CS 436
Cloud Computing

04.04.2024

Load Balancing

- DNS based sol’ns
- HTTP redirection
- Load Balancer
  - Layer 4 Load Balancer (Legacy Load balancing)
  - Layer 7 Load Balancer (Layer 7 Switching)

DNS Based Load Balancing

Name	TTL	Type	IP
www.example.com.	10800	A	192.168.1.2
www.example.com.	10800	A	10.2.54.4

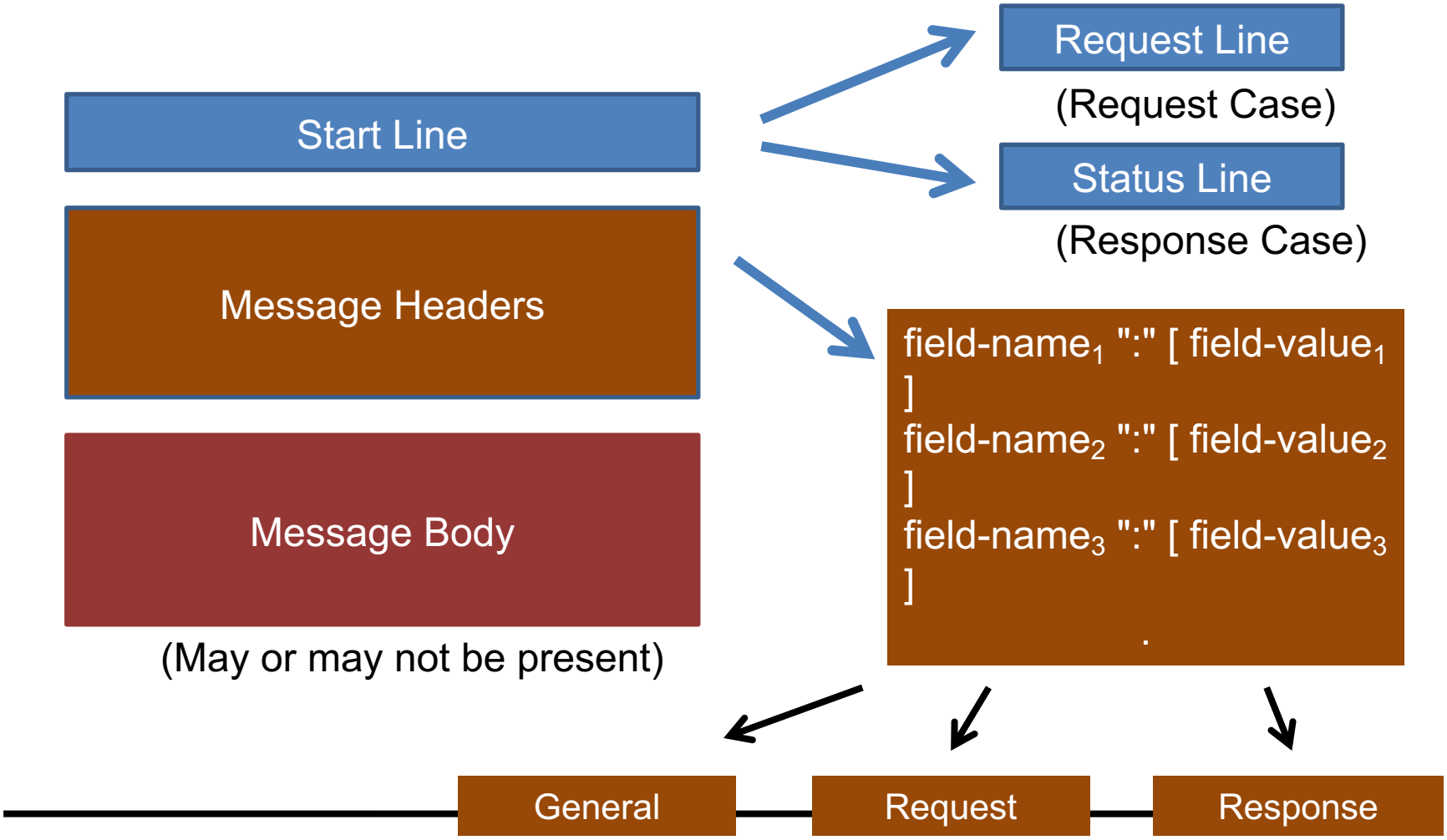
Round robin policy supported by DNS servers

