Improving Data Access in P2P Systems

Karl Aberer and Magdalena Punceva

Swiss Federal Institute of Technology

Manfred Hauswirth and Roman Schmidt

Technical University of Vienna

Outline

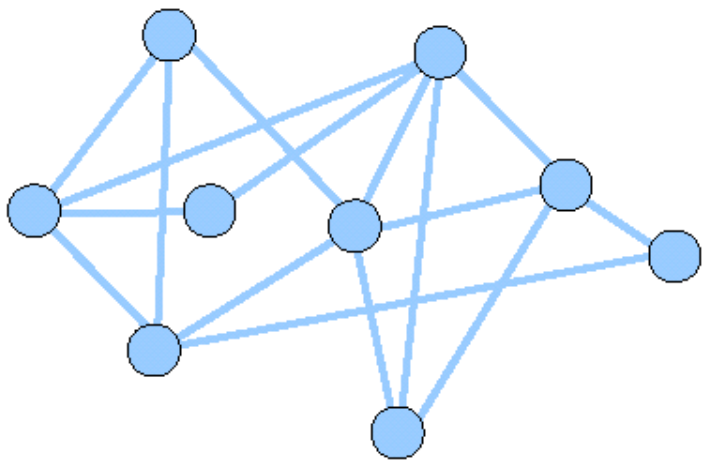
- Introduction
- Gnutella system
- Peer-Grid structure
- Gridella system
- Future works and conclusions

Introduction

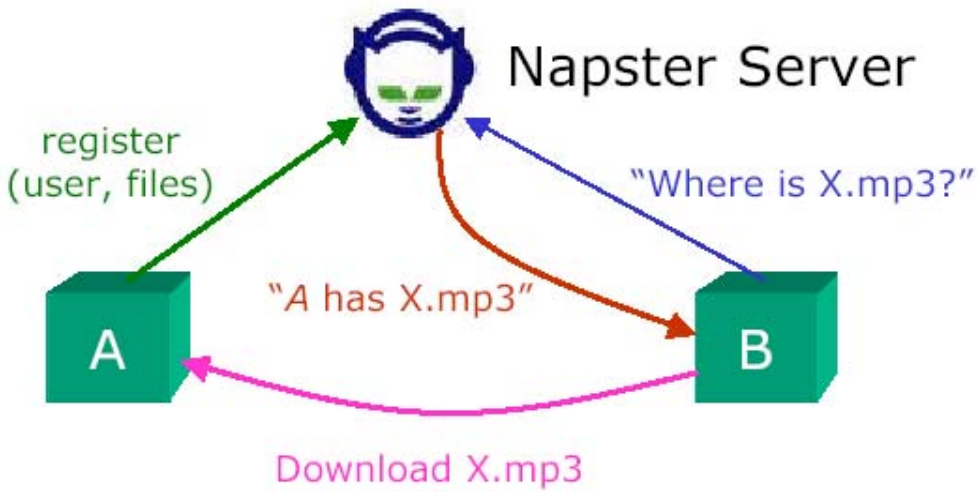
- Client-server-based system
  - Server capacity
  - Network capacity
  - Service reliability
- Some solutions
  - Load balancing
  - Fault tolerance
  - Caching
  - Replication

What is P2P?

- Every participating node acts as both a client and a server ("servent")
- Every node "pays" its participation by providing access to (some of) its resources
- Properties:
  - no central coordination
  - no central database
  - no peer has a global view of the system
  - global behavior emerges from local interactions
  - all existing data and services are accessible from any peer
  - peers are autonomous
  - peers and connections are unreliable



Napster: Communication Model



Introduction

- This paper present Gridella, a P2P system based on Peer-Grid (P-Grid), improves on Gnutella's search performance while reducing bandwidth requirements.

