

# ICN 障害時のコンテンツ可用性の集中的回復制御

上山 憲昭<sup>†</sup>

<sup>†</sup> 福岡大学工学部 〒 814-0180 福岡県福岡市城南区七隈 8-19-1  
E-mail: [†kamiyama@fukuoka-u.ac.jp](mailto:†kamiyama@fukuoka-u.ac.jp)

あらまし インターネットでは動画視聴や Web 閲覧などのコンテンツ配信が全トラフィックの約 90% を占めている。そこでルータにキャッシュを設置しネットワーク (NW) 内でコンテンツの配信元を選択する ICN が、効率的にコンテンツを配信する NW として国内外で精力的に研究されている。ICN が社会インフラとして広く普及するには、震災・人災等による NW 障害時にも残った正常ルータ範囲でコンテンツの取得が継続して行える必要がある。そのため正常ルータへの到達性 (NW 可用性) に加えて、正常ルータにコンテンツが存在し配信要求が届くこと (コンテンツ可用性) が求められる。ICN ではコンテンツのオリジナル提供者のホストに配信要求が届くようルータの転送表 (FIB) が構築されるため、NW 障害時にオリジナルへの到達性が喪失したコンテンツに対する配信要求は転送できず、コンテンツ可用性が喪失する。ところが ICN では可用性が喪失したコンテンツのコピーが正常ルータにキャッシュされている可能性がある。そこで本稿では、キャッシュコピーの 1 つをオリジナルに昇格させ、新たなオリジナルに配信要求が届くよう FIB を更新することで、新たにバックアップ装置を用意することなくコンテンツの可用性を回復させることを提案する。また複数の正常ルータにコピーが存在する場合、正常ルータのオリジナルまでの平均距離が最小となるオリジナル昇格位置の集中的な選択法を提案する。

キーワード ICN, 障害回復, コンテンツ可用性

## Concentrated Recovery Method of Content Availability at Network Failure in ICN

Noriaki KAMIYAMA<sup>†</sup>

<sup>†</sup> Faculty of Engineering, Fukuoka University 8-19-1, Nanakuma, Jonan-ku, Fukuoka, 814-0180  
E-mail: [†kamiyama@fukuoka-u.ac.jp](mailto:†kamiyama@fukuoka-u.ac.jp)

**Abstract** In the Internet, traffic caused by content delivery, e.g., video streaming and web browsing, dominates about 90 % of all the traffic. Hence, ICN (information-centric networking) in which content items are delivered from cache memories provided at routers in the networks has gathered wide attention as a new network architecture efficiently delivering content. To spread ICN widely as a social infrastructure, ICN needs to achieve high sustainability, i.e., enabling users at normally-operated routers to obtain content. Therefore, ICN is required to sustain not only the network availability, i.e., connectivity to normally-operated routers, but also the content availability, i.e., reachability to content. In ICN, FIBs (forwarding information bases) at routers are constructed so that the content requests reach the hosts of content providers having the originals of content items. Hence, the requests for content items whose connectivity to originals is lost cannot be transferred, and the content availability of these content items is lost. However, copies of content whose content availability is lost are possibly cached at one or more normally-operated routers in ICN. Therefore, in this paper, we propose to recover content availability without providing backup systems by promoting one copy at normally-operated routers to original and updating FIBs so that content requests are transferred to the newly promoted original. We also propose a method of selecting one copy with the minimum average distance to other normally-operated routers when copies are cached at multiple normally-operated routers.

**Key words** ICN, fault recovery, content availability

### 1. はじめに

Web コンテンツや, YouTube に代表される UGC (User Generated Content), コンテンツ事業者 (CP) が作成する映画やドラマ等の大容量リッチコンテンツなど, デジタルコンテンツの配信がインターネットのトラフィックの 9 割以上を占めている [4]. インターネットでは IP アドレスを用いてパケットをルーティングするため, 配信に先立ち配信元となるサーバの IP アドレスをコンテンツの名称から解決するオーバーヘッドが発生する。そこでコンテンツをネットワーク (NW) 内のルータでキャッシュし, コンテンツの名称を用いて配信要求を転送し, 転送経路上で要求コンテンツを保持しているルータからコンテンツを配信する Information-Centric Networking (ICN) が注目され, 精力的な研究が進められている [7]. ICN の思想を用いた具体的な NW アーキテクチャとしては, TRIAD [11], Content-Centric Networking (CCN) [13], Data-Oriented Network Architecture (DONA) [14], Named Data

Networking (NDN) [23] など様々なものが提案されている [21]. ICN を用いることで名前解決処理のオーバーヘッドを回避し, ユーザに近い位置からコンテンツを配信することで遅延時間や NW 負荷の低減が期待される。

事故や人為的ミス, サイバー攻撃等により, NW を構成するルータやリンクには常に障害が発生する可能性がある。また社会インフラとしての情報 NW には公共性の高い情報や緊急情報など, 重要な情報を人々に確実に届けることが求められる。そのため ICN が社会インフラとして広く普及するには, NW 障害時にも残った正常ルータでコンテンツの取得が継続して行える高いロバスト性が求められる。ロバスト性を維持するには, 正常ルータ間でパケットが交換できるよう任意の正常ルータへの到達性 (NW 可用性) が求められる。また ICN ではコンテンツの提供者 (Publisher) のホストに存在するオリジナルに配信要求 (Interest) が届くようルータの経路表 (FIB: Forwarding Information Base) が構築されるため, NW 障害時に Publisher への到達性が喪失したコンテンツに対する配信要求は正常ルータ範囲においても転送できない。そのため正常ルータ範囲でコ

コンテンツの取得を継続して行うには、NW 可用性に加えて正常ルータにオリジナルが存在し取得できること (コンテンツ可用性) が求められる。

NW 可用性の回復に関しては、インターネットの経路制御方式である OSPF (Open Shortest Path First) にコンテンツ名を宛先とする拡張を図り、障害時はルータが経路情報の広告により経路を再構築する NLSR (Named-data Link State Routing) が検討されている [12] [13]。また従来のインターネットでは CDN 事業者がコンテンツのオリジナルを管理してバックアップサーバを用意し、障害発生時には CDN 事業者の DNS サーバにおける配信サーバ選択によりバックアップから配信することでコンテンツ可用性の回復が可能である [17]。しかし ICN では Publisher が NW にコンテンツの提供を登録することで Subscriber の Interest が Publisher にルーティングされる [21]。このようにコンテンツのオリジナルは Publisher のホストに存在するため、NW 障害時に接続性が喪失したルータに収容されたオリジナルへの到達性が喪失する。ICN の障害時にコンテンツ可用性を回復させるため、NW がコンテンツの配信元を明示的に選択する PSIRP/PURSUIT を対象に、Al-Naday らは同一コンテンツを提供する他の Publisher のオリジンサーバに配信元を切り替えることを提案している [1]。しかし複数の重複したオリジナルが存在しない場合には可用性を回復できない。また Sourlas らはコンテンツを取得したユーザ端末をオリジンサーバとして活用することで、NW 障害時にコンテンツ可用性を維持することを提案している [19]。しかしルータにユーザ端末に Interest を誘導するためのテーブルが必要になり、またユーザの協力が必要となる。

