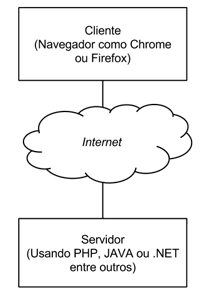
# HTTP: Entendendo a web por debaixo dos panos

## O que é HTTP:

Nesse treinamento focaremos nos fundamentos da web. Isto é importante pois a grande maioria das aplicações hoje em dia a utilizam de alguma forma ou funcionam dentro dela. Não focaremos em nenhuma plataforma específica de desenvolvimento como Java ou PHP. Focaremos nas regras de comunicação da web.

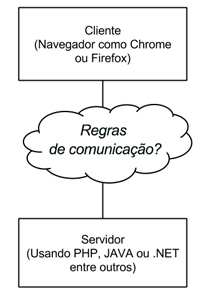
Quando se fala em HTTP, o primeiro pensamento que vem a nossa mente é sobre a utilização da **internet**, é o cenário onde vemos realmente na prática a utilização do HTTP. Nós acessamos sites em que seus endereços iniciam com *http://* e por isso precisamos conhecer o que realmente está acontecendo ao fazer isso.

No momento em que acessou este curso, esta aula, entre o navegador e a Alura aconteceu uma comunicação, e esta comunicação tem duas partes bem conhecidas que chamamos de **Client-Server** ou em português **Cliente-Servidor**. Este é um modelo arquitetural, ou seja, a internet inteira é baseada nesta arquitetura onde há um cliente que solicita e um servidor que responde.

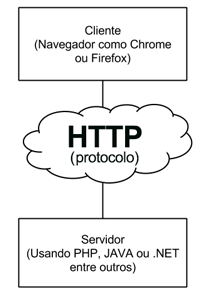


Em qualquer comunicação é preciso existir algumas regras para que as duas partes consigam se entender com sucesso. Pensando na comunicação do seu navegador entre a Alura ou algum outro site esse conjunto de regras é basicamente um **protocolo**, onde neste cenário é o **HTTP**.

Os protocolos são definidos, especificados e disponibilizados para implementação em ambas as partes, para consultar a especificação do HTTP, você pode utilizar o seguinte endereço: <https://tools.ietf.org/html/rfc2616>



Resumindo: O **HTTP é um protocolo** que define as **regras de comunicação** entre **cliente e servidor** na internet. Vamos focar nos próximos vídeos e entender melhor esse protocolo tão importante. mãos a obra!



**O que aprendemos neste capítulo?**

* Na internet sempre tem um cliente e um servidor
* Entre o cliente e o servidor precisam haver regras de comunicação
* As regras são definidas dentro de um protocolo
* HTTP é o protocolo mais importante na internet
* Você já usou torrent para baixar algum arquivo na internet? Caso sim, aproveitou um outro modelo de comunicação, o **P2P** ou Peer-To-Peer!
* O modelo **Cliente-Servidor** não é o único modelo de comunicação na rede, nem sempre o mais adequado. Por exemplo, imagine que precisemos contar as letras de 20 palavras. No caso do modelo Cliente-Servidor, quem fará esse trabalho é o servidor, certo? E se precisar contar as letras de 1 milhão de palavras? Muito trabalhoso para o servidor, não?
* O modelo **Cliente-Servidor** tenta centralizar o trabalho no servidor, mas isso também pode gerar gargalos. Se cada Cliente pudesse ajudar no trabalho, ou seja, assumir um pouco da responsabilidade do servidor, seria muito mais rápido. Essa é a ideia do P2P! Não há mais uma clara divisão entre Cliente-Servidor, cada cliente também é servidor e vice-versa!
* Isto é útil quando você precisa distribuir um trabalho ou necessita baixar algo de vários lugares diferentes. Faz sentido?
* Usando algum aplicativo de Torrent, o protocolo utilizado não é o HTTP, e sim o protocolo **P2P**, como BitTorrent ou Gnutella.

Agora já sabemos que existe um cliente, o navegador, como Chrome e Firefox, e um servidor, a Alura. Para definir as regras de comunicação entre cliente e servidor, existe o protocolo HTTP.

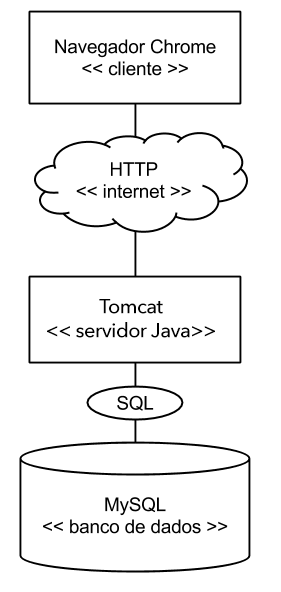
Também já sabemos que o servidor usa alguma plataforma, como PHP, Java, .Net ou outros. Qual plataforma realmente é utilizada? Não é tão fácil de descobrir, pois o HTTP, de propósito, não está focado em alguma plataforma específica e esconde isso de nós. Bom, eu não vou esconder nada e vou contar para vocês que a Alura usa a plataforma Java e o servidor concreto se chama Tomcat.

Também já falamos que o SQL é uma linguagem para consultar o banco de dados. Alura usa SQL para acessar o banco de dados MySQL.

Com essas informações já temos uma breve ideia da arquitetura da Alura!

Cliente <--- HTTP ---> Servidor Java <--- SQL ---> Banco de dados

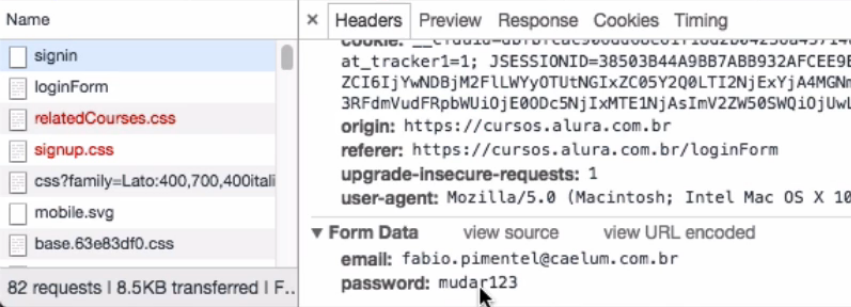
Visualizando:



## HTTPS - a versão segura do HTTP

Sabendo que o HTTP é o protocolo que define as regras de comunicação na web, precisamos observar algumas coisas. Quando usamos o HTTP, todos os dados enviados entre cliente e servidor são *transmitidos em texto puro*, inclusive dados sensíveis, como login e senha!

Quando acessamos a Alura por exemplo, precisamos fornecer informações de autenticação, essas informações são nosso email e senha, que são enviadas e validadas pela plataforma para que assim consigamos assistir as aulas. Estas informações são enviadas em texto limpo e é possível visualizá-las pelas ferramentas do desenvolvedor do navegador. A aba *network* nos possibilita isso.



Mas por que é importante sabermos isso? Quando o navegador pede informações da Alura, nessa comunicação há vários intermediários. Por exemplo, usando uma conexão Wi-Fi, os dados do navegador passam primeiro para o roteador Wi-Fi, e do roteador passam para o modem do provedor, do modem para algum servidor do provedor de internet, como Oi ou NET.

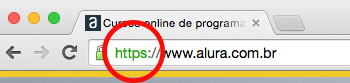
É muito provável que existam outros servidores intermediários no provedor antes que os dados realmente cheguem no servidor da Alura. Com a resposta é a mesma coisa, ela volta passando por esses servidores no meio antes de chegar até nosso navegador. O problema é, quando usamos HTTP, qualquer servidor no meio pode espionar os dados enviados, algo totalmente inseguro! Imagine se essas informações fossem relativas a contas bancárias. Não seria nada seguro!

Para estes outros cenários, existe o HTTP**S**, que basicamente é o HTTP comum, porém com uma camada adicional de segurança/criptografia que antes era SSL, mas posteriormente passou a ser também TLS. É muito comum que estas duas siglas sejam encontradas juntas como SSL/TLS por se tratarem da mesma questão de segurança. Sendo assim, temos dois termos:

1. HTTP: HyperText Transfer Protocol
2. SSL/TLS: Secure Sockets Layer / Transport Layer Security

## Funcionamento do HTTPS:

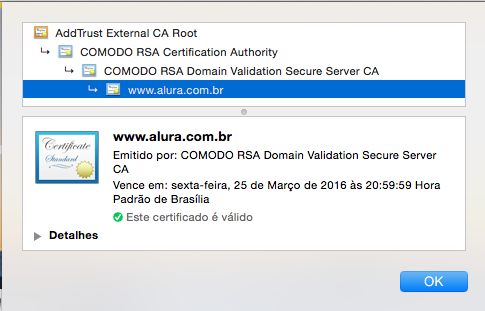
Ao acessarmos o [site da Alura](https://www.alura.com.br/) pelo navegador podemos perceber que ele já usa o protocolo **HTTPS**:



Reparem que no navegador, ao lado do https, aparece um cadeado e que ao clicarmos no cadeado podemos ver mais informações sobre HTTPS. Uma dessas informações indica que a Alura tem uma identidade confirmada. O que isso quer dizer?

O HTTPS para garantir segurança usa criptografia baseada em chaves públicas e privadas e para gerar essas chaves publicas e privadas é preciso garantir a identidade de quem possui essas chaves e isso é feito a partir de um **certificado digital**, ou seja, um certificado digital é utilizado para **identificar** determinada entidade e ainda é utilizada para geração das chaves de criptografia.

