



RESTful Web Services



Web Application

- NCSA (National Centre for Supercomputing Applications) create a way for programs to be executed inside a web server
 - The application could return non static dynamically generated hypertext documents!
 - Content no longer limited by what is uploaded to the web server
- This environment for executing applications inside a web server is call CGI (Common Gateway Interface)
- Early CGIs were Perl and Bash shell scripts



Web Architecture



- Web has a very simple architecture compared to CORBA, DCOM and other distributed architectures
- Dual purpose – serve files and application services
- Web servers are free
- Very easy to develop applications, no specialized tools were require
- HTTP is very easy to understand and debug
- All it requires is an architectural blue print which is to come in 2000



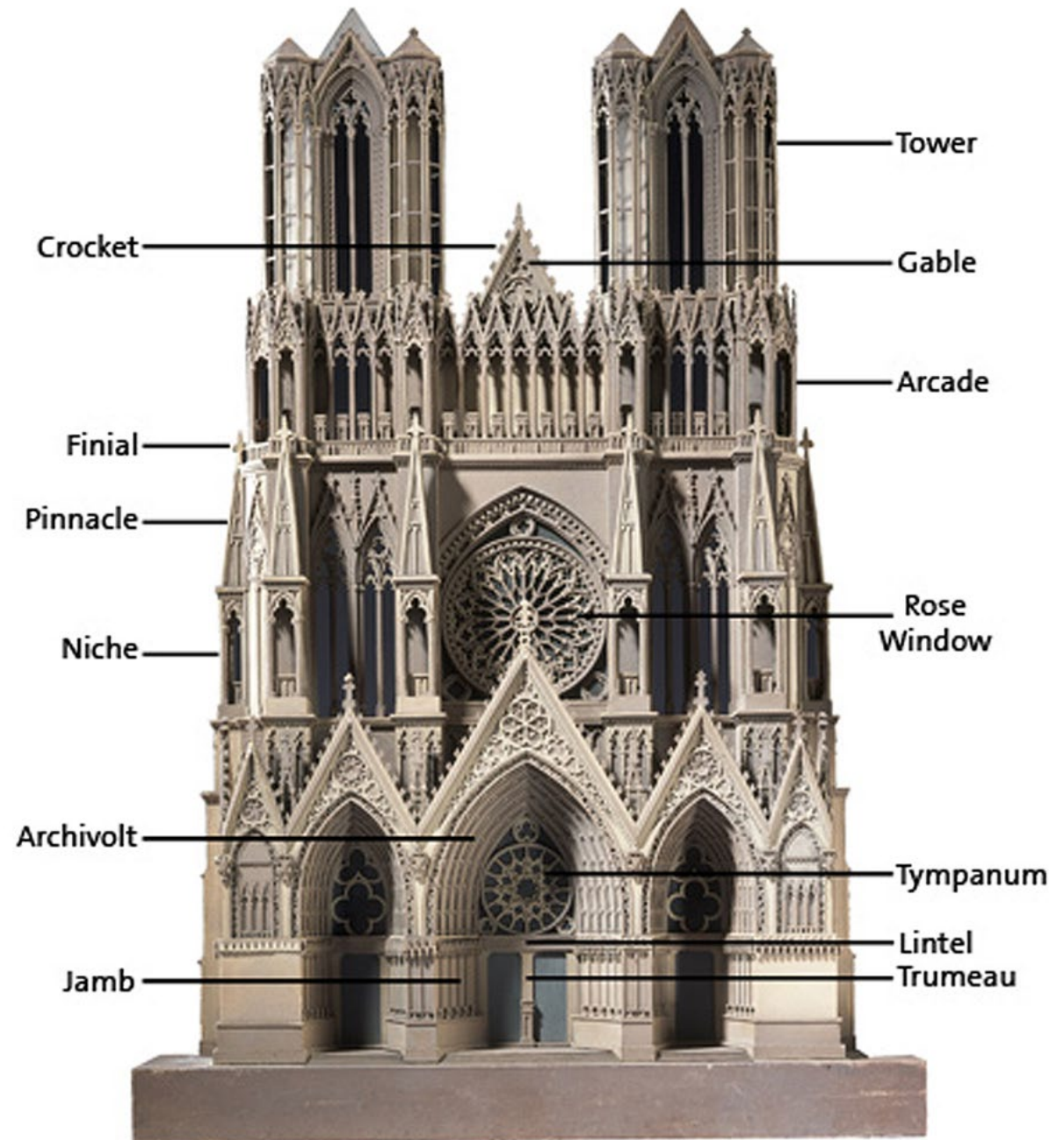
RESTful Web Services



- HTTP based application services allows a way for different computer systems to interoperate
 - The requestor does not have to be a browser
- Roy Fielding formalizes RESTful Web Service in his 2000 PhD thesis
 - Architectural Styles and the Design of Network-based Software Architectures
 - https://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf
- Coined the term **REST**
 - **RE**presentational **S**tate **T**ransfer

RESTful Web Services

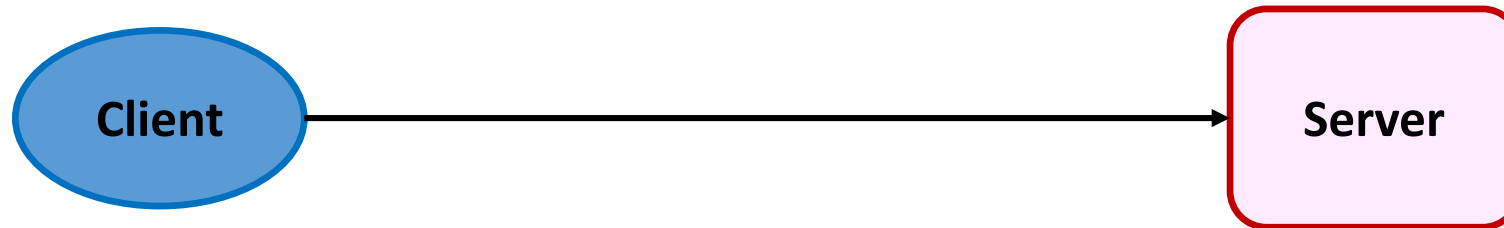
- Architectural style for distributed system
 - Building distributed systems based on a set of principles
- Architectural style consist of a set of constrains applied to elements within the architecture
- Jazz music
 - Music elements: key, notes, timing, phrasing, etc
 - Architectural constrains: syncopation, improvisation, etc.





