
INFOH-511 WEB SERVICES

LECTURE 2: REST & ROA

Seminar division



REST

= REPRESENTATIONAL STATE TRANSFER

REST: some history

At the beginning of our efforts within the Internet Engineering Taskforce to define the existing (HTTP/1.0) and design the extensions for the new standards of HTTP/1.1 and Uniform Resource Identifiers (URI), we recognized the need for a model of how the World Wide Web (WWW, or simply Web) should work. This idealized model of the interactions within an overall Web application—what we refer to as the Representational State Transfer (REST) architectural style—became the foundation for the modern Web architecture, providing the guiding principles by which flaws in the existing architecture could be identified and extensions validated prior to deployment.



- The term **REpresentational State Transfer** was introduced and defined in 2000 by Roy Fielding in his doctoral dissertation.
- Fielding is one of the principal authors of the Hypertext Transfer Protocol (HTTP) specification versions 1.0 and 1.1.

REST: some history

“ A *software architecture* is an abstraction of the runtime elements of a software system during some phase of its operation. A system may be composed of many levels of abstraction and many phases of operation, each with its own software architecture. An architecture determines how system elements are identified and allocated, how the elements interact to form a system, the amount and granularity of communication needed for interaction, and the interface protocols used for communication.



REST is a coordinated set of architectural constraints that attempts to minimize latency and network communication, while at the same time maximizing the independence and scalability of component implementations. This is achieved by placing constraints on connector semantics, where other styles have focused on component semantics. REST enables the caching and reuse of interactions, dynamic substitutability of components, and processing of actions by intermediaries, in order to meet the needs of an Internet-scale distributed hypermedia system.

From “Principled design of the modern Web architecture.”

REST: some history

The name “Representational State Transfer” is intended to evoke an image of how a well-designed Web application behaves: a network of Web pages forms a virtual state machine, allowing a user to progress through the application by selecting a link or submitting a short data-entry form, with each action resulting in a transition to the next state of the application by transferring a representation of that state to the user.



- The modern (human) Web is one instance of a REST-style architecture
- RESTful web services transfer these ideas from the human web to web services, founded by idea that there should be no essential difference between the human web (designed for human consumption) and the “programmable web” designed for consumption by software

ROA: Some History

- REST is actually a [meta-architecture](#): it is a collection of architectures – the current Web is just one instance.
- In 2007, Richardson & Ruby introduced the term [Resource Oriented Architecture \(ROA\)](#) to refer to a set of design guidelines and best practices (adhering to the REST constraints) that should be used to design RESTful Web Services
- We mention the ROA guidelines in what follows.
- Beware, however, that many current web services do not implement or follow all of these guidelines. [Your milage may vary].

Key REST concepts

- Rest consists of 4 key concepts:
 - Resources
 - Resource names (URIs)
 - Resource representations
 - Links between resources
- And 4 key properties:
 - Addressability
 - Statelessness
 - Connectedness
 - The Uniform Interface

Resources

The key abstraction of information in REST is a resource. Any information that can be named can be a resource: a document or image, a temporal service (e.g., “today’s weather in Los Angeles”), a collection of other resources, a nonvirtual object (e.g., a person), a concept and so on.

- Resource examples
 - A historical building
 - The newspaper “le soir”
 - The newspaper “le soir” at a particular date
 - The collection of all Belgian newspapers
 - The Belgian prime minister
 - The preferred newspaper of the prime minister

RESTFUL WEB SERVICE DESIGN GUIDELINE 1:

Every object manipulated by the web service (or web application) should be identified and exposed as a **resource**.

URIs: names for resources

- A URI is the name of a resource
- Examples:
 - <http://www.lesoir.be>
 - [http:// www.lesoir.be/edition/20-01-2012](http://www.lesoir.be/edition/20-01-2012)
 - [http:// www.example.org/newspapers/belgium](http://www.example.org/newspapers/belgium)
 - [http:// www.example.org/newspapers?country=belgium](http://www.example.org/newspapers?country=belgium)

RESTFUL WEB SERVICE DESIGN GUIDELINE 2:

- Every identified resource must be assigned at least one URI. This ensures it is **addressable**
- A URI should never represent more than one resource.
- Resources can have multiple URIs, but should have as few URIs as possible.

