

Title	パイプライン化によるキャッシュの高周波動作の可能性に関する研究
Author(s)	鷓飼, 和歳
Citation	
Issue Date	1999-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1285
Rights	
Description	Supervisor:日比野 靖, 情報科学研究科, 修士

修士論文

パイプライン化による キャッシュの高周波動作の可能性に関する研究

指導教官 日比野靖 教授

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

鵜飼 和歳

1999年2月15日

要旨

マルチスレッド型プロセッサは、キャッシュに対してレイテンシよりも高スループットを求める。本論文では、キャッシュをこのような用途で使用する場合に、パイプライン化によってどの程度の高周波動作が可能となるかを検討する。

パイプライン化によってキャッシュを高周波動作させる場合に、メモリセルアレイの大きさをどの程度にするかが重要となる。メモリセルアレイを実際に設計することにより、メモリセルアレイの分割に対する高速化の割合を求め、適当な大きさを導き出す。また、パイプラインキャッシュという特殊な構成におけるプリチャージ回路の有効性についても検討する。

次に、設計目標とするキャッシュの容量を決定し、細部にわたり設計を行なう。各ステージごとの詳細な回路設計、チップ内の配置についての設計、およびチップ内を通るアドレスパスやデータパス等の設計により、パイプラインキャッシュの段数および動作周波数を明らかにする。また、パイプライン化により増加する回路の面積がどの程度になるかについても明らかにする。

目次

1	はじめに	1
2	キャッシュ機構のパイプライン化	2
2.1	マルチスレッド型プロセッサに適したパイプラインキャッシュ	2
2.2	キャッシュ機構のパイプライン化	2
2.3	キャッシュの容量と構成	4
2.4	階層的デコード法	6
2.5	メモリセルアレイの分割	7
2.5.1	メモリセルアレイの分割方法	7
2.5.2	配線長による遅延について	8
2.5.3	メモリセルアレイの分割による遅延縮小	8
2.6	チップ内の配置	10
2.7	2章のまとめ	11
3	パイプラインキャッシュの設計	12
3.1	ステージ構成	12
3.2	チップ内の配置および配線の設計	14
3.2.1	メモリセル、デコーダ、その他周辺回路のチップ内配置	14
3.2.2	配線抵抗および配線容量	17
3.3	回路設計	18
3.3.1	ラッチの設計	18
3.3.2	メモリセルの設計	19
3.3.3	各ステージの設計	20
3.4	レイアウト設計	23
3.4.1	レイアウト設計ルール	23

3.4.2	メモリセルおよび基本回路のレイアウト	23
3.5	面積の導出	24
3.5.1	基本回路の面積	24
3.5.2	配線による面積	25
3.5.3	各ステージごとの面積	25
4	SPICE シミュレーション	27
4.1	MOSFET モデル	27
4.1.1	パラメータの設定	27
4.1.2	特性	29
4.2	H レベルと L レベルの規定	31
4.3	ラッチの動作タイミング	32
4.4	メモリセルアレイの動作タイミング	34
4.4.1	タグメモリアクセスステージ	34
4.4.2	データメモリアクセスステージ	40
4.5	各ステージのシミュレーション	45
4.5.1	シミュレーション方法	45
4.5.2	シミュレーション結果	47
4.5.3	まとめ	49
5	考察	51
5.1	メモリセルアレイの分割に関する考察	51
5.2	プリチャージ回路に関する考察	54
5.3	現在研究されている高速 SRAM との比較	56
6	結論	57
A	各配線と周辺配線との距離関係	59
B	回路図	62
B.1	ラッチ	63
B.2	メモリセル	64
B.3	デコードステージ	65
B.4	タグメモリアクセスステージ	68
B.5	タグ比較ステージ	70

B.6	データメモリアクセスステージ	71
B.7	出力パスステージ	73
B.8	マルチプレクサステージ	74
B.9	プリチャージ回路を導入した時のビット線	75
C	レイアウト図	76
D	参考図	84

第 1 章

はじめに

マシンサイクルの短縮を追求し、RISC(Reduced Instruction Set Computer) が発展してきているが、配線遅延の影響で速度向上が飽和してきている [1]。それを克服するため、プロセッサに高度のパイプライン処理を行いスループットの向上を図ることが盛んに行われている。私たちの研究室では、パイプラインの段数を増やすことにより高周波で動作させ、ステージ数と同数のスレッドを時分割に並列に処理させることにより様々なハザードを回避するマルチスレッド型パイプラインプロセッサの研究を行っている。このプロセッサを効率的に動作させるには、命令やデータを高周波で供給するパイプライン化されたキャッシュが必要となる。

キャッシュをパイプライン化する場合、クロックサイクルの長さはもっとも遅い操作に依存してくるため、パイプラインの段数を増やすことによりクロックサイクルの短縮、すなわち高周波動作が可能となる。マルチスレッド型プロセッサではキャッシュをかなり深くパイプライン化することが可能であり、そのためどの程度の高周波動作が可能となるかを検討する。また、多段のパイプライン化を行う上でボトルネックとなるメモリセルアレイの配線遅延やゲート遅延について、実際にメモリセルアレイを設計することにより評価する。

メモリセルアレイの分割はデコーダの複雑化をもたらす。これによる配線遅延やゲート遅延の増加を縮小するため、デコーダの分割についても実際に設計を行ない評価する。最後に、メモリセルアレイ・デコーダ・その他周辺回路を組合せ、各ステージごとに動作を確認する。

第 2 章

キャッシュ機構のパイプライン化

まず、マルチスレッド型プロセッサにパイプラインキャッシュが適する理由を説明する。そして、キャッシュ機構のパイプライン化についての概要を示し、高周波動作のパイプラインキャッシュを設計するために考慮しなければならない点を指摘する。設計目標とするパイプラインキャッシュの容量・構成を示し、具体的なデコード方法について検討した。また、キャッシュを高周波動作させるためにはメモリセルアレイの分割について考える必要があるが、設計前に予想されることについての検討を行なった。

2.1 マルチスレッド型プロセッサに適したパイプラインキャッシュ

マルチスレッド型パイプラインプロセッサは、ステージ数と同数のスレッドを時分割に並列に処理させるので、1スレッド当たりのレイテンシは増加する。しかし、全体のスループットは向上する。この多数のスレッドを時分割に並列に処理するという特徴から、キャッシュについてもレイテンシの増加は許される。よって、マルチスレッド型パイプラインプロセッサにおいて、パイプライン化されたキャッシュが有効となってくる。

2.2 キャッシュ機構のパイプライン化

プロセッサの高速化の要求によって、プロセッサのパイプライン処理が行なわれるようになった。パイプライン処理は複数の命令を数段の処理に分割し少しずつずらして同時並行的に実行する実現方式である。これにより命令のスループットを向上させている。本

研究ではこのパイプライン処理をキャッシュ機構に適用することによってキャッシュのスループットの向上をねらう。

パイプラインキャッシュの操作は大まかに次のような順で行なわれる。

1. アドレスデコード：アドレスのインデックス部に基づき対応するキャッシュのブロックを選択する。
2. タグメモリのアクセス：タグおよび有効ビットの読み出しや書き込みを行なう。
3. タグ比較：タグの比較をし、ヒット/ミスを判定する。
4. データメモリのアクセス：タグ比較の結果ヒットであれば、データの読み出しや書き込みを行なう。
5. 出力バス：データメモリから読み出されたデータをマルチプレクサまで伝送するパス。
6. マルチプレクサ：読み出されたデータ(ブロック)から要求されたワードを選択する。

6 ステージ構成パイプラインキャッシュは、上記の各処理の間にラッチが入り、読み出しまたは書き込みの処理が各クロックごとに行なわれる。パイプラインの流れを示した図を図 2.1 に示す。

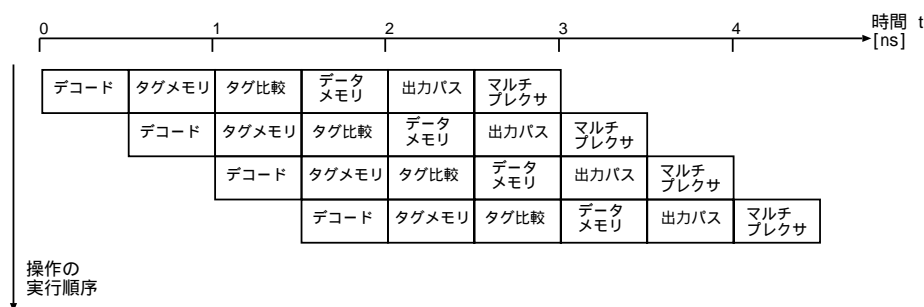


図 2.1: パイプラインステージ

パイプラインキャッシュには6種類の操作があり、タグの比較をデータメモリのアクセスの前に行なっている。しかし、通常のキャッシュでは、データメモリの読み出しと並行してタグの読み出しも行なわれており、タグの比較は最終段で行なわれる。

仮に、パイプラインキャッシュにおいてタグの比較をデータメモリのアクセスの後に行うとする。その場合、データメモリアクセス時には書き込みがヒットしているかわからな

いため、データメモリのアクセスステージは読み出し処理だけが行なわれる。最終段でタグ比較が行なわれ、書き込みがヒットとなり書き込み処理をデータメモリに対して行なう時、次々にやってくる読み出し処理がデータメモリを専有してしまっており、データメモリに対して書き込みを行なうタイミングがない。そのために、タグの比較をデータメモリアクセスの前に行ない、読み出しと書き込みを同一ステージで行なえるようにした。

本研究において設計したパイプラインキャッシュは、処理に時間のかかるステージを分割し、上記の操作をさらに9段にパイプライン化している。なぜ9段のステージ構成としたかは、後章で順に述べていく。

2.3 キャッシュの容量と構成

マルチスレッド型プロセッサからは、キャッシュへのアクセスが各クロックサイクルごとに行なわれる。このため、各スレッド間でキャッシュのブロック競合が起こるのではないかと懸念される。また、マルチスレッド型プロセッサの場合、メインメモリへのアクセス要求は、各スレッドが生起させたキャッシュミスから次々にやってくる。このため、メインメモリへのアクセスペナルティをどのように縮めるかが重要となる。

マルチスレッド型プロセッサに用いるキャッシュには様々な問題があるが、本研究の設計目標とするキャッシュの容量や構成については相原の提案 [2] を参考した。

まず、表 2.1 のベンチマークの中で中間的な局所性を持つ `spice` を基準として考え、`spice` のトレースデータの 90 % を収められるよう、キャッシュのデータ容量は 1 スレッドつき、8K ワードとする。複数のスレッドを扱う場合、スレッド数分の容量を持つスレッド共用キャッシュが必要となる。よって、本研究のキャッシュは 16 スレッド分 (128K ワード) のデータ容量を持つスレッド共用キャッシュとする。

表 2.1: ベンチマークのアドレス分布範囲

	cc	spice	tex
アドレス分布範囲 [ワード]	143K	107K	32K
90 % 範囲 [ワード]	35.3K	6.7K	0.3K

また、キャッシュの構成を以下のようにする。

- 4 ウェイ・セット・アソシアティブキャッシュ。

- ラインサイズは、4ワード。
- キャッシュの書き込み方式は、メモリラヒックを低く抑えることができるライトバック方式とする。ライトバック方式とすることで、キャッシュ機構が複雑化し、レイテンシが増加するが、パイプライン化をさらに進めることにより、スループットを向上させる。
- 書き込み時のミスが生じた場合、ライト・アロケートを使用する。
- キャッシュの総容量は、有効ビット部 32Kb、タグ部 480Kb、データ部 4Mb より、4608Kb。

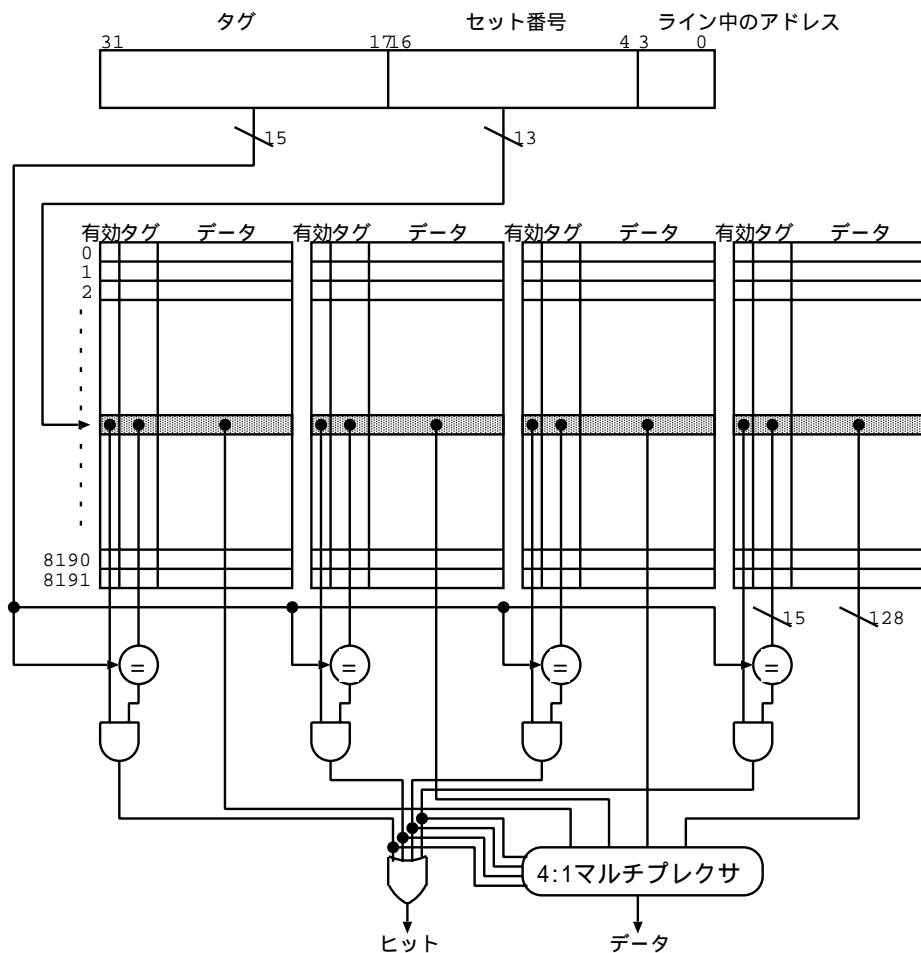


図 2.2: キャッシュの構成

2.4 階層的デコード法

キャッシュ中の操作でも特に遅いものにデコーダがある。そこで、デコーダを3段の処理に分割しパイプライン化を施すことによって階層的にデコードを行なう方法を検討する[4]。

図 2.3のように、ラッチ 1 からラッチ 2 までをデコードステージ 1、ラッチ 2 からラッチ 3 までをデコードステージ 2、ラッチ 3 からラッチ 4 までをデコードステージ 3 に分割し、アドレス(セット番号)のデコードを行なう。このように処理を行なうことにより、入力ラッチからメモリセルアレイまでのアドレス線による遅延やデコーダの論理による遅延を分割することができる。

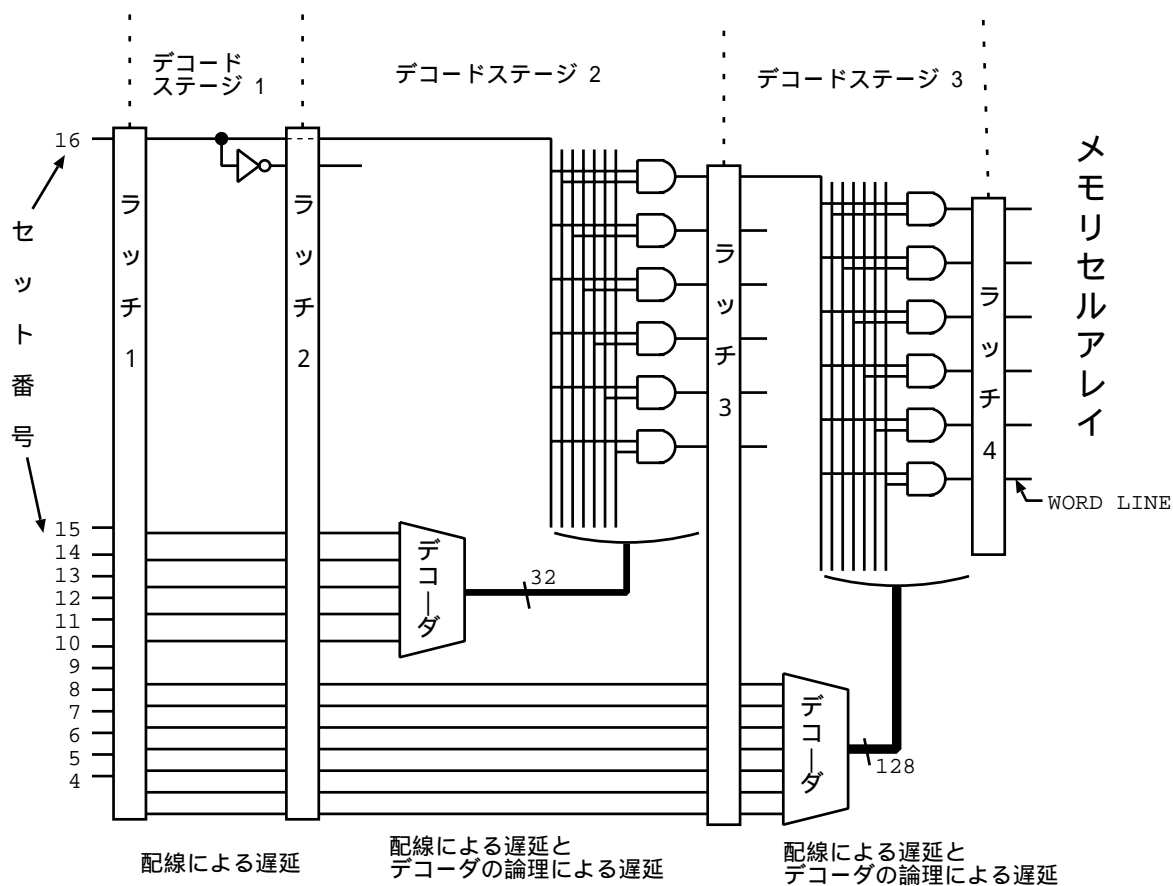


図 2.3: 階層的デコードの流れ

前節で述べたように、本研究で設計したキャッシュは4ウェイセットアソシアティブで、ラインサイズは4ワード、データ容量は128Kワード(4Kb)であるので、セット数は8192

となる。よって、図中のセット番号は 13bit とした。

2.5 メモリセルアレイの分割

本節では、メモリセルアレイの遅延時間を縮小させる方法について示す。まず、なぜメモリセルアレイを分割するか。そして、分割による遅延時間縮小について、設計およびシミュレーションを行なう前に予想される事柄を検討した。

2.5.1 メモリセルアレイの分割方法

図 2.4 のように、デコーダはラッチを挿入することによりパイプラインの段数を増やすことができる。しかし、メモリセルアレイの読み出し操作はパイプライン化を施すことができない。よって、メモリセルアレイは物理的な大きさを小さくすることにより、ワード線およびビット線の配線遅延時間を短くし、クロックサイクルの短縮をはかる。

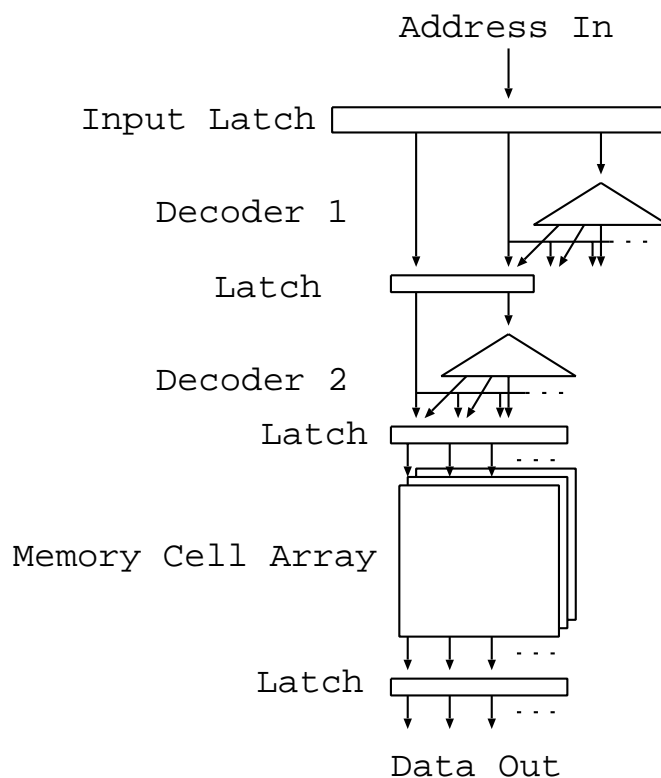


図 2.4: パイプラインキャッシュのブロック図

2.5.2 配線長による遅延について

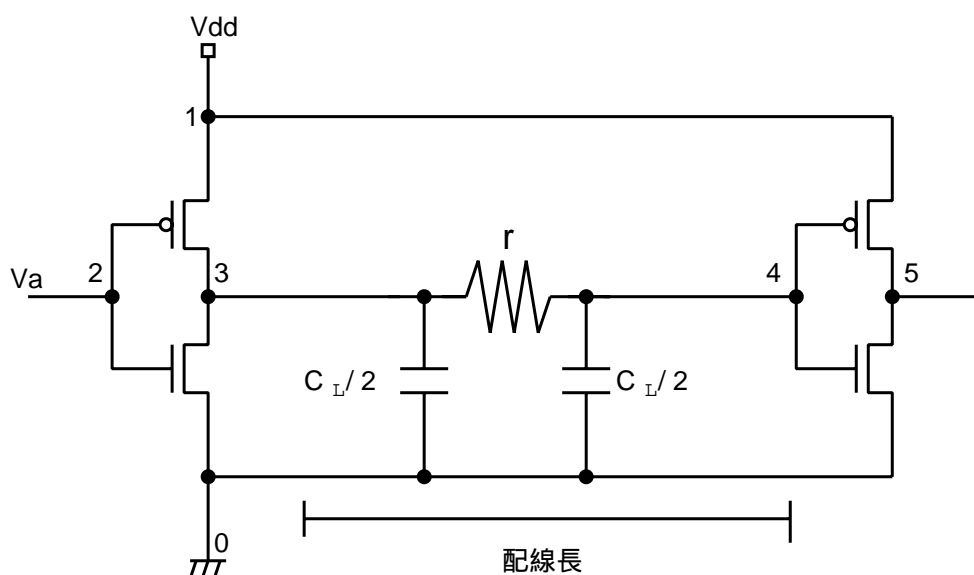


図 2.5: 2 段のインバータと配線

図 2.5の回路の遅延時間は次式で表される。

$$\tau = R_{on}C_L + R_{on}C_g + rC_L + rC_g$$

上式の R_{on} は MOSFET のオン抵抗、 C_g は MOSFET のゲート容量、 r は配線抵抗、 C_L は配線容量である。

配線長を短くすると、 r と C_L を含む項が小さくなる。一般に、上式の 4 項はそれぞれの値が桁違いに異なり、どれか 1 項または数項が遅延時間を支配する。 r と C_L を含む項が上式の中で支配的である時、配線長の短縮により遅延時間を縮小できる。しかし、 r と C_L を含む項が r と C_L を含まない項より一桁小さくなった時、配線長の短縮による遅延時間の縮小がほとんど得られなくなる。

