

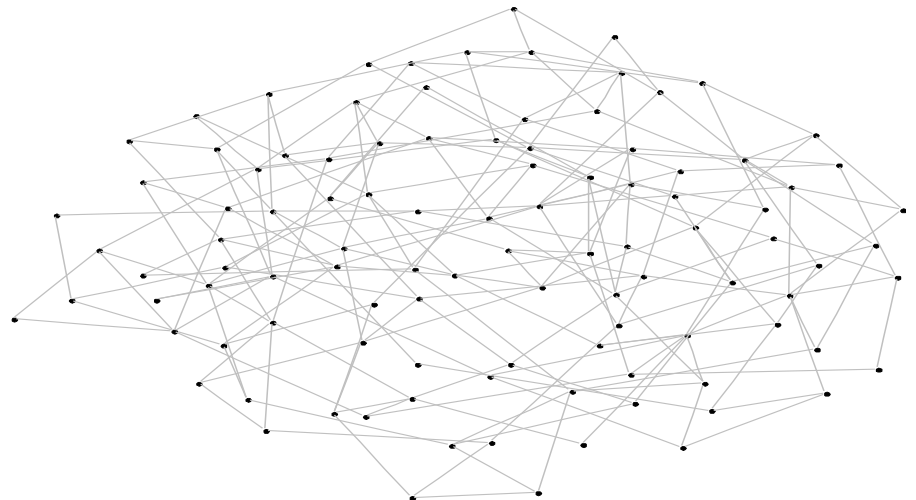
*“Peer-to-Peer Systems: Introduction
and Challenges”*

*“Συστήματα Ομοτίμων: Εισαγωγή και
Προκλήσεις “*

Δημήτρης Ζεϊναλιπούρ

Λέκτορας, Ανοικτό Πανεπιστήμιο Κύπρου

*ΕΠΛ 601: Κατανεμημένα
Συστήματα, Τμήμα
Πληροφορικής,
Πανεπιστημίου Κύπρου, 16
Νοεμβρίου 2007.*



<http://www.cs.ucy.ac.cy/~dzeina/>

Στόχοι Διάλεξης

- Μια γενική εισαγωγή στα Συστήματα Ομοτίμων **(Ελληνικές Διαφάνειες)**.

Eng Keong Lua et al. "A Survey and Comparison of Peer-to-Peer Overlay Network Schemes," IEEE Communications Surveys and Tutorials, Vol 7, No 2 (Second Quarter, 2005), pp. 72-93.

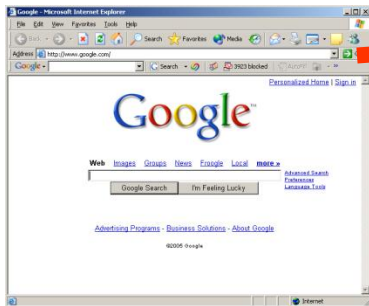
- Αναφορά στο ερευνητικό πρόβλημα: Ανάκτηση Δεδομένων σε Συστήματα Ομοτίμων "Information Retrieval in Peer-to-Peer Systems" **(Αγγλικές Διαφάνειες)**

D. Zeinalipour-Yazti, V. Kalogeraki and D. Gunopulos, "Information Retrieval Techniques for Peer-to-Peer Networks", IEEE CiSE Magazine, Special Issue on Web Engineering, IEEE Publications, pp.12-20., July/August 2004

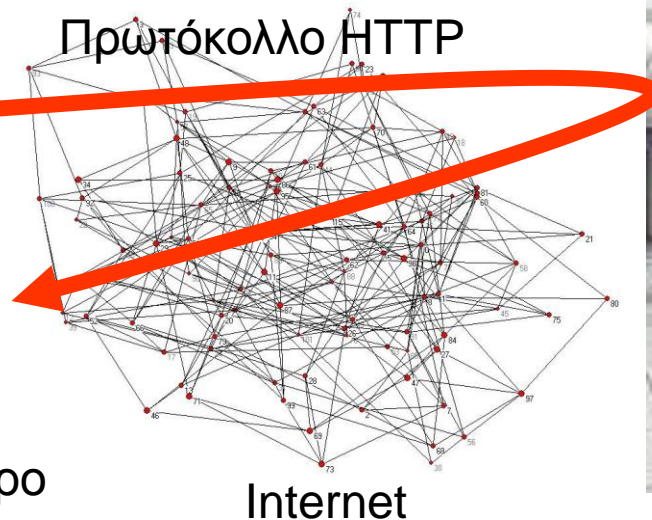
Μέρος 1: Εισαγωγή στα Συστήματα Ομοτίμων

Από συστήματα Πελάτη-Εξυπηρετητή σε Συστήματα Ομοτίμων

- Οι περισσότερες υπηρεσίες στο Διαδίκτυο ακολουθούν το μοντέλο Πελάτη-Εξυπηρετητή (π.χ. HTTP, FTP, DNS, POP3,...)



Κάποιος πελάτης
(χρήστης) στην Κύπρο



Google.com (CA, USA)

Λόγοι Ανάπτυξης του Peer-to-Peer?

(στην μορφή που το ορίζουμε σήμερα *)

- Ο **μεγάλος αριθμός PCs** με πολύ ψηλή υπολογιστική ισχύ διαθέσιμα στα άκρα (edges) του Internet.
- Επίσης, υπάρχει πολύ **ψηλότερο bandwidth** διαθέσιμο στα άκρα του Internet (ADSL, Satellite, Cable, Ethernet LANs, etc.)
- Επομένως, γίνεται εφικτό να αξιοποιήσουμε το ***Storage, Cycles, Content*** και να έχουμε αλληλεπίδραση Η/Υ στα άκρα (edges) του Internet, χωρίς την χρήση κεντρικών servers .
- * Σημειώστε ότι το άλλα Internet Services από το 1980- π.χ. DNS, BGP, Usenet etc, ορίζουν και αυτά ένα P2P μοντέλο ανταλλαγής πληροφοριών

Χαρακτηριστικά συστημάτων Πελάτη-Εξυπηρετητή (Π/Ε)

Πλεονεκτήματα

- Κεντριοποιημένη διαχείριση δεδομένων
- Ασφάλεια δεδομένων
- **Αυτά τα χαρακτηριστικά είναι ιδανικά για πολλές Διαδικτυακές Εφαρμογές (Συστήματα Τραπεζών, Ηλεκτρονικού Ταχυδρομείου, κτλ.)**

Χαρακτηριστικά συστημάτων Πελάτη-Εξυπηρετητή (Π/Ε)

Μειονεκτήματα

- Χρειάζονται ακριβές υπολογιστικές υποδομές (π.χ., το Youtube.com αναλώνει 25TB – 250TB ανά μήνα και πληρώνει \$5M/μήνα ή \$170K/μέρα για το bandwidth)
- Σε περίπτωση βλάβης του εξυπηρετητή χάνουμε την υπηρεσία (**single point of failure**).
- Χρειάζονται συνεχή διαχείριση (**administration**).
- Μπορούν να λογοκριθούν (**copyright**) και να ελεγχθούν (από καθεστώτα, κτλ) π.χ., τα αποτελέσματα αναζήτησης μηχανών αναζήτησης ελέγχονται από το κράτος στην Κίνα.

Τα P2P Systems προσπαθούν να ξεπεράσουν αυτά τα προβλήματα.

Τι είναι Peer-to-Peer (P2P);

- **Φοιτητής:** “Online Gaming” και “Multimedia downloads”
- **Υπάλληλος Εταιρείας:** “Instant Messaging”
- **Warner Bros:** Η φθηνή (για την εταιρεία) διανομή ταινιών σε πελάτες. **Sony:** Φθηνό Online Gaming.
- **Άλλοι:** Αποδοτικό & Φτηνό VOIP (Skype)
- **Διάφοροι Ορισμοί φαίνεται να συμφωνούν στα πιο κάτω:**
 - Ο Διαμοιρασμός Πόρων (**Resource Sharing** - CPU cycles, DISK, Bandwidth, etc.)
 - Απευθείας Επικοινωνία (**Direct Communication**) μεταξύ Ίσων (peers)
 - Αποκέντρωση (**decentralization**) & αυτόματη οργάνωση (**self organization**).

Τι είναι Peer-to-Peer (P2P);

- Οι χρήστες του συστήματος έχουν διπλό ρόλο: δηλ. είναι **Πελάτες** και **Εξυπηρετητές** την ίδια στιγμή.
- Καθώς ο χρήστης Α εξυπηρετείται από τον χρήστη Β, κάποιος χρήστης Γ μπορεί να εξυπηρετείται από τον Α.
- Δεν υπάρχει κεντροποιημένη διαχείριση...το οποίο δημιουργεί ένα αίσθημα ελευθερίας.
- ...Παράλληλα όμως δημιουργεί και πολλά προβλήματα (όπως αυτό της παράνομης ανταλλαγής τραγουδιών)

Πλεονεκτήματα Συστημάτων P2P

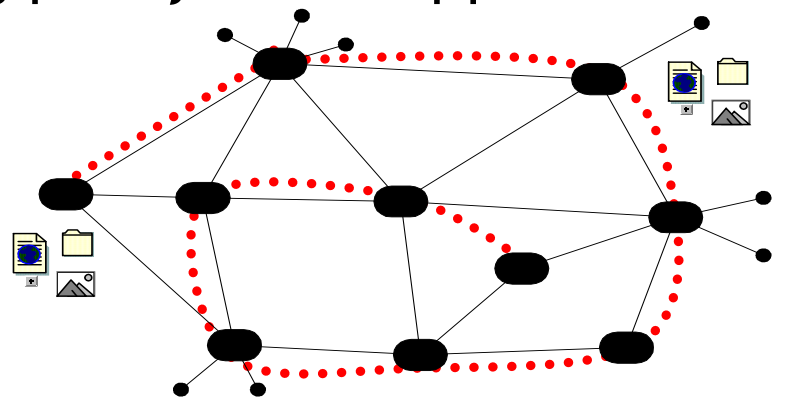
- **Edge-Computing**
 - Αξιοποίηση αχρησιμοποίητου **bandwidth, storage, processing power** στα **άκρα (edge)** του Internet
- **Scalability**
 - Δεν υπάρχει συμφόρηση (bottleneck) σε κάποιο κεντροποιημένο κόμβο. Επομένως τα συστήματα αυτά μπορούν να μεγαλώνουν “απεριόριστα”.
 - Oct 12th Gnutella (Limewire.com): 2,219,539 κόμβοι
- **Reliability (Αξιοπιστία)**
 - No single point of failure, Γεωγραφική Κατανομή Περιεχομένου (CDNs)
- **Ease of administration**
 - Οι κόμβοι οργανώνονται μεταξύ τους αυτόματα (**self-organization**).
 - Αυτόματα επίσης γίνεται το **replication** και το **load balancing** καθώς τέτοια συστήματα παρέχουν **fault tolerance**.
- **Anonymity – Privacy**
 - ...κάτι το οποίο δεν είναι εύκολο σε ένα κεντροποιημένο σύστημα

Εφαρμογές Συστημάτων P2P

- Ανταλλαγή Αρχείων (Napster, Gnutella, Bittorent, ...)
- Διαδικτυακή Τηλεφωνία (Skype)
- Διαδικτυακά Παιχνίδια (Playstation Online Gaming)
- Πάταξη του Spam (SpamNet)
- Instant Messaging (IRC, MSN & Yahoo Msgers)
- Content Distribution Networks (CorelCDN)
- P2P Web Caching (Squirrel)
- Application-Level Multicast (Narada)

P2P Συστήματα και Overlay Networks

- **P2P Συστήματα** είναι δομημένα πάνω από **Overlay Networks** (χρησιμοποιούνται και οι δυο όροι).
- Δηλαδή οι peers, εγκαθιδρύουν TCP ή UDP socket connections με άλλους κόμβους. Αυτό δημιουργεί ένα νοητό “virtual” γράφο διασύνδεσης.
- Ο γράφος αυτός **δεν** αντιπροσωπεύει τις πραγματικές συνδέσεις μεταξύ των κόμβων.



