

動的ミラーリングにおける分散管理を用いたミラーサーバの割当て

03T229 河田 光司（最所研究室）

動的ミラーリング機構において、ミラーサーバ群へのミラーサーバの加入、離脱、ミラーサーバの検索に関する分散管理アルゴリズムを設計した。さらに、そのアルゴリズムに対して、分析とシミュレーションによる評価を行い、その結果、通信量の観点で設計したアルゴリズムが効率的であることが分かった。

1 はじめに

現在、Web サーバの負荷を軽減させるために、ミラーリングという手段が用いられており、DNS ラウンドロビンや、負荷分散装置^[1]などを使ってアクセスをミラーサーバに振り分けている。これらの方法ではミラーサーバのデータは特定の Web サーバのものになっており、その Web サーバのミラーリングしかできない。これに対して、我々の研究室では Web サーバの負荷状況に応じてミラーサーバを割当てる動的ミラーリング機構の研究^[2]を行っており、この機構を運用していくためには機構内のミラーサーバを管理する必要がある。

本研究では、動的ミラーリング機構に属しているミラーサーバ群を分散管理する方法を提案する。提案の管理方法は Web サーバが複数のミラーサーバを用いてミラーリングする時やミラーサーバが複数の Web サーバに対してミラーリングする時、いわゆる多対多の関係を扱うことができる。このときのミラーサーバと Web サーバの対応関係、ミラーリングしているデータの取扱いなどを管理する。論文では、この分散管理におけるミラーサーバ群へのサーバの加入、離脱、ミラーサーバの検索に関するアルゴリズムを設計し、評価している。

2 動的ミラーリングにおける分散管理

ミラーサーバの分散管理は各ミラーサーバが一部のミラーサーバ群の情報を持ち、サーバ情報をお互いに通知し合うことで群全体を管理する方式である。以下の前提条件を満たすように分散管理を設計する。

● 前提条件

- 管理専用のサーバを用意することなく、完全な分散管理とする
- Web サーバ群とミラーサーバ群は流動的であるとする
- Web サーバとミラーサーバは多対多の関係とする
- ミラーサーバの割当ては、ミラーサーバ側で行う（クライアントはリクエストを送る

のみ) とする

- ミラーサーバ側のみ管理情報を持つとする
- サーバ性能ごとにミラーサーバの接続形態を構成するとする

ミラーサーバ群の接続形態は図1に示すように、サーバ性能の高低で親子関係を築き、各ミラーサーバは親と子のリンクを持つ。さらに、図には記述していないが同世代の兄弟サーバとのリンクを持つ。このサーバの管理構造は以下の条件で作られた親子関係を持つ。

1. 各世代に属する最高性能サーバはその上位世代の最低性能サーバよりも性能が低い
2. 各世代に属する最低性能サーバはその下位世代の最高性能サーバよりも性能が高い
3. 横のサーバ数は縦のサーバ数 +2 が上限、縦のサーバ数 -2(最下位階層は縦のサーバ数 -1) が下限である
4. 全階層のサーバ数の差が1以下である
5. 加入時の条件として、登録された世代のサーバ数がその上位世代のサーバ数を上回らない
6. 離脱時の条件として、離脱が起きた世代のサーバ数がその上位世代のサーバ数を下回らない
7. サーバ性能の項目ごとにサーバの管理構造を構成する

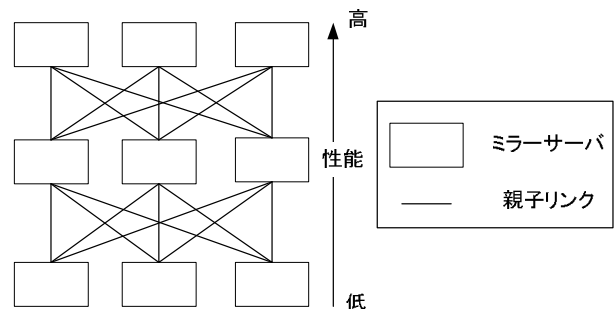


図1 サーバの管理構造

3 アルゴリズム設計

本要旨では紙面の制限のためミラーサーバ群への加入時のアルゴリズムについてのみ説明する。ミラーサーバ群への加入の全体の流れを図2に示す。この図は縦のサーバ数が2から3になるまでの流れである。ミラーサーバ群はサーバの新規加入により横に拡張しようとする。このとき、加入したサーバ性能に基づく再構成が行われる。さらに横のサーバ数が縦のサーバ数+2を超えた場合に、縦のサーバ数を一つ増加させる。図ではサーバ数が、8から9に増加するとき、この処理が行われている。縦のサーバ数が N 、横のサーバ数が $N+2$ とすると、全サーバ数は以下の数式になり縦サーバ数を1つ増加させたときの構造よりも1個少なくなっている。

$$\begin{aligned} N \times (N + 2) &= N^2 + 2N \\ &= N^2 + 2N + 1 - 1 \\ &= (N + 1) \times (N + 1) - 1 \end{aligned}$$

