Designs and Evaluation of a Tracker in P2P Networks

Adele Lu JIA and Dah Ming CHIU

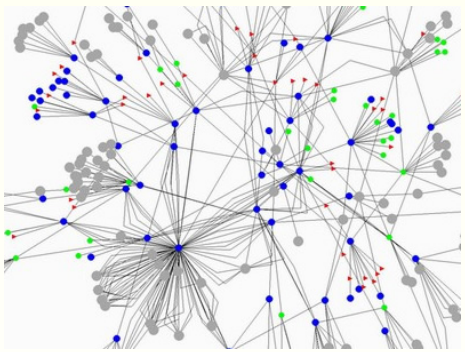
The Chinese University of Hong Kong

September 2008

Outlines

Introduction

- The term *tracker* originated from BitTorrent.
- The most critical function: *peer introduction*.
- Have more partner peers, get better performance.

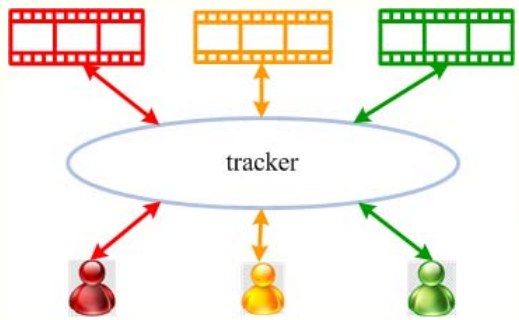

Introduction

- Tow main approaches for tracker design:
  - i server-based approach
  - ii peer-based approach
- Our contributions:
  - i A systematic description of the different designs of the tracker function.
  - ii Models for comparing the reliability of these designs.

A taxonomy of tracker designs

- Tracker function: mapping an object to a set of peers holding this object.
  - i Objects
  - ii Peers

- Three dimensions in tracker design:
  - i Who provide the tracker function?
  - ii How are objects assigned to tracker nodes?
  - iii How are peers assigned to tracker nodes?

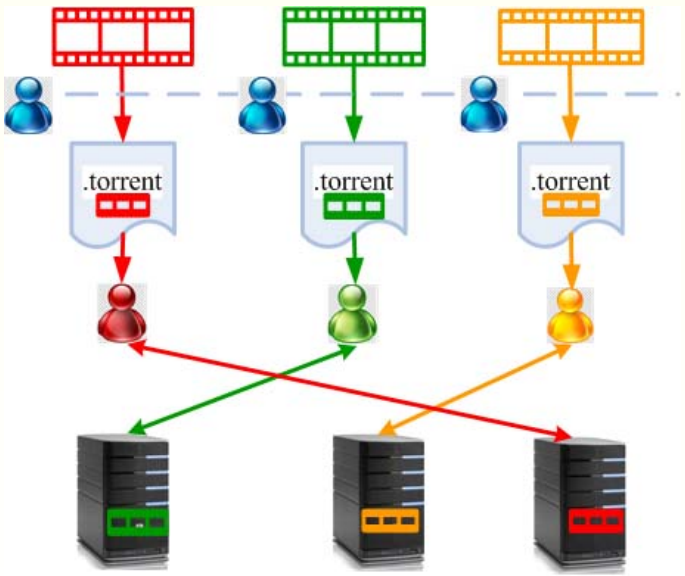


A taxonomy of tracker designs

1. Who provide the tracker function?
   - Deployed server(**DS**)
   - Peers(**P**)
2. How are objects assigned to tracker nodes?
   - By manual configuration(**M**)
   - Via a distributed hash table(**DHT**)
3. How are peers assigned to tracker nodes?
   - Based on user choice(**U**).
   - Automatic(**A**).

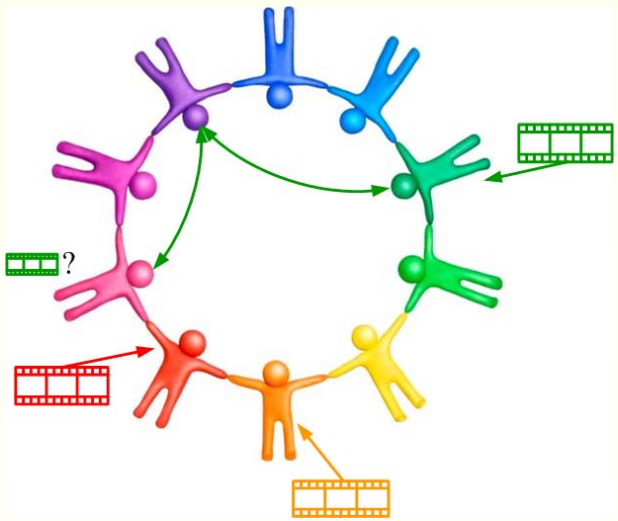
Examples

- Classic BitTorrent(**DS+M+U**)
  - The tracker is a deployed server.
  - The binding of the tracker node to the object is manually configured by file publisher in a meta-file(.torrent file).
  - Peers find the tracker by downloading the torrent file.



