

## PROGETTARE MEMORIE PSEUDO-STATICHE PER LA SOSTITUZIONE DIRETTA DI MEMORIE

Jarrod Eliason  
Senior Design engineer  
Ramtron International

Le memorie pseudo-statiche sono progettate per essere inserite direttamente nello zoccolo di una RAM statica (SRAM) anche se il loro funzionamento interno non è di tipo statico. Il mercato offre due tipologie di memorie pseudo-statiche: PSRAM (pseudo-static RAM) e FRAM (ferroelectric random access memory). Le PSRAM si rivolgono alle applicazioni che impiegano SRAM lente e competono unicamente sulla base del fattore "prezzo per bit". Le FRAM sono orientate alle applicazioni che impiegano SRAM dotate di batteria di backup (battery-backed SRAM, o BBSRAM) e competono sulla base dei costi del sistema e della relativa logistica. Le FRAM si rivolgono anche alle applicazioni che richiedono l'acquisizione di dati da conservare in modo non-volatile, nelle quali offrono prestazioni superiori. Questo articolo descrive il modo in cui questi due tipi di memorie pseudo-statiche riescono con successo a "mascherarsi" da memorie statiche, facilitando così il lavoro dei progettisti di sistema.

Una caratteristica esclusiva delle SRAM asincrone è il rilevamento della transizione dell'indirizzo (address transition detection, ATD). I chip di memoria appartenenti a questa famiglia consentono il cambiamento continuo dello stato dei pin di indirizzo, con qualunque cadenza e per qualunque durata, garantendo l'emissione dei dati corretti entro un tempo  $t_{AA}$  (address access time) dal-

*L'articolo descrive le tecniche e i circuiti necessari per realizzare memorie pseudo-statiche che possano essere inserite direttamente al posto delle memorie statiche. Questa possibilità favorisce l'impiego delle PSRAM nei telefoni cellulari e l'adozione delle FRAM al posto delle BBSRAM*

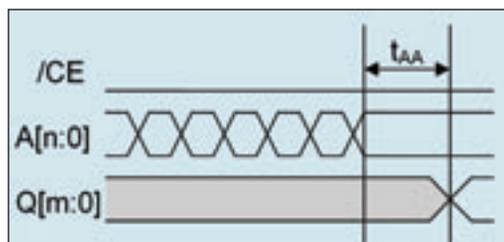


Fig. 1 - L'ATD nelle SRAM

l'istante in cui tali pin raggiungono uno stato stabile (Fig. 1).

Questa flessibilità consente ai progettisti di microcontrollori e microprocessori ("micro", nel resto dell'articolo) di affrontare con tranquillità le relazioni temporali tra i segnali delle interfacce di memoria. In un tipico micro con interfaccia di memoria diretta (cioè con indirizzi/dati non multiplexati) l'uscita chip-select e gli indirizzi sono comandati dallo stesso fronte d'onda del clock (Fig. 2).

Solitamente le relazioni temporali tra CLK e /CS ( $t_{CS}$ ) e tra CLK e A ( $t_{ADR}$ ) sono definite nelle specifiche del dispositivo, mentre la relazione temporale tra /CS e A non viene neppure citata nella documentazione del prodotto, né tanto

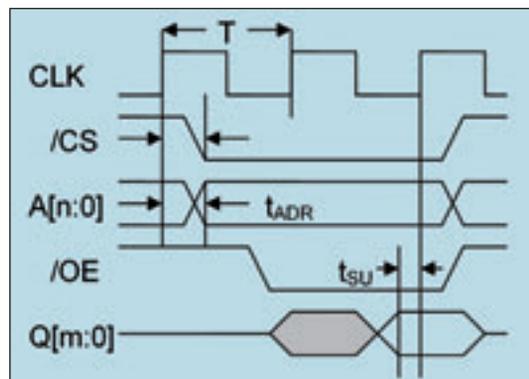


Fig. 2 - Interfaccia diretta tra micro e memoria

meno garantita. Per le SRAM ciò non costituisce un problema; l'unico requisito è che il tempo di accesso degli indirizzi sia sufficientemente breve. Il tempo ammesso dal micro è pari a due periodi di clock ( $T$ ) meno il ritardo di propagazione dell'indirizzo o del chip-select (il valore più alto tra  $t_{ADR}$  e  $t_{CS}$ ) e il tempo di setup dei dati del micro ( $t_{SU}$ ). Nella maggior parte dei casi  $t_{ADR}$  e  $t_{CS}$  sono uguali,

pertanto il risultato del calcolo è lo stesso in entrambi i modi.

