

WWW основывается на использовании гипертекста(Бернерс-Ли, 1989г.)

Http(RFC 2616, 1990г.) .

Html – язык для описания документов и связывания их гиперссылками.

В качестве языка форматирования в html используются декларации SGML и метод определения типа документа DTD.

ASN-основа языка SMI.

DHTML-концепция с набором технологий, позволяющих браузеру динамически изменять загруженные html-документы в ответ на пользовательские действия без соединения с веб-сервером. Основа концепции: html, css, скриптовый язык, интерпрет. Браузером, dom.

DOM- объектная модель html-документа.

Java-апплет- прикладная программа на языке java в формате байт-кода, вып-ся jvm, которая не является интерпретатором байт-кода.

Flash-приложения – Action Script.

Методы HTTP: get, head, put, post, delete, link, unlink, доп. метод.

Множественный get-одинаковый результат, post –разный.

Запрос содержит: метод запроса, URI, версию протокола, сообщение в формате mime(упр. инф., инф. о клиенте, тело сообщения).

URI: URL, URN.

URL-идентификатор URI, который помимо идентификации ресурса предоставляет еще и информацию о ее местонахождении.

URN-идентификатор URI, который идентифицирует ресурс в определенном пространстве имен.

Шаблон идентификатора ресурса: <источник><путь><запрос><фрагмент>/

Для реализации механизма управления сессиями используются два доп. Поля http-заголовка:set-cookie и cookie.

Интерфейс CGI- стандарт сопряжения веб-серверов и программ, основная задача – определить, как передать программе параметры http-запроса и как передать обратно сформированный программой http-ответ.

Стандартный поток ввода процесса-обработчика ассоциируется с потоком вывода веб-сервера.

Станд. поток вывода – с потоком ввода.

Параметры запроса передаются процессу-обработчику через аргументы командной строки и через переменные окружения.

Необходимость подготовки и запуска нового процесса на каждый запрос – дорого.

Консорциум W3C: html, xml, css, http, url, ftp, nntp, sgml.

AJAX – асинхронное взаимодействие с веб-сервером. Уменьшение числа обращений к серверу. Код на JS отправляет xml-запрос, исп-ет dom для обновления. Основа-html(dhtml).

Две схемы аутентификации: Basic, Digest.

Механизм аутентификации реализует веб-сервер.

Basic:браузер делает запрос, сервер возвращает указание на необходимость аутентификации.

Didest:дает пользователю возможность доказать серверу, что он знает правильный пароль без передачи его по сети. MD5, счетчик запросов.

SSL обеспечивает три основных механизма защиты: взаимную аутентификацию, обеспечение конфиденциальности и целостности данных.

Сертификат – документ в электронной форме, содержащий открытый ключ, принадлежащий держателю, и доп. Информацию о его владельце.

Общий закрытый ключ, хеш-функция – в обмене сообщениями между сервером и клиентом.

Ограничения длины ключей:

- 1)Симметричное шифрование – 40 разрядов
- 2)Несимметричное шифрование – 512 разрядов

Прокси – программа, которая передает запросы клиентских программ в Интернет, получает ответы и передает их обратно.

Функции http-прокси:

- 1)Кэширование полученных ресурсов
- 2)Ограничение доступа к ресурсам интернета
- 3)Модификация запрашиваемых html – документов
- 4)Ограничение скорости работы с интернетом
- 5)Биллинг(Журналы, популярность)
- 6)Маршрутизация запросов.

WSDL-какие функции предоставляет веб-сервис

UDDI-поиск существующих веб-сервисов

Веб-сервис: является повторно используемым, определяется одним или несколькими явными интерфейсами, независимыми от технологии, может быть вызван посредством протоколов взаимодействия ресурсов между собой.

SOAP-протокол доступа к простым объектам

WSDL-язык описания веб-сервисов

UDDI- протокол поиска ресурсов в интернете.

UDDI: business entity, business service, binding template, technology model.

На стороне клиента

HTML4, HTML5

CSS

DOM

JavaScript, AJAX

path – путь (логический) к ресурсу

query – аргументы для динамического ресурса

fragment вообще не передается на сервер

•URL-кодирование

разрешенные символы: [A-Za-z0-9] и '.', '-', '~', '_'

пробел кодируется +

все остальные символы кодируются %xx, где xx – код символа в кодировке UTF-8

HTTP-запрос

метод, URL ресурса, версия протокола

заголовки

(опционально) тело запроса

•HTTP-ответ

версия, код и статус ответа

заголовки

тело ответа

•1xx – информационные

o100 Continue

•В запросе должно быть Expect: 100-continue

•2xx – успех

o200 Ok

•3xx – перенаправление

o301 Moved Permanently

o302 Found

o303 See Other

o304 Not Modified

o307 Temporary Redirect

4xx - ошибка со стороны клиента

o400 Bad Request

o401 Unauthorized

- В ответе должен быть проставлен WWW-Authenticate

o403 Forbidden

o404 Not Found

o405 Method Not Allowed

•5xx – ошибка со стороны сервера

o500 Internal Server Error

o501 Not Implemented

o502 Bad Gateway

o503 Service Unavailable

Заголовки HTTP-ответов

- Server

Content-Length (Content-Length: 78)

- Content-Type

- Location

- Set-Cookie

HTTP 1.0 и 1.1: основные отличия

- Поддержка TCP-соединений открытыми

- Передача тела по частям (chunked transfer encoding)

- Кеширование (ETag)

- Заголовок Host

- Аутентификация HTTP digest

- Cookies

Веб-приложение взаимодействует с внешними подсистемами: СУБД, LDAP, интерпретатор ОС, почтовый демон, браузер, файловая система и т.п.

Injection-атаки: внешний пользователь может выйти из контекста данных в контекст команд.

Примеры Injection-атак

- SQL-операторы

данные не обрабатываются при построении SQL-запроса

- HTML-разметку

данные не обрабатываются при построении HTML

- Javascript-операторы

данные не обрабатываются при построении JSON-объекта

- CSS-директивы

данные не обрабатываются при построении CSS-правил

- HTTP-заголовки

данные не обрабатываются при построении ответного заголовка (например, Location)

- XPath-операторы

данные не обрабатываются при построении XPath-запроса

Способы реализации сеансов

- По IP-адресу

- HTTP-аутентификация

- URL-overwriting

- Скрытые поля форм + PostBack

- Cookies

SOP на пальцах: код из одного источника не может читать или изменять DOM-объекты другого источника

- Источник = протокол + домен + порт

Same origin policy & JavaScript

- Window boundary
- Frame boundary
- Embedded scripts
- Images & CSS