ICN ではオリジナルへの到達性が喪失したコンテンツのコピーが正常ルータにキャッシュされている可能性がある。そこで本稿ではこれらのコンテンツに対し、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キャッシュに永続に残すとともに、昇格ルータに Interest が転送されるよう全ての正常ルータの FIB を更新することで、新たにバックアップ装置を用意することなくコンテンツ可用性を回復させることを提案する。可用性喪失コンテンツのコピーが複数の正常ルータにキャッシュされている場合、どのルータのコピーをオリジナルに昇格させるかを選択する必要があるが、NW トラフィック量を削減するには配信フローホップ長を低減することが望ましい。そこで障害ルータに隣接するルータの 1 つが正常ルータのキャッシュコンテンツと隣接ルータの稼働状態に関する情報を収集し、他の正常ルータへの平均ホップ長が最小となる位置に存在するコピーをオリジナル昇格コピーとして選択することを提案する。商用 ISP トポロジを用いた計算機シミュレーション評価により、提案方式の有効性を確認する。例えば各ルータにコンテンツ数の 1% の容量のキャッシュが実装されている場合に、15% 程度のコンテンツの可用性が NW 障害により喪失した際、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が回復し、可用性喪失コンテンツの 30%~60% 程度の配信要求が救済されることを示す。

## 2. ICN と NLSR

### 2.1 ICN ルータの基本機能

本稿では CCN もしくは NDN のアーキテクチャを想定する。コンテンツを取得したいユーザは Interest をオリジナル保有サーバ (オリジンサーバ) に向けて送出し、ルータはコンテンツの名称を用いて Interest を転送する。Interest の転送先はルータに用意された FIB の該当エントリを参照することで決定されるが、IP ルーティングに用いられる OSPF と同様、リンクコストの総和が最小となる経路をダイクストラ法で算出することで FIB を設定する [12] [13]。ただしコンテンツの名称に基づき FIB をルックアップする点が OSPF とは異なる。FIB によって決定された Interest の転送経路は default path と呼ばれる。FIB を用いて Interest を転送する方法以外に、Interest が到着した Face 以外の全ての Face に Interest をブロードキャストすることで OSPF により設定した経路以外のより品質のよい経路を適用的に用いるアプローチも検討されているが [8] [9]、Interest が大量に NW 上に生じる問題がある。

ルータにはコンテンツ<sup>(注1)</sup>をキャッシュするためのメモリ (CS: Content Store) が実装されており、Interest 転送経路上のルータは要求コンテンツをキャッシュしている場合には Interest を次ホップルータに転送しないで棄却し要求コンテンツを要求

元に配信する。ルータには Interest の到着 Face 番号を記録する PIT (Pending Interest Table) と呼ばれるテーブルが管理され、PIT を参照することで Interest が転送された経路を逆方向にコンテンツが要求ユーザまで配信される。PIT のエントリはコンテンツを配信した際に消去される。

コンテンツ配信経路 (default path) 上のルータは転送されてきたコンテンツをキャッシュすることが可能である。実際にキャッシュするか否かを決定するためのキャッシュポリシーとしては、default path 上の全てのルータがキャッシュすることや [13]、配信フローホップ長の逆数の確率で default path 上の各ルータがキャッシュするなど [6]、様々な方式が提案されている。CS の空き容量が不足する場合には、CDN で一般的に用いられている、最後に要求されてからの経過時間が最大のコンテンツを削除する LRU (Least Recently Used) などのキャッシュ置換方式を用いて空き容量を確保し、新たに取得したコンテンツをキャッシュする。

### 2.2 ICN ルータの機能拡張

前節で述べた ICN ルータの基本機能に加え、本稿では新たに (i) Interest 発生数の測定機能、(ii) キャッシュの永続化機能、の 2 つの機能拡張を想定する。ルータ  $n$  は自身が収容するユーザからの任意のコンテンツに対する Interest を受信することに Interest 観測カウンタ  $y_n$  をインクリメントする。固定長  $T_m$  の連続したタイムスロット (TS) を考え、 $T_m$  が経過するごとに  $y_n$  を直前の TS の Interest 発生数  $Y_n$  として記録し、 $y_n$  をゼロに初期化する。このようにしてルータ  $n$  は直前の TS 内の Interest 発生数  $Y_n$  を記憶する。

CS にキャッシュされたコンテンツはキャッシュ置換方式によって削除される可能性がある。そのため CS にキャッシュされたコンテンツの中でオリジナルに昇格したものを永続的に保持するための方策として、(i) オリジナル昇格コピーを永続的に保持するためのメモリを CS とは別に用意、(ii) 新たなメモリを用意せず CS にキャッシュされた各コンテンツに対しキャッシュアウトの有効・無効を表すフラグを付与、のいずれかを用いる。(ii) の場合、通常のキャッシュコンテンツに対しては本フラグをオンにしてキャッシュアウトを許容するが、オリジナル昇格コンテンツに対しては本フラグをオフとしキャッシュアウトを許容しないことで、CS 内での永続的なキャッシュが可能となる。別途、メモリを用意しなくてよいので実装コストが低い反面、CS の一部をオリジナル収容領域として使用するため CS の容量が減少する。

### 2.3 NLSR による FIB の更新

ルータ間で各種情報を交換するためのプロトコルとして、文献 [12] で提案された NLSR による FIB 更新や、各ルータの隣接ルータ稼働性確認機能を想定する。IP ルータは Hello メッセージと呼ばれる制御パケットを周期的に隣接ルータ間で交換することで、各ルータは隣接ルータの稼働状態を確認し NW 全体のトポロジを把握して FIB を設定する [10]。ICN では Interest をコンテンツの名称を用いて転送するため、制御メッセージを隣接ルータ間で交換するには各ルータに名称を付与する必要がある。そのため NLSR では各ルータに名称を付与し、通信相手のルータの名称を宛先とする Interest を用いて隣接ルータ間でメッセージを交換する [12]。NLSR は OSPF と同様、隣接ルータ間でトポロジ情報を交換する Adjacency LSA (Link State Advertisements) と、FIB のエントリ情報を交換する Prefix LSA を用いて FIB を更新する。本稿でもルータ間での各種制御情報の交換のため、NLSR の使用を想定する。以下に NLSR におけるこれらプロトコルの概要を述べる。

