DATA COMMUNICATOIN NETWORKING

Instructor: Ouldooz Baghban Karimi

Course Book: Computer Networking, A Top-Down Approach, Kurose, Ross

Slides:

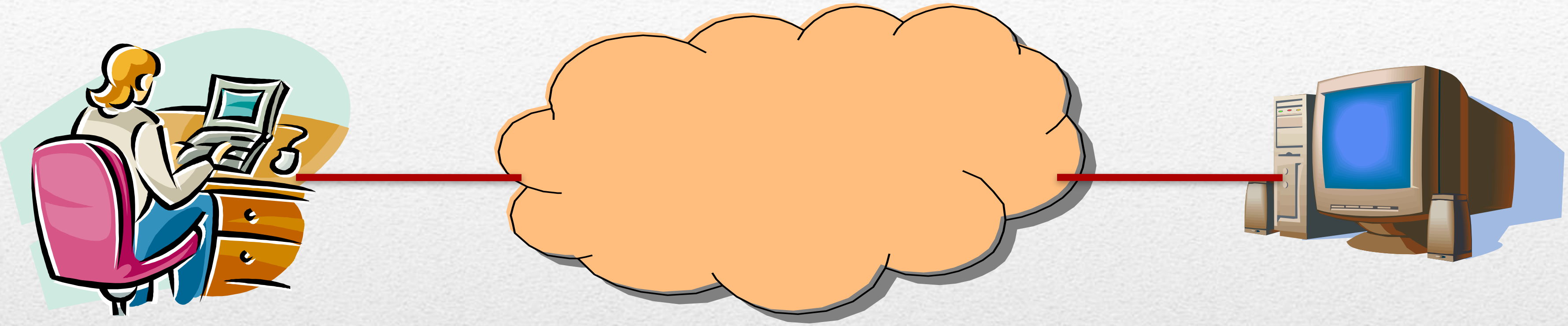
- Course book Slides
- Slides from Princeton University COS461 Spring 2012 offering, Jennifer Rexford

Course Overview

- **Basics of Computer Networks**
  - Internet & Protocol Stack
  - Application Layer
  - Transport Layer
  - Network Layer
  - Data Link Layer

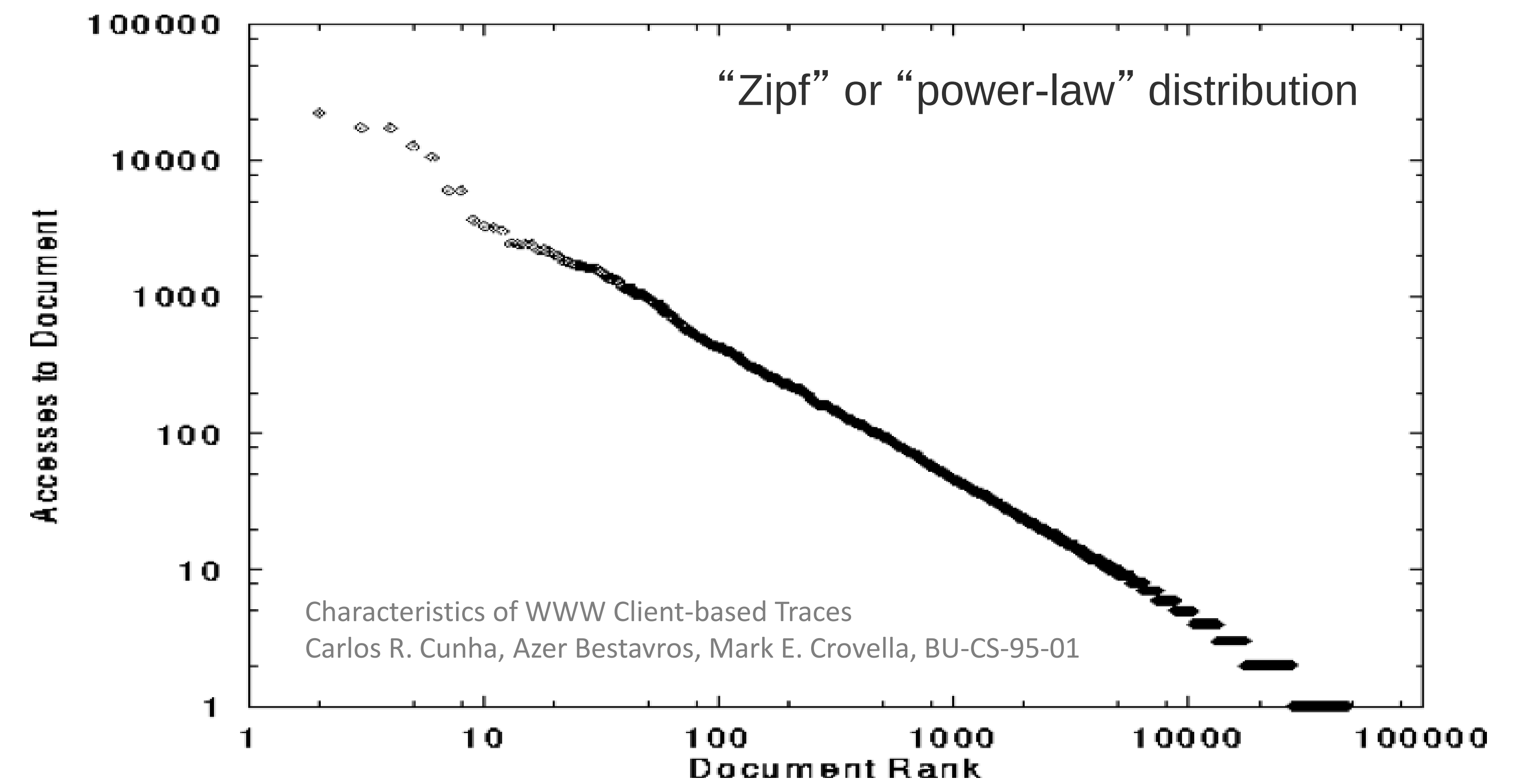
- **Advanced Topics**
  - Case Studies of Computer Networks
  - Internet Applications
  - Network Management
  - Network Security

Content Distribution Networks



- **Single server**
  - Single point of failure
  - Easily overloaded
  - Far from most clients
- **Popular content**
  - Popular site
  - “Flash crowd” (aka “Slashdot effect”)
  - Denial of Service attack

Skewed Popularity of Web Traffic

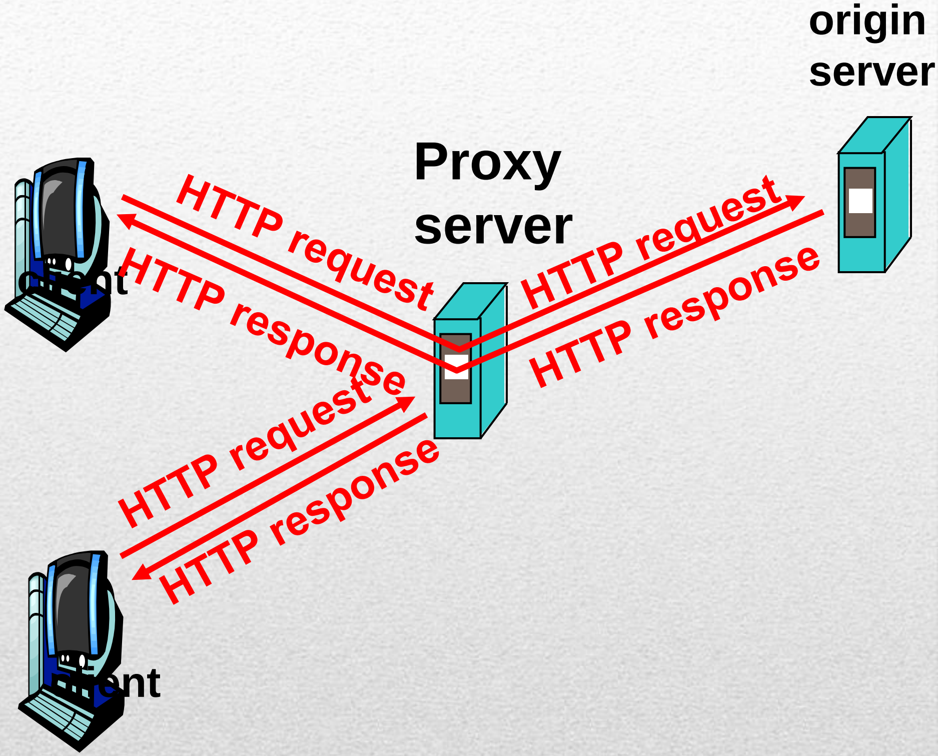


“Zipf” or “power-law” distribution

Characteristics of WWW Client-based Traces
Carlos R. Cunha, Azer Bestavros, Mark E. Crovella, BU-CS-95-01