Representations

- A **representation** is a description of (some part of) the resource.
- A resource can have multiple representations (one in HTML, one in XML, one in a Google protocol buffer, ...).
- Example:
 - <http://www.example.org/newspapers/belgium> could support both HTML and JSON

RESTFUL WEB SERVICE DESIGN GUIDELINE 3:

Specify, for every resource:

- The representations that the service serves to the client
- The representations that the service accepts from the client

Use standard representations whenever possible

On content negotiation and URIs

- Content negotiation can be used to get distinct representations of the same resource.
- By giving these distinct representations their own URIs, however, we also make them **addressable**
- Example:
 - <http://www.example.org/newspapers/belgium>
 - <http://www.example.org/newspapers/belgium.json>
 - <http://www.example.org/newspapers/belgium.xml>
 - <http://www.belgium.be/consitution>
 - <http://www.belgium.be/constitution.nl>
 - <http://www.belgium.be/constitution.fr>
- Addressability is useful, e.g.,
 - <http://myservice.com/rank?site=http%3A%2F%2Fwww.example.org%2Fnewspapers%2Fbelgium.json>

Sidenote: JSON, that other XML

- JSON =JavaScript Object Notation
- A syntactical fragment of JavaScript for describing data
- Due to its increased popularity, libraries for reading & writing JSON exist for virtually every programming language
- Sometimes more compact than XML, but since it does not specify a schema, XML-like data-binding tools (JAXB, ...) are not available
- Mainly used in conjunction with AJAX

```
{ "reportData":  
  [  
    { "year": 2011, "profit": 2000000 },  
    { "year": 2010, "deficit": 1000 },  
  ],  
  "author": "John Doe"  
}
```

The Uniform Interface

- All access to resources happens through HTTP uniform interface (GET, POST, PUT, DELETE, HEAD, OPTIONS).

CRUD	REST	
CREATE	POST	Create a (sub)resource
RETRIEVE	GET	Retrieve a representation of a resource
UPDATE	PUT	Modify a resource/create a new resource
DELETE	DELETE	delete a resource
	OPTIONS	Discover what HTTP methods are supported by the resource
	HEAD	requests headers only (similar to GET but omits representation)

The Uniform Interface (cont.)

RESOURCE RETRIEVAL

HTTP Client



Web Server



Database



GET /books?isbn=122

SELECT * FROM BOOKS
WHERE ISBN=122

HTTP/1.1 200 OK

```
{ "title": "REST",  
  "authors":  
    ["Auth1", "Auth2"]  
}
```

The Uniform Interface (cont)

