

広域 WWW キャッシュサーバネットワークの設計

川本 芳久[†] 中西 通雄[‡] 松田 秀雄[†] 橋本 昭洋[†] 坂本 佳則[§]

[†]大阪大学大学院基礎工学研究科情報数理系専攻

[‡]大阪大学情報処理教育センター

[§]奈良工業技術センター

{kawamoto, naka, matsuda, hasimoto}@ics.es.osaka-u.ac.jp,
yosino-s@niit.pref.nara.jp

概要 WWWの利用者は年々多くなっており、WWWはインターネットにおける主要なアプリケーションとして欠かせないものとなっている。それに伴って、現在インターネットを流れるトラフィックの大部分はWWWが占めており、回線の圧迫や応答の遅延が大きな問題となっている。この問題を解決するため、WWWのトラフィックでは重複するデータが多いことから、WWWキャッシングプロキシサーバ(キャッシュサーバ)を設置し相互接続することが一般的に行われるようになってきた。本報告では、WIDEインターネットで運用されているキャッシュサーバネットワークでの経験を元に、広域ネットワークにおけるWWWキャッシュサーバネットワークの設計について検討を行う。

Design of Wide Area WWW Cache Server Network

Yoshihisa KAWAMOTO[†] Michio NAKANISHI[‡] Hideo MATSUDA[†] Akihiro HASHIMOTO[†]

Yoshinori SAKAMOTO[§]

[†] Dept. of Informatics and Mathematical Science,

Graduate School of Engineering Science, Osaka University

[‡] Education Center for Information Processing, Osaka University

[§] Nara Prefectural Institute of Industrial Technology

Abstract WWW (World-Wide Web) users are growing every year, and WWW service is becoming major application of Internet. As this growth, it is a serious problem that the traffic of WWW occupies the greater part of Internet bandwidth. In order to solve this problem, WWW caching proxy servers (cache servers) are used and referred each other. In this paper, we report the network of cache servers used on WIDE Internet, and discuss with the design of WWW cache servers network on wide area network.

1 はじめに

WWWの利用者は年々多くなっており、WWWはインターネットにおける主要なアプリケーションとして欠かさないものとなっている。それに伴って、現在インターネットを流れるトラフィックの大部分はWWWが占めており、回線の圧迫や応答の遅延が大きな問題となっている。この問題を解決する一番の方法は、インターネット全体の回線容量を十分に大きくすることである。しかし、急速なWWWトラフィックの増加にネットワーク整備が間に合わなかったり、回線容量を増やした分だけWWWトラフィックが増加してしまい、結局問題が解決しない状況がある。特に組織内の回線容量が十分にあっても、組織外との接続の回線容量は不足している組織は多く、その不足している部分での混雑は深刻である。

このような状況を改善するために、WWWのトラフィックでは重複するデータが多いことから、WWWキャッシングプロキシサーバ(キャッシュサーバ)を設置することが一般的に行われるようになってきた。キャッシュサーバは、プロキシサーバの提供する通信の中継機能に加え、通信データをキャッシングする機能を持つものである。キャッシュされたデータを効率よく利用できれば、トラフィックの削減だけでなく利用者への応答速度の改善にも効果がある。

また、データのキャッシングはある程度利用規模がなければ有効に機能しないと考えられるため、複数のキャッシュサーバを使用したキャッシュサーバネットワークも従来から考えられて来た。しかし従来の手法では、管理コストの点や中継による遅延に関して、広域ネットワークに適用したときに問題があり現実的でない。例えば、階層型キャッシュや地理的プッシュ型キャッシュは、広域ネットワークに適用した場合に管理コストが高かったり中継による遅延が大きいなどの問題を持っている。これらの問題を解決するために、特に分散型キャッシュに対して、管理手法やキャッシュサーバ選択手法に対する提案が数多くなされている。

本報告では、WIDEインターネットで運用されているキャッシュサーバネットワークでの経験を元に、広域ネットワークにおけるWWWキャッシュサーバネットワークの設計について検討を行う。

なお、単にキャッシュと呼ぶ場合はキャッシュされたデータあるいはキャッシュサーバ、キャッシュサーバネットワークを指すことが多いが、ここではキャッシュされたデータあるいはキャッシュサーバ