Gnutella

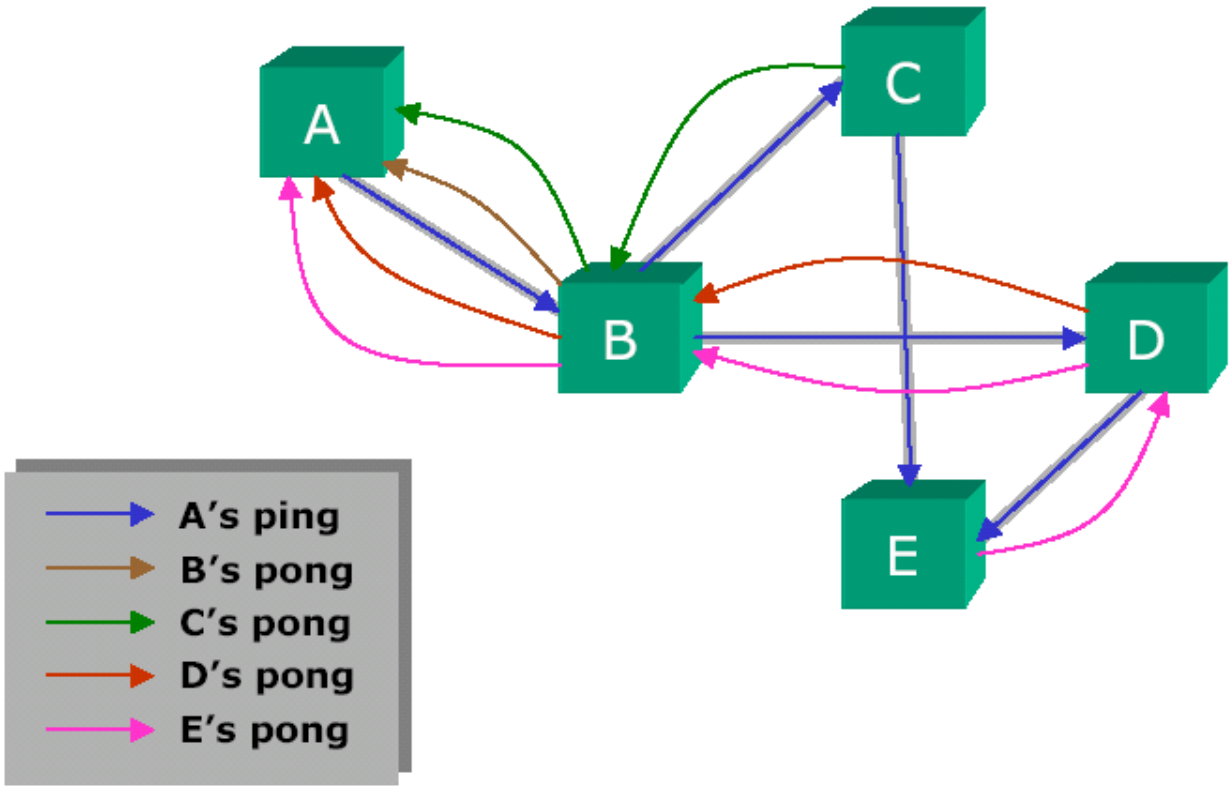
- Gnutella is a decentralized file-sharing system.
- The host communicates peer-to-peer via the Gnutella protocol.

- http://www.clip2.com/GnutellaProtocol04.pdf
- http://www.gnutella.com

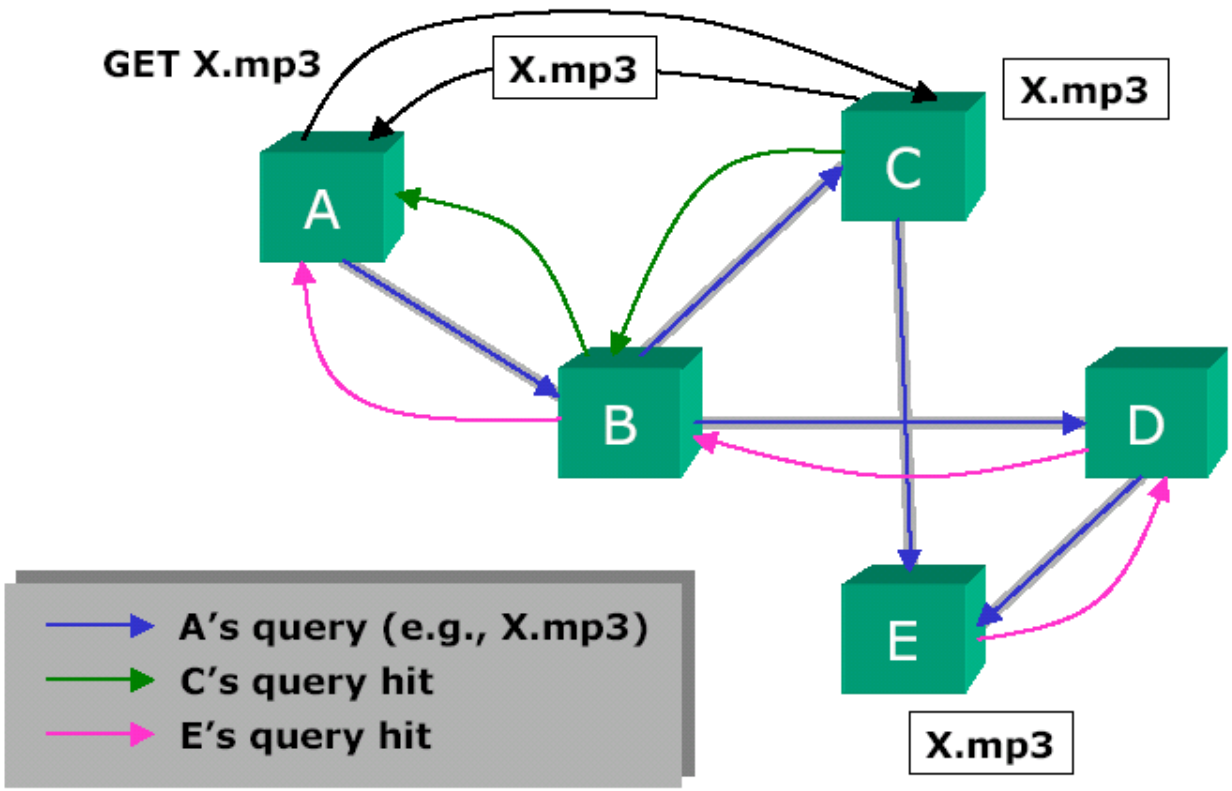
Gnutella: Protocol Message Types

Type	Description	Contained Information
Ping	Announce availability and probe for other servents	None
Pong	Response to a ping	IP address and port# of responding servent; number and total kb of files shared
Query	Search request	Minimum network bandwidth of responding servent; search criteria
QueryHit	Returned by servents that have the requested file	IP address, port# and network bandwidth of responding servent; number of results and result set
Push	File download requests for servents behind a firewall	Servent identifier; index of requested file; IP address and port to send file to