RESOURCE CREATION

HTTP Client



Web Server



Database

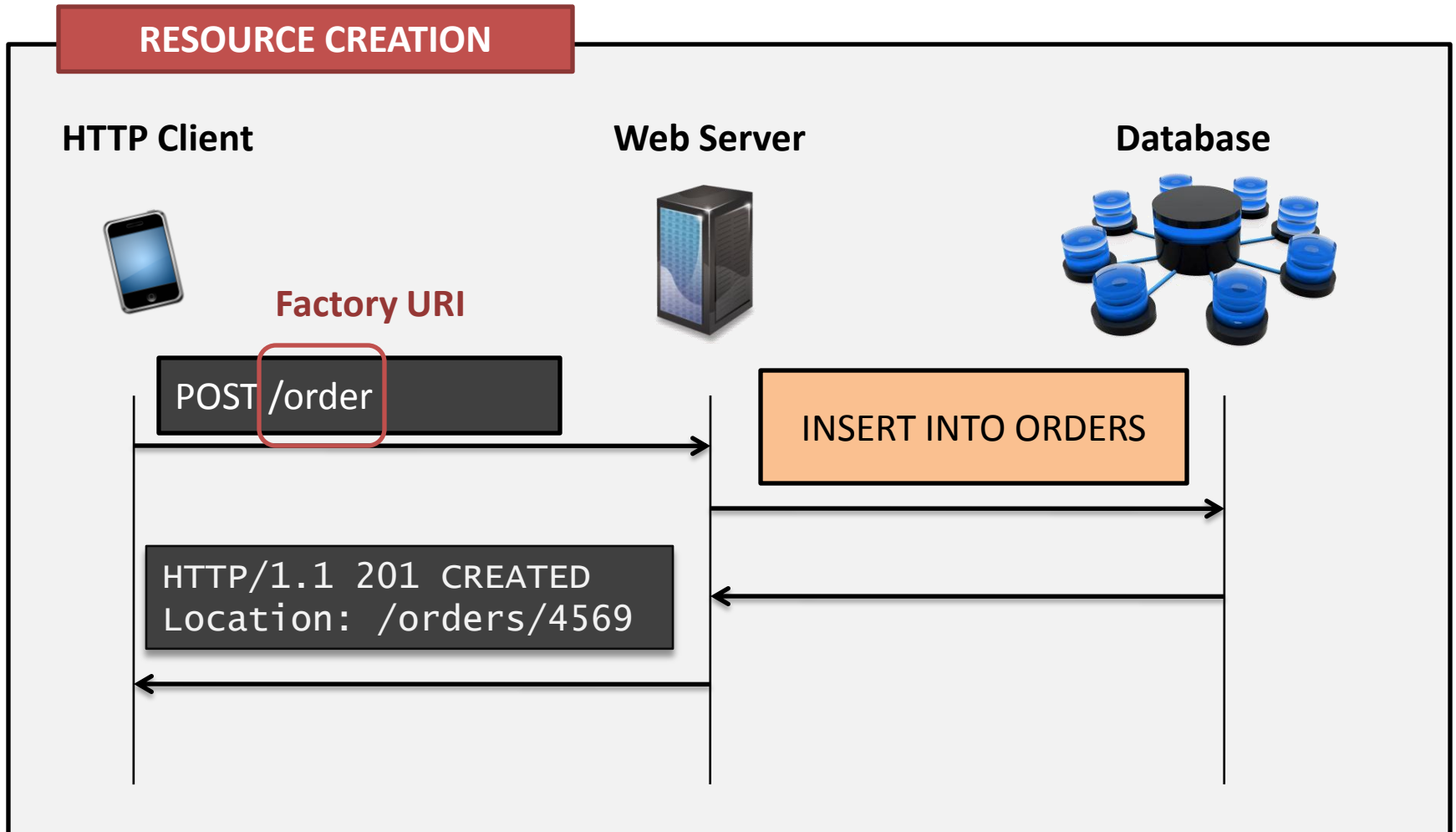


Factory URI

POST /order

INSERT INTO ORDERS

HTTP/1.1 201 CREATED
Location: /orders/4569



The Uniform Interface (cont)

RESOURCE UPDATE

HTTP Client



Web Server



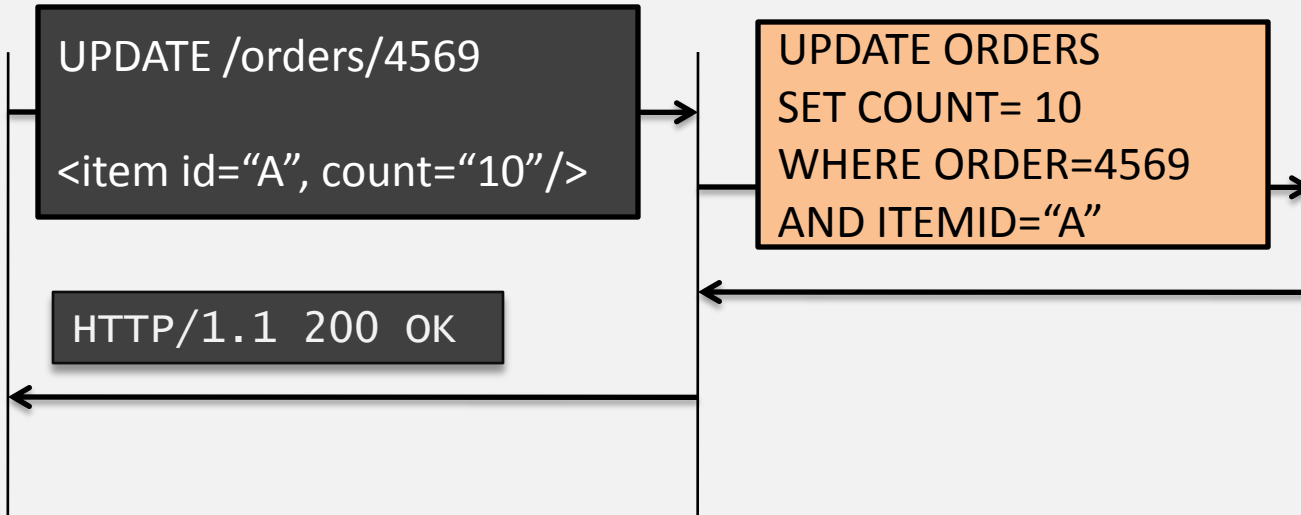
Database



```
UPDATE /orders/4569
<item id="A", count="10"/>
```

```
UPDATE ORDERS
SET COUNT= 10
WHERE ORDER=4569
AND ITEMID="A"
```

```
HTTP/1.1 200 OK
```



