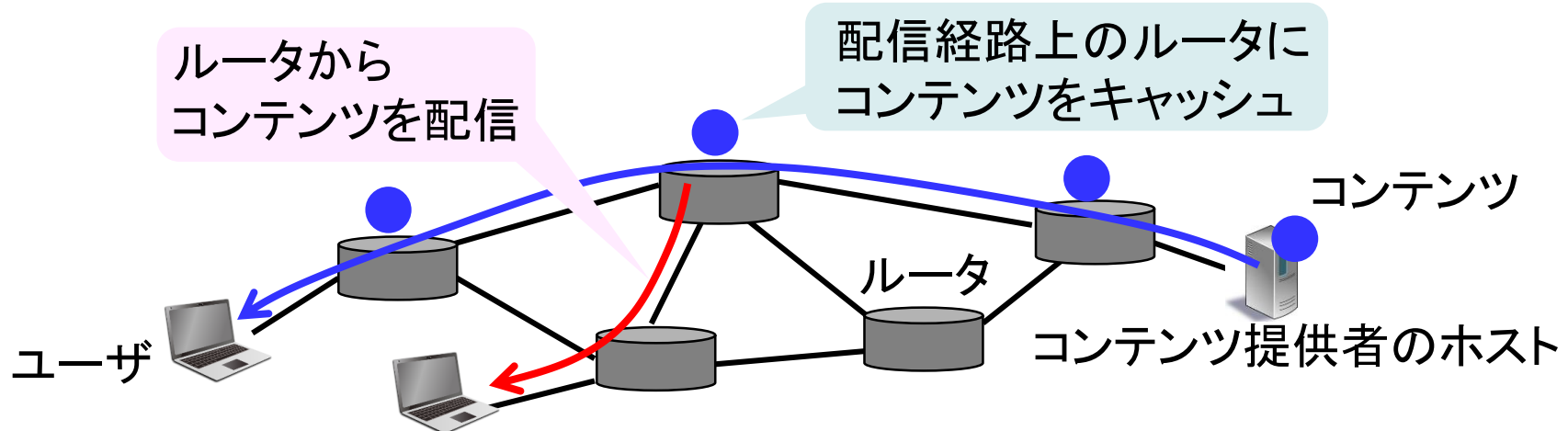

ICNにおける障害時の コンテンツ可用性回復

福岡大学 上山憲昭

ICN (information-centric networking)

- インターネットの主流サービスはコンテンツ配信
- ユーザはコンテンツに関心, しかしNWはホスト名ベース
⇒ 通信開始時に名前解決のオーバヘッドが発生
- NW事業者とは独立したCDN事業者が配信PFを提供
⇒ コンテンツ配置や配信サーバ選択の最適化が困難

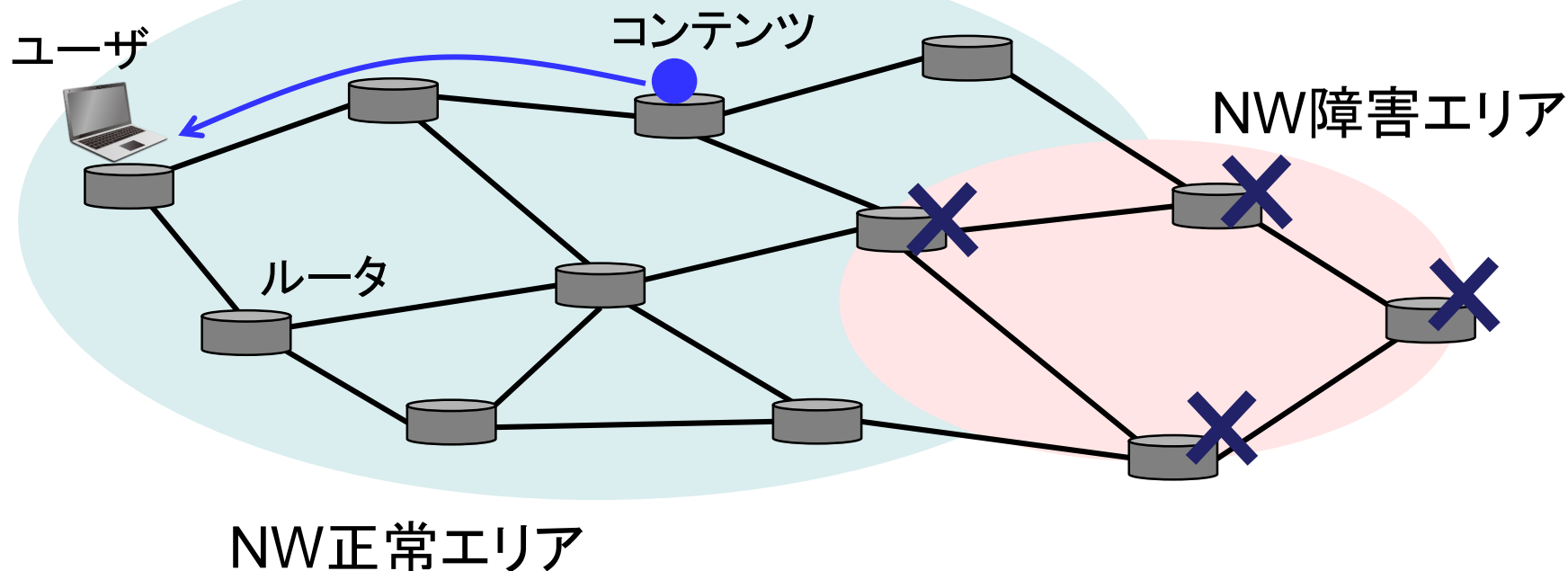
ICN (Information-Centric Network):
ルータがコンテンツをキャッシュし, 名称でルーティング



以後, CCN [Jacobson09], NDN [Zhang10]を想定

ネットワーク障害時のロバスト性

- 災害, サイバー攻撃などにより, NW障害が発生
- ICNが社会インフラとして普及するには, NW障害時にもNWの正常エリアにてコンテンツが取得できるロバスト性が必要