ネットワークのことをキャッシュと呼ぶ。

2 WWWキャッシュサーバネットワーク

WWWプロキシサーバにキャッシング機能が追加されてから、キャッシュされたデータを有効に利用するために、キャッシュサーバ同士を相互接続する方法が数多く考えられてきた。

階層型キャッシュは、キャッシュサーバがツリー状に配置され、下位階層のキャッシュサーバは上位階層のキャッシュサーバへWWWリクエストの中継を依頼する形態のキャッシュサーバネットワークである。これは組織内でキャッシュサーバを運用する場合に有効な方法であるが、以下のような問題点がある。

1. 階層が大きくなるにつれて、上位階層のキャッシュサーバへ負荷が集中し、階層型キャッシュ全体の性能が低下する。
2. 階層が大きくなるにつれて、中継による遅延が大きくなり、利用者への応答速度が低下する。
3. 上位階層のキャッシュサーバが事故などで停止してしまうと、WWWアクセスができなくなる。

特に、2、3の理由により、これをそのまま広域ネットワークで利用することは困難である。

分散型キャッシュは、この階層型キャッシュの欠点を補う方法として有効である。分散型キャッシュでは、キャッシュサーバ同士がICP(Internet Cache Protocol) [3]を用いてキャッシュデータの存在を確認することができる。そのため、応答のないキャッシュサーバの回避や、柔軟な接続形態を取ることが可能となっている。

ただ、どのような接続形態を取るとよいかということについての決まった方針は存在せず、目的に応じた形態を試行錯誤で決定していることがほとんどである。

2.1 キャッシュサーバネットワークの設計

キャッシュサーバネットワークを設計する場合には、どのような目的で運用するのかによって、いくつか考慮すべき点がある。ここでは、運用の目的を以下のように分けて考え、それぞれについて何を考慮して設計する必要があるかを論じる。

1. WWWトラフィックの削減
2. 利用者への応答時間の改善
3. 広域ネットワークへの適用

2.2 WWWトラフィックの削減

現在のインターネットバックボーンにおけるトラフィックの50%から70%は、WWWトラフィックだといわれる。また、通信回線が細くなればなるほど、WWWトラフィックの全体に占める割合は多くなる傾向にある。そのため、通信回線帯域の有効利用を考えると、WWWトラフィックを削減することが非常に重要になる。

一般にWWWトラフィックには重複するデータが多いため、キャッシュサーバはWWWトラフィックの削減には効果的だと考えられている。この場合には、どのようにして重複するWWWトラフィックを減少させるかがポイントとなる。

2.2.1 キャッシュデータのヒット率の向上

キャッシュデータのヒット率は、クライアントへの応答データに対するキャッシュデータの割合として定義される。重複するWWWトラフィックの減少のためには、このヒット率を向上させて各キャッシュサーバから出るWWWリクエストを減少させることが有効である。

個々のキャッシュサーバのヒット率を向上させるための手法としては、キャッシュデータのための記憶領域を大きくとることや、キャッシュデータの廃棄アルゴリズムの改良が考えられる。

キャッシュデータが多くなると不要なデータを廃棄する必要がある。その場合、キャッシュデータの廃棄アルゴリズムは、現在LRU (Least Recently Used) に従ったものがほとんどであるが、データの種類に応じて廃棄することが考えられる。例えば、テキストデータに比べてイメージデータは比較の変更可にいと考えられるため、イメージデータを長くキャッシュしておけばヒット率は向上する。

また、WWWリクエストの回数はZipfの法則に従うという報告[7]があり、記憶領域が不足している場合にLRUでは十分にキャッシュデータが生かされないことが示されている。廃棄アルゴリズムの改良はあまり容易ではないが、ヒット率の向上のためには考えなければいけない問題である。

さらに、ヒット率向上のためにはある程度のWWWリクエストを受ける必要がある。そのため

には、キャッシュサーバの利用を積極的にアナウンスして多くの利用者に使用してもらおう努力が大切である。

2.2.2 組織的なキャッシュサーバネットワークの構築と利用

キャッシュサーバの利用は比較的小規模な組織単位で行われることが多いため、単独ではキャッシュサーバが有効に機能するためのWWWリクエストの量が得られないことがある。その場合には、より大きな組織単位でキャッシュサーバ同士を相互接続し、全体でWWWリクエストを多く確保することが有効だと考えられる。

