

論文 / 著書情報
Article / Book Information

題目(和文)	不揮発/擬似不揮発記憶を用いたCMOSロジックの低消費電力技術に関する研究
Title(English)	Study on low-power technologies for CMOS logic systems using nonvolatile/virtually nonvolatile retention
著者(和文)	北形大樹
Author(English)	Daiki Kitagata
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11428号, 授与年月日:2020年3月26日, 学位の種別:課程博士, 審査員:菅原 聡,中川 茂樹,宮本 恭幸,岡田 健一,杉野 暢彦,宗片 比呂夫, 高木 信一,前島 英雄
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11428号, Conferred date:2020/3/26, Degree Type:Course doctor, Examiner:,,,,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

学位論文

不揮発/擬似不揮発記憶を用いた CMOS ロジックの
低消費電力技術に関する研究

Study on low-power technologies for CMOS logic systems
using nonvolatile/virtually-nonvolatile retention

令和 2 年 3 月 10 日

東京工業大学 工学院
電気電子系 電気電子コース
17D10217 北形 大樹

指導教員 菅原 聡 准教授

目次

第 1 章 序論	1
1.1 CMOS ロジックシステムにおける待機時電力の問題	1
1.2 待機時電力の起源	2
1.2.1 ゲートリーク	2
1.2.2 サブスレッショルドリーク	3
1.3 パワーゲーティングとその課題	5
1.3.1 構成方法	5
1.3.2 性能制約要因	6
1.3.3 CMOS ロジックシステムにおけるパワーゲーティングの課題	7
1.4 キャッシュへの不揮発記憶の導入方法	8
1.5 不揮発記憶を用いたその他の待機時電力削減技術	11
1.6 本研究の目的と本論文の構成	13
1.7 図と表	14
1.8 参考文献	20
第 2 章 不揮発/擬似不揮発性 SRAM の性能評価指標	24
2.1 不揮発/擬似不揮発性 SRAM システムの構成方法	24
2.1.1 用語の説明	24
2.1.2 パワースイッチアーキテクチャ	25
2.1.3 アレイアーキテクチャ	26
2.2 不揮発/擬似不揮発記憶を用いた PG の性能評価指標	29
2.2.1 BET	29
2.2.2 BET の削減方法	33
2.2.3 PG のリーク電力削減率	37
2.2.4 本論文における評価方法	38
2.3 不揮発/擬似不揮発性 SRAM の設計指標	39
2.4 まとめ	43

2.5 Appendix	43
2.6 図と表	45
2.7 参考文献	59
第3章 不揮発性 SRAM	61
3.1 不揮発性双安定回路の実現方法	61
3.2 不揮発性 SRAM の回路構成と動作原理	63
3.3 不揮発性 SRAM の設計	66
3.3.1 セル設計	66
3.3.2 リーク削減アーキテクチャ	70
3.3.3 動作波形	71
3.4 電力・エネルギー性能の評価	72
3.5 ストアエネルギーレイテンシ削減アーキテクチャ	73
3.5.1 階層型ストアフリー電源遮断アーキテクチャ	74
3.5.2 積極的無用データフラッシュアーキテクチャ	76
3.6 試作 TEG による検証	78
3.7 まとめ	79
3.8 Appendix	80
3.9 図と表	84
3.10 参考文献	118
第4章 擬似不揮発性 SRAM	123
4.1 擬似不揮発性 SRAM の実現方法	123
4.1.1 擬似不揮発記憶を用いた PG	123
4.1.2 擬似不揮発性双安定記憶回路の構成方法	125
4.2 擬似不揮発性 SRAM の回路構成と動作原理	126
4.3 擬似不揮発性 SRAM の設計	129
4.3.1 パワースイッチの構成	129
4.3.2 セルの設計方法	130
4.3.3 設計結果	132

4.4	電力・エネルギー性能の評価.....	136
4.5	リーク電力削減アーキテクチャ.....	139
4.5.1	デュアルパワースイッチを用いた低消費電力化技術.....	139
4.5.2	無用データのフラッシュアーキテクチャ.....	140
4.6	擬似不揮発性 SRAM マクロの電力・エネルギー性能.....	141
4.7	先端 CMOS を用いた性能評価.....	142
4.8	まとめ.....	143
4.9	Appendix.....	144
4.10	図と表.....	146
4.11	参考文献.....	180
第 5 章	CMOS ロジックシステムへの応用.....	184
5.1	不揮発/擬似不揮発記憶を用いた PG の性能評価.....	184
5.1.1	CMOS ロジックシステムのリーク電力削減率と面積増加率.....	185
5.1.2	不揮発/擬似不揮発性 FF の導入による効果.....	187
5.1.3	不揮発/擬似不揮発性 SRAM の導入による効果.....	190
5.2	PG の平均リーク電力に対する BET の効果.....	190
5.2.1	非実行時間のモデリング.....	191
5.2.2	キャッシュの平均リーク電力.....	192
5.2.3	システムの平均リーク電力.....	193
5.3	不揮発/擬似不揮発性 SRAM の総合比較.....	195
5.4	まとめ.....	198
5.5	Appendix.....	198
5.6	図と表.....	202
5.7	参考文献.....	221
第 6 章	結論.....	223
	本研究に関する発表.....	226
	謝辞.....	230

第 1 章

序論

本章では、はじめにマイクロプロセッサ (μP) や System-on-a-chip (SoC) などの CMOS ロジックシステムにおける待機時電力に関する課題と、その原因と対策について述べる。次いで、待機時電力の最も有効な削減方法であるパワーゲーティング (PG) について述べ、PG の課題を示す。また、 μP におけるキャッシュの PG に関する課題と、これに不揮発記憶を導入することの意義を述べる。最後に以上の議論を踏まえ、本研究の目的と本論文の構成を述べる。

1.1 CMOS ロジックシステムにおける待機時電力の問題

現代の高度情報化社会は半導体エレクトロニクス技術の進歩とともに発展してきた。特に、各種コンピューティングシステムに搭載させるマイクロプロセッサ (μP)、スマートフォンなどのスマートモバイルデバイスに搭載される System-on-a-chip (SoC) などの CMOS ロジックシステムは高度情報化社会を支える重要な基盤技術の 1 つになっている。CMOS ロジックシステムの CMOS とは Complementary metal-oxide-semiconductor の略語であるが、その意味するところは、デバイスの構造を表すというよりは、相補的に機能する 2 種類の (伝導型の異なる) MOSFET (Metal-oxide semiconductor field-effect transistor) と呼ばれるトランジスタを半導体チップ上に集積して構成した電子回路または電子システムの総称であることが多い。CMOS のそれぞれの MOSFET は図 1.1.1 に示す構造を持ち、ソース (S) とドレイン (D) の間を流れる電流をゲート (G) によって制御する 3 端子のデバイスであるが、この電流はソース-ドレイン間の半導体領域 (ボディ (B) と呼ばれることが多い) の電位 (ボディバイアスまたは基板バイアスと呼ばれる) によって変化させることが可能なため、実際には 4 端子デバイスである (図 1.1.1 の場合ではウェルと呼ばれる領域に印加される電圧によって基板バイアスが制御される)。

CMOS ロジックシステムの発展は、これまで MOSFET の微細化によって達成されてきた。スケーリング則[1]と呼ばれる微細化の指導原理に従って MOSFET のデバイスサイズ(より正確にはソース-ドレイン間の長さを表すチャンネル長)を縮小することで、高密度集積化にともなう様々な高性能化を実現してきた。しかし、微細化によってチャンネル長が 100nm を下回ってくると、消費電力の増加にともなう問題が顕在化してくる。CMOS によって構成される回路の消費電力は動作時電力(ダイナミック電力)と待機時電力(スタティック電力)に分けられる。動作時電力は回路の動作にともなって消費される電力で、主に MOSFET のスイッチングによる配線の充放電によって生じる。一方、待機時電力は回路の待機(スタンバイ)状態時にトランジスタに生じるリーク電流によって消費する電力である。図 1.1.2 は CMOS ロジックシステムの電力密度を MOSFET のゲート長に対してプロットしたものである[2]。動作時電力と待機時電力はチャンネル長の微細化とともに増加するが、特に待機時電力が急激に増加している。このまま何も対策を講じなければ数 10nm の世代以降では待機時電力が動作時電力と同等あるいはそれ以上にまで増大する。したがって、今後の CMOS ロジックシステムの発展には、待機時電力の削減が重要な課題の1つとなる[3]。

1.2 待機時電力の起源

待機時電力は MOSFET のリーク電流に起因する。MOSFET には図 1.2.1 に示すように様々なリーク電流が生じるが、特にゲートリークとサブスレッショルドリークの 2 つが重要である[4]。ゲートリークはゲート酸化膜の薄膜化にともない、ゲート酸化膜をキャリアがトンネルすることで生じる。サブスレッショルドリークは温度でその下限が決まってしまうリーク電流で、チャンネル長の微細化により顕在化する。これは、反転チャンネルを形成していないにも関わらず、キャリアがソース-チャンネル間のポテンシャル障壁を熱的に乗り越えることによって発生するリーク電流である。

1.2.1 ゲートリーク

これまでの CMOS 技術では、信頼性の高いゲート絶縁膜として用いられたものは Si を熱酸化して形成した SiO_2 であった。しかし、この場合では、MOSFET のゲート長が数 10nm を下回るようになると、スケーリングによるゲート酸化膜の薄層化によって、ゲートリ

ークが無視できないほど増大するようになる。特に SiO₂ の膜厚が 1.5nm を切ると、ゲートリークの電流密度が許容できないレベルに到達する[5]。ゲートリークは MOSFET の電流駆動能力の劣化を招き、さらに、先に述べた待機時電力の増大を引き起こす。ゲートリークは次式で表される[6]。

$$I_{ox} = K_G W (V_{ox}/T_{ox})^2 \exp(-\alpha T_{ox}/V_{ox}) \quad (1.1)$$

ここで、 K_G 、 α は材料やプロセスで決まる定数、 W はチャネル幅、 V_{ox} はゲート電圧、 T_{ox} はゲート酸化膜の膜厚である。式(1.1)から、ゲートリークを低減するにはチャネル幅を狭くする、電源電圧を下げる、ゲート酸化膜の膜厚を厚くするという 3 つの方法があることがわかる。このうち、チャネル幅や電源電圧は回路性能やシステムの制約上自由に変えることは難しい。ゲート酸化膜については、膜厚を厚くするほどゲートリークは減少するが、同時に MOSFET のゲート容量の低下により電流駆動能力が劣化する(したがって、この場合も回路性能が劣化する)ため、厚膜化は容易ではない。そのため、ゲート酸化膜を薄膜化せずに十分な MOSFET の電流駆動能力を実現させる技術が必要となる。そこで、high-k 絶縁材料と呼ばれる誘電率の高い絶縁材料でゲート絶縁膜を構成する方法が開発されてきた。この high-k 絶縁材料をゲート絶縁膜に用いれば、厚いゲート絶縁膜でも薄い SiO₂ 膜と同等にチャネルにキャリアを誘起できるため、電流駆動能力を劣化させることなくトンネル電流を低減できる。現在ではゲート酸化膜として HfO₂ 系絶縁膜が用いられている。HfO₂ 系絶縁膜による MOS 構造は堆積膜ではじめて実用化された MOS ゲート構造で、十分なトランジスタ性能を実現できる良好な界面構造を有し、また CMOS プロセスとの親和性もよい。このインパクト(特に堆積膜で実現したという意義)は大きく、その後の Ge や III-V 族化合物半導体などの非 Si チャネルの MOSFET 開発に極めて重要な役割を果たしている。以上のようにゲートリークの低減には high-k 絶縁膜によるデバイス技術(ゲートスタック技術)による対策が有効である。図 1.2.2 に示すように high-k 絶縁膜の導入により、待機時電力におけるゲートリークの問題はその影響がほとんど生じない程度までに大きく改善した[6]。

1.2.2 サブスレッショルドリーク

上述した high-k ゲート絶縁膜の導入後の待機時電力の主な原因は MOSFET のサブスレッショルドリークである。これは次式で表される。

$$I_{sub} = I_S \exp [(V_{GS} - V_{th})/nV_{\theta}] [1 - \exp(-V_{DS}/V_{\theta})] \quad (1.2)$$

ここで、 I_s は温度、デバイスサイズ、基板濃度に依存する定数、 V_{GS} と V_{DS} はそれぞれゲートバイアスとドレインバイアス、 V_{th} はしきい値電圧、 V_θ は熱電圧 ($=k_B T/q$)、 n はデバイスの各種キャパシタンスから決まる定数であるが理想的な場合には 1 となる。図 1.2.3(a) に、ドレイン電流 (I_{DS}) のゲートバイアス (V_{GS}) 依存性の対数プロットを示す。しきい値以下の領域では I_{DS} は V_{GS} に対して指数関数的に減少し、特性の傾きの逆数はサブスレッショルドスイング S と呼ばれ、次式で与えられる。

$$S = (\ln 10 \cdot nk_B T)/q \approx 60n \text{ mV/decade} \quad (1.3)$$

$n \sim 1$ の場合、 S は温度で決まり、デバイス技術では削減できない限界値となる。室温におけるこの限界値は 60 mV/decade である。以上から、図 1.2.3(b) に示すように V_{th} がサブスレッショルドリーク削減のための重要なパラメータとなる。しかし、 V_{th} を大きくするとリークは削減できるが、次式で与えられる遅延時間が劣化する(図 1.2.3(c))。

$$t_{pd} = (C_L V_{DD})/I_D \approx C_L V_{DD}/K(V_{DD} - V_{th})^\alpha \quad (1.4)$$

α は実験から経験的に得られる定数でおよそ 1.3 程度の値であり、 V_{DD} は電源電圧である[6]。したがって、図 1.2.3(b), (c) のように、 V_{th} によるサブスレッショルドリークの削減は、回路性能とトレードオフの関係にあり、サブスレッショルドリークの削減と回路の速度性能の向上の両立は困難である(同様に電源電圧を低電圧化してリークを削減することもできるが、この場合でも遅延時間は劣化する。またシステムの要求から、これを自由に変えることが難しい場合が多い)。

このようなジレンマを解消して、リーク電流の削減と回路性能の向上を両立する方法が考えられている[7]。デュアル V_{th} と呼ばれる方法では、回路内のクリティカルパス(回路の遅延を決定するパス)におけるトランジスタの V_{th} のみを下げて遅延の劣化を防ぎつつ、非クリティカルパスのみ V_{th} を上げてサブスレッショルドリークを削減する[8]。また、VTCMOS という方法では、低い V_{th} で設計しておき、基板バイアスによってダイナミックに実質的な V_{th} を変調してリークを削減する[9]。以上の方法は、回路性能をできるだけ維持しつつ、しきい値を大きくすることでリーク電流を削減する方法である。一方、これとは異なるアプローチとして、待機時に回路の電源遮断を行うことでサブスレッショルドリークを削減する方法もある。これはパワーゲーティング(PG)と呼ばれ、回路をパワーメインと呼ばれるブロックに分割して、パワーメインの電源供給をパワースイッチ(スリープトランジスタ)と呼ばれるトランジスタで制御する方法である[10]。パワーメインが待機状態となるときに、パワースイッチによってパワーメインの電源を遮断する。パワースイ

ッチによって電源遮断されたパワードメインには、遮断状態のパワースイッチとの分圧で決まる低い電圧しかかからないため、遮断時のリーク電流を削減できる。さらに、ウェルの電氣的接続をうまく選んでおくことで、電源遮断時に自動的にリーク電流が減るように基板バイアスがかかるようしておくこともできる。PG は現在 μP や SoC に標準的に搭載されている有効性の高い待機時電力の削減方法である[11, 12].

1.3 パワーゲーティングとその課題

1.3.1 構成方法

PG では CMOS ロジックシステムをパワードメインに分割し、パワードメインごとにパワースイッチを配置することで、各ドメインを独立にパワーマネジメントする。よって、使用していない(待機時の)パワードメインを電源遮断することができるため、待機時電力の削減が可能になる。

図 1.3.1 にパワースイッチの導入方法を示す。初期の PG の提案では同図(a)のようにパワードメインの電源線と接地線をそれぞれ p チャネル MOSFET (以下, pMOS と略記する)と n チャネル MOSFET (以下, nMOS と略記する)で構成されるパワースイッチで電氣的に分離できる方式となっていたが、現在では、同図(b), (c)に示すように電源線側または接地線側のみにパワースイッチを入れる方式が採用されることが多い。これらは、それぞれ仮想電源電圧アーキテクチャ、仮想接地アーキテクチャと呼ばれる。仮想電源電圧アーキテクチャで用いられるパワードメインと電源線の間には挿入されるパワースイッチはヘッダパワースイッチ、仮想接地アーキテクチャの接地線とパワードメインとの間に挿入されるパワースイッチはフッタパワースイッチと呼ばれる。基本的にはヘッダパワースイッチは pMOS, フッタパワースイッチは nMOS で構成する。一般に、パワースイッチの V_{th} はパワードメイン内のトランジスタより大きくなるように設計する。ヘッダパワースイッチとパワードメインの接続点における電位を仮想電源電圧または仮想 V_{DD} (Virtual V_{DD} , 本論文では VV_{DD} と略記する), フッタパワースイッチとパワードメインの接続点の電位を仮想接地電圧 (Virtual ground voltage, 本論文では VV_{GND} と略記)と呼ぶ。GND は接地を表しその電位は V_{GND} と表すが、習慣にしたがって GND を V_{GND} の意味で代用することもある。ただし仮想接地の電位は VV_{GND} と表す。

PG における電源遮断時のリーク電流は、先に述べたように遮断状態にあるパワースイッチとパワードメインの分圧で決まるため、パワースイッチの電源遮断時における抵抗はできるだけ大きいほうが良い。しかし、このためにパワースイッチのサイズを小さくすると、パワースイッチを導通させたときのパワースイッチによる電圧降下のため V_{DD}/V_{GND} が低下/上昇してしまい、回路性能を劣化させてしまう。そこで、パワースイッチを導通させたときに V_{DD}/V_{GND} を劣化させないように大きく設計しておき、電源遮断時にはパワースイッチをできるだけ深く遮断して遮断抵抗をできるだけ大きくするスーパーカットオフ (SCCMOS) という技術が提案されている[13]。仮想電源電圧アーキテクチャでは電源電圧以上でヘッダパワースイッチ (pMOS) を、仮想接地アーキテクチャでは負の電圧でフッタパワースイッチ (nMOS) をドライブして深く遮断する。負バイアスを用いることなくスーパーカットオフを実現できるため、スーパーカットオフには仮想電源電圧アーキテクチャがよく用いられる。これとは対照的に、パワースイッチを電源遮断時に十分にリークが削減できるように設計しておき、導通時にはこのパワースイッチをオーバードライブする BGMOS と呼ばれる技術も提案されている[14]。この技術をヘッダパワースイッチ (pMOS) に用いれば、負バイアスが必要になるが、フッタパワースイッチ (nMOS) では負バイアスを用いることなく実現できる。

1.3.2 性能制約要因

PG の待機時エネルギーの削減効率をさらに向上させるためには、PG の実行頻度を増やす必要がある。PG の実行頻度はその粒度で評価される。時間的・空間的粒度の粗い (実行頻度の少ない) PG を粗粒度 PG、時間的・空間的粒度の細かい (実行頻度の多い) PG を細粒度 PG と呼ぶ。細粒度の PG を実現するためには、短い待機時間であっても PG を実施できる必要があるが、これは技術的に容易ではない (後述)。したがって、多くの場合で比較的容易に実現できる粗粒度の PG が用いられている。待機時エネルギーの削減効率が高い PG の実現には、理想的にはシステム/アプリケーションに合わせて最適化された粒度の PG を行う必要があるため、粗粒度の PG のみならず、細粒度の PG も必要となる。しかし、実際にはロジックシステム内の各種記憶回路の情報が揮発性であることに起因する要因から、細粒度の PG の実現は難しく、一般に PG はその性能を発揮しきれていない。

