Content Delivery and File Sharing in the Modern Internet

Henry M. Levy

Department of Computer Science and Engineering
University of Washington
Seattle, WA

Content Delivery and File Sharing in the Modern Internet

Henry M. Levy

Department of Computer Science and Engineering
University of Washington
Seattle, WA

Work with:

- **Steve Gribble**
- **John Zahorjan**

- Stefan Saroiu
- Krishna Gummadi
- Richard Dunn

Outline and Goals

1. Explore and characterize content delivery in today's Internet
   - Web, Akamai CDN, Kazaa & Gnutella P2P
     - » What is the impact of these systems on the Internet?
     - » What are the characteristics of the new delivery systems?

2. Understand what drives P2P file sharing systems
   - Dynamics of the web are understood
     - » Driven by changes to documents, Zipf popularity
   - Dynamics of P2P are unknown

Outline

1. Characterize content delivery in today's Internet

2. Understand what drives P2P file sharing systems

Surprise: The WWW is changing even more rapidly now!

- Thirst for data (+ new types of data) continues to increase.
- People are using new means to provide and obtain that data.

- The result -- the web is now seeing a mixture of new and old content-delivery mechanisms:
  - Conventional web clients and web servers
  - Global-scale content-delivery networks (e.g., Akamai)
  - Self-organizing peer-to-peer file-sharing systems (e.g., Gnutella, Kazaa)

Quick overview: Peer-to-peer (P2P) systems

- Peers are individually owned computers, most on modems or broadband
- Peers collaborate to exchange content among themselves
- Each peer is both a client and server
- Peers issue or broadcast text queries to the peer network to find content

- Example -- Kazaa:
  - No centralized components
  - Two-level structure – some peers are "supernodes"
  - Supernode indexes content from the peers underneath it
  - Supernodes can communicate with each other to find content
  - Files transferred in segments from multiple peers simultaneously
  - The protocol is proprietary

Peer-to-peer systems

- These systems are technically interesting
- They are autonomous, totally distributed, self-organizing

- There is a *huge* amount of research on P2P
- There is *almost no data* on P2P

- **Questions:**
  - **What are the characteristics of the new P2P systems?**
  - **What is the impact of these P2P systems on the web?**

Methodology

- Data is based on a 9-day trace collected at UW from May 28, 2002 through June 6, 2002.
- We use *passive network monitoring*.
- Our trace machine sees every packet going in and out.
- We classify traffic based on port number and other information in the message headers.
- We anonymize all sensitive information before writing to disk.

- Trace machine is a dual-CPU Dell with 2.0 GHz Xeon processors and a gigabit network card, running FreeBSD.

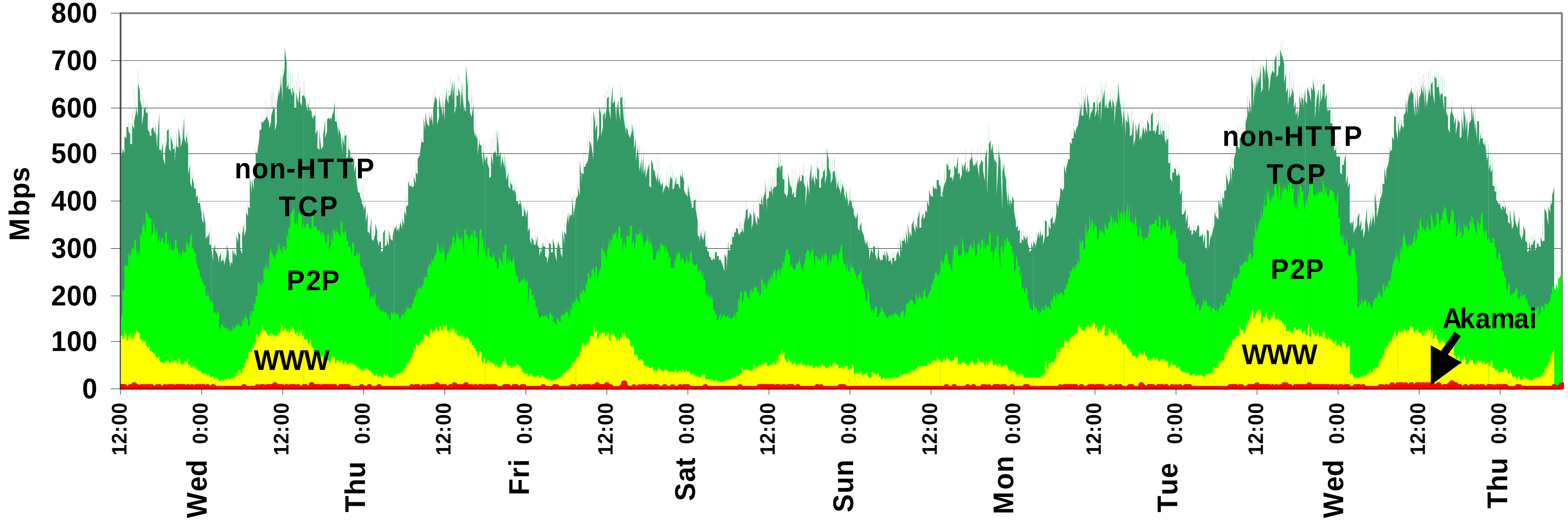
Limitations

- Only studied one population (UW)
- Finite trace period
- Could see data transfers, but not encrypted control traffic
- Cannot see UW-internal traffic

Question 1:

- What is the bandwidth impact of new P2P and content-delivery systems?

Where has all the bandwidth gone?



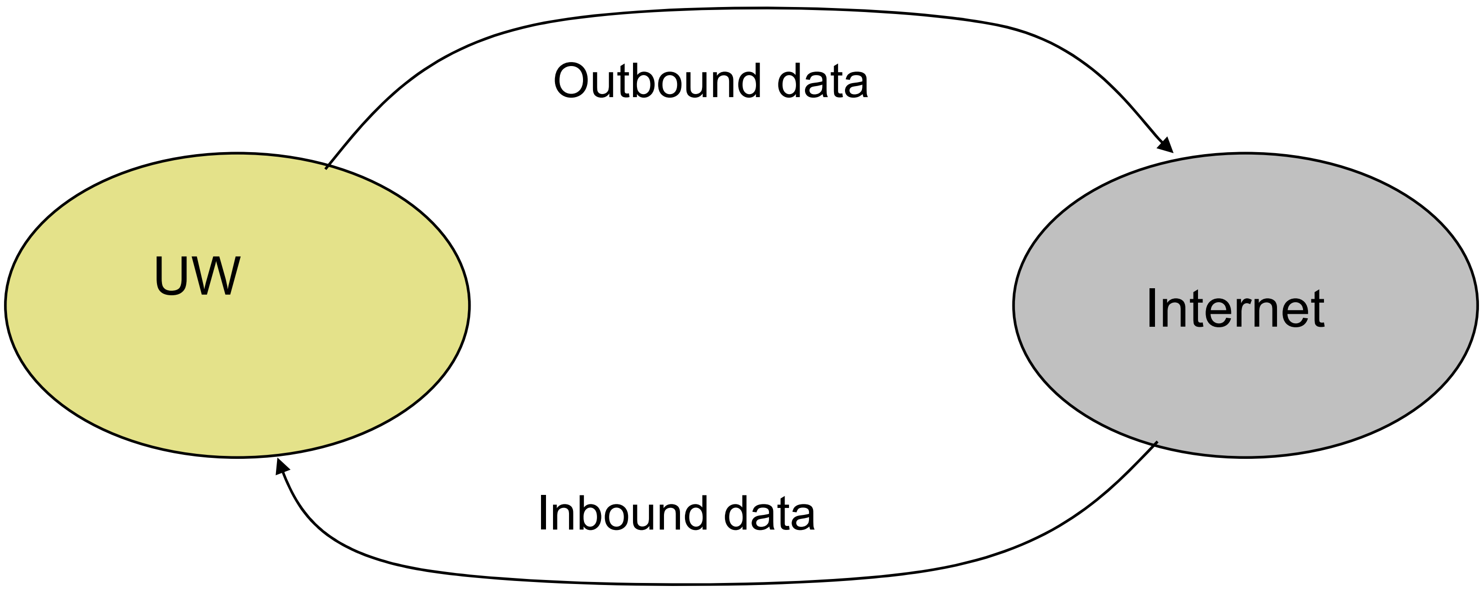
Breakdown of UW TCP bandwidth into HTTP Components (May 2002)

- WWW = 14% of TCP traffic; P2P = 43% of TCP traffic
- **P2P dominates WWW in bandwidth consumed!!**

Definition:

- Inbound traffic: data objects requested by UW clients, transmitted into the UW from an outside source.

- Outbound traffic: data objects requested by a source external to UW, transmitted by a UW server.



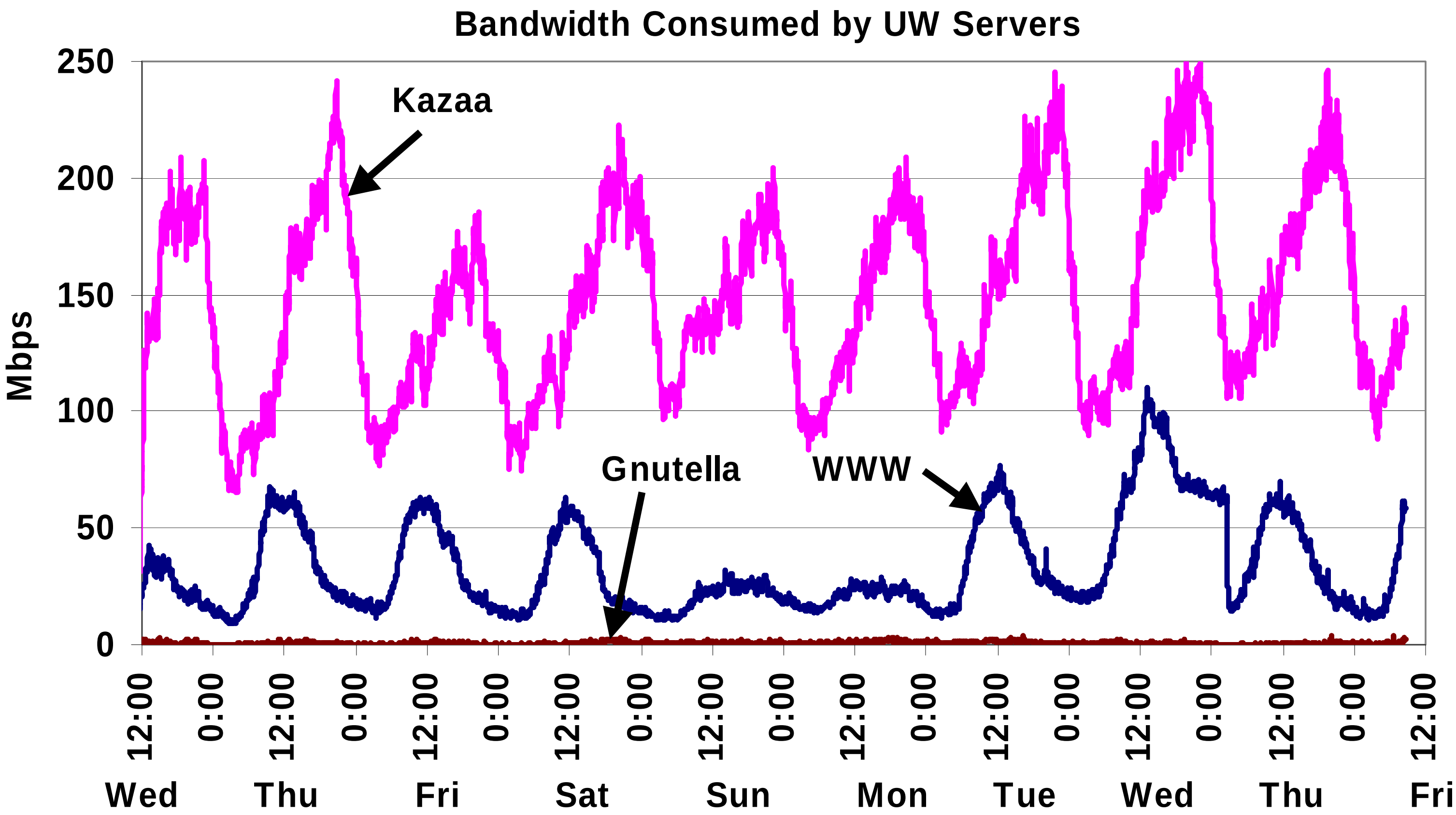
Detailed Trace Statistics

	WWW	
	inbound	*outbound*
HTTP Xactions	**323,072,253**	**73,001,891**
Unique objects	**72,818,997**	**3,412,647**
Clients	**39,285**	**1,231,308**
Servers	**403,087**	**9,821**
Bytes Xferred	**1.51 TB**	**3.02 TB**
Median Object Size	**1,976 B**	**4,646 B**
Mean Object Size	**22,491**	**334,944**

