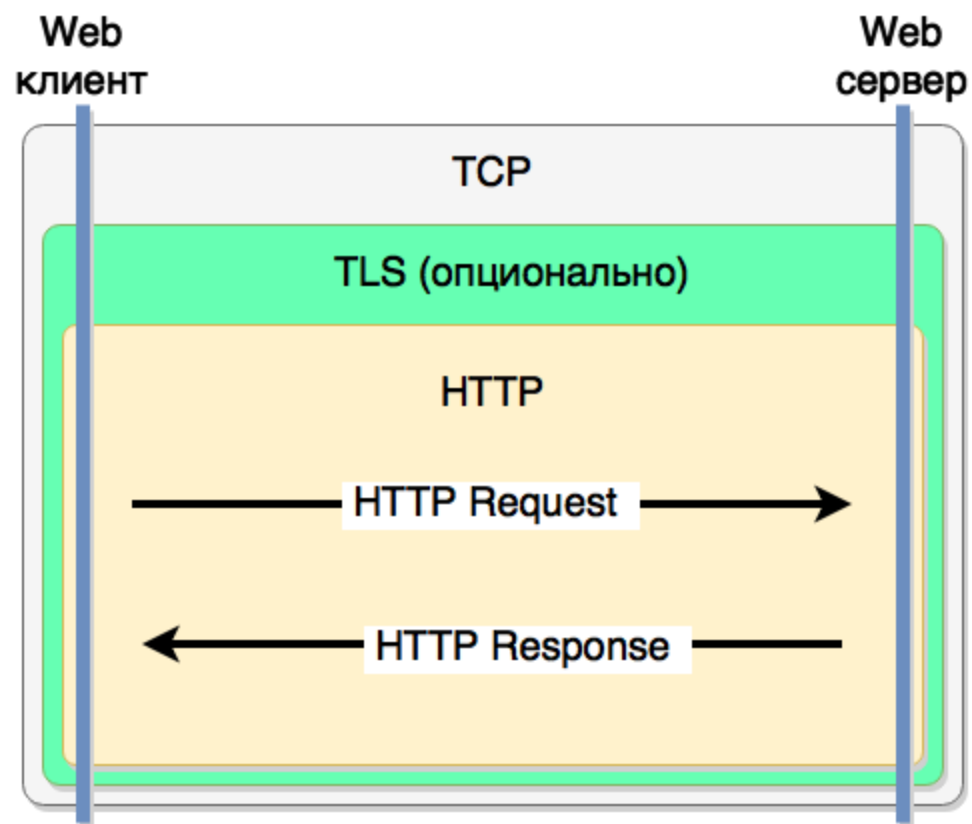


HTTP

Какие задачи решает HTTP?

- Передача документов
- Передача мета-информации
- Авторизация
- Поддержка сессий
- Кеширование документов
- Согласование содержимого (negotiation)
- Управление соединением



Ключевые особенности HTTP

- Работает поверх TCP/TLS
- Протокол запрос-ответ
- Не поддерживает состояние (соединение) - **stateless**
- **Текстовый** протокол
- Расширяемый протокол

HTTP/1.0 запрос

GET http://www.ru/robots.txt HTTP/1.0

Accept: text/html, text/plain

User-Agent: telnet/hands

If-Modified-Since: Fri, 24 Jul 2015 22:53:05 GMT

Перевод строки - \r\n

HTTP/1.1 запрос

GET /robots.txt HTTP/1.1

Accept: text/html,application/xhtml+xml

Accept-Encoding: gzip, deflate

Cache-Control: max-age=0

Connection: keep-alive

Host: www.ru

User-Agent: Mozilla/5.0 Gecko/20100101 Firefox/39.0

HTTP/1.1 OTBeT

HTTP/1.1 404 Not Found

Server: nginx/1.5.7

Date: Sat, 25 Jul 2015 09:58:17 GMT

Content-Type: text/html; charset=iso-8859-1

Connection: close

<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">

<HTML><HEAD>...

HTTP запрос состоит из

- строка запроса
 - метод
 - URL документа
 - версия
- заголовки
- тело запроса

HTTP методы

- **GET** - получение документа
- **HEAD** - получение только заголовков
- **POST** - отправка данных на сервер
- PUT - отправка документа на сервер (*)
- DELETE - удаление документа (*)
- CONNECT, TRACE, OPTIONS - используются редко (*)
- COPY, MOVE, MKCOL - расширения WebDAV (*)

HTTP ответ состоит из

- строка ответа
 - метод
 - URL документа
 - версия
- заголовки
- тела ответа - документ

HTTP коды ответа

- 1xx - информационные
- 2xx - успешное выполнение
- 3xx - перенаправления
- 4xx - ошибка на стороне клиента
- 5xx - ошибка на стороне сервера

HTTP коды ответа (1)

