

Title	冗長ディスクアレイ用耐故障分散ディスクキャッシュに関する研究
Author(s)	小島, 信
Citation	
Issue Date	1997-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1031
Rights	
Description	Supervisor:横田 治夫, 情報科学研究科, 修士

冗長ディスクアレイ用

耐故障性分散ディスクキャッシュに関する研究

北陸先端科学技術大学院大学

情報科学研究科 情報システム学専攻

小島信

keywords RAID、信頼性、ディスクキャッシュ
ヒット率、ディスクアクセス性能

Abstract

1 はじめに

近年、計算機システムの規模は拡大の一途をたどり、それと共に情報機械の処理性能に対してさらなる高速化と信頼性への要望が高まっている。CPU や半導体メモリと言ったマイクロデバイス分野では、アーキテクチャやプロセス技術の進歩によって、さながら日々これ躍進と言った所である。しかし、総合的に見たコンピュータの性能の向上には、I/O サブシステム性能の向上もまた欠く事は出来ないと言えよう。

RAID は、そのような二次記憶装置に対する要求を満足させる為に提案されたディスクアレイシステムである。RAID はデータにパリティ演算を行なう事によって、システムにアベイラビリティを与え、ディスク故障に対するデータの信頼性を高めている。しかし、RAID5 アレイは、データ更新の際に古いデータとパリティの読み込みと新しいデータとパリティの書き込み、合計 4 つのディスクアクセスを行なわなければならないと言う問題を持つ [1][2]。

この更新性能を改善する一般的な方法は、RAID に大容量ディスクキャッシュを用いる事である。しかし、これらはキャッシュ障害について特別な事は行なっていない。

耐故障性を持ったディスクキャッシュの 1 つのアイディアは、キャッシュを 2 重化して同じデータを格納するようにし、そしてそれぞれに独立した電源を持たせる事である [1][3]。

ここで我々は、これとは異なる耐故障性を持ったディスクキャッシュとして、分散ディスクキャッシュ付き RAID という方法を提案する。RAID 用分散ディスクキャッシュとは、1

つのキャッシュに故障が生じた場合でも、他のキャッシュがディスクに維持された冗長データをを用いる事で、自動的にデータを回復する事が可能なキャッシュの集まりである。

2 分散ディスクキャッシュの概要

分散ディスクキャッシュは使用する RAID のディスクドライブと同じ台数のキャッシュを用いる。ホストから入力された更新データに対して、このシステムではデータをキャッシュに格納する前にまずパリティ情報のための演算を行なう必要がある。従って、キャッシュ上にあるデータは既にパリティによる冗長性をもっていると言える。この時、キャッシュ上のデータには耐故障性がある。もし、この時点でキャッシュ障害が発生したとしても、残りの最新のデータの排他的論理和演算を行なう事で、その失われたデータを復元する事が出来る。

更新データと同様に、演算によって求められたパリティ情報もキャッシュに書き込まれる。演算に必要な旧情報は、それぞれのディスクあるいはキャッシュから読み込まれる。両データが共にヒットした場合は、データを素早く更新する事が出来る。一方だけヒットした場合は、ヒットした方が素早くデータをアレイ制御装置に送るので、アレイ制御装置はディスクアクセスの待ち時間なしに、XOR 処理を進める事が出来る。

各キャッシュは回線で互いに結ばれており、それによって互いのキャッシュデータの情報を交換する事が出来る。この情報は、キャッシュ故障の際に失われたキャッシュデータを素早く再構成するのに役に立つ。

3 実験システムを用いた性能実験

本研究では、実験モデルとして複数のトランスピュータで構築した、RAID5 タイプの分散キャッシュモデル、2 重キャッシュモデル、キャッシュ未搭載モデルを作成。ディスク故障とキャッシュ故障が存在しない応答時間やヒット率やキャッシュ方式の比較を行なった。評価実験には、容量 20MB のディスクドライブとキャッシュを 6 台と、キャッシュメモリの合計の容量は、最大 1 MB まで測定した。以下の項目に、行なった評価実験の結果を示す。なお、ここから以下、分散キャッシュのヒットと言った場合、それは更新データに対するデータとパリティ情報の同時ヒットを指すものとする。

