

Estudo de análise de dados em Big Data com foco na otimização de fluxo de veículos em Cidades Inteligentes

Lucas Candido SIMEÃO, Almir Rogério CAMOLESI

lucas_simeao@hotmail.com, camolesi@femanet.com.br

RESUMO: Smart Cities ou Cidades Inteligentes é um conceito que com o início da internet se desenvolveu para trazer melhores técnicas de gerenciamento e melhor desenvolvimento das cidades. Tal conceito possibilitou uma quebra de paradigma das cidades convencionais para uma cidade em que todo o meio pudesse estar conectado e com o auxílio da tecnologia e novas possibilidades foram surgindo. Com essas novas oportunidades chegando, viu-se que um grande volume de dados era criado todos os dias denominado Big Data e precisava ser armazenado em um local onde não teria perda e que se tornasse útil para a cidade. Nesse contexto que foi desenvolvido esse projeto, que teve por objetivo a coleta desses dados e armazená-los em um local que esteja disponível para todos a qualquer momento.

PALAVRAS-CHAVE: Smart; Cities; Big; Data; Nuvem.

ABSTRACT: Smart Cities is a concept that with the beginning of the internet was developed to bring better management techniques and better development of the cities. It became possible that there would be a paradigm shift from conventional cities to a city where all the medium could be connected and with the aid of technology new possibilities were emerging. With these new opportunities coming, it was seen that a large amount of data was created every day called Big Data and needed to be stored in a place where it would not be lost and that would become useful for the city. In this context is the project, which aims to collect these data and store them in a location that is available to everyone at any time.

KEYWORDS: Smart; Cities; Big; Data; cloud.

1. Introdução

Primeiramente para esse estudo, foi necessário entender sobre os conceitos de Smart Cities ou Cidades Inteligentes que tem por objetivo melhorar a vida das pessoas em questões ao gerenciamento de recursos e ao aproveitamento do tempo. Na década de 90, com o avanço da internet, a tecnologia começou a fazer parte do dia a dia das pessoas, possibilitando que esses conceitos pudessem se tornar realidade. Com o aumento

crecente das pessoas nos centros urbanos, muitos problemas foram surgindo de modo que as cidades tiveram que se adaptar a essas mudanças e o uso da tecnologia foi uma das saídas para transformar esse problema de infraestrutura e aglomerações de pessoas em uma oportunidade para novos negócios e um novo mercado foi surgindo em conjunto. Com o passar do tempo vemos que algo que era muito raro se tornou comum nos dias de hoje, antigamente existia 3 grandes cidades que possuía uma população que ultrapassava a marca de um milhão de habitantes. Hoje em dia, isso é uma realidade em quase todos os lugares do mundo, pois muitas pessoas abandonaram os campos e foram para os centros urbanos. E com tantas pessoas juntas assim é normal surgirem problemas que antes não existia. (DELOITTE, 2017).

Com a preocupação crescente de se ter um melhor engajamento da cidade com a tecnologia, alguns países decidiram construir do zero cidades nas quais todas as características da mesma pudessem ser envolvidas de algum modo com novas técnicas e não mais dos velhos conceitos que existem. Alguns exemplos são Songdo na Coreia do Sul e Masdar em Dubai e São Gonçalo no Brasil que possuem um melhor aproveitamento de tudo que a cidade possa oferecer (LAGUNA, 2017).

Segundo André Lemos (2013), as Cidades Inteligentes constituem de um grande volume de dados que estão ligados em redes de nuvens com comunicações autônomas entre vários objetos que fazem parte de sua arquitetura. Logo o termo Inteligente é um sinônimo no qual toda a área da cidade produz, consome e distribui uma grande quantidade de dados em tempo real.

Logo, toda a cidade por ser inteligente ele de algum modo terá uma grande coleta de informações, sendo por meio de sensores ou alguns dispositivos eletrônicos mais conhecidos por Internet das Coisas (IoT), e com isso uma gama de dados serão coletados e nesse contexto entra o Big Data será o lugar onde os mesmos serão armazenados. Os dados possuem uma estrutura adequada para o local onde serão armazenados e um Big Data pode ser implementada de várias maneiras e para isso um local para armazenar todos esses dados é de suma importância para toda a gestão dessa cidade e o meio mais confiável seria a nuvem, que por sua vez pode ter um grande volume de dados e praticamente zero de perda dos mesmos (PORTOBELLO, 2018).

