



GOVERNO DO ESTADO DO RIO DE JANEIRO
SECRETARIA DE ESTADO DE CIÊNCIA E TECNOLOGIA
FUNDAÇÃO DE APOIO À ESCOLA TÉCNICA
CENTRO DE EDUCAÇÃO PROFISSIONAL EM TECNOLOGIA DA INFORMAÇÃO
FACULDADE DE EDUCAÇÃO TECNOLÓGICA DO ESTADO DO RIO DE JANEIRO
FAETERJ/PETRÓPOLIS

Aplicação tempo-real com NodeJS - Chat por localização geográfica

João Pedro Souza Homem

Petrópolis - RJ

Dezembro, 2016

João Pedro Souza Homem

***Aplicação tempo-real com NodeJS - Chat por
localização geográfica***

Trabalho de Conclusão de Curso apresentado à
Coordenadoria do Curso de Tecnólogo em Tec-
nologia da Informação e Comunicação da Fa-
culdade de Educação Tecnológica do Estado do
Rio de Janeiro Faeterj/Petrópolis, como requi-
sito parcial para obtenção do título de Tecnó-
logo em Tecnologia da Informação e Comuni-
cação.

Orientador:

Marlan Kulberg

Petrópolis - RJ

Dezembro, 2016

Folha de Aprovação

Trabalho de Conclusão de Curso sob o título “*Aplicação tempo-real com NodeJS - Chat por localização geográfica*”, defendida por João Pedro Souza Homem e aprovada em 12 de Dezembro de 2016, em Petrópolis - RJ, pela banca examinadora constituída pelos professores:

Prof. Marlan Kulberg
Orientador

Prof. Banca Interna
Faculdade de Educação Tecnológica do Estado
do Rio de Janeiro Faeterj/Petrópolis

Prof. Banca Interna
Faculdade de Educação Tecnológica do Estado
do Rio de Janeiro Faeterj/Petrópolis

Declaração de Autor

Declaro, para fins de pesquisa acadêmica, didática e técnico-científica, que o presente Trabalho de Conclusão de Curso pode ser parcial ou totalmente utilizado desde que se faça referência à fonte e aos autores.

João Pedro Souza Homem
Petrópolis, em 12 de Dezembro de 2016

Agradecimentos

Agradeço a todos que de forma direta ou indireta contribuíram para a elaboração deste trabalho através do investimento em minha capacitação, compartilhamento de conhecimento, disponibilização de ferramentas e componentes para desenvolvimento de *softwares*, entre outros não menos importantes.

Dentre todos os relacionados a este agradecimento, gostaria especialmente de agradecer a toda comunidade *open-source*, da qual pude usufruir de diversos recursos que viabilizaram a implementação da aplicação abordada neste documento. A iniciativa *open-source* abre muitos caminhos para a inovação e inclusão na área de desenvolvimento de *softwares* e deve ser incentivada e reconhecida, atribuindo-se os devidos méritos a cada um que adere a iniciativa em seus projetos.

Epígrafe

"Faça-se simples com ideias complicadas"

(JoPaseho)

Resumo

A evolução da tecnologia vem trazendo para as nossas vida diversas inovações que estão mudando a forma de lidar com inúmeras situações. Alcançamos novos patamares em como tratamos os negócios, relações pessoais, a busca por novos conhecimentos, entre outras coisas. Tudo isso nos levou a buscar entender um tipo de aplicação que tem se tornado cada vez mais presente nas ferramentas que utilizamos.

As aplicações tempo-real são o propósito da realização deste trabalho. Primeiramente, buscamos no mercado algumas referências de aplicações para embasar o desenvolvimento de uma aplicação de *chat* por localização geográfica. Encontramos em nossas referências diversos aspectos que foram considerados para a escolha de como e quais tecnologias usar para construir as funcionalidades.

Tendo como base todas as possibilidades que tínhamos para implementação, delimitamos o problema em construir uma ferramenta com tecnologias modernas e utilizadas no mercado por aplicação com fins semelhantes ao nosso, e que demonstrasse o funcionamento elemental da transmissão de dados em tempo real na *web*.

Sendo assim, subdividimos o desenvolvimento em duas partes principais, que são a aplicação servidor e cliente.

- A aplicação servidor desempenhará o papel de receber as informações da aplicação cliente, processá-las, e distribuí-las para os outros clientes através de uma conexão bidirecional, que nos permite implementar uma comunicação tempo-real entre os diversos usuários da aplicação.
- A aplicação cliente, resumidamente pode ser considerada a interface com a qual os usuários interagem com a aplicação. Através dela o usuário poderá encontrar, por de um mapa terrestre, um outro usuário que esteja utilizando a aplicação no mesmo momento, que será identificado como um ponto marcado no local em que os dados de geolocalização coletados pelo servidor informam. Após a localização do usuário, este poderá ser contactado através de uma janela de comunicação por texto.

A arquitetura utilizada foi inspirada em alguns padrões que moldam a infraestrutura da aplicação para privilegiar ações de escalabilidade, portanto apresentamos parte desse esquema implementado em nosso sistema, e uma outra parte complementar de forma teórica.

Palavras-chave Aplicação tempo-real. Chat por localização. Escalabilidade. NodeJS. SocketIO. Redis. MongoDB. AngularJS. Google Maps.

Abstract

The evolution of technology has brought to our lives several innovations that are changing the way to deal with many situations. We reached new heights in how we treat business, personal relationships, the search for new knowledge, among other things. All this led us to seek to understand a type of application that has become increasingly present in the tools we use.

The real-time applications are the subject of this work. First of all, we seek some applications's references to support us with the development of a geographic location based chat. We found in our references several aspects that were considered for the choice of how and which technologies to use to build the functionalities.

Based on all the possibilities that we had to implement, we delimited the problem of building a tool with modern technology and used in the market for application purposes similar to ours, and to demonstrate the elemental operation of real-time data transmission in the web.

Thus, the development was subdivided into two main parts, which are the application server and client.

- The application server will play the role of receiving the client application information, process them, and distribute them to other clients through a bidirectional connection that allows us to implement a real-time communication between the application's users.
- The client application can be briefly considered the interface which users interact with the application. Through it, you can find, by a terrestrial map, another user who is using the application at the same time, it will be identified as a marked point where the geolocation data collected by the server report. After found the user location, it can be contacted by a text communication window.

The architecture used was inspired by some standards that shape the application infrastructure to favor scalability actions, therefore we present part of this scheme implemented in our system, and another complementary part theoretically.

Keywords Real-time Application. Chat by location. Scalability. NodeJS. SocketIO. Redis. MongoDB. AngularJS. Google Maps.

Lista de Figuras

4.1	Fluxo simplificado de requisições no NodeJS. Fonte: maxroecker [27]	p. 28
4.2	Conexão entre servidor e cliente com SocketIO. Fonte: ifollow [31]	p. 30
4.3	Localização do Nginx na arquitetura de aplicações <i>web</i> . Fonte: Thunder-Boy [45]	p. 36
5.1	Representação UML dos modelos persistidos no banco de dados.	p. 39
5.2	Fluxo de acesso aos recursos com OAuth2. Fonte: jlabusch [54]	p. 40

Lista de Tabelas

Lista de Abreviaturas

PUB/SUB Publish/Subscription

SQL Structured Query Language

NoSQL Not Only Structured Query Language

SPA Single Page Application

MVVM Model-View-ViewModel

REST Representational State Transfer

JSON JavaScript Object Notation

API Application Programming Interface

I/O Input/Output

OS Operation System

RHEL Red Hat Enterprise Linux

JVM Java Virtual Machine

JRE Java Runtime Environment

NPM Node Package Manager

UI User Interface

ACID Atomicity, Consistency, Isolation, Durability

MVW Model-View-Whatever

MVC Model-View-Controller

XML Extensible Markup Language

HTML HyperText Markup Language

UX User Experience

SEO Search Engine Optimization

CSS Cascading Style Sheet

SASS Syntactically Awesome Stylesheets

HTML Hypertext Markup Language

HAML HTML Abstraction Markup Language

SSL Secure Sockets Layer

HTTP Hypertext Transfer Protocol

UTC Coordinated Universal Time

RPM Red Hat Package Manager ou RPM Package Manager

CRUD Create, Read, Update, Delete

UML Unified Modeling Language

Sumário

1	Introdução	p. 15
1.1	Motivação	p. 15
1.2	Objetivos	p. 16
2	Trabalhos Relacionados	p. 17
2.1	Uber	p. 17
2.2	Redis Publish-Subscribe (RedisMVA)	p. 18
2.3	Redis Pub-Sub (Rajaraodv)	p. 18
2.4	FanMappr e RadiusIM	p. 19
2.5	Snapchat, Happn, Tinder e afins	p. 20
3	Problematização	p. 22
3.1	Caracterização do Problema	p. 22
3.2	Solução Proposta	p. 23
4	Hardware e Software Utilizados	p. 25
4.1	Codificação e Testes	p. 25
4.1.1	Ambiente de Desenvolvimento	p. 25
4.1.2	Versionamento de Código	p. 26
4.1.3	Editores de Texto	p. 26
4.2	Back-end	p. 27
4.2.1	NodeJS	p. 27
4.2.2	NPM	p. 29

4.2.3	ExpressJS	p. 29
4.2.4	SocketIO	p. 30
4.2.5	MongoDB	p. 31
4.2.6	Redis	p. 31
4.3	Front-end	p. 32
4.3.1	Bower	p. 32
4.3.2	AngularJS	p. 32
4.3.3	Angular Material	p. 33
4.3.4	Gulp	p. 34
4.3.5	SASS	p. 34
4.3.6	HAML	p. 34
4.3.7	Webpack	p. 35
4.3.8	Babel	p. 35
4.3.9	Nginx	p. 36
5	Implementação	p. 37
5.1	Backend	p. 37
5.1.1	Persistência	p. 39
5.1.2	Autenticação e Autorização	p. 40
5.1.3	Rotas	p. 41
5.1.4	Socket	p. 41
5.2	Frontend	p. 41
6	Resultados	p. 42
7	Trabalhos Futuros	p. 43
8	Conclusão	p. 44

1 Introdução

Os avanços da tecnologia em conjunto com a internet têm ditado aos seus usuários novas formas de consumo de dados e informações. Além disto, um assunto tem ganhado cada vez mais destaque no provimento de dados na rede: a periodicidade. A inclusão dos meios tecnológicos no cotidiano das pessoas e o grande aumento do volume de dados gerados tem desenvolvido uma nova tendência de informações instantâneas, o que nos leva a elaborar meios para atender a esta nova demanda.