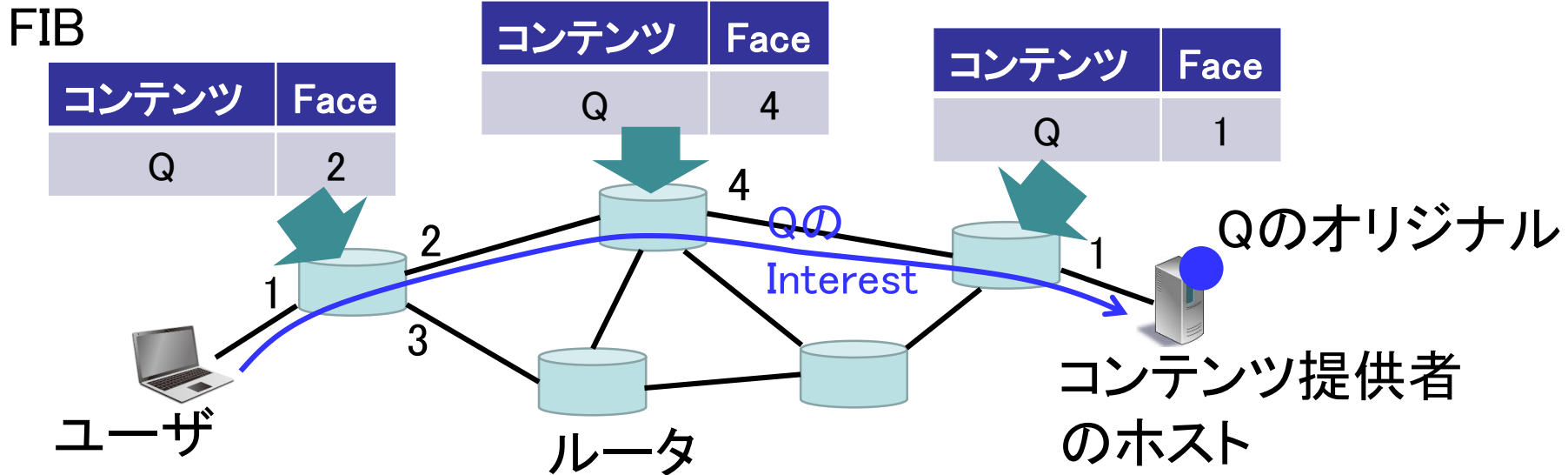


ICNの配信メカニズム

- コンテンツ提供者のホストにコンテンツのオリジナルが存在
- ユーザの配信要求 (Interest) がオリジナルに転送されるようルータの転送テーブル (FIB) を設定 [Hoque13][Jacobson09]



