

クラスタファイルシステムにおいて負荷分散を実現する ファイル配置手法

丈達 翔太[†] 芝 公仁[†] 岡田 至弘[†]
[†] 龍谷大学理工学部

1 はじめに

近年、情報技術の発達により利用されるファイル数およびファイルサイズが増加してきている。これに伴い、ファイルシステムへの負荷が増大している。そのため、ファイルシステムの負荷分散を行う技術としてクラスタファイルシステムが注目されている。

クラスタファイルシステムにはファイル管理にメタデータサーバを用いるものが多い。しかし、このようなクラスタファイルシステム他に、メタデータサーバを持たないクラスタファイルシステム [1] がある。その中のひとつにハッシュテーブルを用いてファイル管理を行うものがある。このようなクラスタファイルシステムでは、ノードの追加や削除による記憶領域の増減時に各ノードが担当するハッシュ空間の割り当て変更が発生し、その結果に応じて格納ファイルの再配置が行われる。このファイル再配置時のファイルの移動はシステムに大きな負荷を与える。また、ファイル再配置後、各ノードが担当するハッシュ範囲の大きさに差があると、ノード間で負荷の偏りが発生する。本稿では、ファイル再配置におけるコストを低減させるためのファイル再配置手法を提案し、その評価を行う。

2 GlusterFS

本研究では、ハッシュテーブルを用いたクラスタファイルシステムとして GlusterFS を用いる。GlusterFS は、Red Hat, Inc により開発されたオープンソースのクラスタファイルシステムである。brick と呼ばれる各ノードの記憶領域を束ね、1 つのファイルシステムを実現する。

GlusterFS では、多くのクラスタファイルシステムにあるメタデータサーバを使用せず、ハッシュ関数を用いてファイル名からハッシュ値を求め、そのハッシュ値よりファイルの管理を行う。すなわち、32bit の範囲を持つハッシュ空間を総 brick 数分に均等に分割し、各 brick に割り当てる。その割り当てられた担当範囲に該当する brick にファイルを格納する。ノードの追加や削除が行われた場合、全てのノードで改めて担当範囲を均等に割り当て直す。このとき、新たに割り当てられた担当範囲に従って格納ファイルを移動させる必要がある。

File placement technique for load balancing in cluster file system

Shota Jotatsu[†], Masahito Shiba[†] and Yoshihiro Okada[†]

[†] Faculty of Science and Technology, Ryukoku University

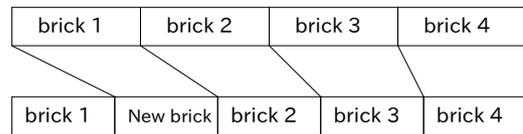


図 1 ハッシュ均等割り当て

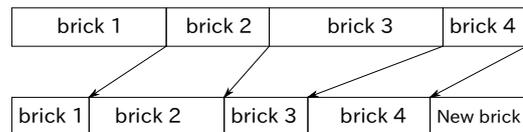


図 2 ランダム割り当て

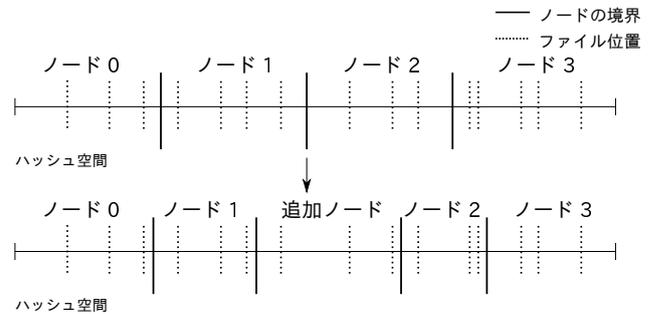


図 3 ファイル数均等割り当て

図 1 のハッシュ均等割り当ては、GlusterFS で使用されているファイル配置手法である。この手法では、全ての brick の担当範囲が変更されるためファイルの移動が起こりやすくなる。また、ハッシュ空間上のファイル位置の偏りから brick に格納されるファイルの偏りが生じる。そこで、その偏りを削減するためにハッシュ空間上のファイルの位置を考慮したファイル配置手法を提案する。

3 ファイル再配置手法

本稿では、ハッシュ空間上のファイル位置を考慮した以下の手法を提案する。

- 部分分割割り当て
- ランダム割り当て
- ファイル数均等割り当て

部分分割割り当ては、最もハッシュの担当範囲が広い brick の半分の範囲を追加 brick に割り当てる。このように、ファイルの再配置が行われる brick を 2 つに限定することにより、ファイル再配置時における処理負荷を低減することができる。

