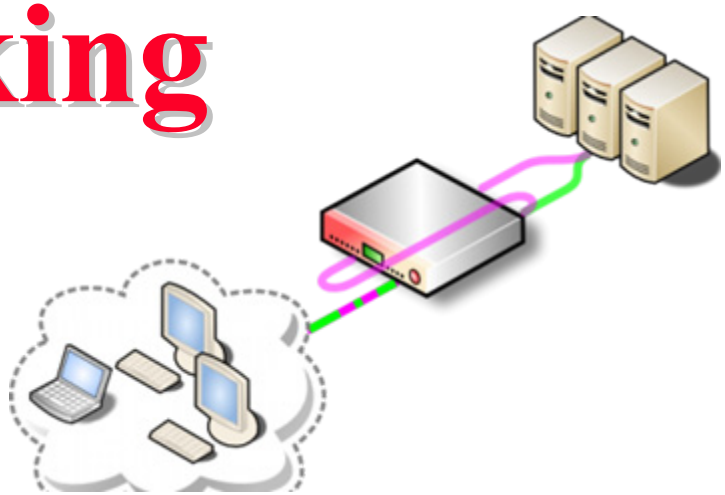
Application Delivery Networking



Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides and audio/video recordings of this class lecture are at:

http://www.cse.wustl.edu/~jain/cse570-13/

©2013 Raj Jain

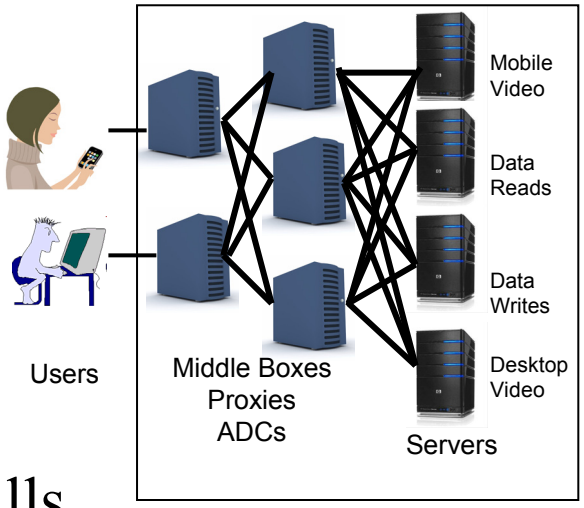
Overview

1. Application Delivery Controllers (ADCs)
2. Load Balancing Concepts and Modes
3. Network Address Translation (NAT)
4. Other ADCs
5. Virtual ADCs

©2013 Raj Jain

Application Delivery in a Data Center

- **Replication**: Performance and Fault Tolerance
  - If Load on S1 >0.5, send to S2
  - If link to US broken, send to UK
- **Content-Based Partitioning**:
  - Video messages to Server S1
  - Accounting to Server S2
- **Context Based Partitioning**:
  - Application Context: Different API calls
    - Reads to S1, Writes to S2
  - User Context:
    - If Windows Phone user, send to S1
    - If laptop user, send to HD, send to S2
- **Multi-Segment**: User-ISP Proxy-Load Balancer-Firewall-Server



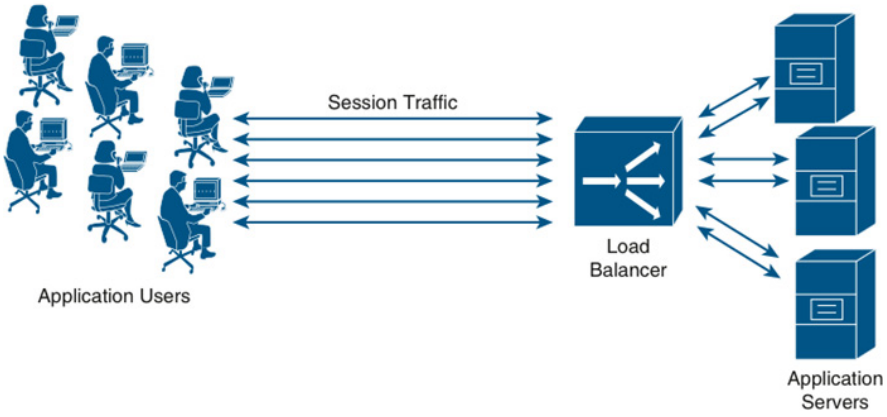
Application Delivery Controllers (ADCs)

- **Networking Service**: Tasks that apply to multiple applications, e.g., security, monitoring, acceleration, offload etc.
- Several applications can share a service
- **ADC**: Appliance that provide network services, e.g., load balancing, firewall, proxy, SSL offload,
- Load balancer is an example of ADC and also the basic component of most ADCs.
- Other ADCs: Firewalls, Reverse Proxies, SSL Offload, TCP Multiplexing, HTTP Compression

©2013 Raj Jain

Load Balancing Concepts

- **Load Balancer**: Allows adding servers as needed.
- **Server Cluster**: A centrally managed set of servers working together on one application. Appear as one server.
- **Server Farm**: A set of independent servers.



Ref: G. Santana, "Datacenter Virtualization Fundamentals," Cisco Press, 2014, ISBN: 1587143240

©2013 Raj Jain

Load Balancing Concepts (Cont)

- **Probes**: Synthetic requests sent by load balancer to see if a server and application is up. E.g., HTTP get request.
- **Virtual IP (VIP)**: Client facing IP address of the load balancer
- **Stickiness Table**: Lists which server a client has been mapped to.
- **Predictor**: Load balancing rules, e.g., round-robin, least connections, hashing, least loaded
  - Round-robin DNS load balancing does not know whether a server is down, overloaded, or appropriate for different types of application traffic (video, voice, data, …)

©2013 Raj Jain

Load Balancing Concepts (Cont)

- **Layer 4 Switching**: L2-L4 fields are used to assign a server. Server can be assigned at TCP Syn.
- **Layer 7 Switching**: L5-L7 fields such as data type or application request type is used.
  ⇒ Load balancer accepts all TCP connections but assigns servers only when the client sends an application request.
  ⇒ *Delayed Binding*
- **Symmetric Connection Management**: Traffic in both direction passes through the load balancer. LB changes the source IP in the requests to server.
- **Asymmetric Connection Management**: Client-to-Server traffic passes through the load balancer. Server-to-Client traffic goes directly. Good for video servers. Used rarely.