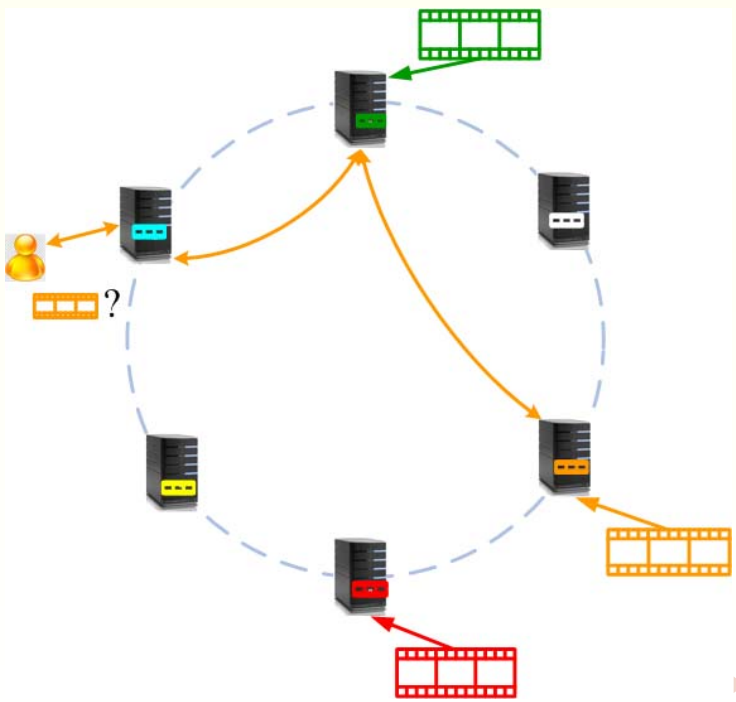
Examples

- eMule KAD(**P+DHT+A**)
  - Tracker function is shared by peers.
  - DHT provides a mapping from an object to a *target node*.
  - Peers find the target nodes via a DHT lookup.



Examples

- PPLive VoD(**DS+DHT+A**)
  - Tracker function is provided by deployed servers.
  - DHT is used to allocate the objects(video files) to tracker servers.



Design considerations

- Ease of implementation.
- Legal liability or management responsibility.
- Costs e.g. server and bandwidth costs.
- Flexibility.
- Security.

Performance models of tracker design

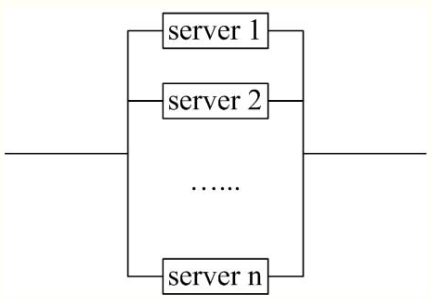
- Two main approaches in tracker design:
  - i server-based approach
  - ii DHT-based approach
- Performance metric: *lookup success probability*(reliability).
  - i Tracker node(s) up and running?
  - ii Tracker node(s) overloaded?

Single node performance factor

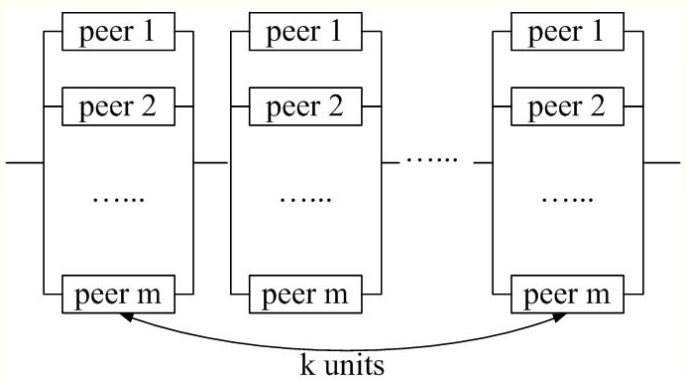
- $R = P_{up} \times P_{queued} \times P_{served}$
  - i R can be R_s or R_p.
  - ii P_{up} =*mean time to failure*/(*mean time to failure*+*mean time to repair*) is the probability that a peer is alive.
  - iii $P_{queued} = 1 - \frac{\rho^h - \rho^{h+1}}{1-\rho^{h+1}}$ is the probability that the tracker node is not overwhelmed.
  - iv P_{served} is the probability that the tracker node stays up till the request is served.