Proxy Caches

- Reactively replicates popular content
- Smaller round-trip times to clients
- Reduces load on origin servers
- Reduces network load, and bandwidth costs
- Maintain persistent TCP connections

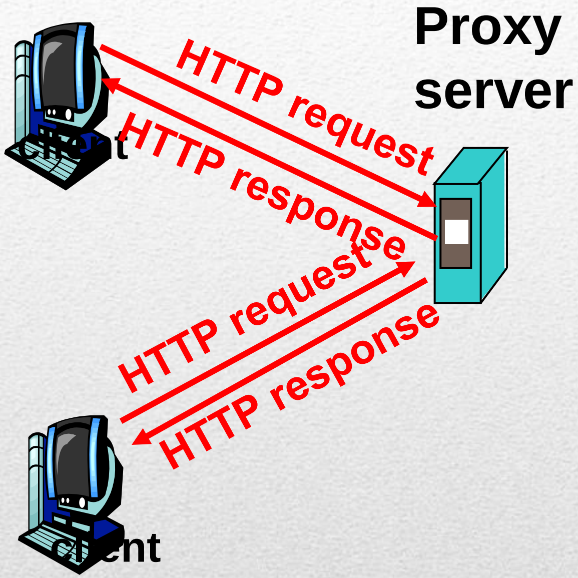


Forward Proxy

- **Cache close to the client**
  - Improves client performance
  - Reduces network provider's costs

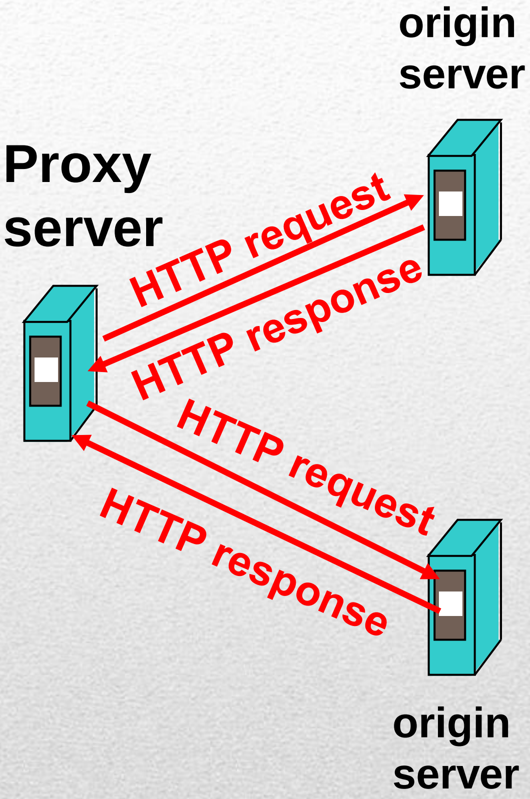
- **Explicit proxy**
  - Requires configuring browser

- **Implicit proxy**
  - Service provider deploys an "on path" proxy
  - … that intercepts and handles Web requests



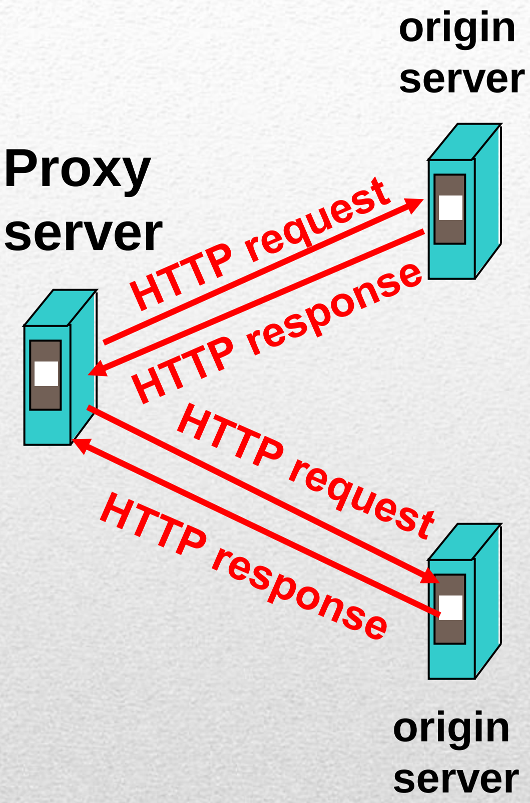
Reverse Proxy

- **Cache close to server**
  - Improve client performance
  - Reduce content provider cost
  - Load balancing, content assembly, transcoding, etc.
- **Directing clients to the proxy**
  - Map the site name to the IP address of the proxy



Reverse Proxy

- Cache close to server
  - Improve client performance
  - Reduce content provider cost
  - Load balancing, content assembly, transcoding, etc.
- Directing clients to the proxy
  - Map the site name to the IP address of the proxy

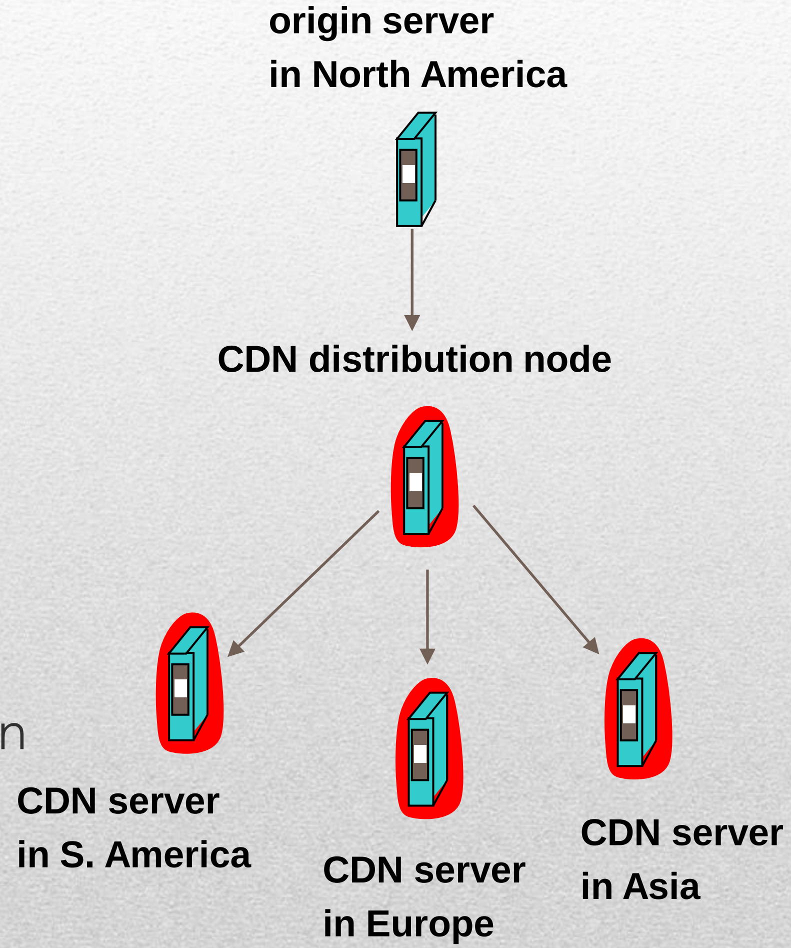


Limitations of Web Caching

- Much content is not cacheable
    - Dynamic data
        - Stock prices, scores, web cams
    - CGI scripts R
        - Results depend on parameters
    - Cookies
        - Results may depend on passed data
    - SSL
        - Encrypted data is not cacheable
    - Analytics
        - Owner wants to measure hits
- Stale data or overhead of refreshing the cached data