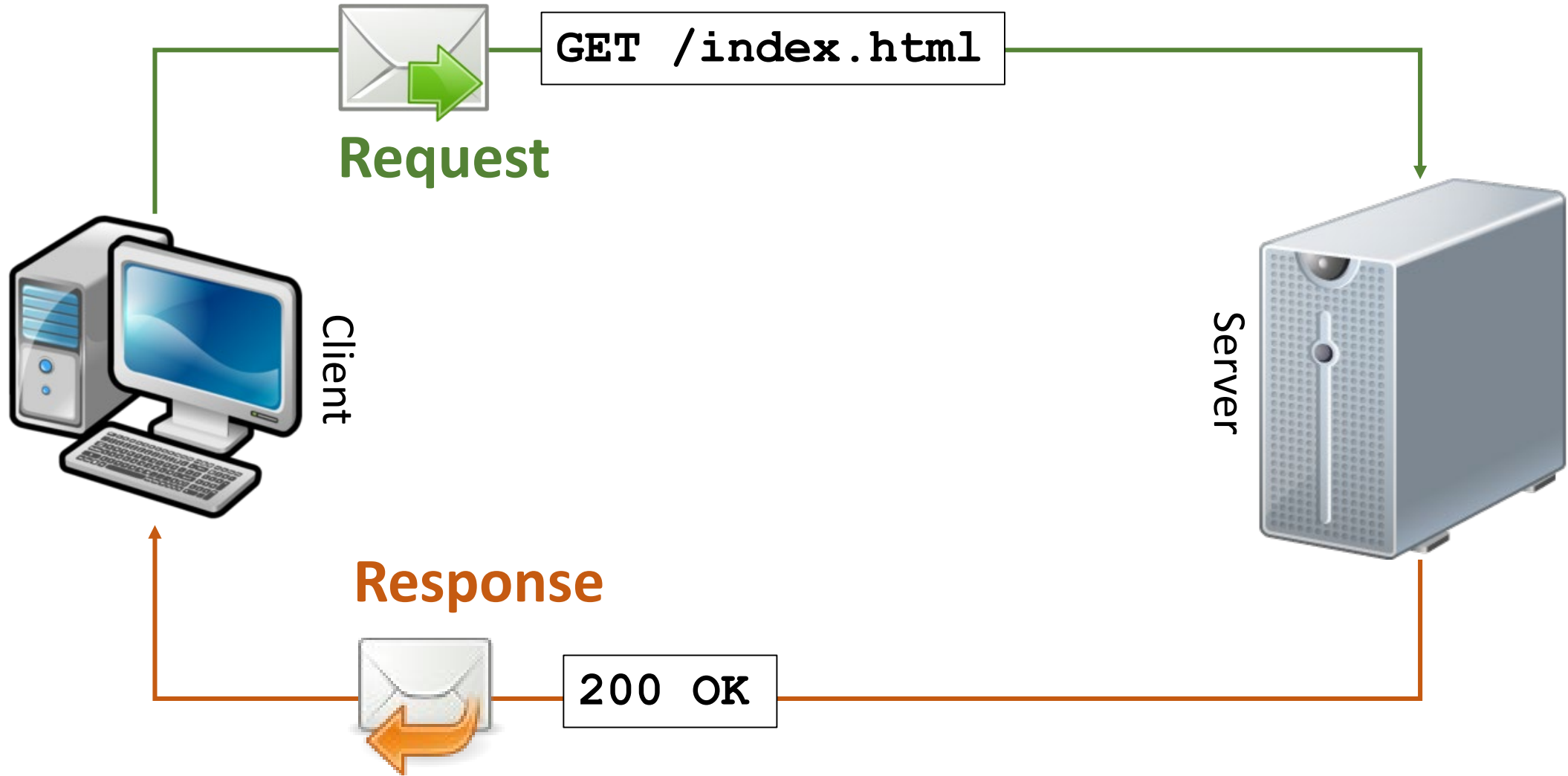
REST Constraint 1 - Client Server

- Allows for separation of concerns
- Separation of client specific concerns with server specific concerns
 - Eg UI, affordance, persistence, etc.





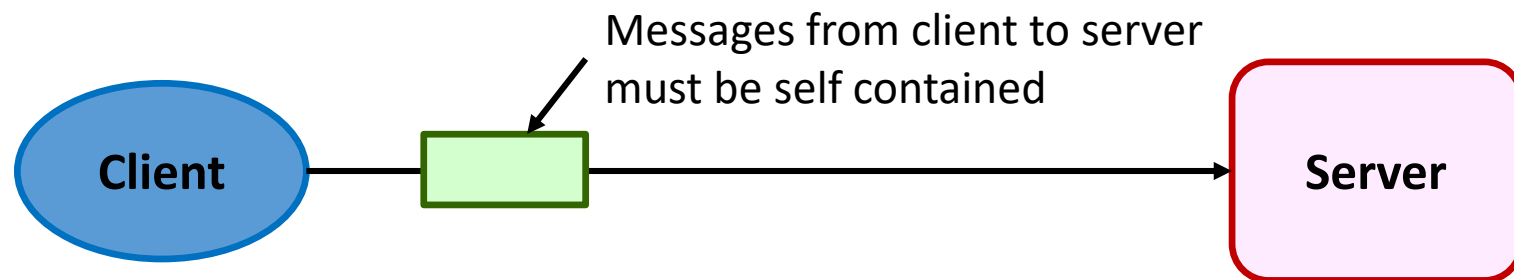
The Web is Client-Server





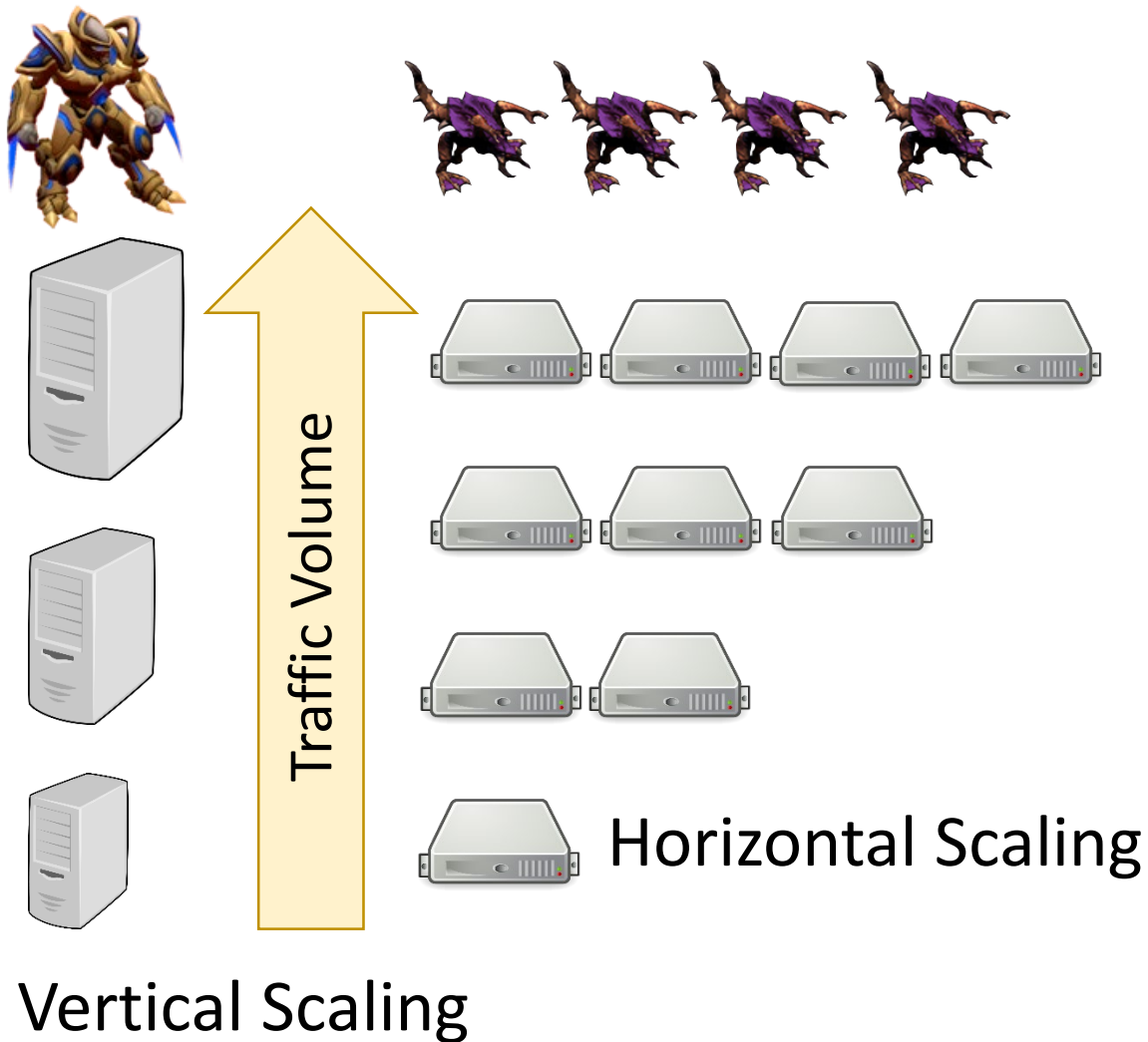
REST Constraint 2 - Stateless

- Communication between client and server must be stateless
 - Viz. Stateless server
- Each request from client to server must contain all information required to understand and to process that request