**Adjacency LSA** ルータは周期的に全ての隣接ルータに対し Info Interest と呼ばれる制御メッセージを送付する。Info Interest を受信したルータは、アクティブな隣接ルータ数と各隣接ルータの名称とその接続リンクのコストを Info Interest を送信したルータに返信する (Adjacency LSA)。もし隣接ルータから応答がない場合、Info Interest を送信したルータは再度 Info Interest を送信するが、一定回数、応答がない場合は該当隣接ルータからその接続リンクのいずれかに障害が発生したものと認識する。OSPF のデフォルトの設定では 10 秒ごとに Hello メッセージが送信され 4 回連続して応答がない場合に隣接ルータが接続リンクの障害を認識する [10]。しかし Hello メッセージ送信間隔を 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とならないことから [2]、隣接ルータや接続リンクの障害発生から 2 秒程度でルータは障害を検出できる [10]。

またルータは隣接ルータや接続リンクの障害発生を検出した場合や新たに NW にルータを接続した場合など、周囲の NW トポロジの変化を検出した際は Info Interest の応答情報と同じ情報を隣接ルータにブロードキャストする (Adjacency LSA)。この Adjacency LSA を受信した隣接ルータはやはり NW ト

(注1): 実際にはコンテンツを複数の部位に分割した chunk を単位にキャッシュされるが本稿ではコンテンツと表記する。

ポロジの変化を認識することになるためさらに隣接ルータに Adjacency LSA を行い、Adjacency LSA が反復されることで NW 上の全てのルータは更新された NW トポロジの情報を共有する。そして新たにダイクストラ法で各ルータに至る最小コスト経路を算出し FIB を更新する。そのため OSPF と同様、ICN においてルータの名称をロケータに用いた Info Interest による NLSR を用いることで、NW トポロジの変化に対して適応的な FIB の更新が可能となる。

**Prefix LSA** OSPF ではルータが転送可能な Prefix を隣接ルータに広告することでルータの FIB にエントリを作成する。ICN では Publisher が NW にコンテンツを登録した際に、Publisher を収容するルータが新たに提供されたコンテンツの名称を隣接ルータに広告 (Prefix LSA) する [21]。そして Prefix LSA を受けた隣接ルータはさらに Prefix LSA を反復することで、新たに提供されたコンテンツのエントリが全てのルータの FIB に作成される。その結果、NW の全てのルータから Publisher に Interest の転送が可能となる。

### 3. コンテンツ可用性の回復制御

本節では本稿で提案する ICN の障害時におけるコンテンツ可用性の回復方式 (COP: Centralized Original Promotion) の詳細を述べる。COP は単一のコピーを各コンテンツに対しオリジナルに昇格させることを想定する。そして障害発生を検出したルータが IR (Initiation Router) として可用性が喪失した全てのコンテンツに対するオリジナル昇格位置を決定する集中制御を想定する。大規模障害により NW が複数の部位に分断された場合、各々の NW 部位において 1 つのルータが IR として COP を行うことで、各々の NW 部位におけるコンテンツ可用性の回復が可能となる。COP の手順は図 1 に示すように、(1)IR による障害発生状態とキャッシュ状態の把握、(2)IR によるオリジナル昇格位置の決定、(3)オリジナル昇格処理、の 3 つの処理から構成される。以下に各処理の詳細を述べる。

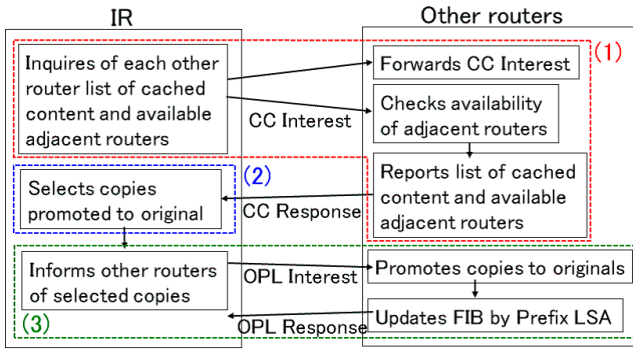


図 1 Procedure of COP

#### 3.1 IR による障害発生状態とキャッシュ状態の把握

Adjacency LSA により隣接ルータもしくは接続リンクにおける障害の発生を検知したルータは IR となり、全ての可用性喪失コンテンツに対するオリジナル昇格位置を集中的に選定する。そのため IR は最初に可用性喪失コンテンツと各正常ルータのキャッシュコンテンツに関する情報を収集する。NW 障害のため、IR からどのような経路を用いても到達できないルータの集合を  $R_f$  と定義し、IR からの接続性が維持されたルータの集合を  $R_a$  とする。全てのルータの集合を  $R$  とすると、 $R_f \cap R_a = \phi$ ,  $R_f \cup R_a = R$  となる。また可用性喪失コンテンツの集合  $M_f$  を、 $R_f$  にオリジナルが収容されたコンテンツの集合と定義する。 $M_f$  を把握するには  $R_f$  を知る必要があるが、IR 自身が検知できるのは自身の隣接ルータか、それらの接続リンクにおける障害発生のみである。そこで IR は全てのルータに対し到達性を確認するための Connectivity Confirmation (CC) Interest を送付することで、 $M_f$  と  $R_a$  の各ルータでキャッシュされているコンテンツのリストを収集する。本処理は、(i)IR の CC Interest 送信、(ii)CC Interest 受信ルータの CC Interest 転送、(iii)CC Interest 受信ルータの CC Response 送信、の 3 つの処理で構成される。

##### (i) IR の CC Interest 送信

隣接ルータへの到達性の喪失を検出したルータは IR となり、直ちに全ての各ルータに CC Interest を送信する。リンク・ルータ障害のため FIB を用いて各ルータを宛先とする CC Interest を転送すると目的ルータに到達しない可能性がある。そこで CC Interest を受信したルータは全ての出力 Face に Interest を転送する Exploration を用いて CC Interest を転送する [8]。

文献 [8] は FIB で設定された経路 (default path) 以外に存在するルータもコンテンツ転送元の候補として探索対象とし、コンテンツ転送の効率性を高めることを目的に Exploration を提案している。ところが経路ルータで Interest をフラディングすると到達性のある全てのルータに Interest が届く。ここでは IR が全てのルータに CC Interest を送付することを目的としていることから、特定のルータの名称ではなくルータが CC Interest であることを認識できるよう CC Interest の識別名称を目的アドレスとする。IR は現在時刻 (タイムスタンプ) を付与した CC Interest を全ての出力 Face に送信する。タイムスタンプは複数のルータが IR としてオリジナル昇格選択を行う問題を回避する目的で付与される。

