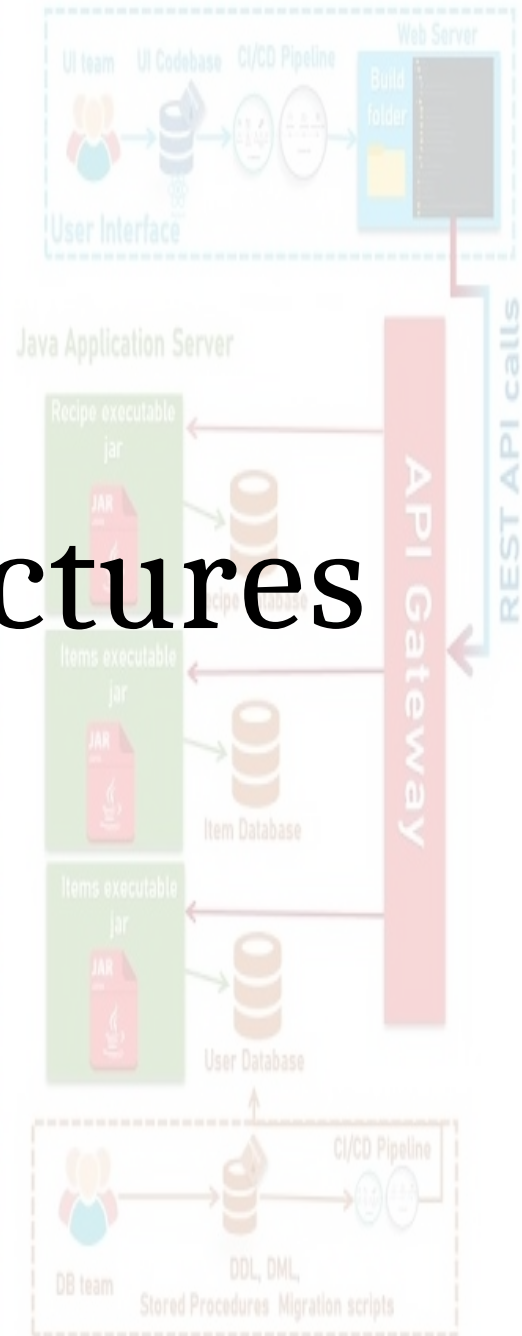
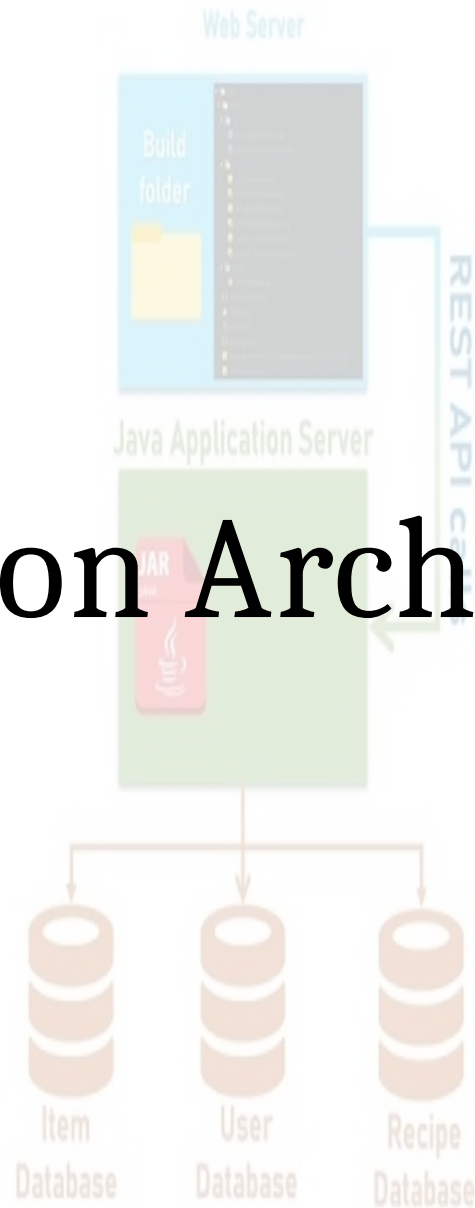
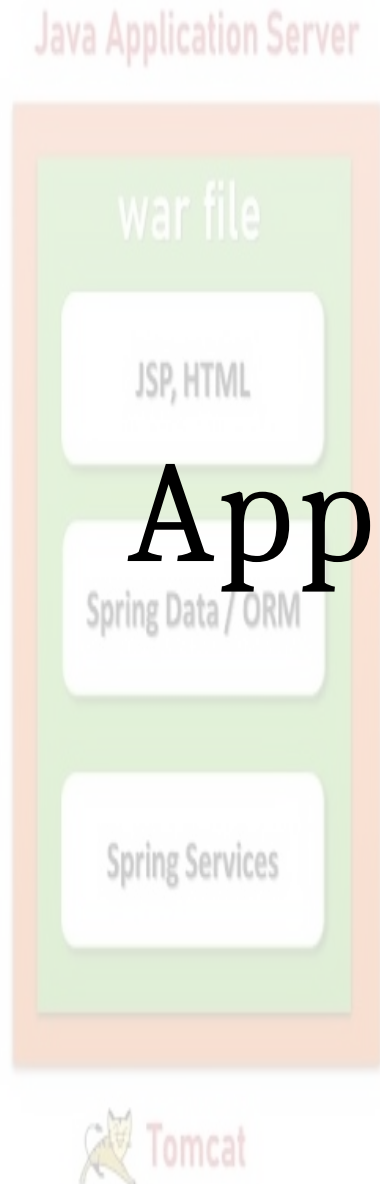


# Application Architectures



# Layered structure

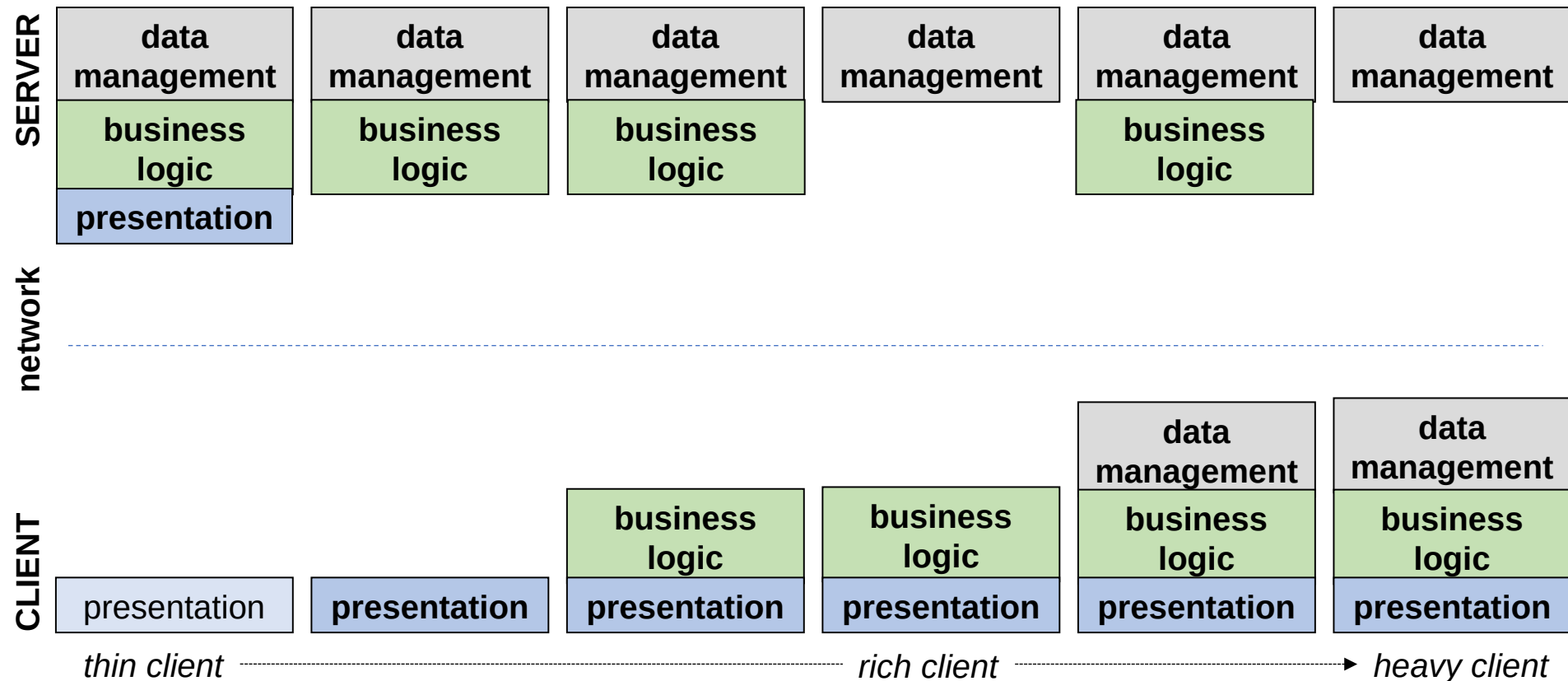
Division of the work of an application into 3 general functions, which can evolve independently:

- Presentation:  
user input and commands, and display
- Business logic:  
business objects, rules, processing logic, processes
- Data:  
storage and logical access



# Distribution onto « Tiers »

Distribution of the layers onto multiple machines (“tiers”) communicating over a network





# Monolithic and Single- tier Applications

# Monolithic application

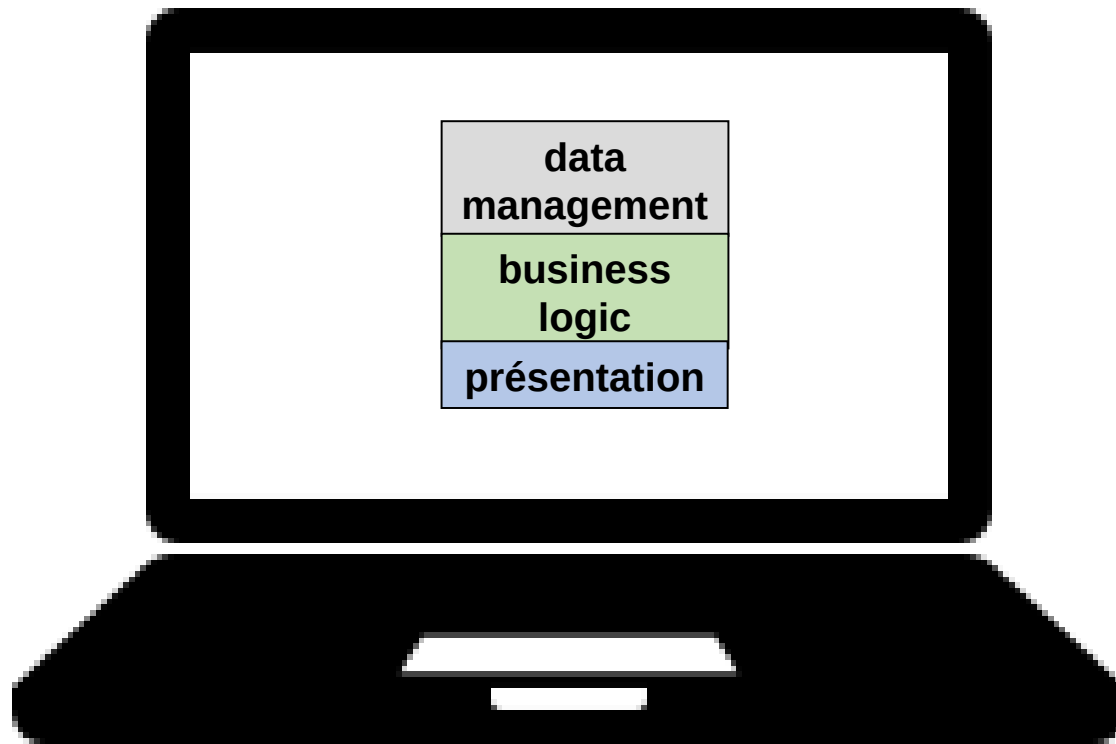
The 3 application layers are intimately interlaced in the same code base

```
import java.io.*;
public class ReadFromFile {
    public static void main(String[] args) throws Exception {
        File file = new File("C:\\Users\\galtier\\Desktop\\test.txt");
        BufferedReader br = new BufferedReader(new FileReader(file));
        String st;
        while ((st = br.readLine()) != null)
            System.out.println(st.toUpperCase());
        encrypt(file, "mySecretKey");
    }
}
```

The diagram illustrates the mapping of code to application layers in a monolithic application. Three colored boxes represent the layers: a grey box for 'data management', a blue box for 'presentation', and a green box for 'business logic'. Blue arrows point from specific lines of code to these boxes: one arrow points from the `File` and `FileReader` objects to the 'data management' box; another points from the `System.out.println` statement to the 'presentation' box; and a third points from the `encrypt` method call to the 'business logic' box.

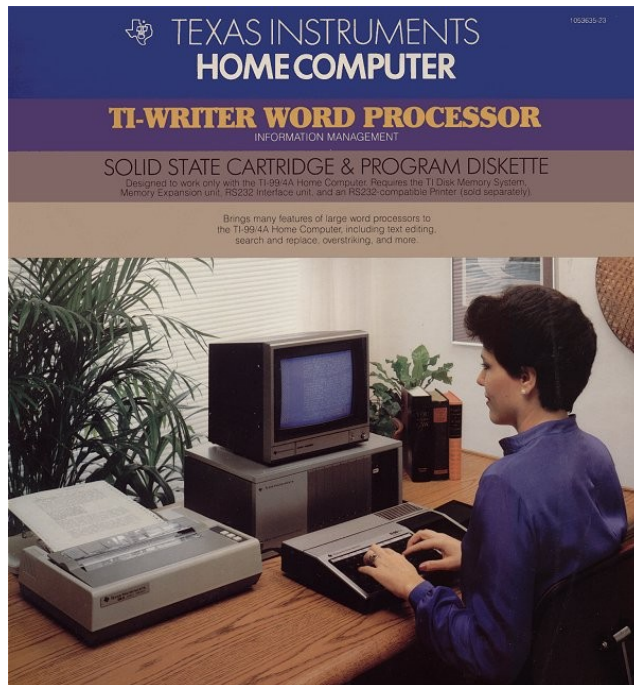
# Single-tier Application

The 3 application layers run on the same computer



# 1<sup>st</sup> architectural style, but still relevant

- The area of “pre-network” PCs (late 70 's – mid 80's)
- Still lots of stand-alone apps



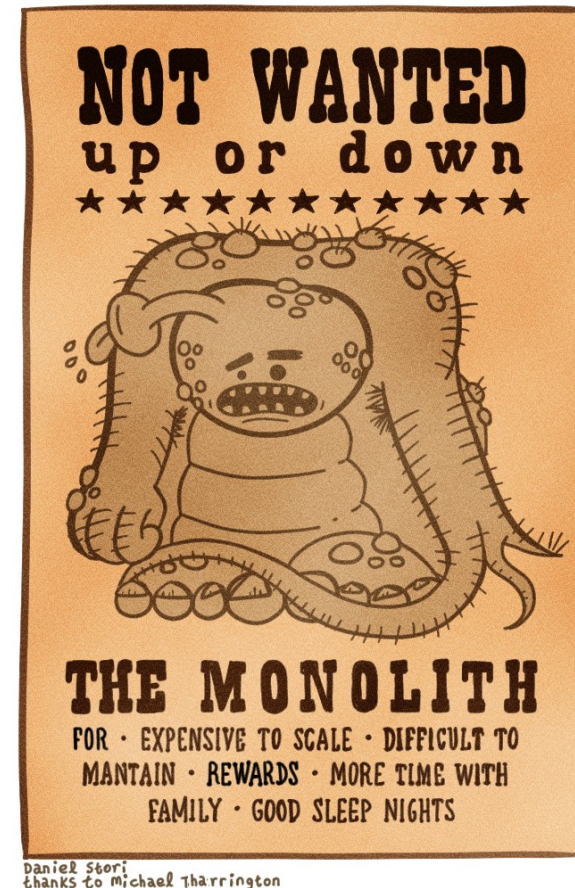
# Advantages of single-tier

- Performance: 0 latency
- Safety by isolation
- Operate even in disconnected mode
- Simplicity (complexity reduced to the one of the code)



# Disadvantages of monolithic applications

- Code is complex to learn, debug and evolve
- Even a minor upgrade requires a complete reinstallation of the entire application
- A failure in one “layer” renders the application completely unusable
- Inability to leverage heterogeneous technologies
- Not cloud-ready



# Disadvantages of single-tier applications

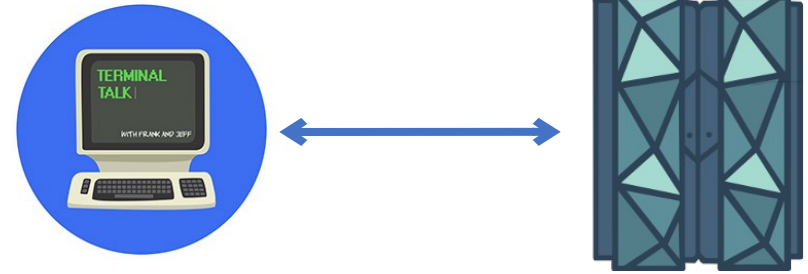


- Performances: depend on the capabilities of the host
- Shared resources impossible, requires duplicates (waste of resources)
- No fault tolerance
- Nomadism is difficult:
  - Access limited to physically logged-in users
  - More difficult (if not impossible) to continue a task from a different workstation
- Deployment is difficult:
  - Requires actions on each terminal
  - To be reinstalled if the underlying system needs to be reinstalled
- From the publisher's point of view:
  - No fix possible without user action
  - Application vulnerable to reverse engineering

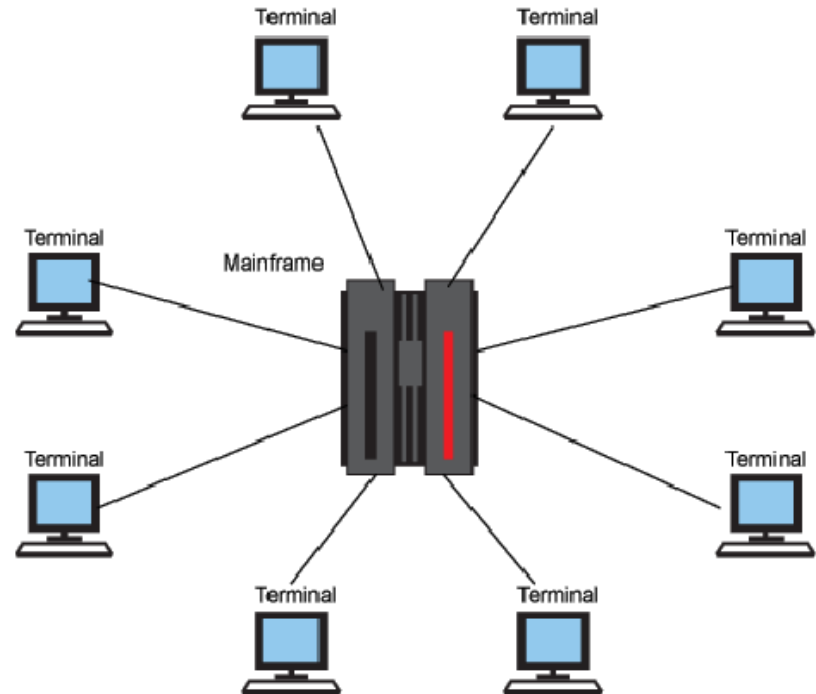


# Mainframe Architectures

# Principle “host” Architecture



- Supercomputer :
  - ensures the data persistence, processing, and presentation
  - proprietary hardware and OS (IBM)
- passive clients :  
thin client visualization application



# Advantages

- Performances: handle a very large number of simultaneous queries on very large databases
- Consistency, stability and long-term support
- Security
- Reliability (IBM Z customers: 99.9999% uptime)

Robustness: <https://www.ibmmainframeforum.com/mainframe-videos/topic10889.html>

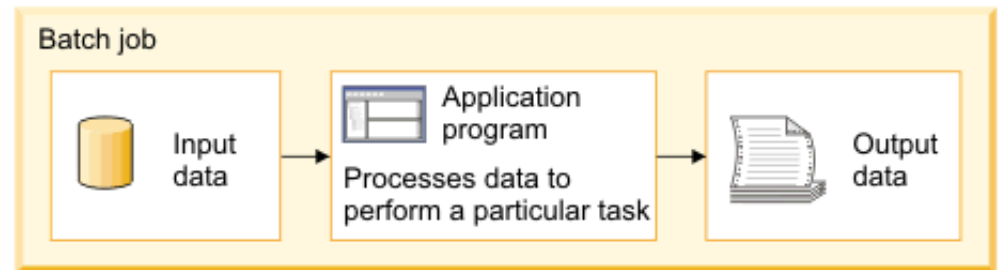


# Performances

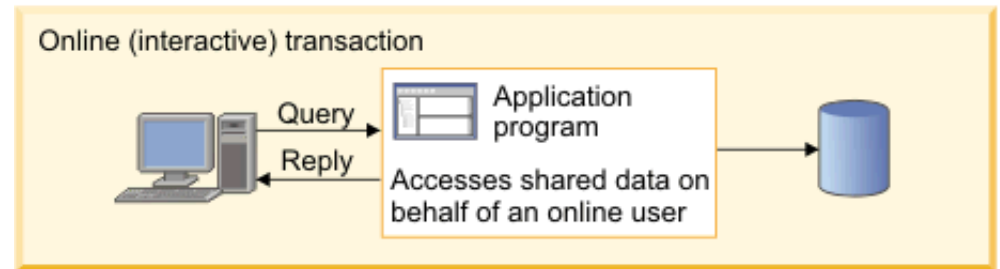
- Ability to process a very large number of simultaneous queries on very large databases

Batch or real time operation:

- Batch back-office



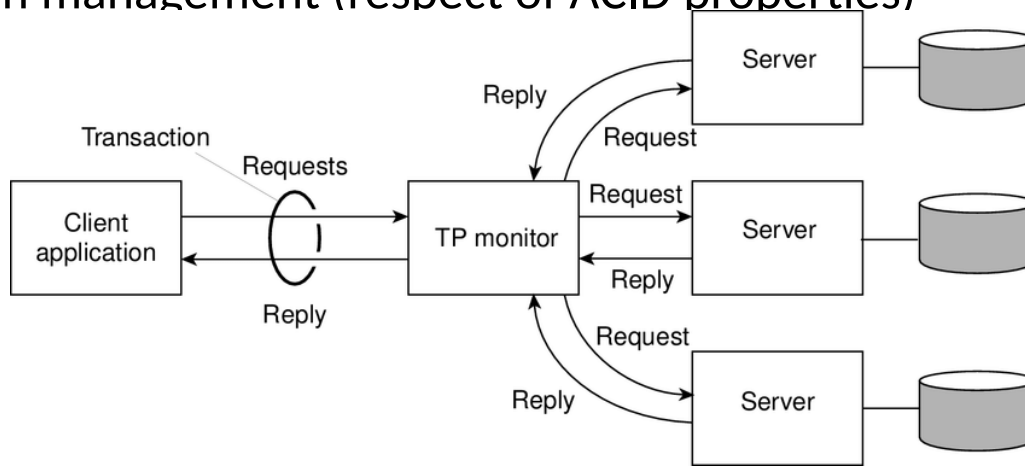
- Transactional



- Used in banks, insurance companies, airlines...