Este estudo foi dividido em 7 seções. Na primeira seção foi introduzida as questões gerais sobre Smart Cities, Big Data e Nuvem que foram os temas principais desse estudo. Na segunda apresentou o estudo sobre Smart Cities, seus conceitos e história. A terceira mostra todo o estudo sobre Big Data e seus conceitos e técnicas que são utilizadas. Na quarta seção é sobre Nuvem, onde suas tecnologias e conceitos são explanados. Já na

seção 5 é apresentado a placa Raspberry Pi, com suas características e funcionalidades. Na seção 6 é o desenvolvimento onde é apresentado todas as técnicas utilizadas e como tudo foi realizado. E por fim, na seção 7 são apresentadas as conclusões geradas desse projeto, juntamente com alguns projetos futuros que podem ser realizados, dando sequência aos estudos.

2. Smart Cities

Desde os tempos antigos os seres humanos veem a necessidade de sair do campo e ir para um local em que possam se reunir e viverem juntos, e com isso vários problemas começaram a surgir e a busca por soluções teve seu início.

No início da década de 90, começou a se difundir o termo cidades digitais, que visava a utilização da tecnologia para ajudar a minimizar os problemas que ocorrem até hoje em dia em nossas cidades.

Alguns anos depois esse com o avanço da internet o termo foi mudado para Smart Cities, que abrange um tema geral e que poderia ser aplicado a toda a cidade e não apenas um aspecto. Antigamente existia poucas cidades que possuíam mais de um milhão de pessoas, hoje existe centenas de cidades que ultrapassa facilmente esse número. Com isso novos problemas foram surgindo e mesmo com a tecnologia do nosso lado e a nosso favor, muitos deles ainda são um desafio para o mundo todo e uma busca para solucionar eles é algo que as pessoas buscam diariamente.

Logo uma busca por uma infraestrutura melhor em que a tecnologia possa ajudar as pessoas a ter uma qualidade de vida em que sejam capazes de desfrutar e ter um melhor proveito de tudo que a cidade pode oferecer é o que buscou-se nesse trabalho.

Para uma cidade se considerar inteligente, todo o meio deve ter sensores captando dados a todos os instantes, como reconhecimento facial, temperatura ou até mesmo sensor de fluxo de veículos ou pedestres. E todos esses dados de certa forma sozinhos não podem fazer nada que ajude a melhorar a vida na cidade.

Com a quantidade abundante de dados que são coletados, a cidade precisa de um meio para armazenar todos esses dados de forma que não haja perda alguma, logo uma nuvem é uma das melhores opções disponíveis hoje em dia, a qual pode de forma dinâmica acompanhar a necessidade atual evitando gastos desnecessários e para isso precisa-se da tecnologia Big Data, que é capaz de guardar todos os dados que são gerados sendo eles estruturados ou não.

E para que isso se torne realidade, existe em alguns lugares do mundo países que estão investindo muito para construir cidades do zero, com o intuito de desde a planta da

cidade até a sua conclusão ela seja inteligente e vários fatores são considerados para que tudo possa dar certo.

Portanto, para que todos os esforços não seja em vão, a cidade necessita de uma coleta eficiente de dados e uma estrutura que seja compatível com os dados coletados e de uma análise desses dados afim de extrair informações e que esteja disponível em qualquer lugar, com o auxílio de uma nuvem.

3. Big Data

O conceito Big Data vem se tornando cada vez mais popular, porém, ainda não está bem claro a sua definição, a sua função e a sua aplicação. Para uma melhor compreensão, é essencial entender o significado dos 3V's: Volume + Variedade + Velocidade e ainda acrescenta 2V's: Veracidade e Valor. O Volume representa a grande quantidade de dados gerados por sistemas corporativos, por mídias sociais, sensores e outros dispositivos; a Variedade representa os dados estruturados e não estruturados, obtidos do Twitter, Facebook, dentre outros, dados de empresas com grandes volumes de geração e movimentação de dados; a Velocidade, que representa a resposta quase que em tempo real para agir no próprio evento gerador das informações. Veracidade, visto que é necessário ter certeza que os dados fazem sentido e são autênticos; Valor porque é absolutamente necessário que a organização que implementa projetos de Big Data obtenha retorno destes investimentos. (TAURION 2013).

Apesar deste grande volume de dados que o Big Data possui a maior preocupação é o que será feito com todos esses dados. Esses dados se forem analisados de forma correta o mesmo conseguirá transformar todo esse dado bruto em informações que poderá ajudar nas tomadas de decisões tornando-as mais inteligentes para as empresas possibilitando que uma nova perspectiva de negócio seja criada, permitindo uma redução de custo e tempo desenvolvendo uma otimização de novos produtos ou serviços.

3.1. Data Mining

Todos os dias um volume descomunal de dados é gerado e armazenados, e a tendência é que aumente a cada dia. Essa quantidade de dados passa a ter algum valor a partir do momento em que é possível extrair conhecimento deles, que é realizado pelo processo de Descoberta de Conhecimento em Banco de Dados ou KDD (Knowledge Discovery in Databases). Neste processo, a Mineração de Dados é a etapa mais significativa que consiste na combinação de métodos de análise de dados com algoritmos sofisticados para que possa ter um processamento de um grande volume de dados,

buscando uma consistência de padrões e relacionamentos sistemáticos entre variáveis (Tan et al. 2006).

