

# Content Distribution Networks and Peer to Peer Systems

Amin Vahdat

CSE 123b

June 6, 2006

# Announcements

- Third assignment
  - Due date June 9, 5 pm
- Final exam, June 14, 11:30-2:30

## More on Denial of Service/BotNets

(thanks to Mark Handley)

# Distributed Denial of Service Attacks

- The Internet does a great job of transmitting packets to a destination
  - Even if the destination doesn't want those packets
  - Overload servers or network links to prevent the victim doing useful work
- Distributed Denial of Service becoming commonplace
  - Automated scanning results in armies of compromised zombie hosts being available for coordinated attacks

# ISP's view of the problem

- ISP1 (very large ISP)
  - 6-7 ongoing DoS attacks at any time.
  - Peak bandwidth seen in UK: 3Gb/s
  - Peak bandwidth known to be seen in US: 5Gb/s (flatlined 2 OC48 links)
- ISP2 (large ISP)
  - >22000 anomalies in May-Sept 2004
  - 5000 high rate
  - 20 real attacks per day - perhaps 1/3 seriously affect customers

# ISP's View of the Problem

- ISP 3: (large international ISP)
  - Sees attacks from 300 to 10000+ simultaneous hosts
  - Sophisticated full spectrum attacks:
    - SYN flood
    - TCP connection flood
    - URL flood
    - UDP flood
    - ICMP flood
    - DNS attacks
    - Malformed packets
  - It's not getting any better

# ISP's View of the Problem

- Major security vendor:

- Lack of data encourages speculation, confusion and hyperbole....
- But trends are worrying:

DoS attacks greater than 10Gbps aggregate

Of 1127 customer-impacting DDoS attacks seen in 2004 on a large network, only 4 employed source address filtering

80K+ node botnet largest seen this year

DoS attack vectors are changing (eg., application level, Ack with simulated sequence numbers)

# ISP's View of the Problem

- ISP 4 (medium size national ISP)
  - **Problem, what problem?**
  - This ISP has no high-profile DoS targets
  - Mostly home users
  - Their backbone and peerings are over-provisioned
  - DoS mostly only noticed when another ISP complains one of their customers is being DoSed
    - Dealt with on a case-by-case basis
    - Not worth them investing in a detection infrastructure

## The Nature of the Attacks

- Broad range
  - Wide range of attacks on end-hosts (CPU, memory exhaustion)
  - Attacks on edge routers (bandwidth exhaustion, forwarding power, CPU cycles)
  - Very little source address spoofing
- Range of possible attacks is much broader, but the simple attacks mostly work well enough

# Motivation of the attackers *today*

- MEECES (Max Kilger, HoneyNet)
  - Money
  - Ego
  - Entertainment
  - Cause
  - Entrance into Social Groups
  - Status

## Profile of attackers today

- Asia-Pacific and South America are main sources
  - Not just Eastern Europe and Russia anymore
  - Mostly poor countries, where a few hundred/thousand dollars is a year's salary
  - Usually good education, but in a country with high unemployment
- Groups communicate mostly in-band (Internet)
  - But most ISPs don't have the resources to analyze TBs/day of IRC logs in many languages
- Many groups are well organized and highly skilled
  - Mostly not for fun on free time anymore

# Bots and Botnets

- *Bot*
  - Application that performs some action on behalf of a remote controller
  - Installed on a victim machine (zombie)
  - Modular (plug in your own functionality/exploit/payload)
- *Botnets*
  - Linkage of “Owned” machines into centrally controlled armies
  - Literally roBOT NETworks
- *Control channel*
  - Method for communicating with an army
- *Herder*
  - Owns control channel, commands botnet army

# Botnets

- Mass acquisition tools used for initial compromise
  - Losing a botnet isn't a tragedy - can quickly re compromise new hosts
- Variety of communication channels used to control botnets, but IRC and P2P protocols are most common
- After compromise, protect host to prevent multiple zombies/agents on the same host

## Botnet Spammer Rental Rates

- >20-30k always online SOCKs4, url is de-duped and updated every >10 minutes. 900/weekly, Samples will be sent on request  
>Monthly payments arranged at discount prices
  - 3.6 cents per bot week
- >\$350.00/weekly - \$1,000/monthly (USD)  
>Type of service: Exclusive (One slot only)  
>Always Online: 5,000 - 6,000  
>Updated every: 10 minutes
  - 6 cents per bot week
- >\$220.00/weekly - \$800.00/monthly (USD)  
>Type of service: Shared (4 slots)  
>Always Online: 9,000 - 10,000  
>Updated every: 5 minutes
  - 2.5 cents per bot week

## What are the Effects?

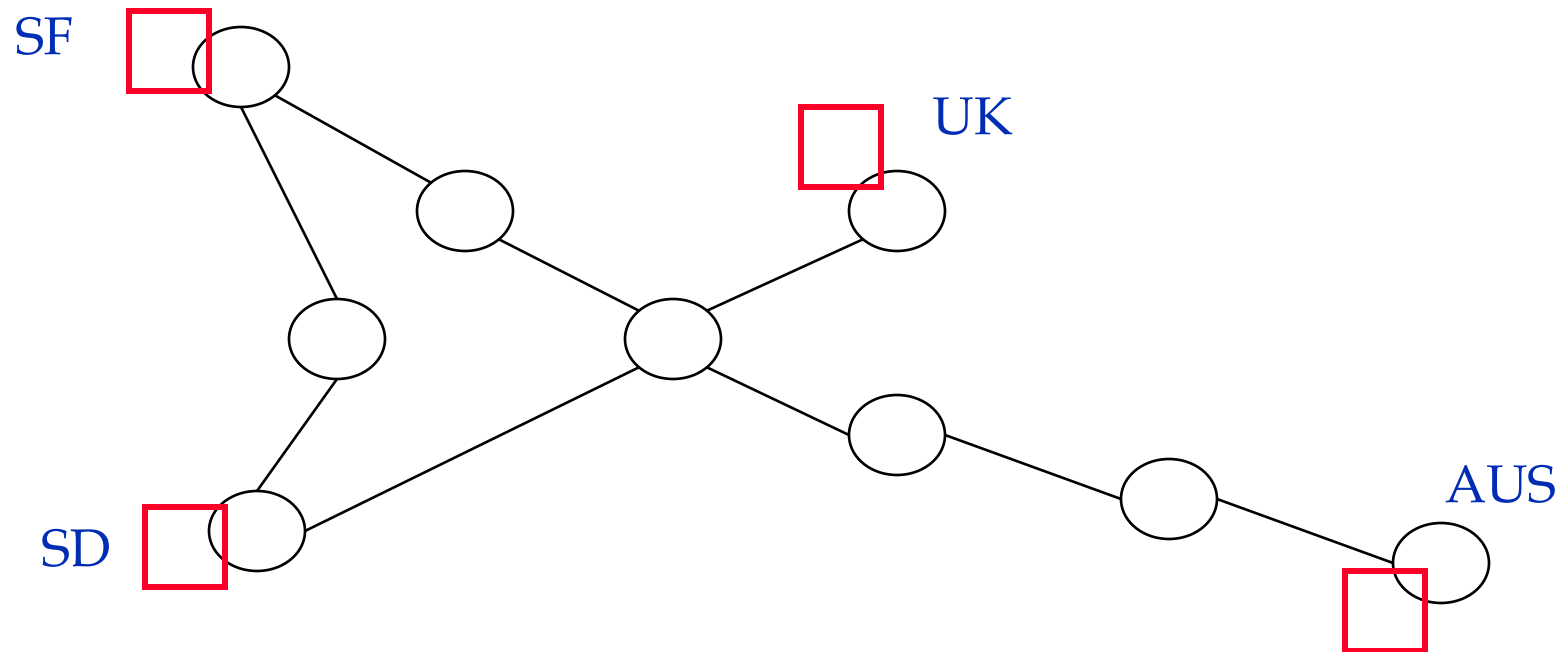
### ■ Application-Level Attacks:

- Use expected behavior of protocols to cause victim to spend resources
- Difficult to filter - looks like real transactions or requests
- Load prevents victim from processing real requests

Attack	Resource Threshold	Requests/Bot	Bots needed to exhaust
Static HTTP GET	60,000/sec	93 requests/sec at 250 bytes/request	645
Dynamic HTTP GET	3,000/sec	93 requests/sec at 250 bytes/request	40
SSL Handshake	600/sec	10 requests/sec	60

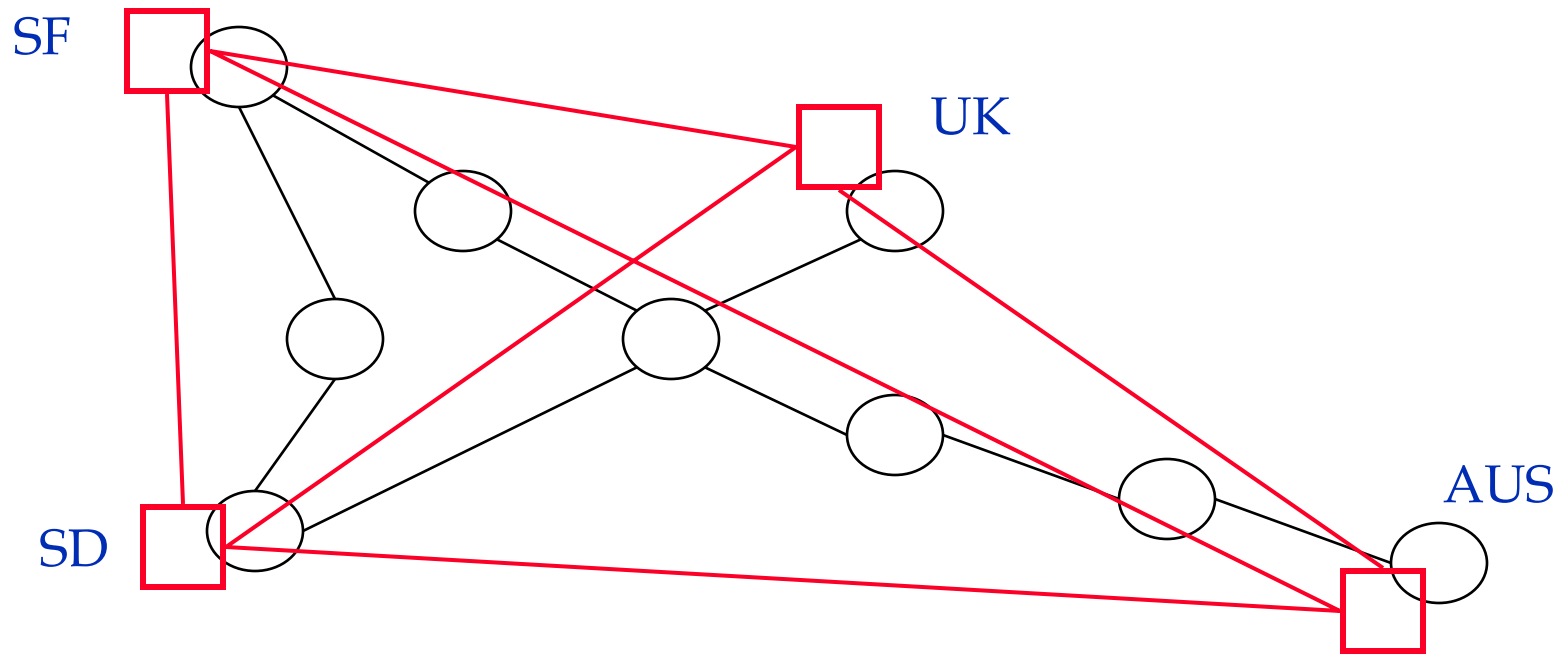
# Content Distribution Networks and Overlay Networks