Alguns questionamentos aos modelos atuais de consumo de dados, que apesar de muito bem sucedido em diversas áreas, surgem como uma necessidade ainda maior pelo consumo de informações. O *e-mail*, tradicional meio de comunicação entre pessoas, oferecendo uma boa razão de confiabilidade e formalidade, já é questionado por *startups* [1] [2] que acreditam no futuro da *web* em tempo real.

1.1 Motivação

As aplicações tempo-real estão surgindo como requisito de muitos dos novos sistemas em desenvolvimento, e buscam atender consumidores que já possuem em seus negócios a necessidade da informação imediata. Neste trabalho iremos desenvolver uma aplicação que explora de forma bem específica essa necessidade da troca de informações entre usuário e aplicação em tempo real.

A aplicação desenvolvida será um *web chat* de comunicação por texto baseado em localização geográfica. As duas principais características do projeto, comunicação por texto e localização geográfica, traz para si aspectos modernos, que exigem a utilização de ferramentas e técnicas atuais de desenvolvimento para a *web*, tendo em vista que deveremos levar em consideração diversos conceitos, como compatibilidade, escalabilidade, infraestrutura de sistemas, entre outros. Tudo isto serve como motivação para atualização e aprendizado de diversos tópicos relacionados ao desenvolvimento de *softwares*.

1.2 Objetivos

A finalidade do projeto será demonstrar como pode ser feita a transferência de dados em tempo real entre servidores e clientes. Para isto, implementaremos uma aplicação que irá realizar a comunicação por texto entre usuários, dos quais serão localizados para comunicação através de localização geográfica em um mapa terrestre real.

O produto mínimo viável da aplicação será elaborado para um ambiente de testes local, não oferecendo um fluxo completo de funcionamento dos servidores em ambiente de produção, nem do fluxo de utilização de um usuário comum em sua interface. Abordaremos apenas alguns conceitos sobre escalabilidade e infraestrutura da aplicação em ambiente de produção, assim como apenas algumas funcionalidades necessárias para interação na interface gráfica.

Iremos explicar com mais clareza os problemas que cercam este projeto no tópico "Problematização", onde falaremos especificamente quais são os ideais de implementação de uma solução em tempo real que tem características como as que nos propusemos a fazer, levando em conta aspectos como produto e serviço, já que ideologicamente um software é feito para realizar uma ou mais funções para um ou mais usuários, mas como citado, o nosso foco é demonstrar fundamentalmente o funcionamento deste tipo de aplicação, e não oferecer um produto ou serviço. A problematização irá se perfazer em torno do estudo de uma forma de se implementar uma solução tempo-real para a *web* envolvendo duas funcionalidades que julgamos serem de boa representatividade para a circunstância.

2 *Trabalhos Relacionados*

Com o objetivo de situar-se melhor em relação ao funcionamento da arquitetura e funcionalidades do sistema, analisamos alguns trabalhos que se assemelham em um ou mais aspectos da aplicação desenvolvida. Dentre estas, estão relacionadas propostas de implementação, que explicam o funcionamento da arquitetura de aplicações tempo-real, e também projetos que foram publicados como produto de mercado.

Iremos fazer uma breve análise de cada uma das plataformas, sendo que algumas análises serão fundamentadas em publicações que falam sobre o funcionamento destas, pois as mesmas não se encontram mais disponíveis para acesso. Resolvemos incluí-las em nossa documentação, pois são ferramentas que tentaram como produto final uma aplicação muito semelhante a que iremos desenvolver como caso de estudo.

2.1 Uber

Iniciando a pesquisa sobre o desenvolvimento de uma aplicação tempo-real, buscamos uma referência de grande escala [3]. O Uber [4] é uma aplicação *web* e *mobile* que oferece serviços de transporte urbano privado através de pessoas que se disponibilizam a serem motoristas e estarem disponíveis na plataforma para alguma oferta de demanda.

Encontramos uma publicação [5] que fez-se de grande valia para o início da escolha das tecnologias que utilizamos para desenvolver a aplicação deste trabalho. A publicação fala sobre como era a arquitetura do sistema, para quais possibilidades ela foi pensada, e sobre a nova organização do sistema e quais tecnologias utilizam para suportar tamanha demanda [3].

É notório ao longo do trabalho que a arquitetura aplicada para desenvolvimento da aplicação não se aproxima em muitos pontos da aplicada no Uber, mas após o estudo da publicação, fomos fomentados a pesquisar algumas das tecnologias utilizadas e como aplicá-las em nosso projeto, onde a principal influência foi a utilização do Redis como banco de dados em memória, e não diretamente uma influencia, mas uma consolidação de ideias, a utilização do NodeJS e um

banco de dados NoSQL para persistência dos dados.

2.2 Redis Publish-Subscribe (RedisMVA)

Esta publicação [6] faz parte de um projeto demonstrativo da utilização do padrão de comunicação pub/sub, com um banco de dados NoSQL, chamado Redis. A publicação fala sobre a implementação em NodeJS de um *chat* de comunicação por texto, utilizando o padrão pub/sub para realizar a comunicação entre processos, onde os processos são as diversas instâncias que podem ser executadas de um mesmo servidor NodeJS distribuídos em diversos locais na internet.

O estudo desta publicação serviu como grande base de aprendizado para a elaboração da arquitetura do sistema que desenvolvemos, pois utilizava tecnologias atuais e que já tínhamos como plano para a implementação do nosso sistema, como citamos no caso do Uber. A utilização da linguagem de programação JavaScript, naturalmente traz um certo intimismo para quem já desenvolve para a *web*, dado que ela pode ser considerada a linguagem da *web* (*client-side*), pois todos os maiores navegadores *web* [7] utilizados no mundo interpretam a linguagem nativamente, sendo muito comum programadores voltados para a área já terem tido algum contato com a linguagem.

O autor da publicação disponibiliza todo o código fonte do projeto de demonstração com licença aberta, sendo assim, ele discorre sobre o tema de forma muito objetiva, apenas abordando em seu texto os pontos de interseção em que a implementação da comunicação pub/sub se insere. O texto oferece uma didática teórica de boa qualidade, demonstrando com diagramas e imagens da aplicação real, o funcionamento de cada passo do qual explica no documento.

A demonstração dos fluxos de dados que ocorrem no sistema ficam bem representadas com os diagramas, embora alguns destes fluxos não estejam implementados no sistema, como o balanceador de carga, visto que não era o foco de esclarecimento, é simples entender como a interface envia os dados para o servidor, que distribui os dados para os outros clientes conectados.

2.3 Redis Pub-Sub (Rajaraodv)

Esta publicação [8] fala sobre o mesmo assunto da citação anterior, que é a implementação de um chat por texto utilizando NodeJS e banco de dados Redis, mas nesta publicação o autor é mais abrangente no quesito de infraestrutura e escalabilidade da aplicação, abordando todos os

passos para uma implantação completa do sistema em um servidor.

O texto da publicação é muito explicativo, é possível até para os mais principiantes no assunto de aplicações tempo-real entenderem a implementação e a finalidade de cada passo relatado. O texto também conta com ilustrações dos fluxos de dados na aplicação e demonstrações do funcionamento da interface.

O que se pode extrair do conteúdo desta publicação além da anterior, é a implementação do balanceador de carga, em conjunto com o *proxy* reverso Nginx, o conceito de *sticky sessions* para traçar requisições a um mesmo servidor, e a integração do sistema em um servidor na nuvem (conjunto de recursos computacionais na internet, dedicados a hospedar aplicações de diversos tipos).

Após o estudo da publicação deste autor, utilizamos em nossa implementação a utilização do Nginx. Entendemos a ideia de centralizar a distribuição de conteúdo estático da aplicação, como a interface, e canalizar as conexões aos servidores através de um ponto em comum, facilitando assim o gerenciamento para futuras implementações de *sticky sessions*, balanceamento de carga, reconexão em casos de *scale-up* ou *scale-out* dos servidores (aumento e diminuição do número de servidores), entre outros benefícios que este tipo de arquitetura pode prover para as aplicações modernas, aspirando o rumo em que a tecnologia vem tomando com o crescimento exponencial de usuários, consumo de dados, globalização, entre outros, que contribuem para a necessidade de se construir sistemas que sejam distribuídos e escaláveis, alcançando altos desempenhos.

2.4 FanMappr e RadiusIM

O FanMappr [9] e o RadiusIM [10] são aplicações baseadas em *chat* por localização geográfica, mas que foram disponibilizadas como produto, com a proposta de criarem relações sociais através da busca de usuários por localização, ou seja, são redes sociais e tinham isso como proposta final. Após uma análise mais técnica nos dois últimos tópicos em relação ao desenvolvimento da aplicação, agora iremos analisar ferramentas que implementam como um produto, aplicações que tem funcionalidades semelhantes as do nosso projeto.

Encontramos o FanMappr e RadiusIM após algumas poucas pesquisas que buscavam especificamente produtos que tivessem funcionalidades muito próximas das que implementamos em nossa aplicação. O FanMappr é uma aplicação com interface e funcionalidades aparentemente modestas, navegando entre os poucos menus que possui, percebemos que as funcionalidades se restringem somente ao *chat* por texto e a um mapa terrestre onde os usuários são localizados.

O RadiusIM não está mais disponível para acesso, mas encontramos uma publicação [10] que possuía algumas imagens sobre a aplicação e uma breve explicação do seu funcionamento, que também é muito semelhante com do FanMapp, mas o RadiusIM, apesar de parecer ter tido mais sucesso, nos anos 2000, presume-se que teve um investimento maior, pois possuía uma interface mais harmoniosa e funcionalidades que eram o de se esperar para sua época.

