

## [招待講演] コンテンツの多様性に基づくキャッシュ制御

上山 憲昭<sup>†</sup>

<sup>†</sup> 日本電信電話株式会社 NTT ネットワーク基盤技術研究所 〒180-8585 東京都武蔵野市緑 3-9-11  
E-mail: †kamiyama.noriaki@lab.ntt.co.jp

あらまし インターネットの 8 割以上のトラフィックはコンテンツ配信であり、コンテンツをユーザの近くから配信することでユーザ品質を向上させネットワーク内のトラフィック量を抑制する CDN (content delivery networks) が広く用いられている。キャッシュサーバのストレージ容量は有限であるため、キャッシュ内に保存されるコンテンツを決めるキャッシュ置換方式が CDN の性能を大きく左右する。本稿では、まず従来のキャッシュ置換方式の代表的なアプローチについて述べ、全てのコンテンツに同じ置換ポリシーを適用する従来のアプローチの中では、LRU (least recently used) が最良の性能を達成することを述べる。そしてキャッシュ制御の新たな視点として、エンドユーザの体感品質 (QoE) に与える効果や、QoE 改善の価値の観点からキャッシュ制御を考えることの重要性を述べる。キャッシュサーバから配信することで得られる CDN の効果は、コンテンツの人気度、Web 応答時間等のサービス全体の性能における効果、コンテンツの種類など、コンテンツの様々な属性に依存する。そこでコンテンツの多様性に基づきキャッシュ制御を行う近年のキャッシュ制御の研究動向を俯瞰し、いくつかの研究事例を紹介する。

キーワード コンテンツ配信, キャッシュ制御, QoE 向上の価値

## [Invited Talk] Cache Control Based on Diversity of Content

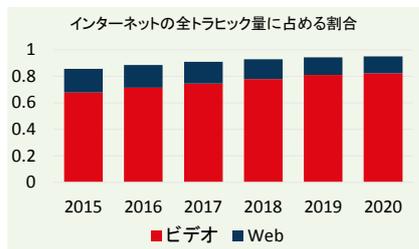
Noriaki KAMIYAMA<sup>†</sup>

<sup>†</sup> NTT Network Technology Laboratories, NTT Corporation 3-9-11, Midori, Musashino, Tokyo 180-8585  
E-mail: †kamiyama.noriaki@lab.ntt.co.jp

**Abstract** More than eighty percent traffic in the Internet is related to content delivery, and CDN (content delivery networks) which improves user quality and reduces the amount of traffic in the networks by delivering content from cache servers located close to users has been widely used. The storage capacity of cache servers is finite, so the cache replacement policy deciding cached content strongly affects the performance of CDN. This paper firstly outlines the main approaches of existing cache-replacement methods which treat all content items with the identical policy and denotes that the LRU (least recently used) achieves the most desirable performance among them. Then, this paper points out the importance of designing cache replacement policy from the viewpoints of the impact on the user-perceived quality, i.e., QoE, and the value of improving the QoE. The effect of CDN obtained by delivering content from cache servers depends on various factors of content including popularity, impact on the overall QoE such as the web response time, and content type. This paper outlines the recent research trend of cache replacement based on diversity of content and introduces some examples of recent studies.

**Key words** content delivery, cache control, value of improving QoE

## コンテンツ配信のトラフィック占有率



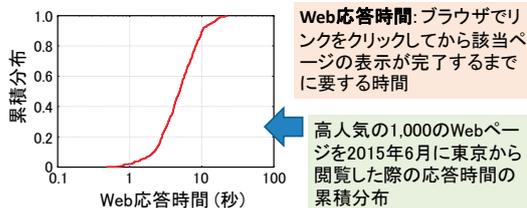
データ引用: [Cisco15]

インターネットの主要トラフィックはコンテンツ配信



Copyright©2016 NTT corp. All Rights Reserved. 1

## Web応答時間の増大



- 応答時間: 50%のページは5秒以上, 10%のページは10秒以上
- 0.1秒の待ち時間の低減によりAmazonの利益は1%向上[Kohavi07]