# Overlay Networks



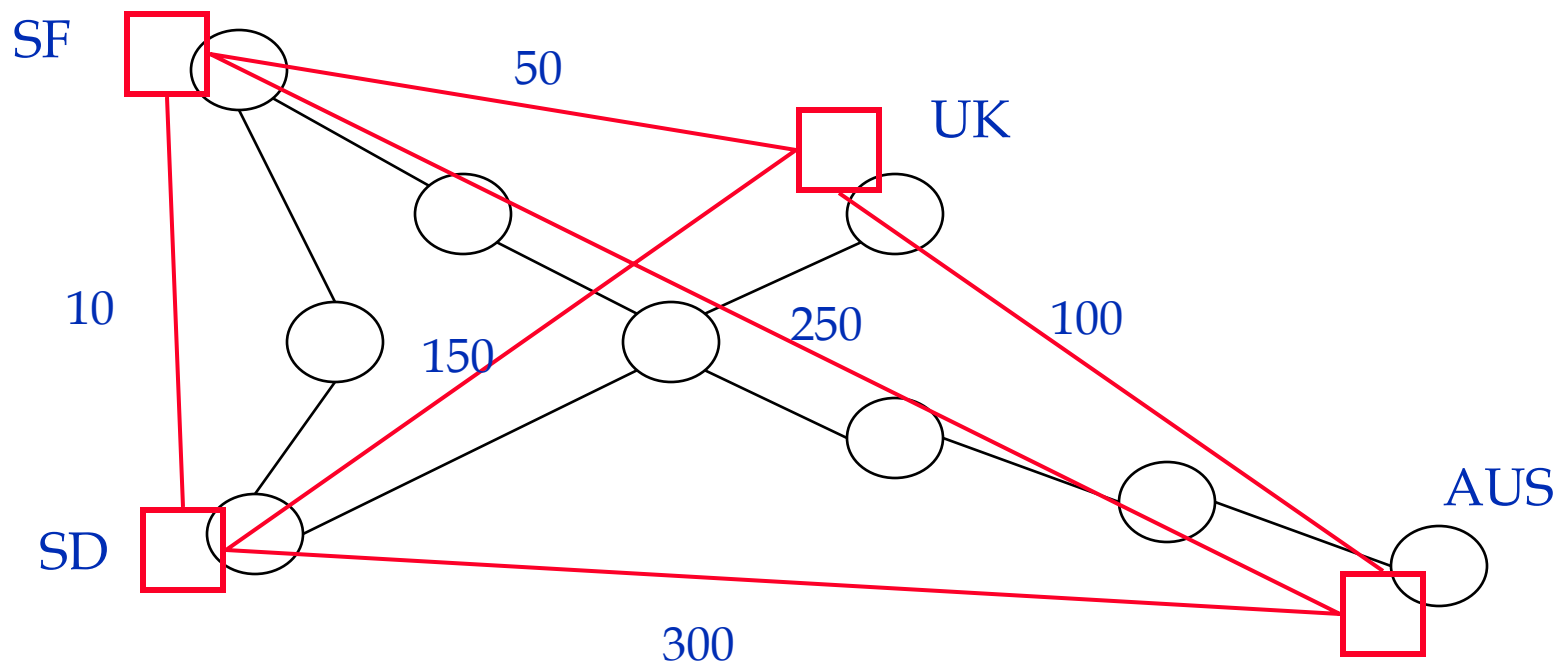
- Underlying network
  - Internet connectivity (IP Routing)

# Overlay Networks



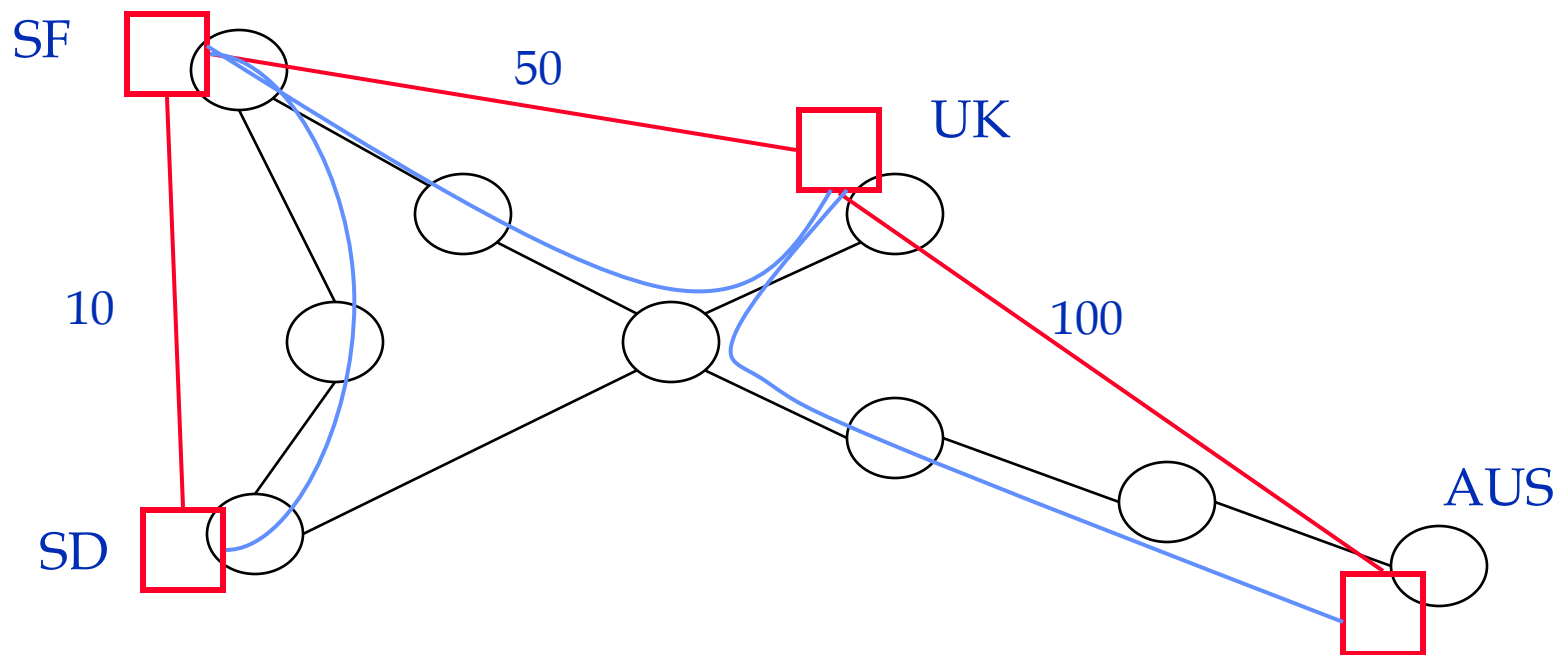
- Potential overlay connectivity
  - SF as root

# Overlay Networks



- Determine edge weights
  - E.g., bandwidth, latency

# Overlay Networks



- Build overlay connectivity
  - An application-layer distribution tree

# Overlay Networks

- We have had overlay networks for at least the past decade
  - Mbone, 6bone, etc.
- Orig. idea: these would be experimental networks that would help with the transition to “production” networks
- Today, overlay networks are being explored as general-purpose networks
  - Driven by content distribution networks and P2P computing

# Challenges to Building Overlay Networks

- What are some of the challenges to building overlays?
  - No central point of control
  - Scalability
  - Network performance tools
  - Building application-level peering that matches the topology of the underlying network
- Familiar story, but different level of abstraction
  - Can account for application-specific information rather than limited information available at network layer
  - Layer 7 versus layer 3 solution

# Content Distribution Networks: Why?

- Offload traffic from content providers
  - Allow sites to focus on core expertise rather than building scalable web services
  - Reduce operational costs for service providers
- Improve client perceived-latency
  - Move content closer to end clients
  - Performance often dominated by network latency
    - Not bandwidth!
- Problem: [www.cnn.com](http://www.cnn.com) (traditionally) maps to a single machine in the network
  - How to redirect client requests to closest/best replica?