Apesar disso, ainda é necessário que uma **autoridade certificadora**, que nada mais é que um órgão ou entidade confiável, garanta não apenas a identidade do site mas também a validade do certificado. No caso da Alura a autoridade certificadora é a COMODO RSA Domain Validation, mas existem outras.



Dito isso, como tudo funciona? Os navegadores em posse da chave pública criptografam as informações e as enviam para o servidor que as descriptografa com a chave privada. É importante notar que apenas a chave privada descriptografa as informações criptografadas com a pública, e também que deve-se manter a chave privada segura.

## O que aprendemos nesse capítulo?

* Só com **HTTPS a web é segura**.
* HTTPS significa usar um **certificado digital** no servidor.
* O certificado prova a **identidade** e tem **validade**
* O certificado possui uma **chave publica**.
* A chave é utilizada pelo navegador para criptografar os dados.

## Certificado Digital

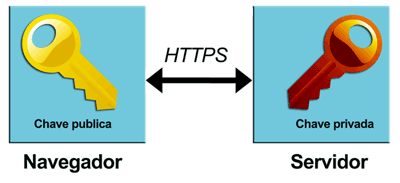
Quando precisamos informar nossos dados a algum servidor, queremos ter certeza que este servidor realmente representa a entidade em questão. Queremos confiar em quem estamos fornecendo nossos dados!

Um certificado digital prova uma identidade para um site, onde temos informações sobre o seu domínio e a data de expiração desse certificado.

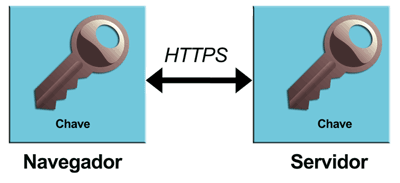
Além disso, o certificado ainda guarda a chave pública que é utilizada para criptografar (cifrar) os dados que são trafegados entre cliente e servidor.

## As chaves do HTTPS

Aprendemos no vídeo que o HTTPS usa uma **chave pública** e uma **chave privada**. As chaves estão ligadas matematicamente, o que foi cifrado pela chave pública só pode ser decifrado pela chave privada. Isso garante que os dados cifrados pelo navegador (chave pública) só podem ser lidos pelo servidor (chave privada). Como temos duas chaves diferentes envolvidas, esse método de criptografia é chamado de **criptografia assimétrica**. No entanto, a criptografia assimétrica tem um problema, ela é **lenta**.



Por outro lado, temos a **criptografia simétrica**, que usa a mesma chave para cifrar e decifrar os dados, como na vida real, onde usamos a mesma chave para abrir e fechar a porta. A criptografia simétrica é muito **mais rápida**, mas infelizmente não **tão segura**. Como existe apenas uma chave, ela ficará espalhada pelos clientes (navegadores) e qualquer um, que tem a posse dessa chave, pode decifrar a comunicação.



Agora, o interessante é que o **HTTPS usa ambos os métodos de criptografia, assimétrica e simétrica**. Como assim? Muita calma, tudo o que aprendemos é verdade! Só faltou o grande final:)

No certificado, vem a chave pública para o cliente utilizar, certo? E o servidor continua na posse da chave privada, ok? Isso é seguro, mas lento e por isso o cliente gera uma chave simétrica ao vivo. Uma chave só para ele e o servidor com o qual está se comunicando naquele momento! Essa chave exclusiva (e simétrica) é então enviada para o servidor utilizando a criptografia assimétrica (chave privada e pública) e então é utilizada para o restante da comunicação.

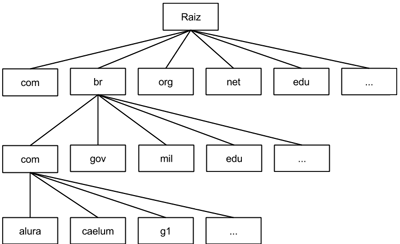
Então, HTTPS **começa** com criptografia **assimétrica** para **depois** mudar para criptografia **simétrica**. Essa chave simétrica será gerada no início da comunicação e será reaproveitada nas requisições seguintes. Bem-vindo ao mundo fantástico do HTTPS:)

## Endereços

Você que está usando a Alura, já conhece então o endereço: **https://www.alura.com.br**. Já sabemos que o endereço começa com http ou https. Repare que depois do nome do protocolo vem **://** seguido pelo nome do site, que é **www.alura.com.br**. No vocabulário de um desenvolvedor o **www.alura.com.br** é o **domínio** (ou domain). A abreviação **www** representa o **world wide web**.

## Analisando o domínio

Vamos dar uma olhada com mais carinho no nome do domínio. Olhando da direita para a esquerda, o domínio começa com br, indicando um site do Brasil. O br representa o top level domain, está na raiz do domínio. Depois vem o com, abreviação de comercial e alura. O com e o alura são sub-domínios.



O www representa também um sub-domínio, no entanto seu uso é opcional, tanto que alura.com.br e www.alura.com.br funcionam e mostram a mesma página. A maioria dos site usam o prefixo www e podemos dizer que isso é algo legado que continua ser popular apesar de não ser necessário.

## Subdomínios

Existe também a ideia de subdomínios, que representam sessões específicas dentro de um site. Por exemplo, no caso do Gmail temos o endereço: mail.google.com , ou ainda no caso do Google Drive: drive.google.com. Tanto Gmail como Drive são subdomínios do domínio Google.

Perceba que esses subdomínios apontam para páginas diferentes dentro do mesmo domínio (Google).

## Endereços IP's

O nome do domínio é organizado em uma hierarquia que foi criada para organizar os sites na internet e para a gente ter algo fácil para se lembrar. Para ser correto, a internet funciona na verdade sem esses domínios. Aqueles nomes são coisas dos humanos, as máquinas na internet têm uma outra forma de se endereçar. Elas usam o que se chama endereços de IP, que nada mais são do que números - muito difícil para gente decorar.

Sendo assim, podemos acessar o Google usando um IP. Para quem é muito curioso, mas muito curioso mesmo e deseja saber qual é o IP do Google, pode digitar na linha de comando a seguinte instrução:

nslookup google.com

...

Name: google.com

Address: 216.58.202.238

Esse comando procura o número IP do Google na internet. Podemos até usar esse endereço no navegador: [http://216.58.202.238](http://216.58.202.238/)

A página principal do Google deve ser exibida. IP's são mais importantes para quem trabalha com rede. O desenvolvedor normalmente não precisa mexer com isso.

Observação: Esse IP pode mudar dependendo do servidor concreto onde o Google foi instalado.

## DNS

Mas a gente não acessa a Google ou a Alura por um número como 52.206.164.173 e sim pela URL.

Ainda bem, não é verdade? Seria inviável decorar todos os serviços e sites que acessamos diariamente apenas por números.

Acontece que por baixo dos panos quando realizamos uma requisição essa URL é transformada em um número por um serviço transparente chamado de **DNS** (***D***omain ***N***ame ***S***ystem).

Esse serviço age como um grande banco de dados de domínios. Portanto quando fazemos uma requisição para alura.com.br o DNS age transformando para um IP e a requisição prossegue.

Podemos inclusive escolher um servidor DNS de preferência na nossa internet. Um bastante usado é o do próprio Google: <https://developers.google.com/speed/public-dns/>

## Portas:

Como já falamos, quando usamos a URL **https://www.alura.com.br** abrimos uma conexão com o servidor que roda em algum lugar na internet. Para estabelecer uma conexão na rede é preciso saber qual é o endereço IP, e já vimos como descobri-lo.

Agora imagine que o servidor é uma casa: dependendo da casa há várias portas disponíveis. O que é preciso saber é qual porta devemos utilizar quando chegarmos na casa. Ou seja, devemos saber qual porta é utilizada para o protocolo HTTP!

## Abrindo portas

A porta reservada para o protocolo HTTP é o 80. Novamente um número, e como o navegador já sabe essa porta padrão, podemos omiti-la, mas nada nos impede adicioná-la no endereço, por exemplo:

http://www.alura.com.br:80

Vai funcionar normalmente, tanto que o navegador esconde automaticamente :80. Vamos tentar uma outra porta, outro numero, por exemplo 81:

http://www.alura.com.br:81

Não funciona, pois essa porta não está aberta no servidor, não podemos estabelecer uma conexão e o tempo de conexão vai esgotar. Igualmente, o protocolo HTTPS possui uma porta padrão, a porta 443, que também podemos omitir ao acessarmos um endereço HTTPS. Podemos testar também e ver que funciona normalmente.

<https://www.alura.com.br:443>

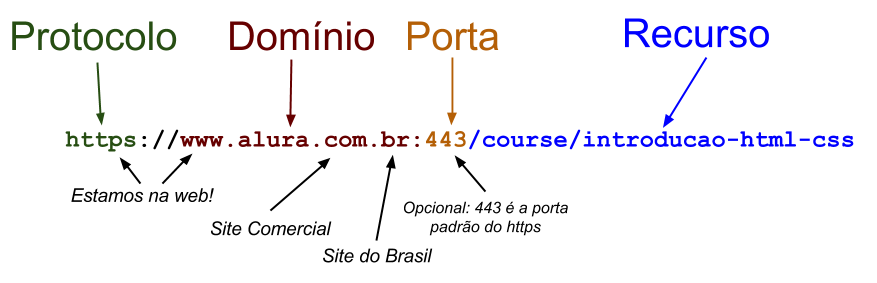
## Recursos

Navegando dentro da Alura, mais informações aparecem depois do nome e do domínio. Por exemplo, para acessar a página principal dos cursos, usamos **https://cursos.alura.com.br/dashboard**. O /dashboard é um recurso (**resource**) do site que gostaríamos de acessar. Existem vários outros recursos na Alura como as carreiras (/careers), ou o fórum de discussões (/forum). O importante é que cada recurso possua o seu nome único.