Detailed Trace Statistics

	WWW		*Kazaa*	
	inbound	*outbound*	*inbound*	*outbound*
HTTP Xactions	**323,072,253**	**73,001,891**	**11,140,861**	**19,190,902**
Unique objects	**72,818,997**	**3,412,647**	**111,437**	**166,442**
Clients	**39,285**	**1,231,308**	**4,644**	**611,005**
Servers	**403,087**	**9,821**	**281,026**	**3,888**
Bytes Xferred	**1.51 TB**	**3.02 TB**	**1.78 TB**	**13.57 TB**
Median Object Size	**1,976 B**	**4,646 B**	**3.75 MB**	**3.67 MB**
Mean Object Size	**22,491**	**334,944**	**29.4MB**	**26.1MB**

Bandwidth consumed by UW servers (outbound traffic)



October 22, 2002

THE DAILY online
UNIVERSITY OF WASHINGTON Seattle

Bandwidth restrictions save almost $1 million

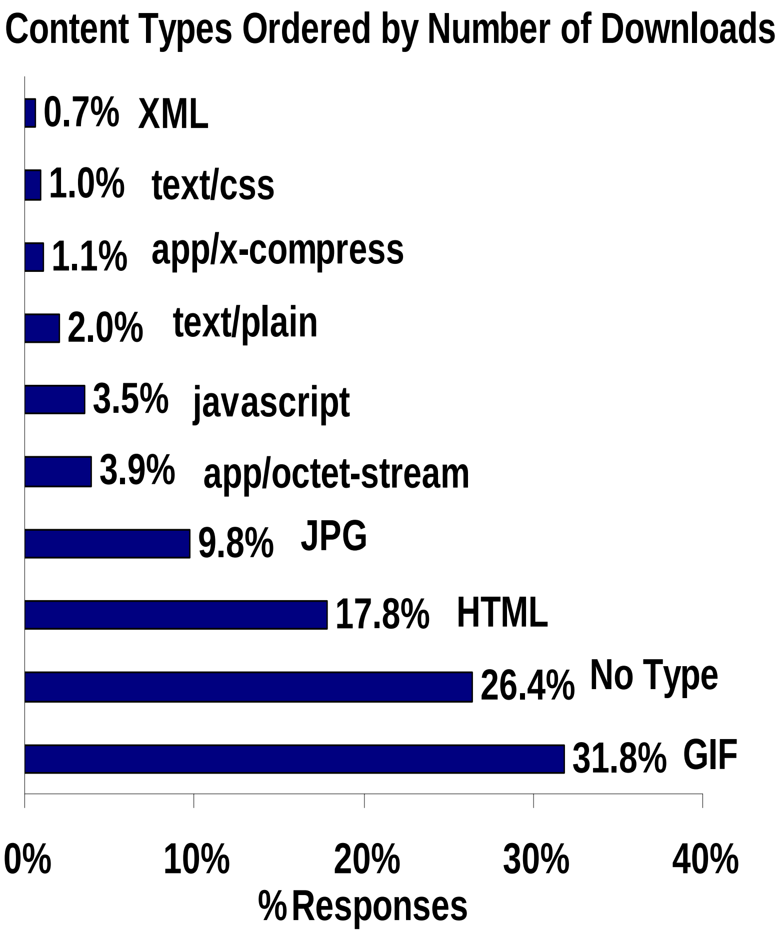
by Steven Friederich
10/22/2002

Those using peer-to-peer software on campus, such as the file-sharing program KaZaA, may notice their network connection has been acting slow lately. New technical restrictions placed within the campus networks have provided a limit to the amount of bandwidth users may access for Web sites and servers.

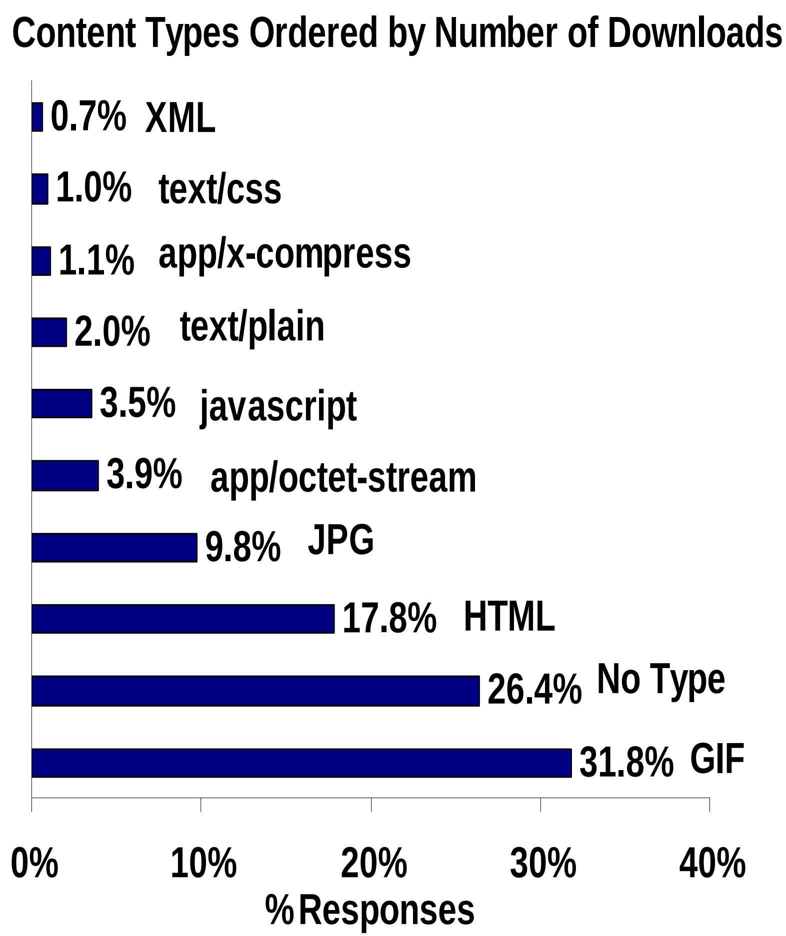
Question 2:

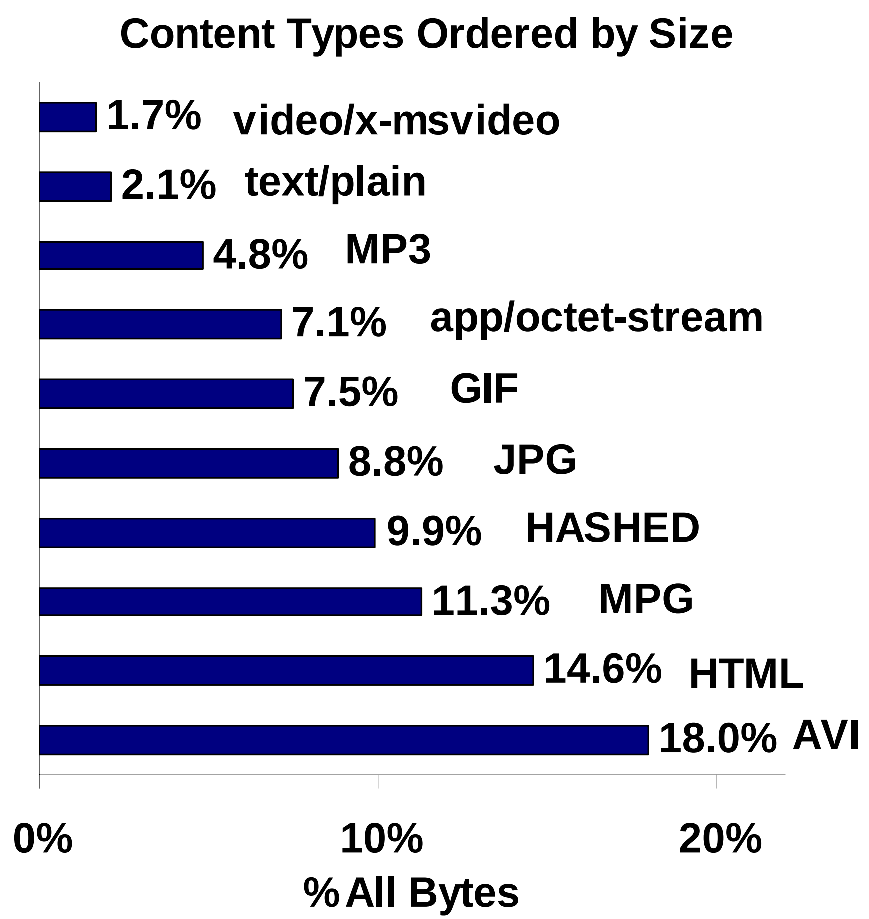
- What are all the the bytes carrying?

What data types are being downloaded (HTTP)?

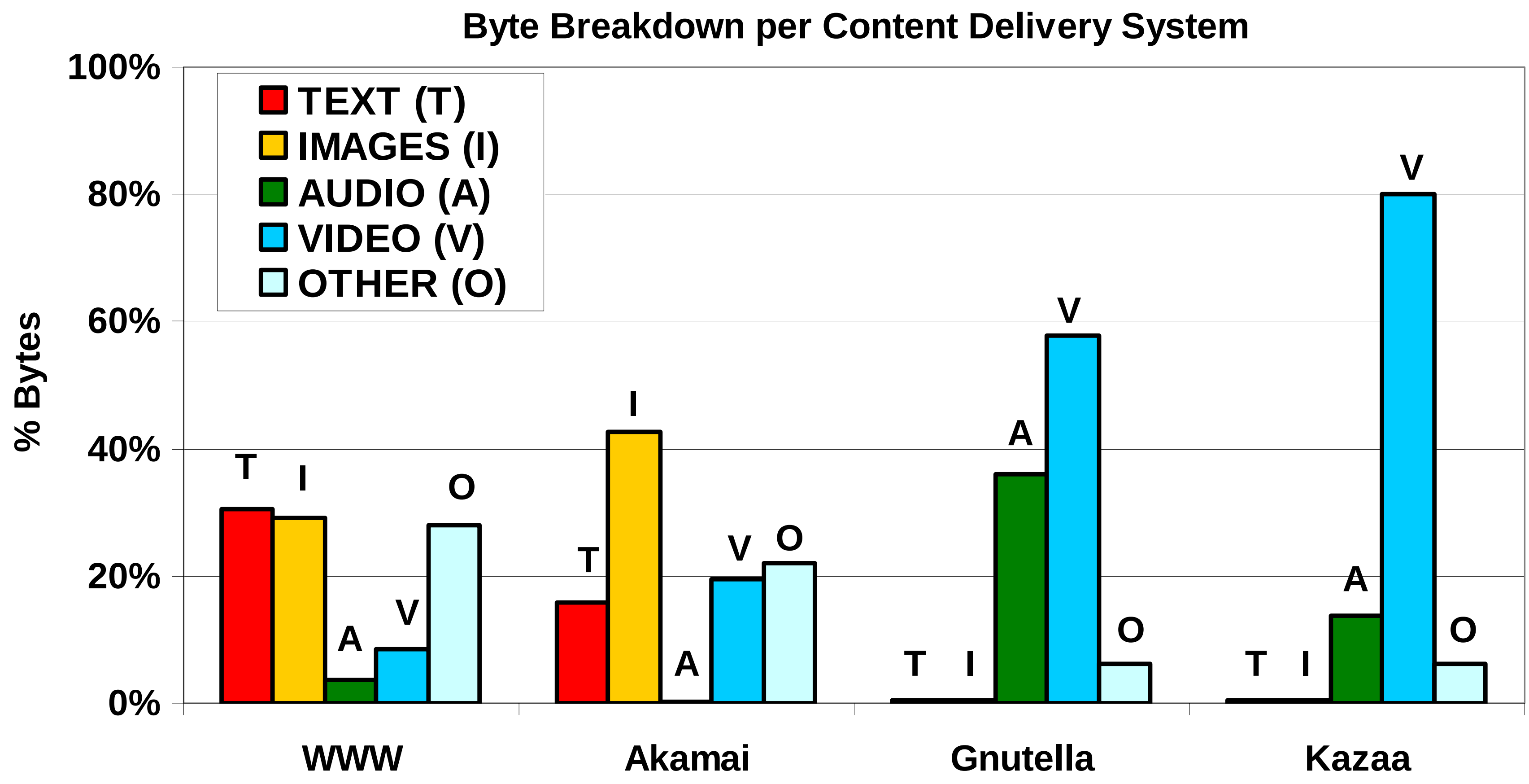


What data types are being downloaded (HTTP)?



