Hierarchical Structures in Peer-to-Peer Overlay Networks

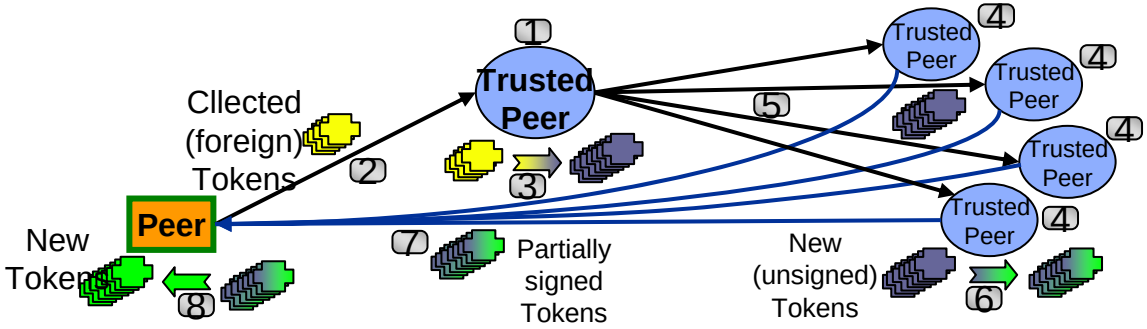
5. Würzburger Workshop
"IP-Netzmanagement, Netzplanung und Optimierung"
18.-19. Juli 2005

Hannes Birck [1] Katrin Tschirpke [2]

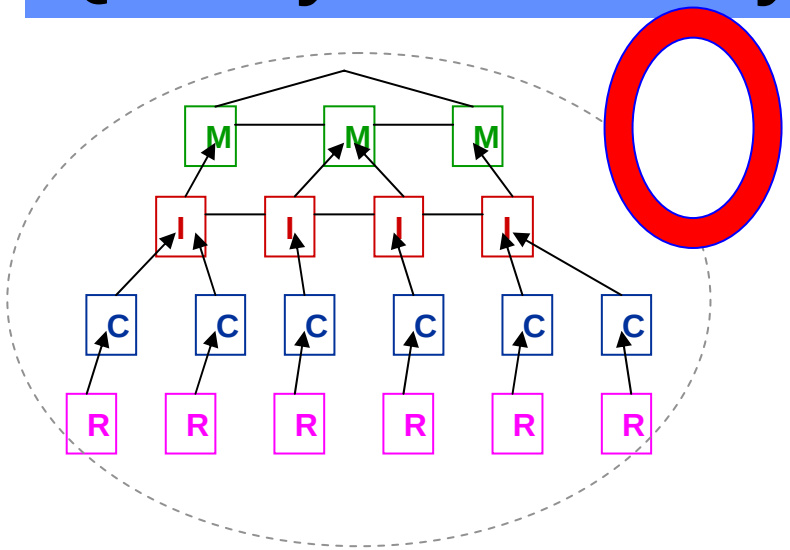
1 ***TU Darmstadt - Darmstadt University of Technology,***
Dept. of Electrical Engineering and Information Technology
KOM - Multimedia Communications,

2 ***FH Aschaffenburg- University of Applied Sciences Aschaffenburg,***
Mathematics, Statistics and Operations Research

Accounting in Autonomous Decentralized Systems



Quality in Overlay Networks



OMICON

A role based P2P protocol

Outline

I. Introduction

II. Analysis of hierarchical overlay structures

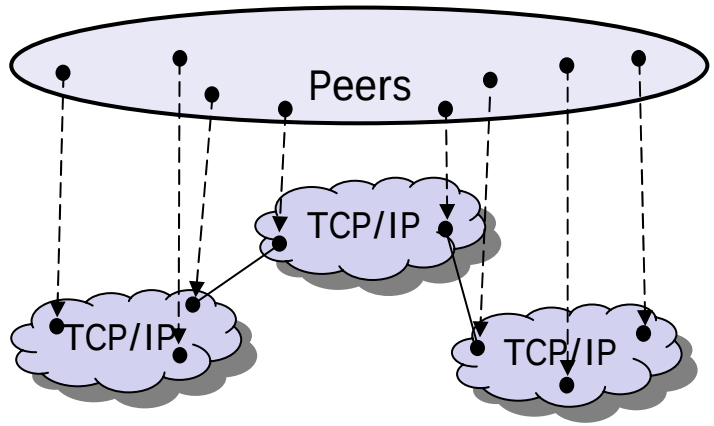
- Application Layer Simulation
- Analysis of the FastTrack Protocol
- The Multi-Layer-Overlay-Protocol

