Multi-layer switch hardware commutation across various layers

Mario Baldi

Politecnico di Torino

http://staff.polito.it/mario.baldi

Based on chapter 10 of:
M. Baldi, P. Nicoletti, "Switched LAN", McGraw-Hill, 2002, ISBN 88-386-3426-2

Copyright notice

Layer 4 Switches

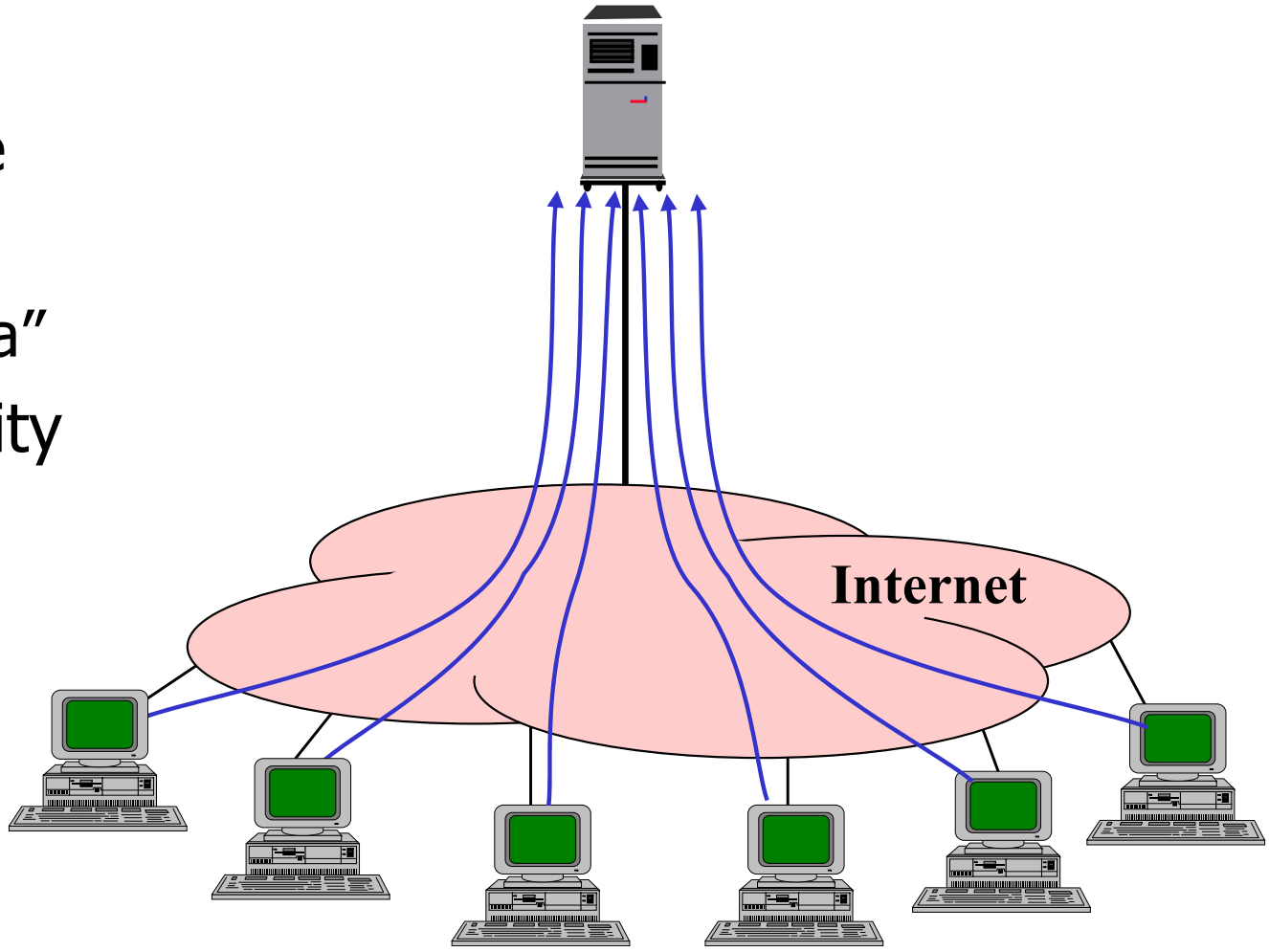
- Hardware processing of layer 4 information
  - Actually TCP and UDP ports
- Often for *filtering* and/or *classification* (not routing)
  - Rules matching to identify packets to process in a specific way
    - Packets are inserted in a specific queue
    - Discard
    - Forwarding using a specific interface (not common)
  - These rules are established during configuration

