

高性能サーバ向けディスクアレイ「DF400」の開発（2） ～I/O性能の予測技術～

4H-5

八木沢 育哉, 松並 直人, 山本 政行, 大枝 高, 高本 賢一*
(株)日立製作所 システム開発研究所, ストレージシステム事業部*

1. はじめに

近年、高性能サーバ向けのストレージとして、複数ディスク並列動作と冗長データ付加により高性能・高信頼を実現するディスクアレイ(RAID)が採用される事例が定着しつつある[1]。高性能サーバ向けディスクアレイに対するサーバベンダからの要求性能は高く、特にデータベースアクセスで発生するトランザクション処理のI/O性能の向上が求められている。

ディスクアレイのトランザクション処理のI/O性能は、コントローラ性能によって決定されるが、コントローラのボトルネックはプロセッサである。したがって、I/O性能の向上には、プロセッサ処理時間の短縮が必要となる。

本稿では、I/O性能予測のベースとなるプロセッサ性能の予測技術について報告する。

2. プロセッサ性能予測の課題

(1)プロセッサ処理時間の内訳 (図1)

(イ)インターフェース制御部

インターフェースコントローラのハードウェア処理に関わるハードウェア依存部①、および、プロセッサ性能に依存するプロセッサ演算部②から成る。

(ロ)非インターフェース制御部

インターフェース制御部以外の部分であり、前述同様のプロセッサ演算部②のみから成る。

(2)プロセッサ処理時間の短縮施策

DF400では従来機に対し、施策(a)でプロセッサ演算部②を、施策(b)でインターフェース制御部を短縮した。本稿では、施策(a)の効果を定量的に評価するためのプロセッサ性能予測技術を開発した。

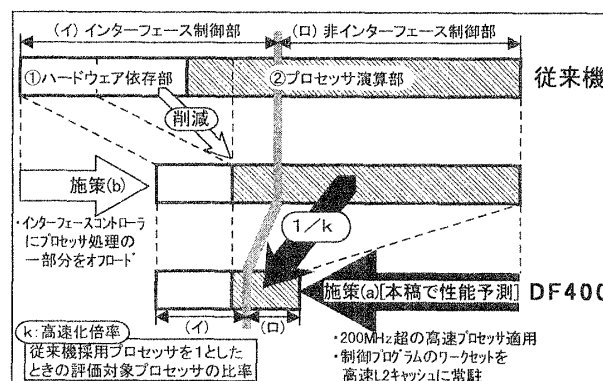


図1 プロセッサ処理時間

(3)目標プロセッサ性能

事前検討により、目標I/O性能は従来機比3倍とした。目標I/O性能達成には、施策(a)でプロセッサ演算部②を1/4にする必要があり、従来機比の高速化倍率 $k \geq 4$ が目標プロセッサ性能となる。

(4)課題

高速化倍率kを予測するには、実測にて従来機採用プロセッサと比較するのが確実だが、現存しない次世代プロセッサ等は評価できない。

3. プロセッサ処理モデル方式

上記の課題解決のため、プロセッサ命令フェッチ動作に着目したプロセッサ処理モデルを考案した。

(1)ディスクアレイ制御プログラムの特徴

- (a)ワークセットは1次(L1)キャッシュより十分大きい。
- (b)シリアルに実行され、局所参照性がない。
- (c)条件分岐が多い。

(2)モデル化の方針

特徴(a)(b)により、命令フェッチ時には1次キャッシュにミスヒットし、プロセッサ内部動作に比べて低速な外部メモリアccessが発生する。モデル化にあたり、命令

Development of Disk Array "DF400" for High Performance Servers(2)
～Technique of Estimate for I/O Performance～
Ikuya Yagisawa, Naoto Matsunami, Masayuki Yamamoto, Takashi Oeda, and Kenichi Takamoto*
Systems Development Laboratory, Data Storage & Retrieval Systems Division*, Hitachi, Ltd.