Ref: http://oldcp.dnsmadeeasy.com/enterprisedns/records.html

HTTP Redirection

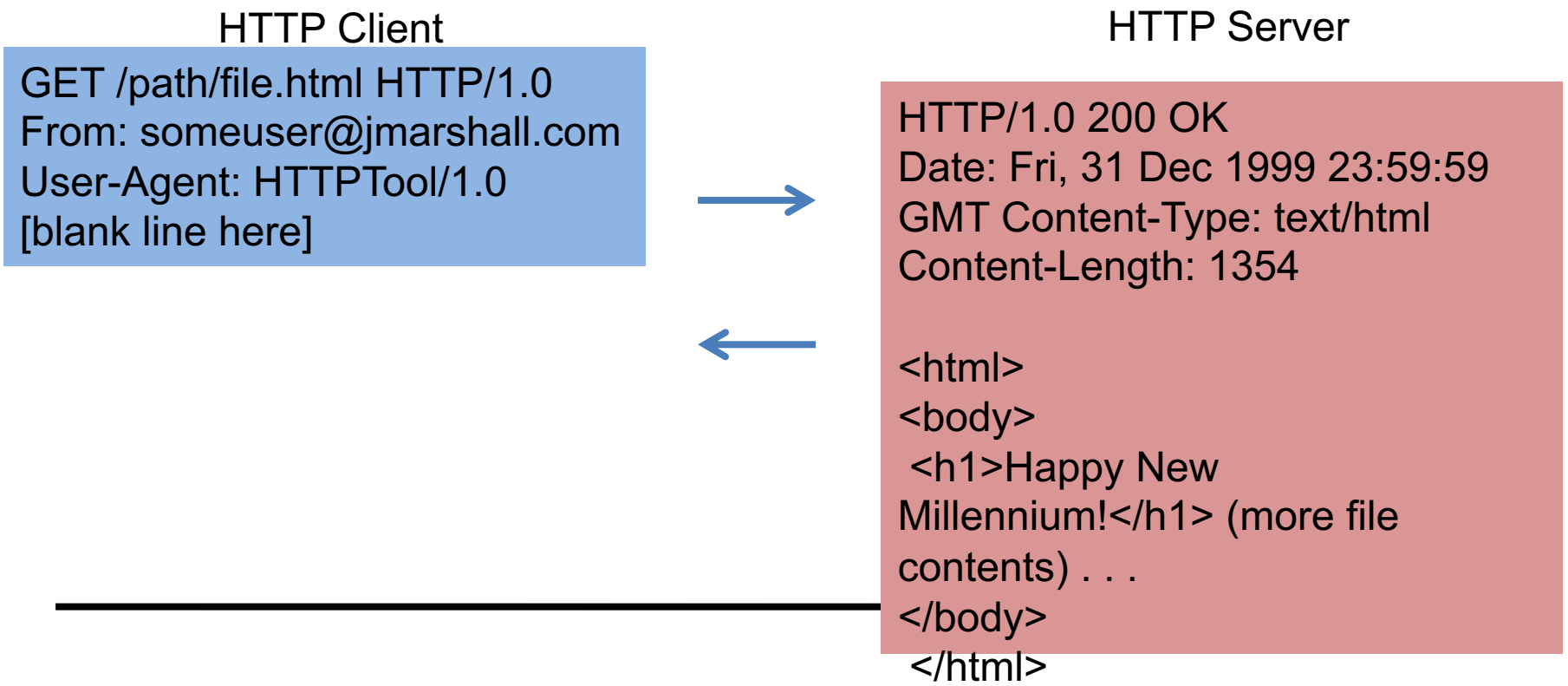
- Advantages: Very simple to implement
- Weaknesses
  - Internals of the server organization becomes visible to the client. Access Transparency violated -> Possible security problems
  - Does not enhance availability: Clients do not automatically switch to another replica in case of failure
  - Scalability: Depends on the delivery speed of the initial document.