Al fine di ridurre le dimensioni dei contenitori, quasi tutti i pin dei moderni micro svolgono più funzioni diverse. Un effetto collaterale di questo multiplexing funzionale è che ogni pin può avere un diverso ritardo interno. A meno di controllare singolarmente ogni ritardo, è probabile che almeno uno dei pin di indirizzo risulti leggermente più lento del chip-select. La questione è resa ancor più complicata da fattori esterni. Ad esempio, una tecnica di progettazione molto comune comporta la condivisione del bus degli indirizzi tra le memorie e/o le periferiche, mentre solitamente i chip-select non sono condivisi tra più dispositivi esterni. Di conseguenza, sebbene sia garantita la precedenza degli indirizzi rispetto al segnale di chip-select emesso dal micro, sui pin della memoria questa relazione temporale può risultare alterata. Una SRAM asincrona vede ogni skew sull'indirizzo come una riduzione del tempo di accesso richiesto, il che non compromette in alcun modo la funzionalità. Per le memorie pseudo-statiche, invece, la relazione temporale tra chip-select e il più lento degli indirizzi è un parametro critico, come verrà spiegato di seguito. Le PSRAM funzionano internamente come RAM dinamiche (DRAM), nel senso che la memoria deve essere continuamente rinfrescata per conservarne lo stato. Inoltre il processo di lettura di una cella PSRAM comporta la degradazione del relativo segnale, pertanto a volte è necessario rilevare lo stato del dato e riportare il segnale al suo valore massimo. Con riferimento allo schema semplificato di figura 3, una cella DRAM conserva il dato intrappolando carica nel nodo di storage (storage node, SN) del condensatore della cella ( $C_{CELL}$ ).

L'accesso alla cella è controllato da un transistor che svolge la funzione di interruttore. Più celle diverse condividono una stessa bit line e la capacità parassita del drain di ogni transistor, sommata alla capacità parassita del metallo, crea una capacità complessiva

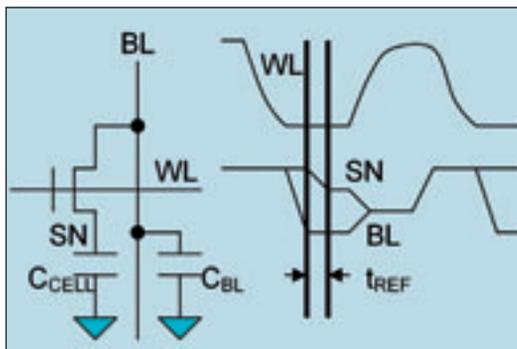


Fig. 3 - Scrittura/lettura del bit 1 in una DRAM

indicata in figura 3 come  $C_{BL}$ . La carica immagazzinata nella cella esce lentamente attraverso il transistor di accesso e il condensatore. La tensione al nodo di storage deve essere rinfrescata prima che il segnale scenda al di sotto del minimo livello rilevabile. Nella figura, questo tempo di refresh è indicato con  $t_{REF}$ . Quando la cella viene letta o rinfrescata, la linea di parola (word line, WL) passa a livello alto e la carica che rimane nel nodo di storage viene condivisa tra  $C_{CELL}$  e  $C_{BL}$ . A questo punto viene rilevata la tensione sulla linea di bit per determinare se la cella conteneva un valore "1" oppure "0" e nel nodo di storage viene ripristinato il valore massimo di tensione associato a quel particolare livello logico.

Dopo la condivisione della carica e prima del ripristino, la tensione rimanente nel nodo di storage può scendere al di sotto del minimo livello rilevabile. Se l'operazione fosse interrotta durante questa finestra temporale critica, il suc-

cessivo accesso allo stesso bit potrebbe fornire un valore errato. Per questo motivo l'azione della lettura di una DRAM viene considerata distruttiva. Una volta che l'accesso alla cella ha avuto inizio, deve essere completato. A differenza della SRAM, la DRAM non può sopportare il continuo cambiamento degli indirizzi illustrato in fig. 1. Tuttavia una DRAM può essere progettata per gestire l'incertezza finita delle temporizzazioni che è propria della maggior parte dei sistemi basati su micro. La FRAM è intrinsecamente non volatile e quindi non richiede un refresh continuo, ma condivide con la DRAM l'aspetto della lettura distruttiva. Si possono quindi utilizzare le stesse tecniche per creare FRAM pseudo-statiche di facile uso.

