

B-047

## 二重キャッシュ環境の下位キャッシュにおける参照の局所性の解析 Analyses of Locality of Reference in L2 Cache

竹内 洸祐<sup>†</sup>      長廻 雄介<sup>†</sup>      山口 実靖<sup>†</sup>  
Kosuke Takeuchi   Yusuke Nagasako   Saneyasu Yamaguchi

### 1. はじめに

ネットワークストレージへのアクセスはサーバ計算機上のキャッシュ(上位キャッシュ)とストレージ機器のキャッシュ(下位キャッシュ)を介して行われる。サーバキャッシュの置換アルゴリズムに LRU が使用されている場合、ネットワークストレージにおいては一度アクセスされたデータが近い将来に再度アクセスされる可能性が低くなり、通常とは逆向きの参照の時間的局所性(負の参照の時間的局所性)が存在する可能性がある。これに関連する既存の研究として、長廻らのシミュレーションによる解析[1]と、Zhouらによる DBMS や DISK I/O 用いた解析[2]があるが、共に実システムの動作を十分に考慮しているとは言えず、実システム上で実アプリケーションを動作させての解析が必要であると考えられる。

本稿では実計算機上で実アプリケーションを動作させ、下位キャッシュへのアクセスパターンの解析を行うことにより実環境における負の参照の時間的局所性の存在の調査および考察を行う。

### 2. 負の参照の時間的局所性

ネットワークストレージへのアクセスは、図 1 のようにサーバ計算機キャッシュとストレージ機器キャッシュを介して行われる。サーバ計算機のキャッシュ置換手法は多くの場合 LRU が用いられており、この場合最近参照されたデータはサーバ計算機のキャッシュに格納される。このため、最近アクセスされたデータへのアクセス要求はサーバ計算機上で処理され、ストレージ機器に要求が届くことはない。従ってネットワークストレージでは“最近アクセスされたデータは近い将来再度アクセスされる可能性が低い”という負の参照の時間的局所性が存在し、従来の参照の時間的局所性を期待している LRU キャッシュ置換手法は効果的に機能しない。

図 2 はシミュレーションにより得たサーバ計算機キャッシュとストレージ機器キャッシュにおける再アクセス時間と再アクセス確率の関係である[1]。再アクセス時間が 0[アクセス]~約 1000[アクセス]のときはネットワークストレージにアクセス要求が来ず、ネットワークストレージキャッシュにおける再アクセス確率が 0[%]になっている。これより、ファイルシステムなどを用いないシミュレーション環境においては負の参照の局所性が存在することがわかる。

