# P2P Traffic Simulation

## 4. Würzburger Workshop

## "IP Netzmanagement, IP Netzplanung und Optimierung"

## Würzburg, Germany

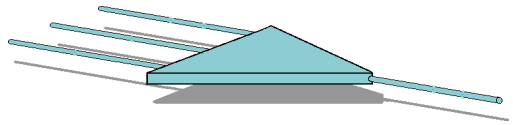
## 27-28 July, 2004

Netzplanung

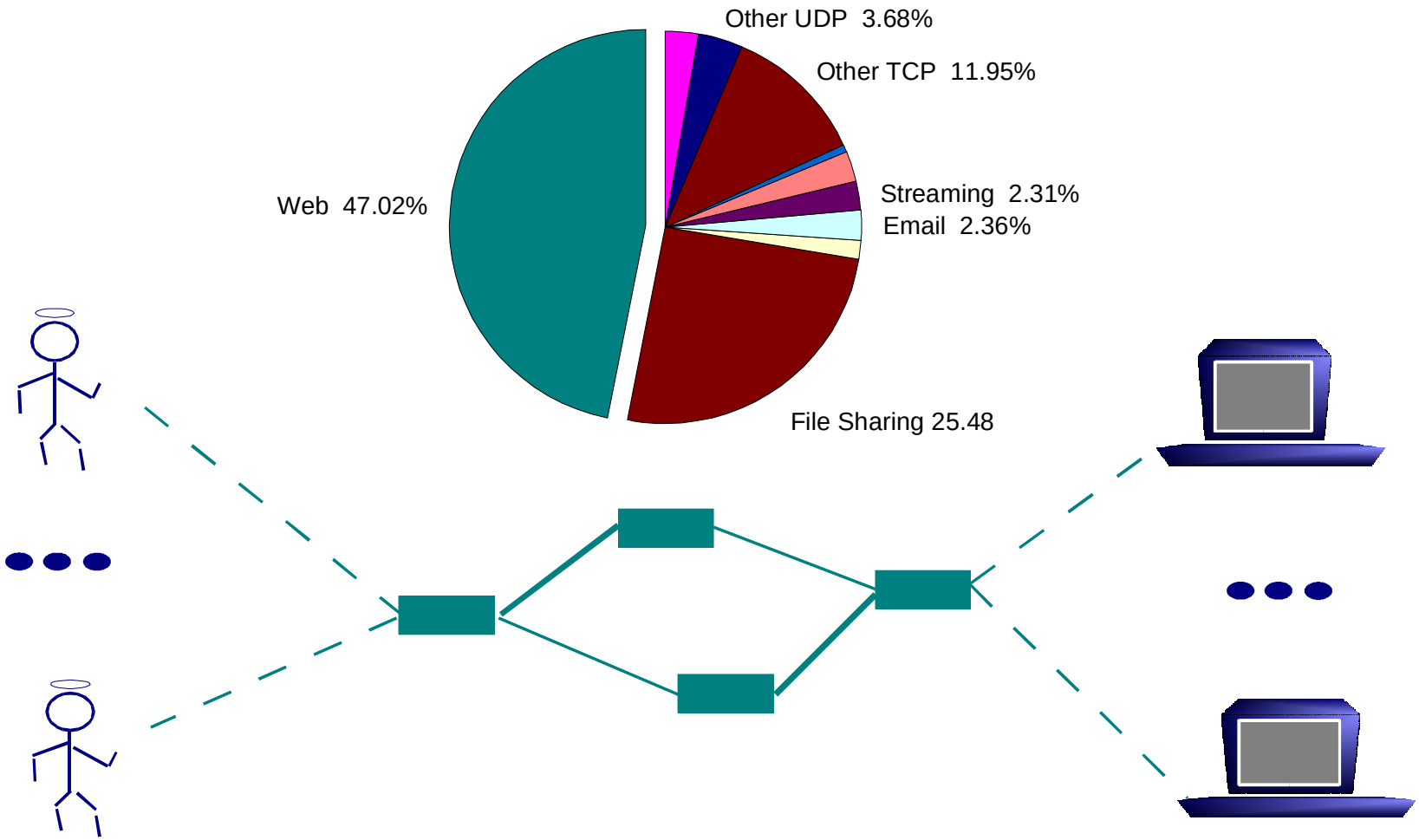
Technische Universität Hamburg-Harburg

*Kommunikationsnetze*

*Sireen Malik[1], Prof. Dr. Ulrich Killat*



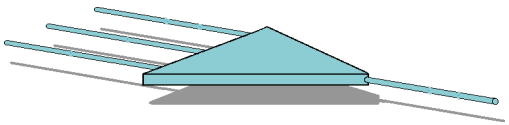
# Internet Traffic



# Why Gnutella 0.4?

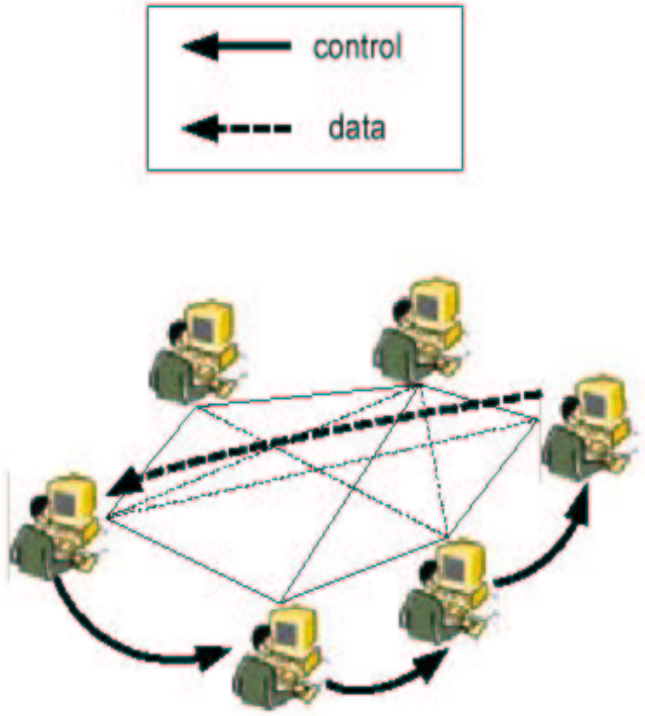
- Kazaa vs. Gnutella
  - Kazaa is dominant but is based upon **FastTrack** proprietary protocol.