次に、PG の各種記憶回路に起因する制約要因について述べる。μP や SoC などのロ

ジックシステムではキャッシュやレジスタといった各種記憶回路に static random access memory (SRAM) や flip flop (FF) といった揮発性の双安定記憶回路が用いられている。したがって、PG によって電源遮断を行うと、これらの記憶回路に記憶されているデータが消失してしまう。記憶回路に記憶されているデータの中には、パワードメインの電源復帰後に必要となる情報も含まれているため、PG を実施する場合にはこのデータを何らかの方法で保持する必要があるが、このためには余計なエネルギーとレイテンシが発生する。これが PG の粒度に大きな制約を与える。

1.3.3 CMOS ロジックシステムにおけるパワーゲーティングの課題

図 1.3.2(a) に PG の観点から見たマルチコアプロセッサの構成例を示す。複数のコアとラストレベルキャッシュ(LLC)から構成され、コア内には各種ステイト/コンフィグレーション用のレジスタと高階層キャッシュ(HLC)が配置されている。近年の μ P では、コア単体の処理速度はほぼ飽和しており、マルチコア化によって高性能化が実現されている。このようなマルチコアのロジックシステムでは、常にすべてのコアを動作させているわけではなく、待機状態となっているコアも少なくない。したがって、このようなコアの待機時電力を削減することが重要になる。コア単位の PG のことをコアレベルの PG と呼ぶ。コアレベルの PG を実行する場合、レジスタについては、電源遮断を行わない電源線を別途用意してデータを保持するリテンション FF を用いるデュアルパワーレイルと呼ばれる方法[15]や、電源非遮断領域に作り込まれたバックアップ用の FF にデータを転送してデータを保持する方法[12]などが用いられている。また、コア内の高階層キャッシュについては、データをすべて破棄して電源遮断を行うフラッシュが用いられる[11]。この場合では、データをすべて破棄してしまうため、PG を実施できるチャンスは限られる。

図 1.3.2(b)に同図(a)に示したプロセッサにおける PG の効果を簡単なモデルを用いて評価した結果を示す。レジスタのステイト/コンフィグレーションの保持にはリテンション FF を用い、HLC, LLC には通常の 6T-SRAM を用いた(その他の詳細は第 5 章参照)。コアレベルの PG を導入することで待機時電力はコアの利用率に応じて効果的に削減できる。ただし、このコアレベルの PG では HLC のデータをフラッシュ(破棄)してしまうため、HLC にあるデータが不要になる場合や、コアの電源復帰後に下層のキャッシュからデータを書き戻しても十分にエネルギーを削減できる場合に限られる(このようなデータの転送には大きなエネルギーとレイテンシを要することに注意)[11]。したがって、PG を実

施できるチャンスは限られる。すなわち PG の粒度は粗く、高効率に待機時電力を削減することが難しい。また、HLC のデータを保持するために SRAM のデータを消失しない程度に電源電圧を下げるスリープという技術[16]を用いることも可能であるが、この場合通常電圧のスタンバイ状態に比べると電力は削減できるものの、上述のフラッシュと比べると電力削減率は小さい。

LLC についても上述のスリープを用いる場合と、HLC と同様にデータを破棄するフラッシュを用いる場合がある。最近の μP や SoC ではチップ上の LLC の占める面積が全体の半分程度にも及び、待機時電力の影響が大きい[17]。このため LLC の待機時電力の削減は極めて重要となる。スリープモードはキャッシュに広く用いられている待機時電力の削減技術であるが、図 1.3.2(b)に示すように電源遮断ほどの待機時電力削減の効果はない。一方、フラッシュを用いる場合、完全な電源遮断による PG を実施できるため、待機時電力の削減効果は大きい。しかし、この場合でも HLC と同様に PG の実行頻度が落ちる。特に、LLC はすべてのコアで共有されているため、データのフラッシュの機会は限られる(粒度は相当に粗い)。以上のようにロジックシステムでは、各種記憶回路内のデータの保持が PG のエネルギー削減効率に大きな制約を与えている。特に、キャッシュのデータ保持が PG の重要な性能の制約要因になっている。したがって、PG による電源遮断を行っても記憶回路、特にキャッシュのデータを失うことなく保持し続けることができれば、PG の実行頻度を高め、高効率に待機時エネルギーを削減することが可能になると考えられる。

1.4 キャッシュへの不揮発記憶の導入方法

上述のように CMOS ロジックシステムにおける PG の待機時エネルギー削減効率の重要な制約要因の 1 つは、キャッシュのデータの保持である。そこで、この課題を解消するために、キャッシュを不揮発化して、電源遮断時にもデータを消失しないようにすることが考えられる。これには、(i) MRAM, PRAM, ReRAM 等の新しい不揮発性メモリ (Emerging nonvolatile memories)を用いる方法、(ii) 不揮発性メモリ素子を用いて不揮発化した SRAM (NV-SRAM)を用いる方法、(iii) CMOS のみで構成し、PG の電源遮断時に実質的に不揮発記憶と見なせる擬似不揮発性 SRAM (VNR-SRAM)を用いる方法が考えられる。 μP や SoC のキャッシュは SRAM で構成され、極めて高速に動作する

記憶回路であることから、不揮発化する際には既存のキャッシュの回路性能を劣化させずに不揮発化する必要がある。以下にキャッシュを不揮発/擬似不揮発化するこれらの技術の概要について述べる。

(i) 既存の不揮発性メモリ

近年, MRAM, PRAM, ReRAM といった抵抗変化型の不揮発性メモリの研究・開発は目覚ましく(近年, MRAM では書き込み方式の物理現象を付記して STT-MRAM などと表記することが多いが本論文では書き込み方式に依らず MRAM という用語を用いる), これまでに, このような不揮発性メモリによるキャッシュ(SRAM で構成)やメインメモリ(DRAM で構成)の置き換えが検討されている[18-21]. これらは基本的には不揮発性メモリ素子と選択用のトランジスタで構成され, 不揮発性メモリ素子の材料・構造や動作原理に違いはあるが, 不揮発性メモリとしての動作は概ね同等である. すなわち, 不揮発に保持できる抵抗値を情報の 0/1 に対応させ, この抵抗の違いを電氣的に読み取る. したがって, 読み出し動作・書き込み動作は常に不揮発性メモリ素子を用いて行うことになるが, 一般にこのような不揮発性メモリ素子の動作速度は, SRAM のような双安定回路に比べて遅い(特に不揮発性メモリ素子の書き込みに要する時間(ストアレイテンシ)が問題となる). MRAM などでは LLC に応用できるような動作速度のものも開発されているが, コア内のキャッシュに用いられるほどの速度性能を有していない. さらに, これらの不揮発性メモリのキャッシュ応用を困難にしているのは, 不揮発性メモリ素子にデータを書き込むときに生じる大きなエネルギー(ストアエネルギー)である. 不揮発性メモリの導入によって待機時電力が削減されても, 動作時電力が大幅に増加する. このため, 不揮発性メモリのストアエネルギーや動作速度(特にストアレイテンシ)を削減する技術が検討されている[21]. 以上から, μ P や SoC などの CMOS ロジックシステムの PG に不揮発記憶を導入する場合, 速度および電力・エネルギー性能から上述の不揮発性メモリを用いることは容易ではない.

(ii) 不揮発性メモリ素子を用いた不揮発性 SRAM

キャッシュを不揮発性メモリ素子に置き換えるのではなく, キャッシュを不揮発化した SRAM(NV-SRAM)で構成する試みがある. 詳細は第 3 章で述べるが, CMOS 双安定記憶回路に不揮発性メモリ素子を接続することで双安定記憶回路を不揮発化すること

が可能となる。ただし、ここで注意をしなければならないのは、不揮発記憶の機能のみを考慮して SRAM を不揮発化すると、その回路性能は大きく劣化してしまうことである [22-25]。PG の応用では、電源遮断の直前と電源復帰直後のみに不揮発記憶を実現できればよく、通常の SRAM 動作に不揮発記憶は不要である。このような構成をとれば、通常の SRAM 動作に不揮発性メモリ素子を用いないことから、上述のストアエネルギーやストアレイテンシの問題を最小限に抑えることができる。しかし、これまで提案された多くの NV-SRAM はこのような観点から PG に適した回路構成になっていない(通常の SRAM 動作と不揮発記憶の分離ができない)。本論文では本著者の指導教員らによって提案された NV-SRAM について検討を行う。この NV-SRAM は通常の SRAM 動作と不揮発記憶の動作を電氣的に分離できることから PG に適していると考えられる。ただし、PG の性能をより高めるためには、不揮発性メモリ素子のストアエネルギーとストアレイテンシを最小限に抑える技術開発が必要となる。

(iii) CMOS のみによる擬似不揮発性 SRAM

NV-SRAM では不揮発性メモリ素子を CMOS ロジックに集積化する必要がある。このようなエンベディッド技術は容易ではなく、コストもかなりかかる。そこで、不揮発性メモリ素子を用いることなく CMOS のみで実質的に不揮発記憶を実現することが必要である。また、不揮発性メモリ素子を用いないことから、上述のストアエネルギー/レイテンシに関する問題は生じないといったメリットもある。この方法には、CMOS のみで如何に擬似的な不揮発記憶を実現するかが重要な課題となる。本論文の提案方法は超低電圧(電源電圧のわずか 15%程度)でデータを保持することで、電源遮断時と同等の待機時電力の削減を実現するもので、PG においては実質的に不揮発記憶と見なすことができる。通常の電源電圧が 1.2V のとき、0.2V 程度の超低電圧でリテンションを行うことで 90%以上の待機時電力を削減できる可能性がある。超低電圧でデータを保持する技術はすでに知られており、シュミットトリガインバータで双安定回路を構成することで実現できる [26]。しかし、この場合では、通常電圧の SRAM 動作時における待機時電力の増大といった問題を生じてしまう。そこで、シュミットトリガインバータのバイアスを切り替えることで、通常の SRAM 動作ではリーク電流を抑えて既存の 6T-SRAM と同等に動作するモードと超低電圧で安定にリテンション動作(データ保持)できるモードを有する擬似不揮発性 SRAM(VNR-SRAM)を提案する(詳細は第 4 章で述べる)。

1.5 不揮発記憶を用いたその他の待機時電力削減技術

以上ではPGにおける不揮発記憶の有用性とその課題について述べた。このような不揮発記憶を用いたPGのことを特に不揮発性パワーゲーティング(NVPG)と呼ぶことがある。また、このNVPGの他に不揮発記憶を用いた待機時電力削減の方法としてノーマリオフ(NOF)と呼ばれるものがある。NOFはMRAM, PRAM, ReRAMなどの新しい不揮発性メモリの出現とともに活発に議論・検討されるようになった。これには、これらのメモリ素子がCMOSロジックシステムにおける待機時電力の問題が顕在化し始めた時期に華々しく登場したことも大きく関係している。以下、 μ PやSoCなどのロジックシステムへの応用を念頭において、NOFの概要と、これと比較してNVPGの特徴と優位性を述べる。

以下、提案当初のNOFアーキテクチャの概要について課題を交えて述べる。図1.5.1(a)にNOFアーキテクチャの概念図を示す。NOFアーキテクチャではロジックシステムの動作を行わない場合には常に回路/システムを電源遮断しておき、待機時電力を削減する技術である。したがって、NOFを実現する場合でも問題となるのはロジックシステム内の記憶回路である。上述のように、通常、ロジックシステム内の記憶回路はCMOS双安定回路で構成されているため、電源遮断を行うとその記憶情報が失われてしまう。よって、ロジックシステム内の記憶回路を不揮発性メモリなどで不揮発化することでこのアーキテクチャを実現することが考えられていた[27]。この場合、ロジックシステム内の記憶回路では、常に不揮発性メモリ素子を用いて、データの読み出し/書き込みを行うことになることから、NOFでは以下のような問題を生じる。

NOFに用いることが想定されていたMRAMなどの抵抗変化型の新しい不揮発性メモリは、開発当初ではSRAM並みの高速動作が期待されていた。しかし、その後の研究開発によって、これはほぼ困難であると理解されるようになった(特にストアレイテンシ)。また、このような不揮発性メモリ素子の書き込みに要するストアエネルギーはSRAMに比べて非常に大きく、通常時の動作にも不揮発性メモリ素子を用いてしまうと待機時電力が削減できても動作時電力が増大するといった問題が生じる。また、短い待機の状態が生じる度に電源遮断/供給を行うと、その動作にともなうエネルギーのオーバーヘッドが大きく、NOFで当初考えられていたような短時間での電源遮断は意味がない(このオーバーヘッドが待機時に削減できるエネルギーを上回る)。さらに、突入電流を低く抑えるため、そもそも電源遮断/供給に要するレイテンシは短くすることは難しい(通

常は数 μs 以上[28, 29]).

最近では MRAM などのストアエネルギーレイテンシは, 各種デバイス技術などの開発によって小さくなってきているが, その限界も理解されるようになった(ごく最近, MRAM ではストアレイテンシを SRAM 並みに短くできることが示されたが, この場合ではストアエネルギーが大幅に増加してしまう[30]). そのため, このような不揮発性メモリのロジックシステムへの応用/検討は概ね LLC, メインメモリ(DRAM)用のキャッシュ, メインメモリに限定されている. したがって, μP や SoC などのロジックシステム分野では, もはや NOF という概念はほとんど使われなくなっている. ただし, NOF は一部のマイクロコントローラなどでは検討されているようである.

図 1.5.1(b)に NVPG の概念を示す[31]. このアーキテクチャは不揮発記憶を効果的に利用して, PG を細粒度化するもので, 本著者の指導教員らによって提案された. したがって, 通常の動作を行っている期間は電源遮断を行わず, 意味のある長さの待機状態について電源遮断を行う. 先に述べたように不揮発記憶を用いることで, その粒度を細かくできると考えられる. NVPG では, 通常の動作時に不揮発記憶を用いる必要はなく, 電源遮断を行うときのみ不揮発記憶を用いればよい. 不揮発性メモリ素子を電源遮断時にしか用いないため, 不揮発記憶に要するストアエネルギーを必要最低限に抑えることができる. また, 不揮発性メモリ素子のストアレイテンシは通常の動作には影響しない. ストアエネルギーレイテンシを最小限に抑えることで, 電源遮断によってエネルギーを削減するために最低限必要となる電源遮断時間 (Break-even time) を短くできる. すなわち, NVPG の細粒度化が可能となる.

NVPG を実現するためには, 通常の動作時には従来通りの高性能な記憶回路として動作し, 電源遮断時にのみ不揮発記憶を行える記憶回路が必要になる. このような記憶回路は CMOS 双安定回路と不揮発性メモリ素子を組み合わせることで実現できると考えられるが, 通常動作では不揮発性メモリ素子の影響を完全に排除できることが重要になる(第 3 章で詳述する). また, NVPG は先に述べた擬似不揮発記憶を用いても実現できる. 以後, 本論文では特に NVPG といった用語は用いず, 不揮発記憶を用いた PG や, 擬似不揮発記憶を用いた PG などと称することにする.

1.6 本研究の目的と本論文の構成

以上から、本論文では不揮発/擬似不揮発記憶を用いた SRAM 技術を開発し、この技術を用いて高効率に待機時エネルギーを削減できる PG 技術を開発する。そして、CMOS ロジックシステムにおける不揮発/擬似不揮発記憶を用いた PG の有用性・有効性を示し、CMOS ロジックの低消費電力化を実現できることを明らかにする。

そのための課題と本論文の構成を以下にまとめる。

- ・不揮発/擬似不揮発記憶を効果的に利用するには、これらの導入によるエネルギー的な損得を評価する必要がある。第 2 章では、この評価指標である損益分岐時間 (Break-even time) の定式化と解析を行い、不揮発/擬似不揮発記憶を用いた PG の性能評価指標を構築する。また、NV-/VNR-SRAM の設計指標を示す。

- ・不揮発記憶を用いた PG では、上述のように不揮発性メモリ素子のストアエネルギーとストアレイテンシを削減する技術開発が課題となる。第 3 章ではこのストアエネルギー/レイテンシを究極的に削減する NV-SRAM 技術を確立する。

- ・擬似不揮発記憶を用いた PG では、超低電圧における安定なデータ保持と不揮発記憶と同等の待機時エネルギーの削減効果を実現することが課題となる。第 4 章では超低電圧での安定性を飛躍的に高め、電源遮断と同程度の電力削減を可能にする VNR-SRAM 技術を確立する。

- ・第 5 章では、NV-/VNR-SRAM を用いて構成した μP における PG 性能を評価し、その有用性・有効性を示す。また、本論文における解析結果から NV-/VNR-SRAM のどちらが今後のプロセッサ応用に適した SRAM 技術であるかを議論する。

なお、本論文では不揮発/擬似不揮発記憶を用いた PG の対象をサーバや PC などに用いる μP に設定し、これに搭載されたキャッシュへの適用を目的とする。このため、キャッシュに用いる NV-/VNR-SRAM としては、GHz 動作が前提となる。したがって、ここで採用するアーキテクチャはリファレンスとする 6T セルを含めて、このような動作が可能なものに限定する。