III. Summary and Conclusion

P2P Overlay Networks

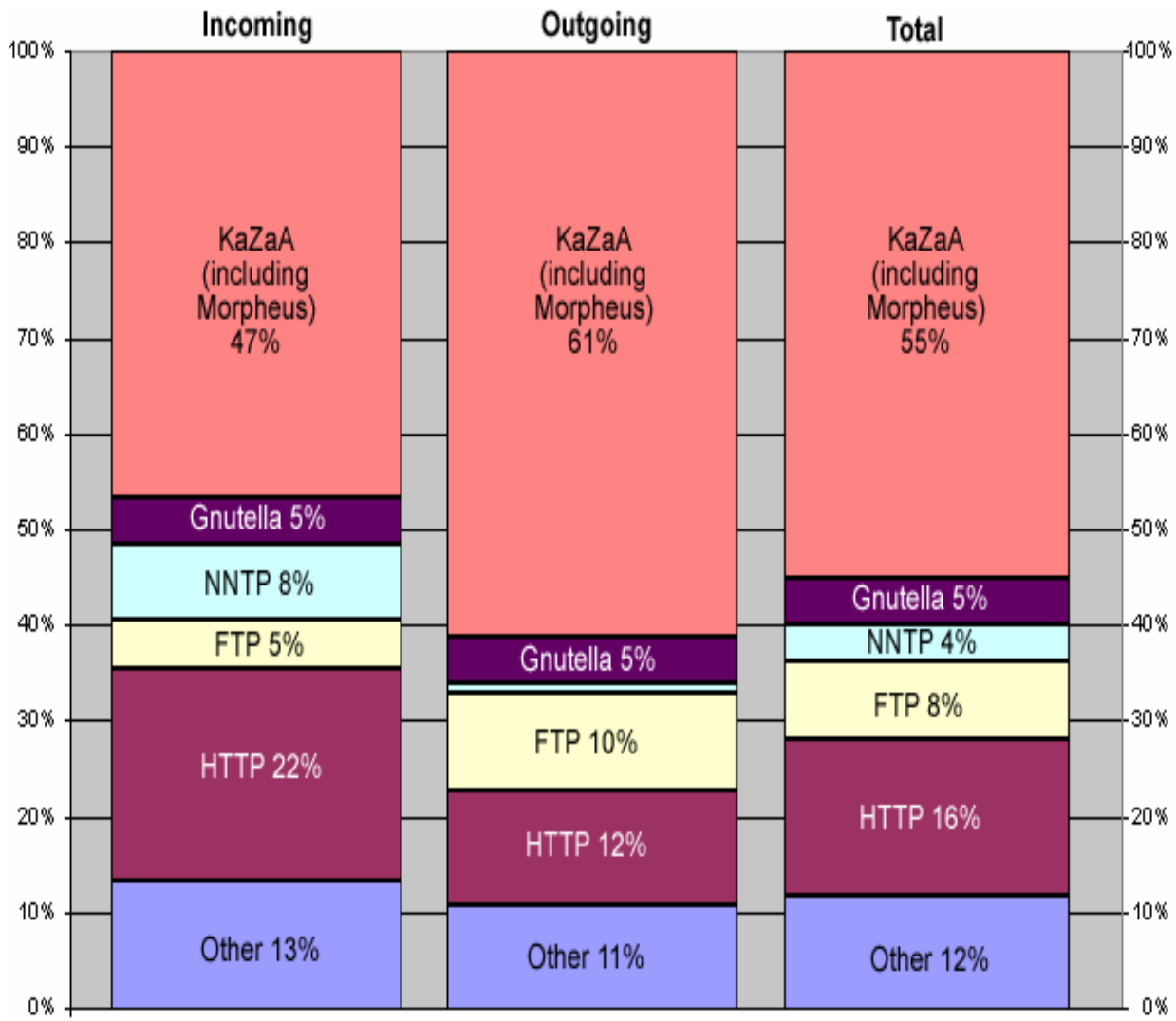
Systems that are considered

- Distributed, Autonomous, Self-organizing
  - P2P (Gnutella, KaZaA, etc.), companies' Intranets
- Assumption
  - Complete decentralization, no central control
- Advantages
  - Robust
    - Failures, Attacks, Censorship
  - Better Resource Utilization
  - Extensibility
    - light weight character
- Problems
  - Huge amount of (signaling) traffic
  - Dynamic behavior of P2P systems
  - Handle large networks



Prevalence of hybrid structures

- KaZaA
- Gnutella



Measurement traffic volume Germany 2004

Hybrid P2P Systems

Hybrid P2P systems have been used by the most popular file-sharing systems, such as KaZaA, Gnutella and Overnet.

- Largest overlay networks ever – very popular
- Cope with extremely dynamical user behavior
- Several millions users online
- Billions of files accessible
- Short search time, good hit rate.

⇒ Why are such Systems so successful?

⇒ Advantages useable for other application areas?

⇒ Concepts improvable?

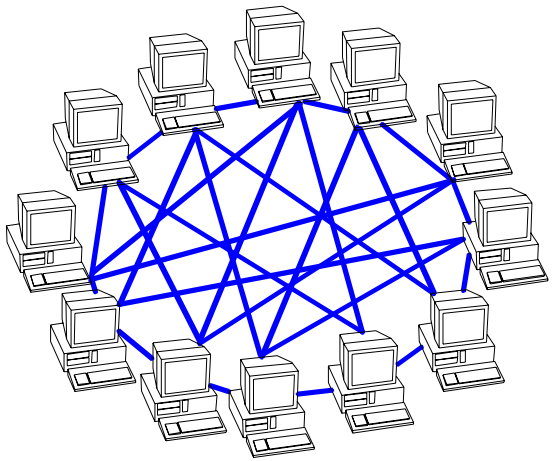
Analysis of hierarchical Overlay Structures

- Why were P2P systems so successful?
  - hybrid structures more effective than others?
- Advantages useable for other structures and systems?
  - Hierarchies for GRID?
  - Mixture with DHT?
- Comparisons of various hybrid structures:
  - Efficiency factor?
  - Application areas (size, user behavior)?

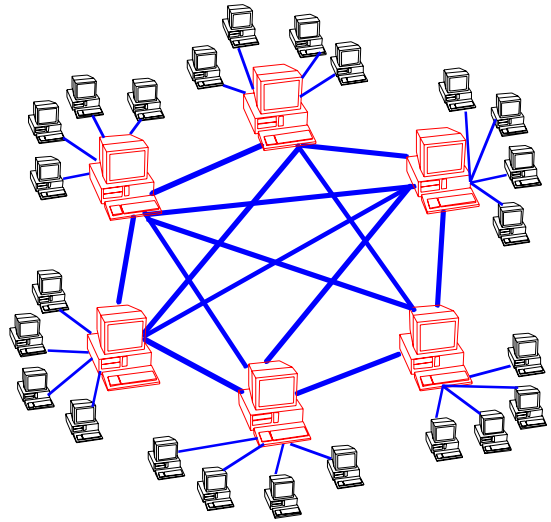
Hybrid and hierarchical Overlay Structures

KaZaA, Gnutella and Overnet:

- This hybrid P2P systems use hierarchical 2-tier structures.