- Gnutella 0.4 vs. Gnutella 0.6
  - The new Gnutella 0.6 is closer to Kazaa.
  - Gnutella 0.4 statistics are widely available.





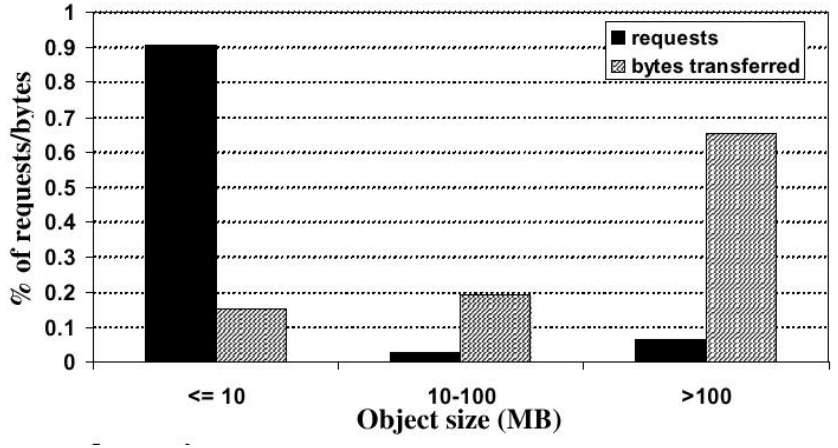
# Gnutella 0.4

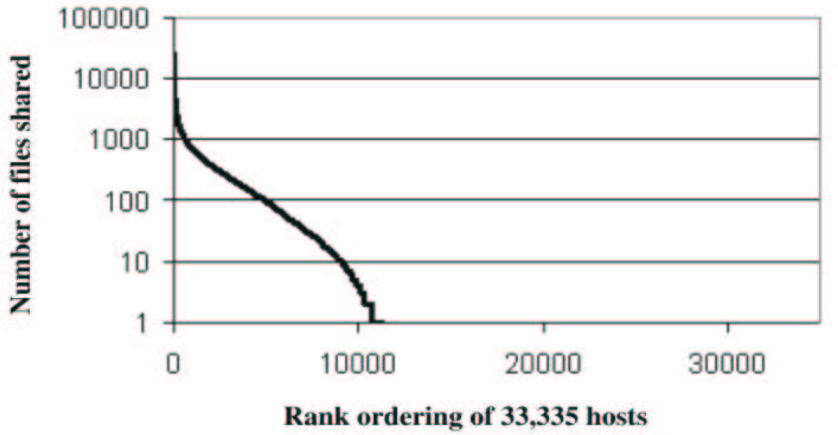
- Distributed Index Flooding Architecture (DIFA)
- Signalling
  - Attempt connection with Bootstrap servers
  - Once connected, **flood** PINGS
  - Reverse routed PONGS from known-hosts
  - PONG-based connection to known hosts
  - **Flood** QUERIES to all known-hosts
  - Reverse routed QUERYHITS
  - File download via HTTP-TCP