Copyright©2016 NTT corp. All Rights Reserved. 2

## トラフィック量の増加



データ引用: [Cisco15]

- コンテンツのリッチ化、動画配信数の増加に伴い、ネットワーク上を流れるトラフィック量が今後も増加
- ネットワーク事業者の設備投資コストが増加



Copyright©2016 NTT corp. All Rights Reserved. 3

## CDNの利用の拡大

- CDN (content delivery network): ユーザ品質の向上、ネットワーク内トラフィック量の削減を目的とした、コンテンツ配信プラットフォーム
- ユーザの近くに設置されたキャッシュサーバからコンテンツを配信



データ引用: [Cisco15]

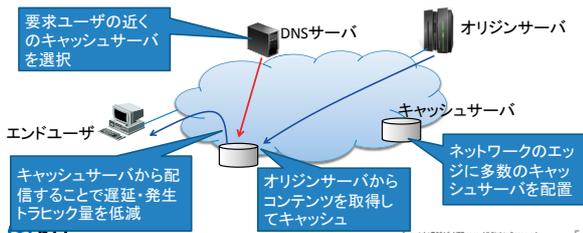
CDNはインターネット上でコンテンツを効率的に配信する技術として広く普及



Copyright©2016 NTT corp. All Rights Reserved. 4

## CDNの概要

- ステークホルダー:
  - CDN事業者: キャッシュサーバとDNSサーバを設定/運用
  - コンテンツ事業者: オリジンサーバを設置しコンテンツを提供
  - エンドユーザ: コンテンツを消費



Copyright©2016 NTT corp. All Rights Reserved. 5

## 主なCDNタイプ

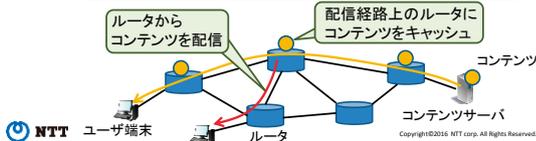
CDNタイプ	インフラ所有者	運用者	顧客	例	特徴
Traditional CDN	CDN事業者	CDN事業者	CP	Akamai	CDN事業者が多数のNWにキャッシュサーバ(CS)を設置
Managed CDN	CDN事業者	CDN事業者	ISP	Akamai Aura MCDN	CDN事業者が特定のISP向けにCDNサービスを提供し、ISPに品質情報を提供
Licensed CDN	CDN事業者	ISP	CP	Akamai Aura LCDN	ISPはCDNのライセンスを受けて自身でCDNを運用
Telco CDN	ISP	ISP	CP	Level 3	ISPが独自にCDNを構築して運用
Virtualized CDN	クラウド事業者	CP	CP	Amazon CloudFront, CloudFlare	商用サービスとして稼働、従量課金でCDNサービスを提供
	ISP	CDN, CP	CP	CDNaaS [Dudouet14], CDN-as-a-VNF [Herbaut16]	研究段階: ISPがCDN事業者やCPに対してCDNを提供

## 情報指向ネットワーク(CCN)

- 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処理が発生
- NW事業者とは独立した事業者が配信システムを提供  
⇒ コンテンツ配置や配信サーバ選択の最適化が困難

CCN (Content-Centric Networking) [Jacobson09]:

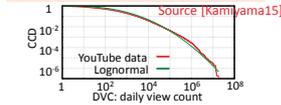
- ルータがコンテンツをキャッシュ
- コンテンツ名称で要求パケットを転送し、発見的にキャッシュルータから配信



## コンテンツの人気の時間・空間的な特長

- 偏り: 少数の人気コンテンツに需要の多くが集中
- 時間的局所性: 時間に対する人気の様々な変化パターン
- 空間的局所性: ローカルな人気度 ≠ グローバルな人気度