Statelessness for Scaling



- Scaling is the capability of the system to handle more workload by provisioning more resources
- Two types of scaling
 - Horizontal scaling - scales by provision more Pods
 - Applications must be stateless allowing the ingress controller to route the request to any Pod
 - Vertical scaling - scaling by giving the application more resources
 - Application must be able to utilize the extra resources eg. more vCPUs or memory

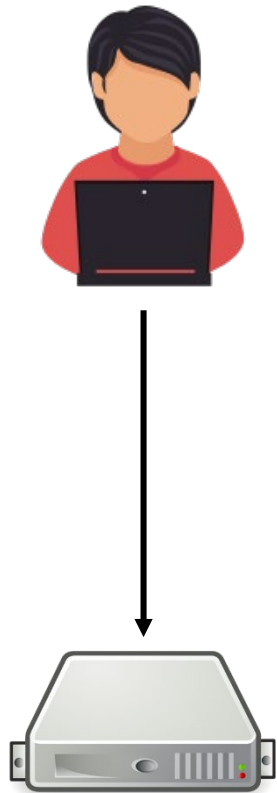


Why is Statelessness Important?

- Predicable - know how much work needs to be done by looking at the request
- Reliable - can easily recover from error because all information required to process a request is contained within that request
 - Can be processed by the next available server if the current one fails
- Scalable - not having to store states allows servers to be allocated when workload is high and scale down when workload is low

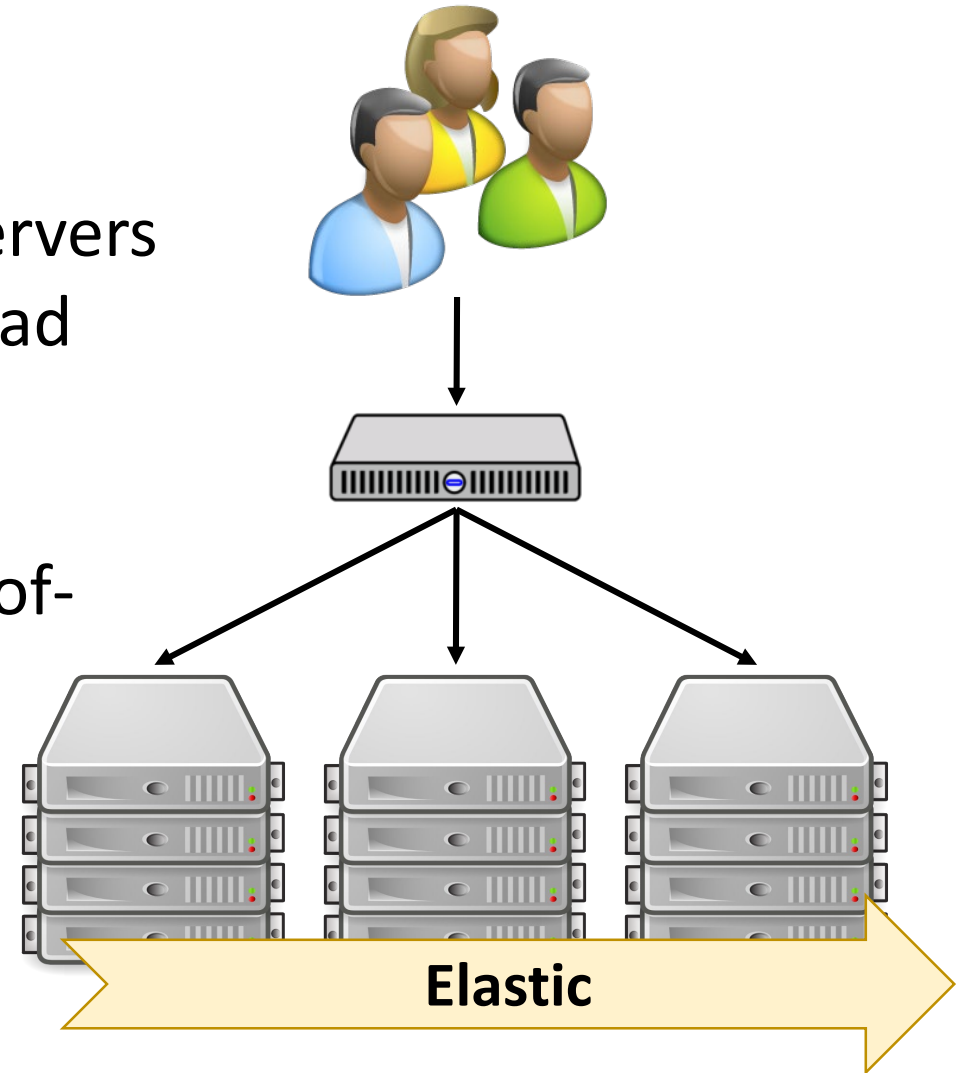


The Web is Stateless



Easily increase the number of servers
(scale out) to handle the workload
Makes the environment elastic

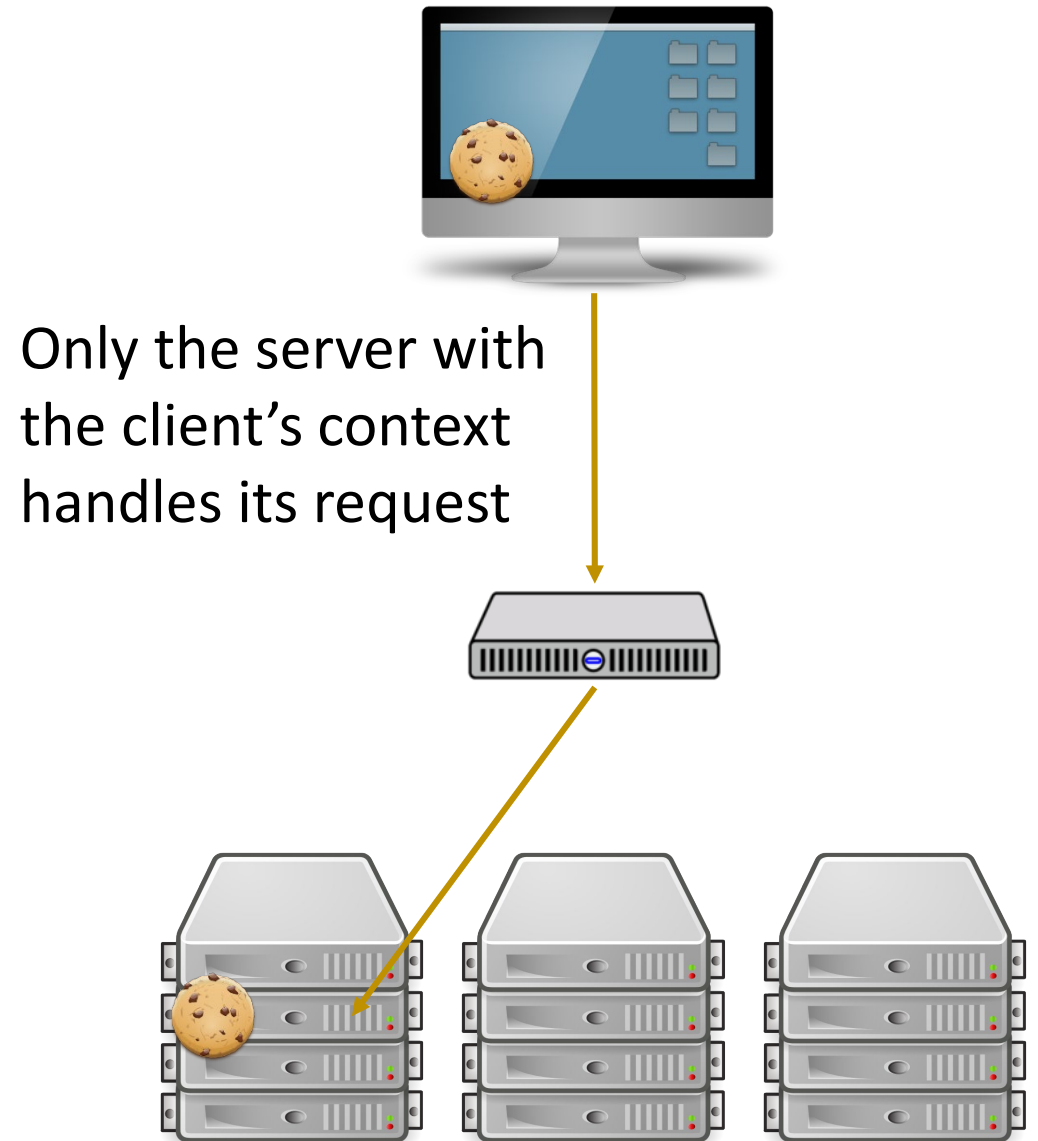
Scaling could be based on time-of-
day, resource utilization or by
anticipation