Content Distribution Network

- Proactive content replication
  - Content provider (e.g., CNN) contracts with a CDN
- CDN replicates the content
  - On many servers spread throughout the Internet
- Updating the replicas
  - Updates pushed to replicas when the content changes



Server Selection Policy

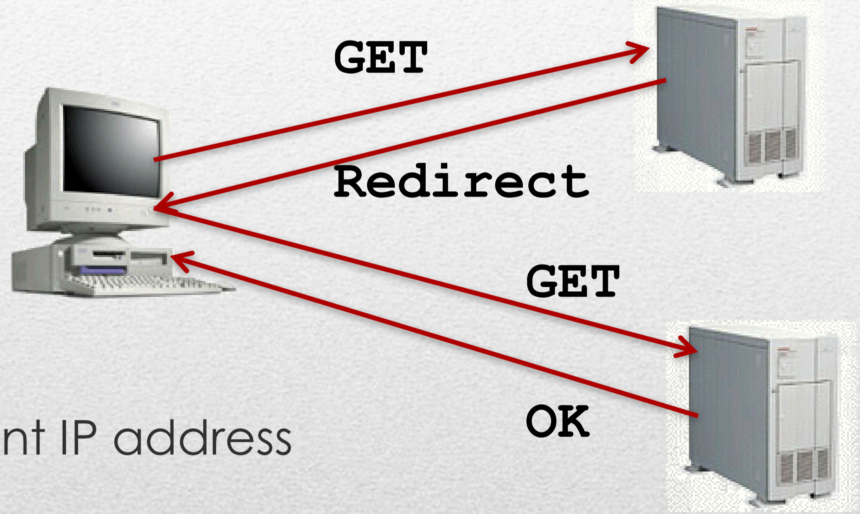
- Live server
    - For availability

Requires continuous monitoring of liveness, load, and performance

- Lowest load
    - To balance load across the servers
- Closest
    - Nearest geographically, or in round-trip time
- Best performance
    - Throughput, latency, …
- Cheapest bandwidth, electricity, …

Server Selection Mechanism

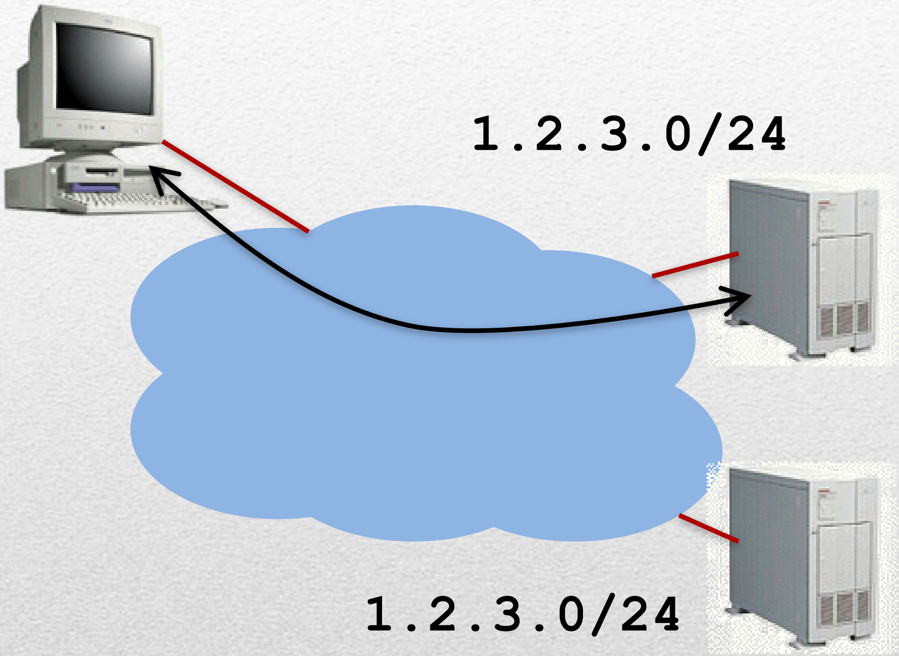
- Application
  - HTTP redirection



- Advantages
  - Fine-grain control
  - Selection based on client IP address
- Disadvantages
  - Extra round-trips for TCP connection to server
  - Overhead on the server

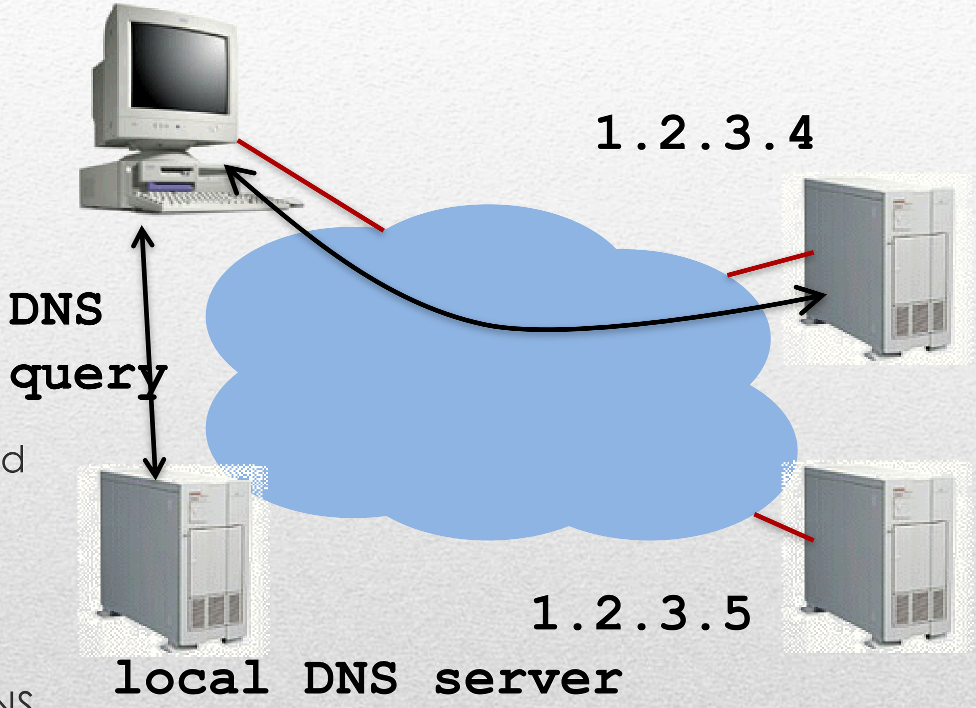
Server Selection Mechanism

- Routing
  - Anycast routing
- Advantages
  - No extra round trips
  - Route to nearby server
- Disadvantages
  - Does not consider network or server load
  - Different packets may go to different servers
  - Used only for simple request-response apps



Server Selection Mechanism

- Naming
  - DNS-based server selection
- Advantages
  - Avoid TCP set-up delay
  - DNS caching reduces overhead
  - Relatively fine control
- Disadvantage
  - Based on IP address of local DNS server
  - “Hidden load” effect
  - DNS TTL limits adaptation



