

IPFSにおける経路キャッシュとピンニング による探索遅延の削減手法

CCS研究会

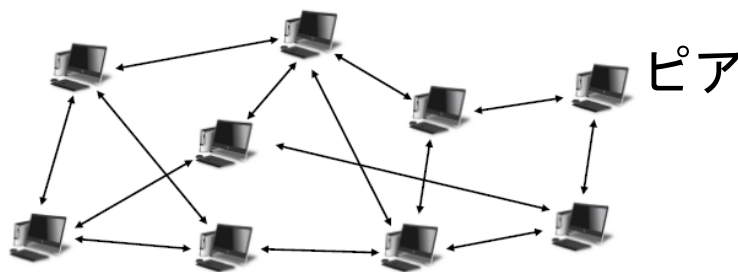
2026年6月12日

立命館大学院 情報理工学研究科

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
 - 中央のサーバを持たず、耐障害性・可用性が高い
 - 各ピアに重複なく一意な識別子(PID)を付与
 - 各ピア(ノード)がコンテンツを保持し、CIDでコンテンツを識別
 - 各ピアが自律的に通信、コンテンツを取得



- 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
- コンテンツのキー(CID)→ピア(PID)の対応を分散して保持
- 各ピアが一部のキーとノード情報を管理

■ Kademlia

- DHTの一種
- IPFSはKademliaの拡張機能

■ DHTウォーク

- IPFSでのPUT操作とGET操作において行われる探索
 - PUT操作…コンテンツのアップロード
 - GET操作…コンテンツの取得
- $O(\log N)$ ステップでノードを探索可能

IPFSのピンニング

■ ピンニング

- IPFS上で特定のコンテンツをノードに永続的に保持
 - GC(garbage collection)による使用頻度の低いコンテンツの削除からコンテンツを保護

■ ピンニングの方法

- 自身のIPFSノードでピンニング
- IPFSピンニングサービスを利用
 - 専用のIPFSノード群を運用
 - 料金を支払うことでコンテンツ固定

■ 料金を支払うことでストレージを提供

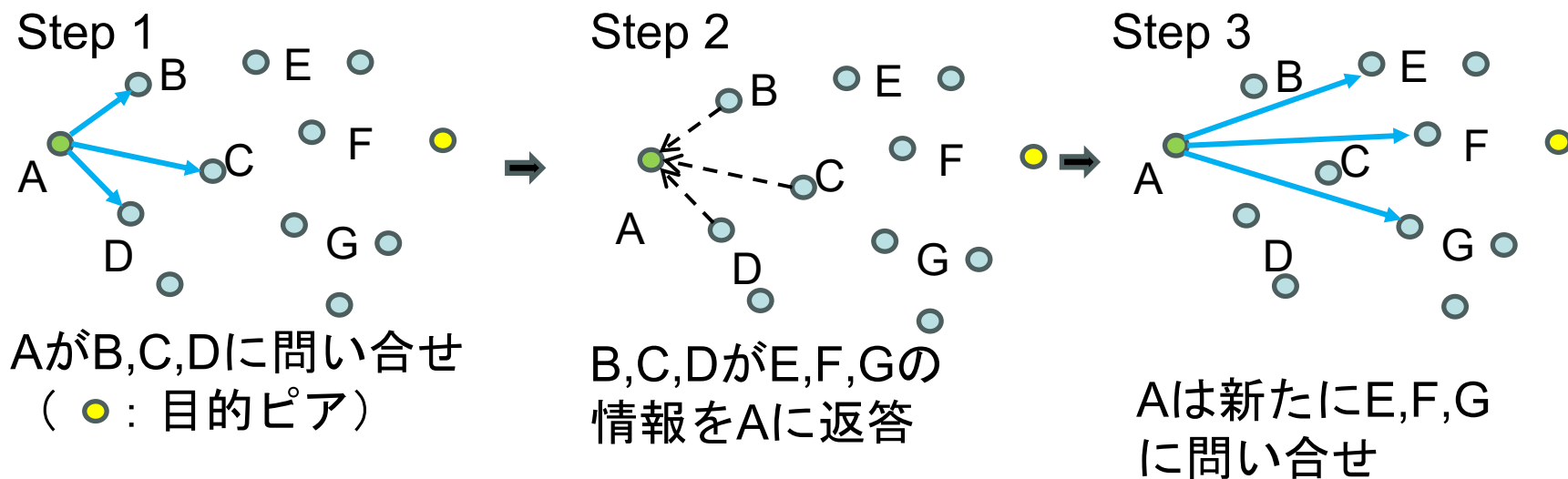
IPFSのコンテンツ取得の流れ(1)