©2013 Raj Jain

Load Balancing Using DNS

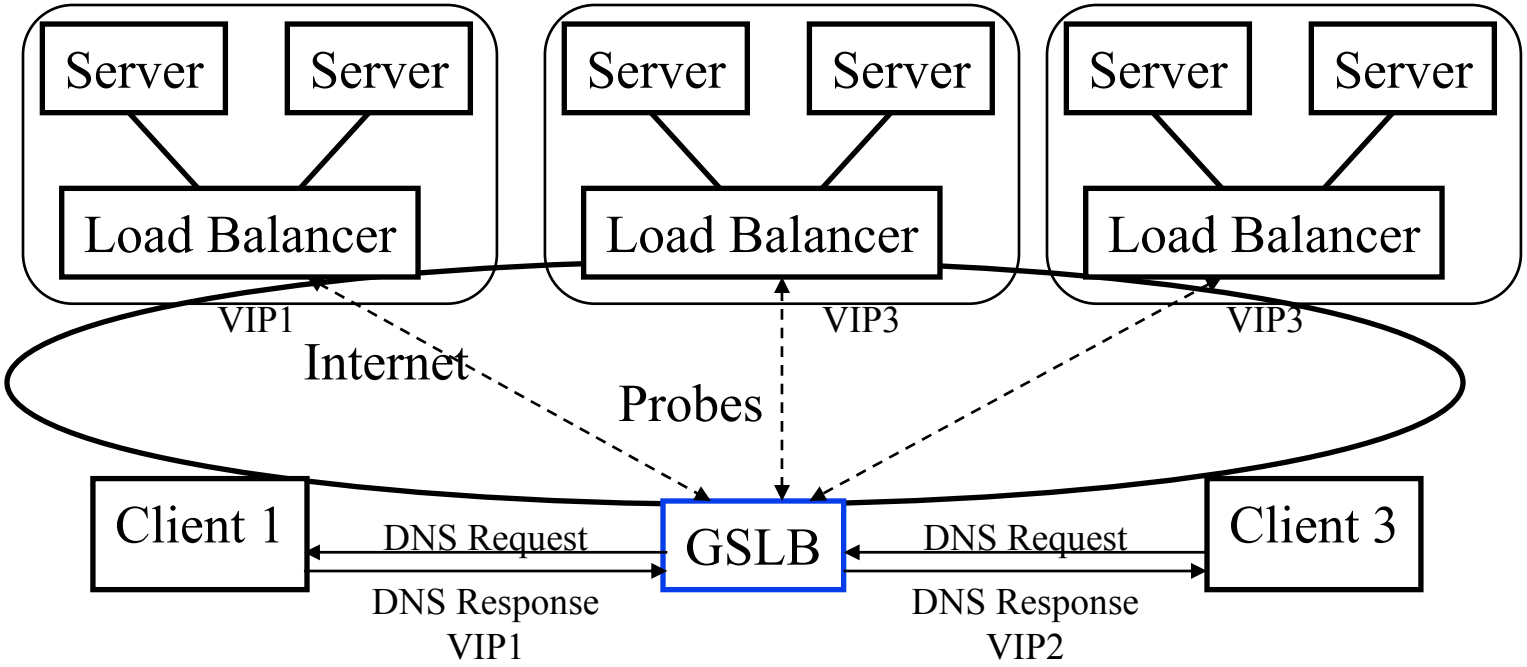
- Domain Name System (DNS) allows multiple IP addresses for a name
- On query, DNS returns a list of addresses. The order of entries is rotated on each subsequent query. For example, abc.com

  1st Time: 74.25.21.201, 74.25.21.202, 74.25.21.203
  2nd Time: 74.25.21.202, 74.25.21.203, 74.25.21.201
  3rd Time: 74.25.21.203, 74.25.21.201, 74.25.21.202

- Rotating DNS is used for load balancing
- Problems:
  - DNS does not know geo-location of servers
  - DNS does not know whether a server is loaded
  - DNS does not whether a server is up
  - Adding or removing a server is too slow

Global Server Load Balancing (GSLB)

- Specialized DNS servers that keep track of server locations and states
  ⇒ Good only for intra-company use.
  Not able to change caches at external servers.



©2013 Raj Jain

Load Balancing Modes

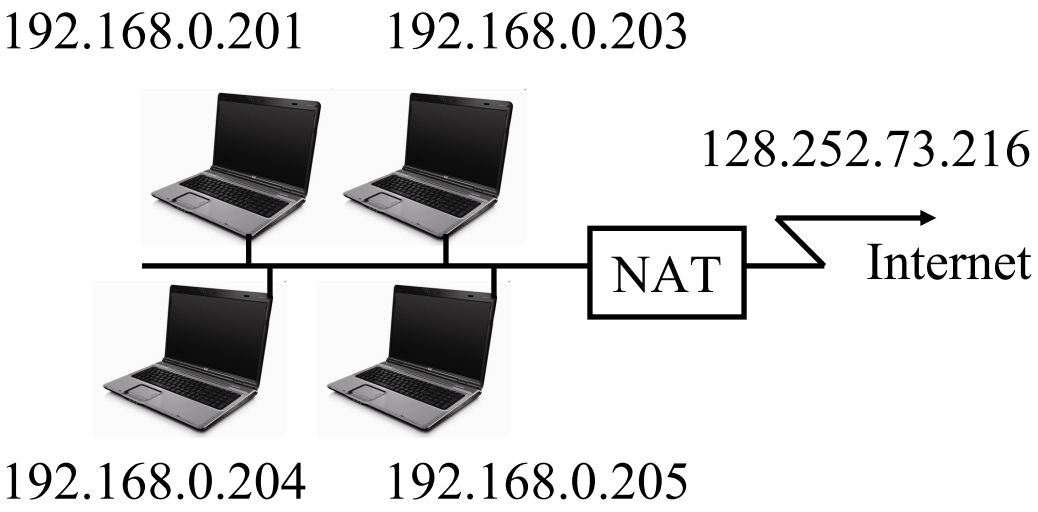
What changes are made to client requests so that the request is routed to the correct server?