Navegando um pouco mais na Alura, podemos perceber que entre o domínio e o recurso podem vir outras informações . Por exemplo, para acessar o curso **HTML5 e CSS3 I: Suas primeiras páginas da Web**, usamos <https://cursos.alura.com.br/course/introducao-html-css>. Ou seja, para acessarmos o recurso /introducao-html-css, usamos um caminho intermediário, o /course. Há vários outros exemplos na Alura que usam caminhos para chegar ao recurso concreto, como por exemplo /courses/mine, e navegando na Alura você encontrará mais.

## Finalmente, a URL

Repare que estamos usando umas regras bem definidas para descrever a localização de um recurso na web. Todos os endereços na web sempre seguem esse mesmo padrão: **protocolo://dominio:porta/caminho/recurso**. Esse padrão, na verdade, segue uma especificação que foi batizada de **Uniform Resource Locator**, abreviada como **URL**. Então, as URLs são os endereços na web!



## URI ou URL

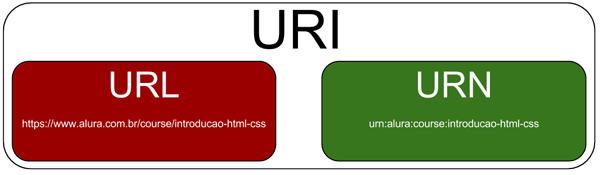
Muitas vezes, desenvolvedores usam a sigla **URI** (***U****niform* ***R****esource* ***I****dentifier*) quando falam de endereços na web. Alguns preferem **URL** (***U****niform* ***R****esource* ***L****ocator*), e alguns misturam as duas siglas à vontade. Há uma certa confusão no mercado a respeito e mesmo desenvolvedores experientes não sabem explicar a diferença. Então, qual é a diferença?

**Resposta 1 (fácil):** Uma **URL** é uma **URI**. No contexto do desenvolvimento web, ambas as siglas são válidas para falar de endereços na web. As siglas são praticamente sinônimos e são utilizadas dessa forma.

**Resposta 2 (mais elaborada):** Uma **URL** é uma **URI**, mas nem todas as **URI's** são **URL's**! Existem **URI's** que identificam um recurso sem definir o endereço, nem o protocolo. Em outras palavras, uma **URL** representa uma *identificação* de um recurso (**URI**) através do endereço, mas nem todas as identificações são **URL's**.

Humm ... ficou claro? Não? Vamos dar um exemplo! Existe um outro padrão que se chama **URN** (***U****niform* ***R****esource* ***N****ame*). Agora adivinha, os **URN's** também são **URI's**! Um **URN** segue também uma sintaxe bem definida, algo assim **urn:cursos:alura:course:introducao-html-css**. Repare que criamos uma outra identificação do curso **Introdução ao HTML e CSS** da Alura, mas essa identificação não é um endereço.

Novamente, a resposta 2 vai muito além do que você realmente precisa no dia a dia. *Normalmente****URL****e****URI****são usados como sinônimos*.



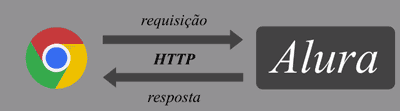
## Modelo Requisição e Resposta

Já descobrimos que o HTTP é um protocolo que define as regras de comunicação entre cliente e servidor e de que as URLs são constituídas. Porém isso não é tudo, vejamos mais alguns detalhes sobre o funcionameto da Web e do HTTP.

Realizaremos um teste **efetuando login** na Alura. Quando preenchemos o formulário e clicamos no botão, o navegador envia o nosso login e a nossa senha para o servidor através do protocolo HTTP! Vamos detalhar um pouco esta ação.

No mundo HTTP, a requisição enviada pelo navegador para o servidor é chamada de **HTTP REQUEST**. Recebemos a página /dashboard como resposta já que enviamos login e senha válidos. No mundo HTTP essa resposta é chamada de **HTTP RESPONSE**.

A comunicação segue sempre esse modelo: o cliente envia uma requisição e o servidor responde. **Requisição e Resposta** ou em inglês: **Request-Response**. Aqui é importante saber que a comunicação sempre começa com o cliente: é ele quem pede as informações. O servidor responde apenas o que foi requisitado e nunca inicia a comunicação!



## Comunicação sem estado

Vamos acessar rapidamente outro site: http://g1.globo.com. Para este acesso estamos enviando uma requisição para g1 e recebemos como resposta a página inicial.

Agora vamos navegar dentro do site e acessar algum artigo. Ao clicarmos enviamos uma nova requisição e percebemos que TODA página foi trocada. Fica mais claro ainda se acessarmos do menu acima algum link (globo esporte ou globo show). Podemos ver que todo o conteúdo do site foi trocado.

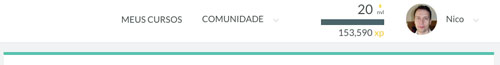
Isso também acontece no caso da Alura (talvez um pouco mais difícil de perceber). Ao acessarmos recursos diferentes todo o conteúdo no navegador foi trocado (apesar do menu parecer o mesmo, ele também foi trocado). A ideia do HTTP é justamente essa, cada recurso é independente do outro e não depende do anterior. Isso também se aplica para os dados enviados na requisição. Cada requisição é independente da outra e ela sempre deve conter todas informações para o servidor responder.

Pense que HTTP funciona como o envio de cartas pelo correio e uma carta representa uma requisição. Você fez uma viagem e gostaria de enviar 3 cartas para sua mãe. Adianta falar para os correios "eu vou colocar o endereço apenas na primeira carta, ai vocês já sabem para onde enviar a segunda e terceira carta"? Não adianta pois os correios tratam cada carta independentemente, e assim também funciona o HTTP. Cada requisição (carta) precisa ter todas as informações. A mesma coisa se aplica para a resposta, precisa ter todas as informações.

Essa característica de cada requisição ser independente é chamada de **stateless**. É esse nome bonito mesmo! O HTTP é um protocolo que não mantém o estado de requisições. Isso significa que só com HTTP não há como se lembrar das requisições anteriores enviadas para o servidor. Por isso precisamos incluir em cada requisição todas as informações, sempre. Para o desenvolvedor este conhecimento é importante pois é justamente essa característica stateless que o atrapalha no dia a dia. Ele precisa preparar a aplicação web para que funcione bem usando o protocolo HTTP, algo que veremos nos treinamentos da Alura.

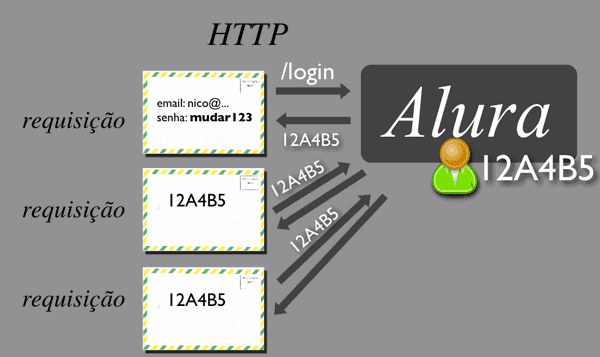
## Lidando com Sessões

Reparem que, mesmo após termos realizado o login e termos enviado várias requisições, aparece o ícone com a minha imagem no menu principal.



Ou seja, a Alura se lembra de alguma forma que eu fiz login em alguma requisição anterior. Como falamos antes, cada requisição deve enviar todas as informações para gerar a resposta. Isso significa que o navegador envia em cada requisição informações sobre o meu usuário! Se cada requisição for independente uma da outra, e não tiver como se lembrar das requisições anteriores, não tem outra explicação a não ser que o navegador envie os dados sobre o meu usuário em cada requisição! Lembre-se da carta postal, ela sempre precisa ter os dados do remetente e aqui não é diferente!

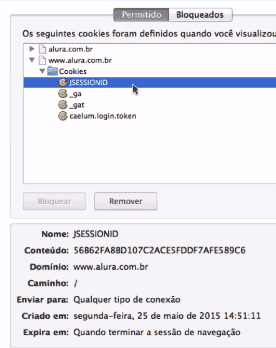
Então o navegador envia o login e senha em cada requisição? Não, não seria muito elegante nem seguro fazer isso. Mas ele faz algo parecido, acreditem ou não. Quando efetuamos o login, a Alura valida os nossos dados, certo? Nesse momento, o servidor tem certeza que o usuário existe e gera uma identificação quase aleatória para o usuário. Essa identificação é um número criado ao vivo e muito difícil de adivinhar. Esse numero é a identificação temporária do usuário e ele será devolvido na resposta.



## Conhecendo cookies

Então onde está esse número? O navegador grava esse número em um arquivo especial para cada site, são os famosos **cookies**. Como acessar esse cookie depende um pouco do navegador em uso. O mais importante é entender o porquê da existência desse número e onde ele foi gravado.

No Chrome podemos ver todos os cookies armazenados nas **Configurações -> Privacidade -> Configurações de conteúdo... -> Todos os cookies e dados de site...**. Se procurarmos por **Alura**, em **cursos.alura.com.br**, lá temos um cookie com o nome **caelum.login.token**, que contém o número da identificação. Se apagarmos esse cookie, perderemos nossa identificação, sendo assim, a Alura exigirá um novo login pois não saberá que já tínhamos logado.