— The virtual P2P topology
..... The physical topology

Κατάταξη P2P Συστημάτων

(Βάση του Βαθμού Αποκέντρωσης)

A) Centralized

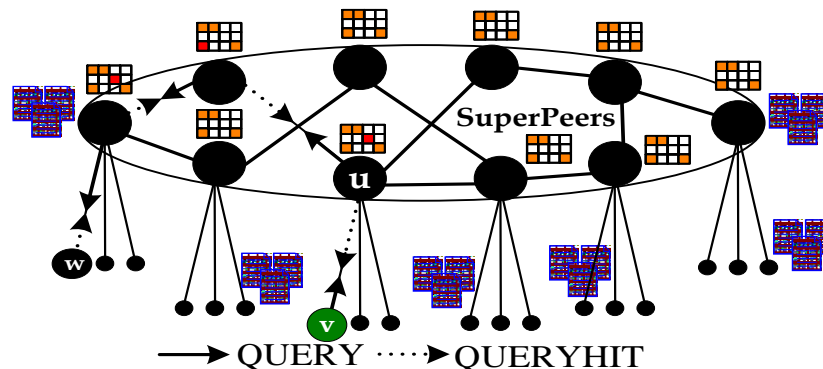
- Κεντρικά Ευρετήρια Αναζήτησης π.χ., Napster, Bittorent

B) Purely Distributed

- Δεν υπάρχουν Ευρετήρια Αναζήτησης π.χ., Gnutella

C) Hybrid

- Κάποιοι επιλεγμένοι peers (η επιλογή γίνεται βάση του διαθέσιμου bandwidth, της ώρα σύνδεσης, κτλ) έχουν μερικά ευρετήρια για τα περιεχόμενα άλλων κόμβων, π.χ., FastTrack (KaZaA), Limewire's Ultrapeers (Superpeers), Skype

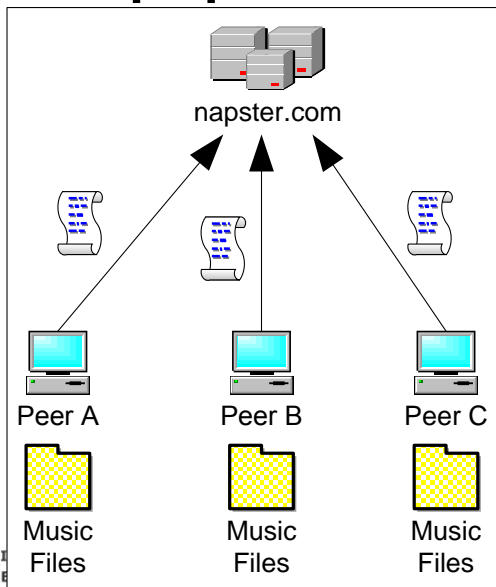


Centralized P2P Συστήματα

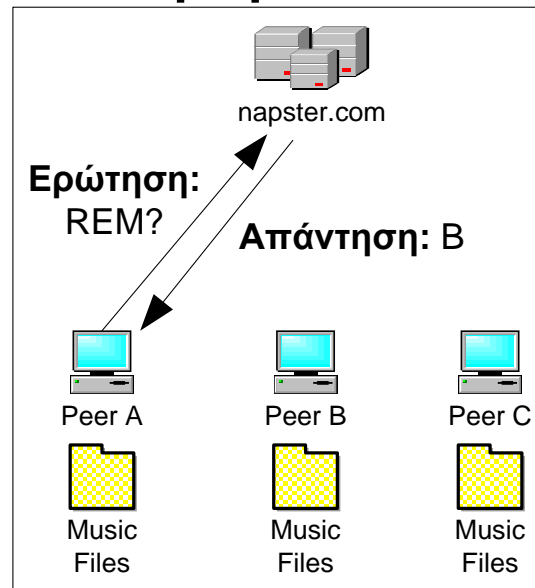
Napster: File Sharing

- Δημιουργείται το 1999 από ένα 18-χρονο φοιτητή.
- Ο κάθε χρήστης εκτελεί τις ακόλουθες τρεις διαδικασίες

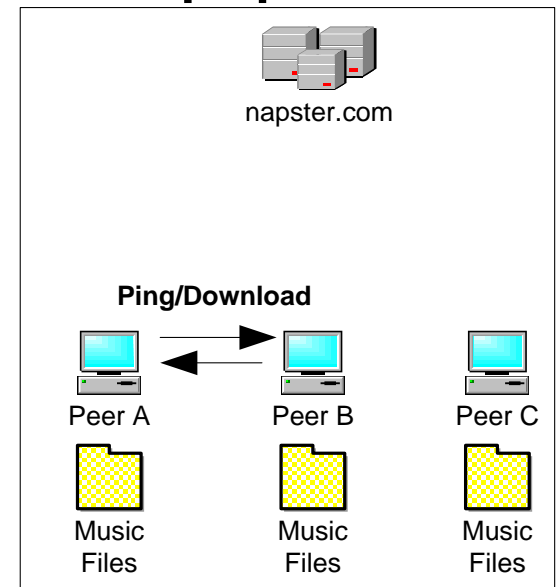
Αποστολή Λίστας Τραγουδιών



Αναζήτηση Τραγουδιών



Ανάκτηση Τραγουδιών



Παράθυρο Αναζήτησης Napster

Napster v2.0 BETA 7

File Actions Help

Home Chat Library Search Hot List Transfer Discover Help

Artist: Find it!

Title: Clear Fields

Max Results: Advanced >>

Filename	Filesize	Bitrate	Freq	Length	User	Connection	Ping
incomplete_other_artist\Tito Puentes Golden Latin Jazz Allstars - Oye Como ...	3,696,640	128	44100	3:51	bdenzler	DSL	343
incomplete_other_artist\[Marty Robbins] The Fastest Gun Around.mp3	542,304	128	44100	0:39	bdenzler	DSL	343
incomplete_other_artist\Ravi Shankar - Chants Of India 04 - Asato Maa.mp3	2,449,408	128	44100	2:35	bdenzler	DSL	343
other_artist\Engelbert Humperdinck - White Christmas.mp3	9,277,648	320	44100	3:52	bdenzler	DSL	343
other_artist\Grateful Dead - Franklin's Tower - Reggae Style.mp3	4,635,458	128	44100	4:48	bdenzler	DSL	343
Unknown Artist - You seriously have to listen to this.mp3	462,848	318	16000	0:17	sam113...	Cable	383
MP3z\artist - 'The Way Life Is' By Drag-On featuring Case.mp3	4,726,784	128	44100	4:54	burg651	Cable	386
MP3z\artist - 'Opposite Of H2O' By Drag-On featuring Jadakiss.mp3	3,540,992	128	44100	3:41	burg651	Cable	386
Various Artist - Perfect Day 97.mp3	3,722,344	128	44100	3:53	falkstad	ISDN-128K	398
Liszt\Liszt - Etude 'Un sospiro' - Cziffra-artist.mp3	2,752,512	128	44100	2:53	lskjdfklj...	Unknown	504
Music\Waiting To Exhale - Original Soundtrack Album - Various Artist - Count...	3,199,083	96	44100	4:26	Jzfork9	56K	511
Track 03_artist.mp3	4,054,332	128	44100	4:13	immusic...	Cable	514
Track 02_artist.mp3	6,228,974	128	44100	6:26	immusic...	Cable	514
Track 01_artist.mp3	4,731,426	128	44100	4:54	immusic...	Cable	514
Track 04_artist.mp3	4,514,505	128	44100	4:41	immusic...	Cable	514
Track 05_artist.mp3	4,105,323	128	44100	4:16	immusic...	Cable	514
mixer in track 01_Artist_0721011750.mp3	180,686	128	44100	0:17	immusic...	Cable	514
Album\Reflex - Keep In Touch-Artist.mp3	7,041,024	160	44100	5:49	rotimca	56K	527

Returned 100 results.

Get Selected Songs Add Selected User to Hot List

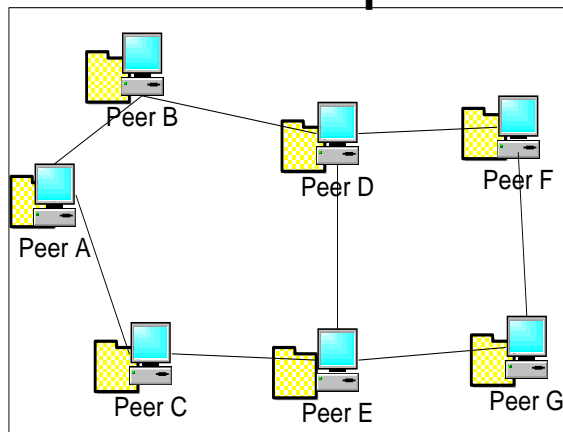
Online (keyscreen): Sharing 491 files. Currently 740,043 files (2,991 gigabytes) available in 5,873 libraries.

Purely Distributed P2P Systems

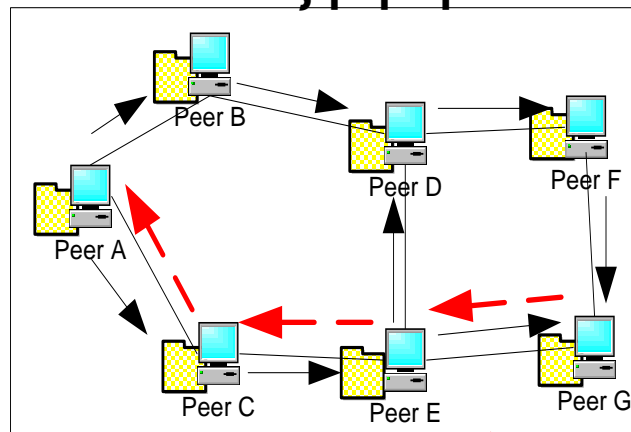
Gnutella: File Sharing

- Το (αρχικό) Napster κλείνει το 2001 μετά από απόφαση του δικαστηρίου για παράνομη ανταλλαγή αρχείων.
- Στις αρχές του 2000 δημιουργείται από τους κατασκευαστές του Winamp, ένα νέο σύστημα στο οποίο δεν υπάρχει κεντροποιημένη διαχείριση της λίστας των αρχείων.
- Με αυτό τον τρόπο δημιουργείτε ένα εντελώς κατακεκομημένο σύστημα (που θεωρητικά δεν μπορεί να κλείσει κανείς...)