1.7 図と表

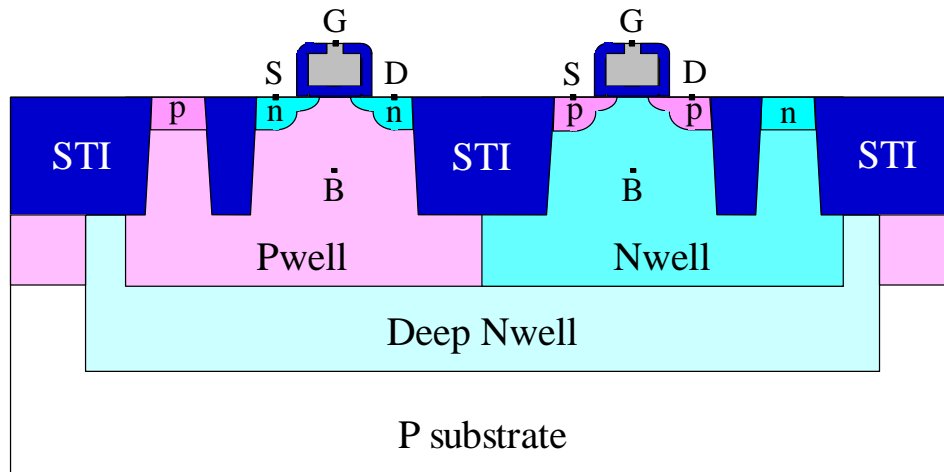


図 1.1.1 MOSFET のデバイス構造

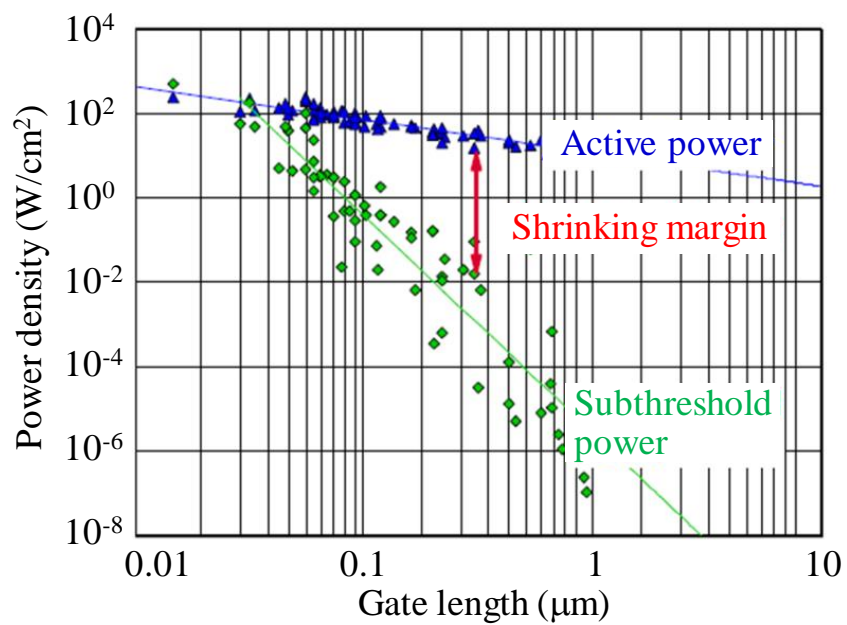


図 1.1.2 CMOS ロジックシステムにおける消費電力密度のゲート長依存性

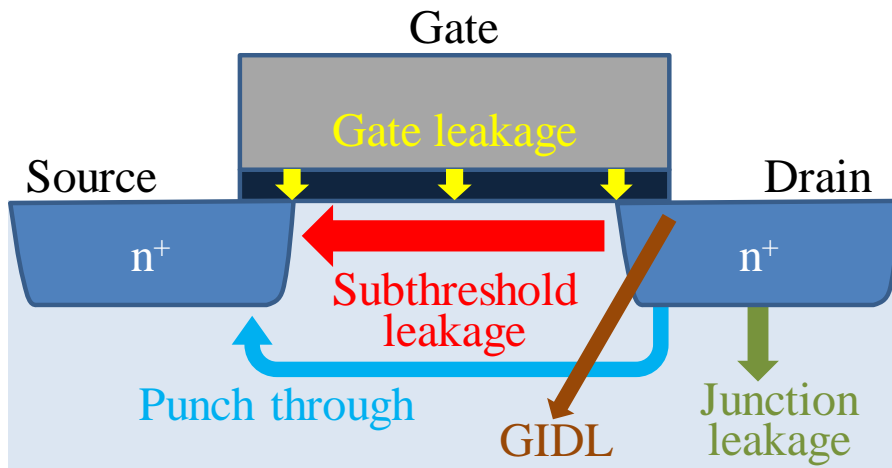


図 1.2.1 MOSFET に生じる各種リーク電流

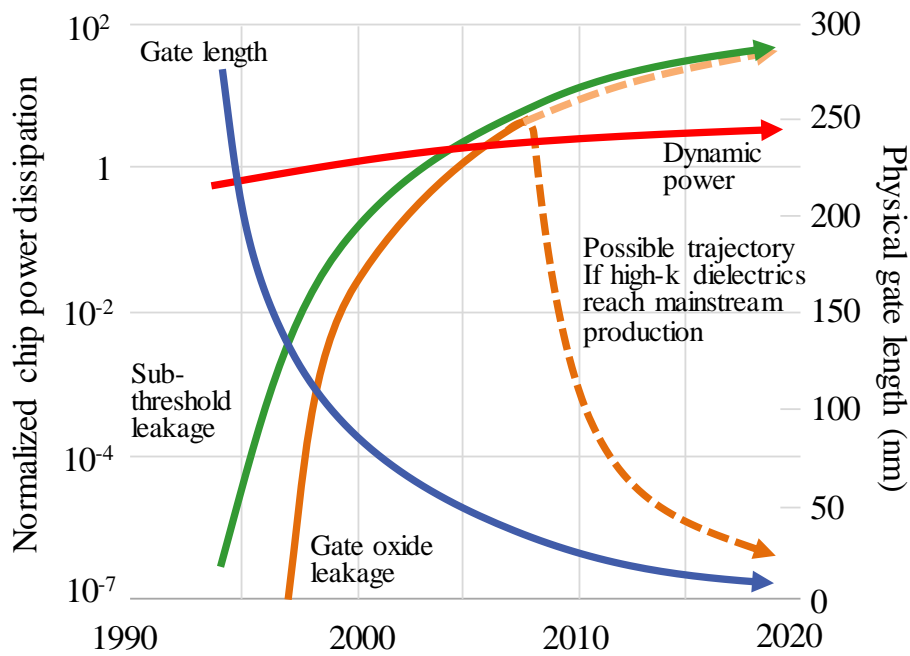


図 1.2.2 ゲートリークとサブスレッショルドリーク

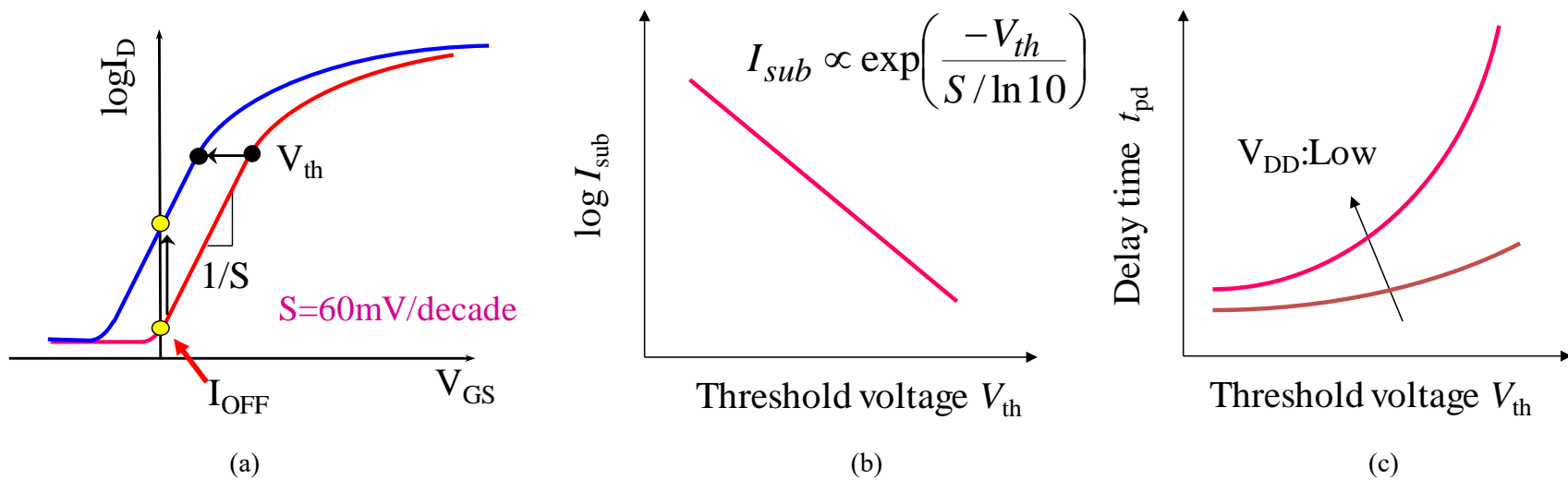


図 1.2.3 (a)ドレイン電流 (I_D) のゲートバイアス (V_{GS}) 依存性, (b)サブスレッショルドリークのしきい値 (V_{th}) 依存性, (c)遅延時間 (t_{pd}) のしきい値 (V_{th}) 依存性

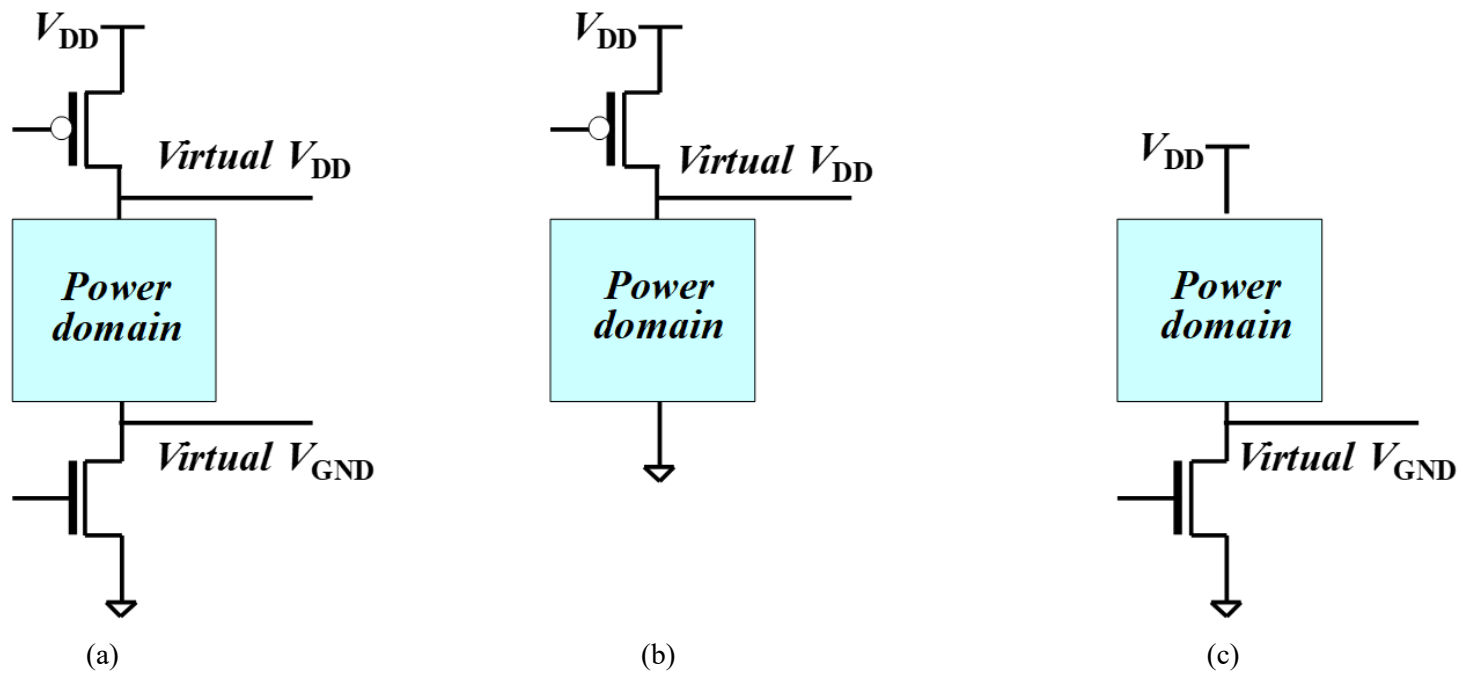
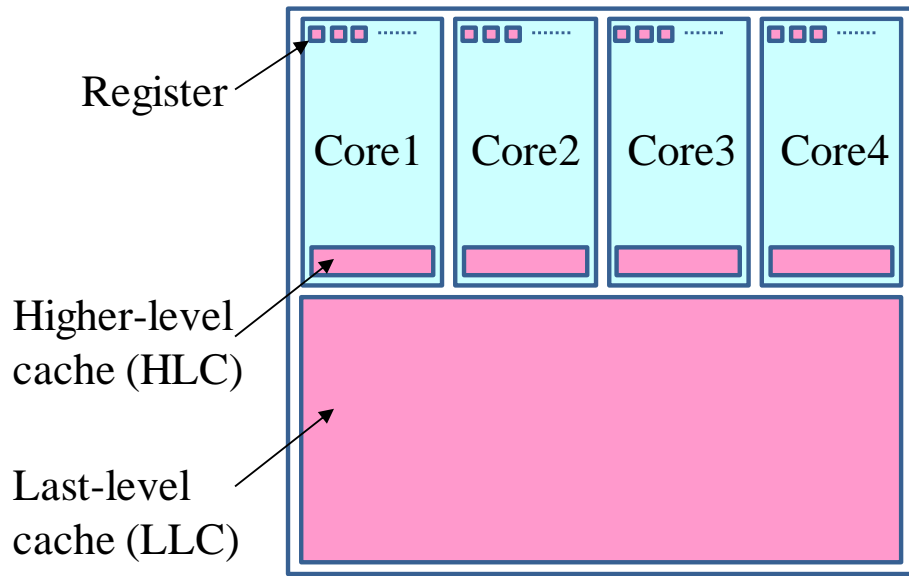
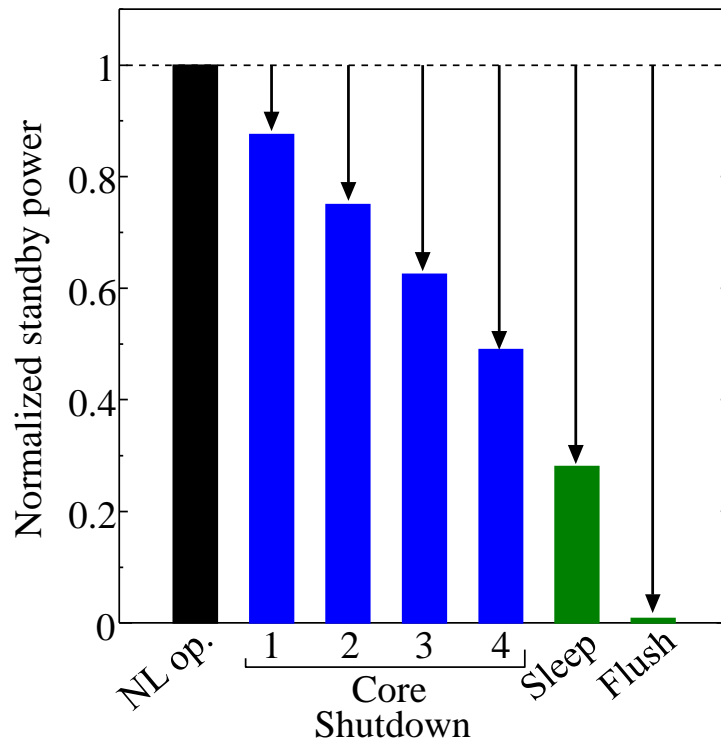


図 1.3.1(a)ヘッダ PS, フッタ PS を両方用いる構成, (b)仮想電源電圧アーキテクチャ, (c)仮想接地アーキテクチャ

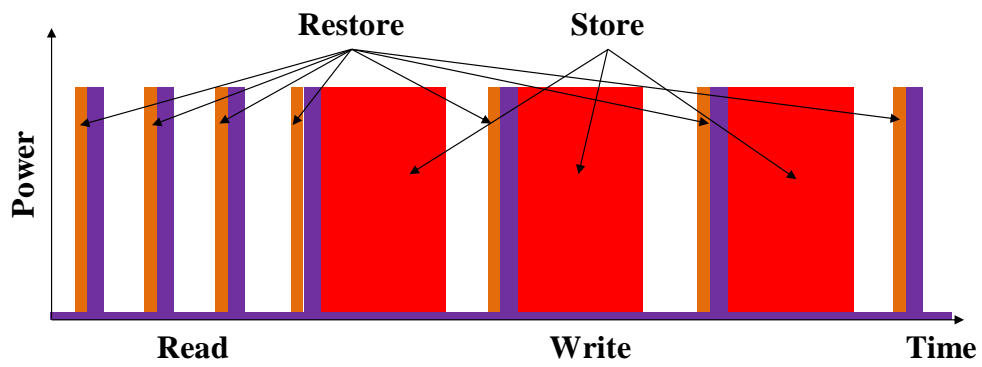


(a)

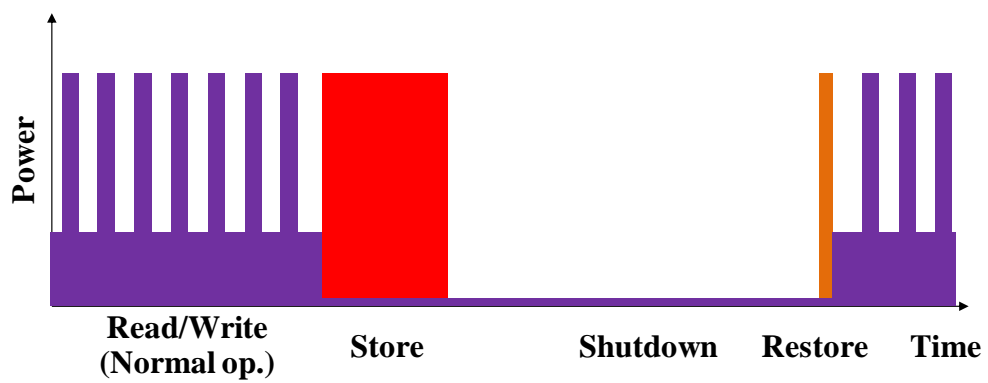


(b)

図 1.3.2 (a)マルチコアプロセッサの概念図と(b)コア、LLCをPGした場合のスタンバイ電力(NL op.は通常のSRAM動作(Normal operation)を、Sleep/Flushはコアを遮断した状態でLLCをスリープ/フラッシュした状態を示す)



(a)



(b)

図 1.5.1 (a)NOF アーキテクチャと(b)NVPG アーキテクチャの概念図

1.8 参考文献

- [1] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc., “Design of ion-implanted MOSFET's with very small physical dimensions”, IEEE Journal of Solid-State Circuits **9**, 256 (1974).
- [2] R. Puri, S. Leon, and B. Subhrajit, “Keeping hot chips cool”, 42nd Design Automation Conference 2005, 285 (2005).
- [3] ITRS2013, International Technology Roadmaps for Semiconductors, 2013.
- [4] K. Roy, M. Saibal, and M. Hamid, “Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits”, Proceedings of the IEEE **91**, 305 (2003).
- [5] S.-H. Lo, D. A. Buchanan, Y. Taur, and W. Wang, “Quantum-mechanical modeling of electron tunneling current from the inversion layer of ultra-thin-oxide nMOSFET's”, IEEE Electron Device Lett. **18**, 209 (1997).
- [6] N. S. Kim, T. M. Austin, D. Blaauw, T. N. Mudge, K. Flautner, J. S. Hu, M. J. Irwin, M. T. Kandemir, and N. Vijaykrishnan, “Leakage Current: Moore’s Law Meets Static Power”, IEEE computer **12**, 68 (2003).
- [7] Y. Shin, J. Seomin, K. M. Choi, and T. Sakurai, “Power gating: Circuits, design methodologies, and best practice for standard-cell VLSI designs.”, ACM Transactions on Design Automation of Electronic Systems **15**, 28 (2010).
- [8] L. Wei, Z. Chen, M. Johnson, K. Roy, and V. De, “Design and optimization of low voltage high performance dual threshold CMOS circuits”, 35th annual Design Automation Conference 1998, 489, (1998).
- [9] T. Kuroda, T. Fujita, S. Mita, T. Nagamatu, S. Yoshioka, F. Sano, M. Norishima, M. Murota, M. Kako, M. Kinugawa, M. Kakumu, and T. Sakurai, “A 0.9 V 150 MHz 10 mW 4 mm² 2-D discrete cosine transform core processor with variable-threshold-voltage scheme”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 166 (1996).
- [10] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, “1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS”, IEEE Journal of Solid-state circuits **30**, 847 (1995).
- [11] V. George, S. Jahagirdar, C. Tong, K. Smits, S. Damaraju, S. Siers, V. Naydenov, T. Khondker, S. Sarkar, and P. Singh, “Penryn: 45-nm next generation Intel® core™ 2 processor”, IEEE Asian Solid-State Circuits Conference 2007, 14 (2007).
- [12] Y. Kanno, H. Mizuno, Y. Yasu, K. Hirose, Y. Shimazaki, T. Hoshi, Y. Miyairi, T. Ishii,

- T. Yamada, T. Irita, T. Hattori, K. Yanagisawa, and N. Irie, “Hierarchical power distribution with power tree in dozens of power domains for 90-nm low-power multi-CPU SoCs”, *IEEE Journal of Solid-State circuits* **42**, 74 (2007).
- [13] H. Kawaguchi, K. Nose, and T. Sakurai, “A CMOS scheme for 0.5 V supply voltage with pico-ampere standby current”, *IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, 192, 1998.
- [14] T. Inukai, M. Takamiya, K. Nose, H. Kawaguchi, T. Hiramoto, and T. Sakurai, “Boosted Gate MOS (BGMOS): Device/circuit cooperation scheme to achieve leakage-free giga-scale integration”, *IEEE Custom Integrated Circuit Conference (CICC) 2000*, 409 (2000).
- [15] V. Zyuban and S. V. Kosonocky, “Low power integrated scan-retention mechanism”, *IEEE International Symposium on Low Power Electronics and Design (ISLPED) 2002*, 98, (2002).
- [16] S. Rusu, S. Tam, H. Muljono, D. Ayers, J. Chang, “A dual-core multi-threaded Xeon processor with 16MB L3 cache”, *IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, 315 (2006).
- [17] N. Kurd, P. Mosalikanti, M. Neidengard, J. Douglas, and R. Kumar, “Next generation intel core™ micro-architecture (Nehalem) clocking”, *IEEE Journal of Solid-State circuits* **44**, 1121 (2009).
- [18] T. Endoh, H. Koike, S. Ikeda, T. Hanyu, and H. Ohno, “An overview of nonvolatile emerging memories—Spintronics for working memories”, *IEEE Journal on emerging and selected topics in circuits and systems* **6**, 109 (2016).
- [19] M. K. Qureshi, V. Srinivasan, “Scalable high performance main memory system using phase-change memory technology”, *J. A. Rivers, ACM SIGARCH Computer Architecture News* **37**, 24 (2009).
- [20] H. Akinaga and H. Shima, “ReRAM technology; challenges and prospects”, *IEICE Electronics Express* **9**, 795 (2012).
- [21] J. Boukhobza, S. Rubini, R. Chen, and Z. Shao, “Emerging NVM: A survey on architectural integration and research challenges”, *ACM Trans. Design Automat. Electron. Syst.* **23**, 1, (2017).
- [22] K. Abe, K. Nomura, S. Ikegawa, T. Kishi, H. Yoda, and S. Fujita, “Hierarchical nonvolatile memory with perpendicular magnetic tunnel junctions for normally-off computing”, *International Conference on Solid State Devices and Materials (SSDM) 2010*, F-9-3 (2010).
- [23] W. Zhao, C. Chappert, V. Javerliac, and J.-P. Noziere, “High speed, high stability and

