CSE 124:
NETWORKING FUNDAMENTALS

George Porter
Sept 27, 2022

UC San Diego

ATTRIBUTION

- These slides are released under an Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0) Creative Commons license

AGENDA

Today:
- Basics of networking, addressing, performance metrics

Tuesday:
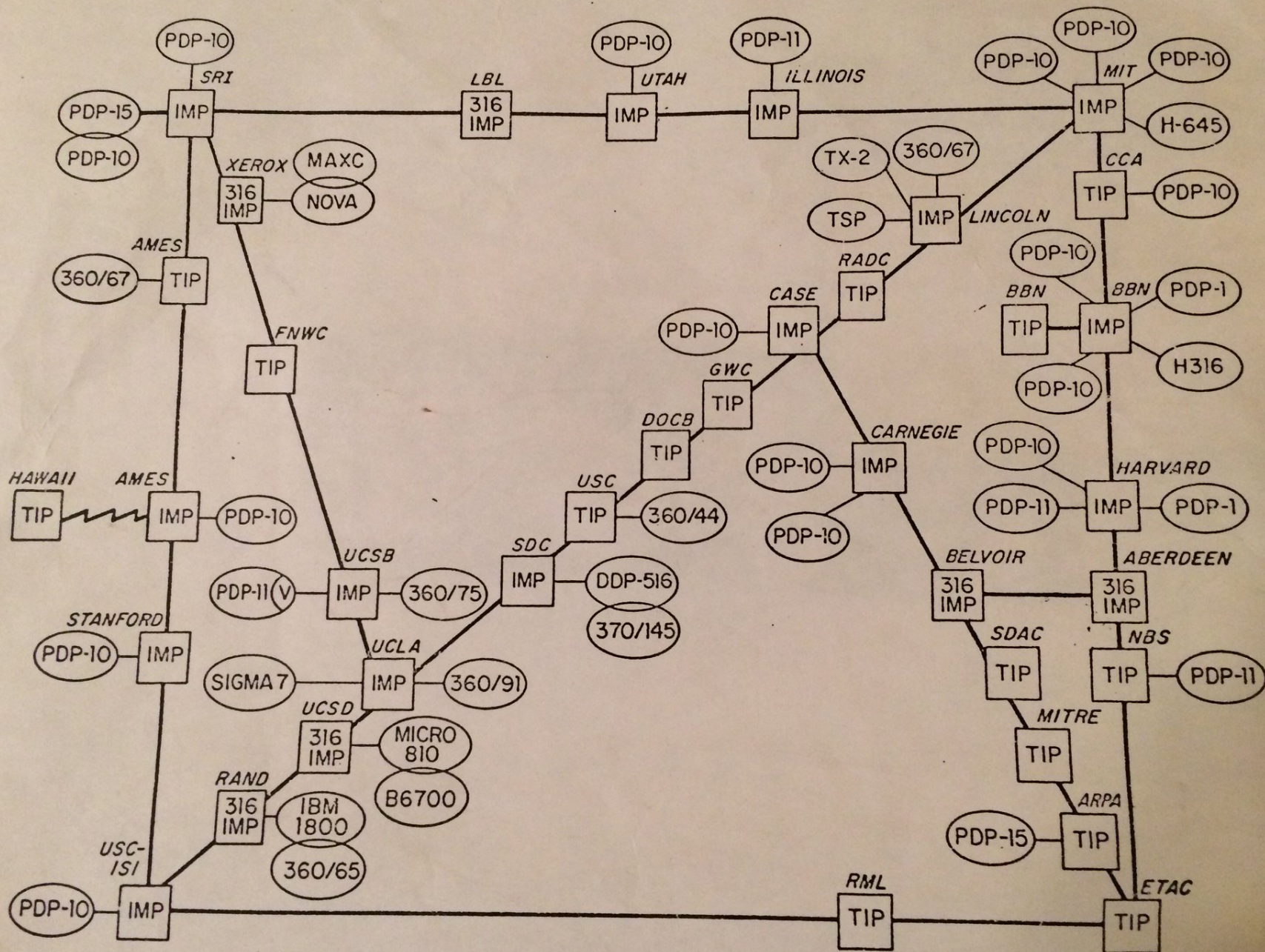- Basics of Go, naming and addressing

TODO THIS WEEK

1. Read Network Programming with Go chapters 1, 2, and 3
2. Project 1 is now out (due Fri Sept 30 at 5pm)

BRIEF HISTORY OF THE INTERNET

- 1968 - DARPA (Defense Advanced Research Projects Agency) contracts with BBN (Bolt, Beranek & Newman) to create ARPAnet

- 1970 - First five nodes:
  - UCLA
  - Stanford
  - UC Santa Barbara
  - U of Utah, and
  - BBN

- 1974 - TCP specification by Vint Cerf

- 1984 – On January 1, the Internet with its 1000 hosts converts en masse to using TCP/IP for its messaging

Data from the Internet Society

ARPA NETWORK, LOGICAL MAP, MAY 1973

SRI — IMP — PDP-10, PDP-15, PDP-10

LBL — 316 IMP

UTAH — IMP — PDP-10

ILLINOIS — IMP — PDP-11

MIT — IMP — PDP-10, PDP-10, PDP-10, H-645

XEROX — 316 IMP — MAXC, NOVA

AMES — TIP — 360/67

FNWC — TIP

LINCOLN — IMP — TX-2, 360/67, TSP

CCA — TIP — PDP-10

RADC — TIP

CASE — IMP — PDP-10

BBN — TIP

BBN — IMP — PDP-10, PDP-1, H316, PDP-10

GWC — TIP

DOCB — TIP

CARNEGIE — IMP — PDP-10, PDP-10

HARVARD — IMP — PDP-10, PDP-11, PDP-1

HAWAII — TIP

AMES — IMP — PDP-10

USC — TIP — 360/44

SDC — IMP — DDP-516, 370/145

UCSB — IMP — PDP-11(V), 360/75

STANFORD — IMP — PDP-10

UCLA — IMP — SIGMA 7, 360/91

BELVOIR — 316 IMP

SDAC — TIP

ABERDEEN — 316 IMP

NBS — TIP — PDP-11

UCSD — 316 IMP — MICRO 810, B6700

MITRE — TIP

RAND — 316 IMP — IBM 1800, 360/65

USC-ISI — IMP — PDP-10

ARPA — TIP — PDP-15

RML — TIP

ETAC — TIP

Outline

1. Performance

2. Layering

3. Addressing

PERFORMANCE METRICS

Bandwidth

Latency

- Bandwidth: number of bits transmitted per unit of time

- Latency = Propagation + Transmit + Queue

  - Propagation = Distance/SpeedOfLight(*)

  - Transmit = 1 bit/Bandwidth

  - Queue = Time waiting in switches/routers behind other traffic (traffic jam)

- Overhead

  - # secs for CPU to put message on wire

- Error rate

  - Probability P that message will not arrive intact

* In that particular medium

BANDWIDTH VS. LATENCY

1 Byte Object

	Latency: 1 ms	Latency: 100 ms
Bandwidth: 1 Mbps	1,008 μs	100,008 μs
Bandwidth: 100 Mbps	1,000 μs	100,000 μs

10 MB Object

	Latency: 1 ms	Latency: 100 ms
Bandwidth: 1 Mbps	80.001 s	80.1 s
Bandwidth: 100 Mbps	.801 s	.9 s

NETWORK PERFORMANCE MEASUREMENT UNITS

Exp.	Explicit	Prefix	Exp.	Explicit	Prefix
10^{-3}	0.001	milli	10^3	1,000	Kilo
10^{-6}	0.000001	micro	10^6	1,000,000	Mega
10^{-9}	0.000000001	nano	10^9	1,000,000,000	Giga
10^{-12}	0.000000000001	pico	10^{12}	1,000,000,000,000	Tera
10^{-15}	0.000000000000001	femto	10^{15}	1,000,000,000,000,000	Peta
10^{-18}	0.000000000000000001	atto	10^{18}	1,000,000,000,000,000,000	Exa
10^{-21}	0.000000000000000000001	zepto	10^{21}	1,000,000,000,000,000,000,000	Zetta
10^{-24}	0.000000000000000000000001	yocto	10^{24}	1,000,000,000,000,000,000,000,000	Yotta

TERMINOLOGY STYLE

- Mega versus Mega, Kilo versus Kilo

  - Computer architecture: Mega → 2^20, Kilo → 2^10

  - Computer networks: Mega → 10^6, Kilo → 10^3

- Mbps versus MBps

  - Networks: typically megabits per second

  - Architecture: typically megabytes per second

- Bandwidth versus throughput

  - Bandwidth: available over link

  - Throughput: available to application

    - E.g. subtract protocol headers, etc.