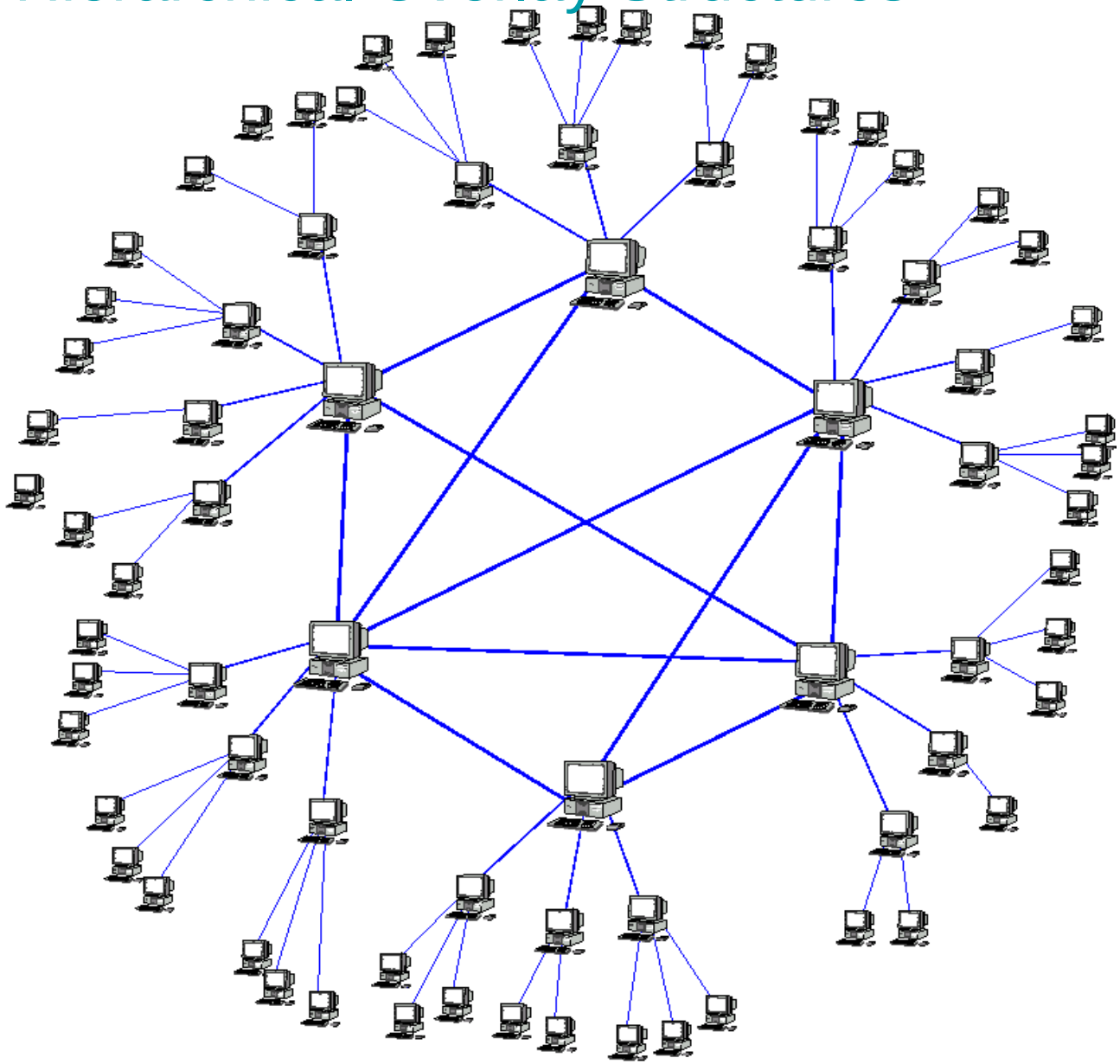


Pure P2P - No central entities



Hybrid P2P – dynamic central entities

Hierarchical Overlay Structures



A Generic Application Layer Simulation

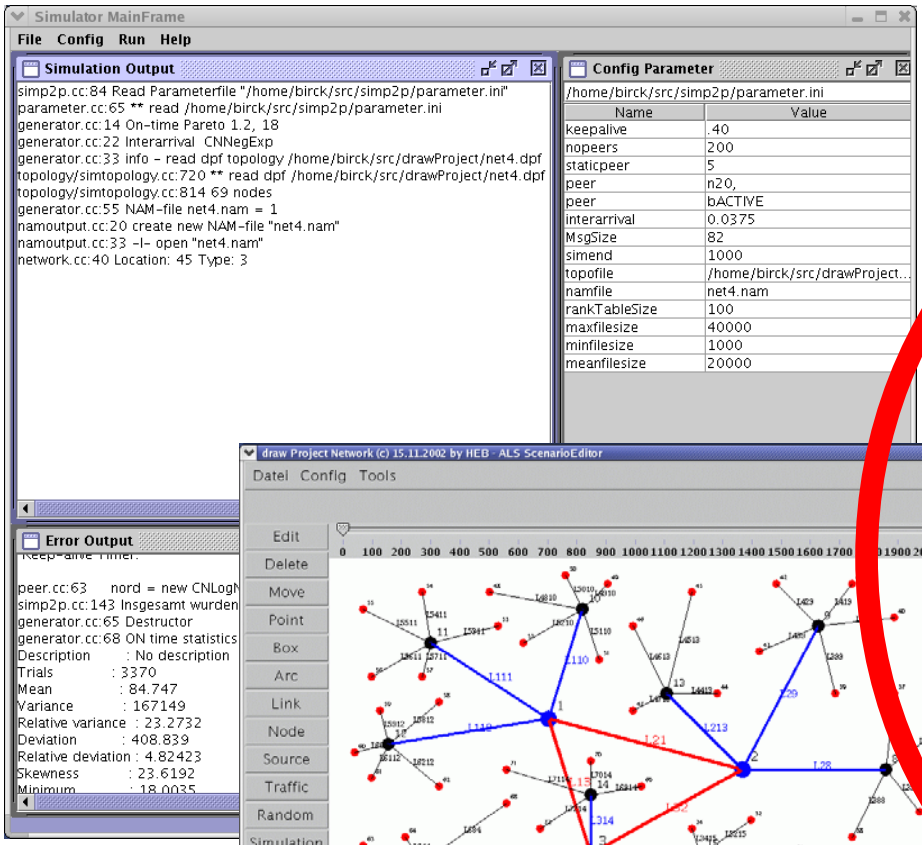
Application Layer Simulation (ALS) Features/Properties