É interessante perceber que estas aplicações não tiveram grande sucesso colocando como produto final uma ferramenta de comunicação com funcionalidades semelhantes a que iremos implementar, o que nos leva a interpretar que a *web* em tempo real não é uma inovação a qualquer custo. É importante considerar que as melhorias que se pode ter com informações em tempo real devem ser mensuradas, e ela é apenas mais uma opção que está ganhando mais importância com as demandas que estão surgindo.

2.5 Snapchat, Happn, Tinder e afins

Neste tópico iremos fazer uma análise, mais filosófica do que técnica, sobre a utilização de aplicações tempo-real na vida dos usuários, pois achamos interessante fundamentar onde estão se rompendo os limites para a necessidade de aplicações deste tipo, e gostaríamos de ressaltar que Snapchat [12], Happn [14] e Tinder [13] foram escolhas que achamos que representam um grupo de aplicações com características próximas, mas que para muitas outras aplicações existentes, caberiam-se a mesma análise.

A seguinte citação [11] nos traz algumas importantes informações sobre a opinião de um usuário comum:

"There's a fundamental difference between location based chat and chat applications like Snapchat. Snapchat changed how people communicate, it did not change how they found people to communicate with... it may have helped better relationships form but it did not create them out of nothing.

Arbitrary things like age and location don't define who we are or what we enjoy, and who we are and what we enjoy is a big part of whether or not we will enjoy being friends with someone. Relationships are an investment, most people don't want to talk to a bunch of new people every day but never form any lasting relationship. If they want to form a lasting relationship it needs to be based on more than "I saw you at whole foods"... maybe their phone will ping in whole foods and maybe it'll be someone interesting and maybe they'll become friends and maybe it'll be a fantastic relationship, but what are the chances of that happening? How many people are willing to go through hundreds and hundreds of

worthless chats just to find one friend?

Maybe one day it will be time for a location based chat app to take off, but I don't think that time is now."

("Do location based chat apps start trending again?"[11], NOTNOTCITRICSQUID, 2012)

A resposta a pergunta do tópico expressa uma opinião que condiz com alguns pontos que pudemos notar ao longo do conteúdo apresentado até aqui, que se relacionam com como as aplicações tem obtido sucesso na inserção da experiência de tempo real na vida dos usuários mais comuns, e aplicações como o FanMapp e o RadiusIM não alcançaram o mesmo resultado. Snapchat, Happn e Tinder são aplicações que funcionam baseando parcialmente e/ou totalmente suas funcionalidades em *chat* por localização geográfica, mas diferem-se com uma certa cautela da experiência exposta aos seus usuários nas suas aplicações, e este é um ponto que acreditamos ser o diferencial para a aceitação das ferramentas, pelo fato não infringir algumas barreiras que existem na interação humana com o mundo virtual.

A possibilidade da informação instantânea tem muito a contribuir em diversas áreas, oferecendo vantagem competitiva para quem está no mundo dos negócios, antecipação de tomada de decisões, entre outros benefícios. Este mundo do tempo-real está a cada vez mais tomando conta do nosso presente como usuários comuns, e a tendência é que essa forma de compartilhamento instantâneo se torne mais natural.

Como as pessoas utilizam e disponibilizam dados, tem uma importância fundamental no impacto em que isso pode ter em suas vidas. Desde as necessidades no mundo dos negócios, às necessidades do cotidiano de cada um. Por parte de quem está inovando no setor, é preciso entender comportamentos e como não romper barreiras que irão fracassar a relação entre tecnologia e pessoas.

Fazemos a ressalva de que a nossa implementação não busca oferecer funcionalidades que atendam aos requisitos que entendemos como necessários em relação a estes conceitos, mas abordamos o assunto pelo interesse em mostrar como a tecnologia também está envolvida com as relações humanas.

3 *Problematização*

3.1 Caracterização do Problema

A internet oferece diversas possibilidades de integração e distribuição de dados em sua rede mundial. A necessidade cada vez maior de se coletar dados que são gerados a todo momento, cria a iniciativa de se laborar novas maneiras de envio e recepção de dados na rede.

A transmissão de dados em tempo real não é uma necessidade recente, desde de que se começaram a desenvolver sistemas críticos de controle na *web*, muitos destes já utilizavam desta forma de se transmitir dados para realizar suas tarefas. As técnicas e tecnologias utilizadas para alcançar tal feito eram diferentes e mais custosas, o que levava este tipo de característica da aplicação ser adotada apenas em casos específicos, onde deveria se contrabalancear com mais perícia a utilização.

Com os avanços das tecnologias de *hardware*, *software* e da internet como um todo, começaram a se criar novas possibilidades da implementação de *features* tempo-real em aplicações cada vez mais próximas do usuário comum. Com tudo isso, podemos observar que muitas das aplicações de sucesso do mercado usufruem desta tecnologia, que permite uma conexão direta e persistente entre cliente e servidor, viabilizando que se possa receber e enviar dados a qualquer momento, atualizando em tempo real todos os interessados em uma determinada informação.

A implementação de funcionalidades tempo-real definitivamente se tornou algo exequível para as aplicações atuais na *web*, como *push notifications*, *stats*, interações em jogos de navegadores, entre outros. A utilização de *WebSockets*, que analogicamente funcionam como os *sockets*, que fazem comunicação entre processos através da rede, os *WebSockets* são especificamente feitos para este tipo de comunicação na *web*, possibilitando uma conexão contínua entre códigos JavaScript executados nos *browsers* e os servidores web. Estes serão base fundamental para a resolução do problema proposto neste trabalho, levando em conta que são a tecnologia mais atual.

3.2 Solução Proposta

Buscando explorar a usabilidade tempo-real na *web*, iremos formular nossa solução no desenvolvimento de uma aplicação de comunicação por texto, também podendo ser chamada de *chat*, onde para que os usuários tenham conhecimento da existência um do outro na aplicação, existirá um mapa terrestre do mundo real que será marcado com pontos identificadores de cada usuário utilizando o sistema. Estes pontos terão como base a coordenada geográfica real do usuário que será fornecida quando o usuário entrar no sistema.

A escolha por uma aplicação deste modelo se deve a ideia de querer tornar o entendimento do emprego da tecnologia mais amigável para as pessoas que não possuem conhecimento técnico muito aprofundado. A troca mensagens explícita em uma conversa entre duas pessoas, essencialmente ilustra a ideia de transferência de dados em tempo real, e em conjunto com a localização geográfica, conseguimos dar um adendo de outra aplicação da tecnologia.

Iremos desenvolver uma interface com a pretensão de ser amigável ao usuário, utilizando conceitos de responsividade para a adaptação do conteúdo mostrado em telas de dispositivos diversos, inclusive móveis, tendo em vista que algumas estatísticas [15] apontam que dispositivos móveis já são responsáveis por pelo menos um quarto do todo o tráfego de dados gerado na internet, e com estimativas [16] [17] de aumento em cada ano. Mesmo não sendo o foco da aplicação ser um produto final, achamos que o aprendizado com este tipo de requisito seria conveniente. Em relação a organização da interface, a aplicação terá duas telas principais, sendo a primeira com as opções de cadastro através do fornecimento de alguns dados pessoais, e autenticação no sistema por *e-mail* e senha. A segunda tela será o ambiente principal da aplicação, onde o usuário poderá navegar por um mapa terrestre em busca de marcações que identifiquem outro usuário, com o qual ele poderá interagir através de uma janela de *chat* por texto que surge neste mesmo ambiente.

O principal componente da solução será o servidor da aplicação. Este será desenvolvido sobre a arquitetura REST, que transmite e recebe dados da interface no formato JSON. Funcionalidades como cadastro de usuário e autenticação farão uso deste modelo. A comunicação tempo-real será provida por um *framework* que utiliza como base a tecnologia de WebSockets para conexões bidirecional de dados entre servidor e cliente. Apesar da tecnologia WebSocket ser a mais apropriada para a este tipo de conexão, o próprio *framework* trata de usar outras técnicas de comunicação tempo-real, como *long polling*, em casos de falta de compatibilidade dos navegadores com a tecnologia de WebSocket. Tudo isso acontece de forma transparente ao usuário e ao desenvolvedor, pois a utilização do *framework* se dá através de API que oferece de

forma simplificada todo o processo de conexão e comunicação.

A persistência dos dados será feita através de um banco de dados no NoSQL. A escolha por este tipo de banco de dados se deve as características da aplicação que exigem muitas operações de I/O em dados com baixo nível de relacionamento com outros dados, onde este tipo de tecnologia NoSQL oferece melhor desempenho e praticidade.

A tecnologia utilizada como ambiente de execução no servidor, utiliza o JavaScript como linguagem padrão, além de oferecer características que também beneficiam o tipo de aplicação que iremos desenvolver, como *single-threaded*, *non-blocking* e *asynchronism*.

Considerando também alguns conceitos de infraestrutura, iremos adotar na arquitetura da aplicação algumas formas de permitir o escalonamento dos serviços através da distribuição dos servidores em diversas instâncias no mesmo servidor físico ou em outros de locais distintos. Para isso utilizaremos tecnologias para fazer a comunicação entre processos, com um banco de dados NoSQL centralizado e em memória, e interações *stateless* entre servidor e cliente.

4 *Hardware e Software Utilizados*

4.1 Codificação e Testes

4.1.1 Ambiente de Desenvolvimento

O ambiente para desenvolvimento da solução foi criado em um único computador, de caráter pessoal, do tipo notebook, com as seguintes especificações técnicas [18]:

- Processador Intel Core i5 3230M - 2.60 GHz (T-Max: 3.20 GHz) - 64 Bits - Smart Cache 3 MB L3
- Chipset Intel ® HM75 Express
- Memória RAM 8 GBs DDR3 SDRAM - 1600 MHz
- Disco rígido 500GB - 5.400 RPM
- Rede 10/100 e Wireless 802.11 b/g/n - Link de conexão com a internet de 15 Mbps

Especificamos apenas algumas informações que julgamos serem de relevância para alguma análise de desempenho da aplicação por parte do leitor. Em nossa solução não iremos realizar levantamentos de dados sobre desempenho, pois este não pode ser considerado um ambiente que simule a realidade em servidores de aplicações, do qual seria de relevância este tipo de medição.