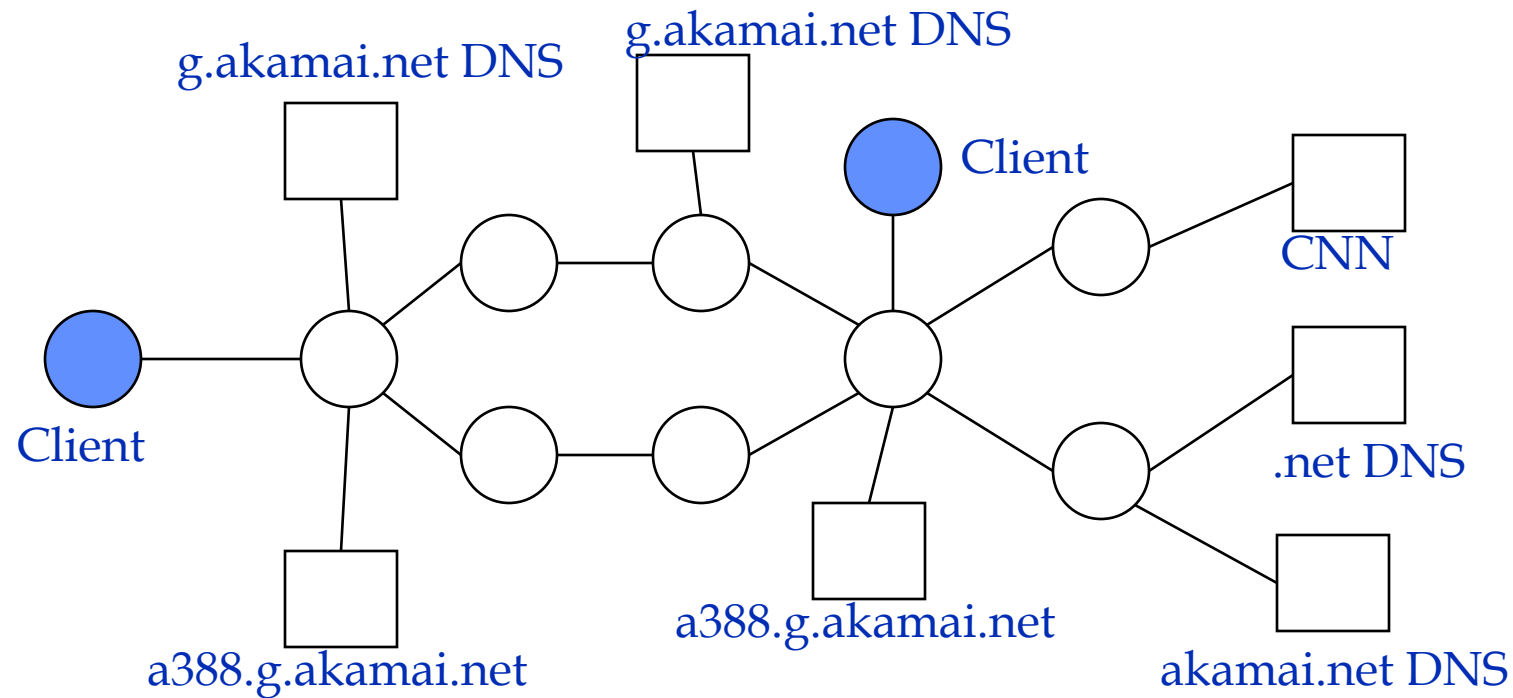
# Edge-service advantages

- Fast
  - Content and applications are served from locations near to end users
- Reliable
  - No single point of failure
  - Automatic failover
- Scalable
  - Global capacity on demand
- Cost effective
  - No over-provisioning
  - No redundant datacenters
  - Simple to manage
- Secure
  - Traffic harder to steal
  - Defense in depth protects central infrastructure

# Content Distribution Networks

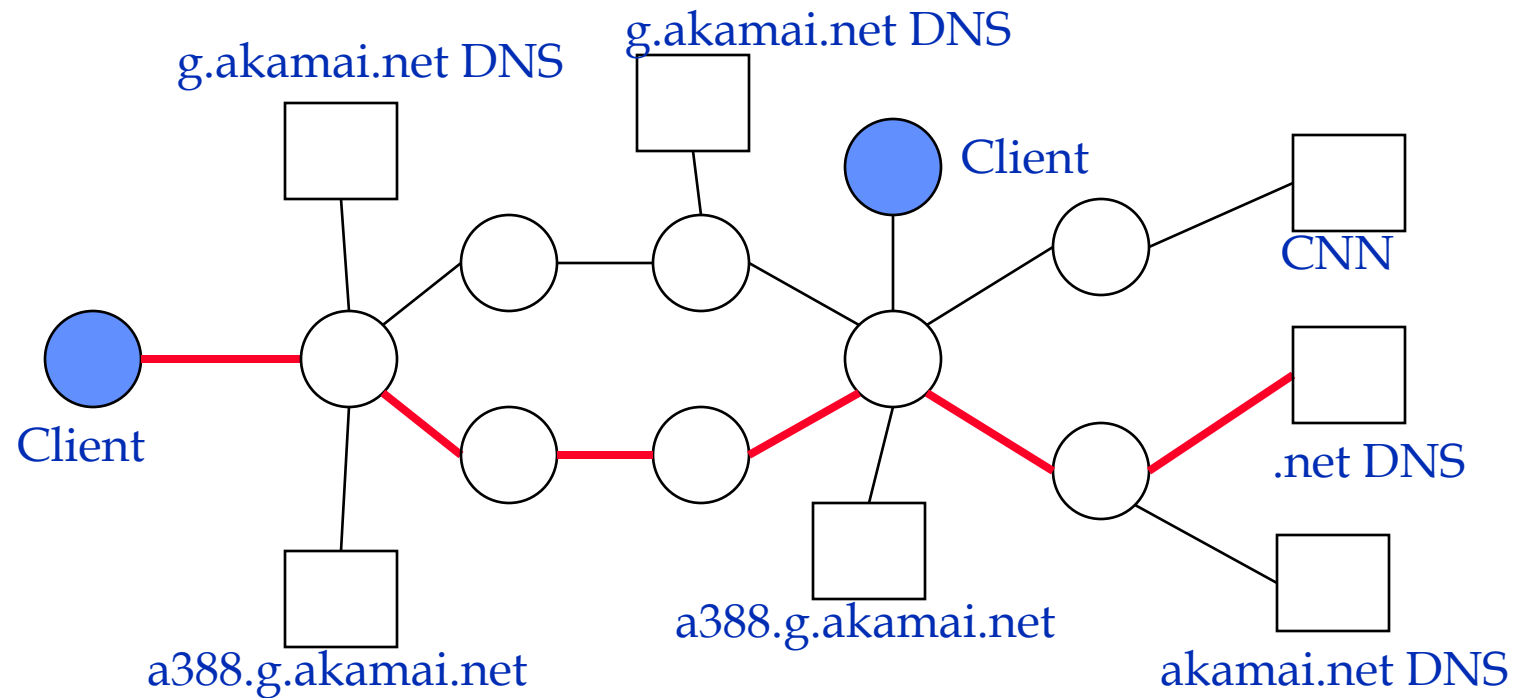
- CDN's propose to host web site images/audio/video
  - Images often make up more than 90% of web traffic
- One problem with replication is *consistency*
  - How do you keep your cached copies up to date?
  - By focusing on multimedia, largely eliminate consistency problems
- The future
  - Web moving increasingly toward *dynamic content*
  - How will CDN's support replication in this case? (ESI's)
    - Consistency, centralized databases, authentication
  - Access logs