- High abstraction level
- Application Layer details
- Easy to learn and use
- Flexible and extensible
- Efficiency

Expected analysis results

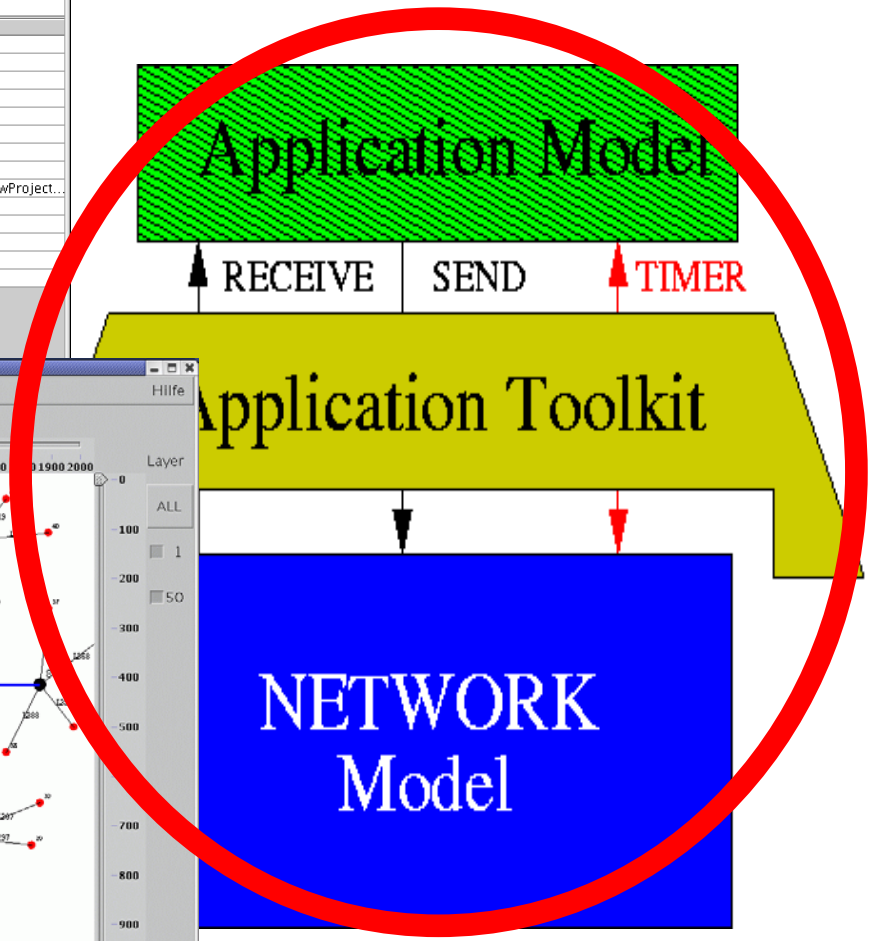
- Traffic characterization
- Traffic volume and overlay topology
- Systems dynamics
- Complete view of overlay traffic
- Proof of concept for overlay protocols