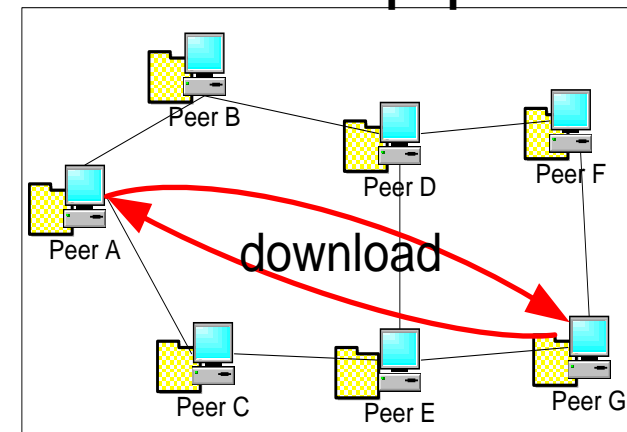
Σύνδεση



Αναζήτηση



Ανάκτηση



Παράθυρο Αναζήτησης Gnutella (LimeWire)

The screenshot displays the LimeWire application window titled "LimeWire: Enabling Open Information Sharing". The interface includes a menu bar (File, View, Navigation, Resources, Tools, Help) and a toolbar with buttons for Search, Monitor, Connection, and Library. The "Monitor" button is circled in red. Below the toolbar, the "Library" section shows a list of files with columns for Name, Size, Type, Path, Uploads, Hits, and Locations. The "MP3 PlayList" section at the bottom shows a list of files with columns for Name, Length, and Bitrate. The status bar at the bottom indicates "Quality: TurboCharged" and "Sharing 6 files".

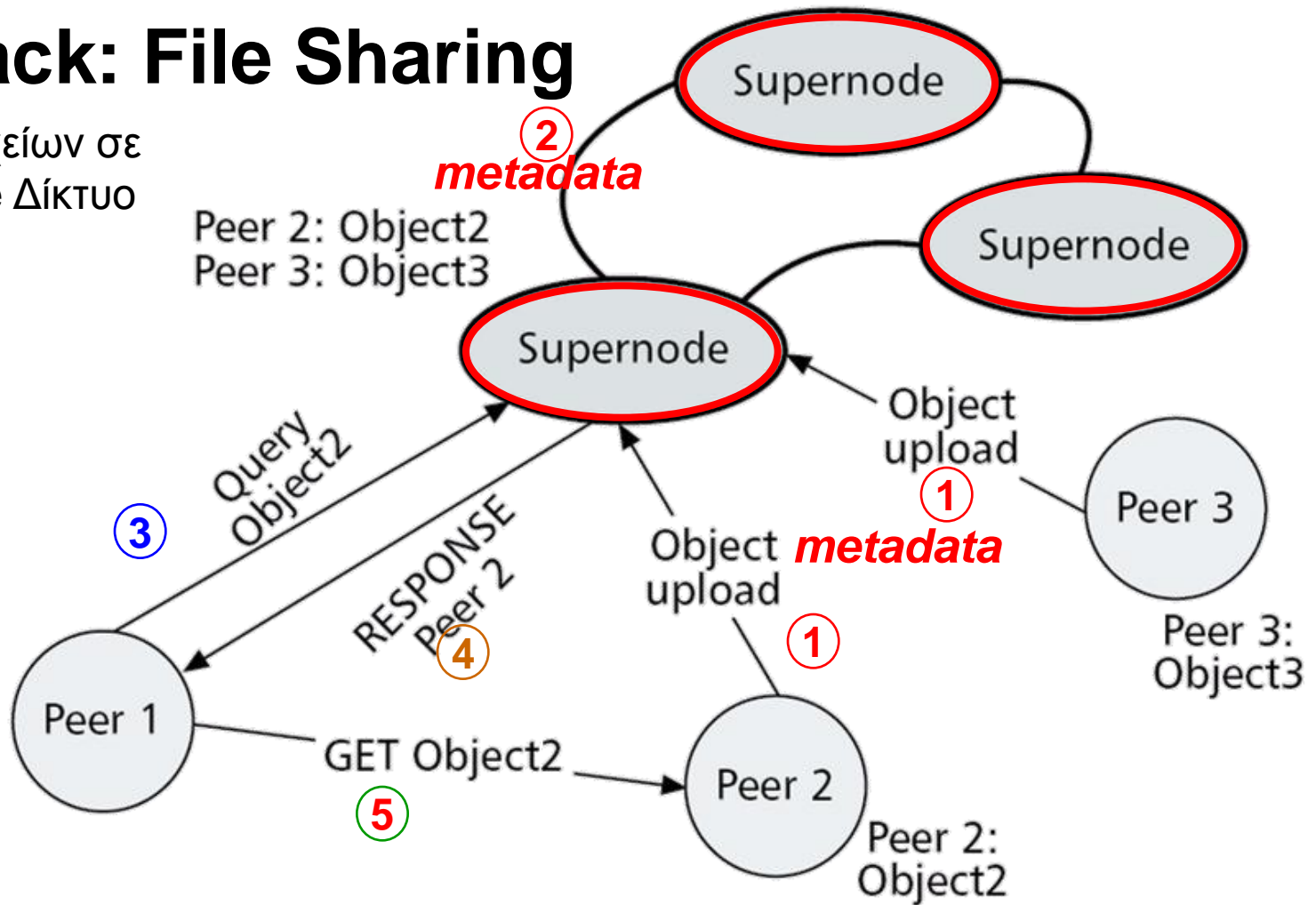
Name	Size	Type	Path	Uploads	Hits	Locations
Alfie_Zappacosta_Start_Again	4,358 KB	mp3	C:\Program Files\Lime... 0 / 0	0	0	0
American_Analog_Set_The_Only_One	2,131 KB	mp3	C:\Program Files\Lime... 0 / 0	0	0	0
great speeches - football - vince lombardi speech	2,149 KB	mp3	C:\Program Files\Lime... 0 / 0	0	0	0
Great Speeches - Malcom X - Black power	1,188 KB	mp3	C:\Program Files\Lime... 0 / 0	0	0	0
John_Vanderslice_Amitriptyline	5,786 KB	mp3	C:\Program Files\Lime... 0 / 0	0	0	0
TheQuickFixKills_Pick_Your_Poison	4,178 KB	mp3	C:\Program Files\Lime... 0 / 0	0	0	0

Name	Length	Bitrate
great speeches - football - vince lombardi speech.mp3	2:17	128
Great Speeches - Malcom X - Black power.mp3	1:16	128
John_Vanderslice_Amitriptyline.mp3	3:44	211
Alfie_Zappacosta_Start_Again.mp3	3:43	160

Hybrid P2P Συστήματα

• Fasttrack: File Sharing

Αναζήτηση Αρχείων σε
ένα SuperNode Δίκτυο

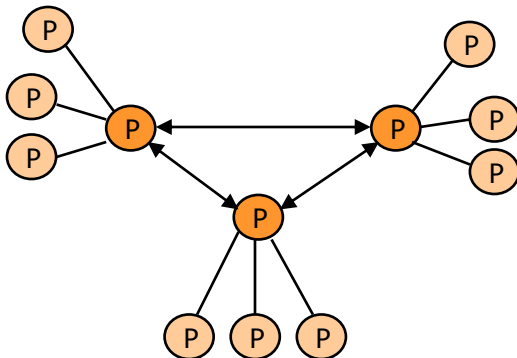


Hybrid P2P Συστήματα

- **Skype : Διαδικτυακή Τηλεφωνία**



- Δωρεάν συνομιλία με άλλους χρήστες οπουδήποτε στον κόσμο.
- Δημιουργήθηκε από τους ιδρυτές του εργαλείου ανταλλαγής αρχείων KaZaA
- Η Αρχιτεκτονική Π/Π χρησιμοποιείται για να δρομολογεί έξυπνα τα πακέτα φωνής μεταξύ των διάφορων Πελατών

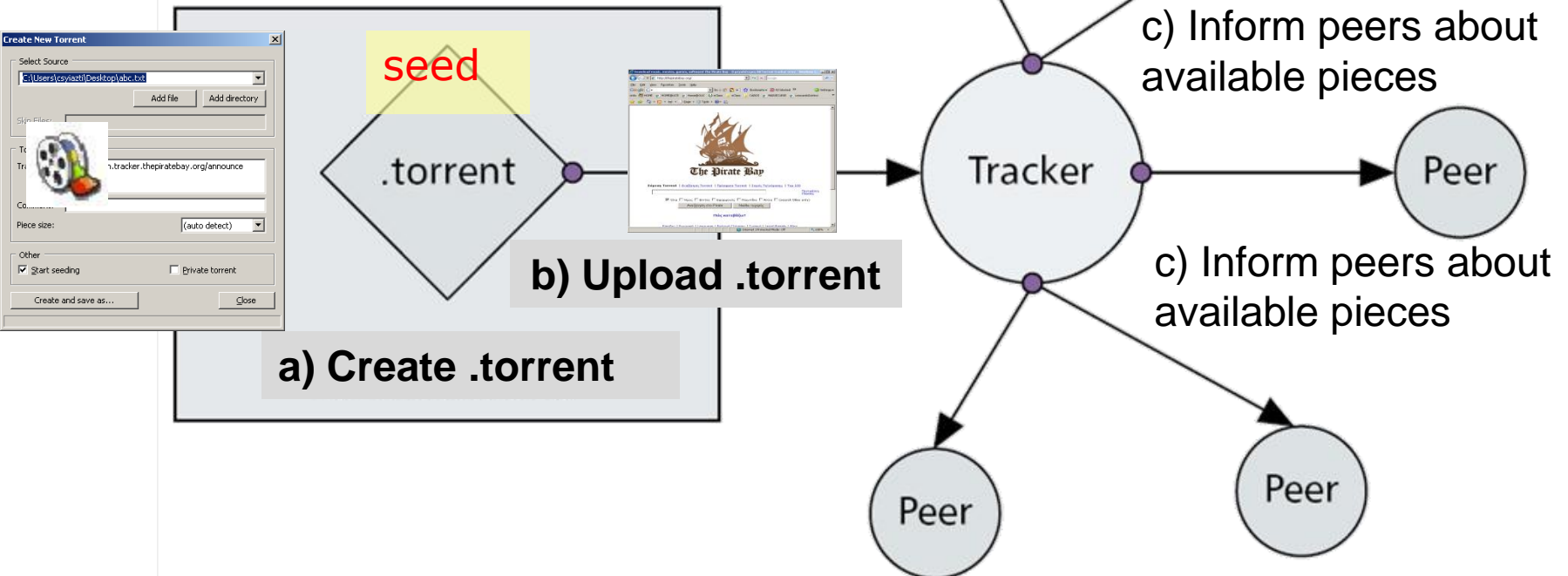


Το πρωτόκολλο επικοινωνίας των superpeers είναι κλειστό (proprietary)