The Uniform Interface (cont)

RESOURCE DELETION

HTTP Client



Web Server



Database



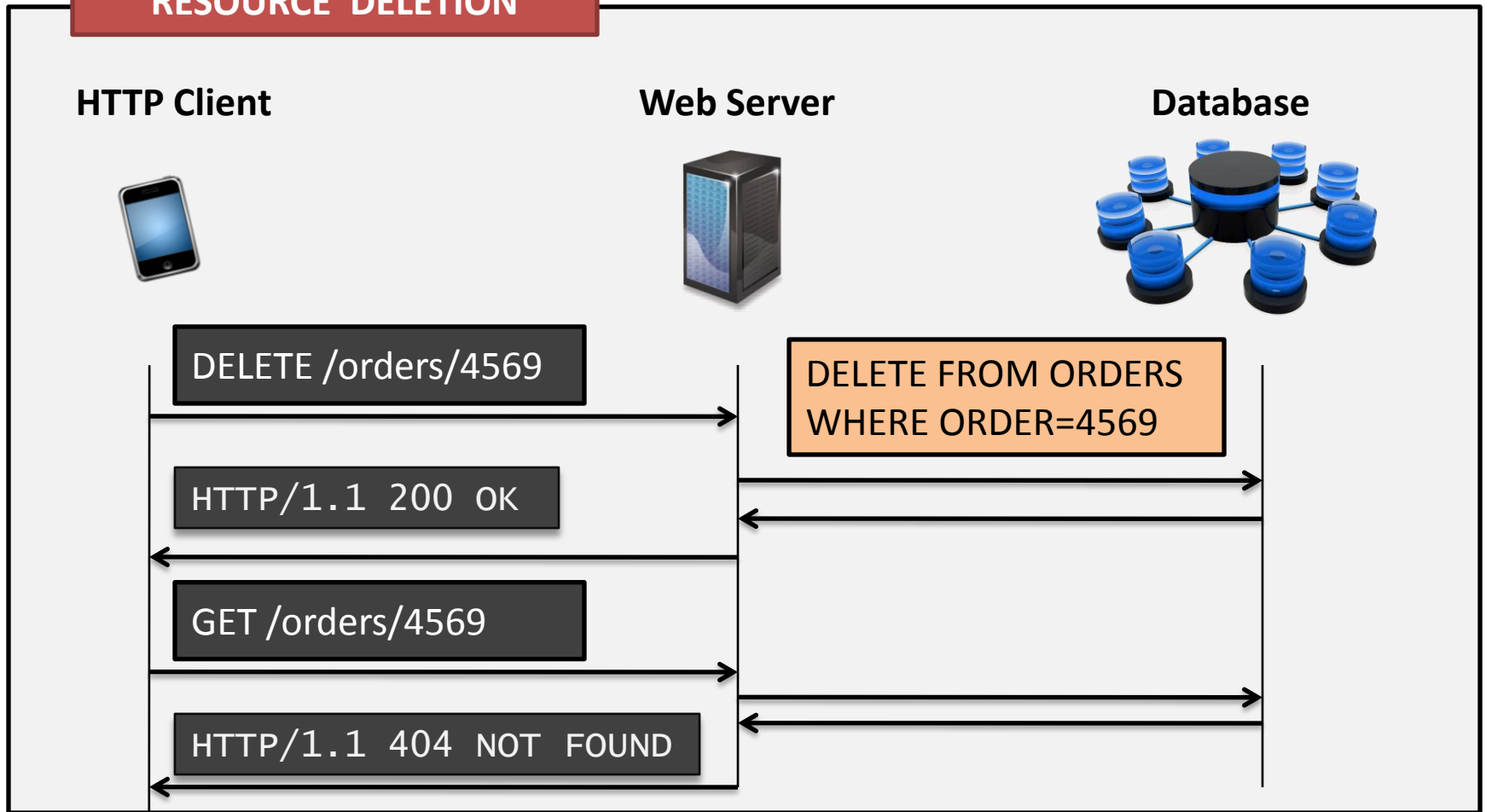
DELETE /orders/4569

DELETE FROM ORDERS
WHERE ORDER=4569

HTTP/1.1 200 OK

GET /orders/4569

HTTP/1.1 404 NOT FOUND



The Uniform Interface (cont)

- All access to resources happens through HTTP uniform interface (GET, POST, PUT, DELETE, HEAD, OPTIONS).
- All information necessary to understand the request must be contained in the request message.