# Transactions

- *Program accessing and/or modifying persistent data*
- A good transaction is
  - **A**tomic
  - **C**onsistent
  - **I**solated
  - **D**urable
- Transactional monitor ("TP monitor")
  - Schedules transactions executed in parallel
  - Multiplexing of requests on system resources
  - Transaction management (respect of ACID properties)





# Extensively used

- 71% of the Fortune 500, 96 of the top 100 banks use mainframes
- process 30 billion business transactions per day, 87% of credit card transactions
- 250 billion lines of COBOL code, and 5 billion new lines each year
- Growth Outlook:
  - demand for HPC
  - increase in the number of banking transactions
  - development of blockchain

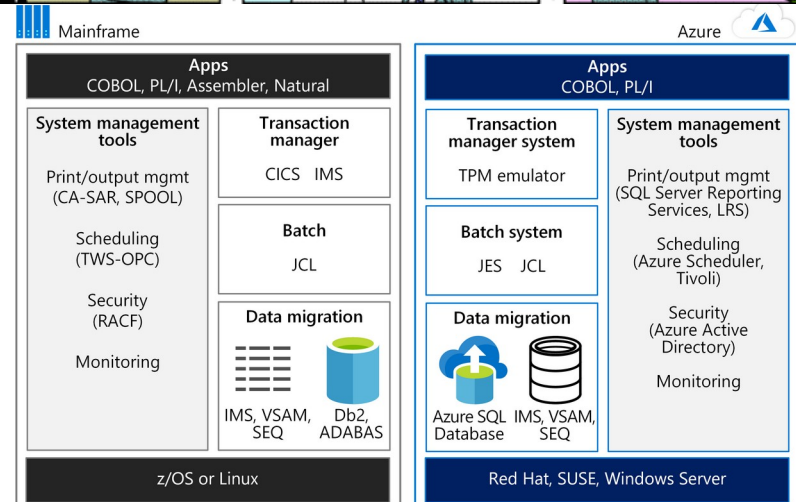


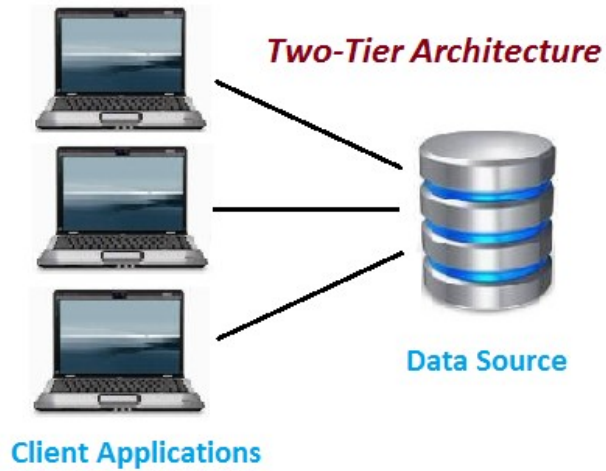
# Obstacles to growth

- Proprietary solutions
- Huge investment
  - but no more than a server farm

(<https://planetmainframe.com/2021/09/the-ibm-mainframe-the-most-powerful-and-cost-effective-computing-platform-for-business/>)

- Shortage of skilled mainframe staff
  - but Cobol is easy to learn
- Real alternatives + migration experience

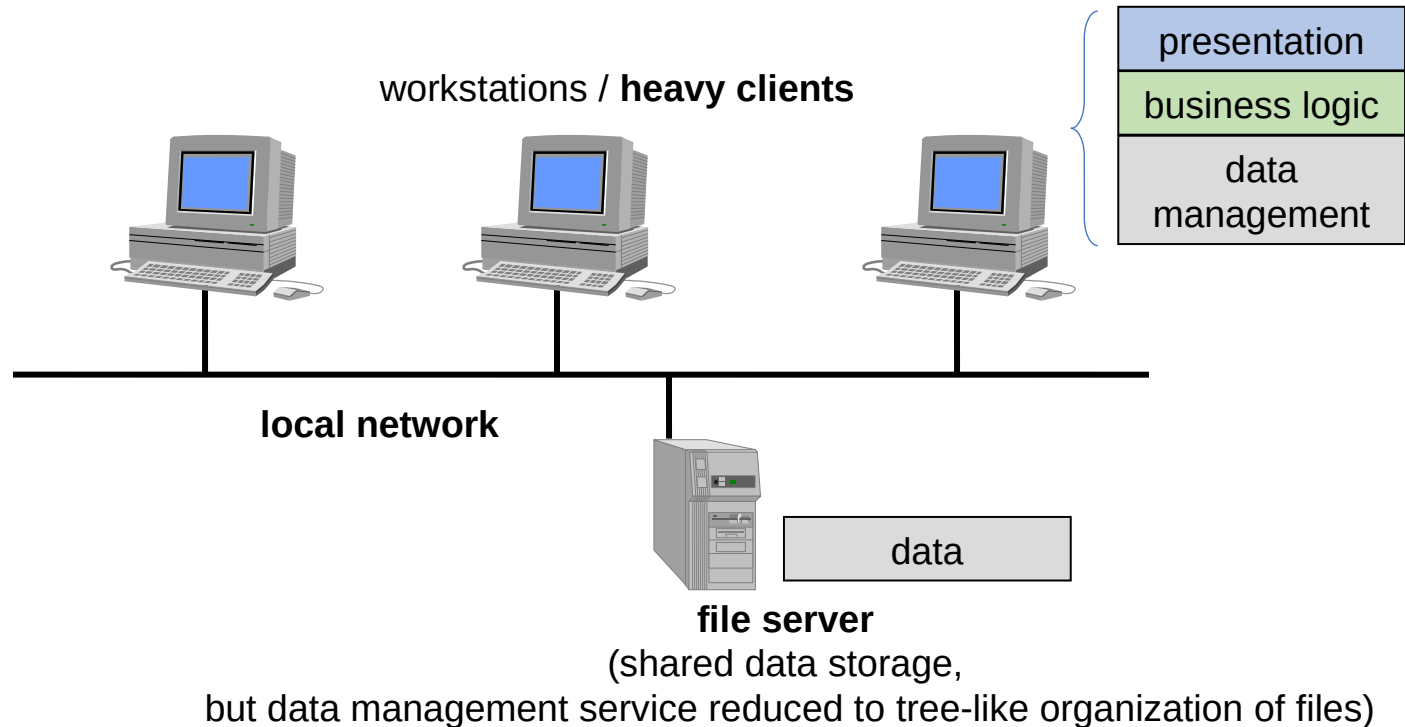




# 2-tier Architecture

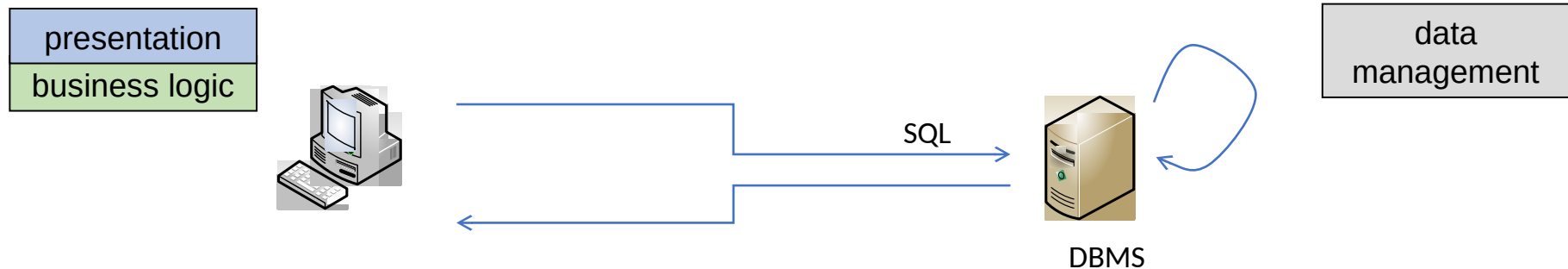
# The origin: “1.5-tier” Architecture

- Development of LANs



- Advantages: information sharing:
  - better communication
  - requires less resources

# 2-tier Architecture



- Central database server
  - Manages physical I/O and provides logical data manipulation
  - Integrity control
  - Secure, optimized, transactional access
- Data handling is decoupled from its representation on disk, closer to the application logic

# 2-tier Architecture limits

- identical problems to single-tier:
  - Not tolerant to client or server failures, updates require user's action...
- excessive use of stored procedures:
  - breaks the principle of single responsibility
  - complex to maintain
  - adherence with the physical model
- performance :
  - Server and access network = bottlenecks

# Thank you, 2-tier Architecture

- Microcomputing (previously confined to office automation) has taken on a growing role in IS
- The DBMS offer has grown, SQL has become widespread
- Has triggered the evolution towards more flexible architectural proposals
- Still relevant for simple applications

Presentation Layer



Business Logic Layer



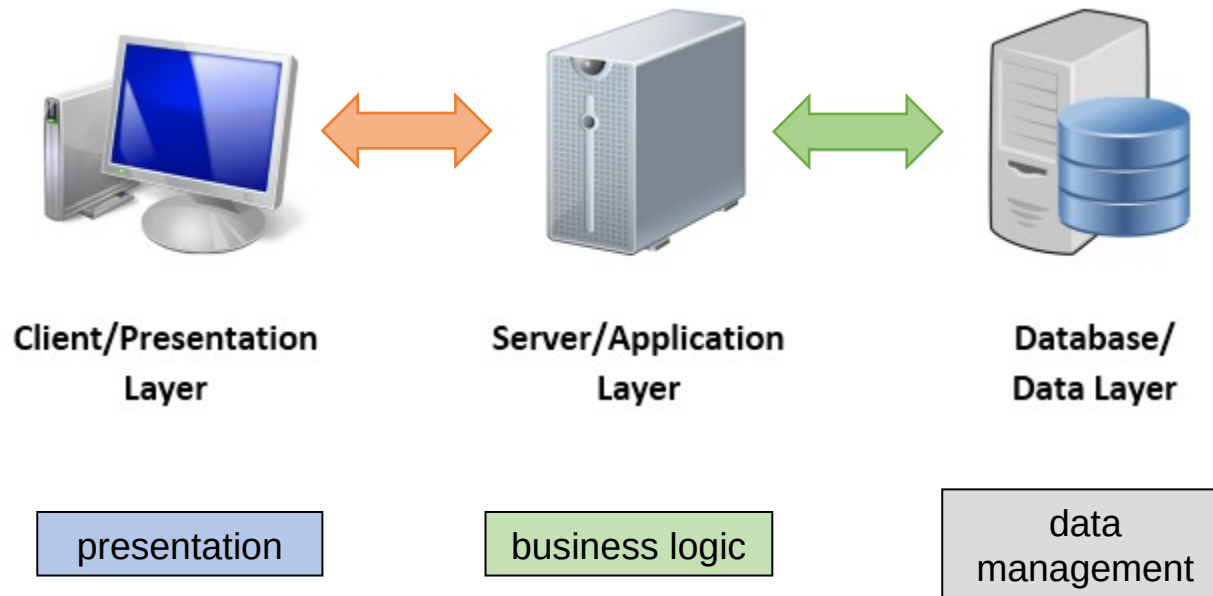
Data Access Layer



Data  
Source

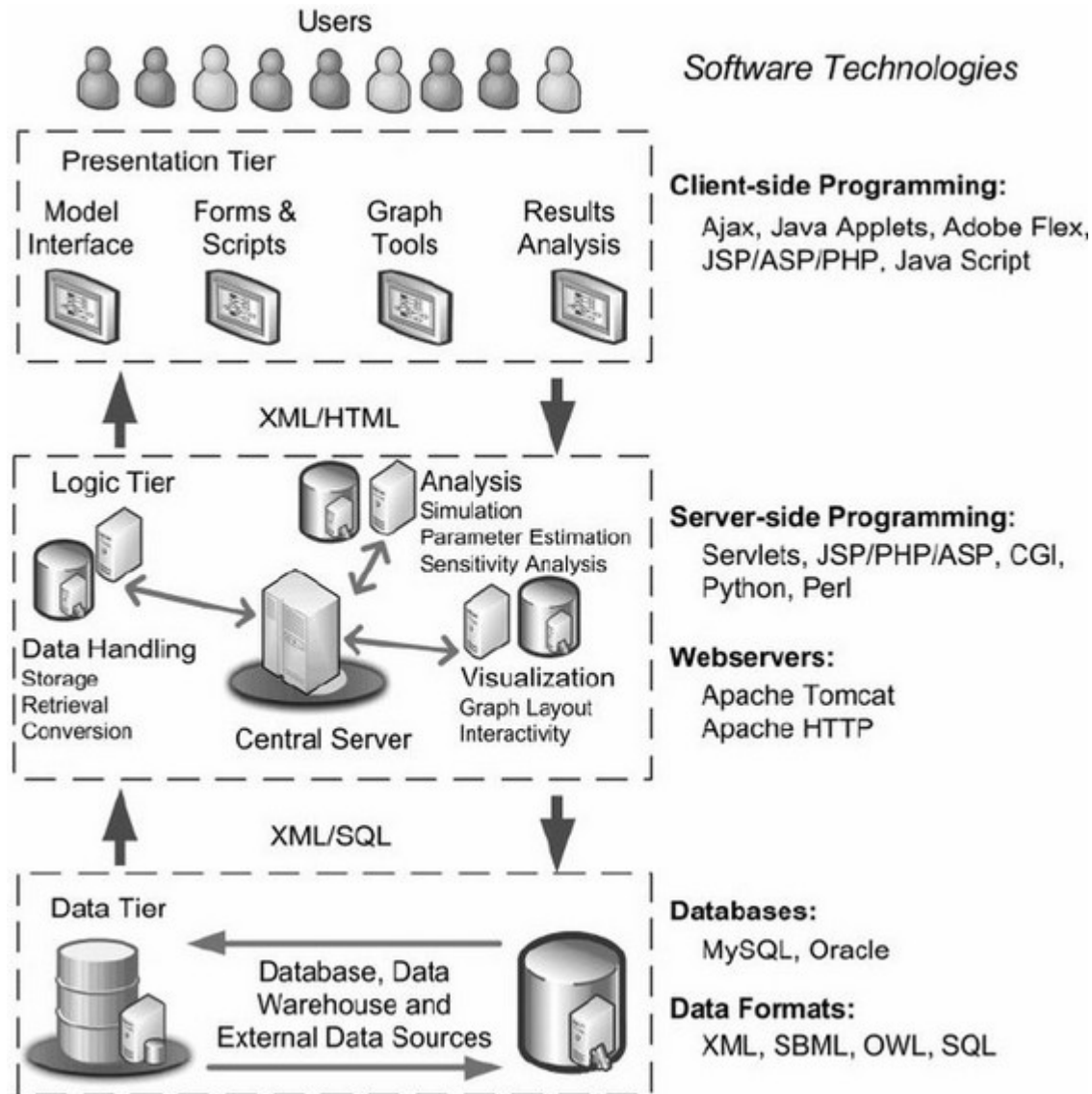
# 3-tier to 5-tier Architectures

# 3-tier

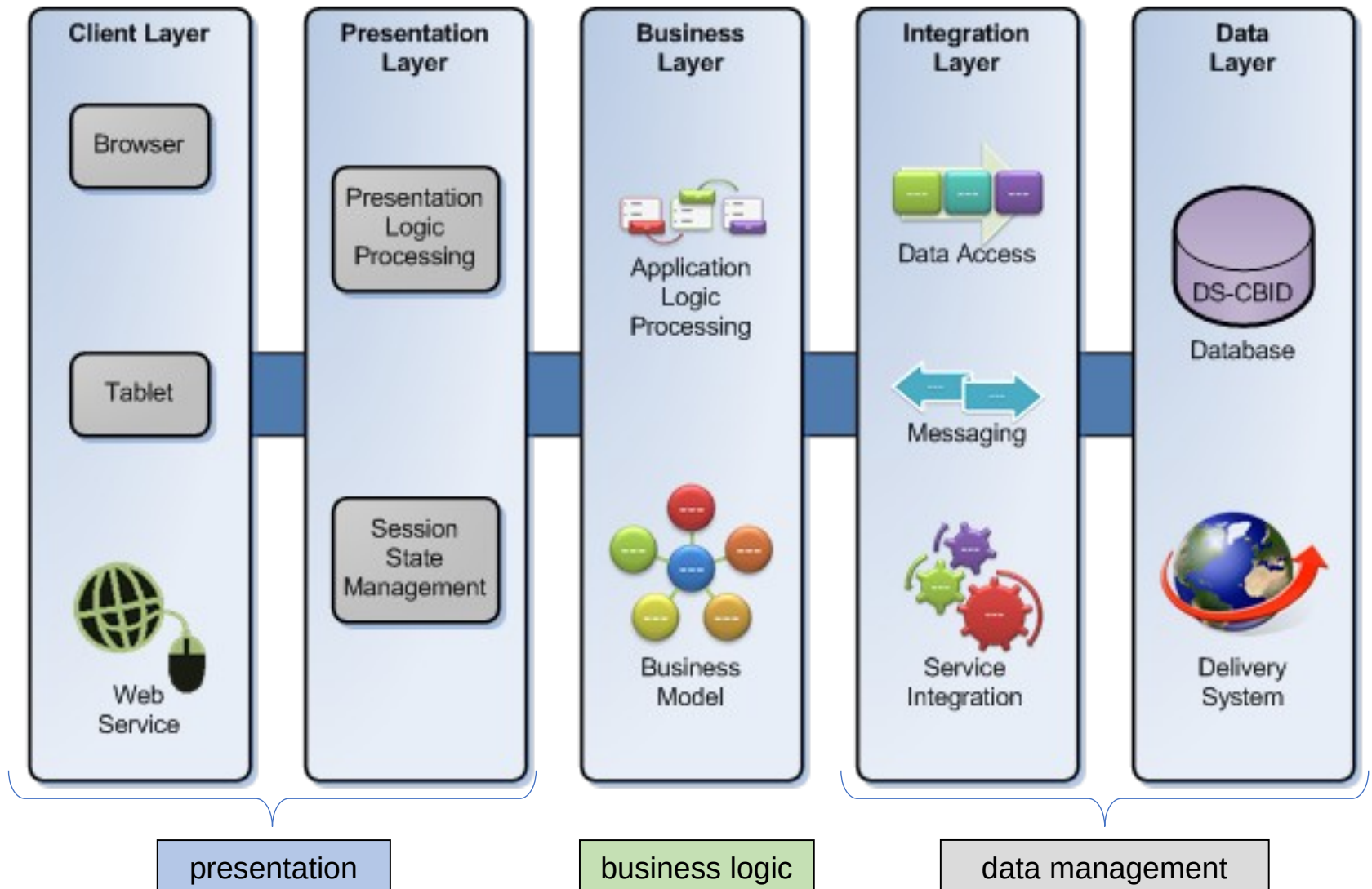




# Example: Classical Web Architecture



# 4-tier, 5-tier



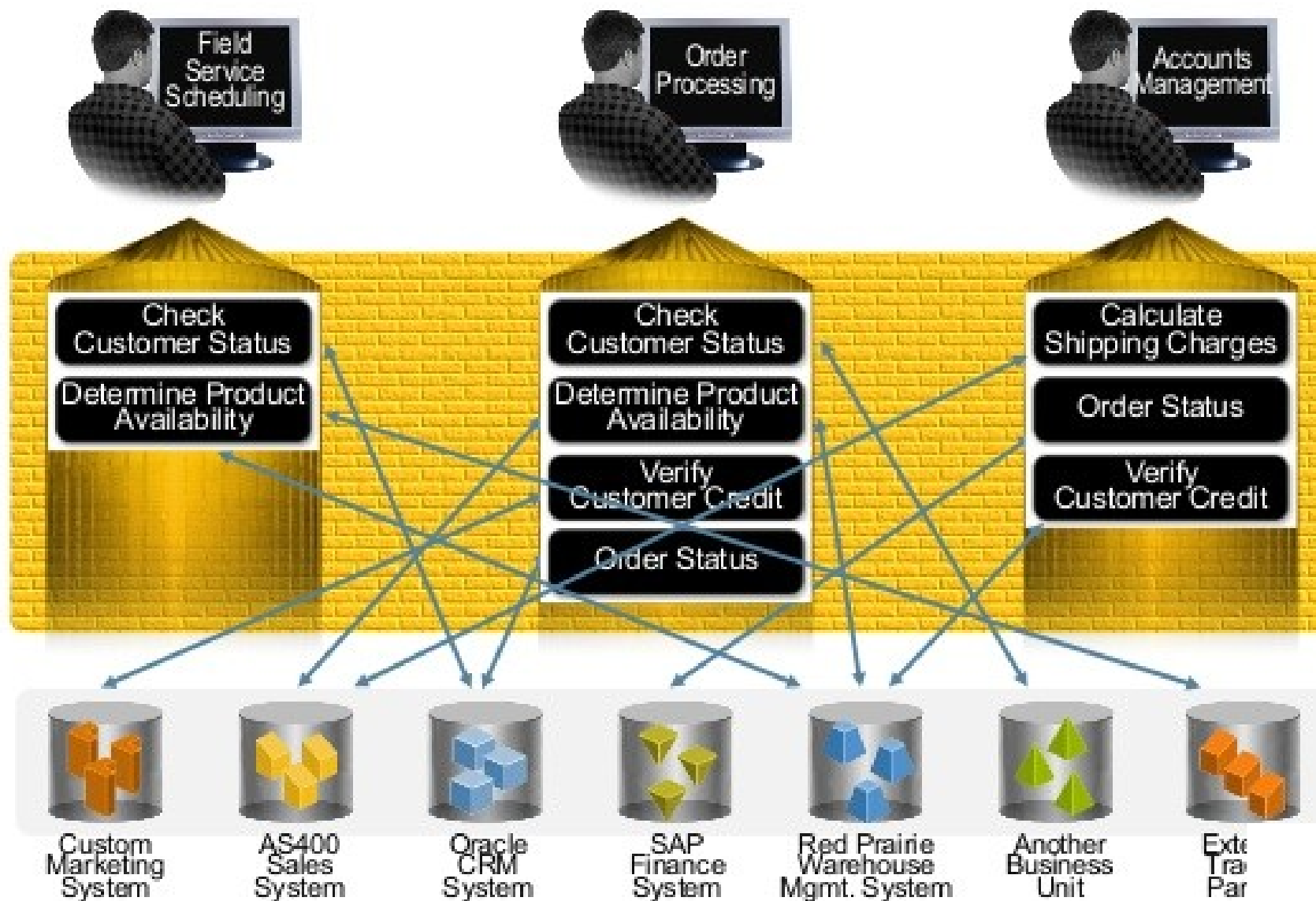
# Perspectives for multi-tier architecture