Centralized P2P Σύστημα

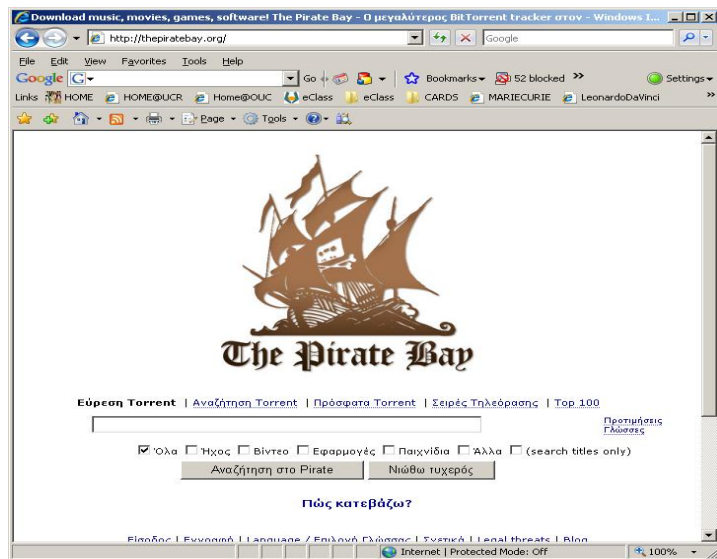
BitTorrent: File Sharing

Ένα πρωτόκολλο μεταφοράς δεδομένων (όχι αναζήτησης δεδομένων)

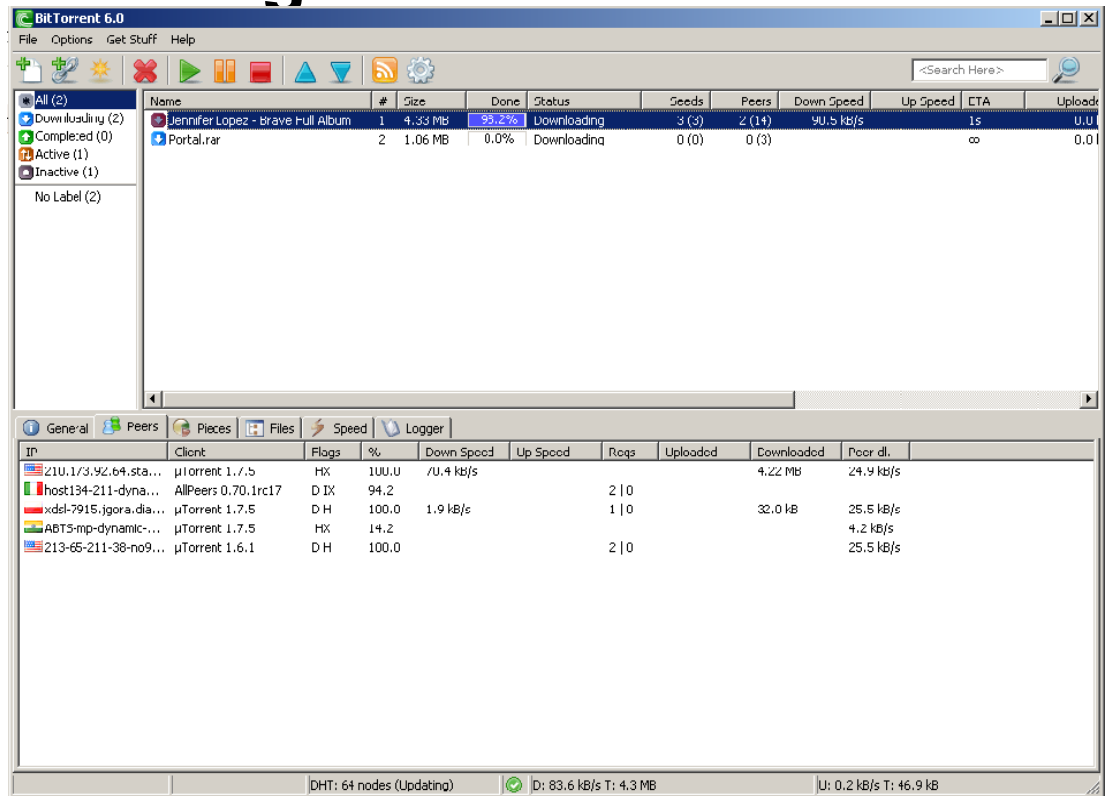


Centralized P2P Σύστημα

BitTorrent: File Searching



Αναζήτηση



Ανάκτηση 256KB από κάθε Peer μέχρι να ανακτηθεί όλο το αρχείο. Τα κομμάτια που ανακτούνται γίνονται παράλληλα διαθέσιμα στους υπόλοιπους μέσω του tracker

Κατάταξη P2P Συστημάτων

(Βάση της Δομής Δικτύου)

Βάση του πως οι κόμβοι συνδέονται μεταξύ τους

A. Αδόμητα P2P Συστήματα (**Unstructured P2P**)

- Δεν υπάρχει αυστηρά ορισμένη δομή του δικτύου
- Παραδείγματα: Gnutella, FastTrack/KaZaA, BitTorrent, Freenet, Overnet/eDonkey

B. Δομημένα P2P Συστήματα (**Structured P2P**)

- Η τοπολογία είναι αυστηρά ορισμένη
- Παραδείγματα : CAN, Chord, Tapestry, Pastry, Kademlia, Viceroy

Κατάταξη P2P Συστημάτων

(Βάση της Δομής Δικτύου)

A) *Unstructured P2P Systems*

Πλεονεκτήματα:

- + Εύκολη δημιουργία, συντήρηση του δικτύου.
- + Κατάλληλα για αναζήτηση δημοφιλών αντικειμένων.

Μειονεκτήματα:

- Η τοπολογία δεν είναι βέλτιστη με αποτέλεσμα οι αναζητήσεις να μην είναι αποδοτικές (π.χ., Gnutella)

B) *Structured P2P Systems – Distributed HashTables (DHTs)*

Πλεονεκτήματα:

- + Κατάλληλα για αναζήτηση συγκεκριμένων αντικειμένων.
- + Γρήγορη αναζήτηση $O(\log n)$

Μειονεκτήματα:

- Οι συχνές αφίξεις και αναχωρήσεις κόμβων (*churn*) δεν επιτρέπουν στο σύστημα να φτάσει ένα *steady state*.

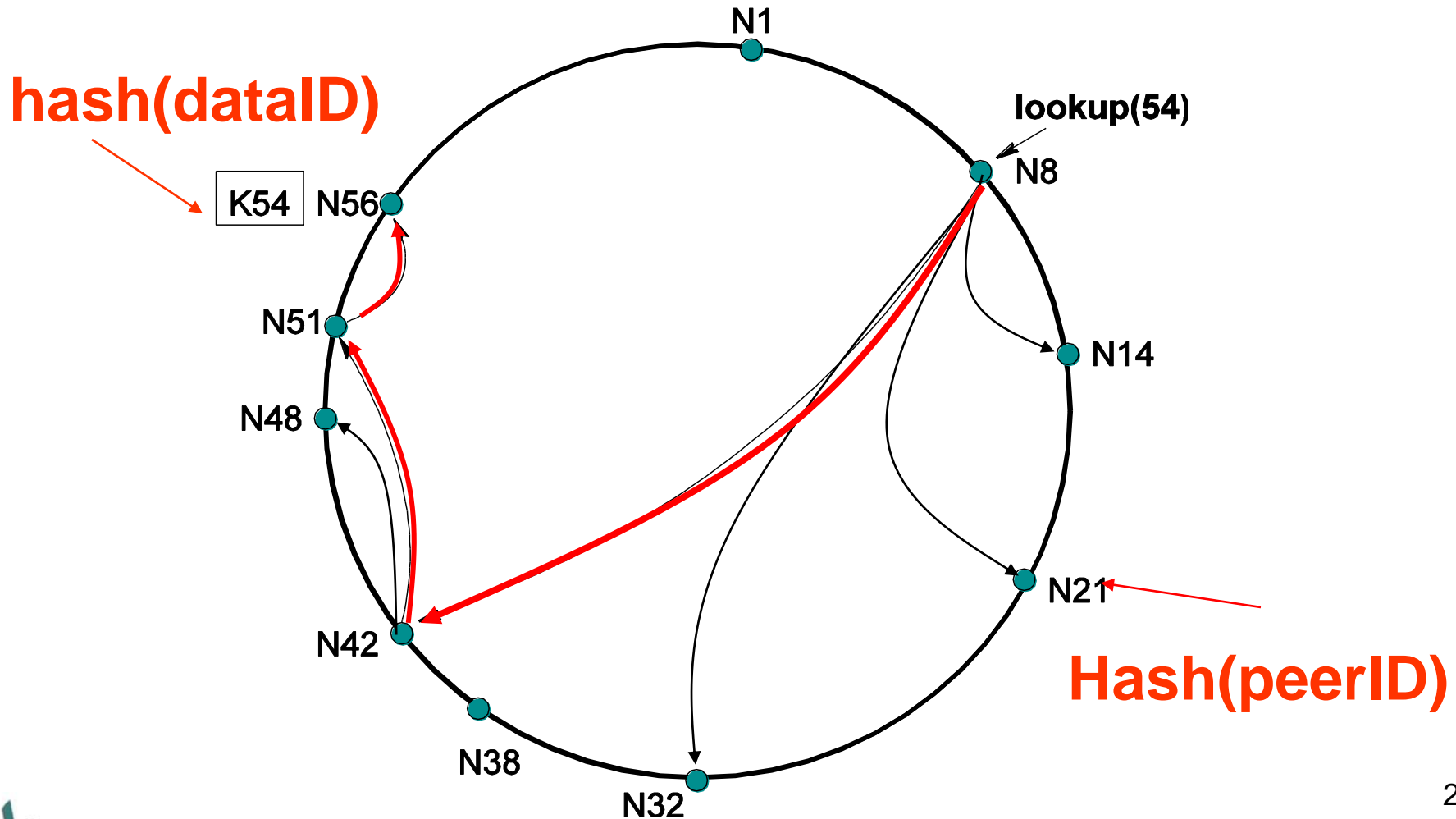
Παράδειγμα Structured P2P : Ο Αλγόριθμος Chord

Βασική Ιδέα: Τοποθέτησε τα δεδομένα σε συγκεκριμένο κόμβο έτσι ώστε η αναζήτηση να είναι αποδοτική $O(\log n)$

Αλγόριθμος σε ψηλό επίπεδο αφαιρετικότητας:

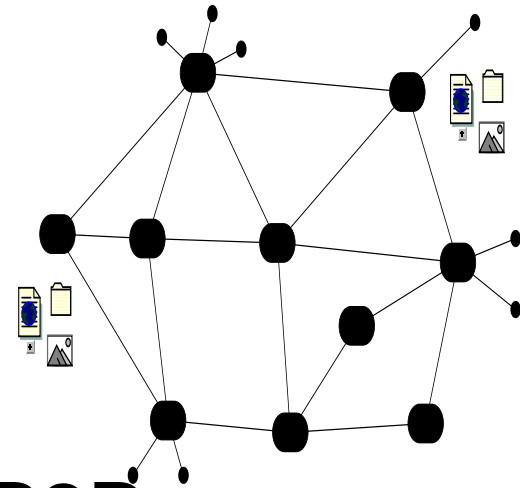
- Για κάθε **peer** και αντικείμενο **data** δημιούργησε το **hash(peerID)** και το **hash(dataID)** με το SHA-1.
- Ανάθεσε σε ένα νοητό δακτύλιο κάθε **hash(dataID)** στον **peerID** με το επόμενο μεγαλύτερο **hash(peerID)**
- Αν ψάξει κάποιος το **hash(dataID)** θα ξέρει ακριβώς σε πιο **peerID** βρίσκεται το **dataID** (θα βρισκείται στο **hash(peerID)**).