# User Behavior Modelling (I)

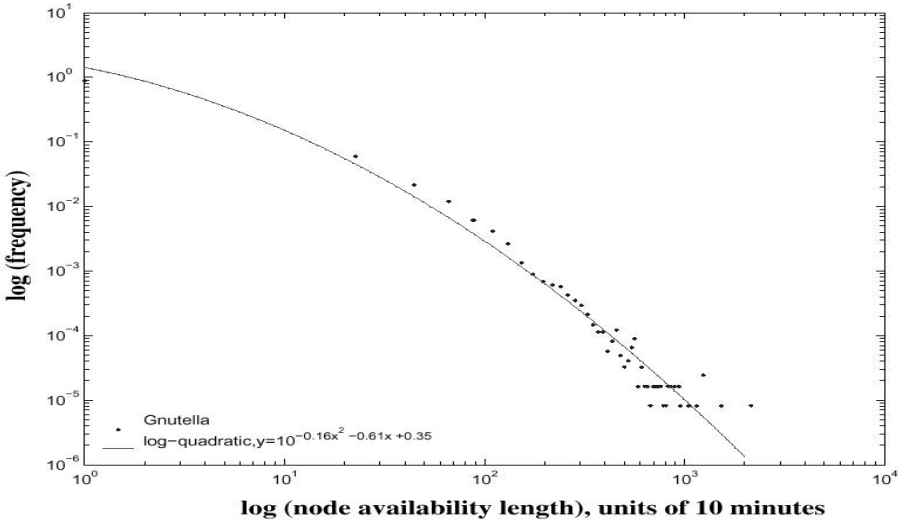
- File size distribution [2]:
  - for < 10MB files (music) less than 15% bytes but more than 88% request
  - for >100MB files (movies) more than 65%  bytes but less than 5% request
- Number of files shared by peers [3]:
  - 66% free-riders
  - 73% share 10 or less files
  - top 1% share 35% files



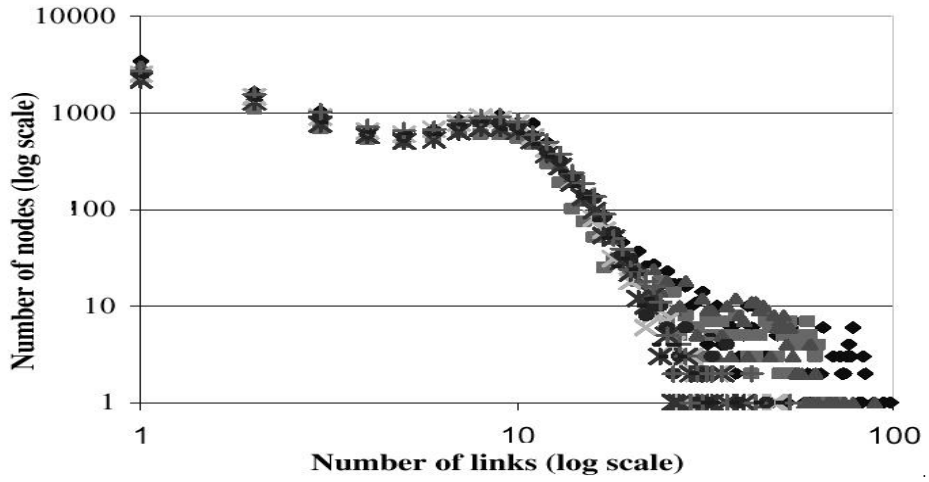



# User Behavior Modelling (II)

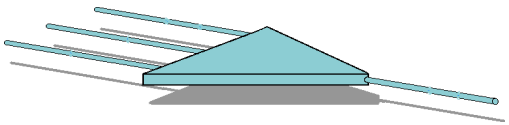
- Peer uptime [4]:
  - 30% for 10 Minutes
  - 20% for > 2 Hours