- Wire speed functionality on 10 Gigabit Ethernet
  - Traditional routers implement filtering functionalities by software with a considerable performance reduction
  - For instance, Access Control Lists implementations
- Commercial devices are Multi-Layer Switches

Motivations for Layer 7 Switches: issues coming from the Internet success

Many simultaneous accesses to a server

- Usually a web server
- Portals
  - Yahoo, Netscape
- Real time news
  - CNN, "La Stampa"
- Software availability
  - Microsoft
- Search engines
  - Google, Yahoo



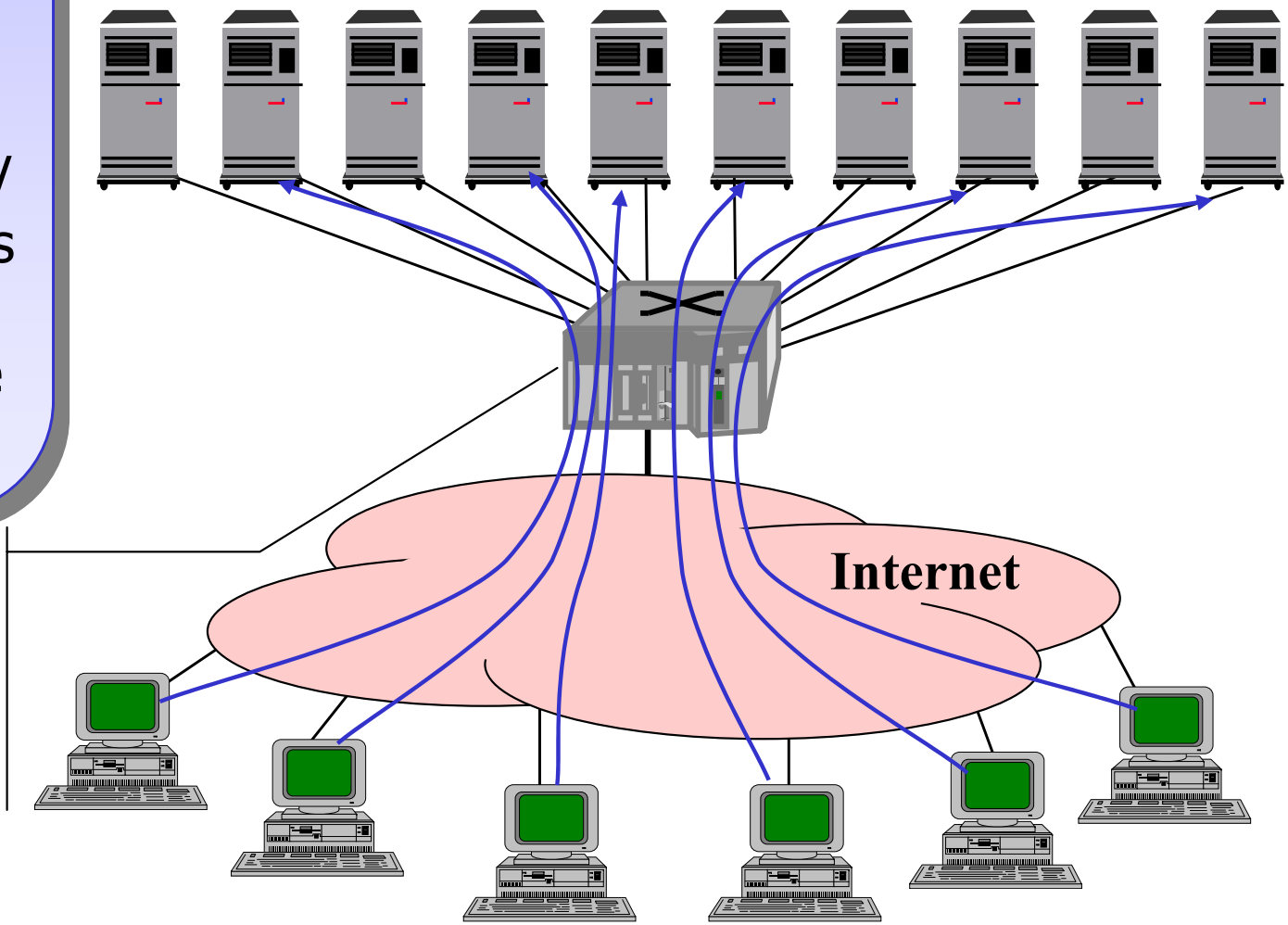
The solution

Server load balancing

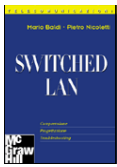
An uniform distribution of requests between servers creates a *virtual server* which capability is equal to the sum of the capabilities of the single servers

Server farm

The usage of different server is cheaper that the realization of a single server having the same capabilities



Layer 7 switch
Application layer switch
Content aware router
Server load balancer



Application Layer Switches

- Requests distribution implies a content based forwarding
  - Forwarding based on application layer information
- Need for high performances
  - The servers in the server farm have GE or 10GE interfaces