- CIDのPR(Provider Record)情報を複数のピアに配布
 - PR情報: 各コンテンツを保持するピアの情報(PID, IPアドレス)
- PR情報が配布されるピアの選択
 - DHTウォークでピアを探索
 - 配布されるPR情報のCIDと, XOR距離が近いPIDを持つピアを20個選出
 - 選出されたピアにPR情報を配布

XOR距離が近いピアに問い合わせることで、
目的データの保有ピアを発見できる可能性が向上

IPFSのコンテンツ取得の流れ(2)

- IPFSはDHTを用いてPUTとGETの対象ピアを選択
- GET: ユーザがIPFSからデータを取得する処理
 - $\alpha = 3$ 個の, CIDに最もXOR距離が近いピアをルーティングテーブルから選択して, 最初に問い合わせ
 - 回答結果の中から, さらにXOR距離が近いピアに問い合わせ
 - 問い合わせを反復し, 目的データを保有するピアを発見
 - 見つからない場合もあり



課題

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

本研究の目的

- 既存研究
 - P2P配信におけるDHTに関する様々なキャッシュ手法
→ IPFSに特化したDHT探索遅延低減のための
キャッシュ方式は未検討
- 本研究の目的
 - PR情報を探索経路上のピアにキャッシュ
→IPFSではノードの離脱が激しい→キャッシュが消滅
 - ピンク専用ピアが特定のCIDのPR情報を保持
→IPFSにおいて、キャッシュの効果を向上し、GET処理
に要する時間を効果的に低減

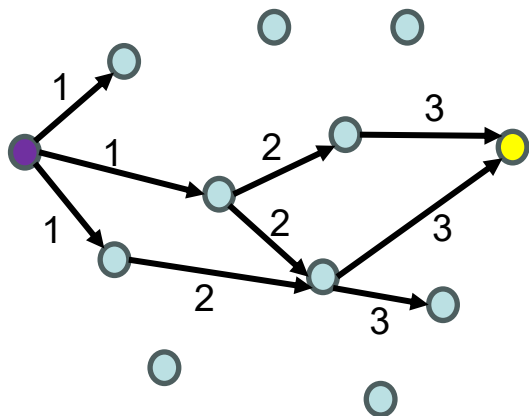
PR情報のキャッシュ機能(1)

■ キャッシュ機能

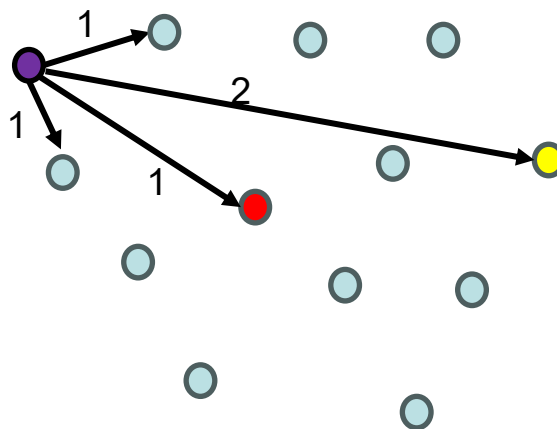
- DHT探索時に探索経路上のノード情報(PID, IP)を記録
- 発見されたPR情報(CID, PID, IP)を経路上の全ノードに通知
- 通知を受けたノードは取得した情報をキャッシュ

- 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

PRキャッシュなし



PRキャッシュあり



- 要求ピア
- 目的ピア
- キャッシュあり
- キャッシュなし

PR情報のキャッシュ機能(2)

- 対象
 - 全コンテンツを対象
- 置換方法
 - LRU (Least Recently Used) 方式
 - キャッシュ容量が上限に達した際、最終参照時刻が最も古いデータから順に破棄