Παράδειγμα Structured P2P : Ο Αλγόριθμος Chord



Βασικά Θέματα που απασχολούν την Ερευνητική Κοινότητα Peer-to-Peer ?

1. **Data Placement:** Που πρέπει να τοποθετηθούν τα δεδομένα
2. **Search Mechanisms:** Με τι μηχανισμούς θα βρούμε την πληροφορία
3. **Overlay Structure:** Πως θα οργανώσουμε την δομή του P2P overlay δικτύου, έτσι ώστε να μπορούμε να πετύχουμε το 1-2 πιο αποδοτικά (χρόνος, χώρος, ανωνυμία, κτλ)

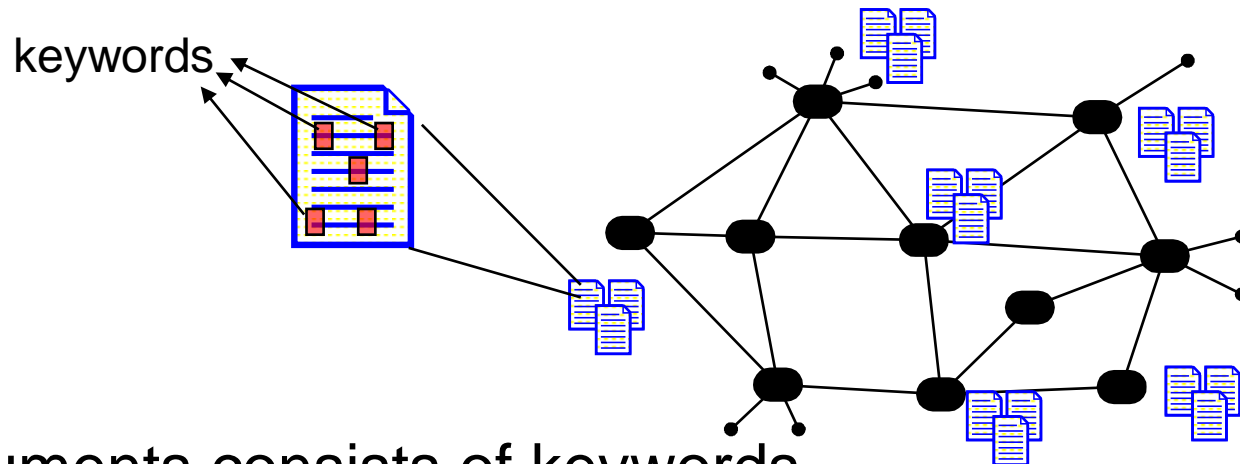


Μέρος 2: Information Retrieval in P2P Systems

Information Retrieval in P2P

Problem:

“How to efficiently retrieve Information in P2P systems where each node shares a collection of documents?”

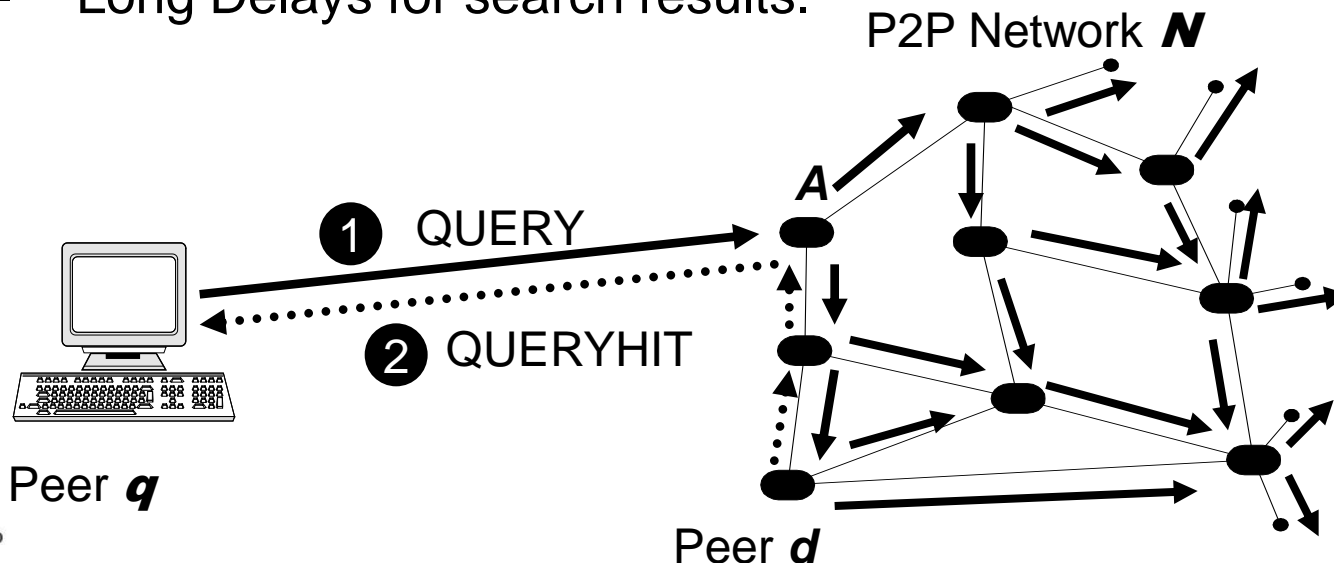


- Documents consists of keywords.
- We are not only interested in “Rem.mp3” but in Google-like keyword searches: “annual growth report cyprus”
- ***“Information retrieval (IR) is the science of searching information in documents (e.g., text, sound or images).***

Search Technique 1 - BFS

Breadth-First Search (Gnutella)

- **Idea:** Each Query Message is propagated along all outgoing links of a peer using TTL (time-to-live).
- TTL is decremented on each forward until it becomes 0
- Technique for I.R in P2P systems such as Gnutella.
- **Highlights**
 - The physical network is overloaded (even with TTL=7)
 - Long Delays for search results.



Search Technique 2 - RBFS

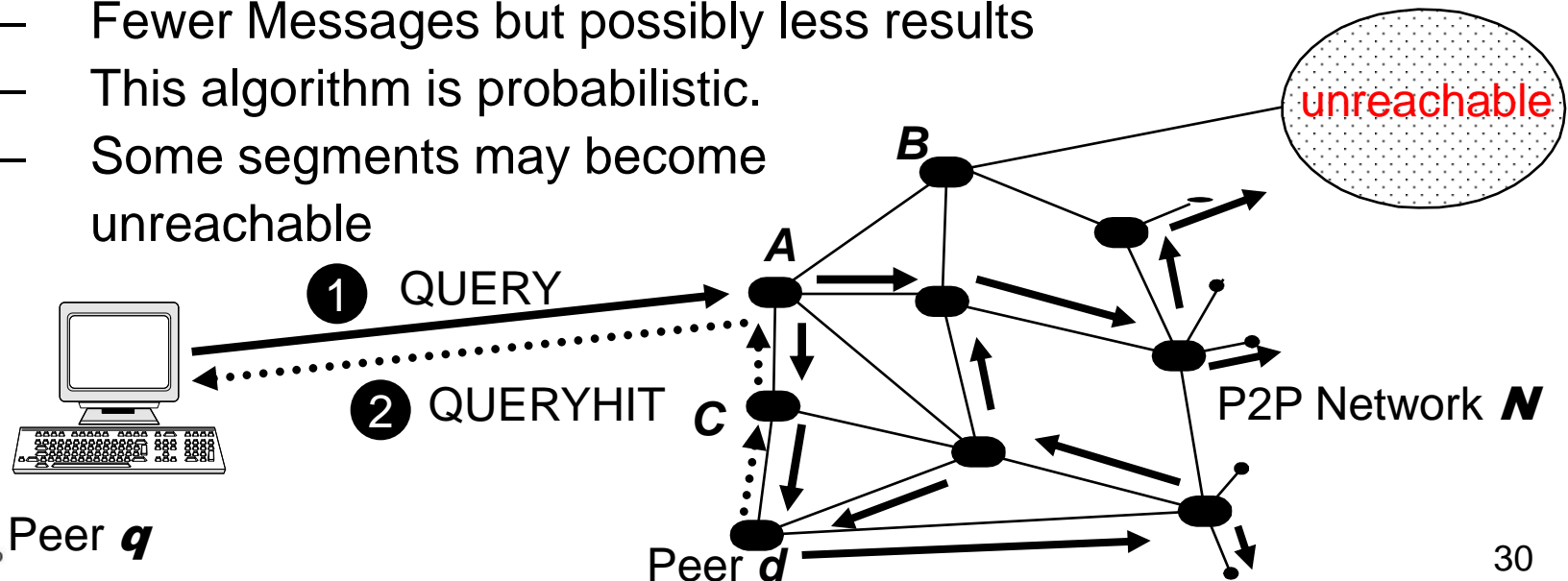
Modified Random BFS

[V. Kalogeraki, D. Gunopulos, D. Zeinalipour-Yazti . CIKM2002]

- **Idea:** Each Query Message is forwarded to only a fraction of outgoing links (e.g. $\frac{1}{2}$ of them).
- TTL is again decremented on each forward until it becomes 0.

- **Highlights**

- Fewer Messages but possibly less results
- This algorithm is probabilistic.
- Some segments may become unreachable

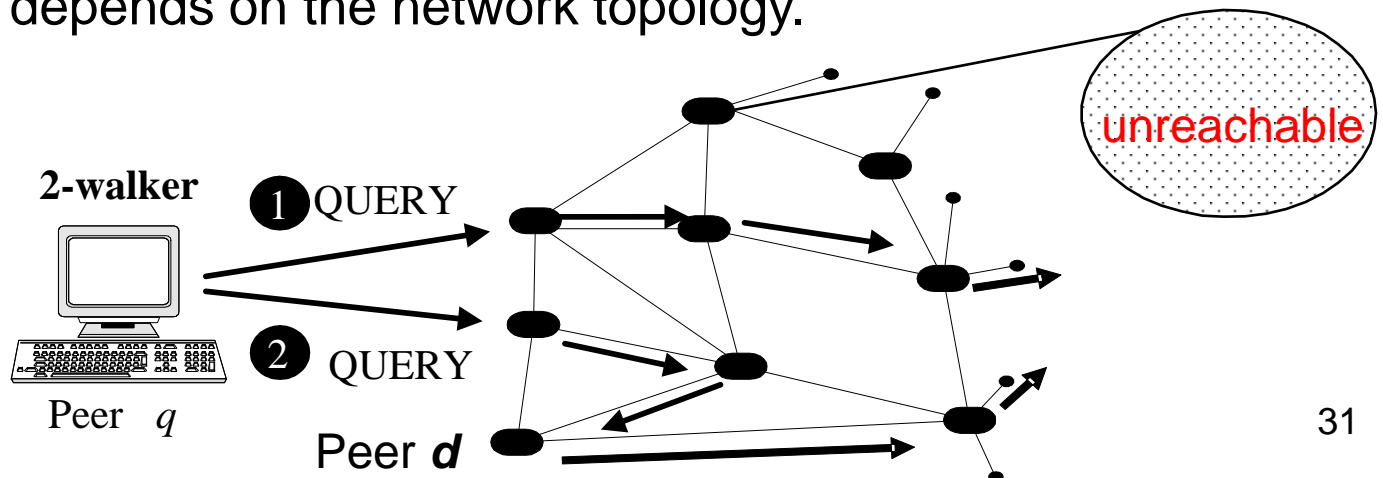


Search Technique 3 - RANDOM

Searching Using Random Walkers

[Q. Lv et al P. Cao, E. Cohen, K. Li, and S. Shenker. ICS2002]