Hardware (ASIC) for layer 7 processing

Layer 7 Switches/Application Layer Switches

The devices on the market are multi-layer switches

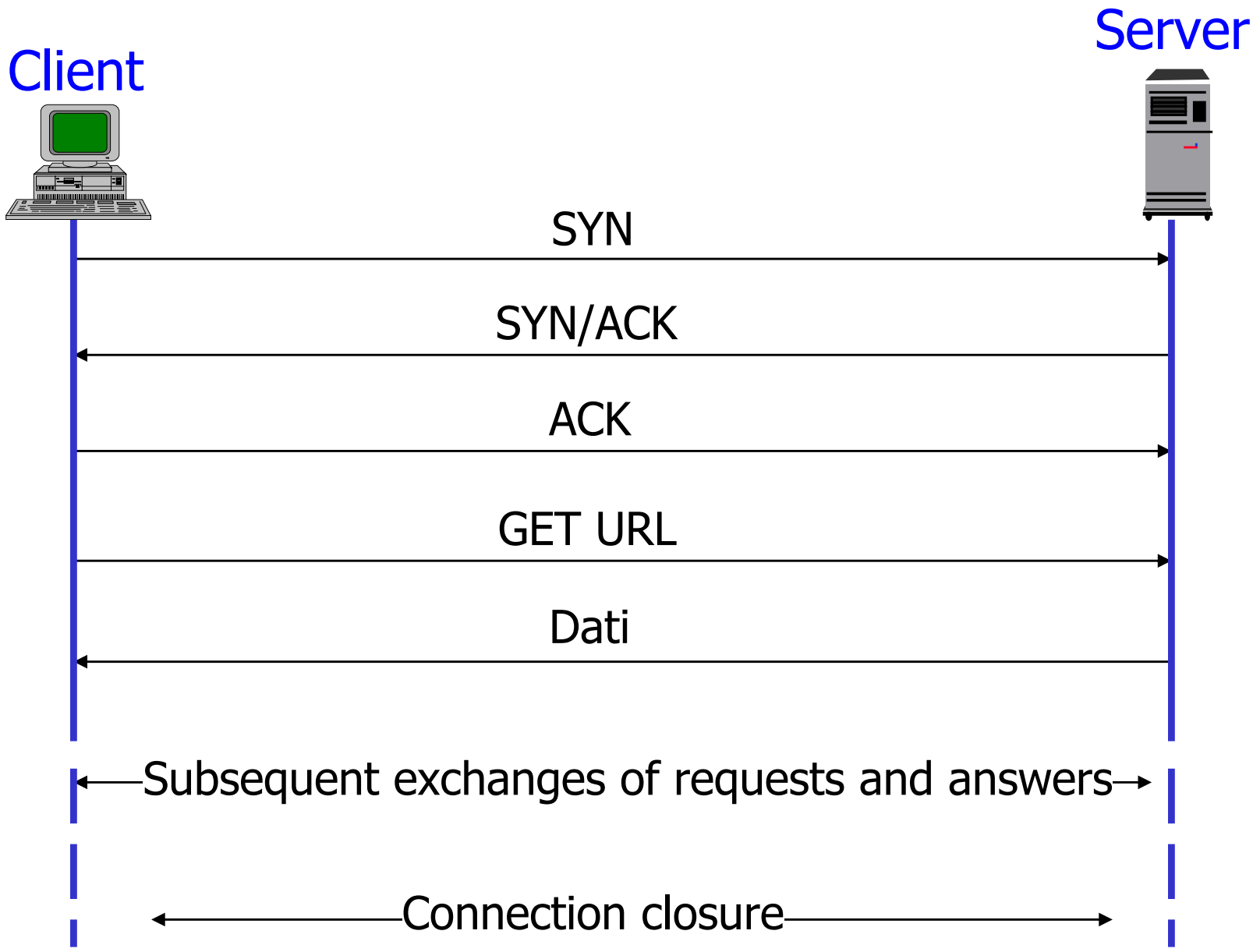
Functionalities

- An IP address identifies the virtual server
- IP packets containing a request reach a layer 7 switch on the path up to the server farm
- A layer 7 switch processes the request and decides to which physical server the related IP packets will be forwarded
- In some applications the server keeps information about the client → following interactions with the same server
  - For instance, shopping chart
  - Sticky conncetions
- Access limitations
  - Implementation of access policies
  - Rules matching and filtering procedures

