Tracker vs. DHT Performance Comparison for P2P Streaming

draft-hu-ppsp-tracker-dht-performance-comparison

Yan Hu, NEC Labs China
Yong Xia, NEC Labs China
Jan Seedorf, NEC Labs Europe
Outline

- Introduction
- Resource discovery performance comparison
- Chunk discovery performance comparison
- Conclusion
Introduction

Different methods for a peer to discover specific resource

- **Tracker-based method**: centralized server
  peer reports its resources to tracker;
  tracker stores and returns resources info to the requesting peer

- **DHT-based method**: fully-distributed lookup
  resources info is stored by many peers in the P2P network

This draft estimates the performance of the two methods

- Assume there are $D$ resources shared by $N$ peers in a P2P system
- For P2P streaming
  - $N$: number of active users in a P2P streaming software
    about 10 million ($10^7$) active users
  - $D$: number of channels (live streaming) or videos (VoD)
    about 100 thousand ($10^5$) resources
Resource discovery

- Two performance comparisons
  - **Resource discovery**: coarse level
    only compare the discovery performance of resource info
  - **Chunk discovery**: grain level
    also compare the discovery performance of chunk info

- Resource discovery performance comparison
  - Tracker-based method:
    tracker stores and returns resource info, chunk info is exchanged using peer gossip
  - DHT-based method:
    resource info is obtained using DHT method, chunk info is exchanged using peer gossip
    (Assumption: DHT nodes are widely distributed on the Internet)
Lookup efficiency

Parameters and assumptions

- **N**: number of peers, \( N = 10,000,000 \)
- **D**: number of resources, \( D = 100,000 \)
- **RTT**: average RTT in the network, \( RTT = 200\text{ms} \)

Lookup efficiency comparison

<table>
<thead>
<tr>
<th></th>
<th>Tracker-based</th>
<th>DHT-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lookup message</td>
<td>( O(1) )</td>
<td>( O(\log(N)) = 23 )</td>
</tr>
<tr>
<td>Lookup operations</td>
<td>( O(1) )</td>
<td>( \log(N) \times O(1) = 23 )</td>
</tr>
<tr>
<td>Lookup latency</td>
<td>( O(1) \times RTT = 200\text{ms} )</td>
<td>( O(\log(N)) \times RTT = 4.6\text{s} )</td>
</tr>
</tbody>
</table>

**Summary**: Tracker-based method is much faster than DHT-based method, the 4.6s lookup latency is relatively high in P2P streaming applications.
Network traffic

Parameters and assumptions

- **N**: number of peers, \( N = 10,000,000 \)
- **T**: each peer requests new resource every \( T \) seconds, \( T = 60 \text{sec} \)
- **S**: average size of one request/response message, \( S = 1\text{KBytes} \)

Network traffic comparison

<table>
<thead>
<tr>
<th></th>
<th>Tracker-based</th>
<th>DHT-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of messages per second</td>
<td>( \frac{N}{T} \times 2 = 3.3 \times 100,000 )</td>
<td>( \frac{N}{T} \times 2 \log(N) = 7.7 \times 1,000,000 )</td>
</tr>
<tr>
<td>Size of messages per second</td>
<td>( \frac{N}{T} \times 2 \times S = 0.33 \text{GBytes} )</td>
<td>( \frac{N}{T} \times 2 \log(N) \times S = 7.7 \text{GBytes} )</td>
</tr>
<tr>
<td>Number of messages in node join/leave</td>
<td>( O(1) )</td>
<td>( O((\log N)^2) = 541 )</td>
</tr>
</tbody>
</table>

**Summary:**
Tracker-based method has smaller network traffic overhead than DHT-based method, both methods are acceptable in P2P streaming applications.
Host requirement

Parameters and assumptions

- **T**: each peer requests new resource every T seconds, T = 60sec
- **S**: average size of one request/response message, S = 1KBytes
- **C**: one peer has C resources, C = 10
- **P**: each peer is represented by P Bytes, P = 20 Bytes

Host requirement comparison

<table>
<thead>
<tr>
<th></th>
<th>Tracker-based</th>
<th>DHT-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory requirement</td>
<td>N<em>C</em>P = 2GBytes</td>
<td>(N*C/D)*P = 20KBytes</td>
</tr>
<tr>
<td>Number of requests</td>
<td>N/T = 1.67*100,000</td>
<td>log(N)/T = 0.4</td>
</tr>
<tr>
<td>received per sec</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of request</td>
<td>N/T<em>2</em>S = 0.33GBytes</td>
<td>2<em>log(N)/T</em>S = 0.8 Kbytes</td>
</tr>
<tr>
<td>response messages per</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sec</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summary:
DHT-based has much less host resources requirement than tracker-based method. For performance considerations, multiple trackers can be used.
Chunk discovery