- low power sensing amplifier for MTJ/CMOS hybrid logic circuits”, *IEEE Trans. Magn.* **45**, 3784 (2009).
- [24] T. Ohsawa, H. Koike, S. Miura, H. Honjo, K. Tokutome, S. Ikeda, T. Hanyu, H. Ohno, and T. Endo, “1Mb 4T-2MTJ nonvolatile STT-RAM for embedded memories using 32b fine-grained power gating technique with 1.0 ns/200ps wake-up/power-off times”, *IEEE symposium on VLSI circuits (VLSIC) 2012*, 46 (2012).
- [25] Y. Shuto, S. Yamamoto, S. Sugahara, “Analysis of static noise margin and power-gating efficiency of a new nonvolatile SRAM cell using pseudo-spin-MOSFETs”, *IEEE Silicon Nanoelectronics Workshop (SNW) 2012*, 1 (2012).
- [26] J. P. Kulkarni, K. Kim, and K. Roy, “A 160 mV robust Schmitt trigger based subthreshold SRAM”, *IEEE Journal of Solid-State Circuits* **42**, 2303 (2007).
- [27] T. Kawahara, “Scalable Spin-Transfer Torque RAM Technology for Normally-Off Computing”, *IEEE Design & Test of Computers* **1**, 52 (2010).
- [28] Y. Kanno, H. Mizuno, N. Oodaira, Y. Yasu, and K. Yanagisawa, “ μ I/O Architecture for 0.13- μ m Wide-Voltage-Range System-on-a-Package (SoP) Designs”, *Symp. on VLSI Circuits Dig. Tech. Pap.*, 168 (2002).
- [29] T. Kamei, M. Ishikawa, T. Hiraoka, T. Irita, M. Abe, Y. Saito, Y. Tawara, H. Ide, M. Furuyama, S. Tamaki, Y. Yasu, Y. Shimazaki, M. Yamaoka, H. Mizuno, N. Irie, O. Nishii, F. Arakawa, K. Hirose, S. Yoshioka, and T. Hattori, “A resume-standby application processor for 3G cellular phones”, *ISSCC Dig. Tech. Pap.*, 336 (2004).
- [30] G. Hu, J. J. Nowak, M. G. Gottwald, S. L. Brown, B. Doris, C. P. D. Emic, P. Hashemi, D. Houssameddine, Q. He, D. Kim, J. Kim, C. Kothandaraman, G. Lauer, H K Lee, N. Marchack, M. Reuter, R. P. Robertazzi, J. Z. Sun, T. Suwannasiri, P. L. Trouilloud, S. Woo, and D. C. Worledge, “Spin-transfer torque MRAM with reliable 2 ns writing for last level cache applications”, *IEEE International Electron Devices Meeting (IEDM) 2019*, 2.6 (2019).
- [31] S. Yamamoto and S. Sugahara, “Nonvolatile static random access memory using magnetic tunnel junctions with current-induced magnetization switching architecture”, *Jpn. J. Appl. Phys.* **48**, 043001 (2009).

第 2 章

不揮発/擬似不揮発性 SRAM の 性能評価指標

本章では、不揮発性 SRAM (NV-SRAM) および擬似不揮発性 SRAM (VNR-SRAM) を用いたパワーゲーティング (PG) の性能評価指標について述べる。はじめに、この定式化のために必要となるパワースイッチのアーキテクチャと NV-/VNR-SRAM のアレイの構成方法と動作の概略を述べる。そして、PG の性能評価指標として用いられる Break-even time (BET) と待機時電力の削減率を定義し、その定式化を行う。また、これらの削減方針や、本論文で用いる NV-/VNR-SRAM のアレイの性能解析モデルとその検証方法についても議論する。最後に NV-/VNR-SRAM の設計指標について述べる。

2.1 不揮発/擬似不揮発性 SRAM システムの構成方法

2.1.1 用語の説明

本節では不揮発性 SRAM (NV-SRAM) および擬似不揮発性 SRAM (VNR-SRAM) の性能評価指標を定義するために必要となるシステムの構成について説明する。はじめに、本論文で用いる略語について説明する。これらは主に図で用いるが本文中にも用いる。通常の SRAM 動作における読み出し、書き込み、スタンバイ、スリープ、電源遮断の各動作モードをそれぞれ、READ, WRITE, SB, SLP, SD と記す。本論文では単にスタンバイ (SB) と書く場合は通常電圧下のデータ保持状態を示し、スリープ (SLP) は双安定回路のデータが失われないうちに電源電圧を絞ったデータ保持状態を表すことにする。これらの各モードを特に指定しないときは単に `mode` と書き、その内容を指定する場合は `mode=READ, WRITE, SB` などと書く。また、通常の SRAM 動作を NL モ

ード (SD を除く上記動作モードの総称 ; normal operation) と書くことがある。NV-/VNR-SRAM では、電源遮断/超低電圧 (ULV) リテンションへの移行に要する動作モード、この状態から電源復帰への移行に要する動作モードをそれぞれ ENT, EXT と記す。特に NV-SRAM ではこれらをストア動作, リストア動作と呼び、それぞれ STR, RST と書くこともある。また、VNR-SRAM の ULV リテンションを ULVR と略記することがある。2つのモードを同時に表示する場合にはスラッシュを用いて SD/ULVR など表記する (これは SD または ULVR を意味する)。ENT, SD/ULVR, EXT をまとめて PG モードと呼ぶ。

2.1.2 パワースイッチアーキテクチャ

本節では本章の議論に必要となる NV-/VNR-SRAM のパワースイッチの構成方法についてその必要事項のみを述べる。詳細は第 3 章, 第 4 章で議論する。第 1 章で述べたように、PG ではパワースイッチを用いてパワードメインの電源供給と遮断を制御する。NV-/VNR-SRAM でも同様にいくつかのアーキテクチャによってパワーマネジメントを行うことができるが、本論文では主にヘッダパワースイッチを用いる。この構成では負バイアスを用いることなく、スーパーカットオフを導入できるといった利点がある (1.3 節参照)。以下では NV-/VNR-SRAM セルにおける通常の SRAM 動作 (NL モード) と PG 動作 (PG モード) を実現するためのパワースイッチのアーキテクチャを示す。図 2.1.1(a) に NV-SRAM のセル領域 (セルアレイ) におけるパワースイッチの構成を示す。図の PS_1 と PS_2 を使い分けることによって、通常電圧の SRAM 動作 (mode=READ, WRITE, SB), スリープ (mode=SLP), 電源遮断 (mode=SD) の 3 つの動作モードに必要な Virtual V_{DD} (VV_{DD}) を生成する。本論文では第 3 章で述べるようにそれぞれ、1.2V, 0.8V, 0.02V を用いる。これらは単一のグローバル配線 (Global V_{DD} line) の電圧 $V_{DD}=1.2V$ から、 PS_1 と PS_2 を用いて生成される。

VNR-SRAM では以上の 3 モードに ULV リテンションを加えた 4 種類の電圧が必要となる。図 2.1.1(b) に VNR-SRAM のセルアレイにおけるパワースイッチの構成を示す。VNR-SRAM では NV-SRAM と同様の PS_1 および PS_2 に加えて、ULV リテンションを実現するために PS_3 を用いる。本論文では第 4 章で述べるように通常電圧の SRAM 動作, スリープ, ULV リテンション, 電源遮断に必要な VV_{DD} にそれぞれ、1.2V, 0.8V, 0.2V, 0.02V を用いる。これらは、 $V_{DDH}=V_{DD}=1.2V$ および $V_{DDL}=0.2V$ の 2 つの電源線 (Global

V_{DDH} line および Global V_{DDL} line)と3つのパワースイッチ PS_1 , PS_2 , および PS_3 から生成される.

一般に SRAM セルにパワースイッチを設けるときは, 複数のセルでパワースイッチを共有してセルアレイの外部に置き, パワースイッチによる占有面積を可能な限り小さく抑える. 本論文でもパワースイッチはこのように構成した.

2.1.3 アレイアーキテクチャ

図 2.1.2(a)に本研究で用いる NV-/VNR-SRAM の入出力に関する信号系統を示す. 電源線(power supply; PWSPL)の他に, 通常の SRAM 動作を行うための制御信号として, クロック信号(CLK)とクロックゲーティング用のクロックイネーブル信号(CLK enable; CLKEN), ライトイネーブル信号(Write enable; WE), アドレス(Address; ADDR)を用いる(電源線は異なる電圧値の2系統を用いることがある). データの入出力はバスを介して行い, この制御にバスイネーブル信号(Bus enable; BE)を用いる. リードイネーブル信号(Read enable; RE)は後述のようにサブアレイ内で WE から生成する. また, これらの通常の SRAM 動作の信号に加えて, パワースイッチ制御信号(power switch control; PSCTRL), セルアレイの電源遮断/復帰移行動作に要する制御信号(Enter/Exit control; EECTRL)を用いる. PSCTRL はパワースイッチのゲート信号から構成され, 本論文では V_{PGn} で表現する(n はパワースイッチの種類を表し, 図 2.1.1 の添え字に対応している). また, そのビット幅は用いているパワースイッチの種類と独立にパワーマネジメントするワードメイン数に依存する. EECTRL は制御方式によってビット幅が決まるが, その種類は複数必要なこともある. また, EECTRL が不要なこともある. NV-SRAM では, ストアとリストアのイネーブル信号(これらから SR 信号および CTRL 信号が生成される)が, VNR-SRAM では CTRL 信号が EECTRL 信号になる. また, 第 4 章で述べる pFBPD 型 VNR-SRAM では EECTRL による制御は不要となる(ただし, この場合一定バイアスは必要). これらの詳細は各章で行う.

次に, NV-/VNR-SRAM のアレイ構成について示す. 通常, SRAM ではセルアレイの領域を比較的小規模のサブアレイに分割して構成する. 図 2.1.2(b)に本論文で用いるサブアレイの構成を示す. サブアレイはセルの集積領域であるセルアレイ(cell array)と, セルの制御を行う周辺回路(peripheral)で構成される(簡単のため本図では周辺回路はセルアレイを取り囲むように配置した. 実際にはサブアレイのセルアレイをさらに分割し

て周辺回路を適宜配置することもあるが、性能指標を定式化するためのモデリングにはこのような簡単な構成で十分である)。サブアレイは独立に電源の供給制御が可能な 8 つのブロックに分割してある。これらを以後ブロックと呼ぶ。周辺回路は通常の SRAM 動作を制御するための回路と不揮発/擬似不揮発記憶を用いて PG を実現するための回路から構成され、前者を通常動作周辺回路 (Peripherals for normal operation mode; NLP), 後者を PG 動作周辺回路 (Peripherals for PG mode using nonvolatile (or virtually nonvolatile) retention; PGP) と呼ぶ。本論文で用いる NV-/VNR-SRAM サブアレイのセルアレイ, NLP, PGP の構成を表 2.1.1 に示す。本論文ではサブアレイの容量は 8kB (256bit×256bit) とし、ブロックの容量は 1kB (64bit×128bit) とした。通常の SRAM の読み出し/書き込み動作のビット幅は 64bit とした。通常動作の周辺回路は後述のようにプリチャージ回路, セクタ, ライトドライバ, センスアンプ, ワードデコーダ, カラムデコーダ, 入出力フリップフロップ (FF) で構成した。PGP の詳細は各章で論ずる。

次に NV-/VNR-SRAM における通常の SRAM 動作について述べる。図 2.1.3 にセルアレイにおけるブロック内の 1 カラム (128bit) と通常動作に用いる周辺回路を抽出した回路構成を示す (後で述べるが、この回路は本論文における各種回路性能の評価回路として用いる)。各セルにはビット線 (BL と /BL) とワード線 (WL) が接続されている (ここでは PGP からの配線は省略してある)。プリチャージ回路 (Precharge) は読み出しおよび書き込み動作の直前に BL と /BL を充電する回路である。ここではビット線を両端から充電できるようにセルアレイの両端に配置した。セクタ (Selector) は BL と /BL をセンスアンプ (Sense amplifier) またはライトドライバ (Write driver) と接続あるいは切り離し (フローティング) を行う回路である。ライトドライバは WE 信号が入力された時、プリチャージされた BL と /BL の一方を放電して、BL と /BL に相補信号を形成する回路である。センスアンプは RE (Read enable) 信号が入力された時、BL と /BL の電位差を増幅して検出する回路で、セルのデータを読み出す時に用いる。RE 信号は図 2.1.3 に示す WE 信号を保持する FF の出力を反転して生成した。ワードデコーダ (Word DEC.) はアドレスを保持し、そのアドレスから読み出しまたは書き込みを行う行のワード線 (WL) を選択する回路である。カラムデコーダ (Column DEC.) は読み出しまたは書き込みを行う列のセクタを選択する回路である。

通常の SRAM 動作における読み出しおよび書き込み動作では、これらの周辺回路は以下のように動作する。

読み出し: RE (/WE)のアサート期間では, クロック信号に同期して行アドレス信号と列アドレス信号がワードデコーダ, カラムデコーダに供給される. 列アドレス信号はセルクタを介して対応する BL および/BL をセンスアンプに接続する. 行アドレス信号はワードデコーダにより WL を 1 本選択しセルにアクセスする. RE 信号を供給するとセンスアンプにより読み出しが可能になる.

書き込み: WL の選択までは読み出しと同様に行い, WE 信号を供給するとデータが BL および/BL を介してセルに供給され, セルは BL および/BL で指定される安定状態に遷移する.

表 2.1.2 に NV-/VNR-SRAM のサブアレイが通常の SRAM 動作モード (NL モード), 電源遮断移行動作モード (ENT モード), 電源復帰移行動作モード (EXT モード), 電源遮断/ULV リテンションモード (SD/ULVR モード) の場合におけるセルアレイと周辺回路の回路状態をまとめた. この表に基づいて次節で性能評価指標の定式化を行う.

NL モード時は, NV-/VNR-SRAM はいずれもセルアレイと NLP に電源供給し, 通常の SRAM 動作を行う. PGP については NV-SRAM では電源遮断するが, VNR-SRAM では動作モードの確定のため電源供給する(ただし, VNR-SRAM の PGP は不要となる場合がある). ENT/EXT モード時は, NV-/VNR-SRAM はいずれも電源遮断/電源復帰移行動作を, サブアレイを一つずつ順次選択しながら実行する. 本論文ではこの際の電流許容値に配線層の Electromigration (EM) を引き起こす電流値を用い, その電流値以下に収めるように設計した. また, この電流値には NV-SRAM ではワーストケースである下層配線における値と, これよりも緩和された上層配線における許容値の場合で評価し, VNR-SRAM では上層配線のみで評価した. なお, 上層配線の許容値を用いた場合にも, 電源線の引き回しの工夫により下層配線の許容値を容易に満たすことができる. 表 2.1.2 には ENT/EXT モード時に, 選択されたサブアレイ (Selected), 動作を待機しているサブアレイ (Waiting), 動作を完了したサブアレイ (Completed) の回路状態を示してある. すべてのサブアレイにおいて NLP は使用しないためあらかじめ電源遮断しておく. ただし, VNR-SRAM は SNM の改善のため ULV リテンション時にビット線を V_{DDL} に充電する(第 4 章参照). この時, プリチャージ回路のみ, その他の回路と異なり, V_{DDL} と接続するためのパワースイッチを使用し, 電源供給する. PGP については, NV-SRAM では NLP と同様にあらかじめ電源遮断した後, 選択したサブアレイのみ電源供給する(この電源供給により生じる周辺回路のラッシュカレントはセルアレイよりも非常に小さいと

め省略した). VNR-SRAM では, 動作モードを保持するために必要であるため電源供給し続ける. SD/ULVR モード時は, NV-SRAM はセルアレイ, NLP, PGP いずれも遮断し, VNR-SRAM は PGP のみ電源供給する.

NLP, PGP はいずれも, 各回路にドライバが必要となる場合には負荷の大きさに応じてサイズを調整した. 比較用の 6T-SRAM のサブアレイは, NV-/VNR-SRAM と同様に表 2.1.1 におけるセルアレイと NLP で構成した. また, スリープモードではビット線をフローティングにし, 十分時間が経過した時点での電力を評価した.

2.2 不揮発/擬似不揮発記憶を用いた PG の性能評価指標

本論文では不揮発/擬似不揮発記憶を用いた PG の性能評価指標として BET と, PG 時(NV-SRAM の場合は電源遮断時, VNR-SRAM の場合は ULV リテンション時)におけるリーク電力($P_{L,SD/ULVR}$)またはこの削減率(γ_{PG})を用いる. 後述のように, BET は PG の粒度の指標であり, $P_{L,SD/ULVR}$ は PG 実行時間が十分に長い時の平均電力である.

各モードにおける物理量はその物理量を表す記号に, 下付きの添え字または括弧で該当するモードを書くことにする. また, 物理量に回路を指定する必要がある場合は, その回路名(セルアレイ(Array), ブロック(Block), NLP, PGP など)を上付きの添え字で示す. NV-SRAM, VNR-SRAM, 6T-SRAM の区別が必要な場合はそれぞれ上付きの添え字 NV, VNR, 6T で明示する. また, 以下では ENT/EXT モードにおける電源遮断移行動作/復帰移行動作をそれぞれ ENT/EXT 動作と呼ぶことにする. ENT/EXT 動作時におけるサブアレイ 1 つあたりの物理量を示す場合, 添え字の動作モードの末尾に 1 を書き足す(例えば 1 つのサブアレイの電源遮断動作に要するレイテンシは τ_{ENT1} , τ_{EXT1} とする).