A Generic Application Layer Simulation



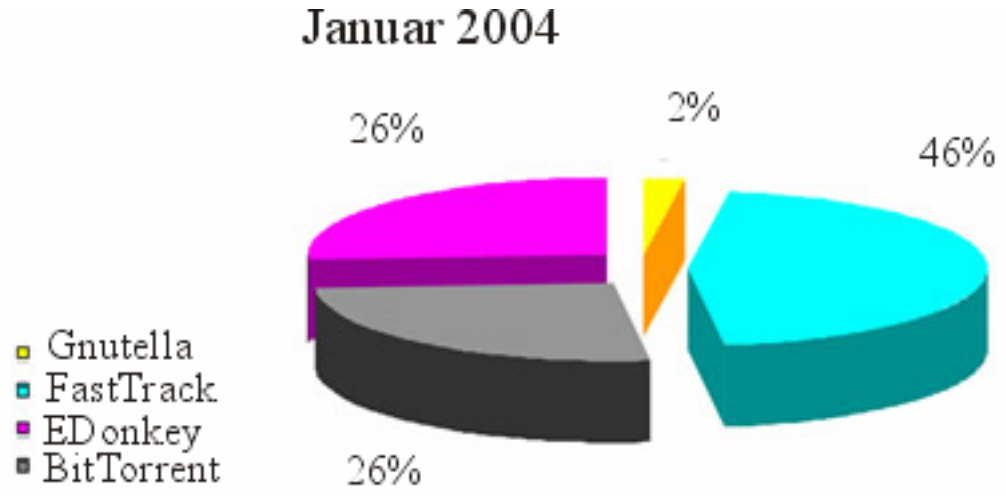
Demo

Architecture



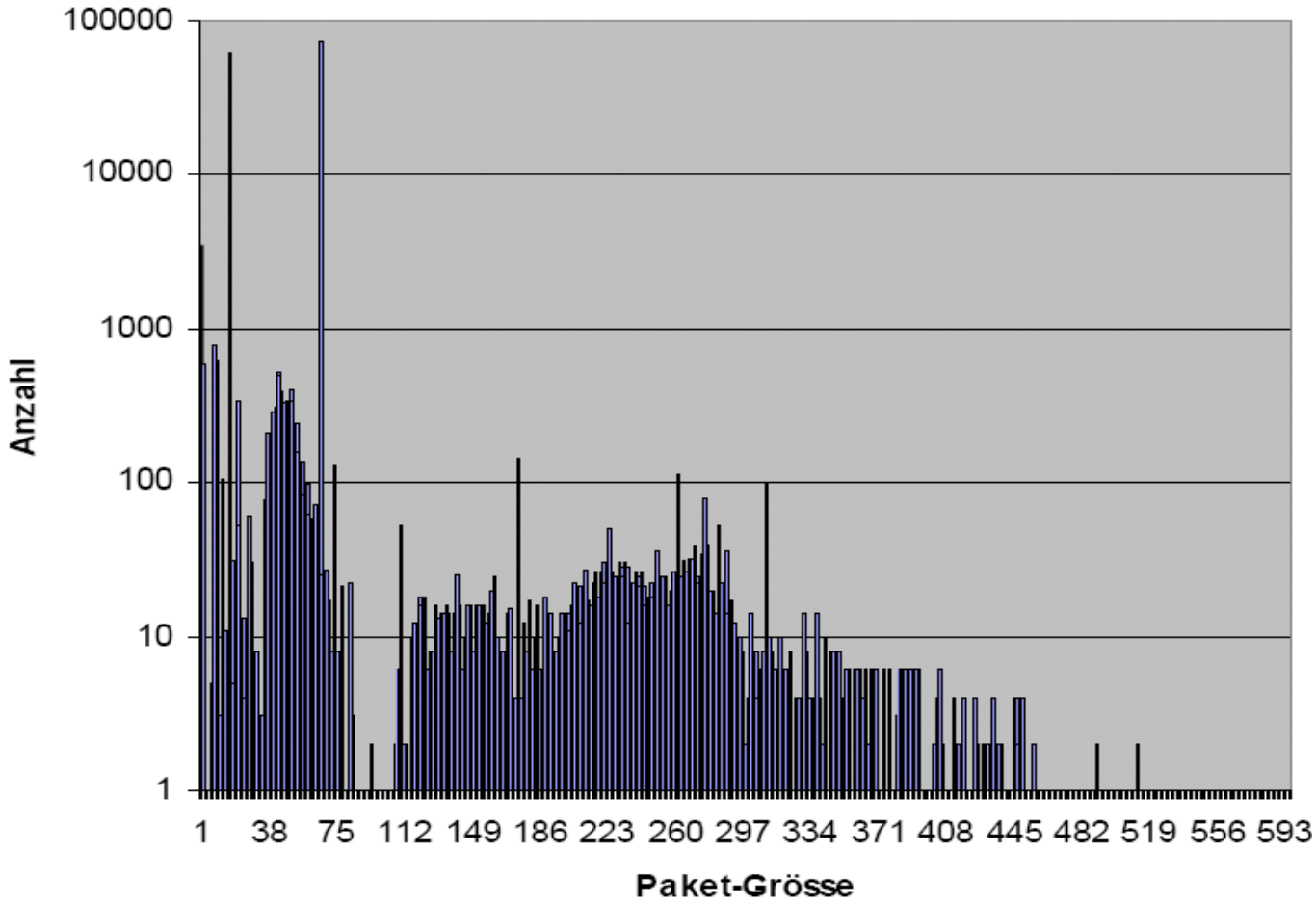
Analysis OpenFastTrack Protocol

- Other peer-to-peer protokolls are already analyzed
  - Gnutella, EDonkey at the TUD
  - FastTrack is one of the last challenges
- FastTrack was the most popular file sharing protocol
  - Responsible for huge amount of network traffic
- Understand the protocol behavior for simulation purpose
  - Function
  - Timing
  - Traffic efficiency



Measurements

Measurement of the traffic behavior
- Join & leave, Retrieval process
- Timing, On/Off times, packet sizes



Our Approach: MLOP

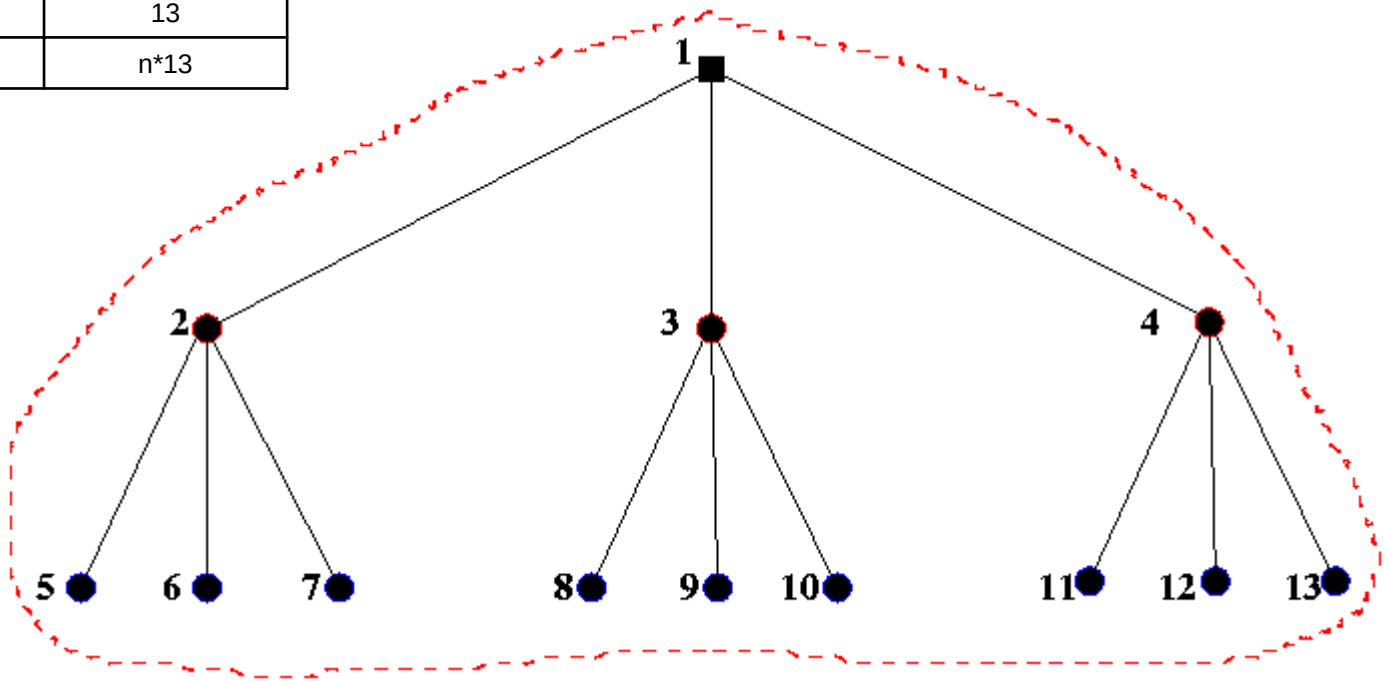
- Conception and implementation of a simple *Multi-Layer-Overlay-Protocol* (MLOP) based on OpenFastTrack protocol