- Node Connectivity [5]:
  - power law for nodes with more than 10 connections

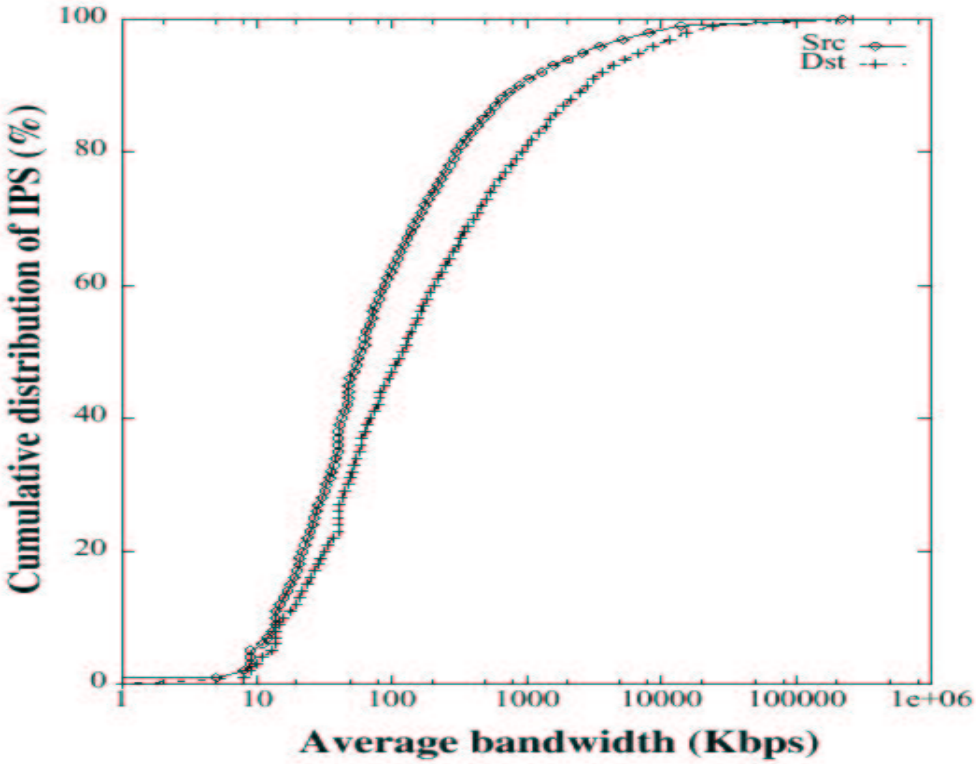




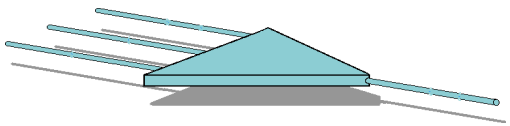
# User Behavior Modelling (III)