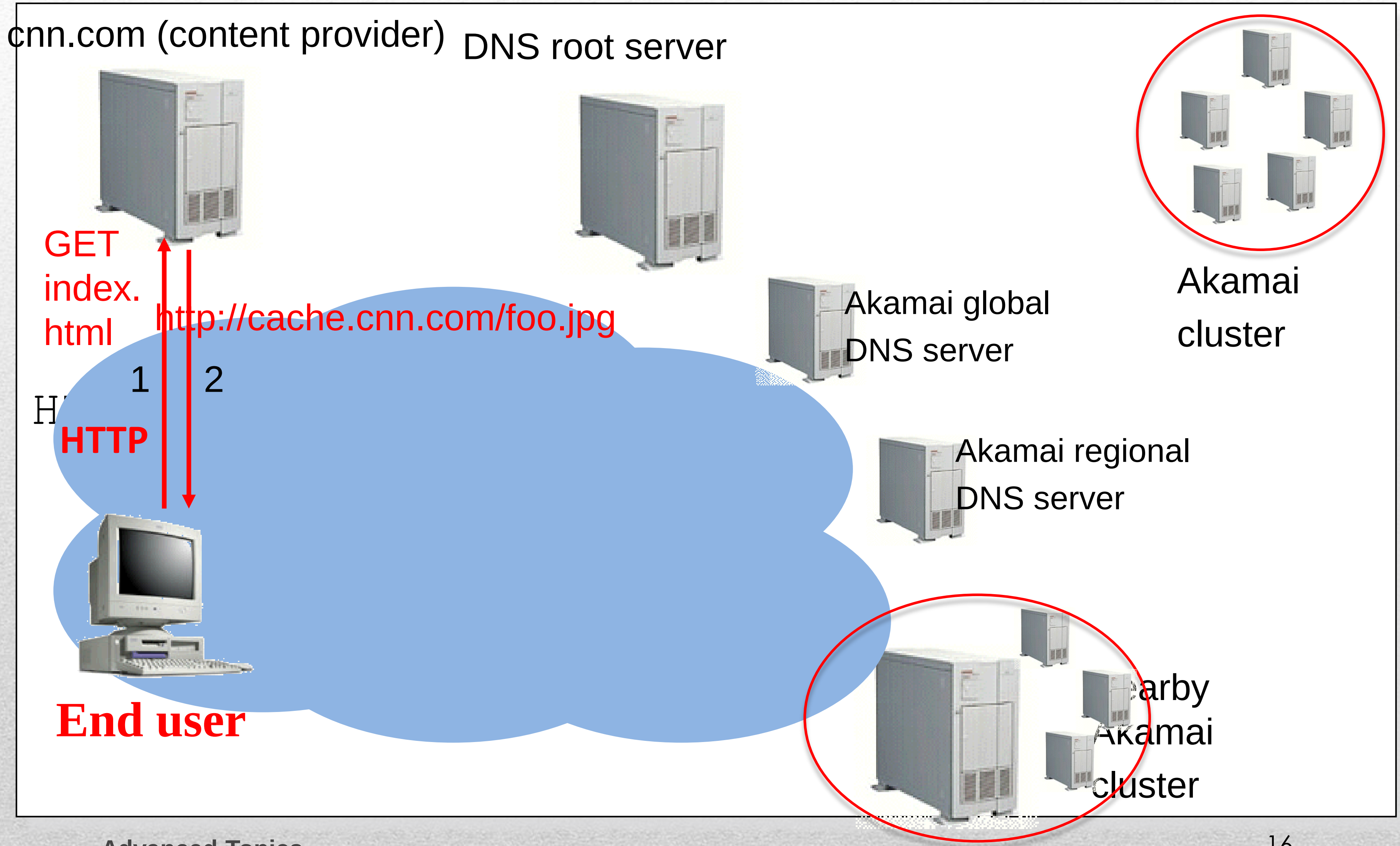
Example: Akamai

- Distributed servers
  - Servers: ~61,000
  - Networks: ~1,000
  - Countries: ~70

- Many customers
  - Apple, BBC, FOX, GM IBM, MTV, NASA, NBC, NFL, NPR, Puma, Red Bull, Rutgers, SAP, …

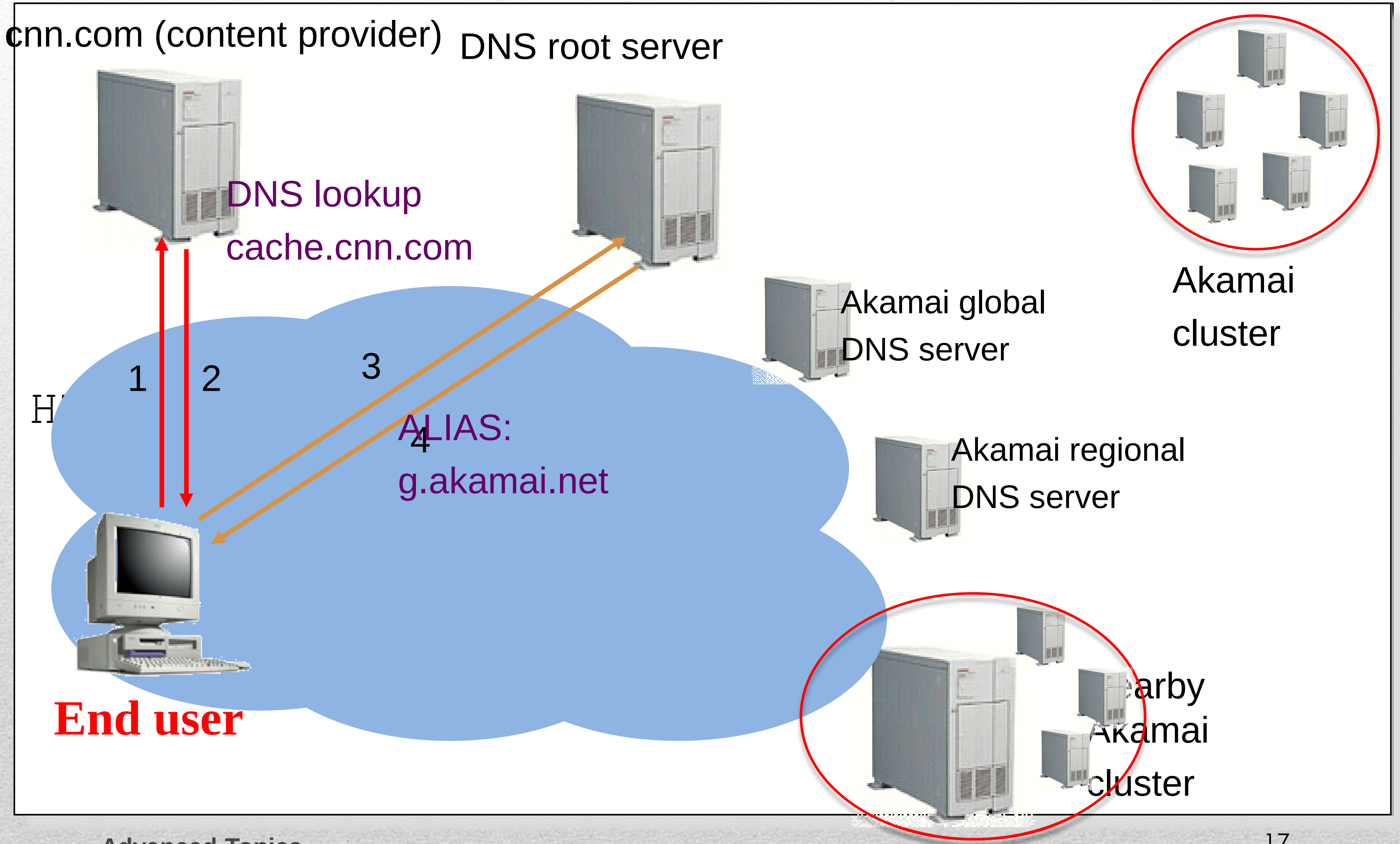
- Client requests
  - Hundreds of billions per day
  - Half in the top 45 networks
  - 15-20% of all Web traffic worldwide