2.2.1 BET

以下, スタンバイ, スリープ, 電源遮断, ULV リテンションの各モードでリーク電流によって消費するスタティック電力をリーク電力(leakage power)と称することにする. 待機時(スタンバイ)電力はモードに依らず(スタンバイモード以外でも)リーク電流によるスタティック電力を表すことがあるが, スタンバイモード以外に用いると紛らわしいため, 本論文ではリーク電力を用いることにする. ただし, 慣習にしたがって, 混乱の恐れのない場合リーク電流によるスタティック電力の総称として待機時電力を用いることがある. また, リ

ーク電力を時間で積分した物理量をリークエネルギー (leakage energy) と呼ぶことがある。

図 2.2.1(a)に 6T-SRAM と NV-/VNR-SRAM について, 通常の SRAM 動作時のリーク電力, ENT 動作における消費電力, 電源遮断/ULV リテンション時のリーク電力, EXT 動作における消費電力の推移を示す. 青線が 6T-SRAM ($p^{6T}(t)$), 赤線が NV-/VNR-SRAM ($p^{NV/VNR}(t)$) の各動作モードにおける電力を示している。

NV-/VNR-SRAM ではデータを失うことなく電源遮断 (SD モード)/ULV リテンション (ULVR モード) を行うことができる. PG を行う場合には通常の SRAM 動作 (NL モード) の終了後に ENT 動作を行い, SD/ULVR モードに突入する (NL モードにおける READ および WRITE 動作の動作時 (ダイナミック) 電力は省略している). その後, EXT 動作を行ってから NL モードに復帰する. 通常の 6T-SRAM は電源遮断/ULV リテンションできないため, この期間はスリープ (SLP) モードとする。

NV-/VNR-SRAM では, これらのセル構成に起因してスタンバイ時のリーク電力が 6T-SRAM に比べて増大し, 図の黄色の領域で示される余剰な消費エネルギー ($\Delta E_{L,NL}$) が発生する (ただし, セル構成によっては, リーク電力が 6T-SRAM を下回る場合もある). これに加え, 図のピンク色の部分で示すように, PG モードでは ENT 動作に要するエネルギー (ΔE_{ENT}) と EXT 動作に要するエネルギー (ΔE_{EXT}) だけ余剰なエネルギーを消費する. しかし, これらの余剰なエネルギーが発生する一方で, 電源遮断/ULV リテンションによって, この期間のリーク電力を大幅に削減することができる (SD/ULVR モード時のリーク電力の削減分を ΔP_{save} とする). NV-/VNR-SRAM では, 電源遮断/ULV リテンションの時間とともに, 図の水色の領域で示す削減エネルギーが増加していく. したがって, 上述の余剰なエネルギー (ΔE_{ENT} と ΔE_{EXT}) は適切に設定された時間以上に電源遮断/ULV リテンションを行うことで補償され, 全体のリークエネルギーは削減されることになる。

NV-/VNR-SRAM における通常の SRAM 動作, ENT/EXT 動作による余剰エネルギーを補填する最小の電源遮断時間を Break-even time (BET) と定義する. 本論文では BET を以後 T_{BE} と表記する. 電源遮断/ULV リテンションの実行時間が T_{BE} よりも短ければ消費エネルギーは 6T-SRAM と比べて増大し, T_{BE} より長ければ削減される. T_{BE} の評価に基づき NV-/VNR-SRAM の PG を実行することが不揮発/擬似不揮発記憶の効果的活用にも必須となる. T_{BE} は上記の定義から以下の式(2.1)を満たす。

$$\Delta E_{L,NL} + \Delta E_{ENT} + \Delta E_{EXT} = \Delta P_{save} \cdot T_{BE} \quad (2.1)$$

式(2.1)から T_{BE} は次式で表せる.

$$T_{BE} = (\Delta E_{L,NL} + \Delta E_{ENT} + \Delta E_{EXT}) / \Delta P_{save} \quad (2.2)$$

T_{BE} の解析で特に重要となるのは NL モードの実行時間 (τ_{NL}) である. 式(2.2)を NL モード時の平均の余剰電力 ($\Delta P_{L,NL}$) を用いて τ_{NL} について整理すると, 次式のように表せる.

$$T_{BE} = \frac{\Delta E_{ENT} + \Delta E_{EXT}}{\Delta P_{save}} + \frac{\Delta P_{L,NL}}{\Delta P_{save}} \cdot \tau_{NL} \quad (2.3)$$

T_{BE} の挙動は τ_{NL} に比例しない成分と比例する成分に分けられ, どちらの成分が変化するかを解析することで詳細なエネルギー解析が可能になる. 以後 τ_{NL} に比例しない成分を T_{BE}^{EE} , τ_{NL} に比例する成分の比例係数を η と呼ぶ. これらはそれぞれ次式のように定義する.

$$T_{BE}^{EE} = \frac{\Delta E_{ENT} + \Delta E_{EXT}}{\Delta P_{save}} \quad (2.4)$$

$$\eta = \frac{\Delta P_{L,NL}}{\Delta P_{save}} \cdot \tau_{NL} \quad (2.5)$$

図 2.2.1(b) は T_{BE} の τ_{NL} 依存性を模式的に示している. T_{BE} を削減するには T_{BE}^{EE} と η をいずれも削減する(すなわち, ΔE_{ENT} , ΔE_{EXT} , $\Delta P_{L,NL}$ を削減し, ΔP_{save} を増加する)ことが重要になる. 式(2.4)と式(2.5)の分母および分子の各項 ($\Delta P_{L,NL}$, ΔE_{ENT} , ΔE_{EXT} , ΔP_{save}) は具体的な回路構成と回路状態を与えることで以下に示すように求めることができる.

(i) NL モード時の余剰電力 ($\Delta P_{L,NL}$)

通常の SRAM 動作時には, すべてのセルアレイは NL モードとする(表 2.1.2 参照). NLP については NV-/VNR-SRAM においていずれも ON 状態とし, PGP については, NV-SRAM では OFF 状態とするが, VNR-SRAM では, 第 4 章で述べる BI モードで動作させるため ON 状態とする. この時, 電力は NL モード時の平均の余剰リーク電力 ($\Delta P_{L,NL}$) と NL モードの実行時間 (τ_{NL}) を用いて次式のように表せる.

$$\Delta P_{L,NL} = \tau_{NL}^{-1} \int_{\tau_{NL}} (p^{NV/VNR}(t) - p^{6T}(t)) dt \quad (2.6)$$

また, 表 2.1.2 に示した回路状態から, $\Delta P_{L,NL}$ はサブアレイ内のセルアレイと周辺回路のリーク電力を用いて次式で表せる.

$$\Delta P_{L,NL} = (\Delta P_{L,NL}^{Array} + \Delta P_{L,OFF/ON}^{PGP}) \quad (2.7)$$

セルアレイの余剰電力 ($\Delta P_{L,NL}^{Array}$) は NV-/VNR-SRAM から 6T-SRAM の電力を引いた

値として定義している. PGP は NV-SRAM/VNR-SRAM の動作モードをそれぞれ OFF/ON と記した(表 2.1.2 参照). NV-/VNR-SRAM と 6T-SRAM で同一の構成としており, NLP により生じる消費電力は相殺されるため, NLP の電力には依存しない. 式 (2.7)から $\Delta P_{L,NL}$ を削減するには NV-/VNR-SRAM においてセルアレイと PGP の電力の削減が重要となる.

(ii) ENT/EXT モード時の余剰エネルギー (ΔE_{ENT} , ΔE_{EXT})

2.1.3 節で述べたように ENT/EXT モードでは, NV-/VNR-SRAM は各サブアレイを 1 つずつ順番に選択して, ENT/EXT 動作を実行し, 動作の完了したサブアレイから順次電源遮断/ULV リテンションに移行する. NV-/VNR-SRAM において, ENT モード時には順番が来るまで動作を待機しているサブアレイは SLP/SB モードとし, 動作を完了したサブアレイは SD/ULVR モードとする. また, EXT モード時には, 待機しているサブアレイを SD/ULVR モード, 動作を完了したサブアレイを SLP/SB モードとする. VNR-SRAM については ENT/EXT 動作に要するレイテンシが NV-SRAM に比べて非常に短いため, 待機しているサブアレイを SB モードにしたとしても, 第 4 章で示すように劇的に T_{BE} を削減できる. ENT/EXT モードにおける NV-/VNR-SRAM と 6T-SRAM の消費エネルギーの差分 (ΔE_{ENT} , ΔE_{EXT}) は, 平均消費電力の差分 (ΔP_{ENT} , ΔP_{EXT}) を用いて次式で表せる.

$$\Delta E_{ENT} = \int_{\tau_{ENT}} (p^{NV/VNR}(t)) dt = \Delta P_{ENT} \cdot \tau_{ENT} \quad (2.8)$$

$$\Delta E_{EXT} = \int_{\tau_{EXT}} (p^{NV/VNR}(t)) dt = \Delta P_{EXT} \cdot \tau_{EXT} \quad (2.9)$$

ここで, すべてのサブアレイの ENT/EXT 動作の完了までに要するレイテンシを τ_{ENT} , τ_{EXT} としている. ENT/EXT モードは 6T-SRAM では不要となるため, ΔE_{ENT} , ΔE_{EXT} は NV-/VNR-SRAM の消費エネルギーに等しい. T_{BE} を削減するためには, 後述のように, この ΔE_{ENT} , ΔE_{EXT} を小さく抑えることが非常に重要となる. ENT/EXT モード時はセルアレイ, 周辺回路でそれぞれ複数の回路状態が混在するため, 2.2.2 節でさらに詳細に定式化を行う.

(iii) SD/ULVR モード時の削減電力 (ΔP_{save})

NV-/VNR-SRAM の SD/ULVR モード時はすべてのサブアレイを電源遮断/ULV リテ

ンションとする. このとき, 平均削減電力 ΔP_{save} は以下のように表せる.

$$\Delta P_{\text{save}} = \Delta P_{\text{L,SD/ULVR}} = \tau_{\text{SD/ULVR}}^{-1} \int_{\tau_{\text{SD/ULVR}}} (p^{6T}(t) - p^{\text{NV/VNR}}(t)) dt \quad (2.10)$$

また, 式(2.10)の ΔP_{save} は, 表 2.1.2 から

$$\Delta P_{\text{save}} = \left| \Delta P_{\text{L,SD/ULVR}}^{\text{Array}} + \Delta P_{\text{L,OFF/ON}}^{\text{PGP}} \right| \quad (2.11)$$

と表せる. この場合も(i)と同様にNLPの電力には依存しない. また, ΔP_{save} を増加するにはNV-/VNR-SRAMにおいてセルアレイとPGPの電力の削減が重要となる.

2.2.2 BETの削減方法

続いて $T_{\text{BE}}^{\text{EE}}$ と η の削減方法について以下に示す. 式(2.4)および式(2.5)に示したように, $T_{\text{BE}}^{\text{EE}}$ は $\Delta E_{\text{ENT}} + \Delta E_{\text{EXT}}$ と ΔP_{save} に, η は $\Delta P_{\text{L,NL}}$ と ΔP_{save} に依存するため, これらの削減が重要となる.

(i) $T_{\text{BE}}^{\text{EE}}$ の削減方針

はじめに, ENTモード時の消費エネルギー(ΔE_{ENT})が $T_{\text{BE}}^{\text{EE}}$ に与える影響について考察する. EXTモード時の消費エネルギー(ΔE_{EXT})も同様に解析できるため, 以下ではENTモードのみ説明する. 図 2.2.2 は ENTモードにおけるNV-/VNR-SRAMのメモリアレイを示す. 同図の正方形はサブアレイを表す. メモリアレイ内のサブアレイの個数を N とし, 各サブアレイにはENT動作を実行する順番($n=1$ から $n=N$)と動作モードを示している. 同図は k 番目のサブアレイでENT動作を実行している状態を表す. 以下では, ENT動作を実行しているサブアレイをENT動作実行サブアレイ, ENT動作を待機しているサブアレイをENT動作待機サブアレイ, ENT動作を完了したサブアレイをSD/ULVRサブアレイとする.

まず, ENTモードにおけるセルアレイの消費エネルギー($\Delta E_{\text{ENT}}^{\text{Array}}$)を導出する. まず, サブアレイ内のセルアレイの消費エネルギーは上述のシーケンスを考慮すると, 以下の3つの成分からなる(図 2.2.2 参照).

- ENT動作実行サブアレイの消費エネルギー($\Delta E_{\text{ENT1}}^{\text{Array}}$)
- ENT動作待機サブアレイのリークエネルギー($\Delta E_{\text{L,SLP/SB1}}^{\text{Array}}$)
- SD/ULVRサブアレイのリークエネルギー($\Delta E_{\text{L,SD/ULVR1}}^{\text{Array}}$)

ENT動作を1番目から N 番目までのすべてのサブアレイについて順次行った場合, セ

ルアレイの総消費エネルギー $\Delta E_{\text{ENT}}^{\text{Array}}$ は以下のように示される.

$$\Delta E_{\text{ENT}}^{\text{Array}} = \sum_{k=1}^N \left(\Delta E_{\text{ENT}1}^{\text{Array}} + \Delta E_{\text{L,SLP/SB}1}^{\text{Array}} \cdot (N - k) + \Delta E_{\text{L,SD/ULVR}1}^{\text{Array}} \cdot (k - 1) \right) \quad (2.12)$$

式(2.12)を計算すると以下のように表せる. ここで $\sum_{k=1}^N k = N(N - 1)/2$ を用いた.

$$\Delta E_{\text{ENT}}^{\text{Array}} = \Delta E_{\text{ENT}1}^{\text{Array}} \cdot N + (\Delta E_{\text{L,SLP/SB}1}^{\text{Array}} + \Delta E_{\text{L,SD/ULVR}1}^{\text{Array}}) \cdot N(N - 1)/2 \quad (2.13)$$

式(2.13)の右辺第 1 項が ENT 動作実行サブアレイのエネルギーの総和を, 第 2 項が ENT 動作待機サブアレイと SD/ULVR サブアレイのリークエネルギーの総和を示す.

$\Delta E_{\text{L,SLP/SB}1}^{\text{Array}}$ は $\Delta E_{\text{L,SD/ULVR}1}^{\text{Array}}$ に対して 1桁程度以上大きいので, 式(2.13)の第 2 項では実質的に $\Delta E_{\text{L,SLP/SB}1}^{\text{Array}}$ が支配的に影響する. すなわち, $\Delta E_{\text{ENT}}^{\text{Array}}$ を削減するには $\Delta E_{\text{ENT}1}^{\text{Array}}$ と $\Delta E_{\text{L,SLP/SB}1}^{\text{Array}}$ の削減が重要となる.

次に, ENT モードにおける周辺回路のエネルギーについて説明する. 周辺回路についても表 2.1.2 より総消費エネルギーは次式で表せる.

$$\Delta E_{\text{L,OFF}}^{\text{NLP}} = \sum_{k=1}^N (\Delta E_{\text{L,OFF}1}^{\text{NLP}} \cdot N) = \Delta E_{\text{L,OFF}1}^{\text{NLP}} \cdot N^2 \quad (2.14)$$

$$\begin{aligned} \Delta E_{\text{ENT}}^{\text{PGP}} &= \sum_{k=1}^N (\Delta E_{\text{ENT}1}^{\text{PGP}} + \Delta E_{\text{L,OFF/ON}1}^{\text{PGP}} \cdot (N - 1)) \\ &= \Delta E_{\text{ENT}1}^{\text{PGP}} \cdot N + \Delta E_{\text{L,OFF/ON}1}^{\text{PGP}} \cdot N(N - 1) \end{aligned} \quad (2.15)$$

PGP は, リークエネルギーの他に ENT 動作に要するエネルギー($\Delta E_{\text{ENT}1}^{\text{PGP}}$)が生じる. これについては NV-SRAM は不揮発性メモリ素子の書き込みエネルギーに比べて非常に小さいため省略し, VNR-SRAM では第 4 章で述べる nFBPD 型のみ CTRL ドライバの切り替えエネルギーを考慮した. 式(2.13)-式(2.15)をすべて足し合わせることで, ENT モードにおける消費エネルギー ΔE_{ENT} は

$$\begin{aligned} \Delta E_{\text{ENT}} &= \Delta E_{\text{ENT}1}^{\text{Array}} \cdot N + (\Delta E_{\text{L,SLP/SB}1}^{\text{Array}} + \Delta E_{\text{L,SD/ULVR}1}^{\text{Array}}) \cdot N(N - 1)/2 \\ &\quad + \Delta E_{\text{OFF}1}^{\text{NLP}} \cdot N^2 + \Delta E_{\text{ENT}1}^{\text{PGP}} \cdot N + \Delta E_{\text{L,OFF/ON}1}^{\text{PGP}} \cdot N(N - 1) \end{aligned} \quad (2.16)$$

と表せる. 同様に EXT モードにおける消費エネルギー ΔE_{EXT} は

$$\begin{aligned} \Delta E_{\text{EXT}} &= \Delta E_{\text{EXT}1}^{\text{Array}} \cdot N + (\Delta E_{\text{L,SLP/SB}1}^{\text{Array}} + \Delta E_{\text{L,SD/ULVR}1}^{\text{Array}}) \cdot N(N - 1)/2 \\ &\quad + \Delta E_{\text{OFF}1}^{\text{NLP}} \cdot N^2 + \Delta E_{\text{EXT}1}^{\text{PGP}} \cdot N + \Delta E_{\text{L,OFF/ON}1}^{\text{PGP}} \cdot N(N - 1) \end{aligned} \quad (2.17)$$

となる. ただし, NV-SRAM では不揮発性メモリ素子の書き込み/復帰動作によりレイテンシ (τ_{ENT}) が非常に長くなるため, サブアレイごとではなく, ブロック

(64 bits×128 lines) ごとに、ストア/リストア動作完了後に電源の遮断/供給を行うものとした (ストアレイテンシが長いため、待機セルのリークエネルギーを削減するための方策である)。一方、VNR-SRAM は NV-SRAM に比べてレイテンシが短いため、サブアレイごとに電源制御する構成とした。

式(2.16)、式(2.17)を式(2.8)、式(2.9)を用いて電力とレイテンシの積に変換し、以上のように導出した T_{BE}^{EE} の各項を式(2.4)に代入すると以下のように表せる。

$$T_{BE}^{EE} = \frac{\left\{ (\Delta P_{ENT1}^{Array} + \Delta P_{ENT1}^{PGP}) \cdot \tau_{ENT1} + (\Delta P_{EXT1}^{Array} + \Delta P_{EXT1}^{PGP}) \cdot \tau_{EXT1} \right\} \cdot N + \left(\Delta P_{L,SLP/SB1}^{Array} + \Delta P_{L,SD/ULVR1}^{Array} \right) \cdot (\tau_{ENT1} + \tau_{EXT1}) \cdot \frac{N(N-1)}{2} + \left(\Delta P_{L,OFF1}^{NLP} \cdot N^2 + \Delta P_{L,OFF/ON1}^{PGP} \cdot N(N-1) \right) \cdot (\tau_{ENT1} + \tau_{EXT1})}{\left| \Delta P_{L,SD/ULVR1}^{Array} + \Delta P_{L,OFF/ON1}^{PGP} \right| \cdot N} \quad (2.22)$$

式(2.22)の分子は次式のように2つの成分に分けられる。

$$\Delta E_{EE} = \left\{ (\Delta P_{ENT1}^{Array} + \Delta P_{ENT1}^{PGP}) \cdot \tau_{ENT1} + (\Delta P_{EXT1}^{Array} + \Delta P_{EXT1}^{PGP}) \cdot \tau_{EXT1} \right\} \cdot N \quad (2.23)$$

$$\Delta E_{leakage} = \left(\Delta P_{L,SLP/SB1}^{Array} + \Delta P_{L,SD/ULVR1}^{Array} \right) \cdot (\tau_{ENT1} + \tau_{EXT1}) \cdot \frac{N(N-1)}{2} + \left(\Delta P_{L,OFF1}^{NLP} \cdot N^2 + \Delta P_{L,OFF/ON1}^{PGP} \cdot N(N-1) \right) \cdot (\tau_{ENT1} + \tau_{EXT1}) \quad (2.24)$$

ΔE_{EE} は ENT/EXT 動作実行サブアレイの消費エネルギーの総和を、 $\Delta E_{leakage}$ は ENT/EXT 動作待機サブアレイと SD/ULVR サブアレイの総リークエネルギーである。式(2.22)において N を増加させた場合、 T_{BE}^{EE} は $\Delta E_{leakage}$ が N の2乗に比例するため、アレイサイズに比例して増大する。以上の議論を踏まえ、NV-SRAM を例にとり、図 2.2.3 に T_{BE}^{EE} のアレイサイズ依存性を示す (VNR-SRAM でも同様の議論が可能である)。図 2.2.3 の黒線は $\Delta E_{leakage}$ を省略し、 ΔE_{EE} のみ考慮した場合の T_{BE}^{EE} を示す。この場合、上述の議論から T_{BE}^{EE} はアレイサイズに関係なく一定となる (すなわち、余剰エネルギーの増加分は比較用の 6T-SRAM のリークエネルギーの増加分で相殺される)。一方、図 2.2.3 の青線と赤線はこれに加えて $\Delta E_{leakage}$ も考慮した場合の T_{BE}^{EE} を示す。青線は ENT 動作待機サブアレイを SB モードとしたときの T_{BE}^{EE} を、赤線は SLP モードとしたときの T_{BE}^{EE} を示す。このエネルギーは 32kB (高階層キャッシュ、特に L1 キャッシュ) 程度の小容量のアレイではほとんど影響を及ぼさないが 2MB (ラストレベルキャッシュ) 程度の大容量のアレイで特に T_{BE}^{EE} に支配的な影響を及ぼし、スリープモードの導入により一定程度の削減効果はあるものの完全には削減できない。以上から T_{BE}^{EE} の削減には特に小容量アレイでは ENT/EXT 動作実行サブアレイのエネルギー (ΔE_{EE}) の削減が、大容