O sistema operacional escolhido para nosso ambiente de desenvolvimento foi o CentOS 7. O CentOS é uma distribuição Linux baseada no Red Hat. A Red Hat é uma empresa privada, que possui um O.S. chamado Red Hat Enterprise Linux (RHEL), e tem como proposta oferecer produtos *enterprise* baseados no GNU/Linux. Tendo em conta que toda distribuição baseada no *core* do Linux deve seguir uma série de regras contidas em sua licença, a Red Hat concentra

seus faturamentos em suporte as distribuições que ela construiu. Ter um sistema compatível com o RHEL traz alguns benefícios que ainda não foram alcançados sem que se tenha um time de pessoas dedicadas ao desenvolvimento de um *software* específico, como no caso da maioria dos sistemas open-sources, que é o alto padrão em testes contra erros e problemas de segurança, *releases* mais curtos, correções de problemas em curto prazo, entre outros benefícios que o CentOS consegue absorver da comunidade Red Hat. Além disso, o CentOS já possui uma considerável comunidade ativa, que facilita na resolução de problemas mais específicos e no desenvolvimento de novas funcionalidades. O sistema também já possui grandes *cases* de sucesso, como Facebook [19] e Google [20], o que passa uma maior confiança para a adoção em outros negócios e uma perspectiva de crescimento contínua.

4.1.2 Versionamento de Código

O versionamento de código é uma tecnologia que permite traçar históricos de mudanças nos arquivos de código escrito através de um gerenciamento multi-usuário. Este gerenciamento permite que uma ou mais pessoas possam modificar um ou mais arquivos simultaneamente ou não, sendo controlados através de funcionalidades que permitem junção, revisão de conflitos, entre outros.

A tecnologia utilizada para versionamento de código neste projeto foi o Git. Este é um sistema de versionamento distribuído que atende a diversos tipos de *workflows* de trabalho. A escolha por esta tecnologia se deve ao fato de ser umas das mais difundidas no mercado, e onde se encontram os maiores e mais confiáveis repositórios de código. Embora o fluxo de desenvolvimento tenha sido muito simples, utilizamos um *workflow* chamado *Gitflow Workflow* [21] que divide o desenvolvimento basicamente em dois *branches* principais, um de produção e outro de desenvolvimento, podendo ser criado *branches* de cada *feature* sobre o *branch* de desenvolvimento.

4.1.3 Editores de Texto

Para o desenvolvimento majoritário do *frontend*, utilizamos o editor de texto SublimeText 3 [22], apoiado de alguns *plugins* para auxílio de escrita e design, como Emmet, ColorHighlighter, AlignTab, JSHint, MaterialTheme entre outros, todos podendo ser encontrado no repositório de pacotes packagecontrol.io [23].

Para o desenvolvimento do *backend* e parte do *frontend*, utilizamos o editor de texto Visual Studio Code 1.6 [24]. A escolha por este outro editor de texto se fez principalmente pela

funcionalidade de *debug*, que apresentou melhores resultados em sua utilização, além de ser uma funcionalidade nativa da aplicação, assim como a ferramenta para versionamento de código. Também foi utilizado *plugins* semelhantes aos escolhidos no SublimeText para auxílio de escrita de código, todos encontrados no repositório de pacotes padrão do editor.

4.2 Back-end

4.2.1 NodeJS

O NodeJS [25] é uma plataforma de desenvolvimento de alta performance construída sobre a *Chrome's Javascript runtime V8* [26]. A V8 é um interpretador Javascript *open source* desenvolvido em C++ pela Google e capaz de executar em sistemas operacionais Windows, Linux e iOS, consequentemente trazendo esta possibilidade de multi plataforma para o NodeJS. Comparando-se a uma tecnologia que está a mais tempo no mercado e se possui um maior entendimento, o funcionamento da *engine* V8 é superficialmente semelhante ao fluxo de execução do conjunto JVM e JRE do Java.

A plataforma NodeJS possui uma arquitetura de funcionamento *single-threaded* com operações orientadas a eventos e I/O não-bloqueantes. A escolha pela linguagem Javascript se molda bem a este tipo de funcionamento, pois por ser uma linguagem dinâmica e trazer características da programação funcional, como *first-class functions* e *function composition*, que auxiliam fortemente no controle do fluxo de execução através *callbacks*, tornam a programação assíncrona mais fácil na plataforma, contribuindo para a criação de sistemas escaláveis e distribuídos.

Para um melhor entendimento de como o NodeJS trata as requisições feitas ao servidor, a imagem a seguir ilustra simplificada o fluxo de tratamento de uma requisição:

Como se pode observar na figura, diferentemente dos servidores *web* comuns, o NodeJS não cria um processo ou *thread* para cada requisição que chega ao servidor. Ao invés disso, cada instância NodeJS é um processo que funciona como um *event loop*, onde para cada operação de I/O requisitada, é registrada uma *callback* que será chamada posteriormente no *event loop* quando os dados estiverem disponíveis. Sendo assim, a thread principal nunca fica bloqueada, ela está sempre aguardando que um evento esteja para acontecer, que pode ser a chegada ou resposta de uma requisição, por exemplo. É possível que surjam questionamentos sobre como o NodeJS faz as requisições de I/O assíncronas e simultaneamente atende outras requisições, mas isso é abstraído ao desenvolvedor, pois o NodeJS utiliza uma biblioteca multi-plataforma, codificada em C, chamada libuv [28], que trata as operações de I/O assincronamente com utilização

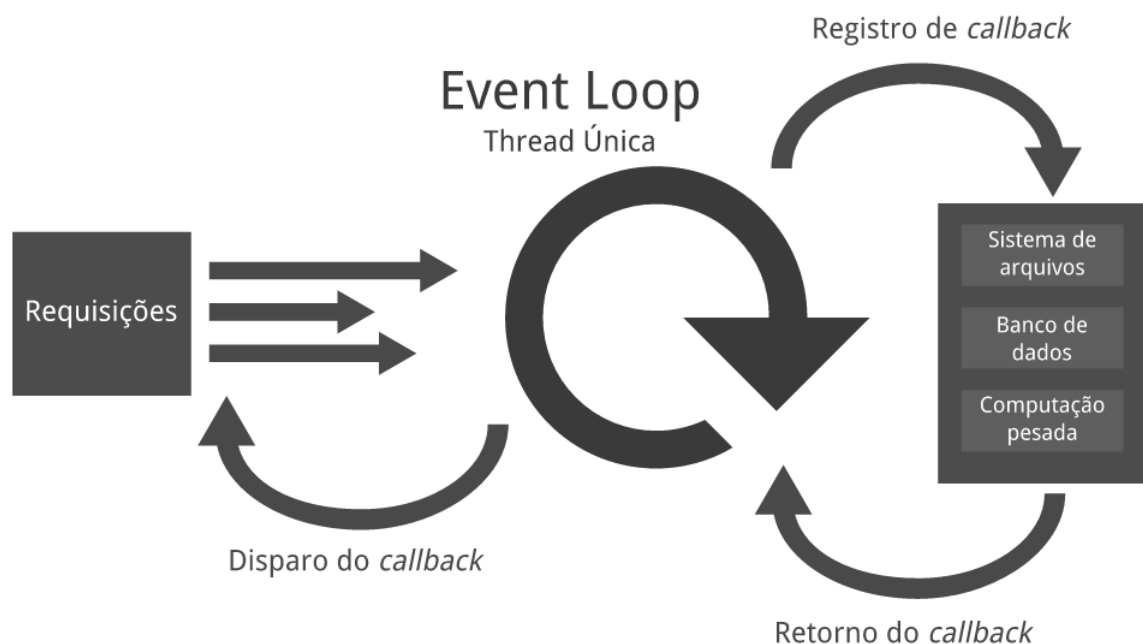


Figura 4.1: Fluxo simplificado de requisições no NodeJS. Fonte: maxroecker [27]

de *thread pool* e afins.

Tendo em mente todos esses conceitos do funcionamento da arquitetura do NodeJS, é possível entender o porque dele conseguir tratar muito mais conexões simultâneas do que as arquiteturas convencionais, pois a criação de processos e *threads* tende a ser um recurso custoso para o sistema operacional, além de reservas de recursos valiosos do sistema, como memória RAM. É importante ressaltar que a programação assíncrona e o ambiente criado pelo NodeJS tem uma curva de aprendizado maior e se faz necessário maiores cuidados por parte do desenvolvedor, pois a utilização incorreta poderá acarretar no travamento da única *thread* da instância da aplicação, bloqueando totalmente o sistema de processar novas requisições.

Por outro lado, quando o entendimento é claro e o desenvolvimento é correto, a utilização dos recursos naturalmente é escalável e distribuída, pois a própria singularidade da plataforma nos faz trabalhar conceitos de sistemas distribuídos, como por exemplo, para que possamos aproveitar de todos os recursos de um processador multi-core com NodeJS, temos que executar mais de uma instância da aplicação, o que nos leva a trabalhar conceitos de clusterização, comunicação de processos e etc, o que facilita no escalonamento horizontal da aplicação em múltiplos servidores.

4.2.2 NPM

O NPM [29] é o gerenciador de pacotes do NodeJS. A função do NPM é oferecer um repositório central de qual se pode baixar diversos módulos distribuídos pela própria equipe que desenvolve o NodeJS ou que outros usuários tenham contribuído.

A evolução do NPM se tornou dominante, sendo considerado o maior repositório de registro de conteúdo Javascript do mundo. Devida a utilização do Javascript para outras finalidades além do desenvolvimento em NodeJS, muitos dos módulos disponibilizados pelo NPM não são parcialmente ou completamente desenvolvidos para a utilização com o NodeJS, sendo que muitos usuários e desenvolvedores utilizam e publicam conteúdo voltado para o desenvolvimento *frontend*, que também utiliza o Javascript como principal linguagem. Talvez este seja um dos grandes motivos de sucesso do NPM em números.

