo de armazenamento do Banco de Dados. De acordo com a Figura 14, a Amazon oferece seis tipos de *engine*, sendo elas Amazon Aurora, MySQL, MariaDB, PostgreSQL, Oracle e Microsoft SQL Server. Assim que algum é selecionado, é possível escolher a versão da *engine* que deseja utilizar.

Figura 14- Engine Type

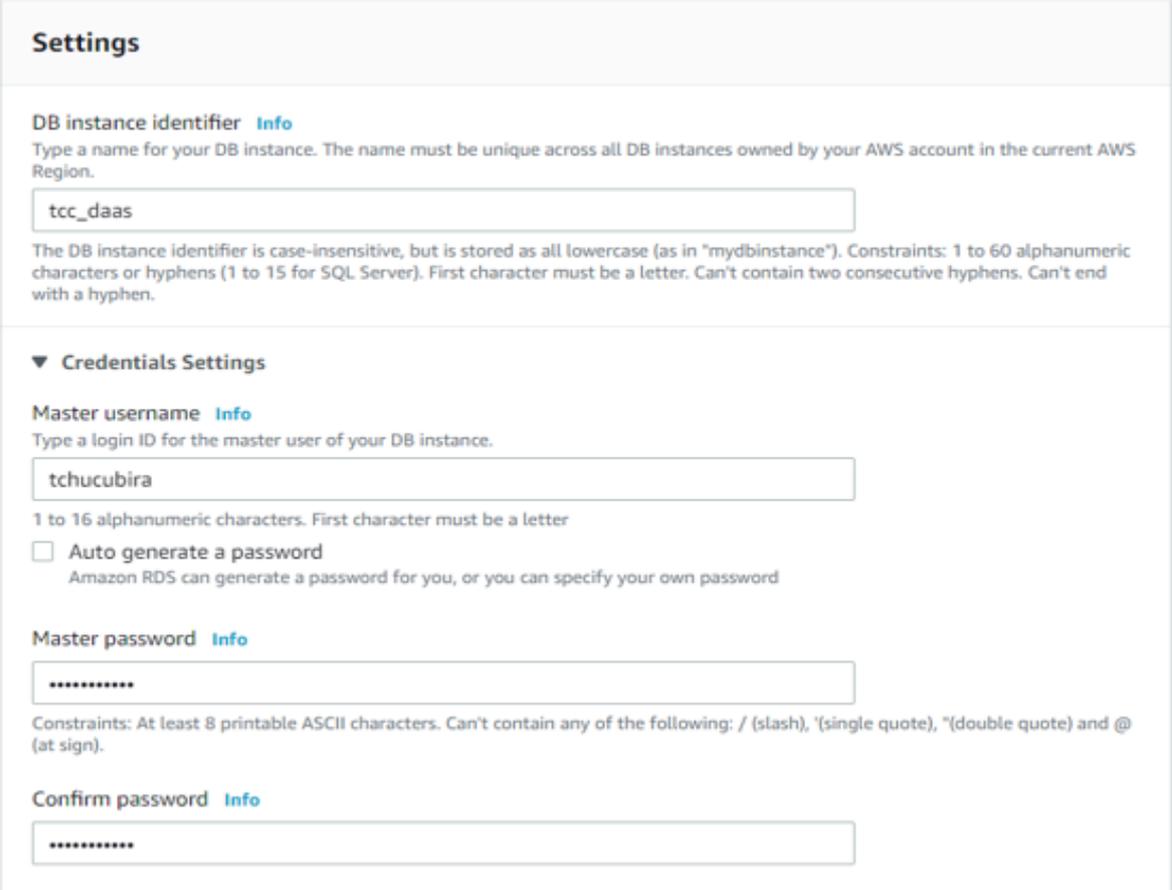


Fonte: Amazon (2020)

Como é mostrado na Figura 15, após a seleção da *engine* é preciso cadastrar credenciais para ter acesso ao banco de dados, pois no caso da Amazon, é necessário a

utilização de um *software* de gerenciamento de banco de dados (SGBD) para realizar o manuseio das informações estruturas que serão inseridas no mesmo.

Figura 15- Configurações do Banco de Dados



**Settings**

**DB instance identifier** [Info](#)  
Type a name for your DB instance. The name must be unique across all DB instances owned by your AWS account in the current AWS Region.

tcc\_daas

The DB instance identifier is case-insensitive, but is stored as all lowercase (as in "mydbinstance"). Constraints: 1 to 60 alphanumeric characters or hyphens (1 to 15 for SQL Server). First character must be a letter. Can't contain two consecutive hyphens. Can't end with a hyphen.

