

IPFSにおけるキャッシュを用いた DHT探索時間の低減

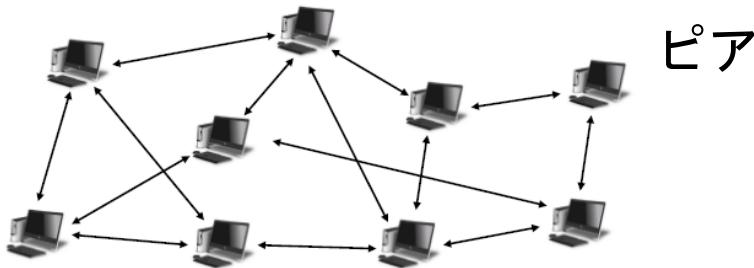
NS研究会
2025年10月10日

立命館大学 情報理工学部
福岡佳奈 上山 憲昭

IPFS (InterPlanetary File System)

■ IPFS (InterPlanetary File System)

- P2Pネットワーク上で、自律的にデータを分散して保存・配布
- 各ピア(ノード)がコンテンツを保持し、ハッシュ値(CID)で識別
- 中央のサーバを持たず、耐障害性・可用性が高い



■ コンテンツ探索

- DHT探索によってCIDに対応するピアを発見
- ネットワーク全体で効率的にルーティング

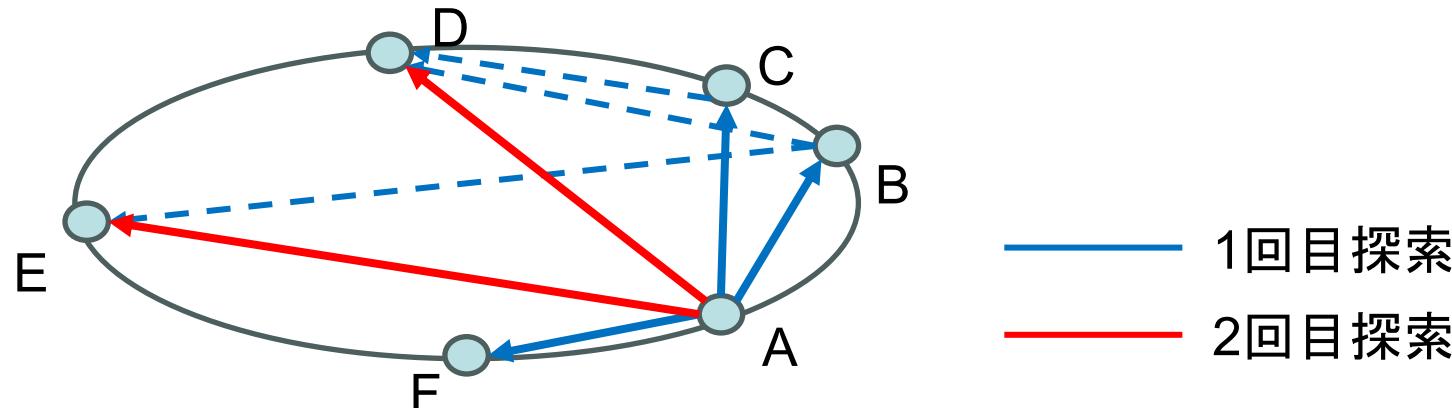
DHT (distributed hash table)

- DHT
 - 分散型ハッシュテーブル
 - 各ピアが一部のキーとノード情報を管理
 - キー(CID)→ピア(PID)の対応を分散して保持
- DHTウォーク
 - ルーティングテーブルを参照し、目的CIDに近づく
 - CIDとPIDのXOR距離が近いピアに問い合わせることで、コンテンツの保有ピアを発見できる可能性が向上
 - $O(\log N)$ ステップでノードを探索可能
- Kademlia
 - DHTの一種
 - IPFSはKademliaの拡張機能

Kademlia

■ 探索

- 初期ピアの選択
 - k-bucketsからCIDに近いピアを α 個選択・探索リストに追加
- 並列問い合わせ
 - 探索リスト内のピアにCIDを並列問い合わせ
- 探索の反復
 - 新しいピアが見つからなくなるか、CIDに最も近い k 個のピアがすべて応答するまで繰り返す
- 探索の完了
 - XOR距離が最も近い k 個のピアの情報が発見