Para Han et al. (2011), a mineração de dados é uma parte do processo do KDD que busca obter ou descobrir padrões a partir de um grande volume de dados, que representem informações úteis. Com os algoritmos certos que utilizem de técnicas estatísticas e de inteligência artificial para classificar os elementos de um conjunto, à procura de erros e possíveis relacionamentos que antes não podiam ser vistas.

3.2. Análise Preditiva

O desejo de prever o futuro é comum entre as pessoas, se fosse capaz de saber com meses de antecedência uma crise que iria assolar o seu país ou o mercado global qual decisão seria a mais apropriada para prevenir ou diminuir o impacto que aquele problema poderia causar. Logo o objetivo da análise é ir além da própria estatística e mostrar, por intermédio dos dados coletados uma visão do que irá acontecer no futuro, tomando melhores decisões e, por consequência, produzindo *insight* (discernimento) que levarão a ações melhores. (PROOF, 2017).

Segundo Hekima (2015), o tema não é novo porem só recentemente tem recebido e ganhado a visibilidade que merece, pois tudo isso se dá a grande evolução tecnológica. As empresas nunca se viram tão expostas ao fator imprevisível como tem sido nos dias de hoje, na era digital; e de semelhante modo a tecnologia nunca foi tão precisa no momento de ajudar a tomar uma decisão com base no histórico do mercado, poderá prever acontecimentos futuros buscando padrões que antes estavam ocultos.

3.3. Análise Prescritiva

Para PROOF (2017), a análise prescritiva consegue obter recomendações inteligentes de modo a otimizar os próximos passos para a estratégia de seus negócios, com uma tática mais efetiva baseada em dados para uma tomada de decisão mais específica oferecendo métodos para alterar o futuro.

Hoje em dia é muito confundida as análises preditivas e prescritivas, de acordo com Hekima (2017), elas trabalham com a mesma lógica, mas com objetivos que se diferem uma da outra. Pois uma trabalha para identificar padrões e tendências no futuro, enquanto a prescritiva irá analisar as consequências que cada ação tomada trará para o meio em que está inserida.

Sendo um dos meios de definir qual escolha seria a mais efetiva em determinadas situações, ela é pouco utilizada pelas organizações muitas vezes pelo seu

desconhecimento do mesmo, e menos de 3% dos usuários utilizam de decisões ou modelagem matemáticas, simulações e otimizações (GARTNER, 2012).

3.4. Análise Descritiva

Segundo HEKIMA (2017), ter uma compreensão em tempo real dos acontecimentos é o que define a análise descritiva com base na mineração de dados na base da cadeia de Big Data. É uma maneira de visualizar os dados, entender como uma base de dados se organiza e o que significa para o presente sem necessariamente relacioná-la com padrões passados ou futuros.

Para PRATES e HOPPEN (2017), análise descritiva ou exploratória de dados é um meio de utilizar métricas e técnicas estatísticas simples ou avançadas para compreender e explicar de que forma são os dados.

3.5. Análise diagnóstica

Segundo PLUGAR (2016), a análise diagnóstica tem como finalidade compreender o impacto de cada ação dentro das estratégias da empresa. Dessa forma, é possível saber tudo que acontece na sua empresa, e para isso tomar medidas para evitar perdas de negócios, deixando de lado o que não traz resultados e mantendo ou ampliando o que for positivo.

A análise diagnóstica tem como objetivo compreender de maneira causal (Quem, Quando, Como, Onde e Por que) todas as suas possibilidades, enquanto a análise descritiva busca detalhar uma base de dados. Se uma empresa executa uma ação de marketing, por exemplo, a análise diagnóstica é o caminho mais curto e eficiente para que os profissionais avaliem os impactos e o alcance dessa ação após sua realização (Hekima, 2017).

4. Cloud Computing

A Computação em Nuvem ou *Cloud Computing* é um termo que vem sendo utilizado amplamente desde que começou a ser popularizado em 2008, tornando-se mais conhecido desde que virou realidade aos usuários da Internet. Seu novo modelo de computação em que os recursos são dinamicamente escaláveis e podendo ser virtualizados e disponibilizados como serviços pela Internet, que consistem nas infraestruturas computacionais contidas nos datacenters pelo mundo (BITTENCOURT e MANOLA, 2011).