Akamai & DNSs



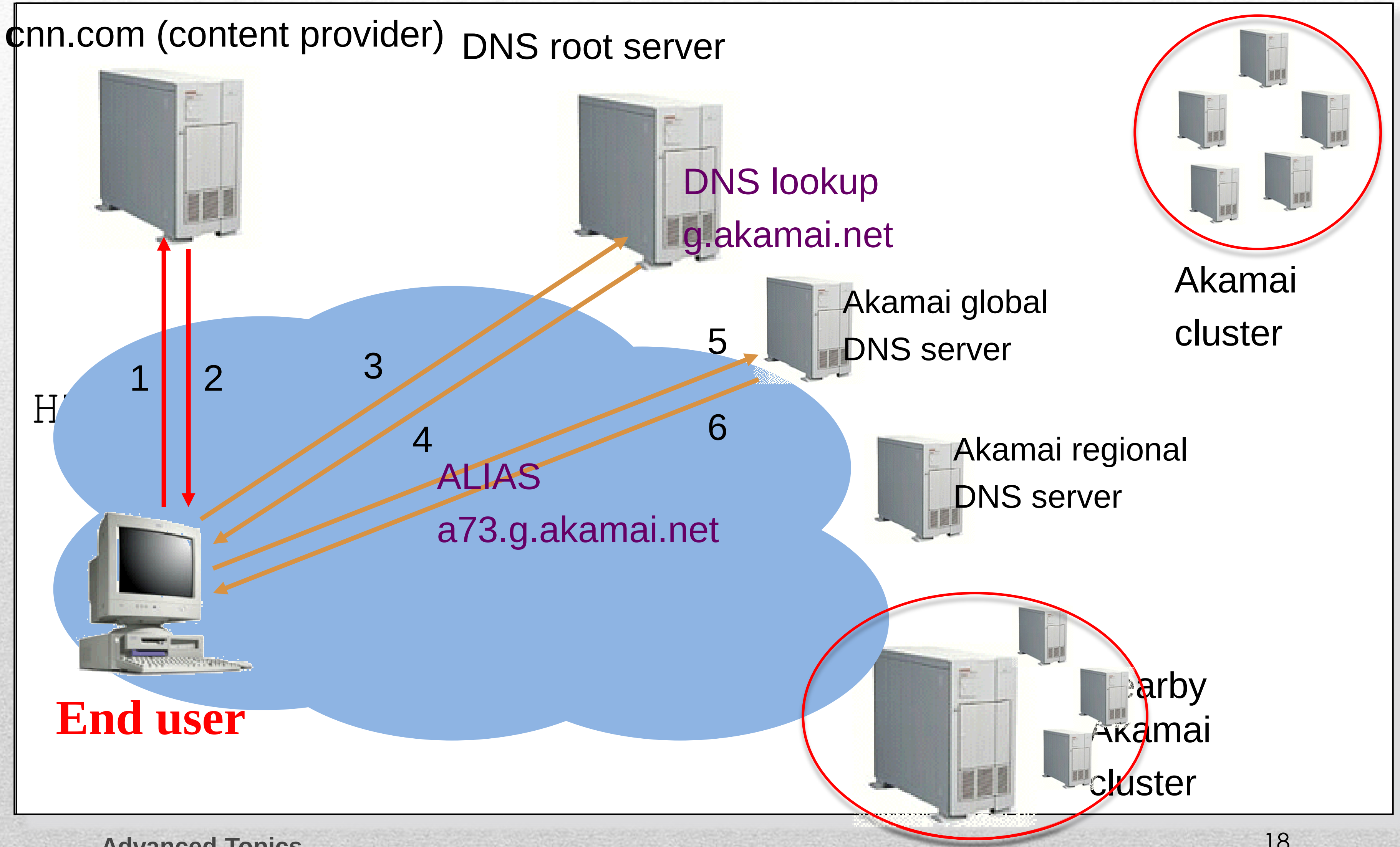
cnn.com (content provider)
DNS root server
Akamai cluster
Akamai global DNS server
Akamai regional DNS server
GET index. html
http://cache.cnn.com/foo.jpg
1
2
H
HTTP
End user
arby Akamai cluster

Akamai & DNSs



cnn.com (content provider)
DNS root server
DNS lookup
cache.cnn.com
1
2
3
4
H
ALIAS:
g.akamai.net
Akamai global
DNS server
Akamai
cluster
Akamai regional
DNS server
arby
Akamai
cluster
End user

Akamai & DNSs



cnn.com (content provider)