- P2P Overlay Network Design Mechanisms in large-scale, heterogeneous, dynamic environments

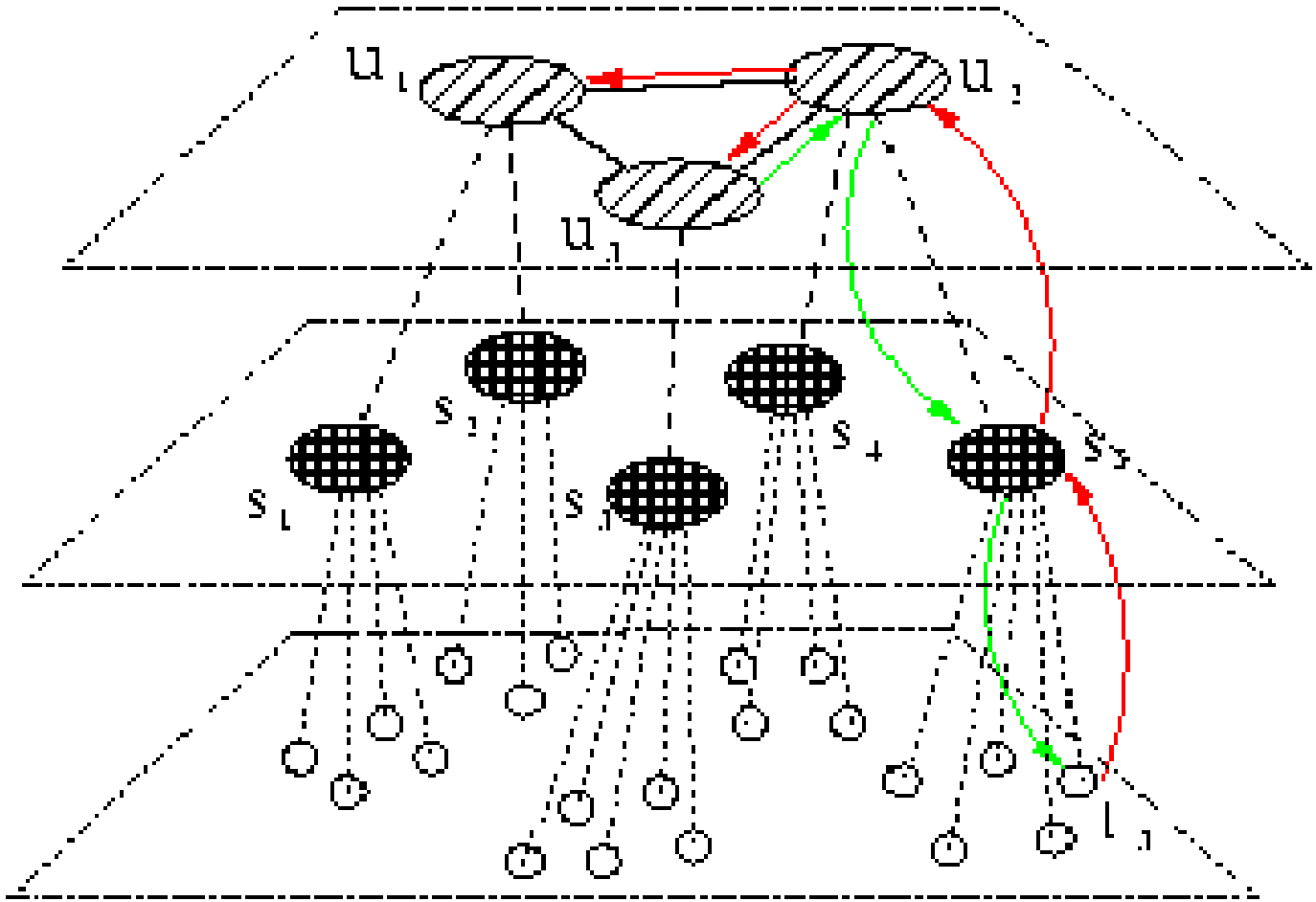
- Design goals for MLOP
  - Analysis protocol
  - Lightweight
  - Reduced signaling traffic
  - Simple routing algorithms

Searching in hierarchical structures

Hop	Nodes
0	1
1	4
2	13
3	n*13



Searching in hierarchical structures (2)