Contudo este termo foi difundido em 2006, pelo Eric Schmidt o então CEO (*Chief Executive Officer* – Diretor Executivo) da Google para descrever os serviços da própria empresa, e posteriormente a Amazon utilizou o mesmo termo para lançar seu serviço EC2 (*Elastic Compute Cloud*) (GILDER, 2006).

De acordo com Rushel, Zanotto e Mota (2010) a computação em nuvem é um ambiente projetado para trabalhar em redundância onde o processamento é centralizado e seus usuários estão movendo suas informações e dados a todos os momentos, possibilitando o acesso ao mesmo em todos os lugares e de forma simples.

4.1. Definição

Para Cearly (2009), computação em nuvem pode ser definido como um paradigma de infraestrutura que permite o estabelecimento do SaaS (software como serviço), sendo um conjunto de serviços com base na *web* e com o objetivo de fornecer funcionalidades, que até o momento precisava de grandes investimentos em *hardware* e *software*, e seu funcionamento era através de um modelo de pagar pelo uso. A definição é um modelo de computação onde os recursos são escaláveis e elásticos podendo ser disponibilizados como serviço para os usuários finais pela Internet.

De acordo com Buyya (2008), a nuvem é um conjunto de computadores virtualizados e interconectados em um tipo de sistema paralelo e distribuído que são providos de forma dinâmica e unificado. Os recursos são fornecidos e regulados através de acordos entre as partes.

Das várias propostas para se definir computação em nuvem, para a agencia governamental dos Estados Unidos *National Institute of Standards and Technology* (NIST 2011) diz ser um modelo que permitir acessos a rede sob demanda de modo adequado a empresa ou usuário, sendo recursos configuráveis que podem ser alterados com poucos ajustes.

4.2. Modelos de Serviços

A Internet é imprescindível para o modelo de computação em nuvem, porem o mesmo é algo muito maior e complexo. Podendo ser tanto um software quanto uma infraestrutura (MOELLER, 2010).

Para Veras (2012), o ambiente de computação em nuvem é composto por serviços, sendo três modelos importantes.

A estrutura da nuvem é definida como mostra a Figura 1, onde um serviço serve de suporte ao outro.



Figura 1 - Representação da Estrutura da Nuvem
Fonte: Adaptado de VERAS (2011), p 31.

4.2.1. SaaS

Software como um serviço, *Software-as-a-Service* (SaaS): basicamente se refere ao *software* na nuvem. Para Reese (2009), é um modelo de *software* implantado na web, onde seu objetivo é ser disponibilizado por um navegador web. Com isso não há necessidades de se preocupar com os requisitos do sistema que esse *software* será processado, porém é de suma importância ter uma conexão com serviços de internet. Um exemplo é o pacote Microsoft Word e o Google Docs, por serem *software* local e web respectivamente. Enquanto o Word deve ser instalado e se preocupar com espaço em disco HD (*Hard Drive*) e sistema operacional, enquanto o Google Docs tem a mesma funcionalidade, porém é necessário apenas conexão com a Internet e um navegador instalado na máquina.

4.2.2. PaaS

Para McGrath(2012), a computação em nuvem PaaS (*Platform-as-a-Service*) significando Plataforma como Serviço. Podendo-se dizer que essa é uma camada intermediária, cujo o objetivo é facilitar o desenvolvimento de aplicações voltadas a usuários de uma nuvem, ocasionando um ambiente de desenvolvimento para as aplicações e fornecendo um sistema operacional e linguagens de programação, com o objetivo de agilizar o processo.

4.2.3. IaaS

Representando a camada inferior de modo a compor a base, pois contém uma plataforma para o desenvolvimento de aplicações e testes. De acordo com Sousa (2009), seu principal intuito é tornar o fornecimento dos recursos como servidores, redes, armazenamento mais fácil e acessível. Seu modelo de infraestrutura é baseado na virtualização de *hardware* que pode ser dinamicamente escalável podendo aumentar ou diminuir seu recurso de acordo com as necessidades das aplicações.

4.3. Modelos de Implantação

O tipo de nuvem depende da necessidade da aplicação a ser utilizada e que serão implantadas. Logo os tipos de acesso e restrições dependem unicamente do tipo de negócio e do nível desejado. Percebe-se que algumas entidades não querem que todos os usuários possam acessar e utilizar determinados recursos em seu ambiente em nuvem. Nos dias de hoje existem vários tipos de implantações de nuvem, porém quatro deles são as mais utilizadas (JADEJA e MODI, 2012).

4.3.1. Nuvem Pública

Pode ser definido como nuvem publica as infraestruturas que são disponibilizadas para o público em geral e que podem ser acessadas por qualquer usuário que conheça a localização do serviço, no qual não possui nenhuma técnica para restrição de acesso. A nuvem publica busca fornecer aos clientes um serviço de TI sem complexidades, no qual o provedor assume todas as obrigações para com a manutenção, gerenciamento e instalação e disponibilidade. Geralmente os serviços são disponibilizados de forma simples com configurações definidas para deixar o mais conveniente para os usuários, portanto, a nuvem publica não é o mais pertinente para usuários que almejam mais segurança e restrições (VERAS, 2011).