▼ **Credentials Settings**

**Master username** [Info](#)  
Type a login ID for the master user of your DB instance.

tchucubira

1 to 16 alphanumeric characters. First character must be a letter

**Auto generate a password**  
Amazon RDS can generate a password for you, or you can specify your own password

**Master password** [Info](#)

\*\*\*\*\*

Constraints: At least 8 printable ASCII characters. Can't contain any of the following: / (slash), '(single quote), "(double quote) and @ (at sign).

**Confirm password** [Info](#)

\*\*\*\*\*

Fonte: Amazon (2020)

Os delimitadores de capacidade de processamento dos dados são definidos de acordo com a *engine* selecionada, suas versões e o plano adquirido. Para o grátis e o MariaDB 10.4.8, é disponibilizado somente uma CPU e 1 GB RAM para o processamento dos dados como exibido na Figura 16.

Figura 16- Delimitadores de Capacidade

**DB instance size**

**DB instance class** [Info](#)  
Choose a DB instance class that meets your processing power and memory requirements. The DB instance class options below are limited to those supported by the engine you selected above.

- Standard classes (includes m classes)
- Memory Optimized classes (includes r and x classes)
- Burstable classes (includes t classes)

db.t2.micro  
1 vCPUs 1 GiB RAM Not EBS Optimized ▼

[i](#) New instance classes are available for specific engine versions. [Info](#)

Include previous generation classes

---

**Storage**

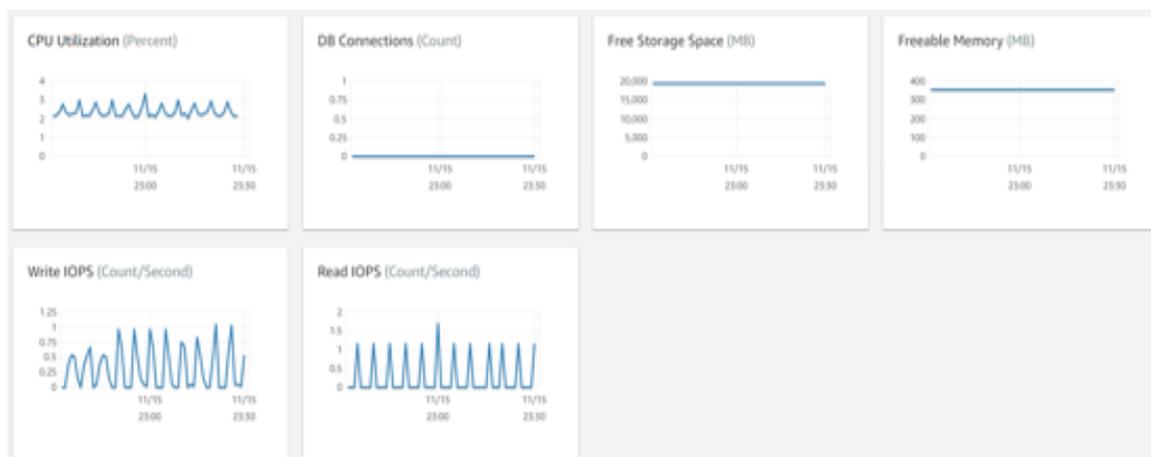
**Storage type** [Info](#)  
General Purpose (SSD) ▼

**Allocated storage**  
20 GiB

Fonte: Amazon (2020)

A *Amazon* também dispõe de ferramentas gráficas de monitoramento dos recursos utilizados, como por exemplo porcentagem de utilização da CPU, contador de requisições de conexões no banco de dados, espaço de armazenamento livre e quantidade (MB) de memória, como apresentado na Figura 17.