Sua utilização é feita basicamente via linha de comando, onde iniciamos um novo repositório quando criamos uma aplicação, e a partir deste comando é criado um arquivo em formato JSON, chamado `package.json`, do qual ficará registrado todas as dependências daquele projeto, através da especificação do nome e versão de cada módulo baixado do repositório oficial.

A utilização de um gerenciador de pacotes traz diversas vantagens, dentre estas, a facilitação no versionamento de código, pois não é necessário manter uma cópia de um conteúdo que poderá ser baixado posteriormente de um local central, o controle das versões de cada biblioteca que se está utilizando no sistema, facilitando a resolução de problemas de compatibilidade e dependências uma das outras, local único para busca e contribuição de soluções na linguagem, entre outras.

4.2.3 ExpressJS

O ExpressJS [30] é o principal *framework* NodeJS para desenvolvimento de aplicações *web* e APIs. A sua proposta é ser simples, performático e *loosely coupled* na construção de aplicações, sem interferir nas características padrões que o NodeJS oferece.

O desenvolvimento com o ExpressJS é bem intuitivo pela organização que ele propõe para a aplicação. Suas abstrações são apenas essenciais, que não interferem na utilização dos artifícios da própria linguagem nativa do NodeJS para a resolução dos problemas, sem muitas automatizações, sendo ainda mais conveniente oferecendo mecanismos de *middlewares* que possibilitam total manipulação das requisições e seus atributos, injeção de *template-engines* e outros *handlers* e *plugins* que podem ser baixados via NPM. A comunidade possui diversas so-

luções empregadas, o que também facilita na curva de aprendizado, produtividade e resolução de problemas. Rapidamente é possível instalá-lo na aplicação e se disponibilizar um servidor HTTP pronto para receber, tratar e responder requisições como API ou UI.

4.2.4 SocketIO

O SocketIO é uma biblioteca para a criação de aplicações tempo-real na *web*, que permite troca de mensagens em formato binário ou texto de forma bi-direcional e instantânea. Ele funciona através de duas bases de código, uma aplicada no servidor NodeJS e outra no *browser* do cliente. A biblioteca faz uso da tecnologia WebSocket para estabelecer uma comunicação bi-direcional e persistente entre servidor e cliente, mas em casos de problemas de compatibilidade, é utilizado o método de *long polling* para estabelecer este tipo de comunicação. As bibliotecas do SocketIO utilizadas no servidor e no cliente possuem chamadas semelhantes, o que facilita na realização da interação dos dois lados.

A figura a seguir demonstra simplificada a forma de conexão entre servidor e cliente descrita anteriormente:

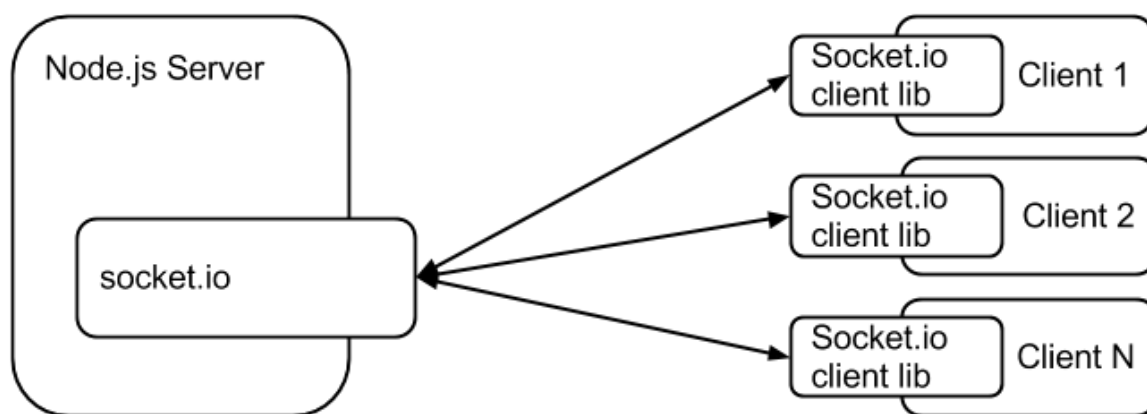


Figura 4.2: Conexão entre servidor e cliente com SocketIO. Fonte: ifollow [31]

Apesar do SocketIO ser um encapsulador da tecnologia *websocket*, as chamadas de sua API são muito semelhantes a utilização do *websocket* diretamente, onde a biblioteca apenas adiciona funcionalidades que são convenientes para o gerenciamento das conexões dos *sockets*, como *broadcast* de mensagens, atrelar informações específicas do usuário de cada *socket*, entre outros.

A tecnologia *websocket* é baseada em canais TCP, o que significa que os dados trafegados por ele tem a garantia de ordem e não perca de informações, o que torna a comunicação mais

confiável e capaz de atender um abrangente número de necessidades. Também é preciso levar em conta os custos que canais TCP podem ter dependendo do tipo de dados que se quer transmitir, onde apesar da possibilidade de transmissão de conteúdo binário por *websockets*, muitas das vezes é mais eficiente a utilização de um canal UDP, como faz uso a tecnologia WebRTC, que é muito utilizada para *streaming* de vídeos, por exemplo.

4.2.5 MongoDB

O MongoDB é um banco de dados de alta performance, enquadrado nos conceitos de bancos de dados NoSQL. A escolha por este tipo de ferramenta se faz pelas características que o modelo de arquitetura escolhido traz para a aplicação. Os bancos de dados NoSQL foram criados para oferecer novas possibilidades para as demandas que exigem escalabilidade horizontal e performance na escrita e leitura de dados, o que se encaixa com os atributos das tecnologias da aplicação desenvolvida.

A forma como este tipo de banco de dados funciona remete diretamente aos ganhos de desempenho e escalabilidade, pois existe uma quebra de conceitos dos bancos de dados relacionais, como o ACID, que trazem flexibilidade em como se tratam os dados armazenados para se ter os benefícios citados. O MongoDB além de adotar estes conceitos, é um banco de dados orientado a documentos, onde a persistência dos dados é realizada em formato JSON. Apesar do conceito NoSQL ser uma grande ruptura de paradigmas, a maior parte da popularidade do MongoDB se deve a sua forma de interação com o usuário, que é amigável a quem está migrando dos conceitos relacionais para o modelo NoSQL. Apesar de não existir a necessidade de criação de tabelas, definição de relações, entre outros, a API do MongoDB oferece a criação de esquemas que remetem a conceitos parecidos, e também habilitam outras possibilidades, como a indexação de campos para buscas mais eficientes. As chamadas de busca no banco também são amigáveis, utilizando palavras chaves como *SELECT*, *WHERE*, *ORDER BY*, entre outros. É importante salientar que algumas das funcionalidades citadas se devem ao uso de um modelador de dados chamado Mongoose [32], ferramenta da qual oferece um *set* de utilidades para se trabalhar com o MongoDB.

4.2.6 Redis

O Redis é um banco de dados NoSQL em memória, que implementa apenas alguns tipos de estruturas de dados, como *strings*, *hashes*, *lists*, *sets* e *sorted hashes*, e funcionalidades para gerência, busca e transmissão dos dados armazenados, o que o torna muito enxuto. Devido ao

Redis ser um banco de dados que funciona em memória, seu desempenho é muito alto, tendo em vista que a memória RAM é um dos componentes de armazenamento de dados mais rápidos em leitura, escrita e transmissão na arquitetura computacional. Por outro lado, a memória RAM é volátil, o que nos leva a buscar entender qual a proposta de um banco de dados em memória.

A utilização principal do Redis é para servir como *cache* e/ou *message broker* de informações em sistemas. Em implementações *stateless* o Redis pode servir como *cache* de dados muito requisitados, pois seu rápido desempenho em memória pode ser vital para atender requisições em tempo hábil. Outra utilidade do Redis, que é o caso deste projeto, é do compartilhamento de dados entre servidores distintos de uma aplicação, onde através das estruturas de dados fornecidas pelo Redis e algumas funcionalidades, como a do protocolo PUB/SUB, é possível que se armazene dados no Redis e outra instância da aplicação sendo executada em outro local possa buscar estes mesmos dados. Devido a simplicidade e robustez que o Redis possui, uma única instância é capaz de realizar milhares de operações, aproximadamente 1.5 milhões por segundo [34], podendo funcionar muito bem sendo este tipo de centralizador, pois o escalonamento de servidores de aplicação ocorre com muito mais frequência do que a de instâncias Redis.

4.3 Front-end

4.3.1 Bower

O Bower [35] é um gerenciador de pacotes, assim como o NPM, mas com conteúdo focado para o *frontend*. Através dele é possível fazer *download* de diversas bibliotecas, como AngularJS, jQuery e Bootstrap. Sua utilização se dá através de linha de comando, onde ao criar um novo projeto *frontend* o usuário deve iniciar o Bower no mesmo diretório, através de linha de comando, e este criará um arquivo chamado `bower.json`, que conterá informações do projeto e todas as nomes e versões de bibliotecas baixadas do seu repositório oficial, permitindo-se assim o gerenciamento de novos pacotes e resolução de dependências entre eles.

O NPM também oferece muitas das bibliotecas disponíveis no Bower, e o contrário muitas vezes não é válido, sendo assim, é comum a utilização dos dois gerenciadores de pacotes no projeto *frontend*, o Bower e NPM.

4.3.2 AngularJS

O AngularJS [36] é um *framework* Javascript para desenvolvimento de interfaces *web*. Mais do que apenas interfaces, o AngularJS auxilia na criação de aplicações complexas no lado cli-

ente da comunicação. Com o avanço das tecnologias e formas de se construir sistemas, ferramentas como o AngularJS vieram para ajudar na organização e qualidade dos sistemas que estavam se desenvolvendo.

