

Self-organisation in P2P Systems

Dipl. Inf. Ivan Dedinski

University of Passau
Institute of Computer Networks and
Computer Communications, Prof.
Dr. H. de Meer

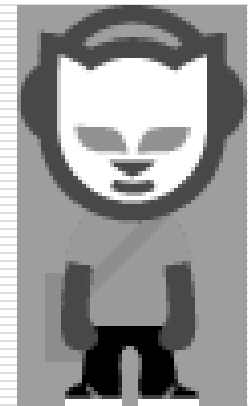


Overview

- The successful story of P2P
 - n 1st, 2nd and 3rd generation P2P – from Gnutella and Napster, over eDonkey, to Chord
 - Research on 3rd generation P2P
 - P2P in new network environments – Mobile P2P, MoPi
 - Controlling P2P – EPSRC P2POpt
-

The successful story of P2P - Napster

- Napster - the original P2P application that popularized the P2P concept to millions.
 - Based on central server
 - Indexed all the files that each Napster user had
 - Users searched for files, using the server
 - Found files were downloaded directly from the user computers.
-



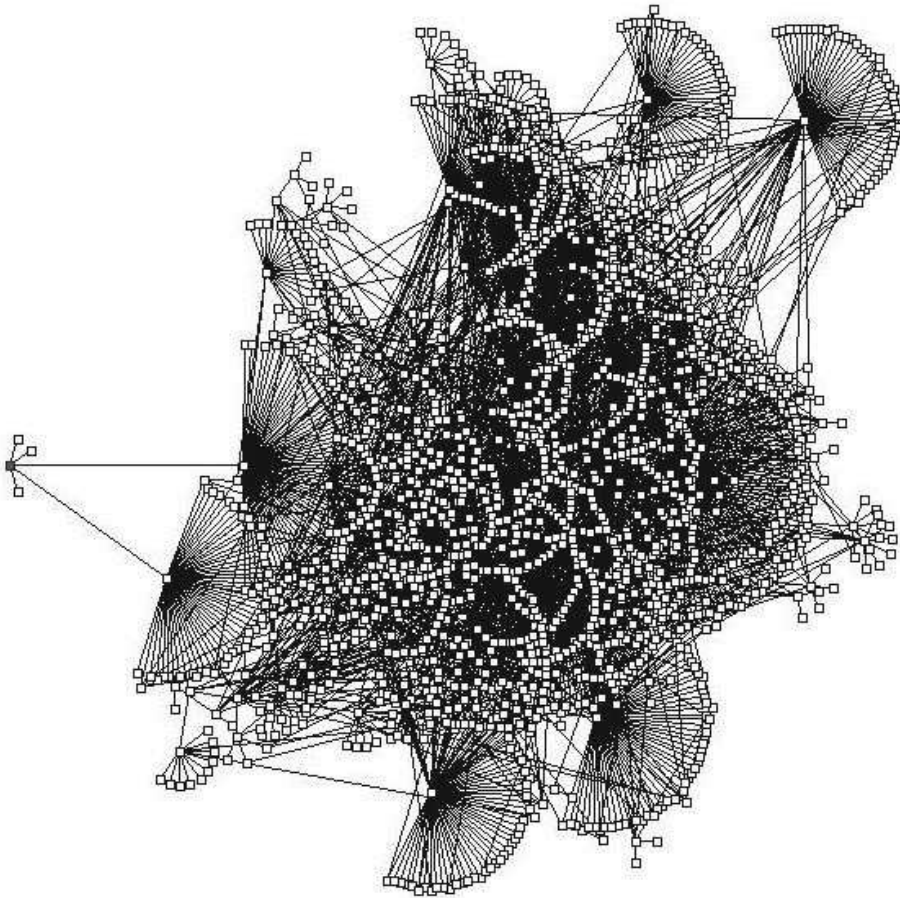
The successful story of P2P – Napster II

- The central server approach was an efficient way to organize the network, but:
 - n It was not fault tolerant
 - n It was attackable by hackers
 - n It was attackable by law
 - >Napster was shut down by the music industry.
-