YouTubeビデオの日視聴数の累積補分布



YouTubeビデオの最頻度地の平均視聴集中度

Region	Videos	Med. V <sub>i</sub>	Avg. F <sub>i</sub>
USA	37.7%	0.513	0.75
Brazil	6.6%	0.497	0.90
UK	4.9%	0.460	0.69
Germany	4.1%	0.937	0.73
Japan	3.2%	1.003	0.85
Spain	2.9%	0.534	0.75
France	2.5%	0.630	0.73
Mexico	2.5%	0.647	0.70
Canada	2.5%	0.323	0.72
Italy	2.4%	0.859	0.81

Source [Brodersen12]

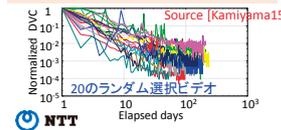
V<sub>i</sub>: 動画iの総視聴数  
V<sub>ik</sub>: 動画iの地域kの視聴数

視聴集中度 F<sub>i</sub>:  
V<sub>ik</sub> の最大値の V<sub>i</sub> に対する比率

$$F_i = \frac{1}{V_i} \max_k V_{ik}$$

日本が最頻度地のビデオの視聴の85%は国内需要

YouTubeビデオの正規化日視聴数の時系列



## CDN上のコンテンツ展開

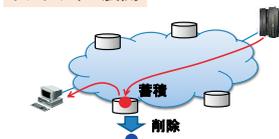
- プロアクティブ展開: 事前にコンテンツをキャッシュサーバ上に配置
- リアクティブ展開: 配信要求に応じて動的にコンテンツを蓄積・削除

プロアクティブ展開



長所: 全体最適化が可能  
短所: 正確な需要予測が必要, 組合せ最適化問題を解くための計算負荷が大

リアクティブ展開

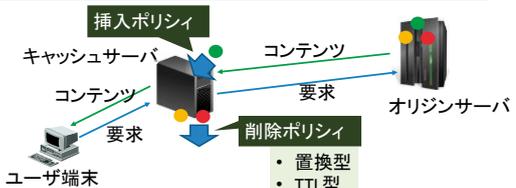


長所: 簡易, コンテンツ需要の時間・空間的な局所性を反映  
短所: 全体最適化が困難

多くのCDNは、簡易で適応的なリアクティブ展開を採用

## リアクティブ展開におけるキャッシュ制御

- 挿入ポリシー: オリジンサーバからコンテンツを取得した際に、コンテンツをキャッシュに保存するか否かを決定
- 削除ポリシー: キャッシュから削除するコンテンツを選択



多くのキャッシュ制御は置換型の削除ポリシー

## キャッシュ置換の主なアプローチ(1)

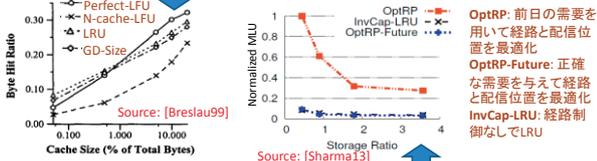
- Recency型ポリシー: 人気の時間変化が激しい場合に有効
  - 最近: 要求があったコンテンツを優先的に保持
  - LRU (least recently used), LRUのアレンジ版
  - 長所: 人気の時間変動に対応, 簡易
  - 短所: 要求頻度を非考慮 ⇒ 長期間需要のあるコンテンツを削除
- Frequency型ポリシー: 人気が時間安定な場合に有効
  - 要求頻度が高いコンテンツを優先的に保持
  - LFU (least frequently used), LFUのアレンジ版
  - 長所: 人気が安定している場合に効果的
  - 短所: 複雑, 過去に人気のコンテンツが残留 (cache pollution)

