

# レポジトリを用いたIPFS可用性向上技術

## 1. 研究背景

- IPFS (InterPlanetary File System)
  - 分散型のピアツーピア (P2P) ファイルストレージプロトコルファイルをハッシュ値 (CID) で一意に識別し、内容自体をアドレスとして検索
  - ノードに長期間データを保持させることが可能
  - どのデータを保持するかは各ノードの自律的な判断で決まる
  - ノードが常にオンライン状態ではなく、オフラインになったら保持しているデータも取得できない
- ピニング
  - ピンされるデータを永続的に保持
  - 各ノードの自由意志で決まる
- Filecoin連携
  - ストレージ提供者への報酬システム「Filecoin」と連携し、ネットワーク参加のインセンティブを提供
  - ノードに長期間データを保持させることが可能

### 既存問題

- データの保存状態は予測不可能
- 災害や故障への耐障害性が低い
- 地理的に離れたノードにデータを要求する可能性があり、遅延が増加

## 2. 研究目的と提案手法

### ◆ 研究目的

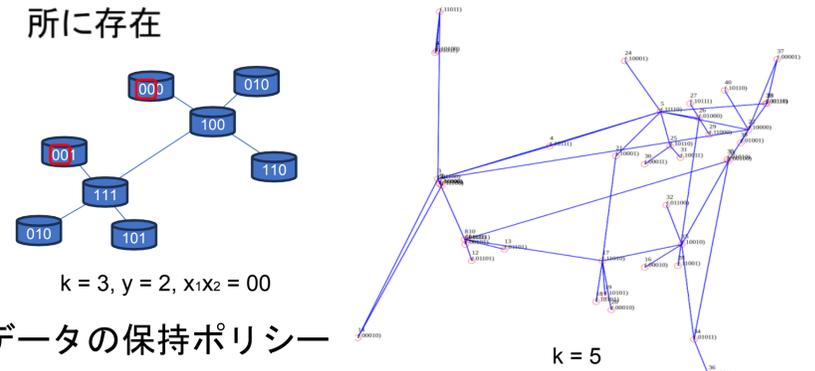
- 耐障害性
  - 同一データを地理的に離れた場所に存在するノードに重複して保持
  - 地震等の大規模災害に対する耐障害性を向上
- コンテンツの遅延時間を短縮
  - 同じデータを相対平均的にネットワーク全体に分布
  - どこから特定のデータを要求しても短い時間で入手できる

### ◆ 提案手法

- レポジトリを用いた可用性向上技術
  - authority やデータ所有者が明示的に、データを永続的に保持するノードの数や、それらの位置を制御
  - 複数の事業者が永続的にデータを保持するノード (レポジトリ) を提供
  - アルゴリズムで各レポジトリを保持させるデータを自律的に決定
- 小さい負荷でIPFSの可用性を柔軟に制御
  - レポジトリに対するNID付与技術
  - 地理的に分散した位置のレポジトリに上位ビットが同一のNIDを付与
  - NIDとCIDの一致性で保持させるデータを決定

## 3. 提案方式の動作

- NID
  - ▶ 複数  $k$  位のバイナリ数字 (例: 10011)、 $k$  はノード数に依存する ( $2^k \leq \text{ノード数} < 2^{k+1}$ )
  - ▶ 上位のビットから順番に可能な限り空間的に離れた場所に存在
- データの保持ポリシー
  - データのCIDとレポジトリのNIDの上位  $y$  ビットが一致している場合、そのデータを永続的に保持
  - $y$  の値はデータの所有者に決まる
  - $y$  が増加すると、保持するノード数が減少
- NID付与アルゴリズム
  - 貪欲法で  $S_k(X_k)$  を上位  $k$  ビットに  $X_k = (x_1, x_2, \dots, x_{1k})$  が割当てられたレポジトリの集合とすると、 $S_k(X_k)$  に含まれる各レポジトリに第  $k+1$  ビットとして 0 か 1 の値を割当てる
  - 総合コスト = ホップコスト + 距離コスト \* 重みと考えると、ネットワークトポロジーや物理距離双方を考慮し、NIDを割当てる

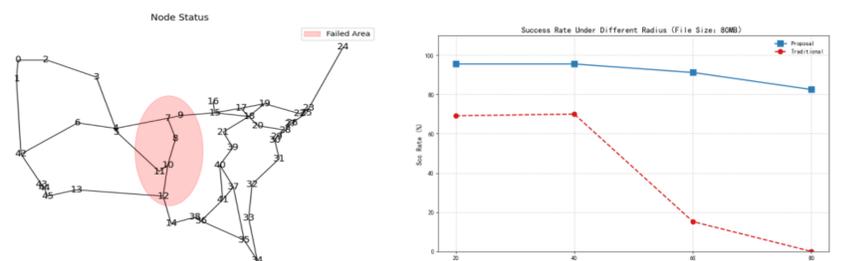


## 4. 評価方式

### ■ 可用性の測定方法

M	シミュレーション回数
N	ノード数
$\mathbb{I}_{success}$	故障半径 指示関数

$$P(R) = \frac{1}{M \cdot N} \sum_{m=1}^M \sum_{n=1}^N \mathbb{I}_{success}(m, n, R) \times 100\%$$



シミュレーターで損害範囲、提案手法と従来手法のデータ要求成功率を輸出

## 5. 今後の予定

- 違う状況で従来手法と比べると、性能がどの程度向上するか、または劣化するか、その境界線は
- 新しいノードがネットワークに参加する時の参加方式を討論
- ベースとしてアプリの実現可能性の分析を行う