量アレイではこれに加えて、ENT/EXT 動作の待機サブアレイのリークエネルギー ($\Delta E_{\text{leakage}}$) の削減も重要となる。

(ii) ENT 動作の省略による $T_{\text{BE}}^{\text{EE}}$ の削減

上述の $T_{\text{BE}}^{\text{EE}}$ を削減するには、NV-/VNR-SRAM のメモリアレイの ENT モードにおいて一部のサブアレイの ENT 動作を省略して電源遮断するのが有効である (VNR-SRAM についても、2.2 節で示したようにパワースイッチをすべて遮断することで完全に電源遮断することが可能である)。図 2.2.4 に ENT 動作を省略した場合の NV-/VNR-SRAM のメモリアレイを示す。 $J+1$ 番目から N 番目までのサブアレイについては ENT 動作を省略して電源遮断し、1 番目から j 番目までのサブアレイについて上述のように順次 ENT 動作を実行するものとする。簡単のために ENT 動作を省略するサブアレイをまとめて示しているが、これらはメモリアレイ内に散在していてもよい。また、ここではサブアレイ単位での省略を示しているが、ブロック単位の粒度で省略しても下記の定式化は同様にできる。この場合の $T_{\text{BE}}^{\text{EE}}$ は全サブアレイに対する、ENT 動作を省略するサブアレイの割合を R_{skip} とすると、式(2.22)の ENT 動作に関する項の N を $N' (=N(1-R_{\text{skip}}))$ に変換し、電源遮断したサブアレイのリークエネルギーを足し合わせることで次式のように表せる。

$$T_{\text{BE}}^{\text{EE}} = \frac{\begin{aligned} & (\Delta P_{\text{ENT1}}^{\text{Array}} + \Delta P_{\text{ENT1}}^{\text{PGP}}) \cdot \tau_{\text{ENT1}} \cdot N' + (\Delta P_{\text{EXT1}}^{\text{Array}} + \Delta P_{\text{EXT1}}^{\text{PGP}}) \cdot \tau_{\text{EXT1}} \cdot N \\ & + (\Delta P_{\text{L,SLP/SB1}}^{\text{Array}} + \Delta P_{\text{L,SD/ULVR1}}^{\text{Array}}) \cdot \left(\tau_{\text{ENT1}} \cdot \frac{N'(N'-1)}{2} + \tau_{\text{EXT1}} \cdot \frac{N(N-1)}{2} \right) \\ & + \Delta P_{\text{L,SD1}}^{\text{Array}} \cdot \tau_{\text{ENT1}} \cdot \frac{N \cdot R_{\text{skip}}(N \cdot R_{\text{skip}} - 1)}{2} \\ & + (\Delta P_{\text{L,OFF1}}^{\text{NLP}} \cdot N^2 + \Delta P_{\text{L,OFF/ON1}}^{\text{PGP}} \cdot N(N-1)) \cdot (\tau_{\text{ENT1}} + \tau_{\text{EXT1}}) \end{aligned}}{\left| \Delta P_{\text{L,SD/ULVR1}}^{\text{Array}} + \Delta P_{\text{L,OFF/ON1}}^{\text{PGP}} \right| \cdot N} \quad (2.25)$$

式(2.25)から ENT 動作の消費エネルギーおよび ENT 動作を実行していないサブアレイのリークエネルギーはいずれも R_{skip} の増加にともない削減される。また、ENT 動作を省略するサブアレイは ENT 動作にかかるレイテンシも削減される。一方、ENT 動作を実行せずに電源遮断/ULV リテンションに移行したサブアレイのリークエネルギーが生じるが、この寄与は比較的小さいため全体として $T_{\text{BE}}^{\text{EE}}$ は効果的に削減される。なお、この省略が有効な状況には、もともと ENT 動作を実行する必要がないサブアレイに ENT 動作を行ってしまっている場合と、ENT 動作を実行する必要があるサブアレイのデータだとしても積極的に破棄する場合がある。これらの詳しいアーキテクチャについては第 3 章で説

明する.

(iii) η の削減

η の削減には上述の通常動作時における NV-/VNR-SRAM セルの余剰なリーク電力と通常動作時の PGP のリーク電力の削減, および電源遮断/ULV リテンション時におけるリーク電力の削減が重要となる. NV-/VNR-SRAM セルではトランジスタ数が増加するため余剰なリーク電力が生じるが, これはNV-SRAM ではバイアス制御を, VNR-SRAM ではモード切替や新型セル(第4章で示す pFBPD 型 VNR セル)を導入することで削減できる. また, PGP の余剰リーク電力は, NV-SRAM では PGP にもスーパーカットオフを導入することで削減でき, VNR-SRAM では pFBPD 型 VNR セルを用いて PGP を省略することで削減できる. 電源遮断/ULV リテンション時のリーク電力については, NV-SRAM ではセル・周辺回路へのスーパーカットオフの導入により, VNR-SRAM では4.5節で述べるデュアルパワースイッチアーキテクチャや, PUDF アーキテクチャを用いることにより削減できる. これらの詳細は各章で示す.

2.2.3 PG のリーク電力削減率

続いて, PGにおけるリーク電力削減率の評価について説明する. 本論文では電源遮断/ULV リテンションのリーク電力($P_{L,SD/ULVR}$)とその削減率(γ_{PG})を用いて評価する. 図2.2.1(a)にはNV-/VNR-SRAMの各動作時のリーク電力も示している. NV-/VNR-SRAMの平均電力($P_{L,ave}$)はENT/EXT動作に関するエネルギーを含む1サイクルにおいて, ΔE_{ENT} , ΔE_{EXT} , $P_{L,NL}^{NV/VNR}$, $P_{L,SD/ULVR}$, τ_{NL} , $\tau_{SD/ULVR}$, τ_{cyc} を用いて次式で表せる.

$$P_{L,ave} = \frac{P_{L,NL}^{NV/VNR} \cdot \tau_{NL} + \Delta E_{ENT} + \Delta E_{EXT} + P_{L,SD/ULVR} \cdot \tau_{SD/ULVR}}{\tau_{cyc}} \quad (2.26)$$

6T-SRAM の場合は通常の SRAM 動作時に通常電圧のスタンバイモード(図2.2.1(a)の $P_{L,NL}^{6T}$), 電源遮断/ULV リテンション時にスリープモード($P_{L,SLP}^{6T}$)を用い, ΔE_{ENT} , ΔE_{EXT} は不要となる. 式(2.26)において $\tau_{SD/ULVR}$ が十分長ければ, $P_{L,ave}$ は $P_{L,SD/ULVR}$ に漸近する. そのため, 本論文で検討する各種モード時における NV-/VNR-SRAM を用いた PG による電力削減率 γ_{PG} は $P_{L,SD/ULVR}$ (6T-SRAM では $P_{L,SLP}^{6T}$) を従来の 6T-SRAM の通常電圧におけるリーク電力である $P_{L,NL}^{6T}$ と比較して評価した(上記の電力にはセルのみで評価する場合とセルアレイと周辺回路の寄与をいずれも含む場合がある). 以上

から NV-/VNR-SRAM の γ_{PG} は以下のように定義される.

$$\gamma_{PG} = \frac{P_{L,SD/ULVR}}{P_{L,NL}^{6T}} \quad (2.27)$$

なお, 図 2.2.1(a)の通常の SRAM 動作(NL モード)時にスリープモードによる短い待機状態を入れることは可能であるが, 以後, 簡単のため通常の SRAM 動作内にこのようなスリープモードは入れていない(以上の定式化には, この効果を容易に組み込むことができる).

2.2.4 本論文における評価方法

(i) ベンチマークシーケンス

図 2.2.5 に本論文で T_{BE} の評価に用いたベンチマークシーケンスを示す. NV-/VNR-SRAM では全セルについての読み出し, 書き込みを 1 回ずつ行い, これを n_{RW} サイクル行う. 1GHz で動作することを前提に通常の SRAM の読み出し動作, 書き込み動作に要する時間はいずれも 1ns とし, これを n_{RW} 回繰り返したときの通常動作時間を τ_{NL} とした. PG 実行時には ENT 動作を行ってから電源遮断/ULV リテンションを開始する. その後, EXT 動作を行い通常の SRAM 動作を再び開始する. 以上を NV-/VNR-SRAM における 1 サイクルとした. 6T-SRAM については NV-/VNR-SRAM と同様に通常動作を実行した後, NV-/VNR-SRAM の PG モード時には, SLP モードとした. その後同様に再び通常動作に戻る.

(ii) 不揮発/擬似不揮発性 SRAM の性能解析モデル

NV-/VNR-SRAM アレイの評価・解析には, 図 2.1.2(b)に示したサブアレイのうち, セル, アレイ, 周辺回路の電力・エネルギー性能をいずれも評価する必要がある. セル及びセルアレイの性能評価には前述の図 2.1.3 を用いた. 通常電圧における SRAM 動作時のリーク電力の算出には, 評価回路の 1GHz 動作における 1 クロックのリーク電力を平均して算出した. また, スリープ, 電源遮断, ULVリテンション時のリーク電力については, 十分時間が経過し, 評価回路内の全ノードのステイトが確定した状態の電力を評価した. 周辺回路については既存の SRAM の設計を参考にロジック, FF, ライトドライバ, センスアンプなどを設計し, これらのリーク電力を足し合わせることで評価した. また, ENT/EXT 動作時のエネルギーオーバーヘッドについては, NV-SRAM は不揮発性メ

メモリ素子 (Magnetic tunnel junction; MTJ) への書き込み/復帰にかかるエネルギーが支配的であり, VNR-SRAM はセルアレイの充放電によるラッシュカレントによるエネルギーが支配的であるため, これら进行评估した. NV-SRAM は, MTJ への書き込み/復帰に要する電流値から 1 セルのエネルギー进行评估し, アレイサイズに応じて定数倍した. VNR-SRAM は, 図 2.1.3 の回路を単位としてラッシュカレントを計算し, こちらもアレイサイズに応じて定数倍して评估した. このとき, セル, V_{DD} , ビット線, ワード線については, 第 4 章で示す 8kB の VNR-SRAM マクロから抽出した寄生抵抗・寄生容量を使用した. なお, グローバルの電源配線 (V_{DD}) には, 一般に電源系のノイズを軽減するためにデカップリングキャパシタを接続するが, パワーゲーティングでは V_{DD} のみを考慮すればよい. デカップリングキャパシタ (およびその充放電の影響) は省略している (典型的なデカップリングキャパシタの容量 (VNR-SRAM アレイ全体のキャパシタンスの 10 倍以上[1]) においてラッシュカレント等への影響は 0.1% 程度以下である). また, 以上の性能解析モデルの検証のため, NV-SRAM では試作した 1kb マクロの TEG 実測値を用いた解析結果との比較を (第 3 章), VNR-SRAM では 8kB マクロの抽出ネットリストを使用した高速 SPICE 大規模シミュレーションによるポストレイアウト解析との比較を行った (第 4 章).

2.3 不揮発/擬似不揮発性 SRAM の設計指標

(i) ばらつきを考慮した設計

最後に本論文における NV-/VNR-SRAM の設計指標について述べる. NV-/VNR-SRAM におけるデータの保持は, CMOS インバータをループ状に接続した双安定回路により実現される. 図 2.3.1 に通常のインバータで構成した双安定回路の回路図と典型的な (ノイズやばらつきのない理想的な状態の) 双安定回路の各 CMOS インバータ入出力特性を同一のグラフ上に示す (このようなグラフをバタフライカーブと呼ぶ). 実線と点線がそれぞれ, 上側と下側のインバータの入出力特性を示しており, 赤点は 2 つの安定点, 青点は論理しきい値 (または準安定点) を示している. 準安定点と 1 つの安定点を含む 2 つのインバータの入出力特性で構成される領域をローブという (すなわち準安定点を挟んで左右に 2 つのローブがある). 動作点が一方のローブ上にあるとき, そのローブのもつ安定点に速やかに移行して安定状態となり, データが保持される. 以後, 2 つの記憶ノード (Q, /Q) のうち, 電位の高い方を H レベル, 電位の低い方を L レベ

ルと呼ぶことにする。このローブに内包される最大の正方形の一辺の長さで雑音余裕 (Static noise margin; SNM) を定義する。

SNM を超えるノイズが入力されたとき、データのフリップなど誤動作を生じることがある。図 2.3.2 に Q、 \bar{Q} に電圧振幅 V_n のノイズ源を挿入した双安定回路と $V_n=0$ と $V_n \neq 0$ の場合のバタフライカーブを示す。黒の実線と点線は図 2.3.1 と同様の、ノイズが生じていない状態のバタフライカーブを示している。このとき、 $(V_Q, V_{\bar{Q}})=(0, V_{DD})$ の安定点でデータが保持されていたとする。ノイズが発生すると上下のインバータのインバータ特性が同図の矢印の方向に平行移動し、バタフライカーブは緑の実線と点線のように変化する。平行移動により $(V_Q, V_{\bar{Q}})=(0, V_{DD})$ の安定点が消失し、 $(V_Q, V_{\bar{Q}})=(V_{DD}, 0)$ の安定点に遷移する。SNM は安定点が消失する瞬間の V_n と等しいので、SNM が大きければノイズが入力されても (安定点の消失による) データの破壊が起こりにくく、SNM が小さければデータの破壊が起こりやすいといえる。このように SNM はノイズ耐性 (すなわち動作安定性) の評価指標として用いることができる。

近年では、トランジスタサイズの縮小にともなってばらつきの影響が深刻化しており、SNM に大きな影響をもたらす。そのため、十分な SNM を確保するにはばらつきを考慮した設計が非常に重要となる。本論文では特に影響の大きいプロセスばらつきに注目して設計を行った。プロセスばらつきは、主にグローバル (チップ間) ばらつきとローカル (チップ内) ばらつきに分けられる。

図 2.3.3 にグローバルばらつきの概念図を示す。図 2.3.3 はウェハと、ウェハ上の異なる位置から切り出した二つのチップを示しており、色の違いはグローバルばらつきにより生じたチップ上のトランジスタのしきい値のずれを示している。グローバルばらつきはウェハ上のチップの位置が異なることによってチップ間のトランジスタのエッチング量などに差が生じた結果発生するばらつきである。各チップ内のすべてのトランジスタはそのしきい値が一様にこのばらつきの影響を受ける。図 2.3.4 にグローバルばらつきを考慮した際の代表的なトランジスタのしきい値の状態を示す。この分布は一般に正規分布で表すことができる。正規分布の中心値は典型的なしきい値を持つトランジスタ (Typical; T) であり、そこから 3σ 程度ずれたしきい値を持つトランジスタはしきい値の低い方を (Fast; F)、高い方を (Slow; S) と呼ぶ。図中の各点はプロセスコーナーと呼ばれ、nMOS、pMOS の順で TT, FF, SS, FS, SF と表記する。

一方、ローカルばらつきはチャンネルにおける不純物ドーピング量のゆらぎなどにより

生じるばらつきである。これにより各チップ内で局所的に目標値からずれたしきい値を持つトランジスタが発生することがある。図 2.3.5 にローカルばらつきの概念図を示す。この場合、グローバルばらつきとは異なり、チップ内のすべてのトランジスタが影響を受ける訳ではなく、一部のトランジスタがランダムに影響を受ける。

一般に SRAM はグローバルばらつきを考慮したすべてのプロセスコーナーで上述の SNM を十分に満たすように設計する。しかし、ローカルばらつきについてはグローバルばらつきよりも大きなばらつき分布を引き起こすため、設計の目標値を下回るセル(不良セル)が生じる可能性がある。この不良セルについては、チップ上のすべてのトランジスタに一樣に発生するものではないため、設計プロセスに含めるのではなく冗長セル(リダンダンシセル)を配置してチップの製造後に不具合のあるセルを置換して補償することが多い。これらを踏まえ、本論文でも同様の方法で NV-/VNR-SRAM を設計する。この設計方法を以下にまとめる。

①グローバルばらつきのプロセスコーナーで目標値を満たすように設計

本論文ではすべてのプロセスコーナーで評価指標の目標値を満たすよう NV-/VNR-SRAM を設計する。SNM を用いて安定性を評価するには、通常 SNM の目標値にある程度の余裕をもたせて設計を行う。しかし、この目標値はシステムやプロセス、使用環境に依存し、一意に定義できるものではない。そのため、本論文では低電圧 SRAM の文献値を参考にし、ワーストケースの SNM の値を参考にした[2,3]。表 2.3.1 にワーストケースの SNM をまとめた。この SNM の文献値は概ね $\sim 30\text{mV}$ ($\sim k_B T/q$)、 $\sim 50\text{mV}$ ($\sim 2k_B T/q$)、 $\sim 80\text{mV}$ ($\sim 3k_B T/q$) の(熱電圧の倍数程度の)3 通りに分類することができる。そこで、本論文ではこれを参考に最も厳しい条件である熱電圧の 3 倍の値である 78mV を用いた。また、NV-SRAM については MTJ の書き込み電流も評価するが、これについては文献値を参考に臨界電流 (J_c) の設計マージンを 20%として設計した(第 3 章参照)。

②ローカルばらつきにより生じた不良セルを冗長セルで補償

次に、グローバルばらつきを考慮して最適設計したセルについてローカルばらつきの影響を評価する。NV-/VNR-SRAM セルのそれぞれで、最もばらつきの影響を受けやすい動作(NV-SRAM における MTJ の書き込み電流、VNR-SRAM における ULV リテ

ンションの SNM など)を選び, ローカルばらつきによるモンテカルロシミュレーションを行った. これにより目標値を下回ったセル(不良セル)についてはそれを補償する冗長セルの割合を評価することで許容できるか否かを評価した. 図 2.3.6 にローカルばらつきを考慮したモンテカルロシミュレーションの解析結果を模式的に示す. 横軸は評価する物理量(x)を, 縦軸はその出現確率を示す. このとき, ある基準値(x_{target})を下回る領域の積分値が不良セルの存在率(P_{fail})となる. つまり, P_{fail} は以下の式で表せる.

$$P_{\text{fail}} = \int_0^{x_{\text{target}}} F(x) dx \quad (2.28)$$

必要な冗長セルの割合(P_{redun})は, 冗長セルの中にも一定の割合で不良セルが存在することから, 以下のように P_{fail} の等比級数の和で表せる.

$$P_{\text{redun}} = \sum_{k=1}^{\infty} (P_{\text{fail}})^k = \frac{P_{\text{fail}}}{1 - P_{\text{fail}}} \quad (2.29)$$

P_{redun} はチップの一般的な歩留まり(例えば 90%程度)に対して, 十分小さい割合におさまっていれば, 従来と同じ割合で冗長セルを配置することで十分に補償できる.

(ii)シミュレーション条件

本論文では Synopsys 社の HSPICE を用いてシミュレーションを行った. シミュレーションには, ルネサスエレクトロニクス社の 65nm CMOS プロセスの process design kit (PDK) [9]を用い, 温度は室温 (25°C)を仮定した. トランジスタモデルは BSIM4 ベースの LP モデルを用いた. プロセスコーナーにおける解析には, しきい値に関して 3σ のばらつきを想定したデバイスパラメータセットが支給されているため, これを用いた. また, ローカルばらつきを与えるパラメータについては, LSTP モデルのばらつきパラメータを LP モデルに適用して評価した.

2.4 まとめ

本章では NV-/VNR-SRAM の性能評価指標および設計方法・指標について以下の項目を詳細に説明した.