Reliability

- Tracker design can be modeled as a *parallel* or *series* set of units.
- Each unit is composed of a parallel set of units(as replication or backup).



(a) Server-based



(b) DHT-based

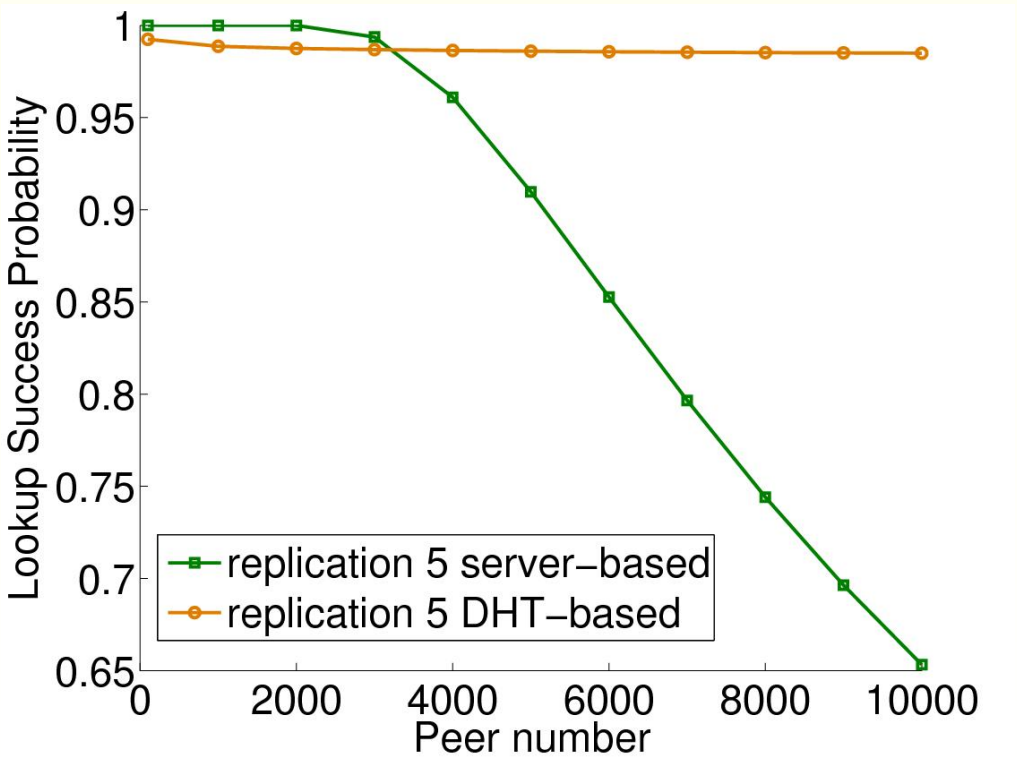
Figure: Reliability system model

Reliability

- $R_{server} = 1 - (1 - R_s)^n$
- $R_{DHT} = (1 - (1 - R_p)^m)^k$
  - i R_{server} and R_{DHT} denote the system reliability.
  - ii R_s and R_p denote the single server(peer) reliability.
  - iii n and m are replication number.
  - iv k is the expected path length visited by DHT(depending on the system population).
- Assumptions:
  - i Constant population.
  - ii Exponential distributed life time.
  - iii Exponential distributed lookups.