Figura 17- Ferramentas gráficas de monitoramento

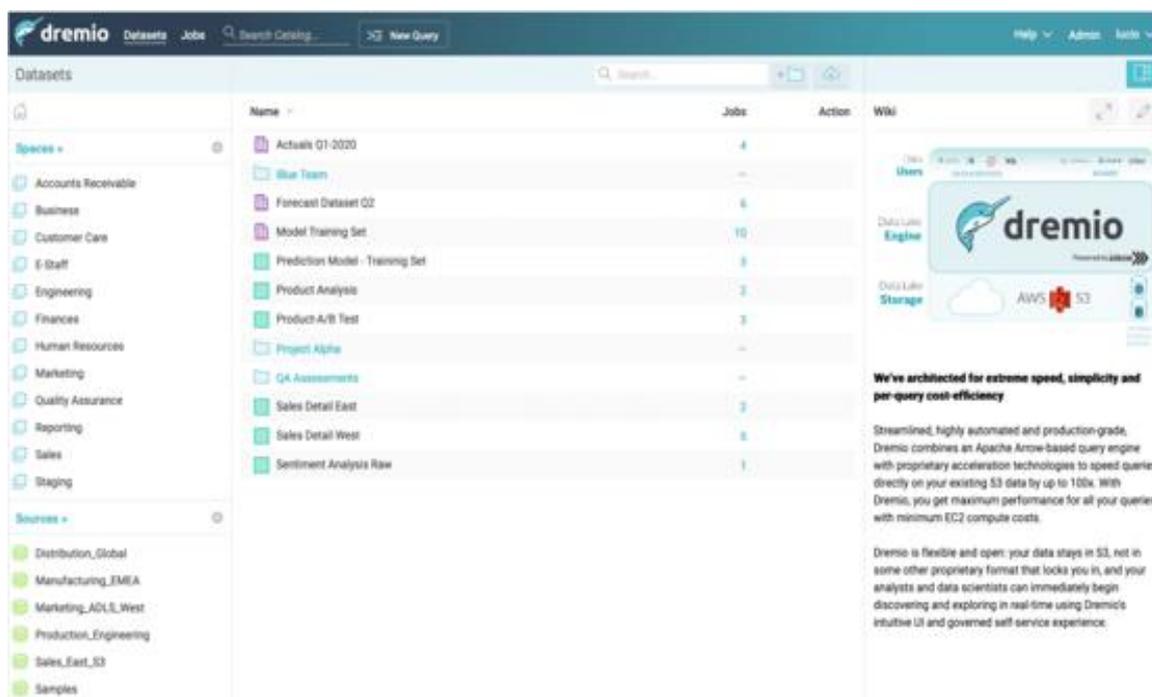


Fonte: Amazon (2020)

Os bancos de dados da *Amazon* por si, não possuem um mecanismo de processamento de dados para refiná-los e constituir uma base que pode ser fornecida como *Data as a Service*. Portanto necessita de uma aplicação externa.

Dremio é uma API que pode fazer conexão com um banco de dados localizado no servidor S3 da *Amazon*. Assim que vinculados, a API está apta a realizar o refinamento dos dados desejados. A Figura 18 mostra uma breve interface da Dremio API.

Figura 18- Dremio API



Fonte:(Amazon, 2020)

### 3.3 Snowflake

Ao acessar o site da *Snowflake*, na tela inicial fica explícito que podemos utilizar seus serviços de Banco de Dados na Nuvem por um cadastro *Free Trial* por 30 dias.

O cadastro *Free Trial* fornece 400 reais de crédito para serem usados gratuitamente. O armazenamento e a CPU são ilimitados, porém a sua utilização consome créditos, e consequentemente quando esses créditos grátis se esgotarem a utilização do *Free Trial* é encerrada. Para continuar utilizando os serviços da *Snowflake*, uma recarga de créditos ou assinatura de um plano é necessária.

A *Snowflake* dispõe por padrão, de três Bancos de Dados relacionais já criados, sendo um deles já preenchido com alguns dados para teste. É possível também criar outros Bancos de Dados além dos pré-existentes. A administração desses bancos é feita pelo SGBD

da própria *Snowflake*. Na criação<sup>1</sup> do Banco, inicialmente foi preciso digitar somente o nome do banco do usuário para identificação posterior. No momento de criação das tabelas, depois de atribuir um nome a ela, podemos adicionar suas colunas. As colunas são criadas somente com quatro características, sendo elas, nome, tipo de dado, valor padrão e um *checkbox* para informar se o campo da coluna pode ou não ser NULL conforme pode ser observado na Figura 20.

Figura 19 - Criação das Tabelas

**Create Table**

Table Name \*

Comment

Columns \*

Name	Type	Not Null	Default
Coluna 1	STRING	<input checked="" type="checkbox"/>	S
C2	STRING	<input type="checkbox"/>	
C3	STRING	<input type="checkbox"/>	