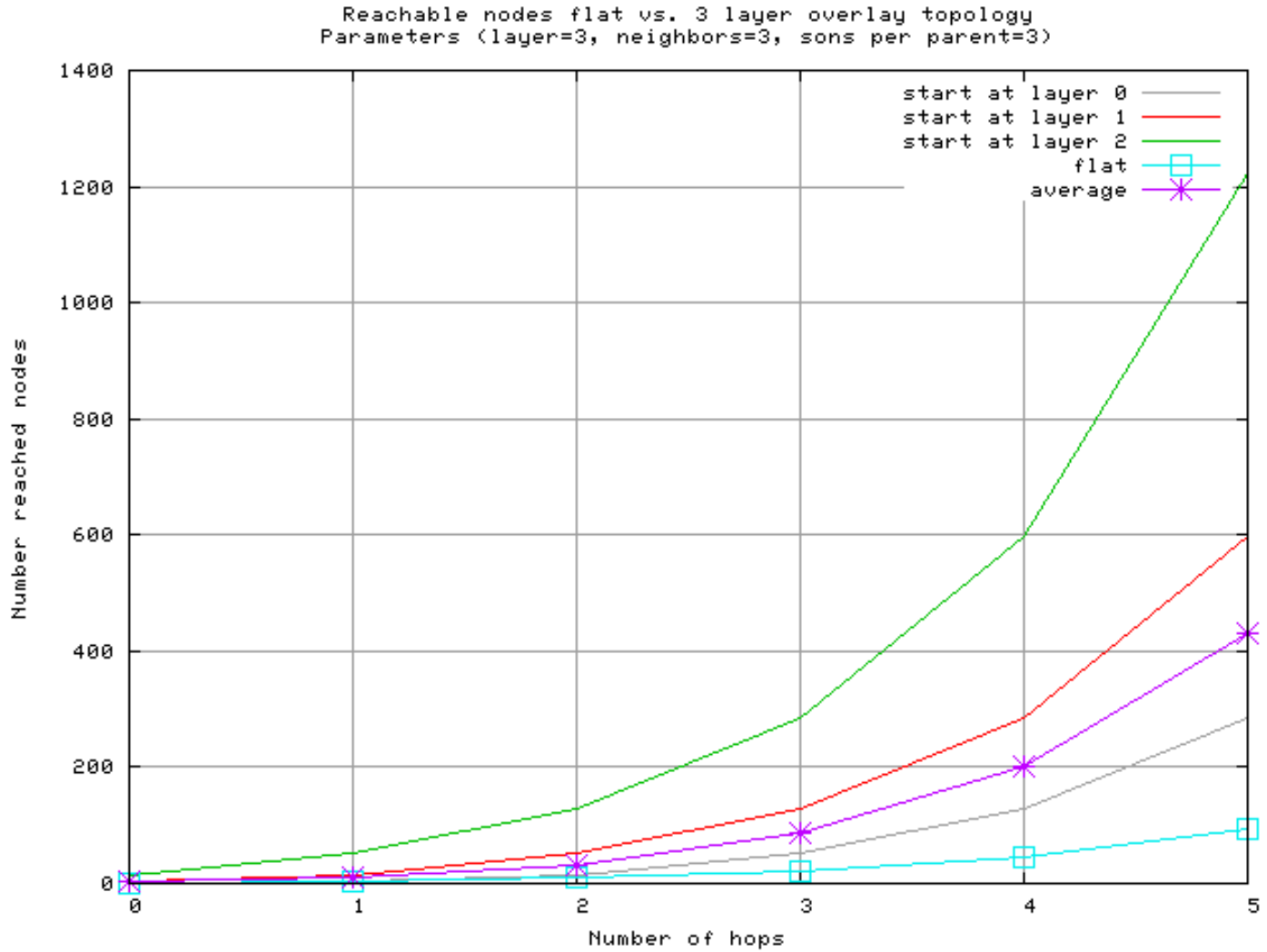
Experiments

"General statement for the (dis-) advantages of hierarchical structures in dynamic overlay networks."

- Application area
  - Performance vs. signaling traffic
  - Network size
  - Scalability
  - Topologies (clustering)

- Experiments of overlay networks with various …
  - Dynamics (on/off, join& leave, queries)
  - Topologies (logical + physical)
  - Hierarchies (1-2-3-4 layer)

Theory of Reachable Nodes



Theory of Reachable Nodes (2)

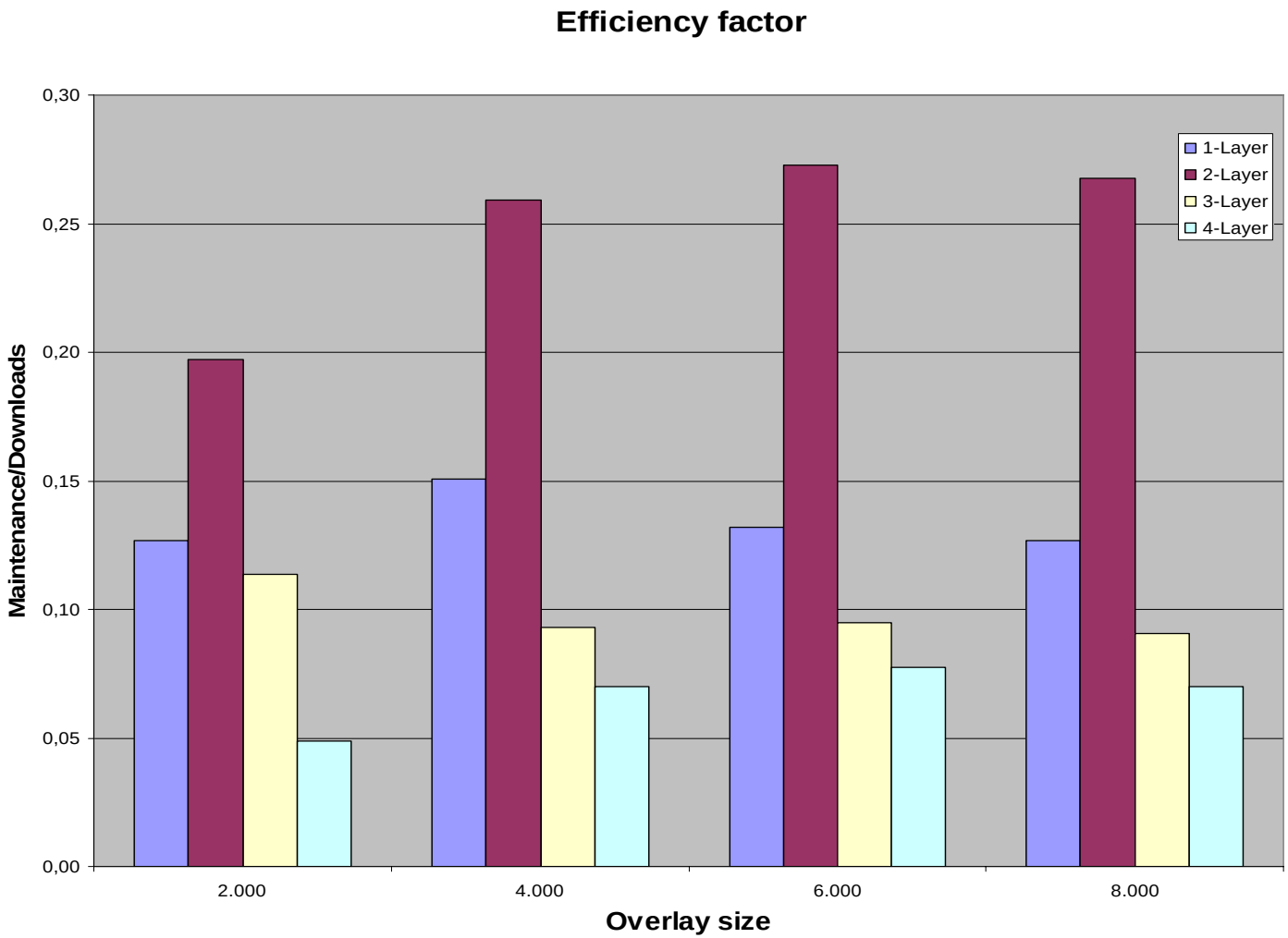
1. Reachable nodes for flat topologies (flooding)