La figura 4 illustra un altro modo tipico in cui i micro utilizzano la funzionalità ATD delle SRAM. In molti micro le linee degli indirizzi e dei dati sono multiplexate, al fine di ridurre il numero di pin necessari per l'interfacciamento con la memoria esterna. Tra il micro e la memoria deve essere inserito un latch esterno, che viene controllato tramite un segnale di abilitazione (address latch-enable, ALE) fornito dal micro. Quando il segnale ALE scende a livello basso, i dati all'ingresso del latch (A/D[n:0]) saranno conservati e mantenuti agli ingressi della memoria (AL[n:0]) per l'intera durata dell'operazione di accesso. Una volta che l'indirizzo è stato fissato nel latch, il micro lascia il bus A/D e - nel caso di un'operazione di lettura -

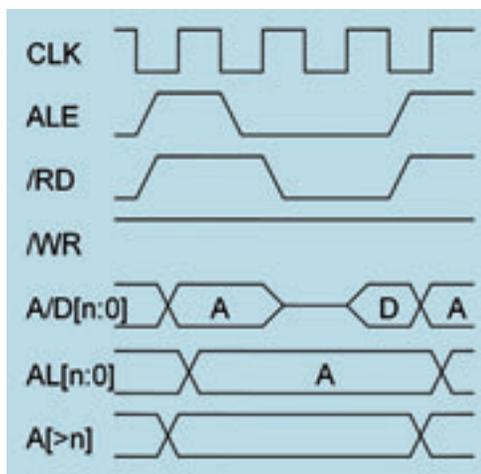


Fig. 4 - Interfaccia multiplexata

comanda il segnale /RD a livello basso, come illustrato in figura 4. Nel caso di un'operazione di scrittura, il micro invia i dati sul bus A/D e comanda il segnale /WR a livello basso. Non tutti gli indirizzi in figura 4 sono trattati dal latch; il ritardo di propagazione di quest'ultimo, pertanto, aggiunge ulteriore skew agli indirizzi stessi.

Molti dei micro che impiegano l'approccio basato sul latch degli indirizzi non forniscono un segnale di chip-enable dedicato. In questo caso la selezione tra diversi chip di memoria o altre periferiche può essere effettuata sfruttando la parte superiore dell'indirizzo, tramite alcuni circuiti logici. Il pin /CE della memoria può essere mantenuto a livello basso per tutto la durata dell'accesso; la fine di un'operazione di accesso e l'inizio della successiva, pertanto, sono indicate unicamente dai cambiamenti degli indirizzi. Affinché una memoria pseudo-statica possa essere inserita direttamente al posto di una SRAM è necessario considerare le seguenti situazioni riguardanti lo skew degli indirizzi:

- 1) Skew tra l'indirizzo e il segnale di chip-select (tempo di setup degli indirizzi negativo).
- 2) Skew tra indirizzi diversi negli accessi ATD.

Il primo caso può essere gestito facilmente al prezzo di un aumento del tempo di accesso e del tempo di ciclo. La soluzione per il secondo caso non ha alcun effetto negativo, nell'ambito di un'implementazione pseudo-statica della funzionalità ATD.

La figura 5 mostra un semplice circuito asincrono di ritardo che può essere usato per ritardare internamente il fronte di discesa del segnale chip-enable, al fine di permettere agli indirizzi di continuare a cambiare - per un certo tempo - anche dopo la discesa del segnale di chip-enable esterno.

Solo in rari casi la documentazione del micro indica l'ammontare del tempo di setup negativo necessario, che pertanto deve essere determinato per ogni singolo sistema. Tuttavia per la maggior parte dei progetti questo tempo è infe-