PERFORMANCE TOOLS

- Ping

  - Test if other side is "alive"

  - Measures round-trip latency

- iperf3

  - Times how long it takes to send N bytes to the other endpoint

  - Used to calculate bandwidth

Outline

1. ~~Performance~~

2. Layering

3. Addressing

NETWORKING AND LAYERING

"All problems in computer science can be solved by another level of indirection."
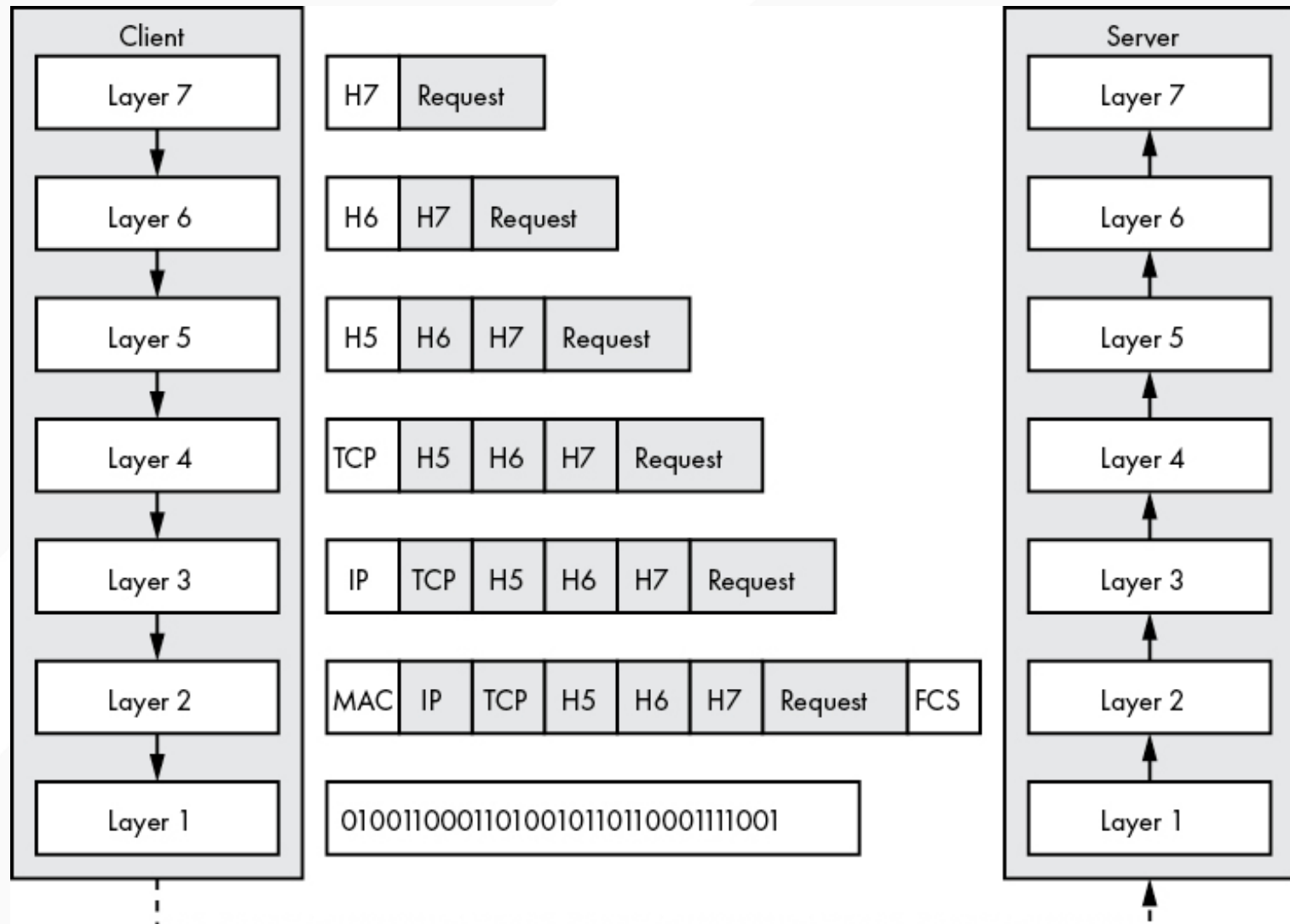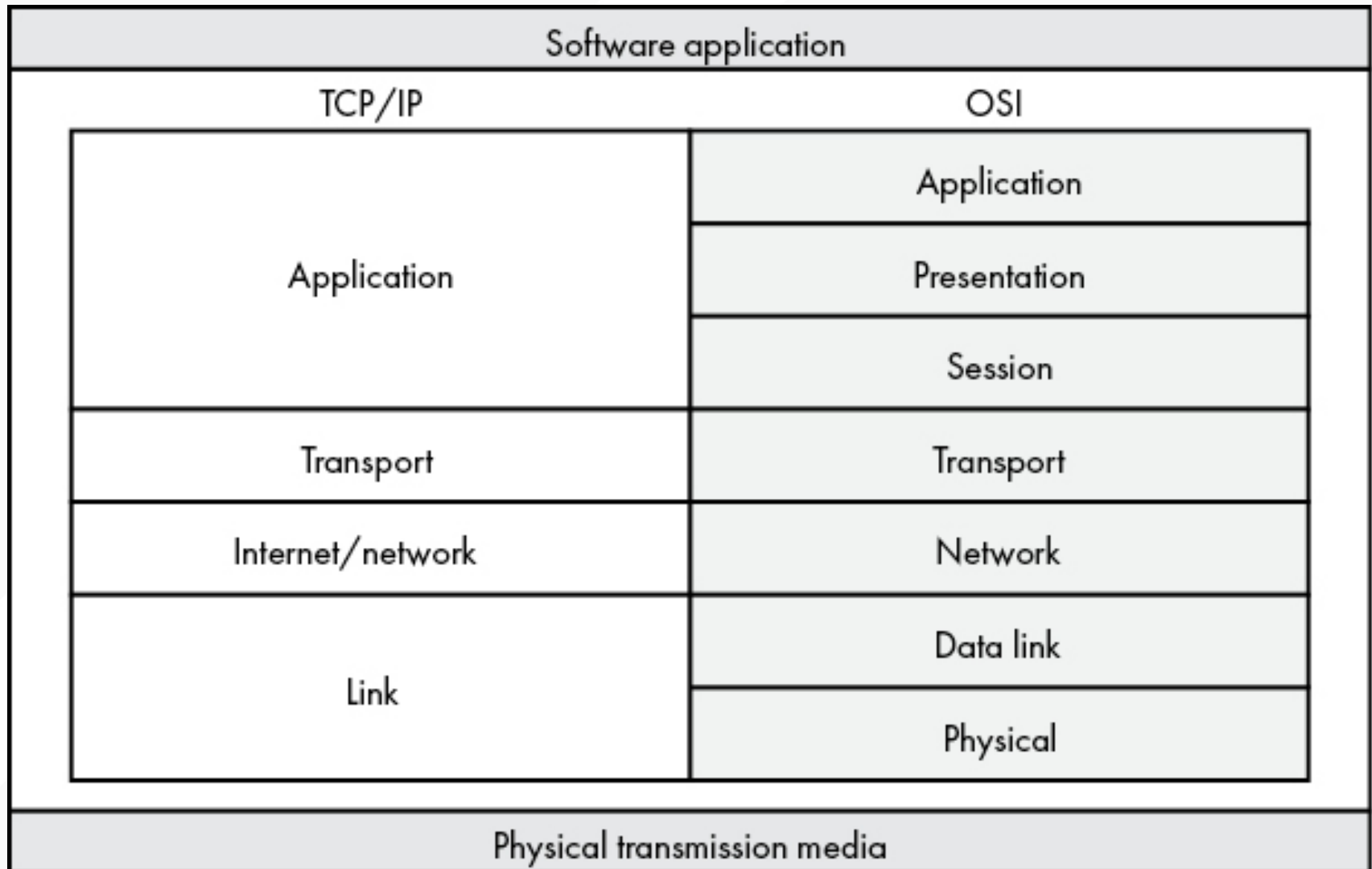
David J. Wheeler