$$r_1 = 1 + n\frac{(n-1)^h - 1}{n-2}$$

2. Reachable nodes for hierarchical topologies

$$r_l = \begin{cases} \sum_{j=0}^{l-1} s^j + \frac{n}{n-2}\sum_{j=0}^{l-1} s^j (n-1)^{h-j} - s^j & : l \geq h \\ \sum_{j=0}^{h} s^j + \frac{n}{n-2} s^h \sum_{j=0}^{h} (\frac{h-1}{s})^j - (\frac{1}{s})^j & : l > h \end{cases}$$

Comparison

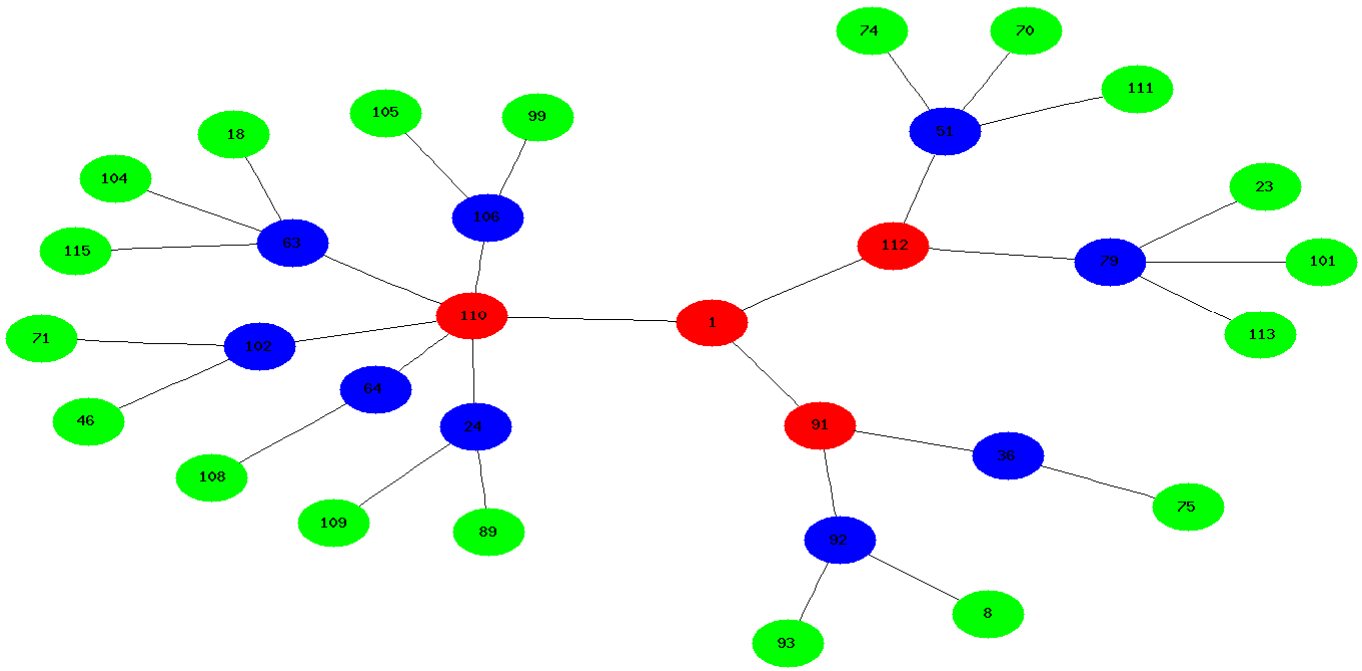


Efficiency factor:

maintenance costs/number successful download

Conclusion

Quality of achieved Network topology



Ongoing Work/Future

- Challenges
  - Clustering/Cliquish
  - Quality of graph structure
- Improvements of MLOP
  - Redundancy
  - Replication index
  - Link concept
- Future trends
  - Technology mixture (DHT + hierarchies).

Summary

- Hierarchical Overlay Structures
  - Prevalence of hybrid systems
- Analysis Framework
  - Application Layer Simulation
  - Analysis OpenFastTrack Protocol
- Multi-Layer-Overlay-Protocol
  - Lightweight protocol for analysis purpose
- Evaluation
  - Strong impact of dynamic behavior

Thank you!

- Contact:
  - Hannes Birck
  - birck@kom.tu-darmstadt.de
- Download ALS/MLOP:
  - http://www.kom.tu-darmstadt.de/~birck