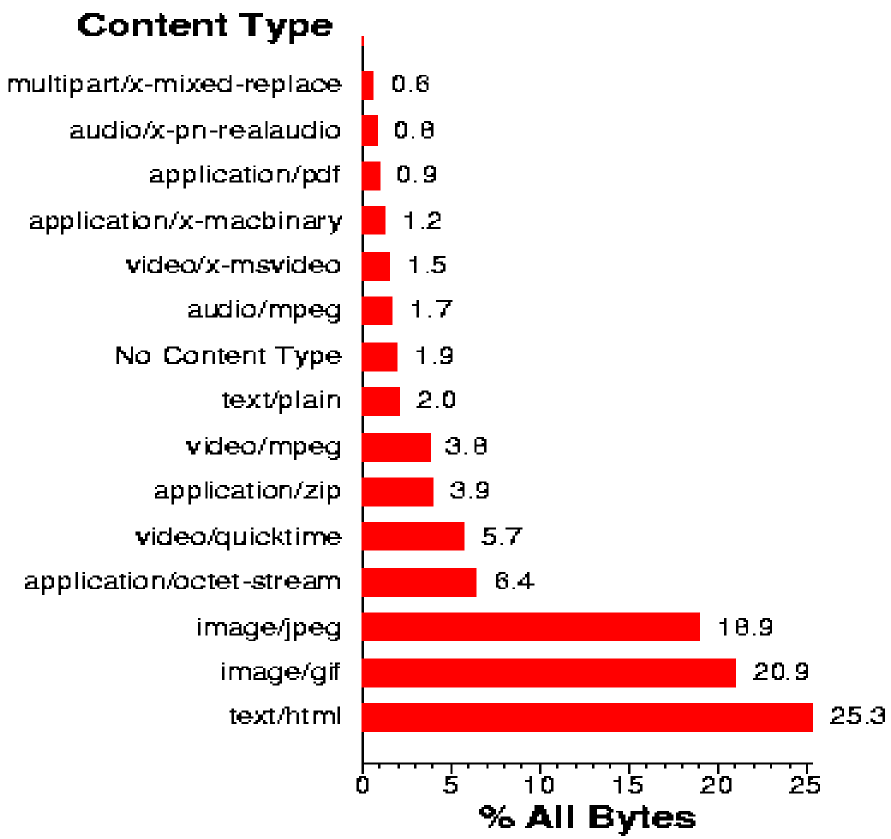


Object type for different systems



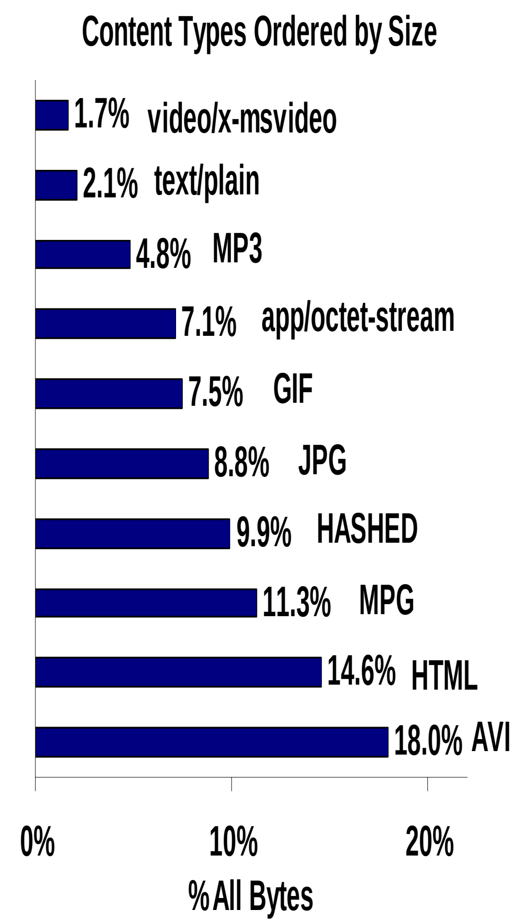
HTTP download content type: now vs. then

May 1999



(b) By Bytes

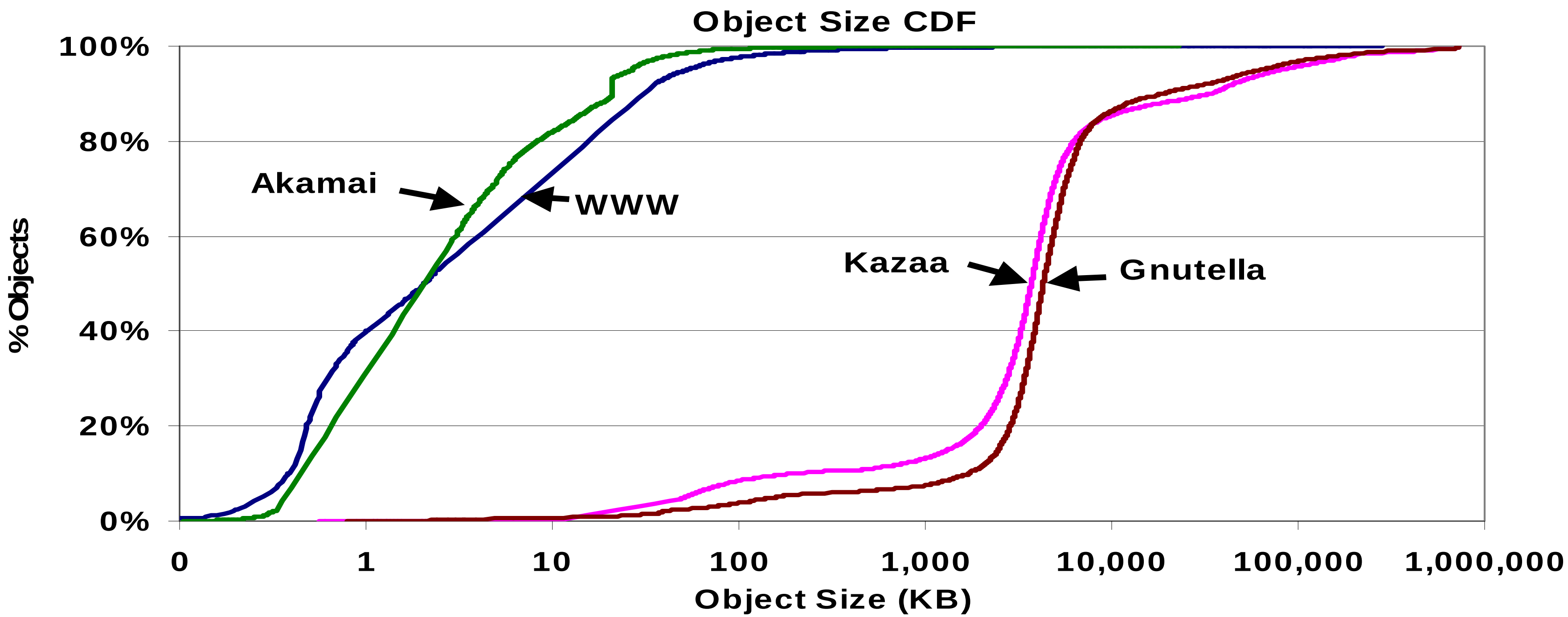
May 2002



Question 3:

- How do workload characteristics differ in P2P systems, compared to www?

Object size

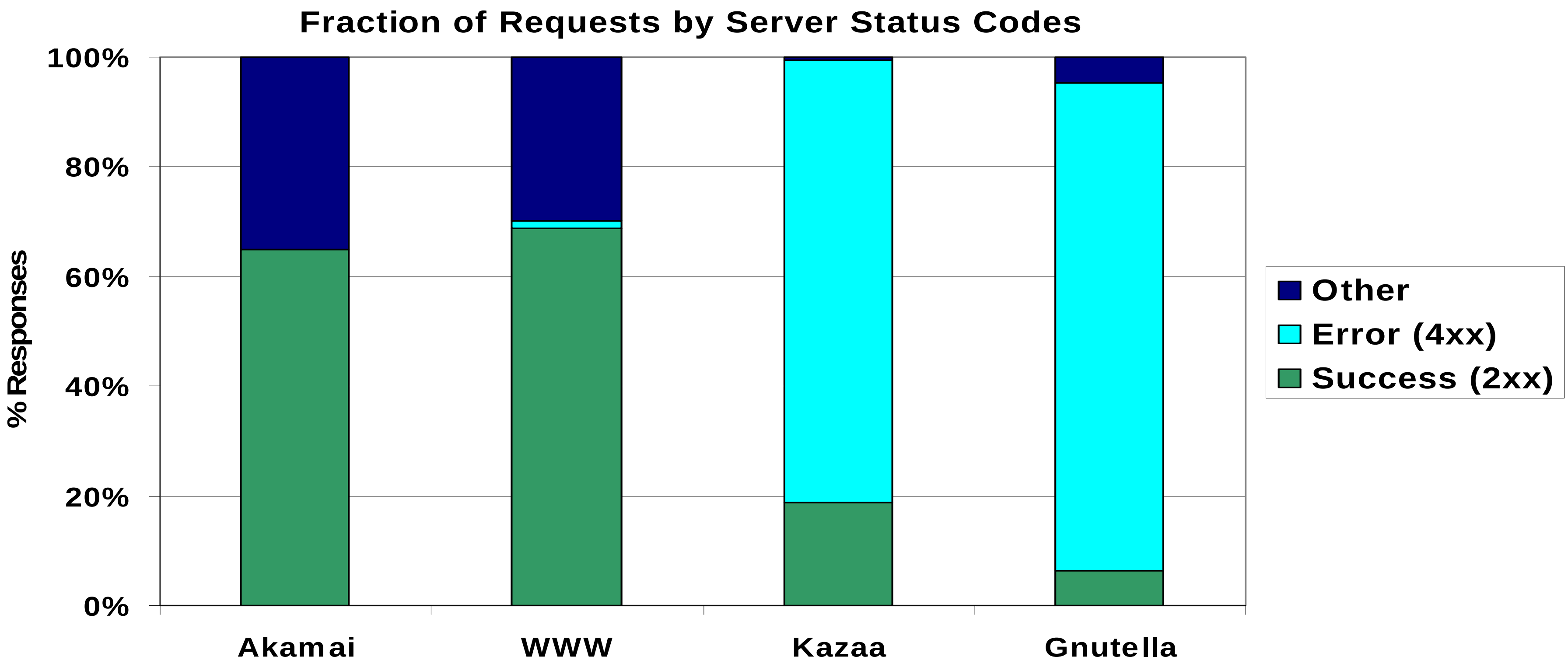


The median Kazaa object is **1000 times larger** than the median WWW object!