相互接続を行う場合に組織内でよく用いられる構成として、階層型キャッシュがある。階層型キャッシュではキャッシュサーバをツリー状に配置し、一番上位のキャッシュサーバがWWWサーバに対してリクエストを出す。こうすることで、上位のキャッシュサーバは多くのWWWリクエストを確保できるため、組織全体でのキャッシュデータのヒット率が向上し、WWWトラフィック削減の効果が期待できる。

ただし階層型キャッシュには前述したような問題点があるため、以下のような対策を施す必要がある。

1. 上位階層で使用するキャッシュサーバは、高負荷に耐えることができ、高速に処理ができる計算機、OS、ソフトウェアを用意する。
2. 上位階層のキャッシュサーバは複数台の計算機を用意して、計算機が停止してもサービスが継続できるようにする。
3. キャッシュサーバの性能を常に監視して、期待している性能が出ていない場合には調整を行えるようにする。

このような問題を回避する他の手法に、分散型キャッシュがある。分散型キャッシュでは、ICPを利用して他のキャッシュサーバへキャッシュデータの存在を確認することによって、階層型キャッシュより柔軟な形態を取れるようにしたものである。

ICPを用いた代表的なキャッシュサーバプログラムとして現在よく利用されるものに、Squid[1]がある。Squidの接続形態としては親(parent)接続と兄弟(sibling)接続がある。クライアントから要求があった場合に、キャッシュサーバはまずsiblingに対してICPでキャッシュデータの有無を問い合わせ

せ、あればその sibling へ WWW リクエストを出し、なければ parent へ、parent の指定がなければ直接 WWW サーバへ WWW リクエストを出す。ICP を用いることによって階層構造を取らずに他のキャッシュサーバのデータを利用できるため、階層型キャッシュの欠点である上位キャッシュサーバの高負荷をある程度抑えることができる。

2.3 利用者への応答時間の改善

対話的処理を行っている場合に利用者が許容できる待ち時間は、10 秒もないといわれる。通信の遅延が大きい WWW サーバに対するアクセスでは、長い通信時間やタイムアウト等、利用者の許容限度を越えることも少なくない。

利用者への応答時間を短くするためにも、キャッシュサーバを利用することができる。キャッシュサーバを利用すると、キャッシュされたデータが存在する場合に利用者との通信が短い時間で済む。また、WWW サーバへの通信経路が複数ある場合に、キャッシュサーバを利用することで速い通信経路を選択することも可能である。この場合には、どれくらい利用者への応答時間を短くできるかがポイントとなる。

2.3.1 複数の通信経路の選択

インターネットにおける通信で経路が複数ある場合に、応答時間が短い経路が選ばれるとは限らず、また利用者が経路の選択を行えることはほとんどない。しかし、キャッシュサーバを利用することによって、ある程度経路を選択することができる。あらかじめそれぞれの経路を利用するキャッシュサーバを導入しておけば、より応答時間が短いほうの、あるいはより帯域の広いほうの経路を利用することが可能である。

ただし、全ての WWW サーバについてどの経路が適切か判断することは不可能であるため、どのようにしてキャッシュサーバと WWW サーバを対応させるかが問題となる。Squid においては、WWW サーバのドメイン名、IP アドレス等に応じてどのキャッシュサーバへリクエストを出すか設定することができるため、.com、.org、.ac.jp 等のドメインごとにキャッシュサーバと WWW サーバを対応させることはできる。この場合の問題点は、いくらかうまく対応づけができたとしても、静的な対応関係であるために定期的な見直しが必要だということである。

動的に対応づけるための研究としては、ネットワークの経路情報を用いる方法 [8] や、エージェントを用いた方法 [4][6] などがあるが、導入はまだ容易ではないと思われる。動的な対応を行うためには、キャッシュサーバ相互の協調が必要であるため、WWW の枠組み自体の見直しも必要だと考えられる。

2.3.2 データの先読み

WWW では、データ中に他のデータへのリンクが存在することが多く、利用者はそのリンクを利用して次のデータをリクエストすることがある。キャッシュサーバではデータ中にリンクが存在することがわかるため、そのリンクで示されるデータを先読みしてキャッシュしておくことで、利用者が次のデータをリクエストした場合の応答時間を短縮することができる。