4.3.2. Nuvem Privada

Este modelo de infraestrutura é proprietário ou alugado por alguma entidade no qual ela é unicamente operada pela mesma, podendo ser local ou remota possuindo suas políticas de acessos.

De acordo com Taurion (2009), a peculiaridade que diferencia as nuvens privadas são as restrições de acesso, pois cada uma se encontra por de trás do *firewall* de cada organização, adotando a tecnologia e desfrutando de suas vantagens, todavia mantendo as regras de segurança da empresa. O custo e os obstáculos para o

estabelecimento de uma nuvem privada são as vezes excessivo, e as operações constantes da nuvem pode ultrapassar os custos de uma nuvem pública. Contudo as vantagens de uma nuvem privada é a opção de maior detalhamento sobre os recursos que a constituem permitindo a empresa acesso as configurações possíveis.

4.3.3. Nuvem Comunitária

A nuvem comunitária tem por objetivo o compartilhamento de uma mesma nuvem para inúmeras empresas, que dispõem interesses em comuns como os quesitos de segurança e políticas da mesma. O modelo aceita acesso local e remoto, normalmente é gerenciado por um grupo da comunidade ou terceiros (JADEJA e MODI, 2012).

4.3.4. Nuvem Híbrida

A infraestrutura desta nuvem é composta por pelo menos duas nuvens, que mantem as características originais de seu modelo, contudo estarão interligadas por uma tecnologia padronizada ou que permite a portabilidade dos dados e aplicações (VERAS, 2011).

Uma nuvem híbrida bem construída poderia atender processos críticos e seguros, como pagamento de clientes ou de serviços secundários, tais como processamento da folha de pagamento de funcionários. Porém a restrição desta nuvem é a dificuldade de se criar e administrar algo desse porte. Os serviços de várias fontes devem ser recebidos e fornecidos como se fossem do mesmo local de origem, e a comunicação entre os componentes públicos e privados tendem a tornar a implementação ainda mais complexa.

5. Raspberry Pi

Atualmente as empresas querem criar um dispositivo que possa fazer o mesmo que um computador comum, mas com um preço que seja relativamente acessível a todos, e para atrair jovens interessados. Com isso, o desenvolvimento do microcomputador Raspberry Pi¹ revolucionou toda a indústria com Eben Upton e sua equipe em 2006, pois em uma placa do tamanho de um cartão possui toda a tecnologia de um desktop e ainda uma flexibilidade superior. É neste contexto que as empresas perceberam que um novo mercado de dispositivos estava se formando, logo o Raspberry Pi veio conquistar o seu lugar em um meio totalmente novo. A primeira diferença marcante entre a arquitetura é a memória ram, que para o Raspberry Pi seus modelos variam de 256Mb a 1Gb com

¹ <https://www.raspberrypi.org>

tecnologia DDR2 e conta com a tecnologia *Swap*. As conexões USB possuem uma diferença também pois ele conta com 4 entradas sendo 2 de 2.0 e 2 de 3.0. A saída HDMI possui uma diferença também nas versões aceitas, o Raspberry aceita versões 1.3 e 1.4. O periférico de rede RJ45 possui barramento de 10/100 e suporta operações *full-duplex* e *half-duplex*. Possuindo o processador Cortex-A53 baseado no Atmel ATmega organização que produz os processadores ARM, mas possuem algumas peculiaridades, pois o seu *clock* chega a 1.2 GHz com uma placa gráfica Broadcom VideoCore IV. Possuindo também saída de áudio (3.5 mm) e *slot* para cartão para aumentar a memória interna e possui 40 pinos Pi-2. A fonte de alimentação de 5V conectada a uma entrada micro-USB. E um grande diferencial do Raspberry Pi é que ele possui Wi-Fi e Bluetooth integrados na placa. Após essa análise detalhada pode-se perceber que a arquitetura é semelhante a um computador, mas com algumas diferenças. Contudo o Raspberry Pi por ter sido o primeiro a sair no mercado ele conta com um número grande de pesquisas e trabalhos relacionados a ele tanto no meio acadêmico quanto os próprios usuários finais (UPTON, 2013).