Top BW-consuming Kazaa objects

	Kazaa (inbound)				Kazaa (outbound)			
	object size (MB)	GB consumed	clients	srvers	object size (MB)	GB consumed	clients	srvers
1	694	8.1	20	164	696	119.0	397	1
2	702	6.4	14	91	699	110.5	1000	4
3	690	6.1	22	83	699	78.7	390	10
4	775	5.6	16	105	700	73.3	558	2
5	698	4.7	14	74	634	64.9	540	1

Availability

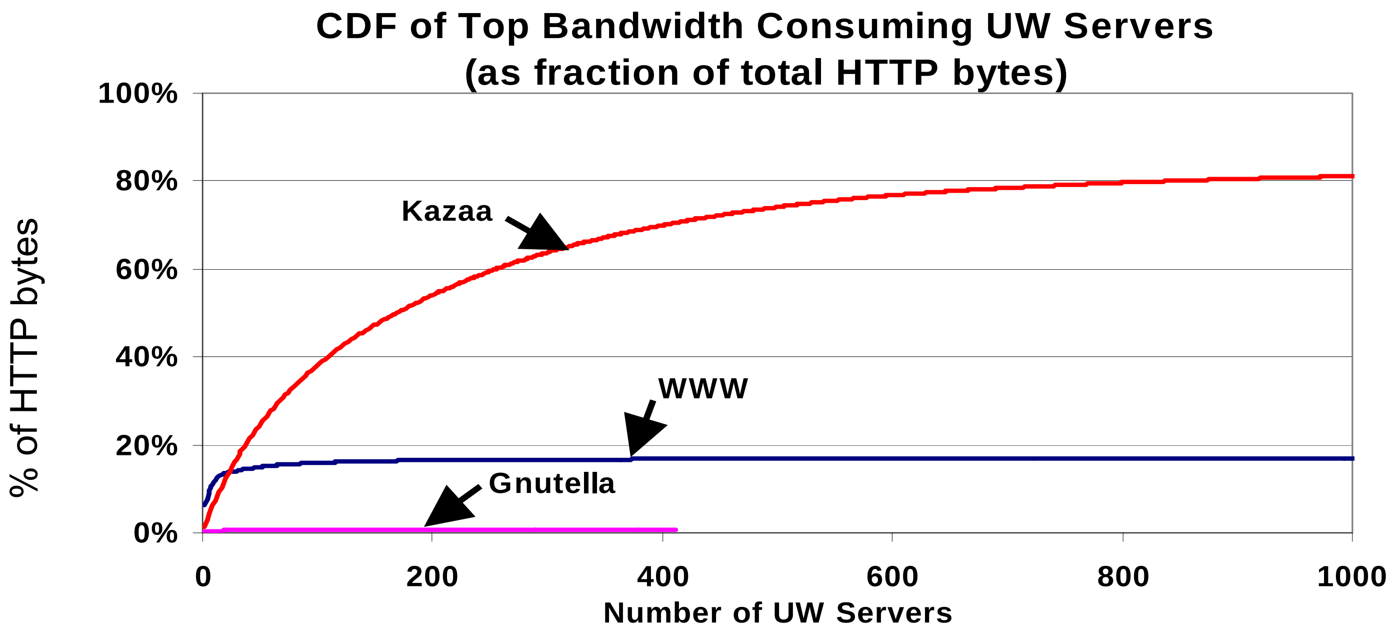


It's mostly unavailable, but it's free (and as long as you get it eventually, you don't care)

Question 4

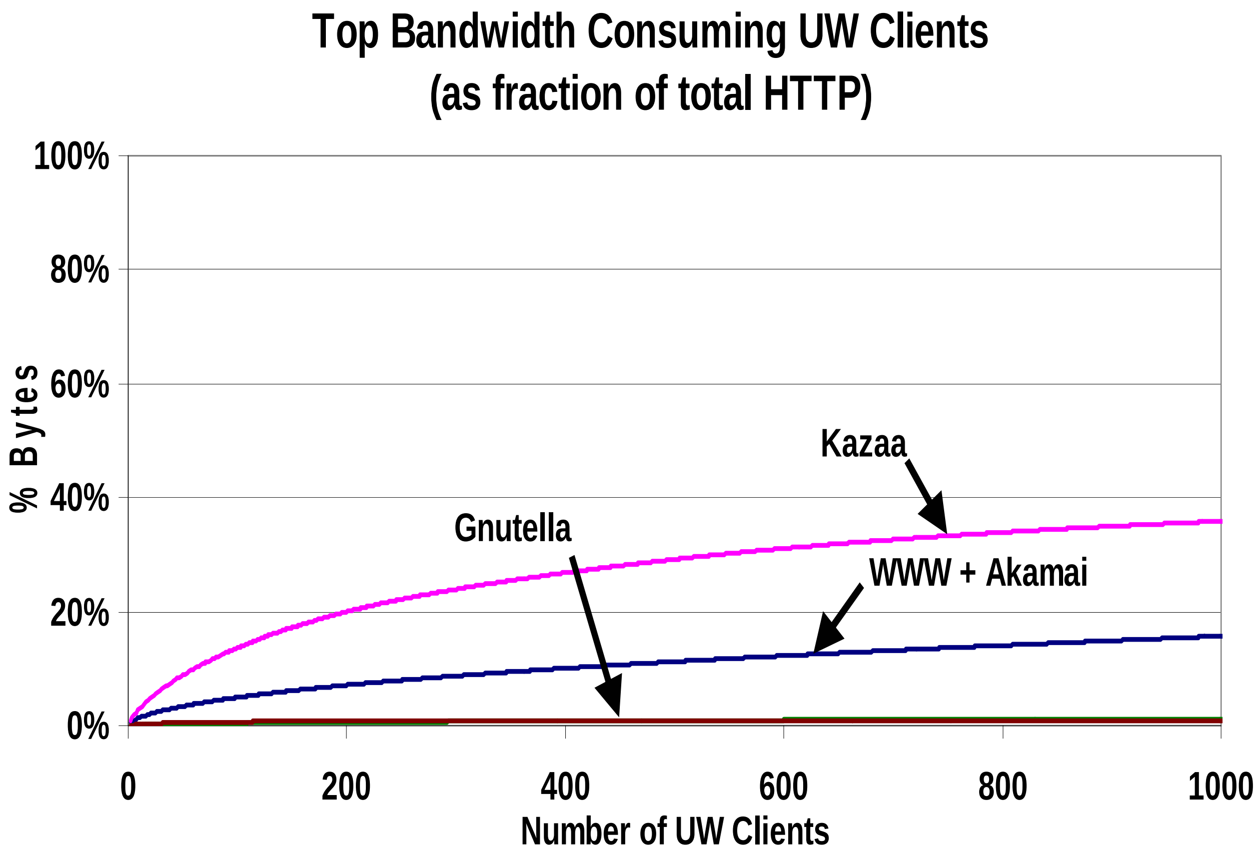
- How is bandwidth use distributed among Kazaa clients and servers?

Top bandwidth producing UW servers



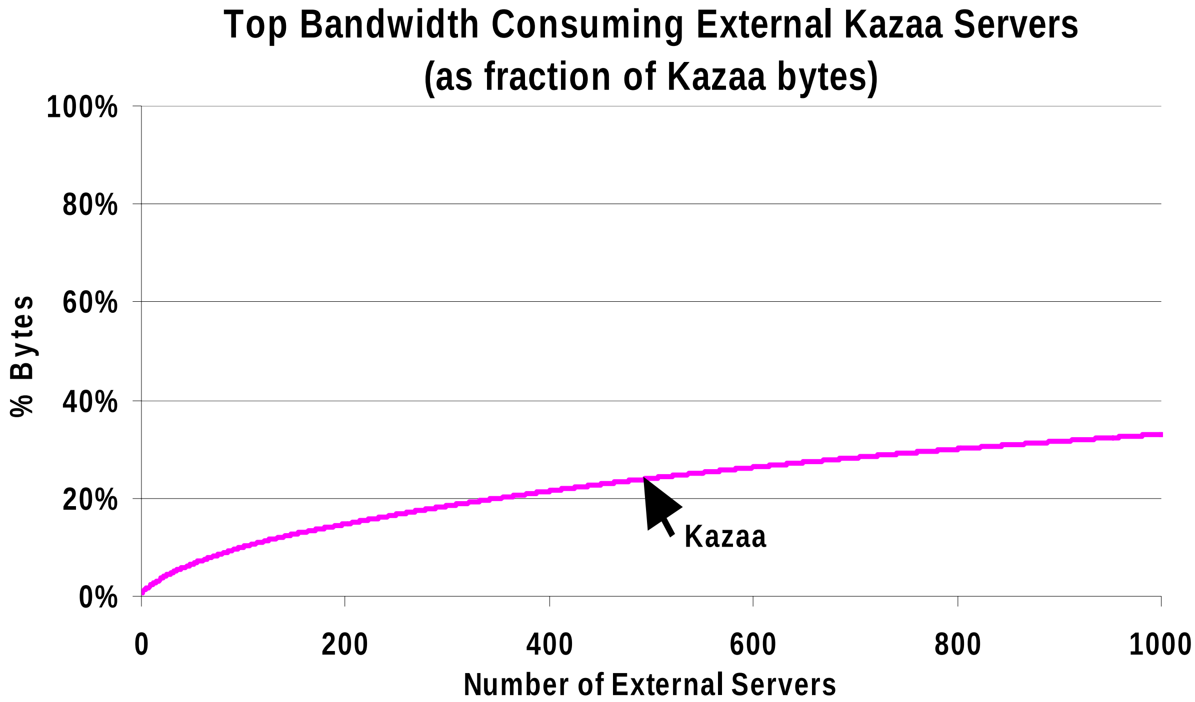
- 20 WWW servers supply 20% of HTTP bytes served
- 400 Kazaa peers generate 70% of all HTTP outgoing traffic!

Top bandwidth consuming UW clients



- Top 200 Kazaa clients are responsible for 20% of all HTTP bytes downloaded

Top bandwidth producing external Kazaa servers



- 600 external Kazaa peers (out of 281K) serve 26% of the Kazaa bytes to UW.
- Surprising -- given the “scalability based design” of P2P systems and Kazaa.

Summary of CDN/P2P study

- Internet has undergone a huge qualitative change in only a few years.
- We've moved from interactive transfer of small files (10s of KBs) to batch transfer of enormous files (100s of MBs).
- P2P now accounts for the majority of HTTP bytes, exceeding WWW traffic by nearly 3X at UW
- P2P documents avg. 3 orders of magnitude larger than WWW docs
- A small number of huge objects are responsible for an enormous fraction of transfers (300 Kazaa objects used 5.6TB of BW!)
- A small number of P2P clients are causing much of the traffic

- **On average, a single UW Kazaa peer uses <u>90 times</u> the bandwidth of a single WWW client!**

Outline

- Characterize content delivery in today's Internet

- Understand what drives Kazaa P2P file-sharing system
  1. Some observations about Kazaa
  2. A model for studying P2P multimedia systems
  3. Locality-aware request distribution

Methodology: trace characteristics

- 6-month Kazaa trace gathered at UW's border router

start date	May 28th, 2002
end date	December 17th, 2002
trace length	203 days, 5 hours, 6 minutes
# of requests	1,640,912
# of transactions	98,997,622
# of unsuccessful transactions	65,505,165 (66.2%)
# of clients	24,578
# of unique objects	633,106 (totaling 8.85TB)
bytes transferred	22.72TB
content demanded	43.87TB

Outline

- Characterize content delivery in today's Internet

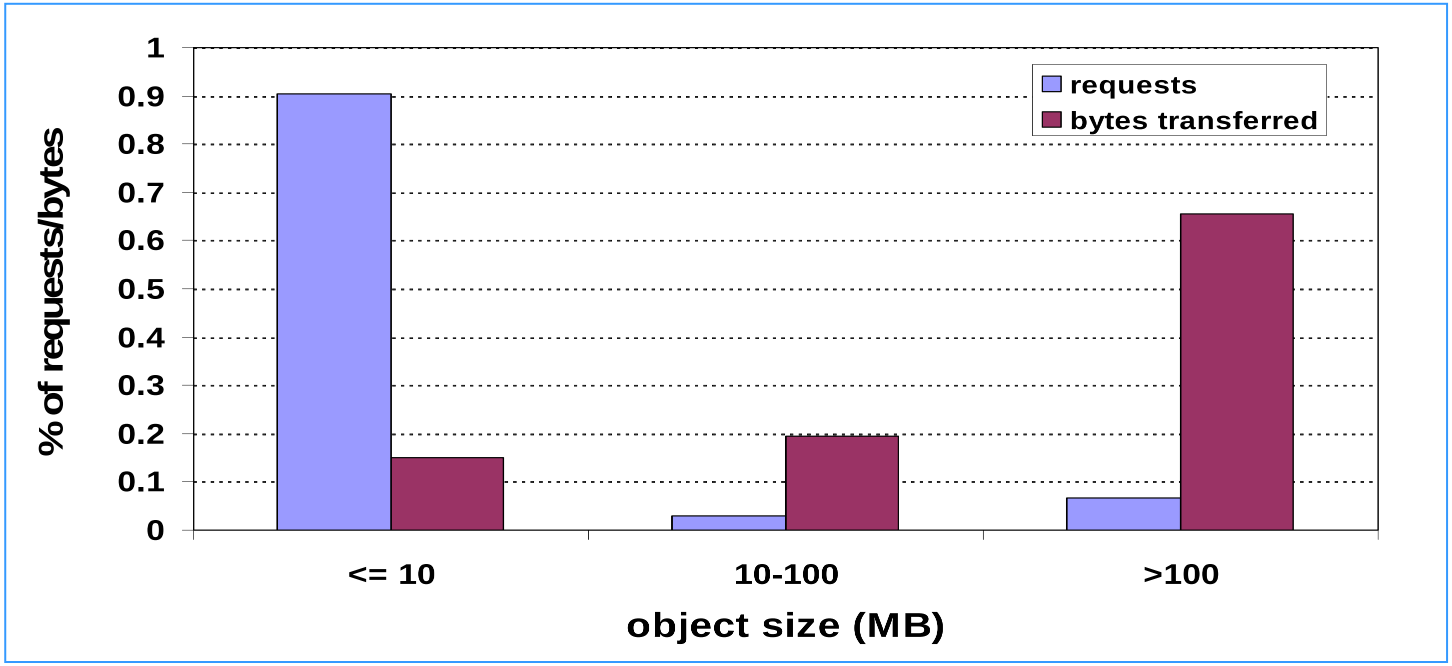
- Understand what drives the Kazaa P2P file-sharing system
  1. Some observations about Kazaa
  2. A model for studying P2P multimedia systems
  3. Locality-aware request distribution