● ヒット率の差

まず書き込みの場合。キャッシュにフルアソシアティブ方式を用いた場合は、RAID5 の 2 重キャッシュと分散キャッシュの間にヒット率の差がない事が判った。ヒット率は、キャッシュ容量にほぼ比例して増えていった。キャッシュにダイレクトマップ方式を用いた場合は、フルアソシアティブ方式に比べ、2 重キャッシュ、分散キャッシュ共に、ヒット率は半分から 4 分の 3 程度の大きさであった。

次に読み込みの場合。分散キャッシュと 2 重キャッシュはフルアソシアティブ方式では、読み込みのヒット率は書き込みの時と近い値であった。分散キャッシュのダイ

レクトマップ方式の場合は、ヒット率は、書き込み時のフルアソシアティブ方式の場合に近い値であった。2重キャッシュのダイレクトマップ方式では、ヒット率の伸び値とも書き込み時と似ていた。

- ヒット-ミスそれぞれの処理時間

まず、書き込みについて。分散キャッシュ方式は、どの場合でもパリティ計算を行なうため、処理時間はヒット時でもキャッシュ無しの RAID の 3 分の 1 程度にしかならなかった。2重キャッシュ方式は、ヒット時の処理時間は極めて低い。ミス時の性能は、論理演算を待たずにデステージ作業が行なえる分、2重キャッシュ方式に比べ、分散キャッシュ方式が有利である。

読み込みでは、論理演算を行なう必要がない。その為、ヒット時の処理時間は、分散キャッシュも 2重どちらも低い値に留まった。

- 運転性能の評価（5万回動作試験）

次にシステムを実際に動かし続けた時の、トータルなアクセス時間を測定してみた。この測定は読み書きランダムにそれぞれ約 2 万 5 千回づつ行ない、測定値として各合計のアクセス時間の集計をとった。

まず、書き込みについて説明する。分散キャッシュは、アクセス時間の総計は、常にキャッシュなしの RAID より低い値を記録した。そして、ダイレクトマップ方式を用いた時にキャッシュ容量の増加と共に、アクセス時間の減少が確認できた。

2重キャッシュの場合、ダイレクトマップ方式では徐々にアクセス時間は減少していったが、フルアソシアティブ方式では、逆に増加していた。この実験ではキャッシュをソフトウェアで動作させているせいであると考えられる。

ダイレクトマップ方式を用いた場合の、分散キャッシュ方式と 2重キャッシュの間のアクセス時間の減少率の差は小さかった。

次に読みの場合。どの方式もアクセス時間は書き込みの場合に比べ、少なかった。分散キャッシュの場合は、キャッシュ容量が増えるにつれ、フルアソシアティブ方式はアクセス時間は増え、ダイレクトマップ方式では減った。

4 結論

本研究で、RAID 上での耐故障性を持った分散ディスクキャッシュシステムの仕組みを考え無故障時のヒット率と速度性能を測定した。

その結果、キャッシュ未搭載の RAID モデルより、低い処理時間を実現できる事が判った。2重キャッシュ方式と比較した場合、ヒット率が低い場合では、分散キャッシュ方式の方が処理が速い事が判った。これは、分散キャッシュ方式が、キャッシュのデータのディスクへの書き込みをシステムの XOR 操作の終了を待たずに行なえるためである。

また、キャッシュ方法としては、アソシアティブ方式が、分散キャッシュに高いヒット率を与える事が判明した。今回の実験では、総キャッシュ容量が増えてもアクセス時間の短縮はみられなかった。しかし、より効率的なキャッシュ探索方法を用いればさらなる高性能化が期待できるものと思われる。

今後の課題としては、ディスク、キャッシュでの障害発生時の性能測定を行なう事が考えられる。

参考文献

- [1] Menon.J and Cortny.J, “The architecture of fault-tolerant cached RAID controller”, *IEEE.NEW York*,76-8, 1993.
- [2] Peter.M.Chen etal, “RAID:High-Performance,Reliable Secondary Storage”, *ACM computing Surveys Vol.26 No2* , June,1994.
- [3] JAI.MENON, “Performance of RAID5 Disk Arrays with Read and Write Caching”, *Distributed and Parallel Database*,2,261-293, 1994.