Normalmente o nome do cookie é algo como session-id, dependendo da plataforma de desenvolvimento utilizada ele pode se chamar de PHPSESSID ou ASP.NET\_SessionId ou JSESSIONID ou outro nome que foi inventado! O Cookie será gerado de forma transparente pela tecnologia que você for utilizar para criar aplicativos web. Esse nome, PHPSESSIONID, JSESSIONID ou outro, é gerado pela ferramenta de gerenciamento de Sessão. Por isso ela muda o nome. Se você está usando PHP, então o PHP gerará o nome do Cookie e seu identificador (número aleatório) e chamará o cookie PHPSESSIONID. No Java já será usado o nome JSESSIONID.

Resumindo, todas as plataformas ajudam a gerar esse número e a criar o cookie de maneira transparente. É dessa forma que as plataformas gerenciam as **SESSÕES** com o usuário. Como isso funciona de modo concreto você aprenderá nos cursos e carreiras específicas.

A ideia de manter dados entre requisições é algo muito comum no desenvolvimento de aplicações na web. Um usuário que se loga no sistema web causa a criação de uma sessão. Uma sessão então é útil para guardar informações sobre o usuário e ações dele. Um exemplo clássico é um carrinho de compras. Entre várias requisições estamos usando o mesmo carrinho de compras que guarda os nossos produtos escolhidos (fizemos uma sessão de compras online).

Resumindo teremos:

* O HTTP usa sessões para salvar informações do usuário
* Sessões só são possíveis por uso de Cookies
* Cookies são pequenos arquivos que guardam informações no navegador
* O HTTP é stateless, não mantem estado.

## Cookies

Quando falamos de **Cookies** na verdade queremos dizer **Cookies HTTP** ou **Cookie web**. Um cookie é um pequeno arquivo de texto, normalmente criado pela aplicação web, para guardar algumas informações sobre usuário no navegador. Quais são essas informações depende um pouco da aplicação. Pode ser que fique gravado alguma preferência do usuário. Ou algumas informações sobre as compras na loja virtual ou, como vimos no vídeo, a identificação do usuário. Isso depende da utilidade para a aplicação web.

Um cookie pode ser manipulado e até apagado pelo navegador e, quando for salvo no navegador, fica associado com um domínio. Ou seja, podemos ter um cookie para www.alura.com.br, e outro para www.caelum.com.br. Aliás, um site ou web app pode ter vários cookies!

Podemos visualizar os cookies salvos utilizando o navegador. Como visualizar, depende um pouco do navegador em questão:

No Chrome: **Configurações -> Privacidade -> Configurações de conteúdo... -> Todos os cookies e dados de site... -> Pesquisar alura**

No Firefox: **Preferências -> Privacidade -> remover cookies individualmente -> Pesquisar alura**

## Depurando o método HTTP

Vamos fazer um teste e usar o navegador para mostrar mais detalhes sobre a comunicação HTTP. Os navegadores mais populares como Google Chrome, Mozilla Firefox ou Microsoft Edge possuem ferramentas e plugins que visualizam como o navegador trabalha e usa o HTTP.

Para habilitar as ferramentas do desenvolvedor no **Chrome** vá ao menu à direita (as reticências na vertical): **Mais ferramentas -> Ferramentas do desenvolvedor**, ou no menu superior: **Ferramentas -> Ferramentas do desenvolvedor**. Após isso, selecionamos a aba **Network**.

No **Firefox** vá ao menu superior: **Ferramentas -> Desenvolvedor web -> Exibir/Ocultar ferramentas**.

Para o **Internet Explorer** aperte a tecla **F12** para abrir o console do desenvolvedor e selecione a aba Rede (ou Network).

# Método GET do HTTP

Vamos abrir o console de desenvolvedor e acessar o [http://www.alura.com.br](http://www.alura.com.br/). Aqui usaremos o navegador Chrome, mas nos outros navegadores o comportamento é bem parecido.

No console podemos ver todas as requisições HTTP executadas pelo Chrome. Mas não só isso, também aparecem alguns códigos e métodos, além do tempo de execução para cada requisição. Repare que chamamos apenas o http://www.alura.com.br, mas foram feitas várias outras requisições em seguida.



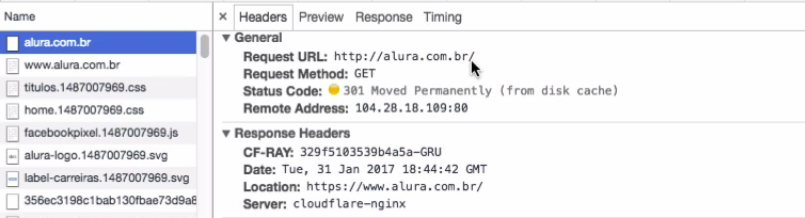
Na primeira coluna aparece a URL (o endereço) e na segunda coluna o método HTTP. O método HTTP indica qual é a **intenção ou ação dessa requisição**. Enviamos uma requisição com o método **GET**. Queremos receber informações, sem modificar algo no servidor, que é justamente a ideia do método GET.

# Primeiro código da resposta

Como resposta recebemos o código de status 301. O protocolo HTTP define alguns códigos padrões para esclarecer a resposta. Indo com o mouse em cima do 301 o Chrome mostra o significado desse código: **Moved Permanently**. Ou seja, o site Alura foi movido para outro lugar! Eis a questão: Onde então está o site Alura?

A localização ou a URL concreta está na resposta HTTP. Vamos clicar em cima do código de status 301 para receber mais informações. Aqui o Chrome mostra todos os cabeçalhos da requisição e da resposta. São muitos (nem tantos) mas o que nos interessa é a nova localização do site. Dentro do item Response Headers podemos ver todos os cabeçalhos que o servidor devolveu e logo logo apareceu um com o nome **Location**. Esse cabeçalho indica a nova URL, só que agora usando **https**.

Quando o navegador recebe o status 301 ele já sabe que é preciso enviar uma nova requisição e procura a nova URL no cabeçalho de resposta Location.



# Redirecionando entre sites

Se alguém acessa a Alura usando http (lembrando, inseguro) automaticamente é chamado o site seguro (https). Isto é um comportamento muito comum para garantir que usamos https sempre. Se esquecermos de usar https, o servidor devolve o status 301 com a nova localização, mas agora usando https.

O navegador, ao receber 301, chama automaticamente a nova URL. No mundo de desenvolvimento web este comportamento é chamado de Redirecionamento pelo navegador, ou **Redirecionamento no lado do cliente**. Fomos redirecionados para o recurso correto. A tarefa do desenvolvedor é definir o código de resposta e, no caso em que algum recurso tenha mudado a URL, o código 301 será usado com o cabeçalho Location.

# O código 200

Continuando no console, a segunda requisição foi para https://www.alura.com.br, novamente usando o método HTTP GET com a intenção de receber dados. Agora o código de resposta foi 200. Este é um dos códigos mais comuns e significa que tudo deu certo! Dessa vez não foi preciso fazer um redirecionamento (não tem o cabeçalho Location na resposta) e não deu nenhum outro problema. A requisição foi aceita e processada corretamente - código 200. Perfeito.

# Tipos de dados diferentes

No console podemos ver que aparecem mais requisições (cada linha representa um novo request). Quando o servidor Alura devolve a resposta para o navegador vem o conteúdo da página inicial em um formato especial, chamado de HTML. O HTML define a estrutura da nossa página, define os menus, botões, links, rodapé etc. Mas dentro do HTML não vêm as imagens e outros arquivos necessários para deixar o site perfeito. Dentro dele vem apenas a URL (endereço) desses outros recursos.

Então, ao receber o HTML, o navegador dispara várias outras requisições para carregar as imagens, fontes e outros dados. Como também são requisições HTTP, o console mostra suas informações. Podemos ver que na resposta vem o tipo do conteúdo, por exemplo text/html, text/css, image/svg+xml, entre outros.

O importante é saber que o protocolo HTTP não está preso em algum formato específico. Podemos trafegar qualquer informação com ele, seja texto ou binário!

## Depurando os códigos de resposta HTTP

Vamos fazer um teste e acessar um recurso que não existe, por exemplo: https://www.alura.com.br/nao-existe. Bom, a Alura mostra uma imagem que indica que não achou o que procuramos, mas vamos dar uma olhada no console. Repare que o código agora é 404. No mundo HTTP 404 significa que o servidor não encontrou o recurso (Not Found).



Durante o desenvolvimento de uma aplicação web podem acontecer problemas no lado do servidor. Isto é normal pois alguma lógica pode falhar, erros acontecem no desenvolvimento! A notícia boa é que o protocolo HTTP vem preparado para isso. Quando algum problema no servidor acontecer, podemos avisar o cliente através do protocolo HTTP. O código mais comum para este tipo de problema é o 500 que significa: "deu pau no servidor :)". Talvez um ou outro aluno já tenha visto um erro 500 na Alura. Isso não deveria acontecer, mas onde há humanos também há problemas, não é mesmo?

# Categorias de códigos

Existem muitos códigos de resposta definidos no protocolo HTTP. Há tabelas disponíveis na web que mostram esses códigos, descrevendo o significado de cada um deles. No entanto, no dia a dia, o desenvolvedor não precisa decorar todos esses códigos. Basta consultar quando for necessário, por exemplo [aqui](http://www.w3schools.com/tags/ref_httpmessages.asp).