DNS root server

DNS lookup
g.akamai.net

Akamai global
DNS server

Akamai
cluster

Akamai regional
DNS server

1

2

3

4

5

6

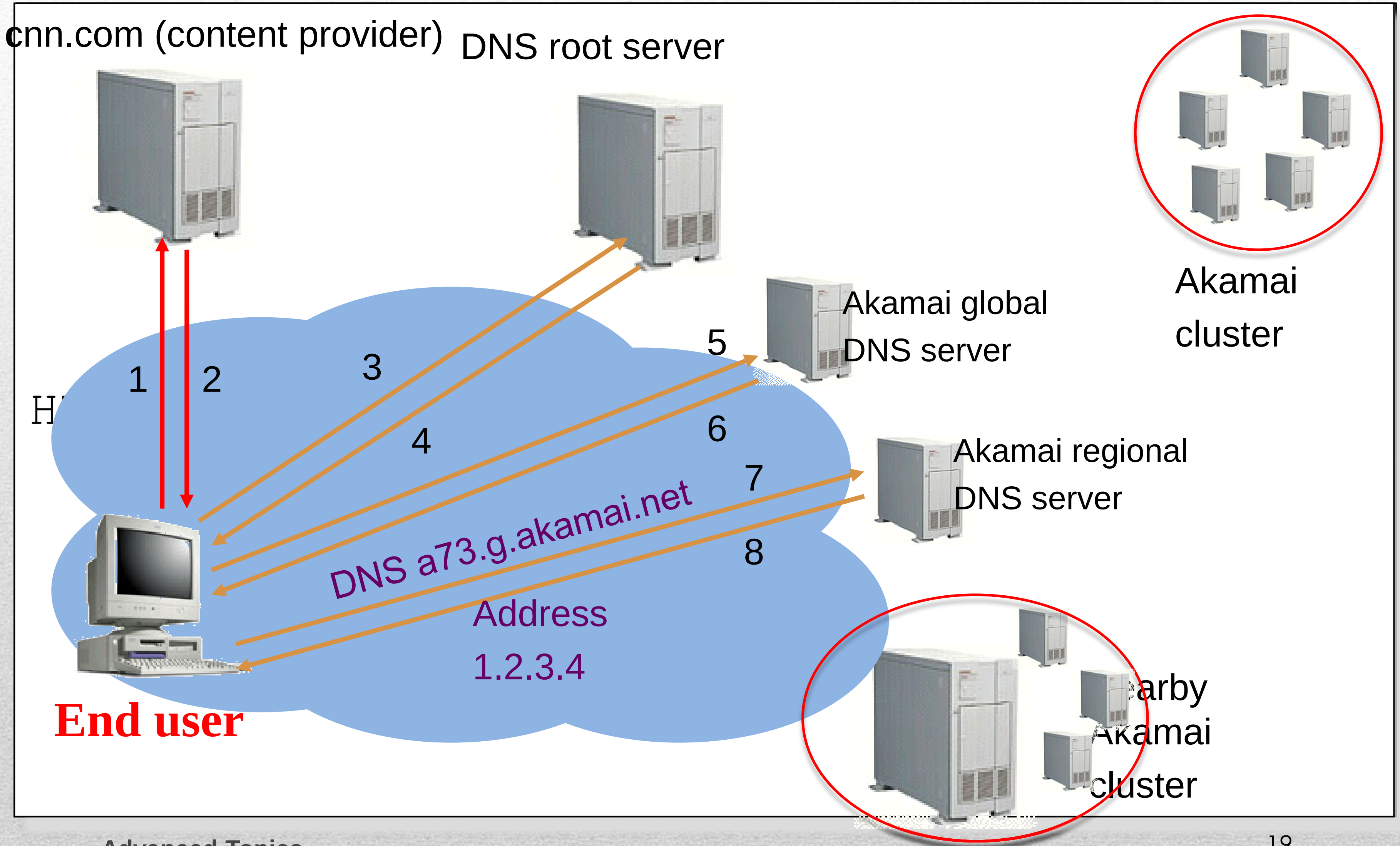
H

ALIAS
a73.g.akamai.net

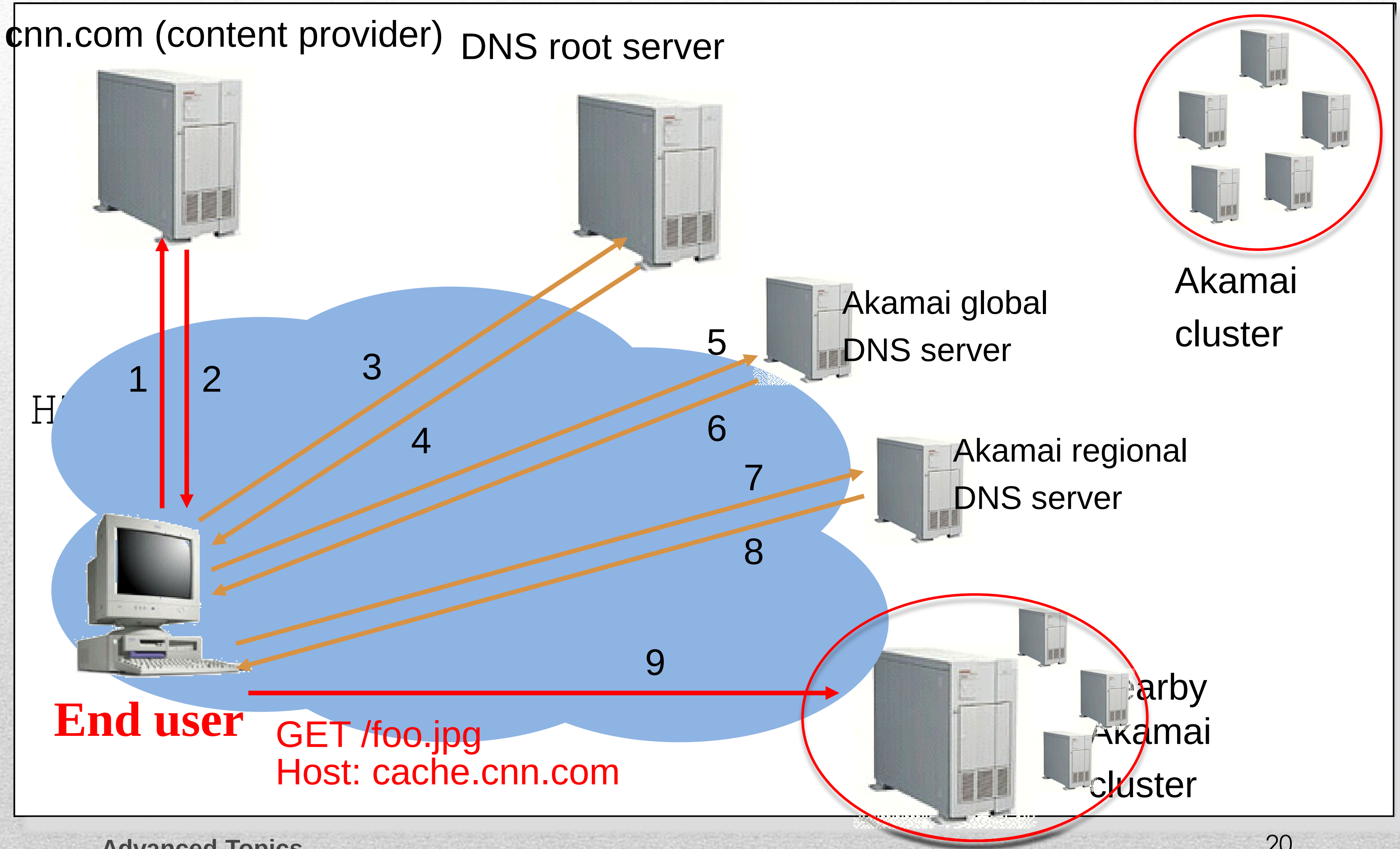
End user

arby
Akamai
cluster

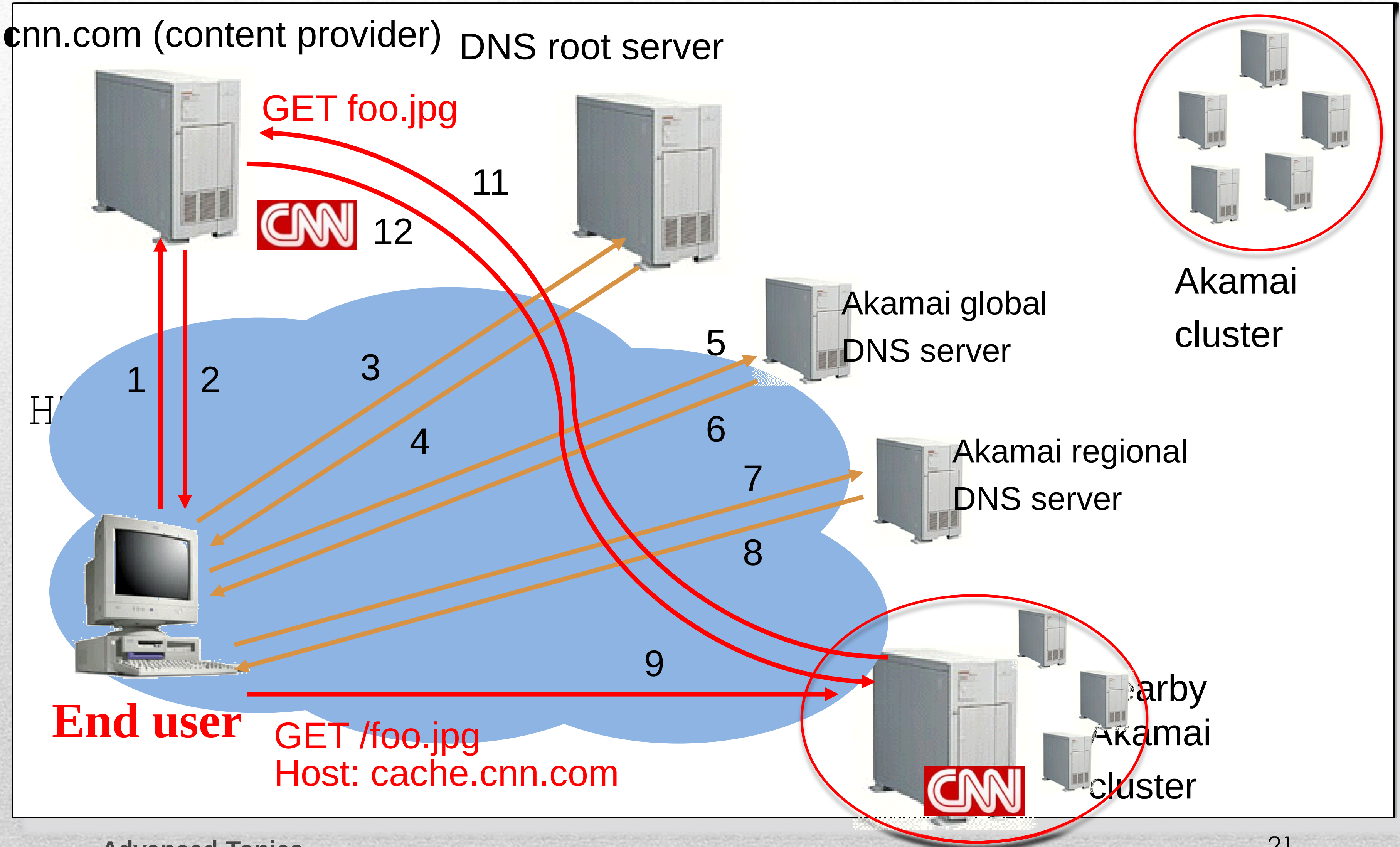
Akamai & DNSs


cnn.com (content provider)
DNS root server
Akamai cluster
Akamai global DNS server
Akamai regional DNS server
1
2
3
4
5
6
7
8
H
DNS a73.g.akamai.net
Address 1.2.3.4
End user
arby
Akamai
cluster