The successful story of P2P-Gnutella

- Suddenly, all Napster users needed a new file sharing network!
 - Gnutella was designed as a network without any central instance.
 - n A kind of self-organisation was needed
 - n Simplicity was very important
 - > "Chaotic" self-organisation, flooding was used for searching the content
-

The successful story of P2P- Gnutella II

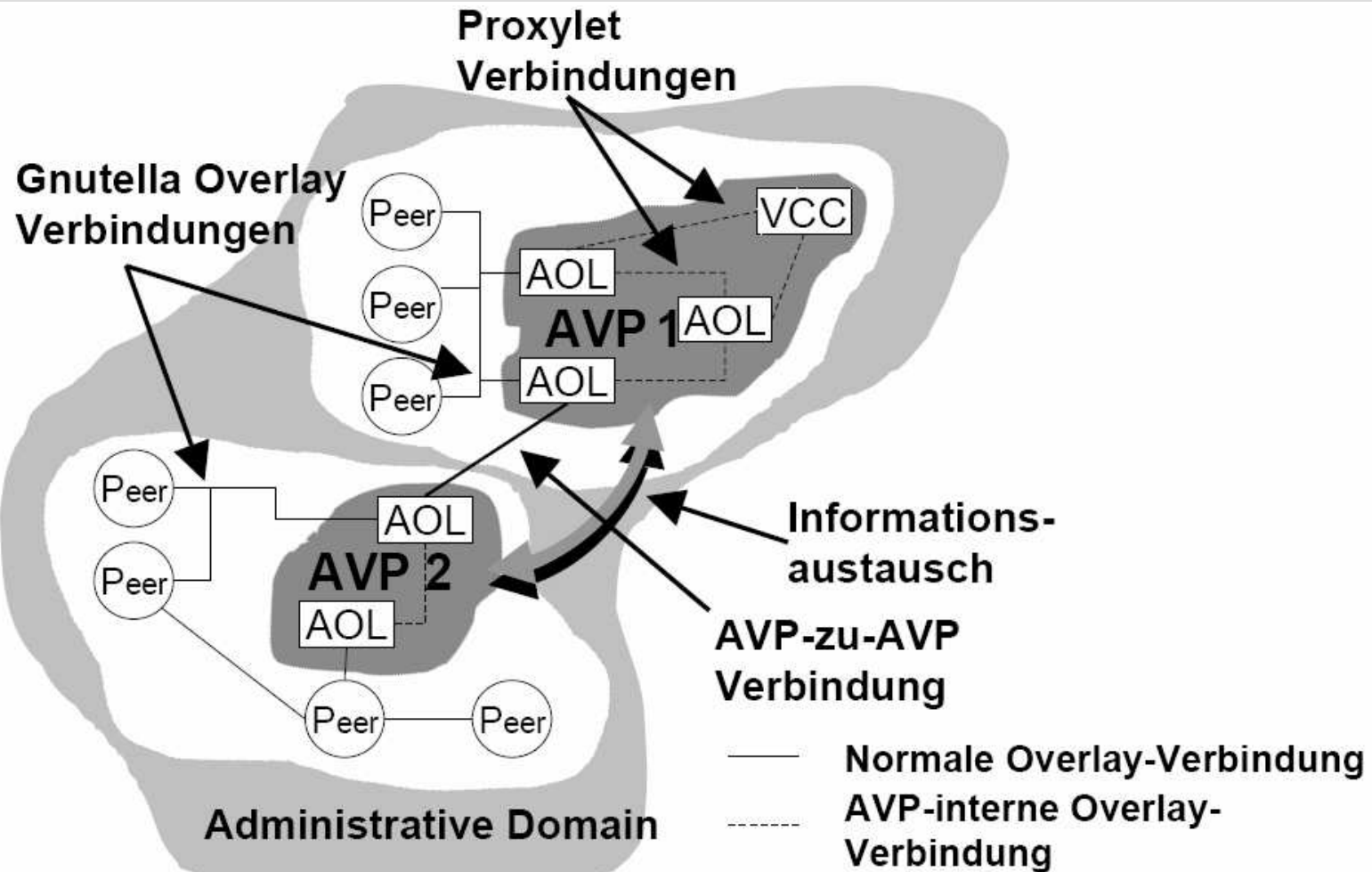


- Each Gnutella peer chooses his neighbours autonomously
- Peers send periodic keep-alive messages
- Hundreds of connections assure high connectivity
- Joining and leaving the network is simple
- But:
 - Flooding is very inefficient
 - Keep-alives overload the network
 - > With the growth of Gnutella, the signalling traffic becomes larger than the download traffic!!!