1. Kazaa is really 2 phenomena

- Kazaa is ***peer-to-peer***
  - *How* files are exchanged
  - Content delivery system built out of volunteers
  - Goal: understand behavior of the peers and implications

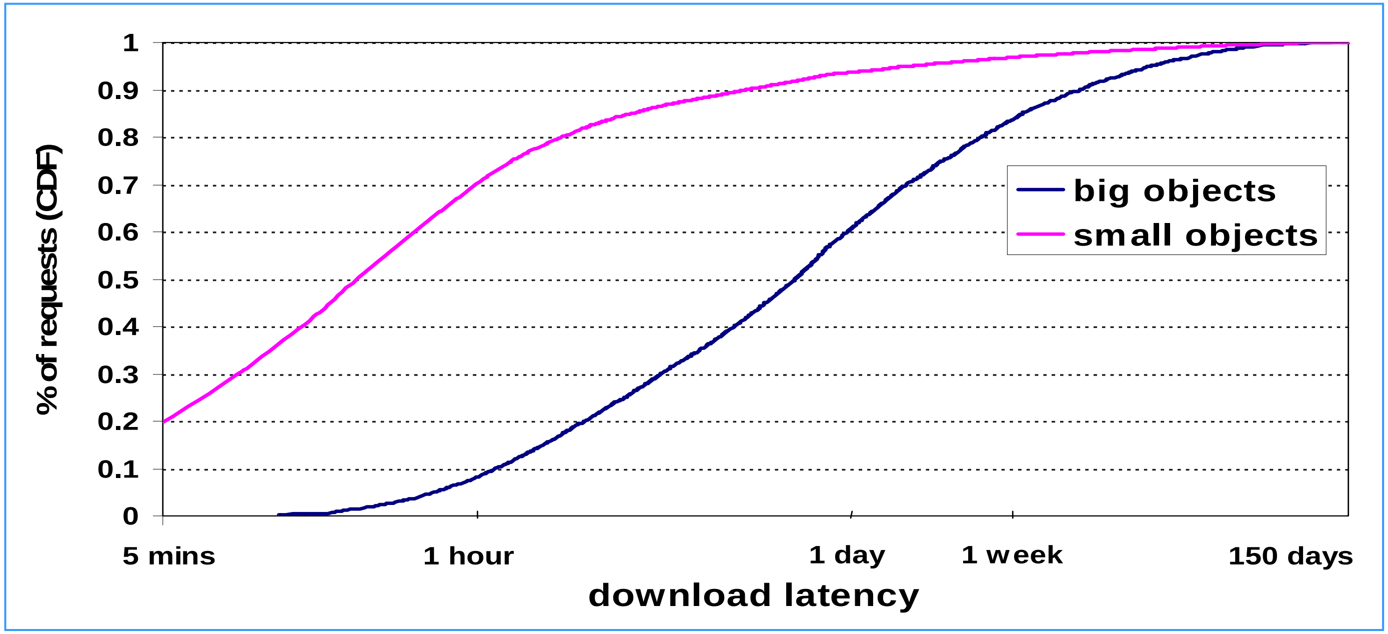
- Kazaa is a ***multimedia*** *workload*
  - *What* is exchanged
  - Broadband, MP3, huge disks sparked a new workload
  - Goal: understand implications of widespread multimedia on Internet, regardless of P2P vs. central, legality

2. Multimedia is really 2 workloads



- If you care about:
  - making users happy: make sure audio arrives quickly
  - making IT dept. happy: cache or rate limit video

3. Users are remarkably patient



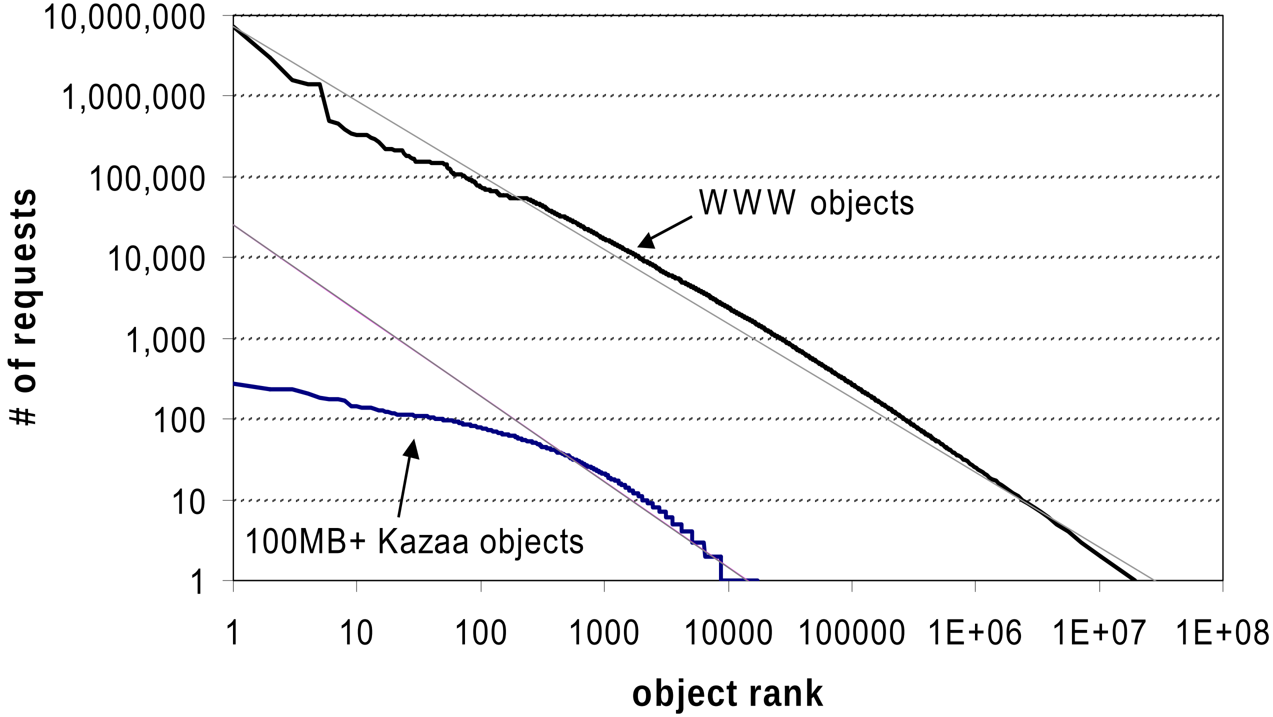
- **audio files take about an hour, video files a day**
  - but in either case, people will wait weeks!
- **Web is an interactive system; Kazaa is a batch system!**

4. Objects are immutable

- the Web is driven by **object change**
  - users revisit popular sites as their content changes
  - rate of change limits Web cache effectiveness [Wolman et al. 99]

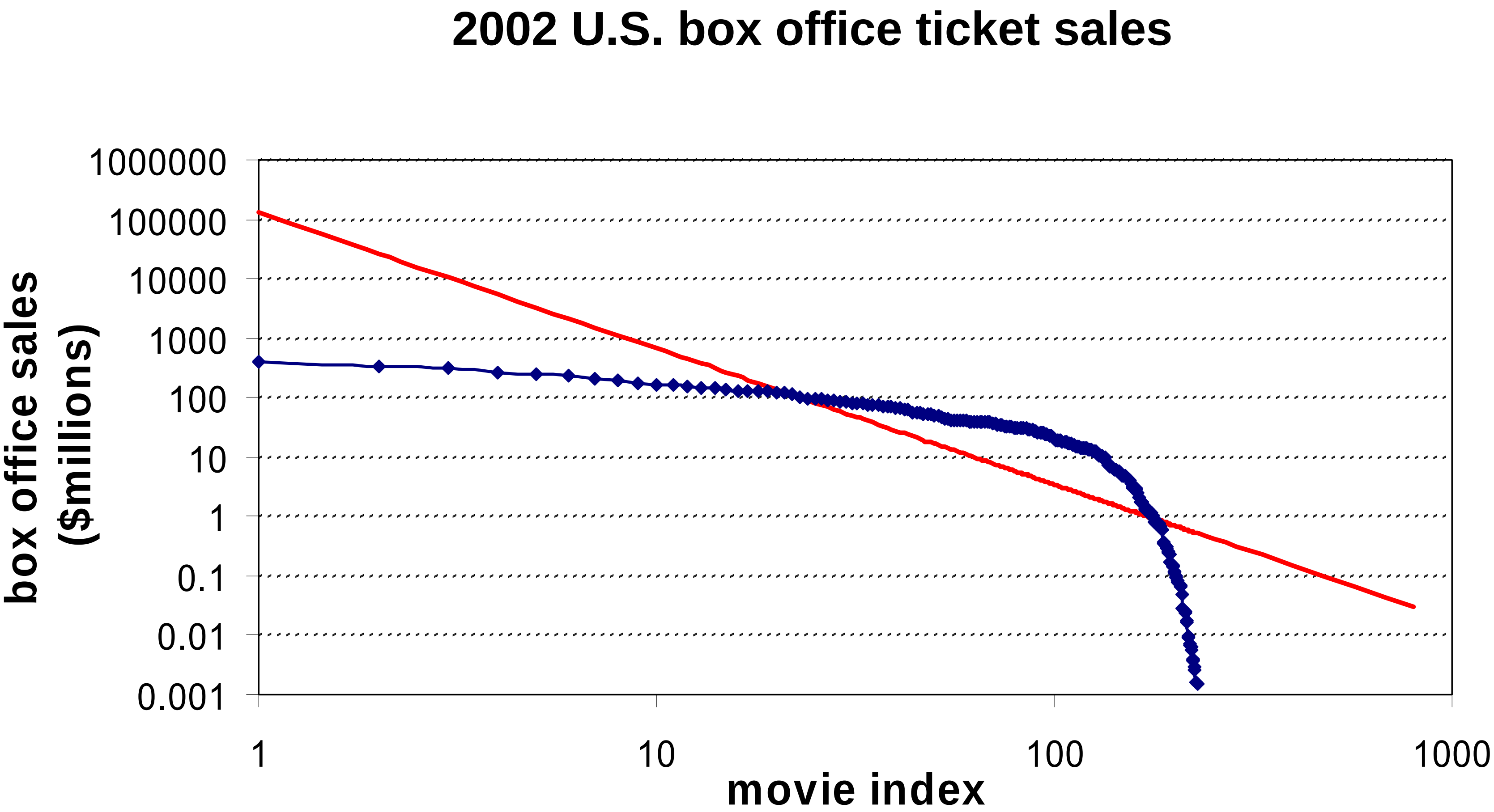
- In contrast, Kazaa (multi-media) objects **never change**
  - users rarely re-download the same object
    - 94% of the time, a user fetches an object *at most once*
    - 99% of the time, a user fetches an object *at most twice*
  - Implication:
    - popularity of most popular object is bounded by user population size

5. Kazaa does *not* obey Zipf's law



- Zipf: popularity(n^{th} most popular object) ~ $1/n^{\alpha}$
- max # requests bounded by population size
- the *most* popular objects are *much* less popular than zipf would predict

Movie sales data (U.S.)



6. Objects have quick turnover

- Popularity is short lived
  - only 5% of the top-100 audio objects stayed in the top-100 over our entire trace [video: 44%]

- Newly popular objects tend to be recently born
  - of audio objects that “broke into” the top-100, 79% were born a month before becoming popular [video: 84%]

Outline

- Characterize content delivery in today's Internet
- Understand what drives the Kazaa P2P file-sharing system
  1. Some observations about Kazaa
  2. A model for studying P2P multimedia systems
  3. Locality-aware request distribution

Modeling P2P Workloads

- Question: can we use a model to gain more insight into the forces driving P2P workloads?