##### (ii) ルータの CC Interest 転送

CC Interest を受信したルータ  $n$  は、PIT に該当 CC Interest のエントリが存在する場合は既にその CC Interest を転送済みなので棄却し、そうでない場合は初めての到着なので (a)PIT のエントリ作成と (b)CC Interest の転送の 2 つの処理を行う。(a)に関しては CC Interest 送信元ルータの名称と、その到着 Face 番号をルータ  $n$  の PIT に新しいエントリとして作成する。(b)に関しては CC Interest の到着 Face を除く全ての Face に CC Interest を転送する。このように同一 CC Interest の重複したフラディングを避けることで CC Interest のループ生成を回避する。また IR から到達可能な総リンク数を  $E_a$  とすると、CC Interest の総転送回数  $E_a$  で IR から  $R_a$  の全ルータに CC Interest が届く。

##### (iii) ルータの CC Response 送信

CC Interest を受信したルータ  $n$  は全ての隣接ルータに対して Adjacency LSA により接続性を確認する。NLSR の接続性確認に要する時間として OSPF の場合を考えると、2.3 節で述べたように最大で 2 秒程度で接続性喪失の検出が可能である。そしてルータ  $n$  は自身の CS にキャッシュしている全てのコンテンツの名称のリストと、接続性が確認された隣接ルータの名称のリストと、直前の TS の Interest 発生数  $Y_n$  を CC Response として IR に返信する。CC Response を受信したルータ  $n'$  は PIT に記載された Face に CC Response を転送する [13] [23]。ICN では Data chunk を転送した際に PIT の該当エントリを削除するが、CC Interest は NW の  $R_a$  の全てのルータに送付されるため、同一の PIT エントリが異なるルータから送信された複数の CC Response に対して利用される。そのため CC Response が最初にルータに到着した際に、直ちに PIT エントリを削除できない。そこで CC Response が初めてルータ  $n'$  に到着してから、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する。リンクの伝搬遅延時間は隣接ルータの稼働性確認に要する最大時間  $D_A = 2$  秒程度と比較して遙かに小さいことから、例えば  $T_P$  にはマージンを含めて 3 秒程度の時間を設定することが考えられる。

#### 3.2 IR によるオリジナル昇格位置の決定

IR は CC Interest を送信後、一定時間  $T_C$  以内に 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する。NW 全体の総リンク数  $E$ 、リンクの最大伝搬遅延時間  $\delta_L$ 、ルータの CC Interest 転送に要する最大時間  $\delta_I$ 、ルータの CC Response 転送に要する最大時間  $\delta_R$ 、ルータの稼働性確認に要する時間  $D_A$  に対し、 $T_C$  は  $E(\delta_L + \delta_I + \delta_R) + D_A$  以上を抑えられる。例えば総リンク数を 100、 $\delta_L$  を 5 ミリ秒 (最大リンク長 1,000km)、 $\delta_I$  と  $\delta_R$  を数 100 マイクロ秒 [20]、 $D_A$  を 2 秒とすると、 $T_C$  を 3 秒程度に設定すればよい。

正常時の Prefix LSA によって各ルータは全ての各コンテンツのオリジナルが提供されているルータを把握している。そのため IR は CC Response を  $R_a$  の全てのルータから受信することで、IR から見て可用性が喪失したコンテンツ  $M_f$  と、 $R_a$  の各ルータが CS に保持しているコンテンツを全て把握できる。 $C_m$  を  $R_a$  の中でコンテンツ  $m$  をキャッシュしているルータの集合と定義する。 $M_f$  の各コンテンツ  $m$  に対し、複数のルータが  $C_m$  に含まれる場合、その中からオリジナルに昇格させるルータを 1 つ選択する。

NW の稼働範囲において Interest はコンテンツのオリジナルを提供するオリジンサーバに向けて転送されるため、オリジナル昇格コピーの選択が可用性回復後の配信フロー長やリンク・ルータ負荷に影響を与える。リンク・ルータ負荷を低減するには配信フロー長が短いことが望ましい。そこで配信要求元とオリジンサーバとの間の最短ホップ経路上のホップ長 (最短ホップ長) の平均値が最小となるよう、 $M_f$  の各コンテンツ  $m$  に対しオリジナル昇格ルータ  $x_m$  を次式で選択する。

$$x_m = \arg \min_{x \in C_m} \sum_{n \in R_a} h_{n,x} r_n \quad (1)$$

ただし  $h_{n,x}$  はルータ  $n$  からルータ  $x$  への最短ホップ長で CC Response で通知された隣接ルータ情報から構築した NW 稼働範囲のトポロジからダイクストラ法を用いて算出する。また  $r_n$  は  $\mathbf{R}_a$  におけるルータ  $n$  の Interest 生成比率で、 $\mathbf{R}_a$  の各ルータから CC Response で通知された  $Y_n$  を用いて  $r_n = Y_n / \sum_{j \in \mathbf{R}_a} Y_j$  より算出する。複数のルータが  $x_m$  の候補となった場合は、その中の 1 つをランダムに選択する。

IR は  $\mathbf{M}_f$  の全てのコンテンツに対し  $x_m$  を計算する必要があるが、各コンテンツの  $x_m$  は他のコンテンツのオリジナル昇格位置とは独立しており、 $R_a \equiv |\mathbf{R}_a|$  と  $C_m \equiv |\mathbf{C}_m|$  に対し、 $O(R_a C_m)$  の計算量で  $x_m$  を得ることができる。

### 3.3 オリジナル昇格処理

IR は  $\mathbf{R}_a$  の全てのルータに対し、 $\mathbf{M}_f$  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 (OPL: original promotion list) を OPL Interest を用いて通知する。OPL Interest は CC Interest と同様、Exploration を用いて転送される。OPL Interest を受信した各ルータ  $n$  は OPL を検査し、 $x_m = n$  の全てのコンテンツ  $m$  の CS 内にあるコピーをオリジナルに昇格させる。すなわちオリジナルコンテンツ用のメモリが別途用意されている場合にはコピーをキャッシュメモリからオリジナル収容メモリに移動し、キャッシュアウト許容フラグを用いる場合には本フラグをオフにする。