- Corrects some of the problems of 2-tier architecture
  - Maintainability, evolvability, deployment
- Very popular model for non-intensive systems
- But to be completed to meet the challenges of reliability, performance, and scalability



# Micro-services Architecture

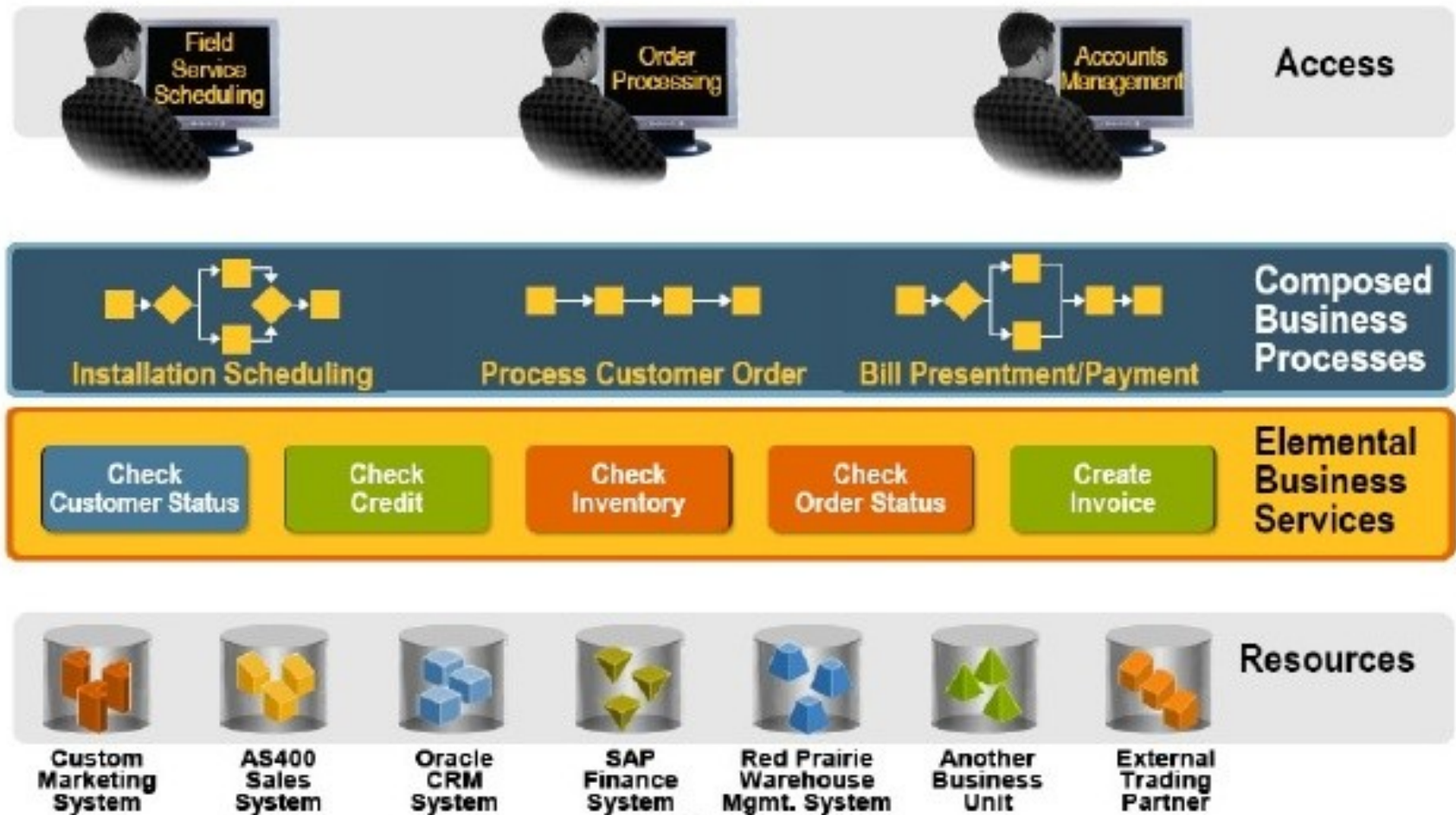
# Siloed Architecture



# Problems with siloed architecture

- Waste of resources
- Complex maintenance
- Lack of data sharing and consistency
- Complexity of IAM (Identity and Access Management)
- Difficult to scale up
- ...

# Microservices Architecture



# (Micro)Service Concept

- Black box performing 1 specific task (business or technical function)
- Can be used via an API (= contract between the customer and the supplier)
- Can call on other services
- Designed to be duplicated → *stateless*:
  - *No application state*
  - *Or client-specific state provided in the request*
  - *Or state on external storage shared with other services*

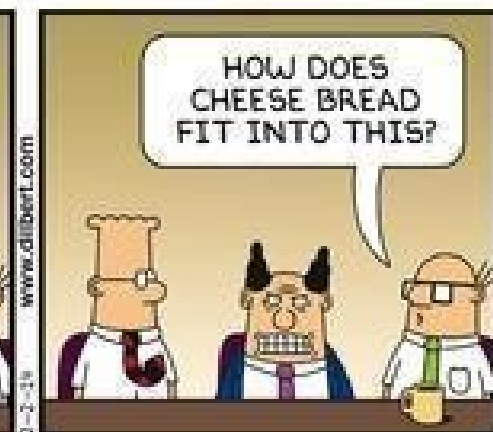
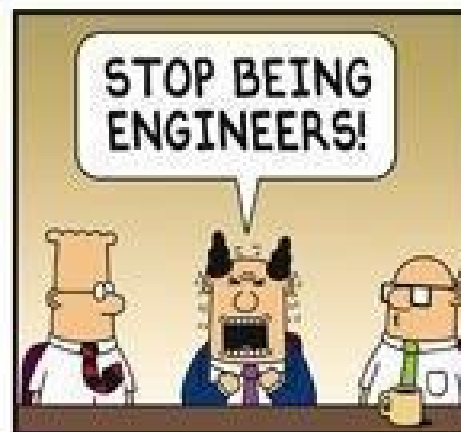
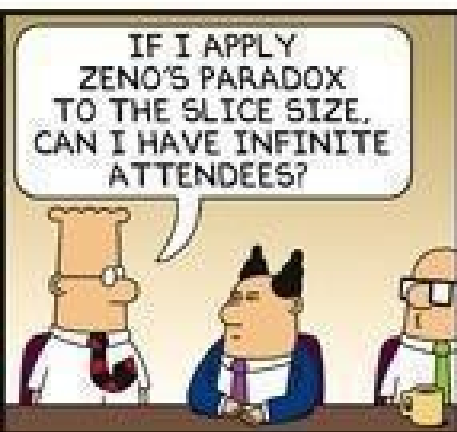


# Advantages of the microservice architecture

- Reuse
- Scaling and fault tolerance thanks to easy duplication
- Fault isolation
- Independent development and deployment
- Ability to use the most appropriate technology for each module
- Small development teams



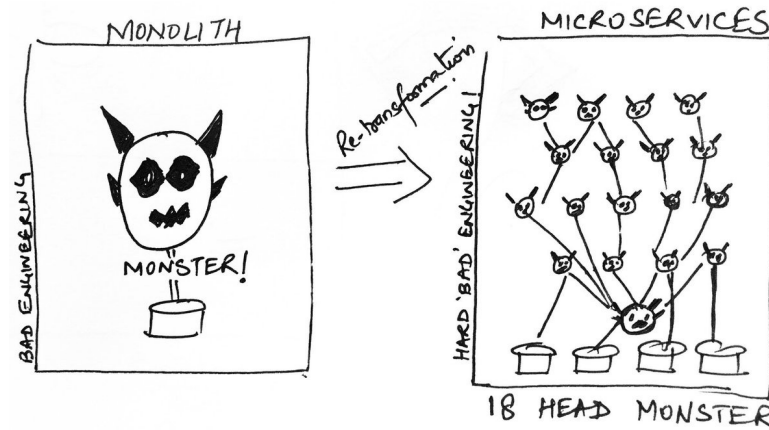
**DILBERT**



**BY SCOTT ADAMS**

# No silver bullet...

- The entropy of the IS increases as well!
- Several examples of strategic retreats on a monolithic solution!
- Microservices do not correct design errors.



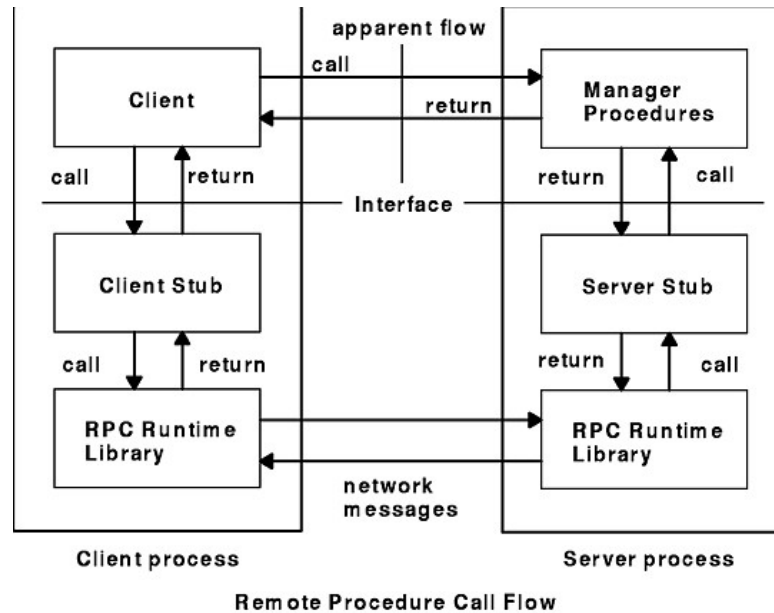
- Intercommunication between services can lead to a higher latency of the application and the network quality becomes crucial



# Middleware

Solutions to ease the connection between services:

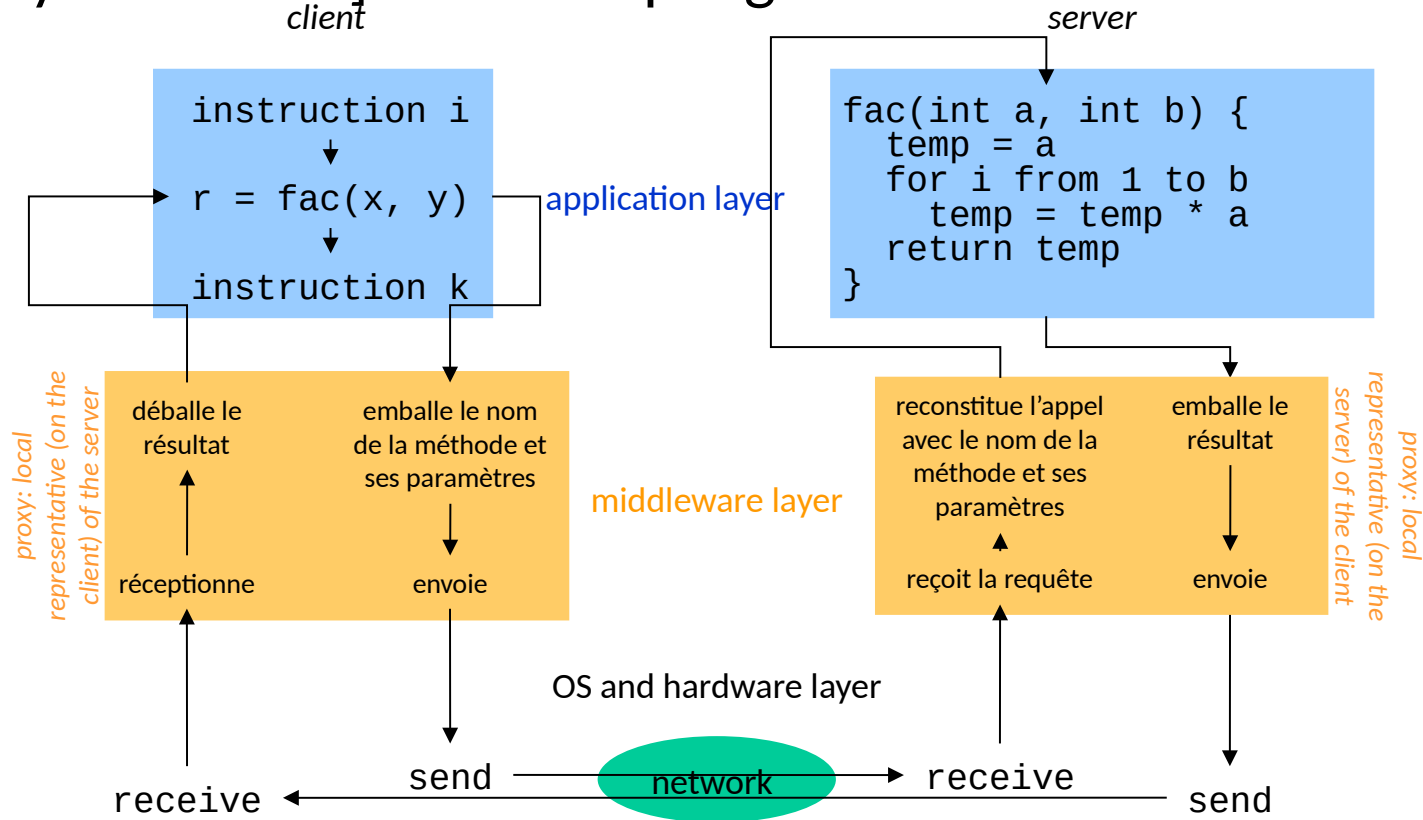
- Locally:
  - Inter-process communication: system, MPI, Unix Domain Socket, etc
- Across the network:
  - Synchronous Remote Procedure Call
  - Asynchronous Messages



# Remote Procedure Call (RPC) and Object Request Broker (ORB)

# RPC

- [asynchronous] loose coupling between client and server



- The proxies handle:
  - network calls
  - format transformations between the client and server

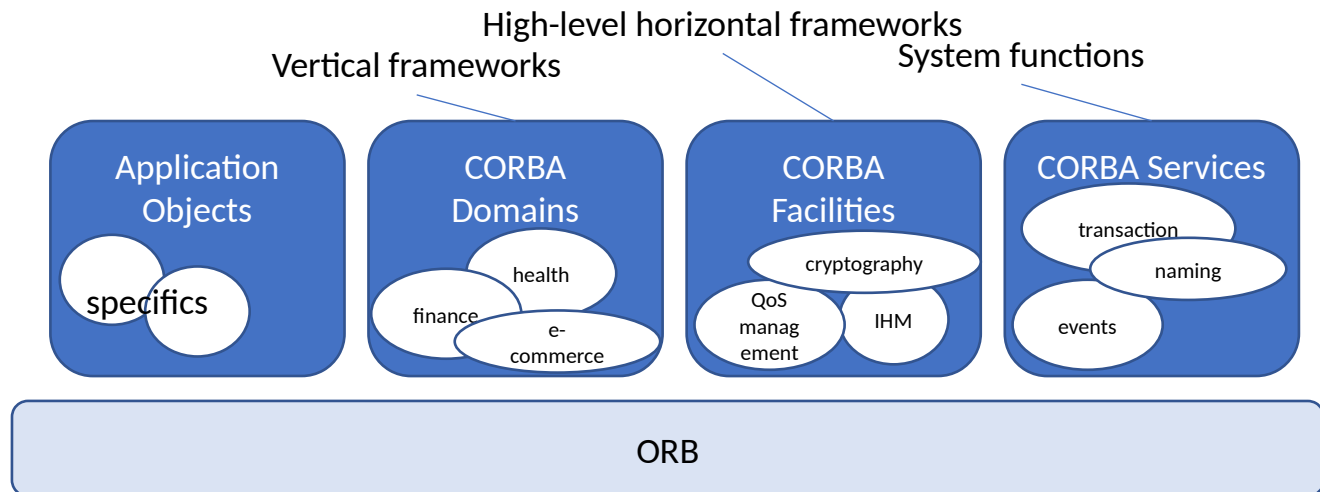
# (some) RPC implementations and frameworks

- Rise:
  - 80's: Sun RPC (as part of NFS protocol): simple, limited to Unix systems
  - 90's: DCE RPC (Open Software Foundation): platform-independent, rich set of functionalities (transactions, encryption...), more complex to use
- Fall:
  - 94: RPC is “fundamentally flawed”: communication latency, partial failures and concurrency issues...
  - Message passing alternatives
- Rise, again: more features, more supported formats/transport...
- 98: XML-RPC: data are XML-formatted and exchanged over HTTP -> SOAP
- 2005: JSON-RPC, lightweight
- 2007: Apache Thrift (init. Facebook): support for multiple serialization format (including binary), support for multiple transport protocols, complete stack for creating clients and servers
- 2009: Avro (Apache Hadoop)
- 2016: gRPC (Google, open source): messages serialized using Protocol Buffers (binary), transported by HTTP/2, multiple features
- 2021: Cap'n Proto (now developed by Cloudflare): performances!