Stateful

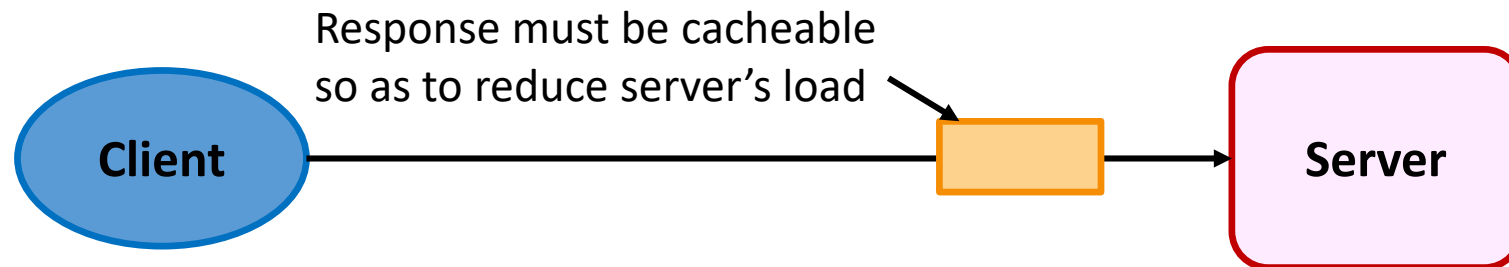
- Some process require state to work, viz. require to remember the context
 - Context - what has transpired between the application and the client
 - Eg. shopping session
- To uniquely identify a client and make sure that only the server that has the context of that client process its request
 - Eg. using cookies to associate with a server side session





REST Constraint 3 - Cache

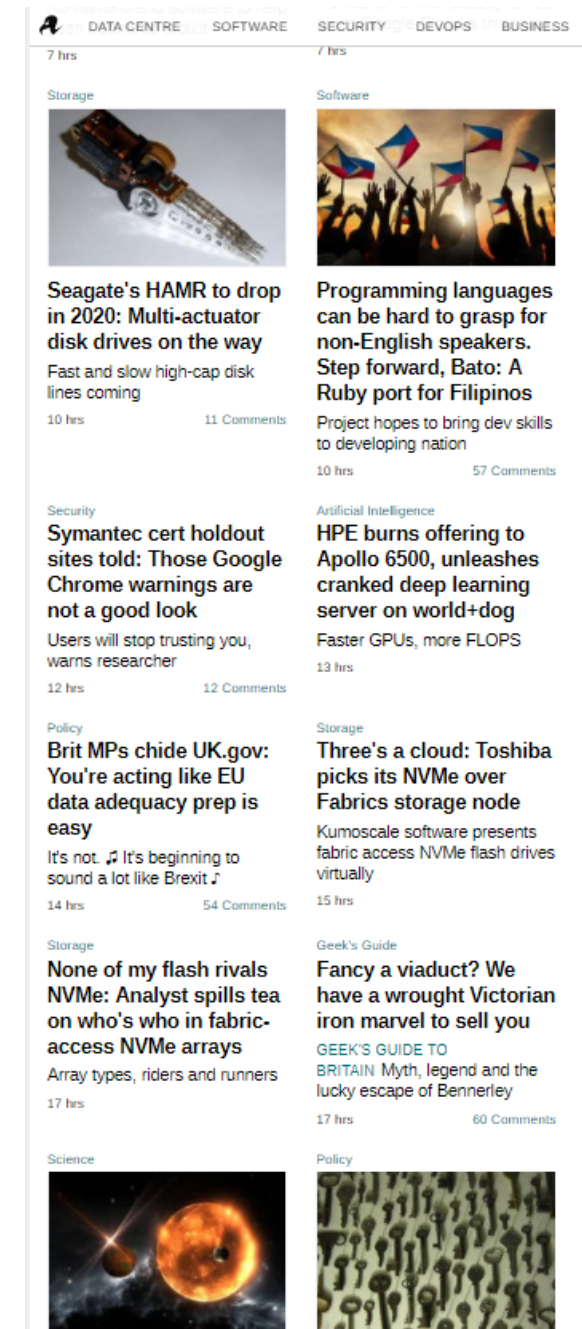
- Data from a response may be implicitly or explicitly labelled as cacheable or non cacheable
 - A cache has the right to return a cache data for an equivalent request
 - Avoid full trip back to the server





Cache

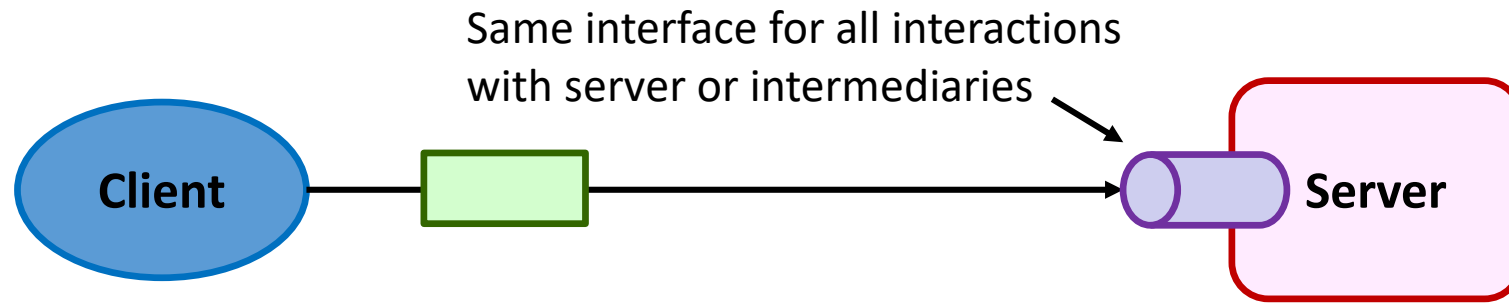
- Caching improves response time
 - Response can be returned by web caches (on the Internet)
 - Response can be returned by your browser (on your computer)
- What to cache?
 - Static resources - resources that are not updated frequently
 - Eg. images, pdf, audio, etc
 - Hard/expensive to create resources
 - Eg. news paper front page
 - Often used resources
- How to cache?
 - By time - the response will be stale in 30 minutes
 - By content - need to retrieve if the content changes





REST Constraint 4 - Uniform Interface

- Provide a general interface for client to invoke the server on all layers of the system