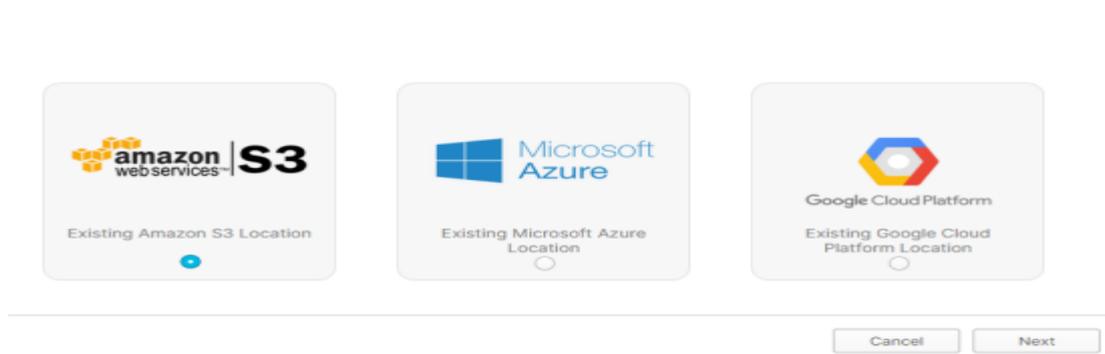
[Show SQL](#)

Fonte:(*Snowflake*, 2020)

*Snowflake* permite que sejam criadas *Stages* que são separadores de partimentos do banco de dados , vinculando dados inseridos no SGBD da *Snowflake* a outros provedores como *Amazon*, *Microsoft Azure* e *Google Cloud Platform*.

<sup>1</sup> <https://kza86317.us-east-1.snowflakecomputing.com/console#/data/stages?databaseName=HEHE>

Figura 20- Stages



Fonte:(Snowflake, 2020)

Ao selecionar um provedor, é necessário inserir as credenciais dos servidores para que os dados sejam redirecionados.

Figura 21 - Credenciais dos servidores

S3 Amazon	Microsoft Azure	Google Cloud
<p><b>Create Stage</b></p> <p>Staged files will be stored in the specified S3 location</p> <p>Name * <input type="text"/></p> <p>Schema Name PUBLIC</p> <p>URL * s3://</p> <p>AWS Key ID <input type="text"/></p> <p>AWS Secret Key <input type="text"/></p> <p>Encryption Master Key <input type="text"/></p> <p>Comment <input type="text"/></p> <p>Show SQL</p>	<p><b>Create Stage</b></p> <p>Staged files will be stored in the specified Azure location</p> <p>Name * <input type="text"/></p> <p>Schema Name PUBLIC</p> <p>URL * azure://</p> <p>Azure SAS Token <input type="text"/></p> <p>Encryption Master Key <input type="text"/></p> <p>Comment <input type="text"/></p> <p>Show SQL</p>	<p><b>Create Stage</b></p> <p>Staged files will be stored in the specified GCS location</p> <p>Name * <input type="text"/></p> <p>Storage Integration * <input type="text"/></p> <p>URL * gcs://</p> <p>Comment <input type="text"/></p> <p>Show SQL</p>

Fonte:(Snowflake, 2020)

Em relação à exportação dos dados, é possível customizar a maneira que os dados serão compactados de acordo com o formato do arquivo selecionado para exportação. Como por exemplo, se for selecionado a exportação para um arquivo do tipo CSV, podemos personalizar o separador de colunas, o separador de dados, método de compressão (ZIP, etc), quantidade de linhas para pular antes dos dados serem colocados no arquivo, entre outros.

Figura 22- Exportação dos dados

Format Type	CSV	▼
Compression Method	Auto	▼
Column separator	Comma	▼
Row separator	New Line	▼
Header lines to skip	0	▲ ▼
Field optionally enclosed by	None	▼
Null String	\\N	▼

Fonte:(Snowflake, 2020)

Caso seja selecionado a exportação para um tipo de arquivo JSON, podemos escolher o formato dos dados que compõem o arquivo exportado. Dentre as opções, é configurável para arquivo exportado a inclusão de objetos duplicados, remoção valores vazios (*NULL*), remoção de erros ocasionados pelo *encoding UTF-8*, entre outros.

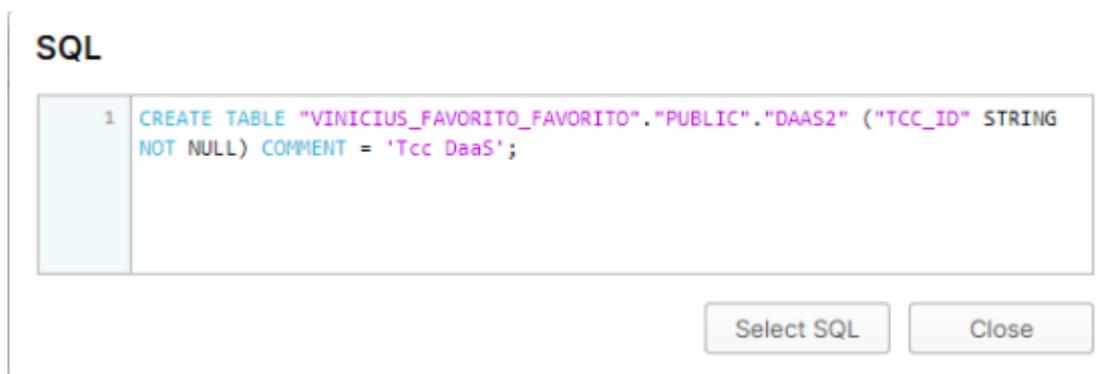
Figura 23- Exportação dos dados em arquivo JSON