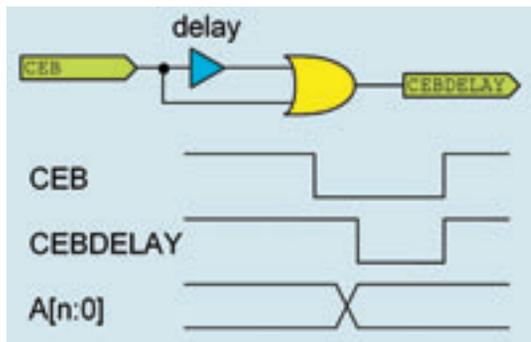


Fig. 5 – Ritardo sul fronte di discesa

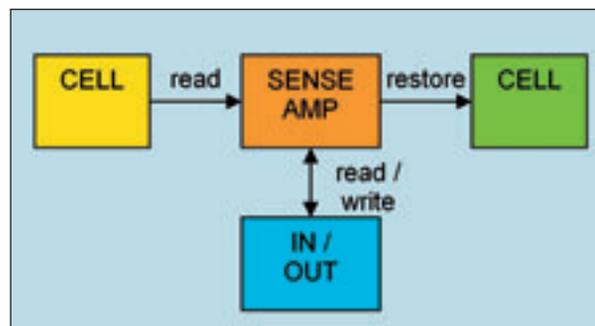
riore a 5 ns. Lo svantaggio di questo semplice approccio consiste nel ritardato inizio dell'accesso alla memoria, che si traduce direttamente in un aumento del tempo di accesso. Il ritardo necessario per garantire -5 ns di tempo di setup in condizioni veloci può comportare un allungamento del tempo di accesso di 12 ns in condizioni lente. Di conseguenza, in alcuni casi i produttori di memorie pseudo-statiche offrono dispositivi caratterizzati da gradi di velocità diversi (speed grade) in funzione della presenza o assenza del ritardo sul fronte di discesa del segnale chip-enable. Ad esempio, il dispositivo FM22L16 di Ramtron (una FRAM parallela da 4 megabit) funziona normalmente con un tempo di accesso di 55 ns quando è garantito un tempo di setup degli indirizzi di 0 ns; su richiesta del cliente può essere attivata un'opzione programmabile in fase di test per portare il tempo di setup degli indirizzi a -5 ns, con l'effetto collaterale di portare il tempo di accesso a 70 ns.

L'implementazione pseudo-statica della funzionalità ATD è in effetti molto più semplice e più sicura della versione completamente asincrona propria delle

SRAM. Tutti i progettisti di memorie, infatti, hanno sentito raccontare qualche storia spaventosa riguardante un'implementazione ATD non funzionante. Considerando che nelle memorie pseudo-statiche tutti gli accessi che vengono iniziati devono essere completati, e che la necessità di un setup negativo degli indirizzi può essere soddisfatta, l'implementazione di ATD si riduce a un semplice confronto tra l'indirizzo presente agli ingressi quando la lettura interna è completata e l'indirizzo fissato nel latch interno nel momento in cui l'accesso ha avuto inizio.

In una FRAM l'operazione di accesso è divisa in una fase di lettura e una fase di riscrittura. Questa soluzione ha due motivazioni. La prima è il requisito, proprio delle SRAM, della compatibilità con la scrittura ritardata (late-write), riferito alla possibilità che l'ingresso /WE scenda a livello basso dopo /CE e che i dati cambino dopo la discesa di /WE. La seconda ragione consiste nel fatto che solitamente il bus dati interno ha una larghezza superiore a quella del bus esterno. Se il bus interno ha una larghezza di 64 bit e quello esterno di 16 bit, ci sono 48 bit aggiuntivi che devono

Fig. 6 – L'accesso alla FRAM



essere letti e ripristinati anche durante i cicli di scrittura. Per tenere conto di questo, tutti gli accessi alla FRAM iniziano come operazioni di lettura e funzionano come tali fino al momento in cui i dati fissati negli amplificatori interni di rilevamento (sense amps) sono connessi al datapath esterno. Nelle caso delle operazioni di lettura i dati degli amplificatori di rilevamento vengono inviati ai pin dati della memoria. Nel caso delle operazioni di scrittura, i dati presenti sui pin dati della memoria vengono inviati agli amplificatori di rilevamento.

L'unica differenza tra un'operazione di lettura e una di scrittura, pertanto, è la direzione del flusso dei dati.

L'operazione di ripristino o riscrittura funziona allo stesso modo in entrambi i casi, con l'invio alla cella dei dati fissati nell'amplificatore di rilevamento. La figura 6 descrive il flusso dei dati nel corso dell'accesso a una FRAM.

Nelle memorie FRAM di vecchia concezione il passaggio tra le due fasi era controllato unicamente tramite il pin /CE. Il fronte di discesa di questo segnale determinava l'inizio della fase di lettura e il fronte di salita l'inizio della fase di ripristino. L'aggiunta della funzionalità ATD richiede che anche gli indirizzi possano comandare il passaggio tra le due fasi. La figura 7 mostra come la versione interna del segnale di chip-enable (cebint) venga modificata per tenere conto della funzionalità ATD. Nella figura 7 i segnali esterni sono indicati con lettere maiuscole e i segnali interni con lettere minuscole.

