

空間的局所性と時間的局所性を再現する ストレージワークロードの自動生成の検討

丸山 哲太郎 *

河場 基行 †

概要

近年のストレージにおける性能向上に寄与する機能は負荷の偏りを利用するものが多いため、これら機能の評価のためには実システムが生じる負荷の時間的・空間的な偏りを再現することが必要である。しかし既存的方式ではこの偏りを忠実に再現することは出来ない。本論文は長期的なアクセス頻度の偏りを再現する Zipf 分布による負荷生成法と、短期的なアクセス頻度の偏りを再現する Hierarchical Reuse Random Walk による負荷生成法によって生成された負荷の特性を解析し、実システムが生じる負荷の偏りを模擬する人工的なワークロード生成方式を考察する。

1 はじめに

近年のストレージが持つ機能の中には、キャッシュ機能や階層ストレージ機能などアクセスの偏り（局所性）を利用して性能を向上させるものが多い。これら機能の評価するためには、局所性を持つ負荷（ワークロード）を実際にストレージに与え、性能を測定することが必要である。この性能測定手法として実ストレージアクセス履歴を用いる手法があるが、アクセスサイズを変えられないなど適用面で問題がある。そこで局所性を持つ負荷を自動生成する方式が望まれている。

本論文ではまず、実ワークロードが長期的なアクセス頻度の偏り（空間的局所性）と短期的なアクセス頻度の偏り（時間的局所性）を併せ持つことを示す。次に、既存方式はこれら局所性のうち一方のみしか持たないことを示す。最後に、実ワークロードにおける局所性を再現するために必要な空間的・時間的局所性を併せ持つ新しい方式への考察を示す。

2 実ワークロードの局所性

ワークロードの長期的なアクセス頻度の偏りである空間的局所性は、ストレージの全領域を単位領域に分割し、全測定時間における単位領域毎のアクセス頻度を頻度順にソートした際の分布の形で判断する。特定の単位領域へのアクセス確率が大きいほど空間局所性が高い。

また短期的なアクセス頻度の偏りである時間的局所性は、ある単位領域にアクセスしてから再びアクセスするまでの回数である再参照間隔の分布の形で判断する。再

参照間隔が短い確率が大きいほど時間的局所性が高い。

空間的・時間的局所性はそれぞれ分布の形から判断できるが、これらを端的に表す値を定義することは非常に難しいため、本論文では便宜的に空間的局所性は最も頻度が高い単位領域の確率、時間的局所性は再参照間隔が 0 である確率を各々の指標とする。

本論文では実ワークロードとして SNIA (Storage Network Industry Association) にて公開されている Microsoft Research 社の実ワークロード (MSR) [1] を用いる。MSR に含まれるワークロードのうちアクセス数が多い上位 10 個のワークロードを用い、各ワークロードがアクセスする領域を 1024 個に分割して空間的・時間的局所性を計算したものを図 1 に示す。

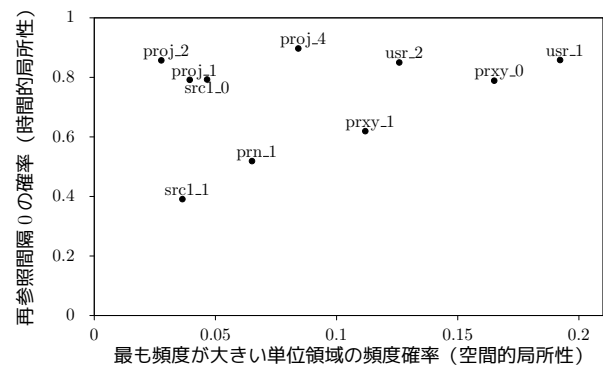


図 1 実ワークロードにおける局所性

図 1 から、MSR におけるどのワークロードも空間的局所性と時間的局所性を併せ持つことが分かる。

3 局所性を再現する既存方式

様々な実ワークロードの空間的局所性は Zipf 分布に従うことが知られている [2]。N 個の単位領域に対する

* (株)富士通研究所 データプラットフォーム研究部
tetsu.maruyama@jp.fujitsu.com

† (株)富士通研究所 データプラットフォーム研究部
kawaba@jp.fujitsu.com

Zipf 分布は式 1 で表され, s の値が大きいほど空間的局所性が高いワークロードとなる.