A abordagem do AngularJS na construção de aplicações *frontend* é completa, oferecendo suporte para manipulação do DOM, *data binding*, gerenciamento de rotas, padrões de projeto (MVW, MVVM, MVC), entre outros. Apesar da fragmentação de alguns recursos em módulos separados, como as rotas, o produto AngularJS consegue oferecer uma solução completa para o desenvolvimento de interfaces de usuário.

O desenvolvimento de uma aplicação AngularJS, normalmente segue as diretivas de uma aplicação SPA com o estilo de programação declarativa, tendo em vista a forma com que se desenvolve interfaces para os *browsers*, as tecnologias envolvidas, como o próprio HTML, em um ambiente orientado a eventos, se faz mais adequada este tipo de abordagem declarativa, que segue uma lógica mais comum as interfaces.

A arquitetura moldada ao um formato de API também é um fator determinante para a utilização de um *framework* como o AngularJS, pois ele já oferece módulos e serviços para o consumo de dados através de *endpoints* que forneçam dados em formato XML ou JSON.

4.3.3 Angular Material

O Angular Material [37] é uma outra solução do produto Angular, também para interfaces, mas focado em conceitos e características de *design*, UX, reutilização, acessibilidade, compatibilidade *cross-browsers*, responsividade, entre outros.

A ferramenta oferece diversos componentes visuais de interface com efeitos e estilizações próprios, levando em conta diversos conceitos abstratos de *design*, que enriquecem a experiência e agradabilidade do usuário ao ver e interagir com eles. Os componentes do Angular Material são totalmente compatíveis com as implementações AngularJS, facilitando na integração entre os componentes. Um grande problema que é evitado utilizando um *framework* como este, são os de problemas de compatibilidade entre os diversos navegadores no mercado, onde cada um possui especificidades que podem causar comportamentos inesperados nos comentários de interface. Acessibilidade e outras otimizações também são importantes para a inclusão do site e SEO.

Por último, um benefício importante que o Angular Material dispõe, é o seu *grid system*, baseado no CSS3 *flexbox*. Esta característica possibilita que todo o conteúdo apresentado no site se redimensione de acordo com o tamanho da tela do dispositivo que está exibindo o conteúdo,

possibilitando assim uma experiência viável para dispositivos móveis, do qual já citamos a importância em números [15] [16] [17].

4.3.4 Gulp

O Gulp [38] é uma ferramenta de automatização de tarefas Javascript, comumente utilizado para automatizações de fluxos de trabalho no desenvolvimento *frontend*. Sua utilização se dá por linha de comando e um arquivo de configuração chamado `gulpfile.js`. Neste arquivo Javascript podemos criar *tasks* que irão realizar tarefas diversas, de acordo com a programação realiza, como por exemplo a execução de testes unitários, concatenação e minificação de arquivos, compilação de linguagens intermediárias, como o SASS, *autoprefixer*, *lint*, entre outros.

Existem diversos *plugins* disponíveis no NPM que executam tarefas específicas, como as citadas anteriormente, bastando ao usuário baixá-las e chamá-las em suas *tasks*, que podem executar uma ou mais funcionalidades organizadas através de chamadas síncronas pelo método *pipe* de cada *task*.

4.3.5 SASS

O SASS [39] é uma linguagem de *script* que pode ser interpretada para CSS. O CSS tem uma sintaxe muito simples e com poucas funcionalidades para estruturação, reaproveitamento e modularização de código, onde o SASS oferece uma série de funcionalidades, como variáveis, operadores, herança e modularização, que podem ser aplicadas para a geração de folhas de estilos.

Todas as funcionalidades do SASS facilitam a construção de interfaces complexas, onde se exigem muitos componentes, classes, hierarquização de seletores, entre outros. A interpretação do código SASS para CSS pode ser feita por diversos interpretadores, inclusive o próprio Gulp oferece *plugins* para este fim.

4.3.6 HAML

O HAML [40] é uma linguagem que tem a proposta de facilitar a escrita do HTML através de uma sintaxe mais curta, indentação de código e alguns conceitos de HTML dinâmico. Neste projeto utilizamos mais especificamente o HAMLJS [41], que é uma derivação da versão principal do HAML, que habitualmente é utilizado com Ruby on Rails. Essa derivação do HAML oferece uma implementação em Javascript, assim possibilitando sua utilização em diversas fer-

ramentas que utilizam a linguagem. O funcionamento do HAML é parecido com o do SASS e a interpretação dos arquivos HAML para HTML pode ser feita por diversas ferramentas, inclusive *plugins* do próprio Gulp podem realizar esta tarefa.

Apesar da escolha de utilização de uma ferramenta transcrita de outra linguagem, existem diversas opções de *template engines* feita exclusivamente para a utilização com Javascript, como o Jade e o EJS, e que podem trazer melhor qualidade das funcionalidades e *releases* mais curtos com novas implementações.

4.3.7 Webpack

O WebPack [42] é um Javascript *module bundler*. Devido a linguagem Javascript não possuir um sistema para isolamento de escopo dos diversos blocos de código que podem existir em uma aplicação, o WebPack proporciona essa funcionalidade baseando-se nos principais padrões de módulos Javascript utilizados, como o AMD e CommonJS, onde ele é capaz de gerenciar todas as dependências entre módulos e gerar *assets* estáticos para a utilização.

A utilização do WebPack se faz principalmente no desenvolvimento *frontend*, pois o NodeJS, principal ambiente de desenvolvimento com Javascript do lado servidor, já adota o sistema de módulos CommonJS, o que aproxima ainda mais os conhecimentos de desenvolvimento *frontend* e *backend* com a adoção de módulos CommonJS nos dois lados, além dos benefícios na organização e isolamento de códigos. Também é possível a utilização de *plugins* intermediários com o WebPack, estes podem oferecer diversas funcionalidades, como interpretação de ECMAScript 2015 ou TypeScript.

4.3.8 Babel

O Babel [43] é um compilador Javascript que permite a utilização das novas funcionalidades do ES2015. Devido a esta grande atualização do Javascript não ser compatível com uma grande maioria dos navegadores do mercado, o Babel oferece a opção de se escrever o código Javascript com a sintaxe ES2015 e ter a geração de outro arquivo traduzido para o ES5, que é a implementação com maior compatibilidade entre navegadores.

A utilização do Babel pode ser feita utilizando diversas ferramentas, como o Gulp ou o WebPack. Quando se utiliza o WebPack no projeto, é preferível que se faça a integração do Babel com ele, pois como ele cuida da resolução de dependências entre os módulos, ele consegue tratar no momento certo a tradução do código em ES2015 para ES5, evitando problemas de interpretação e compatibilidade.

4.3.9 Nginx

O Nginx [44] é um servidor *web* de alta performance *single-threaded*, uma arquitetura parecida com a do NodeJS, com funcionalidades que o permitem atuar como *proxy* reverso e/ou *load balancer* de aplicações. Para demonstrar onde o Nginx se encaixa na infraestrutura de uma aplicação *web*, podemos observar a seguinte imagem:

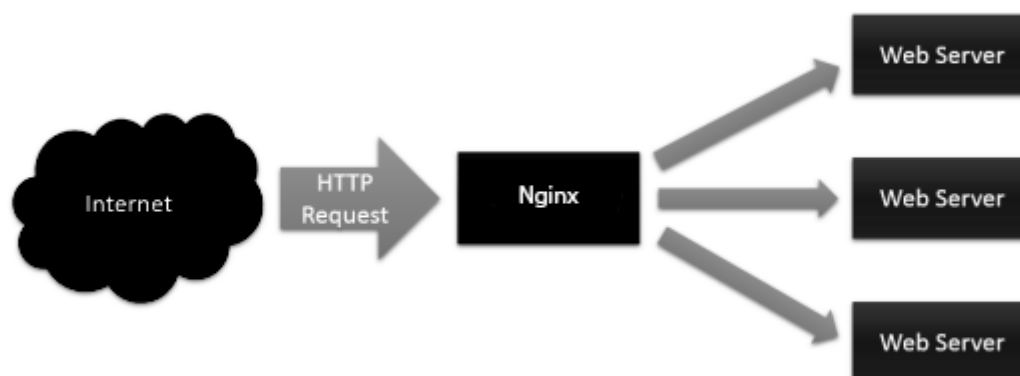


Figura 4.3: Localização do Nginx na arquitetura de aplicações *web*. Fonte: Thunder-Boy [45]

O conceito de *proxy reverso* se dá pela capacidade do Nginx encaminhar requisições de entrada para os servidores da rede privada, ao contrário do que ocorre com os *proxys* comuns, que encaminham o tráfego de saída da rede. Sendo assim, o Nginx está na frente dos servidores de aplicação e gerencia todas as conexões que chegam de um determinado domínio, possibilitando assim que também gerencie controles de *cache*, conexões SSL, *streaming*, fornecimento de conteúdos estáticos da aplicação, como arquivos HTML, folhas de estilo, scripts, imagens, entre outros.

O Nginx também pode funcionar como um *load balancer*, tendo em vista que ele recebe todas as conexões e conhece todos os caminhos para os servidores, bastando apenas a implementação de um ou mais algoritmos de balanceamento de carga, como Round Robin, Least Traffic, Least Latency, URL, entre outros. Uma funcionalidade importante que também pode ser implementada com a ajuda do Nginx é o conceito de *sticky sessions*, que por exemplo, é uma opção para o funcionamento de conexões *websockets* em aplicações distribuídas em múltiplos servidores, devido a necessidade de se manter as requisições traçadas ao mesmo servidor que contém suas informações de conexão.

5 *Implementação*

A implementação do sistema segue todas as diretrizes de tecnologias e conceitos apresentados até este tópico. A descrição de como a implementação foi feita será dividida em duas partes principais: *backend* e *frontend*. Sendo a parte de *backend* a implementação da infraestrutura e servidor da aplicação, e na parte do *frontend* a implementação de interface com o usuário. No desenvolvimento elegemos um nome fictício para a aplicação, chamado Orb. Este nome será utilizado para remeter a aplicação em alguns textos, logos, implementações, entre outros.