課題

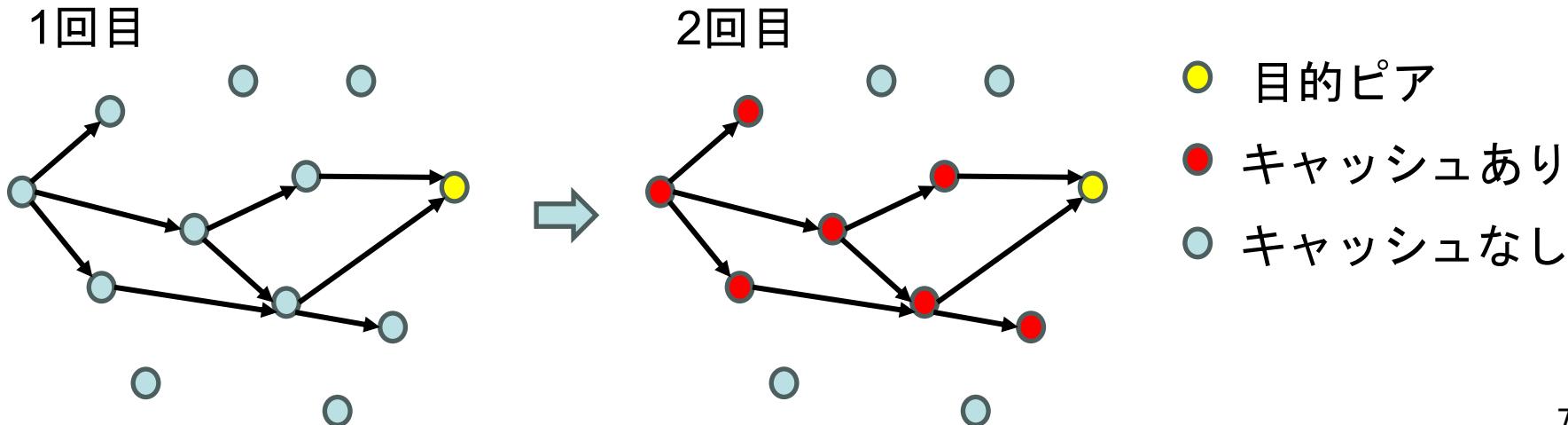
- ピア選択時に物理的な位置を非考慮
 - CIDとXOR距離が近いPIDを有するピアを選択
 - 各ピアが存在する位置が考慮されない ⇒ 遅延の増大
- ピア数が増加するほど、DHT探索に要する遅延や処理負荷が増大

研究目的

- 既存研究
 - DHTに関する様々なキャッシング手法の研究
→IPFSに特化したDHT探索遅延低減の未評価
- 本研究の目的
 - Publisher情報を探索経路上にキャッシングすることで、IPFSのGET処理(コンテンツ取得)に要する時間を低減
 - Publisher情報…コンテンツを保持しているノードの情報

提案方式

- キャッシュ機能
 - DHT探索時に探索経路上のノードの情報(PID, IP)を記録
 - 発見されたPublisher情報(CID, PID, IP)を経路上の全ノードに通知
 - 通知を受けたノードは取得した情報をキャッシュ
- 2回目以降、同じCIDに対する要求受信時はキャッシュ情報を利用して探索をショートカット