# Object Request Broker

- Object oriented RPC: method calls on remote objects
- Most popular technologies:
  - CORBA (Common Object Request Broker Architecture) (1991)
    - OO-RPC for heterogeneous objects
    - but also a set of services



- DCOM (Distributed Component Object Model) (1995), .Net Remoting
  - Microsoft-equivalent to CORBA
- Java RMI (Remote Method Invocation) (1998)
  - for Java objects

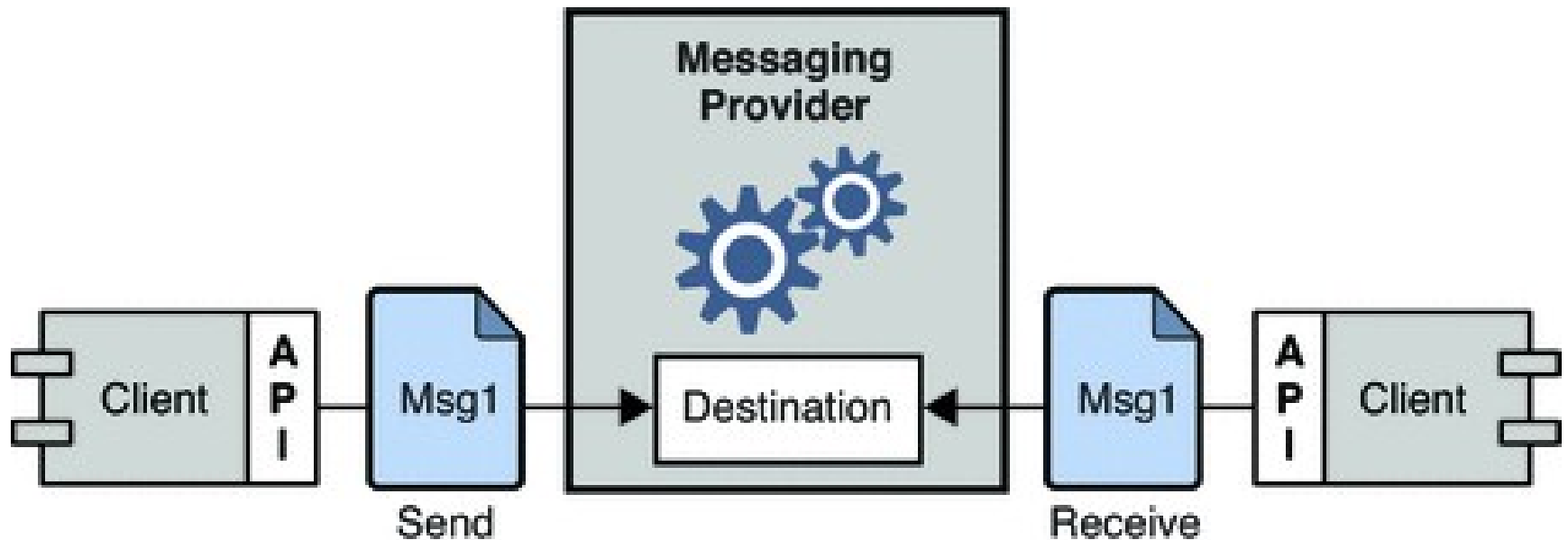
# CORBA perspectives

- Limitations:
  - local calls are treated the same as remote calls → inefficient
  - complex standard
  - difficult to have different versions of a service coexisting
  - fewer and fewer experts
- Why hasn't it disappeared?
  - still important legacy
  - one of the few candidates (with DDS) when there are strong real time constraints

*Alcatel-Lucent network management system, communications between military planes and ESA satellites, air control systems, Siemens electrical power plant management system...*

# Service call

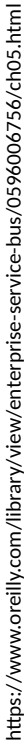
- 1st generation Web Services:
  - Requests and responses transported by SOAP messages, usually on top of HTTP
  - 4 patterns supported by WSDL:
    - Request - response
    - One way request
    - Notification
    - Request - response
  - WS-\*: myriad of specifications to complete the messaging service
- Web service in a REST architecture:
  - URI-addressed resources
  - Requests and responses typically carried over HTTP, exploiting the semantics of HTTP methods



# Message Oriented Middleware

<https://www.oreilly.com/library/view/enterprise-service-bus/0596006756/ch05.html>

- <https://www.oreilly.com/library/view/enterprise-service-bus/0596006756/ch05.html>



# Optional Features

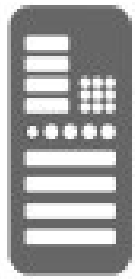
- Strict FIFO (, guaranteed delivery of messages in the right order) or hierarchical organization of messages, priority levels
- Point-to-point: a message read by a destination is no longer available for the others, or Publish-Subscribe : all subscribers to the queue receive a copy of each message (guaranteed delivery: at least once or exactly once)
- message filtering
- encryption/decryption functions, compression/decompression, format transformation
- message retention for offline consumers
- message expiration or validity date
- persistence (on physical media)
- reliability (Ack from MOM to sender and Ack from receiver to MOM)
- transactions
- ...

# Evolution of MOMs

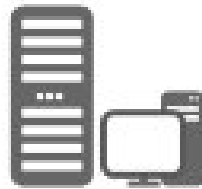
- 95-2010: Earlier versions
  - 1994: IBM MQSeries (now IBM MQ): pioneer commercial MOM
  - 1994: TIBCO Rendezvous: high performance
  - 1996: Microsoft MSMQ, part of Microsoft Windows Server platform
  - 1998: Oracle MQ, now open source
  - 1999: FioranoMQ: HP for trading and finance
  - 2004: Apache ActiveMQ (open-source, java-based)
  - 2007: RabbitMQ (open-source, Erlang-based)
- 2010: Additional features:
  - 2011: Kafka: HA, replicate...
- 2010's: Integration with cloud technologies:
  - 2011: Amazon Simple QS
  - 2015: Google Cloud Pub/Sub
  - 2018: IBM Event Stream (based on Kafka), easily integrates with IBM cloud services
  - 2018: Azure Service Bus
  - 2019: CloudAMQP (based on RabbitMQ): automatic scaling

# Overview

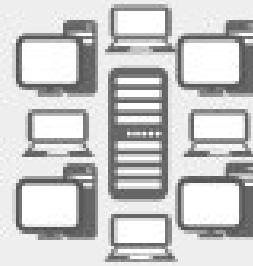
Infrastructure



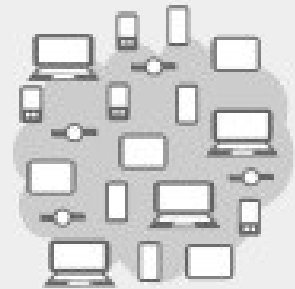
Mainframe



PCs & Servers

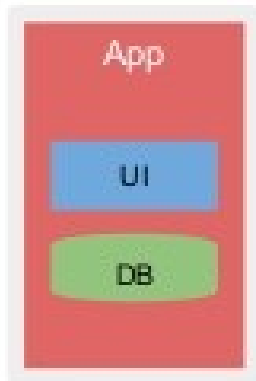


Web



Cloud

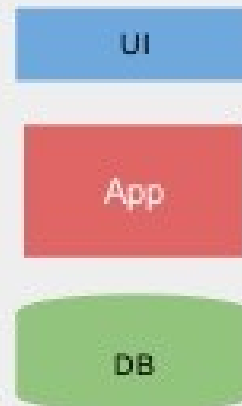
Applications



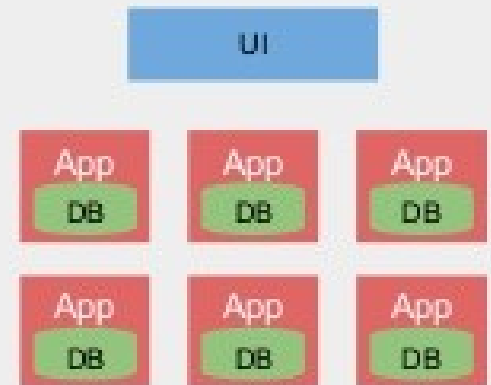
Monolithic



Client Server



N-Tier



Service Oriented



[illegible]

# Generic Definition

- Software function (tool, resource, data...)
- Accessed via the network (remote, deployed, @)
- Offered to other software units (M2M)
- Platform- and language-independent
- Can be described and advertised
- 2 roles:
  - Service requester = client
  - Service provider = server

# What is an API?

- a means of exposing business/enterprise resources via the Internet to external or internal software consumers
  - Well-defined interface: contract
  - Easily accessible by third parties
  - Use of standard protocol(s)

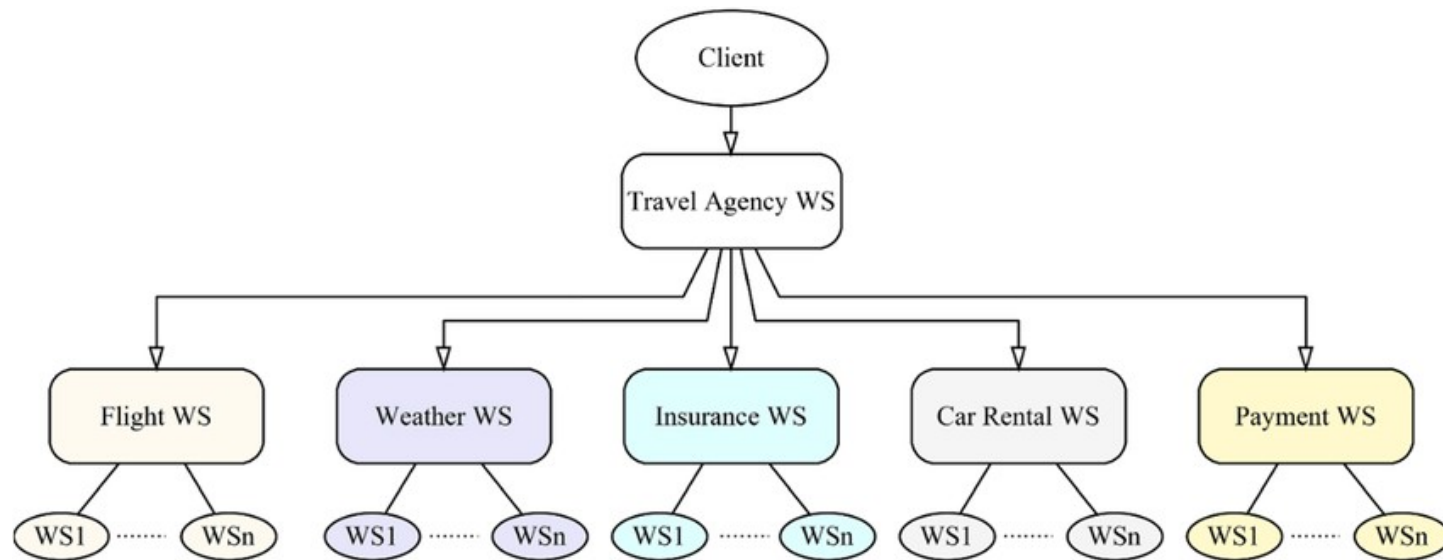
≈ Web Service

# Usage

- Services are used as software libraries to build applications
- 2 contexts:
  - External services
  - Internal services

# External Services (Partner or Public Services)

- open to the partners of the organization (B2B, B2C)

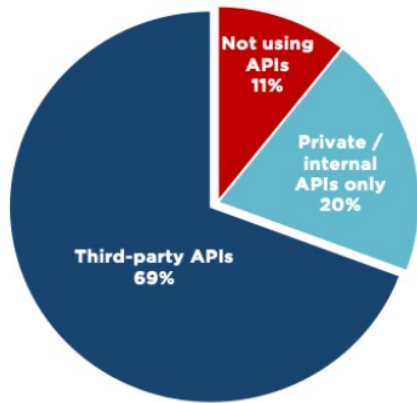


# API becomes more of a priority than UI

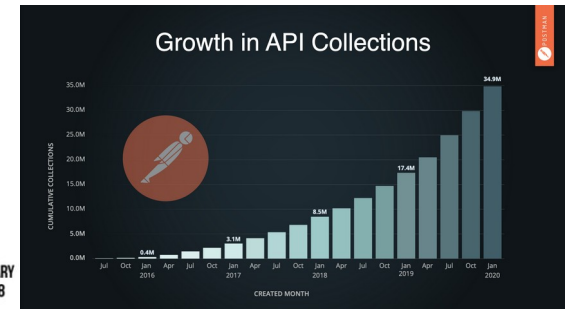
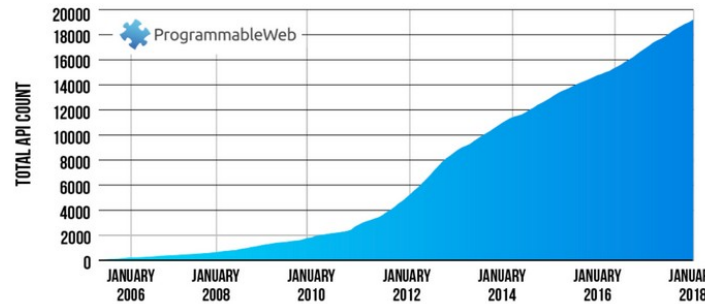
Nearly 90% of developers use APIs

% of developers (Q3 2020 n=15,299)

<https://nordicapis.com/apis-have-taken-over-software-development/>



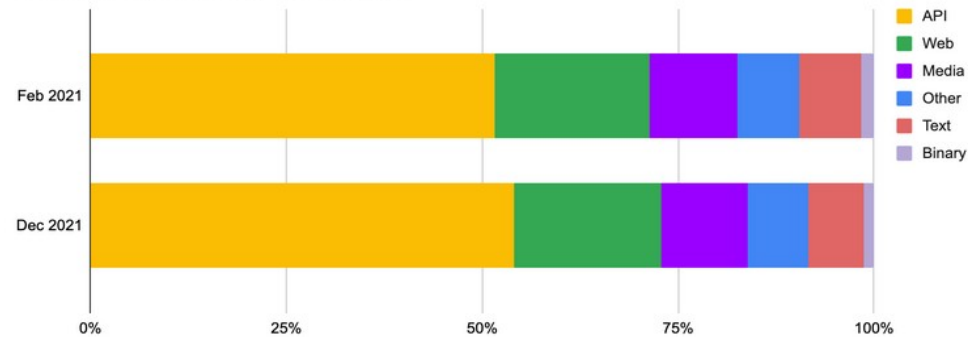
Source: SlashData Developer Economics survey 19th edition / DATA



Cloudflare traffic: API use in 2021

<https://blog.cloudflare.com/landscape-of-api-traffic/>

Traffic composition by content type



Programmatic access is considered at least as vital as human access, if not more so.

<https://www.postman.com/state-of-api/>



gravitee.io

Platform

How it works

Use Cases

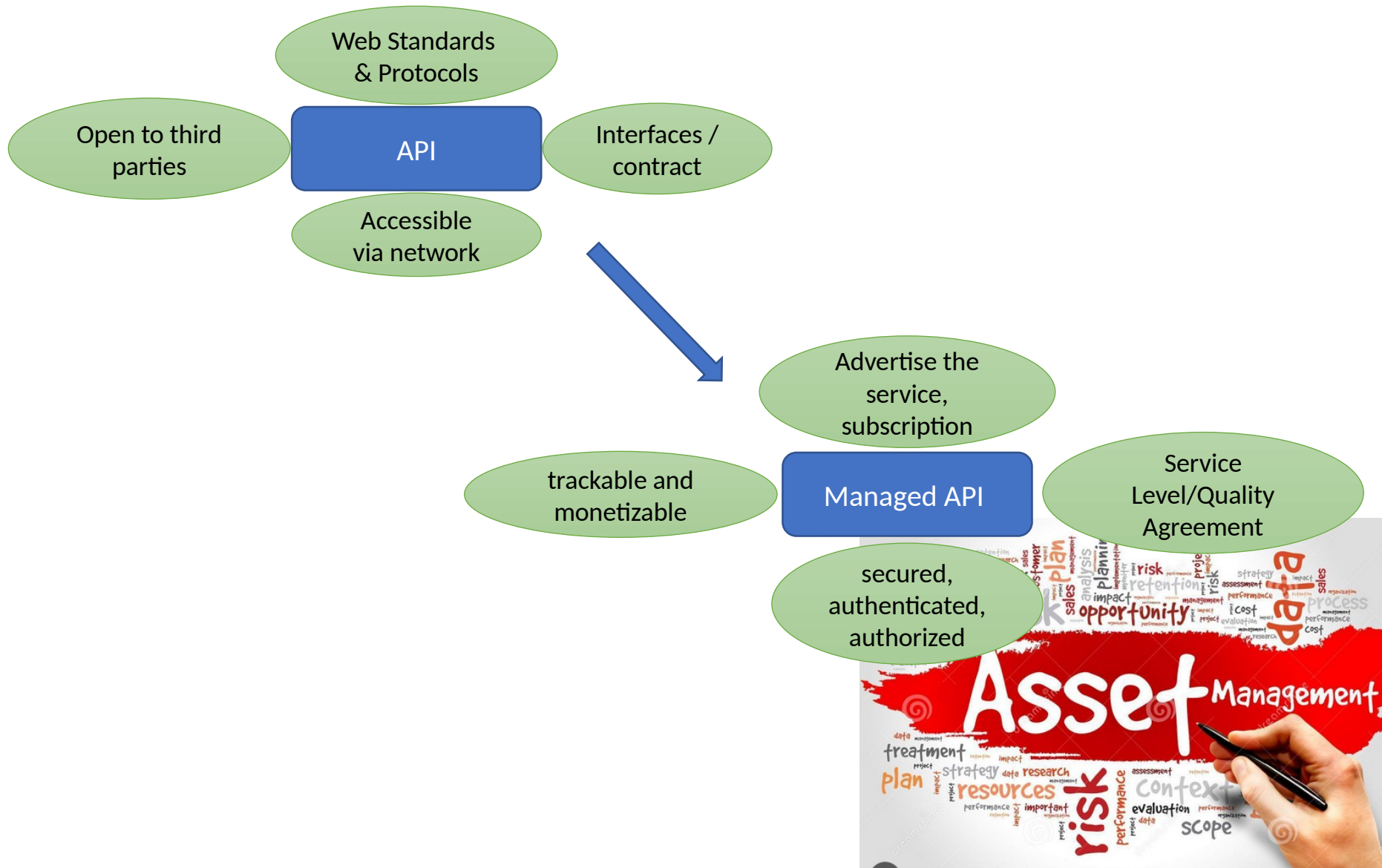
Why Gravitee

Resources

## API-First As The Norm

Adopting an API-first strategy will be increasingly common in the future. In fact, Postman's 2021 found that 39.2% of teams have already designed and defined APIs and schema before they ever

# From Basic WS to Managed API



# Benefits of Web Services from the Client's Point of View

- Take advantage of third-party data or programs without having to:
  - develop, test, update and maintain code
  - acquire and maintain a hosting infrastructure
- Easily compose services and replace one component by an alternative



# Trade-offs for the Client

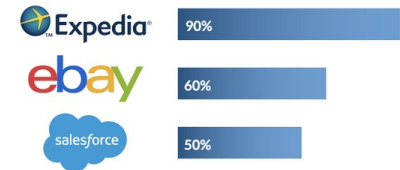
Developers lost control of the services and the services are remote →

- A service might be temporally unavailable
- Performances might become poor
- Data of the client can get lost, divulged, corrupted...
- A service might not longer be maintained
- The service fee might increase
- ...

# From the Service Provider's Point of View

- Benefits
  - Increases revenue
  - Extends customer reach
    - New form of marketing : B2D “business to developer”
  - Stimulates innovation
- Risks
  - Decreases ad revenue
  - No more control on the final user's experience

Percentage of Revenue Generated Through APIs

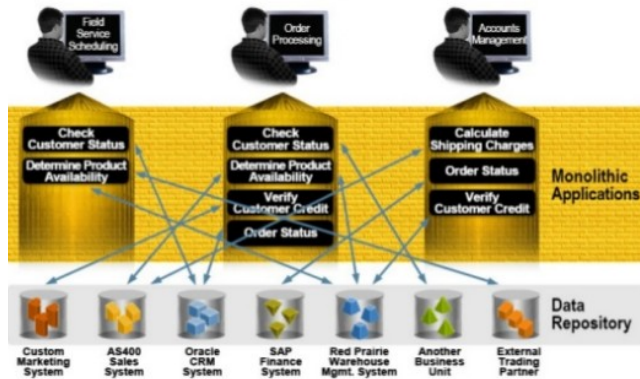


Source: Harvard Business Review, The Strategic Value of APIs, 2015.

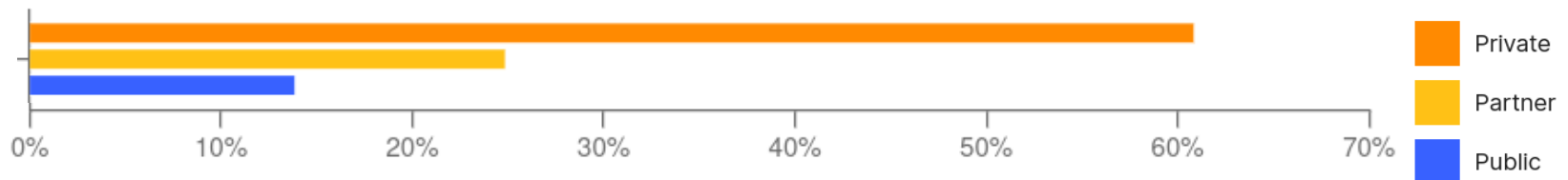
# Internal Services (Private Services)

- access restricted to the organization

Monolithic Systems



Reuse Services via Re-composition



# Internal Services: an injunction!

- Jeff Bezos's mandate (2002)



1. All teams will expose their data and functionality through service interfaces.
2. Teams must communicate with each other through these interfaces.
3. The only communication allowed is via service interface calls over the network.
4. It doesn't matter what technology they use.
5. All service interfaces, without exception, must be designed from the ground up to be externalizable. No exceptions.
6. Anyone who doesn't do this will be fired.
7. Thank you; have a nice day!

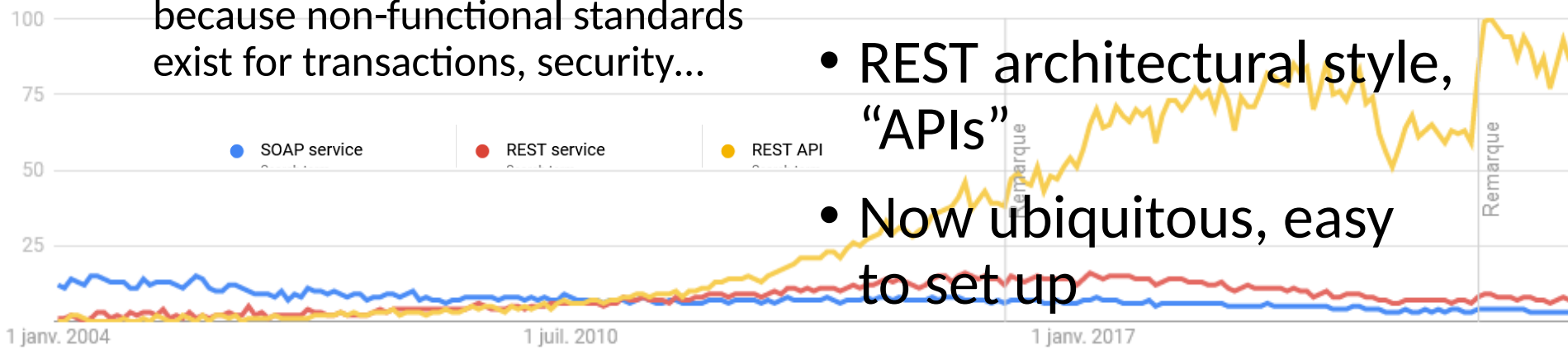
# 2 “Flavors” of Web Services

## Process-Oriented Services

- Distributed Information Systems required middleware, RPC and Object Brokers, were poorly adapted to B2B → use web protocols for transport and XML as IDL and format
- SOAP + WSDL standards
- “first-generation” web services, still used for complex applications because non-functional standards exist for transactions, security...