$$f(x; s, N) = \frac{1/x^s}{\sum_{n=1}^N 1/n^s} \quad (1)$$

また時間的局所性を binary tree 上の移動で模擬するものが, 業界標準のベンチマークである SPC-1[3] にも採用されている Hierarchical Reuse Random Walk[4] (HRRW) である. 2^H 個の単位領域を高さ H の binary tree の leaf に配置し, i 番目にアクセスした単位領域を表す leaf を x_i とする. HRRW は binary tree を登る一定値 k と更に 1 段登る確率 ν の 2 つのパラメータを持ち, これらから binary tree を登る段数 $h (0 \leq h \leq H)$ を求め, 式 2 に基づき x_i から x_{i+1} を決定するものである. R は $0 \leq R < 1$ を満たす一様乱数であり, h の値が小さいほど時間的局所性が高いワークロードとなる.

$$x_{i+1} = 2^h \lfloor \frac{x_i}{2^h} \rfloor + \lfloor R2^h \rfloor \quad (2)$$

2 章において局所性が最も高い usr_1 の空間的局所性を Zipf 分布, 時間的局所性を HRRW によって再現することを試みた. Zipf 分布のパラメータを $s = 1.14$, HRRW のパラメータを $k = 0, \nu = 0.3$ とした時の空間的・時間的局所性のグラフを図 2, 図 3 に示す.

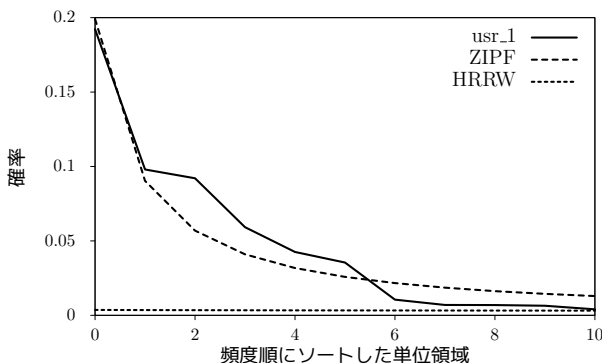


図 2 usr_1 の空間的局所性を再現

これらのグラフから, 空間的・時間的局所性は各々 Zipf 分布・HRRW にて再現できることが分かる. しかしどちらも空間的・時間的局所性を併せ持たず, 実ワークロードの局所性を再現することは出来ない.

4 局所性を再現するモデルへの考察

Zipf 分布により生成されたワークロードは図 2 のように空間的局所性を持つが, 式 1 に過去の位置情報が無

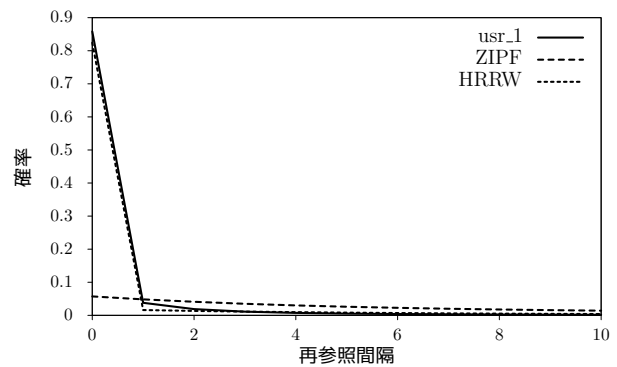


図 3 usr_1 の時間的局所性を再現

いことから各々のアクセスは独立であるため, 図 3 のように時間的局所性を持たない.

一方 HRRW は式 2 の第一項から, h の値が小さい場合に直前の位置情報が保存されるため, 図 3 のように時間的局所性を持つ. しかし式 2 の第二項により, 一様乱数によってアクセスが全領域に均等に分散するため, 図 2 のように空間的局所性が失われてしまう.

実ワークロードは図 1 のように一般的に空間的・時間的局所性を併せ持つ. これを忠実に再現するためには, 式 2 の第二項を一様乱数ではなく, binary tree の形に依存する乱数に置き換える新しいモデルが必要である.

5 まとめ

実ワークロードが空間的・時間的局所性を併せ持ち, 既存方式ではこれら局所性を同時に再現できないことを示した. また HRRW をベースにした空間的・時間的局所性を併せ持つ新しいモデルの可能性を示唆した.

今後は, 新しいモデルを定式化し, 実際のストレージ装置上において実ワークロード及び新しいモデルによって生成された実ワークロードを模擬する負荷を用いて性能実測を行い, 人工生成の質を検証する予定である.

参考文献

- [1] <http://iotta.snia.org/traces/388>
- [2] H. Yu, et. al.: "Understanding User Behavior in Large-Scale Video-on-Demand Systems". *Proc. of the 1st ACM SIGOPS*, 2006
- [3] <http://www.storageperformance.org>
- [4] Bruce McNutt: "The Fractal Structure of Data Reference - applications to the memory hierarchy", *Springer Science+Business Media, LLC*