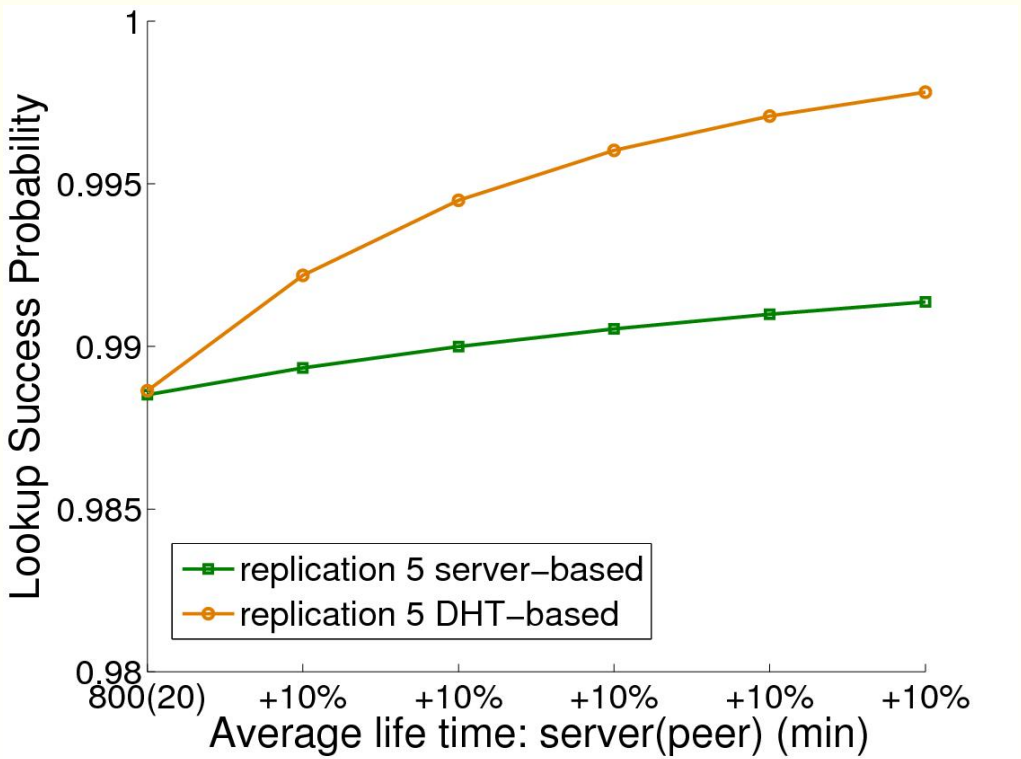
Observations from modeling

- The reliability of server-based tracker design can deteriorate quickly under heavy load.
- DHT-based tracker design has better scalability.



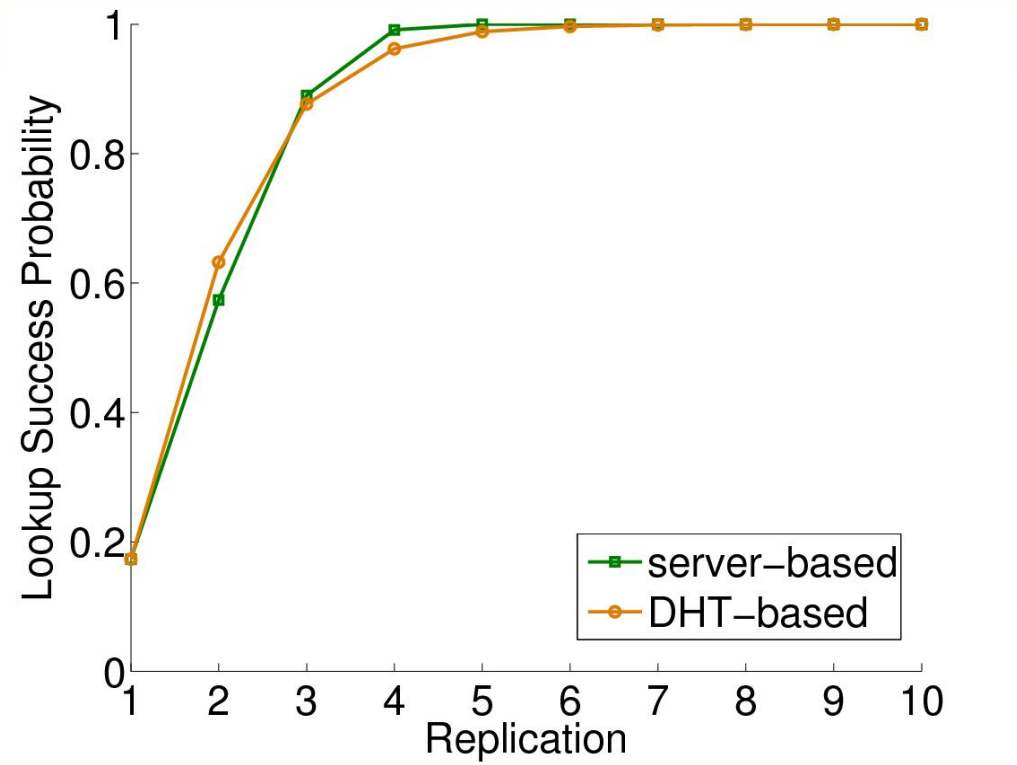
Observations from modeling

- More stable nodes can improve the reliability.
- DHT-based tracker design is more sensitive to peer's stability.



Observations from modeling

- More replication can improve the reliability.
- Sensitivity is about the same for both designs.



Conclusions and future work

- A preliminary study of tracker design in P2P systems.
- A taxonomy of tracker design.
- Simple models to analyze the reliability.
- More detailed modeling and analysis is in progress.
- We are also interested in building a capacity planning tool and studying a comparison of the overheads incurred by the different tracker designs.

Thank you.

Questions are welcome.