Gnutella: Meeting Peers (Ping/Pong)



Gnutella: Searching (Query/QueryHit/GET)



Gnutella's Problems

- High bandwidth consumption
- High cost of database search on all requested hosts
- No estimate of query duration and probability for success of query
- Free-riding problem

Gnutella: Summary

- Completely decentralized
- Hit rates are high
- High fault tolerance
- Adopts well and dynamically to changing peer populations
- Protocol causes high network traffic (e.g., 3.5Mbps). For example:
  - 4 connections C / peer, TTL = 7
  - 1 ping packet can cause $2 * \sum_{i=0}^{TTL} C * (C-1)^i = 26{,}240$ packets
- No estimates on the duration of queries can be given
- No probability for successful queries can be given
- Topology is unknown ⇒ algorithms cannot exploit it
- Free riding is a problem
- Reputation of peers is not addressed
- Simple, robust, and scalable (at the moment)

Free-riding on Gnutella [Adar00]

- 24 hour sampling period:
  - 70% of Gnutella users share no files
  - 50% of all responses are returned by top 1% of sharing hosts

- Verified hypotheses:
  - H1: A significant portion of Gnutella peers are free riders.
  - H2: Free riders are distributed evenly across domains
  - H3: Often hosts share files nobody is interested in (are not downloaded)

- Can a set of peers without central coordination provide efficient search and small storage space?

CAN

- Content Addressable Network(*SIGCOMM'01*)

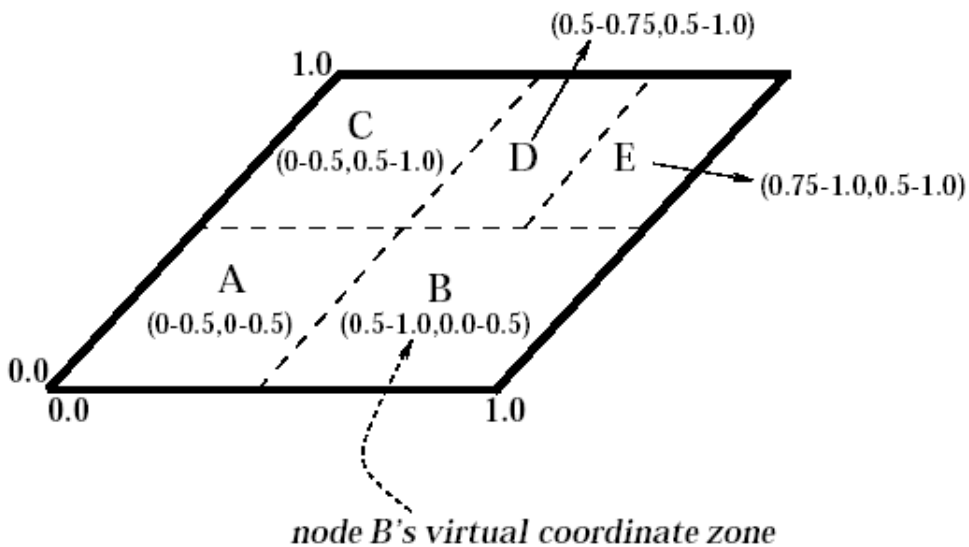
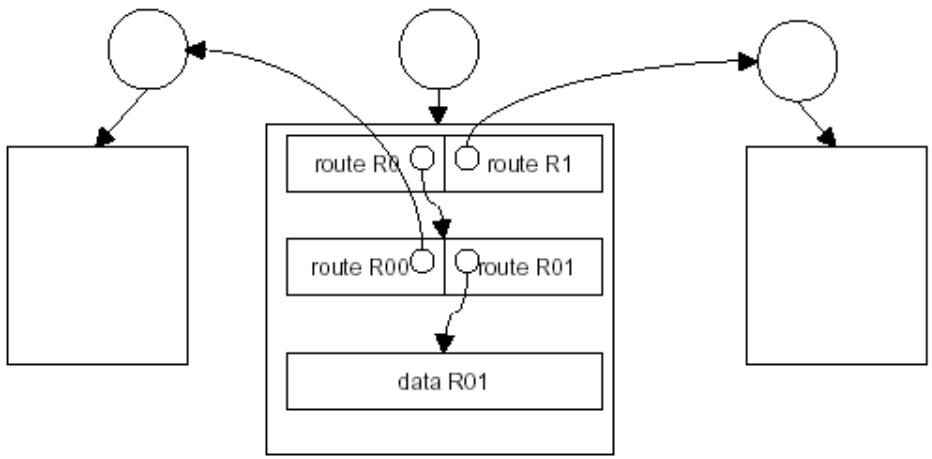