Il segnale cebint interno è leggermente ritardato rispetto al segnale /CE esterno, come spiegato precedentemente. Quando la fase di lettura ha inizio, gli indirizzi vengono fissati nel latch interno. Nella figura il segnale di controllo del latch è indicato con alatch e gli indirizzi fissati nel latch sono indicati con la[n:0]. Prima del completamento della fase di lettura, ogni cambiamento del-

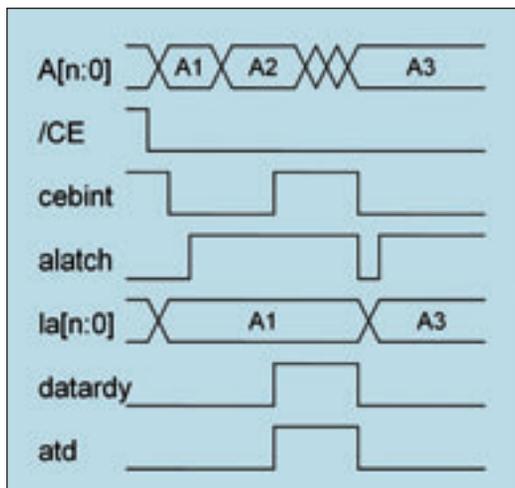


Fig. 7 - L'ATD pseudo-statico

l'indirizzo esterno viene ignorato. Quando la fase di lettura è completata, come indicato dal segnale datardy (data ready), il segnale atd passa a livello alto se il bus esterno A[n:0] assume uno stato diverso dal bus interno la[n:0]. Nell'esempio illustrato dalla figura 7 l'indirizzo esterno cambia prima che il segnale datardy salga a livello alto, pertanto il segnale atd passa a livello alto immediatamente dopo datardy.

Una volta che la transizione dell'indirizzo è stata rilevata come indicato dal segnale atd, il segnale interno cebint viene forzato a livello alto per iniziare la fase di ripristino. Quando questa è completata, il controllo di cebint viene restituito al pin esterno /CE. Se quest'ultimo

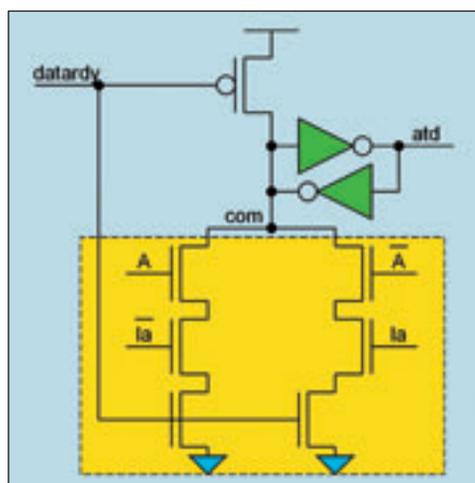
si trova ancora a livello basso, avrà inizio un nuovo accesso all'indirizzo che è presente nell'istante in cui la fase di ripristino viene completata. In questa implementazione, il tempo di ripristino della memoria è utilizzato per tenere conto del tempo di skew degli indirizzi. La funzione ATD viene attivata dall'indirizzo più veloce, mentre l'indirizzo più lento ha a disposizione l'intero tempo di ripristino della memoria per giungere a destinazione. La maggior parte delle memorie pseudo-statiche può tollerare un tempo di skew tra indirizzi di oltre 10 ns o anche 20 ns.

Anche il circuito di comparazione degli indirizzi è molto semplice, come appare dalla figura 8. Per semplicità l'immagine mostra la comparazione di un solo bit dell'indirizzo; il circuito all'interno della linea tratteggiata viene replicato per tutti i bit. Il segnale di comparazione comune (com) viene precaricato a livello alto e il percorso di scarica risulta interrotto quando il segnale datardy è a livello basso. Quando datardy passa a livello alto, il percorso di scarica è abilitato. Se A e la hanno lo stesso valore, un transistor NMOS per ogni lato si trova nello stato "off", pertanto il segnale com rimane a livello alto e il segnale atd a livello basso.

Se invece A e la assumono valori diversi, tutti i tre transistor NMOS su uno stesso lato commutano allo stato "on" permettendo la scarica del nodo com, con il conseguente passaggio del segnale atd a livello alto.

Le tecniche e i circuiti descritti in questo articolo consentono la progettazione di memorie pseudo-statiche che possono sostituire direttamente le memorie statiche. Il successo di questi progetti è dimostrato dal crescente impiego delle PSRAM nei telefoni cellulari e dall'adozione delle FRAM in luogo delle BBSRAM.

Fig. 8 - Circuito di comparazione degli indirizzi



Ramtron International  
readerservice.it n. 2