

インターネットにおけるドメインキャッシュサーバの実現

1U-6

高品 智一 渡辺 尚 水野 忠則
静岡大学 工学部

1 はじめに

近年、インターネットを使用した広域な情報提供サービスが行われてきている。それは、点在するサーバに情報を持たせ、多数存在するクライアントからアクセスできるようにしたものである。しかし、1つのサーバに対して多くのクライアントが同時にアクセスするため、1つのクライアントに対する応答時間が長くなる傾向にある。また、サーバがダウンしたときには、そのサーバが持つ情報を得ることができなくなる。

我々は分野や意味が同じ情報の集合であるデータベースなどを、どのように分割しどのサイトに配置すれば、もっとも効率よくかつ信頼性を高くできるかを検討してきた[1,2]。本研究では、クライアントがサーバからとってきた情報を記憶するドメインキャッシュサーバ(DCS)を提案する。これによって、サーバの負担を減らし、さらにはインターネット全体の負荷を減らし、サービスを向上させることを目的としている。また、サーバがダウンした場合、DCSに蓄積された情報によってサービスを継続し、信頼性を高めることも目的の1つである。

2 従来研究

本研究のモデルとなるインターネットの情報提供サービスの処理は、クライアント-サーバ方式をとっている。即ち、クライアントとなるユーザがサーバを選択し、サーバの中の欲しい情報を引き出す方式である。

インターネット上の広域情報サービスで利用されている情報のキャッシュ方式について、Proxyサーバ方式とWWFS方式[3]の2つについて特徴を述べる。

2.1 Proxyサーバ

Proxyサーバは、security上の理由からクライアントが自由に外のネットワークにアクセスできない場合に、代理のサーバの役目を果たすものである。このProxyサーバを使うことによって、IP reachableでないクライアントからもWWWを使うことができるようになる。また、マッキントッシュのクライアントのようにクライアント側で漢字変換をサポートしていない場合、Proxyサーバ側でShift-JISに変換することで、マッキントッシュ

上で日本語が読めるようになる。

Proxyサーバにはキャッシュの概念が存在し、1度アクセスした情報は記憶しておいて次からはその情報を利用することができる。ただし、キャッシュ情報の管理とクライアントとの通信により負荷が集中するといった問題がある。

2.2 WWFS

WWFSの特徴は広域ネットワーク環境における情報共有を支援する分散ファイルシステムを実現していることである。そのシステムの基本構成(図1)では、既存のプロトコルとの整合性をとるために、クライアントからサーバへのファイルアクセスの中間にcsd(cluster server daemon)と呼ぶサーバを置き、クライアント(LAN)側と広域ネットワーク側にある情報提供サーバ群とのアクセス・プロトコルの変換を行う。

WWFSにもキャッシュの概念が存在するが、Proxyサーバと同様の問題がある。

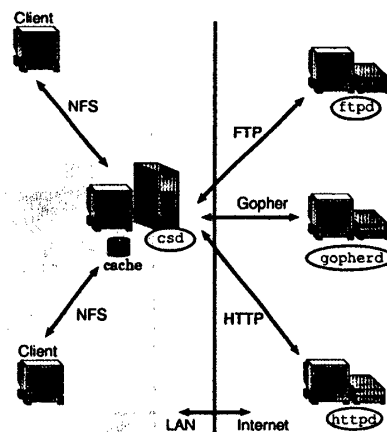


図1: WWFSの構成

3 ドメインキャッシュサーバの導入

ProxyサーバとWWFSの両システムにはキャッシュの概念が存在する。しかし、どちらの方式においても“1台の端末に負荷が集中する”といった問題点が存在する。そこで我々は、

- 負荷を分散する
- データの一貫性を考慮する
- 応答時間を短くする

という3点に重点を置き、これらを満たす方式(DCSシステム)を検討した。

図2に我々が提案するDCSシステムの構成を示す。クライアントとサーバ間はFTP等の通常のプロトコルを拡張したもので通信を行う。クライアントが取得した情報をDCSが記憶したり提供するために、クライアントとDCSとの間に新しい通信プロトコルを設けている。