The successful story of P2P- Optimizing Gnutella

- As in the real world, some peers are more equal than others:
 - n Ultrapeers
 - n Active Virtual Peers (AVPs)
 - <http://www.acu.rl.ac.uk/msn2002/Talks/HermannDeMeer.pdf>
 - Bring intelligence into the chaotic network, without changing the protocol
 - n Interest Groups, based on:
 - Location
 - Content

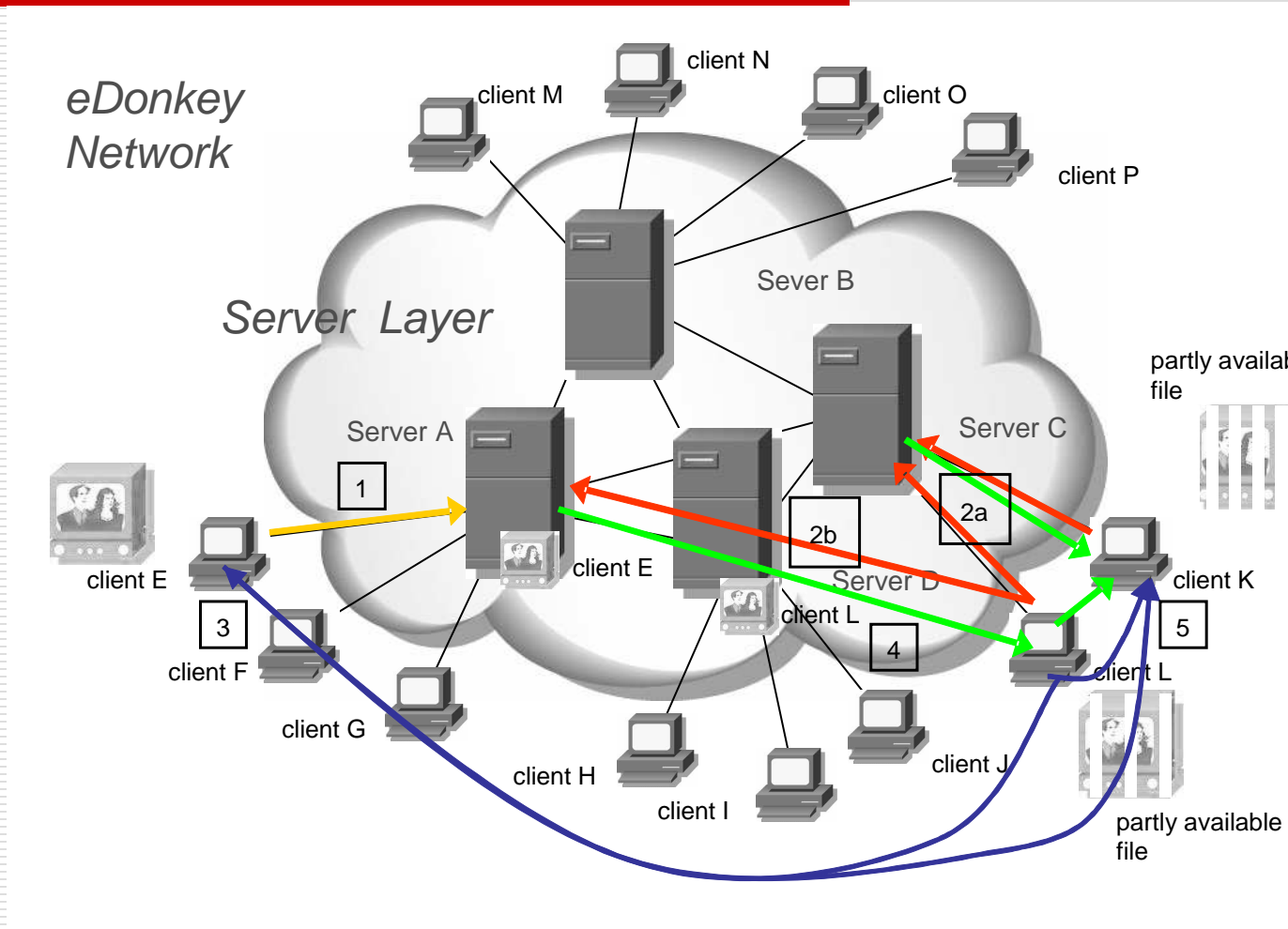
The successful story of P2P-Active Virtual Peer