## キャッシュ置換の主なアプローチ(2)

- Function型ポリシー:
  - 各コンテンツに対しキャッシュ有効性を表す関数 H<sub>i</sub> を定義
  - H<sub>i</sub> が最小のコンテンツを削除
  - Recency型やFrequency型を包括する汎用性
  - GD (greedy dual)-Size [Cao97]:
 
$$H_i = \frac{c_i}{s_i} + L$$
 c<sub>i</sub>: オリジンサーバからコンテンツ i を取得するコスト  
s<sub>i</sub>: コンテンツ i のサイズ  
L: 最後に削除されたコンテンツの H<sub>i</sub> の値
  - 長所: 人気度以外の様々な観点を考慮
  - 短所: 遅延に基づく尺度は測定が困難, 複数尺度を考慮する場合は重みの設定が困難

## 最適なキャッシュ置換法

- DECのWebプロキシトレースデータを用いて、LRU、LFU、GD-Sizeのバイトキャッシュヒット率を比較 [Breslau99] ⇒ LRU: best, LFU: worst



- Akamai CDNのNewsビデオを取得したトレースデータを用いて、配信経路とコンテンツ配信位置を最適化する効果を分析 [Sharma13]
  - 前日のデータを元に経路と配信位置を同時に最適化(OptRP)した場合の最大リンク使用率は、単にLRUを用いた場合よりも大

人気の時間変化が激しい場合、LRUが最良

## キャッシュ置換法の新たな視点

従来の評価の観点 ⇒ 多くの場合、LRUで十分

- (バイト)キャッシュヒット率
- 最大(平均)リンク負荷

1. キャッシュの効果は、これら観点だけでは計れない

- エンドユーザの体感品質(QoE)に与える効果
    - Webページの表示待ち時間
    - 動画ストリーミング配信品質
  - ユーザQoE改善の「価値向上」に対する貢献
    - 例えばWeb応答時間が1秒改善することの価値は、Webページによって異なる。
2. プロアクティブ展開(レプリケーション)との併用



Copyright©2016 NTT corp. All Rights Reserved. 14

## キャッシュ制御の方向性

- 従来のキャッシュ制御: 全てのコンテンツを同等に扱う。
- しかし需要の時間的・空間的パターンやキャッシュ配信の効果は、コンテンツごとに大きく異なる。

### 1. ユーザQoE改善効果の向上

- Web応答時間に与える影響の高いコンテンツを優先
- 配置の空間的局所性が高いコンテンツを優先

### 2. キャッシュ配信の「価値」の向上

- キャッシュ配信で得られる効用の高いコンテンツを優先

### 3. プロアクティブ展開との併用

- 需要量が時間的に安定しているコンテンツをプロアクティブ展開
- コンテンツ事業者別に最適配置(仮想化CDN)

コンテンツの差別化(NW中立性とも関係)

## 方向性1: ユーザQoE改善効果の向上

- キャッシュ配信のユーザQoE改善に対する効果が高いコンテンツを優先

### Web表示待ち時間に与える影響に基づく差別化

- HTML, CSS (cascading style sheets), JavaScriptなど、表示待ち時間に与える影響の大きなオブジェクトを優先 (SPDY, HTTP/2に実装)
- オブジェクトの依存関係グラフから優先度を決定 [Narayanan16]

### コンテンツの空間的配置状態に基づく差別化

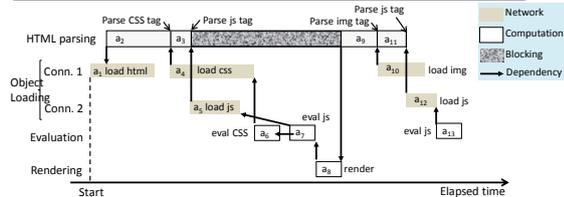
- Webカテゴリ間の優先制御 [Kamiyama16]: Reference (Stack overflow, Yahoo Answersなど)やNews (CNN, Yahoo Newsなど)等のコンテンツのグローバル性が高いカテゴリを優先