RESTFUL WEB SERVICE DESIGN GUIDELINE 4:

Specify, for every URI (and hence, resource): the HTTP methods supported (e.g., GET and POST, but not DELETE, PUT)

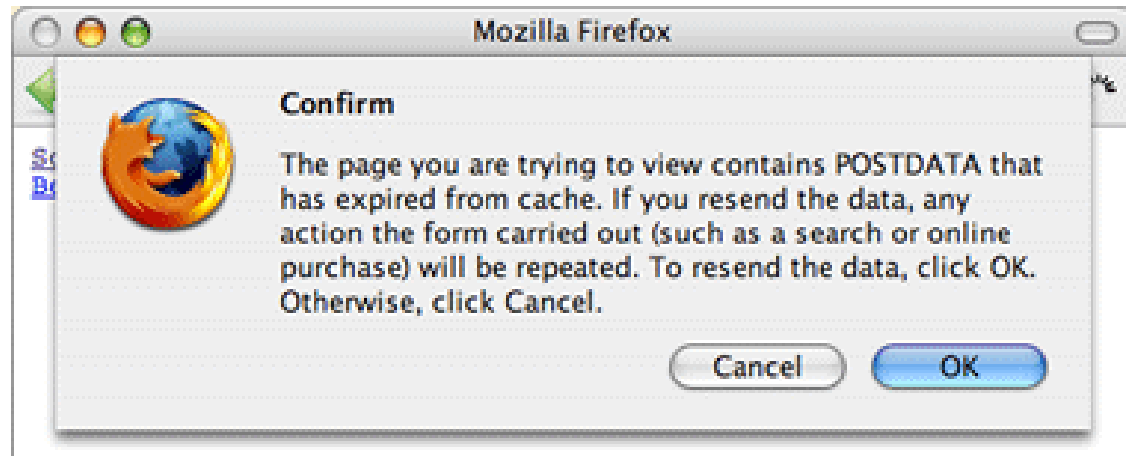
Allow querying of these operations by supporting OPTIONS

Statelessness

- Each request from client to server must contain all of the information necessary to understand the request, and cannot take advantage of any client state stored on the server.
- Application state is therefore kept entirely on the client.
- Resource state is of course still kept on the server
- Statelessness:
 - Improves reliability because it makes it easier to recover from partial failures
 - Improves scalability because:
 - Servers can quickly free computing resources after each request
 - Different requests can be handled by different servers (load balancing) exactly because the server doesn't have to manage resource usage across requests.

Safety & idempotency

- GET, HEAD, OPTIONS are read-only operations but PUT, POST, DELETE are read-write operations with side effects.
- An operation f is called **idempotent** if
$$f(f(x)) = f(x)$$
- PUT and DELETE are idempotent.
- Idempotent and read-only operations can safely be re-executed multiple times (e.g., network timeouts) without risking errors
- POST is not idempotent nor read-only, and is not safe to re-execute



Why the Uniform Interface matters

- Consider a GET of <http://api.del.icio.us/posts/delete>
- This misuses GET and does not adhere to the uniform interface
- But software programs don't know this. Programs that follow a link by GETTING it may hence (inadvertly) delete data. [e.g., Google Web Accelerator]