RAID: Redundant Arrays of Inexpensive Disks
I/O: Input and Output

実行動作において支配的となる外部メモリアクセスの影響が反映されるようにした。

また、特徴(c)のため、バースト転送でフェッチした命令のいくつかは実行されず、パイプライン処理効果が薄くなる。本稿では、命令の1バースト目がフェッチされた段階までを外部メモリアクセス動作とした。

(3)モデル

命令実行動作のモデル例を図2に示す。命令実行は、プロセッサ内部動作と外部メモリアクセスから成る。プロセッサ内部動作は、(i)~(iv)の各ステージであり、命令フェッチの前段階と命令フェッチ後の演算・ストアから構成する。外部メモリアクセス動作は、(I)外部メモリのアドレスを指定してから、(II)1バースト目の命令をフェッチするまでとした。

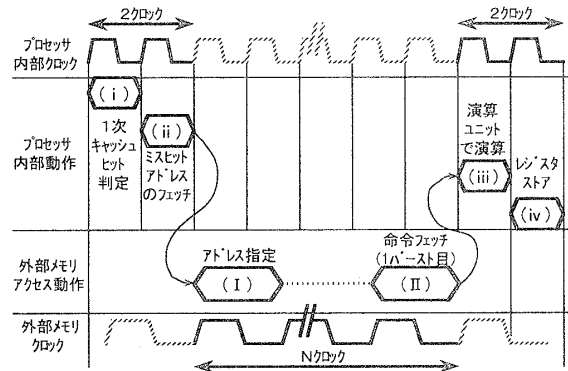


図2 プロセッサ処理モデル

本モデルにより命令実行時間の定義は式(1)とし、命令実行時間の相対比を高速化倍率kとした。

[命令実行時間]

$$=[\text{プロセッサ内部動作時間}] + [\text{外部メモリアクセス時間}]$$

$$=[\text{プロセッサ内部動作サイクル数}] \times [1/\text{プロセッサ内部周波数}] + [\text{外部メモリアクセスサイクル数}N] \times [1/\text{外部メモリ周波数}] \dots \text{式(1)}$$

4. モデルの妥当性の検証

モデルの検証のため、現存するプロセッサについて実測値との比較評価を行った。実測用プログラムとしては、ディスクアレイ制御プログラムからプロセッサ演算部②を抽出した自製ベンチマークテストを採用した。

図3に、各種プロセッサと外部メモリを組合わせたプロセッサ構成方式A~C、Y、Zにおける予測値と実測値を

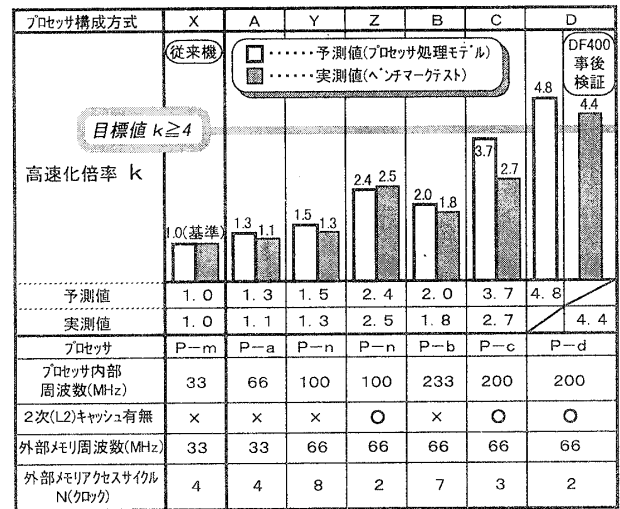


図3 プロセッサ性能予測

示す。構成方式Xは、基準となる従来機のものである。

各構成方式において、予測値は実測値とほぼ一致し、本モデルの妥当性を検証できた。

また、DF400で採用した構成方式Dはプロセッサ選定段階において現存しなかったが、実機での事後検証にて、誤差率+9%の精度で予測できたことを確認した。

5. プロセッサ性能予測結果

モデルによるプロセッサ性能予測の結果、構成方式Dは目標の高速化倍率 $k \geq 4$ を満足しており、DF400において採用した。実機での事後検証においても、高速化倍率 $k = 4.4$ の実測結果を得た。

6. まとめ

プロセッサ命令フェッチ動作に着目したプロセッサ処理モデルにより、ディスクアレイのプロセッサ性能予測技術を開発した。また、制御プログラムベースのベンチマークテストによる実測でモデルの妥当性を確認した。

参考文献

[1]David A.Patterson, et al.: "A Case for Redundant Arrays of Inexpensive Disks (RAID)", Report no.UCB / CSD 87 / 391, Computer Science Division Department of Electrical Engineering and Computer Science, University of California, Berkeley, 1987.