- Access speeds [6]:
  - measured for Kazaa
  - 66% are modem users



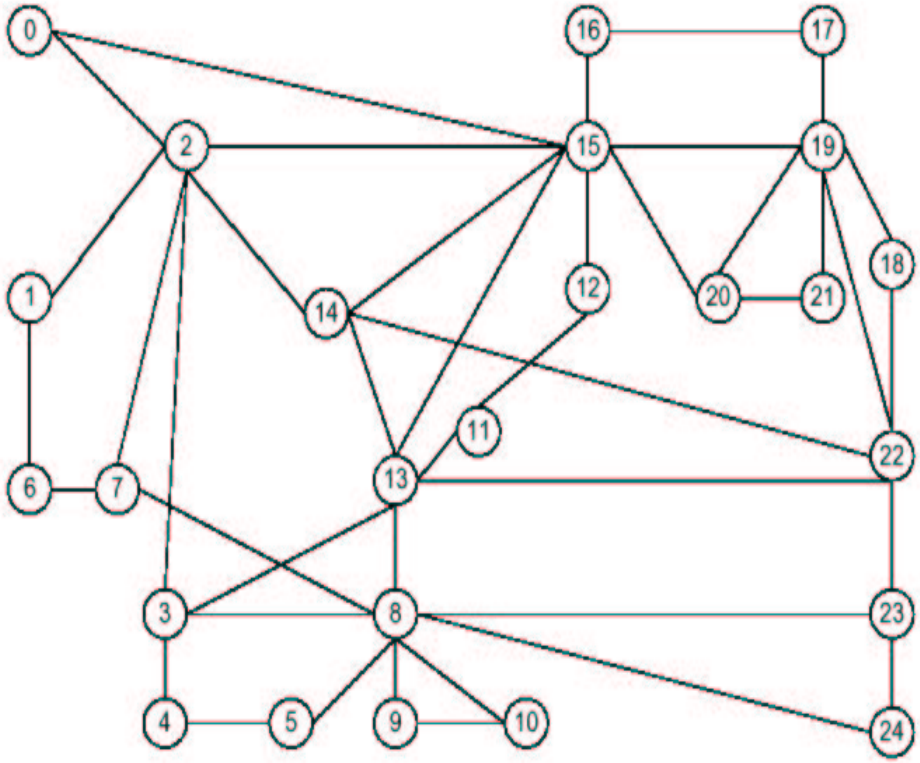
# Implementation Details

- Implementation in NS2
- At start-up user access speeds adapted according to access-speeds distribution
- Maximum number of connections adapted according to node-connectivity distribution (e.g. Modem-user = 10 conn.)
- File-size distribution implemented globally; each peer reads from it on start-up by first picking a number from number-of-files distribution.
- File-popularity distribution also implemented globally; each peer reads from it before generating a Query.
- Peer keeps On for a time from peer-uptime distribution.

# Simulation

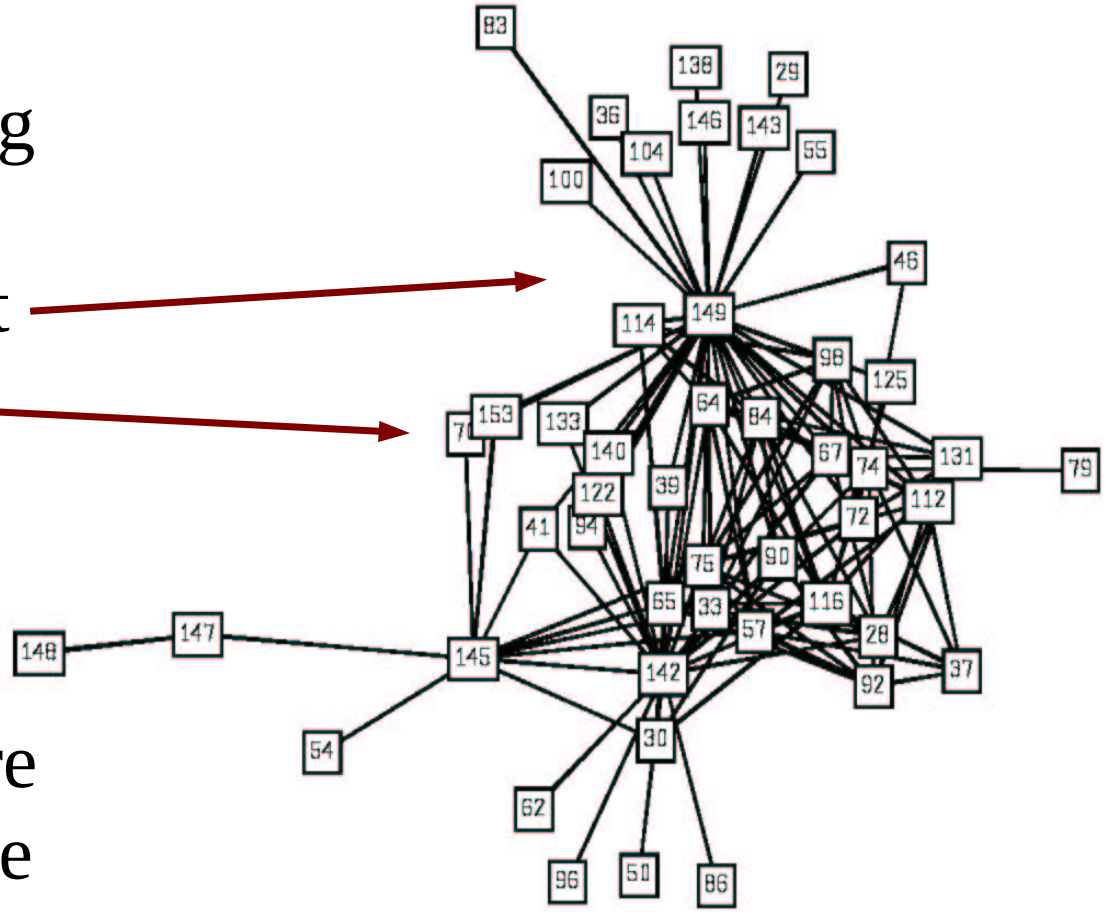
- 25 nodes ATT&T network
- The number of peers attached to each core node is picked randomly from a Perato-distribution (shape parameter=1.5)
- Total number of *M*= 1000 peers.