2.5.3 メモリセルアレイの分割による遅延縮小

メモリセルアレイの分割について、 $16 \times 16bit$ の小さなモデルでわかり易く示した図を図 2.6 に示す。メモリセルアレイの遅延には、ドライバやセンスアンプによるゲート遅延と、ワード線やビット線による配線遅延がある。配線遅延は分割により比例的に小さくな

ると考えられるが、ゲート遅延は一定のままであるため、メモリセルアレイの分割を進めると遅延縮小の限界が現れてくる。

しかし、メモリセルアレイの分割によりセンスアンプにつながるセルの数が少なくなるため、センスアンプの感度を下げることが可能となる。また、ドライバに関して同様に小さくすることができ、ゲート遅延の減少が可能となる。

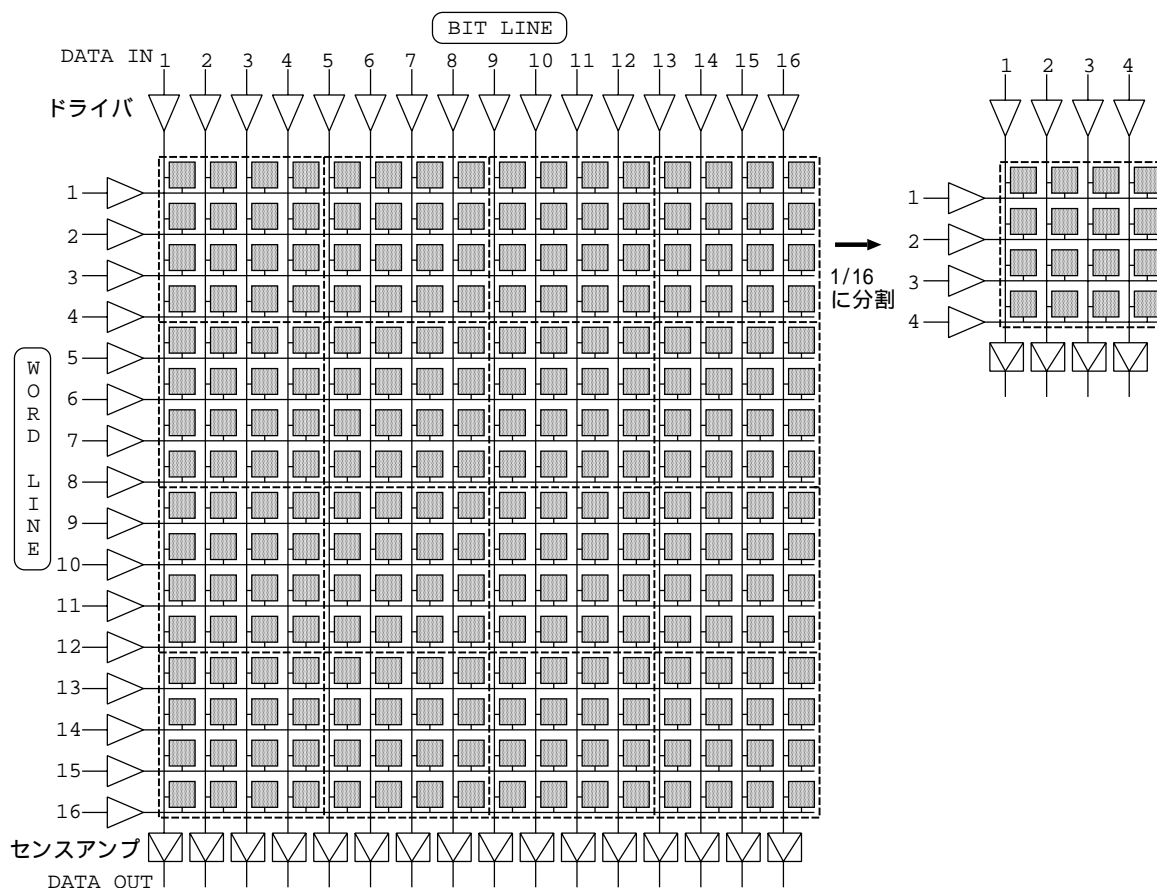


図 2.6: メモリセルアレイの分割

2.6 チップ内の配置

パイプラインキャッシュにおいて、各ステージのラッチのタイミングがうまく合うように、キャッシュを構成する各回路をチップ上にどのように配置するかが問題となる。特に、キャッシュの面積の大部分を占めるメモリセルアレイの配置が重要となる。

本設計では、各メモリセルアレイを図2.7のように配置し、各メモリセルアレイまでの長い配線をH木状に配線した。このように配置および配線することにより、長い配線の途中のラッチまでのクロックを均等に配分できる。

また、高周波動作のためには、長い配線の途中にラッチを複数入れなければならないかも知れない。H木状に配線することにより、長い配線の任意の位置にラッチを挿入することができる。

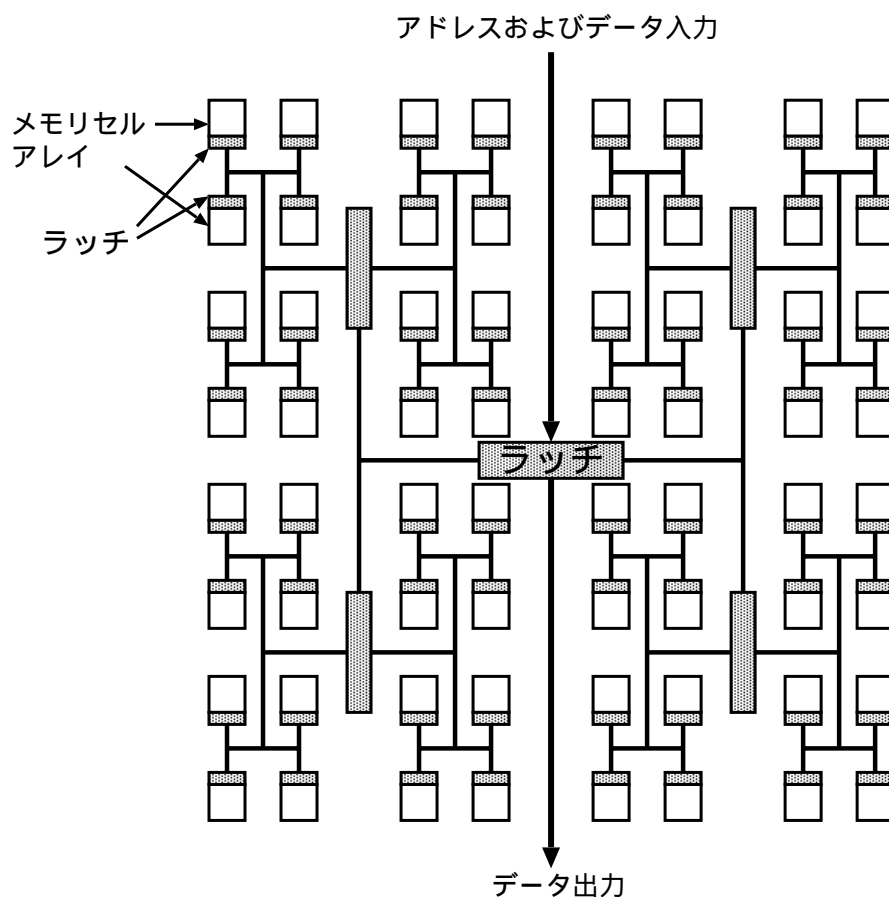


図 2.7: H 木による配置

2.7 2章のまとめ

本章で述べた事柄を簡単に以下にまとめた。

- キャッシュ機構のパイプライン化

読み出しまたは書き込みという異なった処理を同一のパイプラインで行なわなければならない。2つの処理を資源競合なく行なうよう、読み出しの場合もタグの比較が終了してからデータメモリへアクセスを行なう構成とした。

- 階層的デコード法

デコードを数段のステージにわけて行なうため、階層的デコード法を使用する。

- メモリセルアレイの分割

遅延時間を縮小するためには、メモリセルアレイは物理的な大きさを分割する必要がある。しかし、メモリセルアレイは分割を進めると、それに比例して遅延時間が縮小せず、ある曲線になると考えられる。

- H木による配置・配線

各ステージのラッチまでのクロックを均等に配分するため、メモリセルアレイをH木状に配置する。メモリセルアレイまでの長い配線上に、任意の位置にラッチを入れることが可能。

H木状に配線することにより、長い配線上の任意の位置にラッチを入れることが可能である。また、デコーダもラッチを挿入することによりパイプラインの段数を増やしていける。そのため、分割を進めると遅延時間の短縮の限界が現れるメモリセルアレイが高周波動作のネックになると予想される。

3章以降では、各ステージごとに詳細な回路の設計、チップ内の配置についての設計、およびチップ内を通るアドレスバスやデータバス等の配線の設計を行ない、各ステージの最小クロックサイクル時間を求める。そして、メモリセルアレイの大きさと遅延時間の関係を求め、パイプラインキャッシュの高周波動作の可能性を検討する。

第 3 章

パイプラインキャッシュの設計

0.25 μm テクノロジーにおけるパイプラインキャッシュの設計を行なった。本章の各節および次章のシミュレーションは、設計をする上で密接に関連しており、それぞれ何度も繰り返し最適な設計を行なった。特にラッチを挿入する箇所については、シミュレーションを繰り返し、最適な位置にできるだけ小さなラッチを挿入するよう努めた。

3.1 ステージ構成

はじめに、パイプラインキャッシュのステージ構成を図 3.1 に示す。図のステージ構成について以下に簡単に説明する。まず、デコードステージは、入力ラッチからチップ上に並べられたメモリセルアレイへアドレスを分配するパスおよびデコーダが存在する。この長い配線による遅延およびデコーダによる遅延を分割するため 3 ステージで処理を行なう。次に、通常の SRAM では並行して行なうタグ/データメモリのアクセスを、タグメモリアクセス、タグ比較、データメモリアクセスと 3 ステージに分けた。パイプラインキャッシュでは、読み出しまたは書き込みの処理がクロックサイクルごとに連続してやってくるため、このように分けることで資源競合なく 2 つの異なる処理が行なえるようになる (2.2 節 参照)。出力パスステージについても、デコードステージと同様に、チップ上に並べられたメモリセルアレイからチップの中央へ出力データを集めてくるための長い配線が存在する。この遅延を分割するため 2 ステージに分けた。

最終的には、後節および後章で述べる設計およびシミュレーションを繰り返し図 3.1 に示す 9 段構成が最適であると導き出した。

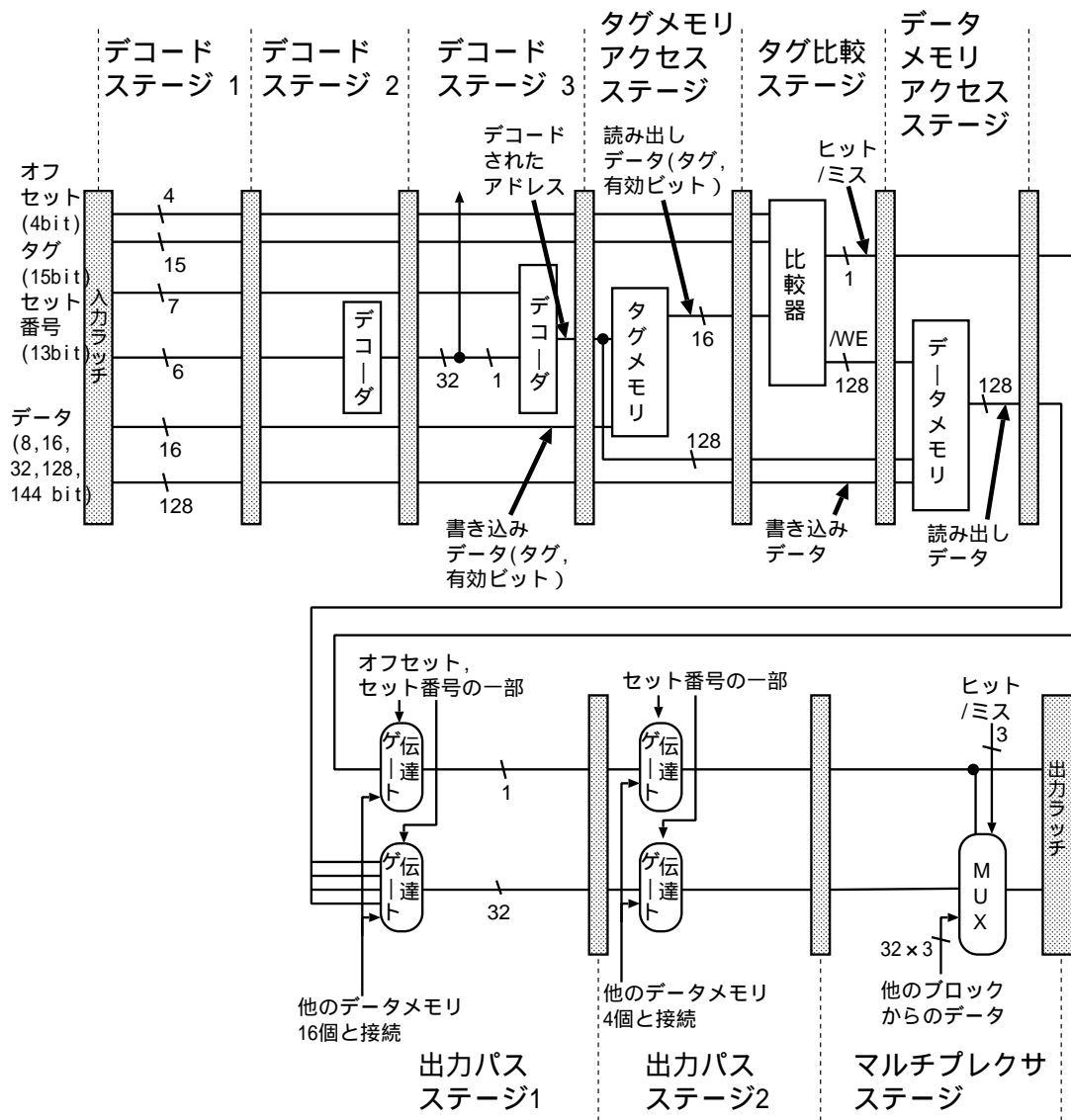


図 3.1: 9 ステージ構成のパイプラインキャッシュ

3.2 チップ内の配置および配線の設計

本節では、メモリセルやデコーダ、その他の回路のチップ内の最適な配置を検討し、配線長を求める。また、周辺配線の位置関係を考慮し、各層の金属配線の配線容量および配線抵抗を導き出す。

3.2.1 メモリセル、デコーダ、その他周辺回路のチップ内配置

チップ全体の容量は、4608Kb である。これを $128 \times 144\text{bit}$ のメモリセルアレイ 256 個に分割し、チップ内の配置を考えた。

まず、図 3.2 はチップ全体のメモリセルアレイ、デコーダ、その他周辺回路の配置を示す。図の 1/4 の部分には、それぞれ $128 \times 144\text{bit}$ のメモリセルアレイが 64 個入る。

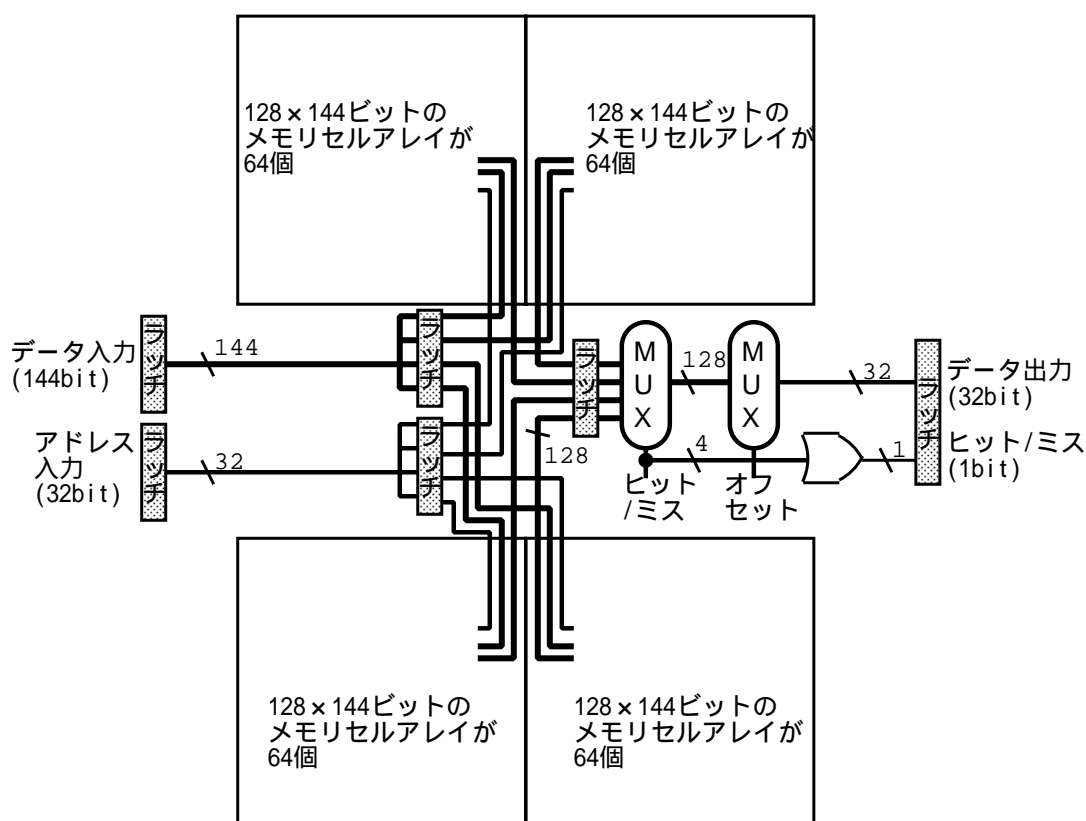


図 3.2: メモリセルアレイ、デコーダ、その他周辺回路のチップ内配置

次に、図 3.3 にチップの 1/4 の部分のメモリセルアレイおよびデコーダの配置を示す。H 木を用いることにより各メモリセルアレイに均等の遅延時間でアドレスやデータが供給されるようにした。また、中央部にラッチおよびミドルデコーダを配置することにより長距離配線を避け、クロックサイクル時間を短縮し、他のステージとクロックサイクル時間の大きな差が生じないように設計を行なった。

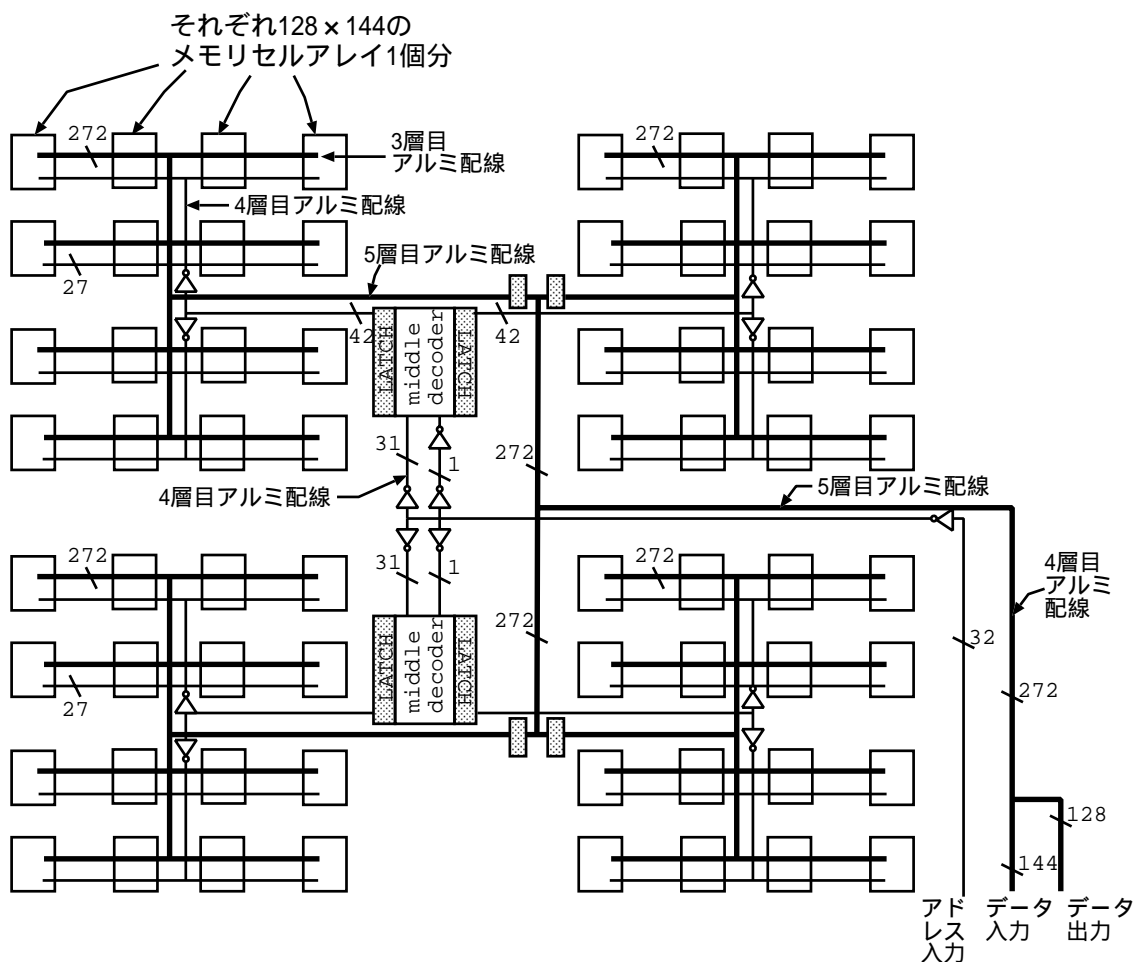


図 3.3: 1/4 の部分のチップ内配置

図 3.4に $128 \times 144\text{bit}$ のメモリセルアレイ 2 個の配置を示す。

$128 \times 144\text{bit}$ のメモリセルアレイは、 $128 \times 16\text{bit}$ のタグメモリと $128 \times 128\text{bit}$ のデータメモリに分けられる。データメモリは、さらにローカルデコーダによりワード線が2分割される。このようにワード線を2分割することにより、配線抵抗と配線容量はそれぞれ半減し遅延は $1/4$ に短縮する [5]。

図 3.4の中にタグメモリアクセスステージ、タグ比較ステージ、データメモリアクセスステージの3ステージが入るため、この図はラッチの数が多くデータの流れがわかりにくい。そのため、データが流れていく順番にラッチに色を付けた図を付録 D に示した。