OSI NETWORK STACK

	Software application		
	Network protocol stack		
Layer 7	Application		
Layer 6	Presentation		
Layer 5	Session		
Layer 4	Transport		
Layer 3	Network		
Layer 2	Data link	Logical link control	
		Media access control	
Layer 1	Physical		
	Physical transmission media		
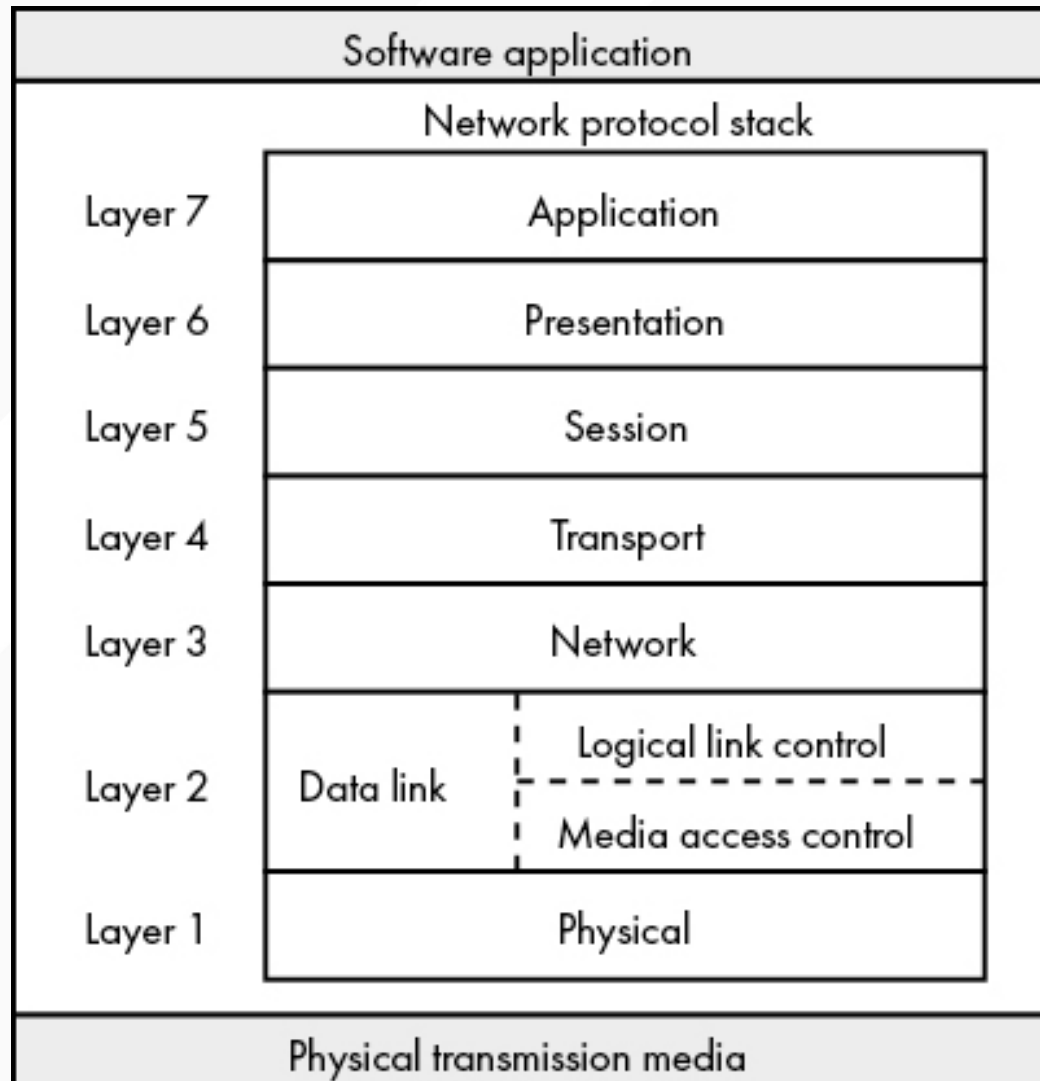
PROTOCOLS GOVERN MESSAGES EXCHANGED WITHIN A SINGLE LAYER
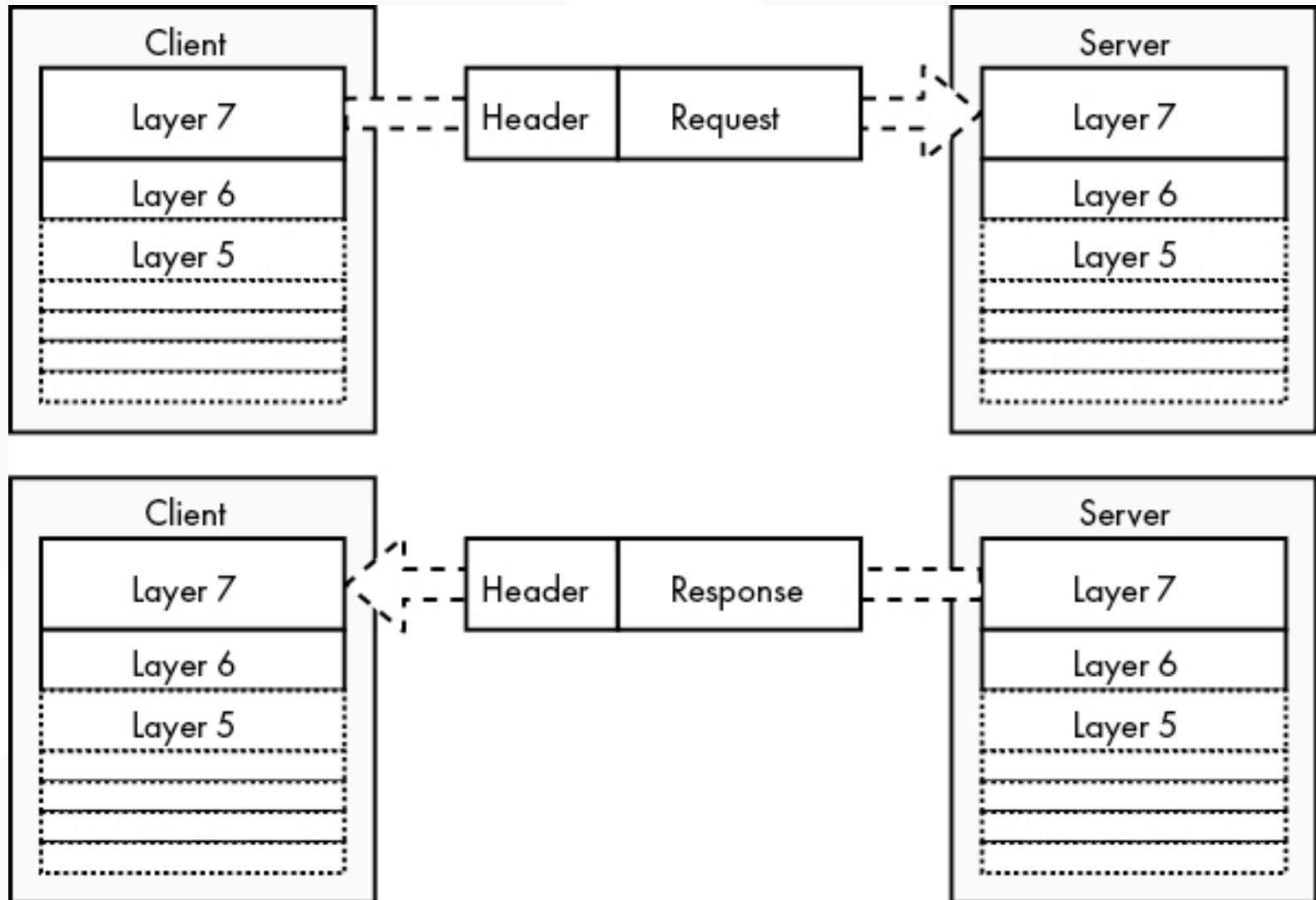
LAYERING AND ENCAPSULATION

TCP/IP MODEL (VS OSI MODEL)

Software application	
TCP/IP	**OSI**
Application	Application
	Presentation
	Session
Transport	Transport
Internet/network	Network
Link	Data link
	Physical
Physical transmission media	

TCP/IP MODEL (VS OSI MODEL)

Forwarding packets between subnetworks (E.g. IPv4 or IPv6)

Software application	
TCP/IP	**OSI**
Application	Application
	Presentation
	Session
Transport	Transport
Internet/network	Network
Link	Data link
	Physical
Physical transmission media	

Forwarding packets (frames) within one subnetwork (E.g. Ethernet)
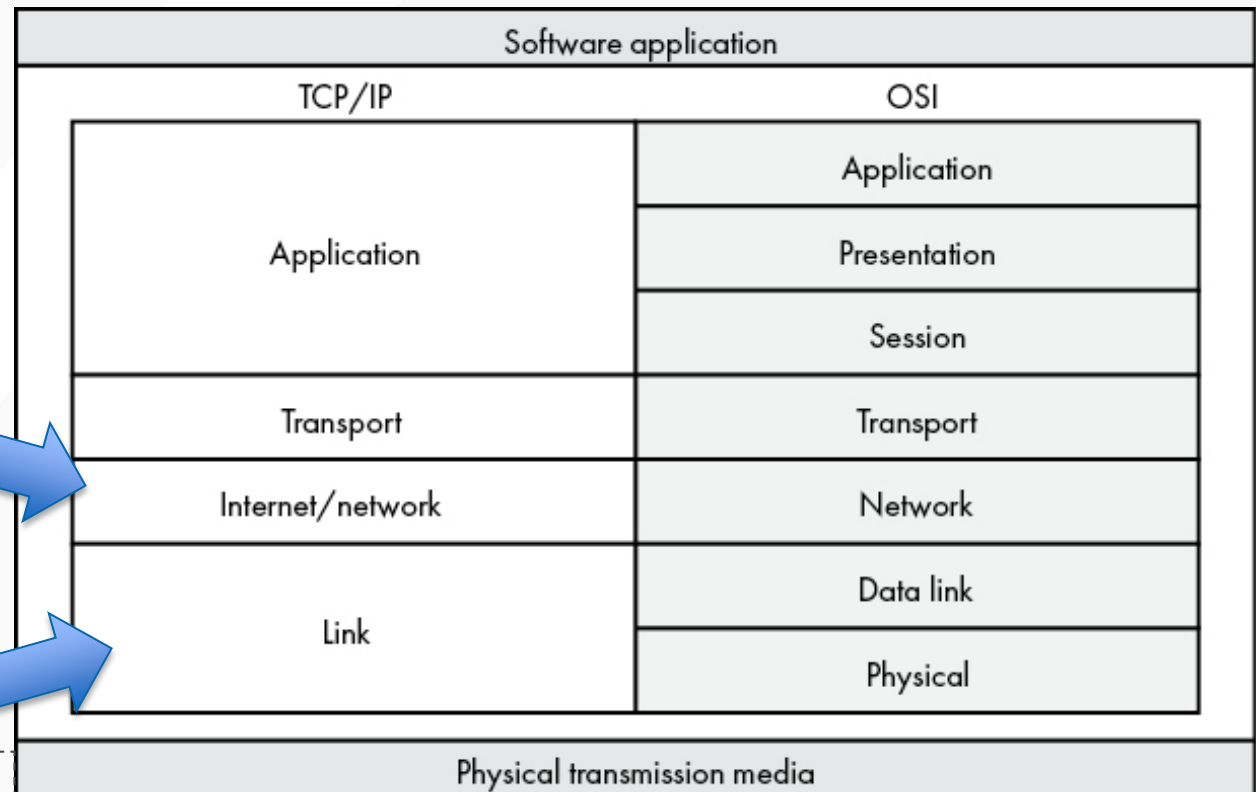
Outline

1. ~~Performance~~

2. ~~Layering~~

3. Addressing

MOTIVATION FOR OUR DISCUSSION OF ADDRESSING

1. When you implement a *client* application, you will typically need to communicate with a remote system/service via its *IP Address*

2. When you implement a cloud-hosted *network service*, you will typically be assigned a *subnet* of IP addresses to use for your various server programs, and will need to use and manage those appropriately

IP PROTOCOL

- Scenario: An internet connected device wants to send a message to another internet connected device anywhere in the world

  - IP protocol handles this

  - Prepend an "IP header" to the message, set the destination IP address, set the source IP address, and Internet routers will forward it along the shortest path till it gets to the destination network

  - From there, the destination network forwards the packet to the intended host

* Simplified model but sufficient for our purposes

IP VERSION 4 (IPV4)

| 11000000 | . | 10101000 | . | 00000001 | . | 00001010 | (Binary) |

| 192 | . | 168 | . | 1 | . | 10 | (Decimal) |

- 32-bits

- Usually represented as four 8-bit bytes (octets)