# Content Distribution Networks



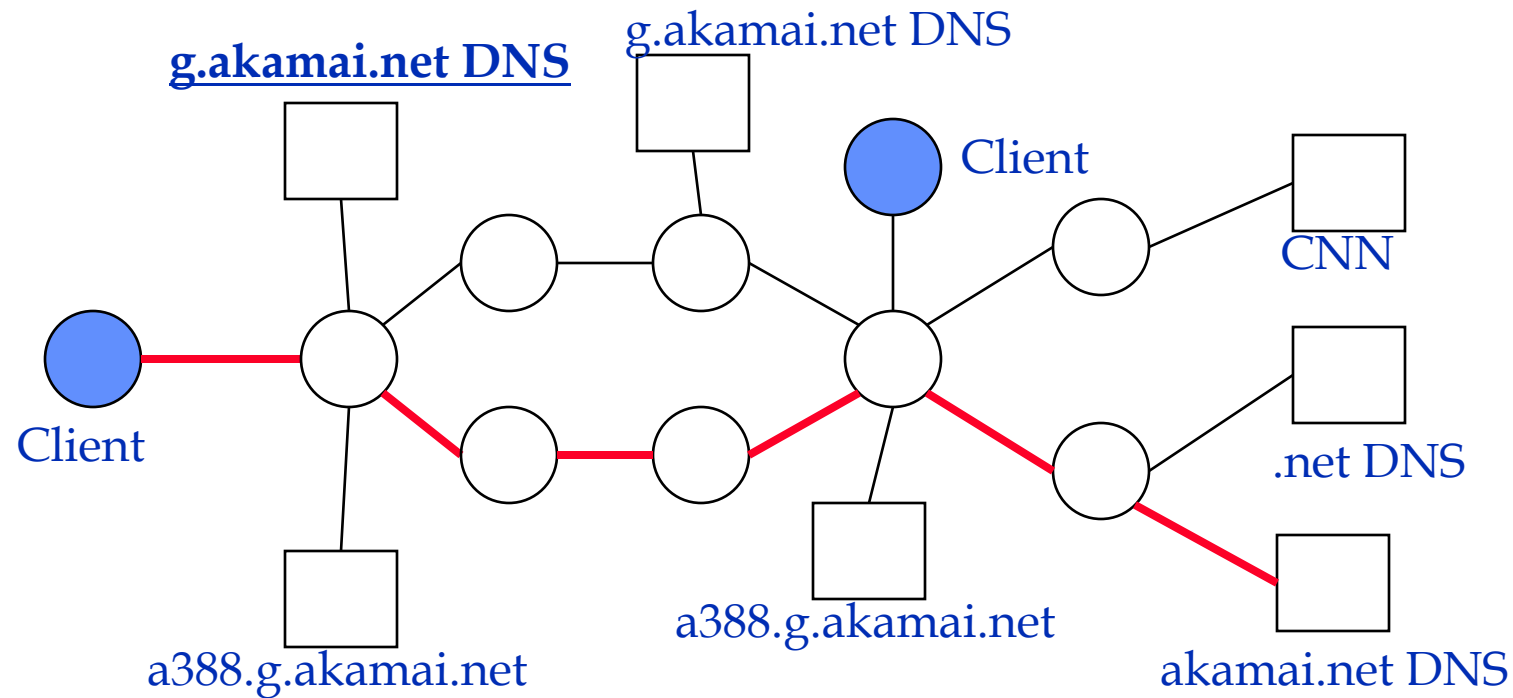
- Goal: quickly/efficiently deliver static content to clients
  - E.g., Akamai, Adero

# Content Distribution Networks



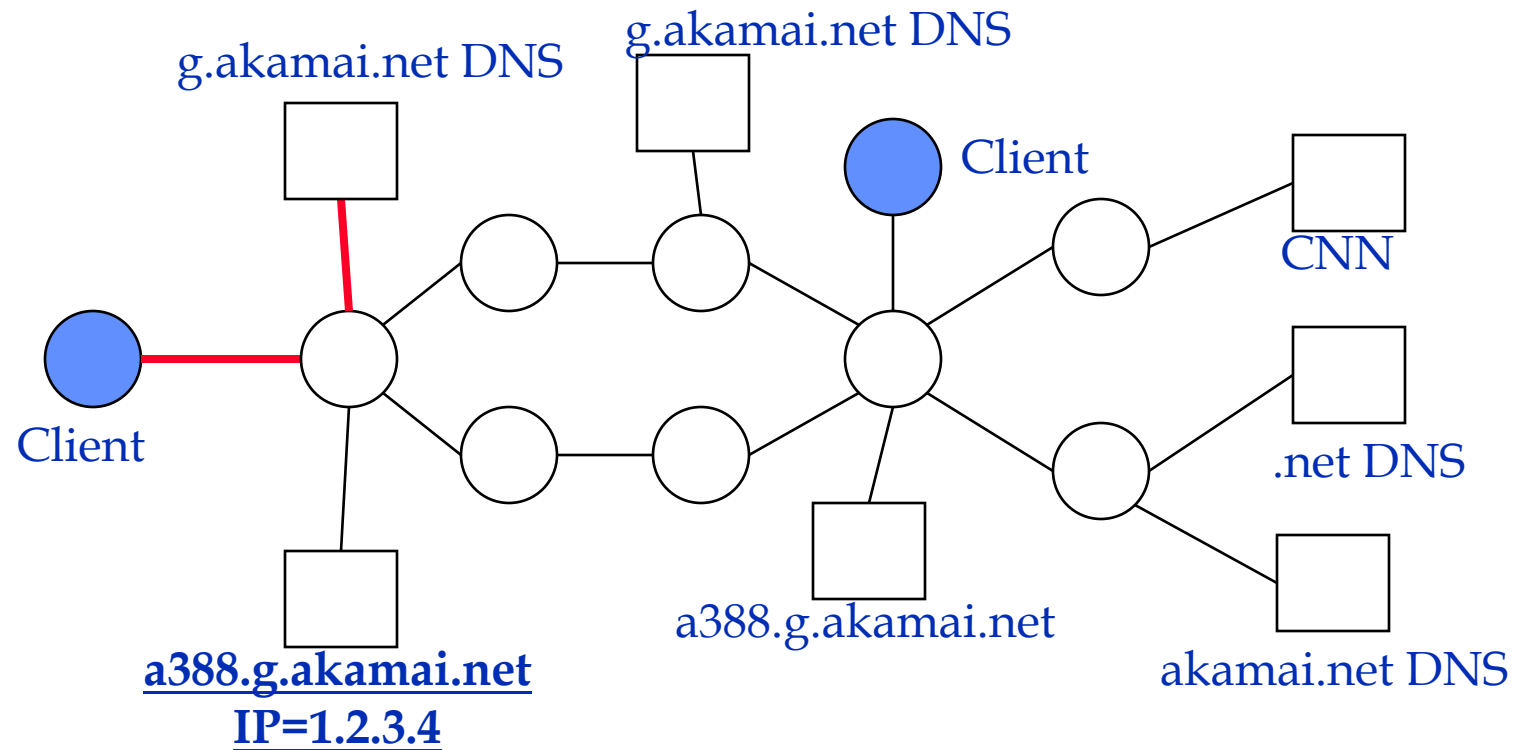
- Lookup akamai.net from .net DNS server
  - Cache reply (48 hour timeout)