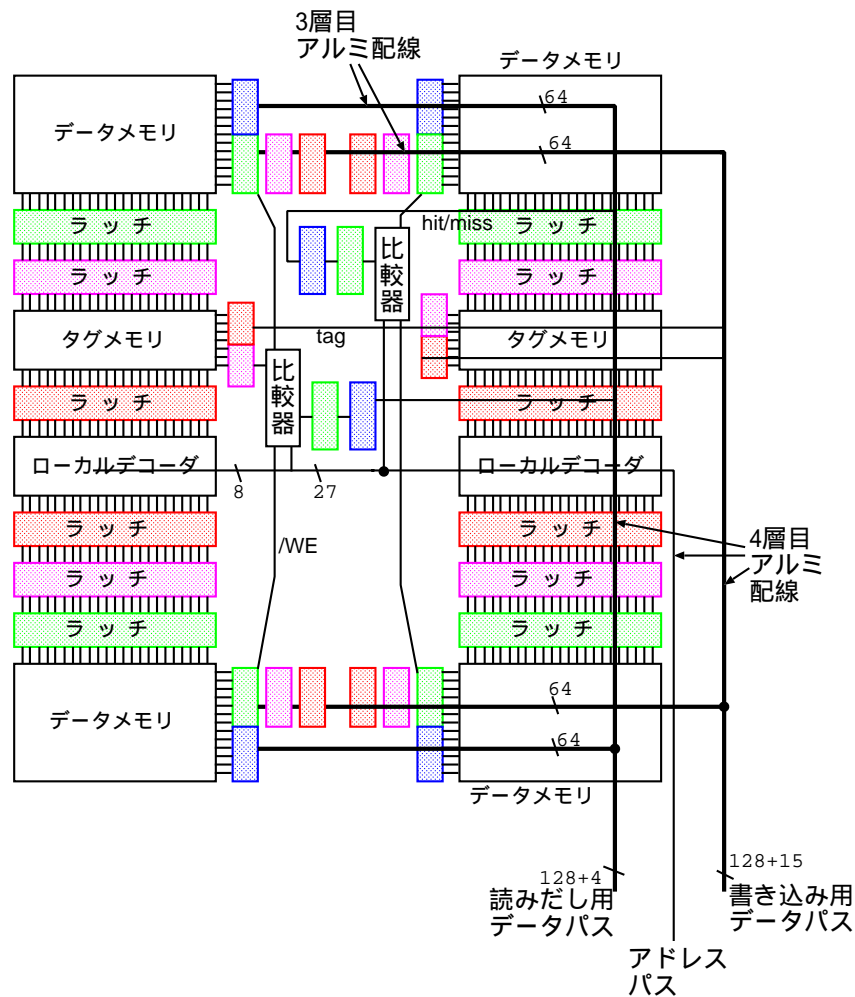


図 3.4: 128×144 のメモリセルアレイ 2 個分の配置

3.2.2 配線抵抗および配線容量

本稿ではビット線、ワード線および各層のアルミ配線の配線抵抗および配線容量について求める。まず、配線抵抗および配線容量の求め方について一般的な式を示し、そして各配線ごとに求めた結果を表 3.1 に示す。周辺配線をどのように考慮したかについては、周辺配線との距離関係の図を付録 A に示した。

配線抵抗の求め方

長さ l [cm]、幅 w [cm]、厚さ t [cm]、配線の体積抵抗率 ρ [$\Omega \cdot cm$] の配線を仮定する。シート抵抗値 ρ_s [Ω] は次式より求まる。

$$\rho_s = \frac{\rho}{t} [\Omega]$$

このシート抵抗値 ρ_s を用い、配線抵抗 R を次式より求めることができる。

$$R = \rho_s \frac{l}{w} [\Omega]$$

アルミの比抵抗 $\rho = 3[\mu\Omega \cdot cm]$ 、ポリシリコンの比抵抗 $\rho = 1000[\mu\Omega \cdot cm]$ である。

配線容量の求め方

次式に単位面積あたりの酸化膜容量 C_o の求め方を示す。(SiO_2 の比誘電率 $\varepsilon_{ox} = 3.82$, 真空の誘電率 $\varepsilon_0 = 8.85 \times 10^{-12}[F/m]$)

$$C_o = \varepsilon_{ox} \varepsilon_0 / \text{酸化膜厚} [F/m^2]$$

この単位面積あたりの酸化膜容量 C_o を用いて、配線容量 C を次式より求めることができる。

$$C = C_o \times \text{面積} [F]$$

表 3.1: 各配線のシート抵抗値および単位面積あたりの酸化膜容量

		シート抵抗値 [Ω]	酸化膜容量 [$\mu F/m^2$]	配線幅 [μm]	厚さ [μm]	配線間ピッチ [μm]
ワード線	ポリ配線	120	420	0.25	0.083	/
	裏打ちアルミ	0.12	96.38	0.75	0.25	/
ビット線		0.36	153	0.25	0.083	/
3層目アルミ配線		0.091	88.4	1.0	0.33	2.0
4層目アルミ配線		0.06	74.9	1.5	0.5	3.0
5層目アルミ配線		0.06	51.7	1.5	0.5	3.0

3.3 回路設計

各ステージごとに詳細な回路設計を行なった。回路図はカラーで描かれているため、付録 B に掲載した。本節で設計した回路図をもとに、次章では SPICE による遅延時間のシミュレーションを行なう。

3.3.1 ラッチの設計

ステージごとに最適なラッチの大きさが異なるため、pMOS および nMOS のゲート幅を 3 倍ずつ変化させた 3 種類のラッチを設計した。各ステージを設計する際、これらの 3 種類のラッチのうちできるだけ小さなラッチを使用するよう努めた。図 3.5 に一番小さなゲート幅の latch-ss の回路図を示す。latch-s、latch-m については、付録の図 B.2 に示す。

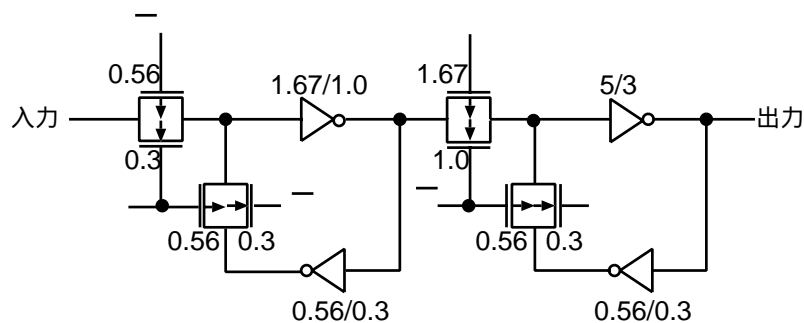


図 3.5: ラッチの回路図 (latch-ss)

3.3.2 メモリセルの設計

メモリセルは6トランジスタで構成した。図 3.6に回路図を示す。各 MOSFET のゲート幅、ゲート長および拡散領域の面積を表 3.2に示す。また、メモリセル内の配線の抵抗・容量を表 3.3に示す。拡散領域の面積および配線の抵抗・容量についてはレイアウト図(図 C.12参照) から導き出した。

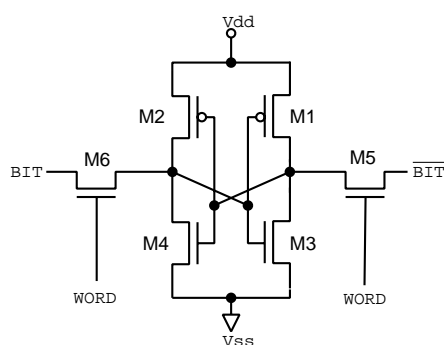


図 3.6: メモリセルの回路図

表 3.2: $0.25\mu\text{m}$ のメモリセル内の各 MOSFET の大きさ

	M1	M2	M3	M4	M5	M6
ゲート長 [μm]	0.3	0.3	0.25	0.25	0.25	0.25
ゲート幅 [μm]	1.0	1.0	1.0	1.0	0.75	0.75
ソース拡散領域 [μm^2]	0.59375	0.59375	0.46875	0.46875	0.5625	0.28125
ドレイン拡散領域 [μm^2]	0.625	0.625	0.5	0.5	0.5	0.5625

表 3.3: メモリセル内の配線の抵抗・容量

	シート抵抗値 [Ω]	単位面積あたりの酸化膜容量 [$\mu\text{F}/\text{m}^2$]
ゲート上のポリシリコン配線	120.48	6760
ポリシリコン配線	120.48	338
アルミ配線	0.36	67.6

3.3.3 各ステージの設計

• デコードステージ

デコード部には、アドレスの入力ラッチからそれぞれのメモリセルアレイまでの配線 $13920\mu m$ と、13 ビットのセット番号のデコードを行なうデコーダがある。この配線による遅延とデコーダの論理による遅延を分割するため、3ステージで処理を行なうように設計した。

– 1ステージ目

1ステージ目は、アドレスの入力ラッチからチップ中央に位置するラッチまでの部分である。このステージには、5層目アルミ配線 $4736\mu m$ がある。回路図を図 B.4に示す。

– 2ステージ目

2ステージ目は、ラッチから中間デコーダを通り次のラッチまでの部分である。このステージには、 $3780\mu m$ の4層目アルミ配線、 $2368\mu m$ の5層目アルミ配線およびデコーダの論理回路6段がある。回路図を図 B.6に示す。

– 3ステージ目

3ステージ目は、ラッチからローカルデコーダを通り次のラッチまでの部分である。このステージには、 $1184\mu m$ の5層目アルミ配線、 $1260\mu m$ の4層目アルミ配線、 $592\mu m$ の3層目アルミ配線およびデコーダの論理回路7段がある。回路図を図 B.8に示す。

• タグメモリアクセスステージ

本設計では、タグメモリとデータメモリに分ける。このように分けることにより、データメモリのアクセスの前にタグ比較が行え、データメモリの読み出しとヒット時の書き込みが同一ステージで行なえる。本設計にはプリチャージ回路を使用していないが、プリチャージ回路に関する議論は考察で行なう。回路図を図 B.10, B.11, B.12に示す。

- タグ比較ステージ

タグ比較ステージでは、タグメモリから読み出されたタグおよび有効ビットの比較を行なう。そして、ヒット時の書き込みの時、与えられたオフセットアドレスのバイト単位の書き込みが行なえるよう/WE 信号の出力を行なう。回路図を図 B.13 に示す。

- データメモリアクセスステージ

128 × 128bit のデータメモリのワード線をローカルデコーダで 2 つに分割し、ワード線の長さと同じワード線に接続されるメモリセルのゲートの数を 1/2 に減らした(図 3.4 参照)。また、ポリシリコンのワード線の上からアルミ配線を裏打ちすることにより、配線抵抗を低く抑えた。このワード線部分の回路図を図 B.16 に示す。また、ビット線部分の回路図を図 B.17, B.18 に示す。

- 出力パスステージ

出力パスステージの遅延要因には、9184 μm の配線によるものとトランスミッションゲートによるものがある。トランスミッションゲートの代わりにトライステートインバータを使用する設計も考えられるが、本設計では面積の小さいトランスミッションゲートを使用した。9184 μm の配線を分割するために、デコードステージでラッチを挿入した同じ位置にラッチを挿入する。このように配置することにより、中間デコーダでデコードされたアドレスを使い、容易にトランスミッションゲートのコントロールができる。

- 1 ステージ目

1 ステージ目は、メモリセルアレイの出力ラッチからデータパスの中間に位置するラッチまでである。このステージには、592 μm の 3 層目アルミ配線、1260 μm の 4 層目アルミ配線、1184 μm の 5 層目アルミ配線、16 個のトランスミッションゲートおよび 2 段の論理回路がある。回路図を図 B.19 に示す。

- 2 ステージ目

2 ステージ目は、データパスの中間に位置するラッチからマルチプレクサステージの入力ラッチまでの部分である。このステージには、

6148 μm の配線および4個のトランスマッションゲートがある。回路図を図 B.20に示す。

- マルチプレクサステージ

マルチプレクサステージには、セット内のブロックを選択するマルチプレクサとブロック内のワードを選択する2つのマルチプレクサがある。ステージの論理回路の段数は6段である。回路図を図 B.21に示す。

表 3.4に、各ステージで最大遅延を形成する経路の配線長および論理回路の段数をまとめた。タグメモリアクセスステージおよびデータメモリアクセスステージは除く。

表 3.4: 各ステージの配線長および論理段数

	配線長 [μm]	論理段数
デコードステージ 1	4736(5 層)	/
デコードステージ 2	3780(4 層)、2368(5 層)	8 段
デコードステージ 3	592(3 層)、1260(4 層)、1184(5 層)	8 段
タグ比較ステージ	/	11 段
出力パスステージ 1	592(3 層)、1260(4 層)、1184(5 層)	2 段
出力パスステージ 2	3780(4 層)、2368(5 層)	/
マルチプレクサステージ	/	6 段

各ステージの遅延要因には、表 3.4でまとめた配線と論理回路以外に、ラッチおよびトランスマッションゲートによる遅延がある。各遅延要因による遅延時間については表 4.19,4.20を参照。

3.4 レイアウト設計

3.4.1 レイアウト設計ルール

本設計で用いるレイアウトルールを表 3.5 に示す。テクノロジーの $1/2$ の値を最小線幅とした。最小線幅を a で表す。

表 3.5: レイアウト設計ルール

マスク	加工形状	寸法
Thinox	最小 Thinox 幅	$2a$
	最小 Thinox 間隔	$2a$
	p-thinox と n-thinox との最小間隔	$6a$
ポリ	最小ポリ幅	$2a$
	最小ポリ間隔	$2a$
	ポリと Thinox との最小間隔	a
	最小ポリゲートと突出長	$2a$
コンタクト	最小コンタクト領域	$2a \times 2a$
	最小コンタクト間隔	$2a$
	コンタクトへのポリまたは thinox の最小オーバーラップ	a
	ゲートポリまでの最小間隔	$2a$
金属	最小金属幅	$2a$
	最小金属間隔	$2a$
	コンタクトへの最小金属オーバーラップ	a

3.4.2 メモリセルおよび基本回路のレイアウト

レイアウトルールに基づき、次に示す基本回路のレイアウトを行なった。ゲート長は $0.25\mu m$ である。NAND ゲート、NOR ゲート、EXOR ゲート、インバータ、トランスミッションゲート、カレントミラー（センスアンプ）、アウトプットインバータ（センスアンプ）、プリチャージ回路、パストランジスタ、ラッチ、メモリセル。

付録 C に各基本回路のレイアウト図を示す。

3.5 面積の導出

本節では、前節の基本回路のレイアウトから面積を導出する。

3.5.1 基本回路の面積

前節の基本回路のレイアウト設計ではゲート幅の変化に対応して面積を求められるよう、次のような工夫を施している。

ゲート長 (L) の方向のサイドを $fixside$ 、ゲート幅 (W) の方向のサイドを $flexside$ とすることにより、基本回路の面積 (S) を次の式により求めることができる。ここで α は、ゲート幅の変化分。

$$S = fixside \times (flexside + \alpha)$$

例としてラッチを取り上げ、ゲート幅 (W) の変化に対応して面積を導出する方法を示す。図 C.11 のレイアウト図から図 B.2 の 4 種類のラッチのうち latch-m の面積を求める。

$$\begin{aligned} S &= fixside \times (flexside + (15 - 1) + (10 - 1) + (1.25 - 0.5) + (0.75 - 0.5)) \\ &= 8.25 \times (8.5 + 24) = 268.1 \mu m^2 \end{aligned}$$

他の基本回路もこのようにゲート幅の変化に対応して面積を求めた。表 3.6 に前節でレイアウトを行なった基本回路の $fixside$ と $flexside$ の値を示す。

表 3.6: 基本回路の $fixside$ と $flexside$ の値

基本回路	$fixside$ [μm]	$flexside$ [μm]
NAND ゲート	2.75	2.25
NOR ゲート	2.75	2.25
EXOR ゲート	4.5	3.625
インバータ	1.75	3.0
トランスマッションゲート	1.75	3.5
カレントミラー (センスアンプ)	4.75	2.25
アウトプットインバータ (センスアンプ)	4.375	2.875
プリチャージ回路	3.75	1.375
パストランジスタ	1.75	0.75
ラッチ	8.25	8.5
メモリセル	4.4 (word-line-side)	4.625 (bit-line-side)

3.5.2 配線による面積

配線には基本回路どうしをつなぐ 1, 2 層目配線、入出力ラッチからメモリセルアレイまでの 3,4,5 層目配線がある。各配線の面積の評価は次のように行なった。

- 基本回路どうしをつなぐ 1, 2 層目配線

前項で求めた基本回路の面積 3 に対して配線チャンネルの面積 7 を加えた。

- 入出力ラッチからメモリセルアレイまでの 3,4,5 層目配線

すべて 3,4,5 層で配線できるように設計した。

3.5.3 各ステージごとの面積

パイプラインキャッシュは、先に求めた基本回路により構成される。配線チャンネルを考慮し、各ステージごとに面積を求めた。そして、パイプライン化により増加する回路の面積がどの程度になるかについて考察した。

ラッチとメモリセルは、配線チャンネルを必要としないで隣合う回路どうしを並べることができるよう、レイアウトに工夫を施した(図 C.11, C.12 参照)。ラッチとメモリセル以外の論理回路には、論理回路の面積 3 に対して配線チャンネルの面積 7 を加えた。

表 3.7 に各ステージごとの面積を示す。それぞれのステージには、そのステージの入出力ラッチの面積を含む。出力ラッチの面積は、次のステージに含まれる。

表 3.7: 各ステージごとの面積 (パイプラインキャッシュ)

	面積 [mm^2]
デコードステージ 1	0.0472
デコードステージ 2	0.274
デコードステージ 3	7.24
タグメモリアクセスステージ	38.5
タグ比較ステージ	7.55
データメモリアクセスステージ	157
出力パスステージ 1	19.4
出力パスステージ 2	0.58
マルチプレクサステージ	0.0888

表 3.7 より、パイプラインキャッシュ全体の面積は 231mm^2 となり、 15.2mm 角のチップとなる。純粋なメモリセル 4608Kb 分の面積は 96.0mm^2 であるので、パイプラインキャッシュ全体の面積の 41.58% をメモリセルが占めることになる。

ここで求めたパイプラインキャッシュ全体の面積には、クロックの制御回路の面積および出力バッファの面積は含まれていない。

考察

パイプライン化されていない通常のキャッシュは、パイプラインキャッシュにラッチがなく、またタグの比較器が各メモリセルアレイごとに必要なく、各ブロックに一つあれば良いことになる。よって、表 3.8 に各ステージのラッチを抜いた面積および比較器 4 個分の面積を示す。但し、入力ラッチだけは含む。

表 3.8: 各ステージごとの面積 (通常のキャッシュ)

	面積 [mm^2]
デコードステージ 1	0.0472
デコードステージ 2	0.0857
デコードステージ 3	6.48
タグメモリアクセスステージ	29.8
データメモリアクセスステージ	146
出力パスステージ 1	15.4
出力パスステージ 2	0.331
マルチプレクサステージ	0.0269
比較器 4 個分の面積	0.00185

上記の各部分の面積をすべて加えると 198mm^2 となる。よって、パイプラインキャッシュ (面積 231mm^2) は通常のキャッシュより 32.9mm^2 大きくなる。これはメモリセル 1.6Mb 分である。また、パイプライン化による面積の増加率は 1.17 である。

本研究で設計したメモリセルは 6 トランジスタ型で $4.4 \times 4.625\mu\text{m}^2$ と比較的大きめのセルである。メモリセルの大きさが小さくなった場合、通常のキャッシュとパイプラインキャッシュは同じ比率で面積が小さくなるため、パイプライン化による面積の増加率も今回導き出した結果より大きくなると考えられる。

第 4 章

SPICE シミュレーション

本章では、前章で設計を行なったパイプラインキャッシュのステージごとのシミュレーション結果を示す。まず、シミュレーションで用いる MOSFET モデルを決定し、pMOS および nMOS の電流-電圧特性を求めた。また、評価を行なうために必要な H レベル入力電圧と L レベル入力電圧を規定した。

4.1 MOSFET モデル

4.1.1 パラメータの設定

表 4.1,4.2を利用し、 $0.25\mu\text{m}$, $0.10\mu\text{m}$ MOSFET モデルのパラメータを電界一定のスケールリングにより求めた [6][7]。本研究で使用する MOSFET モデルを表 4.3に示す。

表 4.1: $0.5\mu\text{m}$ デザインルールの各種パラメータと移動度の標準値

	nMOSFET	pMOSFET
ゲート酸化膜厚 $t_{ox}(\text{nm})$	10	10
ゲート酸化膜容量 $C_o(fF/\mu\text{m}^2)$	3.38	3.38
不純物濃度 $N_A, N_D(\text{cm}^{-3})$	10^{-16}	10^{-16}
移動度 $(\text{cm}^2/V \cdot \text{sec})$	1350	480

表 4.2: 4 μm デザインルールの拡散容量値

	n+ 拡散容量	p+ 拡散容量
単位ゼロバイアスパルク 接合底面の容量 C_j	$1 \times 10^{-4} pF/\mu m^2$	$1 \times 10^{-4} pF/\mu m^2$
単位ゼロバイアスパルク 接合側面の容量 C_{jsw}	$9 \times 10^{-4} pF/\mu m$	$8 \times 10^{-4} pF/\mu m$

表 4.3: MOSFET モデル

	4 μm	0.5 μm	0.25 μm	0.10 μm
$N_{sub}[cm^{-3}]$		1E17	2E17	5E17
$C_{gso}[F/m]$		169p	169p	169p
$C_{gdo}[F/m]$		169p	169p	169p
$C_j[F/m^2]$	1E-4		1E-4	1E-4
$C_{jsw}(nMOS)[F/m]$	9E-10		0.5625E-10	0.225E-10
$C_{jsw}(pMOS)[F/m]$	8E-10		0.5E-10	0.2E-10
$L_d[m]$		5E-8	2.5E-8	1E-8
$P_b[V]$		0.75	0.75	0.75
$T_{ox}[m]$		10n	5n	2n
$U_o(nMOS)[cm^2/V \cdot s]$		1350	1350	1350
$U_o(pMOS)[cm^2/V \cdot s]$		480	480	480
$V_{max}[m/s]$			1E5	1E5
$X_j[m]$	9 ~ 8E-6	1.125 ~ 1E-6	0.5625 ~ 0.5E-6	0.225 ~ 0.2E-6

$N_{sub}[cm^{-3}]$	基板不純物濃度
$C_{gso}[F/m]$	単位チャンネル幅あたりのゲート・ソースオーバーラップ容量
$C_{gdo}[F/m]$	単位チャンネル幅あたりのゲート・ドレインオーバーラップ容量
$C_j[F/m^2]$	単位接合面積あたりのゼロバイアスパルク接合底面の容量
$C_{jsw}[F/m]$	単位接合周囲長あたりのゼロバイアスパルク接合側面の容量
$L_d[m]$	ゲートと拡散層とのオーバーラップ長
$P_b[V]$	バルク接合電位
$T_{ox}[m]$	ゲート酸化膜厚
$U_o[cm^2/V \cdot s]$	表面移動度
$V_{max}[m/s]$	キャリアの最大ドリフト速度
$X_j[m]$	ドレイン・ソース拡散深さ