- E.g. 206.109.1.6, 127.0.0.1, 192.168.1.10

IP VERSION 6 (IPV6)

An IPv6 address   (in hexadecimal)
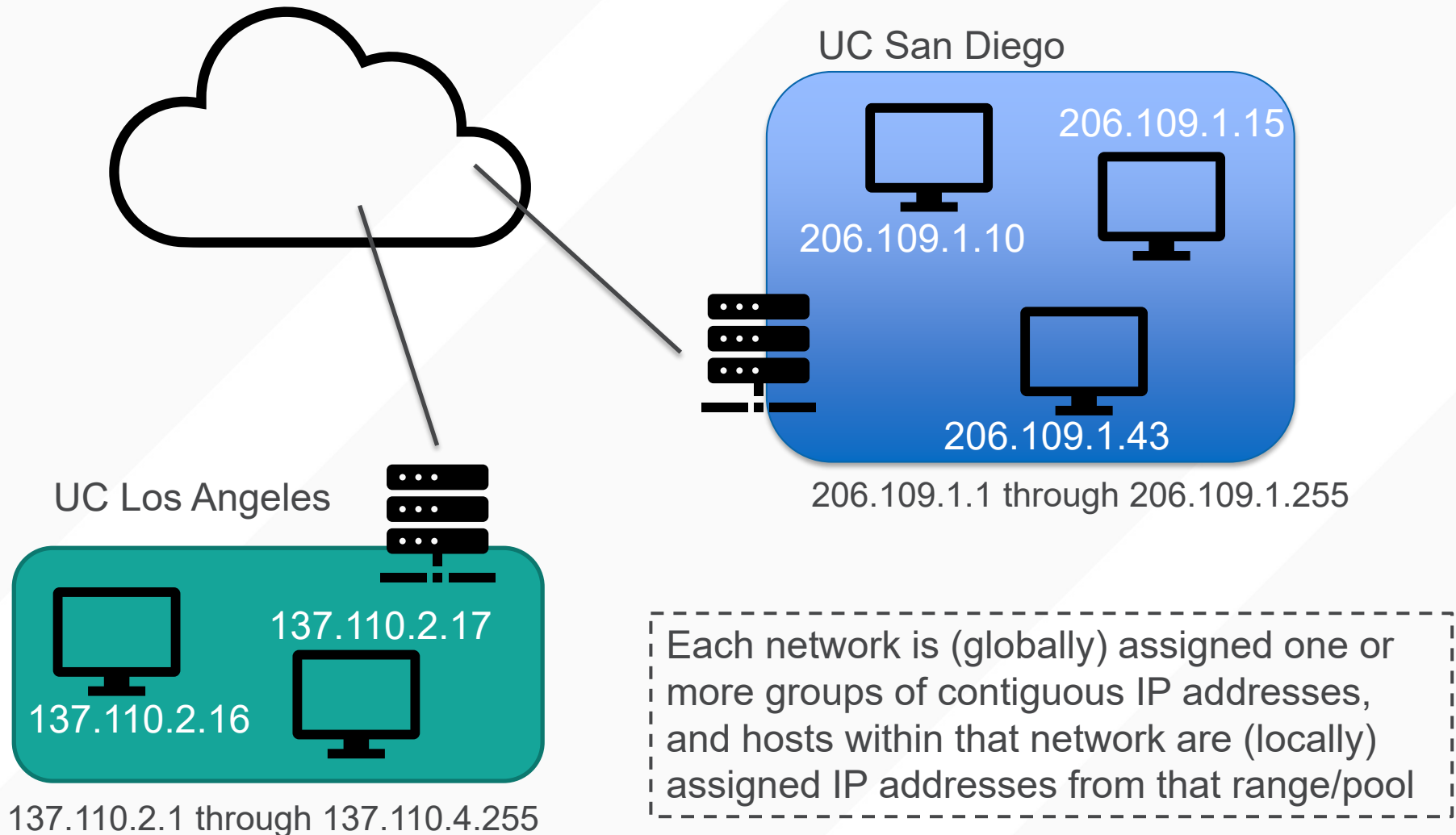
2001:0DB8:AC10:FE01:0000:0000:0000:0000

⬇   ⬇   ⬇   ⬇

2001:0DB8:AC10:FE01::   Zeroes can be omitted

0010000000000001:0000110110111000:1010110000010000:1111111000000001:

0000000000000000:0000000000000000:0000000000000000:0000000000000000

- 128-bits

- Represented as 8 16-bit blocks, in hex, separated by colons

- E.g. 2001:4860:4860::8888, ::1, 2345:0425:2CA1:0000:0000:0567:5673:23b5

THE INTERNET: A NETWORK OF SUB-NETWORKS

UC San Diego

206.109.1.15

206.109.1.10

206.109.1.43

206.109.1.1 through 206.109.1.255

UC Los Angeles

137.110.2.17

137.110.2.16

137.110.2.1 through 137.110.4.255

Each network is (globally) assigned one or more groups of contiguous IP addresses, and hosts within that network are (locally) assigned IP addresses from that range/pool

ROUTE AGGREGATION

- 2^32 possible (IPv4) addresses spread across ~100,000 independent networks

- Each router keeps the "next hop" on a per-network basis, not per-host basis

- But networks can be different sizes (e.g. UCSD is bigger than a small startup)

- Each router has to keep a list of networks (and their next hops) + how "big" each network is

CLASS-BASED ADDRESSING (NOT REALLY USED ANYMORE)

- Most significant bits determines "class" of address

| Class A | 0 | Network | Host | | 127 nets, 16M hosts |

| Class B | 1 | 0 | Network (14) | Host (16) | 16K nets, 64K hosts |

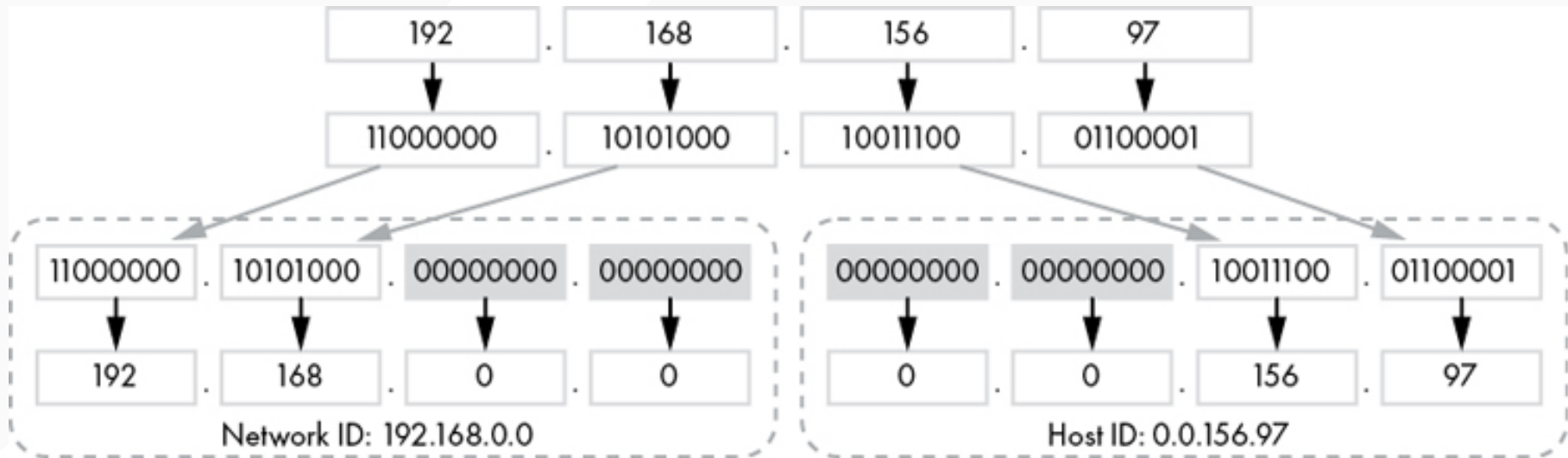| Class C | 1 | 1 | 0 | Network (21) | Host (8) | 2M nets, 254 hosts |

- Special addresses

  - Class D (1110) for multicast, Class E (1111) experimental

  - 127.0.0.1: local host (a.k.a. the loopback address)

  - Host bits all set to 0: network address

  - Host bits all set to 1: broadcast address

CLASS-BASED ADDRESSING EXAMPLES

Network ID	First octet		Second octet		Third octet		Fourth octet	Host ID
8 bits	Network		Host		Host		Host	24 bits
	10		1		2		3	
16 bits	Network		Network		Host		Host	16 bits
	172		16		1		2	
24 bits	Network		Network		Network		Host	8 bits
	192		168		1		2	

ADDRESSING EXAMPLE
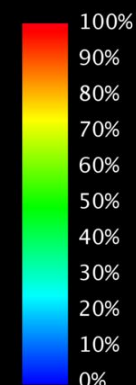
IP ADDRESS PROBLEM (1991)

- Address space depletion

  - In danger of running out of classes A and B

- Why?

  - Class C too small for most organizations (only ~250 addresses)

  - Very few class A – very careful about giving them out (who has 16M hosts anyway?)