- ・本論文で構築した NV-/VNR-SRAM のシステムモデル(パワースイッチ, アレイ)の各

種動作の実現方法について説明した。

- 本論文で用いる性能評価指標である BET (T_{BE})とリーク電力削減率(γ_{PG})について定式化し, 特に PG の粒度(実行頻度)の指標となる T_{BE} については詳細な解析により増大の要因を考察し, NV-/VNR-SRAM の電力・エネルギー性能を高めるための設計指針を得た。

- NV-/VNR-SRAM アレイにおける電力・エネルギー性能の解析モデルを構築した。

- 本論文で用いるばらつきを考慮したセルの設計方法と設計指標をについて述べた。

以上から不揮発/擬似不揮発性 SRAM の性能評価指標を構築した。

2.5 Appendix

•周辺回路の構成と設計

以下に本論文で用いた周辺回路の構成と設計を示す。図 2.5.1 に(a)プリチャージ, (b)セクタ, (c)ライトドライバ, (d)センスアンプ, (e)FF の構成を示す。同図の数字は pMOS/nMOS のチャンネル幅のチャンネル長に対する比 (W/L)を示している。例えば, pMOS が 30 倍, nMOS が 20 倍なら 30/20 と表記する。また, pMOS, nMOS のいずれかのみの場合には単に数字のみ示している。各ゲートの設計については何も指示がない場合は原則として 30/20, それ以外の場合は数字を記載した。デコーダのアドレスラッチに用いた FF は, 行デコーダのみ速度性能を考慮して出力ドライバが大きい FF を用い, それ以外の FF は電力性能を優先して出力ドライバが小さい FF を用いた。プリチャージ, ライトドライバ, FF は既存の SRAM 設計を参考にして設計した。セクタは通常 SRAM の書き込み動作時にビット線電位を駆動できるよう微調整した。センスアンプはアームラッチ型の構成[5]を用い, 各トランジスタはスタンダードセルを参考にしてチャンネル幅を決定した。図 2.5.2 にワードデコーダとそのデコーダセルの詳細を示す。ワードデコーダはアドレスラッチ用の FF とデコーダセルで構成した。256 行のワード線を 8bit アドレスで指定する。デコーダセルは 8 ビットアドレスのデコードに必要なデコーダツリーに, BE 信号とクロック信号による制御用のゲートと出力ドライバを接続している(出力ドライバは既存の SRAM 設計からワード線の負荷に応じて拡大/縮小した)。図 2.5.3 に NV-SRAM の SR デコーダを示す。基本的には図 2.5.3 と同様だが, クロック信号はストアイネーブル信号(STEN)により制御できる。また, リストアイネーブル信号(RSTEN)を用いてリストア動

作時にデコーダドライバを複数導通できるような構成とした。図 2.5.5 にカラムデコーダの構成を示す。サブアレイは列方向に 4 ワードの切り替えができるような構成としているため、2 ビットの FF(CDF)と 4 つのデコーダセル(CDC)が必要である。デコーダセルの構成も図 2.5.5 に示す。

2.6 図と表

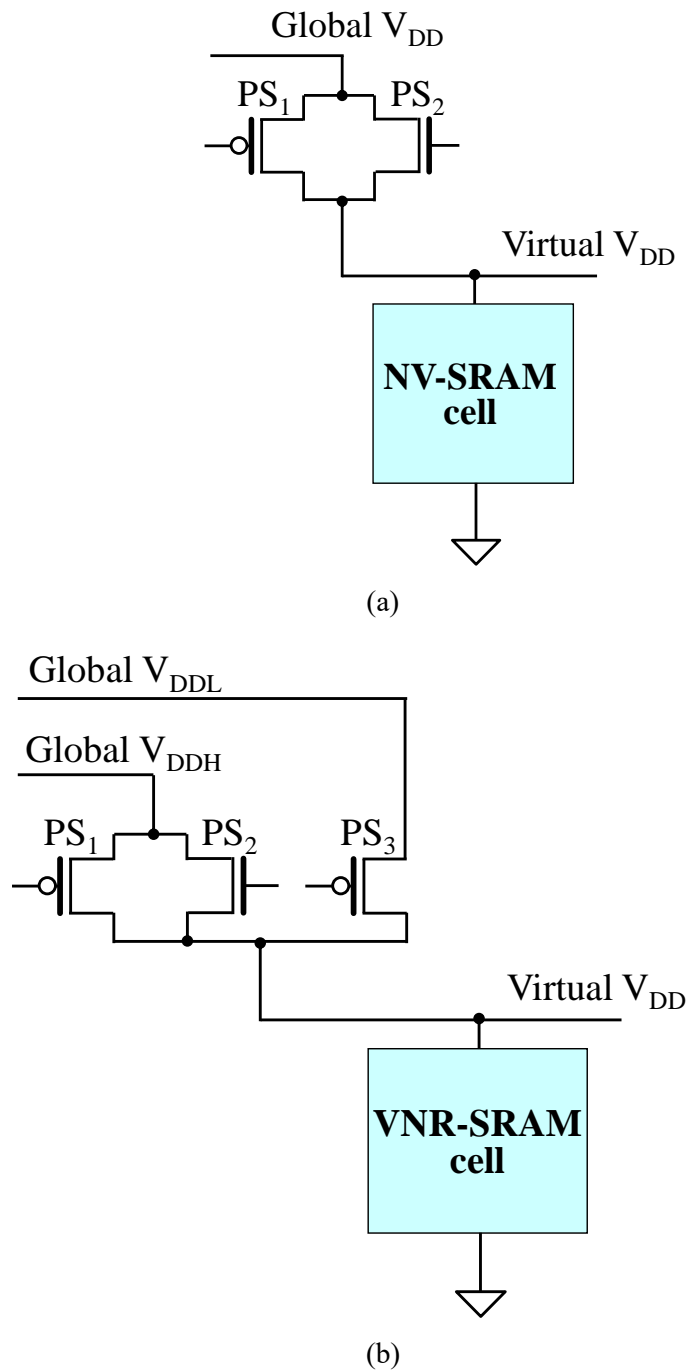
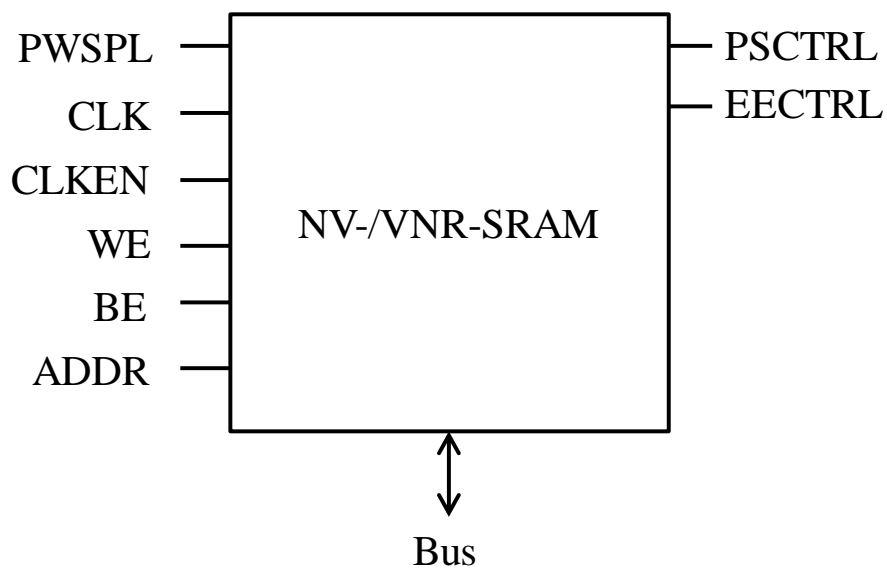
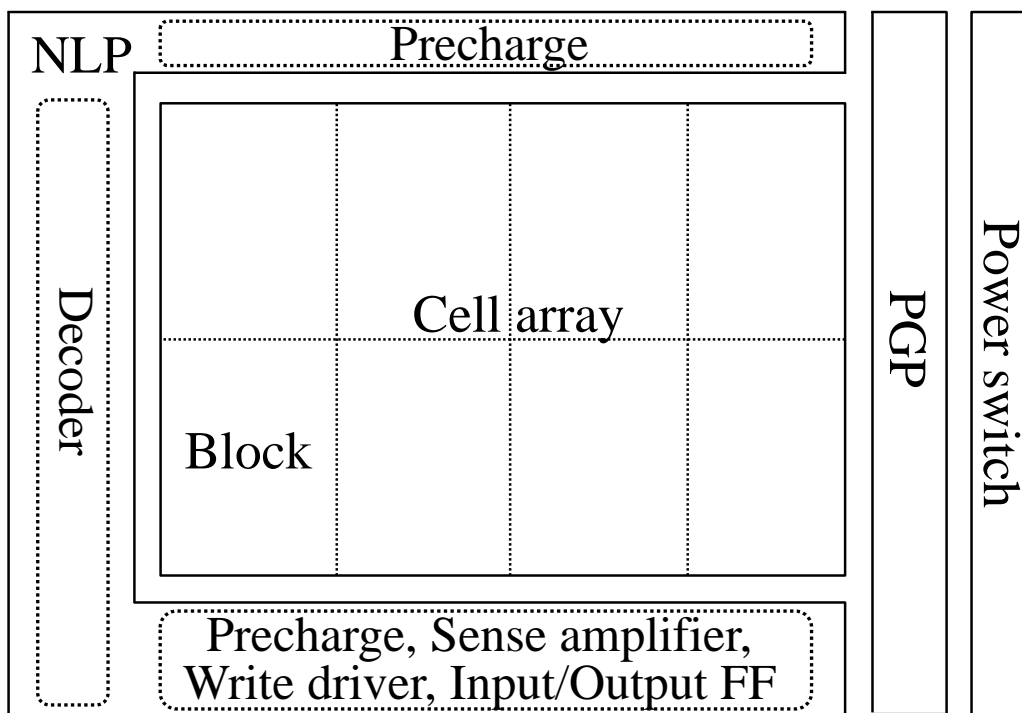


図 2.1.1 (a)NV-SRAM セルのパワースイッチ構成と(b)VNR-SRAM セルのパワースイッチ構成



(a)



(b)

図 2.1.2 (a)NV-/VNR-SRAMの各種制御信号と(b)NV-/VNR-SRAMサブアレイの回路構成

表 2.1.1 本論文におけるサブアレイの標準構成

Cell array	256bit × 256bit (8kB)
Block	64bit × 128bit (1kB)
NLP	Precharge, Selector, Write driver Sense amplifier, Word decoder, Column decoder, input/output FF
PGP NV-SRAM	SR decoder, CTRL driver
VNR-SRAM	CTRL driver

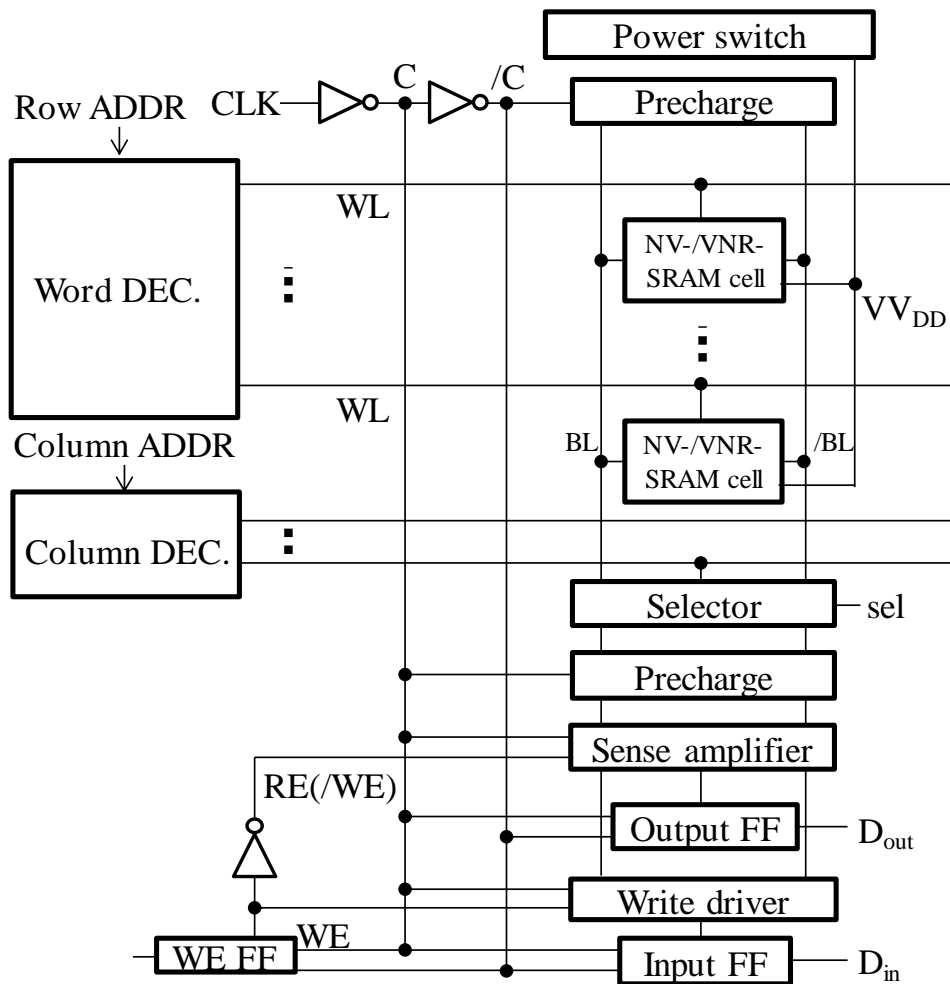
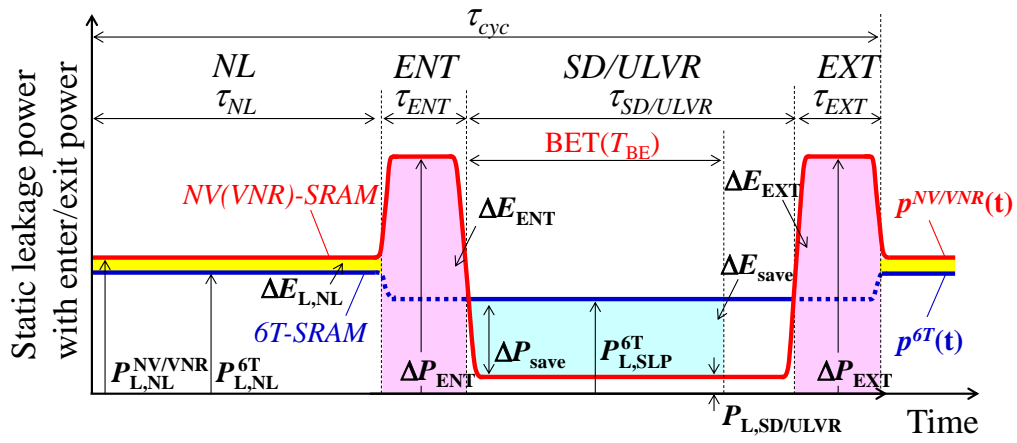


図 2.1.3 ブロック内の 1 カラムにおけるセル, パワースイッチ, NLP の構成

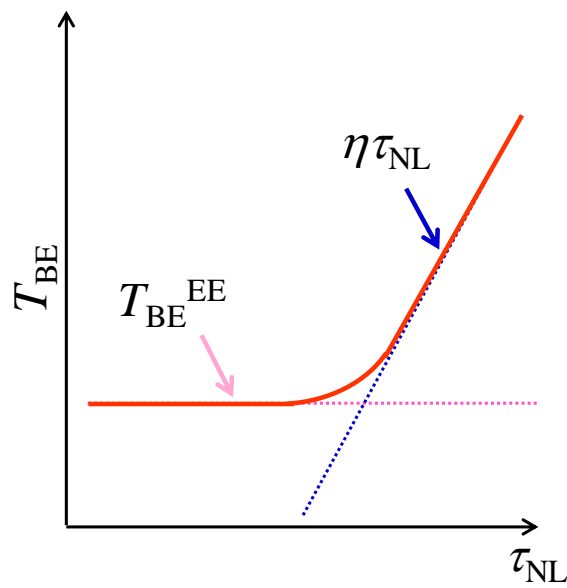
表 2.1.2 NV-/VNR-SRAM の各動作モードにおけるセルアレイ, NLP, PGP の回路状態

	Circuits	Operation modes			
		Normal operation	Enter operation	Exit operation	Shutdown/ ULV retention
NV-SRAM	Cell array	NL	Selected: ENT Waiting: SLP Completed: SD	Selected: EXT Waiting: SD Completed: SLP	SD
	NLP	ON	OFF	OFF	OFF
	PGP	OFF	Selected: ENT Waiting: OFF Completed: OFF	Selected: EXT Waiting: OFF Completed: OFF	OFF
VNR-SRAM	Cell array	NL	Selected: ENT Waiting: SB Completed: ULVR	Selected: EXT Waiting: ULVR Completed: SB	ULVR
	NLP	ON	OFF*	OFF*	OFF*
	PGP	ON	ON	ON	ON

*プリチャージのみ電源遮断を行わない。



(a)



(b)

図 2.2.1 (a) NV-/VNR-SRAM と 6T-SRAM の各動作モードにおける消費電力の推移
(b) T_{BE} の τ_{NL} 依存性の模式図