- 200 OK - запрос успешно выполнен
- 204 No Content - запрос успешно выполнен, но документ пуст
- 301 Moved Permanently - документ сменил URL
- 302 Found - повторить запрос по другому URL
- 304 Not Modified - документ не изменился, использовать кеш

HTTP коды ответа (2)

- 400 Bad Request - неправильный синтаксис запроса
- 401 Unauthorized - требуется авторизация
- 403 Forbidden - нет доступа (неверная авторизация)
- 404 Not Found - документ не найден
- 500 Internal Server Error - неожиданная ошибка сервера (application)
- 502 Bad Gateway - проксируемый сервер отвечает с ошибкой
- 504 Gateway Timeout - проксируемый сервер не отвечает

Заголовки HTTP (общие)

Для управления соединением и форматом сообщения (документа)

- Content-Type - MIME тип документа
- Content-Length - длина сообщения
- Content-Encoding - кодирование документа, например gzip-сжатие
- Transfer-Encoding - формат передачи, например, chunked
- Connection - управление соединением
- Upgrade - смена протокола

Заголовки HTTP запросов

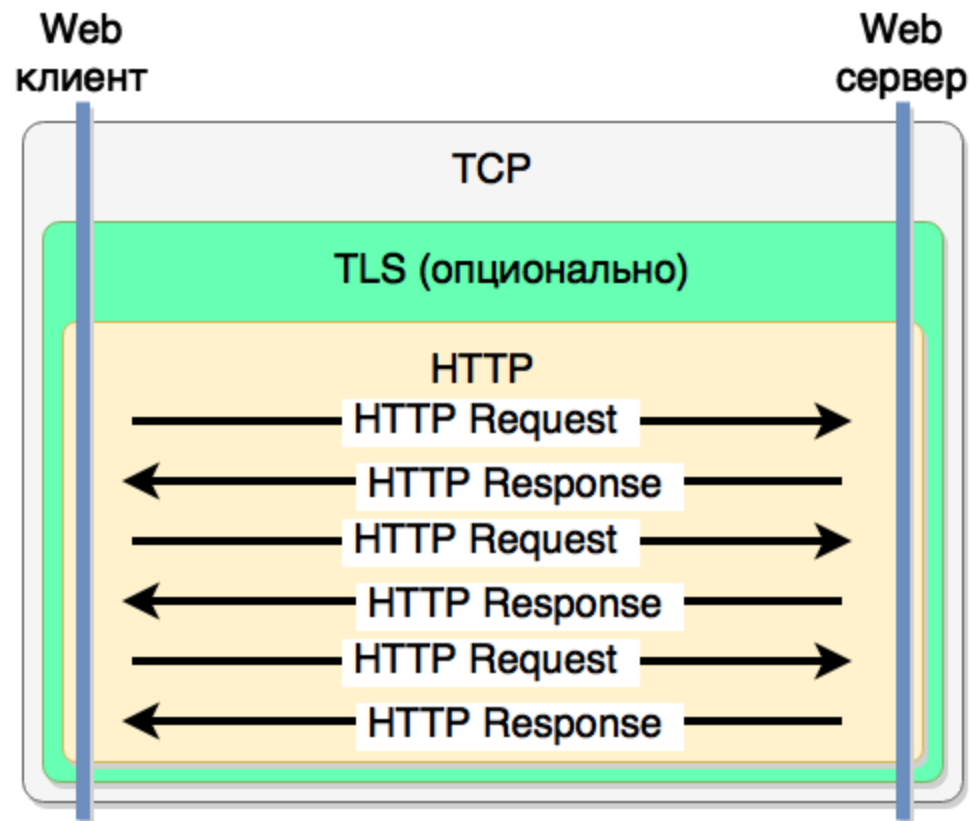
- Authorization - авторизация, чаще всего логин/пароль
- Cookie - передача состояния (сессии) на сервер
- Referer - URL предыдущего документа, контекст запроса
- User-Agent - описание web-клиента, версия браузера
- If-Modified-Since - условный GET запрос
- Accept-* - согласование (negotiation) содержимого

Заголовки HTTP ответов

- Location - новый URL документа при перенаправлениях
- Set-Cookie - установка состояния (сессии) в браузере
- Last-Modified - дата последнего изменения документа
- Date - Дата на сервере, для согласования кешей
- Server - описание web-сервера, название и версия

HTTP/1.1

управление соединением



Общие принципы

Протокол HTTP/1.0 предполагает закрытие TCP соединения сразу после ответа сервера.

Протокол HTTP/1.1 предполагает удержание TCP соединения, если не было заголовка `Connection: close`.

Логика управления в HTTP/1.1

Соединение должно быть закрыто, если:

- сервер или клиент использует HTTP младше 1.1
- сервер или клиент передал заголовок `Connection: close`
- по истечении таймаута (обычно небольшой, около 10с)

Иначе соединение остается открытым для последующих запросов.