  - Class B – greatest problem
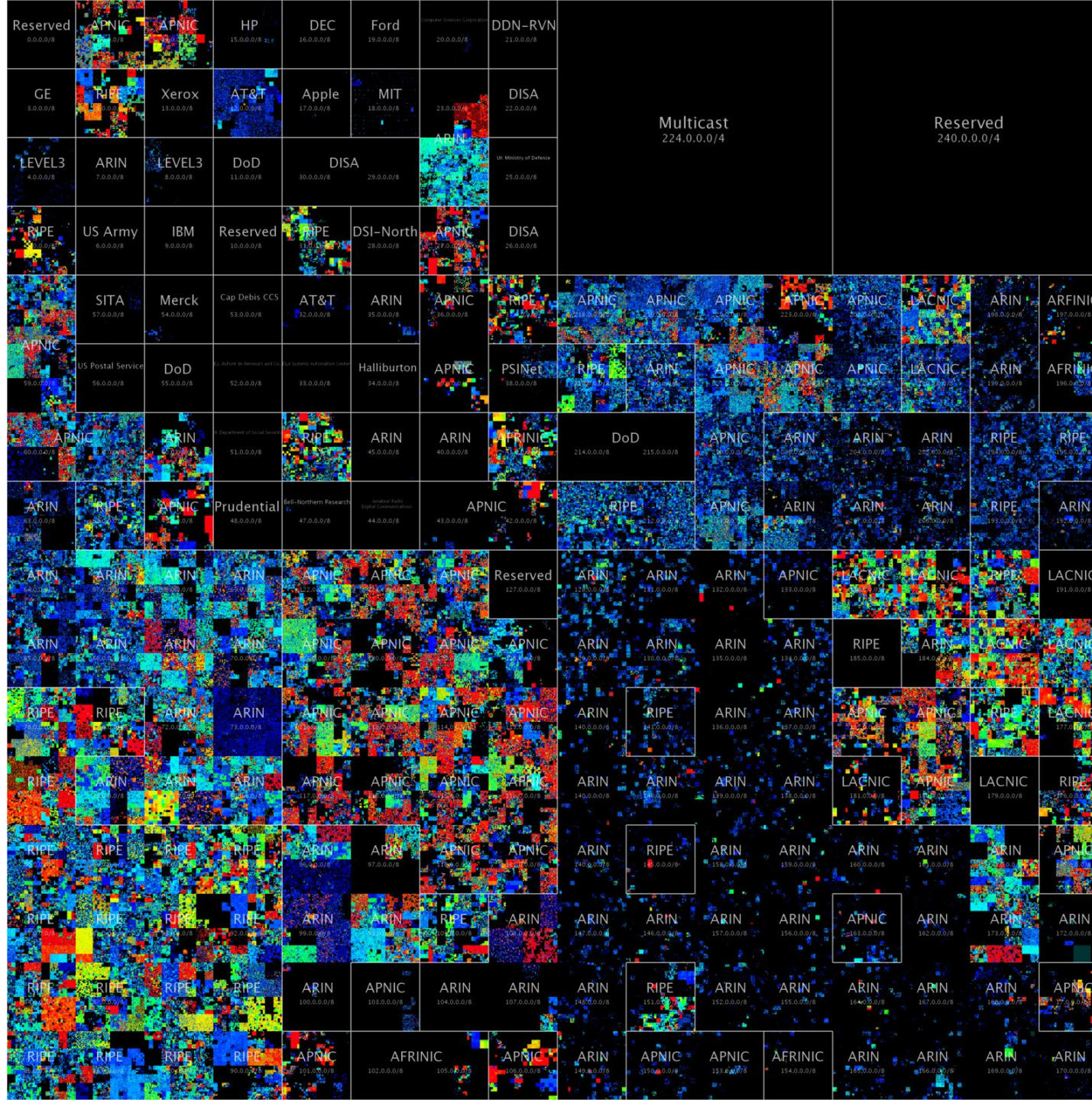
IPv4 Census Map
June – October 2012

Utilization

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%

Prefix Sizes

= /8

= /12

= /16

= /20

= /24

420 Million hosts that responded to ICMP Ping
at least 2 times between June and October 2012
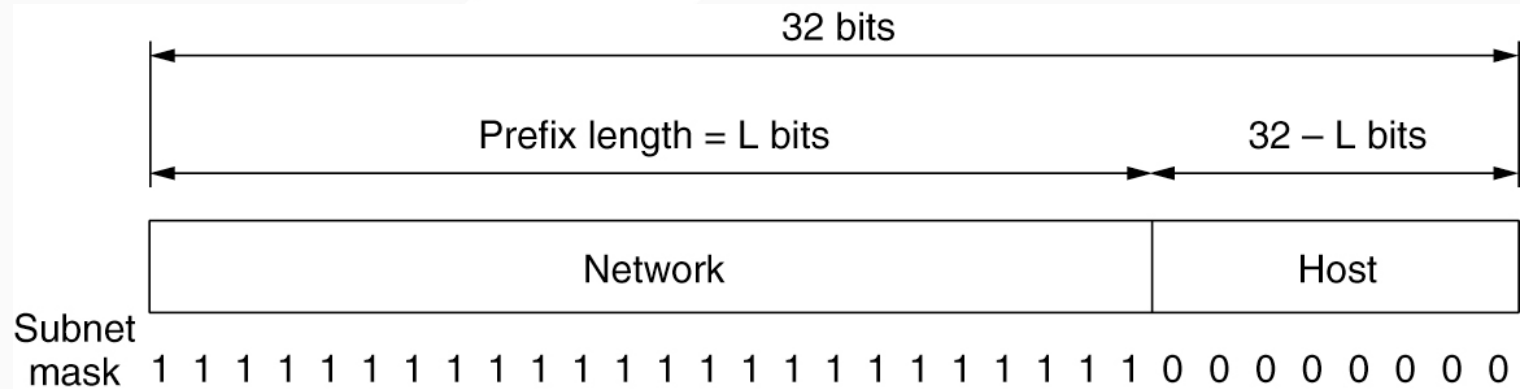Source: Carna Botnet

CIDR

- Classless Inter-Domain Routing (1993)

  - Networks described by variable-length prefix and length

  - Allows arbitrary allocation between network and host address

Network	Host

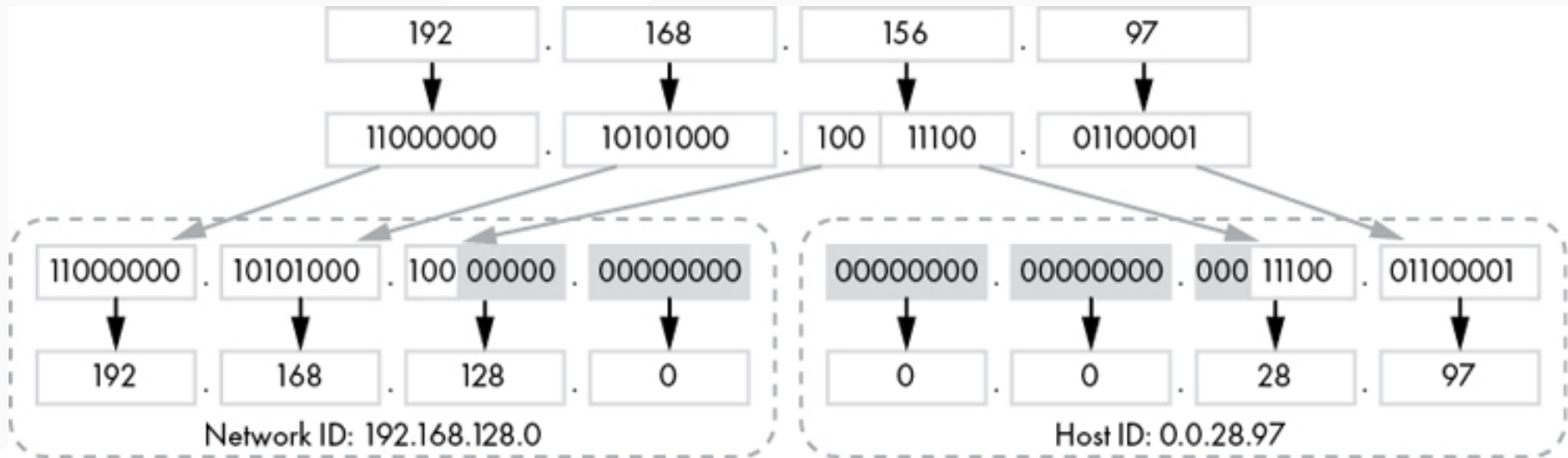 Prefix   Mask=# significant bits representing prefix

  - e.g. 10.95.1.2 contained within 10.0.0.0/8:

    - 10.0.0.0 is network and remainder (95.1.2) is host

- Pro: Finer grained allocation; aggregation
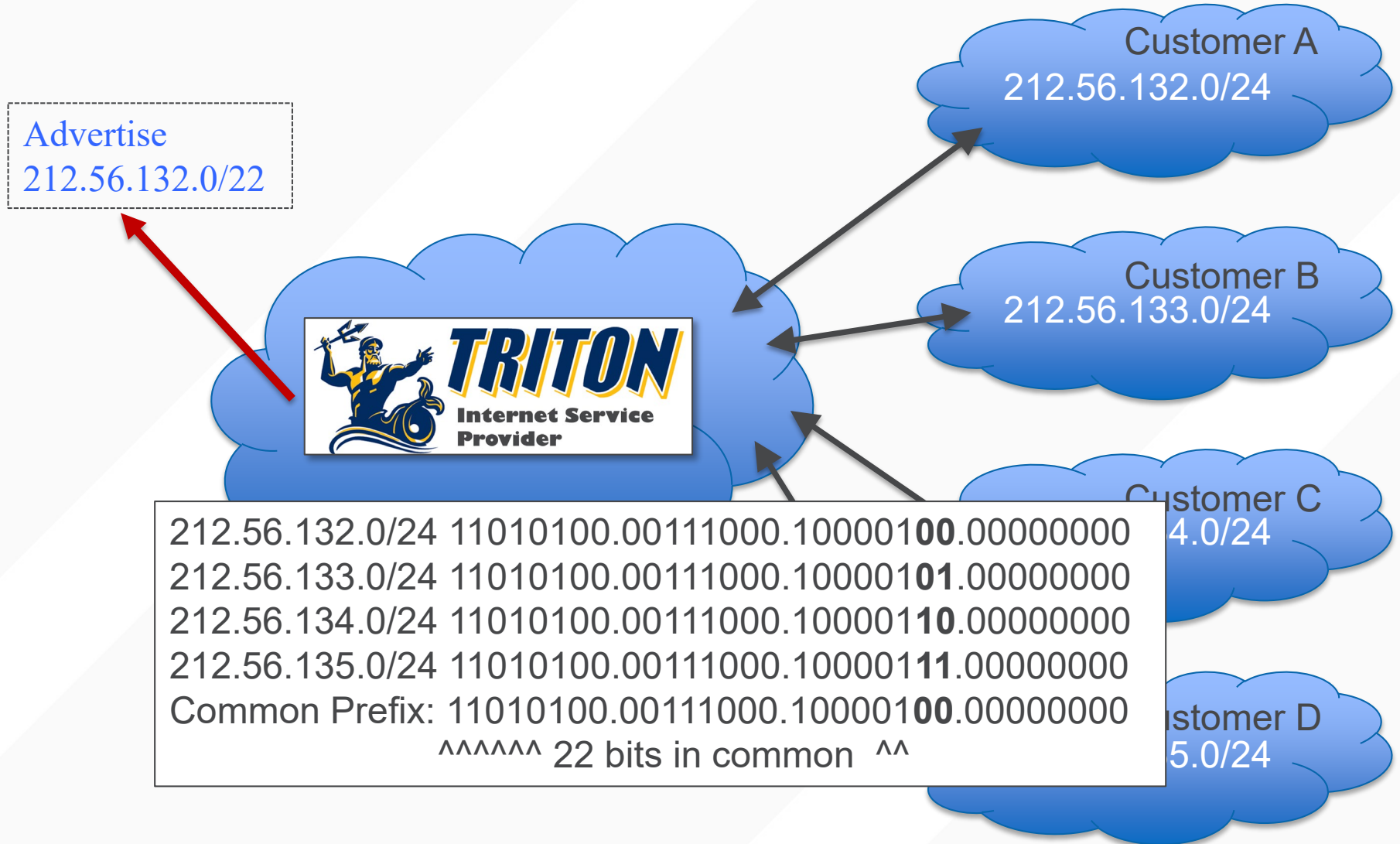
- Con: More expensive lookup: longest prefix match

SUBNETS AND NETMASKS

192.168.156.97/19

| 192 | . | 168 | . | 156 | . | 97 |

| 11000000 | . | 10101000 | . | 100 | 11100 | . | 01100001 |

Network ID: 192.168.128.0

| 11000000 | . | 10101000 | . | 100 | 00000 | . | 00000000 |
| 192 | . | 168 | . | 128 | . | 0 |

Host ID: 0.0.28.97

| 00000000 | . | 00000000 | . | 000 | 11100 | . | 01100001 |
| 0 | . | 0 | . | 28 | . | 97 |

ADDRESS AGGREGATION EXAMPLE

Advertise
212.56.132.0/22

Customer A
212.56.132.0/24

Customer B
212.56.133.0/24

Customer C
.4.0/24

Customer D
.5.0/24

TRITON
Internet Service
Provider

212.56.132.0/24 11010100.00111000.100001**00**.00000000
212.56.133.0/24 11010100.00111000.10000**01**.00000000
212.56.134.0/24 11010100.00111000.10000**10**.00000000
212.56.135.0/24 11010100.00111000.10000**11**.00000000
Common Prefix: 11010100.00111000.10000**00**.00000000
^^^^^^ 22 bits in common ^^

ROUTING TABLE EXAMPLE (R2)

- Packet to 10.1.1.6

- Matches 10.1.0.0/23

Destination	Next Hop
127.0.0.1	loopback
Default or 0/0	10.1.16.1
10.1.8.0/24	interface1
10.1.2.0/23	interface2
10.1.0.0/23	10.1.2.2
10.1.16.0/24	interface3
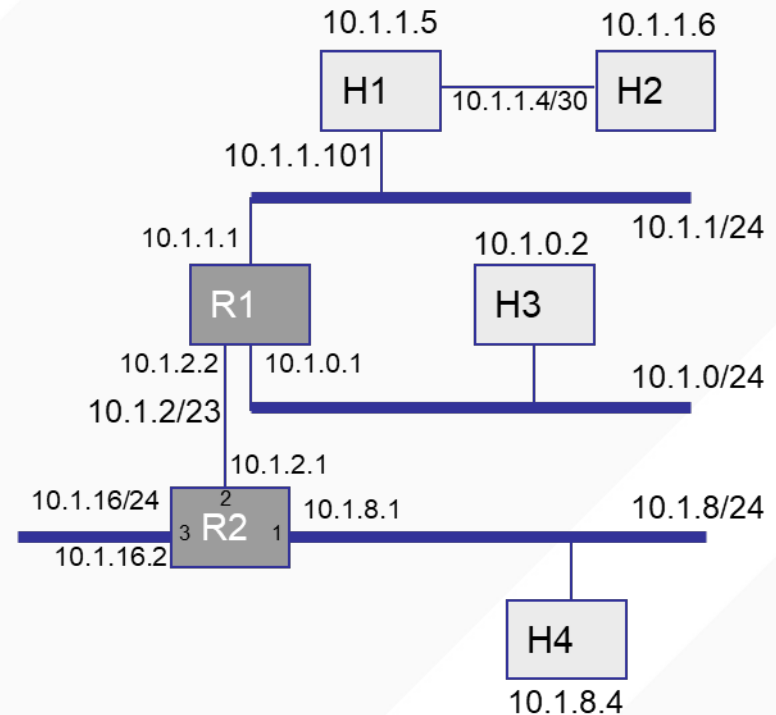
- Packet to 10.1.1.6

- Matches 10.1.1.4/30

  - Longest prefix match

Routing table at R1

Destination	Next Hop
127.0.0.1	loopback
Default or 0/0	10.1.2.1
10.1.0.0/24	interface1
10.1.1.0/24	interface2
10.1.2.0/23	interface3
10.1.1.4/30	10.1.1.101

10.1.1.5 10.1.1.6
H1 10.1.1.4/30 H2
10.1.1.101
10.1.1/24
10.1.1.1 10.1.0.2
R1 H3
2
3 1
10.1.2.2 10.1.0.1 10.1.0/24
10.1.2/23
10.1.2.1
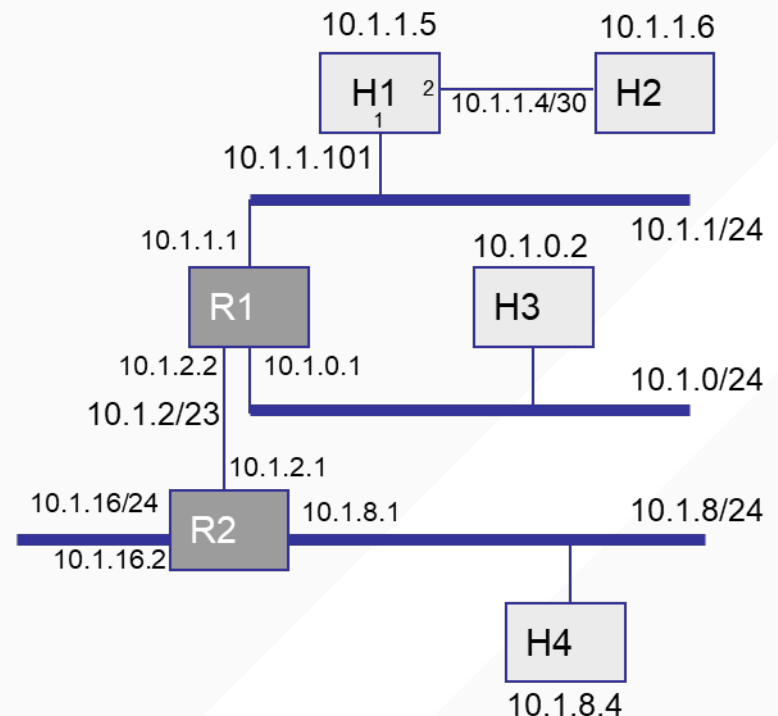10.1.16/24 R2 10.1.8.1 10.1.8/24
10.1.16.2
H4
10.1.8.4
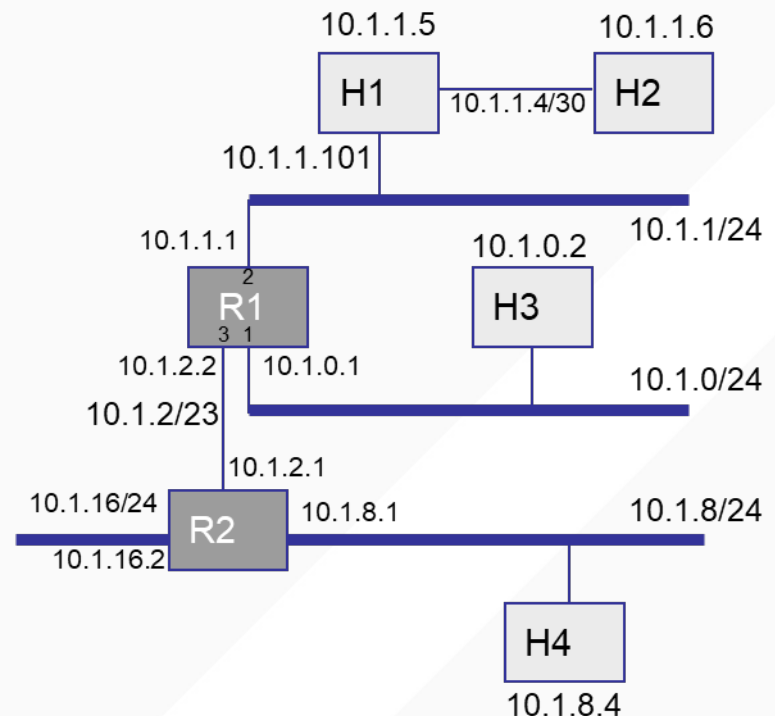
ROUTING TABLE EXAMPLE 3 (H1)