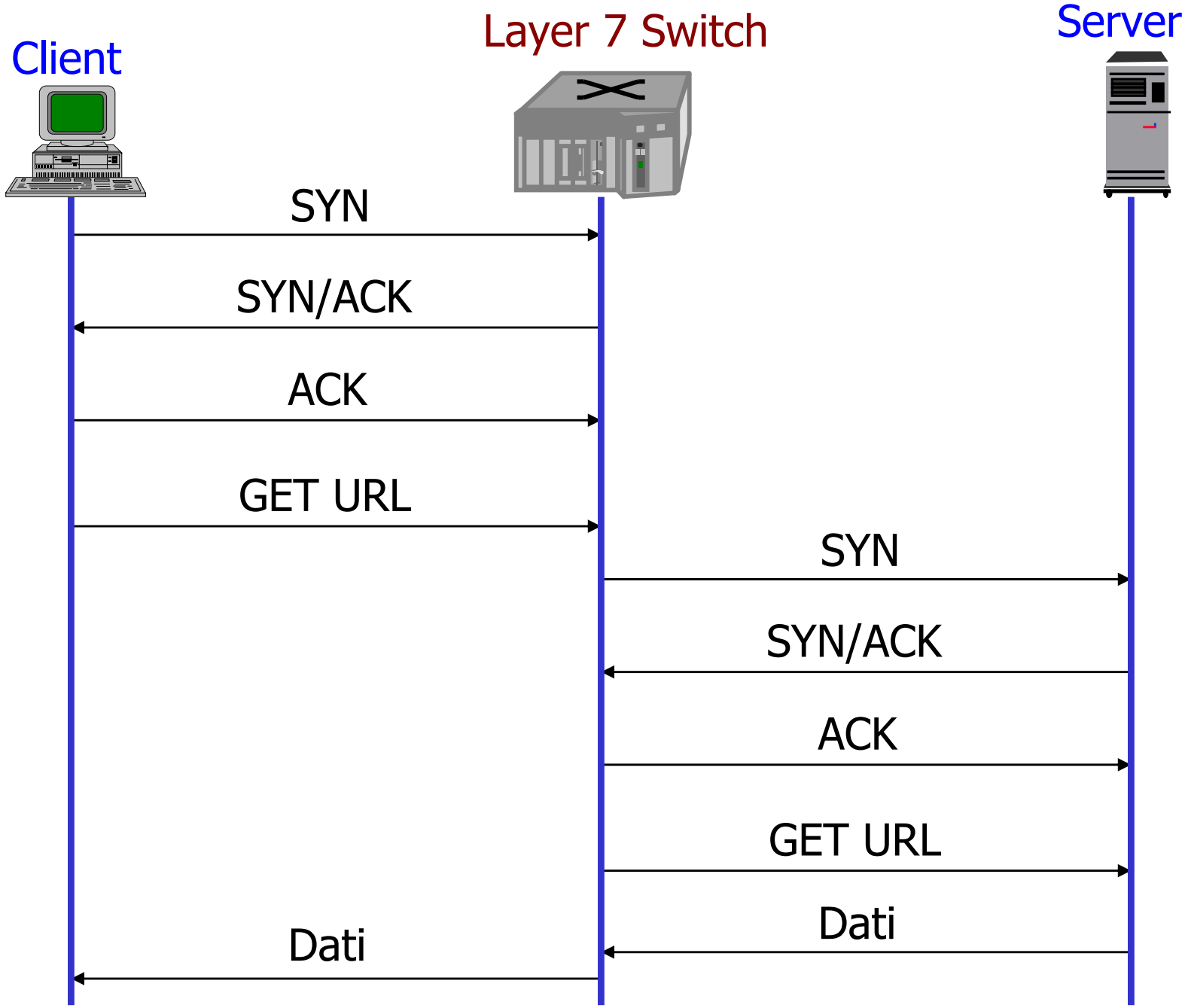
Choice of the physical server

- Identical physical servers
  - Each server keeps a copy of the same data
  - Weighted round robin
    - New requests are distributed over servers by turns
  - Weighted least connections
    - A new request is assigned to the server that is processing the smallest number of requests
  - A weight helps to keep in count different performances
- Different physical servers
  - The data stored in the virtual server is distributed over different servers
    - Welcome page replicated on different servers
    - Other pages, distributed over different servers
  - The choice of the servers depends on the content
  - Requests distribution results in load balancing

Layer 7 processing: not only a new header to analyze



Connection closure



Consequences

- Heavy work to keep TCP connections
  - Acknowledgements processing
  - Storage of state information
  - Storage of the packets in the windows
    - Sent, but still not confirmed
  - Hard to implement in hardware

Performance indexes

- Maximum number of opened connections
  - About a few hundred thousands