Format Type	JSON	▼
Compression Method	Auto	▼
	<input checked="" type="checkbox"/> Enable Octal	?
	<input checked="" type="checkbox"/> Allow Duplicate	?
	<input checked="" type="checkbox"/> Strip Outer Array	?
	<input checked="" type="checkbox"/> Strip Null Values	?
	<input checked="" type="checkbox"/> Ignore UTF-8 Errors	?

Fonte:(Snowflake, 2020)

Na tela de criação dos bancos, tabelas, *schemas*, arquivos de exportação, *Snowflake* apresenta botão que possibilita a visualização dos *scripts* SQL executados por trás da interface. Conforme a imagem abaixo, exibe o *script* que é executado na criação de uma tabela.

Figura 24- Criando uma tabela



The image shows a screenshot of the Snowflake SQL editor interface. At the top left, the word "SQL" is displayed in a bold, black font. Below it, a text area contains a single line of SQL code: `1 CREATE TABLE "VINICIUS_FAVORITO_FAVORITO"."PUBLIC"."DAAS2" ("TCC_ID" STRING NOT NULL) COMMENT = 'Tcc DaaS';`. The code is color-coded: "1" is in light blue, "CREATE TABLE" is in light blue, the table name and schema are in purple, and the column definition and comment are in black. Below the text area, there are two buttons: "Select SQL" and "Close", both in a light gray color.

Fonte:(*Snowflake*, 2020)

Uma das características *DaaS* que podemos identificar na API *Snowflake* é o compartilhamento de dados. Porém não é possível utilizá-lo pela versão gratuita.

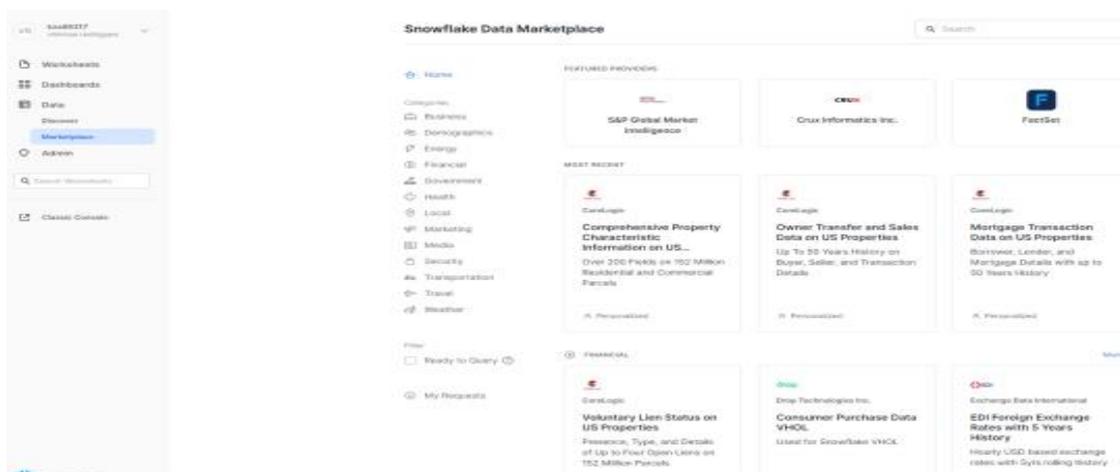
Figura 25- Compartilhamento de dados



Fonte:(*Snowflake*, 2020)

Outra característica identificada é o *Snowflake Data Marketplace*, onde podemos ver vários grupos de dados de várias categorias.