- ## Packet to 10.1.1.6

- ## Direct route

  - ### Longest prefix match

Routing table at H1

Destination	Next Hop
127.0.0.1	loopback
Default or 0/0	10.1.1.1
10.1.1.0/24	interface1
10.1.1.4/30	interface2

```
10.1.1.5                    10.1.1.6
  H1   2                      H2
     1    10.1.1.4/30
10.1.1.101
                                        10.1.1/24
10.1.1.1              10.1.0.2
                     H3
  R1
10.1.2.2    10.1.0.1              10.1.0/24
10.1.2/23
         10.1.2.1
10.1.16/24   R2   10.1.8.1              10.1.8/24
  10.1.16.2
                            H4
                          10.1.8.4
```

TCP/IP MODEL (VS OSI MODEL)
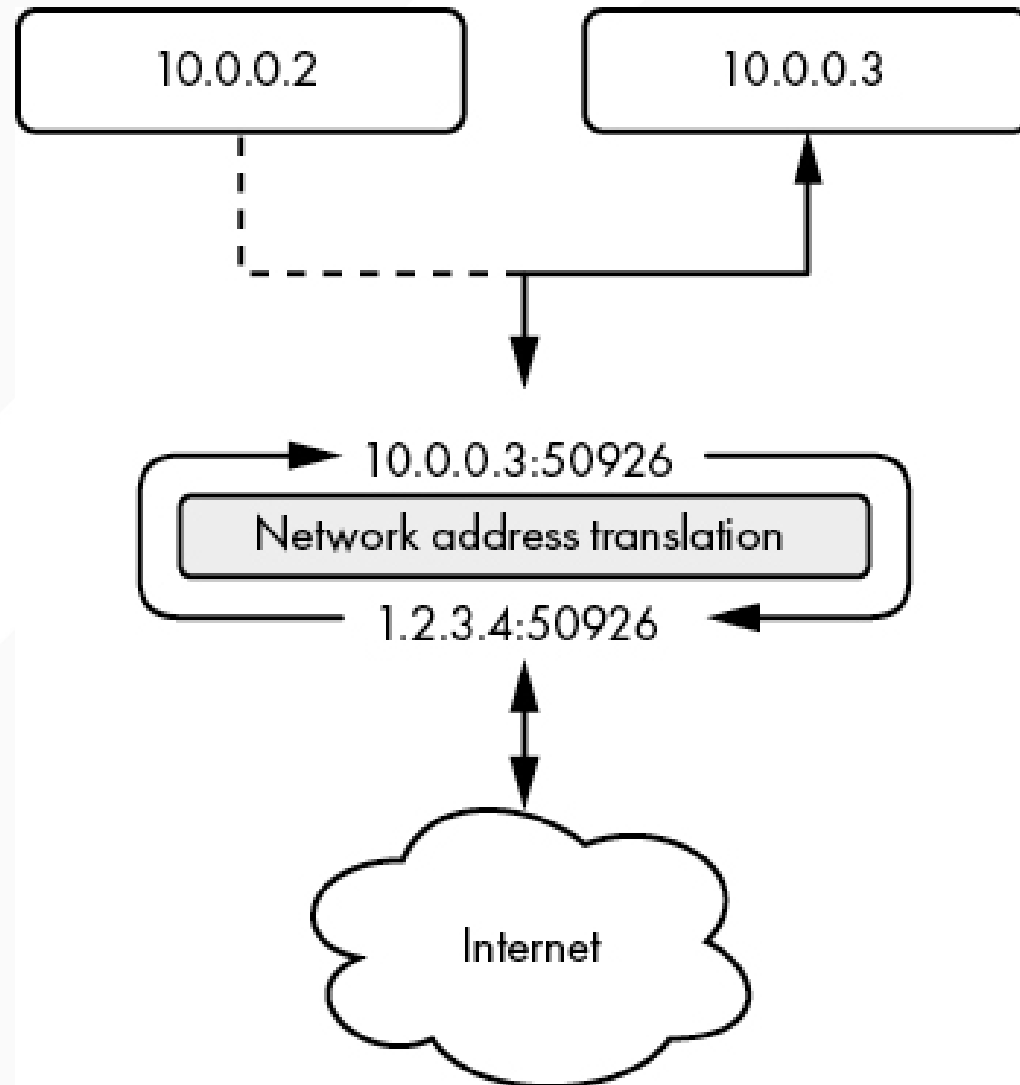
Which *process* on the destination machine? What if I want to send a stream of bytes, not just a limited-size message?

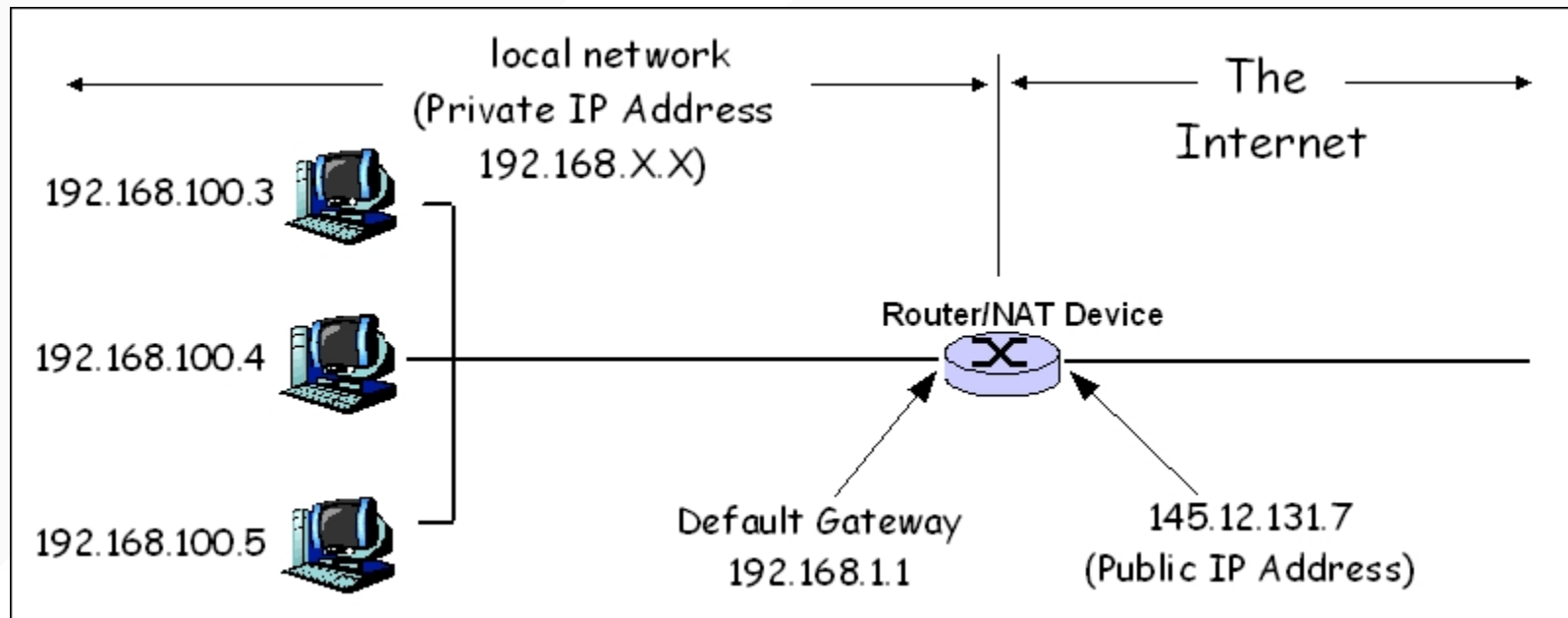| Software application	
TCP/IP	**OSI**
Application	Application
	Presentation
	Session
Transport	Transport
Internet/network	Network
Link	Data link
	Physical
Physical transmission media	
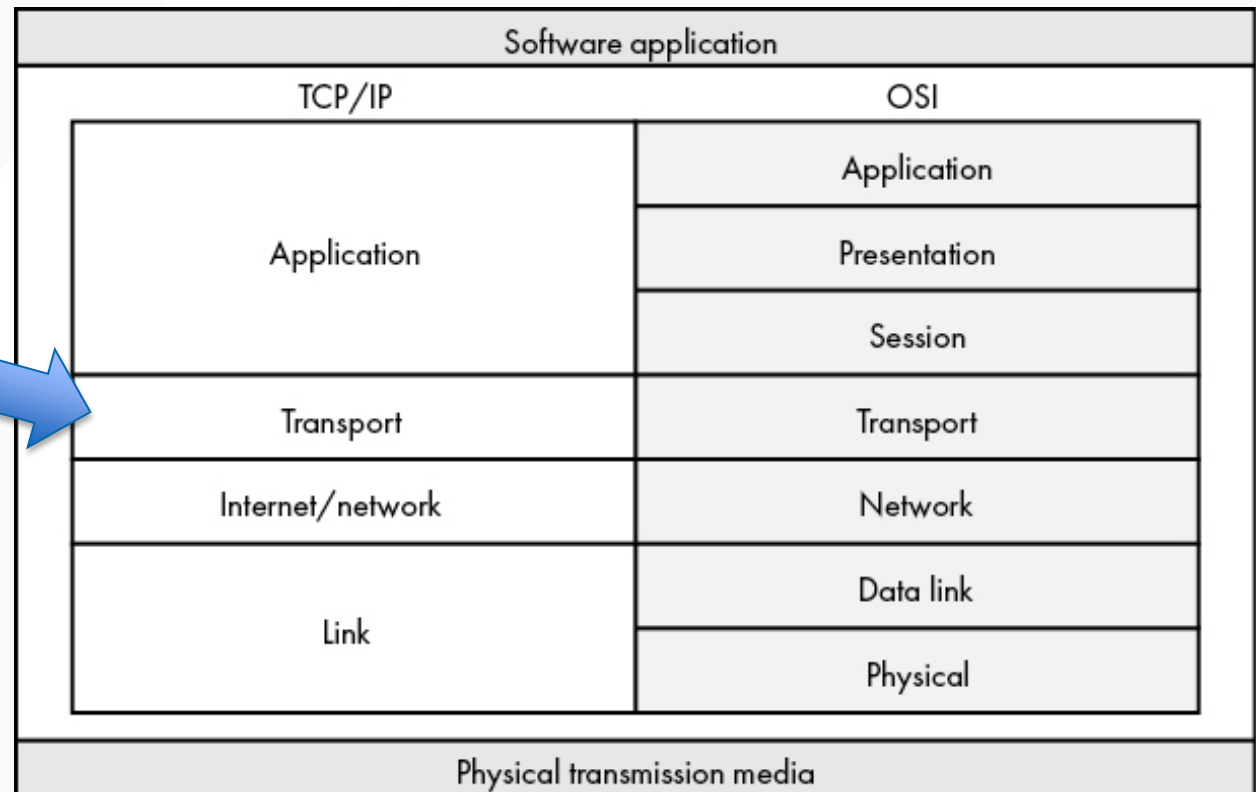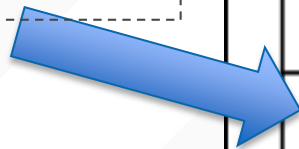
PORTS