Netzplanung

# Results(I)

- Snapshot of an evolving topology:
  - Star topology at start
  - Evidence of Pong-based connections.
- Stable topology had power law overlay network $f_d \propto \frac{1}{d^{1.05}}$, where $d$ is degree and $f_d$ is the frequency of $d$.



# Results(II)

- Stable simulation:

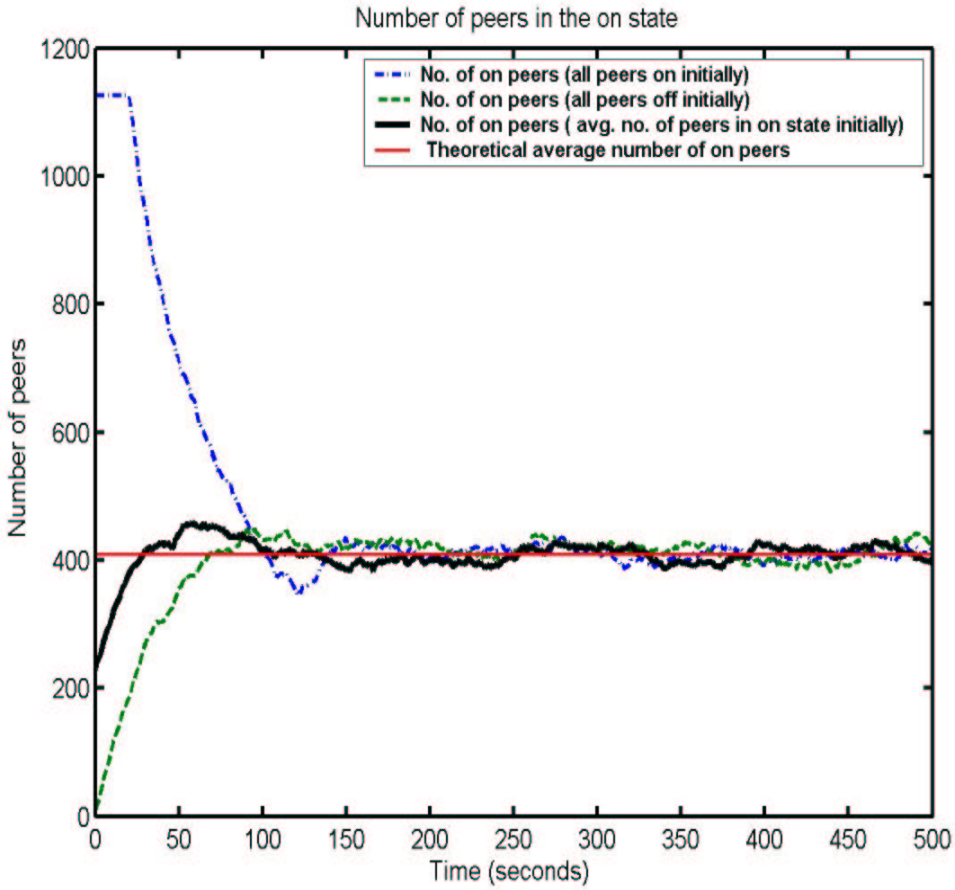
$$E[on\ peers] = \frac{M * E[T_{on}]}{[T_{on}] + [T_{off}]},$$