5.1. Sistemas Operacionais

Além das várias diferenças dos *hardwares*, uma outra composição importante também é o sistema operacional, que por sua vez é o que controla o equipamento. Hoje em dia a maioria dos *desktops* e *notebooks* utilizam os sistemas operacionais Microsoft Windows ou Apple OS X, contudo, essas plataformas são conhecidas por terem os códigos fechados. Todavia os sistemas operacionais de código fechado que guardam seus segredos a sete chaves, podem apenas demonstrar seus produtos finais sem poder verificar como são feitos. Enquanto o Raspberry Pi foi planejado para utilizar o sistema operacional Linux, que diferente dos de código fechado é possível fazer o download do código-fonte pertencente ao sistema operacional e fazer seja qual for as alterações necessárias. Logo com essa oportunidade de modificação, foi possibilitado que o Linux se adaptasse para que pudesse ser executasse no Raspberry Pi, tendo algumas versões como o Debian, Fedora Remix e Arch Linux foram utilizados. Cada uma atendendo uma necessidade específica, porem todas com uma semelhança, a de terem o código aberto. A essência do Raspberry Pi é a de proporcionar e estimular o ensino de computação básica, pois o dispositivo disponibiliza facilidades, pois muitos projetos o estão utilizando, inclusive este trabalho (HALFACREE e UBTON, 2013).

6. Desenvolvimento

O presente projeto foi um estudo sobre cidades inteligentes que envolve duas tecnologias e por serem muito abrangentes foi dividido em duas partes no qual a primeira parte foi responsável pela Computação em Nuvem e a outra pela armazenamento dos dados em um Big Data.

A primeira parte do projeto teve por objetivo o desenvolvimento de uma nuvem privada utilizando a tecnologia Owncloud, pois o mesmo ofereceu uma maior disponibilidade de recursos que o próprio administrador da nuvem pode optar por usar ou não, logo, este processo resultou em um ganho significativo em relação aos outros tipos de nuvem que o usuário só tem acesso ao produto final.

Para a implementação da nuvem foi utilizada a placa Raspberry Pi, pois seu baixo custo financeiro e baixo consumo de energia foi uma das características que mais se destacaram. Esta placa possui um processamento semelhante à de um computador convencional, e sua portabilidade é incrível, pois o mesmo possui o tamanho de um cartão, o que permite o seu deslocamento sem muito esforço, tornando tudo mais fácil.

Logo, para a placa funcionar é necessário a instalação de um sistema operacional, com isso o S.O. Linux Raspbian Stretch, foi escolhido para este projeto por ser de código aberto e podendo ser modificado posteriormente conforme as necessidades que forem surgindo. Em sequência foi instalado o servidor Nginx e o PHP7, juntamente com todos os pacotes necessários para seu funcionamento como mostra a figura 2.

```
root@raspberrypi:/home/pi# sudo apt-get install nginx openssl ssl-cert php7.0-cl
i php7.0-sqlite php7.0-gd php7.0-common php7.0-cgi sqlite3 php-pear curl libapr1
libtool curl libcurl4-openssl-dev php7.0-xml php7.0 php7.0-dev php7.0-curl php7
.0-gd php7.0-fpm memcached php7.0-memcache php7.0-zip php7.0-intl php7.0-mbstrin
g varnish
```

Figura 2 - exibição dos comandos de instalação do servidor Nginx e o PHP7 com todos os pacotes necessários.
Fonte: do autor

Após a instalação do servidor nginx e do php7 foram realizadas algumas configurações para que se adeque a ferramenta Owncloud. Posteriormente foi instalado a ferramenta Owncloud na sua versão 10.0.8, como ilustra a figura 3.

```
root@raspberrypi:/home/pi# wget https://download.owncloud.org/community/owncloud
-10.0.8.tar.bz2
```

Figura 3 – exibição do comando de download da ferramenta Owncloud
Fonte: do autor

Depois que as configurações necessárias foram feitas, para que a ferramenta Owncloud funcione corretamente, foi adicionado o certificado de segurança SSL, para

que toda a conexão e acesso fique mais seguro e protegido de possíveis ataques ou ameaças. Por fim, foi testado em laboratório, realizando as validações necessárias. Para a instalação e configuração do servidor foi usado o manual oficial² disponível no site da comunidade Owncloud e alguns conhecimentos adquiridos durante todo o curso de Ciências da Computação.

Logo, o objetivo da segunda parte foi de realizar uma pesquisa para conhecer melhor a área de Big Data e algumas das formas de unir toda essa estrutura de dados com uma análise de forma a tirar proveito desses dados. A figura 4 ilustra como seria o modelo da base de dados que seria responsável por armazenar todos os dados que serão coletados pelo módulo de IoT, no trabalho de JUBRAN e CAMOLESI, 2018. Neste trabalho são descritas toda a parte de sensores e placas controladoras que foi realizada para a coleta de todos esses dados, enviando para o módulo de Big Data para ser armazenado e posteriormente analisados.