- Objective: study three key issues identified by traces
  1. "fetch-at-most-once" behavior
  2. object birth rate
  3. client birth rate

Model basics

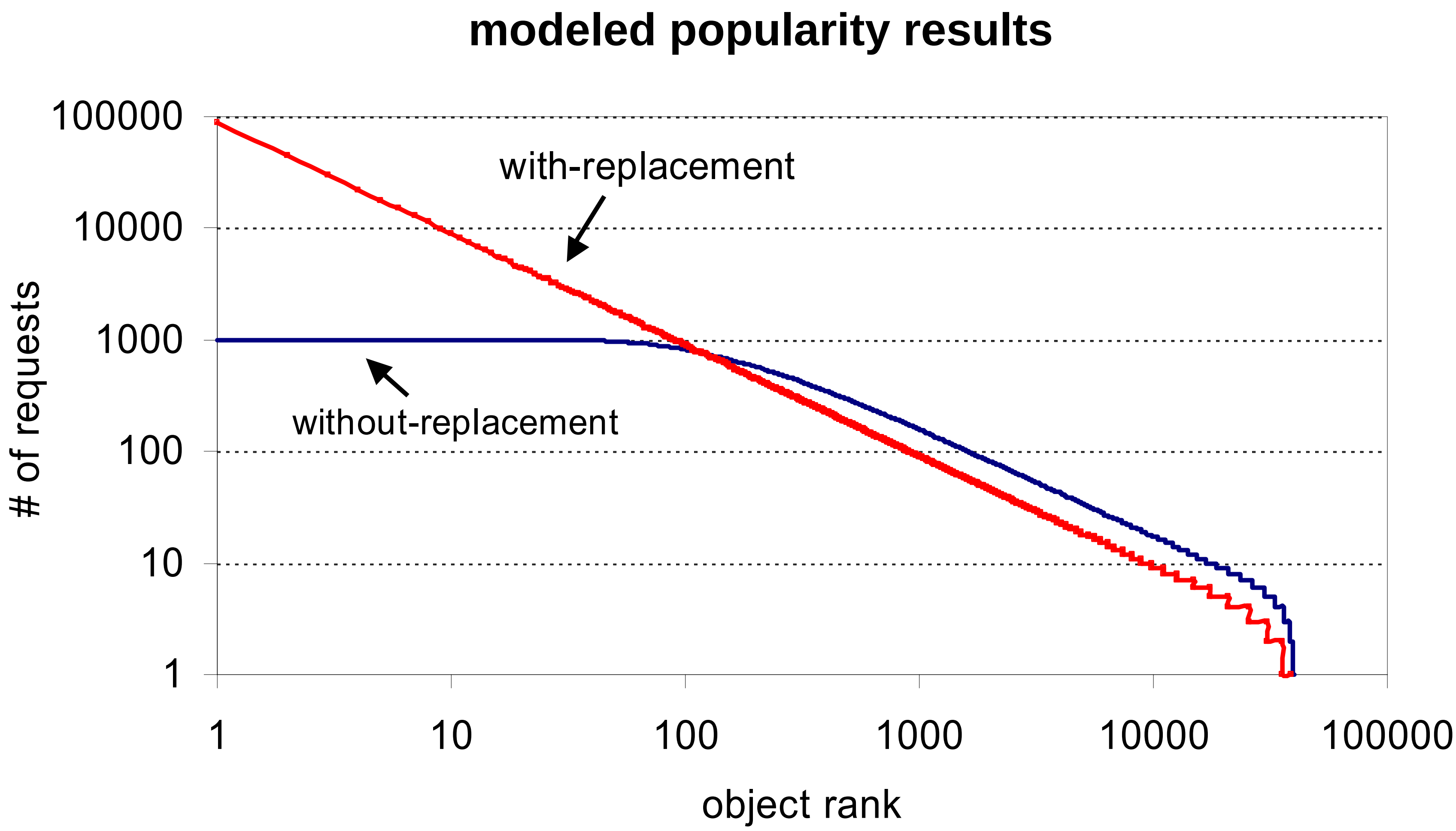
- Objects are chosen from an underlying Zipf curve, but assuming *“fetch-only-once”* behavior per client.
- Over time, users “coast” down the Zipf curve towards less popular objects

- New objects are inserted from the original Zipf (objects below pushed down)
- New users begin with a fresh Zipf curve

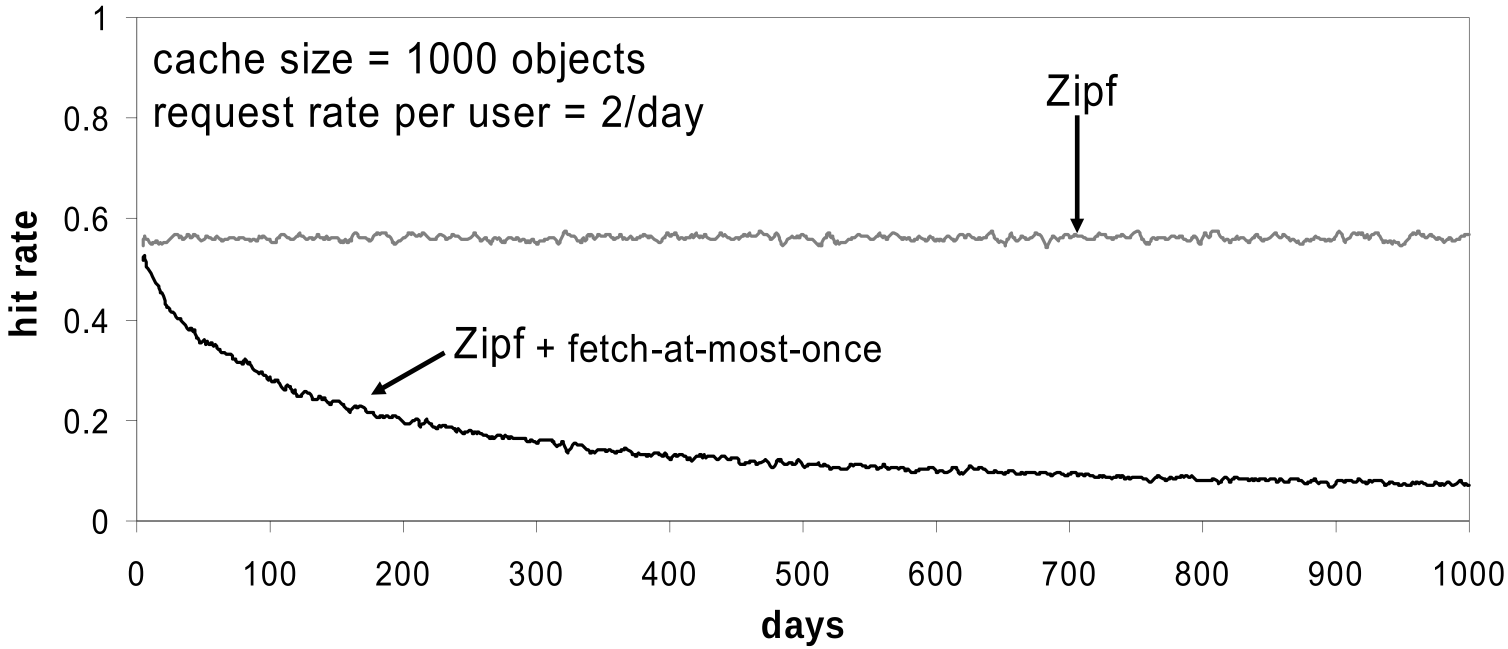
Model parameters

C	# of clients	1,000
O	# of objects	40,000
λ_R	client req. rate	2 objects/day
α	Zipf param driving obj. popularity	1.0
P(x)	prob. of client req. object of pop rank x	Zipf (1.0) + fetch-at-most-once
A(x)	prob. of new object inserted at pop rank x	Zipf (1.0)
M	cache size (frac. of obj)	varies
λ_O	object arrival rate	varies
λ_c	client arrival rate	varies

Fetching without replacement flattens Zipf curve head



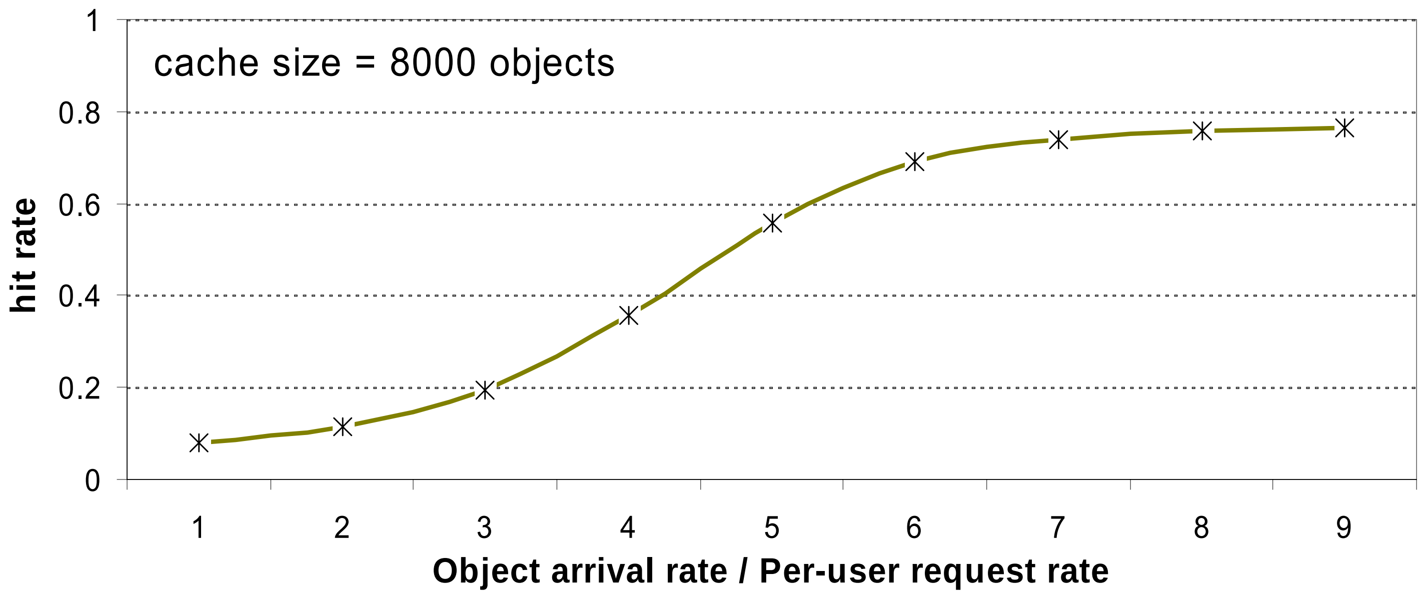
Caching implications



Given a fixed object and client population:

- with replacement → hit rate stays flat over time
- without replacement → hit rate degrades over time

New objects help, not hurt

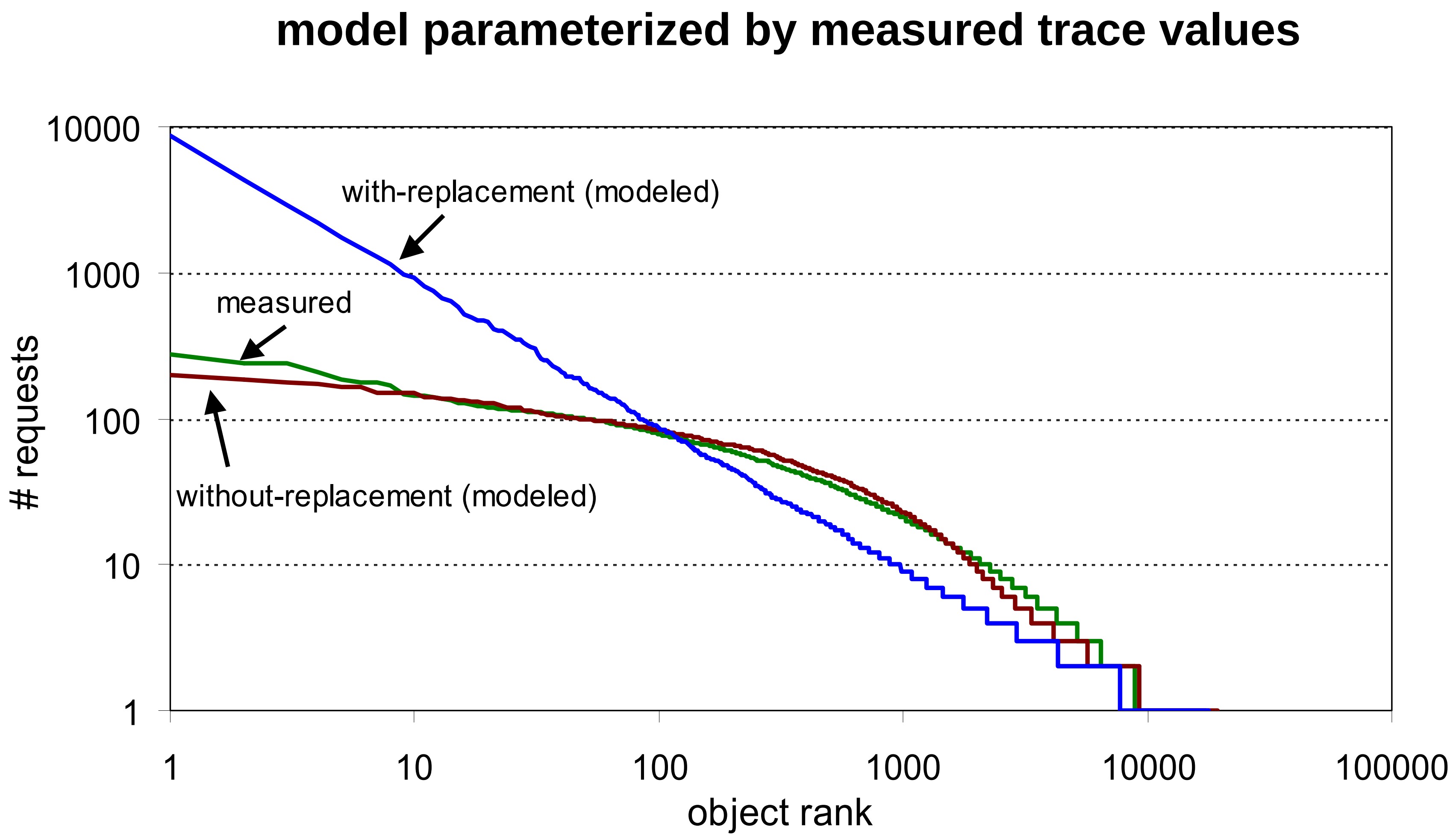


- New objects do cause cold misses
  - but they replenish the highly cacheable part of the Zipf curve
- A slow, constant arrival rate stabilizes performance
  - rate needed is proportional to avg. per-user request rate