4.1.2 特性

0.25 μm テクノロジーにおける nMOS および pMOS の電流-電圧特性を示す。

nMOSFET

上図は $I_{\text{ds}}-V_{\text{ds}}$ 特性、下図は $I_{\text{ds}}-V_{\text{gs}}$ 特性である。ゲート長は 0.25 μm 、ゲート幅は 10 μm とした。

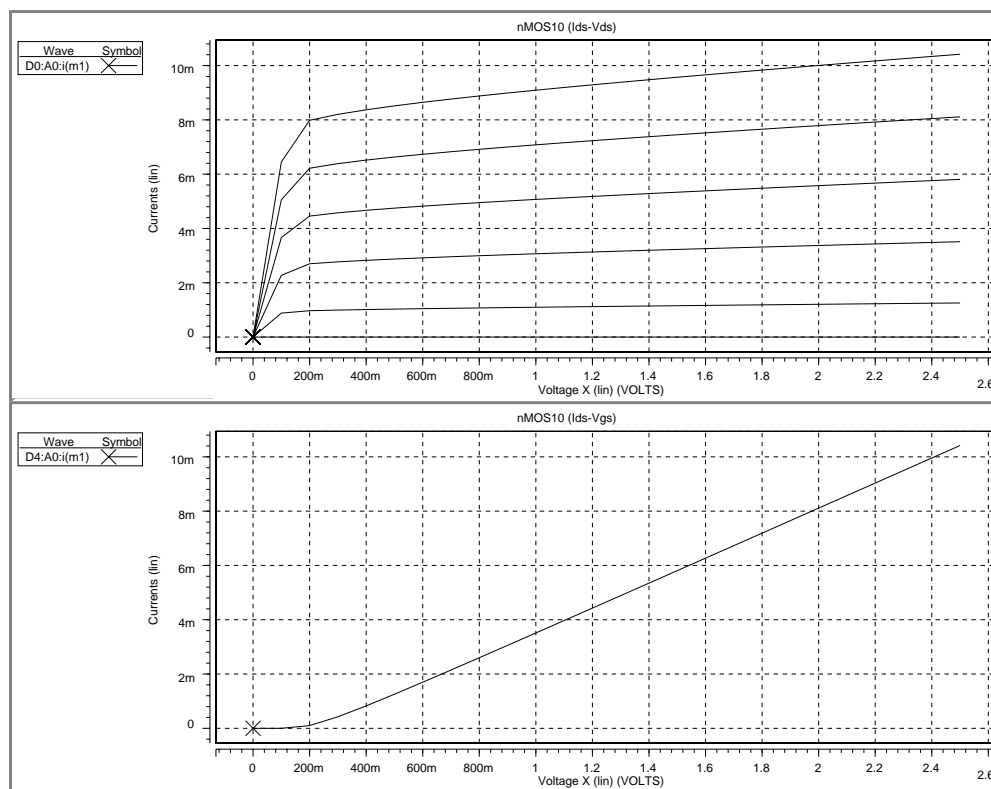


図 4.1: nMOSFET の電流-電圧特性

上図の V_{gs} の電圧は、上から 2.5V, 2.0V, 1.5V, 1.0V, 0.5V, 0V である。下図の V_{ds} の電圧は 2.5V である。

pMOSFET

上図は I_{ds} - V_{ds} 特性、下図は I_{ds} - V_{gs} 特性である。ゲート長は $0.30\mu m$ 、ゲート幅は $10\mu m$ とした。

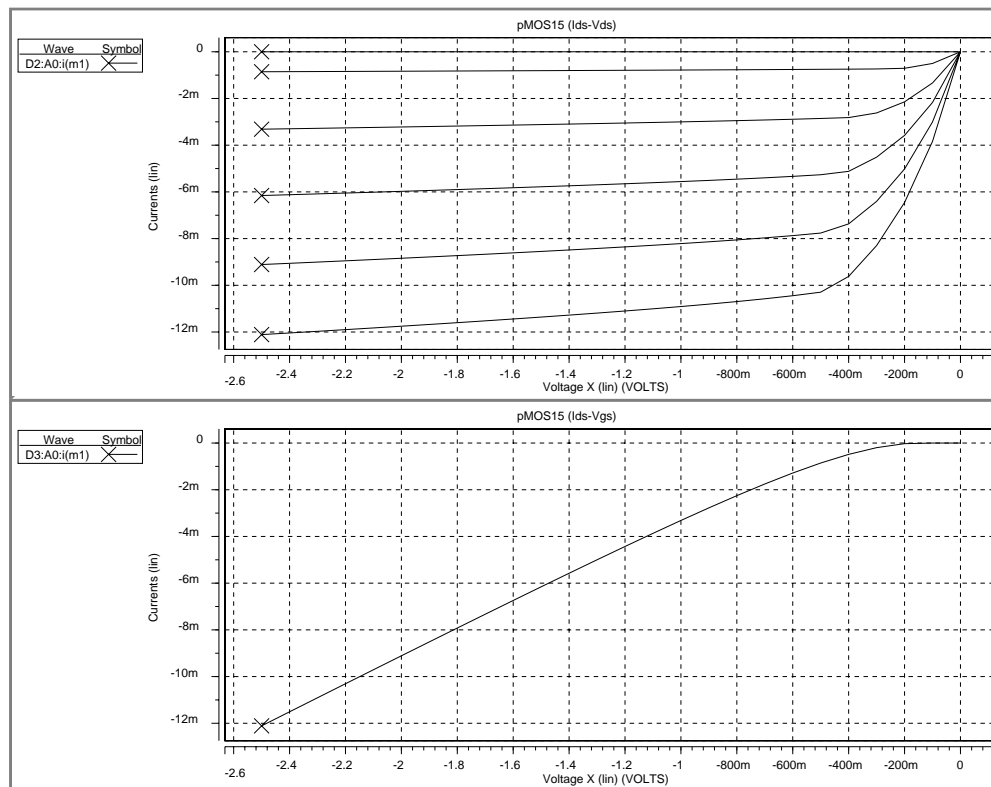


図 4.2: pMOSFET の電流-電圧特性

上図の V_{gs} の電圧は、上から 0V、-0.5V、-1.0V、-1.5V、-2.0V、-2.5V である。下図の V_{ds} の電圧は -2.5V である。

4.2 H レベルと L レベルの規定

設計や評価を行なうために、H レベル入力電圧と L レベル入力電圧を規定する必要がある。

本研究では、一般的な CMOS の IC の規格に従い、電源電圧の 30 % 以下を L レベル、70 % 以上を H レベルと規定する。

表 4.4: H レベルと L レベルの規定

	電源電圧	
	2.5V (0.25 μ m 時)	1.0V (0.10 μ m 時)
H レベル	1.75V	0.7V
L レベル	0.75V	0.3V

本研究で使用する伝播時間（遅延時間）は、求めたい回路の入力電圧が上記の電圧に達した時間から、その回路の出力電圧が上記の電圧に達するまでの時間差とする（図 4.3 参照）。

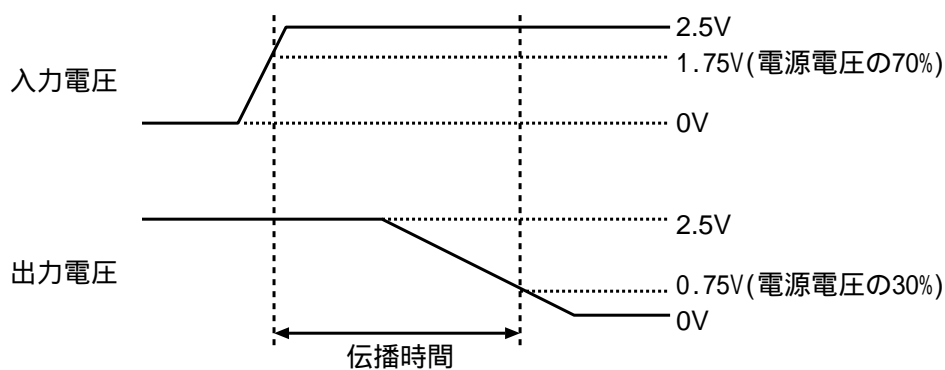


図 4.3: H レベルと L レベルの規定

4.3 ラッチの動作タイミング

本研究のラッチは、立ち下がりのエッジ・トリガ方式を用いる。動作タイミングを図 4.4 に示す。

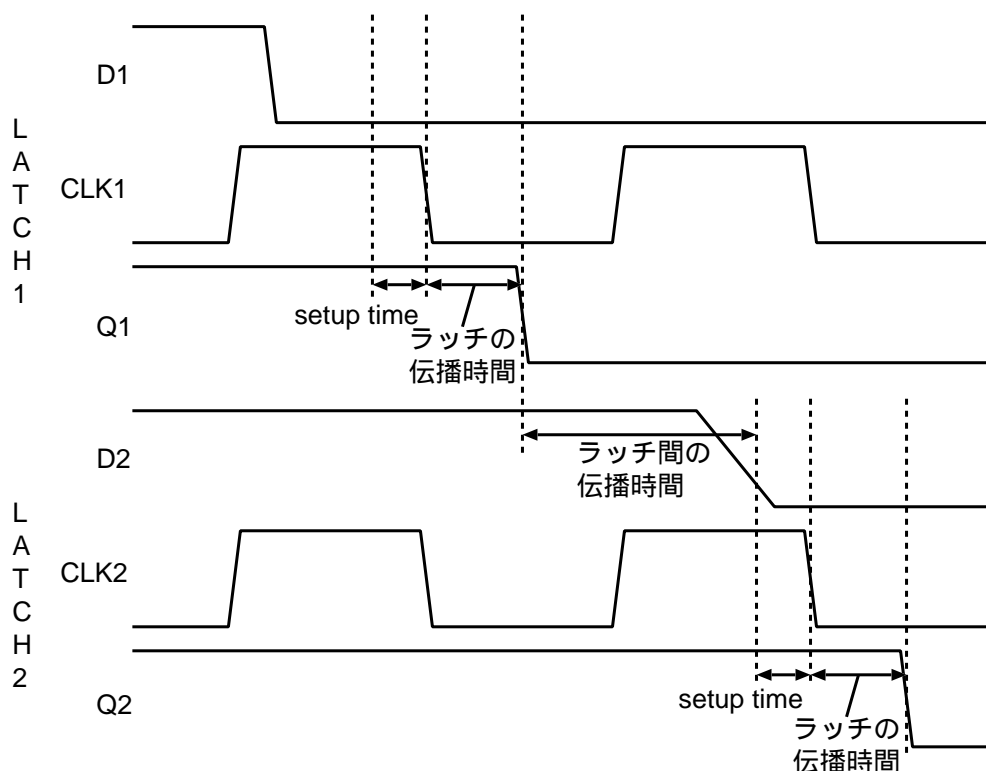


図 4.4: ラッチの動作タイミング

図 4.4の動作タイミングを持つラッチを使用したシステムでは、クロックサイクル時間は下記の 3 つの遅延時間の最悪値の合計以上でなければならない。

$$\text{ラッチの伝播時間} + \text{ラッチ間の伝播時間} + \text{setup time}$$

ここでは、ホールド時間の要件は満足しているものと仮定している。通常的设计では、ホールド時間の要件が問題になることはない。ラッチの伝播時間の方がホールド時間よりも常に長いからである。

前章で説明したように、本研究では大きさの異なる 3 種類のラッチを使用した (図 B.2 参照)。表 4.5 にそれぞれのラッチの setup time のシミュレーション結果を示す。

表 4.5: ラッチの setup time ($0.25\mu m$)

ラッチの大きさ	setup time [psec]	
	H レベル入力時	L レベル入力時
latch-ss	28	28
latch-s	25	24
latch-m	24	23

表 4.6: ラッチの setup time ($0.10\mu m$)

ラッチの大きさ	setup time [psec]	
	H レベル入力時	L レベル入力時
latch-ss	11	12
latch-s	10	10
latch-m	9	10

各ステージの最小クロックサイクル時間を求める際は、表 4.5, 4.6の値をラッチの setup time として使用する。

4.4 メモリセルアレイの動作タイミング

メモリセルアレイには、 $128 \times 16\text{bit}$ のタグメモリと $128 \times 128\text{bit}$ のデータメモリがある。2つのメモリの違いは、ワード線の長さがタグメモリは16メモリセル分、データメモリは128メモリセル分であることである。ビット線の長さは、128メモリセル分と同じである。

本設計では、ビット線にプリチャージ回路を導入しなかった。理由は、パイプライン動作のため、通常のSRAMのように他の処理をしている時にプリチャージを行なうことができず、プリチャージ時間がそのままステージの処理時間に加わるからである。また、 128×128 の大きさのメモリセルアレイではプリチャージによる利益がそれほど得られないからである。プリチャージに関する議論については、考察で行なう。

4.4.1 タグメモリアクセスステージ

読み出しおよび書き込み動作の処理時間

読み出し動作 図4.5に読み出し動作時の各信号のタイミングを示す。この図はHレベルのデータが読み出された時のタイミングである。

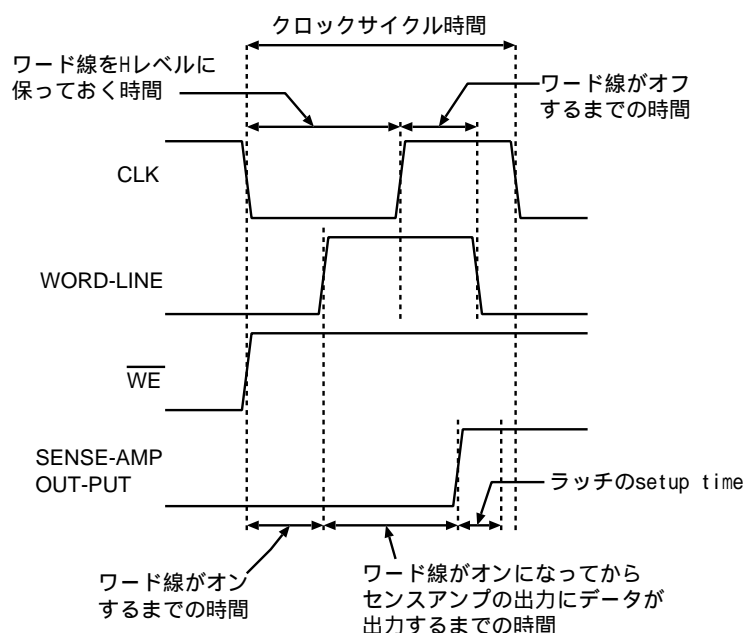


図 4.5: 読み出し動作

読み出し動作の処理時間は次式のようになる。

$$\text{読み出し動作の処理時間} = \text{ワード線がオンするのに必要な時間} + \text{ワード線がオンしてからデータが出力するまでの時間} + \text{ラッチの setup time}$$

表 4.7,4.8にシミュレーション結果を示す。

表 4.7: タグメモリ読み出し動作の処理時間 ($0.25\mu m$)

		ワード線オン [psec]	データ出力 [psec]	setup time [psec]	合計 [psec]
タグメモリ	H	83.8	193	25	302
	L	83.8	168	24	276

表 4.8: タグメモリ読み出し動作の処理時間 ($0.10\mu m$)

		ワード線オン [psec]	データ出力 [psec]	setup time [psec]	合計 [psec]
タグメモリ	H	33	69.2	10	113
	L	33	62.0	10	105

$0.25\mu m$ の読み出し処理時間は $302[\text{psec}]$ 、 $0.10\mu m$ の読み出し処理時間は $113[\text{psec}]$ となる。

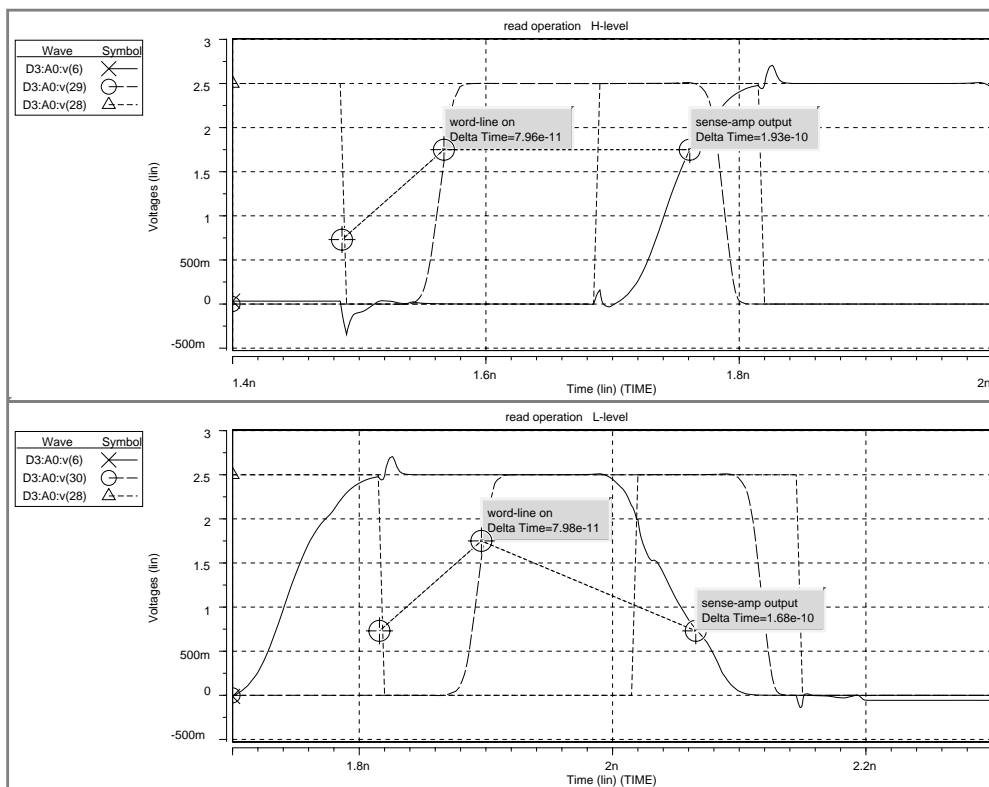


図 4.6: タグメモリの読み出し動作 ($0.25\mu m$)

書き込み動作 図 4.7に書き込み動作時の各部のタイミングを示す。この図はHレベルのデータを書き込む時のタイミングである。

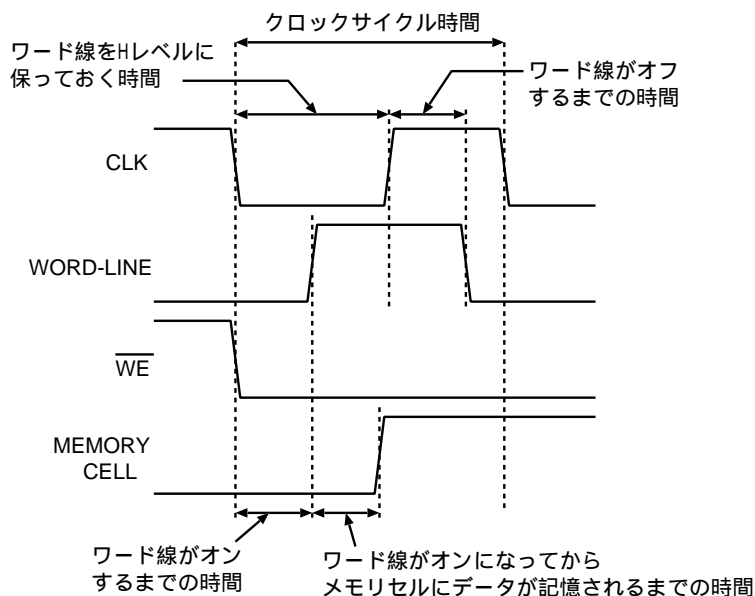


図 4.7: 書き込み動作

書き込み動作の処理時間は次式のようになる。

書き込み動作の処理時間 = ワード線がオンするのに必要な時間 + ワード線がオンしてからデータが書き込まれるまでの時間

表 4.9,4.10 にシミュレーション結果を示す。

表 4.9: タグメモリ書き込み動作の処理時間 (0.25 μ m)

		ワード線オン [psec]	データ書き込み [psec]	合計 [psec]
タグメモリ	H	83.8	92.9	177
	L	83.8	107	191

表 4.10: タグメモリ書き込み動作の処理時間 ($0.10\mu m$)

		ワード線オン [psec]	データ書き込み [psec]	合計 [psec]
タグメモリ	H	33	34.5	67.5
	L	33	44.7	77.7

$0.25\mu m$ の書き込み処理時間は $191[psec]$ 、 $0.10\mu m$ の書き込み処理時間は $77.7[psec]$ となる。

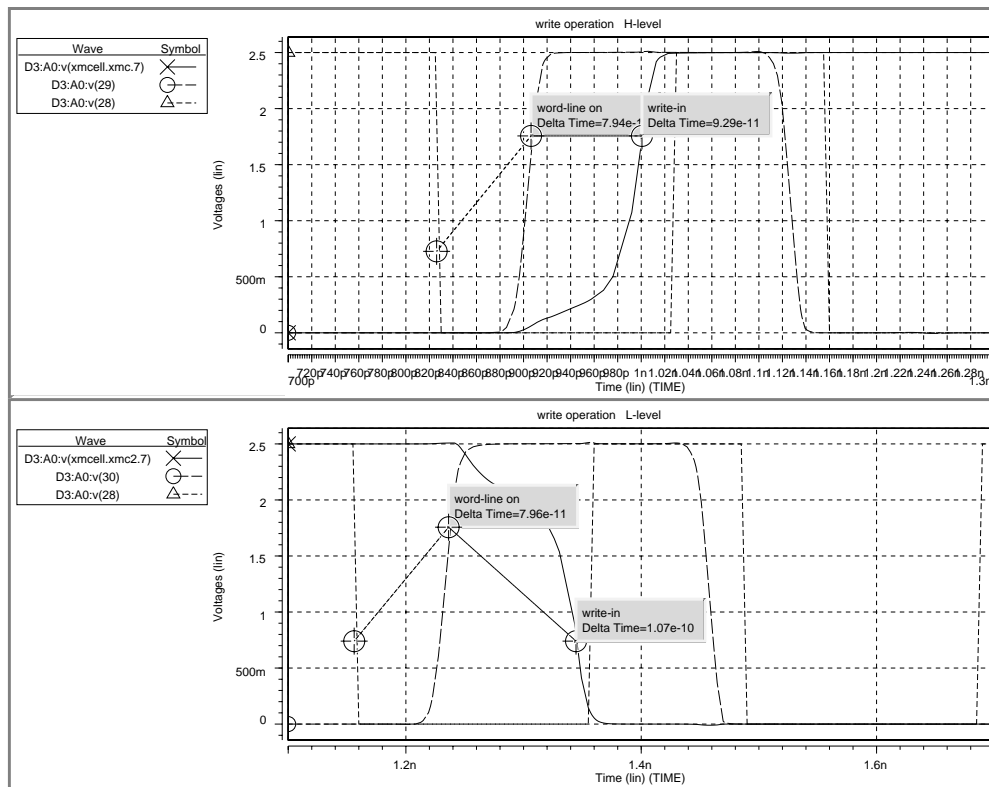


図 4.8: タグメモリの読み出し動作 ($0.25\mu m$)