1. Dual NAT (One-Arm mode)
2. Server NAT (Routed Mode)
3. Transparent Mode (Direct Routed Mode)

Type	Change Source IP Address?	Change Source TCP Port?
Dual NAT	Y	Y
Server NAT	N	Y
Transparent	N	N

©2013 Raj Jain

Network Address Translation (NAT)



- Private IP addresses 192.168.x.x cannot be used on the public Internet
- NAT overwrites source addresses and port on all outgoing packets. Makes a note. Writes corresponding destination addresses and port on all incoming packets
- Only outgoing connections are possible

NAT Example

192.168.0.201 192.168.0.202

128.252.73.216

NAT → Internet

Outgoing Source Address	Outgoing Source Port	New Source Port
192.168.0.201	8001	3001
192.168.0.201	3002	3002
192.168.0.202	8001	3003

IP	Dst: 74.125.224.210 Src: 192.168.0.201
TCP	Dst: 80 Src: 8001

IP	Dst: 74.125.224.210 Src: 128.252.73.216
TCP	Dst: 80 Src: 3001

Public Address

New Port #

Source ⇄ NAT ⇄ Destination

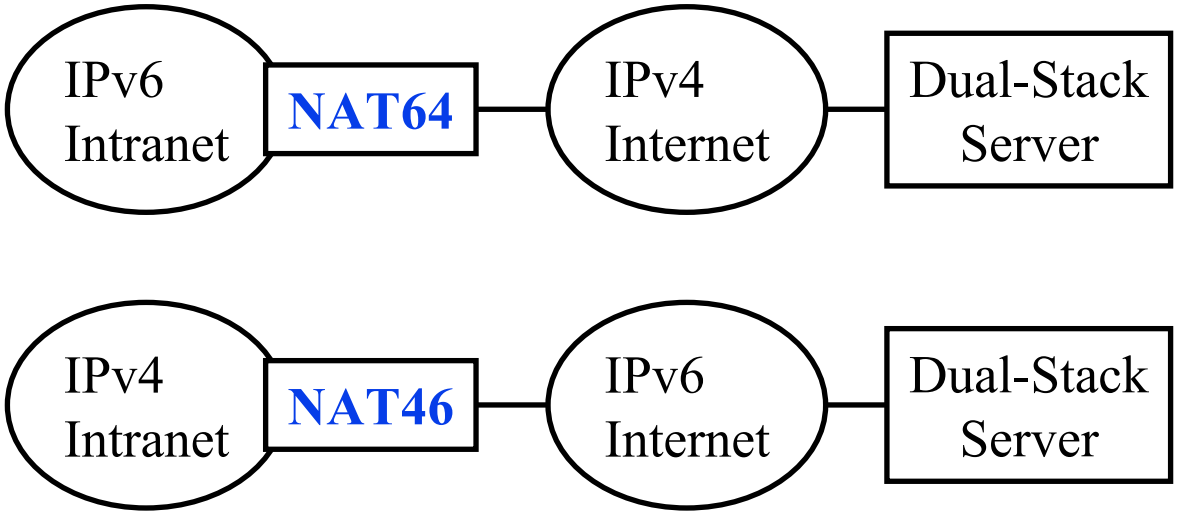
IP	Dst: 192.168.0.201 Src: 74.125.224.210
TCP	Dst: 8001 Src: 80

IP	Dst: 128.252.73.216 Src: 74.125.224.210
TCP	Dst: 3001 Src: 80

©2013 Raj Jain

NAT64 and NAT46

- NAT described so far is called NAT44 (both internal and external addresses are IPv4).
- A NAT allows internal addresses to be IPv6 even if the external Internet is still IPv4 and vice-versa.
  Will also need to encapsulate traffic.
- This allows organizations and/or carriers to transition to IPv6 now.