- Maximum number of new connections for second
  - About a few ten thousands

Shortcuts

- They are possible when requests forwarding is not really based on content
  - Weighted round robin, weighted least connections
  - Destination port
- Isolate the messages that open the connection
  - SYN
- Forward to the real server chosen
- The connection is directly opened with the real server
- Control if the opening of the connection is successful
- Control if the closure of the connection is successful
  - Useful mostly when in weighted least connections mode

Advanced functionalities

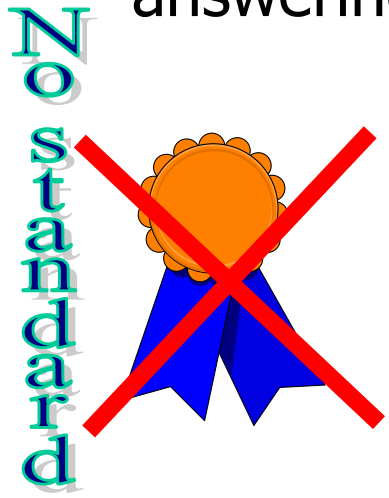
- Periodical checks to verify if the real servers are working
  - Proper answer to SYN messages
  - Collect and process statistical data
- Faulty real servers are excluded
  - Periodical checks to verify if the faulty servers work
- Pending requests assigned to faulty servers are assigned to other servers
  - Fault protection
- Controls and statistics based on SYN messages are used to avoid Denial of Service (DoS) attacks
- Layer 7 switch: load balancing
  - More than one devices working simultaneously
  - Coordination and division of requests load