## Resources-Oriented Services

- From static web pages to dynamic web pages to web applications (UI = web browser, business processes are executed on the server) to web services
- REST architectural style, “APIs”
- Now ubiquitous, easy to set up



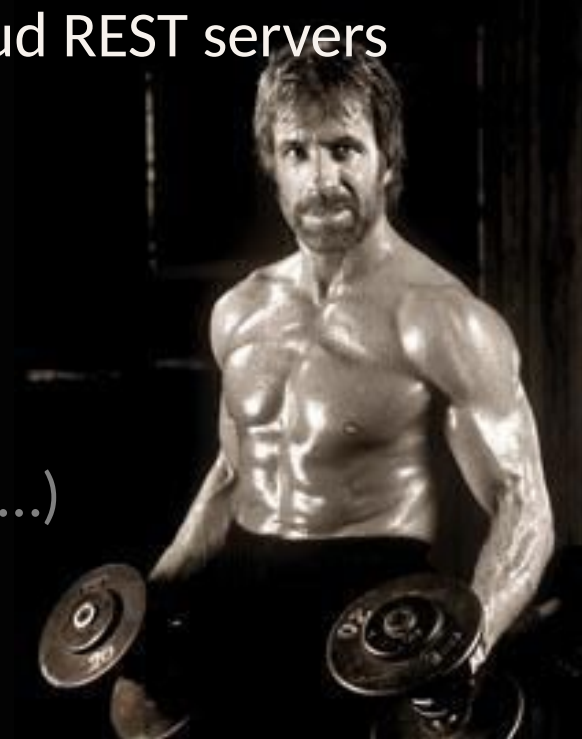
# After this Course...

You will be able to design, set up and take advantage of a Service-Oriented Architecture

- find Web Services and understand their interfaces, including GraphQL
- write well-designed and documented APIs
- implement in Python and deploy on the cloud REST servers
- write Python clients
- cite several Chuck Norris's facts

Not covered:

- SOAP and WSDL
- DevOps (deployment, mock tests, load tests...)
- security
- scripted composition of services

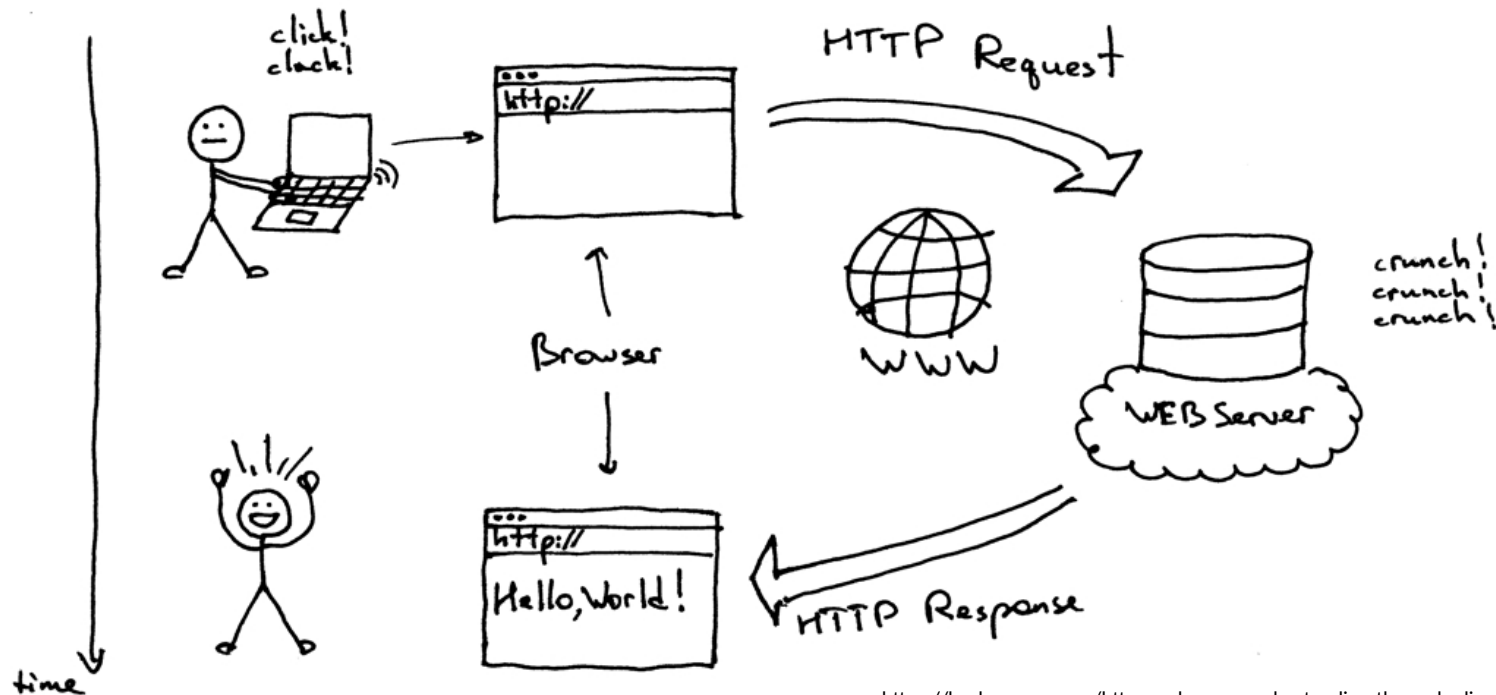


http://www.



# \*The\* Web Services Protocol

- Application-layer protocol
- Client sends service requests using HTTP messages
- Server replies using HTTP messages



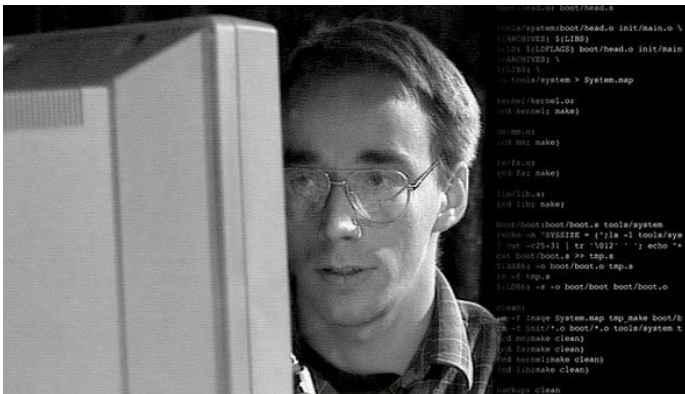


# When was HTTP first specified?



**1981**

(IBM PC 5150)



**1991**

(Linus Torvalds introduces Linux)



**1986**

(Brain: first computer virus for MS-DOS)

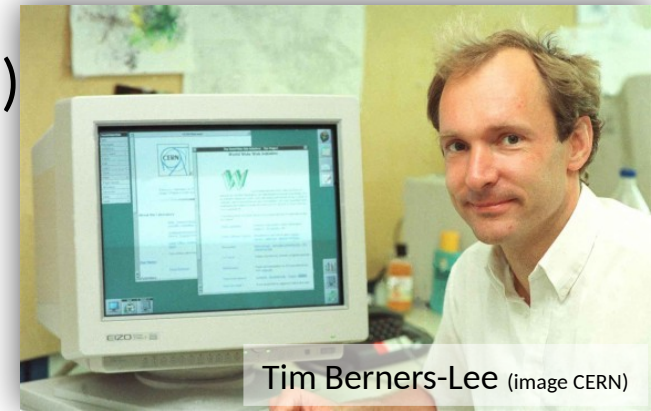


**1996**

(Google search engine)

# Hypertext Transport Protocol

- 1989-90, Tim Berners-Lee's problem at CERN:  
how to integrate and exchange information held on different computers in scattered places?
- Already exist:
  - TCP: reliable transport of information on the Internet
  - DNS: domain name ("www.centralesupelec.fr") ↔ IP @ ("138.195.9.117")
    - Human friendly
    - Computer friendly
  - object in a database that *references* others
- Put them all together: HTTP
  - Retrieve linked documents (resources)
  - Accessible via the Internet



Tim Berners-Lee (image CERN)

# Client-Server Protocol request / response

**Client**

(web browser or other application)

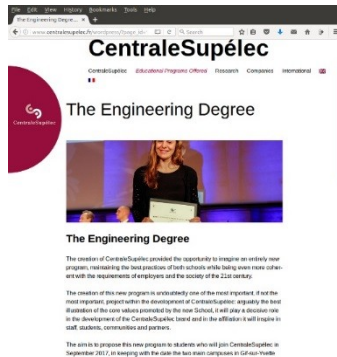


1. user clicks on hyperlink

2. HTTP request message

4. HTTP response message

5. display file



**web server**



back-end  
(database...)

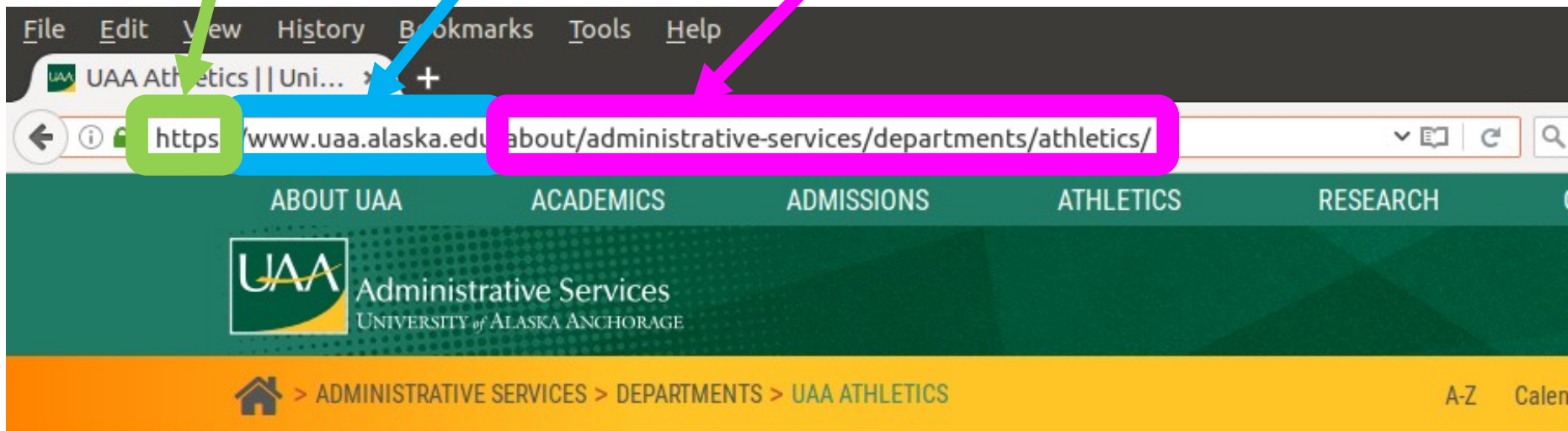


3. create or retrieve file



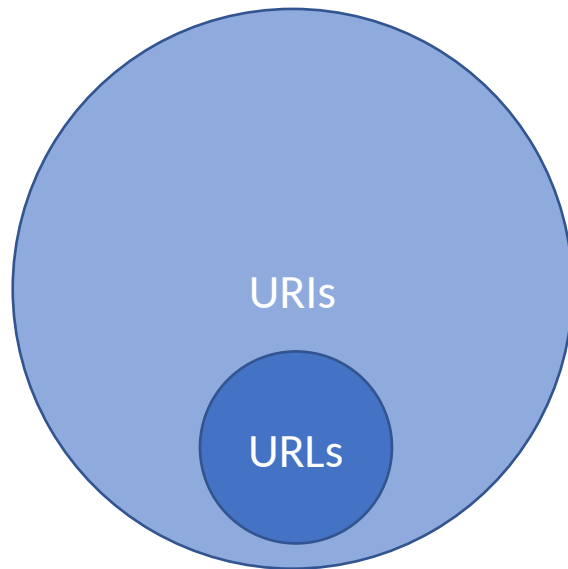
# Resources = addressable files

- Any kind of file: HTML file, JPG image file, binary file...
- URL (Uniform Resource Locator)  
= protocol + server host name + path on server



# Side note about URL and URI

- URI: identifier (name of a restaurant)
- URL: locator (GPS coordinates of the restaurant)



All URLs are URIs:

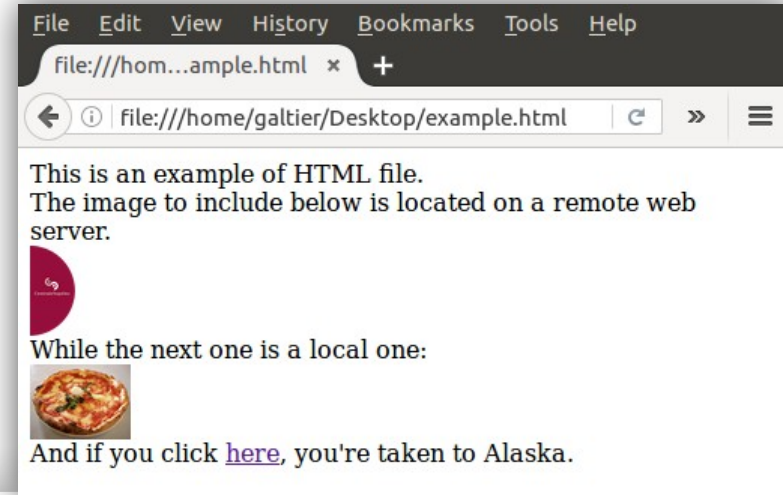
with the GPS coordinates I arrive to the right restaurant

Not all URIs are URLs:

the name of the restaurant gives no information on its location

# HTML file may include references to others resources

- 3 resources are required to display this web page:
  - HTML file
  - CentraleSupelec logo image
  - Pizza image



```

Edit View Search Tools Documents Help
Open [icon] Save

<html>
<head><title="HTML Example"/></head>
<body>
This is an example of HTML file.</br>
The image to include below is located on a remote web server.</br>
</br>
While the next one is a local one:</br>
</br>
And if you click <a href="https://www.uaa.alaska.edu/about/administrative-services/departments/athletics/">here</a>, you're
taken to Alaska.
</body>
</html>
```