オリジナルへの到達性が消失すると、Interestが転送できない



2つの可用性

- レベル1：ネットワーク可用性
 - 正常ルータ間の到達性
- レベル2：コンテンツ可用性
 - 正常ルータからコンテンツへの到達性

従来の研究：ネットワーク可用性の回復

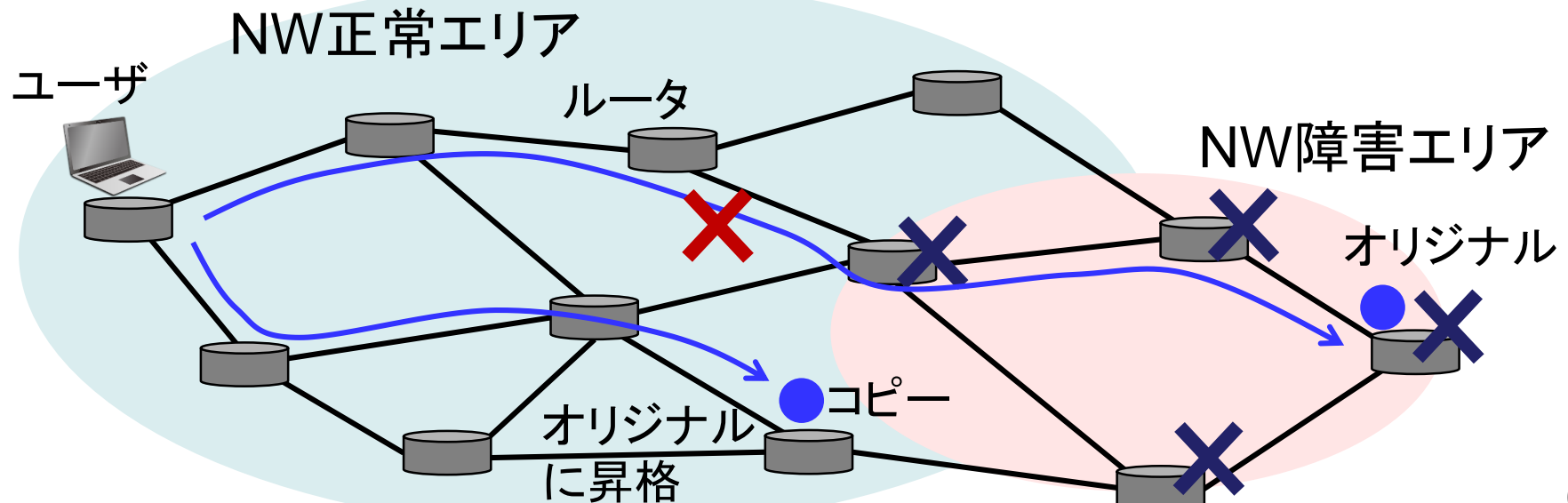


正常ルータに收容されたユーザがコンテンツを取得するには、コンテンツ可用性が必要

本研究のアイデア

- ICNではルータがコンテンツをキャッシュ
- 一部の正常ルータに、可用性が喪失したコンテンツのコピーがキャッシュされている可能性あり

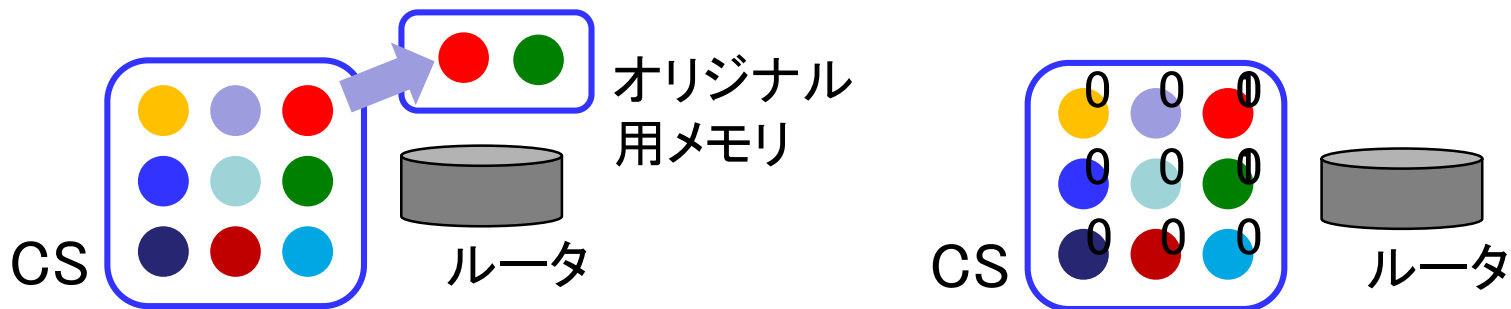
正常ルータにキャッシュされたコピーの1つをオリジナルに昇格させFIBを更新することで、コンテンツ可用性を回復



ICNルータの機能拡張

■ キャッシュの永続化機能

- オリジナルに昇格したコピーを永続的に保持する必要あり
- 2つの方法:
 - 昇格オリジナルを保持するメモリをCSとは別に用意
 - CS内の各コンテンツに対し、キャッシュアウトの有効・無効を表す永続化フラグを付与



■ Interest発生数の測定機能

- 各ルータは直前のタイムスロット内のInterest発生数を測定
⇒ オリジナル昇格位置を選択する際に使用

オリジナル昇格処理のアプローチ

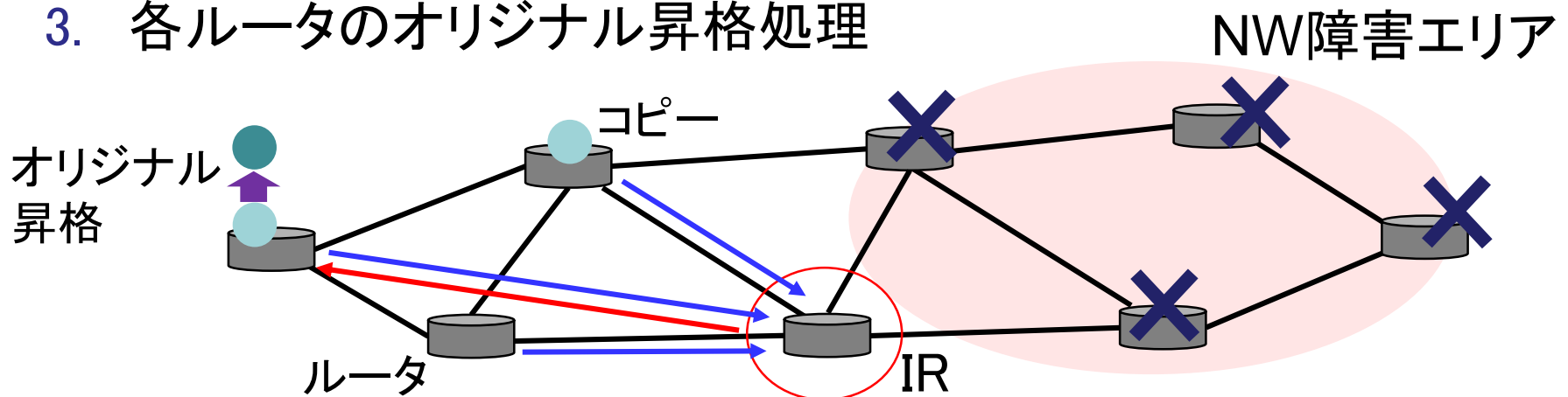
- 昇格コピー数: 単一 vs 複数
 - 単一: オリジナル収容メモリの消費量が小さくFIBのエントリ選択処理負荷が低い
 - 複数: さらなるNW障害に対するロバスト性が高い
- 昇格位置の決定者: 集中 vs 分散
 - 集中: 唯一の正常ルータで, すべての可用性喪失コンテンツのオリジナル昇格位置を選択 ⇒ 最適性が向上するが, 情報収集オーバーヘッドが必要
 - 分散: 各正常ルータが自律的にオリジナル昇格処理
- 昇格処理の開始時期: 即時 vs 経路更新後
 - 分散 ⇒ 即時
 - 集中 ⇒ ルータ間でInterestの転送が必要なため経路更新後?

本研究のオリジナル昇格処理のアプローチ

- 昇格コピー数：単一
- 昇格位置の決定者：集中
 - IR (Initiation Router): 障害発生個所に隣接する任意の1ルータ
- 昇格処理の開始時期：即時
 - InterestをフラッディングすることでFIBを用いずに情報収集

提案方式COP (Centralized Original Promotion)の手順

1. IRの情報収集(障害発生状況と各ルータのキャッシュ状態)
2. IRのオリジナル昇格位置決定
3. 各ルータのオリジナル昇格処理



隣接ルータへの接続性喪失の検知

- NLSR (Named-data Link State Routing)[Hoque13]
 - 隣接ICNルータ間で情報を交換するためのプロトコル
 - OSPFにコンテンツ名を宛先とする拡張
- Adjacency LSA (Link State Advertisements)
 - 周期的(10秒)に隣接ルータに対しInfo Interestを送付
 - 一定回数(4回), 応答がない場合に, 接続性の喪失を認識
 - 確認間隔を500ミリ秒程度に減らしてもルータの処理負荷の増加は問題なく[Basu01], 2秒程度で障害の検知が可能[Goyal12]
 - 障害検知時はInfo Interestを隣接ルータにブロードキャストし, 各受信ルータはInfo Interestのブロードキャストを反復することで, すべてのルータがNWTポロジの変化を認識