New clients don't help

- they have some potential…
  - they have a "fresh" Zipf curve to draw from
  - therefore will have higher hit rate
- but new clients grow old too
  - ultimately they increase the size of the old population
  - to offset, must add clients at exponentially increasing rate
    - » not sustainable in long run…

Model validation



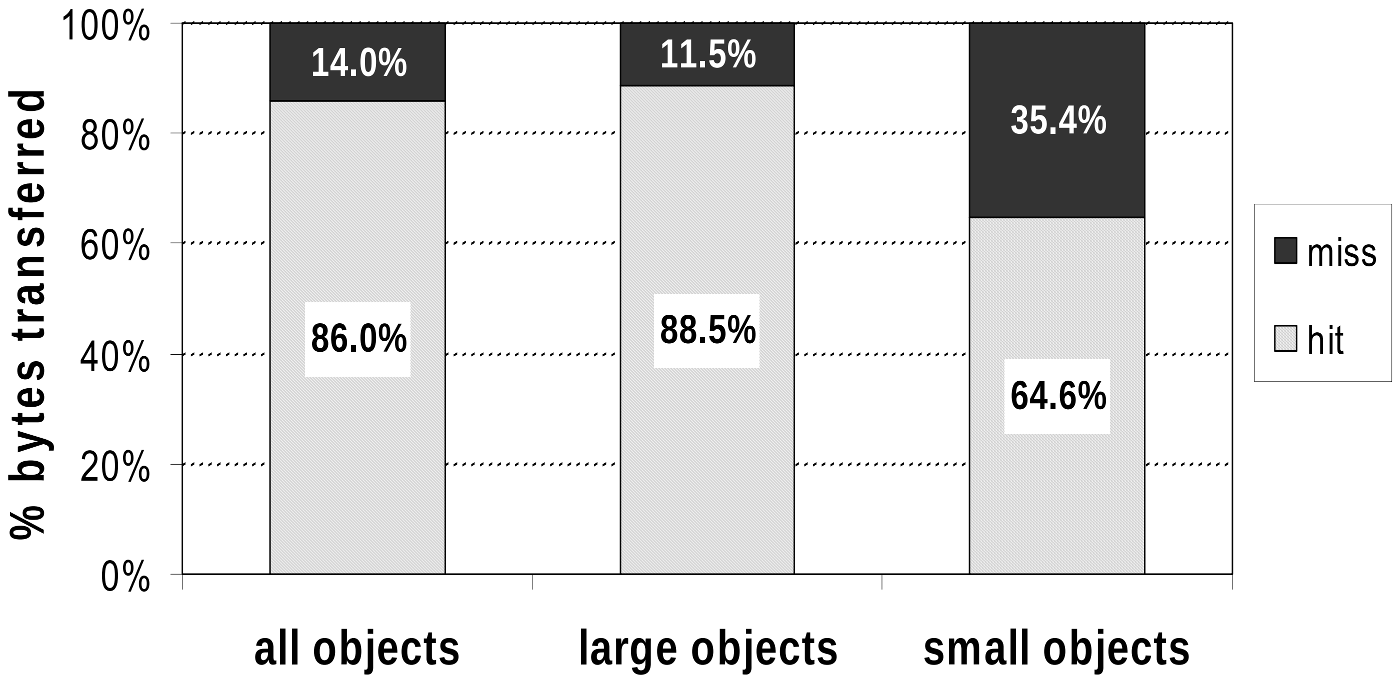
Outline

- Characterize content delivery in today's Internet

- Understand what drives the Kazaa P2P file-sharing system
  1. Some observations about Kazaa
  2. A model for studying P2P multimedia systems
  3. Locality-aware request distribution

Kazaa has untapped locality

simulated proxy cache hit rate for UW P2P environment



implication → 86% of bytes ***already exist*** within UW when they're downloaded externally!

Locality-aware request routing

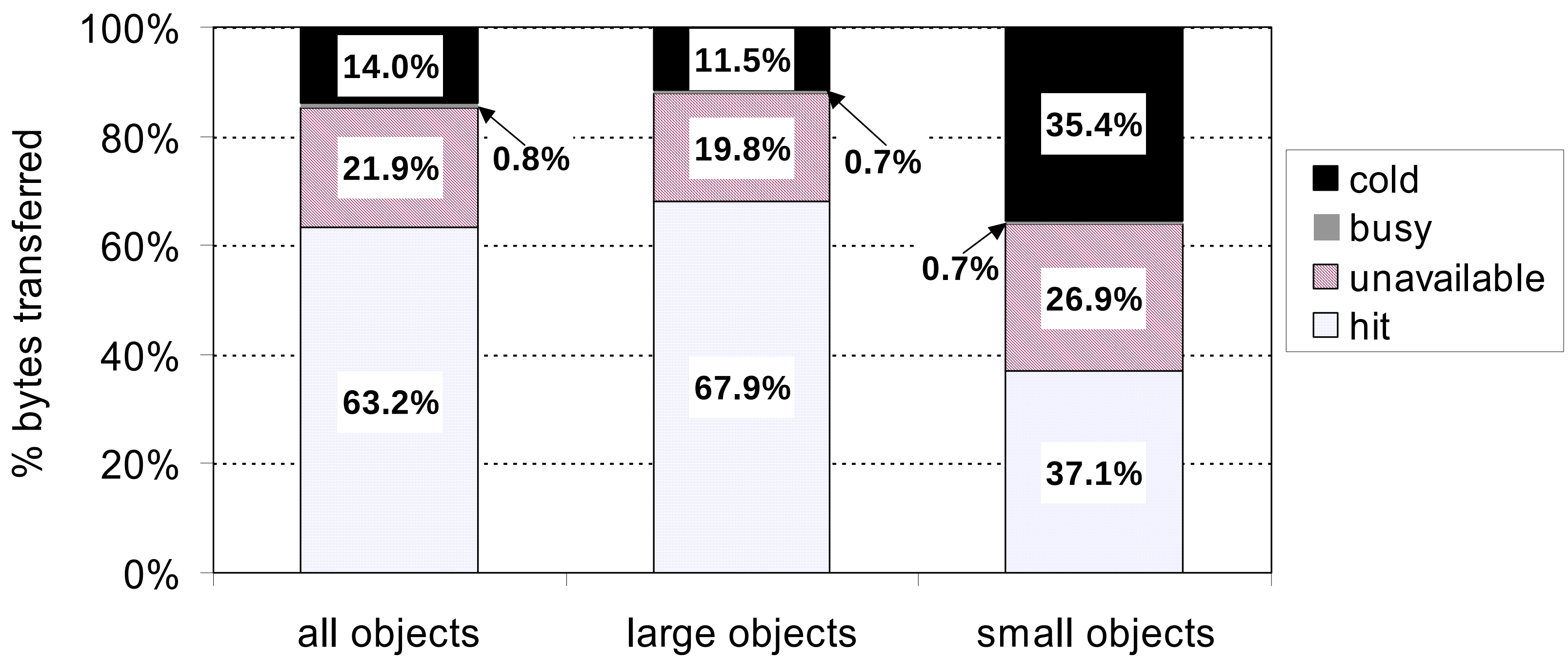
- Alternative: make better use of local peers

- Scheme 1: use a *redirector* instead of a cache
  - Redirector sits at university border and watches traffic
  - It indexes content and *redirects* requests to local peers that can serve it

- Scheme 2: decentralized request distribution
  - Use location information in the P2P protocols

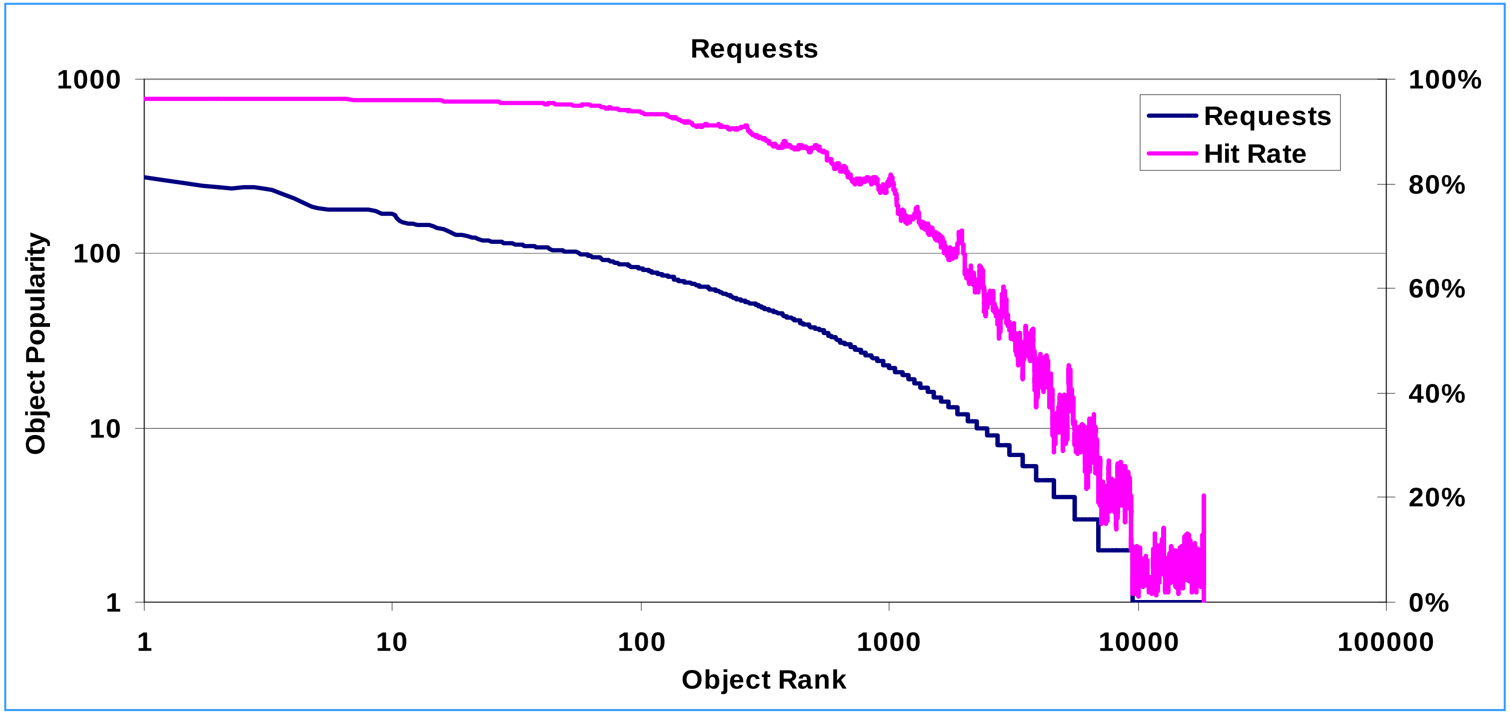
- We simulated locality-aware routing using our trace data
  - (note that both schemes are identical w.r.t. the simulation)

Basic locality-aware result



- **Locality-aware schemes works!**
  - bars are a *lower bound* (assume worst availability from trace)
  - savings is huge (9.6 TB for large objects in our trace)
  - but there is still opportunity for improvement
  - need to figure out what causes unavailability misses

How can we eliminate availability misses?