そしてルータ  $n$  は隣接ルータに対し、自身がオリジナルに昇格させた  $x_m = n$  の全てのコンテンツに対する Prefix LSA を NLSR プロトコルにより実施する。その結果  $\mathbf{R}_a$  のルータの FIB にこれらコンテンツに対するエントリが作成され、これらコンテンツに対し  $\mathbf{R}_a$  のルータからの Interest 転送が可能となり、これらコンテンツの可用性が回復する。オリジナル昇格処理が完了したルータは、オリジナル昇格処理が成功したコンテンツの名称のリストを OPL Response として CC Response と同様の方法で IR に返信する。IR は各ルータから返信された OPL Response を確認し、何等かの理由でオリジナル昇格に失敗したコンテンツに対し、他のルータでコピーが存在する場合には再度、前節で述べた方法でオリジナル昇格位置を選択し OPL Interest を送信する。

### 3.4 複数の IR による重複したオリジナル昇格処理の回避

障害発生ルータ・リンクに隣接した正常ルータが複数存在する場合、複数のルータが IR としてコンテンツ可用性の回復処理を行う可能性がある。本稿では、ただ 1 つのルータが IR として全ての可用性喪失コンテンツ  $\mathbf{M}_f$  のオリジナル昇格位置を選択することを想定している。そのため他のルータが送信した CC Interest を受信したルータは、隣接ルータとの接続性の喪失を検出した場合にも IR にはならず CC Interest を送信しない。しかし最初に障害発生を検出したルータが送信した CC Interest が全ての正常ルータに到達するまでの間に、他のルータが障害発生を検出し CC Interest を送信した場合には、複数のルータが IR として CC Interest を送信する状態となる。そこで 3.1 節で述べたように IR はタイムスタンプを付加した CC Interest を送信する。そして 3.2 節で述べたように、IR は他の IR から CC Interest を受信した場合は、タイムスタンプを自身の CC Interest のタイムスタンプと比較することで、自身のタイムスタンプが最も古く最初に CC Interest を送信した場合のみ、オリジナル昇格ルータの選択と通知を行う。その結果、常に唯一つの正常ルータが  $\mathbf{M}_f$  の全てのコンテンツに対しオリジナル昇格位置を決定する。

## 4. 性能評価

### 4.1 評価条件

#### 4.1.1 NW トポロジ

CAIDA で公開されている米国の 4 つの商用 ISP のバックボーン NW のトポロジを評価に用いる [3]。各ノードにはルータが 1 つ設置されており、 $p_n$  をノード  $n$  の人口比とする。Interest 発生比率  $r_n$  は  $p_n$  と一致し、Interest の default path はオリジンサーバから要求ユーザへの最短ホップ経路とする。表 1 に各トポロジのノード数  $N$ 、リンク数  $E$ 、人口比で重みづけたノード間の平均ホップ距離  $H$  をまとめる。ノード集合  $\mathbf{N}$  に対し  $H$  を  $H = \sum_{i,j \in \mathbf{N}, i \neq j} p_i p_j h_{i,j}$  で定義する。 $H$  はオリジンサーバから配信した場合の平均ホップ長に相当する。4 つの NW の中で、 $H$  が小さくハブノードを経由することで少ないホップ長で他のノードに到達可能な Allegiance Telecom と Verio をハブ・スポーク (H&S) 型に分類する。一方、 $H$  が大きく他のノードに到達するには多数の中継ノードを経由する必要がある

At Home Network と CAIS Internet をラダー型に分類する。

#### 4.1.2 コンテンツ

全てのルータにおけるコンテンツの人気順位と配信要求比率が従う分布は同一とする。YouTube, VoD, Web のコンテンツごとの要求数分布は、両端を除き Zipf 分布で近似できることが知られていることから [5] [15] [22]、コンテンツ  $m$  の配信要求比率  $q_m$  をパラメタ  $\theta$  の Zipf 分布で与え、特に断らない限り  $\theta = 0.8$  に設定する [22]。コンテンツ数を  $M = 10,000$ 、全コンテンツは同一サイズとし、各ルータの CS は最大で  $B$  個のコンテンツを収容可能でありキャッシュ容量比率  $b$  を  $b = B/M$  で定義する。特に断らない限り  $B = 100$  ( $b = 0.01$ ) に設定する。シミュレーション開始時に各コンテンツのオリジンサーバ設置ルータを  $p_n$  に比例する確率でランダムに選択する。 $r_n = p_n$  に比例する確率でランダムに選択したルータから  $q_m$  に比例する確率でランダムに選択したコンテンツ  $m$  に対する配信要求を発生させる。ルータ  $u$  から生じたコンテンツ  $m$  への配信要求に対して、Interest が  $u$  から  $m$  のオリジンサーバ収容ルータに default path 上を転送され、 $m$  をキャッシュしている default path 上の  $u$  に最も近いルータ  $s$  からコンテンツ  $m$  が default path 上をユーザに配信される。評価結果はオリジンサーバ配置の影響を受けるため、ランダムに設定した 10 個のオリジンサーバ配置パターンで評価を行い、その平均値で評価する。

表 1 Summary of properties of four networks evaluated

| Network            | $N$ | $E$ | $H$  | Type         |
|--------------------|-----|-----|------|--------------|
| At Home Network    | 46  | 55  | 6.83 | Ladder       |
| CAIS Internet      | 37  | 44  | 5.49 | Ladder       |
| Allegiance Telecom | 53  | 88  | 2.81 | Hub & Spokes |
| Verio              | 35  | 72  | 2.23 | Hub & Spokes |

#### 4.1.3 キャッシュ方式

経由ルータでのコンテンツのキャッシュ判断を行うキャッシュ方式に関し、以下の 5 つの方式を想定する。

**AllCache:** 発ルータ  $s$  から目的ルータ  $u$  に至る default path 上の全てのルータでキャッシュ [6]

**EdgeCache:** default path 上の最終ホップノード (ルータ  $u$ ) でのみキャッシュ [19]

**UniCache:** default path 上の各ルータにて確率  $1/h_{s,u}$  でキャッシュ [6] (平均的には default path 上の 1 つのルータにてキャッシュ)

**ProbCache:** default path 上の各ルータ  $c$  にて確率  $h_{s,c}/h_{s,u}$  でキャッシュ [18] (ルータ  $s$  からの距離に比例し、ルータ  $u$  に近いほど高い確率でキャッシュ)

**LCD (leave copy down):** ルータ  $s$  から次ホップのルータでのみキャッシュ [16] (オリジンサーバの近隣にコピーが存在する傾向が高く、段階的に NW 上をコピーが拡散)

全てのキャッシュ方式においてキャッシュ置換法としては LRU を用い、オリジンサーバ収容ルータにはコピーをキャッシュしない、またオリジナル昇格コピーを永続的に保持するためのメモリを CS とは別に用意するものとし、障害発生後も正常ルータの CS 容量は不変とする。各ルータのオリジン昇格コンテンツ数に制約は設けない。