Figure 1: ***Example 2-d space with 5 nodes***

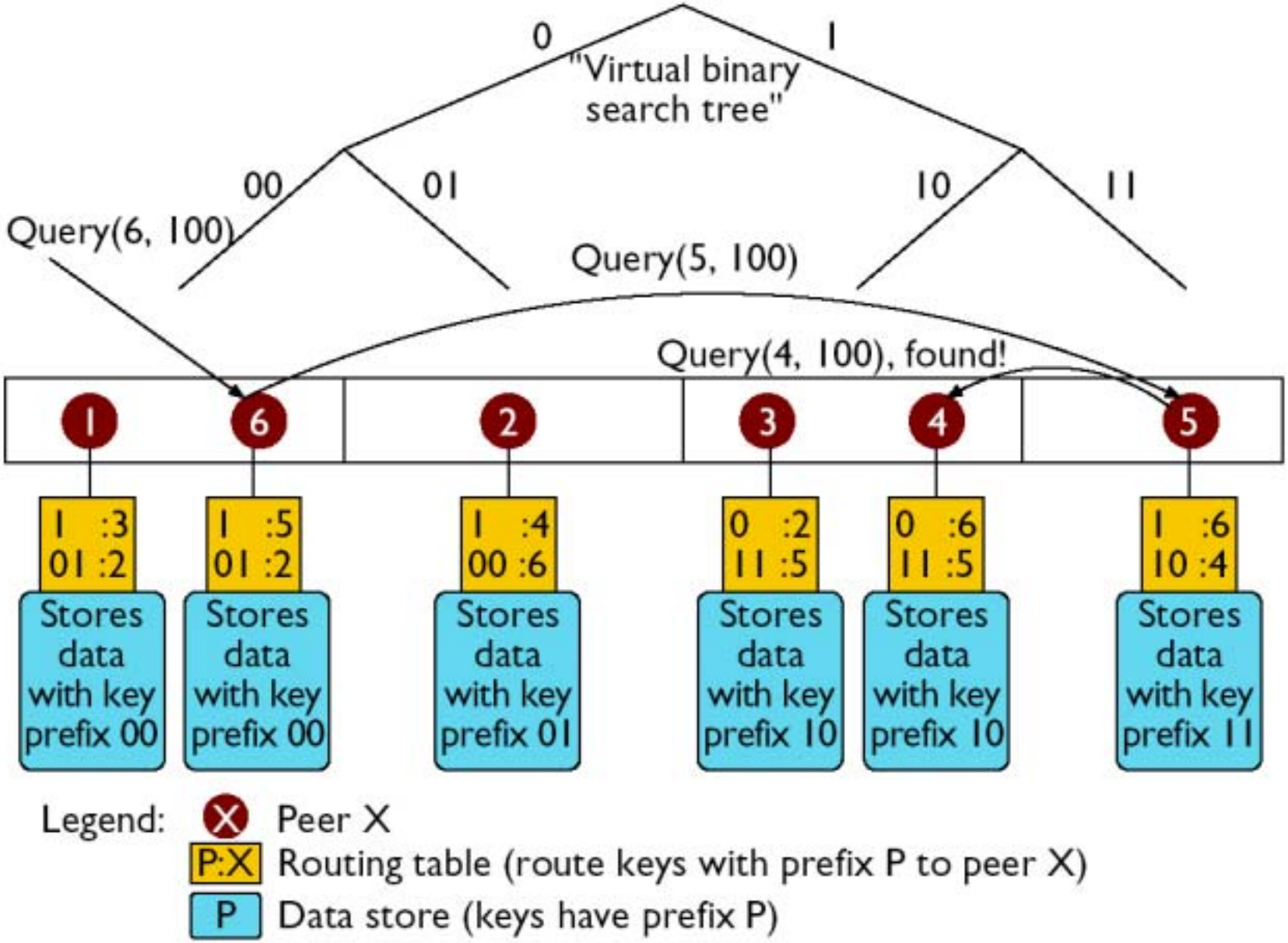
P2P Data Access Structures

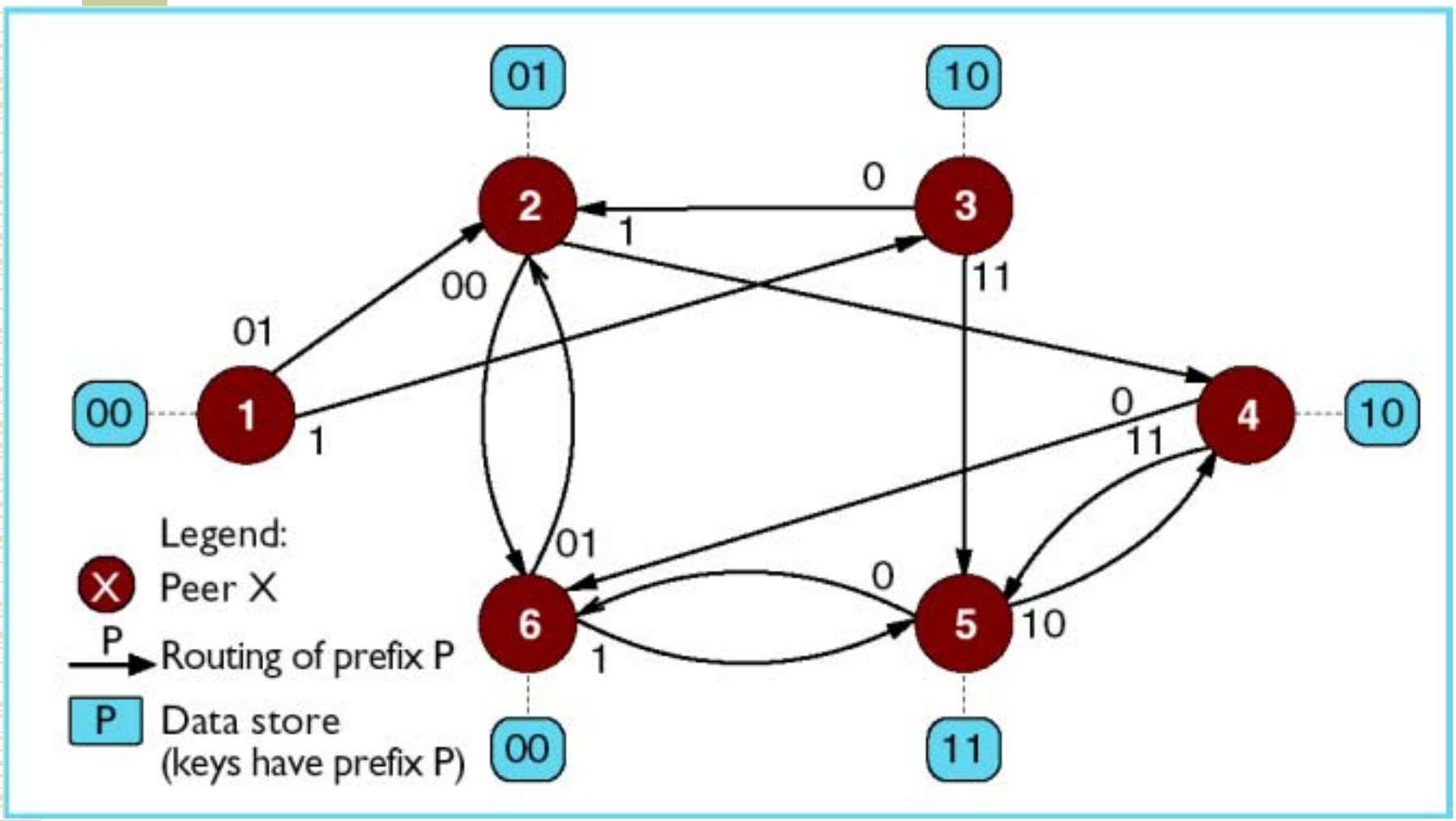
- Every peer maintains a small fragment of the database and a routing table
- The peers implement a routing strategy
- Replication can be used to increase robustness



Peer-Grid

- P-Grid is a **virtual binary search tree** that distributes replication over a community of peers and supports efficient search.
- Peers in P-Grid perform construction and search/update operations without any central control or global knowledge in an unreliable environment.


0
1
"Virtual binary search tree"
00
01
10
11
Query(6, 100)
Query(5, 100)
Query(4, 100), found!
1
6
2
3
4
5
1 :3
01 :2
1 :5
01 :2
1 :4
00 :6
0 :2
11 :5
0 :6
11 :5
1 :6
10 :4
Stores data with key prefix 00
Stores data with key prefix 00
Stores data with key prefix 01
Stores data with key prefix 10
Stores data with key prefix 10
Stores data with key prefix 11
Legend:
Peer X
P:X
Routing table (route keys with prefix P to peer X)
P
Data store (keys have prefix P)


01
10
2
3
0
1
00
11
01
00
1
1
4
10
0
11
Legend:
X
Peer X
P
Routing of prefix P
P
Data store
(keys have prefix P)
01
6
0
5
10
1
00
11

