

# キャッシュコヒーレンスプロトコルに対する一考察

5 T-3

山本 伸介\*, 中嶋 卓雄\*, 中村 良三\*

\*熊本大学 工学部

## 1 はじめに

コンピュータの小型化, 高性能化にともない, 複数の異機種計算機をネットワーク上で有機的に結合し, 計算機資源を有効に活用し, より複雑な問題を解決する分散処理技術が活発に研究されている.

分散処理技術の中で, 分散共有メモリ (DSM) やキャッシュプロトコルに関する研究が活発に行われている. 特にキャッシュプロトコルの中で, スヌープキャッシュは, キャッシュ-主記憶間がバス結合型のアーキテクチャ上で共有するバスを監視することによってキャッシュの内容の一致を保証する代表的なキャッシュである. キャッシュの内容の一貫性を保持するプロトコルはキャッシュコヒーレンスプロトコルと呼ばれ, (1) ライトスルー, (2) ライトバックと (a) データ無効化型, (b) データ更新型とによって分類される. さらに複製されたデータブロックにオーナーシップの概念を導入し, 高速化が図られている.

本稿では, データのキャッシュへの読み込みに対して高速な, ライトバック, データ無効化型プロトコルである Berkeley モデルについて改良する.

## 2 Berkeley モデル

### 2.1 オーナーシップ

Berkeley モデルにおけるオーナーは, そのブロックに対する様々な要求・一貫性に最終的な責任を持つ. あるブロックのオーナーになったキャッシュ/メモリは, そのブロックへの他からの要求に対しブロックを供給し, また主記憶への書き戻しについても責任を持つ.

### 2.2 状態

Berkeley モデルでは, 各ブロックが「オーナーシップ」を持ち, キャッシュの状態を次の4つに分類する.

- I (Invalid). 目的のデータブロックを持たない.
- EO (Exclusive-Owner). 他にデータブロックのコピーがなく, そのオーナーである.
- SO (Shared-Owner). データブロックのコピーを持ち, そのオーナーである.
- SN (Shared-Nonowner). データブロックのコピーを持ち, そのオーナーでない.

### 2.3 状態遷移

Berkeley モデルにおける状態遷移を図1に示す.

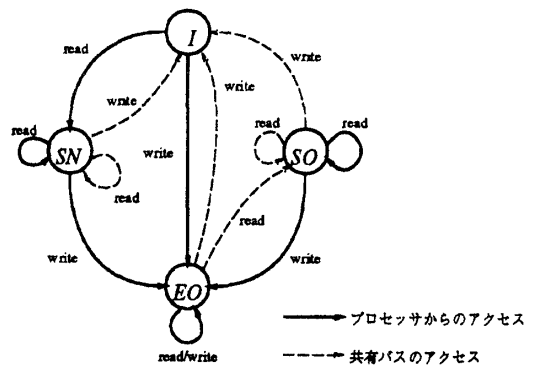


図1: Berkeley モデルの状態遷移

### 2.4 問題点

一般的に, キャッシュ-主記憶間の転送は, キャッシュ-キャッシュ間の転送より時間がかかる. Berkeley モデルでは, オーナーシップを用いることで, 他のプロトコルより主記憶への書き戻しの頻度を低くし, プロトコル全体の効率向上を図っている.

Discussion about Cache Coherence Protocol  
Shinsuke Yamamoto\*, Takuo Nakashima\*, Ryozyo Nakamura\*

\*Faculty of Engineering, Kumamoto University,

しかし、オーナーを持つキャッシュにおいて、ブロックの追い出しなどによりオーナーが主記憶に移動した場合、そのブロックに対する読み込み要求があるたびにキャッシュ-主記憶間の転送が繰り返し発生することとなる。

### 3 提案するモデル

本稿では、Berkeley モデルの次の状態、

- SO と複数の SN からなる状態での、SO におけるキャッシュの追い出し
- 複数の SN からなる状態での、I における読み込み

という主記憶へのアクセスに対する高速化を図るため、SN 状態を

- SNO (Shared-Nonowner-Other)  
データブロックのコピーを持ち、オーナーではないが、他のキャッシュにブロックのオーナーが存在する。
- SNN (Shared-Nonowner-No)  
データブロックのコピーを持ち、オーナーでなく、他のキャッシュにブロックのオーナーが存在しない。

の2つの状態に分類する。以下に分類した2つの状態の遷移を示す。

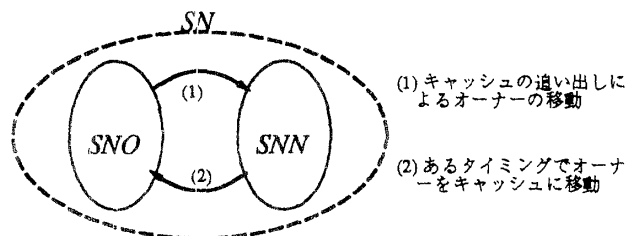


図2：SNO と SNN の状態遷移

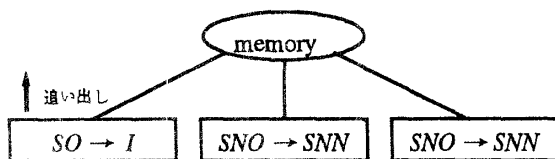


図3：キャッシュ追い出しによるオーナー移動

システム全体の流れとして、図2における状態遷移(1)を図3に、状態遷移(2)を図4に示す。

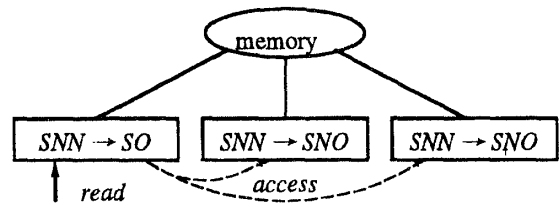


図4：主記憶からのオーナー移動

### 4 評価

一般的にキャッシュにおけるアクセス速度を向上させるためには、出来るだけ多くのコピーブロック(SN 状態)が存在すればよい。従来の Berkeley モデルにおいて、SN 状態が増加する場合は、

1. 主記憶から直接読み込む (SO 状態が存在しない場合)
2. SO 状態のキャッシュから読み込む

の2通りである。今回、提案したモデルでは、(1)の場合における SN 状態のキャッシュから読み込みが発生した時点で(2)の状態に移行でき、(1)の場合のキャッシュ-主記憶間の転送の頻度を低く抑えることでシステム全体の速度向上が期待できる。

### 5 おわりに

今回は、キャッシュ-主記憶間のデータ転送を、キャッシュの状態を増加させることによって減少させるモデルを提案した。このモデルでは、他のキャッシュのオーナーシップを意識した処理を実現している。

今後は、キャッシュからのデータブロックの追い出しアルゴリズムとオーナーシップの移動との対応関係について考察し、より高速なキャッシュコピーレンスプロトコルを検討していきたい。

### 参考文献

[1] 天野英晴, “並列コンピュータ (Parallel Computers)”, 情報系教科書シリーズ第18巻, 昭晃堂, 1996.  
[2] R.H.Katz, et.al, “Implementing a Cache Consistency Protocol”, ISCA85, pp.276-283, 1985.