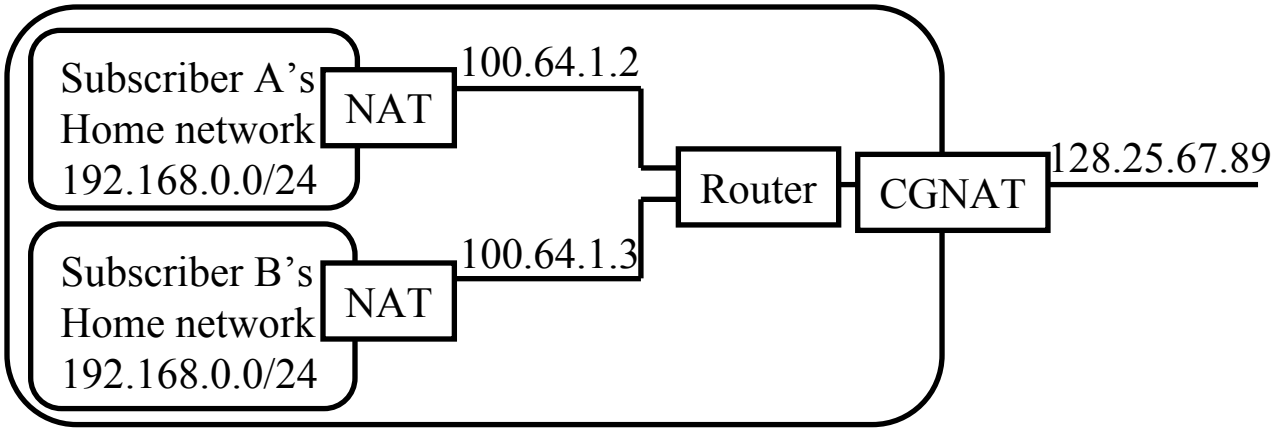


Ref: http://en.wikipedia.org/wiki/NAT64

©2013 Raj Jain

Carrier Grade NAT

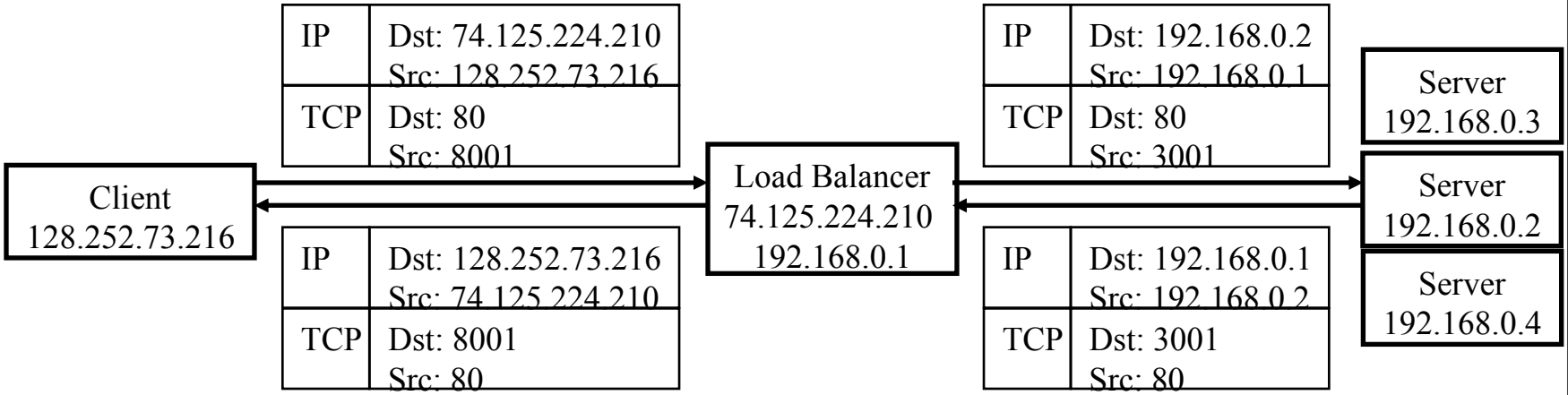
- Internet Addressing and Naming Authority (IANA) has recorded 100.64.0.0/10 as **shared address range**
  Like private address space but reserved for use by carriers
- Carriers can allocate an address from this range to their subscriber and then use a small number of public IPv4 or IPV6 addresses using a "**Carrier Grade NAT** (CGNAT)"
  Also known as **NAT444**, Large Scale NAT (LSN)
- These addresses can be used inside a carrier's own network but not on public internet. These addresses can be re-used inside other carrier's network ⇒ Shared.



©2013 Raj Jain

Dual NAT

- Also known as "**One Arm Mode**"
- Load balancer changes both source and destination addresses and destination port numbers on client requests. Like a regular NAT.
- Server sends the response to load balancer, Load balancer uses the port # to find the client address and forwards it to clients
- Client and servers can be on the same subnet. Servers can not see real client addresses.

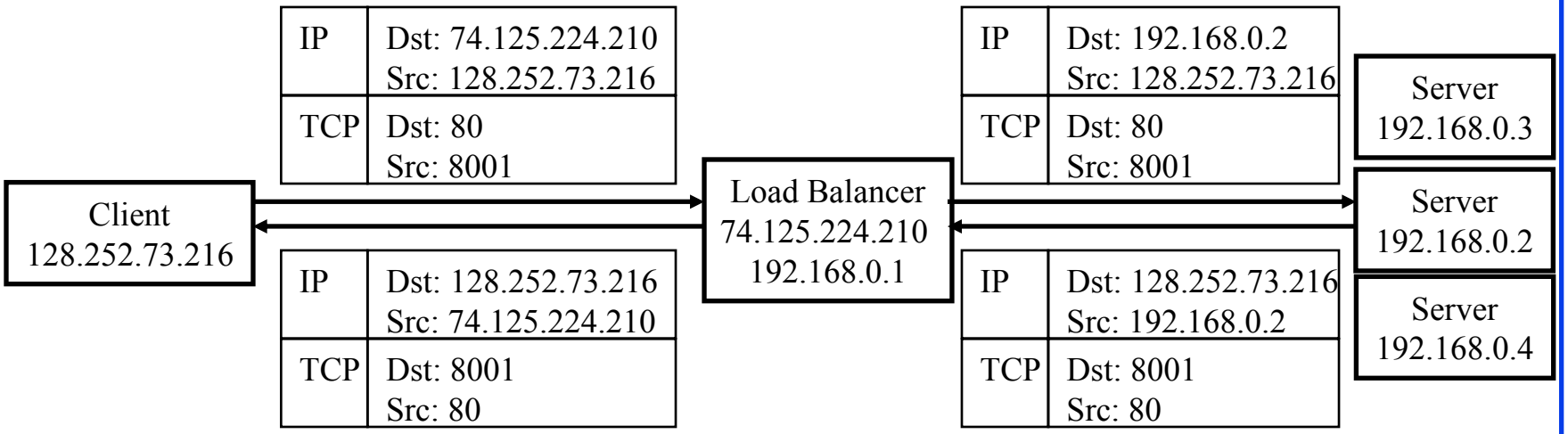


Ref: P. Srisuresh and D. Gan, "Load Sharing using IP NAT (LSNAT)," RFC 2391, Aug 1998, http://tools.ietf.org/html/rfc2391

