Empowered by Innovation NEC

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 = 200ms
- Lookup efficiency comparison

	Tracker-based	DHT-based
Lookup message	O(1)	O(log(N)) = 23
Lookup operations	O(1)	log(N)*O(1) = 23
Lookup latency	O(1)*RTT = 200ms	O(log(N))*RTT = 4.6s

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 = 60sec
  - **S**: average size of one request/response message, S = 1KBytes
- Network traffic comparison

	Tracker-based	DHT-based
Number of messages per second	N/T*2 = 3.3*100,000	N/T*2*log(N) = 7.7*1,000,000
Size of messages per second	N/T*2*S = 0.33GBytes	N/T*2*log(N) *S = 7.7GBytes
Number of messages in node join/leave	O(1)	$O((\log N)^2) = 541$

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

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

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”)

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 = 200ms
  - **M**: each peer gossip with M neighbors, M = 20
- Lookup efficiency comparison

	Tracker-based		DHT-based
	Tracker side	Peer side	
Lookup message	O(1)	M*O(1) = 20	O(log(N)) = 23
Lookup operations	O(1)	O(1)	log(N)*O(1) = 23
Lookup latency	O(1)*RTT = 200ms	O(1)*RTT = 200ms	O(log(N))*RTT = 4.6s

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

	Tracker-based		DHT-based
	Tracker side	Peer side	
Number of messages per sec	N/T*2 = 3.3*100,000	M*N/I*2 = 4*10,000,000	N*(R/Z)*2log(N) = 1,000,000,000
Size of messages per sec	N/T*2*S = 0.33GBytes	M*N/I*2*S = 40GBytes	N*(R/Z)*2log(N)*S = 1TBytes

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 = 1KBytes
  - **H**: number of chunks in one resource, H = 10000
- Host requirement comparison

	Tracker-based		DHT-based
	Tracker side	Peer side	
Memory requirement	N*C*P = 2GBytes	M*Bm = 20KBytes	(N*C/D)*P*(D*H/N) = 2MBytes
Number of requests received per sec	N/T = 1.67*100,000	M/I = 2	(R/Z)*log(N) = 46
Size of req/resp messages per sec	N/T*2*S = 0.33GBytes	M/I*2*S = 4KBytes	(R/Z) *log(N)*2*S= 92 Kbytes

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 short 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!