O importante é saber que algo que começa com **2xx** significa coisa boa, a requisição foi executada com sucesso. Quando recebemos algo como **3xx** normalmente significa que o navegador precisa fazer algo a mais (o cliente precisa agir) pois algo mudou ou um recurso não existe mais. **4xx** significa que o navegador enviou dados errados, como por exemplo uma URL errada. Caso o servidor gere algum problema, serão utilizados os códigos **5xx**.

No dia a dia os códigos 200, 404 e 500 são de longe os mais utilizados!

## Parâmetros na requisição com métodos GET e POST

Vamos ver um outro exemplo de uma URL e acessar [http://www.youtube.com](http://www.youtube.com/). Como tem muito conteúdo no YouTube, vamos pesquisar, por exemplo por Ayrton Senna. Repare que, ao pesquisar no YouTube, a URL mudou um pouco. O recurso acessado pela busca se chama **/results** (os resultados da pesquisa) mas agora temos um **parâmetro da requisição**, indicado pela **?**: **https://www.youtube.com/results?search\_query=Ayrton+Senna**

O parâmetro se chama **search\_query** com o valor **Ayrton+Senna**. Esses parâmetros da URL normalmente são chamados de **Query Params**. O HTTP permite enviar mais de um parâmetro, basta concatenar o próximo parâmetro através do caractere **&**.

Por exemplo, a busca avançada do Google usa vários parâmetros para refinar a pesquisa como o idioma, o país ou data. Veja como o Google concatena os Query Params: https://www.google.com.br/?gws\_rd=ssl#lr=lang\_pt&tbs=lr:lang\_1pt&q=neymar

## Parâmetros com GET

Reparem que no console sempre aparece o tipo (ou método) da requisição, sendo GET. Uma característica da requisição GET é enviar os parâmetros pela URL! Isso é útil quando queremos deixar os parâmetros visíveis. Assim podemos facilmente guardar a URL com os parâmetros para repetir a requisição algum momento depois. Mas será que isso também é uma boa opção na hora de enviar credenciais como login e senha? Queremos que apareça a senha de um usuário na URL?

Imagina que você efetue o login no seu banco e na URL apareça: https://www.bb.com.br/login?login=nico&password=supersecreto

Não parece ser uma boa solução, não é mesmo?

## O método HTTP POST

Vamos efetuar um login na Alura para ver como esse sistema envia dados do usuário para o servidor.



Repare que a URL para enviar o login e senha se chama https://www.alura.com.br/signin. Repare também que o método HTTP utilizado mudou. Estamos usando o **HTTP POST**! Usando o POST, o navegador envia os dados do formulário no corpo da requisição e não na URL! Se fosse um GET, todos os dados seriam enviados através da URL. Como a Alura não deseja que os dados do login e senha apareçam na URL do navegador, foi utilizado um HTTP POST. Faz sentido?

## GET para receber, POST para criar?

Um outro exemplo de um método POST na Alura é quando criamos uma pergunta no forum. Nesse momento o usuário submete um formulário com dados para criar um novo recurso na Alura (a pergunta do aluno se torna um recurso). O método POST foi inicialmente pensado para criar algo novo no servidor como acabamos de fazer. Ou seja, ao enviar uma requisição POST para o servidor, a nossa intenção é criar algo novo. No entanto, nem sempre isso é realmente utilizado dessa maneira. Por exemplo, acabamos de usar um POST para verificar o login, ou seja, não alteramos ou adicionamos nada na Alura. Nossa motivação para o POST era esconder os parâmetros apenas.

Como o servidor realmente reage quando recebe uma requisição POST depende da implementação, depende da lógica atrás. Os métodos como GET e POST definem uma intenção mas o que realmente será executado depende do servidor.

No dia a dia, vocês vão ver códigos usando GET para fazer pesquisas mas também para alterar ou remover algo no servidor. A mesma coisa podemos dizer sobre POST. Vocês vão usar o POST para inserir e alterar dados, mas também para pesquisar. As aplicações podem adaptar o significado dos métodos HTTP quando for necessário.

## Serviços Web-REST

á estudamos como o modelo de requisição-resposta funciona usando o browser para esse estudo. Mas será que toda a requisição HTTP sempre tem como origem um navegador? E toda resposta só possui conteúdo que ele entende: HTML, CSS, Javascript e imagens?

Nesse capítulo veremos como as aplicações conseguem se comunicar e responder as questões levantadas anteriormente.

##Serviços WEB

Hoje existem milhões de softwares rodando ou sendo desenvolvidos em várias linguagens de programação e frameworks. Tais softwares não vivem necessariamente isolados e podem querer se comunicar de alguma forma.

Um exemplo clássico é o login via rede social que estamos cada vez mais habituados. Essa conversa acaba sendo transparente para nós, usuários, já que exige uma autorização de acesso às nossas informações.

Em outros momentos as aplicações conversam sem que nada visual seja implementado ou mesmo uma autorização do cliente seja pedida.

As aplicações que disponibilizam serviços para outras são chamadas de webservices. E uma API de utilização é documentada para uma integração eficiente entre sistemas.

Temos serviços web para trabalhar com pagamentos(Paypal é um exemplo famoso), upload de imagens, transformação de CEP em endereços textuais e diversos outros. Tudo isso é feito através do poderoso protocolo **HTTP**.

##Cenário de trabalho

Você muito provavelmente já teve uma péssima experiência quando estava sem conexão com a internet usando um aplicativo móvel. Alguns apps não funcionam sem um acesso a rede porque as principais funcionalidades são feitas via requisições **HTTP**.

Essas requisições são implementadas programaticamente pelo desenvolvedor. Podemos implementá-las em várias linguagens de programação: Java, PHP, Javascript etc.

Para exemplificar vamos imaginar que temos aqui na Alura uma divisão responsável por uma aplicação de pedido de refeições que é a AluraFood.

A AluraFood tem duas equipes em ação: a do serviço web(ou simplesmente API web) e a dos apps mobile(Android e iOS).

Os desenvolvedores responsáveis pela tela de listagem de restaurantes vão precisar receber do serviço os detalhes de cada restaurante. Felizmente o pessoal responsável pelo webservice já documentou exatamente o que seria necessário:

Listagem de todos os restaurantes --> GET - http://alurafood.com/api/restaurante

Como retorno dessa requisição poderíamos receber o seguinte conteúdo HTML:

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="UTF-8">

<title>AluraFood</title>

<link rel="stylesheet" type="text/css" href="principal.css">

<script type="text/javascript" src="build.js"></script>

</head>

<body>

<table>

<tr>

<td>Bob's</td>

<td>8</td>

<td>R. do México 100</td>

<td><img src="http://alurafoods.com/uploads/logos/bobs.png"/></td>

</tr>

<tr>

<td>Subway</td>

<td>8.5</td>

<td>Av. Rio Branco 202</td>

<td><img src="http://alurafoods.com/uploads/logos/subway.png"/></td>

</tr>

<tr>

<td>Experimenta Lanches</td>

<td>9</td>

<td>R. do Brasil 545</td>

<td><img src="http://alurafoods.com/uploads/logos/e-lanches.png"/></td>

</tr>

</table>

</body>

</html>

Perceba que temos uma listagem de restaurante sendo apresentada dentro de uma tabela(elemento **table** do HTML) e cada linha(elemento **tr**) possui 4 colunas(**td**). Dentro de cada coluna temos as informações dos restaurantes: nome, nota de avaliação, endereço e logo.

Logo percebemos que os responsáveis precisarão realizar uma análise do conteúdo HTML e extrair dele somente as informações necessárias. Esse ato de analisar o documento é chamado de realizar um parsing do arquivo.

Perceba que o HTML tem muito mais do que o necessário para essa equipe, é um formato de verbosidade considerável: temos cabeçalhos e tags diversas. Para piorar estamos trafegando muito mais informações do que o necessário e onerando até mesmo a banda do nosso usuário. Péssimo cenário não é mesmo?

Pensando nessa deficiência do HTML temos outros formatos que fazem mais sentido quando uma representação de um recurso (um restaurante) se faz necessário. Temos como exemplo mais legível o **XML(eXtensible Markup Language)** que poderia ser devolvido como resposta e ter o seguinte conteúdo:

<?xml version="1.0"?>

<restaurantes>

<restaurante>

<nome>Bob's</nome>

<avaliacao>8</avaliacao>

<endereco>R. do México 100</endereco>

<logo>http://alurafoods.com/uploads/logos/bobs.png</logo>

</restaurante>

<restaurante>

<nome>Subway</nome>

<avaliacao>8.5</avaliacao>

<endereco>Av. Rio Branco 202</endereco>

<logo>http://alurafoods.com/uploads/logos/subway.png</logo>

</restaurante>

<restaurante>

<nome>Experimenta Lanches</nome>

<avaliacao>9</avaliacao>

<endereco>R. do Brasil 545</endereco>

<logo>http://alurafoods.com/uploads/logos/e-lanches.png</logo>

</restaurante>

</restaurantes>

Outro famoso formato e onerando menos ainda a rede, por ser mais leve, é o **JSON(JavaScript Object Notation)**:

[

{

"nome": "Bob's",

"avaliacao": "8",

"endereco": "R. do México 100",

"logo": "http://alurafoods.com/uploads/logos/bobs.png"

},

{

"nome": "Subway",

"avaliacao": "8.5",

"endereco": "Av. Rio Branco 202",

"logo": "http://alurafoods.com/uploads/logos/subway.png"

},

{

"nome": "Experimenta Lanches",

"avaliacao": "9",

"endereco": "R. do Brasil 545",

"logo": "http://alurafoods.com/uploads/logos/e-lanches.png"

}

]

Mas como especificar à aplicação de serviço que gostaríamos de receber em um formato JSON? Via cabeçalho **HTTP**!

