**17.3- Sự kết hợp bộ nhớ cache và giao thức MESI**

- Ghi lại: Thao tác ghi thường **chỉ được thực hiện cho bộ đệm**. Bộ nhớ chính chỉ được cập nhật **khi line bộ đệm tương ứng bị xóa khỏi bộ đệm** 🡪 có thể dẫn đến **sự không nhất quán**

- Ghi qua: Tất cả các thao tác ghi **được thực hiện vào bộ nhớ chính cũng như vào bộ đệm**, đảm bảo rằng bộ nhớ chính luôn hợp lệ. Ngay cả với cách này, **sự không nhất quán có thể xảy ra** trừ khi các bộ nhớ cache khác **giám sát lưu lượng bộ nhớ** hoặc nhận được một số thông báo trực tiếp về bản cập nhật.

🡺 **Giao thức MESI (4 trạng thái của dữ liệu trong bộ đệm: sửa đổi / độc quyền / chia sẻ / không hợp lệ) được khuyến nghị ở đây.**

# Giải pháp phần mềm

* Giải pháp kết hợp bộ nhớ đệm phần mềm tránh sự cần thiết phải lắp thêm mạch phần cứng và logic bổ sung bằng cách dựa vào trình biên dịch và hệ điều hành để giải quyết vấn đề (không muốn thêm phần cứng).
* Hấp dẫn bởi vì việc phát hiện các **vấn đề tiềm ẩn** được chuyển từ **thời gian chạy** sang **thời gian biên dịch** và **độ phức tạp** của thiết kế được chuyển **từ phần cứng** sang **phần mềm**.
  + *Các cơ chế kết hợp dựa trên trình biên dịch này sẽ thực hiện phân tích dựa trên mã code để xác định các mục dữ liệu nào có thể* ***trở nên không an toàn*** *cho bộ đệm và chúng đánh dấu chúng. Sau đó Hệ điều hành hoặc phần cứng sẽ* ***ngăn*** *các mục không thể lưu trong bộ nhớ cache* ***được lưu*** *vào bộ nhớ cache.*
  + Tuy nhiên, các phương pháp tiếp cận **phần mềm về thời gian biên dịch** thường phải đưa ra các quyết định thận trọng, dẫn đến việc **sử dụng bộ đệm không hiệu quả**

# Giải pháp về mặt phần cứng

* Các giải pháp dựa trên phần cứng Thường được gọi là **giao thức kết hợp bộ đệm**.
* Những giải pháp này cung cấp **nhận dạng động** tại thời điểm chạy của **các điều kiện không nhất quán** *tiềm năng.*
* Bởi vì vấn đề chỉ được giải quyết khi nó thực sự phát sinh nên **việc sử dụng bộ nhớ cache hiệu quả hơn**, dẫn đến **hiệu suất được cải thiện** so với cách tiếp cận phần mềm.
* Các cách tiếp cận là trong suốt, **không thể nhận ra đối với người lập trình** và trình biên dịch, giảm gánh nặng phát triển phần mềm.
* Có thể chia thành hai loại:
  + Giao thức thư mục: dữ liệu được chia sẻ, được đặt trong một thư mục chung duy trì sự gắn kết giữa các bộ đệm
  + Giao thức snoopy (rình mò): mọi bộ điều khiển bộ đệm đều **giám sát các bus**, theo dõi các trạm phát sóng mà có thể khiến nó vô hiệu hóa line bộ đệm của nó

# Directory Protocols (Giao thức thư mục)

* *~~Có một bộ điều khiển tập trung, nó là một phần của bộ điều khiển bộ nhớ chính~~*
* *~~Trong một hệ thống dựa trên thư mục, dữ liệu được chia sẻ được đặt trong một thư mục chung duy trì sự gắn kết giữa các bộ đệm. Thư mục~~* ***~~hoạt động như một bộ lọc,~~*** *~~thông qua đó~~* ***~~bộ xử lý phải xin phép~~*** *~~tải một mục từ bộ nhớ chính vào bộ đệm. Khi một mục được thay đổi, thư mục sẽ cập nhật hoặc làm mất hiệu lực các bộ đệm khác chứa mục đó.~~*

- Các giao thức thư mục **thu thập và duy trì thông tin về các bản sao của các line cache**. *Thông thường, có một* ***bộ điều khiển tập trung*** *là một phần của bộ điều khiển bộ nhớ chính và một thư mục được lưu trữ trong bộ nhớ chính*.

*- Thư mục chứa thông tin trạng thái toàn cầu về nội dung của các bộ nhớ cache cục bộ khác nhau.*

- Khi một bộ điều khiển bộ đệm riêng lẻ thực hiện một yêu cầu, bộ điều khiển tập trung sẽ kiểm tra và đưa ra các lệnh cần thiết **để truyền dữ liệu giữa bộ nhớ và bộ đệm hoặc giữa các bộ đệm**. Nó cũng có trách nhiệm giữ thông tin cập nhật; do đó, mọi hành động cục bộ có thể ảnh hưởng đến trạng thái toàn cầu của một line phải được báo cáo cho bộ điều khiển trung tâm.

- Thông thường, **bộ điều khiển duy trì thông tin về bộ xử lý nào có bản sao của line nào**. Trước khi bộ xử lý có thể ghi vào một bản sao cục bộ của một line, nó phải yêu cầu quyền truy cập độc quyền vào line từ bộ điều khiển. Trước khi cấp quyền truy cập độc quyền này, bộ điều khiển sẽ gửi một thông báo tới tất cả các bộ xử lý với một bản sao được lưu trong bộ nhớ cache của line này, **buộc mỗi bộ xử lý phải vô hiệu hóa bản sao của nó.** Sau khi nhận được xác nhận từ mỗi bộ xử lý như vậy, bộ điều khiển cấp quyền truy cập độc quyền cho bộ xử lý yêu cầu.

- Khi **một bộ xử lý cố đọc một line được cấp riêng cho bộ xử lý khác**, nó sẽ gửi một thông báo **bỏ lỡ đến bộ điều khiển**. Bộ điều khiển sau đó đưa ra lệnh cho bộ xử lý giữ line đó, yêu cầu bộ xử lý thực hiện ghi lại (**write back**) vào bộ nhớ chính. Line này có thể được chia sẻ để đọc bởi bộ xử lý gốc và bộ xử lý yêu cầu.

Các lược đồ thư mục chịu các **nhược điểm của nút cổ chai** trung tâm *và chi phí liên lạc giữa các bộ điều khiển bộ đệm khác nhau và bộ điều khiển trung tâm*. Tuy nhiên, **chúng có hiệu quả trong các hệ thống quy mô lớn** liên quan đến nhiều bus hoặc một số sơ đồ kết nối phức tạp khác.