©2013 Raj Jain

Server NAT

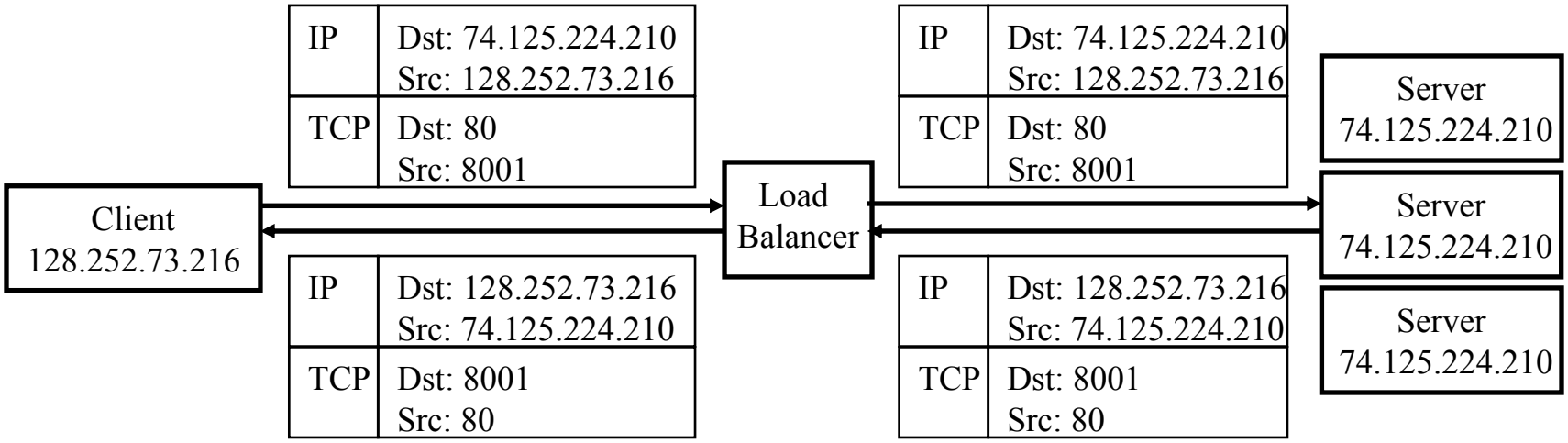
- Also known as "**Routed Mode**"
- Load balancer changes destination IP on client requests
  No changes to port numbers.
- Server sends the response directly to client
- Load balancer is made the default gateway.
  ⇒ All server responses go through the load balancer
  Load balancer changes the source IP addresses on responses.
- Works only if clients and servers are on different subnets.
  Servers can see the real client addresses (Good for security).



©2013 Raj Jain

Transparent Mode

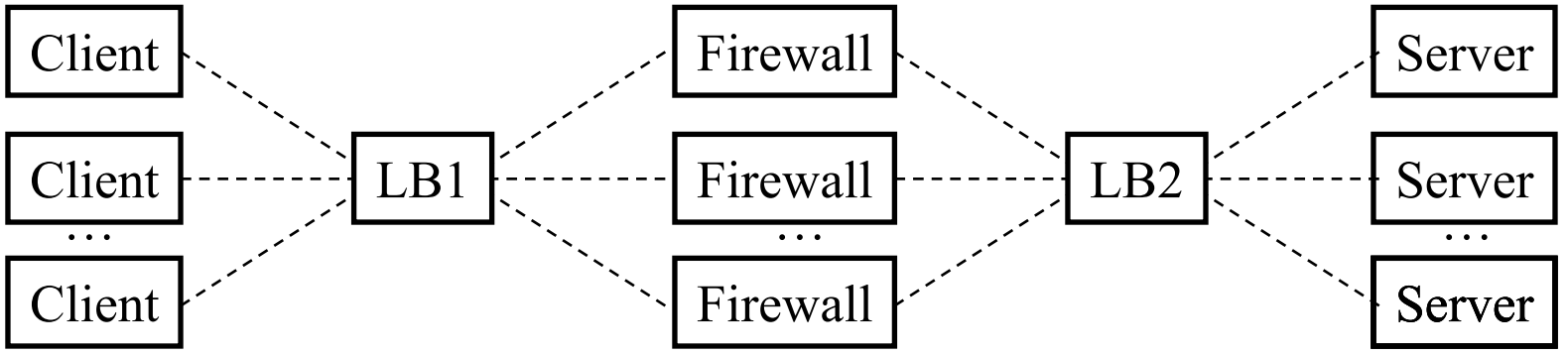
- Also known as "**Direct Routing Mode**"
- All servers are programmed to respond to the *same* Virtual IP adr but not respond to ARP for that VIP. They have different real IP adrs for other purposes and respond to ARP for that RIP.
- Load balancer is the sole layer 2 switch to server subnet.
  It does not changes anything at L3 or L4 ⇒ *Transparent*
  It simply changes the MAC address and directs the frame to the selected server.