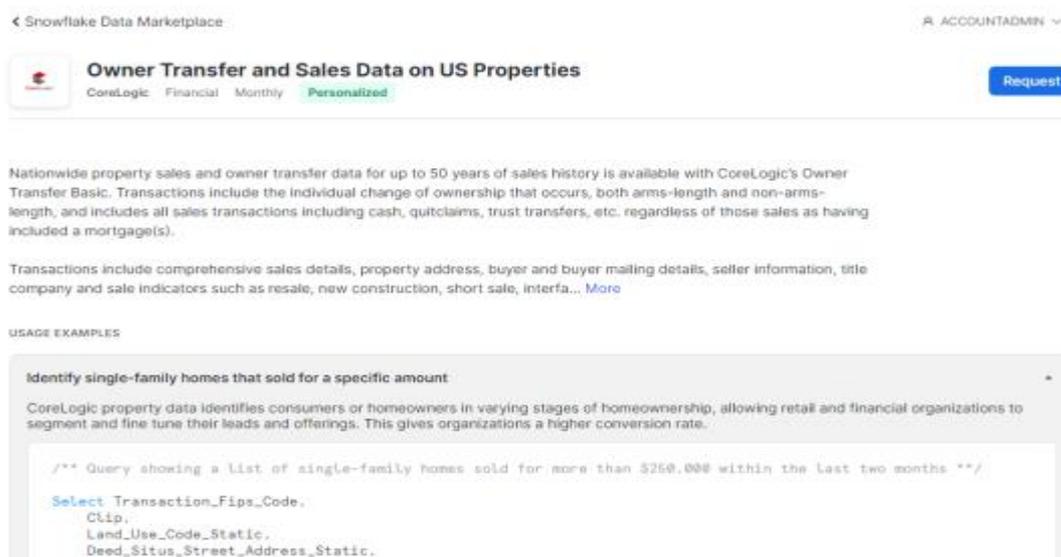
Figura 26- *Snowflake Data Marketplace*



Fonte:(Snowflake, 2020)

Ao selecionar um item, uma breve descrição sobre a empresa que está disponibilizando os dados e sobre os próprios dados disponíveis. Diante disso, por uma interação na interface, podemos solicitar ou adquirir diretamente esses dados.

Figura 27- Breve descrição sobre a empresa que está disponibilizando os dados



Fonte:(Snowflake, 2020)

Quando os dados podem ser adquiridos diretamente, basta adicionarmos o nome do banco de dados que será criado em seu domínio e selecionar as permissões necessárias para acessá-lo. Caso contrário, é necessário fazer uma solicitação para que a empresa disponibilize os dados, conforme a Figura 28.

Figura 28- Permissões para ter os dados

Fonte:(Snowflake, 2020)

### 3.5. Comparação entre os Provedores DaaS

Diante da descrição do funcionamento de três provedores de serviços *DaaS*, *Snowflake*, *Oracle* e *Amazon*, foram selecionadas características gerais desses provedores e listados em um *checklist*. No intuito de visualizar e esclarecer diferenças e semelhanças entre os provedores analisados, identificar a melhor situação para agregá-los na arquitetura de um sistema que necessita de dados como um serviço.

Tabela 1 - Checklist Proposto

-	<i>Snowflake</i>	<i>Oracle</i>	<i>Amazon</i>
O processamento dos dados é feito diretamente por uma			x

API externa do provedor.			
Depende de API ou serviço à parte para refinar os dados.	<i>x</i>	<i>x</i>	<i>x</i>
Possível compartilhamento dos dados sem necessidade de estarem inseridos no Marketplace por uma ferramenta direta do SGBD.	<i>x</i>		
Os dados refinados podem ser exportados ou incluídos em um Marketplace.	<i>x</i>	<i>x</i>	<i>x</i>
Banco de dados possui um provedor próprio.		<i>x</i>	<i>x</i>
Dados refinados são enviados automaticamente de	<i>x</i>		

volta para o banco de dados.			
Depende de um SGBD a parte para manusear os dados.			x
Utiliza banco de dados relacional.	x	x	x
O Refinamento funciona com dados estruturados em bancos NoSQL.	x	x	x

## 4. Limitações

Diante da listagem de provedores *DaaS*, observamos que em sua maioria, para que possam ser utilizados, seria necessário adquirir uma assinatura mensal ou inserção de cartões de créditos. Portanto, como não temos condições de realizar tais ações para usufruir da tecnologia, nossa listagem comparativa foi reduzida para análise somente de provedores que fornecem serviços gratuitos. Abaixo está algumas das ferramentas que encontradas que não poderiam agregar na análise:

- *WhoisXML API Enterprise API*
- *SAP HANA Cloud*
- *Spectrum Geocoding on Demand*
- *Wizdom.ai*
- *Fiserv Wealth Management Network*
- *Nexis data as a service*

## 5. Considerações Finais

Este trabalho teve como objetivo realizar pesquisas exploratórias e esclarecer qual seria a melhor situação e provedor que se encaixa na atual situação da empresa dos leitores interessados em aderir os serviços *DaaS*. Com um embasamento teórico de conceitos de armazenamento de dados em nuvem, sobre as áreas que envolvem o *DaaS*, sobre como o serviço *DaaS* auxiliará as empresas.