Copyright©2016 NTT corp. All Rights Reserved. 16

## Webブラウザにおける処理のクリティカルパス

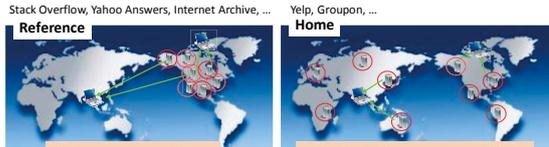
プロセス間の依存性のため、特定のオブジェクトの取得や演算が全体の表示待ち時間を決めるクリティカルパスが存在



- Starts HTML parsing ( $a_2$ )
- Continues HTML parsing while downloading CSS object ( $a_4$ )
- Suspends HTML parsing while downloading JS object ( $a_2$ )
- Downloads CSS and JS objects in parallel
- Starts eval of CSS obj after obtain ( $a_4$ )
- Starts eval of JS after eval of CSS ( $a_5$ )
- Starts render after obtaining CSS and JS objects ( $a_4$ )
- Resumes HTML parsing when completing rendering ( $a_4$ )

## CDNの応答時間削減効果

スポーツ、ニュース等のWebページのカテゴリごとに、オブジェクトの地理的な配置傾向が異なり、CDNの効果は異なる可能性あり



### アプローチ [Kamiyama16]

- Webカテゴリやアクセス位置別に、CDNの効果を推定
- 世界の12地点から1,000のWebページにアクセスした際の実測データから、カテゴリ別のエッジ配信の応答時間削減効果を試算

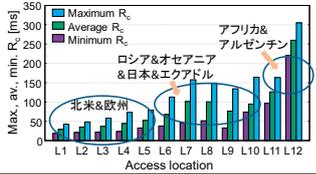


Copyright©2016 NTT corp. All Rights Reserved. 18

## 各カテゴリの平均RTT

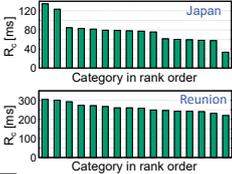
$R_c$ : カテゴリ c の平均RTT

各測定地点の平均RTTの最大/平均/最小値



L1	Massachusetts	L4	Ireland	L7	Australia	L10	Ecuador
L2	Wisconsin	L5	Germany	L8	New Zealand	L11	Argentina
L3	California	L6	Russia	L9	Japan	L12	Reunion

各カテゴリの平均RTT (降順)



- ・ アフリカとアルゼンチン以外の地域: カテゴリ間の差異が大
- ・ アフリカとアルゼンチン: カテゴリ間の差異は小さく、どれも遠い

NTT

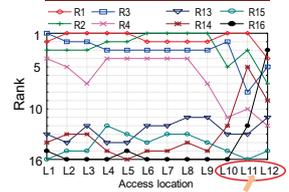
Copyright©2016 NTT corp. All Rights Reserved. 19

## 平均RTTのカテゴリ順位

全測定地点(AI)と3測定地点のカテゴリ順位

Rank	All	California	Japan	Reunion
R1	Reference	Reference	Adult	Regional
R2	Adult	News	Reference	Business
R3	News	Adult	News	Shopping
R4	Games	Society	Science	Reference
R5	Computers	Business	Computers	News
R6	Science	Science	Society	Arts
R7	Society	Games	Games	Adult
R8	Regional	Kids&teens	Kids&teens	Computers
R9	Arts	Computers	Arts	Recreation
R10	Business	Regional	Health	Society
R11	Kids&teens	Health	Sports	Sports
R12	Health	Sports	Business	Games
R13	Sports	Recreation	Regional	Kids&teens
R14	Recreation	Arts	Home	Science
R15	Home	Home	Recreation	Home
R16	Shopping	Shopping	Shopping	Health