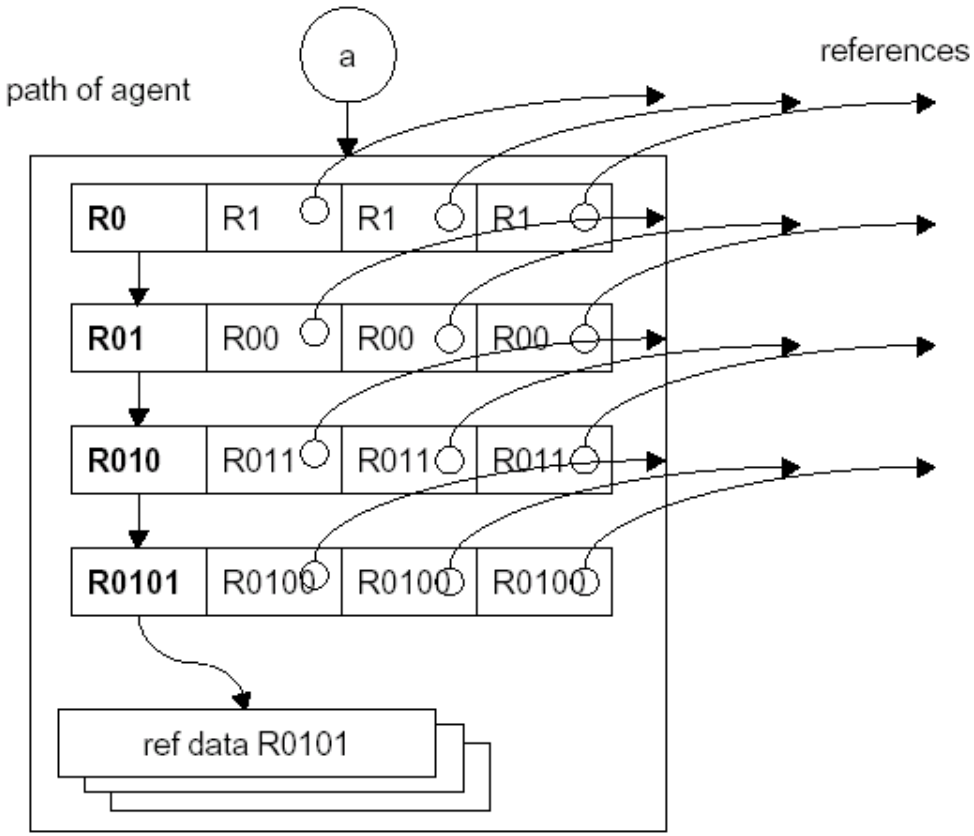
Peer-Grid Properties

- It is completely decentralized
- All peers serve as entry points for search
- Interactions are strictly local
- It uses randomized algorithms for access and search
- Probabilistic estimates of search request success can be given
- Search is robust against node failures
- It scales gracefully in the total number of nodes and data items

Routing

- Every participating peer is represented by its **path**, the binary bit string representing the subset of the tree's overall information that the peer is responsible for.
- For each bit in its path, a peer stores the address of at least one other peer that is responsible for the other side of the binary tree at that level.

Data Structure of an Agent



P-Grid Construction Algorithm (Bootstrap)

- When agents meet (randomly)
  - Compare the current search paths p and q
- Case 1: p and q are the same
  - If the maximal path length is not reached extend the paths and split search space, i.e. to p0 and q1
- Case 2: p is a subpath of q, i.e. q = p0…
  - Extend p by the complement of q, i.e. p1
- Case 3: only a common prefix exists
  - Forward to one of the referenced peers
  - Limit forwarding by *recmax*

0,0 -> 00,01

0,01 -> 00,01

00,011 -> r01,011 -> 010,011

4. Search and Update

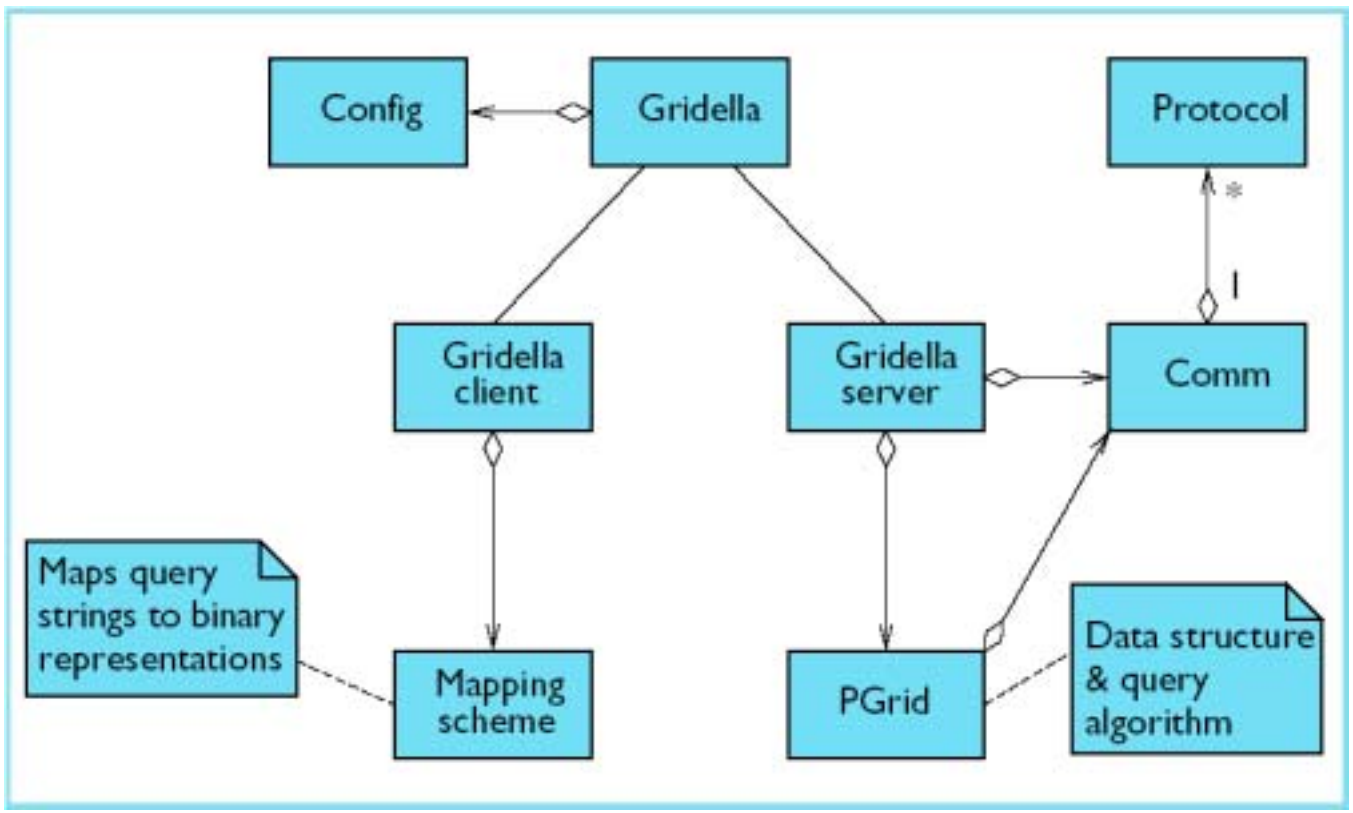
- Search straightforward
  - Follow own path or references
  - At most k steps
  - If multiple references are online, select randomly
- Updates
  - All replicas need to be found
  - Repeated searches
    - Breadth first (limited recursion breadth)
    - Depth first
    - Depth first and contact buddies with same key