©2013 Raj Jain

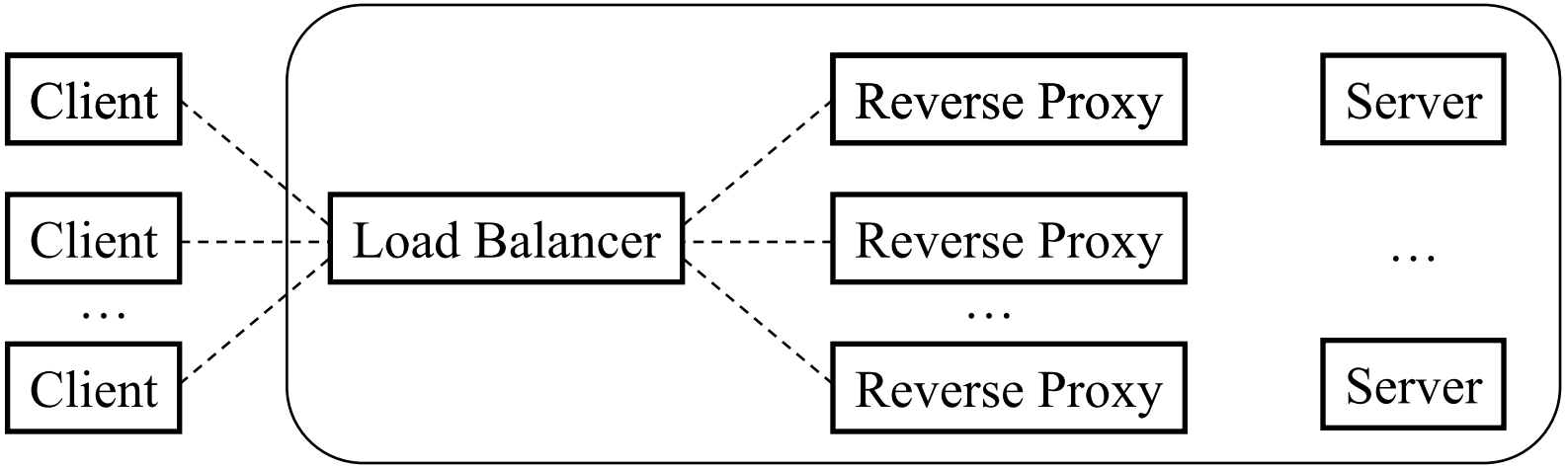
Firewall Load Balancing

- They need to see real IP addresses of clients
  ⇒ Transparent mode (balance via MAC addresses)
- Firewalls are stateful ⇒ Return traffic through the same firewall
- Solution: Complementary hashing using return load balancer.
  LB1 hashes source IP and/or Port to select firewall instance
  LB2 hashes destination IP and/or Port to select firewall instance.
  This is also known as a "**Firewall Sandwich**"



Reverse Proxy Load Balancing

- Proxies used for security, caching, application acceleration
- Proxies are usually used for *outgoing* requests
- Reverse proxies are used for *incoming* requests
- Usually transparent mode is used
- Load balancing using predictors such as hashing or round robin



©2013 Raj Jain

SSL Offload

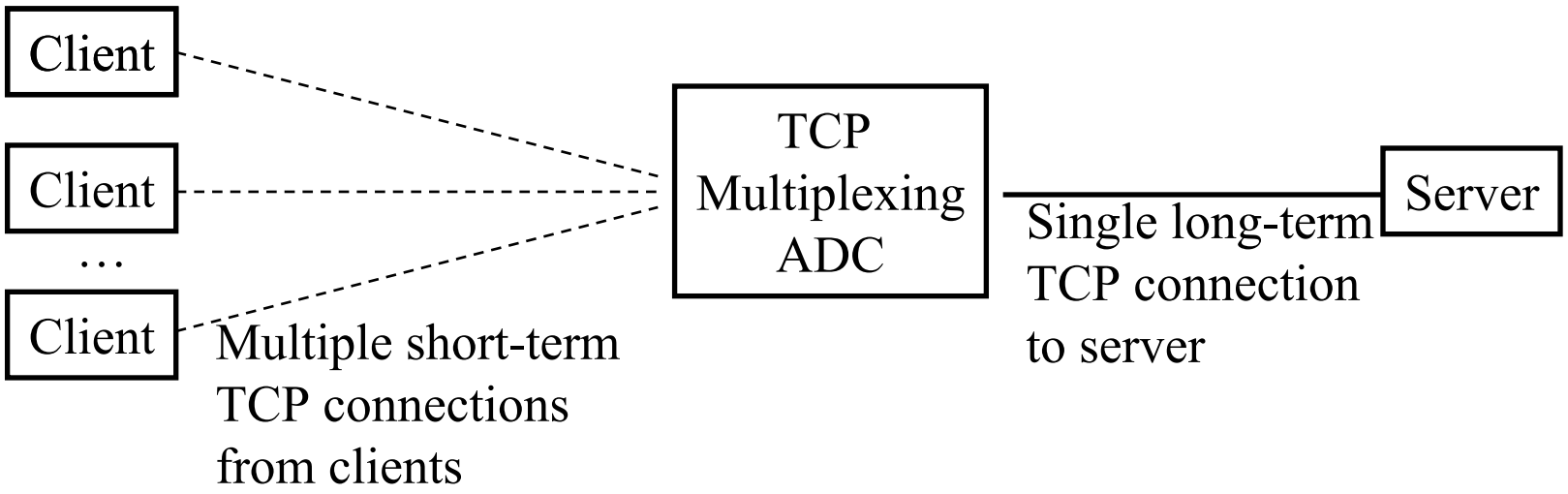
- Secure Socket Layer (SSL) and Transport Layer Security (TLS) are used for secure connections, e.g., https://
- Secure ⇒ Application data is encrypted.
- Some load balancer provide SSL offload

1. **SSL Termination**: Client to LB is secure. LB to Servers is clear. LB can select servers based on application data.
2. **SSL Initiation**: Client to LB is clear. LB to servers is secure.
   - Used when the client is local but the server is remote.
   - Helps clients by SSL offload.
   - One certificate can be used for all clients ⇒ Save cost

3. **End-to-end SSL**: Client-to-LB, LB-to-servers both secure. Internal encryption can be simpler. One Certificate for all servers.