しかし、リクエストの数が多くなると、先読みすべきデータが大量に発生し、また一度も参照されない先読みデータが増加するといった問題がある。これを解決するには、利用者の要求にあった先読みの戦略を取る必要があるが、うまい方法がないのが現状である。

先読みを行う研究 [9] に基づくキャッシュサーバとして Wcol [2] があるが、まだ一般的に使われるまでにはなっていない。

2.4 広域ネットワークへの適用

分散型キャッシュ等のキャッシュサーバネットワークは、主に組織内 LAN 等の高速で広帯域のネットワークでの利用を前提としているため、広域ネットワークに適用した場合に問題となることがある。以下に、広域ネットワークにおける分散型キャッシュの問題について述べる。

2.4.1 ICP の限界

ICP では応答が返ってくるのが保証されないため、Squid ではタイムアウトによって応答受け付けを打ち切っている。そのため、どのキャッシュサーバにもキャッシュデータがなく、応答の遅いキャッシュサーバがあった場合に、キャッシュの有無の確認だけに長い時間がかかってしまう。特に、広域ネットワークで ICP を利用すると、この遅延による利用者への応答時間の悪化が問題となる。また、ICP による接続を行っているキャッシュサーバ間の回線帯

域が狭い場合に、ICPバケットによってバケット数が増加しルータの負荷が増大することがある。

さらに、ICPでのヒット率はかなり小さいことがわかっており、広域ネットワークにおけるICPによる問い合わせはコストが高いといえる。

2.4.2 異なるデータ参照ポリシー

広域でキャッシュサーバネットワークを構築した場合には、複数の組織が参加することが一般的である。その場合、ある組織内でしか参照を許可していないデータでも、その組織のキャッシュサーバを経由すると他の組織でも参照できてしまうことがある。逆に、アクセス拒否のメッセージがキャッシュされてしまった場合には、他の組織では見られるはずのデータが参照できないということになる。

また、ICP問い合わせに対するアクセス許可とキャッシュデータ参照に対するアクセス許可が異なっている場合、ICPでキャッシュデータの存在を確認してもキャッシュデータが参照できないという問題が起こる。

これらは、個々のキャッシュサーバの設定の問題であるが、気づきにくく、データが漏洩してしまうから対応しても手後れであることも多い。

3 WIDEプロジェクトキャッシュサーバネットワーク

WIDEプロジェクト W4C ワーキンググループでは、広域キャッシュサーバネットワークの挙動と有用性についての研究環境を構築するため、1997年4月からWIDEインターネット内でキャッシュサーバネットワークの構築を行ってきた。WIDEインターネットのWWWキャッシュのバックボーンという意味合いを含め、この環境をWIDE CacheBoneと名づけている。

キャッシュサーバプログラムはSquidを用い、ICPを利用した分散型キャッシュの構築を目標とした。

3.1 CacheBone の設計

当時はまだ、広域キャッシュサーバネットワークに対する指針等がなかったため、設置と並行して以下のような設計を行った。

- WIDEインターネットの中継点であるNOC (Network Operation Center) を通過するICP

は極力少なくする。そのために、できる限りNOC上にキャッシュサーバを設置する。

- NOCでない組織のキャッシュサーバは、一番近いNOCのキャッシュサーバのみとsibling接続する。
- NOCのキャッシュサーバは、2ホップまでのNOCのキャッシュサーバとsibling接続する。
- 接続の多いNOCのキャッシュサーバは、ドメイン別のparentとなってもらう。

構築を開始してから約半年かかって、図1のような形態となった。CacheBoneには11組織が参加しており、ネットワーク組織が3、大学組織が8である。

3.2 問題

1年程度運用を行って、各キャッシュサーバのアクセスログを解析した結果、前述したようなICPの限界やデータ参照のポリシーについての問題が明らかとなっている。

ICPの限界に関しては、現在sibling接続を一切中止し、対外接続を行っているNOCへのparent接続に変更することによって、問題をある程度回避している。ただ、この形態も問題の解決にはなっておらず、今後の課題である。

また、CacheBoneの設計当時に比べ、現在ではWIDEインターネットのバックボーンの回線帯域が太くなっている。これに伴い、設計当時はバックボーンのWWWトラフィックの削減が大きな目標であったが、現在では利用者への応答時間の改善がより大きな目標となってきている。そのため、目標の変化にあわせたCacheBoneの再設計を行う必要がある。