MODELO DE BASE DE DADOS									
Nº	ID DO RASPBERRY	ID DO ARDUINO	NOME DA RUA	NUMERO DA RUA	DATA	HORARIO	SEMAFORO 1	SEMAFORO 2	MEDIA
0	0001	0001	Rua A	1 a 80	20/05/2018	10:30	15	10	12,5
1	0002	0002	Rua A	80 a 160	23/06/2018	11:20	20	15	17,5
2	0003	0003	Rua A	160 a 241	25/06/2018	15:30	19	14	16,5
3	0001	0004	Rua B	13 a 50	27/06/2018	18:10	12	17	14,5
4	0002	0005	Rua B	50 a 90	10/07/2018	12:30	14	16	15
5	0003	0006	Rua B	90 a 120	15/07/2018	16:50	16	15	15,5
6	0001	0007	Rua C	250 a 280	20/07/2018	18:20	17	13	15
7	0002	0008	Rua C	280 a 370	25/07/2018	22:45	19	14	16,5
8	0003	0009	Rua C	370 a 400	05/08/2018	13:40	13	18	15,5
9	0004	0010	Rua C	400 a 430	10/08/2018	19:25	18	12	15

Figura 4 – modelo da base de dados

Fonte: do autor

Com o estudo percebeu-se que precisaria fazer uma mineração em todos esses dados bruto, de forma que com a ajuda de algumas técnicas que são usadas hoje em dia um conhecimento fosse adquirido. Foram realizadas pesquisas e há 4 tipos de análises com propósitos diferentes, mas com um objetivo em comum, transformar esses dados em informações.

7. Conclusões & Trabalhos Futuros

Este estudo serviu de grande aprendizado em como desenvolver uma Cloud Computing através de plataformas já existentes, estudo geral sobre Big Data, Data Mining entre outros conceitos relacionados a funções desenvolvidas para dar suporte a todo o sistema.

² https://doc.owncloud.org/server/latest/admin_manual/installation/source_installation.html

Para a exemplificação deste trabalho foi desenvolvido um estudo de caso, com base na tecnologia Owncloud. Para tal foi estudada a forma utilizada para criação de uma nuvem particular que ao se deparar com qualquer tipo de mudança do meio ela automaticamente realiza um backup de forma dinâmica de forma que todos os dados que forem gravados na base não sejam perdidos.

Com isso a tecnologia Big Data é de grande valia pois o mesmo possui uma estrutura que está apta a receber todos os tipos de dados, sendo eles estruturados ou não e nas cidades inteligentes vários tipos de sensores são instalados como os de temperatura, presença entre outros. Neste caso, foi utilizado nesse projeto o sensor de proximidade para detectar cada veículo que passa pelos semáforos para posteriormente os dados coletados serem analisados.

Logo, cidades inteligentes possuem inúmeros aspectos ainda que devem ser pesquisados e muitos outros ainda irão surgir conforme a necessidade que as cidades forem tendo pois o número de pessoas tende a aumentar a cada dia e novos desafios surgirem. Todo este trabalho e um dos meios de proporcionar uma pequena melhora para as pessoas que vivem em grandes centros, é a dos semáforos inteligentes com base nos dados coletados trazendo um ganho de tempo para as pessoas, mesmo que o tempo seja mínimo a longo prazo pode se tornar considerável.

Como trabalho futuro sugere o uso das técnicas estudados neste trabalho aplicando os conceitos de Big Data e Nuvem, e realizando a comunicação dessas duas tecnologias e uma aplicação das técnicas de análises estudadas, transformando um dado em uma informação que ajudará a trazer um futuro melhor para as nossas cidades.

8. REFERÊNCIAS BIBLIOGRÁFICAS

BITTENCOURT, A. A.; MANOLA, R. Segurança em computação em nuvem, Vitória, 2011. 18.

BUYYA, R.; YEO, C.; VNUGOPAL, S. Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities, 2008.

CEARLEY, D.; AL, E. Hype Cycle for Applications Development. **Gartner Group Reporter**. Disponível em: <<http://www.gartner.com>>. Acesso em: 17 Fevereiro 2018.

DELOITTE. **Deloitte's 360° Smart City Framework: From Vision to Reality**. Disponível em: <<http://smartcity.deloitte.com/about/>>. Acesso em: 17 Setembro 2018.

GARTNER. **Gartner Taps Predictive Analytics as Next Big Business Intelligence Trend**. Disponível em: <<http://www.enterpriseappstoday.com/business->

intelligence/gartner-taps-predictive-analytics-as-next-big-business-intelligence-trend.html >. Acesso em: 04 Outubro 2018.

GILDER, G. Wired Magazine, 2006. Disponível em: <<https://www.wired.com/2006/10/cloudware/>>. Acesso em: 05 Março 2018.