TCP Multiplexing ADC

- TCP connections require 3-way handshake, ack for segments, sliding window flow control, congestion control, and termination for each connection
- A ADC can be used to reduce the number of TCP connections.
- Also, helps reduce total latency since LB to server connection runs at high window levels.



©2013 Raj Jain

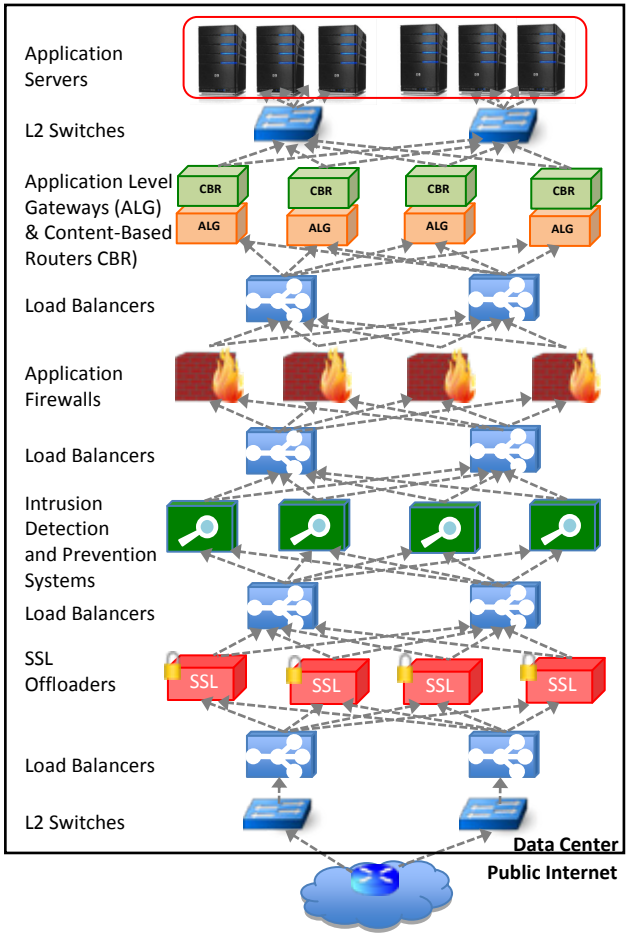
HTTP Compression

- Most http responses are compressed to reduce bandwidth usage
- A load balancer can offer this service and relieve the servers
- This is just one example of numerous other application acceleration techniques

©2013 Raj Jain

ADC Proliferation in Data Centers

- Application logic in servers
- Security (firewall, intrusion detection, SSL offload) in middle boxes
- Performance optimization (WAN optimizers, content caches) middleboxes
- Application-level policy routing (APR): Partitioning and replication middleboxes



©2013 Raj Jain

ADC Proliferation (Cont)

- Number of middleboxes (Application Delivery Controllers) is comparable to the number of routers
- Market size of optimization ADCs will grow from 1.5B in 2009 to $2.24B in 2013
- Security appliances will grow from $1.5B in 2010 to $10B in 2016

Appliance Type	Number
Firewalls	166
NIDS	127
Conferencing/Media Gateways	110
Load Balancers	67
Proxy Caches	66
VPN devices	45
WAN optimizers	44
Voice Gateways	11
Middleboxes total	636
Routers	~ 900

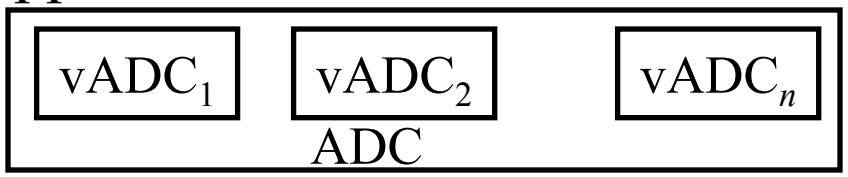
Ref: Technavio, "Global Application Delivery Controllers Market in Datacenters 2009-2013," March, 2010, http://www.technavio.com/content/global-application-delivery-controllersmarket-datacenters-2009-2013

Ref: Vyas Sekar, *et. al.,* "The Middlebox Manifesto: Enabling Innovation in Middlebox Deployments," ACM HotNets 2011.

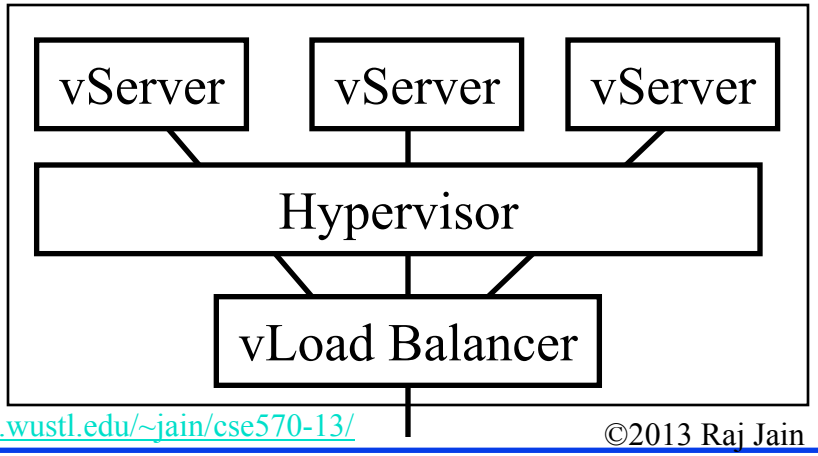
©2013 Raj Jain

Virtual ADCs

- **Hardware Based vADCs**:
  - A single physical ADC presents multiple virtual contexts
  - Each virtual ADC can be used by a different tenant or different application