The successful story of P2P- eDonkey

- Napster was shut down by law, Gnutella is shutting down itself, what to do?
- Compromise between Napster and Gnutella: eDonkey, or hybrid file sharing
- Consists of “many Napsters”
- Hard to shut down, because many central instances exist.
- Efficient, because not completely decentralized.

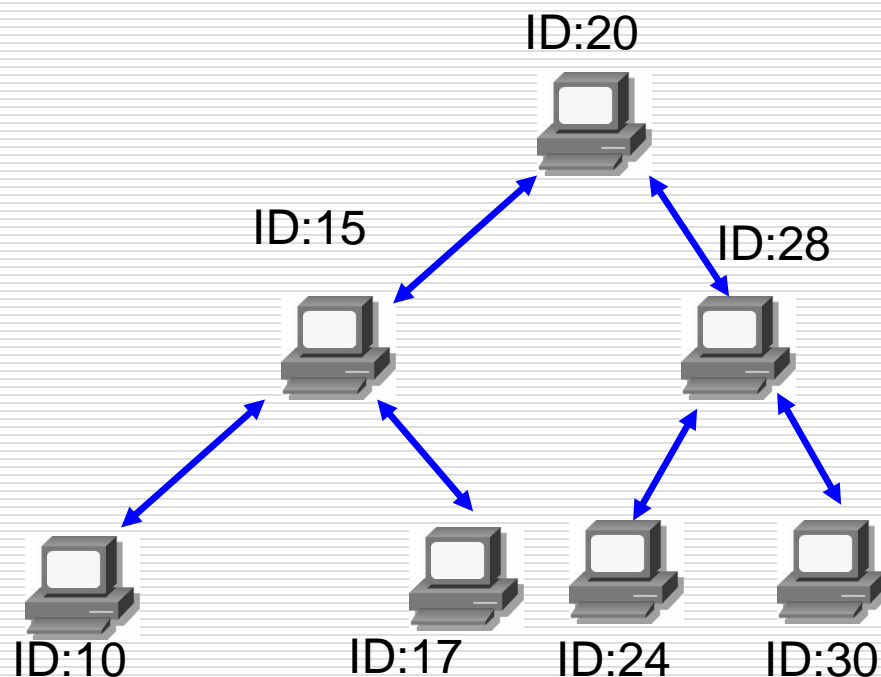