HAN, J., KAMBER, M. and PEI, J. (2011). **Data Mining Concepts and Techniques**. The Morgan Kaufmann Series in Data Management Systems. Elsevier Science, 3ª edition. São Francisco.USA

HEKIMA. **Tipos de análise de Big Data: você conhece todos os 4?**. Disponível em: <<http://www.bigdatabusiness.com.br/conheca-os-4-tipos-de-analises-de-big-data-analytics/>>. Acesso em: 05 Novembro 2018.

JADEJA, Y.; MODI, K. Cloud Computing - **Concepts, Architecture and Challenges**, Kumaracoil, 22 Mar 2012. 4. Disponível em: <<https://sci-hub.tw/10.1109/ICCEET.2012.6203873>>. Acesso em: 14 Jan 2018.

JUBRAN, L. K. Z. e CAMOLESI, A. R. **O uso de Internet das Coisas para gestão do controle de veículos em Cidades Inteligentes**. Relatório Final do Programa de Iniciação Científica. Fundação Educacional do Município de Assis, 2018

LOPES, F. R. B.; MARTINS, H. P.; ANDRADE, V. S. **Sistema de Arquivos Distribuídos: Comparativo Entre Ambiente Local e em Nuvem**. Faculdade de Tecnologia de Bauru. Bauru, p. 20. 2016.

LEMONS, Andre (2014). **Cidades Inteligentes: Lugar, Territorialização Informacional e Inteligência**. Disponível em <<http://www.lab404.ufba.br/?p=2491>>. Acesso em 01 nov 2017.

MCGRATH, M. P. **Understanding PaaS**. Sebastopol: O'Reilly, 2012.

MOELLER, R. R. **IT Audit, Control, and Security**. Canada: John Wiley & Sons, v. 13, 2010.

PLUGAR. **Quatro faces da análise de dados: análise diagnóstica** Disponível em: <<https://www.plugar.com.br/quatro-faces-da-analise-de-dados-analise-diagnostica/>>. Acesso em: 20 Outubro 2018.

PRATES, W, R.; HOPPEN, Joni. **Análise descritiva, preditiva, prescritiva e cenarização: como gerar valor nos negócios**, 2018. Disponível em: <<https://www.aquare.la/analise-descritiva-preditiva-prescritiva-e-cenarizacao/>>. Acesso em: 20 Setembro 2018.

PROOF. **Coloque a inteligência e os insights para trabalhar por meio da análise preditiva**. Disponível em: <<https://www.proof.com.br/blog/coloque-a-inteligencia-e-os-insights-para-trabalhar-por-meio-da-analise-preditiva/>>. Acesso em: 12 Novembro 2018.

REESE, G. **Cloud Application Architectures - Building Applications and Infrastructure in the Cloud**. Sebastopol: O'Reilly, 2009.

RUSCHEL, H.; ZANOTTO, M. S.; MOTA, W. C. **Computação em Nuvem**, Curitiba, 2010. Disponível em: <<http://www.ppgia.pucpr.br/~jamhour/RSS/TCCRSS08B/Welton%20Costa%20da%20Mota%20-%20Artigo.pdf>>. Acesso em: 20 Janeiro 2018.

SOUSA, F.; MOREIRA, L.; MACHADO, J. Computação em Nuvem. **Conceitos, Tecnologias, Aplicações e Desafios.**, Ceará, 20 Setembro 2009. 150-175. Disponível em: <https://www.researchgate.net/profile/Javam_Machado/publication/237644729_Computacao_em_Nuvem_Conceitos_Tecnologias_Aplicacoes_e_Desafios/links/56044f4308aea25fce3121f3/Computacao-em-Nuvem-Conceitos-Tecnologias-Aplicacoes-e-Desafios.pdf>. Acesso em: 10 Fevereiro 2018.

TAN, P., STEINBACH, M., and KUMAR, V. (2006). **Introduction to Data Mining**. Pearson international Edition. Pearson Addison Wesley.

TAURION, C. **Computação em nuvem**: transformando o mundo da tecnologia da informação. Rio de Janeiro: Brasport, 2009.

TAURION, Cezar. **Coletânea de posts publicados no Blog developerWorks em2012developerWorks Brasil**. Disponível em: <<http://www.ibm.com/developerworks/blogs/page/ctaurion> >.

UPTON, E.; HALFACREE, G. **Raspberry Pi Manual do Usuário**. São Paulo: Novatec, v. 1, 2013. Disponível em: <<https://s3.novatec.com.br/capitulos/capitulo-9788575223512.pdf>>. Acesso em: 20 Jun 2018.

VERAS, M. **Cloud computing**: nova arquitetura da TI. 1. ed. Rio de Janeiro: Brasport, 2012.