このように、ミラーサーバ数に応じて、縦と横のサーバ数を均等に拡大していくことができる。

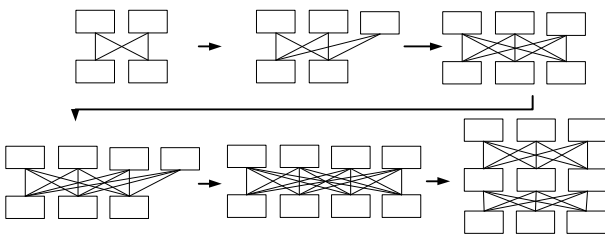


図2 加入の全体の流れ (四角はミラーサーバ)

4 アルゴリズム評価

加入時のアルゴリズムは最初に加入する場所や、ミラーサーバ群の配置状況によって変化する。このアルゴリズムを分析すると、最大ステップ数、最大通信量および最大通信量のオーダーは表1のようになる。 a は適切な階層と加入した階層の差、 b は上に移動する階層と最上位階層との差、 s は縦のサーバ数、 n は全サーバ数である。この表の式から得られる値を理論値と呼ぶ。

表1 加入における最大ステップ数と最大通信量

最大ステップ数	$a + 2 + 1 + 3b + s + 1$
最大通信量	$a + 2 + b \times (2 + \text{親子兄弟数} \times 2) + \sum_{l=2}^{s+1} l + (s - 1) + \sum_{l=2}^{s+1} l \times (2 + \text{親子兄弟数} \times 2)$
通信量のオーダー	$n\sqrt{n}$

次に、シミュレーションにより評価する。シミュレ-

ータはLinuxの1つであるFedora core 4の上でC言語を用いて開発した。このシミュレータはプロセスをミラーサーバと見立て、ミラーサーバ間の通信にはTCP/IPを用い、ポート番号でミラーサーバを識別して通信した。通信量はsend関数の使用回数に比例するので、その回数を測定することにした。

具体的には、ミラーサーバの加入により、縦のサーバ数が増加が起こる場合について測定した(ステップ数と通信量が最大になる)。その結果を図3に示す。図の横軸は、再構成が起こった後の縦のサーバ数を示す。図の結果からシミュレーション値は、最大通信量の理論値を超えることがなかった。これは理論値の計算において、親子兄弟数を最大値にしているためである。オーダーの $n\sqrt{n}$ にそってシミュレーション値は増加しており、このオーダーを超える通信量が発生することはなかった。通信量の観点において、 $n\sqrt{n}$ のオーダーは実際に運用していくのには効率的なアルゴリズムと考える。

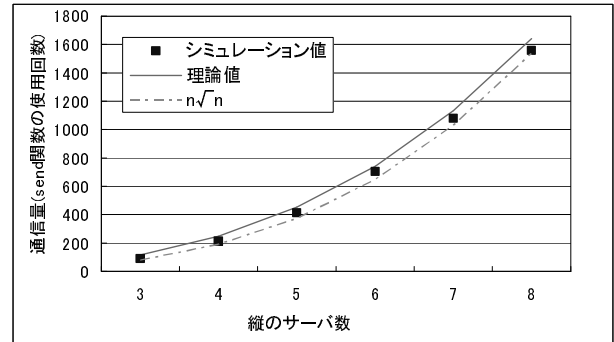


図3 加入時の最大通信量

5 まとめ

本研究では、動的ミラーリングにおけるミラーサーバの分散管理法を提案し、ミラーサーバ群への加入、離脱、ミラーサーバ検索に関するアルゴリズムを設計し評価した。その結果は通信量の観点で設計したアルゴリズムが効率的であることを示している。今後の課題としては分散管理におけるミラーサーバダウンやネットワークエラーに関する処理を考える必要がある。

参考文献

- [1] Tony Bourke, 「サーバの負荷分散技術」オライリー・ジャパン 2001年
- [2] 入江正行, 「Webサーバの負荷状況に応じた動的ミラーリング機構」, 香川大学工学部, 2005年