### 3. 下位キャッシュにおける参照の時間的局所性の解析方法

実計算機上で FFSB, pgbench, ブロックデバイスへのランダムアクセスを実行し、下位キャッシュへのアクセス要求

<sup>†</sup>工学院大学

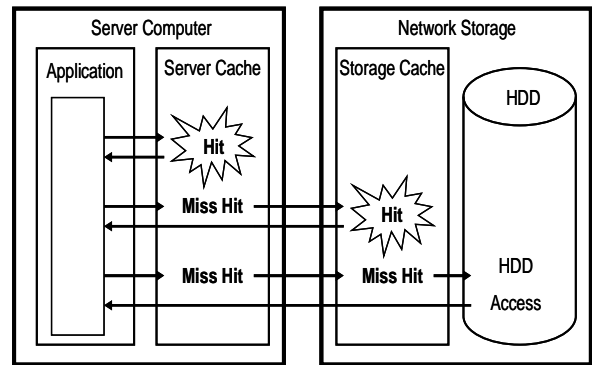


図 1 .二重キャッシュ構造

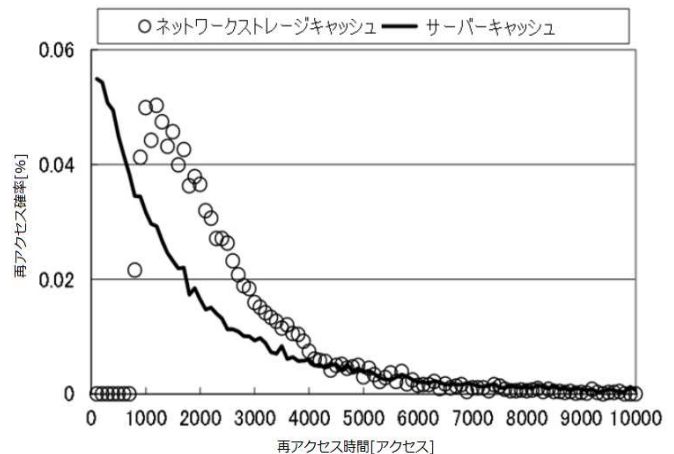


図 2 .負の参照の時間的局所性

をモニタリングすることにより、負の時間的局所性の調査を行った。OS には Linux2.6.18.8(CentOS5.5)を使用し、測定計算機は 2.0[GB]のメモリを搭載しており、そのうち 1.77[GB]がキャッシュに使用されている。モニタリングは OS の SCSI 層にて行い、ファイルシステムは ext2,ext3,NTFS を用いた。SCSI 層は LRU で管理される OS のキャッシュの下位に位置している。

### 4. 解析結果

再アクセス間隔とアクセス発生確率(確率密度関数)の関係を図 3~11 に示す。図 3~8 は FFSB を用いた解析であり、図 3 は総データサイズ 16[GB]、図 4 は総データサイズ[64GB]となっており、図 5~8 は総データサイズ[4GB]でオペレーションサイズが順に 4[KB],16[KB],64[KB],[1MB]である。図 9~11 は pgbench による解析結果であり、総データ量は順に [4GB],[16GB],[64GB]である。図 12 はファイルシステムを用いずに 4[GB]のブロックデバイス領域に 4[MB]のランダムアクセスを繰り返した時の解析結果である。

図 3 の解析結果より、ext2 では 8500[アクセス]以下の再アクセス確率が、NTFS では 7000[アクセス]以下の再ア

セス確率が極めて低く負の参照の時間的局所性の存在が確認された。同様に図4の ext2 にて 9100[アクセス]以下, NTFS にて 6000[アクセス]以下, 図5の ext2 にて 8000[アクセス]以下, NTFS にて 274000[アクセス]以下, 図6の ext2 にて 130000[アクセス]以下, NTFS にて 295000[アクセス]以下, 図7の ext2 にて 27000[アクセス]以下, ext3 にて 27000[アクセス]以下, NTFS にて 35000[アクセス]以下, 図8 ext2, ext3 にて 9000[アクセス]以下, NTFS にて 25000[アクセス]以下, 図9の ext2 にて 4000[アクセス]以下, NTFS にて 4000[アクセス]以下, 図10の ext2 にて 21000[アクセス]以下, NTFS にて 3000[アクセス]以下, 図11の ext2 にて 2000[アクセス]以下, NTFS にて 4000[アクセス]以下, 図12の 15000[アクセス]以下にて再アクセス確率が極めて低いことが確認された。以上より, ファイルシステムに ext2, NTFS を用いた場合およびファイルシステムを用いない場合は負の参照の時間的局所性が存在し, ext3 を用いた場合は必ずしも存在しないことが確認された。

## 5. おわりに

本稿では, 多段キャッシュ環境において下位キャッシュへのアクセスパターンの解析を行い, 負の参照の時間的局所性の存在の調査を行った。

今後は負の参照の時間的局所性を考慮したキャッシュ置換アルゴリズムの提案を行っていく予定である。

## 参考文献

- [1] Yusuke Nagasako and Saneyasu Yamaguchi, "A Server Cache Size Aware Cache Replacement Algorithm for Block Level Network Storage," In Proceedings of The 4th Ad Hoc, Sensor and P2P Networks Workshop (AHSP) 2011
- [2] Yuanyuan Zhou and James F. Philbin, Kai Li, "The Multi-Queue Replacement Algorithm for Second Level Buffer Caches," In Proceedings of the 2001 USENIX Annual Technical Conference, 2001.