k: key(path) length

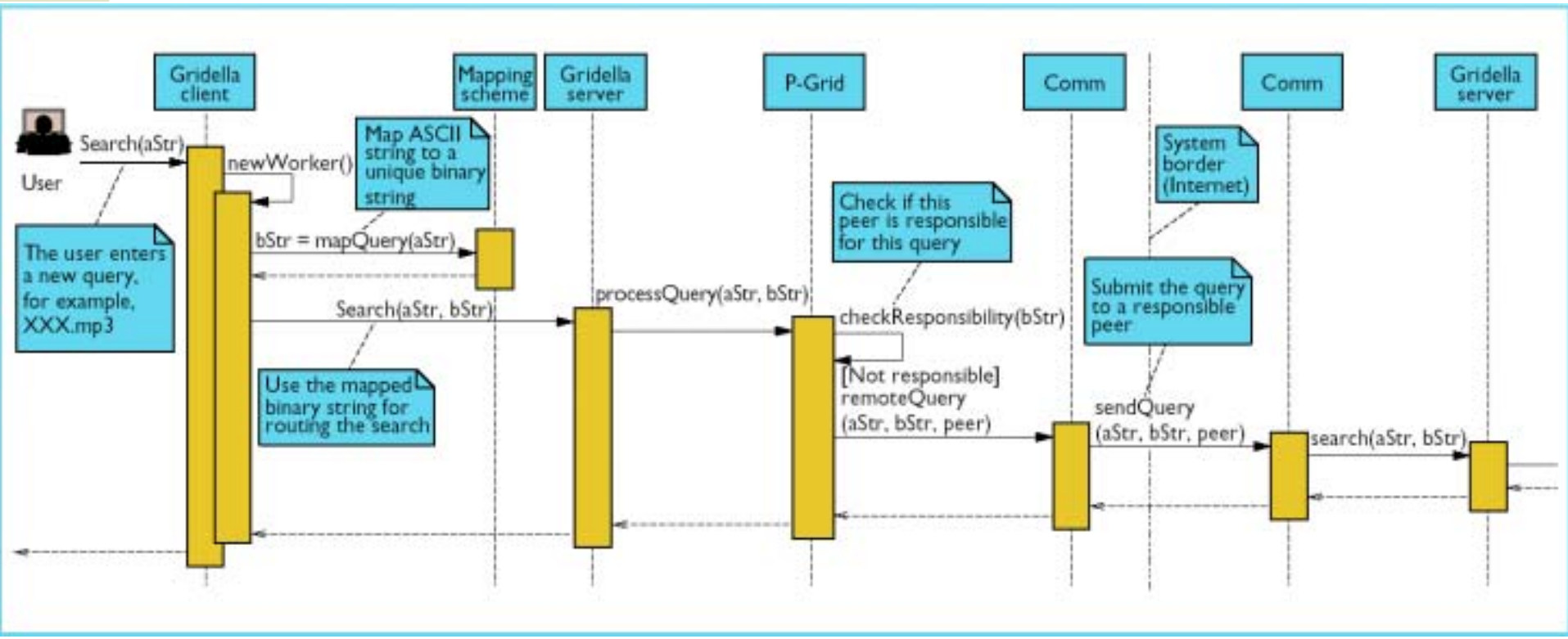
Gridella

- Gridella is a P-Grid-based, Gnutella-compatible P2P system.
- Gridella is written in Java and will be released under the GNU general public license.

Gridella Components



Gridella Interaction



Performance Comparison

Table 1. Performance comparison of Gridella and Gnutella.

Peers	Gridella messages	Gnutella messages
20,000	61	8,744
40,000	63	26,240
60,000	65	26,240
80,000	65	78,728
100,000	68	78,728
120,000	69	78,728
140,000	68	78,728
160,000	69	78,728
180,000	69	78,728
200,000	72	78,728

Future Works

- New e-commerce models:
  - Use P-Grid for decentralized storage of reputation data to address security issues such as authenticity and confidentiality.
- Micro-payment system
  - Force people to pay for service to avoiding free riding.
  - Offers and downloads from a peer would earn credits, which could be used to pay for services requested.