- A uniform interface has the following sub-constraints
 - Identify resources thru identifiers
 - Manipulation of resources through its representation
 - Self describing message
 - Hypermedia as the engine of application states



Uniform Interface - Resources Identification

Every resource can be uniquely identified

A resource name
within the server

GET /doc/test.html



Uniform Interface – Manipulate Resources thru its Representation

Representation



Canon in D

As music sheet

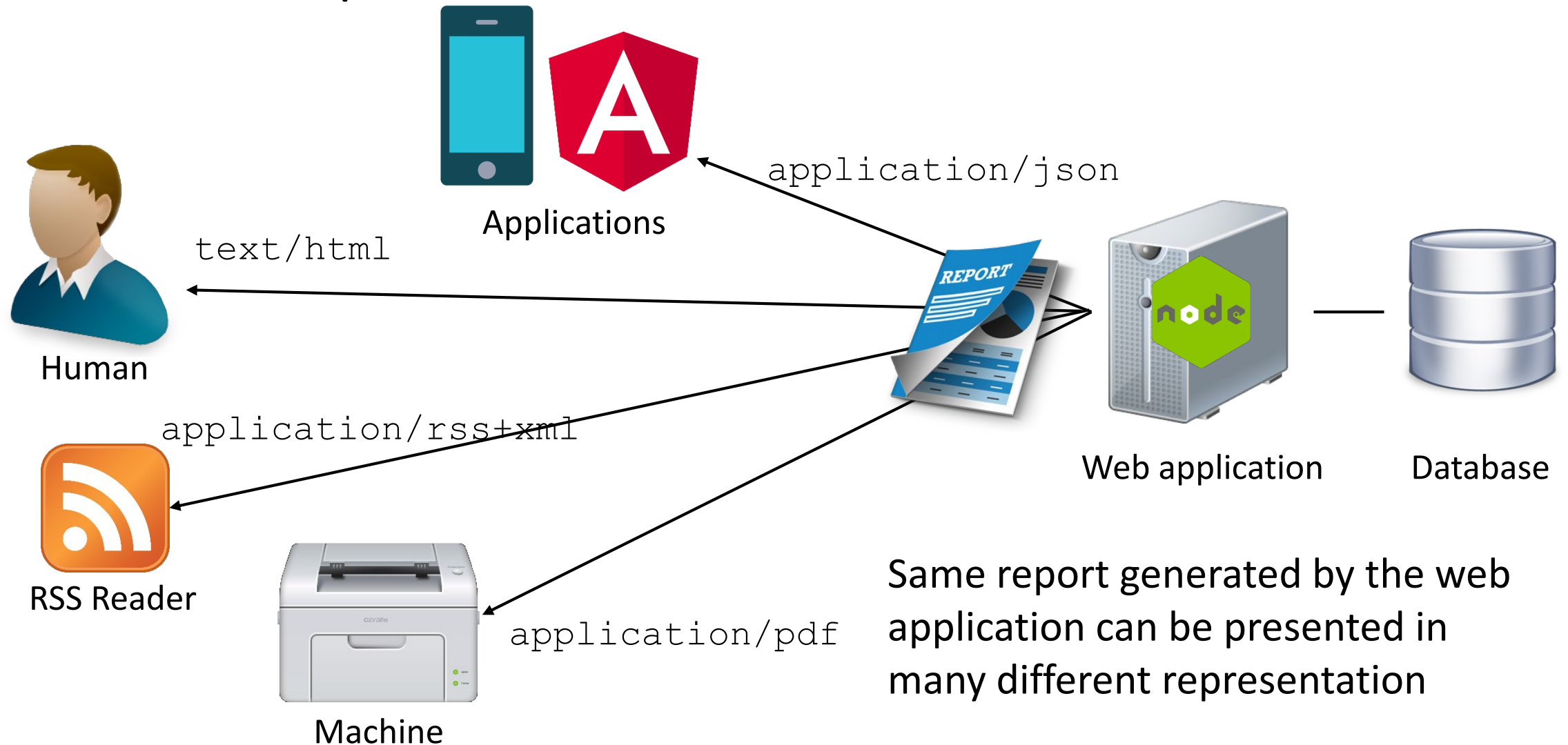


As MP3 file





Uniform Interface – Manipulate Resources thru its Representation



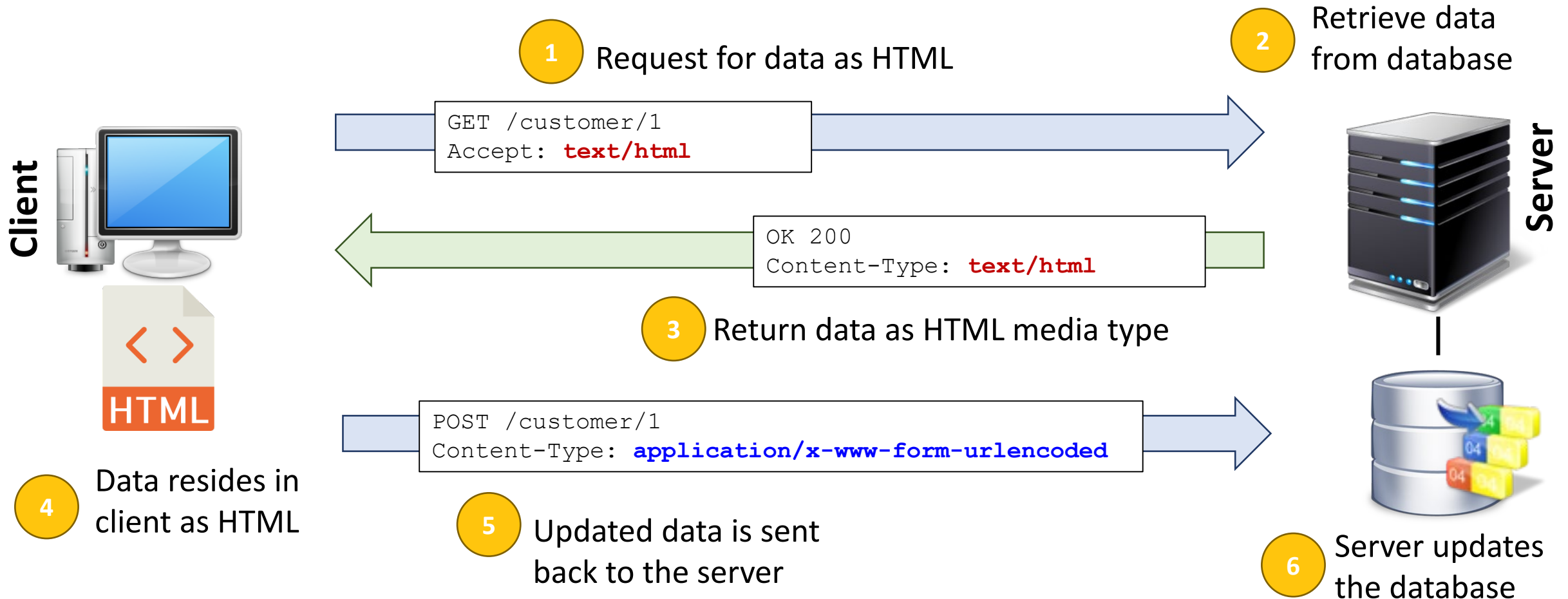


Uniform Interface – Manipulate Resources thru its Representation

- A single resource can be represented by many different types of media
- Content negotiation - a client can potentially present a list of media types to the server
 - Indicating its choice and preference for a particular resource



Uniform Interface – Manipulate Resources thru its Representation



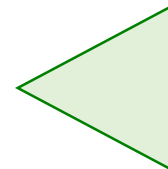
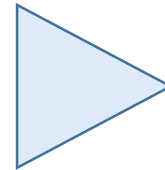