O código-fonte da aplicação está hospedado no GitHub [46], como *software open-source*, liberado para utilização através da licença MIT [47]. A abordagem utilizada para a descrição da implementação feita fará referências pontuais a trechos do código-fonte que poderão ser consultados pelo leitor no repositório informado [46], evitando-se ao máximo a exposição de código neste formato de texto que não é favorável a análise deste tipo de conteúdo.

5.1 Backend

Primeiramente iremos descrever a criação do ambiente de desenvolvimento do projeto, onde inicialmente foi instalado o sistema operacional e todas as ferramentas que decidimos utilizar para a construção da aplicação, e posteriormente iniciamos o processo de codificação.

Iniciando-se pelo sistema operacional, foi criada uma partição separada no disco rígido do computador que dispúnhamos, com volume de 60 GBs. O sistema operacional escolhido para o ambiente de desenvolvimento foi de uma distribuição Linux chamada CentOS [49], versão 7. A instalação do sistema operacional foi feita através de um *pendrive* e foram seguidas todas as configurações padrões do *wizard* de instalação do sistema.

Após a instalação do sistema operacional, fizemos a instalação de todas as ferramentas, a nível de O.S., que foram escolhidas para o desenvolvimento da aplicação, como o NodeJS, MongoDB, Redis, editores de texto e navegadores. Todas estas ferramentas foram instaladas através do repositório de pacotes padrão do sistema CentOS 7, chamado YUM. Para algumas

ferramentas foram necessário a atualização das fontes do repositório YUM e outras foram instaladas através de pacotes RPM disponibilizados no próprio site da ferramenta.

A implementação do *backend* iniciou-se com a criação de um diretório chamado *backend* na pasta raiz do projeto, do qual a partir dele foram seguidas recomendações de boas práticas [48] para a organização da estrutura de diretórios em aplicações ExpressJS. A seguinte listagem identifica o nome destes diretórios e suas respectivas funcionalidades:

controllers Este diretório contém os arquivos de código relacionados as rotas e suas lógicas de implementação. Rotas são os *endpoints* que podem ser acessados através de requisições HTTP enviadas por uma determinada URL que a identifica. Neste diretório contém os arquivos das rotas de cliente, usuário e *token*.

domain Este diretório contém os arquivos relacionados a lógica de negócio da aplicação. Os *handlers* que tratam as requisições *websocket* da aplicação estão implementados em arquivos contidos nesta pasta.

helpers Este diretório contém arquivos que implementam funcionalidades que podem ser utilizadas em diversos pontos do sistema, este diretório funciona como um *cross-cutting layer* da aplicação. Nesta pasta estão implementados alguns métodos que ajudam na geração de *random strings* e retorno de *datetime* em formato UTC.

middlewares Este diretório contém arquivos relacionados a *middlewares* de tratamento de requisições que chegam ao servidor, como realização de autenticação e autorização de acesso as rotas da aplicação.

models Este diretório contém arquivos relacionados aos objetos que podem ser persistidos na aplicação. Devido ao Javascript ser uma linguagem dinâmica, estes objetos não representam um contrato de como estes eles devem existir na aplicação, mas como eles devem ser persistidos no banco de dados, que no caso desta aplicação é o MongoDB. Neste diretório serão encontrados os arquivos de persistência dos *tokens*, *chat*, *messages*, *users*, entre outros.

Após a definição da estrutura de diretórios da aplicação, foi realizada a instalação e inicialização do NPM, responsável por gerenciar todos os pacotes que iremos utilizar no desenvolvimento da solução.

5.1.1 Persistência

A persistência dos dados foi feita em um banco de dados MongoDB, onde para sua utilização foi feito o *download* do pacote Mongoose pelo NPM. A conexão com o banco foi feita no arquivo de inicialização da aplicação, através do método `connect` do próprio Mongoose, passando-se como parâmetros uma *connection string* com o endereço local do banco de dados. Por ser um ambiente de testes, não utilizamos autenticação com o banco de dados, mas caso fosse necessário, bastava-se apenas adicionar os dados de *username:password* na própria *connection string*.

No diretório *models* foram criados arquivos para cada objeto que fosse ser persistido no banco. As implementações da persistência utilizaram o Mongoose como modelo, exportando de cada módulo o objeto *schema* de cada persistência, podendo através deste realizar todas as operações de CRUD.

A seguinte imagem ilustra a representação dos objetos persistidos no banco de dados:

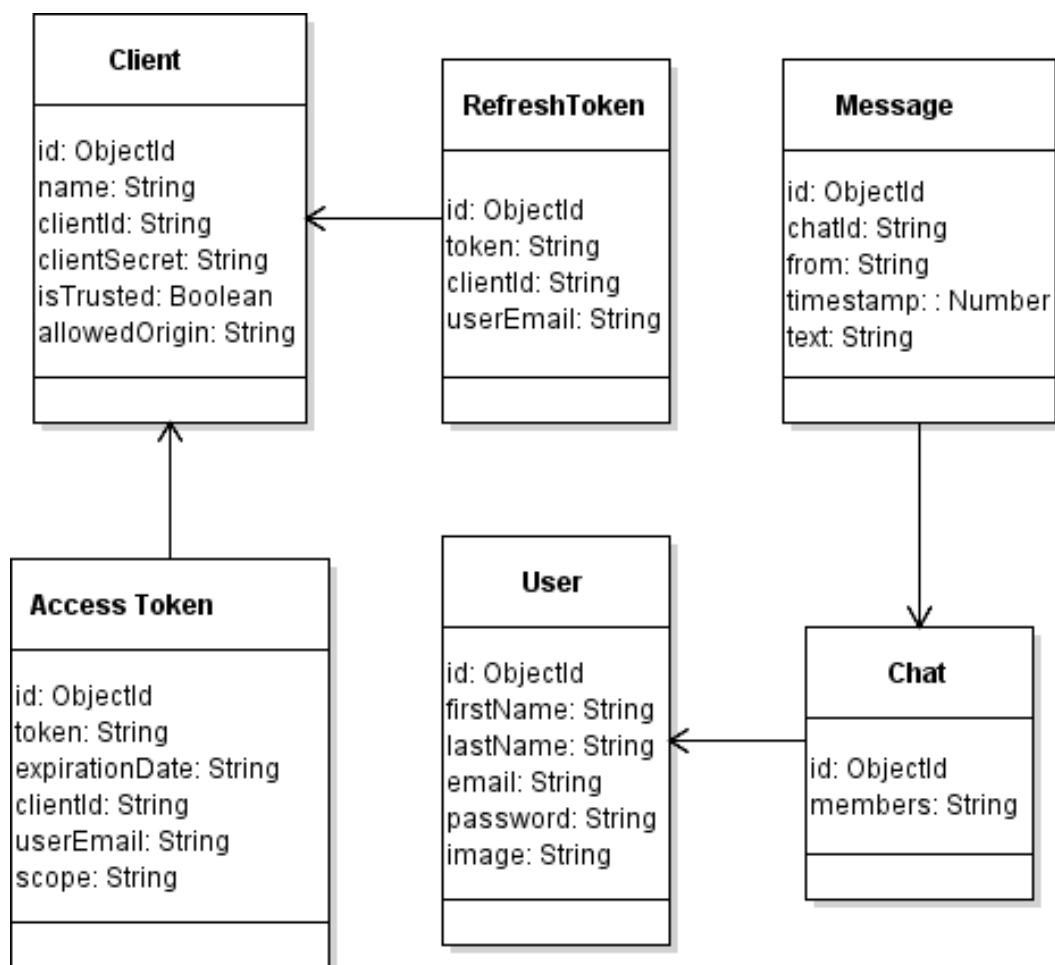


Figura 5.1: Representação UML dos modelos persistidos no banco de dados.

5.1.2 Autenticação e Autorização

Para a implementação do sistema de autenticação, cujos códigos estão no diretório *middlewares*, utilizamos três pacotes NPM, que funcionam como *middlewares* que extraem de um *request* as informações de autenticação, e através de uma *callback* é possível validar essas informações e sinalizar se as informações estão corretas para concretização da autenticação. Para o sistema de autorização, utilizamos o OAuth2 com algumas modificações em seu padrão para atender aplicações *trusted*, como o nosso *frontend app* desenvolvido em AngularJS, que é um tipo de aplicação que não necessita de autorização do usuário da conta para acessar suas informações confidenciais.

Como estamos utilizando um sistema de *tokens*, para realizar acesso ao *endpoint* token, o qual é utilizado no sistema de *login* para se conseguir um Access Token e um Refresh Token através do fornecimento de *e-mail* e senha, é realizada a autenticação pelos padrões Basic [50] e Client [51]. Após esta autenticação, o usuário é encaminhado para um *exchange* do OAuth2 que irá gerar, persistir e retornar um Access Token e um Refresh Token para o solicitante, que poderá usá-lo para acessar os demais *endpoints* da aplicação.

Como já introduzido, o restante das rotas da aplicação são protegidos através do padrão de autenticação Bearer [52], que procura em cada requisição feita um Access Token, que será verificado, e caso seja válido permitirá o acesso ao *endpoint* solicitado. Em caso do Access Token não ser válido, o servidor irá emitir uma resposta de erro, HTTP 401, informando o problema ocorrido, e se o problema for a expiração do Access Token, ele poderá utilizar o Refresh Token para conseguir um novo Access Token válido por mais um determinado tempo, evitando de ter que realizar toda o processo de autenticação com *e-mail* e senha novamente.