Generic HTTP Message



Typical Messaging Sequence

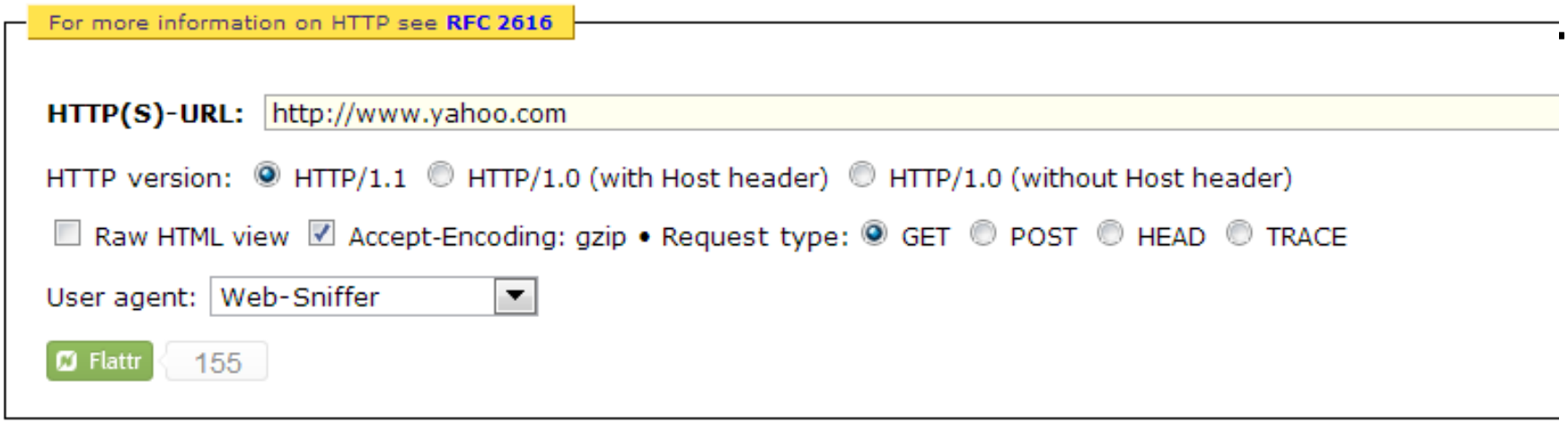
- We want to access: http://www.somehost.com/path/file.html



Response Classes (RFC2616)

- 1xx: Informational - Request received, continuing process
- - 2xx: Success - The action was successfully received, understood, and accepted
- - 3xx: Redirection - Further action must be taken in order to complete the request
- - 4xx: Client Error - The request contains bad syntax or cannot be fulfilled
- - 5xx: Server Error - The server failed to fulfill an apparently valid request

Yahoo Redirection



HTTP Request Header

```
Connect to 87.248.112.181 on port 80 ... ok

GET / HTTP/1.1[CRLF]
Host: www.yahoo.com[CRLF]
Connection: close[CRLF]
User-Agent: Web-sniffer/1.0.44 (+http://web-sniffer.net/)[CRLF]
Accept-Encoding: gzip[CRLF]
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8[CRLF]
Accept-Language: en-GB,en-US;q=0.8,en;q=0.6,tr;q=0.4[CRLF]
Accept-Charset: ISO-8859-1,UTF-8;q=0.7,*;q=0.7[CRLF]
Cache-Control: no-cache[CRLF]
Referer: http://web-sniffer.net/[CRLF]
[CRLF]
```

Yahoo Redirection - II

HTTP Response Header

Name	Value
Status: HTTP/1.1 302 Found	
Date:	Tue, 21 May 2013 12:00:46 GMT
P3P:	policyref="http://info.yahoo.com/w3c/p3p.xml", CP="CAO DSP COR CUR ADM DEV TAI PSA PSD IVAi IVDi CONi T
Cache-Control:	private
X-Frame-Options:	SAMEORIGIN
Set-Cookie:	fpc=d=GM9G2jOUTowmMDxTEZqMwNkz7vR8j_iZCWLz93qFqrHHpl0n.mVPm1UslQRCE2gXuBL81_4biWISePOkM0QAr &v=2; expires=Wed, 21-May-2014 12:00:46 GMT; path=/; domain=www.yahoo.com
Location:	http://de.yahoo.com/?p=us
Vary:	Accept-Encoding
Content-Type:	text/html; charset=utf-8
Age:	0
Connection:	close
Server:	YTS/1.20.13

Commercial Load Balancer Spec

	ALB-X E4	ALB-X E5
Chassis	1U Half Depth	1U Full Depth
CPU	Quad Core	2 X Hex Core
Memory	8GB	12GB - 64GB
Network Interface	4x 10/100/1000	8x 10/100/1000
Max Interface	8x 10/100/1000	16x 10/100/1000 or 2 x 10GbE
Power Consumption	71W to 100W maximum	150W to 200W maximum
HTTP Max Throughput	8 Gbps	25 Gbps
Layer 7 – HTTP Requests Per Second	70,000 rps	400,000 rps
SSL Transaction Per Second	6000 tps	20,000 tps