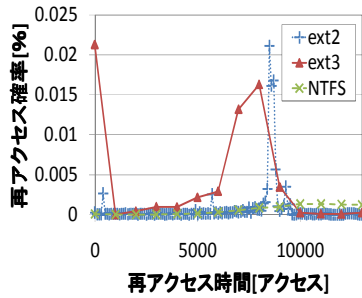


図3. FFSB 測定結果 (データサイズ 16GB)

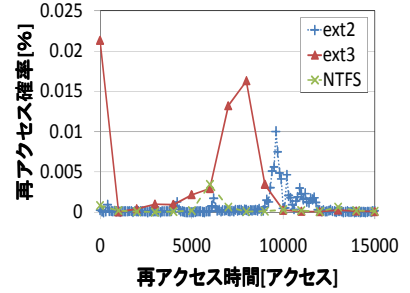


図4. FFSB 測定結果 (データサイズ 64GB)

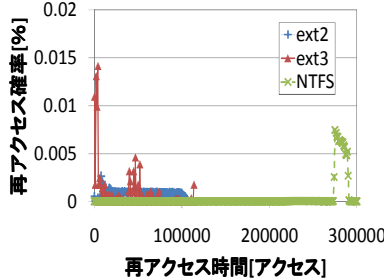


図5. FFSB 測定結果 (オペレーションサイズ 4KB)

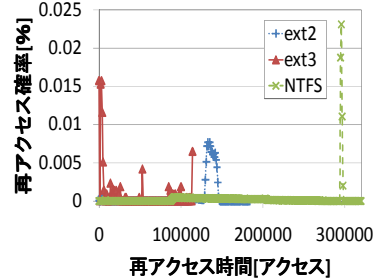


図6. FFSB 測定結果 (オペレーションサイズ 16KB)

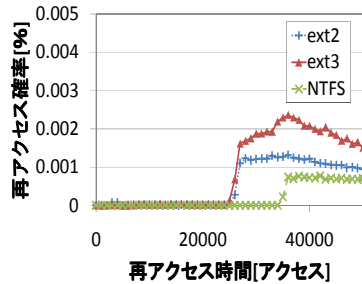


図7. FFSB 測定結果 (オペレーションサイズ 64KB)

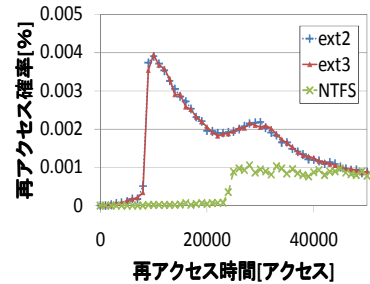


図8. FFSB 測定結果 (オペレーションサイズ 1MB)

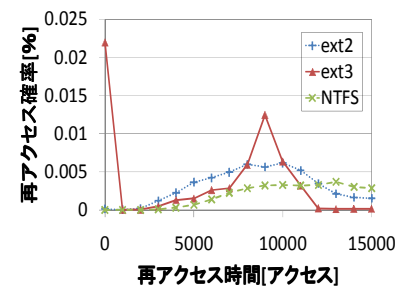


図9. pgbench 測定結果 (データサイズ 4GB)

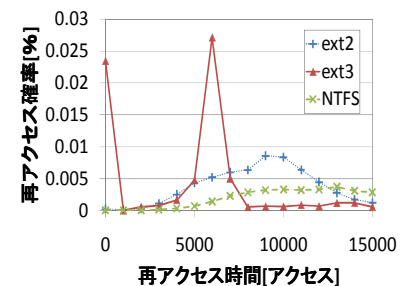


図10. pgbench 測定結果 (データサイズ 16GB)

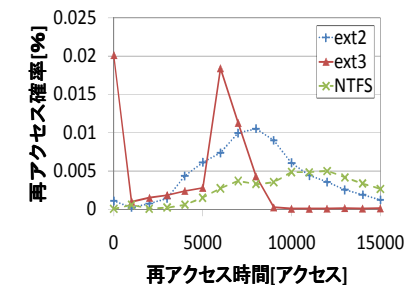


図11. pgbench 測定結果 (データサイズ 64GB)

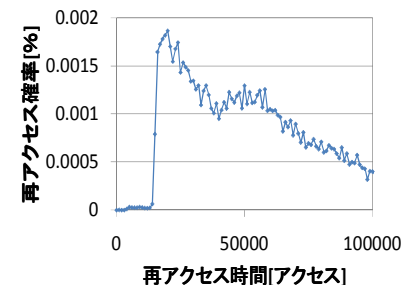


図12. ブロックデバイスへのランダムアクセス測定結果