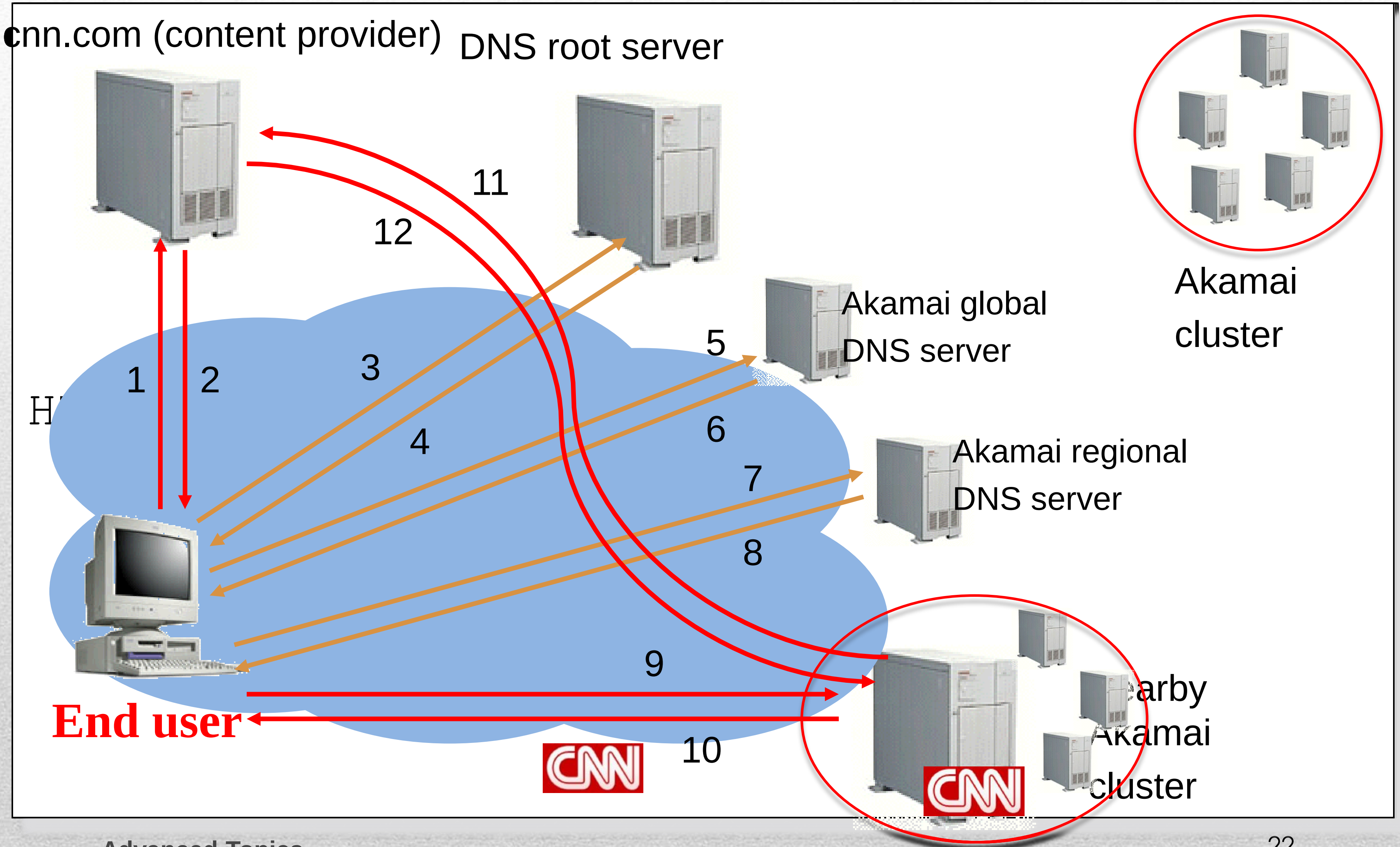
Akamai & DNSs


cnn.com (content provider)
DNS root server
Akamai cluster
Akamai global
DNS server
Akamai regional
DNS server
H
1
2
3
4
5
6
7
8
9
End user
GET /foo.jpg
Host: cache.cnn.com
arby
Akamai
cluster