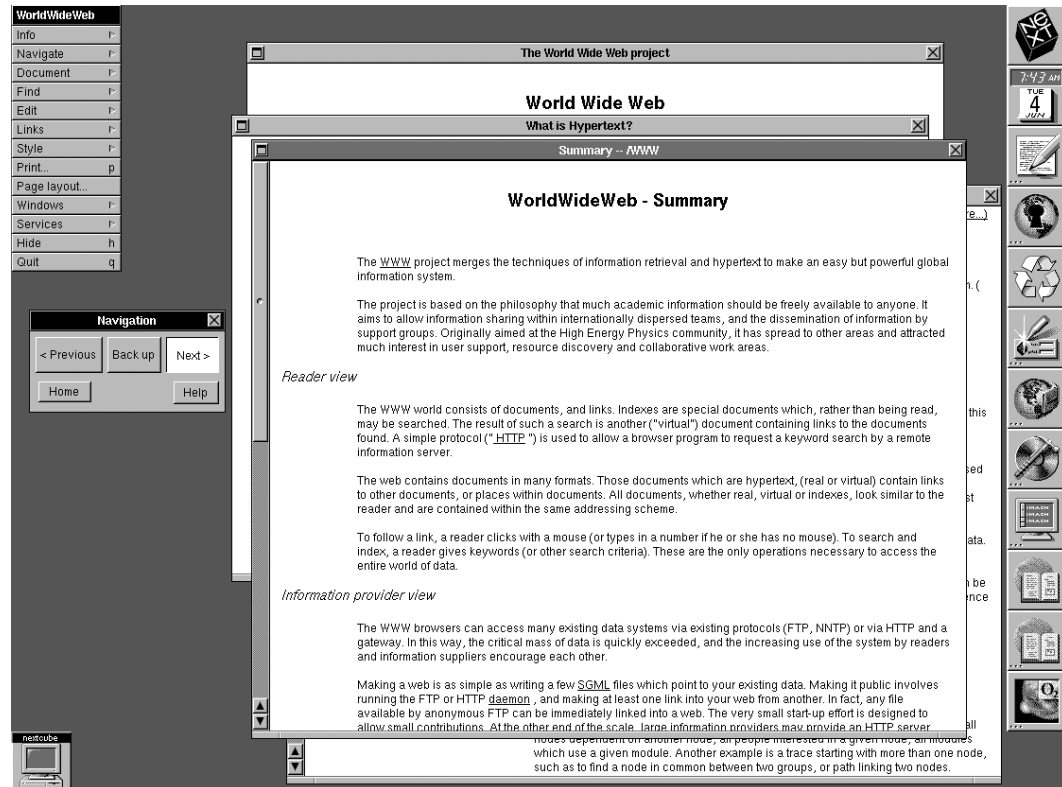
# HTTP Versions

HTTP/0.9

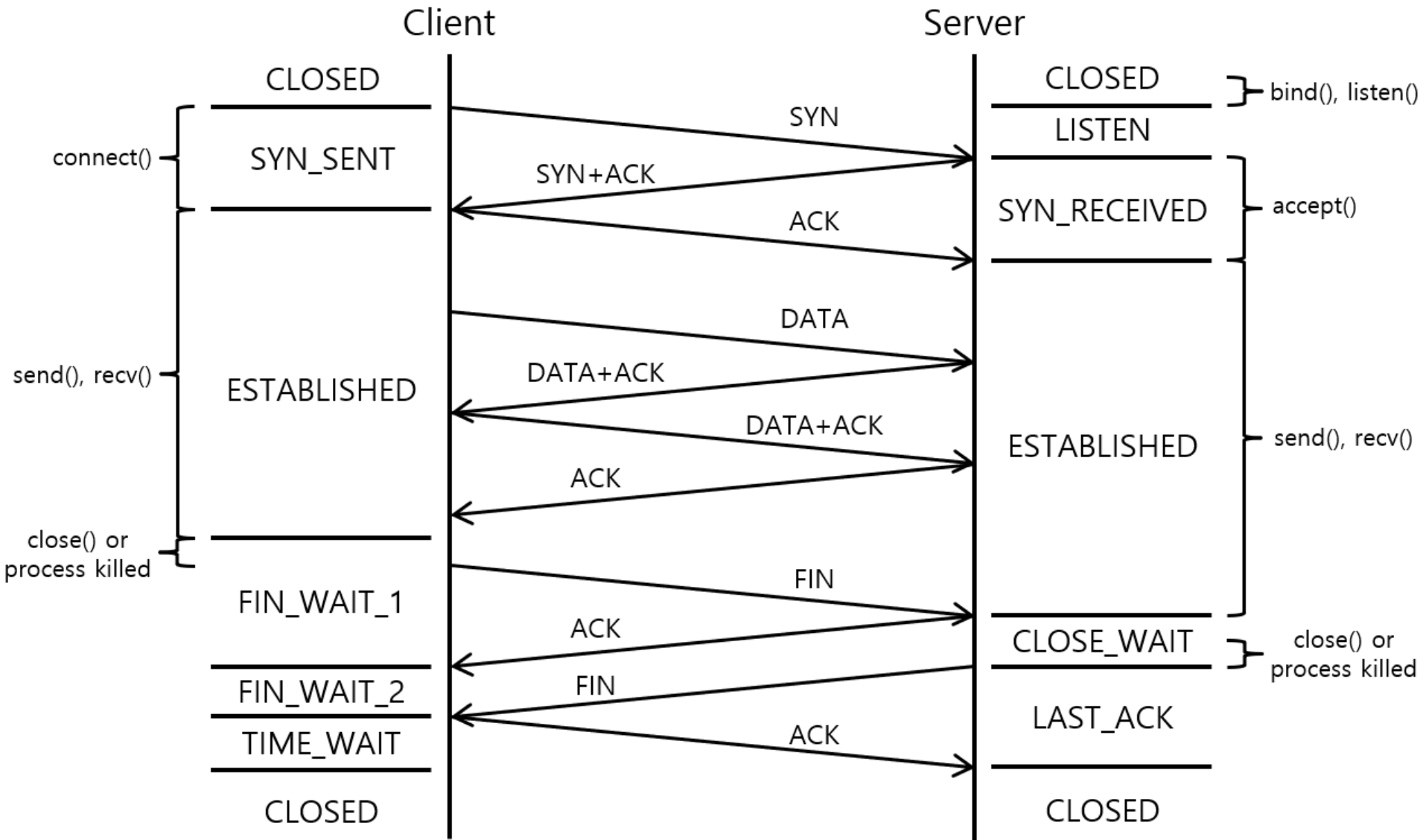


1991

- 1991 – v0.9
  - First documented version
  - First web browser

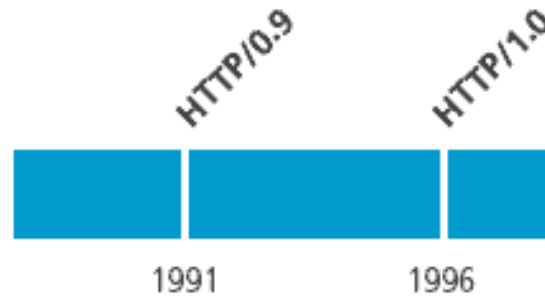


# Side note on TCP 3-way handshake

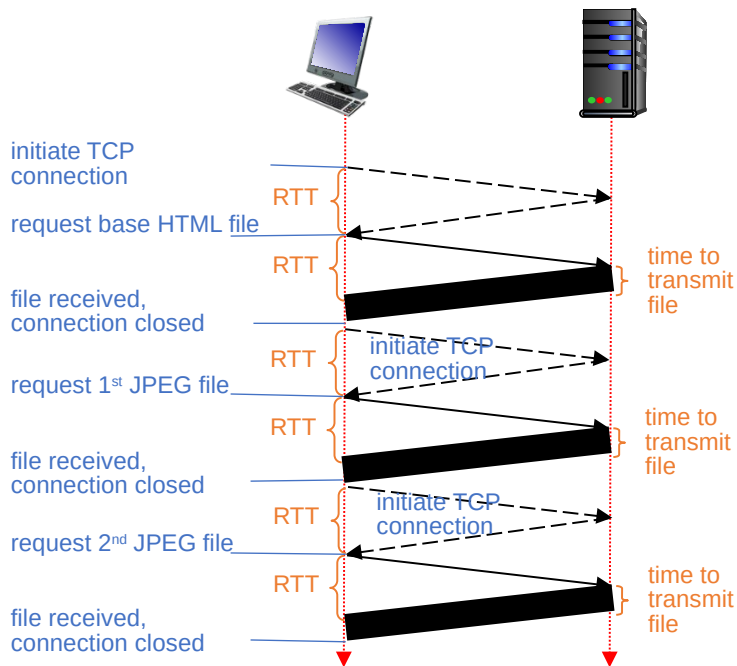




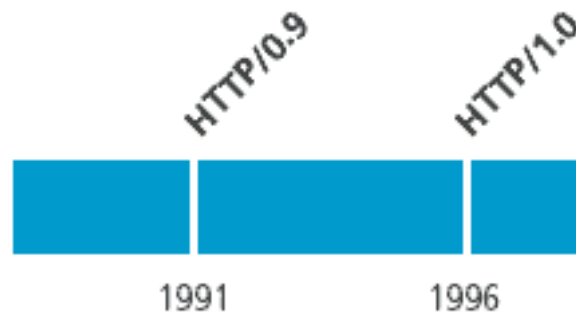
# HTTP Versions



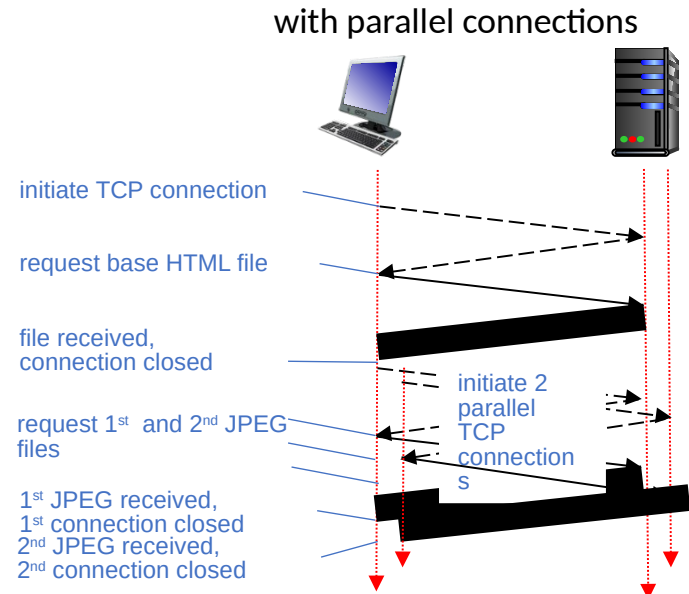
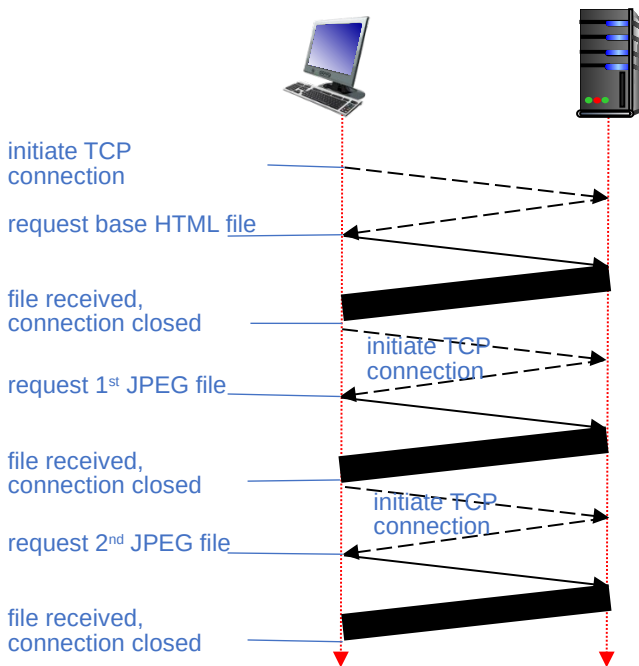
- 1996 – v1.0
  - One TCP connection per resource



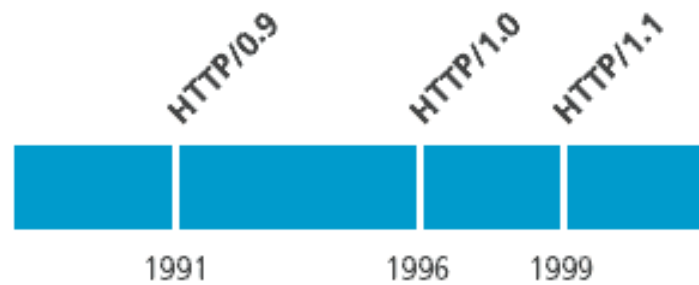
# HTTP Versions



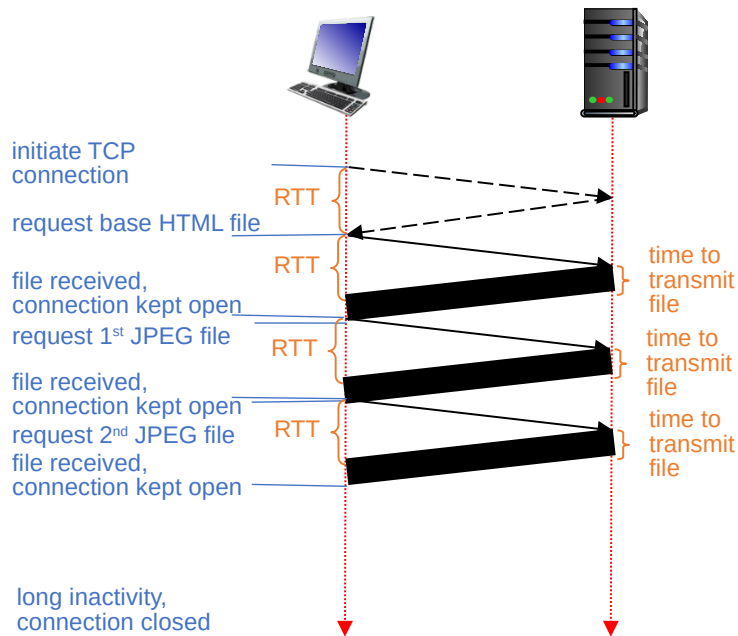
- 1996 – v1.0
  - One TCP connection per resource



# HTTP Versions



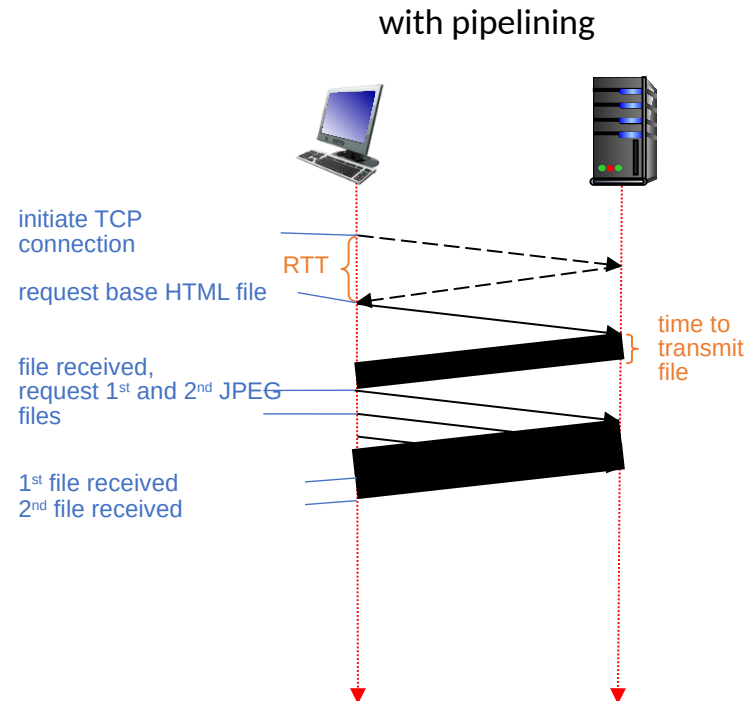
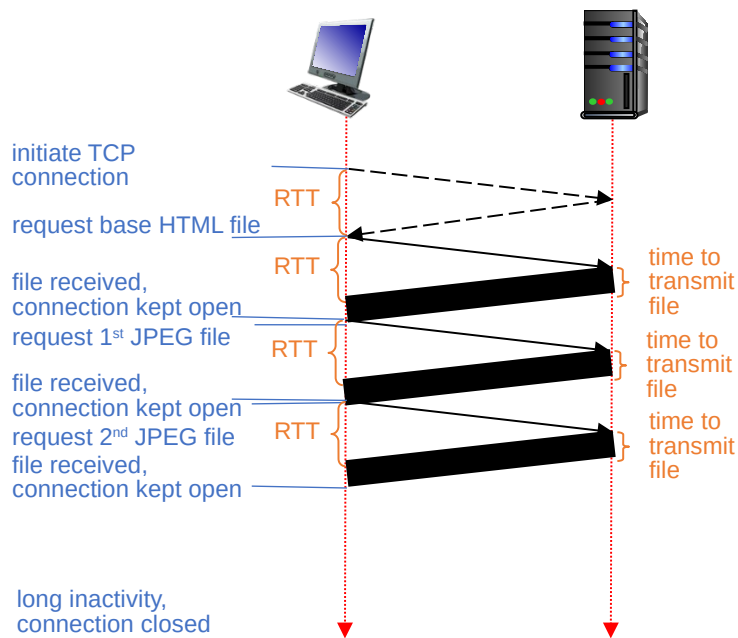
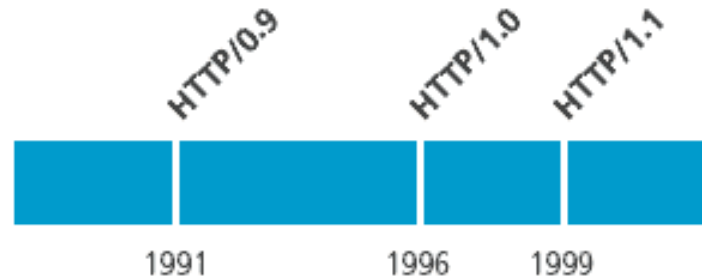
- 1999 – v1.1
  - Persistent connection



```
GET https://www.centralesupelec.fr/ HTTP/1.1
Host: www.centralesupelec.fr
User-Agent: Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:109.0) Gecko/20100101 Firefox/109.0
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate, br
Connection: keep-alive
Upgrade-Insecure-Requests: 1
```

# HTTP Versions

- 1999 – v1.1
  - Persistent connection

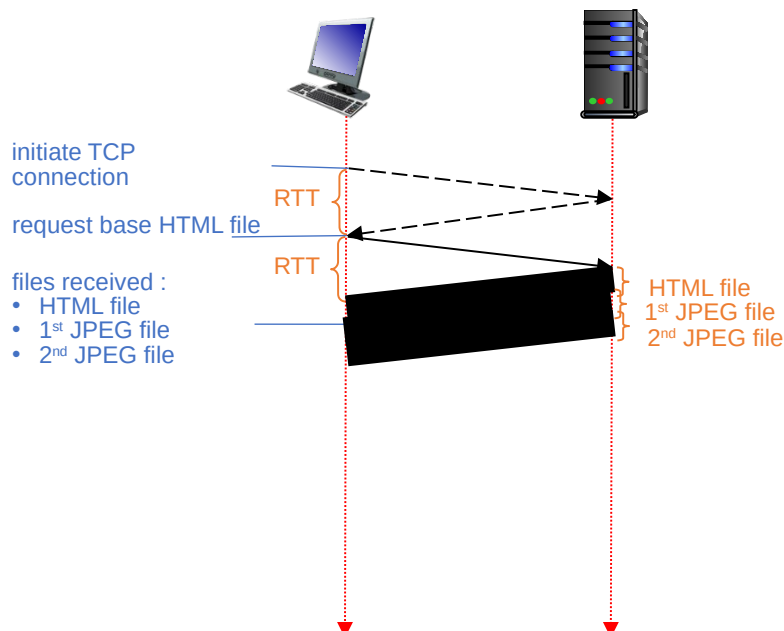


# HTTP Versions



- 2015 – v2
  - Server “pushes” content
  - [and other optimizations]

The screenshot shows a web browser window with the Akamai logo and the headline "HTTP/2 is the future of the Web, and it is here!". Below this, it says "Your browser supports HTTP/2!". A paragraph explains that this is a demo of HTTP/2's impact on downloading many small tiles for the Akamai Spinning Globe. Two side-by-side images of the globe are shown, each with a list of metrics: HTTP/1.1 (Latency: 27ms, Load time: 3.23s) and HTTP/2 (Latency: 27ms, Load time: 1.17s). At the bottom, it says "Demo content inspired by Galaxie".



# Optional Reading Exercise

- Find the document which describes HTTP/2.
- What is the “head-of-line blocking” (HOL blocking) problem observed in HTTP/1.1?
- Read the beginning of the FAQ at <https://http2.github.io/faq/>

# Reading: Results

- HTTP/2 is defined in RFC 7540.
- HOL blocking:
  - Imagine a HTTP client that sends to a server 2 requests over the same TCP connection, and that the first response is "large" in content length while the second response is "small" in content length.
  - Due to the nature of the HTTP 1.x protocol, the second response must wait for the first response to complete: the second response is *head-of-line blocked* by the first response.
  - HTTP/2 is fully multiplexed (instead of ordered and blocking), allowing multiple request and response messages to be in flight at the same time (it's even possible to intermingle parts of one message with another on the wire).

# HTTP Messages

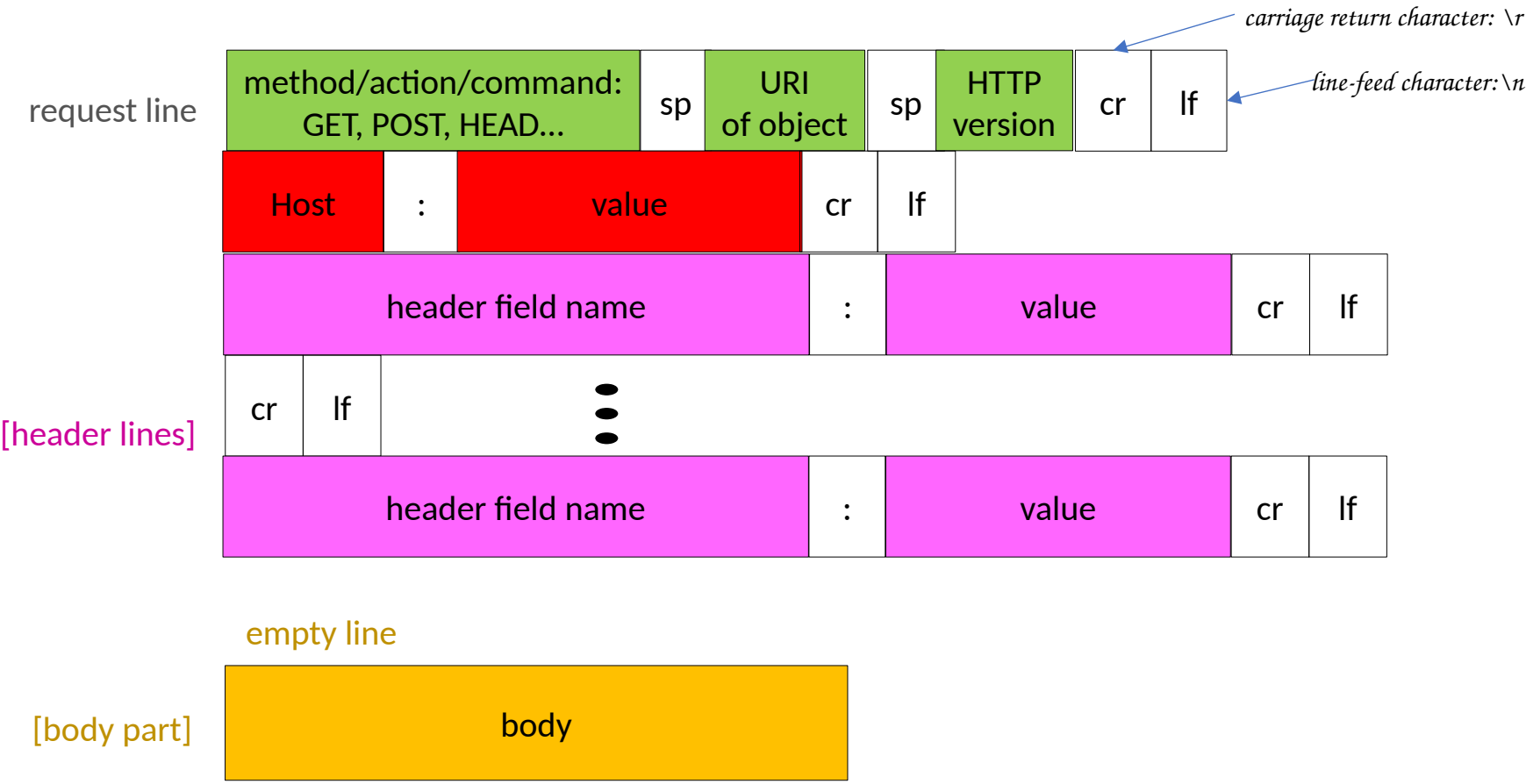
- 2 kinds of messages
  - Request
  - Response
- In ASCII (HTTP 1.x)



# HTTP Requests Commands

- GET
  - retrieves an object
  - no request body
- HEAD
  - same response as GET but empty response body (used to test the access to or the "freshness" of the object without actually downloading it)
- POST
  - results in the creation of a new resource on the server
  - usual request: contains data
  - usual response: URL of the created resource
- PUT
  - updates an existing resource
  - request usually contains data
- DELETE
  - deletes a resource

# HTTP Request Format



# HTTP GET Request Example

GET /node/44 HTTP/1.1\r\n

Host: mapi.centralesupelec.fr\r\n

User-Agent: Mozilla/5.0 (X11; Ubuntu; Linux x86\_64; rv:50.0) Gecko/20100101 Firefox/50.0\r\n

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

Accept-Language: en-US,en;q=0.5\r\n

Accept-Encoding: gzip, deflate\r\n

Connection: keep-alive\r\n

\r\n

# HTTP POST Request Example

POST /post.php HTTP/1.1\r\n

Host: posttestserver.com\r\n

User-Agent: Mozilla/5.0 (X11; Ubuntu; Linux x86\_64; rv:50.0) Gecko/20100101 Firefox/50.0\r\n

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

Accept-Language: en-US,en;q=0.5\r\n

Accept-Encoding: gzip, deflate\r\n

Content-Type: text/xml\r\n

Content-Length: 27\r\n

Connection: keep-alive\r\n

\r\n

firstname=John\r\nlastname=Doe

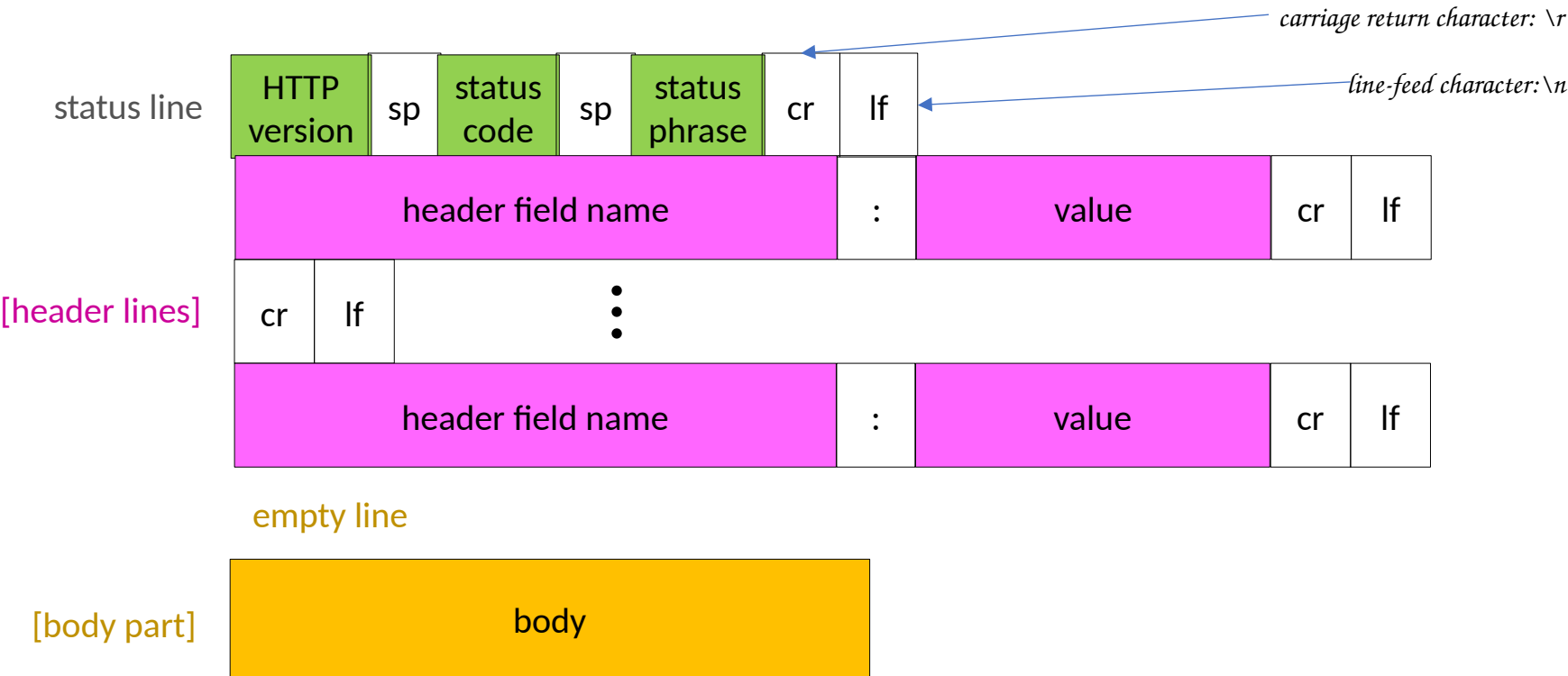
# Request Parameters

3 symbols to add parameters to an URL:

- **?** concatenates the URL and the string of parameters
- **&** separates multiple parameters
- **=** assigns a value to a parameter

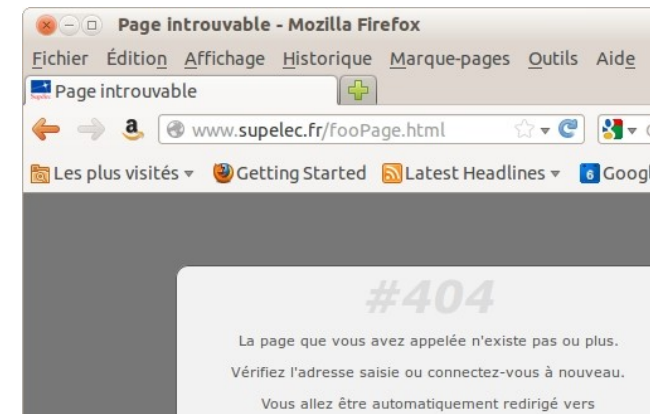
GET /products?priceMin=10&priceMax=40

# HTTP Response Format



# Status Codes

- 2xx: success
  - 200 OK
- 3xx: further action required
  - 301 Moved Permanently: the new URL is specified in a header field
- 4xx: client error
  - 400 Bad Request: badly formulated query
  - 404 Not Found: object does not exist on the server
- 5xx: server-side error
  - 505 HTTP Version Not Supported



# HTTP Response Example

HTTP/1.1 200 OK

Date: Wed, 01 Feb 2017 12:48:22 GMT

Server: Apache/2.4.10 (Debian)

[...]