NLSRの利用を想定

最初に隣接ルータとの接続性喪失を検知したルータがIR

IRの情報収集

- R_a : IRからの到達性がある正常ルータの集合

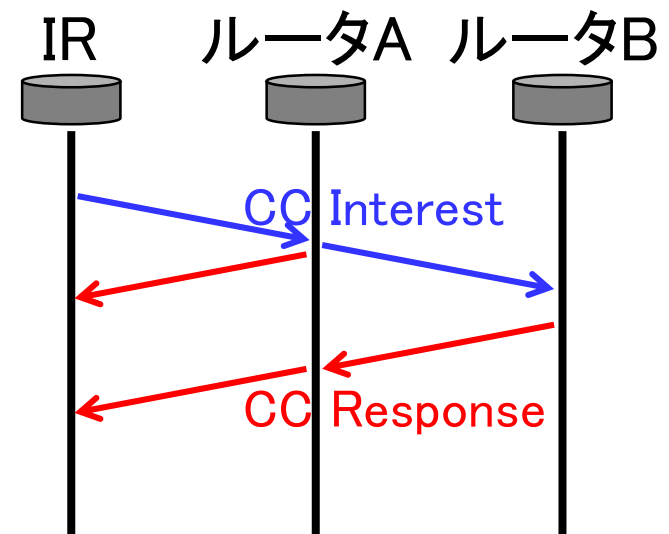
- R_f : IRからの到達性が喪失したルータの集合

⇒ IRは R_f を知ることによって可用性喪失コンテンツ M_f を把握

IRは R_a の各ルータでキャッシュされたコンテンツと、 M_f (R_f)に関する情報を収集

IRが情報を収集するための処理手順

1. IR⇒他ルータ: CC (Connectivity Confirmation) Interest送信
2. 他ルータ⇒IR: CC Response返信



ExplorationによるCC Interestの転送

- 各ルータを宛先にFIBを用いてCC Interestを転送すると、NW障害のため宛先ルータに届かない可能性あり



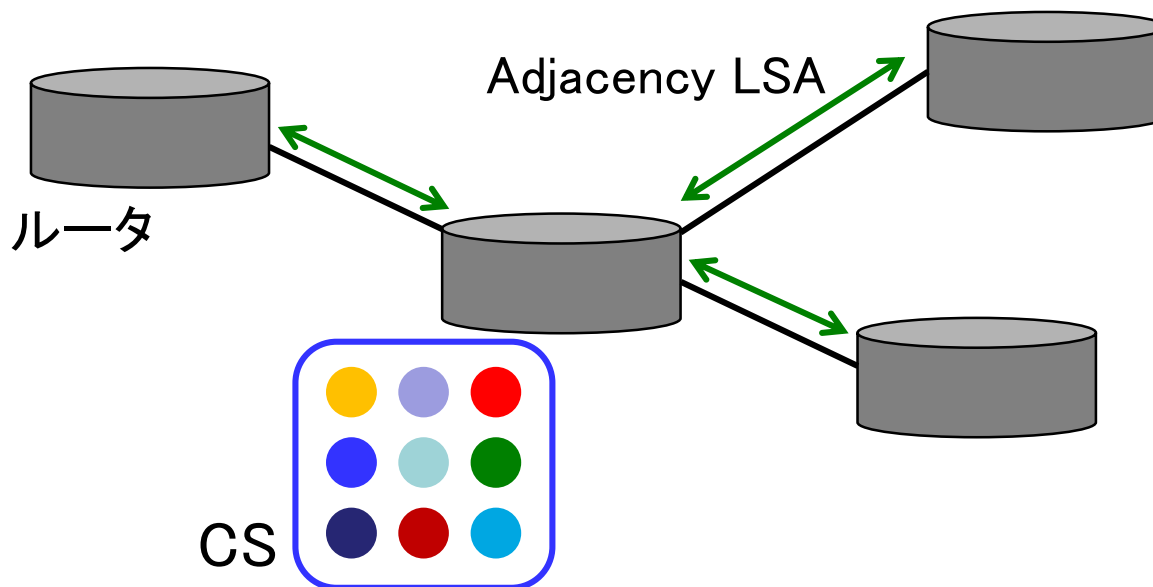
■ Explorationを用いてCC Interestを転送 [Chiocchetti12]

- IRは識別名称と現在時刻(タイムスタンプ)を付与したCC InterestをすべてのFaceに送信
- CC Interestを受信したルータの処理
 - PITに該当CC Interestのエントリが存在する場合は棄却
 - PITに該当CC Interestのエントリが存在しない場合:
 - CC Interestの送信元ルータの名称と、到着Face番号をPITの新しいエントリとして作成
 - CC Interestの到着Face以外のすべてのFaceにCC Interestを転送

同一CC Interestの重複した転送を回避することで、CC Interestのループ生成を回避

CC Responseの送信

- CC Interestを受信したルータの処理
 - Adjacency LSAにより, 全隣接ルータとの接続性を確認
 - CC ResponseをIRに向けて送信(以下の情報を含む)
 - 自身のCS内にキャッシュしているコンテンツのリスト
 - 接続性が確認された隣接ルータのリスト
 - 直前のタイムスロット内の発生Interest数 Y_n



CC Responseの転送

- CC ResponseはICNのデータチャンクに相当
- ICNルータのデータチャンク到着時の処理
 - PITの該当エントリに記載されたFaceにチャンクを転送し, PITの該当エントリを削除
 - CC Interestは全ルータに送付され, 同一PITエントリは複数のルータが送出したCC Responseに対し共通して使用

一定期間, PITのエントリを存続

- 他のルータが送出したCC Responseを受信したルータの処理
 - 初めてCC Responseが到着してから一定時間 T_p が経過した時点でPITの該当エントリを削除
 - ルータの隣接ルータとの接続性確認に要する時間(2秒程度)と比較して, CC Interest/Responseの転送遅延時間は遥かに小さいことから, マージンを含めて T_p を3秒程度に設定

IRによるオリジナル昇格位置の選定

- IRはCC Interestの送信から一定時間 T_c 内にCC Responseが到着しないルータを, IRからの到達性が喪失したルータと判断
 - 例えば, 総リンク数100, リンクの最大長1,000km, ルータのCC Interest/Response転送に要する時間100 μ 秒程度の場合, T_c を3秒程度に設定
- IRは R_a の各ルータでキャッシュされたコンテンツと, 可用性喪失コンテンツ M_f を認識 $\Rightarrow R_a$ の中で可用性喪失コンテンツ m をキャッシュしているルータの集合 C_m を把握