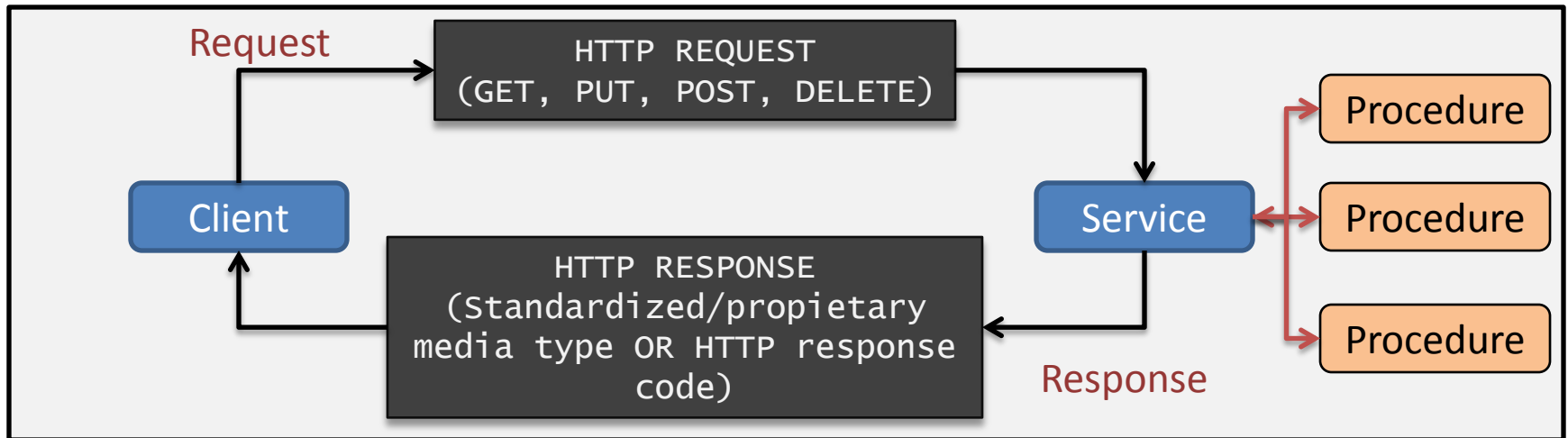
RESTFUL WEB SERVICE DESIGN GUIDELINE 5:

Use the HTTP methods correctly when designing web services.

Connecting resources

- Server response representations should include links to other relevant resources
- This makes a web service self-documenting (the representation can be parsed to see what other resources can be accessed).
- [This will become more clear with the example that follows]

Summary: REST = Resource-based API



- In a resource-based API, all procedures, instances of domain data, and files are given a URI.
- HTTP is used as a complete application protocol to define standard service behavior.
- Information is exchanged based on standardized media types (JSON, XML, ATOM, ...) and HTTP response codes where possible
- Clients manipulate the state of resources through representations (e.g., a database table row may be represented as XHTML, XML, or JSON).

Discussion

- In order to allow clients to cache representations that do not change frequently, the server should include the following headers:
 - Last-Modified
 - Etags
- This allows clients to use conditional get (using e.g., If-Modified-Since and If-None-Matches)

AN EXAMPLE

Design procedure

- Richardson and Ruby propose the following design method to obtain a resource-oriented architecture

The ROA procedure

- Figure out the data set
- Split the data set into resources
- For each kind of resource
 - Expose a subset of the uniform interface
 - Design the representation(s) accepted from the client
 - Design the representation(s) served to the client
 - Integrate this resource into existing resources, using hypermedia links
 - Consider the typical course of events: what's supposed to happen?
 - Consider error conditions: what might go wrong?

The example scenario

- Example taken from Richardson and Ruby *RESTful web services*.
- Suppose we want to construct an (imaginary) web service that serves maps of different kinds:
 - Political maps
 - Physical maps
 - Road maps
 - Geological mapsof different planets and at different scales

Step 1: Figure out the data set

- Maps are made out of *points* (specific longitude & latitude)
- A map concerns a certain *planet* (Earth, Venus)
- Some points on a map are *places* (Brussels, the Himalaya, ...)
- Places can be of different *types*

Step 2: split the data into resources

- Web services commonly expose three kinds of resources:
 - **Predefined resources** for especially important aspects of the application
 - (e.g., top-level directory of other available resources)
 - A **resource** for every **object** exposed through the service
 - **Resources** representing the results of **algorithms** applied to the data set
 - (e.g., a search resource <http://google.com/search?q=books>)

Step 2: our resources so far

- For planets and entire maps:
 - The list of planets
 - Mars
 - Earth
 - The satellite map of Mars
 - The radar map of Venus
 - The road map of Earth

Step 2: our resources so far

- For parts of maps
 - 24.9195N 17.821E on Earth
 - 24.9195N 17.821E on the political map of Earth
 - 24.9195N 17.821E on Mars
 - 44N 0W on the geologic map of Earth
- For places:
 - The Cleopatra crater on Venus
 - Campus Solbosch of ULB in Brussels, Belgium on Earth
 - The place called Springfield in Massachusetts, in the United States of America, on Earth