The successful story of P2P-eDonkey II



The successful story of P2P-DHTs

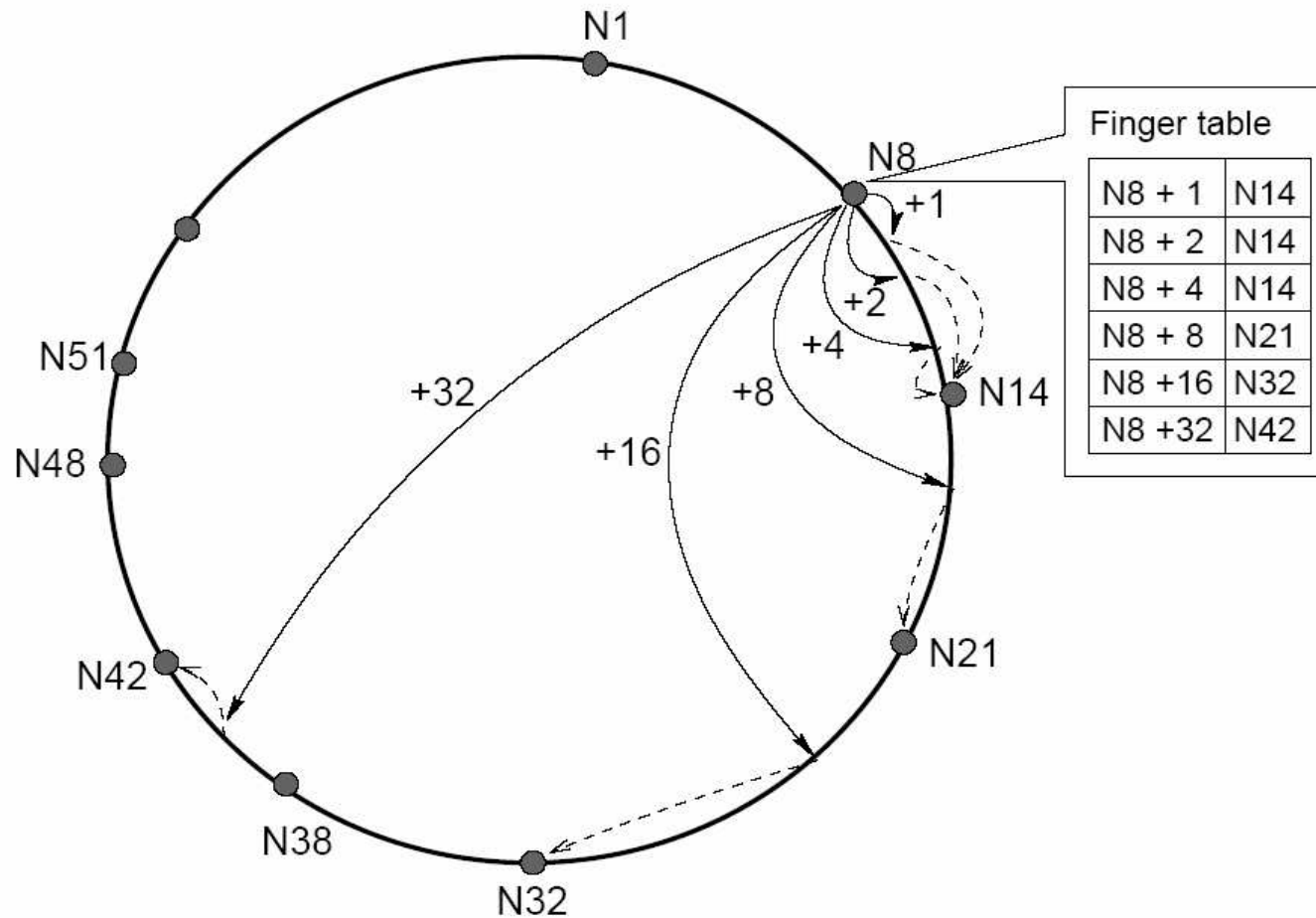
-
- Up to now, all discussed systems were focusing peer autonomy.
 - One single peer does not need to obey any connection rules
 - No real “superior” structure exists.
 - DHT – use big number of peers to build up a strongly ordered P2P system.
-