上下4カテゴリの各測定地点の順位



- ・ 南米とアフリカ以外の地域は平均RTTのカテゴリ順位が世界で同様
  - ・ Reference (Stack overflow, Yahoo Answers等), News (CNN, Yahoo News等) など、コンテンツが普遍的なサイトのオブジェクトは北米に集中 ⇒ 平均RTTが大
  - ・ Shopping (Amazon, Ebay等), Home (Yelp, Groupon等) など、コンテンツの地域性が高いサイトのオブジェクトは各地域で固有 ⇒ 平均RTTが小
- ・ 南米とアフリカは全カテゴリのオブジェクトが遠方にあり、平均RTTのカテゴリ順位が固有

NTT

Copyright©2016 NTT corp. All Rights Reserved. 20

## Web応答時間削減量上限値の推定式の導出

仮定

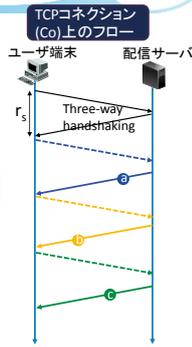
- ・ 全サーバの全TCP Co上で同時にObj取得を開始
- ・ Objを確立TCP Co間で均等に取得
- ・ 各TCP Co上で連続して間欠なくObjを取得
- ・ エッジノードとユーザ端末間のRTTはゼロ
- ・ DNS解決/サーバ応答/Obj転送の時間は不変

エッジ配信による応答時間削減量の推定

$$\max_{s \in S_x} \left[ \left\lceil \frac{m_s}{P} \right\rceil + 1 \right] r_s$$

- ・  $S_x$ : Webページ x の配信サーバの集合
- ・  $m_s$ : 配信サーバ s からの取得オブジェクト数
- ・  $r_s$ : ユーザ端末と配信サーバ s 間の平均RTT

実測値を適用



NTT

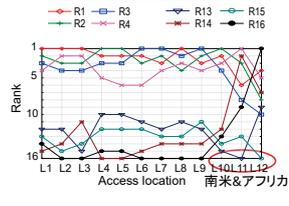
Copyright©2016 NTT corp. All Rights Reserved. 21

## 平均応答時間削減量のカテゴリランキング

全測定地点(AI)と3測定地点のカテゴリ順位

Rank	All	California	Japan	Reunion
R1	News	News	Adult	Shopping
R2	Reference	Society	Reference	Business
R3	Adult	Reference	News	Sports
R4	Society	Adult	Society	News
R5	Sports	Kids&teens	Health	Society
R6	Games	Business	Kids&teens	Regional
R7	Health	Games	Games	Recreation
R8	Business	Health	Science	Reference
R9	Kids&teens	Science	Sports	Arts
R10	Science	Sports	Home	Adult
R11	Regional	Recreation	Arts	Kids&teens
R12	Arts	Regional	Computers	Health
R13	Home	Home	Business	Home
R14	Recreation	Computers	Recreation	Science
R15	Computers	Arts	Regional	Games
R16	Shopping	Shopping	Shopping	Computers

上下4カテゴリの各測定地点の順位



- ・ アフリカと南米: 平均RTTのカテゴリ間の格差が小さいため、平均RTTと応答時間削減量との相関性は小
- ・ 他地域: 平均RTTが大きなカテゴリの応答時間削減効果が大

アフリカと南米以外の地域は、応答時間削減効果の高いカテゴリが共通

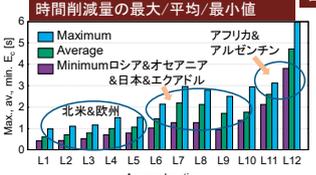
NTT

Copyright©2016 NTT corp. All Rights Reserved. 22

## 各カテゴリの平均Web応答時間削減量

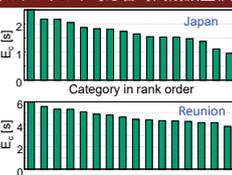
$p=2$

各測定地点の各カテゴリの平均応答時間削減量の最大/平均/最小値



L1	Massachusetts	L4	Ireland	L7	Australia	L10	Ecuador
L2	Wisconsin	L5	Germany	L8	New Zealand	L11	Argentina
L3	California	L6	Russia	L9	Japan	L12	Reunion