複数のルータが C_m に含まれる場合,
その中の1つを選択

オリジナル昇格位置の選定方法

- オリジナル昇格位置が、可用性回復後の配信フロー長やリンク・ルータ負荷に影響
- オリジナル昇格位置と R_a のルータへの配信要求数比で重みづけた平均ホップ長が最小となるよう、 M_f の各コンテンツ m のオリジナル昇格ルータ x_m を次式で選択

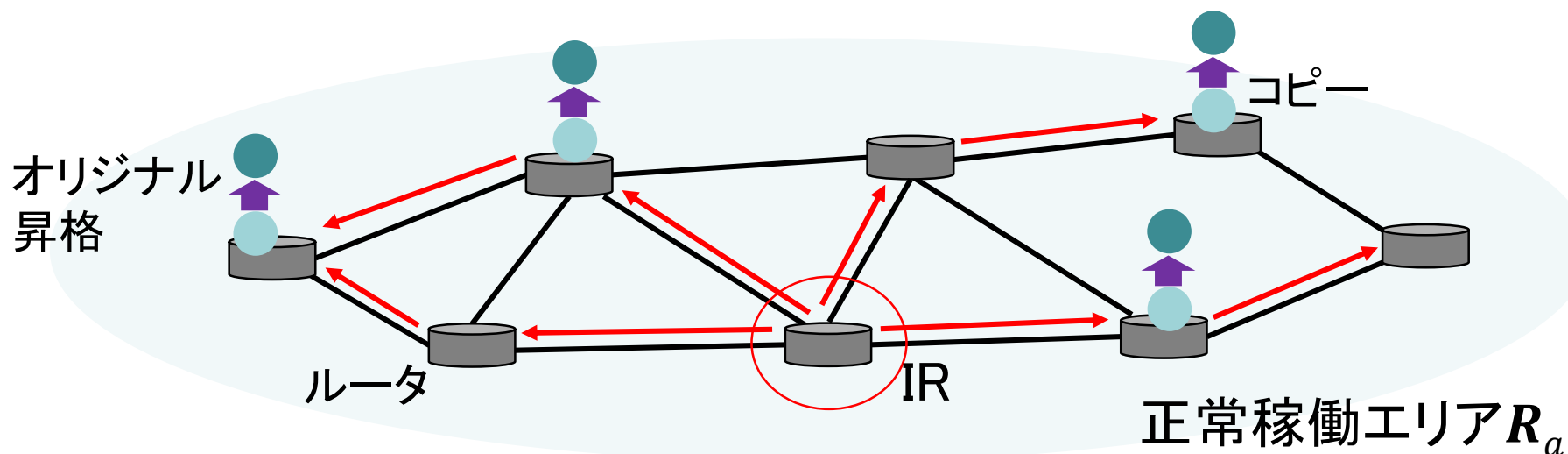
$$x_m = \arg \min_{x \in C_m} \sum_{n \in R_a} h_{n,x} r_n$$

- $h_{n,x}$: ルータ n からルータ x への最短ホップ長
- r_n : ルータ n のInterest生成比率 (Y_n より算出)

$$r_n = Y_n / \sum_{j \in R_a} Y_j$$

オリジナル昇格処理

- IRは, M_f の各コンテンツに対するオリジナル昇格ルータ x_m の決定内容を, OPL (Original Promotion List) Interestを用いて, R_a の全ルータに通知 (Explorationにより転送)
- OPL Interestの受信ルータ n の処理:
 - $x_m = n$ の各コンテンツ m をオリジナルに昇格 (専用メモリに移動 or 永続化フラグをセット)
 - Prefix LSAを広告 \Rightarrow オリジナル昇格コンテンツに対し R_a のルータからInterestの転送が可能となり, **コンテンツ可用性が回復**



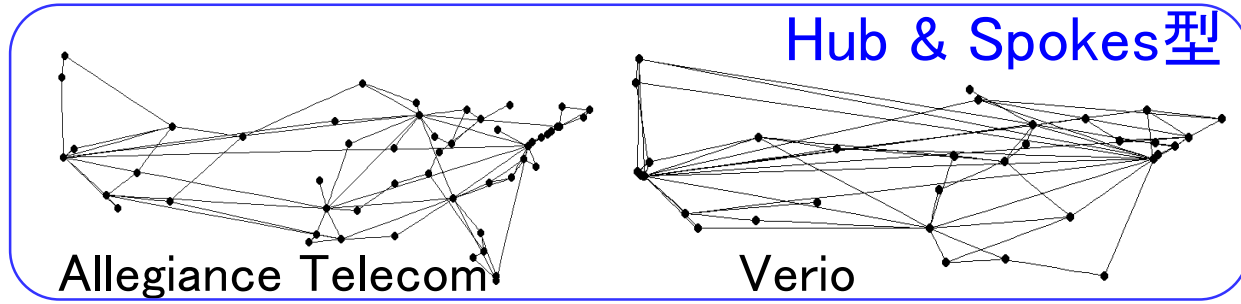
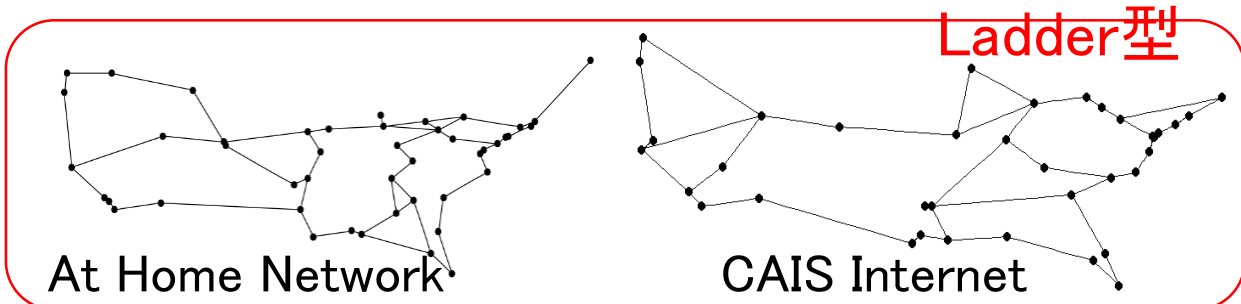
複数IRの重複処理の回避