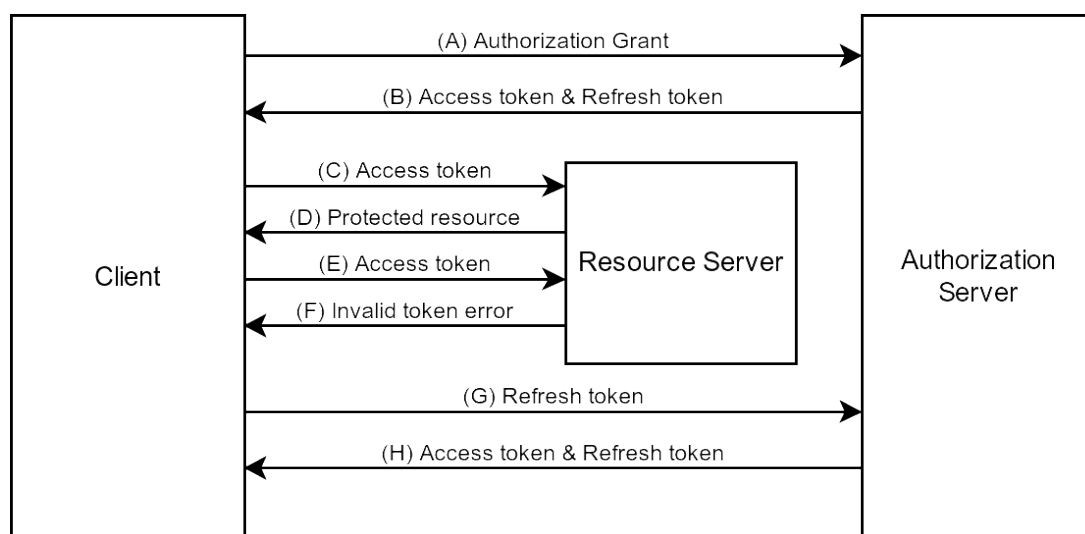


Figura 5.2: Fluxo de acesso aos recursos com OAuth2. Fonte: jlabusch [54]

5.1.3 Rotas

As rotas da aplicação estão implementadas

5.1.4 Socket

5.2 Frontend

6 *Resultados*

7 Trabalhos Futuros

8 *Conclusão*

Referências

- [1] Liveclicker Delivers Dynamic Messaging Capability with RealTime Email Solution. Disponível em: <http://www.destinationcrm.com/Articles/CRM-News/CRM-Featured-News/Liveclicker-Delivers-Dynamic-Messaging-Capability-with-RealTime-Email-Solution-9.aspx>. Acesso em: 19 Out. 2016.
- [2] Liveclicker RealTime Email Solution. Disponível em: <http://www.realtime.email/>. Acesso em: 19 Out. 2016.
- [3] 50 Amazing Uber Statistics (October 2016). Disponível em: <http://expandedramblings.com/index.php/uber-statistics/>. Acesso em: 20 Out. 2016.
- [4] Uber. Disponível em: <https://www.uber.com/>. Acesso em: 20 Out. 2016.
- [5] How Uber Scales Their Real-Time Market Platform. Disponível em: <http://highscalability.com/blog/2015/9/14/how-uber-scales-their-real-time-market-platform.html>. Acesso em: 20 Out. 2016.
- [6] Redis Publish-Subscribe (RedisMVA). Disponível em: https://github.com/sayar/RedisMVA/blob/master/module6_redis_pubsub/README.md#redis-publish-subscribe. Acesso em: 20 Out. 2016.
- [7] Web browsers usage table. Disponível em: <http://caniuse.com/usage-table>. Acesso em: 20 Out. 2016.
- [8] Redis Pub/Sub (Rajaraodv). Disponível em: <https://github.com/rajaraodv/redispubsub>. Acesso em: 20 Out. 2016.
- [9] FanMappr. Disponível em: <http://h.fanapp.mobi/modules/fanmappr/fanmappr.php?fid=1056666>. Acesso em: 20 Out. 2016.
- [10] RadiusIM – Social Messenger Com Busca Por Geotagging. Disponível em: <http://br.wwwhatsnew.com/2008/09/radiusim-%E2%80%93-social-messenger-por-com-busca-por-geotagging/>. Acesso em: 20 Out. 2016.
- [11] NOTNOTCITRICSQUID. Do location based chat apps start trending again?. Disponível em: https://www.reddit.com/r/startups/comments/1tkc3f/do_location_based_chat_apps_start_trending_again/. Acesso em: 21 Out. 2016.
- [12] Snapchat. Disponível em: <https://www.snapchat.com/>. Acesso em: 21 Out. 2016.

- [13] Tinder. Disponível em: <https://www.gotinder.com/>. Acesso em: 21 Out. 2016.
- [14] Happn. Disponível em: <https://www.happn.com/>. Acesso em: 21 Out. 2016.
- [15] How Much Web Traffic Is On Mobile Devices (Really)? Disponível em: <https://keriganmarketing.com/freelunch/view/how-much-web-traffic-is-on-mobile-devices/>. Acesso em: 22 Out. 2016.
- [16] Internet stats and facts for 2016. Disponível em: <https://hostingfacts.com/internet-facts-stats-2016/>. Acesso em: 22 Out. 2016.
- [17] Mobile Marketing Statistics compilation. Disponível em: <http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/>. Acesso em: 22 Out. 2016.
- [18] Notebook HP 1000-1460BR. Disponível em: <http://www.lojahp.com.br/Voce/Notebook/Notebook-HP-1000-1460BR-com-Intel-Core-i5-3230M-Windows-8-4GB-500GB-Gr.html>. Acesso em: 24 Out. 2016.
- [19] Facebook Linux, What Distro is it?. Disponível em: <http://www.internetnews.com/blog/skerner/facebook-linux-CentOS.html>. Acesso em: 24 Out. 2016.
- [20] Red Hat Enterprise Linux Atomic Host expands to Google Compute Engine. Disponível em: <https://cloudplatform.googleblog.com/2014/07/red-hat-enterprise-linux-atomic-host-expands-to-google-compute-engine.html?m=1>. Acesso em: 24 Out. 2016.
- [21] Gitflow Workflow. Disponível em: <https://www.atlassian.com/git/tutorials/comparing-workflows/gitflow-workflow>. Acesso em: 24 Out. 2016.
- [22] Sublime Text. Disponível em: <https://www.sublimetext.com/>. Acesso em: 24 Out. 2016.
- [23] Package Control. Disponível em: <https://packagecontrol.io/>. Acesso em: 24 Out. 2016.
- [24] Visual Studio Code. Disponível em: <https://code.visualstudio.com/>. Acesso em: 24 Out. 2016.
- [25] NodeJS. Disponível em: <https://nodejs.org/>. Acesso em: 24 Out. 2016.
- [26] Chrome V8. Disponível em: <https://developers.google.com/v8/>. Acesso em: 24 Out. 2016.
- [27] Introdução ao Node.js. Disponível em: <http://maxroecker.github.io/blog/introducao-ao-nodejs/>. Acesso em: 24 Out. 2016.
- [28] Libuv. Disponível em: <https://github.com/libuv/libuv>. Acesso em: 24 Out. 2016.
- [29] NPM. Disponível em: <https://www.npmjs.com/>. Acesso em: 24 Out. 2016.
- [30] ExpressJS. Disponível em: <https://expressjs.com/>. Acesso em: 24 Out. 2016.

- [31] Criando Salas com Socket.IO. Disponível em: <http://blog.ifollow.com.br/2015/06/01/criando-salas-com-socket-io/>. Acesso em: 25 Out. 2016.
- [32] Mongoose. Disponível em: <http://mongoosejs.com/>. Acesso em: 25 Out. 2016.
- [33] Redis. Disponível em: <http://redis.io/>. Acesso em: 25 Out. 2016.
- [34] Using New Relic to Understand Redis Performance: The 7 Key Metrics. Disponível em: <https://blog.newrelic.com/2015/05/11/redis-performance-metrics/>. Acesso em: 25 Out. 2016.
- [35] Bower. Disponível em: <https://bower.io/>. Acesso em: 25 Out. 2016.
- [36] AngularJS. Disponível em: <https://angularjs.org/>. Acesso em: 25 Out. 2016.
- [37] Angular Material. Disponível em: <https://material.angularjs.org/>. Acesso em: 25 Out. 2016.
- [38] Gulp. Disponível em: <http://gulpjs.com/>. Acesso em: 25 Out. 2016.
- [39] SASS. Disponível em: <http://sass-lang.com/>. Acesso em: 25 Out. 2016.
- [40] HAML. Disponível em: <http://haml.info/>. Acesso em: 26 Out. 2016.
- [41] HAMLJS. Disponível em: <https://github.com/creationix/haml-js>. Acesso em: 26 Out. 2016.
- [42] WebPack. Disponível em: <https://webpack.github.io/>. Acesso em: 26 Out. 2016.
- [43] Babel. Disponível em: <https://babeljs.io/>. Acesso em: 26 Out. 2016.
- [44] Nginx. Disponível em: <https://www.nginx.com/>. Acesso em: 26 Out. 2016.
- [45] NAT 1:1 ou Proxy Reverso, qual a melhor alternativa? Disponível em: <http://jff.eti.br/nat-11-ou-proxy-reverso-qual-a-melhor-alternativa/>. Acesso em: 26 Out. 2016.
- [46] Orb. Disponível em: <https://github.com/joaopsh/orb>. Acesso em: 26 Out. 2016.
- [47] MIT License. Disponível em: <https://opensource.org/licenses/MIT>. Acesso em: 26 Out. 2016.
- [48] Best practices for Express app structure. Disponível em: <https://www.terlici.com/2014/08/25/best-practices-express-structure.html>. Acesso em: 27 Out. 2016.
- [49] CentOS. Disponível em: <https://www.centos.org/>. Acesso em: 27 Out. 2016.
- [50] Passport-HTTP. Disponível em: <https://github.com/jaredhanson/passport-http>. Acesso em: 27 Out. 2016.
- [51] passport-oauth2-client-password. Disponível em: <https://github.com/jaredhanson/passport-oauth2-client-password>. Acesso em: 27 Out. 2016.
- [52] passport-http-bearer. Disponível em: <https://github.com/jaredhanson/passport-http-bearer>. Acesso em: 27 Out. 2016.

-
- [53] OAuth2orize. Disponível em: <https://github.com/jaredhanson/oauth2orize>. Acesso em: 27 Out. 2016.
- [54] OAUTH 2.0 SERVER. Disponível em: <http://jlabusch.github.io/oauth2-server/>. Acesso em: 27 Out. 2016.