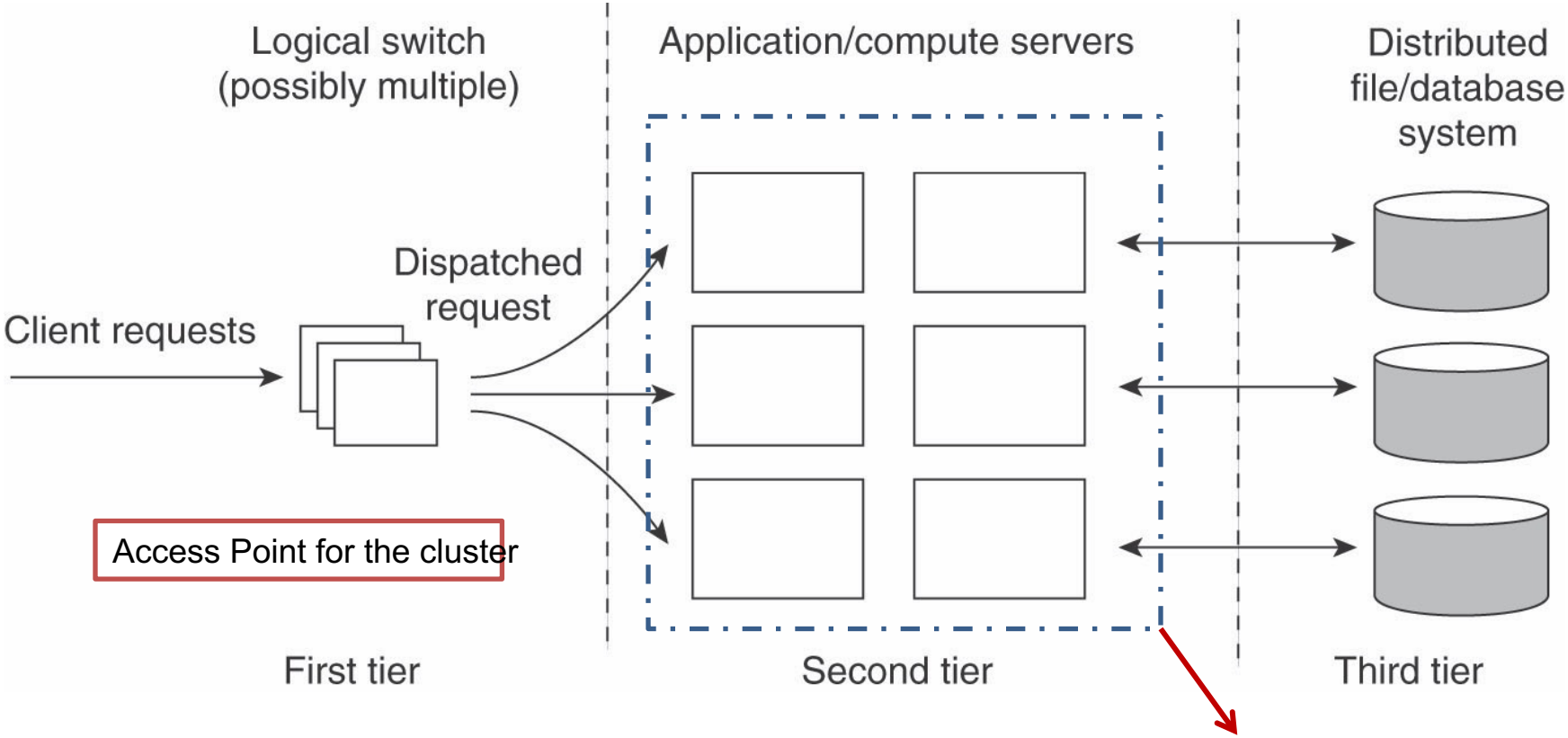
Scheduling Methods (Layer 4 Balancer)

- Round Robin
- Weighted Round Robin
- Least Connection
- Weighted Least Connection

Layer 4 vs. Layer 7 Load Balancing

- İdentical functionality should be served in L4 Load Balancing (No dedicated PHP, image etc. server can exist)
- L7 load balancing allows partial replication and dedicated servers.

Load Balancing



Homogeneous vs. Heterogeneous Application Servers