Foram selecionados alguns *softwares* de *DaaS* disponíveis no mercado é apresentado um estudo comparativo de suas características. Os critérios utilizados para criação do *checklist* serão desenvolvidas de acordo com as características apresentadas no estudo comparativo entre os provedores, descrevendo alguns aspectos relacionados ao *DaaS*, e que por sinal também serão respondidos com base na mesma referência.

Como trabalho futuro, é desejável expandir o *checklist* para abordar perguntas referentes à evolução da tecnologia, visando melhorar a qualidade da utilização do *DaaS*.

## 6. Referências Bibliográficas

BEDGOOD, Larisa. Real-Life Examples of Companies Using Data-as-a-Service to Boost Revenue. [S. l.], 15 mar. 2015. Disponível em: <https://datafloq.com/read/real-life-examples-of-companies-using-DaaS/961>. Acesso em: 19 nov. 2020.

BINGHAM, Richard. **Platform Services: Integrating With The Oracle Data-as-a-Service Cloud**. 17 mar. 2017. 886 x 666. Disponível em: <https://blogs.oracle.com/fadevrel/platform-services%3a-integrating-with-the-oracle-data-as-a-service-cloud>. Acesso em: 15 nov. 2020.

DATA as a Service (*DaaS*). Direção: Joel Battistoni. <https://www.youtube.com/watch?v=ayS5yvdg2Tg>: [s. n.], 2015. Disponível em: <https://www.youtube.com/watch?v=ayS5yvdg2Tg>. Acesso em: 16 nov. 2020.

DATA SCIENCE ACADEMY (org.). **BIG DATA E CLOUD COMPUTING – DESAFIOS E OPORTUNIDADES**. [S. l.], 6 set. 2017. Disponível em: <http://datascienceacademy.com.br/blog/big-data-e-cloud-computing-desafios-e-oportunidades/>. Acesso em: 21 abr. 2020.

DEVMEDIA, (Rio de Janeiro - RJ) (org.). **Big Data Tutorial**. [S. l.], 2014. Disponível em: <https://www.devmedia.com.br/big-data-tutorial/30918>. Acesso em: 19 maio 2020.

DREMIO'S DATA LAKE ENGINE (ed.). What is Data-as-a-Service?. *In* : **What is Data-as-a-Service?**. [S. l.], 2019. Disponível em: <https://www.dremio.com/what-is-data-as-a-service/>. Acesso em: 1 abr. 2020.

DREMIO'S DATA LAKE ENGINE (ed.). **Deploy an Unparalleled AWS Data Lake Solution**. 2019. 886 x 666. Disponível em: <https://www.dremio.com/aws/>. Acesso em: 15 nov. 2020.

EQUIPE DE REDAÇÃO (org.). **AFINAL, O QUE É CLOUD COMPUTING?**. [S. l.], 3 set. 2019. Disponível em: <https://blog-br.softwareone.com/afinal-o-que-e-cloud-computing>. Acesso em: 22 maio 2020.

EQUIPE DE REDAÇÃO. **A Internet completa 50 anos cercada de polêmicas e desafios**. Disponível em: <https://www.bahiadevalor.com.br/2019/11/internet-completa-50-anos-cercada-de-polemicas-e-desafios/>. Acesso em 28 de maio de 2020.

FELDMAN, DAMON. **DaaS Success Factors and Risks**. [S. l.], 2020. Disponível em: <https://www.marklogic.com/blog/data-as-a-service-success-factors-risks/>. Acesso em: 24 maio 2020.

GIRALDO, Valentina. **Descubra finalmente o que é o cloud computing e para que serve a computação em nuvem**. Disponível em: <https://rockcontent.com/blog/cloud-computing/>. Acesso em 23 de abr. de 2020.

GUIMARÃES, Leandro. **Big Data VS: você sabe quais são os Vs do Big Data?**. [S.l.], 22 out. 2018. Disponível em: <https://www.knowsolution.com.br/big-data-vs/>. Acesso em: 24 abr. 2020.

KANTARCI, Atakan. **Data as a Service (DaaS): What, Why, How, Use cases & Tools**. [S. l.], 13 nov. 2020. Disponível em: <https://research.aimultiple.com/data-as-a-service/#what-is-dataasaservice>. Acesso em: 19 nov. 2020.