Step 2: our resources so far

- Algorithmic resources
 - Places on Earth called Springfield
 - Container ships on Earth
 - Craters on Mars more than 1 km in diameter
 - Places on the moon named before 1900

Step 2: in conclusion, our resources:

1. The list of planets
2. A place on a planet—possibly the entire planet—identified by name
3. A geographic point on a planet, identified by latitude and longitude
4. A list of places on a planet that match some search criteria
5. A map of a planet, centered around a particular point

Step 3: name the resources

Some guidelines, born of collective experience:

- Use path variables to encode hierarchy:

```
/parent/child
```

- Put punctuation characters in path variables to avoid implying hierarchy where none exists:

```
/parent/child1;child2
```

- Use query variables to imply inputs into an algorithm:

```
/search?q=jellyfish&start=20
```

Step 3: name the resources

1. The list of planets at the root URI
<http://maps.example.com>
2. A place on a planet, possibly the entire planet, identified by name
<http://maps.example.com/Venus>
<http://maps.example.com/Earth/France/Paris>
3. A geographic point on a planet, identified by latitude and longitude
<http://maps.example.com/Earth/24.9195,17.821>
4. A list of places on a planet that match some search criteria
<http://maps.example.com/Earth?show=Springfield>
<http://maps.example.com/Mars?show=craters+bigger+than+1km>
5. A particular map of a planet, centered around a particular point
<http://maps.example.com/geographic/Venus/24.5,17.2>

Step 4: design representations

- An XML representation of the list of planets

```
<?xml version="1.0"?>
<planets>
<planet href="http://maps.example.com/Earth" name="Earth" />
<planet href="http://maps.example.com/Venus" name="Venus" />
...
</planets>
```

- Note: Richardson and Ruby choose XHTML in their examples (using the “class” attribute to denote semantics), but this is inferior to choosing XML.

Step 4: design representations

- An XML representation of a given planet

```
<?xml version="1.0"?>
<planet>
  <name> Earth </name>
  <maps>
    <map href="/road/Earth"> Road </map>
    <map href="/satellite/Earth">Satellite</map>
    ...
  </maps>
  <description> Third planet from the So ... </description>
</planet>
```

Step 4: design representations

- An XML representation of a point on a map

```
<?xml version="1.0"?>
<point>
  <coordinate> 37.0,-95</coordinate>
  <tile src="/road/Earth/images/37.0,-95.png" />
  <nav direction="north" href="46.0518,-95.8" />
  <nav direction="northeast" href="46.0518,-89.7698" />
  <nav direction="south" href="36.4642,-84.5187" />
  <nav direction="southeast" href="32.3513,-95.8" />
  ...
  <zoom direction="in" href="..." />
  <zoom direction="out" href=".../" />
</point>
```

Step 4: design representations

- An XML representation of results of search “list of places called springfield in the US”

```
<?xml version="1.0"?>
<result>
  <description> ... </description>
  <places>
    <place href="/Earth/USA/IL/Springfield">Springfield, IL</place>
    <place href="/Earth/USA/MA/Springfield">Springfield, MA</place>
    ...
  </places>
</result>
```


Expose a subset of the Uniform Interface

- Let's say that the service is read-only for now
 - We only provide GET and HEAD
- What is supposed to happen? What can go wrong
 - GET:
 - 200 OK if resource exists + representation
 - 404 Not Found if resource does not exist ...
 - ...or 303 See other if we think we have an alternate solution
 - HEAD: same, but no representation

What about read&write resources?

- See example in handouts available on website

DISCUSSION

Discussion

- REST = synchronous
- Asynchronous can be crafted on top by splitting asynchronous requests into multiple synchronous requests, but this essentially defeats the uniform interface:
 - POST to a URL to submit a new asynchronous request
 - Server replies with 202 “Created” + URI at which the status of the request can be queried (GET) or by which the request can be deleted (DELETE)
 - [Note that asynchronous operations hence cannot be GETten because they create new suboperations]

Discussion

- Lots of web services claim RESTfulness, but actually overloaded POST:
 - They send data to a particular process to call a particular function = RPC call
 - If possible, restrict POST to the use of factory URI's!

References

- R. T. Fielding and R. N. Taylor. *Principled design of the modern Web architecture*. ACM Transactions on Internet Technology (TOIT), 2(2), May 2002, pp. 115-150.
- L. Richardson and S. Ruby. *RESTful Web Services*. O'Reilly, 2007.