ランダム割り当てでは、図2のように、全ての brick に無作為に担当範囲を範囲を割り当てる。このとき、全ての brick の担当範囲の合計がハッシュ値の最大値を越えることがないように、担当範囲は1から $(0xFFFFFFFF / \text{brick 数})$ とする。また、追加 brick は最後尾に配置する。

ファイル数均等割り当てでは、図3のようにファイル再配置後に全ての brick で格納ファイル数が均等になるようにハッシュの担当範囲を割り当てる。全ての brick で格納ファイル数が均等になるため、ファイル再配置後の負荷が分散されと考えられる。

4 評価

各ファイル再配置手法でのファイル再配置時のファイルの移動数、ファイル再配置後の負荷コストの測定を行った。実験は以下の条件で行った。

- 10000 個のファイルを格納した brick を1つ作成する
- brick を1個ずつ追加し、追加する度にファイル再配置を行う
- brick が10個になるまで繰り返す

本稿では、ハッシュ均等割り当て、ランダム割り当て、ファイル数均等割り当てでの実験を行う。また、ファイル再配置におけるコストの評価を行うために以下の式を定義する。

$$S = F + D + M + H \quad (1)$$

式1における各項について説明する。

F は、各 brick が総ファイル数の何割のファイルを格納しているのかを算出し、その割合における標準偏差を算出したものである。

D は、各 brick が総ファイルの合計容量の何割を格納しているのかを算出し、その割合における標準偏差を算出したものである。

M は、ファイル再配置時に移動されたファイルの割合である。

H は、各 brick のハッシュの担当範囲がハッシュ値の最大値の何割であるかを算出し、その割合における標準偏差を算出したものである。

S は、各項の合計値であり、ファイル再配置におけるコストを示す。この Score の値が小さいほど、低コストな手法であるといえる。

ハッシュ均等割り当てにおけるファイル再配置時のスコアを図4に示す。図4からスコア値がほぼ横ばいになっていることがわかる。これは、全ての項において偏りが一定であるためだと考えられる。

ランダム割り当てにおけるファイル再配置時のスコアを図5に示す。図5からファイル再配置を繰り返す度にスコア値が増加していることがわかる。これは、ランダムな割り当てを行ったことにより、各項における偏りが増加したためである。

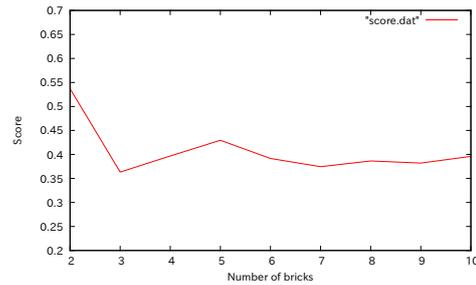


図4 ハッシュ均等割り当て

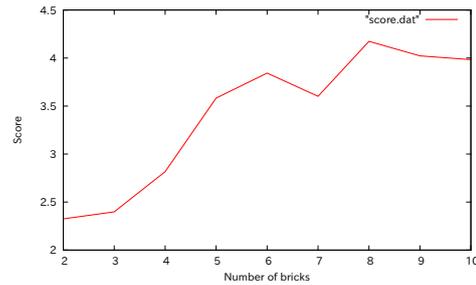


図5 ランダム割り当て

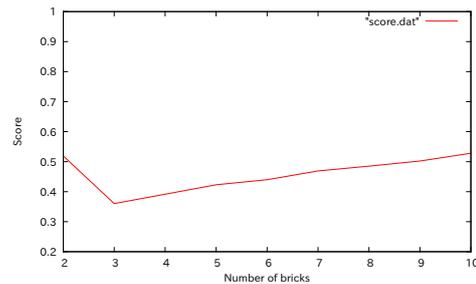


図6 ファイル均等割り当て

ファイル数均等割り当てにおけるファイル再配置時のスコアを図6に示す。図6からスコア値がゆるやかに増加していることがわかる。ファイル数均等割り当てでは、各 brick の格納ファイル数の偏りを低減させたが、他の項において偏りが増加してしまったためと考えられる。

5 おわりに

本稿では、GlusterFSにおけるファイル再配置について述べ、ハッシュ空間上のファイル位置を考慮したファイル配置手法を提案した。また、ファイル再配置における負荷コストを評価する式を定義し、各手法について評価を行った。

参考文献

- [1] Levy, Eliezer and Silberschatz, Abraham "Distributed File Systems: Concepts and Examples", ACM Computing Surveys, pp. 321-374 (1990)