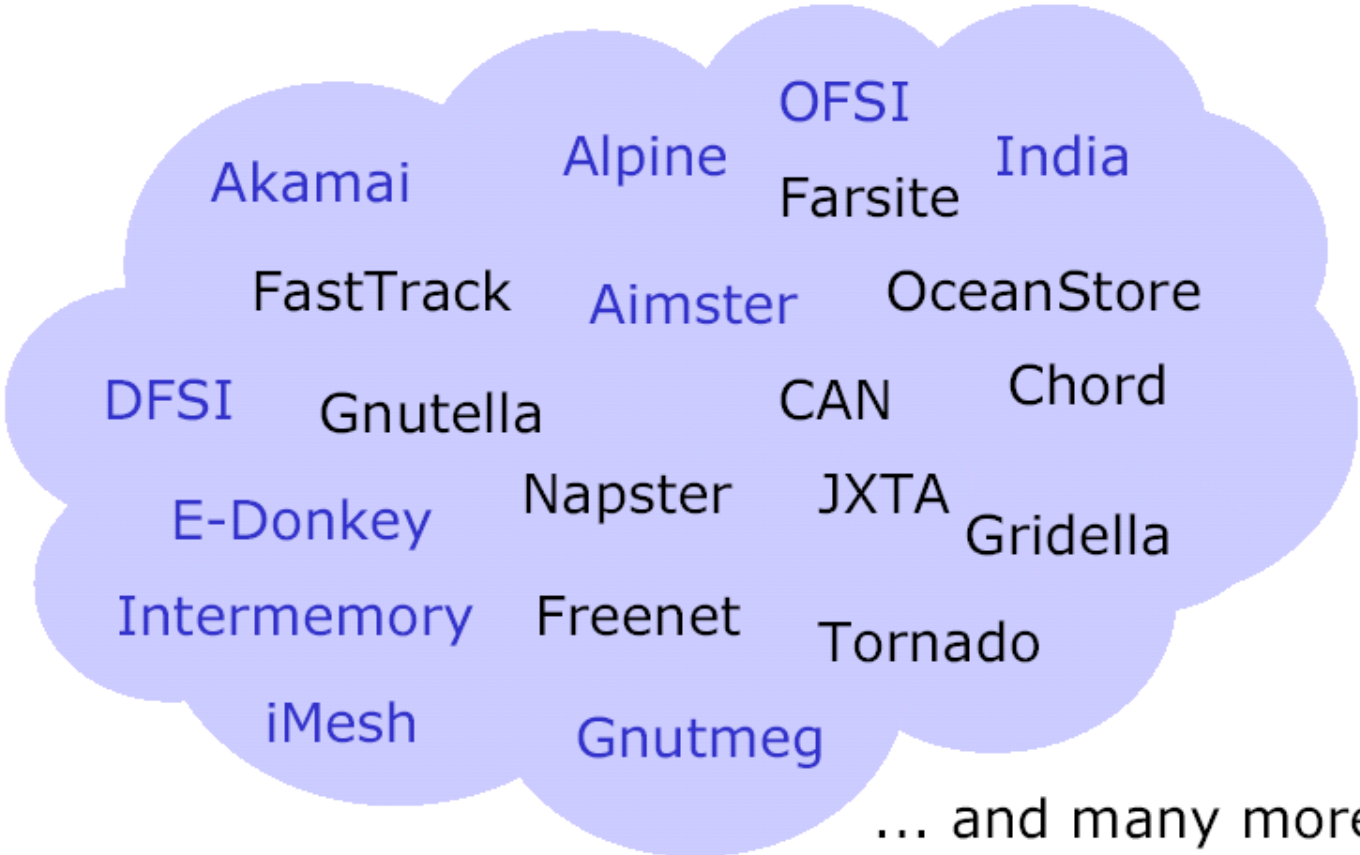
Research Issues

- P2P for reliable E-Commerce
  - dynamic business models
  - trust establishment
  - peer-to-peer transactions
  - decision making
- Quality of service
  - improved fault tolerance
  - archival storage
  - quality guarantees
- Richer data model
  - Relational, XML
  - meta-data model
  - improved search
- Multimedia (MPEF7, MPEG2000)
- Message-based applications
  - scalability
  - improved search capabilities
- Mobility
- P2P Search Engines

Conclusion Remarks

- Considerable research and experimentation remains to make P2P systems feasible for application domains beyond mere MP3 and image exchange, such as creating a new paradigm for decentralized **e-commerce systems**, or for new types of network infrastructures such as **mobile ad hoc networks**.

2. The P2P Cloud



Summary and Comparison of Approaches

	Paradigm	Search Type	Search Cost (messages)	Fixed Data Assignment	Common Network Origin
Gnutella	Breadth-first search on graph	String comparison	$2*\sum_{i=0}^{TTL} C*(C-1)^i$	no	no
Freenet	Depth-first search on graph	Equality	O(Log n) ?	no	no
Chord	Implicit binary search trees	Equality	O(Log n)	yes	yes
CAN	d-dimensional space	Equality	O(d n^(1/d))	no	yes
Tapestry	Prefix trees	Equality	O(Log n)	yes	no ?
P-Grid	Binary prefix trees	Prefix	O(Log n)	no	no

Listen — P2P is around

- P2P systems receives currently a lot of attention
  - File-sharing systems
    - Napster, Gnutella, Freenet, etc.
  - Conferences
    - O'Reilly P2P conference 2001 (conferences.oreilly.com/p2p/)
    - 2001 International Conference on Peer-to-Peer Computing (P2P2001) (www.ida.liu.se/conferences/p2p/p2p2001/)
    - Peer-to-Peer Computing Workshop at Networking 2002 conference (http://www.cnuce.pi.cnr.it/Networking2002/)
    - March-issue 2002 Scientific American: D. Anderson, J. Kubiatowicz: The World-wide Computer
    - etc.
- P2P is nothing new – see Arpanet