# Content Distribution Networks



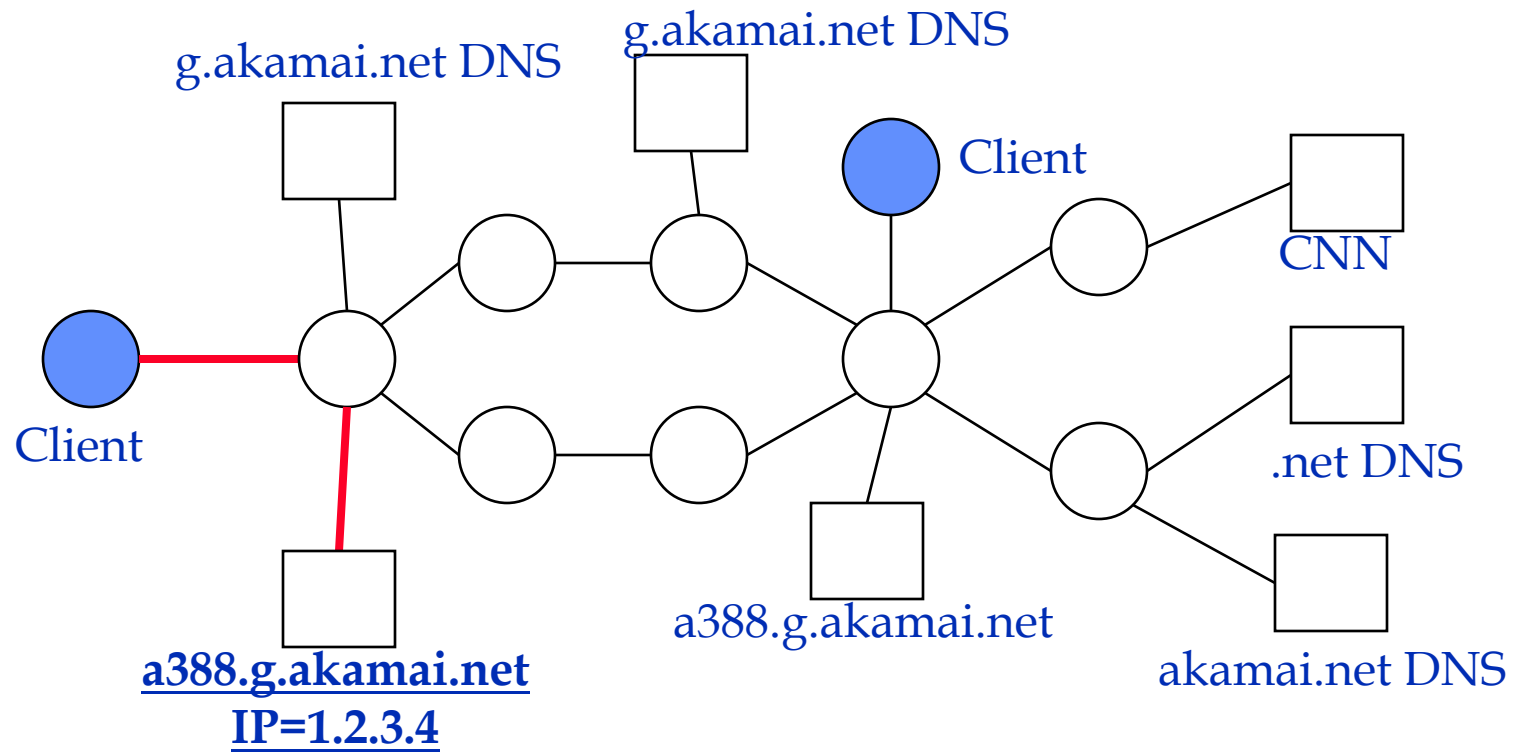
- Lookup g.akamai.net from akamai.net DNS server
  - Get nearby DNS server; Cache reply (30 minute timeout)