Para indicar que queremos resposta no formato JSON usa-se um **Accept: application/json** como cabeçalho HTTP. Por outro lado já na resposta uma indicação desse conteúdo é especificado pelo cabeçalho **Content-Type: application/json**.

## O que é Rest

Tudo certo para a listagem de restaurantes. Mas será que o app AluraFood se resume a listar restaurantes? Provavelmente não, dado que o usuário efetua pedidos, um restaurante tem cardápio que poderia sofrer alterações e por aí vai.

Algumas funcionalidades específicas aos responsáveis de um restaurante podem ser necessárias. E para isso o webservice deveria estar preparado também para lidar com essa necessidade:

Listagem de todos os restaurantes --> GET - /restaurante

Adicionar um restaurante --> POST - /restaurante

Perceba que no exemplo fictício as duas primeiras URIs são idênticas e a funcionalidade muda completamente a partir do método HTTP usado:

GET -> Listagem

POST -> Criação

A atualização poderia ter um outro endpoint como por exemplo:

PUT/PATCH - /restaurante/1

O número 1 ao final da URI indica um identificador a um restaurante específico.

A remoção de um restaurante poderia seguir o mesmo modelo:

DELETE - /restaurante/1

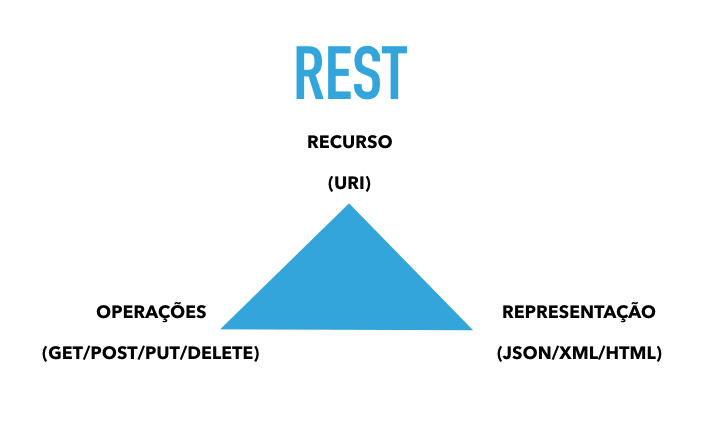
Para a busca do cardápio de um restaurante específico o endpoint gerado poderia seguir o modelo:

GET - /restaurante/1/cardapio

## O padrão REST

Logo podemos perceber que o padrão usado pela equipe do webservice define que uma requisição web tem três tipos de componentes importantes: recursos (URI), operações (GET, POST, PUT, DELETE/...) e representação de dados(XML, JSON, ...).

Esses três componentes em conjuntos seguindo algumas práticas são a base para o modelo arquitetural **REST(Representational State Transfer)** ou em português **Transferência de Estado Representacional**.



## Recurso

Ao criar as URIs do nosso sistema devemos levar em conta que elas representam recursos, não ações.

Em sistemas REST, nossas URIs devem conter apenas substantivos, que são nossos recursos: /restaurante/adiciona não é uma boa URI, pois contém um verbo e não está identificando um recurso, mas sim uma operação.

Para representar a adição de um restaurante podemos usar a URI /restaurante com o método HTTP POST, que representa que estamos adicionando alguma informação no sistema.

## Operações

O protocolo HTTP possui operações através de métodos como: GET, POST, PUT e DELETE.

Cada método tem uma semântica diferente e juntando o método à URI deveríamos conseguir representar todas as ações do nosso sistema.

As semânticas principais são:

* GET - recupera informações sobre o recurso identificado pela URI. Ex: listar restaurante, visualizar o restaurante 1. Uma requisição GET não deve modificar nenhum recurso do seu sistema, ou seja, não deve ter nenhum efeito colateral, você apenas recupera informações do sistema.
* POST - adiciona informações usando o recurso da URI passada. Ex: adicionar um restaurante. Pode adicionar informações a um recurso ou criar um novo recurso.
* PUT - adiciona (ou modifica) um recurso na URI passada. Ex: atualizar um restaurante.
* DELETE - remove o recurso representado pela URI passada. Ex: remover um restaurante.

## Representação

Quando fazemos uma aplicação não trafegamos um recurso pela rede, apenas uma representação dele. E essa representação pode ser feita de diferentes formas como JSON, XML ou HTML.

## Conclusão

Nossas URIs devem representar recursos, as operações no recurso devem ser indicadas pelos métodos HTTP e podemos falar qual é o formato em que conversamos com o servidor com o Content-Type e Accept que são cabeçalhos do HTTP.

## HTTP2

Até agora sempre usamos o browser para realizar uma requisição. Mas podemos realizar fora dele usando a linha de comando por exemplo. Um programa famoso para isso é o **CURL**. No Linux e MacOS ele já vem instalado por padrão.

Caso esteja usando o Windows é necessário a instalação dele. O download deve ser feito por aqui: <https://curl.haxx.se/download.html>

Para realizar e depurar uma requisição via CURL podemos simplesmente executar no terminal o seguinte comando:

curl -v www.caelum.com.br

Uma saída típica dele seria:

Fabios-MacBook-Pro:~ fabiopimentel$ curl -v www.caelum.com.br

\* Rebuilt URL to: www.caelum.com.br/

\* Trying 172.217.29.51...

\* Connected to www.caelum.com.br (172.217.29.51) port 80 (#0)

> GET / HTTP/1.1

> Host: www.caelum.com.br

> User-Agent: curl/7.49.1

> Accept: \*/\*

>

< HTTP/1.1 200 OK

< Content-Type: text/html; charset=utf-8

< Vary: Accept-Encoding,User-Agent

< Content-Language: pt-br

< Content-Type: text/html;charset=UTF-8

< X-DNS-Prefetch-Control: on

< X-Cloud-Trace-Context: 3e5e270ee3ab1e79f81b10d2cdef53cd

< Date: Fri, 24 Mar 2017 19:20:12 GMT

< Server: Google Frontend

< Content-Length: 95776

<

<!DOCTYPE html>

<html class="no-js"lang="pt-br"> <head> <title>Caelum | Cursos de Java, .NET, Android, PHP, Scrum, HTML, CSS e JavaScript </title>

…

Pode-se notar pela saída que temos logo no começo as informações do request efetuado:

> GET / HTTP/1.1

> Host: www.caelum.com.br

> User-Agent: curl/7.49.1

> Accept: \*/\*

`

E após essas infos temos o cabeçalho da resposta obtida pelo servidor:

< HTTP/1.1 200 OK

< Content-Type: text/html; charset=utf-8

< Vary: Accept-Encoding,User-Agent

< Content-Language: pt-br

< Content-Type: text/html;charset=UTF-8

< X-DNS-Prefetch-Control: on

< X-Cloud-Trace-Context: 3e5e270ee3ab1e79f81b10d2cdef53cd

< Date: Fri, 24 Mar 2017 19:20:12 GMT

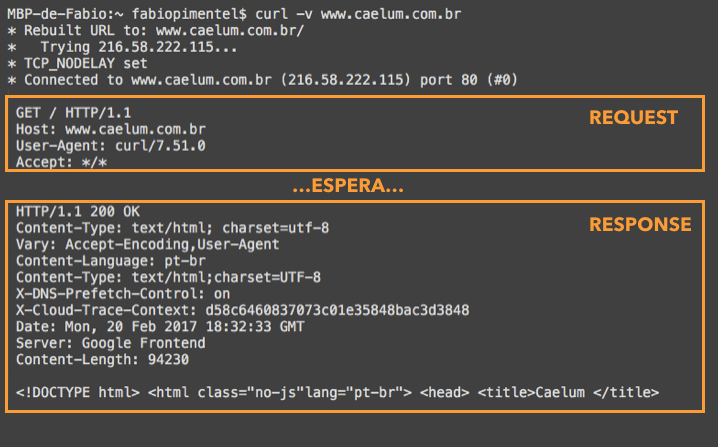
< Server: Google Frontend

< Content-Length: 95776

Logo depois vem o corpo da resposta (HTML da página requisitada):

<!DOCTYPE html> <html class="no-js"lang="pt-br"> <head> <title>Caelum | Cursos de Java, .NET, Android, PHP, Scrum, HTML, CSS e JavaScript </title> <meta name="viewport"content="width=device-width,initial-scale=1"> <meta name="format-detection"content="telephone=no"> <meta name="referrer"content="origin"> <meta name="description"content="A Caelum tem os cursos de Java, Scrum, Web, Front-end, PHP, .NET e Mobile mais reconhecidos no mercado, com didática diferenciada e instrutores qualificados."> <link rel="canonical"href="https://www.caelum.com.br/"> <style>.calendario .sem-turmas,.calendario-compacto .mais-turmas,.fm-message.fm-warning{font-style:italic}

Em resumo o output apresentando pelo CURL possui essa divisão:



# HTTP/2

O protocolo que estamos trabalhando até agora foi especificado na década de 90 e de lá até hoje muitas alterações foram feitas até na forma como usamos a internet.