- 障害エリアの隣接正常ルータが複数ある場合、複数のルータがIRとしてオリジナル昇格位置を決める可能性あり
- 以下の方法でIRを単一ルータに限定
 - CC Interest送出**前**に他のルータからCC Interestを受信した場合、CC Interest送出を中止
 - CC Interest送出**後**に他のルータからCC Interestを受信した場合、そのタイムスタンプ自身のものより古い場合は、オリジナル昇格処理を中止

最初にCC Interestを送信したルータのみがIRとしてオリジナル昇格処理を実施

性能評価条件

■ NWTポロジ: 米国の4つの商用ISPバックボーンNW



Network Name	N	E	D_0
At Home Network	46	55	6.83
CAIS Internet	37	44	5.49
Allegiance Telecom	53	88	2.81
Verio	35	72	2.23

N :ノード数, E :リンク数
 D_0 :人口比で重みづけた平均ノード間ホップ距離

- コンテンツ数: $M = 10,000$
- 各ノードの配信要求数比率: 各ノード n の人口比 p_n
- コンテンツの要求数比率: パラメタ $\theta = 0.8$ のZipf分布 (全ノードで同一)
- 各ノードのキャッシュ容量: 同一容量 $B = 100$ (M に対する比率 $b = 0.01$)
- オリジナル配置: 各コンテンツを確率 p_n で選択したノードに配置(異なるシードで10回反復した平均値で評価)

比較キャッシュ方式

AllCache

配信経路上の全ルータでキャッシュ



EdgeCache

配信要求ルータでキャッシュ



UniCache

配信経路ホップ数Dに対し確率1/Dでキャッシュ



ProbCache

配信ルータからの距離に比例する確率でキャッシュ



LCD

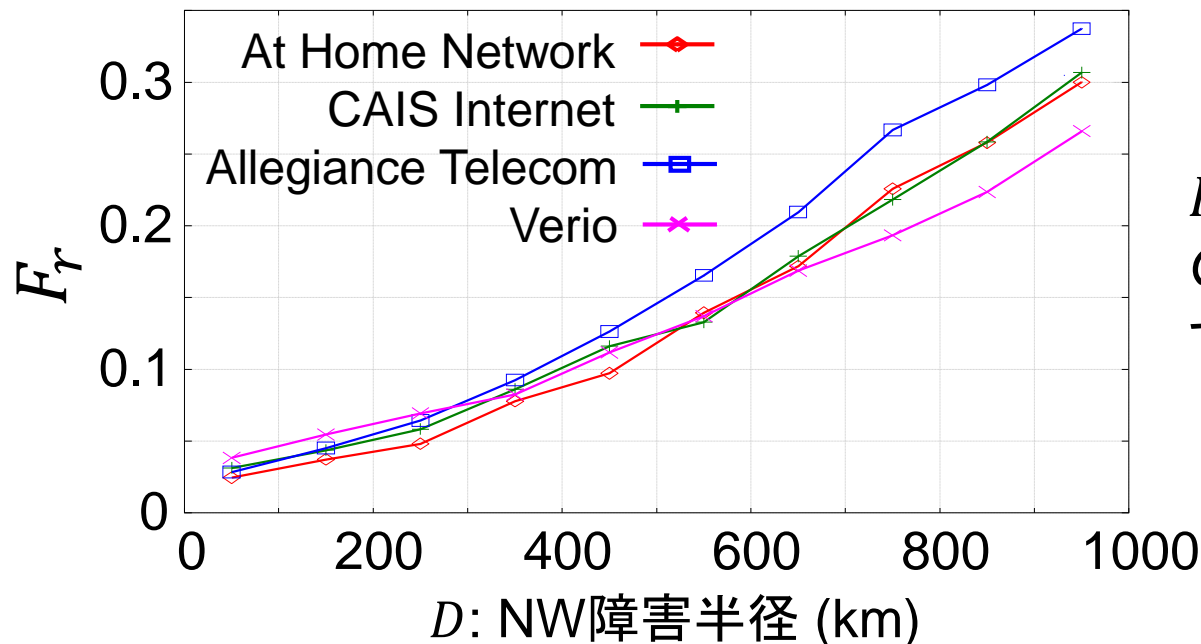
配信ルータの1ホップ下流のルータでキャッシュ



(Leave Copy Down)

ネットワーク障害

- 配信要求を10万回発生させた時点で、ルータ n を中心とした半径 D kmの円内に存在する全てのルータに障害を発生
- 複数の領域にNWが分断された場合は、正常ルータ数が最大のNW領域のみを評価対象 $\Rightarrow R_a$ の構成ルータ
- 与えられた D に対し、 N の各ルータを中心としたNW障害を発生させる処理を反復し、平均値で評価 ($D = 500$ km)