- **Software Based vADCs**: Inside or outside a physical server.
  - Run on standard processors (pADCs generally run on monolithic hardware)

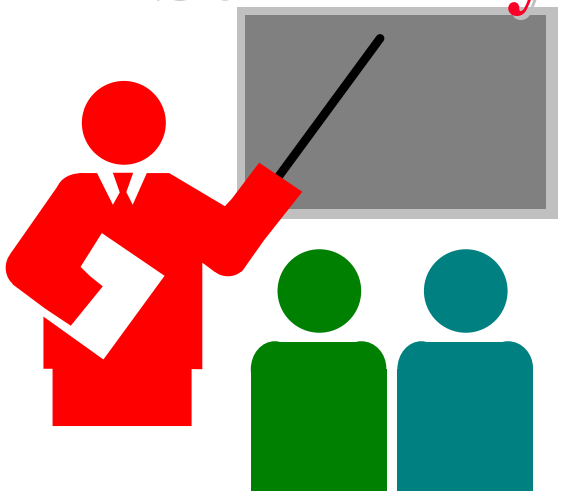


Multi-Function ADCs

- Combine in one ADC:
  - Protocol optimization
  - Location based DNS
  - Load balancing
  - Security
  - Data compression

©2013 Raj Jain

Summary



1. Application delivery involves partitioning an application based on server context, user context, application context
2. Load balancer was the first application delivery controller. Now ADCs are used for many common services such as firewalls, TCP multiplexing, proxy, etc.
3. Many different flavors of NAT are used by ADCs.
4. ADCs are becoming virtual and multi-function.

©2013 Raj Jain

Reading List

- G. Santana, "Datacenter Virtualization Fundamentals," Cisco Press, 2014, pp. 110-139, ISBN: 1587143240 (Safari Book) (*Must Read*)

©2013 Raj Jain

Wikipedia Links

- http://en.wikipedia.org/wiki/Address_Resolution_Protocol
- http://en.wikipedia.org/wiki/Application_delivery_controller
- http://en.wikipedia.org/wiki/Application_delivery_network
- http://en.wikipedia.org/wiki/Carrier-grade_NAT
- http://en.wikipedia.org/wiki/Delayed_binding
- http://en.wikipedia.org/wiki/Domain_Name_System
- http://en.wikipedia.org/wiki/HTTP_compression
- http://en.wikipedia.org/wiki/Load_balancing_(computing)
- http://en.wikipedia.org/wiki/Middlebox
- http://en.wikipedia.org/wiki/Network_address_translation
- http://en.wikipedia.org/wiki/Proxy_server
- http://en.wikipedia.org/wiki/Reverse_proxy
- http://en.wikipedia.org/wiki/Round-robin_DNS
- http://en.wikipedia.org/wiki/Secure_Socket_Layer
- http://en.wikipedia.org/wiki/SSL_acceleration
- http://en.wikipedia.org/wiki/Virtual_IP_address

©2013 Raj Jain

Acronyms

- ADC Application Delivery Controller
- ALG Application Level Gateways
- API Application Programming Interface
- APR Application-level Policy Routing
- ARP Address Resolution Protocol
- CBR Content Based Router
- CGNAT Carrier Grade Network Address Translation
- DNS Domain Name System
- GSLB Global Service Load Balancer
- HTTP Hypertext Transfer Protocol
- IANA Internet Addressing and Naming Authority
- IP Internet Protocol

©2013 Raj Jain

Acronyms (Cont)

- IPv4 Internet Protocol version 4
- IPv6 Internet Protocol version 6
- ISP Internet Service Provider
- LB Load Balancing
- MAC Media Access Control
- NAT Network Address Translation
- NAT44 NAT IPv4 private to IPv4 public
- NAT444 NAT IPv4 private to IPv4 public via IPv4 Carrier
- NAT46 NAT IPv4 private to IPv6 public
- NAT64 NAT IPv6 private to IPv4 public
- NIDS Network Intrusion Detection Systems

©2013 Raj Jain

Acronyms (Cont)

- pADC Physical Application Delivery Controller
- RIP Real IP Address
- SSL Secure Socket Layer
- TCP Transmission Control Protocol
- TLS Transport Layer Security
- vADCs Virtual Application Delivery Controller
- VIP Virtual IP address
- VPN Virtual Private Network
- vServer Virtual Server
- WAN Wide Area Network

Quiz 8

1. Video messages going to one server group while accounting to another is an example of _____-based partitioning
2. Windows users being served by a different server group than iphone user is an example of _____-context based partitioning.
3. A centrally managed set of servers working together on one application is called server _____
4. _____ are synthetic request sent by load balancer to see if a server is up.
5. Load balancing rules are called _____
6. Layer _____ switching requires delayed binding.
7. One problem with load balancing using _____ is that it does not know whether a server is up.
8. Global server load balancing is good mostly for _____ use.
9. In a NAT64 system, the private network is _____ and the public network is _____.
10. 100.64.0.0/10 has been allocated as _____ address range.

©2013 Raj Jain

Solution to Quiz 8

1. Video messages going to one server group while accounting to another is an example of *content*-based partitioning
2. Windows users being served by a different server group than iphone user is an example of *user*-context based partitioning.
3. A centrally managed set of servers working together on one application is called server *Cluster*
4. *Probes* are synthetic request sent by load balancer to see if a server is up.
5. Load balancing rules are called *Predictor*
6. Layer *7* switching requires delayed binding.
7. One problem with load balancing using *rotating DNS* is that it does not know whether a server is up.
8. Global server load balancing is good mostly for *intra-company* use.
9. In a NAT64 system, the private network is *IPv6* and the public network is *IPv4*.
10. 100.64.0.0/10 has been allocated as *shared* address range.

©2013 Raj Jain