クロックサイクル時間の決定

クロックサイクル時間は、読み出しおよび書き込みの処理時間をそのままクロックサイクル時間とすることができず、クロックの立ち上がり立ち下りのタイミングを考える必要がある。

- クロックの立ち下りから立ち上がりまでの時間でワード線を H レベルに保っておく時間を決めている。

ワード線を H レベルに保っておかなければならない時間は、ワード線がオンとなってからデータが読み出される時間またはデータが書き込まれる時間である。表 4.7,4.9 のシミュレーション結果より、この時間は 193[psec] 以上としなければならない。0.10 μ m では、表 4.8,4.10 より、69.2[psec] 以上としなければならない。

- クロックの立ち上がりから立ち下りまでの時間は、ワード線をオフするのに必要な時間とラッチの setup time 以上取らなければならない。

ワード線をオフするのに必要な時間を表 4.11,4.12 に示す。

表 4.11: ワード線のオン/オフ時間 (0.25 μ m)

	ワード線がオンする までの時間 [psec]	ワード線がオフする までの時間 [psec]
タグメモリ	83.8	76

表 4.12: ワード線のオン/オフ時間 (0.10 μ m)

	ワード線がオンする までの時間 [psec]	ワード線がオフする までの時間 [psec]
タグメモリ	33	29.8

表 4.11,4.12 の結果にラッチの setup time を加え、0.25 μ m では 101[psec] 以上、0.10 μ m では 39.8[psec] 以上取らなければならない。

上記のシミュレーション結果に余裕を見て、0.25 μ m ではクロックの L レベルの時間を 200[psec]、H レベルの時間を 110[psec] とする。0.10 μ m ではクロックの L レベルの時間

を 70[psec]、H レベルの時間を 50[psec] とする。よって、タグメモリアクセスステージのクロックサイクル時間は $0.25\mu m$ では 310[psec]、 $0.10\mu m$ では 120[psec] となる。

4.4.2 データメモリアクセスステージ

読み出しおよび書き込み動作の処理時間

読み出し動作 各信号のタイミングについてはタグメモリアクセスステージの所で示した。(図 4.5 参照)

読み出し動作の処理時間は次式のようになる。

$$\text{読み出し動作の処理時間} = \text{ワード線がオンするのに必要な時間} + \text{ワード線がオンしてからデータが出力するまでの時間} + \text{ラッチの setup time}$$

表 4.13, 4.14 にシミュレーション結果を示す。

表 4.13: データメモリ読み出し動作の処理時間 ($0.25\mu m$)

		ワード線オン [psec]	データ出力 [psec]	setup time [psec]	合計 [psec]
データメモリ	H	110	188	25	323
	L	110	160	24	294

表 4.14: データメモリ読み出し動作の処理時間 ($0.10\mu m$)

		ワード線オン [psec]	データ出力 [psec]	setup time [psec]	合計 [psec]
データメモリ	H	44.6	67.0	10	122
	L	44.6	59.9	10	115

$0.25\mu m$ の読み出し処理時間は 323[psec]、 $0.10\mu m$ の読み出し処理時間は 122[psec] となる。

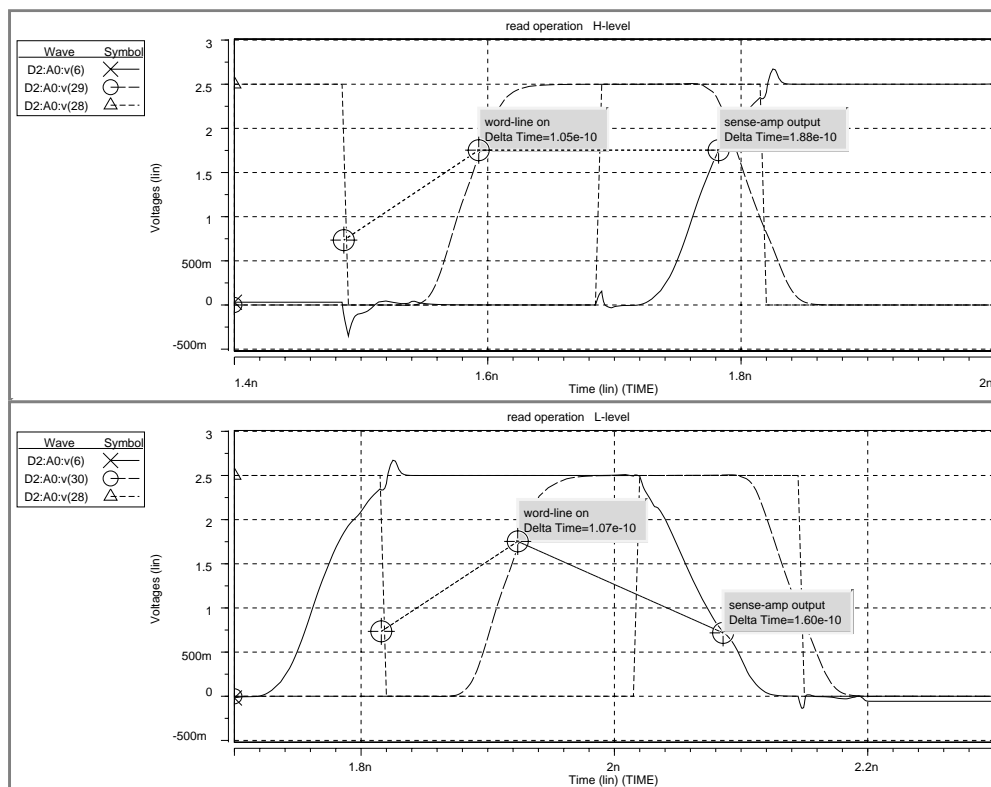


図 4.9: データメモリの読み出し動作 (0.25 μ m)

書き込み動作 各信号のタイミングについてはタグメモリアクセスステージの所で示した(図 4.7参照)

書き込み動作の処理時間は次式のようになる。

書き込み動作の処理時間 = ワード線がオンするのに必要な時間+ ワード線がオンしてからデータが書き込まれるまでの時間

表 4.15,4.16 にシミュレーション結果を示す。

表 4.15: データメモリ書き込み動作の処理時間 (0.25 μ m)

		ワード線オン [psec]	データ書き込み [psec]	合計 [psec]
データメモリ	H	110	111	221
	L	110	106	216

表 4.16: データメモリ書き込み動作の処理時間 (0.10 μ m)

		ワード線オン [psec]	データ書き込み [psec]	合計 [psec]
データメモリ	H	44.6	45.0	89.6
	L	44.6	43.5	88.1

0.25 μ m の書き込み処理時間は 221[psec]、0.10 μ m の書き込み処理時間は 89.6[psec] となる。

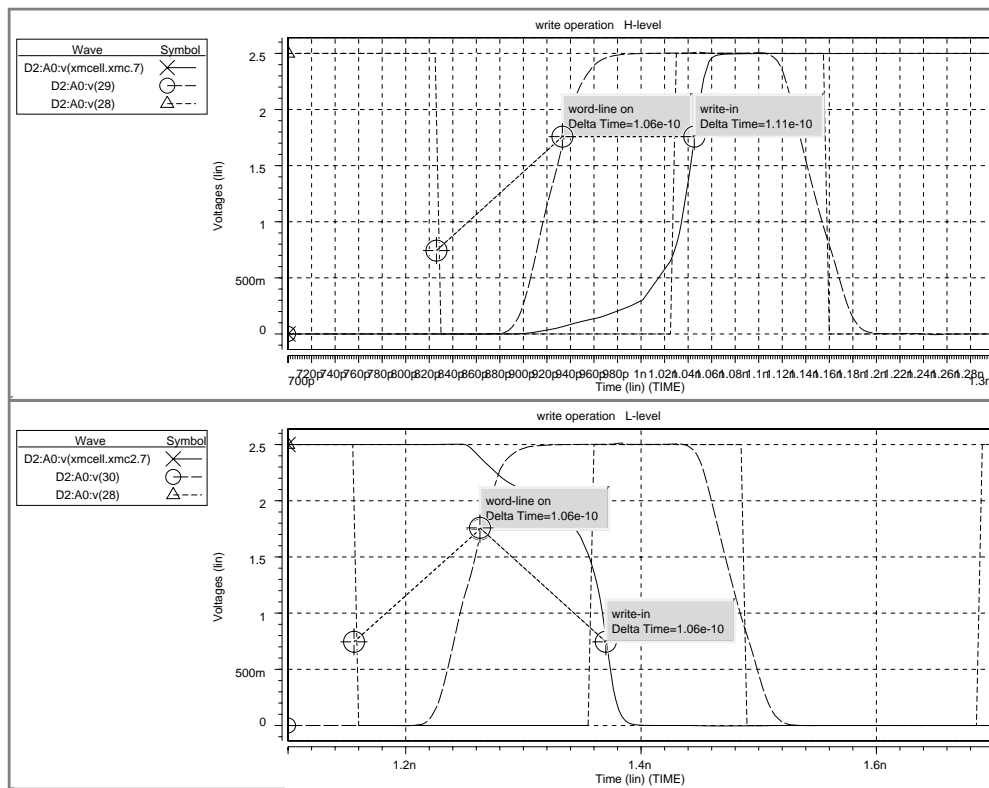


図 4.10: データメモリの書き込み動作 ($0.25\mu m$)

クロックサイクル時間の決定

タグメモリアクセスステージの所でも述べたように、読み出しおよび書き込みの処理時間をそのままクロックサイクル時間とすることができず、クロックの立ち上がり立ち下りのタイミングを考える必要がある。

- クロックの立ち下りから立ち上がりまでの時間でワード線を H レベルに保っておく時間を決めている。

ワード線を H レベルに保っておかなければならない時間は、ワード線がオンとなってからデータが読み出される時間またはデータが書き込まれる時間である。表 4.13, 4.15 のシミュレーション結果より、この時間は $188[\text{psec}]$ 以上としなければならない。 $0.10\mu m$ では、表 4.14, 4.16 より、 $67.0[\text{psec}]$ 以上としなければならない。

- クロックの立ち上がりから立ち下がりまでの時間は、ワード線をオフするのに必要な時間とラッチの setup time 以上取らなければならない。

ワード線をオフするのに必要な時間を表 4.17,4.18 に示す。

表 4.17: ワード線のオン/オフ時間 ($0.25\mu m$)

		ワード線がオンする までの時間 [psec]	ワード線がオフする までの時間 [psec]
データメモリ	最小時間	104	101
	最大時間	110	108

表 4.18: ワード線のオン/オフ時間 ($0.10\mu m$)

		ワード線がオンする までの時間 [psec]	ワード線がオフする までの時間 [psec]
データメモリ	最小時間	39	38.5
	最大時間	44.6	43.7

表 4.17,4.18の結果にラッチの setup time を加え、 $0.25\mu m$ では 133[psec] 以上、 $0.10\mu m$ では 53.7[psec] 以上取らなければならない。

上記のシミュレーション結果に余裕を見て、 $0.25\mu m$ ではクロックの L レベルの時間を 190[psec]、H レベルの時間を 140[psec] とする。 $0.10\mu m$ ではクロックの L レベルの時間を 70[psec]、H レベルの時間を 60[psec] とする。よって、データメモリアクセスステージのクロックサイクル時間は $0.25\mu m$ では 330[psec]、 $0.10\mu m$ では 130[psec] となる。

4.5 各ステージのシミュレーション

4.5.1 シミュレーション方法

デコードステージ

1 ステージ目 アドレスの入力ラッチからチップ中央に位置するラッチまでの部分である（図 B.4 参照）。このステージはラッチ間に組合せ論理回路がなく、長いアドレスパスにおける伝播時間のみである。そのためラッチの出力端から入力端までを通してシミュレーションを行なった。

2 ステージ目 ラッチから中間デコーダを通り次のラッチまでの部分である（図 B.6 参照）。このステージのラッチ間には、長いアドレスパスとデコーダがある。このアドレスパスの配線抵抗と配線容量により波形がなまるため、アドレスパスの出力端までの部分と、そこからデコーダを通り次のラッチまでの部分に 2 つに分けてシミュレーションを行なった。

3 ステージ目 ラッチからローカルデコーダを通り次のラッチまでの部分である（図 B.8 参照）。このステージのラッチ間には、長いアドレスパスとデコーダがある。2 ステージ目と同様、このアドレスパスの配線抵抗と配線容量により波形がなまるため、アドレスパスの出力端までの部分と、そこからデコーダを通り次のラッチまでの部分に 2 つに分けてシミュレーションを行なった。

タグメモリアクセスステージ

本ステージのシミュレーション方法は、まずワード線とビット線とに分けてシミュレーションを行ない、最小クロックサイクルの目安をつけた。そして、前節で求めた読み出しおよび書き込みの動作タイミングに従ってステージ全体を通したシミュレーションを行ない、その時に正常な動作が検証されたクロックサイクル時間を本ステージの最小クロックサイクル時間とした。

タグ比較ステージ

本ステージには組合せ論理回路しかなく、ラッチの出力端から入力端までを通してシミュレーションを行なった（図 B.13 参照）。

データメモリアクセスステージ

本ステージのシミュレーション方法は、タグメモリアクセスステージと同様である。まず、ワード線とビット線とに分けてシミュレーションを行ない、最小クロックサイクルの目安をつけた。そして、前節で求めた読み出しおよび書き込みの動作タイミングに従ってステージ全体を通したシミュレーションを行ない、その時に正常な動作が検証されたクロックサイクル時間を本ステージの最小クロックサイクル時間とした。

出力パスステージのシミュレーション

1 ステージ目 メモリセルアレイの出力ラッチから中間ラッチまでの部分である（図 B.19 参照）。このステージのラッチ間には、長いデータパスと組合せ論理回路がある。このデータパスの配線抵抗と配線容量により波形がなまるため、データパスの出力端で2つの部分に分けてシミュレーションを行なった。

2 ステージ目 中間ラッチから判定部の入力ラッチまでの部分である（図 B.20参照）。このステージはラッチ間に組合せ論理回路がなく、長いデータパスにおける伝播時間のみである。そのためラッチの出力端から入力端までを通してシミュレーションを行なった。

マルチプレクサステージのシミュレーション

本ステージには組合せ論理回路しかなく、ラッチの出力端から入力端までを通してシミュレーションを行なった（図 B.21参照）。

4.5.2 シミュレーション結果

表 4.19,4.20に $0.25\mu m$ と $0.10\mu m$ の各ステージのシミュレーション結果を示す。それぞれのステージにおいて、ステージの出力が H レベルとなる最大時間、L レベルとなる最大時間を求めている。各ステージにおける最小クロックサイクル時間は、この時間の大きい方となる。表 4.19,4.20より得られた各ステージの最小クロックサイクル時間をグラフ化した図を図 4.11に示す。

$0.10\mu m$ に微細化時のシミュレーションは、 $0.25\mu m$ のパラメータをすべて電界一定の比例縮小により行なった (MOSFET モデルの節 参照)。しかし、メモリセルは配線容量や抵抗の変化によりうまく動作しなかったため、図 B.12,B.18のメモリセルのパストランジスタのゲート幅だけを $0.35\mu m$ に変更した。

表 4.19: 各ステージのシミュレーション結果 ($0.25\mu m$)

		ラッチの 伝播時間 [psec]	配線部の 伝播時間 [psec]	論理回路部 の伝播時間 [psec]	ラッチの setup time [psec]	合計 [psec]
デコードステージ 1	H	41.7	108	/	24	174
	L	44.1	104	/	23	171
デコードステージ 2	H	41.7	105	120	25	292
	L	43.4	106	108	24	282
デコードステージ 3	H	40.6	72.6	192	28	333
	L	40.7	69.3	171	28	309
タグメモリアクセス						310
タグ比較	H	31.8	/	113	25	170
	L	31.4	/	259	24	315
データメモリアクセス						330
出力パス 1	H	46.0	174	6.05	25	251
	L	43.5	172	9.28	24	249
出力パス 2	H	74.7	161	/	25	261
	L	100	139	/	24	263
MUX	H	39.2	/	143	24	207
	L	58	/	124	23	205

表 4.20: 各ステージのシミュレーション結果 (0.10 μ m)

		ラッチの 伝播時間 [psec]	配線部の 伝播時間 [psec]	論理回路部 の伝播時間 [psec]	ラッチの setup time [psec]	合計 [psec]
デコードステージ 1	H	15.3	112	/	9	137
	L	15.4	111	/	10	137
デコードステージ 2	H	14.7	80.1	41.7	10	147
	L	15.0	83.5	38.9	10	148
デコードステージ 3	H	14.1	53.2	67.6	11	146
	L	13.3	52.9	57.7	12	136
タグメモリアクセス						120
タグ比較	H	12.8	/	47.5	10	70.3
	L	12.5	/	96.6	10	120
データメモリアクセス						130
出力パス 1	H	18.2	116	2.51	10	147
	L	16.8	121	3.66	10	152
出力パス 2	H	24.8	117	/	10	152
	L	28.8	109	/	10	148
MUX	H	14.7	/	50.0	9	73.7
	L	22.0	/	43.6	10	75.6

4.5.3 まとめ

図 4.11 に $0.25\mu\text{m}$ および $0.10\mu\text{m}$ の各ステージの最小クロックサイクル時間をまとめたグラフを示す。

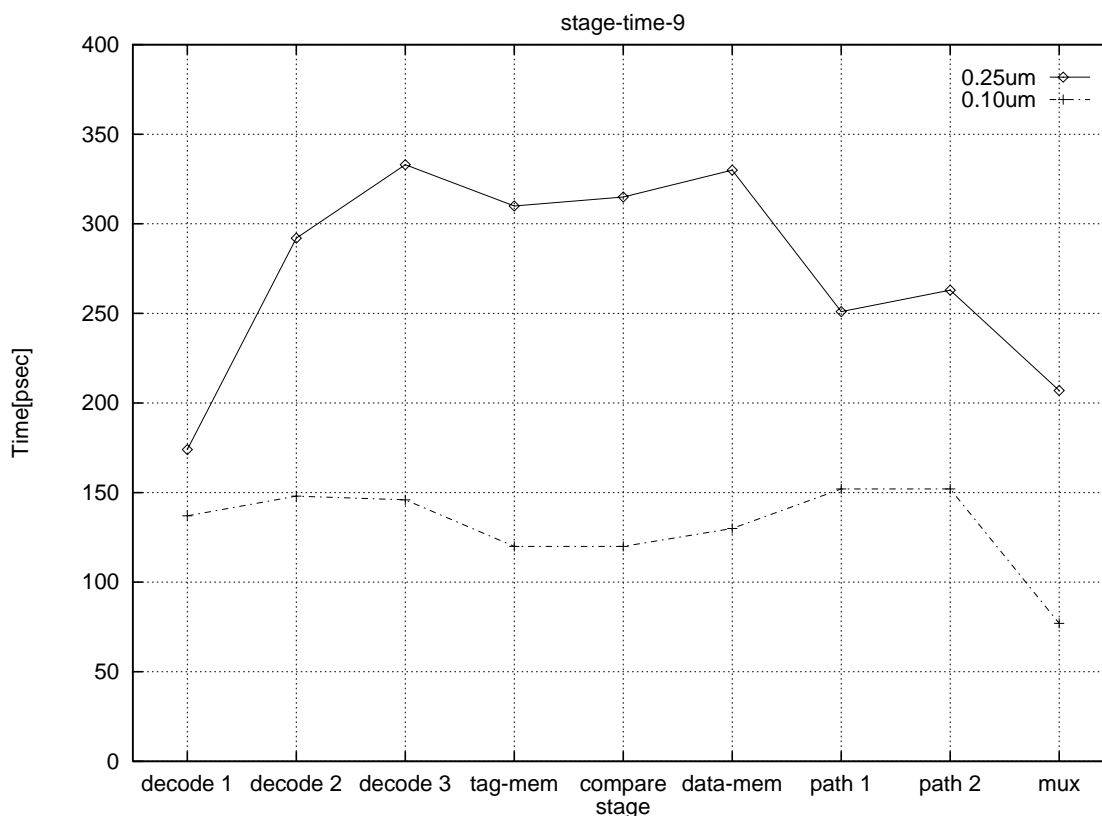


図 4.11: 各ステージの最小クロックサイクル時間

0.25 μm

最小クロックサイクル時間が最大となるステージは、デコードステージ3の $333[\text{psec}]$ およびデータメモリアクセスステージの $330[\text{psec}]$ である。よって、9段構成のパイプラインキャッシュでは 3GHz で動作可能である。

現在研究されている $0.25\mu\text{m}$ テクノロジーのパイプライン化されていないSRAMの動作周波数は 500MHz 程度であり、9段にパイプライン化することにより6倍に高速化することがわかる [14][15]。

本設計で使用した $128 \times 128\text{bit}$ のメモリセルアレイをさらに分割し、パイプラインのス

テージ数を増やしたとしてもそれほど高速化されない。その理由については、メモリセルアレイの分割に関する考察の所で論じる。

0.10 μm

0.25 μm から 0.10 μm に微細化することにより、タグ比較ステージやマルチプレクサステージのように論理回路の遅延のみが影響するステージは、電界一定の比例縮小則にしたがい 2.5 倍程度の高速化ができた。しかし、配線遅延が影響するステージはそれほど高速化が達成されなかった。特に、ラッチと配線のみで構成されているデコードステージ 1 は 1.27 倍の高速化しか達成できなかった。

本構成では、0.10 μm に微細化しても突出して遅くなるステージが現れなかったが、メモリセルアレイにアクセスするステージが他のステージよりも高速に動作が行なえるため、他のステージをさらに分割することによりさらに高速化が可能である。

9 段構成のままだと動作周波数は 6.5GHz であるが、メモリセルアレイにアクセスするステージ以外をさらに分割することにより、7.7GHz の動作周波数が実現できる。

第 5 章

考察

5.1 メモリセルアレイの分割に関する考察

1つのメモリセルアレイの大きさを $64 \times 64bit$ 、 $128 \times 128bit$ 、 $256 \times 256bit$ と変化させ、読み出しおよび書き込みにかかる時間を求めた。

シミュレーションを行なった回路

$64 \times 64bit$ および $256 \times 256bit$ のメモリセルアレイのシミュレーションを行なった回路は、 $128 \times 128bit$ のメモリセルアレイの回路図 (図 B.16, B.17, B.18) をそれぞれ次のように変更した。