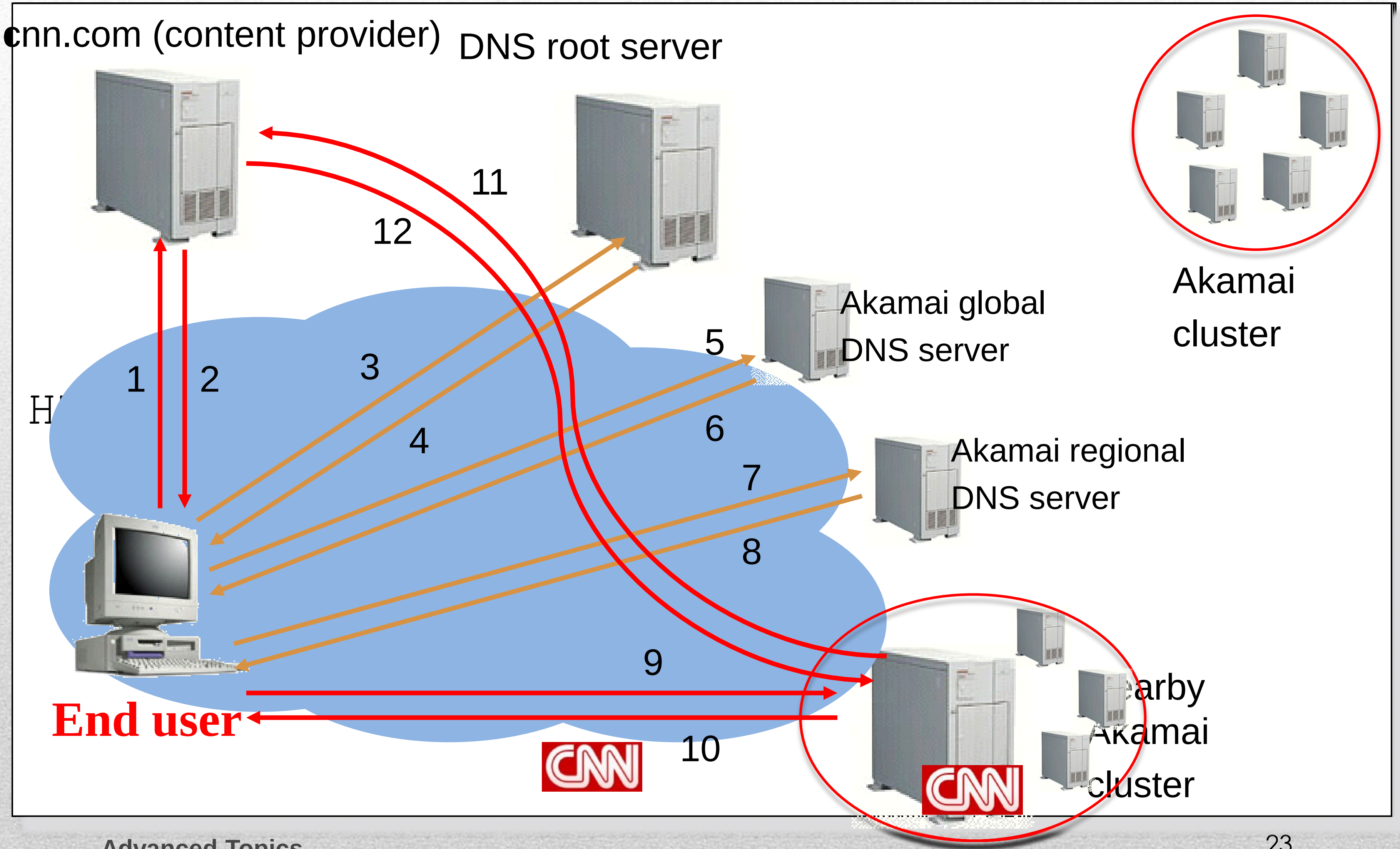
Akamai & DNSs


cnn.com (content provider)
DNS root server
GET foo.jpg
11
12
Akamai cluster
Akamai global DNS server
5
1
2
3
H
4
6
Akamai regional DNS server
7
8
9
End user
GET /foo.jpg
Host: cache.cnn.com
arby
Akamai
cluster

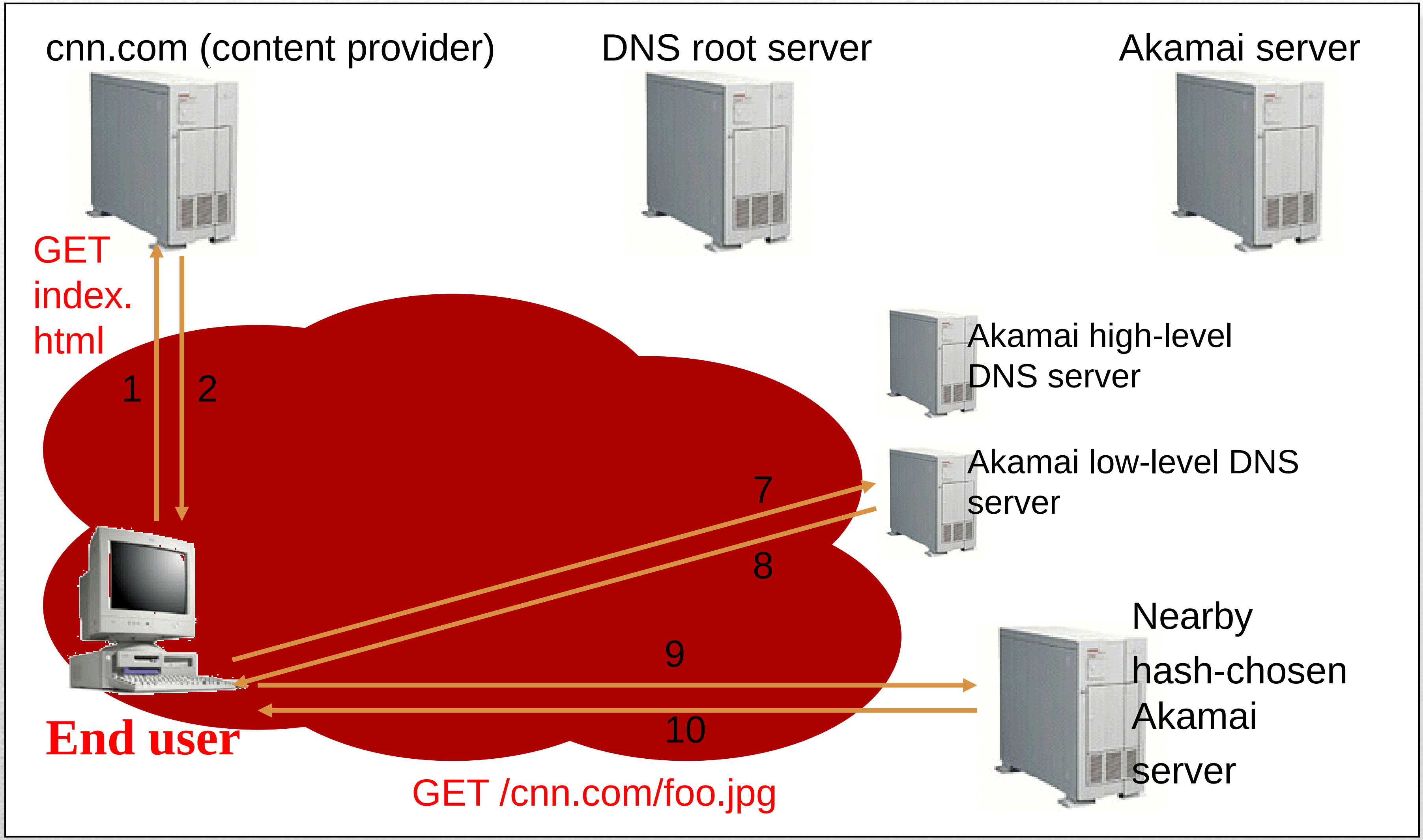
Akamai & DNSs


cnn.com (content provider)
DNS root server
11
12
Akamai global
DNS server
Akamai
cluster
1
2
3
H
4
5
6
7
Akamai regional
DNS server
8
9
End user
10
earby
Akamai
cluster

Akamai & DNSs


cnn.com (content provider)
DNS root server
Akamai cluster
11
12
Akamai global
DNS server
5
1
2
3
H
4
6
7
Akamai regional
DNS server
8
9
End user
10
CNN
CNN
arby
Akamai
cluster

Akamai & DNSs


cnn.com (content provider)
DNS root server
Akamai server
GET index.html
1
2
Akamai high-level DNS server
Akamai low-level DNS server
7
8
Nearby hash-chosen Akamai server
9
10
End user
GET /cnn.com/foo.jpg

Mapping System

- Equivalence classes of IP addresses
  - IP addresses experiencing similar performance
  - Quantify how well they connect to each other

- Collect and combine measurements
  - Ping, traceroute, BGP routes, server logs
    - E.g., over 100 TB of logs per days
  - Network latency, loss, and connectivity

- Map each IP class to a preferred server cluster
  - Based on performance, cluster health, etc.
  - Updated roughly every minute

- Map client request to a server in the cluster
  - Load balancer selects a specific server
  - E.g., to maximize the cache hit rate

Adapting to Failures

- Failing hard drive on a server
  - Suspends after finishing “in progress” requests

- Failed server
  - Another server takes over for the IP address
  - Low-level map updated quickly

- Failed cluster
  - High-level map updated quickly

- Failed path to customer’s origin server
  - Route packets through an intermediate node

Akamai Transport Optimizations

- Bad Internet routes
  - Overlay routing through an intermediate server
- Packet loss
  - Sending redundant data over multiple paths
- TCP connection set-up/teardown
  - Pools of persistent connections
- TCP congestion window and round-trip time
  - Estimates based on network latency measurements

Akamai Application Optimizations

- Slow download of embedded objects
    - Prefetch when HTML page is requested

- Large objects
    - Content compression

- Slow applications
    - Moving applications to edge servers
    - E.g., content aggregation and transformation
    - E.g., static databases (e.g., product catalogs)
    - E.g. batching and validating input on Web forms

Conclusion

- Content distribution is hard
  - Many, diverse, changing objects
  - Clients distributed all over the world
  - Reducing latency
- Contribution distribution solutions
  - Reactive caching
  - Proactive content distribution networks