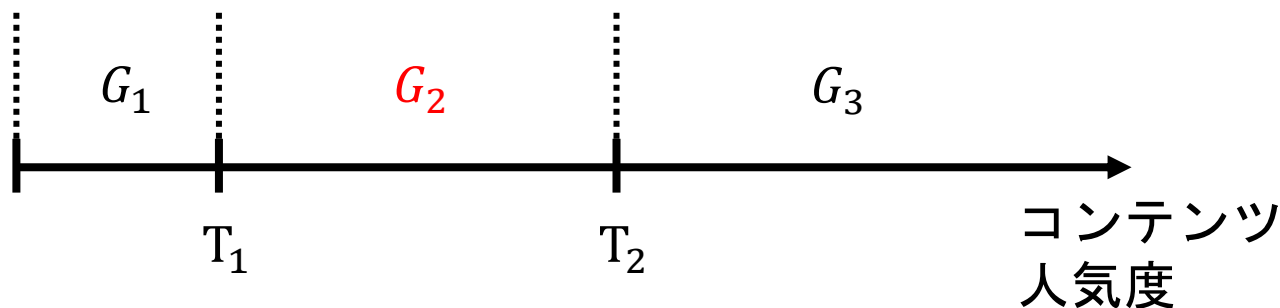
→頻繁にリクエストされるPR情報を優先的に保持

提案PRピンング機能（1）

- IPFSではピアの離脱が頻繁に発生
 - コンテンツの保持ピアが全てなくなると、取得できない
- 従来のピンング機能
 - IPFSで永続的にコンテンツを保持
 - コンテンツが取得できる割合を向上
- 提案手法
 - 従来のピンングとは異なり、PR情報をピンング
 - ピンングノード
 - IPFS上にランダムに配置
 - 常にオンライン

提案PRピンニング機能 (2)

- 人気順位 T_1 , T_2 を境に, CIDを3つのグループに分類
 - 第1層: 超人気コンテンツグループ G_1
 - 第2層: 準人気コンテンツグループ G_2
 - 第3層: 低人気コンテンツグループ G_3
- 第2層 G_2 のCIDのみ, PRのピンニング対象
 - G_1 : 多数のピアがPR情報をキャッシュ保持→ピンニング不要
 - G_2 : 要求数が多いがキャッシュにPR情報が残りにくい層
 - G_3 : 要求数が少ないので, PRピンニングの対象外



数値評価: 構成IPFS

- 一般ノード数: 10,000
- ピンニング専用ピア数: 1,000
- コンテンツ数: 10,000
- リクエスト数: 5,000
- 各ピアのキャッシュ容量: 20
- 各ピンニングの容量: 20
- ピアのオフライン率: 50%
- コンテンツの需要分布: パラメタ0.8のZipf分布
- $T_1 = 5$
- $T_2 = 205$