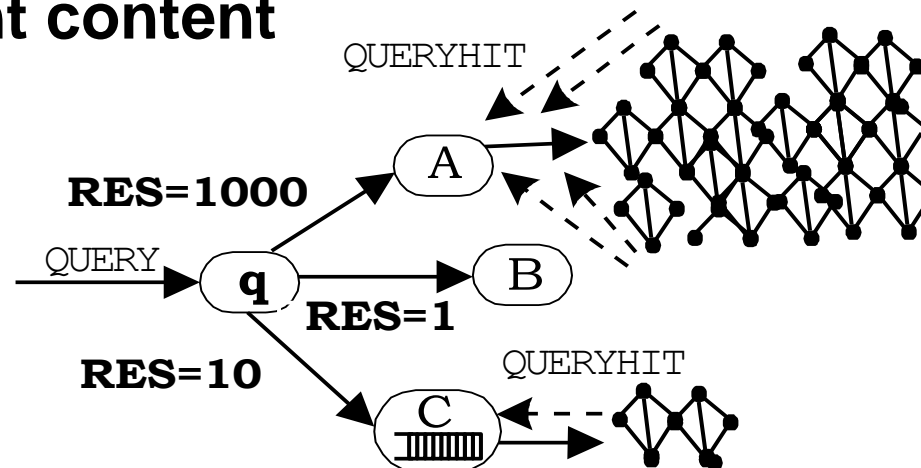
- **Idea:** Each Query Message is forwarded to 1 neighbor
- With k walkers after T steps we reach \approx the same nodes as 1 walker after kT steps. (They use 16-64 walkers)
- **Highlights**
 - Network Traffic reduced (from BFS) by 2 orders of magnitudes
 - Increases the user-perceived delay (from 2-6 hops to 4-15 hops)
 - This algorithm is probabilistic and the likelihood to locate the objects depends on the network topology.



Search Technique 4 - >RES

Directed BFS and the >RES Heuristic

- >RES: The Most Results Heuristic.
- **Idea:** Forward the query to your neighbors based on aggregate statistics (e.g. num of results a peer returned, shortest queue, shortest response time)
- **>RES works well because:**
 - It captures stable/large network segments.
 - Potentially less overloaded peers
- **Drawback: >RES doesn't route queries to the most relevant content**



Search Techniques: Remark

- On 1st June 2004 we crawled the Gnutella P2P Network for 5 hours with 17 workstations.
- We analyzed 15,153,524 query messages.
- **Observation:** High locality of specific queries...similar to the WWW.
- *We try to exploit this property for more efficient searches*

#	Query	Occurrence	%	#	Query	Occurrence	%
1	divx avi	588,146	3,88%	11	divx	24,363	0,16%
2	spiderman avi	50,175	0,33%	12	spiderman	23,274	0,15%
3	p__ mpg	39,168	0,25%	13	xxx avi	22,408	0,14%
4	star wars avi	38,473	0,25%	14	capture the light	21,651	0,14%
5	avi	29,911	0,19%	15	buffy mpg	20,365	0,13%
6	s__ mpg	27,895	0,18%	16	g__ mpg	20,251	0,13%
7	Eminem	27,440	0,18%	17	buffy avi	19,874	0,13%
8	eminem mp3	25,693	0,16%	18	t__ mpg	19,492	0,12%
9	dvd avi	25,105	0,16%	19	seinfeld vivid	18,809	0,12%
10	b_____	24,753	0,16%	20	xxx mpg	18,686	0,12%

Search Technique 5 - ISM

Intelligent Search Mechanism (ISM) [CIKM'02, IS'05]

Query	GUID	Connection & Hits	Timestamp
Athens 2004 Olympics	G439ID	(peer1,20), (peer4,50),...	100002222
VLDB Canada 2004	F549QL	(peer2,10)	100065652
***	***	***	***
Florida storm	PN329D	NULL	100022453

|L|-dim space: {athens,2004,olympics,vldb,canada,florida, storm}
e.g. If q ="athens olympics" $\Rightarrow \vec{q}$ (vector of q) = [1,0,1,0,0,0,0]

b) Cosine Similarity – The Similarity Function

$$\text{sim}(q, q_i) = \cos(q, q_i) = \frac{\sum(\vec{q} * \vec{q}_i)}{\sqrt{\sum(\vec{q})^2} * \sqrt{\sum(\vec{q}_i)^2}}$$

c) RelevanceRank – Ranking Neighbors by similarity

$$RR(\text{peer}_i, q) = \sum_{q_j = \text{"Queries answered by } \text{peer}_i\text{"}} \text{sim}(q_j, q)^\alpha * \text{results}(q_j)$$



Search Technique 6 - PlanetP

Using Randomized Gossiping to Replicate

Global State [F.M Cuenca-Acuna, Thu D. Nguyen HPDC-12]

- **Idea: Advertise a summary of a Peer's content to the neighborhood (using Bloom Filters).**
- **Bloom Filters are used for Membership Queries**

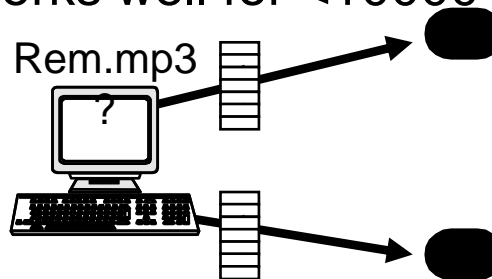
PeerA: Does PeerB maintain “rem.mp3”?

PeerA: `lookup_locally(BloomfilterB, “rem.mp3”)`

PeerA: Answer NO (definite) or MAYBE (False Positives are possible)

- **Highlights**

- No Data Replication Required
- Fairly Scalable (Works well for <10000 nodes)



Συμπαγείς Κατάλογοι (Compact Membership Directories - Bloom Filters)

Πρόβλημα

- Έστω ότι ο **A** θέλει να στείλει στον **B** μία συλλογή $D=\{d_1,d_2,\dots,d_n\}$ από n στοιχεία.
- Αυτή η συλλογή μπορεί να είναι για παράδειγμα τα ονόματα όλων των αρχείων που έχει ο **A**, έτσι ώστε να ξέρει και ο **B** τι αρχεία έχει ο **A**.
- Αντί να στείλει όλη την λίστα ο **A** στον **B** (το οποίο είναι ακριβό), ο **A** μπορεί να στείλει ένα συμπαγή κατάλογο (Bloom Filter)

Bloom Filter: είναι ένα διάνυσμα **V** (**vector**) από **m bits**, στο οποίο μπορεί να συμπεριληφθεί το περιεχόμενο της συλλογής **D**, με τον ακόλουθο τρόπο:

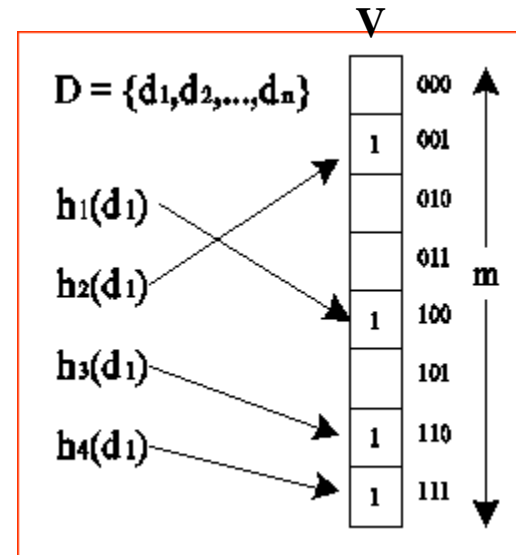
Συμπαγείς Κατάλογοι (Compact Membership Directories - Bloom Filters)

Bloom Filter

1. Θέτουμε σε ένα διάνυσμα V μεγέθους m όλα τα στοιχεία ίσο με «0».
2. Χρησιμοποιώντας K διαφορετικά hash functions, κάνουμε hash όλα τα n στοιχεία της συλλογής $D=\{d_1,d_2,\dots,d_n\}$.
3. Όπου γίνεται hash θέτουμε την τιμή του πίνακα ίσο με «1»
4. Ο A αποστέλλει το V (όχι το D) στον B .
5. Αν ο B θέλει να δει αν ο A έχει ένα αρχείο «test.c», τότε χρησιμοποιεί τα ίδια K hash functions πάνω στο όνομα του αρχείου και ελέγχει αν όλα τα στοιχεία του διανύσματος V είναι ίσο με «1».

Αν όχι τότε ο A σίγουρα δεν έχει το αρχείο,
Ειδ' αλλιώς πολύ πιθανόν ο A το έχει!

- Για μεγάλες τιμές του m και K τα bloom filters δουλεύουν πολύ αποδοτικά
- Έχουν εφαρμογές σε καταναμημένα συστήματα (web caches, p2p systems, etc)



Search Techniques: Remark

- All techniques discussed so far do not take into account **anonymity**.
- While A downloads (or conducts queries) through B, B knows exactly what we are looking for
- Freenet
 - Designed to provide anonymity.
 - ...although it is not widely used today.

Search Technique 7 - Freenet

Depth-First-Search and Freenet

[I. Clarke O. Sandberg, B. Wiley, and T.W. Hong, LNCS 2009]

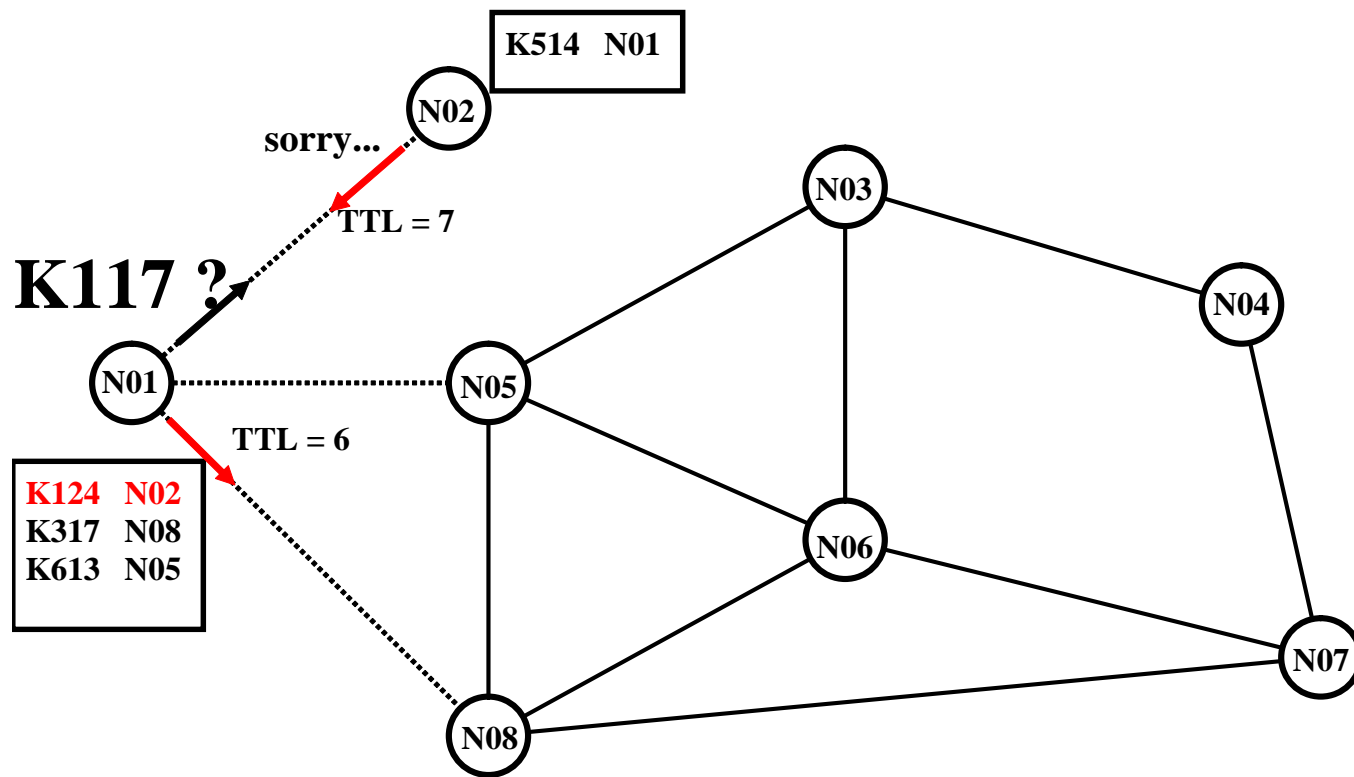
Idea: Looking for A? Search for hash(A).
Queries are routed based on the “key closeness” in a DFS manner.