*where,*

$$E[T_{on}] = average\ peer\ on\ time$$

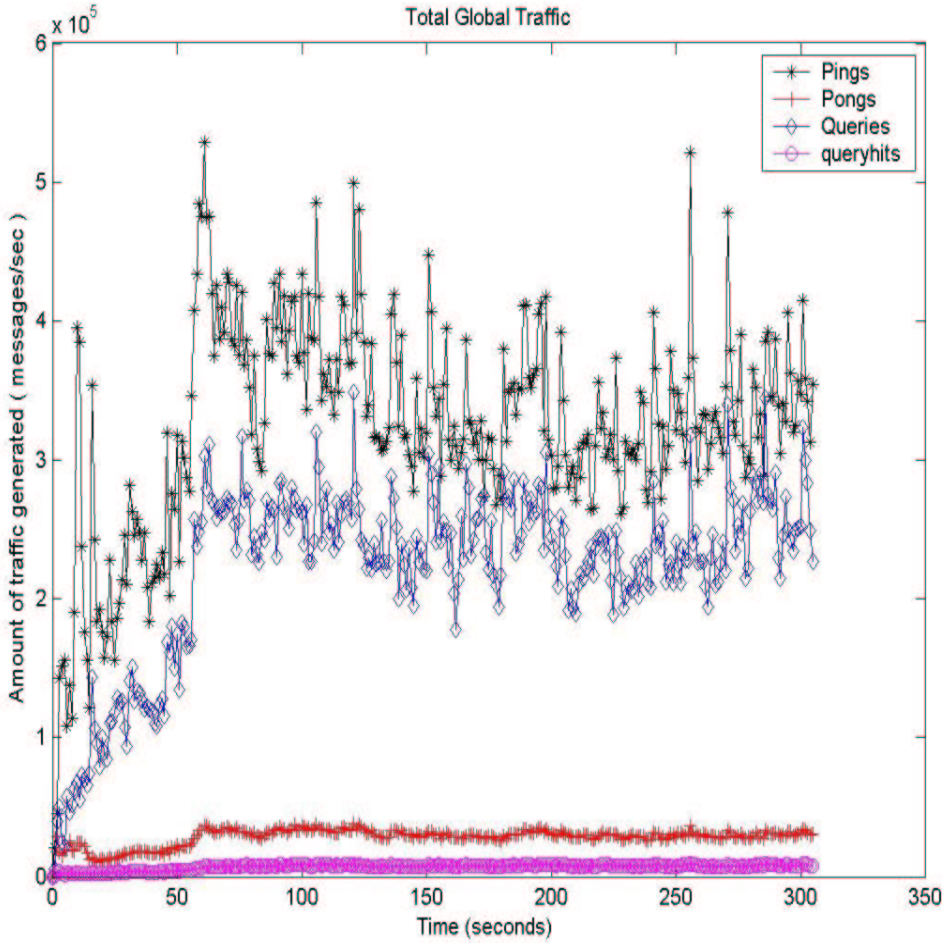
$$E[T_{off}] = average\ peer\ off\ time$$

- A start with average number of peers yields the fastest simulations.



# Results(III)

- Ping and Query traffic is high because of the flooding/broadcast.
- Shows strong correlation with average number of active peers.

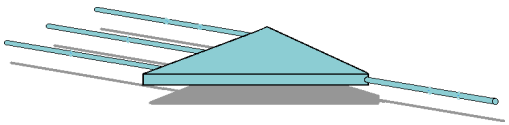



Netzplanung

# Conclusion

- User behavior was identified and implemented.
- Gnutella 0.4 implemented.
- Some results verified.
- Road open for more studies.
- Future work: Gnutella 0.6.





# References

[1] *Sireen Malik is a member of the Peer-to-Peer Task Force at the department of Digital Communication Networks at the Technical University of Hamburg-Harburg, Germany. Carsten Bürmeister and Jan Kraaier are the other two members of the team.*
[2] "Measurements, Modeling, and Analysis of a Peer-to-Peer File-Sharing Workload", Krishna P. Gummadi, Richard J. Dunn, Stefan Saroiu, Steven D. Gribble, Henry M. Levy, and John Zahorjan, Department of Computer Science and Engineering, University of Washigton.
[3] "Free Riding on Gnutella", E.Adar and B.Huberman, 2000
[4] "Availability and Locality Measurements of Peer-Peer Files", J.Chu, K.Lebonte, B. Levine, 2002.
[5] "Mapping the gnutella network: properties of large scale peer-to-peer systems and implications", Matei, et.al.
[6]"Analyzing peer-to-peer traffic across large networks", S.Jen. J.Wang, IMW 2002