# Content Distribution Networks



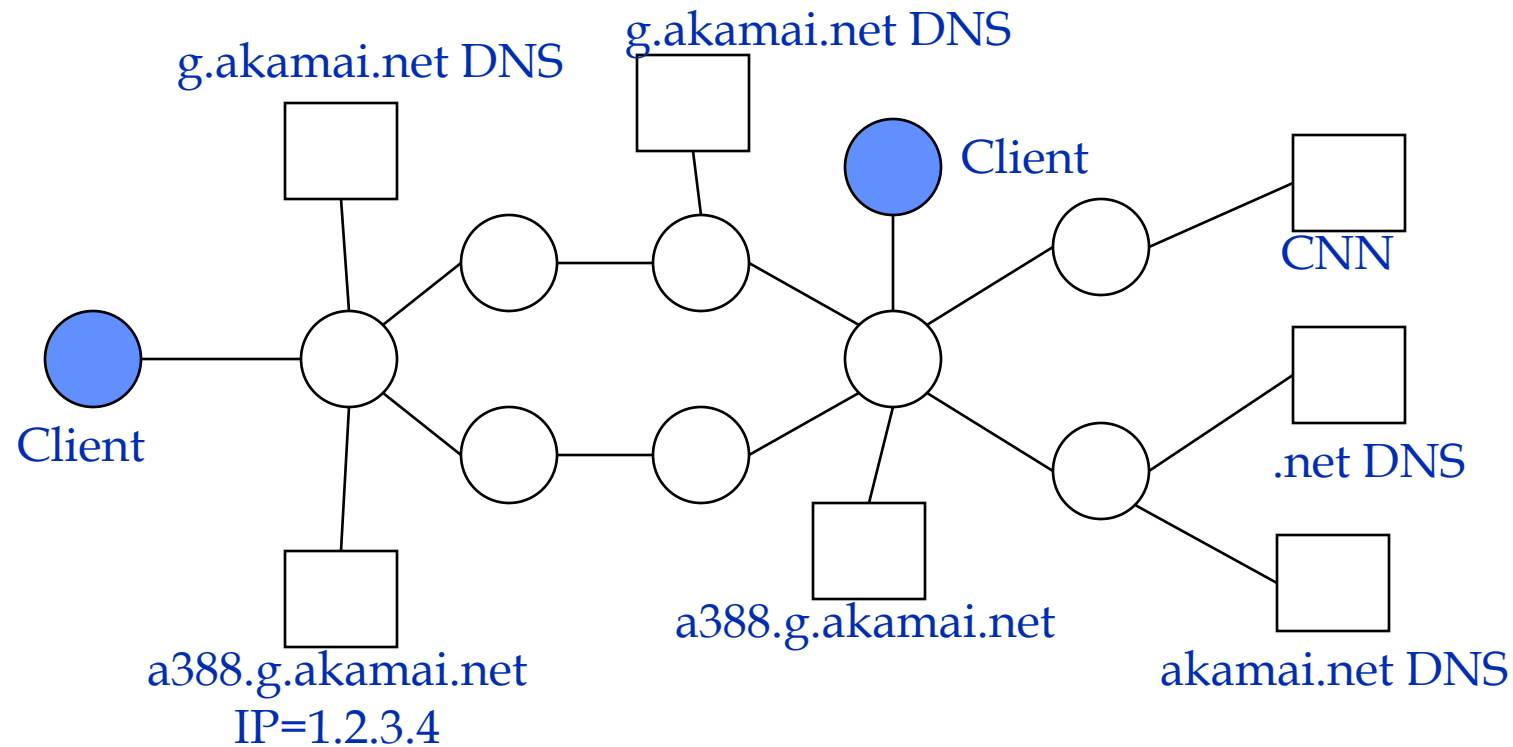
- Lookup a388.g.akamai.net from g.akamai.net DNS server
  - Get nearby replica location; cache reply (30 second timeout)