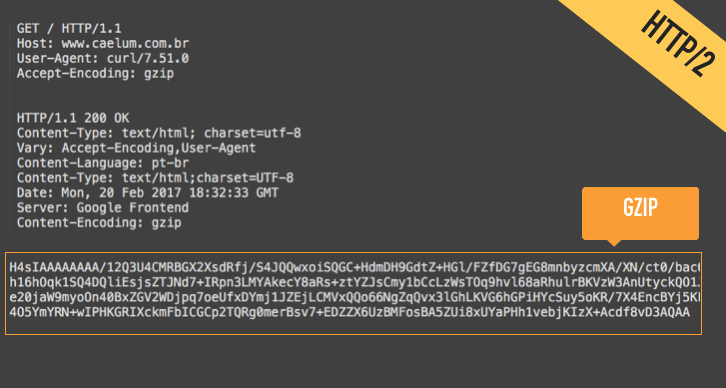
Com a chegada do mundo mobile novas preocupações apareceram e otimizações são cada vez mais necessárias para uma boa performance. Por isso uma mudança foi necessária e em 2015 depois de alguns anos de especificações e reuniões surgiu a versão 2 desse protocolo.

A nova versão é batizada de **HTTP/2** e tem como página principal de documentação e referência essa: https://http2.github.io/.

A nova versão do protocolo HTTP traz mudanças fundamentais para a Web. Recursos fantásticos que vão melhorar muito a performance da Web além de simplificar a vida dos desenvolvedores.

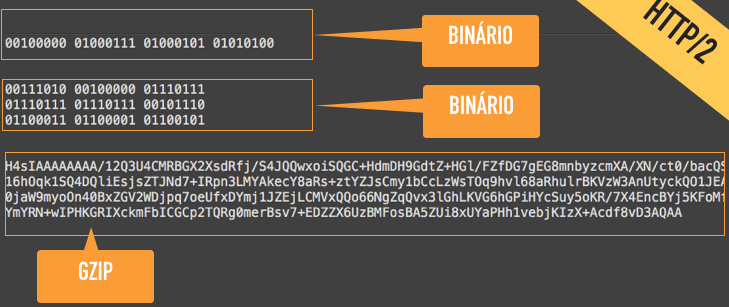
No HTTP 1.1, para melhorar a performance, habilitamos o **GZIP** no servidor para comprimir os dados das respostas. É uma excelente prática, mas que precisa ser habilitada explicitamente. No HTTP/2, o **GZIP é padrão e obrigatório**.

É como se a gente passasse a ter a resposta assim:

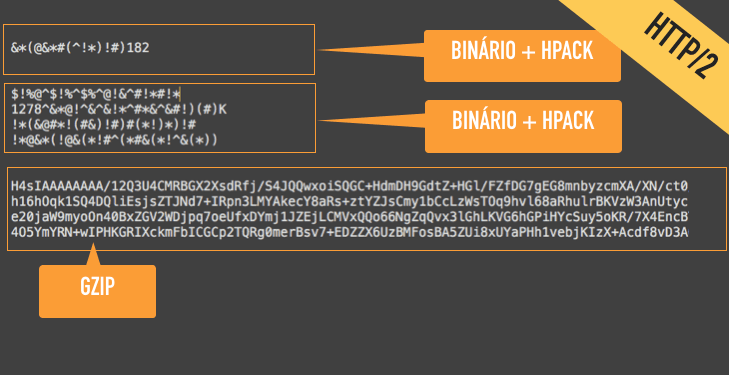


Mas, se você já olhou como funciona uma requisição HTTP, vai notar que só GZIPar as respostas resolve só metade do problema. Tanto o request quanto o response levam vários cabeçalhos (headers) que não são comprimidos no HTTP 1.1 e ainda viajam em texto puro.

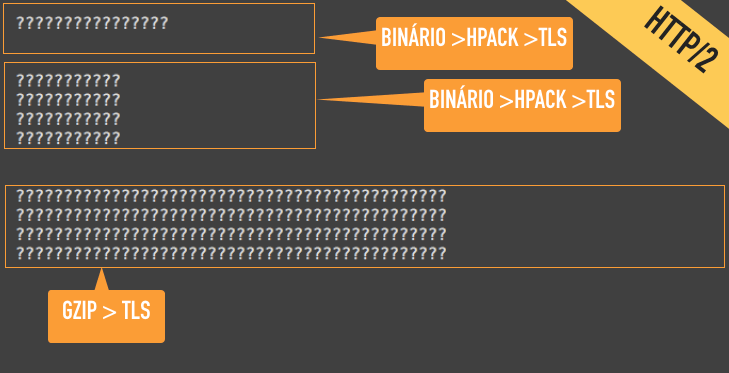
Já na nova versão, os headers passam a ser binários:



Além de binários eles são comprimidos usando um algoritmo chamado **HPACK**. Isso diminui bastante o volume de dados trafegados nos headers.



Além de todas essas otimizações para melhorar a performance ainda houve uma preocupação com segurança exigindo TLS por padrão também.



## HTTP2 – Cabeçalhos Stateful

Agora, queremos representar uma requisição. No código abaixo, estamos fazendo uma requisição através do método GET, que já conhecemos. Essa requisição está sendo feita para a raiz, bem parecido com o que fizemos no CURL no vídeo anterior.

GET /

Host: www.caelum.com.br

User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10.12; rv:34.0)

Accept: text/html,application/xhtml+xml,application/xml;q=0.9,\*/\*;q=0.8

Accept-Language: pt-BR,pt;q=0.8,en-US;q=0.5,en;q=0.3

Accept-Encoding: gzip, deflate

Quando realizamos uma requisição para essa URL via Firefox, por exemplo, o navegador envia alguns cabeçalhos que são padrões. Por exemplo, no cabeçalho Host, temos o domínio para onde estamos realizando essa requisição, que é www.caelum.com.br. Como estamos realizando um GET para /, o path para onde estamos realizando a requisição é www.caelum.com.br/.

Além disso, vemos uma informação de User-Agent, que é basicamente a fonte dessa requisição, neste caso é o navegador, Mozilla. Quando realizarmos uma requisição pelo CURL, aparecerá CURL, e se for num Safari, Chrome, qualquer outro navegador, irá aparecer a informação de onde estamos realizando de fato a requisição. Ou seja, nesse cabeçalho a gente especifica quem é o usuário.

Nele também é dito que é aceito, por padrão, o HTML, na linguagem tanto português quanto inglês, e que estamos aceitando uma codificação GZIP, já que no HTTP1 podemos especificar que tipo de compressão nossa requisição está aceitando.

## Precisamos repetir os cabeçalhos enviados em uma requisição anterior?

Agora vamos realizar uma outra requisição, mas dessa vez para o arquivo **principal.js**. Então, quando a requisição para página principal foi feita, provavelmente recebemos um HTML, e desse HTML foi necessário realizar uma outra requisição, porque era um recurso importante para a página ser exibida, como por exemplo um arquivo JavaScript.

GET /principal.js

Host: www.caelum.com.br

User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10.12; rv:34.0)

Accept: text/html,application/xhtml+xml,application/xml;q=0.9,\*/\*;q=0.8

Accept-Language: pt-BR,pt;q=0.8,en-US;q=0.5,en;q=0.3

Accept-Encoding: gzip, deflate

Então, na próxima requisição teremos que enviar novos parâmetros, novos dados. Então, mais uma vez, é uma nova requisição, agora para /principal.js, um recurso do nosso sistema. Qual o Host? **www.caelum.com.br**. Qual o Agent? Mozilla, versão 5.0. O que é aceito? qual linguagem é aceita? que tipo de compressão estamos aceitando?

Podemos perceber que o que colocamos nessa segunda requisição é exatamente o mesmo que fizemos na primeira? Os mesmos dados estão sendo trafegados nas duas requisições. Seria ótimo se só trafegássemos isso uma única vez, pois se estamos enviando mais dados, oneramos ainda mais nosso usuário, usando mais banda, deixando essa requisição mais lenta.

Seria muito bom se só pudéssemos ou só tivéssemos que enviar uma única vez, e é exatamente isso que o HTTP2 faz. A partir do HTTP2, não precisamos mais repetir os **Headers**, os cabeçalhos que já enviamos em uma requisição anterior. Logo, quando fazemos uma requisição para o **principal.js**, onde teríamos os cabeçalhos exatamente iguais aos da requisição passada, nós não precisamos enviar novamente esses dados.

## Cabeçalhos diferentes

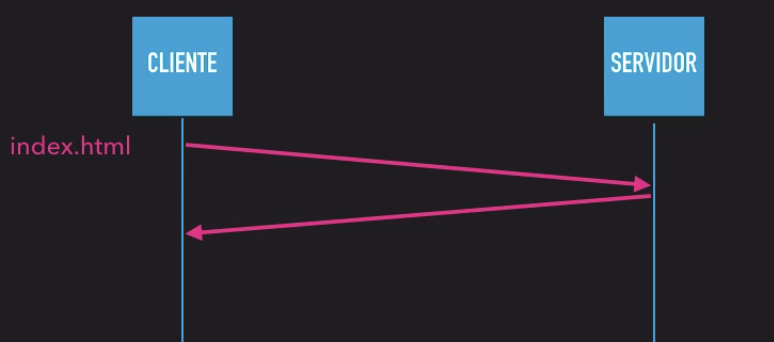
Agora, se temos uma imagem, os cabeçalhos podem mudar, por exemplo, o Host, que pode estar especificado na página principal. Logo, na primeira requisição, o conteúdo HTML especificou que tem que buscar uma imagem do Host, que é **image.caelum.com.br**, um subdomínio dentro da nossa aplicação. Então, esse cabeçalho terá que ser alterado, logo enviaremos apenas os cabeçalhos que são diferentes.

Isso está especificado no HTTP2, para que uma requisição fique mais leve e não onere tanto o usuário. Isso é conhecido como **Headers Stateful**.