Content-language: fr

Content-Encoding: gzip

Content-Length: 4740

Keep-Alive: timeout=5, max=100

Connection: Keep-Alive

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

.....;r.8...W...-{(.(KO...!K..-.\$..q;.(...D.4..T.v.\.....q.{...Z.2\_....b.....(l....Df"..Hn...|....}.WQ..m....  
WD.sR)..J.....L:9.C..MC...X.I...  
J..(...'".....J. D....d%bN,. \$..Y.....z.....y(.MS....#.qV.....>.9.j.0  
s&...v.M...'').....m8..<=.i..%B.....S.x}.J.:V..{."..HM..4b..!YJ.....X{i...l.;.T.X}....N.r .<d...#.....S..  
..#Oa. ...V..EPj..G...A..D.K...Z1..c.h,b.4..b.3...l.6..La..>.L8#l.U.\.....2..y!...S,,.....%.....>..ID...  
N..^



# HTTP Response Example

HTTP/1.1 404 Not Found

Date: Wed, 01 Feb 2017 13:14:55 GMT

Server: Apache/2.4.10 (Debian)

[...]

Content-language: fr

Keep-Alive: timeout=5, max=100

Connection: Keep-Alive

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

305d

<!DOCTYPE html>

<html lang="fr" dir="ltr" prefix="content: http://purl.org/rss/1.0/modules/content/ dc: ht

<div id="block-zircon-content" class="block block-system block-system-main-block">

La page demand..e n'a pas pu ..tre trouv..e.

</div>

# Optional Reading

- What is HTTP error code 418?



# Hyper Text Coffee Pot Control Protocol

Article [Talk](#)

[Read](#) [Edit](#) [View history](#)

From Wikipedia, the free encyclopedia

The **Hyper Text Coffee Pot Control Protocol** (**HTCPCP**) is a facetious [communication protocol](#) for controlling, monitoring, and diagnosing [coffee pots](#). It is specified in [RFC 2324](#), published on 1 April 1998 as an [April Fools' Day RFC](#),<sup>[2]</sup> as part of an [April Fools prank](#).<sup>[3]</sup> An extension, HTCPCP-TEA, was published as RFC 7168 on 1 April 2014<sup>[4]</sup> to support brewing teas, which is also an April Fools' Day RFC.

## Protocol [\[edit\]](#)

RFC 2324 was written by [Larry Masinter](#), who describes it as a satire, saying "This has a serious purpose – it identifies many of the ways in which [HTTP](#) has been extended inappropriately."<sup>[5]</sup> The wording of the protocol made it clear that it was not entirely serious; for example, it notes that "there is a strong, dark, rich requirement for a protocol designed [espressoly](#) *[sic]* for the brewing of coffee".

Despite the joking nature of its origins, or perhaps because of it, the protocol has remained as a minor presence online. The editor [Emacs](#) includes a fully functional client side implementation of it,<sup>[6]</sup> and a number of bug reports exist complaining about [Mozilla](#)'s lack of support for the protocol.<sup>[7]</sup> Ten years after the publication of HTCPCP, the *Web-Controlled Coffee Consortium* (*WC3*) published a first draft of "HTCPCP Vocabulary in [RDF](#)"<sup>[8]</sup> in parody of the [World Wide Web Consortium](#)'s (W3C) "HTTP Vocabulary in RDF".<sup>[9]</sup>

On April 1, 2014, RFC 7168 extended HTCPCP to fully handle teapots.<sup>[4]</sup>

## Commands and replies [\[edit\]](#)

HTCPCP is an extension of [HTTP](#). HTCPCP requests are identified with the [Uniform Resource Identifier](#) (URI) scheme `coffee` (or the corresponding word in any other of the 29 listed languages) and contain several additions to the HTTP methods:

<b>BREW</b> or <b>POST</b>	Causes the HTCPCP server to brew <a href="#">coffee</a> . Using POST for this purpose is deprecated. A new <a href="#">HTTP request header field</a> "Accept-Additions" is proposed, supporting optional additions including Cream, Whole-milk, Vanilla, Raspberry, Whisky, Aquavit, etc.
<b>GET</b>	"Retrieves" coffee from the HTCPCP server.
<b>PROPFIND</b>	Returns <a href="#">metadata</a> about the coffee.
<b>WHEN</b>	<a href="#">Says "when"</a> , causing the HTCPCP server to stop pouring <a href="#">milk</a> into the coffee (if applicable).

It also defines two [error responses](#):

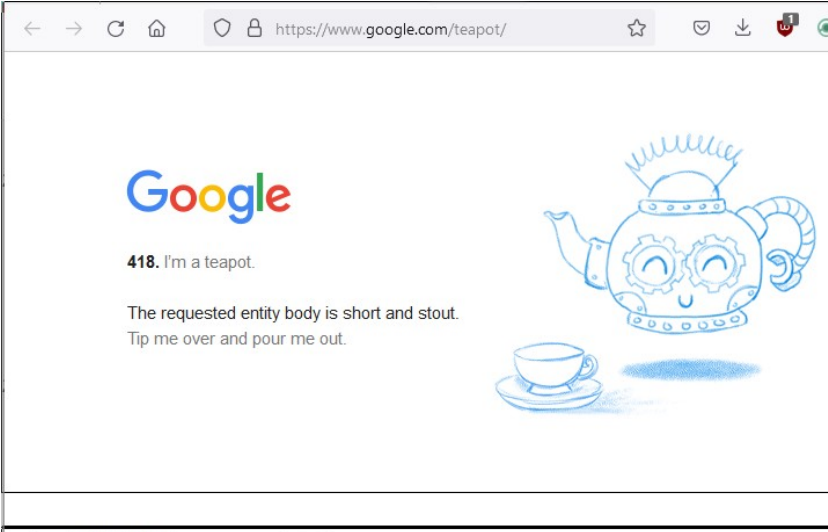
<b>406 Not Acceptable</b>	The HTCPCP server is unable to provide the requested addition for some reason; the response should indicate a list of available additions. The RFC observes that "In practice, most automated coffee pots cannot currently provide additions."
<b>418 I'm a teapot</b>	The HTCPCP server is a <a href="#">teapot</a> ; the resulting entity body "may be short and stout" (a reference to the song " <a href="#">I'm a Little Teapot</a> "). Demonstrations of this behaviour exist. <sup>[1][10]</sup>

### Hyper Text Coffee Pot Control Protocol



[Back-end infrastructure](#) of [error418.net](#), which implements HTCPCP

<b>International standard</b>	<a href="#">Internet Engineering Task Force</a>
<b>Developed by</b>	<a href="#">Larry Masinter</a>
<b>Introduced</b>	April 1, 1998
<b>Website</b>	<a href="#">rfc2324</a>



# Optional Lab Exercise

- Use putty or telnet to connect to port 80 of a web server (<http://www.columbia.edu> for instance) and issue HTTP/1.x requests (get /~fdc/sample.html). Observe the responses.

```
telnet serverName 80
```

- For HTTPS, use:

```
openssl s_client -connect  
serverName:443
```

(note: this exercise is limited to HTTP/1.x because HTTP/2 is no longer textual but uses binary format commands)

# Lab: Results

```
telnet www.columbia.edu 80
```

```
Trying 128.59.105.24...
```

```
Connected to source.failover.cc.columbia.edu.
```

```
Escape character is '^['.
```

```
HEAD /~fdc/sample.html HTTP/1.1
```

```
Host: www.columbia.edu
```

opens a TCP connection on web server port 80 and sends everything that is typed

typed out request

```
HTTP/1.1 200 OK
```

```
Date: Sun, 19 Feb 2023 09:15:26 GMT
```

```
Server: Apache
```

```
Last-Modified: Fri, 17 Sep 2021 19:26:14 GMT
```

```
Accept-Ranges: bytes
```

```
Content-Length: 34974
```

```
Vary: Accept-Encoding,User-Agent
```

```
Content-Type: text/html
```

```
Set-Cookie: BIGipServer~CUIT~www.columbia.edu-80-pool=1764244352
```

received response

# Lab: Results

```
openssl s_client -connect edition.cnn.com:443
```

```
CONNECTED(00000003)
```

```
...CERTIFICATE STUFF...
```

```
---
```

```
GET /travel HTTP/1.1
```

```
Host: edition.cnn.com
```

```
HTTP/1.1 200 OK
```

```
Connection: keep-alive
```

```
Content-Length: 220918
```

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

```
cache-control: max-age=60
```

```
Date: Sun, 19 Feb 2023 09:22:49 GMT
```

```
[...]
```

```
<!doctype html><html lang="en"><head><meta http-equiv="x-ua-compatible"
content="ie=edge"/><title data-rh="true">CNN Travel | Global Destinations, Tips
& Video</title><meta data-rh="true" name="theme-color"
content="#31315b"/><meta data-rh="true" charSet="utf-8"/><meta data-rh="true"
```

opens an SSL connection on web server port 443 and sends everything that is typed

typed out request

received response

# Lab: Results

```
openssl s_client -connect edition.cnn.com:443
```

```
CONNECTED(00000003)
```

```
...CERTIFICATE STUFF...
```

```
---
```

```
GET /travel HTTP/1.1  
host edition.cnn.com
```

```
HTTP/1.1 400 Bad Request  
Connection: close  
Content-Length: 11  
content-type: text/plain; charset=utf-8  
x-served-by: cache-cdg20763
```

```
Bad Requestclosed
```

typed out request  
(correct syntax is  
Host: edition.cnn.com)

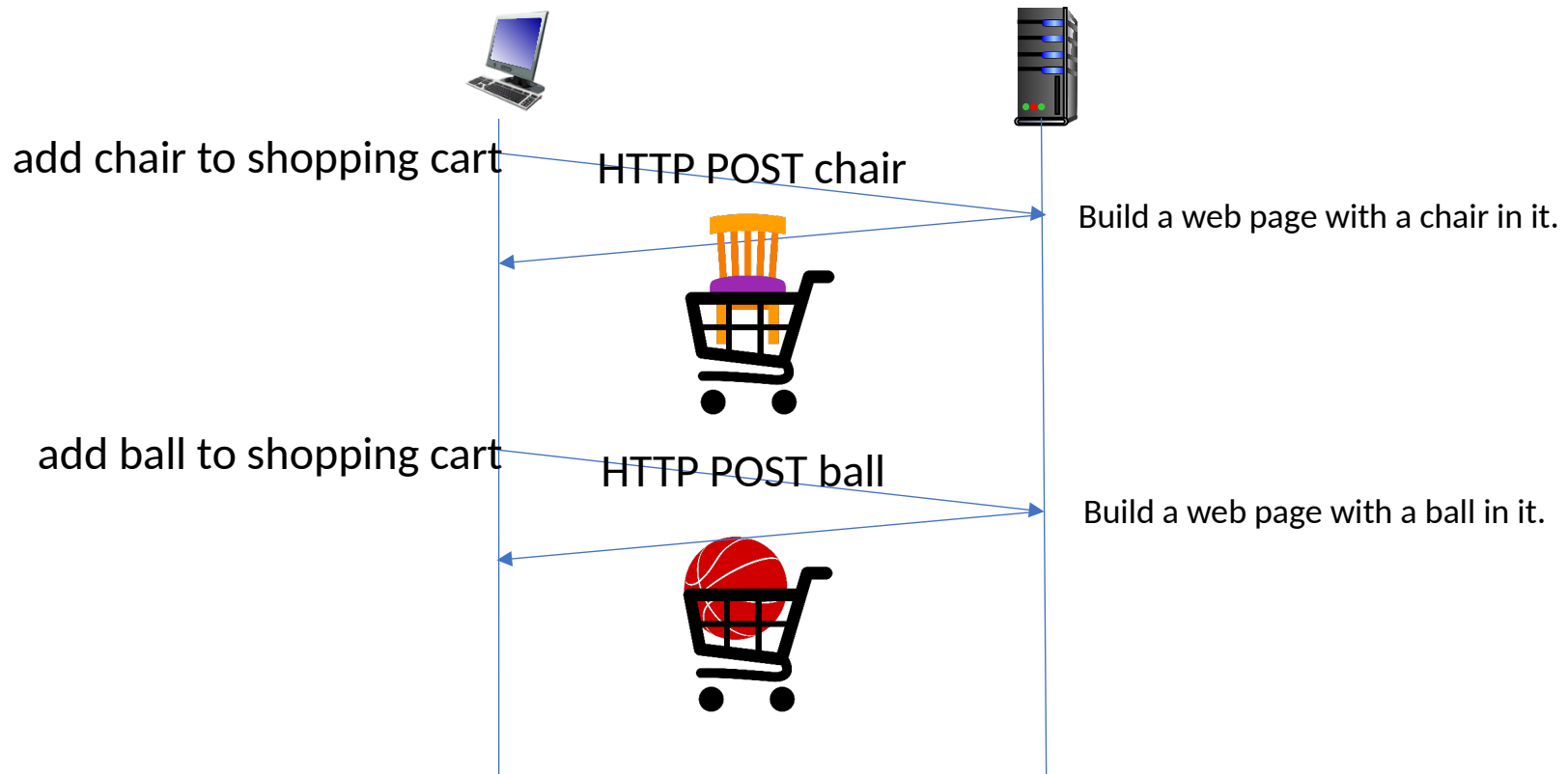
received response

# HTTP Server is Stateless

- A stateless protocol does not require the server to retain information or status about each user for the duration of multiple requests.
- Successive requests from a given client to a server are not treated as a chain but rather as separate requests, independent from the previous ones.



# What we get is not what we want.



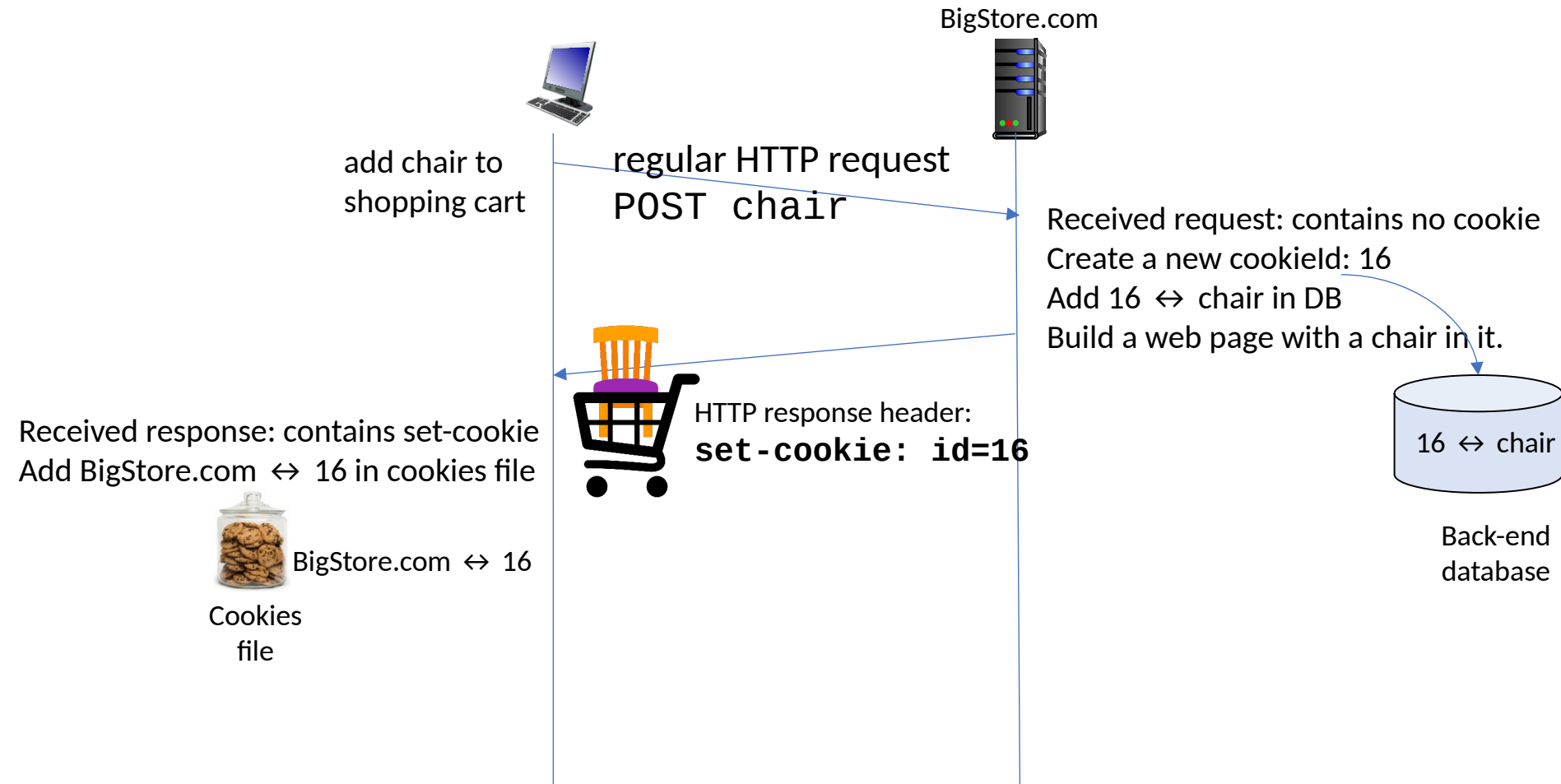
# Cookies

© Original Artist  
Reproduction rights obtainable from  
[www.CartoonStock.com](http://www.CartoonStock.com)

