

利用者管理エッジサーバを用いた CDN システムの試作

2018SE029 加藤大雅

指導教員：宮澤元

1 はじめに

Web コンテンツをインターネット経由で配信するために最適化されたネットワークとして、Content Delivery Network(CDN) が用いられている。CDN では、管理している複数の代理サーバを用いて Web サイト上のコンテンツを利用者に配信する。CDN を用いることで利用者からのアクセスを複数の代理サーバで分散処理でき代理サーバを利用者からのネットワーク距離が近い場所に設置することで利用者からのアクセスに伴う通信遅延を低減できる。

しかし、既存の CDN では利用者からのアクセス要求に対する応答時間が十分短いとは言えない。なぜなら、代理サーバと利用者間のネットワーク距離は常に十分短いとは限らないからである。また、代理サーバに利用者からのアクセスが集中しないことも保証できない。

本稿では利用者が管理するエッジサーバを代理キャッシュサーバとして利用する CDN について述べる。試作したシステムを用いた実験結果についても示す。

研究課題は以下の 2 点である。

- タグ情報に基づくプリフェッチ機能を持つ CDN 代理キャッシュサーバの試作
- 試作システムを用いた実験による CDN 代理キャッシュサーバの有効性の確認

2 研究の背景

CDN とは、Web コンテンツをインターネット経由で配信するために最適化されたネットワークである。コンテンツを一つのサーバで配信しようとする、多数のアクセスがあった際に処理の集中によって遅延が増加してしまう。そこで利用者により近い位置の別のサーバにコンテンツをキャッシュし、そのサーバからもコンテンツを配信することで、アクセス負荷を分散させ通信遅延の増加を防ぐことができる。

インターネット上の情報量の増加によって、実装された CDN というシステムだが、未だ通信遅延の問題はまだ完全には解決していない。従来の CDN の代理サーバは、元のサーバと比べて利用者に近い位置にあるとはいえ利用者のすぐそばに存在していない。そのため、距離による通信遅延が発生する。また、代理サーバにアクセスが集中してしまった場合、処理の遅延が発生してしまう。

このような問題を解決するために様々な手法が提案されている。村瀬らは Web の代理サーバを用いて HTML リンクを解析して Web ページのプリフェッチを行う手法を提案している [1]。この手法は、動画のようなリンク情報を持たないデータに対して直接適用することはできない。

3 利用者管理エッジサーバを用いた CDN

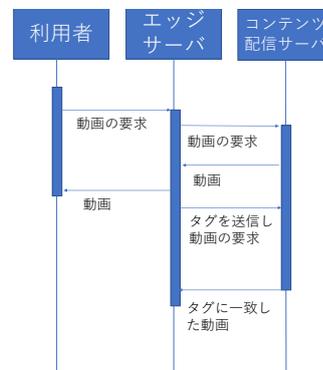


図 1: 提案 CDN のシーケンス図

本研究で提案する CDN は、コンテンツ配信サーバと利用者管理エッジサーバからなる。利用者は自らが管理するエッジサーバに接続してコンテンツにアクセスする。このエッジサーバはプリフェッチ機能を備えた代理キャッシュサーバとして働く。利用者管理エッジサーバと利用者間のネットワーク距離は非常に短いので、ネットワーク距離に起因する通信遅延を改善できる。これらのエッジサーバは利用者に管理されており、利用者がアクセス権限を承認した複数のクライアントから同時利用したり、キャッシュされたコンテンツを共有できる。これにより不特定多数のクライアントからアクセスされない、アクセス時間の増大を防ぐことができる。エッジサーバでのプリフェッチはコンテンツのタグ情報に基づいて行われる。

図 1 に提案する CDN を用いてコンテンツにアクセスする際のシーケンス図を示す。利用者からのキャッシュされていないコンテンツ要求に対して、エッジサーバはコンテンツ配信サーバにコンテンツを要求する。コンテンツ配信サーバからコンテンツが配信されるとエッジサーバはこれを利用者に提供する。一方で、エッジサーバはキャッシュされたコンテンツのタグをコンテンツ配信サーバに送信し、関連コンテンツを要求し、プリフェッチを行う。利用者が管理するエッジサーバと利用者間のネットワーク距離は非常に短く、ネットワーク距離に起因するコンテンツへのアクセス時間を改善できる。

4 キャッシュサーバの実装

前節で述べた CDN システムを実装した。nginx[2] をコンテンツ配信サーバとして利用した。エッジサーバを代理キャッシュサーバとして利用するために、リバースプロキシサーバとしての設定を行った。またキャッシュされたコンテンツのタグ情報を用いてエッジサーバとコンテンツ配

信サーバ間でコンテンツのプリフェッチを行うプログラムを実装した。エッジサーバ上のプログラムを実行すると、キャッシュされた動画からタグを取得してコンテンツ配信サーバに送信する。コンテンツ配信サーバ側から送られたタグと一致するタグを持つ動画が配信されるので、エッジサーバにキャッシュすることでプリフェッチを行う。