Uniform Interface - Self Describing Messages

- Every message passed from a client to a server must be self contained
 - Like a postcard or letter
- Everything to understand and decode the message is in the message
- Supports stateless servers
- The message semantics is also visible to intermediaries (if any, layer system) to understand and cache it

`GET /search?limit=50&q=hello`

`Accept: application/json,text/csv=0.3,*/*;q=0.1`



`200 OK`

`Content-Type: text/csv`



Uniform Interface - HATEOAS

- HATEOAS - Hypermedia as the engine of application states
- The hypermedia/text provides hyperlinks
 - Some link navigate to other hypermedia

```
<a href=http://...>
```

- Some links provide ways to change the state of the hypermedia

```
<form method="POST" action="/update">  
  <input type="text" name="first_name">  
  ...  
  <button type="submit">Update</button>  
</form>
```

Uniform Interface - HATEOAS

1

Wish to update a particular resource

2

Retrieve the resource

GET /resource/1

3

Requested resource is returned as a HTML form. The state of the resource is transferred to the client

HTML

5

Modified state of the resource is send back to the server

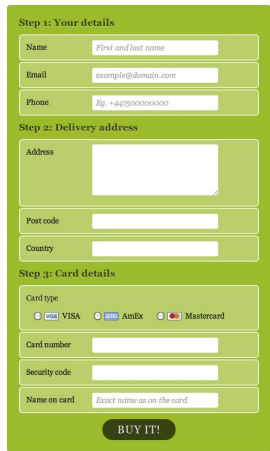
POST /resource/1

6

State of the resource is updated by the web application

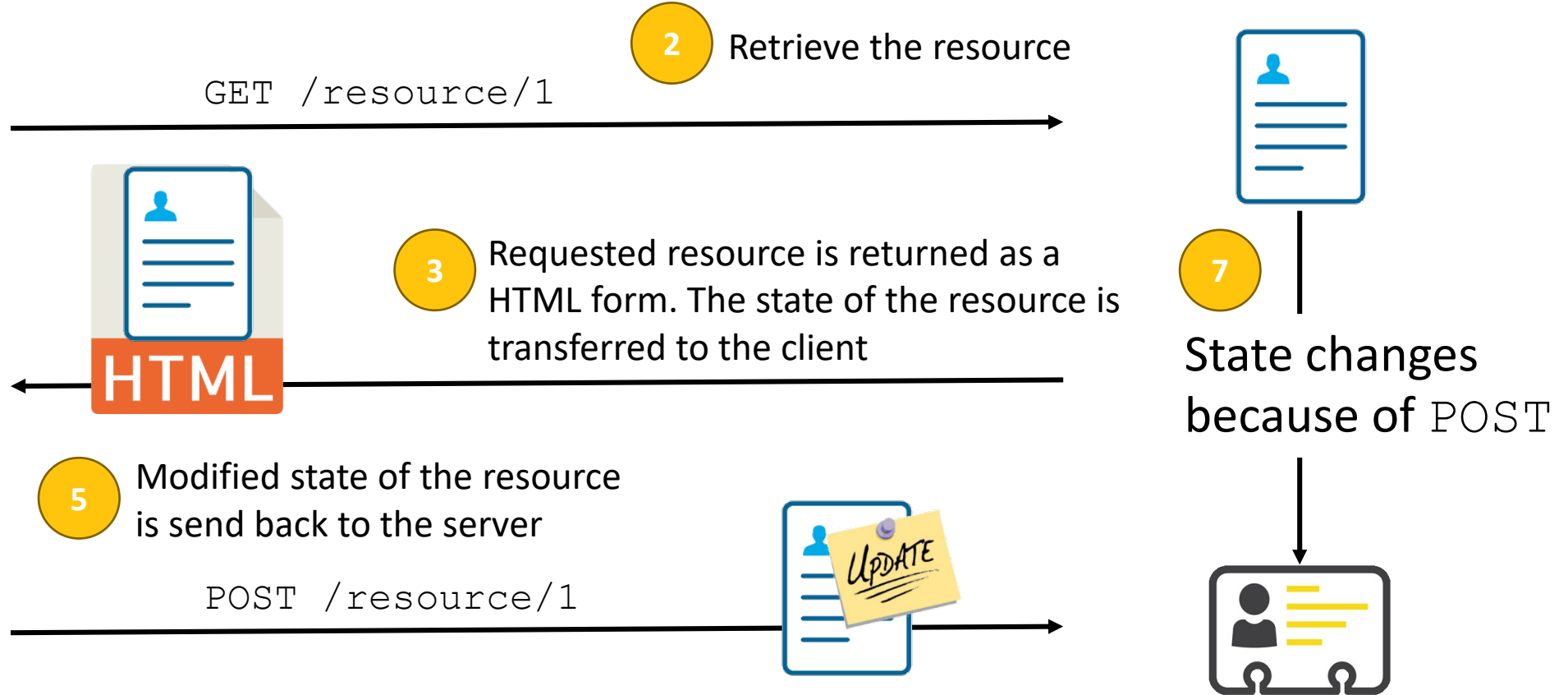
7

State changes because of POST



4

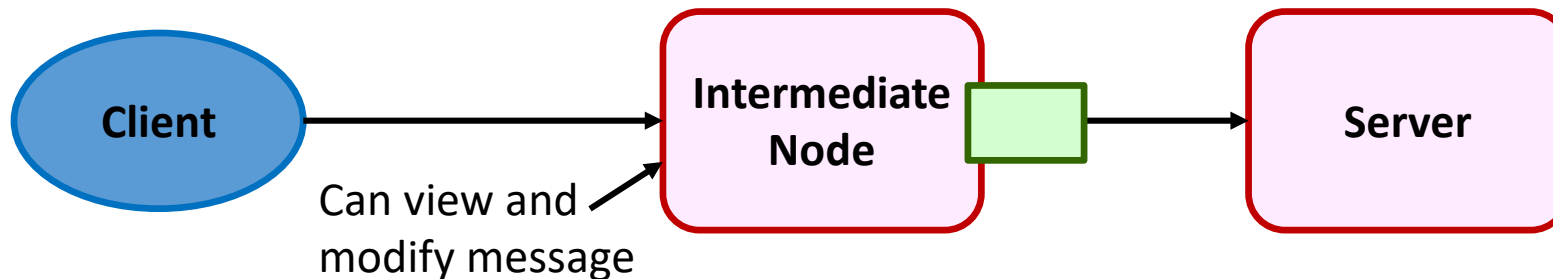
Form is the representation of the resource's state. Also contains hyperlinks to operate/change the resource





REST Constraint 5 - Layered System

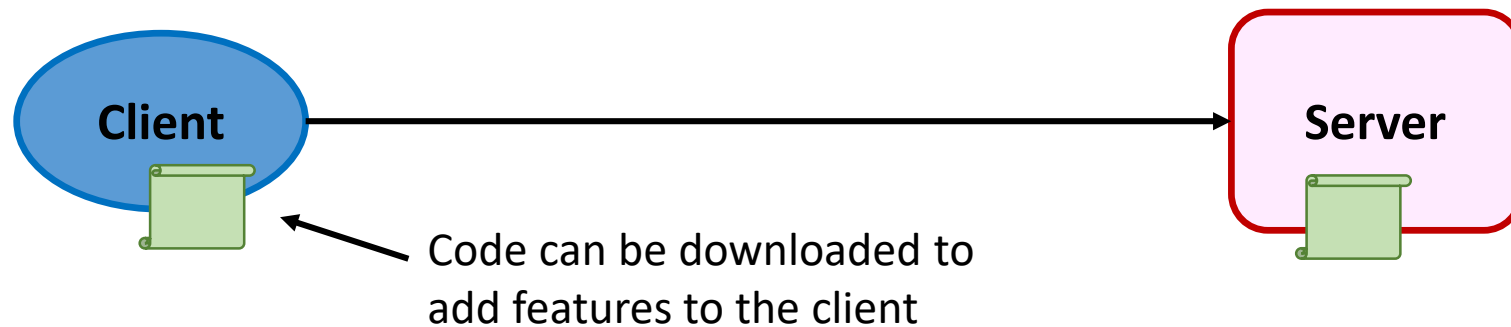
- Application from one layer cannot “see” beyond the intermediate layer
- Promotes decoupling, reduce the complexity of the overall system
 - Wrap legacy application eg. main frame application by presenting modern and simpler ways to access it
- Intermediate layers can modify data