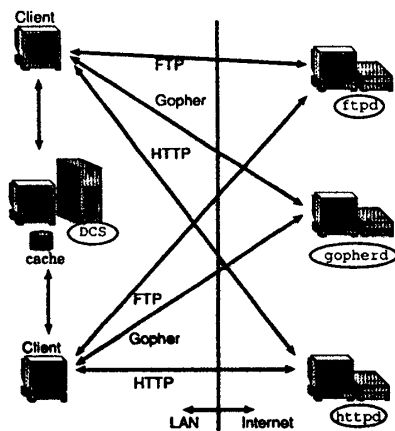


図2: DCSシステムの構成

次に、DCSシステムの処理の流れを述べる。まず、クライアントは通常どおりデータを送ってくれるようにサーバに要求を送る。サーバ側はクライアントから送られてきたメッセージをもとに、必要なファイルを探す。その際、もしそのファイルがDCSに存在するならば、クライアント側にそこからデータを取ってくるようにする。これはサーバ側が、どのファイルの複製がどこのDCSに存在するかといった情報を持つことによって実現する。そして、存在しないならば、サーバがクライアントにそのデータを送り、クライアントは都合のいいときに(一貫性を考慮して)DCSにそのデータを送る。

以上によって、同じドメイン内のクライアントから同じサーバに同じデータを取りに行く場合には、そのサーバへのアクセスは1回ですむことになる。即ち、クライアントとサーバの間のネットワークの負荷を減らしたり応答時間を短縮することが可能になる。

DCSシステムがProxyサーバやWWFSと違うのは、前者はftpdやhttpdやgopherd等のサーバとは別に新たにDCSというキャッシュのみを扱うサーバを導入したことである。従来方式はクライアントとサーバとの中継を行う端末を使ってキャッシュを実現しており、その端末に負荷が集中する可能性があるといった問題点が存

在する。DCSは中継を行う端末が存在しないため、負荷が集中するといった問題を避けることができる。

4 ドメインキャッシュサーバの実装

現在、前の章で述べたシステムが有効であることを証明するために、DCSシステムのプロトタイプを作成しているところである。これはWWWをターゲットにして、このシステムに合うようにクライアント(Mosaic)側とサーバ(NCSA httpd)側を変更し、さらにDCSを作成するものである。

主な変更点は、

- クライアントがデータを要求するときにhttpでは“GET”を用いるが、拡張したプロトコルでは“DGET”や“RGET”を用いる。これによって新しいhttpdと従来のものとの互換性を保つ。
- サーバにドメインごとのファイル参照履歴を管理する機能を取り付ける。これによってDCSのキャッシュの状態を管理する。
- コネクションキャッシュ[4]機能を取り付ける。これによってコネクション確立時間を短縮し応答時間を短くする。

といった点である。

5 おわりに

本論文でDCSシステムを導入した。このシステムではProxyサーバやWWFSなどで問題となる負荷の集中を解決するために、クライアントとサーバに負荷を分散している。今後の課題としては、実装に基づくシステムの有効性について検討し、さらに、それらを世の中に出ているシステムに導入してもらえるようにしていきたい。

参考文献

- [1] 宮西, 高品, 渡辺, 水野: 分散システムにおける独立性の観点からの資源分散方法の提案, 情報処理学会第48回全国大会 7D-1, (1994)
- [2] 高品, 宮西, 渡辺, 水野: 分散環境における部分データベースの複製配置法, 情報処理学会 マルチメディア通信と分散処理, pp.7-12(1994.7)
- [3] 山口 英: WWFS(前編), UNIX MAGAZINE, 株式会社アスキー, pp34-42(1994.11)
- [4] Costales, B. with Allman, E. and Rickert, N.: *sendmail*, O'Reilly & Associates, Inc., pp.493-494(1993.11)