The successful story of P2P-DHT Simple Example

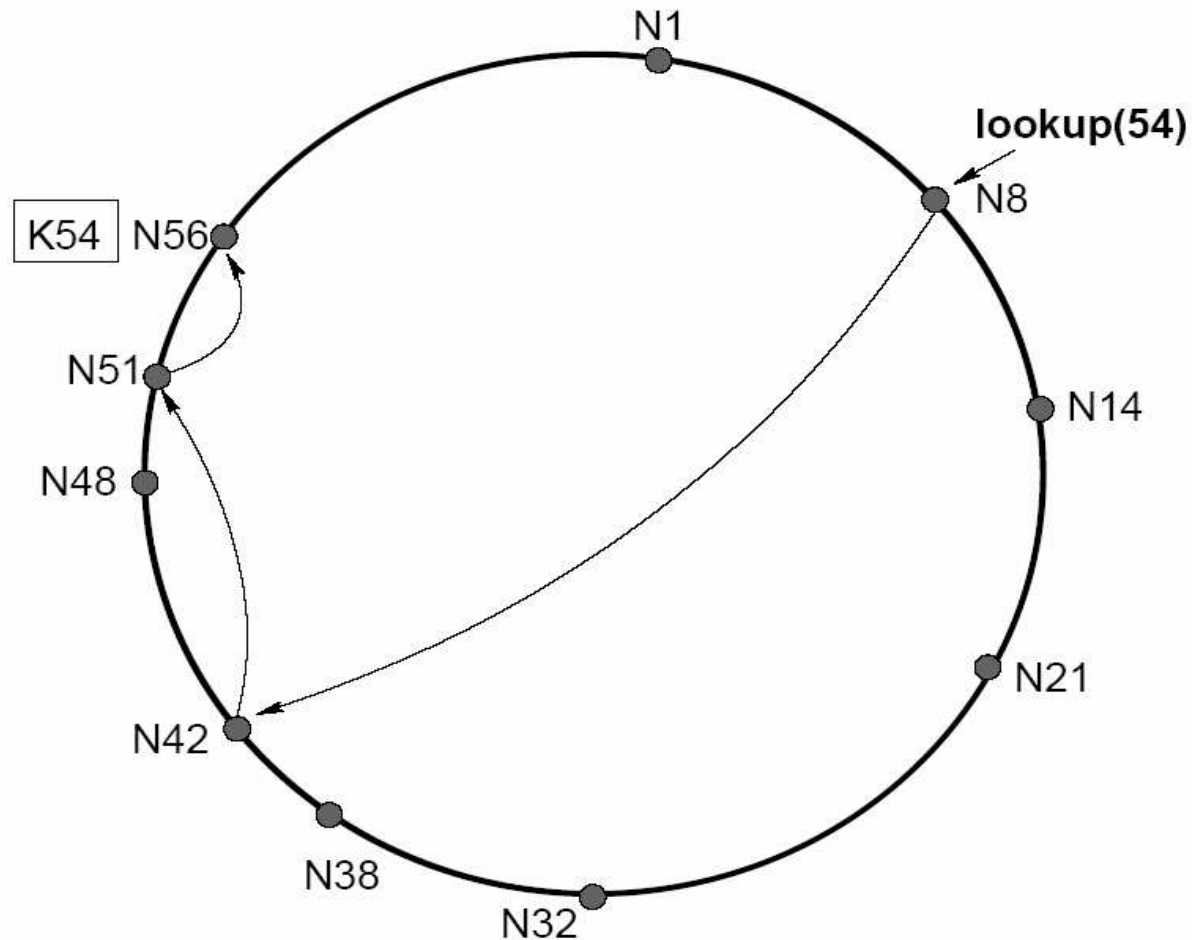


- Each peer has an unique ID
- Each piece of content has also an ID
- Each peer is allowed to connect only to 3 other peers – 2 children and one parent
- A peer is responsible for the content with ID < than his own ID and > than the ID of the parent
- Efficient, but not fault tolerant