5 実験

利用者管理エッジサーバを用いた CDN の有用性を確認するために実験を行った。

5.1 実験環境

コンテンツ配信サーバには AWS(Amazon Web Service) の EC2 インスタンス, エッジサーバには Raspberry Pi 4 を利用して CDN を構成した。この CDN にクライアントとして実験用 PC からアクセスを行う。利用した計算機の仕様を表 1 に示す。コンテンツ配信サーバには AWS(Amazon Web Service) の EC2 インスタンス, エッジサーバには Raspberry Pi 4 を利用して CDN を構成した。この CDN にクライアントとして実験用 PC からアクセスを行う。利用した計算機の仕様を表 1 に示す。

表 1: 実験用 PC とエッジサーバとコンテンツ配信サーバの性能

役割	実験用 PC	エッジサーバ	コンテンツ配信サーバ
ハードウェア	dynabookRX73/FBE	Raspberry Pi 4 Model B	t2.micro
OS	Windows 10 home	RaspbianGNU/Linux10	Ubuntu 16.04
メモリ	8GB	4GB	1GiB

5.2 コンテンツアクセス実験

実験用 PC で curl コマンドを動作させ、コンテンツへのアクセス時間を計測した。以下の 3 通りのアクセス先のそれぞれに、動画へのインデックスと、20 分の動画の 2 通りのコンテンツにアクセスした。計測は 10 回ずつ行った。

- コンテンツ配信サーバに直接アクセスする
- エッジサーバを経由してアクセス (キャッシュなし)
- エッジサーバを経由してアクセス (プリフェッチ後)

5.3 実験結果

インデックスへのアクセス時間を表 2 に、動画へのアクセス時間を表 3 に示す。10 回の実行のうち、最初の 2 回と 10 回の結果の平均と標準偏差を掲載した。インデックスへのアクセスではコンテンツ受信にかかる時間がすべて 0 であるが、接続時間はエッジサーバにアクセスした方が短かった。動画へのアクセスでは 1 回目を除きエッジサーバにアクセスした場合の方がコンテンツ受信にかかる時間が短かった。しかし、1 回のみは逆であった。またプリフェッチの有無で結果はほぼ変化しなかった。

5.4 考察

エッジサーバに接続した際の 1 回目だけ時間がかかるのは、コンテンツがキャッシュされていないので、コンテンツ配信サーバからデータを転送する必要があるからで

表 2: インデックスへのアクセス時間 (msec.)

サーバ種類	実行回数	接続時間	コンテンツ受信		
			開始時刻	完了時刻	時間
コンテンツ配信サーバ	1	188	485	485	0
	2	219	610	610	0
	平均	188	472	485	0
	標準偏差	18	58	58	0
エッジサーバ	1	15	521	521	0
	2	94	94	94	0
	平均	51	141	141	0
	標準偏差	69	153	153	0
プリフェッチ後のエッジサーバ	1	63	586	586	0
	2	31	54	54	0
	平均	38	130	130	0
	標準偏差	28	164	164	0

表 3: 動画へのアクセス時間 (msec.)

サーバ種類	実行回数	接続時間	コンテンツ受信		
			開始時刻	完了時刻	時間
コンテンツ配信サーバ	1	203	437	142296	141859
	2	188	407	174547	174140
	平均	213	439	185022	184584
	標準偏差	38	64	28230	28254
エッジサーバ	1	93	391	583125	582734
	2	1	16	67541	67525
	平均	36	69	121681	121614
	標準偏差	39	117	162251	162138
プリフェッチ後のエッジサーバ	1	16	453	570922	570469
	2	16	16	72815	72799
	平均	29	128	117055	116979
	標準偏差	33	135	159601	159469

ある。一方で 2 回目以降はアクセス時間をかなり短縮でき、エッジサーバの有用性を示すことができた。こうした結果になったのは、エッジサーバの性能がクラウドサーバよりも低いことによる処理遅延よりも、通信距離が長くなることによる通信遅延の方が影響が大きかったからである。プリフェッチの有無で結果がほぼ変化しなかったことから、プリフェッチが期待通りに動作しなかった事がわかる。これはプリフェッチプログラムが作成するキャッシュが nginx のキャッシュ形式に合っていなかったことが原因と考えられる。

6 おわりに

本研究では、利用者が管理するエッジサーバをキャッシュサーバとして利用する CDN を提案した。試作システムを用いた実験では、コンテンツアクセス時間を向上させることができた。これにより、CDN にエッジサーバを加えることの有用性を示すことができた。今後の課題は、実装を改善してプリフェッチの効果を検証することと、様々なエッジサーバやクライアントを用いた多様な環境で実験を行い、提案手法の有効性を示すことである。

参考文献

- [1] 村瀬茂樹他. “Proxy サーバを利用した新たなサービスの可能性について” 情報処理学会研究報告グループウェアとネットワークサービス (GN) ,2000 巻 6 号,pp.19-24,2000.01.20
- [2] NGINX(エンジンエックス) | 日本公式サイト <https://www.nginx.co.jp/>