#### 4.1.4 NW 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が生じたのみとし、IR からの非到達ルータ  $\mathbf{R}_f$  とする。ただし正常ルータが 2 つ以上の分断された NW 領域に分断された場合には、含まれる正常ルータ数が最大の NW 領域のみを評価対象とし、 $\mathbf{R}_a$  の構成ルータとする。それ以外の分断された NW 領域に属する正常ルータは  $\mathbf{R}_f$  に含める。 $\mathbf{R}_a$  の中での IR の位置は明示的に考慮せず、3. 節で述べた COP のコンテンツ可用性回復処理のうち、3.2 節で述べた IR によるオリジナル昇格位置の決定処理のみを計算機シミュレーションで実施する。与えられた障害半径  $D$  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、その平均値で評価する。

図 2 に 4 つの各 NW に対し、障害半径  $D$  を変化させたときの非到達ルータ  $\mathbf{R}_f$  の全ルータに対する比率  $F_r$  と、これら非到達ルータにオリジナルが存在し可用性が喪失したコンテンツ  $\mathbf{M}_f$  の全コンテンツに対する比率  $F_m$  を各タブロットする。 $D$  の増加に伴い非到達ルータ数と可用性喪失コンテンツ数は線形に増加するが、例えば  $D = 500$  km の場合、全体の 10%~15% 程度のルータの到達性が喪失し、全体の 15% 程度のコンテンツの可用性が喪失する。以後、特に断らない限り  $D = 500$  km に設定する。



#### 4.2 コンテンツ可用性の回復効果

本節では提案方式 (COP) のコンテンツ可用性の回復効果を評価する。全ての評価において At Home Network と Verio は CAIS Internet と同様の結果が得られたことから、At Home Network と Verio の結果は省略する。図 3 に各キャッシュ方式を用いた場合の COP のコンテンツ可用性回復率  $\Phi_c$  を、各ルータのキャッシュ容量比率  $b$ 、コンテンツ人気分布の Zipf パラメタ  $\theta$ 、NW 障害半径  $D$  に対して各々プロットする。ただし  $\Phi_c$  を NW 障害によって可用性が喪失したコンテンツの中で、1 つ以上の正常ルータ  $R_a$  にコピーが存在し、COP によって可用性の回復が可能なコンテンツの割合で定義する。 $b$  の増加に伴い各コンテンツのキャッシュコピー数が増加し  $\Phi_c$  は向上する。 $\theta$  の増加に伴いコンテンツの人気の偏りが強まるため高人気コンテンツのコピー数が増加し低人気コンテンツのコピー数が減少する結果、 $\Phi_c$  は低下する。 $D$  の増加に伴い正常ルータ  $R_a$  が減少するため、 $R_a$  のルータに 1 つ以上のコピーが存在する可能性が低下し  $\Phi_c$  は減少する。また最大オリジン・コピー間距離が大きいキャッシュ方式ほど  $\Phi_c$  が低い。すなわち LCD はオリジン周辺にコンテンツがキャッシュされるため、オリジンが喪失したコンテンツは同時にキャッシュコピーも喪失している可能性が高く  $\Phi_c$  が低い。EdgeCache はオリジンとは離れた位置のルータにコンテンツをキャッシュするため、5 つのキャッシュ方式の中では最も  $\Phi_c$  が低い。AllCache は、Allegiance Telecom においては EdgeCache に次いで  $\Phi_c$  が低い。CAIS Internet においては LCD に次いで  $\Phi_c$  が低い。例えば  $b = 0.01$ ,  $\theta = 0.8$ ,  $D = 500\text{km}$  の場合、COP 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が回復する。

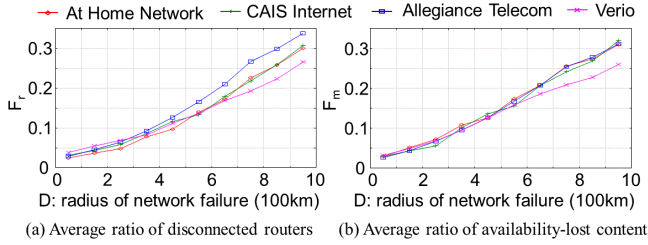


図 2 Ratio of routers and content items with losing availability

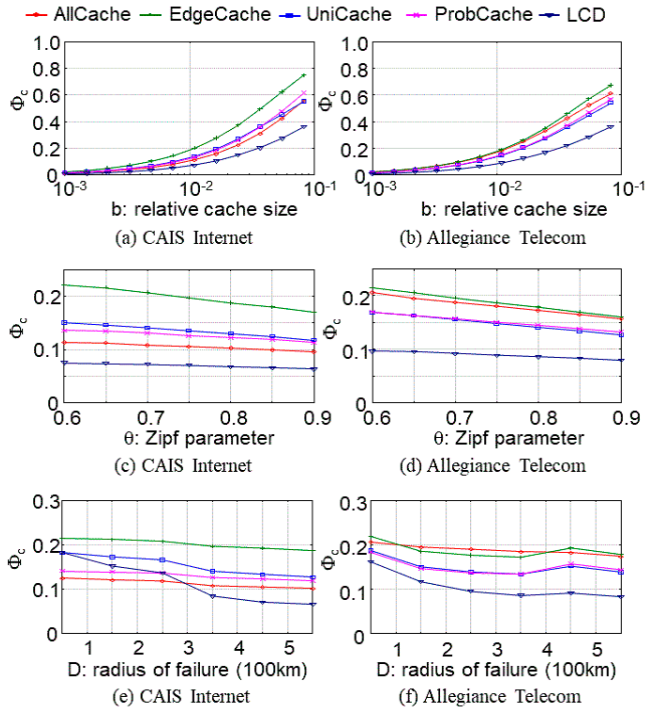


図 3 Ratio of availability-recovered content among availability-lost content at network failure

コンテンツを配信要求比率  $q_m$  に応じて 5 つの人気度グループ (PG: popularity group) に分類する。 $q_m$  が最も高い 20 個のコンテンツを PG1, 21 番から 100 番目までを PG 2, 101 番から 400 番目までを PG 3, 401 番から 2,000 番目までを PG

4, 2,001 番から 10,000 番目を PG 5 に分類する。図 4 に、各人気グループ  $g$  の可用性喪失コンテンツの中で COP により可用性が回復したコンテンツの割合  $\phi_{c,g}$  をプロットする。人気の高いコンテンツほど多数のルータにコピーが存在するため  $\phi_{c,g}$  が高い。