4 まとめ

本報告では、WIDEインターネットで運用されているキャッシュサーバネットワークでの経験を元に、広域ネットワークにおける分散型キャッシュの設計について述べた。広域キャッシュサーバネットワークを設計するにあたっては、どのような目的でキャッシュサーバを設置するのかを明らかにして、明確に目標を設定することが必要である。また、広域ネットワークにおいて、ICPによるsibling接続

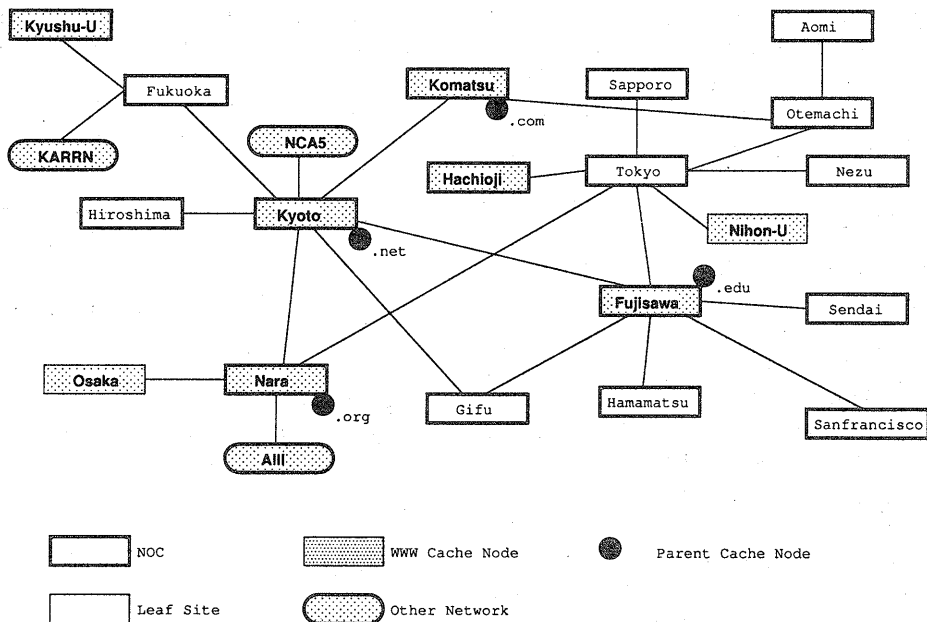


図 1: WIDE CacheBone

はコストが高く、他の手法を用いる必要があると思われる。

W4C ワーキンググループでは、今後もこの CacheBone を用いて、より目標を達成できる広域キャッシュサーバネットワークについて研究を進めていく予定である。

参考文献

- [1] D. Wessels et al., "Squid Internet Object Cache", <http://squid.nlanr.net/Squid/>
- [2] 知念賢一, "WWW Collector - The Prefetching Proxy Server for WWW", <http://shika.aist-nara.ac.jp/products/wcol/>
- [3] D. Wessels and K. Claffy, "Internet Cache Protocol (ICP), version 2", Request For Comments 2186, September 1997.
- [4] 富川裕樹, 高井昌章, 山本強, "アプリケーション層での動的な WWW アクセス経路選択", 情報処理学会研究報告 Vol.98-DSM-11, pp. 1-5, 1998.
- [5] 鍋島公章, "WWW キャッシュサーバの性能評価とチューニング", 情報処理学会研究報告 Vol.97-DSM-8, pp. 49-54, 1997.
- [6] 田中友英, 篠田陽一, "WWW における動的経路制御を用いた多段キャッシュシステム", 情報処理学会研究報告 Vol.97-DPS-83, pp. 49-54, 1997.
- [7] 西川記史, 細川貴史, 辻洋, 森靖英, 吉田健一, "WWW トラフィック分析と分散キャッシュ", 情報処理学会研究報告 Vol.97-DSM-5, pp. 7-12, 1997.
- [8] C. Grimm, M. Neitzner, H. Pralle, J. Voেকler, "Request Routing in Cache Meshes", 3rd WWW Caching Workshop, <http://wwwcache.ja.net/events/workshop/>, 1998.
- [9] 知念賢一, 山口英, "WWW 先読み代理サーバにおける先読み対象決定戦略", 情報処理学会研究報告 Vol.96-DPS-77, pp. 55-60, 1996.