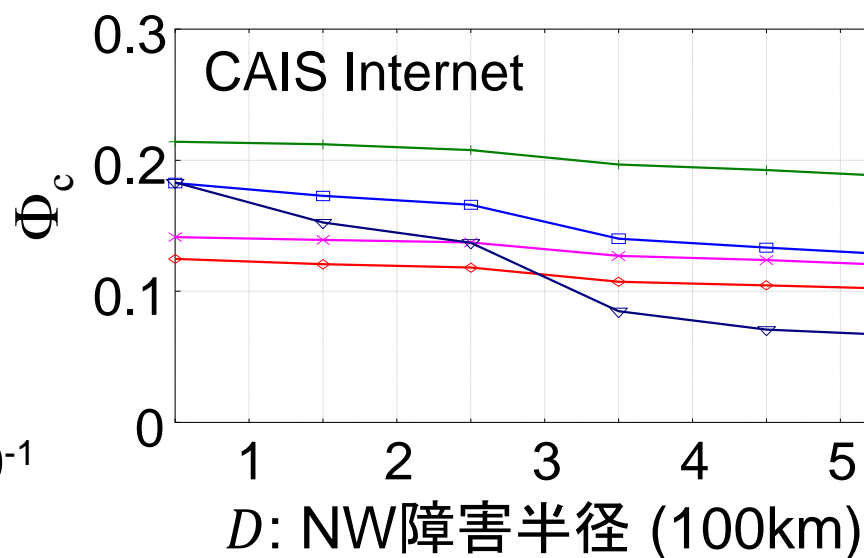
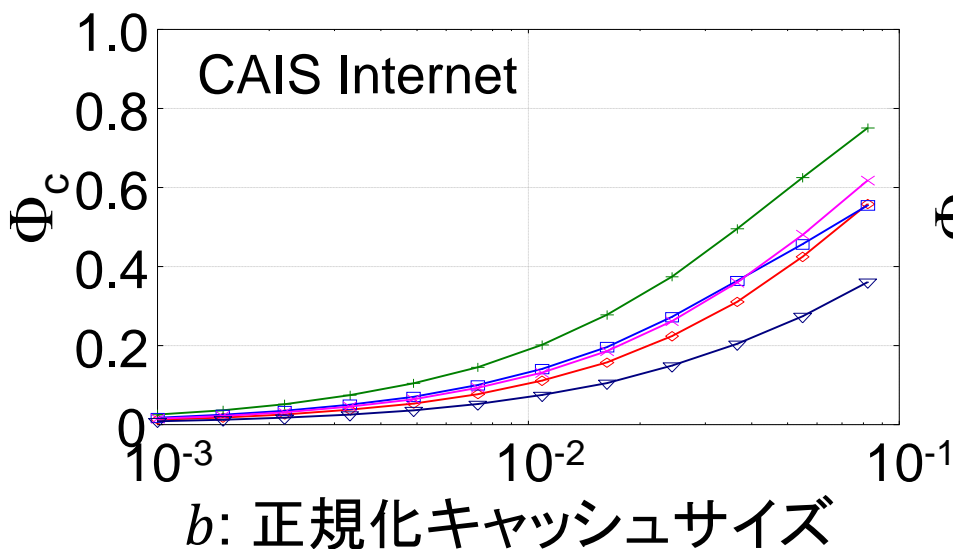


F_r : 非到達ルータ数の全ルータ数 N に対する比率

コンテンツ可用性の回復効果

Φ_c : 可用性喪失コンテンツの中で, 1つ以上の R_a でキャッシュされているコンテンツの割合

—●— AllCache —■— EdgeCache —□— UniCache —×— ProbCache —▽— LCD

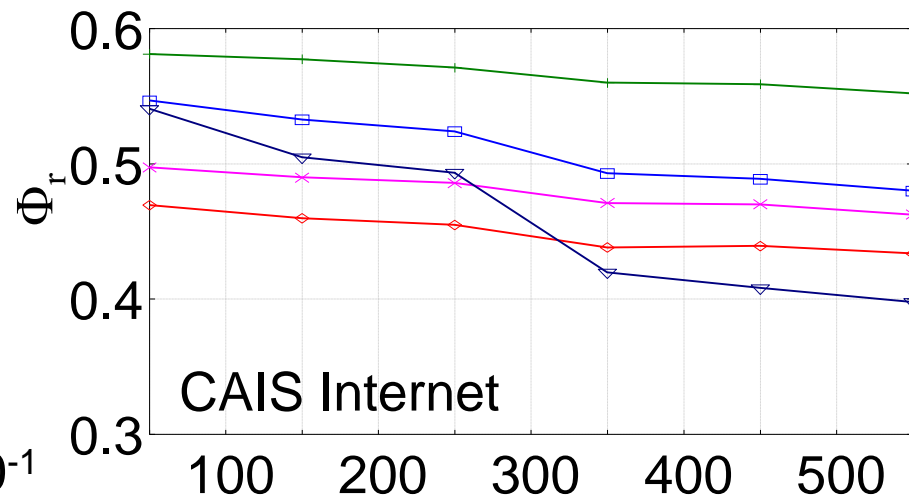
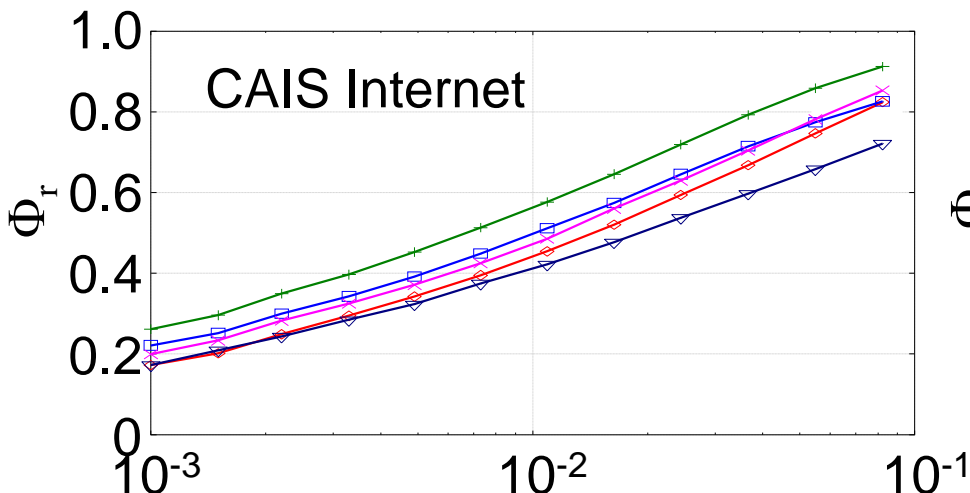


- LCD: オリジナル周辺でキャッシュするため Φ_c が小さい
- EdgeCache: ユーザ収容ルータでキャッシュするため Φ_c が大きい
- 例えば $b = 0.01$, $\theta = 0.8$, $D = 500$ km の場合, 提案方式を用いることで, **5%~20%程度の可用性損失コンテンツの可用性を回復**