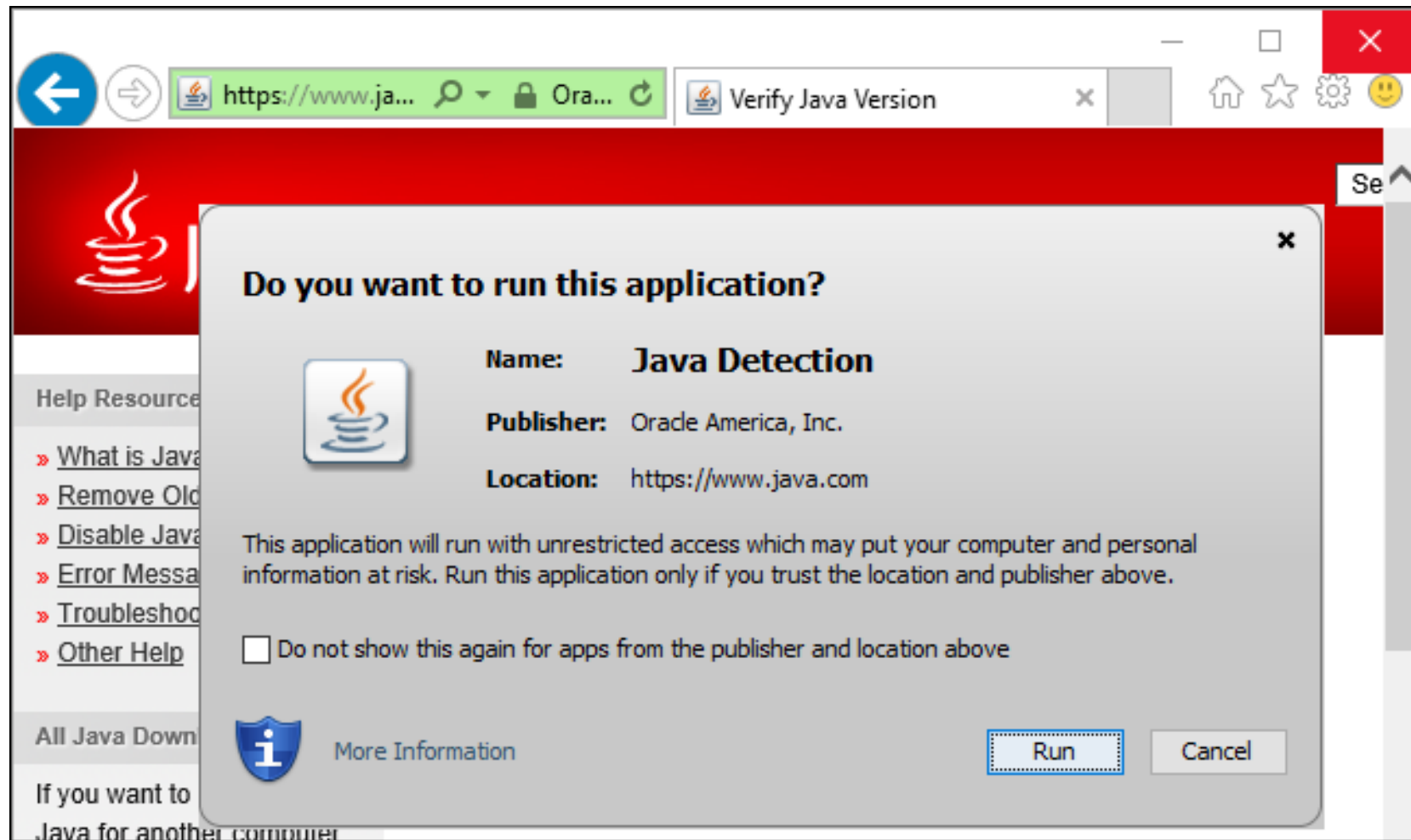
REST Constraint 6 - Code on Demand

- Client functionality can be extended by download code from the server
 - Progressively adds functionality to client





Download Code !





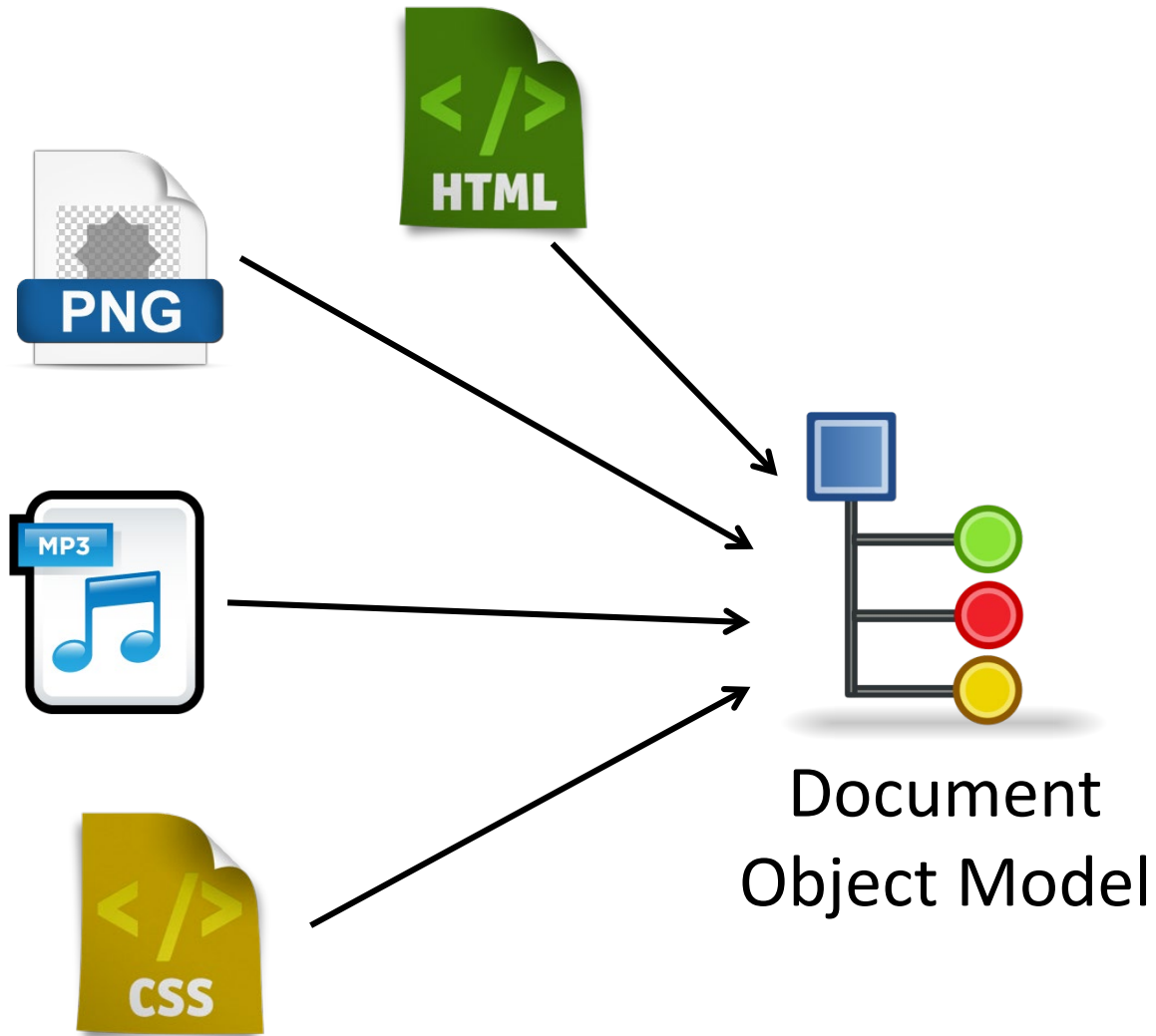
HTML Document

```
<html>
  <head>
    <link rel="stylesheet" href="styles.css">
    <script src="script.js">
  </head>
  <body>
    <h1>hello, world</h1>
    
    ...
  </body>
</html>
```