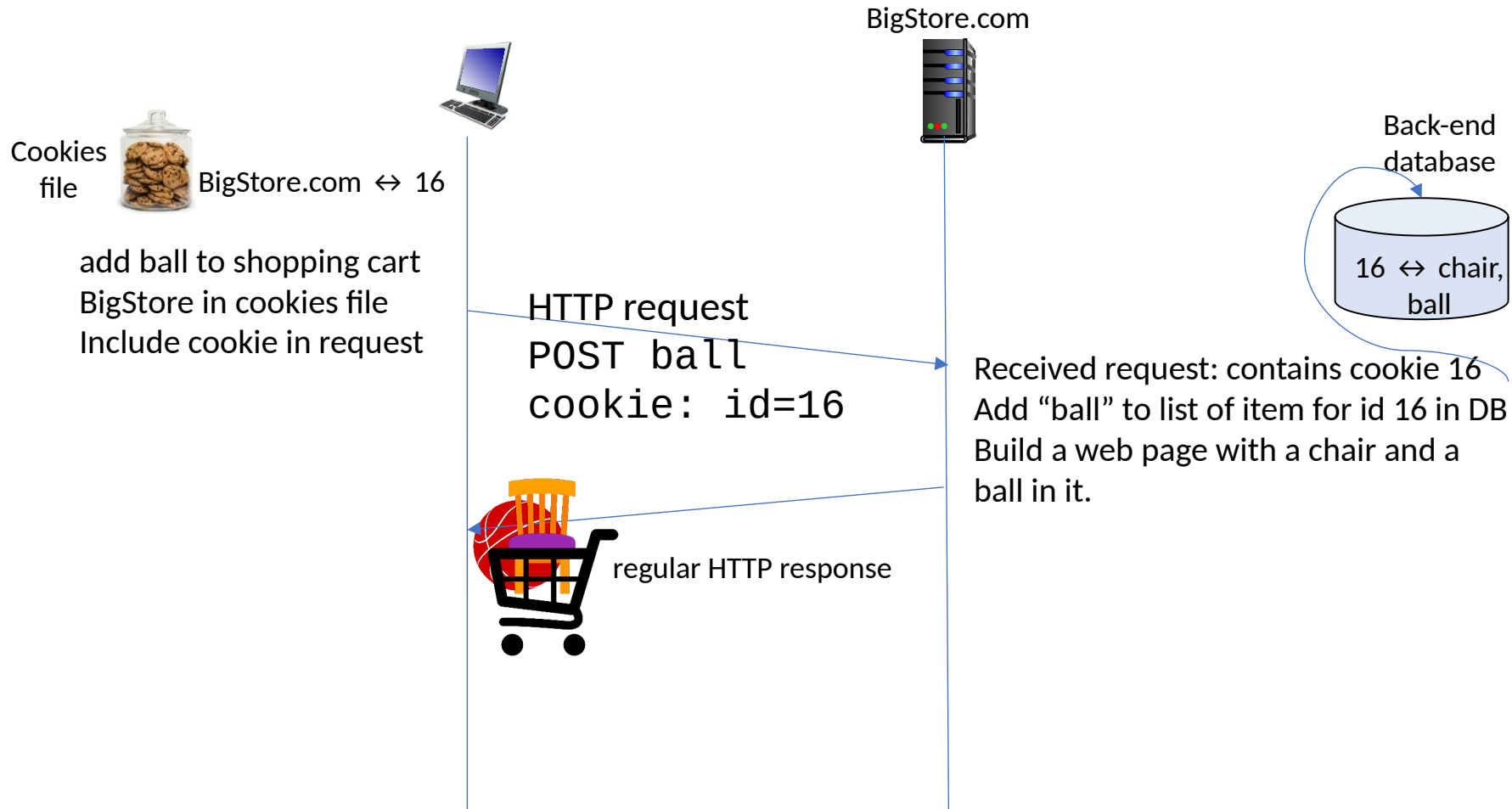
*Shhhhh... they're in here  
somewhere. Dad's always  
talking about the cookies  
in his computer.*



# Cookie Example



# Cookie Example



# Uses

**create a user session layer on top of stateless HTTP**

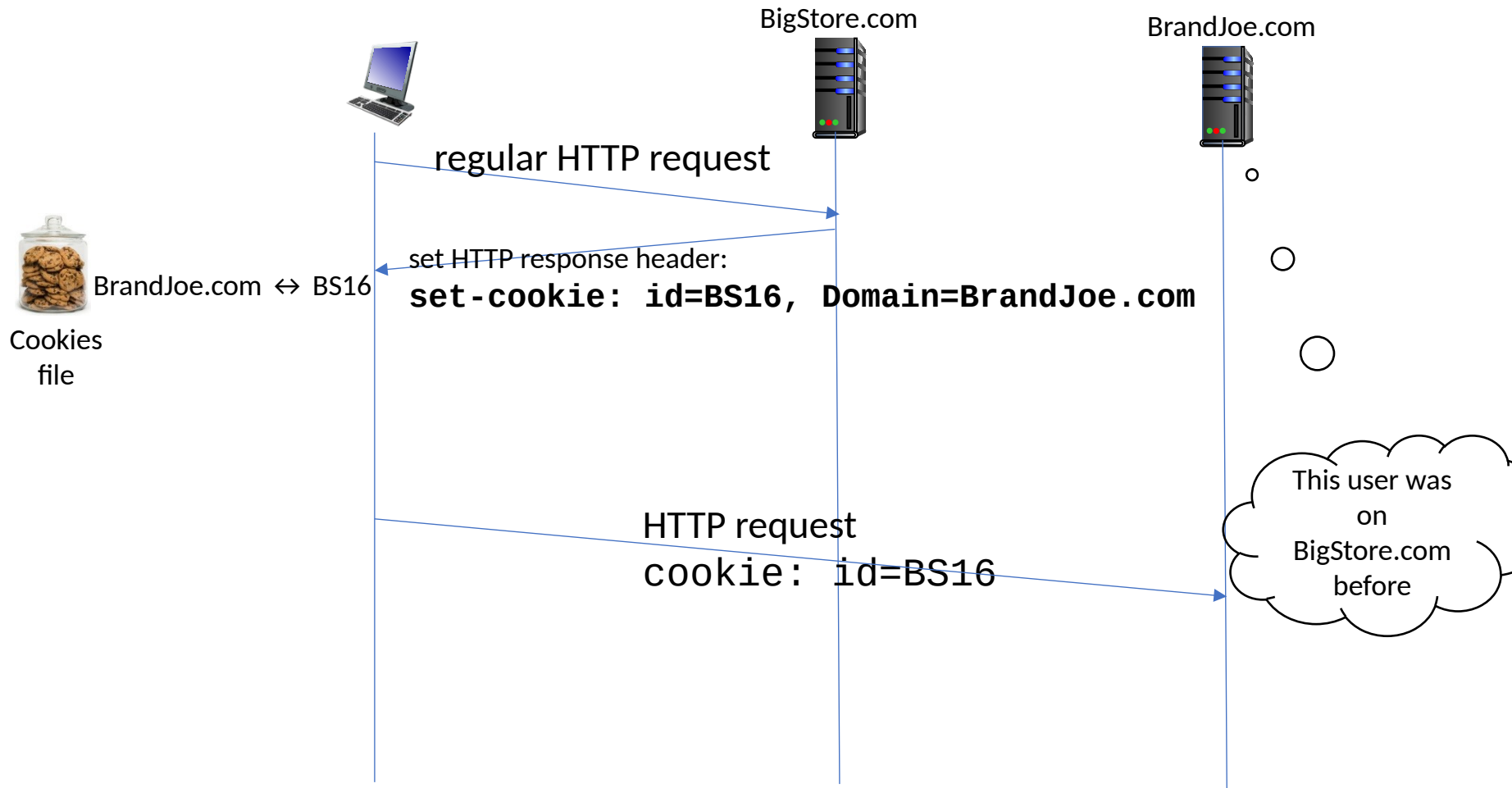
- content adaptation (recommendation based on previous visits etc.).
- shopping carts (e-business)
- session definition at application layer (Web mail)
- authorization
- ...

# Suspicious

- Invasion of privacy

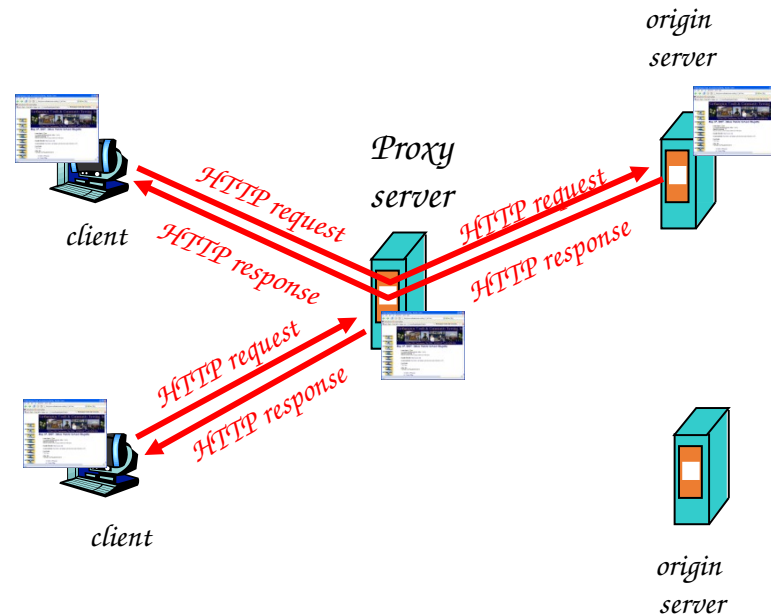


# Third-party advertising cookies



# Web Cache (proxy server)

- to satisfy the requests without involving the real server
- browser must be configured to send all HTTP requests to cache
- reduced traffic on Internet, improved response time





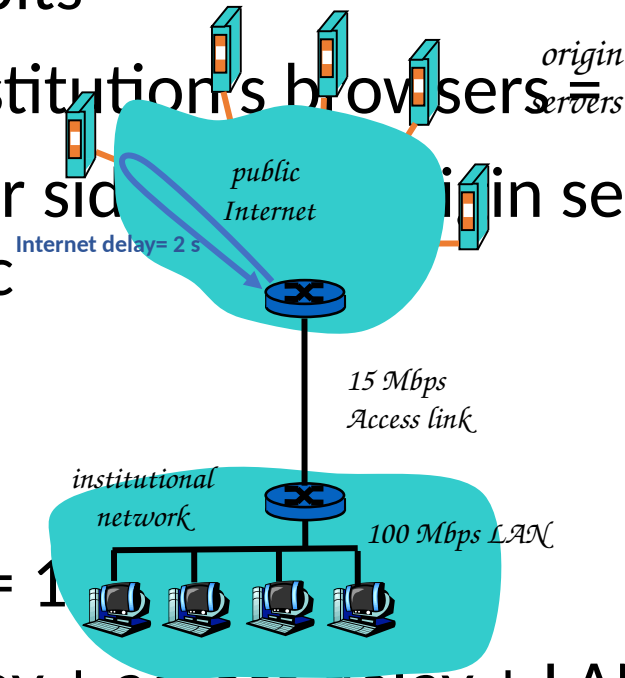
# Caching Example

## assumptions

- average object size = 1Mbits
- avg. request rate from institution's browsers = 15/sec
- delay from Internet router side in server and back to router = 2 sec

## consequences

- utilization on LAN = 15%
- utilization on access link = 1
- total delay = Internet delay + access delay + LAN delay  
= 2 sec + minutes + milliseconds



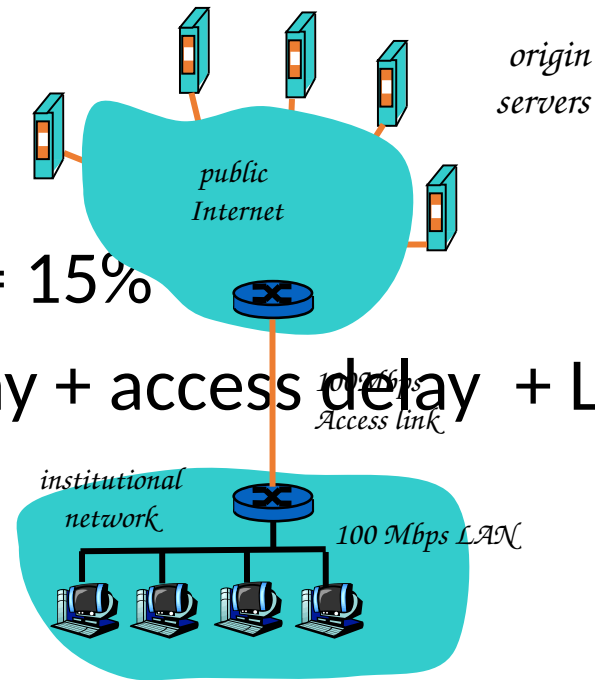
# Caching Example (cont)

possible solution

- increase bandwidth of access link to 100 Mbps

consequence

- utilization on LAN = 15%
- utilization on access link = 15%
- total delay = Internet delay + access delay + LAN delay  
= 2 sec + msec + msec  
often a costly upgrade



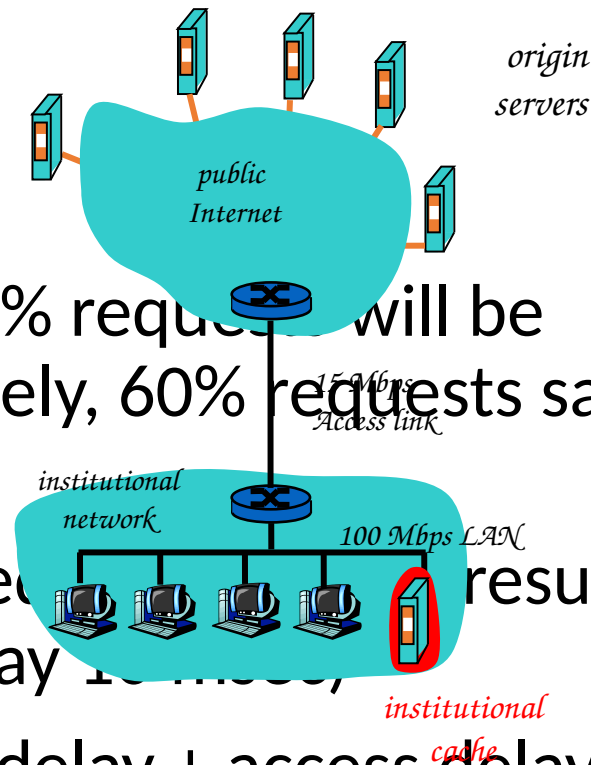
# Caching Example (cont)

possible solution:

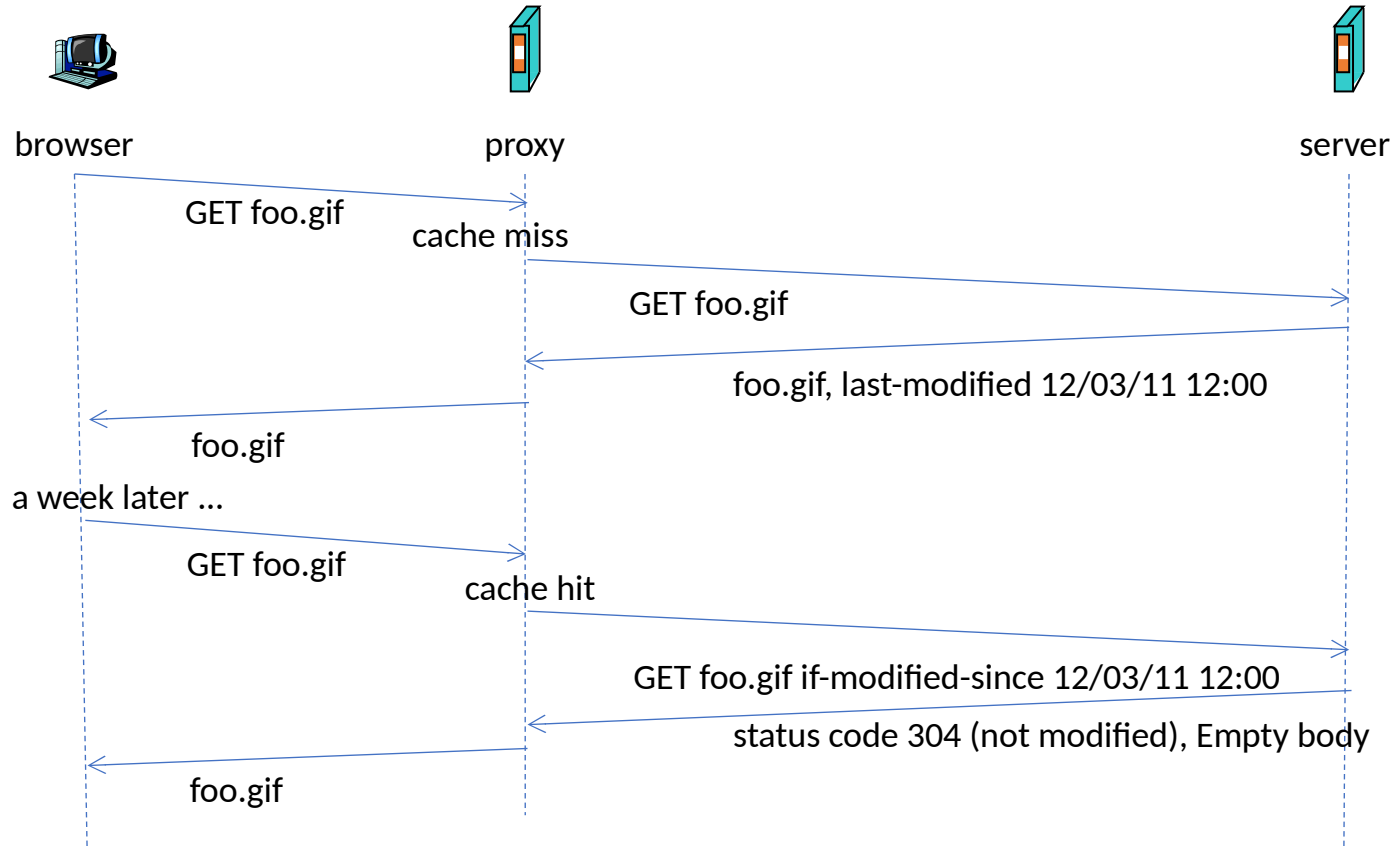
- install cache

consequence

- suppose hit rate is 0.4 (40% requests will be satisfied almost immediately, 60% requests satisfied by origin server)
- utilization of access link reduced, resulting in negligible delays (say 10 msec)
- avg total delay = Internet delay + access delay + LAN delay  
$$= 0.6 * 2.01 \text{ secs} + 0.4 * 10 \text{ millisecs} < 1.3 \text{ secs}$$



# Conditional GET



# Other Uses

- Allow multiple users to get a resource which access is limited to the proxy.
- Track and log web accesses.
- Deny access to a list of web sites.



# Origins

- *Representational State Transfer* – REST: defined in 2000 Roy Fielding's PhD dissertation (after he worked on HTTP 1.1 and URI RFCs)
- Web application =
  - network of Web resources (a virtual state-machine)
  - where the user progresses through the application by selecting resource identifiers and resource operations (application state transitions), resulting in the next resource's representation (the next application state) being transferred to the end user for their use.
- An architectural style, not a standard nor a protocol

# Principles of RESTful Architecture (1/2)

- A resource
  - is identified using an URI,
  - references
    - one entity (eg. user Paul) or
    - a set of entities (eg. all male users)
  - URI doesn't change (but the referenced entity might)
  - and can have multiple representations (JSON, XML...).
- The representation of a resource contains enough information for the client to request a change to its state.
  - Messages include enough information to describe how to process them (eg. Content type)
  - HATEOS (*Hypermedia as the Engine of Application State*)



# HATEOAS Example

request

```
GET /accounts/12345 HTTP/1.1
Host: bank.example.com
Accept: application/xml
...
```

response if balance > 0

```
HTTP/1.1 200 OK
Content-Type: application/xml
Content-Length: ...

<?xml version="1.0"?>
<account>
  <account_number>12345</account_number>
  <balance currency="usd">100.00</balance>
  <link rel="deposit" href="https://bank.example.com/accounts/12345/deposit" />
  <link rel="withdraw" href="https://bank.example.com/accounts/12345/withdraw" />
  <link rel="transfer" href="https://bank.example.com/accounts/12345/transfer" />
  <link rel="close" href="https://bank.example.com/accounts/12345/status" />
</account>
```

response if balance < 0

```
HTTP/1.1 200 OK
Content-Type: application/xml
Content-Length: ...

<?xml version="1.0"?>
<account>
  <account_number>12345</account_number>
  <balance currency="usd">-25.00</balance>
  <link rel="deposit" href="https://bank.example.com/accounts/12345/deposit" />
</account>
```

# Principles of RESTful Architecture (2/2)

- Separation of concerns between the client (user interface concerns) and the server (data storage and processing concerns)
- Stateless communication: the server only stores resources states while the client is in charge of providing the application state.
- Responses should define the extent to which they can be cached.
- A client may not be directly connected to the end-server: there can be proxies, an additional security layer, and the server might call other servers to complete the service.

# Semantics of HTTP methods

HTTP method	Operation on the resource	URIs: examples	HTTP response status	<i>location</i> header	safe	idem potent
GET	read	GET /serv/users GET /serv/users/34	200 OK	no	yes	Yes
POST	create	POST /serv/users # body { name: "Toto" }	201 Created	Yes	no	no
PUT	update	PUT /serv/users/34 # body { name: "Jacques" }	200, 204 No Content	no	no	Yes
PATCH	partial update					
DELETE	delete	DELETE /serv/users/34	200, 204, 202 Accepted	no	no	yes

# Example of scenario

- Book a room:

POST [http://myhotel.com/reservations?date="12/03/2021"&nights=2&persons=4](http://myhotel.com/reservations?date=12/03/2021&nights=2&persons=4)

Server replies with reservation number 123

- Display reservation:

GET <http://myhotel.com/reservations/123>

- Update the reservation:

PATCH <http://myhotel.com/reservations/123?persons=3>

- Cancel the reservation:

DELETE <http://myhotel.com/reservations/123>

# Best Practices for well-designed RESTful APIs

- Use only nouns for a URI:  
~~/getAllReservations~~ GET /reservations
- Use plural nouns:  
GET /reservations for all reservations  
GET /reservations/123 for a specific reservation
- GET method should not alter the state of a resource
- Use sub-resources for relationships between resources  
GET /reservations/123/persons/1: first occupant of the reservation #123
- Use “content-type” and “accept” HTTP headers to specify input/output format
- Provide proper HTTP status codes

# Best Practices for well-designed RESTful APIs

- Offer filtering and paging capabilities for large data sets

`GET /reservations?date=28/02/2021`

`GET /reservations?from=5&to=25`

- Version the API

2 strategies:

- In the URI: `GET /api/v2/reservations/123`



Easy to use with a web browser



Non-compliant with REST principle “one resource = one URI”

- In the accept header:

`GET /api/reservations/123 accept: application/v2`



More complex for the client



More REST-compliant