The successful story of P2P-Chord



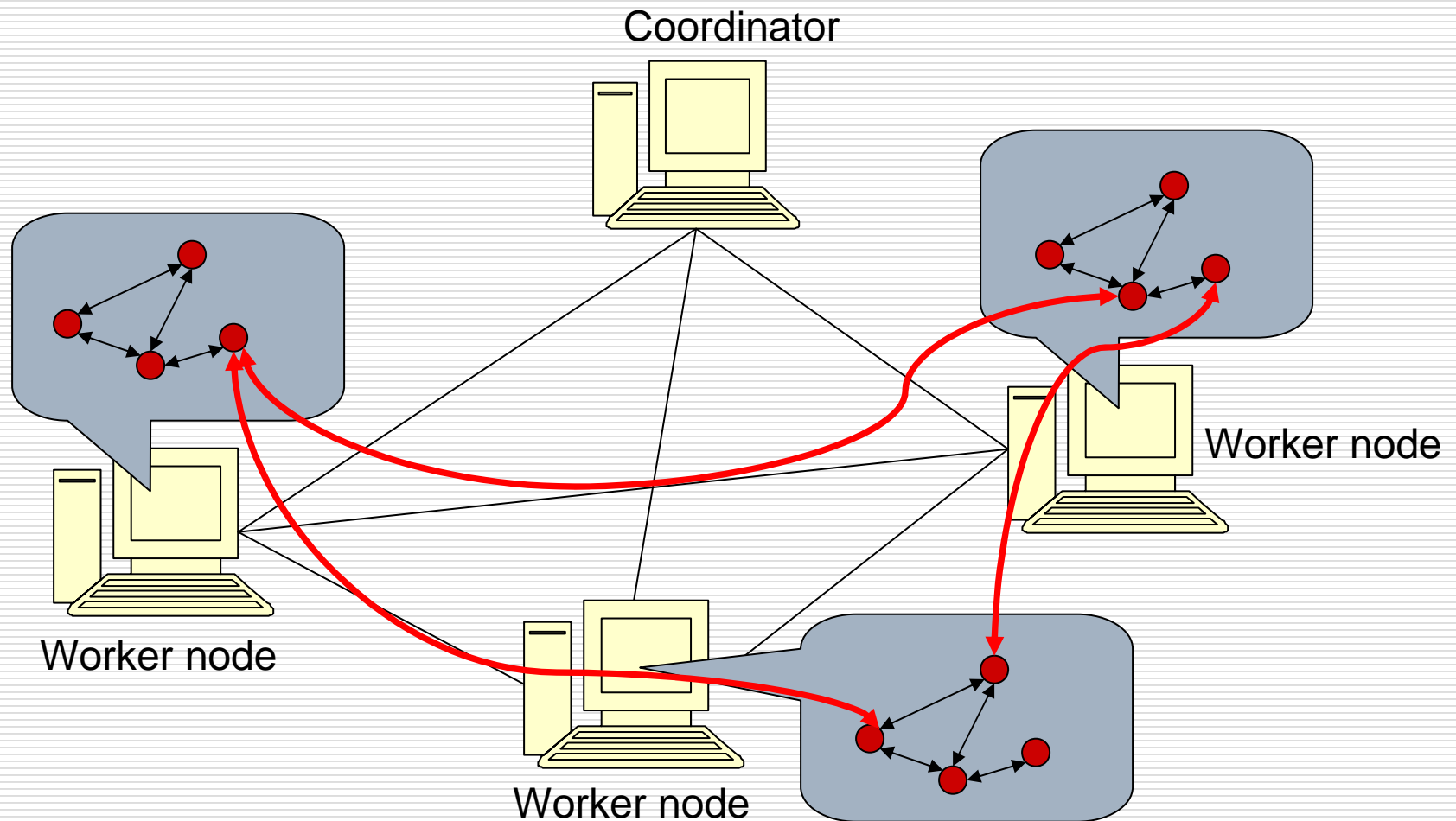
The successful story of P2P- Chord II



Research on 3rd generation P2P

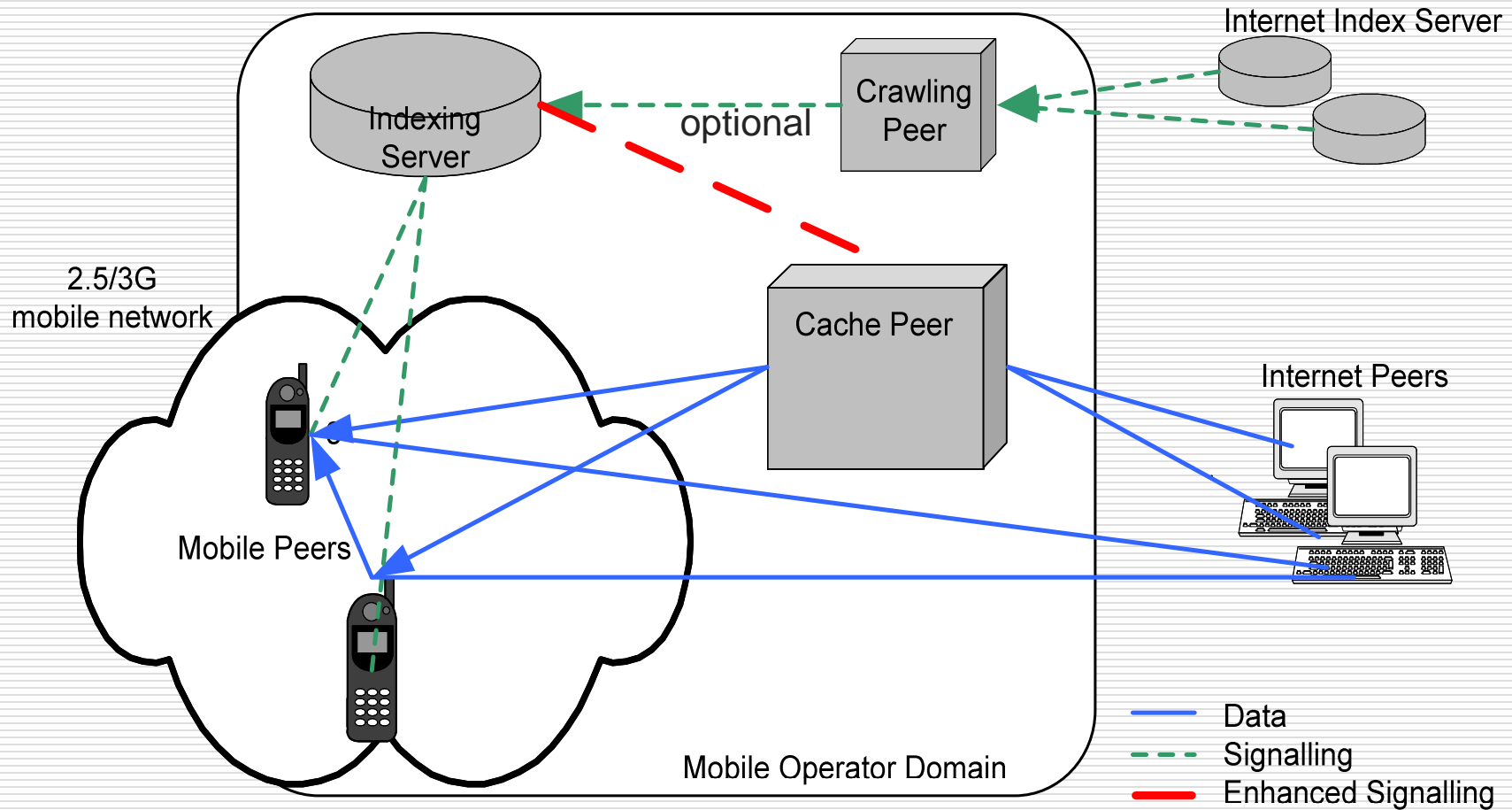
- Only exact match queries possible
 - Physical network properties are not taken into account
 - Load balancing may not be optimal, frequent queries could overload single nodes
 - Relatively complex joining and disconnecting from the network
 - Currently no replication concept
- n > for all these topics we have ideas and ongoing research activities.

The JNS Network Simulator – Parallel Extension



P2P in new network environments

Mobile P2P, MoPi



P2P in new network environments

Mobile P2P, MoPi II

Contributors:

Prof. Herman de Meer, Ivan Dedinski, Jens Oberender
Universität Passau
Department of Computer Science and Mathematics



Dr. Kurt Tutschku
Bayerische-Julius-Maximilians Universität Würzburg
Department of Distributed Systems



Frank-Uwe Andersen, Dr. Cornelia Kappler
Siemens ICM

Controlling P2P – EPSRC P2POpt:

- P2P is currently rather chaotic and hard to control. Many different protocols exist.
- P2P protocols are changing rapidly.
- P2P traffic is currently more than 70% of the total Internet traffic.
- P2P is making impossible for the network providers to sell resources, which do not exist in fact.

Controlling P2P – EPSRC P2POpt II

- Need for controlling P2P traffic by the network operators emerges.
 - To be controlled, P2P traffic has to be identified first. This is a hard task, because P2P protocols are flexible – tunneling, proxies, etc.
 - Classification based on pattern recognition possible.
 - After P2P traffic has been identified, it can be controlled and optimized – AVP, caches, proxies, etc.
-

Controlling P2P – EPSRC P2POpt III

Contributors:



Engineering and Physical Sciences Research Council



Prof. J. Sventek



Prof. D. Hutchison, Dr. L. Mathy



Advancing and supporting the
UK's education and research network



Prof. Dr. Ing. Hermann De Meer, Ivan Dedinski

<http://www.dcs.gla.ac.uk/p2popt/>

Conclusion

- P2P is a relatively new concept and research area.
 - There are many open questions and challenging topics.
 - The concept of self-organisation has proven to be promising, but still has to be further optimized. However this does not seem impossible.
-