各カテゴリの平均応答時間削減量 (降順)



$E_c$ : カテゴリ c の平均応答時間削減量

- ・ 北米・欧州以外の地域ではエッジ配信の効果が大きく、ロシア・オセアニア・日本・南米が1~3秒程度、アフリカが4~6秒程度、応答時間を削減
- ・ これら地域では、カテゴリ間で応答時間削減効果に1~2秒程度の差異

NTT

Copyright©2016 NTT corp. All Rights Reserved. 23

## 方向性2: キャッシュ配信の「価値」の向上

- ・ コンテンツ事業者がキャッシュ配信によって得られる効用値を各コンテンツに対して定義
- ・ キャッシュ制御で効用値の総和を最大化させることが目標

■ Data Sponsoring (ゼロ・レーティング) (※無線配信が対象)

- ・ 無線NW事業者が特定コンテンツの通信料を無償化(T-MobileのポケモンGoトラヒック無料化など)
- ・ コンテンツ事業者がユーザの料金を負担してでも配信する価値があると判断 ⇒ 配信の価値大

■ Contract-based TTL cache [Ma15]

- ・ コンテンツ事業者からキャッシュ保持時間(TTL)に比例する料金を徴収

■ Optimum TTL cache [Dehghan16]

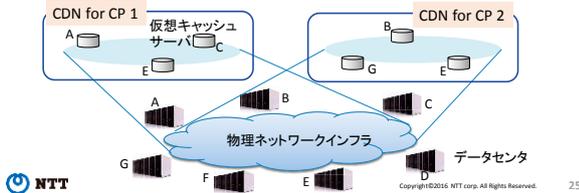
- ・ ヒット率 $h_i$ の関数でコンテンツ i の効用値  $U_i(h_i)$  を定義
- ・ キャッシュ保持時間(TTL)の値を、 $U_i(h_i)$ が最大化するよう設定

NTT

Copyright©2016 NTT corp. All Rights Reserved. 24

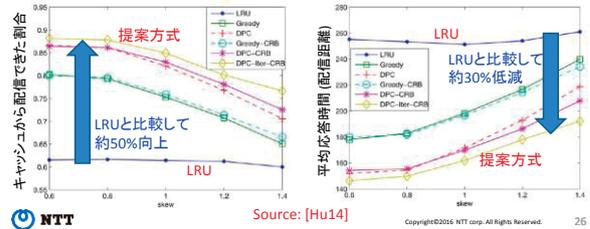
### 方向性3: プロアクティブ展開との併用

- 需要予測が容易なコンテンツ, 需要が安定したコンテンツに限定すれば, プロアクティブ展開が有効
- 仮想化CDN ⇒ 特定のコンテンツ事業者にフォーカス
  - 需要の予測精度の向上
  - プロモーション戦略とリンク
  - 対象コンテンツが限定されるため最適化問題を解く上で有利



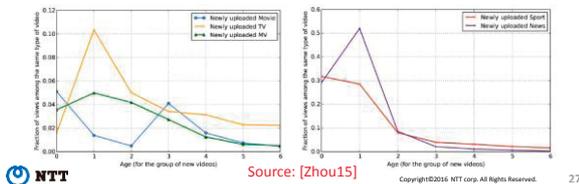
### 仮想化CDNにおけるプロアクティブ展開

- 貪欲算法に基づくプロアクティブ展開の研究例 [Hu14]
  - オリジン配信コスト, ストレージ資源調達コスト, 配信資源(VM)調達コスト, ユーザとキャッシュサーバの距離制約が付与
  - 総コストを最小化させる二つのGreedy型アルゴリズム



### 需要化変パターンに応じたキャッシュ制御

- 中国のオンラインVoD事業者Tencent Videoのアクセスログを分析
- アップロード日から1週間の視聴数に対する各日の視聴数の割合をNews, Sports, TV, MV (music video), Moviesのカテゴリごとに平均値をプロット
  - Sports, News: 初期の視聴数が多く, その後は急減
  - TV, MV, Movies: 緩やかに推移
- キャッシュ容量を二つに分割し, 一方をFIFOで, 他方をLFUで更新
- Sport, NewsはFIFOキャッシュに, TV, MV, MoviesはLFUキャッシュに収容