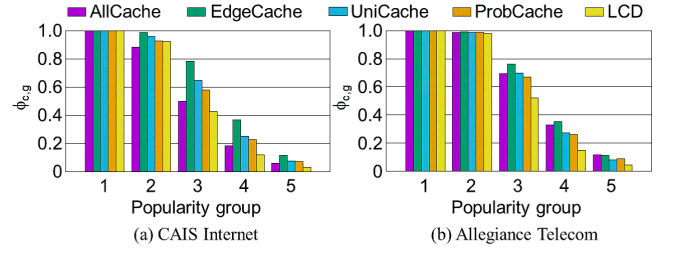


図 4 Ratio of availability-recovered content among availability-lost content at network failure in each popularity group

#### 4.3 配信要求の受理確率の向上効果

本節では COP のコンテンツ可用性回復で得られる、配信要求の受理確率の向上効果を評価する。NW 障害発生後、COP を用いて可用性喪失コンテンツの可用性を回復させた後、100,000 回の配信要求を 4.1 節で述べた方法で発生させたときに、可用性喪失コンテンツに対する配信要求の中で、COP によりコンテンツ可用性が回復しコンテンツ取得が可能となったものの割合を  $\Phi_r$  とする。図 5 に、CAIS Internet と Allegiance Telecom において  $b$  を変化させたときの  $\Phi_r$  を示す。 $\Phi_r$  の  $b$  に対する傾向やキャッシュ方式間の優劣に関しては前節で述べた  $\Phi_c$  と同様の傾向が確認できた。ただし人気の高いコンテンツほど COP による可用性回復確率が向上することから、 $\Phi_r$  は  $\Phi_c$  より大きく、例えば  $b = 0.01$ ,  $\theta = 0.8$ ,  $D = 500\text{km}$  の場合、COP により可用性喪失コンテンツの 30%~60% 程度の配信要求に対するコンテンツ配信が可能となる。

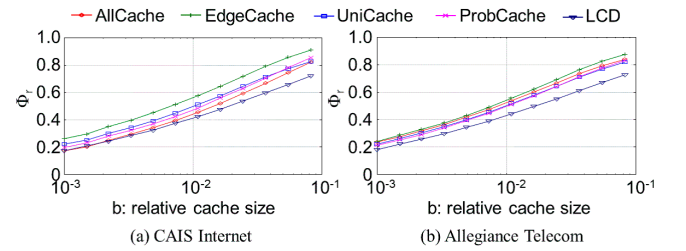


図 5 Ratio of accepted requests for availability-lost content at network failure

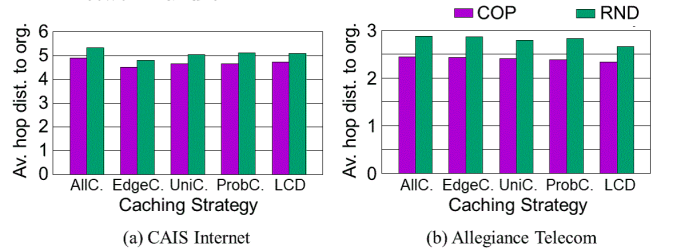


図 6 Average hop distance from users to originals of availability-recovered content

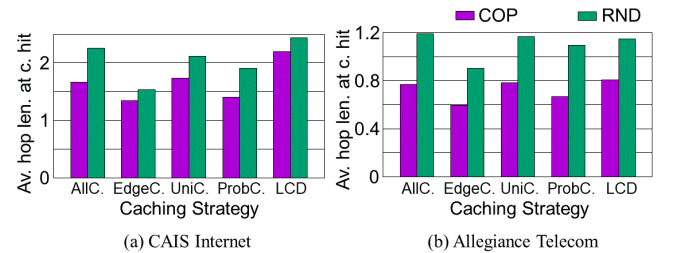


図 7 Average hop distance from users to closest cached copies on default path of availability-recovered content when cache hit

#### 4.4 オリジナル昇格位置選択法の比較

3.2 節で述べた COP のオリジナル昇格位置の選択法の有効性を確認するため、ランダムにオリジナル昇格位置を選択した

場合 (RND) と、コンテンツ可用性回復後の平均フロー長などの各種特性を比較する。図 6 に、CAIS Internet と Allegiance Telecom におけるコンテンツ可用性の回復に成功したコンテンツの配信要求元とオリジナルとの間の平均ホップ長を示す。ただし RND では各可用性喪失コンテンツ  $m$  に対し、 $R_a$  の  $m$  をキャッシュに保持するルータの中からランダムに 1 つ選択したルータを  $m$  のオリジナル昇格位置として選択する。RND と比較して COP を用いることで、ユーザとオリジナルとの平均ホップ距離を、CAIS Internet においては 6%~9% 程度、Allegiance Telecom においては 12%~16% 程度、低減が可能である。他の 2 つの NW においても同様の結果が得られたが、高次数のハブノードにキャッシュされたコピーをオリジナルに昇格させることで効果的なオリジナル距離の低減が可能な H&S 型のトポロジの方が COP の効果が高い。本節の全ての結果において、At Home Network は CAIS Internet と、Verio は Allegiance Telecom と同様の結果が得られた。いずれの NW においてもキャッシュ方式による差異は小さい。

オリジナル収容ルータに至る Interest 転送経路上の任意のルータで要求コンテンツがキャッシュされている場合を、キャッシュがヒットしたと見なす。図 7 に COP と RND におけるコンテンツ可用性回復コンテンツに対するキャッシュヒット時の配信フローの平均ホップ長をプロットする。CAIS Internet においては 10%~26% 程度、Allegiance Telecom においては 30%~39% 程度、低減が可能である。COP はオリジナル距離に加えキャッシュヒット時のフロー長も低減させるが、低減効果はキャッシュヒット時の方が大きい。図 8 に COP と RND におけるコンテンツ可用性回復コンテンツに対するキャッシュヒット率を示す。COP は RND と比較してユーザ・オリジナル間の距離が低減するため、各コンテンツ配信における default path 長が低減しキャッシュ候補となるルータが減少する結果、キャッシュヒット率が低下する。

最後に図 9 に、コンテンツ可用性回復コンテンツに対する全ての配信の平均フロー長を示す。COP は RND と比較してキャッシュヒット率は低下するものの、キャッシュヒット時、およびキャッシュミス時の配信フロー長の低減効果が大きく、COP の平均フロー長は RND の場合より CAIS Internet においては 3%~6% 程度、Allegiance Telecom においては 6%~8% 程度、低減する。