- IP addresses identify a *machine*

  - Actually they identify a network interface on a machine

- How to identify different programs on the machine?

  - Process ID/PID? (no... why not?)

  - Instead we use a port (which is a 16-bit number)

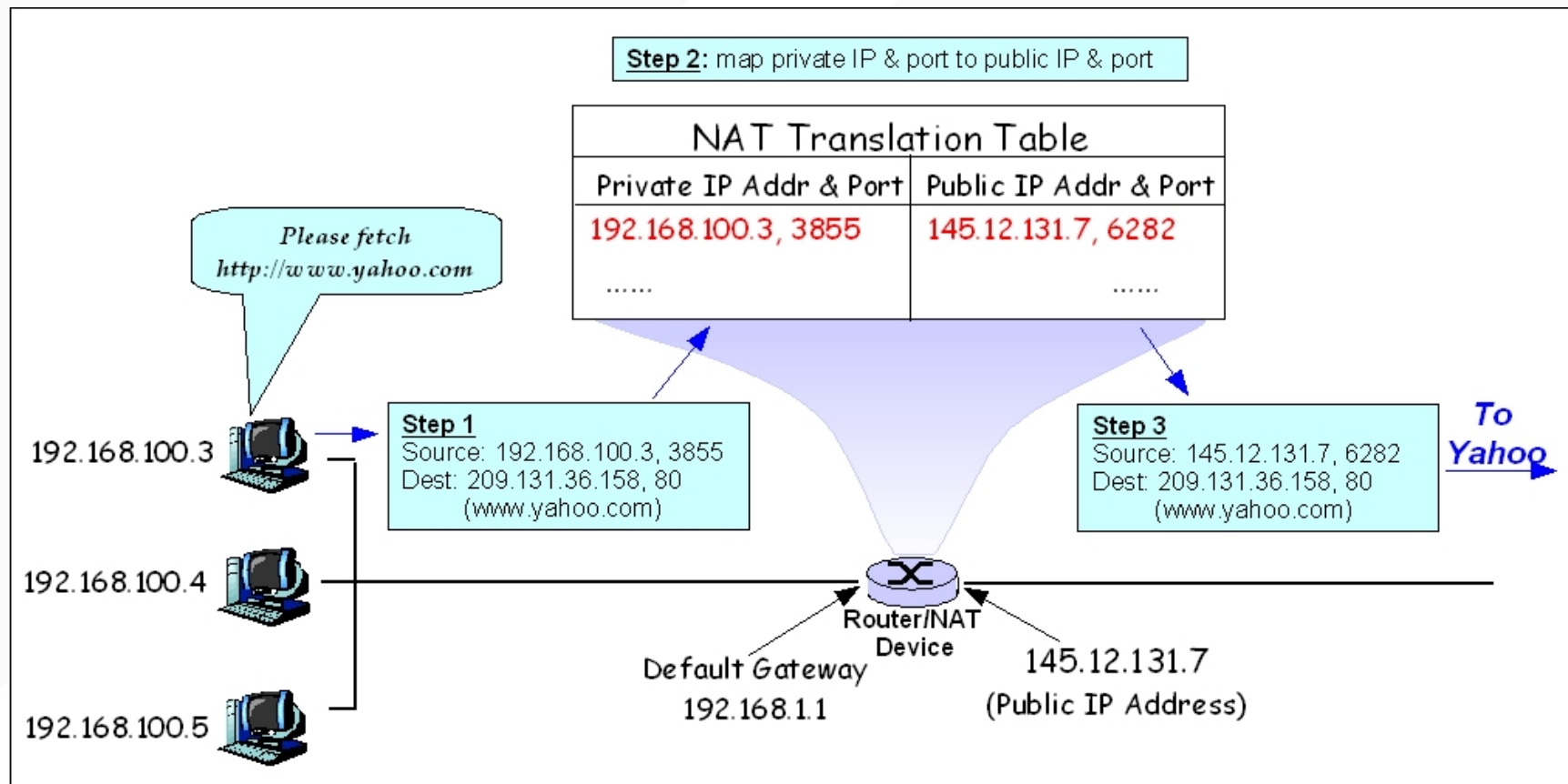 - 0-1024 reserved for the OS, you can use 1025-65535
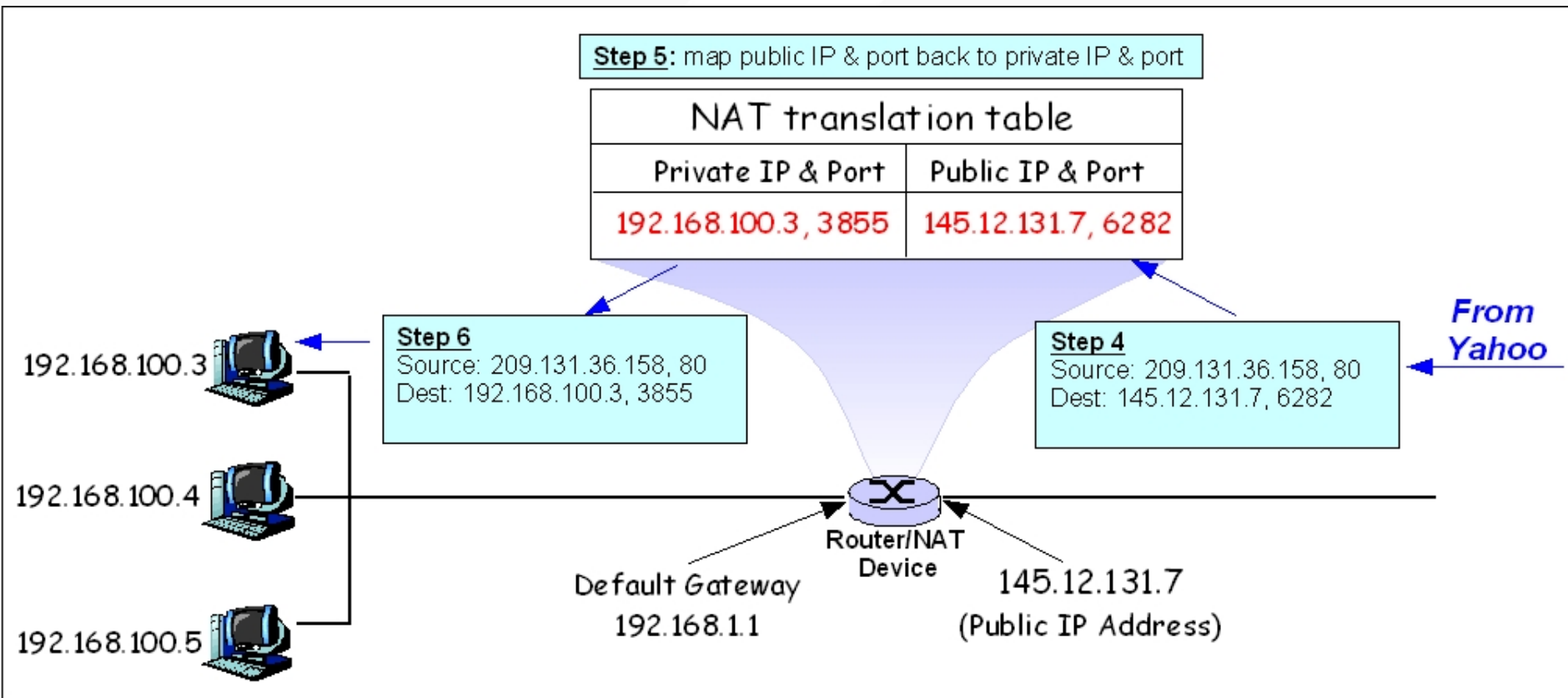
NETWORK ADDRESS TRANSLATION

LOAD BALANCING VIA NETWORK ADDRESS TRANSLATION

LOAD BALANCING VIA NETWORK ADDRESS TRANSLATION

LOAD BALANCING VIA NETWORK ADDRESS TRANSLATION

Step 5: map public IP & port back to private IP & port

NAT translation table

Private IP & Port	Public IP & Port
192.168.100.3, 3855	145.12.131.7, 6282

From Yahoo

Step 6
Source: 209.131.36.158, 80
Dest: 192.168.100.3, 3855

Step 4
Source: 209.131.36.158, 80
Dest: 145.12.131.7, 6282

192.168.100.3

192.168.100.4

192.168.100.5

Router/NAT Device

Default Gateway
192.168.1.1

145.12.131.7
(Public IP Address)