- $64 \times 64bit$ のメモリセルアレイ

ビット線、ワード線の長さをそれぞれ $1/2$ 。
メモリセルの数を $1/2$ 。

- $256 \times 256bit$ のメモリセルアレイ

ビット線、ワード線の長さをそれぞれ 2 倍。
メモリセルの数を 2 倍。

図 B.17の入力ラッチを latch-s に変更。

また、同図の WRITE 回路のインバータを $30/20$ 、パストランジスタを 22.5 に変更。

読み出しおよび書き込みの処理時間の求め方

読み出し動作および書き込み動作の処理時間は次式のようになる。

読み出し動作の処理時間 = ワード線がオンするのに必要な時間 + データが出力するまでの時間 + ラッチの setup time

書き込み動作の処理時間 = ワード線がオンするのに必要な時間 (またはビット線の電位がある程度変化するのに必要な時間) + データがメモリセルに記憶されるまでの時間

書き込み動作時は、ビット線の電位の変化とワード線の電位の変化が並行して行なわれており、 $256 \times 256bit$ のメモリセルアレイでは次のような現象が起こる。ビット線の長さが2倍になるため、ビット線の電位が十分変化しないうちにワード線がオンし、メモリセルに間違った値が書き込まれる。したがって、 $256 \times 256bit$ メモリセルアレイの書き込み動作の処理時間を求める場合は上式の括弧内の時間を使用した。

シミュレーション結果

3種類のメモリセルアレイの読み出しおよび書き込みの処理時間を表 5.1に示す。また、グラフを図 5.1に示す。

表 5.1: メモリセルアレイの分割による処理時間の変化

	$64 \times 64bit$	$128 \times 128bit$	$256 \times 256bit$
読み出しの処理時間 [psec]	271	323	491
書き込みの処理時間 [psec]	163	221	409

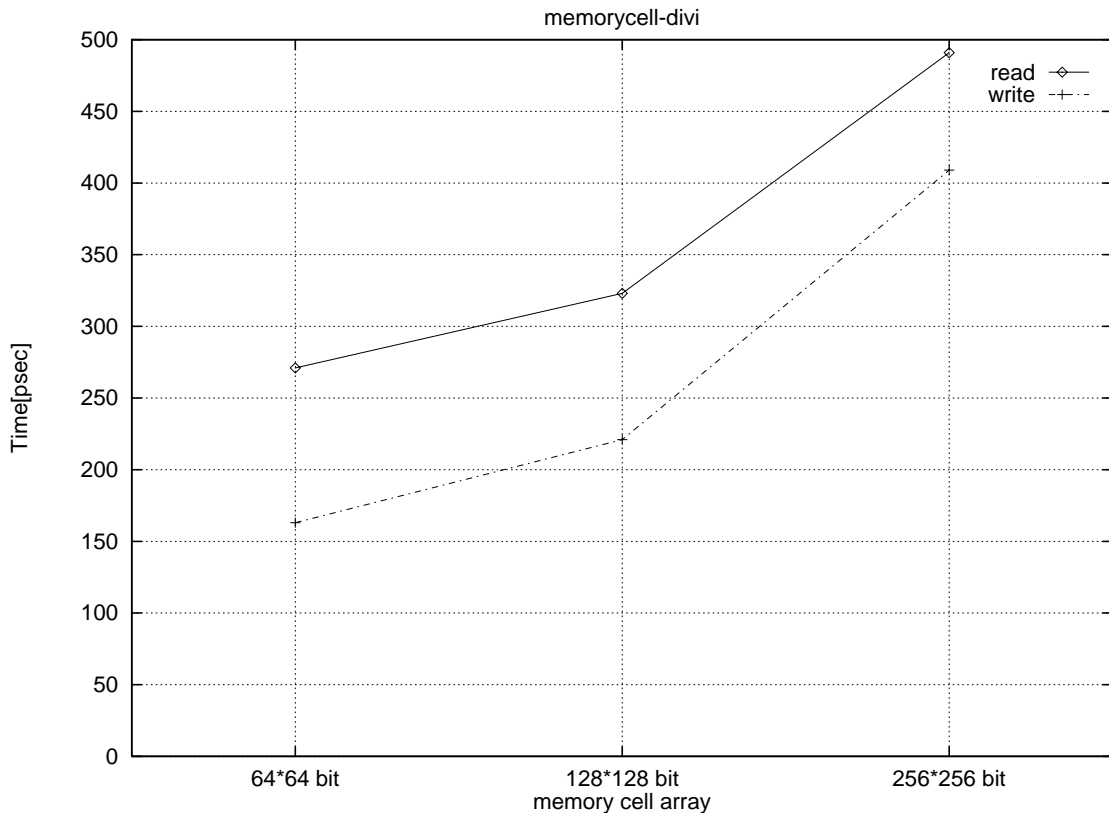


図 5.1: メモリセルアレイの分割による処理時間の変化

考察

ステージのクロックサイクル時間を決める読み出しの処理時間は、 $256 \times 256bit$ から $1/4$ の大きさの $128 \times 128bit$ に分割した場合 $170[psec]$ 程度短くなった。速度向上比は 1.52 となる。さらに $1/4$ の大きさの $64 \times 64bit$ に分割すると、 $50[psec]$ 程度短くなり、速度向上比は 1.19 となる。

$128 \times 128bit$ から $64 \times 64bit$ に分割した場合、 1.19 倍の速度向上にとどまった理由は、 $128 \times 128bit$ の大きさのメモリセルアレイの全遅延時間に占める配線遅延の割合がゲート遅延の割合より小さかったからであると考えられる。

$1/4$ の分割により速度は向上するが、ドライバやセンスアンプの数が 2 倍に増える。よって、 1.19 倍の速度向上では分割のメリットはあまりない。本研究では、 $1/4$ の分割により 1.52 倍の速度向上が達成できた $128 \times 128bit$ の大きさのメモリセルアレイを用いてデータメモリを設計した。

ところで、データメモリ以外にタグメモリの設計も必要となる。タグメモリに関しては、ワード線の長さをデータメモリの $1/8$ にした $128 \times 16bit$ のメモリセルアレイを使用した。ワード線の長さを $1/8$ にしたが、表 4.11, 4.17 を比較してもらえばわかるように 1.3 倍程度しか速度向上されなかった。これはワード線の遅延要因が、配線によるものよりもゲートによるものが支配的であったからであると考えられる。

本構成では、タグメモリの比較をデータメモリの近くで行わないために、 $128 \times 16bit$ のメモリセルアレイを使用することになった。ワード線をドライブするドライバーの数を減らすためには、タグメモリも $128 \times 128bit$ 構成の方が良いと考えられる。

5.2 プリチャージ回路に関する考察

$128 \times 128bit$ のデータメモリにおいて、プリチャージ回路を導入した時とプリチャージ回路を省いた時の、読み出しおよび書き込みにかかる時間の比較を行なった。プリチャージ回路を導入した時の回路図については、図 B.22 に示す。

読み出しおよび書き込みの処理時間の求め方

プリチャージ回路を導入した時の読み出しおよび書き込み動作の処理時間は以下のようになる。

読み出し動作の処理時間 = ワード線がオンするのに必要な時間 + データが出力するまでの時間 + ラッチの setup time + プリチャージ時間

書き込み動作の処理時間 = ワード線がオンするのに必要な時間 + データがメモリセルに記憶されるまでの時間 + プリチャージ時間

また、プリチャージ回路を省いた時の処理時間は以下のようになる。

読み出し動作の処理時間 = ワード線がオンするのに必要な時間 + データが出力するまでの時間 + ラッチの setup time

書き込み動作の処理時間 = ワード線がオンするのに必要な時間 + データがメモリセルに記憶されるまでの時間

シミュレーション結果

表 5.2,5.3 に、プリチャージ回路を導入した時と省いた時の読み出しおよび書き込みの処理時間を示す。

表 5.2: プリチャージ回路を導入した時の処理時間

		ワード線オン [psec]	データ出力/書き込み [psec]	setup time [psec]	プリチャージ [psec]	合計 [psec]
読み出し	H	110	51.0	25	100	286
	L	110	2.82	24	100	237
書き込み	H	110	52.7	/	100	263
	L	110	43.6	/	100	254

表 5.3: プリチャージ回路を省いた時の処理時間

		ワード線オン [psec]	データ出力/書き込み [psec]	setup time [psec]	合計 [psec]
読み出し	H	110	188	25	323
	L	110	160	24	294
書き込み	H	110	111	/	221
	L	110	106	/	216

考察

ステージのクロックサイクル時間を決める読み出しの処理時間は、プリチャージ回路を導入した時には 37[psec] 速くなり、1.13 倍の速度向上が可能である。しかし、プリチャージのオン/オフを制御するクロックが必要となり、現実にはクロックスキューのため 1.13 倍以下の速度向上になる考えられる。

ところで、1.13 倍の速度向上にとどまった理由は、 $128 \times 128bit$ という小さなメモリセルアレイのために、プリチャージ回路により抑えることができる配線遅延の割合が少なかったことが原因であると考えられる。プリチャージによる速度向上も、メモリセルアレイの分割と同様に全遅延時間に占める配線遅延とゲート遅延の割合が重要になってくる。

また、1.13 倍程度の速度向上にとどまったもう一つの理由に、パイプラインキャッシュという特殊な構成であることが上げられる。通常の SRAM では、メモリセルアレイのアクセス以外の処理を行なっている間にプリチャージを行なうことができ、プリチャージ時間を隠蔽できる。しかし、パイプラインキャッシュではプリチャージ時間がそのまま処理時間に加わる。したがって、ステージのクロックサイクル時間を 1.13 倍にしか短縮できなかった。

5.3 現在研究されている高速 SRAM との比較

表 5.4 に現在研究されている高速 SRAM をまとめた [14][15][16][17]。

表 5.4: 現在研究されている高速 SRAM との比較

	容量 [Mb]	周波数 [MHz]	ゲート長 [μm]	面積 [mm^2]	セル面積 (構造) [μm^2]	金属層	電圧 [V]
Intel98	4	450	0.35	222	(6T)	4 層	2.5V
Hitachi98	4.5	550	0.25	128	9.84	4 層	1.8V
IBM97	4	350	0.30	145	18.7(6T)	4 層	2.5V
NEC97	4	500	0.25	132	12.6(6T)	3 層	2.5V
本設計	4.5	3000	0.25	231	20.35(6T)	5 層	2.5V

表 5.4 より、現在研究されている高速 SRAM の動作周波数は 500MHz 程度であり、9 段にパイプライン化することにより 6 倍の高速化が実現した。また、セル面積を同じ 6T 構造のものと比較すると 1.6 倍程度大きい。セルを小さくすれば配線長が短くなり、さらに高速動作が可能である。

パイプライン化されていない SRAM は、テクノロジーが微細化されたとしても配線遅延の影響で動作周波数がそれほど向上しないと考えられる。しかし、パイプラインキャッシュの場合は、長い配線を分割しパイプラインの段数を増やすことにより、微細化しても一定である配線遅延を多数のステージに分割させることができる。メモリセルアレイにアクセスするステージの遅延時間は、メモリセルアレイの大きさをゲート遅延と配線遅延の割合の最適となる大きさに分割することにより、遅延時間を短縮させることが可能である。0.10 μm のパイプラインキャッシュでは、0.25 μm で最適な段数の 9 段からさらに段数を増やすことにより 7.7GHz の周波数で動作可能である。本設計では、縮小比と同じ 2.5 倍に動作周波数を向上させることができた。

第 6 章

結論

マルチスレッド型プロセッサは、キャッシュに対してレイテンシよりも高スループットを求める。キャッシュをこのような用途で使用する場合に、パイプライン化によってどの程度の高周波動作が可能となるかを研究した。

パイプライン化によってキャッシュを高周波動作させる場合に、メモリセルアレイの大きさをどの程度にするかが重要となる。メモリセルアレイを小さく分割していくと、全遅延時間に占める配線遅延の割合がゲート遅延の割合よりも小さくなり、分割による高速化の割合が小さくなる。本研究では、 $0.25\mu\text{m}$ テクノロジーにおいては、 $128 \times 128\text{bit}$ の大きさが適当であると導き出した。

また、 $128 \times 128\text{bit}$ のメモリセルアレイにプリチャージ回路を導入した時と省いた時の比較を行なった。そして、パイプラインキャッシュという特殊な構成における $128 \times 128\text{bit}$ 程度の小さなメモリセルアレイは、プリチャージ回路により高速化されないことがわかった。

本研究では、 4Mb のデータ容量を持つキャッシュの設計を細部にわたり行なった。各ステージごとに詳細な回路の設計、チップ内の配置についての設計、およびチップ内を通るアドレスパスやデータパス等の配線の設計を行ない、各ステージの最小クロックサイクル時間を求めた。 $0.25\mu\text{m}$ テクノロジーでは、9段のステージ構成とすることにより 3GHz の動作周波数で動作可能であることを導き出した。 $0.10\mu\text{m}$ テクノロジーでは、 $0.25\mu\text{m}$ と同じ9段のステージ構成では 6.5GHz 、さらにステージを分割することにより 7.7GHz の動作が可能であることを導き出した。 $0.25\mu\text{m}$ と $0.10\mu\text{m}$ の動作周波数を比較すると、配線遅延を多数のステージに分割することにより、縮小比と同じ 2.5 倍に動作周波数が向上している。このことから、配線遅延の影響で速度向上が飽和してきているテクノロジーでは、パイプラインキャッシュは理想的な構成であるといえる。

また、パイプラインキャッシュを構成する基本回路のレイアウトを行ない面積を評価した。

その結果、パイプライン化による面積の増加率は1.17倍であった。面積にして 32.9mm^2 であり、これはメモリセル1.6Mb分である。

最後に、私たちの研究室ではマルチスレッド型プロセッサにウェーブパイプラインを導入する研究も行なっている。ウェーブパイプラインを導入するとプロセッサの動作周波数はかなり高くなると考えられる。プロセッサの動作周波数が3GHzまでは、パイプラインキャッシュにより対応できることがわかった。

第 A 章

各配線と周辺配線との距離関係

本章では、配線の抵抗および配線の容量を求めるために使用した各配線と周辺配線との距離関係のモデル図を示す。図中の青色で示した数値は、単位面積あたりの酸化膜容量 C_0 である。また、括弧内の数値はとなりあう2配線間の酸化膜容量である。

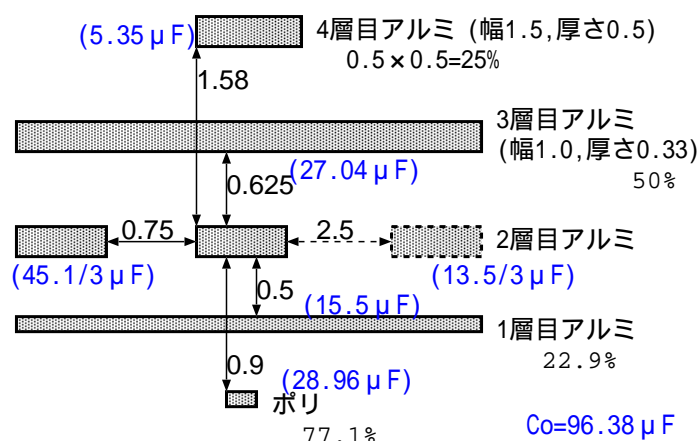


図 A.1: ワード線 (2層目アルミ配線) と周辺の配線との距離関係

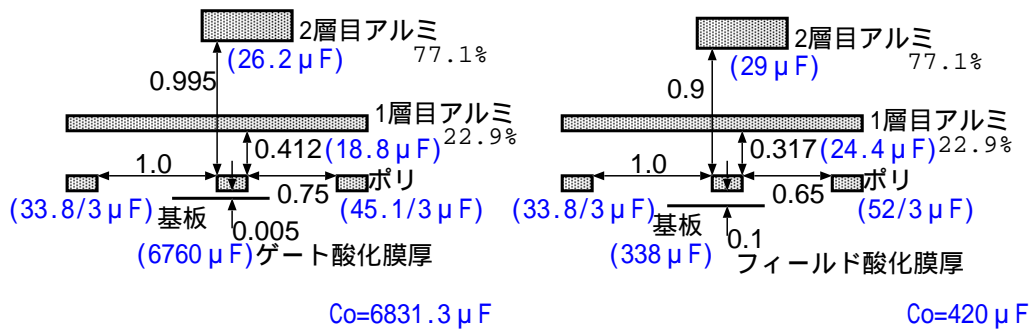


図 A.2: ワード線 (ポリシリコン配線) と周辺の配線との距離関係

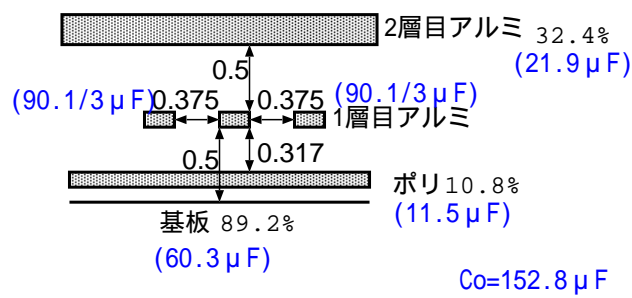
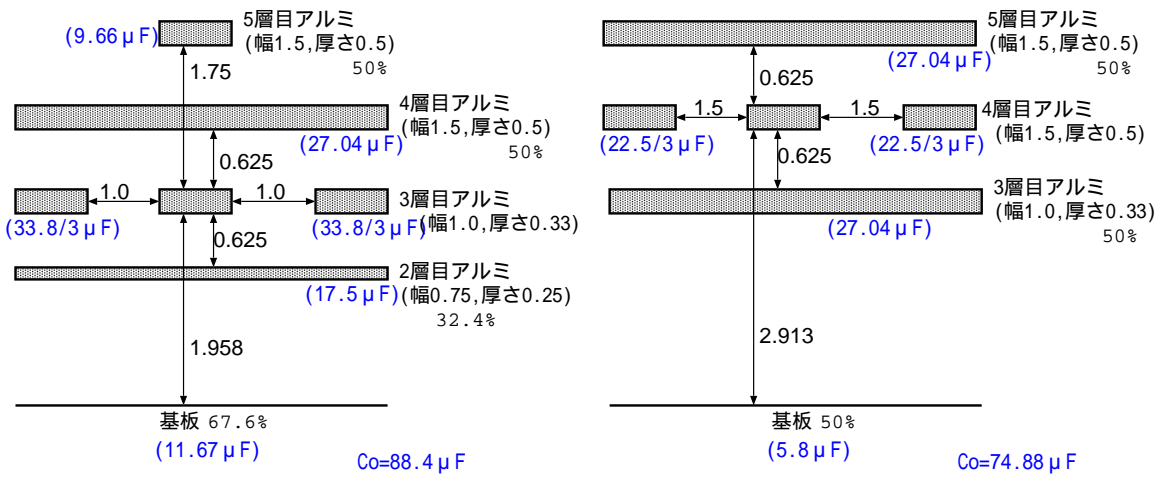
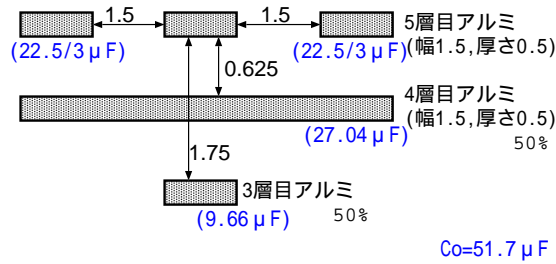


図 A.3: ビット線 (1層目アルミ配線) と周辺の配線との距離関係



(a) 3層目アルミ配線 (配線間ピッチ2 μm) (b) 4層目アルミ配線 (配線間ピッチ3 μm)



(c) 5層目アルミ配線 (配線間ピッチ3 μm)

図 A.4: データ・アドレスバス配線 (3,4,5 層目アルミ配線) と周辺配線との距離関係

第 B 章

回路図

本章では、SPICE シミュレーションに用いた各ステージの回路図を示す。

論理ゲートに表記した数値は W_p/W_n 、pMOSFET、nMOSFET に表記した数値はそれぞれ W_p 、 W_n である。ここで、 W_p は pMOSFET のチャンネル幅、 W_n は nMOSFET のチャンネル幅であり、その単位は (μm) である。ゲート長は pMOSFET が $L_p = 0.30\mu m$ 、nMOSFET が $L_n = 0.25\mu m$ に固定している。

図中の番号は、SPICE プログラムにおける端子番号である。また、図の赤色または緑色の点線で囲んだ部分は、sub-circuit を示す。緑色は sub-circuit の中の sub-circuit であり、わかり易くするために色を変えた。

図中に青色の数値と矢印で示したものは、端子間の配線長である。記号 ϕ はクロックである。

本研究は、各ステージにおける最大遅延を求めることが目的であり、必要でない回路は省略した。

赤	SPICE ネットリストの端子番号 (main-circuit)	ピンク	sub-circuit	緑	sub-sub-circuit
青	配線長	水色	各ステージにおける入力ラッチの数		