LOMBARDO, Marcelo. **Big Data para leigos – entenda em 5 minutos por que isso muda a forma como vemos o mundo**. [S. l.], 26 set. 2013. Disponível em: [https://marcelo.omie.com.br/2013/09/big-data-para-leigos-1.html?utm\\_source=Blog+ML&utm\\_campaign=da0960b41e-Blog\\_ML\\_BigData26set13&utm\\_medium=email&utm\\_term=0\\_a9c6809433-da0960b41e-25251137](https://marcelo.omie.com.br/2013/09/big-data-para-leigos-1.html?utm_source=Blog+ML&utm_campaign=da0960b41e-Blog_ML_BigData26set13&utm_medium=email&utm_term=0_a9c6809433-da0960b41e-25251137). Acesso em: 20 maio 2020.

MCDANIEL, Stacey. **What is Data as a Service?**. [S. l.], 30 set. 2019. Disponível em: <https://www.talend.com/resources/what-is-data-as-a-service/>. Acesso em: 24 maio 2020.

NOGUEIRA, Luiz. **Dados mostram o crescimento impressionante da internet em 10 anos**. [S. l.], 17 maio 2019. Disponível em: <https://olhardigital.com.br/noticia/dados-mostram-o-crescimento-impressionante-da-internet-em-10-anos/85914>. Acesso em: 28 maio 2020.

NORMANDEAU, Kevin. Beyond Volume, Variety and Velocity is the Issue of *Big Data* Veracity. [S.l.], 13 set. 2013. Disponível em: <https://insidebigdata.com/2013/09/12/beyond-volume-variety-velocity-issue-big-data-veracity/>. Acesso em: 26 abr. 2020.

NOHEJL, Vít. **What You Need to Know: Cloud Services Pricing - vCPU vs. OCPU?**. 2020. 886 x 666. Disponível em: <https://www.industry-era.com/Cloud-Services-Pricing.php>. Acesso em: 15 nov. 2020.

OPTIDATA. **Entenda a relação entre Cloud Computing e Big Data**. [S.l.], 2018. Disponível em: <https://www.optidatacloud.com/blog/entenda-relacao-entre-cloud-computing-e-big-data/>. Acesso em 29 maio 2020

REDHAT. **O que é Big Data?**. [S.l.], 2020. Disponível em: <https://www.redhat.com/pt-br/topics/big-data/>. Acesso em 29 maio 2020

RUSSO, BRUNO. **Computerworld - Nuvem pública, privada ou híbrida? Entenda as diferenças**. Disponível em: <https://computerworld.com.br/2017/06/14/nuvem-publica-privada-ou-hibrida-entenda-diferencas/>. Acesso em 24 de abr. de 2020.

SANTANNA, Miguel. **Entenda o que é e como funciona o software como serviço (SaaS)**. [S. l.], 23 fev. 2018. Disponível em: <https://rockcontent.com/blog/saas/>. Acesso em: 24 maio 2020.

SANTOS, Pedro. **Big Data: o que é, para que serve e como aplicar?**. [S. l.], 5 out. 2018. Disponível em: <https://inteligencia.rockcontent.com/big-data/>. Acesso em: 23 abr. 2020.

SAS INSIGHTS (org.). **Big Data: What it is and why it matters**. [S. l.], 2019. Disponível em: [https://www.sas.com/pt\\_br/insights/big-data/what-is-big-data.html](https://www.sas.com/pt_br/insights/big-data/what-is-big-data.html). Acesso em: 23 abr. 2020.

SOARES, Guilherme. **Para onde vai o armazenamento de dados com o *Big Data*?**. [S. l.], 30 nov. 2014. Disponível em: <https://cio.com.br/para-onde-vai-o-armazenamento-de-dados-com-o-big-data/>. Acesso em: 19 maio 2020.

TAKE, Administrator. **O que são os 5 V's do *Big Data*?**: Conceito + Aplicação. [S.l.], 18 jan. 2017. Disponível em: <https://take.net/blog/inovacao/5-vs-big-data>. Acesso em: 24 abr. 2020.

TERZO, Olivier; BUCCI, Enrico; RUIU, Pietro; XHAFI, Fatos. *Data as a Service (DaaS) for Sharing and Processing of Large Data Collections in the Cloud*. **Seventh International Conference on Complex, Intelligent, and Software Intensive Systems**, [s. l.], p. 2-7, 2013. DOI 10.1109/CISIS.2013.87. Disponível em: [https://www.researchgate.net/publication/261524559\\_Data\\_as\\_a\\_Service\\_DaaS\\_for\\_Sharing\\_and\\_Processing\\_of\\_Large\\_Data\\_Collections\\_in\\_the\\_Cloud](https://www.researchgate.net/publication/261524559_Data_as_a_Service_DaaS_for_Sharing_and_Processing_of_Large_Data_Collections_in_the_Cloud). Acesso em: 10 nov. 2020.