Highlights:

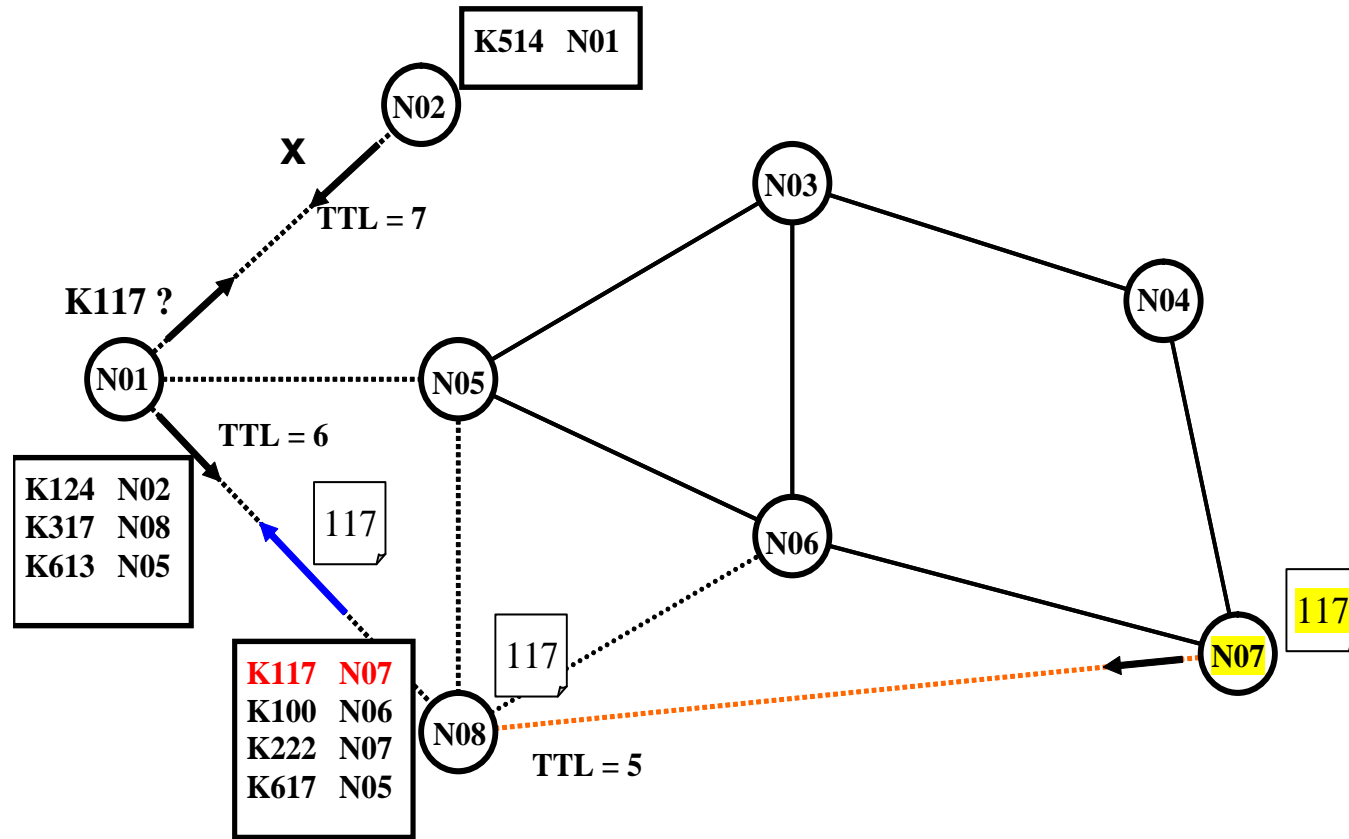
- Uses caching of key/object for future requests.
- Data Replication along the QueryHit path provides **availability**
- Anonymity of Searcher and Publisher.
- Drawbacks:
 - Searches ONLY based on Object Identifier.**
 - The user-perceived delay is high

Search Technique 7 - Freenet

Search: Queries are routed based on the “key closeness” in a DFS manner.



Search Technique 7 - Freenet



Item Found:

- The key K117 is added to the key table of N08, while the answer is routed back.
- Document 117 is replicated along the Queryhit path

Search Technique 7 - Freenet

- pros:
 - complete decentralization
 - fault tolerance/self-organization
 - anonymity
 - scalability (to some degree)
- cons:
 - questionable efficiency & performance
 - rare keys disappear from the system

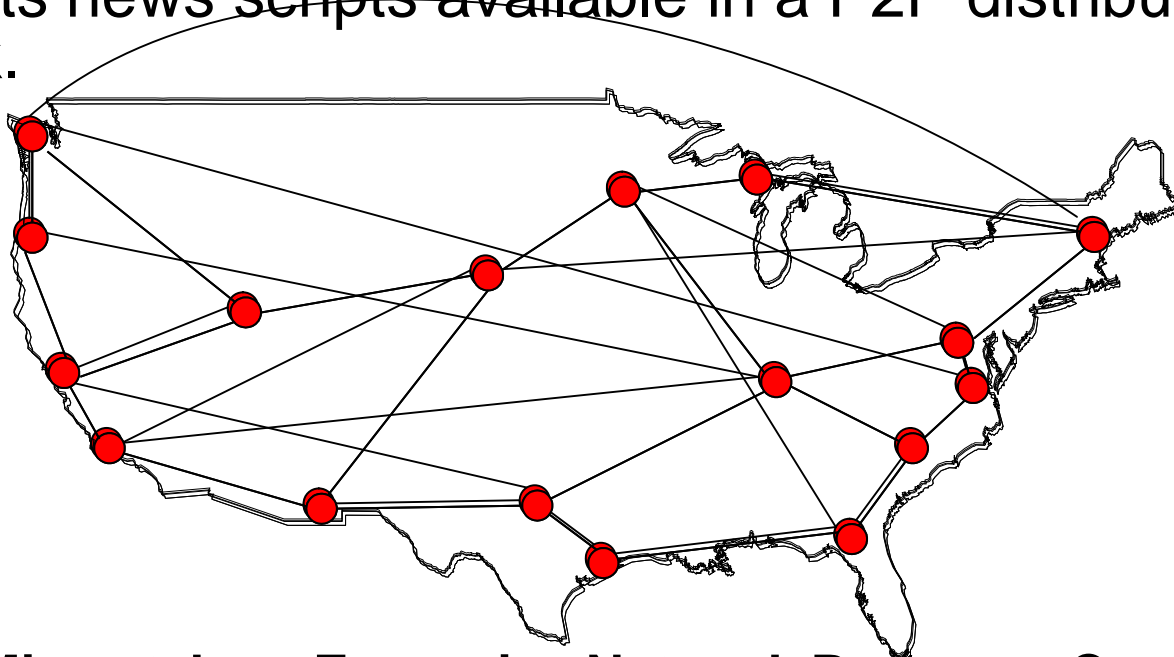
Network-Efficient Topologies

- P2P Networks are usually network-agnostic. Therefore they don't take into regards the **efficiency of the overlay topology**.

Example

- Assume that we have a US Newspaper Agency that makes its news scripts available in a P2P distribution network.

The Agency
Topology
Nodes
Network
Network



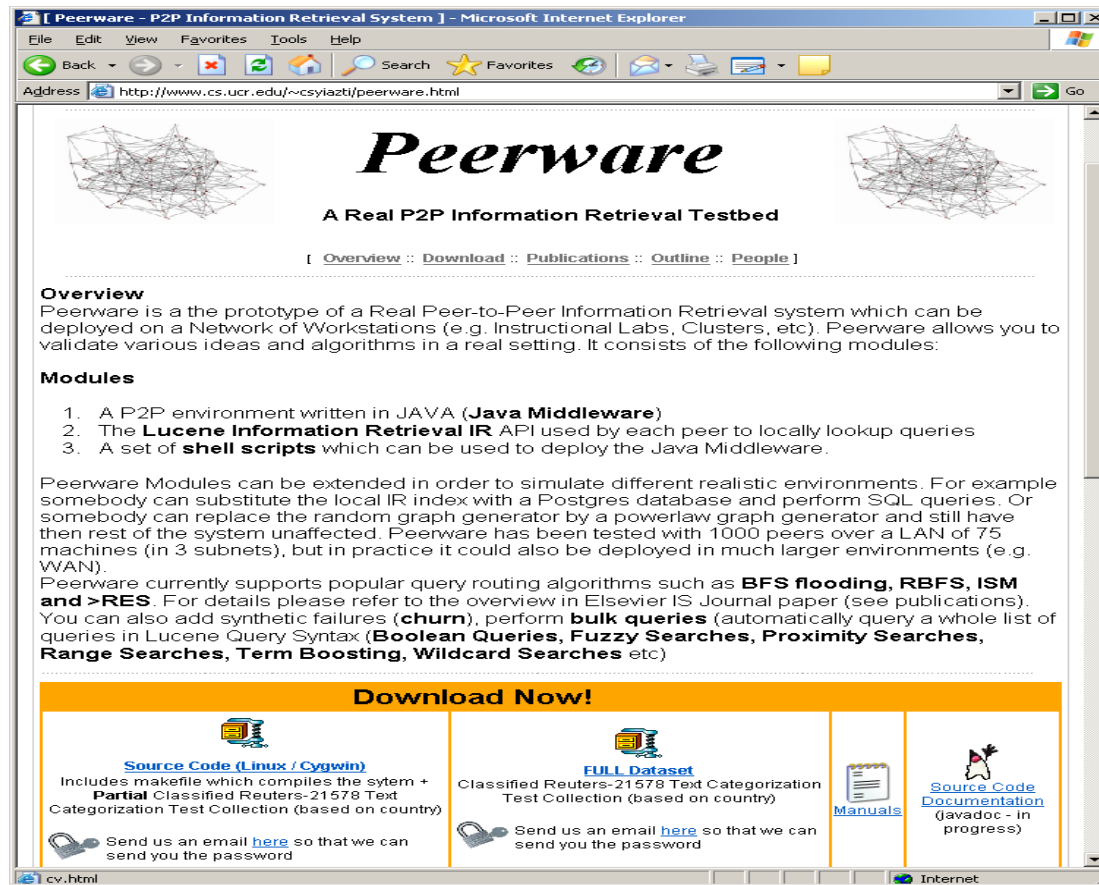
Network Mismatch => Excessive Network Resource Consumption

Open-Source Software

The Peerware system:

<http://www.cs.ucr.edu/~csyiazti/peerware.html>

(An open-source Distributed Content-Retrieval System)



The screenshot shows a web browser window titled "Peerware - P2P Information Retrieval System" with the address bar displaying "http://www.cs.ucr.edu/~csyiazti/peerware.html". The page features a central heading "Peerware" in a large, stylized font, flanked by two network graph visualizations. Below the heading is the subtitle "A Real P2P Information Retrieval Testbed" and a navigation menu with links for "Overview", "Download", "Publications", "Outline", and "People".