比較方式と評価項目

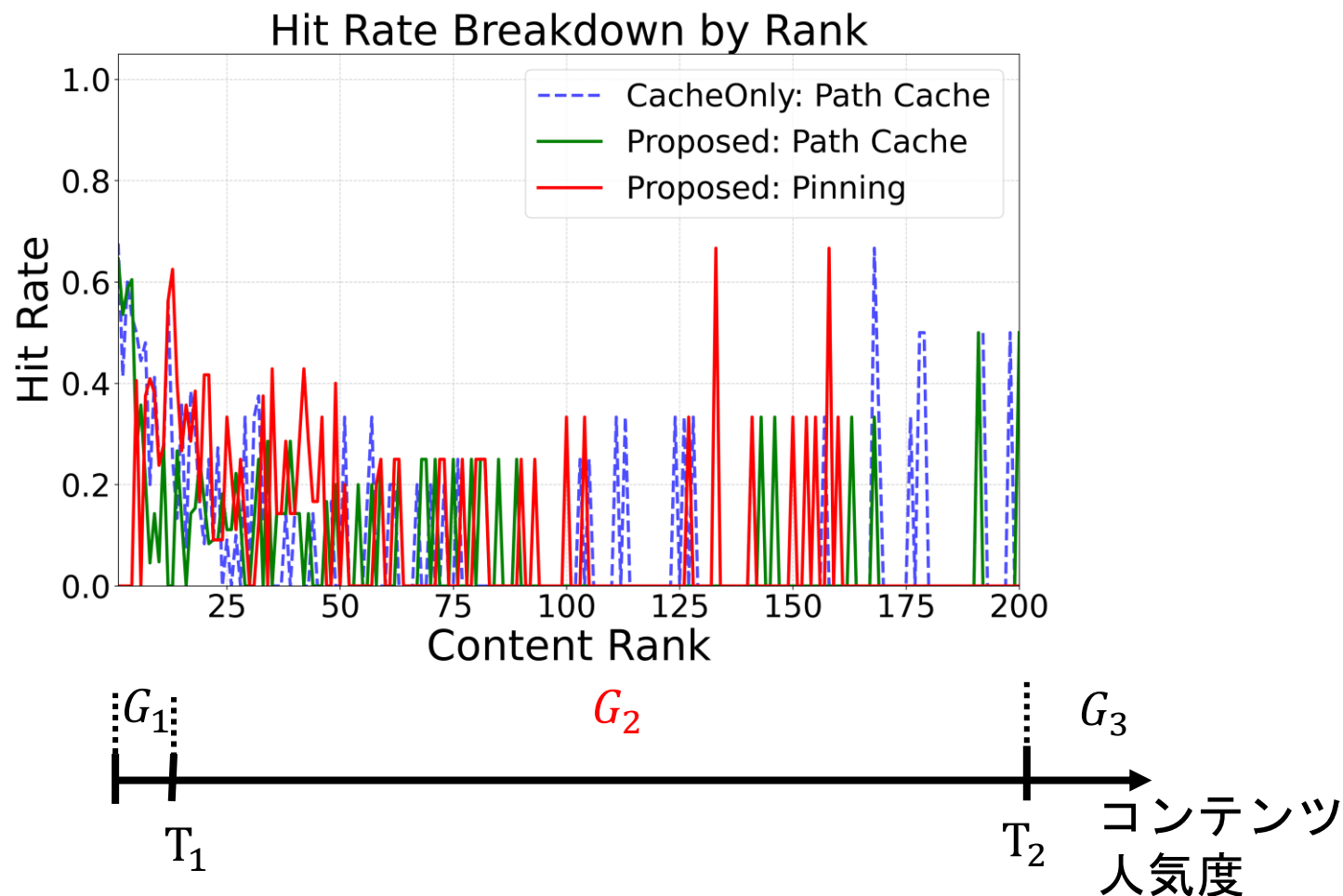
■ 比較方式

- 従来のIPFS
- 経路キャッシュ機能のみのIPFS
- 経路キャッシュ機能＋ピンング機能のIPFS (提案方式)

■ 評価項目

- キャッシュヒット率
 - 経路キャッシュとピンングの各ヒット率
- 平均ホップ長
- ネットワーク規模を拡大したことによる平均ホップ長とピンングヒット率の変化
- リクエスト分布の変化による平均ホップ長とピンングヒット率の変化

キャッシュヒット率



- 超人気コンテンツ: 経路キャッシュのみと同等のヒット率
- 準人気コンテンツ: 安定して高いヒット率を維持

平均ホップ長

■ 全体の平均ホップ長の比較

従来手法	キャッシュ	提案手法
5.23	4.99	4.75

■ 準人気コンテンツの平均ホップ長の比較

従来手法	キャッシュ	提案手法
5.49	4.85	4.22

- 提案手法は従来のIPFS, 経路キャッシュのみのIPFSと比較して, 平均ホップ長を低減
- 特にピンング対象層の平均ホップ長を大きく低減

ネットワーク規模の拡大の影響

- 1倍：一般ノード：10000, ピンング専用ノード：1000, CID数：5000
- 2倍：一般ノード：20000, ピンング専用ノード：2000, CID数：10000
- 平均ホップ長

	1倍	2倍
全体	3.6	4.23
TOP1	3.10	3.91
G_2	3.16	3.36

- キャッシュヒット率

	1倍	2倍
全体	5.59	2.02
TOP1	46.9	20.4
G_2	4.52	2.31

ネットワーク規模の拡大の影響

■ ピニング対象である G_2 のピンングヒット率

	1倍	2倍
G_2	13.7	14.2

- ノード数が増加したことによって平均ホップ長が増加
- CID数が増加したことでキャッシュヒット率が低下
- **ピンングヒット率はネットワーク規模によって変化なし**
 - 一般ノード数とピンング専用ノードの増加によるCIDのPR数が比例して増えているため(割合は変化なし), ピンング専用ノードのPRに遭遇する確率を維持

CIDリクエストのパラメータ変更

- CIDリクエストとピンング数の割当のZipf分布のパラメータを変更
 - パターン: 0.4, 0.8, 1.2の3通り
- 全体の平均ホップ長の比較

	従来手法	キャッシュ	提案手法
0.4	4.27	4.3	4.15
0.8	4.25	4.06	3.68
1.2	4.04	2.96	2.76

- 全体のキャッシュヒット率の比較

	キャッシュ	提案手法
0.4	0.58	0.706
0.8	4.90	5.59
1.2	28.16	24.9

CIDリクエストのパラメータ変更

■ ピニング対象である G_2 のピンングヒット率

	提案手法
0.4	7.44
0.8	13.7
1.2	17.7

■ キャッシュヒット率

- 提案手法とキャッシュのみの手法がほぼ同じように増加
- 人気の偏りが大きくなるほど、人気の高いCIDのリクエスト数が増加しキャッシュ数が多くなるため

■ ピニングヒット率

- 人気の偏りが大きくなるほど、 G_2 もリクエスト数が増加しピンングヒット率が増加

■ 人気の偏りが小さいほど特定の層(キャッシュ:超人気コンテンツ, ピニング:準人気コンテンツ)を対象としている提案手法は効果が低下

まとめ・今後の予定

■ まとめ

- IPFSにおいて、経路情報のキャッシュとピンングを用いた、経路平均ホップ長の削減方式を提案
- 平均ホップ長の削減とキャッシュヒット率の向上を確認
- ネットワーク規模とリクエストの人気度を変更した際の提案手法の有効性について確認

■ 今後の予定

- 経路キャッシュとピンングによるネットワーク負荷の評価