図 B.1: 回路図に示した色の意味

B.1 ラッチ

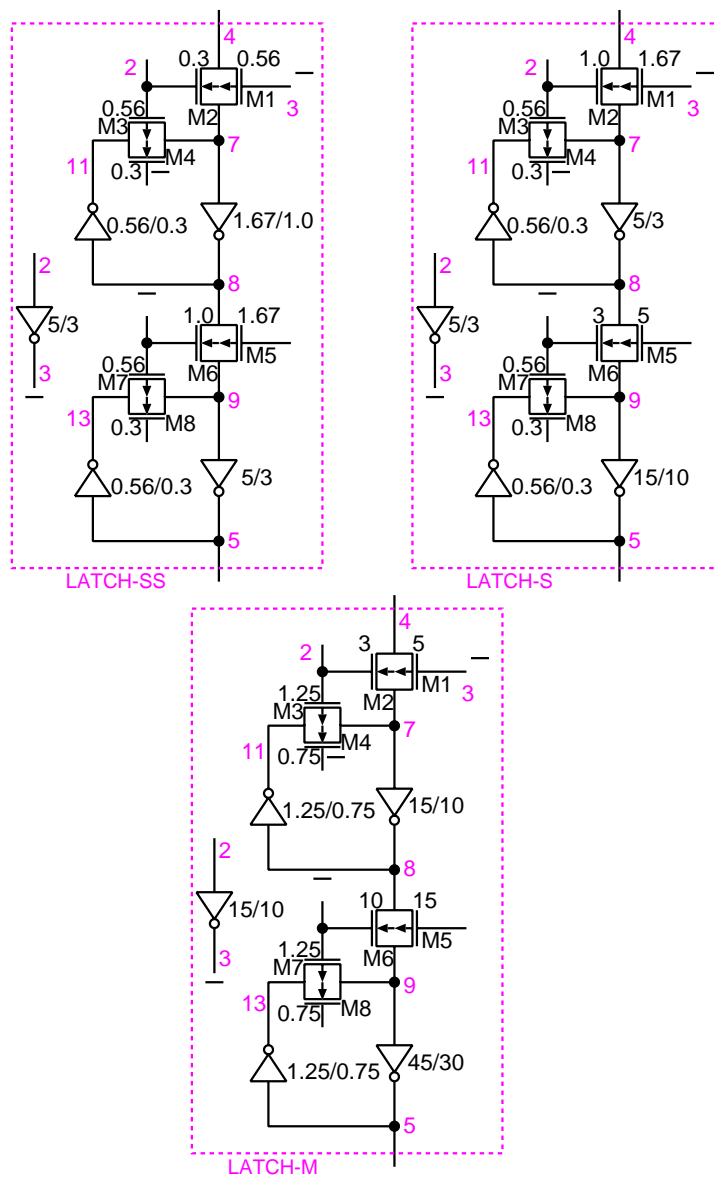


図 B.2: 大きさの異なる 3 種類のラッチ

B.2 メモリセル

メモリセルの回路図と、レイアウト図 (図 C.12参照) から求めた MOSFET の大きさや配線に関する情報を取り入れた SPICE ネットリストを示す。

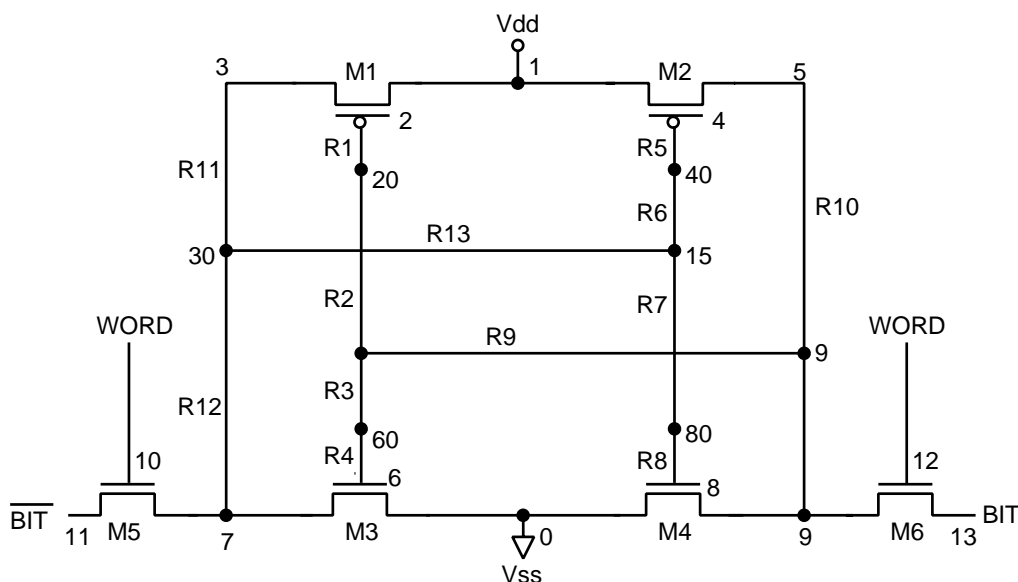


図 B.3: メモリセルの回路図

下記に $0.25\mu\text{m}$ のメモリセル部分の SPICE ネットリストを示す。記号の意味は次のようになる。L はトランジスタのゲート長または配線長。W はトランジスタのゲート幅または配線幅。AD、AS はドレイン、ソース拡散領域の面積。PD、PS はドレイン、ソース領域の周囲の長さ。COX は単位面積あたりの酸化膜容量。RSH はシート抵抗値。

SPICE ネットリスト (メモリセル)

```
M1 3 2 1 1 PMOS1 L=0.3U W=1.0U AD=0.625p AS=0.59375p PD=3.25U PS=2.125U
M2 5 4 1 1 PMOS1 L=0.3U W=1.0U AD=0.625p AS=0.59375p PD=3.25U PS=2.125U
M3 7 6 0 0 NMOS1 L=0.25U W=1.0U AD=0.5p AS=0.46875p PD=2.5U PS=2.5U
M4 9 8 0 0 NMOS1 L=0.25U W=1.0U AD=0.5p AS=0.46875p PD=2.5U PS=2.5U
M5 11 10 7 0 NMOS1 L=0.25U W=0.75U AD=0.5p AS=0.5625p PD=2.25U PS=2.5U
M6 9 12 13 0 NMOS1 L=0.25U W=0.75U AD=0.5625p AS=0.28125p PD=2.5U
PS=1.495U
R1 2 20 RE-PG-M W=0.3U L=0.375U
R2 20 14 RE-P-M W=0.3U L=1.0U
R3 14 60 RE-P-M W=0.25U L=0.5U
R4 60 6 RE-PG-M W=0.25U L=0.5U
```

R5 4 40 RE-PG-M W=0.3U L=0.375U
 R6 40 15 RE-P-M W=0.3U L=0.5U
 R7 15 80 RE-P-M W=0.25U L=1.0U
 R8 80 8 RE-PG-M W=0.25U L=0.5U
 R9 14 9 RE-AL1-M W=0.25U L=1.5U
 R10 9 5 RE-AL1-M W=0.25U L=1.625U
 R11 3 30 RE-AL1-M W=0.25U L=0.75U
 R12 30 7 RE-AL1-M W=0.25U L=0.875U
 R13 30 15 RE-AL1-M W=0.25U L=1.625U

配線モデル

.MODEL RE-PG-M R COX=6760U RSH=120.48	ゲート上のポリシリコン配線
.MODEL RE-P-M R COX=338U RSH=120.48	ポリシリコン配線
.MODEL RE-AL1-M R COX=67.6U RSH=0.36	アルミ配線

B.3 デコードステージ

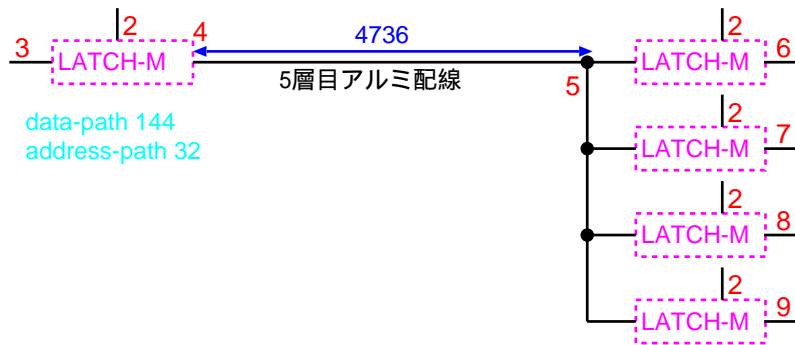


図 B.4: デコードステージ 1 (データパス、アドレスパス)

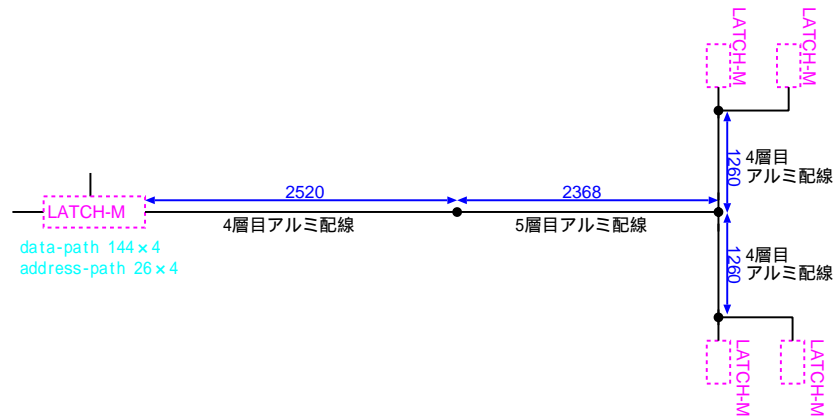


図 B.5: デコードステージ2 (データパス、アドレスパス)

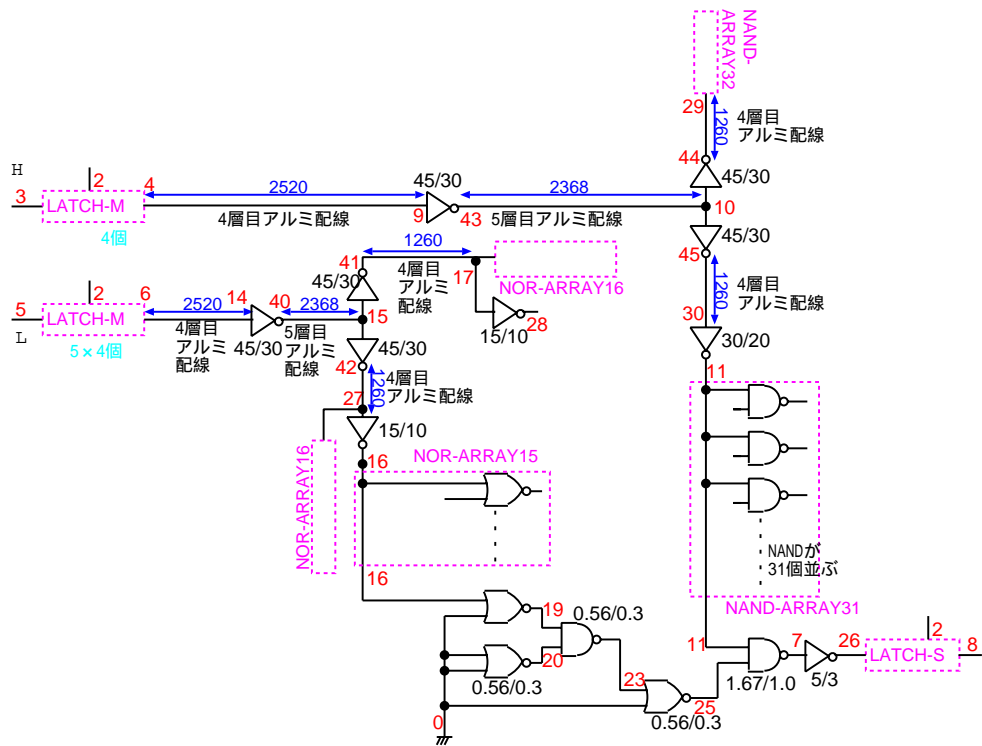


図 B.6: デコードステージ2 (アドレスパス)

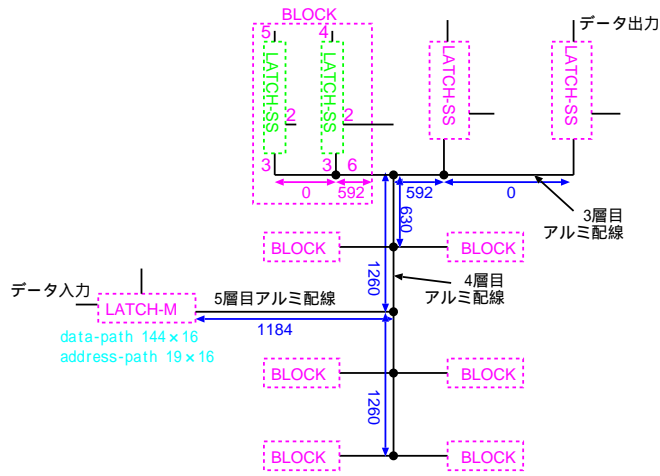


図 B.7: デコードステージ3 (データパス、アドレスパス)

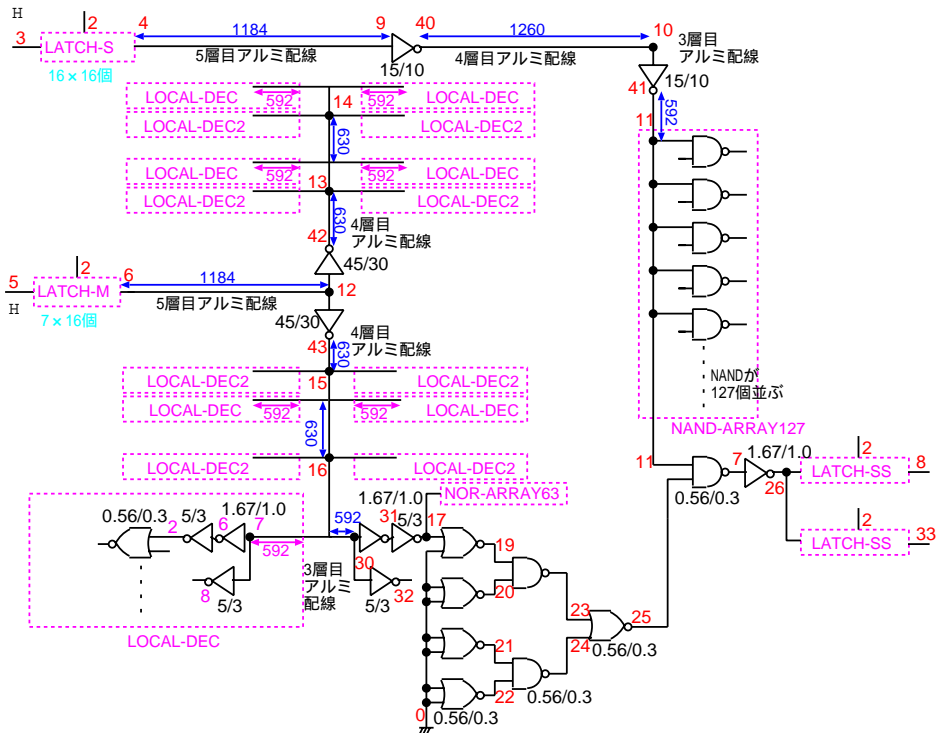


図 B.8: デコードステージ3 (アドレスパス)

B.5 タグ比較ステージ

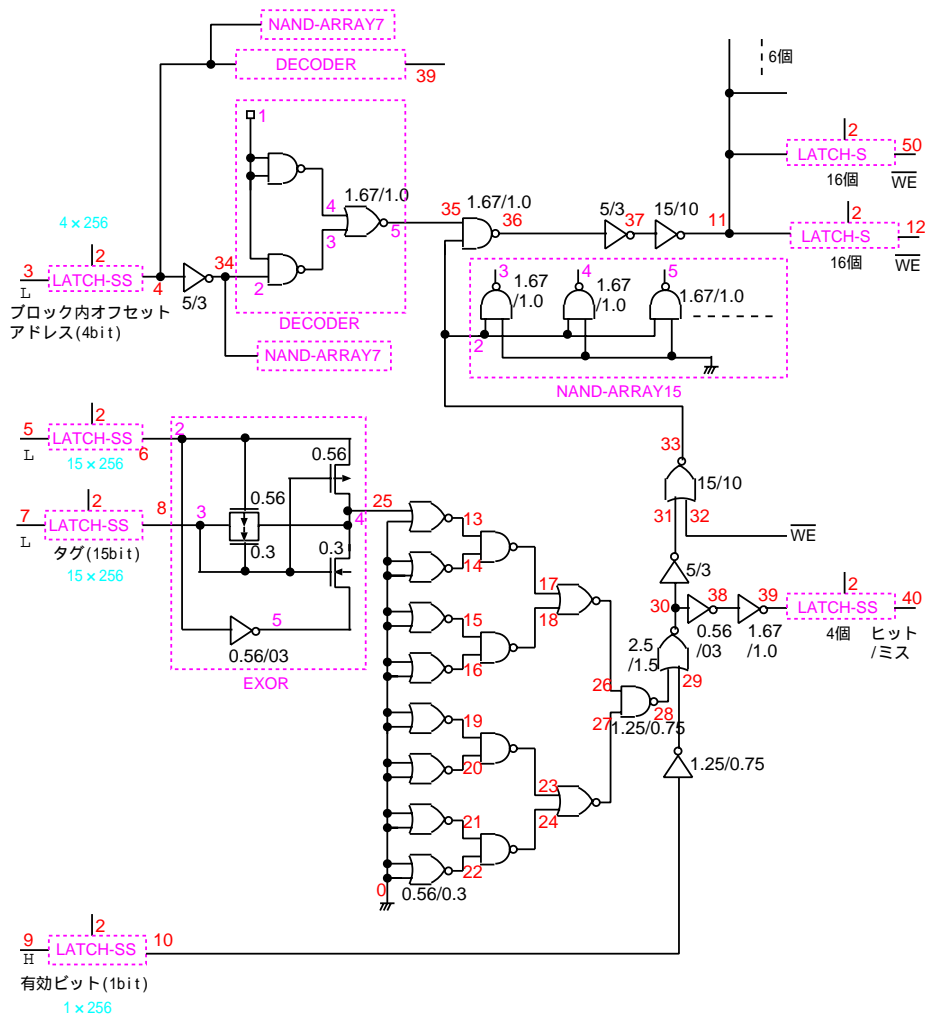


図 B.13: タグ比較ステージ

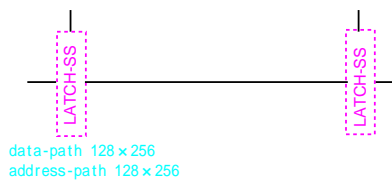


図 B.14: タグ比較ステージ (データパス、アドレスパス)