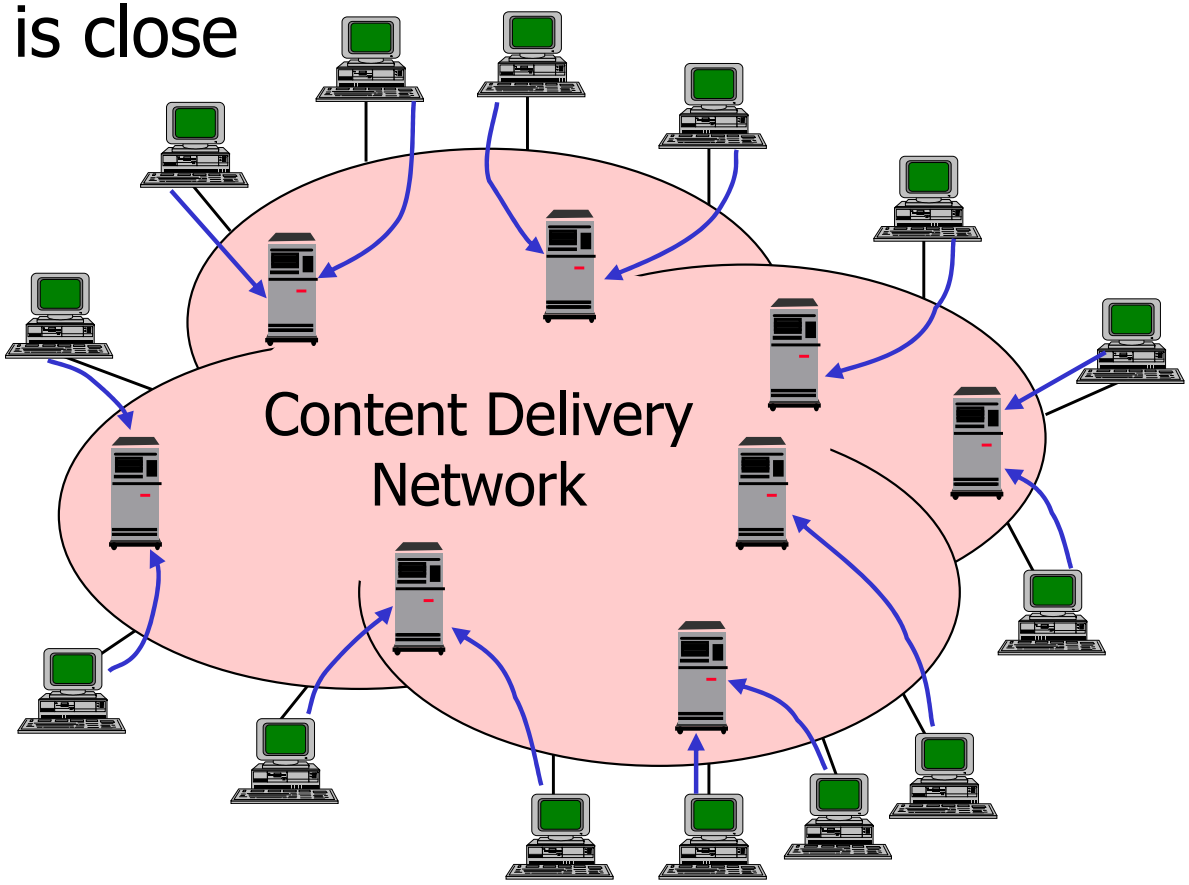
Advanced functionalities

- Layer 7 Switches redundancy
  - If a device stops working or is isolated, the other ones replace it
  - *Stateless failover*: connections must be initialized again
  - *Stateful failover*: the backup switch knows the state of the primary one: no need to initialized connections again
    - Very complex
- Verification of the balancing efficiency
  - Traffic analysis to estimate the answering time of the server
  - An agent executed on the server keeps state information and sends it to the layer 7 switch
    - Non-standard protocol used to communicate with the server

Content Delivery Network (CDN)

- Generalization of the load balancing concept
- The Server farm is distributed over the network
- The forwarding is based on the content of the requests
- The content delivered is close to the users
  - Reduction of the answering time





Something more than a different application mode

- At first Layer 7 switches were the basis of the CDNs
- Advanced routing functionalities are needed
  - Routing protocols to spread information about contents position
  - Almost brand new functionalities
- Current solutions based on tricks
  - Lower complexity
  - Lower efficiency
- Modifications to DNS (Domain Name Service)
  - A name is related to several addresses
  - The address provided is the one of the closest server to the requester
- Modified URLs (Universal Resource Locator)
  - Solution adopted by Akamai