### まとめ

- コンテンツ配信の品質改善, トラフィック量削減のためCDNが広く普及
- ヒット率やリンク負荷抑制の観点からは, 需要のローカルリティに適切可能なLRUがキャッシュ置換制御法として良好
- 従来のキャッシュ制御は全コンテンツを同等に扱うが, 需要の時空間的パターンやキャッシュ配信の効果はコンテンツごとに大きく異なる.
- キャッシュ制御の方向性
  1. ユーザQoE改善効果の向上
  2. キャッシュ配信の「価値」の向上
  3. プロアクティブ展開との併用

### References

- [Breslau99] L. Breslau, et al., Web Caching and Zipf-like Distributions: Evidence and Implications, IEEE INFOCOM 1999.
- [Brodersen12] A. Brodersen, et al., YouTube Around the World: Geographic Popularity of Videos, WWW 2012.
- [Cao09] P. Cao, et al., Cost-Aware WWW Proxy Caching Algorithms, USENIX 1997.
- [Cisco15] Cisco Visual Networking Index: Forecast and Methodology, 2015–2020.
- [Dehghan16] M. Dehghan, et al., A Utility Optimization Approach to Network Cache Design, IEEE INFOCOM 2016.
- [Dudouet14] F. Dudouet, et al., A Case for CDN-as-a-Service in the Cloud: A Mobile Cloud Networking Argument, IEEE ICACCI 2014.
- [Herbaut16] N. Herbaut, et al., Content Delivery Networks as a Virtual Network Function: a Win-Win ISP-CDN Collaboration, GLOBECOM 2016.
- [Hu14] M. Hu, et al., Practical Resource Provisioning and Caching with Dynamic Resilience for Cloud-Based Content Distribution Networks, IEEE Trans. Parallel and Distributed Systems, Vol.25, No.8, pp.2169-2179, Aug. 2014.
- [Jacobson09] V. Jacobson, et al., Networking Named Content, CoNext 2009.
- [Jin12] Y. Jin, et al., CoDas: An Experimental Cloud-Centric Content Delivery Platform for User-Generated Contents, IEEE ICNC 2012.

- [Kamiyama15] 上山憲昭, 他, 時系列モデルを用いたYouTubeビデオ視聴数分布の再現, 信学技報, NS2015-81, 2015年9月.
- [Kamiyama16] N. Kamiyama, et al., Analyzing Effect of Edge Computing on Reduction of Web Response Time, IEEE GLOBECOM 2016.
- [Kohavi07] R. Kohavi and R. Longbotham, Online Experiments: Lessons Learned, IEEE Computer, Vol.40, No. 9, pp.103-105, Sep. 2007.
- [Ma15] R. Ma, et al., Cashing in on Caching: On-demand Contract Design with Linear Pricing, ACM CoNEXT 2015.
- [Marchetta16] P. Marchetta, et al., MC3: A Cloud Caching Strategy for Next Generation Virtual Content Distribution Networks, IFIP Networking 2016.
- [Narayanan16] S. Narayanan, et al., Reducing Latency Through Page-aware Management of Web Objects by Content Delivery Networks, ACM SIGMETRICS 2016.
- [Otto12] J. Otto, et al., Content Delivery and the Natural Evolution of DNS, ACM IMC 2012.
- [Sharma13] A. Sharma, et al., Distributing Content Simplifies ISP Traffic Engineering, ACM SIGMETRICS 2013.
- [Zhou15] Y. Zhou, et al., Video Popularity Dynamics and its Implication for Replication, IEEE Trans. Multimedia, Vol.17, No.8, pp.1273-1285, Aug. 2015.