Downloading code to
your browser

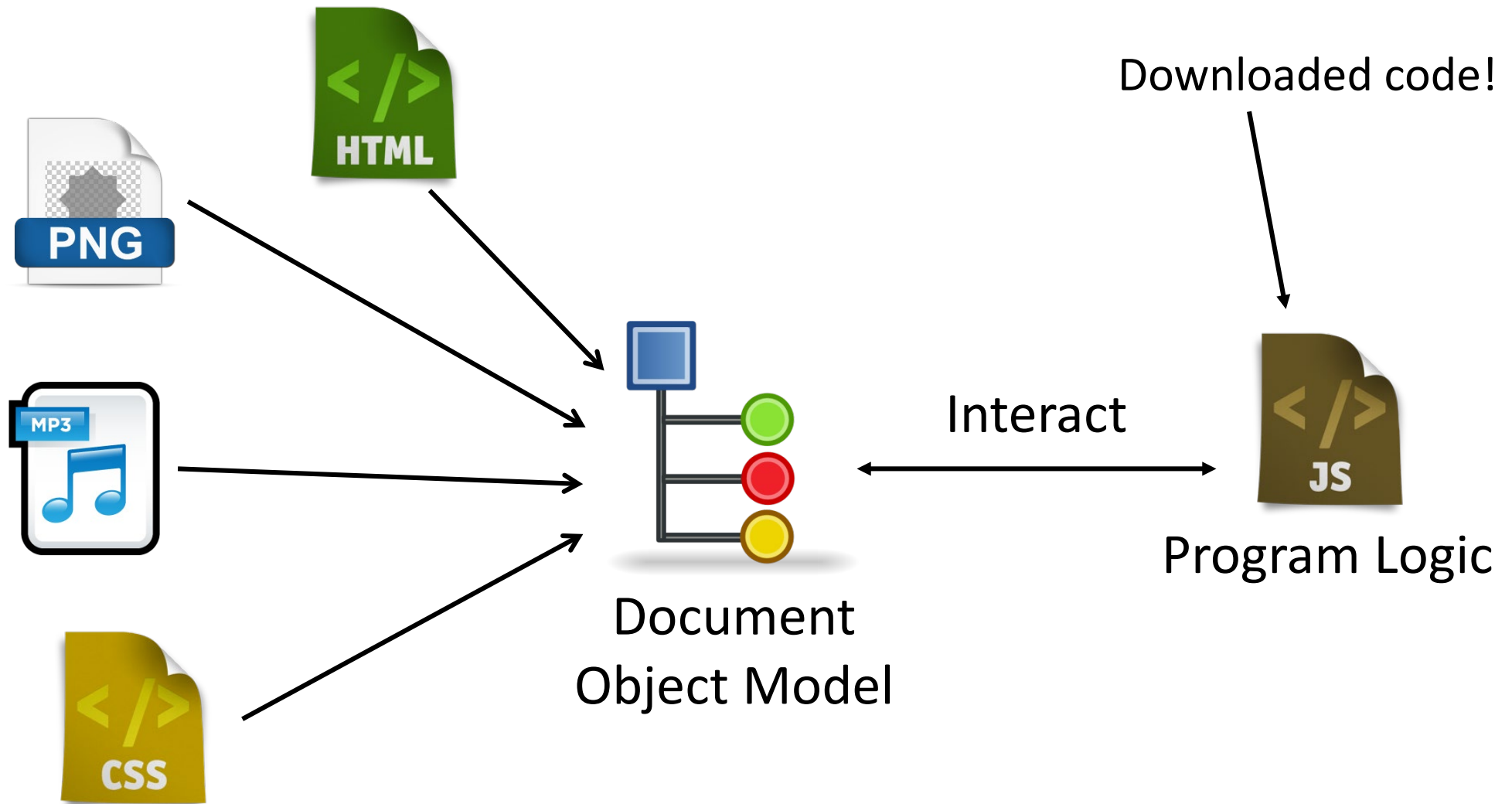


HTML5 Application





HTML5 Application

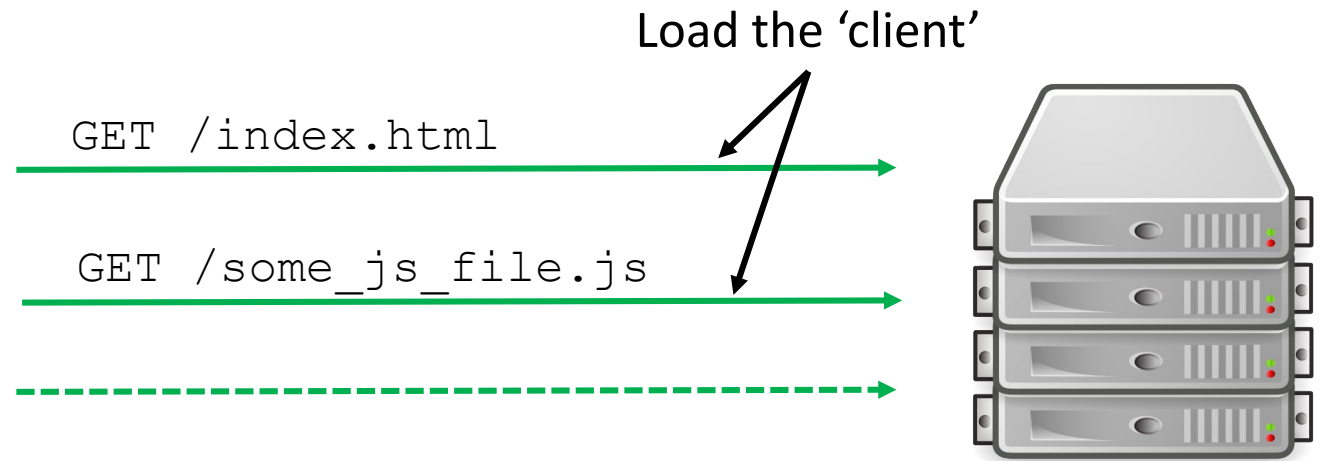




Code on Demand

- The website that provides that API also supply the client to use the API
 - Some API are public some are internal

Downloaded
JavaScript
application
running on the
browser





Code on Demand and HATEOAS

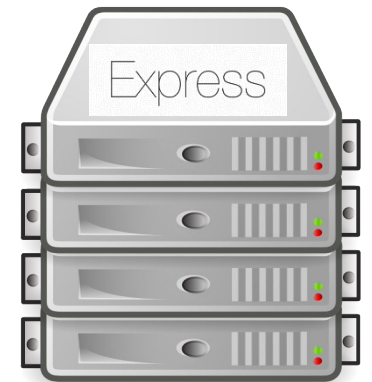
GET /api/heros
loaded the master list

Selecting a hero on the
master's list resulted in
an AJAX call to retrieve
the details



GET /api/heros

GET /api/hero/15



Selection of a hero causes a
change in application state