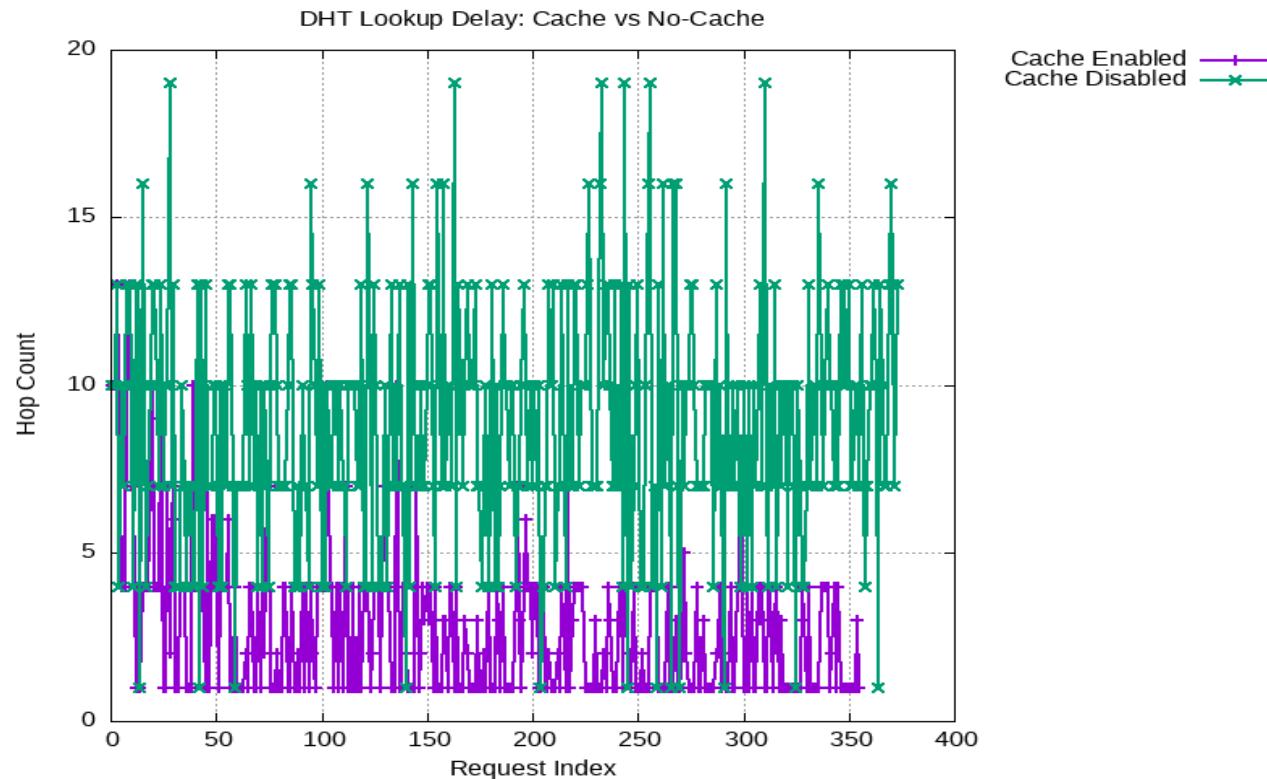
評価条件

- 比較方式
 - キャッシュありの場合と、なしの場合の平均ホップ数とキャッシュヒット率を比較
- 本研究で構成したIPFS
 - ノード数: 10000
 - キャッシュサイズ: 20
 - Zipf分布のパラメータ: 0.8

キャッシュ機能の性能評価 (1)

■ リクエスト数TOP1のCIDのホップ数の推移

- 紫: キャッシュあり
- 緑: キャッシュなし
- 縦軸: リクエストごとのホップ数
- 横軸: リクエスト数



キャッシュ機能の性能評価 (2)

- リクエスト数TOP1のCIDによる平均ホップ数の比較(CID数: 1000)

リクエスト数(回)	1000	5000	10000
キャッシュあり	4.06	2.73	2.27
キャッシュなし	9.47	9.03	9.15

- キャッシュヒット率

リクエスト数(回)	1000	5000	10000
キャッシュヒット率(%)	30.2	52.4	62.2

- リクエストが増加するほど、キャッシュありのホップ数が減少、キャッシュヒット率が増加
 - 探索経路上のノードに人気のコンテンツのキャッシュが配置されるため

キャッシュ機能の性能評価 (3)

- リクエスト数のTOP1のCIDによる平均ホップ数の比較(リクエスト数:10000)

CID数(個)	100	1000	5000
キャッシュあり	1.88	2.27	2.72
キャッシュなし	8.94	9.15	8.80

- キャッシュヒット率

CID数(個)	100	1000	5000
キャッシュヒット率(%)	87.3	62.2	34.3

- コンテンツが増加するほど、キャッシュありのホップ数が増加、キャッシュヒット率が減少
 - コンテンツが多様化するほど、キャッシュの効果が減少

まとめ

- 本研究では、IPFSのDHT探索時間の低減を目的として経路上のノードへのキャッシュを提案
 - IPFSのDHT探索経路上のノードにキャッシュの効果について、ホップ数を計測
 - リクエスト回数が増加するほど平均ホップ数の減少、キャッシュヒット率の向上
 - キャッシュの有効性を確認
- 今後について
 - キャッシュの配置の検討
 - ピニングとキャッシュの組み合わせ
 - ピニング…IPFS上で特定のコンテンツを永続的に保持