B.7 出力パスステージ

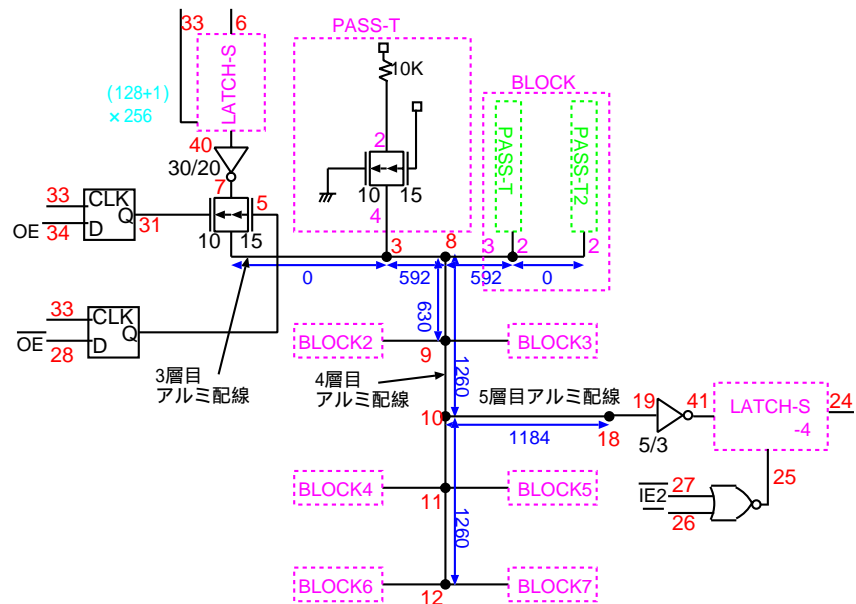


図 B.19: 出力パスステージ 1

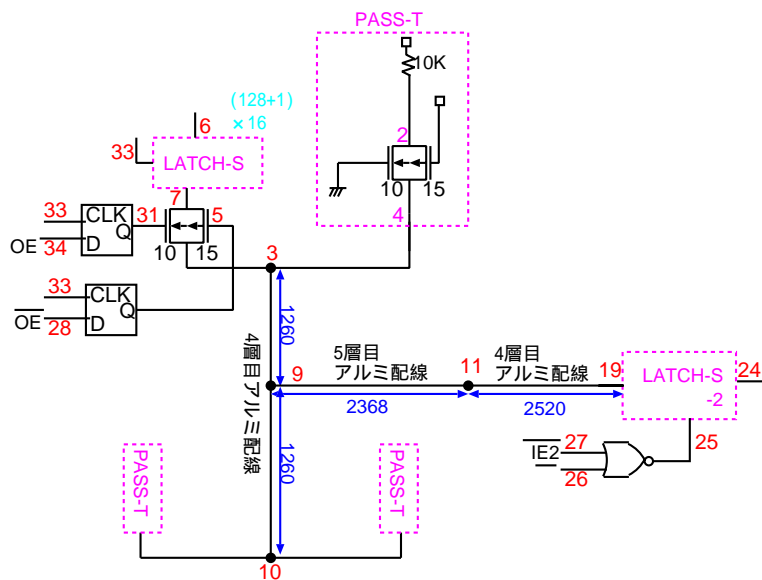


図 B.20: 出力パスステージ 2

B.8 マルチプレクサステージ

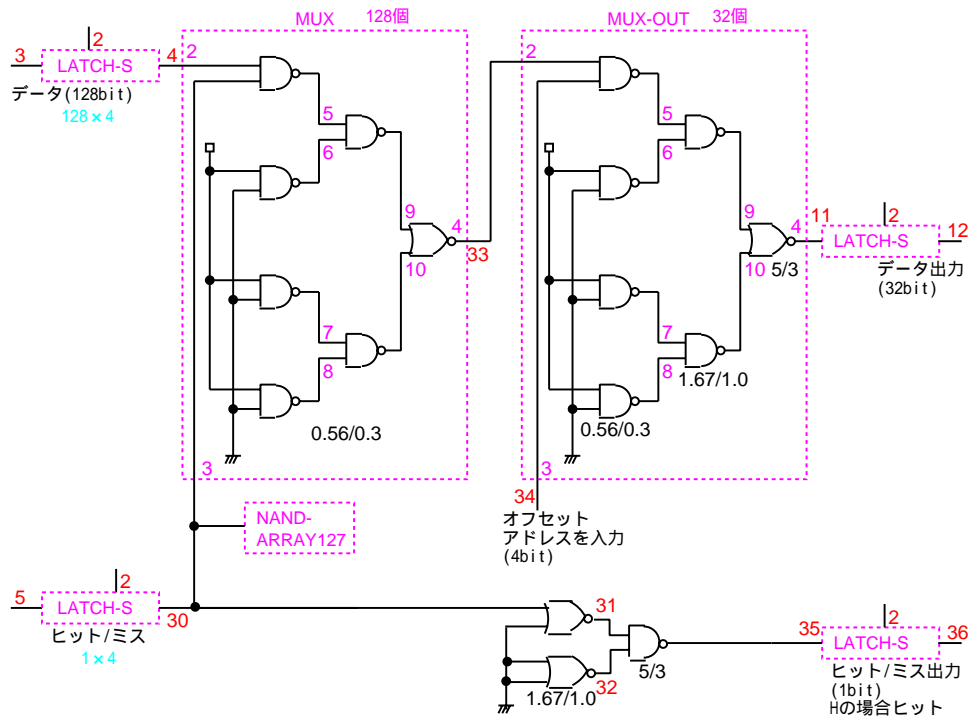


図 B.21: マルチプレクサステージ

第 C 章

レイアウト図

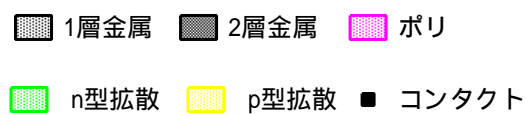


図 C.1: 層の種類

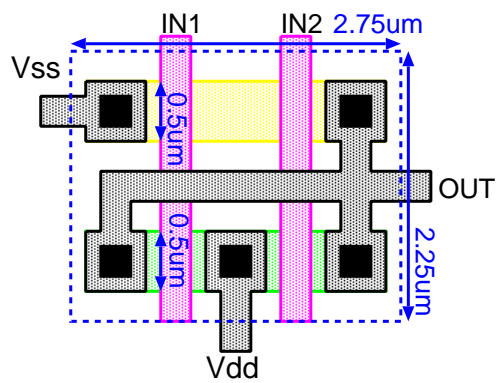


図 C.2: NAND ゲートのレイアウト

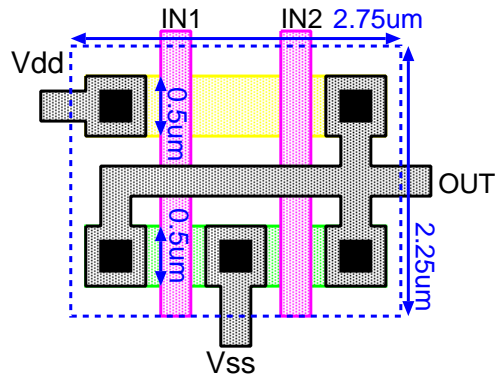


図 C.3: NOR ゲートのレイアウト

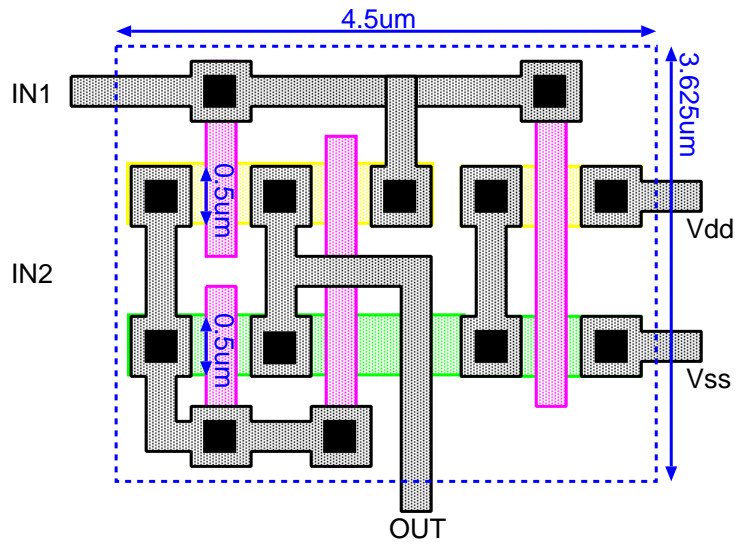


図 C.4: EXOR ゲートのレイアウト

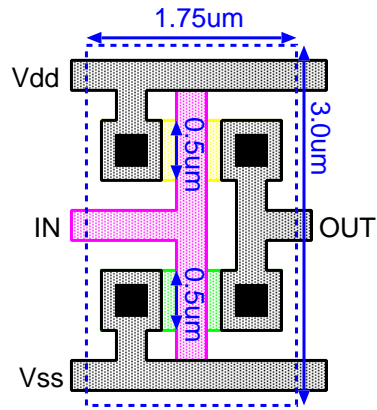


図 C.5: インバータのレイアウト

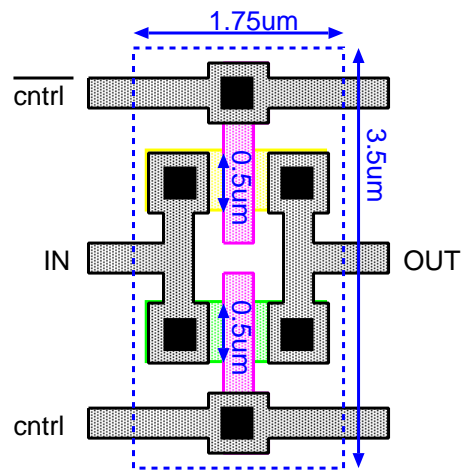


図 C.6: トランスミッションゲートのレイアウト

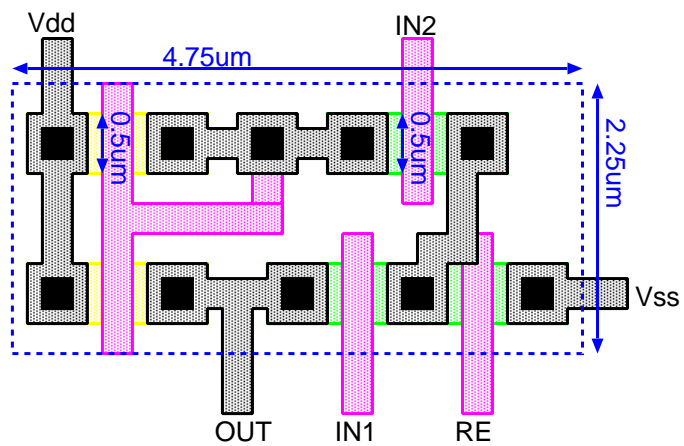


図 C.7: カレントミラー (センスアンプ) のレイアウト

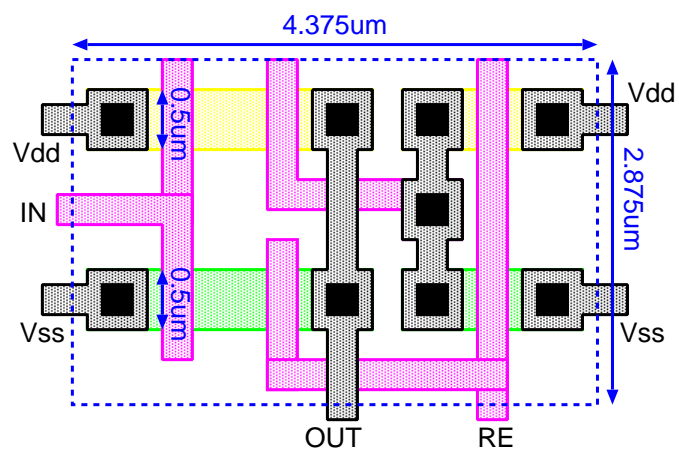


図 C.8: アウトプットインバータ (センスアンプ) のレイアウト

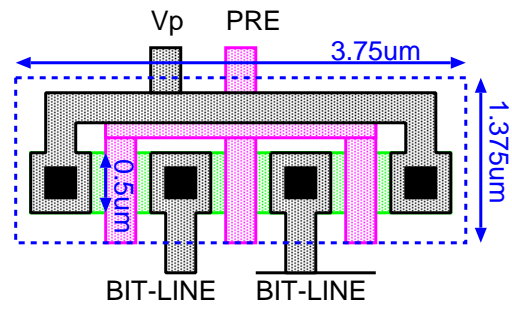


図 C.9: プリチャージ回路のレイアウト

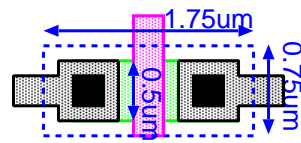


図 C.10: パストランジスタのレイアウト

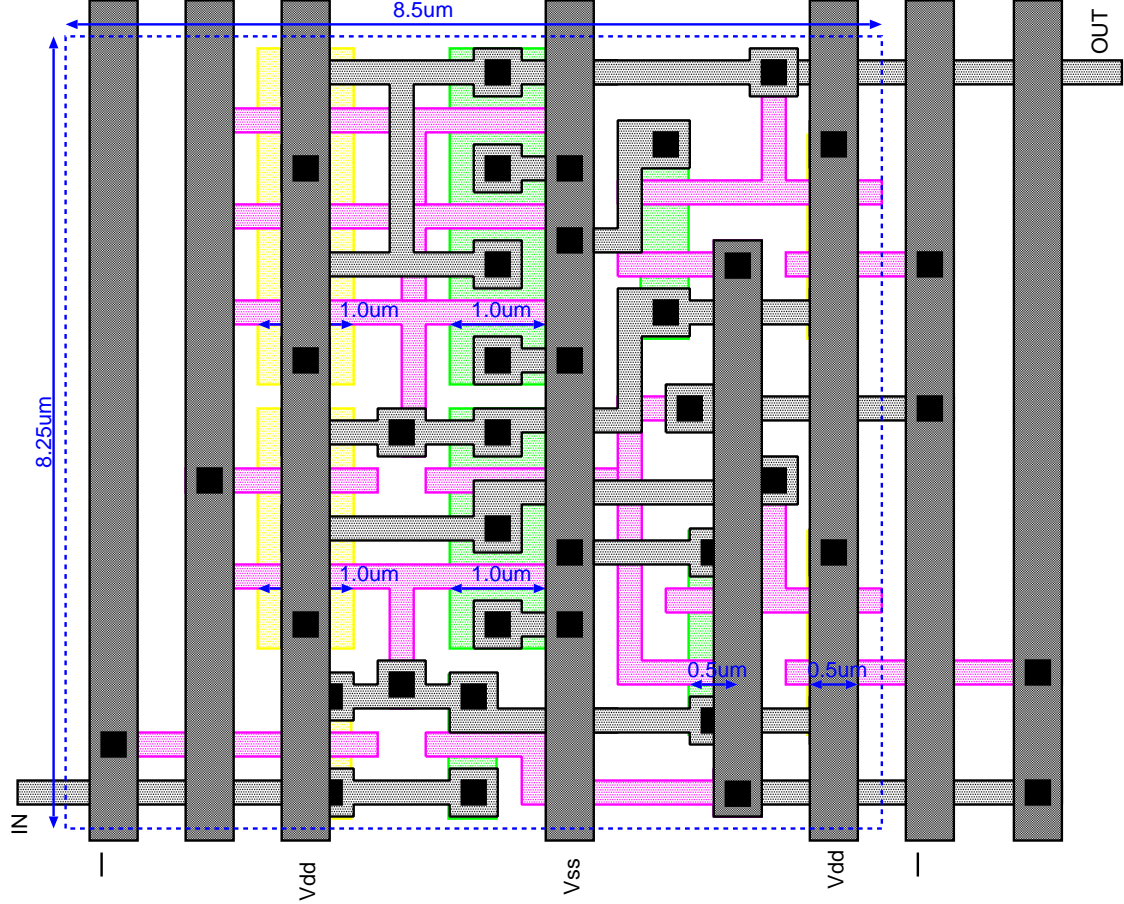


図 C.11: ラッチのレイアウト

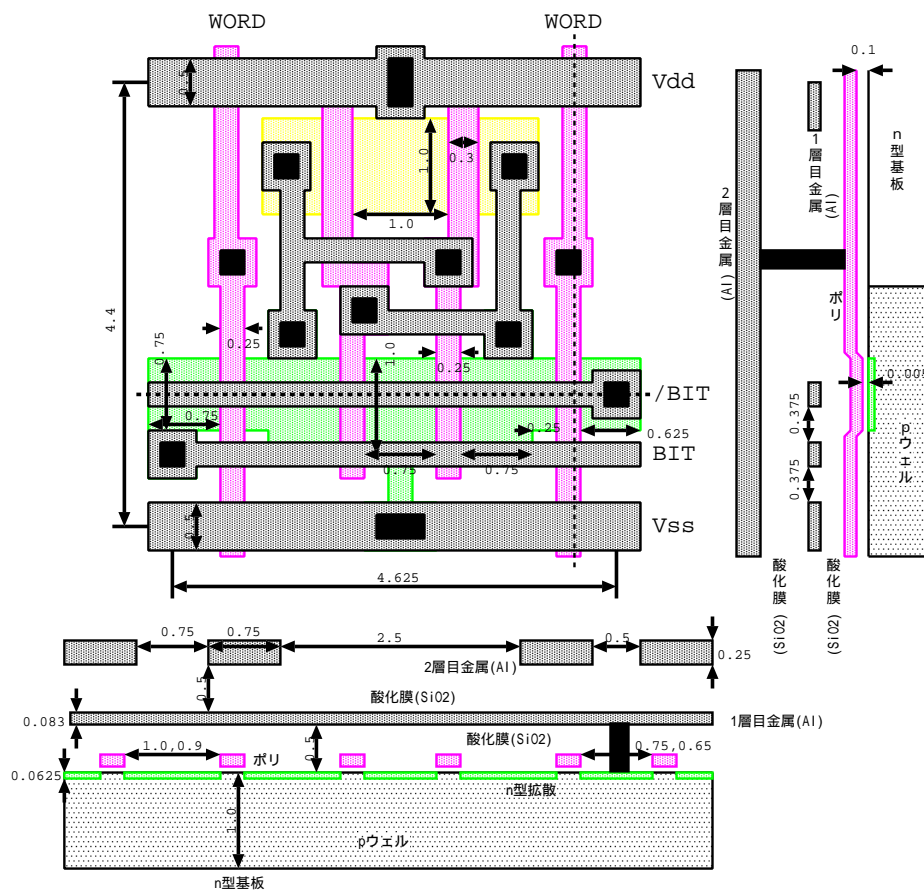


図 C.12: メモリセルのレイアウト

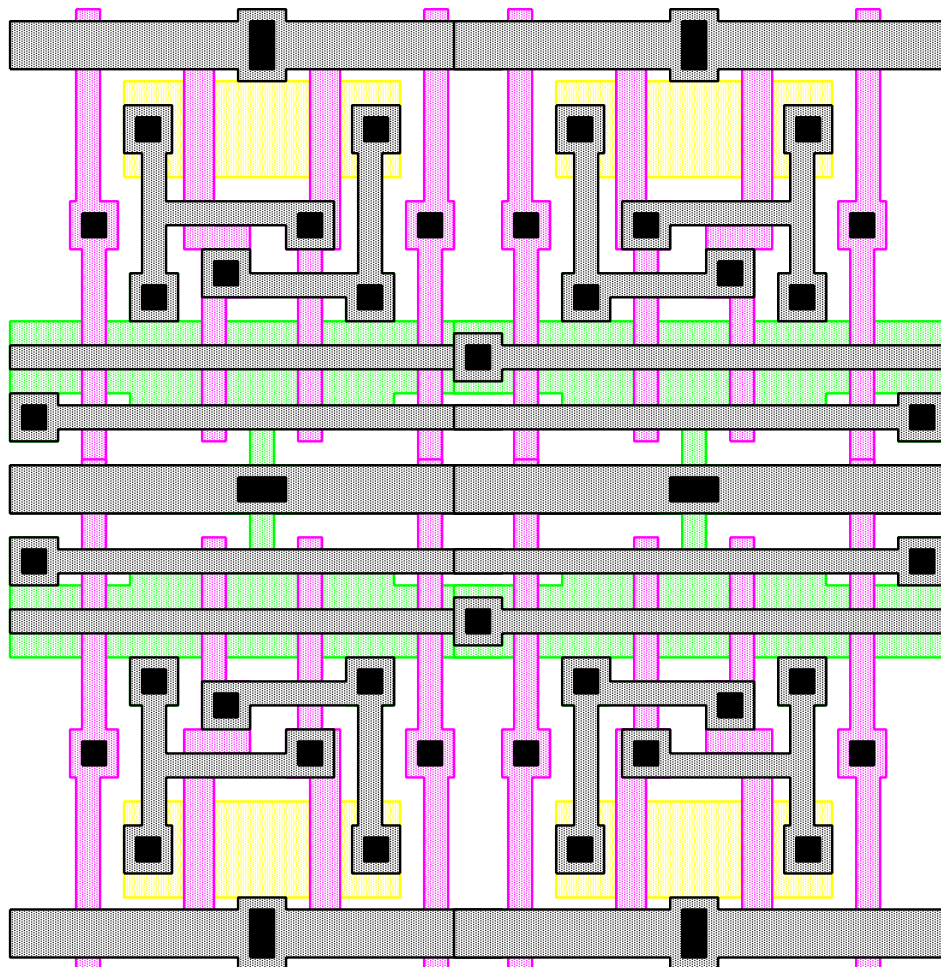


図 C.13: メモリセル4個分のレイアウト

第 D 章

参考図

本章では、本文に掲載した図をカラーで示した。

図 D.1は、データが流れていく順番にラッチに色をつけた。データは赤・ピンク・緑・青の順番に流れる。また、同じ色のラッチは、同じ時間にデータが並行して入るラッチである。

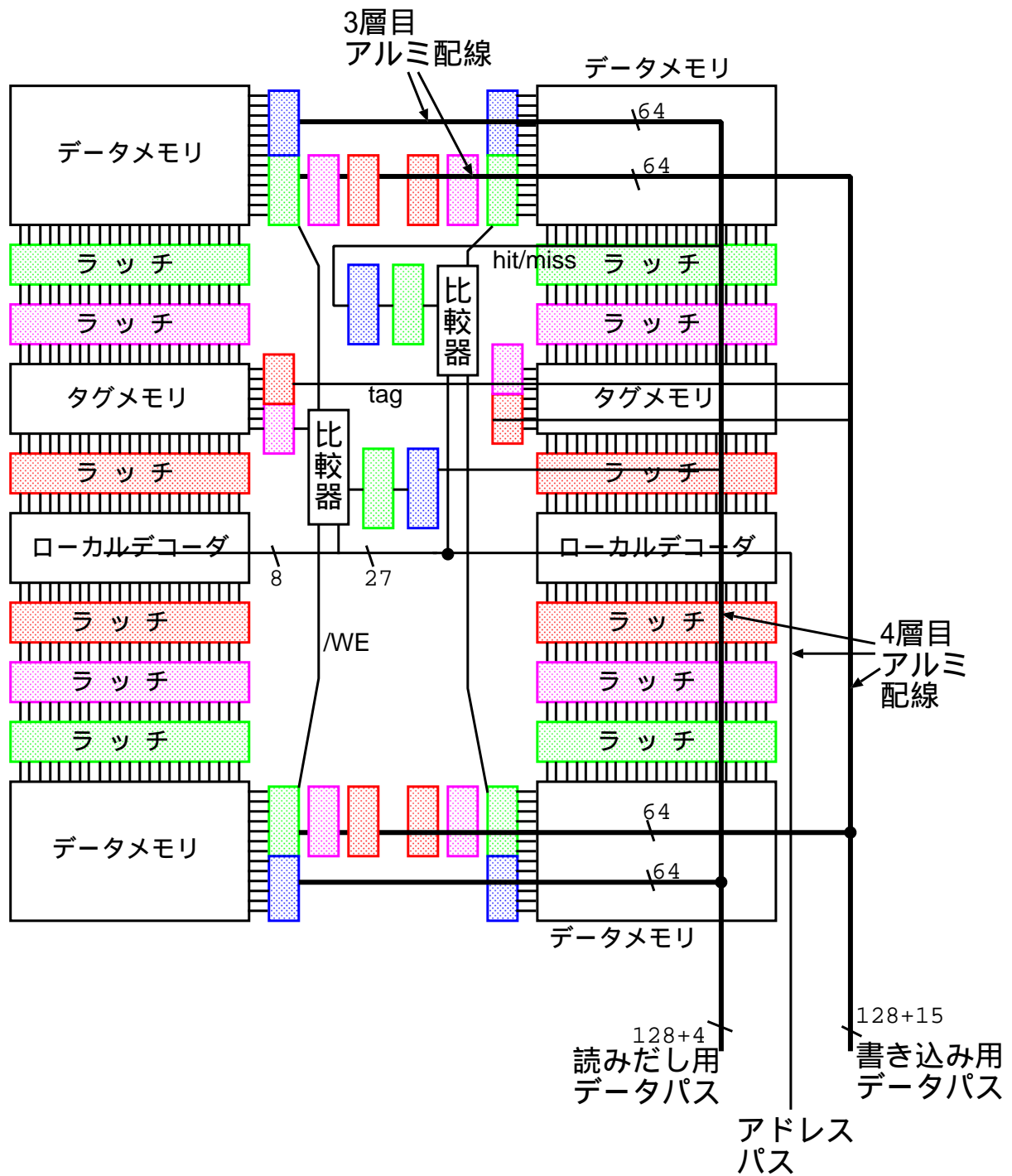


図 D.1: 128 × 144 のメモリセルアレイ 2 個分の配置

謝辞

本研究を進めるにあたり、終始熱心かつ寛容な御指導を賜りました日比野 靖 教授に心から感謝いたします。

また、適切な御助言をして頂きました堀口 進 教授、篠田 陽一 助教授、丹 康雄 助教授、宮崎 純 助手に深く感謝いたします。

さらに、マルチスレッド型プロセッサに関する御助言を頂いた杉本 朗子さん、池田 吉朗君に感謝いたします。

その他、貴重な御意見、御討論を頂きました日比野・中島研究室の皆様をはじめ、多くの方々の御助言に対し厚く御礼申し上げます。

参考文献

- [1] 井口秀之, 次世代集積技術による RISC アーキテクチャプロセッサの設計と評価, 北陸先端科学技術大学院大学修士論文, 1995.
- [2] 相原孝一, マルチスレッド型プロセッサ向けのキャッシュ機構のパイプライン化に関する研究, 北陸先端科学技術大学院大学修士論文, 1997.
- [3] 伊藤英治, 関数型プログラムの実行に適したマルチスレッド型プロセッサ・アーキテクチャに関する研究, 北陸先端科学技術大学院大学修士論文, 1997.
- [4] Toshihiko Hirose, et al., A 20ns 4Mb CMOS SRAM with Hierarchical Word Decoding Architecture, ISSCC DIGEST OF TECHNICAL PAPERS, pp.132-133, 1990.
- [5] 伊藤清男, 超 LSI メモリ, 培風館, 1994.
- [6] 榎本忠儀, CMOS 集積回路, 培風館, pp.90, 1996.
- [7] Neil H.E. Weste and Kamran Eshraghian, 富沢 松山 監訳, CMOS VLSI 設計の原理 -システムの見点から-, 丸善, pp.111, 1988.
- [8] 菅野卓雄 監修, 飯塚哲哉 編, CMOS 超 LSI の設計, 培風館, pp.17, 1989.
- [9] Marco Annaratone, 平野浩太郎 富田昌宏 訳, デジタル CMOS 回路設計, コロナ社, 1992.
- [10] Betty Prince, Semiconductor Memories, Second edition, John Wiley & Sons, 1995.
- [11] Betty Prince, High Performance Memories, John Wiley & Sons, 1996.
- [12] Daniel Foty, MOSFET modeling with SPICE, Prentice Hall, 1997.
- [13] 西久保靖彦, ASIC 設計回路シミュレータ SPICE 入門, 日本工業技術センター, 1988.

- [14] B.Bateman, et al., A 450MHz 512kB Second-Level Cache with a 3.6GB/s Data Bandwidth, ISSCC DIGEST OF TECHNICAL PAPERS, pp.358-359, 1998.
- [15] H.Nambu, et al., A 1.8ns Access, 550MHz 4.5Mb CMOA SRAM, ISSCC DIGEST OF TECHNICAL PAPERS, pp.360-361, 1998.
- [16] Georie Braceras, et al., A 350MHz 3.3V 4Mb SRAM Fabricated in a 0.3um CMOS Process, ISSCC DIGEST OF TECHNICAL PAPERS, pp.404-405, 1997.
- [17] Kazuyuki Nakamura, et al., A 500MHz 4Mb CMOS Pipeline-Burst Cache SRAM with Point-to-Point Noise Reduction Coding I/O, ISSCC DIGEST OF TECHNICAL PAPERS, pp.406-407, 1997.
- [18] Koichiro Ishibashi, et al., A 300MHz 4-Mb Wave-Pipeline CMOS SRAM Using a Multi-Phase PLL, ISSCC DIGEST OF TECHNICAL PAPERS, pp.308-309, 1995.