No início do curso, comentamos que o HTTP era stateless, ou seja, ele não guarda informações das requisições passadas. E isso continua valendo, mas no caso dos cabeçalhos, existe um ambiente que guarda estado.

## HTTP2 – Server Push

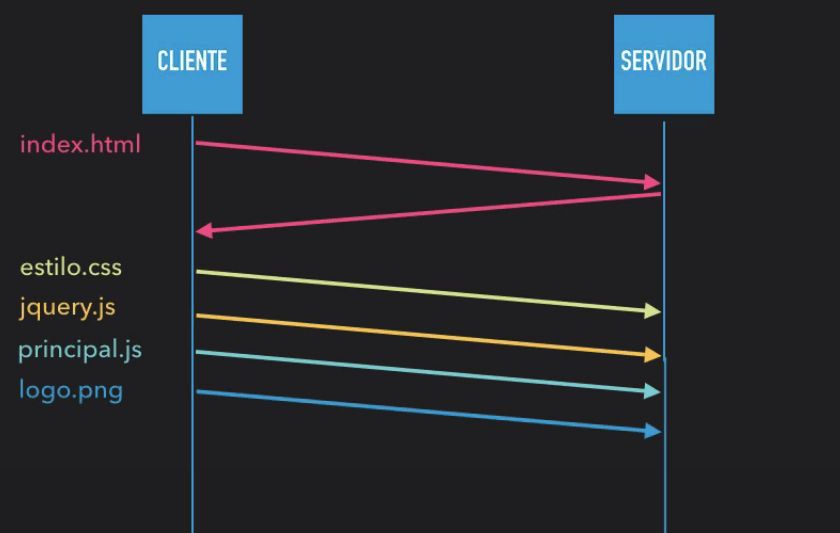
Temos o cliente e um servidor sendo representados. Podemos imaginar que estamos fazendo uma requisição para uma página principal, a **index.html**. Essa requisição bate no servidor e o servidor nos traz o conteúdo HTML.





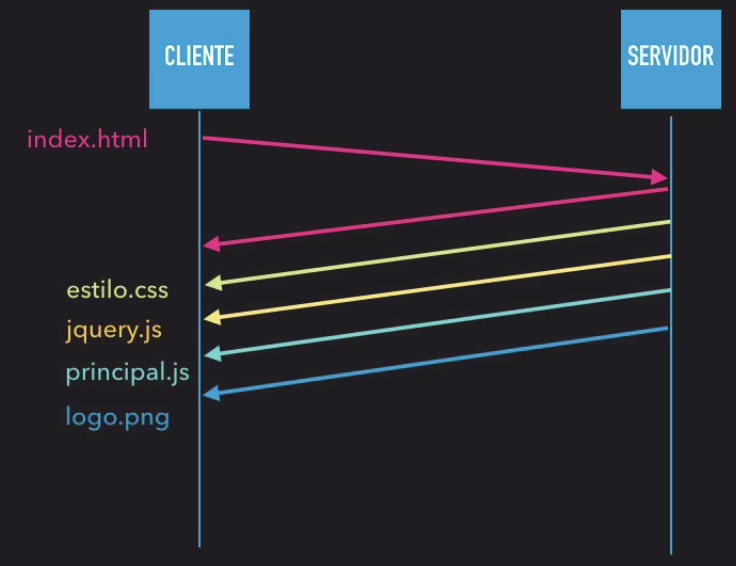
O HTML retornado pode ter o título **Caelum**, e então vai aparecer no nosso navegador essa informação. Além disso, temos um arquivo CSS, de estilização da página, que é o **estilo.css**, e dois arquivos JavaScript necessários para a página ser executada, o **jquery.js** e o **principal.js**. Além disso, no meio do corpo do HTML, tem um recurso que é de imagem, temos a imagem **logo.png**. Mas além desses, podemos ter vários outros recursos na nossa página.

Então, ao receber esse conteúdo, o browser tem que sair fazendo requisições de tudo o que é necessário para que ele renderize a página. O navegador interpreta esse conteúdo HTML de cima pra baixo, verifica que o primeiro recurso necessário é o **estilo.css**, aí ele vai lá buscar. O segundo recurso necessário, **jquery.js**, que é uma biblioteca JavaScript. E além disso, precisamos do principal.js e do logo.png:



Todos esses recursos especificados no HTML são novas requisições que o browser precisa fazer, e nosso cliente precisa executar. O servidor vai recebendo essas requisições, mas o cliente fica ali esperando até que essas respostas cheguem e o nosso browser consiga de fato renderizar o conteúdo para o usuário. Então há uma espera até essas respostas chegarem de fato, pois o servidor devolve as respostas das requisições na mesma sequência que foram geradas.

A partir do HTTP2, isso ficou um pouco diferente. Agora temos uma conversa mais paralela. Anteriormente estávamos trabalhando com conceitos de requisições seriais, fazíamos uma requisição e esperávamos receber, fazíamos outra requisição e esperávamos receber e por aí vai. No HTTP2, quando o cliente realiza uma requisição para \*index.html, o servidor devolve a página, mas ele já pode passar para o browser as informações necessárias para que essa página possa ser, de fato, exibida. Ou seja, ele consegue dar um passo além:



Isso é uma outra abordagem que surgiu no HTTP2, muito mais interessante. Mas quando o browser for interpretar essa página HTML, vai ter que passar pelo conteúdo que especifica o arquivo CSS? Sim, mas quando ele passar pelo **estilo.css**, vai verificar que já recebeu. Ou seja, ele percebe que já recebeu essas informações.

Este é o conceito de **Server Push**, ou seja, o **server** envia dados para o cliente sem que o cliente tenha solicitado, tornando o tráfego de dados muito mais otimizado.

## HTTP2 – Multiplexing

Outra coisa importante de requisição é que temos o conceito de **request** e **response**. Cada requisição e cada resposta no HTTP1.1 são únicos.

“Por baixo dos panos”, antes dessa requisição de fato ser feita, há uma conexão, comunicação entre cliente e servidor, que chamamos de **TCP**. Para que consigamos realizar uma requisição via HTTP, antes existe um modelo de TCP, que é um protocolo de transporte. Isso é mais a nível de infraestrutura, pois quando trabalhamos com desenvolvimento, acabamos deixando isso pra lá, já que ficamos na camada acima dessa conexão.

Queremos mostrar é que quando fazemos uma requisição, ela é única. No HTTP, cada requisição deveria abrir uma conexão TCP, executar e fechar.

Mas isso seria muito ruim porque conexão TCP é recurso caro, é um recurso que demora a ser alocado. Claro que é muito rápido a nível computacional, mas é mais um passo antes da requisição HTTP prosseguir e recebermos uma resposta.

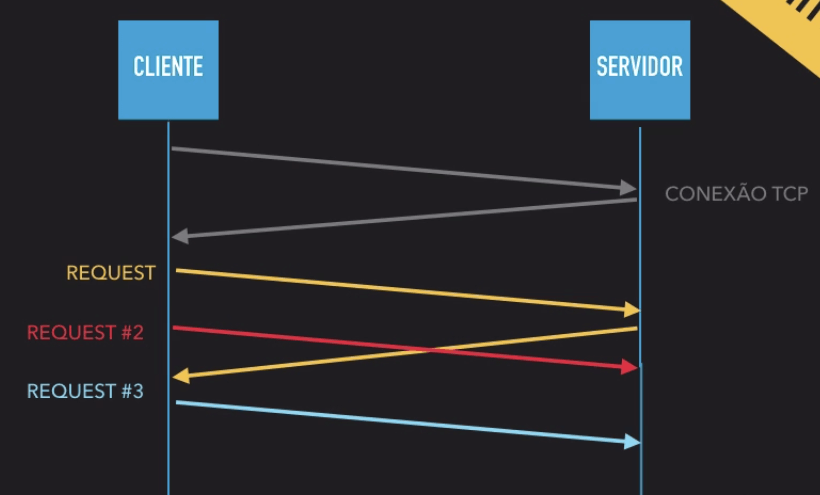
Então o que acontece, no HTTP1 existe um mecanismo chamado de **Keep-Alive**. O **Keep-Alive** determina quanto tempo, por exemplo, a nossa conexão pode ficar ativa. Ou seja, não encerra essa conexão TCP. Portanto, conseguimos realizar várias requisições com a mesma conexão TCP.

Hoje, na maioria dos browsers, temos um número entre 4 e 8 de conexões simultâneas por domínio. Significa que se fizermos uma requisição para a página da Caelum e a página da Caelum tiver mil recursos, o browser tem 4 a 8 conexões TCP ativas para conseguir realizar essas requisições em paralelo, e não serial. Mas isso na versão 1.1.

## Keep-Alive no HTTP2

O Keep-Alive continua existindo no HTTP2, só que ele trouxe uma novidade. Por exemplo, se temos uma conexão TCP aberta e realizamos uma requisição, poderíamos já dar prosseguimento às próximas requisições, isso em paralelo, sem de fato ficar esperando o resultado dela, de maneira assíncrona, e vamos recebendo essas respostas à medida em que o servidor for conseguindo processar.

Na imagem abaixo, fizemos a requisição 1 e requisição 2, quando íamos fazer requisição 3, já recebemos uma resposta:



Então, essas requisições e respostas vão chegando a todo tempo. É totalmente paralelo. A mesma coisa acontece com o servidor, não precisamos esperar uma resposta para enviar outra. Se já está pronta para ser enviada, ele já envia diretamente.

Esse conceito que surgiu no HTTP2 é chamado de **Multiplexing** e traz uma performance bastante relevante para o nosso HTTP.