- Two performance comparisons
  - **Resource discovery**: coarse level
    only compare the discovery performance of resource info
  - **Chunk discovery**: grain level
    also compare the discovery performance of chunk info

- Chunk discovery performance comparison
  - Tracker-based method:
    tracker stores and returns resource info, chunk info is exchanged using peer gossip
  - DHT-based method:
    both resource info and chunk info are obtained using DHT method
    (i.e., the first solution in “Chunk Discovery for P2P Streaming”)

©NEC Labs China 2009
Lookup efficiency

Parameters and assumptions
- **N**: number of peers, $N = 10,000,000$
- **D**: number of resources, $D = 100,000$
- **RTT**: average RTT in the network, $RTT = 200$ms
- **M**: each peer gossip with M neighbors, $M = 20$

Lookup efficiency comparison

<table>
<thead>
<tr>
<th></th>
<th>Tracker-based</th>
<th>DHT-based</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tracker side</td>
<td>Peer side</td>
</tr>
<tr>
<td>Lookup message</td>
<td>$O(1)$</td>
<td>$M \times O(1) = 20$</td>
</tr>
<tr>
<td>Lookup operations</td>
<td>$O(1)$</td>
<td>$O(1)$</td>
</tr>
<tr>
<td>Lookup latency</td>
<td>$O(1) \times RTT = 200$ms</td>
<td>$O(1) \times RTT = 200$ms</td>
</tr>
</tbody>
</table>

**Summary:**
Tracker-based method is much faster than DHT-based method, the 4.6s lookup latency is relatively high in P2P streaming applications.
Network traffic

Parameters and assumptions

- **T**: each peer requests new resource every T seconds, T = 60sec
- **S**: average size of one request/response message, S = 1KBytes
- **I**: peer sends gossip messages every I seconds, I = 10 sec
- **R**: video rate, R = 32 KBytes/sec; **Z**: chunk size, Z = 16 KBytes

Network traffic comparison

<table>
<thead>
<tr>
<th></th>
<th>Tracker-based</th>
<th>DHT-based</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tracker side</td>
<td>Peer side</td>
</tr>
<tr>
<td>Number of messages per sec</td>
<td>N/T<em>2 = 3.3</em>100,000</td>
<td>M<em>N/I</em>2 = 4*10,000,000</td>
</tr>
<tr>
<td>Size of messages per sec</td>
<td>N/T<em>2</em>S = 0.33GBytes</td>
<td>M<em>N/I</em>2*S = 40GBytes</td>
</tr>
</tbody>
</table>

Summary:
Tracker-based method has smaller network traffic overhead than DHT-based method, both methods are acceptable in P2P streaming applications.
Host requirement

Parameters and assumptions
- **C**: one peer has C resources, $C = 10$
- **P**: each peer is represented by P Bytes, $P = 20$ Bytes
- **Bm**: bitmap size, $Bm = 1$KBytes
- **H**: number of chunks in one resource, $H = 10000$

Host requirement comparison

<table>
<thead>
<tr>
<th></th>
<th>Tracker-based</th>
<th>DHT-based</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tracker side</td>
<td>Peer side</td>
</tr>
<tr>
<td>Memory requirement</td>
<td>$N<em>C</em>P = 2GBytes$</td>
<td>$M*Bm = 20KBytes$</td>
</tr>
<tr>
<td>Number of requests received per sec</td>
<td>$N/T = 1.67*100,000$</td>
<td>$M/I = 2$</td>
</tr>
<tr>
<td>Size of req/resp messages per sec</td>
<td>$N/T<em>2</em>S = 0.33GBytes$</td>
<td>$M/I<em>2</em>S = 4KBytes$</td>
</tr>
</tbody>
</table>

Summary:
DHT-based has much less host resources requirement than tracker-based method. For performance considerations, multiple trackers can be used.
Conclusion

- This draft compares resource discovery and chunk discovery performance of Tracker-based and DHT-based method.
- Tracker-based method has much shorter response time than DHT-based method.
- DHT-based method’s response time can be long, not suitable for delay sensitive streaming applications.
- Per-host requirement of tracker is higher than DHT nodes, but still within reach of a small number of commodity PCs.
Thanks!