配信要求成功確率の向上効果

Φ_r : コンテンツ可用性回復後に, 可用性喪失コンテンツに対する配信要求の中で配信が可能となったものの割合

—●— AllCache —+— EdgeCache —□— UniCache —×— ProbCache —▽— LCD

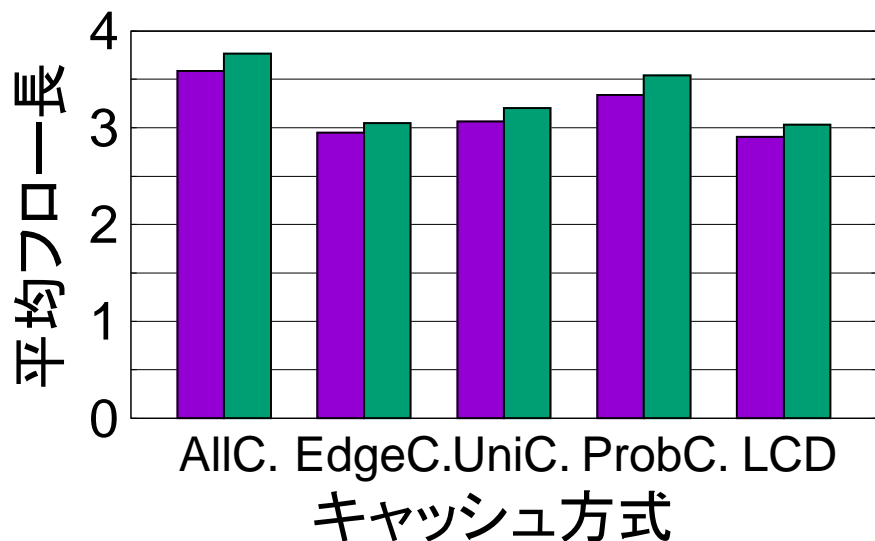


- 高人気コンテンツほど, 多数のルータでキャッシュされるため NW障害時に R_a のルータにコピーが存在する可能性が高い
- Φ_r は Φ_c より大きく, $b = 0.01$, $\theta = 0.8$, $D = 500$ kmの場合, 提案方式を用いることで, **30%~60%程度の配信要求を救済**

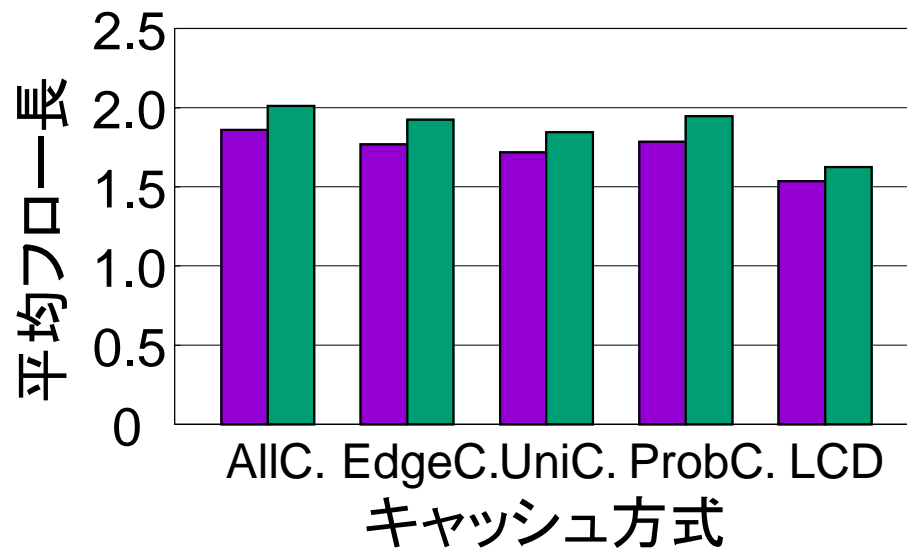
可用性回復後の平均フロー長

■ 提案方式

■ ランダム方式: C_m の中からランダムに1つ選択



CAIS Internet



Allegiance Telecom

■ ランダム方式と比較して、平均フロー長を3%~8%程度低減

まとめ

- ICNでは、ユーザの配信要求 (Interest) がオリジナルに転送されるようルータの転送テーブル (FIB) を設定
- オリジナルへの到達性が消失するとInterestが転送できない
- 本研究の目的:
 - 正常ルータにキャッシュされたコピーの1つをオリジナルに昇格させFIBを更新することで、コンテンツ可用性を回復
 - 単一のルータが全てのオリジナル昇格位置を決定するプロセスとアルゴリズムを提案
- 提案方式の効果:
 - 5%~20%程度の可用性損失コンテンツの可用性を回復
 - 30%~60%程度の配信要求を救済
 - ランダム方式と比較して平均フロー長を3%~8%程度低減

関連研究

- [Basu01] A. Basu and J. Riecke, Stability issues in OSPF routing, ACM SIGCOMM 2001
- [Chiocchetti12] R. Chiocchetti, D. Rossi, and G. Carofiglio, Exploit the Known or Explore the Unknown? Hamlet-Like Doubts in ICN, ICN 2012
- [Goyal12] M. Goyal, M. Soperi, E. Baccelli, G. Choudhury, A. Shaikh, H. Hosseini, and K. Trivedi, Improving Convergence Speed and Scalability in OSPF: A Survey, IEEE Communications Surveys & Tutorials, Vol. 14, No. 2, pp. 443–463, Second Quarter 2012
- [Hoque13] A. Hoque, S. Amin, A. Alyyan, B. Zhang, L. Zhang, and L. Wang, NLSR: Named-data Link State Routing Protocol, ACM ICN 2013
- [Jacobson09] V. Jacobson, D. K. Smetters, J. D. Thornton, M. Plass, N. Briggs, and R. L. Braynardet, Networking Named Content, ACM CoNEXT 2009
- [Zhang10] L. Zhang, D. Estrin, J. Burke, V. Jacobson, J. D. Thornton, D. K. Smetters, B. Zhang, G. Tsudik, K. Claffy, D. Krioukov, D. Massey, C. Papadopoulos, T. Abdelzaher, L. Wang, P. Crowley, and E. Yeh, Named Data Networking (NDN) Project, Technical Report NDN-0001, 2010

謝辞：本研究の成果は、JSPS科研費18K11283により得られたものである。