# Content Distribution Networks



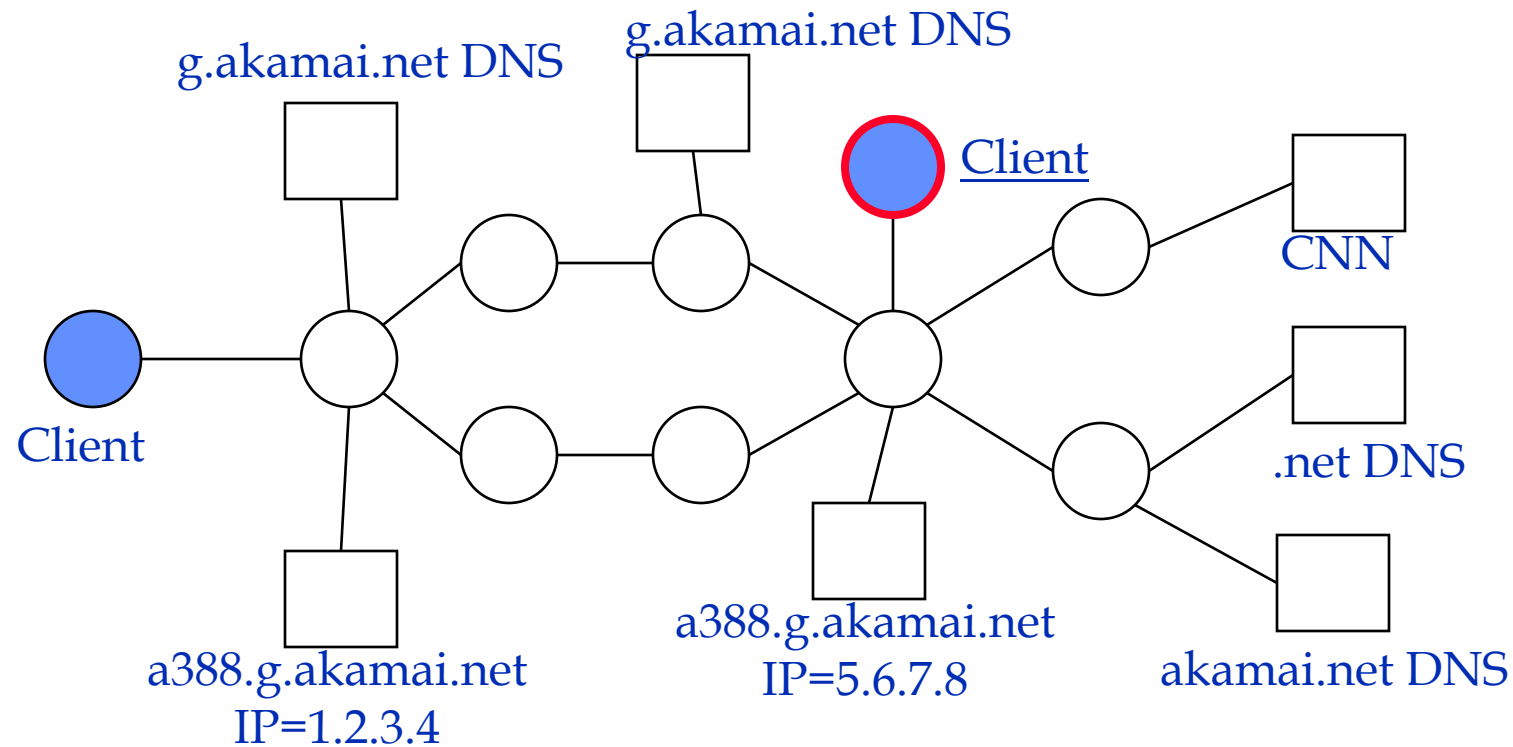
- Retrieve content from nearby server

# Content Distribution Networks

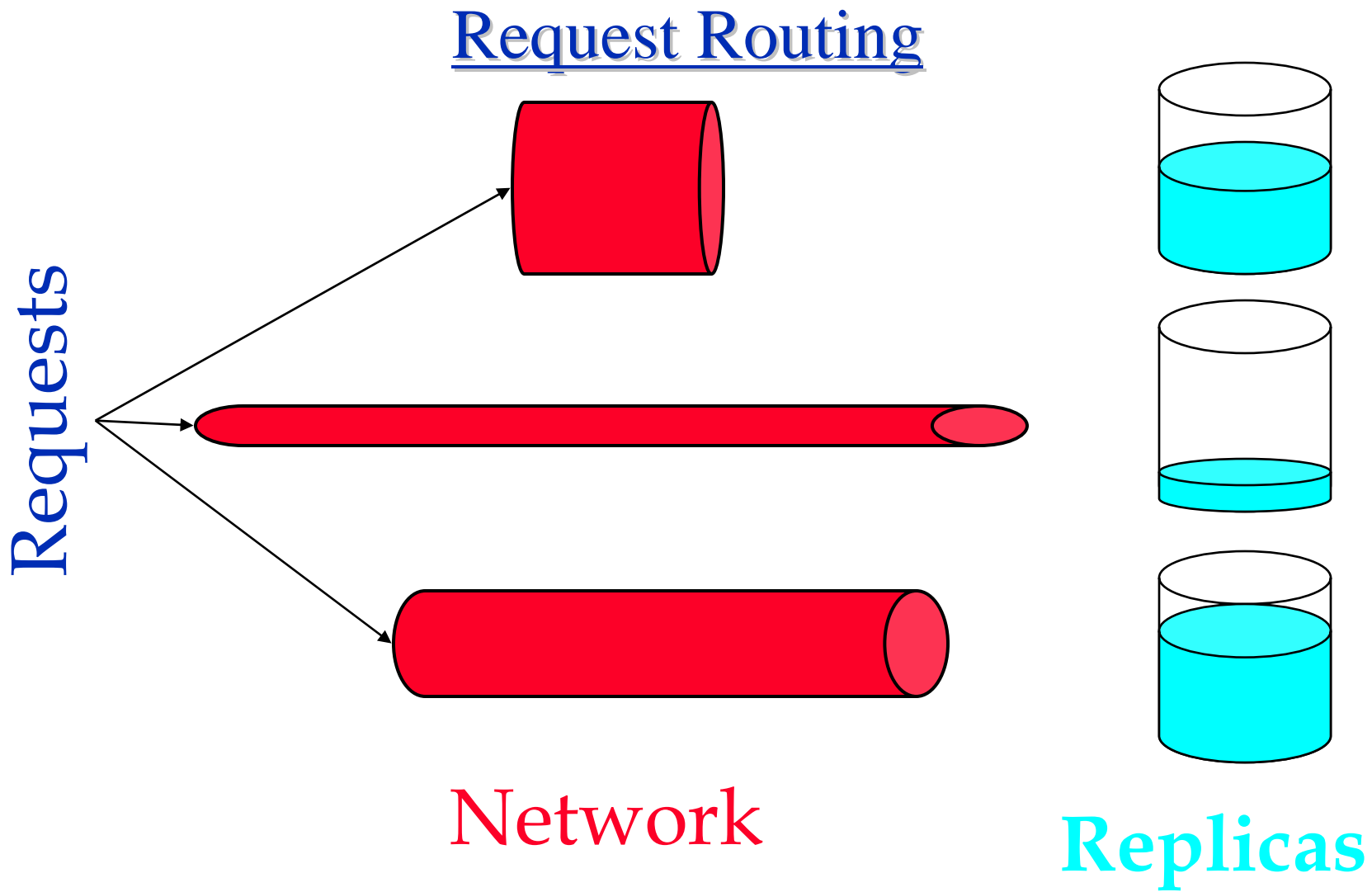


- A few minutes later, Client wishes to retrieve another page
  - What happens if IP 1.2.3.4 is overloaded?

# Content Distribution Networks



- What would happen if a client in another part of the network requested CNN content?



## Architectural Discussion

- Typically cache DNS lookups for days
  - Akamai overloads DNS lookup for load balancing
  - Allows for fine-grained redirection of client requests to appropriate replicas
- Akamai turns one round trip request into two
  - But... the two round trips could be to closer sites
  - Relies upon retrieval of relatively large objects to amortize cost of first round trip (for DNS lookup)
- What is Akamai optimizing for?
  - Consumed bandwidth? Client-perceived latency?
  - Who is paying Akamai?

# Content Distribution Networks: Motivation

- Four bottlenecks
  - The first mile
  - Peering points
  - Network backbones (bandwidth is not free)
  - The last mile
- Where is traffic coming from (2000)
  - How many ISP's make up 95% of network traffic?
  - What are the top ISP's as measured by traffic?

# Content Distribution Networks: Motivation

- Four bottlenecks
  - The first mile
  - Peering points
  - Network backbones (bandwidth is not free)
  - The last mile
- Where is traffic coming from (2000)
  - 95% of traffic from 7000 ISP's (Zipf distribution again)
  - Heavy tailed distribution (sum of top 10 is less than 20%)
  - UUNET (6%), at home (4%), AOL (3.6%)
  - Only 10 networks have more than 1%

## Akamai Measurements (2000)

- Peak backbone bandwidth ~200 Gbps
  - Small cable channel with 100k subscribers would consume 30 Gbps
- Akamai Network deployment
  - 6200 servers, 400 networks, 54 countries

## Akamai Measurements (2005)

- 15,000+ servers
- 2,400+ Locations
- 1,000+ networks
- 70 countries
- 80+ gigabits/sec of traffic
- 50+ billion hits per day
- 10-15% of the world's web traffic
- 18,000 web sites, 1,300 recurring revenue customers

## Building out the Akamai Network (2000)

- Akamai goal: convince all 7000 ISP's that make up 95% of traffic to deploy 1 or more Akamai servers
- Values for ISP's
  - Increased QoS for subscribers
  - Decreased cost for transit
  - Easy revenue stream from delivery
  - Satellite transport (11 mb/s usenet feed)
  - Network monitoring tools
  - (free)

## Performance (2000)

- Average speedup of 2-46 times
  - Median of 7x, 86% reduction in download time
  - Peak speedup is 3-136 with a median of 13
- Consistent performance
  - For all 94 sites, the CDN reduced the standard deviation of download by 30-92%
- Availability and Reliability
  - CDN reduced the error rate by 30-100% for all sites

# Components of a CDN

- Network performance monitoring
  - What to measure? active/passive, how to measure?
  - More than 10 million data points a second
- Server performance monitoring (disks, CPU, network, processes, hardware)
  - Usage and capacity monitoring (usage by object, storage usage per server/region, network utilization)

## Resource Management/Server Selection

- Takes as input data about system state and needs to connect users to servers
- Optimize performance while managing:
  - Traffic, server load, fault tolerance and reliability, storage, hit rate, load for multiple types of traffic, multiple server resources
- Must be truly distributed and fault tolerant (must tolerate failure of multiple components without ever disrupting service)
- Must work with imperfect information and in an unreliable environment
  - >150,000 requests/sec at peak

# Log Collection and Processing

- Scale to handle 10s of billions of hits per day
- Distributed collection and filtering
- Scalable processing and data mining infrastructure
  - Needs to have short delay (cannot wait a day to deliver data to customers)
  - How many unique users? What are the paths through their web sites?
- Loading into server 300 GB/day

## Akamai HTTP Server

- Event kernel for performance
- Native SSL, cookie support, authentication, DRM, dynamic content generation, XSLT processing
- Streaming Servers (streaming qt, real g2, windows media)
  - A few Gb's traffic in each of the 3 streaming formats