Overview
Peerware is the prototype of a Real Peer-to-Peer Information Retrieval system which can be deployed on a Network of Workstations (e.g. Instructional Labs, Clusters, etc). Peerware allows you to validate various ideas and algorithms in a real setting. It consists of the following modules:





Modules

1. A P2P environment written in JAVA (**Java Middleware**)
2. The **Lucene Information Retrieval IR** API used by each peer to locally lookup queries
3. A set of **shell scripts** which can be used to deploy the Java Middleware.

Peerware Modules can be extended in order to simulate different realistic environments. For example somebody can substitute the local IR index with a Postgres database and perform SQL queries. Or somebody can replace the random graph generator by a powerlaw graph generator and still have the rest of the system unaffected. Peerware has been tested with 1000 peers over a LAN of 75 machines (in 3 subnets), but in practice it could also be deployed in much larger environments (e.g. WAN).

Peerware currently supports popular query routing algorithms such as **BFS flooding, RBFS, ISM and >RES**. For details please refer to the overview in Elsevier IS Journal paper (see publications). You can also add synthetic failures (**churn**), perform **bulk queries** (automatically query a whole list of queries in Lucene Query Syntax (**Boolean Queries, Fuzzy Searches, Proximity Searches, Range Searches, Term Boosting, Wildcard Searches** etc))

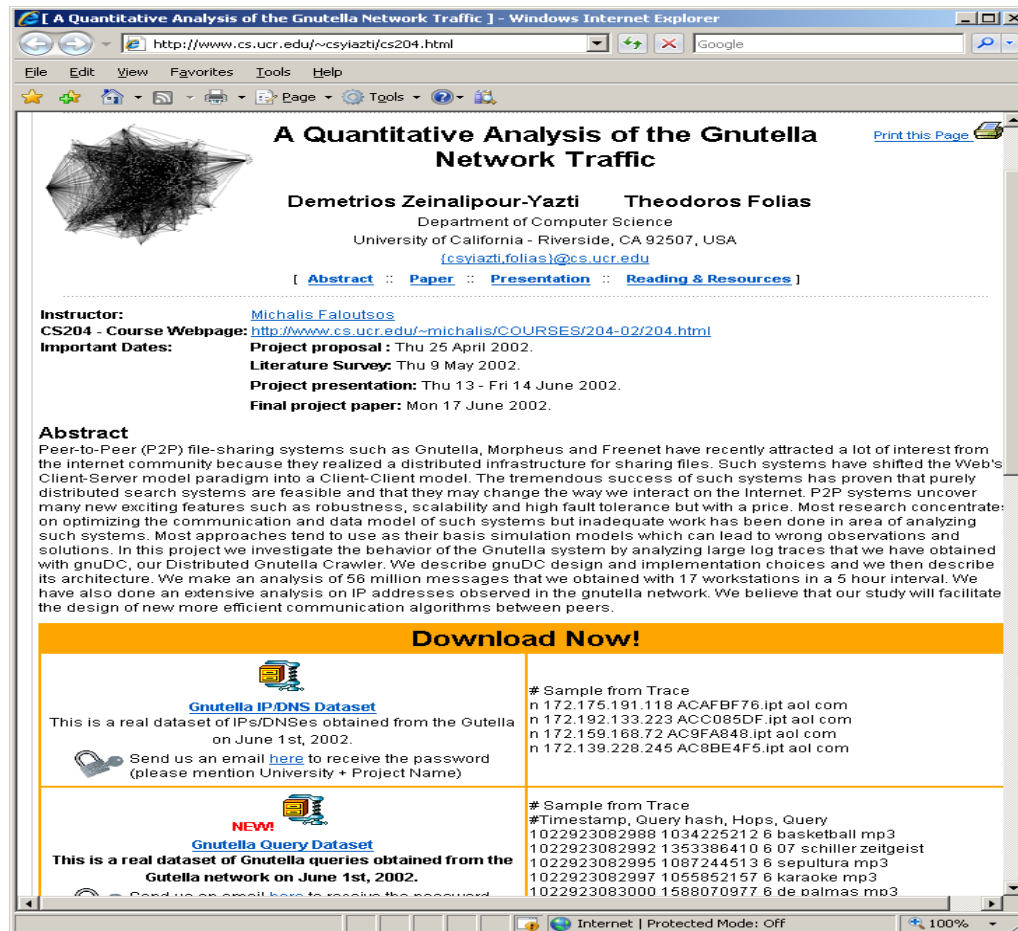
Download Now!

 Source Code (Linux / Cygwin) Includes makefile which compiles the system + Partial Classified Reuters-21578 Text Categorization Test Collection (based on country)	 FULL Dataset Classified Reuters-21578 Text Categorization Test Collection (based on country)	 Manuals	 Source Code Documentation (javadoc - in progress)
---	--	--	---

Send us an email [here](#) so that we can send you the password

Open-Source Software

Publicly-Available traces from Gnutella:
<http://www.cs.ucr.edu/~csyiazti/cs204.html>



The screenshot shows a Windows Internet Explorer browser window. The address bar contains the URL <http://www.cs.ucr.edu/~csyiazti/cs204.html>. The page title is "A Quantitative Analysis of the Gnutella Network Traffic". The authors listed are Demetrios Zeinalipour-Yazti and Theodoros Folias, from the Department of Computer Science at the University of California - Riverside, CA 92507, USA. The page includes links for "Abstract", "Paper", "Presentation", and "Reading & Resources". Below this, there is information about the instructor, Michalis Faloutsos, and the course CS204. A section titled "Abstract" describes the project's focus on analyzing Gnutella network traffic. At the bottom, there are two "Download Now!" sections. The first section is for the "Gnutella IP/DNS Dataset" and provides a list of IP addresses. The second section is for the "Gnutella Query Dataset" and provides a list of query timestamps and hashes.

A Quantitative Analysis of the Gnutella Network Traffic





Demetrios Zeinalipour-Yazti Theodoros Folias
Department of Computer Science
University of California - Riverside, CA 92507, USA
csyiazti.folias@cs.ucr.edu

[[Abstract](#) :: [Paper](#) :: [Presentation](#) :: [Reading & Resources](#)]

Instructor: [Michalis Faloutsos](#)
CS204 - Course Webpage: <http://www.cs.ucr.edu/~michalis/COURSES/204-02/204.html>
Important Dates: **Project proposal:** Thu 25 April 2002.
Literature Survey: Thu 9 May 2002.
Project presentation: Thu 13 - Fri 14 June 2002.
Final project paper: Mon 17 June 2002.

Abstract
Peer-to-Peer (P2P) file-sharing systems such as Gnutella, Morpheus and Freenet have recently attracted a lot of interest from the internet community because they realized a distributed infrastructure for sharing files. Such systems have shifted the Web's Client-Server model paradigm into a Client-Client model. The tremendous success of such systems has proven that purely distributed search systems are feasible and that they may change the way we interact on the Internet. P2P systems uncover many new exciting features such as robustness, scalability and high fault tolerance but with a price. Most research concentrate on optimizing the communication and data model of such systems but inadequate work has been done in area of analyzing such systems. Most approaches tend to use as their basis simulation models which can lead to wrong observations and solutions. In this project we investigate the behavior of the Gnutella system by analyzing large log traces that we have obtained with gnuDC, our Distributed Gnutella Crawler. We describe gnuDC design and implementation choices and we then describe its architecture. We make an analysis of 56 million messages that we obtained with 17 workstations in a 5 hour interval. We have also done an extensive analysis on IP addresses observed in the gnutella network. We believe that our study will facilitate the design of new more efficient communication algorithms between peers.

Download Now!

 Gnutella IP/DNS Dataset This is a real dataset of IPs/DNSes obtained from the Gutella on June 1st, 2002.  Send us an email here to receive the password (please mention University + Project Name)	# Sample from Trace n 172.175.191.118 ACAFBF76.ipt aol com n 172.192.133.223 ACC085DF.ipt aol com n 172.159.168.72 AC9FA848.ipt aol com n 172.139.228.245 AC8BE4F5.ipt aol com
 NEW! Gnutella Query Dataset This is a real dataset of Gnutella queries obtained from the Gutella network on June 1st, 2002.  Send us an email here to receive the password	# Sample from Trace #Timestamp, Query hash, Hops, Query 1022923082988 1034225212 6 basketball mp3 1022923082992 1353386410 6 07 schiller zeitgeist 1022923082995 1087244513 6 sepultura mp3 1022923082997 1055852157 6 karaoke mp3 1022923083000 1588070977 6 de palmas mp3

Σχετικά Ερευνητικά Άρθρα

- **"pFusion: An Architecture for Internet-Scale Content-Based Search and Retrieval"** by D. Zeinalipour-Yazti, V. Kalogeraki, D. Gunopulos, IEEE Transactions on Parallel and Distributed Systems, (**IEEE TPDS**), accepted, 2006.
- **"Structuring Topologically-Aware Overlay Networks using Domain Names"**, D. Zeinalipour-Yazti, V. Kalogeraki, **Computer Networks** (Comnet), Elsevier Publications, Volume 50, Issue 16 , pp. 3064-3082, 2006.
- **"Exploiting Locality for Scalable Information Retrieval in Peer-to-Peer Systems"**, D. Zeinalipour-Yazti, V. Kalogeraki and D. Gunopulos, Information Systems (InfoSys), Volume 30, Issue 4, Pages 277-298, 2005.

*“Peer-to-Peer Systems: Introduction
and Challenges”*

*“Συστήματα Ομοτίμων: Εισαγωγή και
Προκλήσεις “*

Δημήτρης Ζεϊναλιπούρ

Λέκτορας, Ανοικτό Πανεπιστήμιο Κύπρου

Βιβλιογραφία

- Eng Keong Lua et al. "A Survey and Comparison of Peer-to-Peer Overlay Network Schemes," IEEE Communications Surveys and Tutorials, Vol 7, No 2 (Second Quarter, 2005), pp. 72-93.
- D. Zeinalipour-Yazti, V. Kalogeraki and D. Gunopulos, "Information Retrieval Techniques for Peer-to-Peer Networks", IEEE CiSE Magazine, Special Issue on Web Engineering, IEEE Publications, pp.12-20., July/August 2004