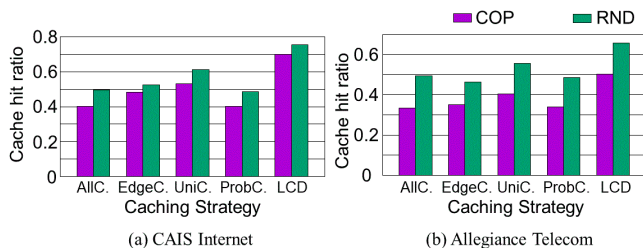


図 8 Cache hit ratio of availability-recovered content

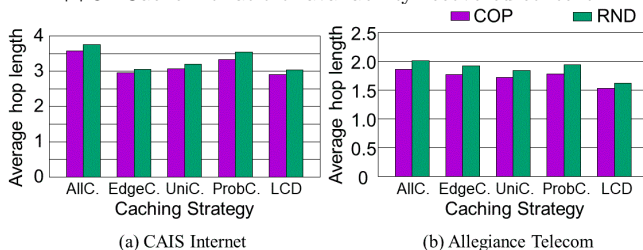


図 9 Average hop length from source routers to users of availability-recovered content

## 5. まとめ

ICN が社会基盤として普及するには、震災、サイバー攻撃などにより生じる大規模 NW 障害の際にも、NW の正常範囲においてコンテンツを継続的に取得できるコンテンツ可用性の維持が求められる。しかし ICN ではコンテンツのオリジナルに向け Interest が転送されるようルータの FIB が設定されるため、オリジナルへの到達性が喪失したコンテンツの可用性が喪失する。そこで本稿では、正常ルータにキャッシュされたコンテンツの中から 1 つを、障害ルータに隣接するルータの 1 つが集中的に選択し、オリジナルに昇格させることでコンテンツ可用性を回復する COP を提案した。実際の ISP の NW トポロジを用いた計算機シミュレーション評価により COP の有効性を確認した。例えば各ルータにコンテンツ数の 1% の容量

のキャッシュが実装されている場合に、15% 程度のコンテンツの可用性が NW 障害により喪失した際、COP 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が回復し、可用性喪失コンテンツの 30%~60% 程度の配信要求が救済される。またランダムに可用性回復位置を選択する場合と比較して、可用性回復コンテンツの平均配信フロー長を数% 程度、低減できることを明らかにした。オリジナル昇格コピーが特定のルータに集中した場合、そのルータや周辺の NW に輻輳が生じることが予想される。そこで今後は、負荷分散の観点も考慮したオリジナル昇格位置の決定法を検討する。また複数のルータが自律分散的にオリジナル昇格処理を行う場合についても検討する予定である。ところで NW 障害時にコンテンツの可用性を回復するには、正常ルータにコンテンツのコピーが存在している必要がある。そのため障害時のコンテンツ可用性の回復率の向上を目的に、オリジナルと離れた位置にコピーを優先的にキャッシュするキャッシュ制御法を検討する予定である。

謝辞 本研究成果は、JSPS 科研費 18K11283 の助成を受けたものである。ここに記して謝意を表す。

## 文 献

- [1] M. F. Al-Naday, M. J. Reed, D. Trossen, and K. Yang, "Information Resilience: Source Recovery in an Information-Centric Network," IEEE Network 2014.
- [2] A. Basu and J. Riecke, "Stability issues in OSPF routing," ACM SIGCOMM 2001.
- [3] CAIDA web page, <http://www.caida.org/data>
- [4] Cisco Visual Networking Index: Forecast and Methodology, 2016-2021.
- [5] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon, "Analyzing the Video Popularity Characteristics of Large-Scale User Generated Content Systems," IEEE/ACM Trans. Networking, Vol. 17, No. 5, pp. 1357-1370, Oct. 2009.
- [6] K. Cho, et al., "WAVE: Popularity-based and Collaborative In-network Caching for Content-Oriented Networks," IEEE NOMEN 2012.
- [7] J. Choi, J. Han, E. Cho, T. Kwon, and Y. Choi, "A Survey on Content-Oriented Networking for Efficient Content Delivery," IEEE Commun. Mag., vol. 49, no. 3, pp. 121-127, Mar. 2011.
- [8] R. Chiocchetti, D. Rossi, and G. Carofiglio, "Exploit the Known or Explore the Unknown? Hamlet-Like Doubts in ICN," ACM ICN 2012.
- [9] R. Chiocchetti, D. Perino, G. Carofiglio, D. Rossi, and G. Rossini, "INFORM: a dynamic Interest Forwarding Mechanism for Information Centric Networking," ACM ICN 2013.
- [10] 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 and Tutorials, Vol. 14, No. 2, pp. 443-463, 2012.
- [11] M. Gritter and D. R. Cheriton, "An architecture for content routing support in the Internet," USENIX USITS 2001.
- [12] A. Hoque, S. Amin, A. Alyyan, B. Zhang, L. Zhang, and L. Wang, "NLSR: Named-data Link State Routing Protocol," ACM ICN 2013.
- [13] V. Jacobson, et al., "Networking Named Content," ACM CoNEXT 2009.
- [14] T. Koponen, M. Chawla, B. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, "A data-oriented (and beyond) network architecture," ACM SIGCOMM 2007.
- [15] S. A. Krashakov, A. B. Teslyuk, and L. N. Shchur, "On the universality of rank distributions of website popularity," Elsevier Computer Networks, Vol. 50, No. 11, pp. 1769-1780, Aug. 2006.
- [16] N. Laoutaris, H. Che, and I. Stavrakakis, "The LCD interconnection of LRU caches and its analysis," Elsevier Performance Evaluation, Vol. 63, Issue 7, pp. 609-634, July 2006.
- [17] E. Nygren, R. Sitaraman, and J. Sun, "The Akamai Network: A Platform for High-Performance Internet Applications," ACM SIGOPS 2010.
- [18] I. Psaras, W. K. Chai, and G. Pavlou, "Probabilistic In-Network Caching for Information-Centric Networks," ACM ICN 2012.
- [19] V. Sourlas, L. Tassioulas, I. Psaras, and G. Pavlou, "Information Resilience through User-Assisted Caching in Disruptive Content-Centric Networks," IFIP Networking 2015.
- [20] Y. Wang, Y. Zu, T. Zhang, K. Peng, Q. Dong, B. Liu, W. Meng, H. Dai, X. Tian, Z. Xu, H. Wu, and D. Yang, "Wire Speed Name Lookup: A GPU-based Approach," USENIX NSDI 2013.
- [21] G. Xylomenos, et al., "A Survey of Information-Centric Networking Research," IEEE Communications Survey and Tutorials, Vol. 16, No. 2, pp. 1024-1049, 2014.
- [22] H. Yu, D. Zheng, B. Y. Zhao, and W. Zheng, "Understanding User Behavior in Large-Scale Video-on-Demand Systems," ACM EuroSys 2006.
- [23] L. Zhang, et al., "Named Data Networking (NDN) Project," Technical Report NDN-0001, Oct. 2010.