- a kind of "natural replication" is driven by popularity
  - this is descriptive but also predictive
  - focus on "middle" popularity objects when designing systems

Summary of Part 2

- Kazaa/P2P is driven by completely different forces than the Web
  - multimedia content dominates the workload
  - system usage is “batch,” not interactive
  - object immutability leads to *fetch-at-most-once* behavior
  - workload is driven by object and client births
  - multimedia is not “Zipf”
- Current file sharing architectures miss opportunity
  - Kazaa makes poor or no use of locality
  - locality-aware architectures can save significant amounts of external bandwidth
- We have a model that captures behavior of this workload

Final note

- The internet continues to reinvent itself every few years and can change dramatically at a moment's notice.

- Multi-media and peer-to-peer file sharing are going to be a huge component of the internet in the future -- *independent* of current legal issues with P2P.

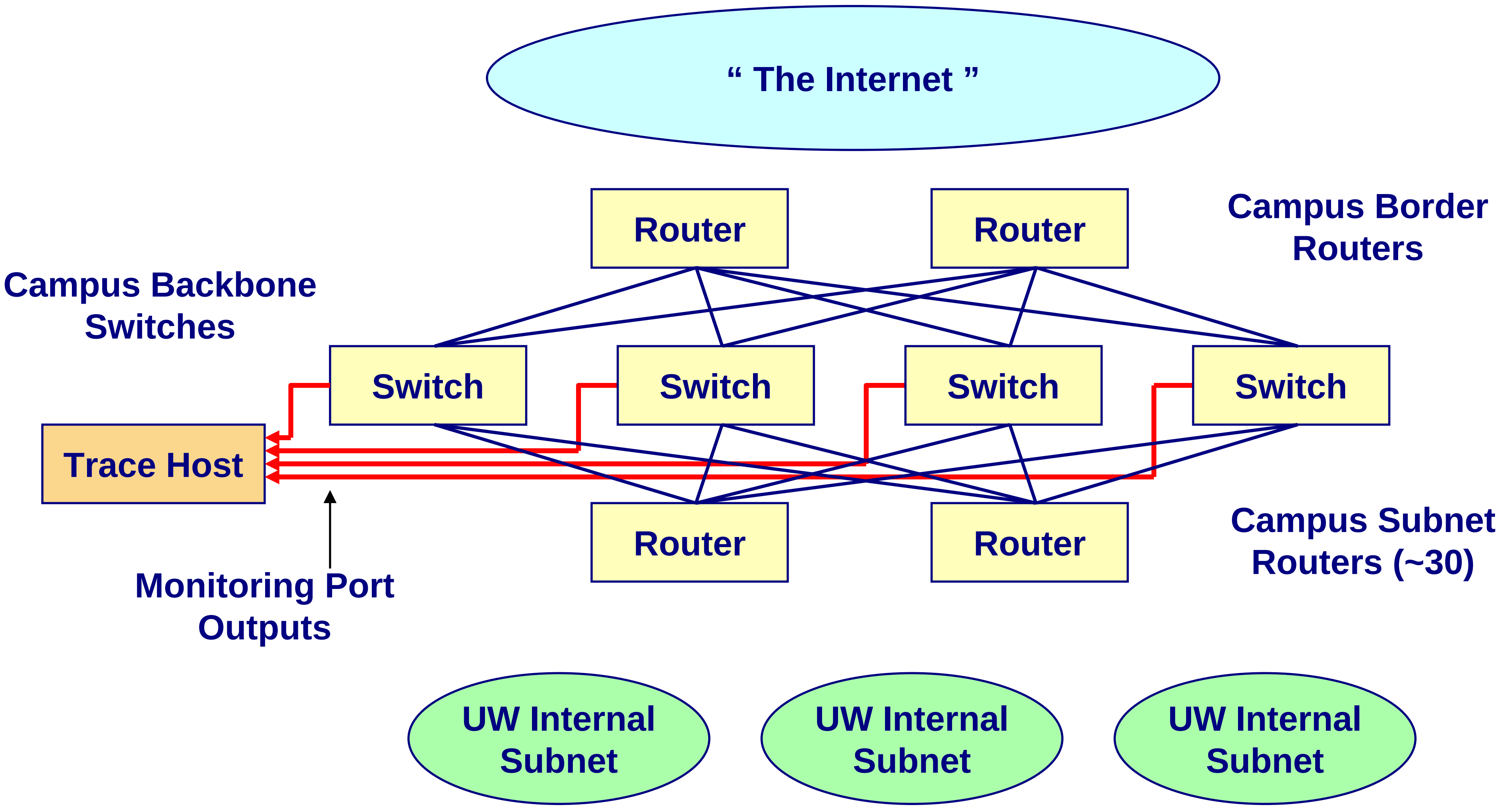
Pubs related to this talk

- An analysis of internet content delivery systems. *Proc. 5th Conf. on Operating Systems Design and Implementation* (OSDI), December 2002.

- Measurement, modelling, and analysis of a peer-to-peer file-sharing workload. *Proc. 19th ACM Symp. on Operating Systems Principles* (SOSP), October 2003.

http://www.cs.washington.edu/research/networking/websys

UW Trace Environment

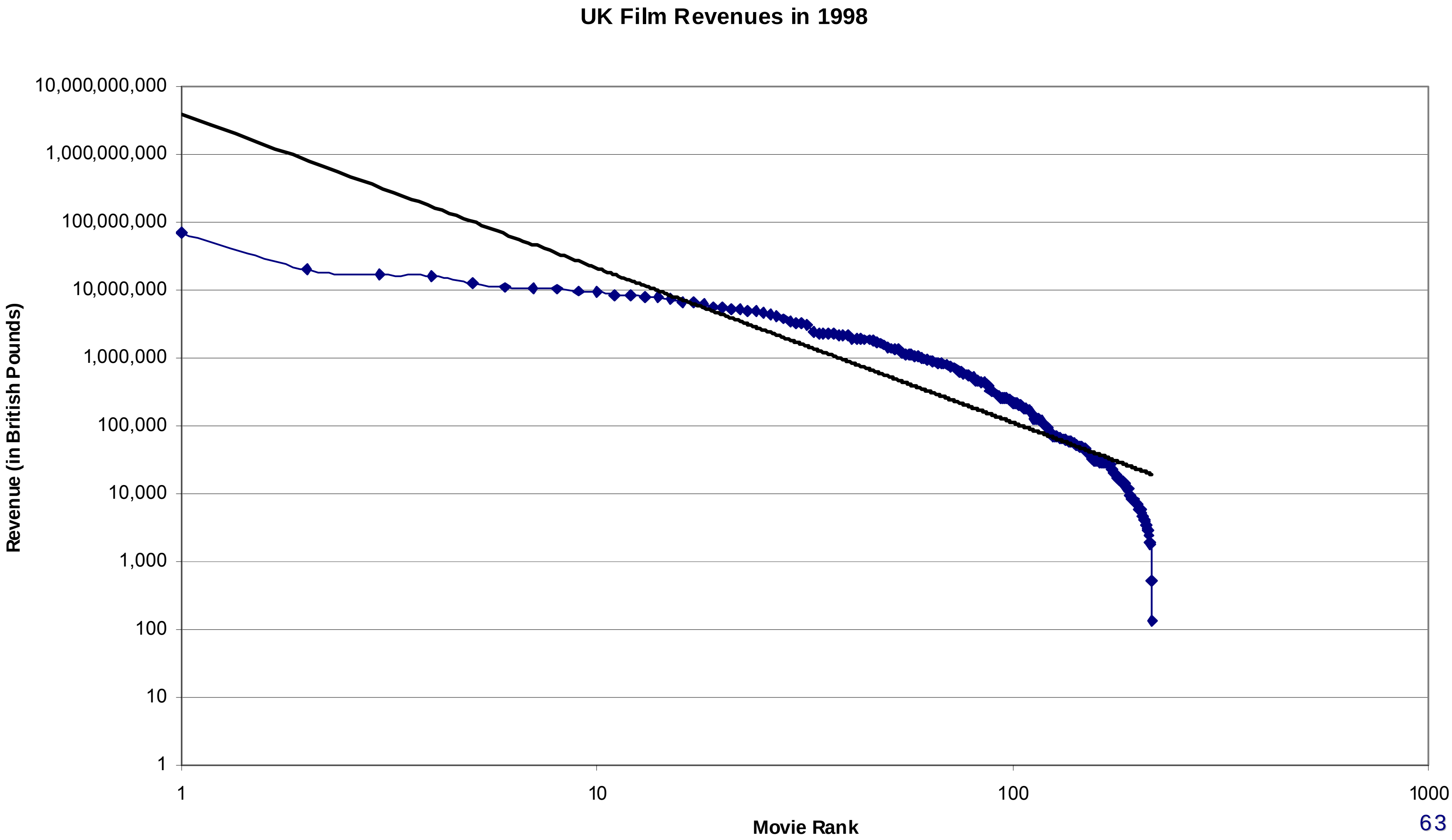


Kazaa workload has quick turnover

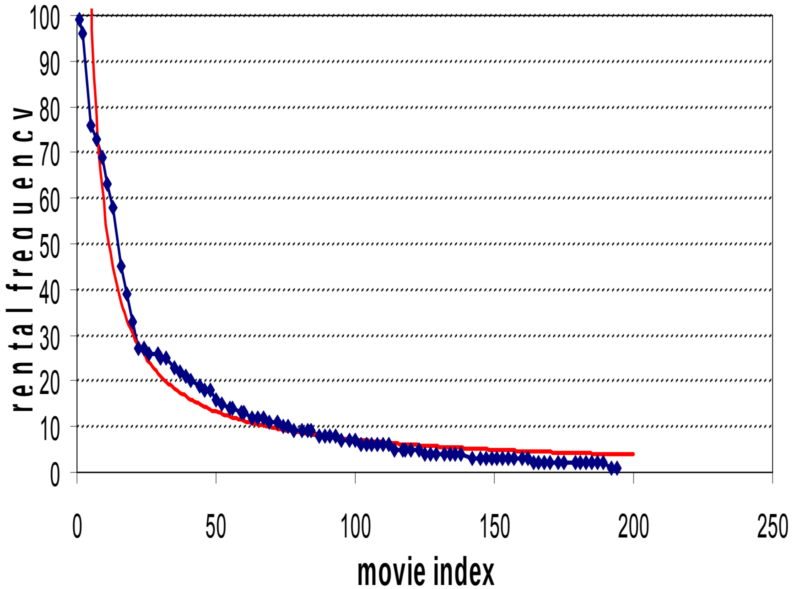
- Popularity is short lived
  - only 5% of the top-100 audio objects stayed in the top-100 over our entire trace [video: 44%]

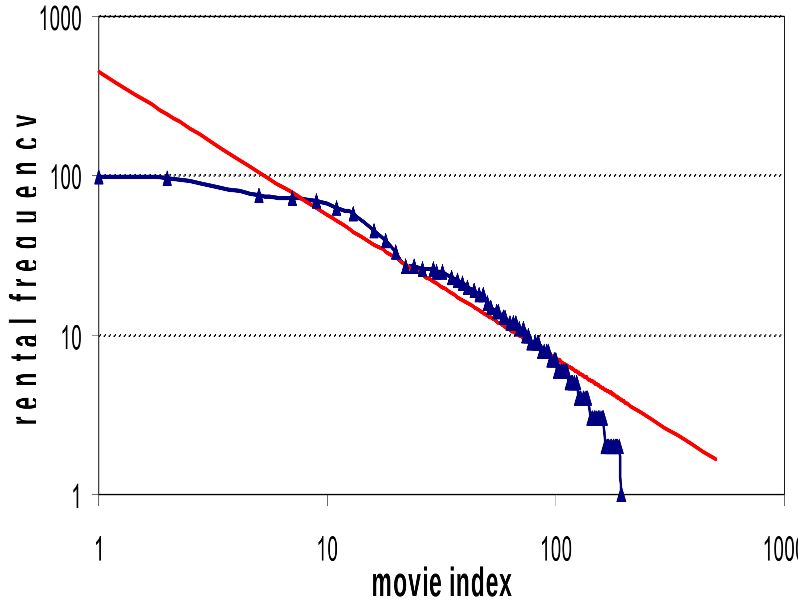
- Newly popular objects tend to be recently born
  - of audio objects that “broke into” the top-100, 79% were born a month before becoming popular [video: 84%]

Movie sales data (UK)



Previously reported Zipf results





- Left side: graph from Dan, Sitaram, and Shahabuddin, "Scheduling policies for an on-demand video server with batching," *Proc. of ACM Multimedia 1994*, Oct. 1994, "demonstrating" Zipf fit of 1992 video rental data.
- Right side: same data plotted on a log/log scale.