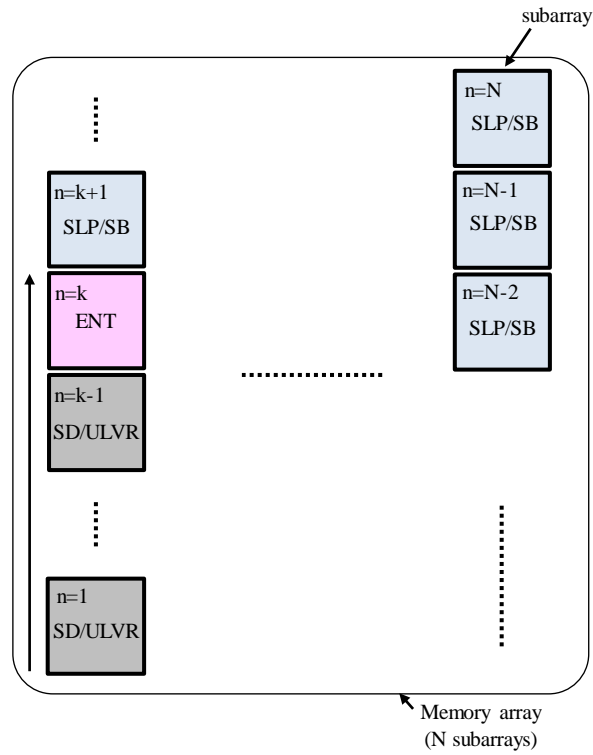


図 2.2.2 ENT モードにおける NV-/VNR-SRAM の各サブアレイの回路状態

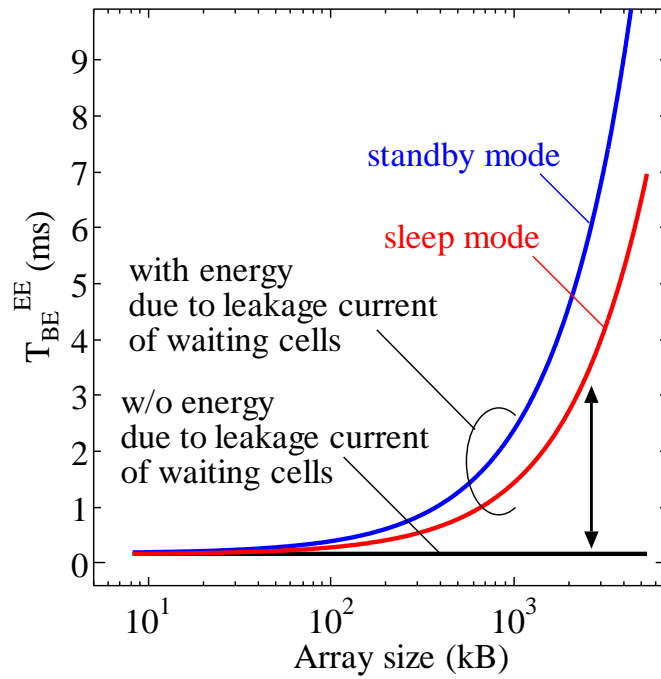


図 2.2.3 T_{BE}^{EE} のアレイサイズ依存性

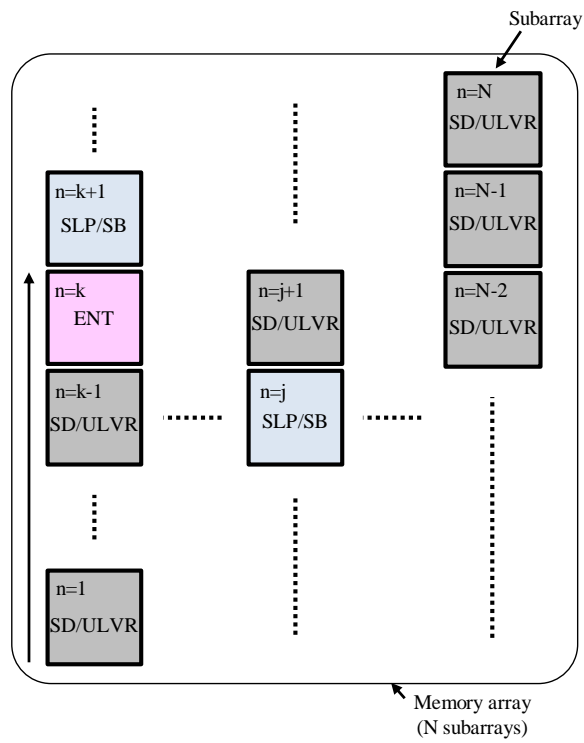


図 2.2.4 一部の ENT 動作を省略する場合の NV-/VNR-SRAM の各サブアレイの回路状態

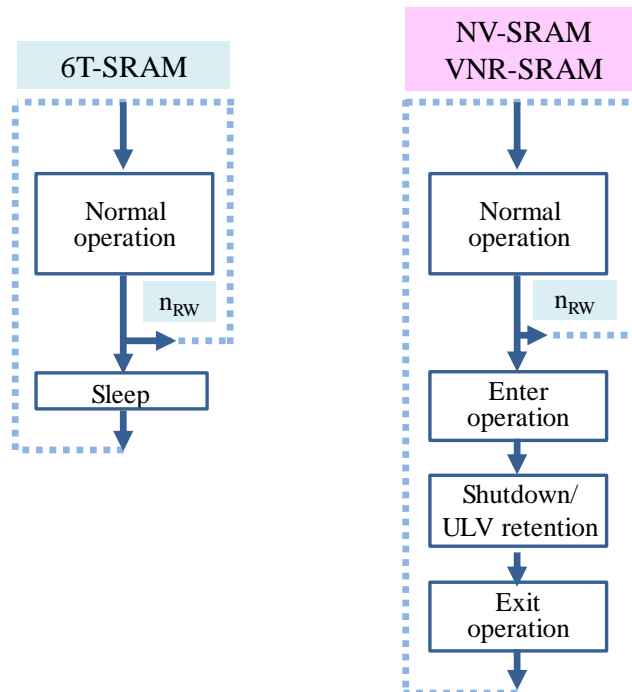


図 2.2.5 本論文における 6T-SRAM と NV-/VNR-SRAM のベンチマークシーケンス

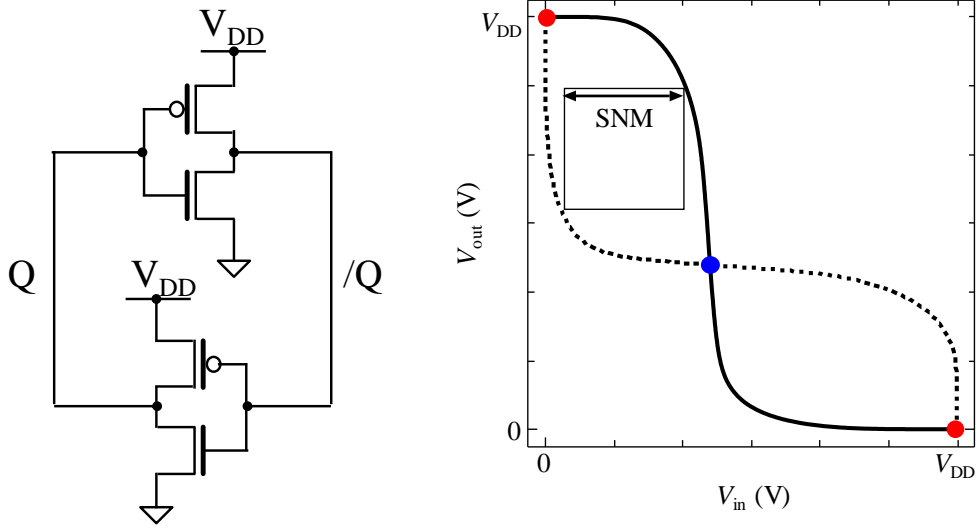


図 2.3.1 通常のインバータで構成した双安定回路とバタフライカーブ

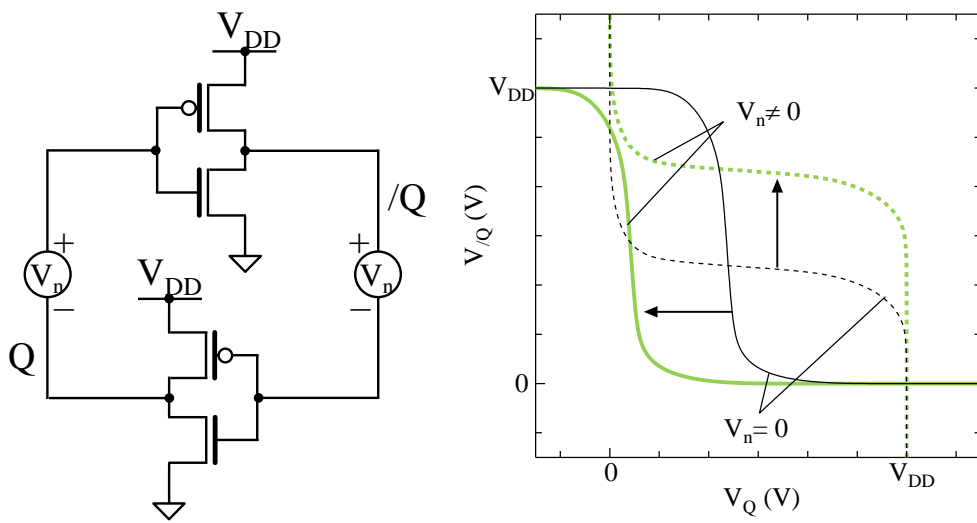


図 2.3.2 記憶ノードにノイズ源を挿入した場合の双安定回路とバタフライカーブ.

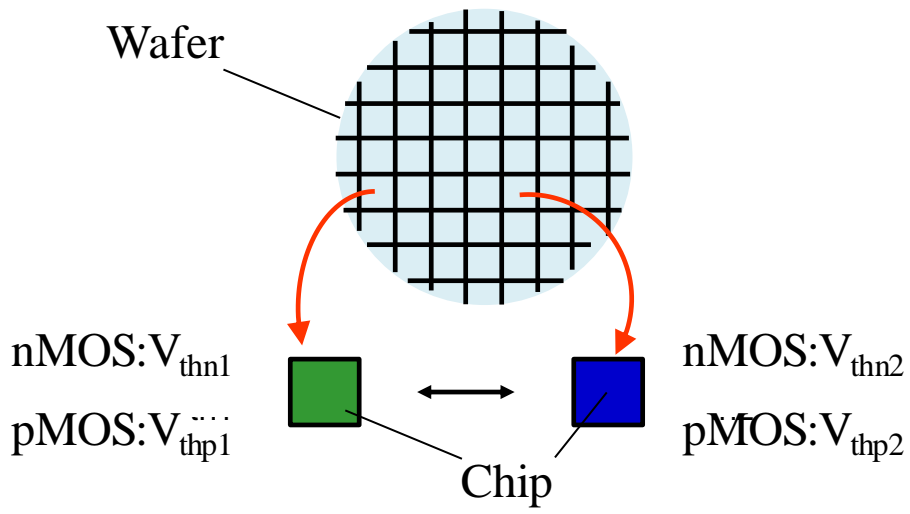


図 2.3.3 グローバルばらつきの概念図

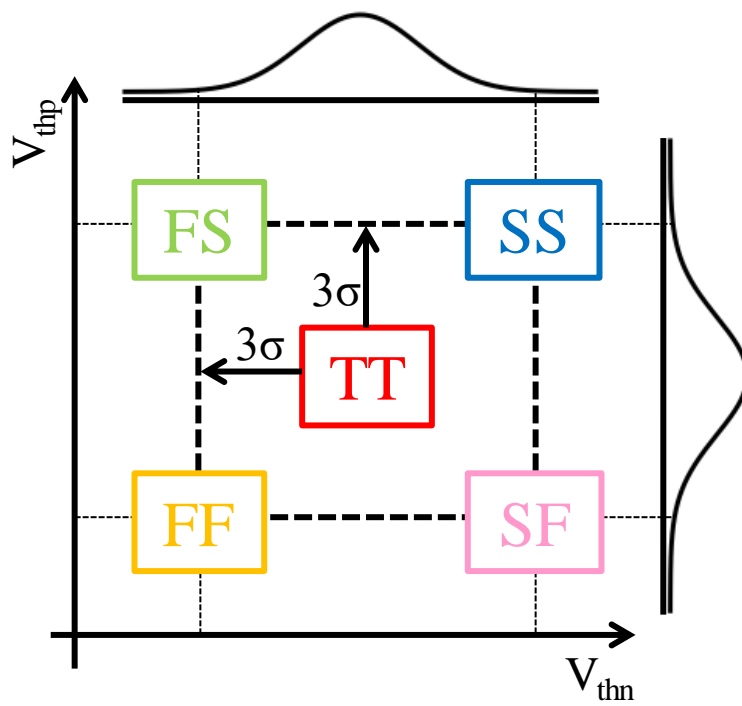


図 2.3.4 プロセスコーナー

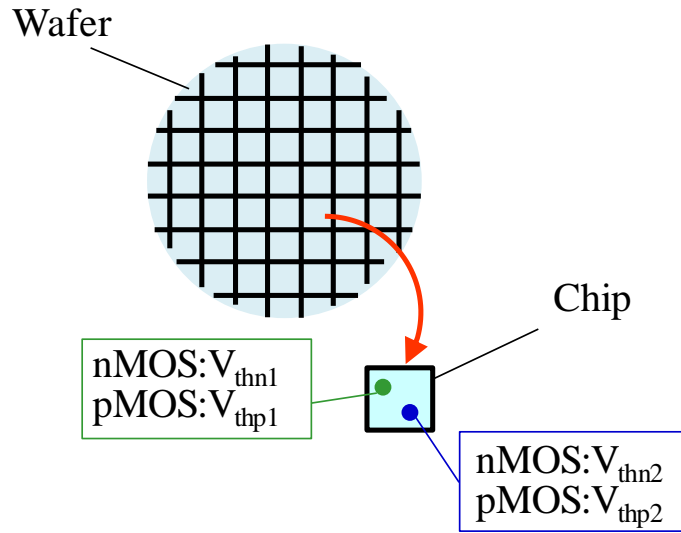


図 2.3.5 ローカルばらつきの概念図

表 2.3.1 各種低電圧 SRAM におけるワースト SNM の報告値

Minimum SNM (mV)	Average (mV)
36	36 ($\sim 1.5k_B T/q$)
53	54.3 ($\sim 2k_B T/q$)
53	
57	
68	83.3 ($\sim 3k_B T/q$)
87	
95	

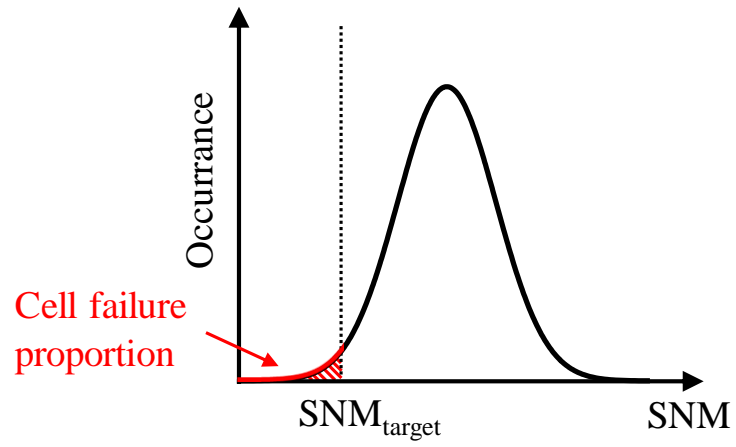


図 2.3.6 ローカルばらつきを考慮したモンテカルロシミュレーションの解析結果の模式
図

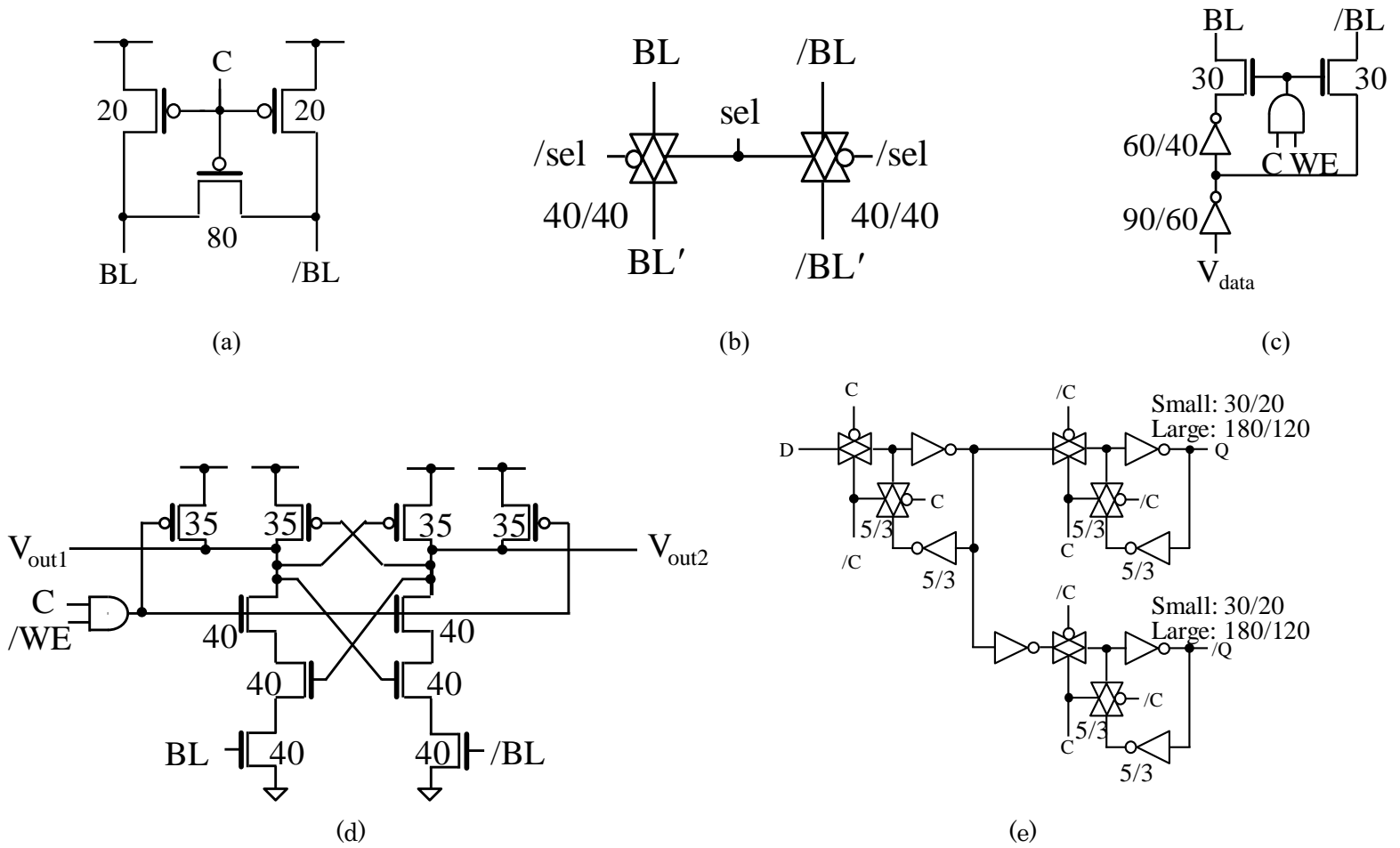


図 2.5.1 (a)プリチャージ, (b)セレクタ, (c)ライトドライバ, (d)センスアンプ, (e)FF の構成

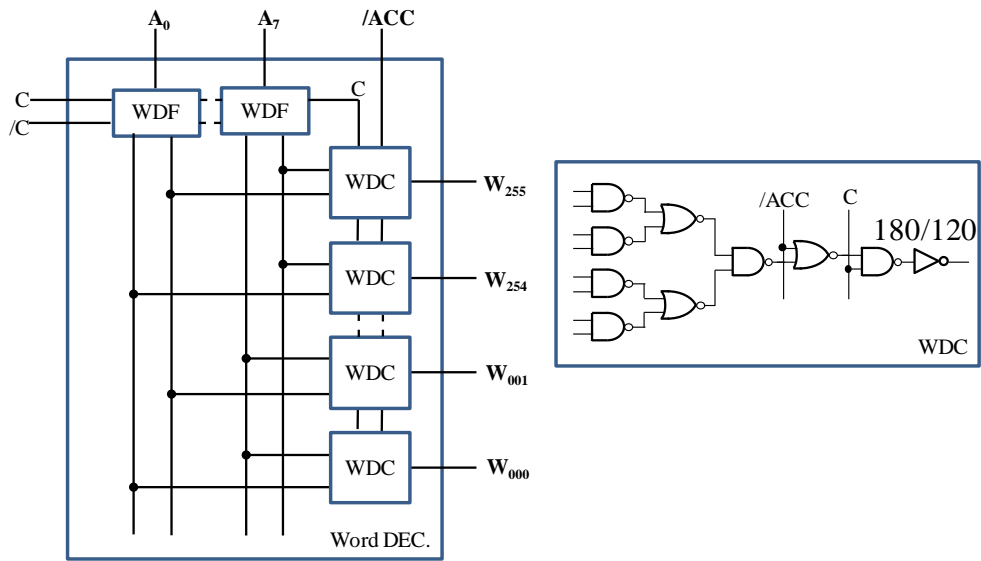


図 2.5.2 ワードデコーダとデコーダセルの構成

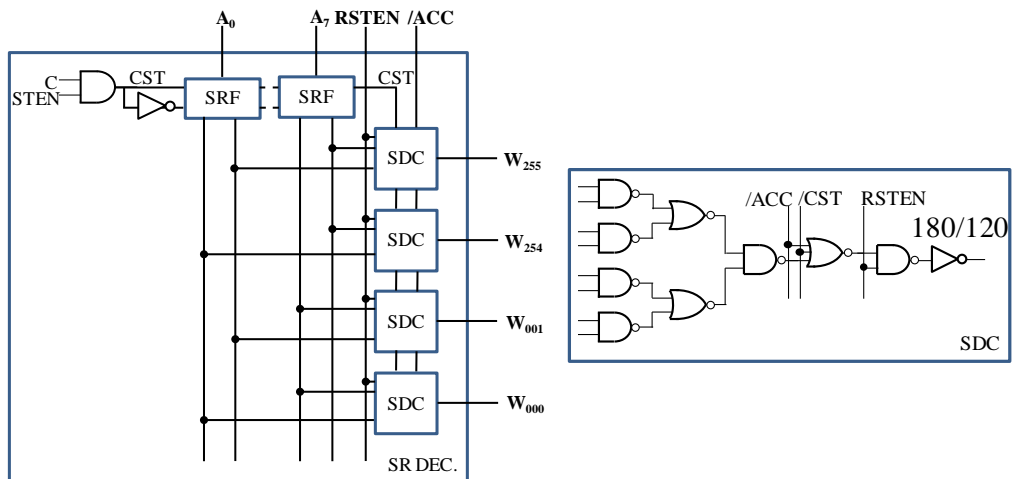


図 2.5.3 SR デコーダとデコーダセルの構成

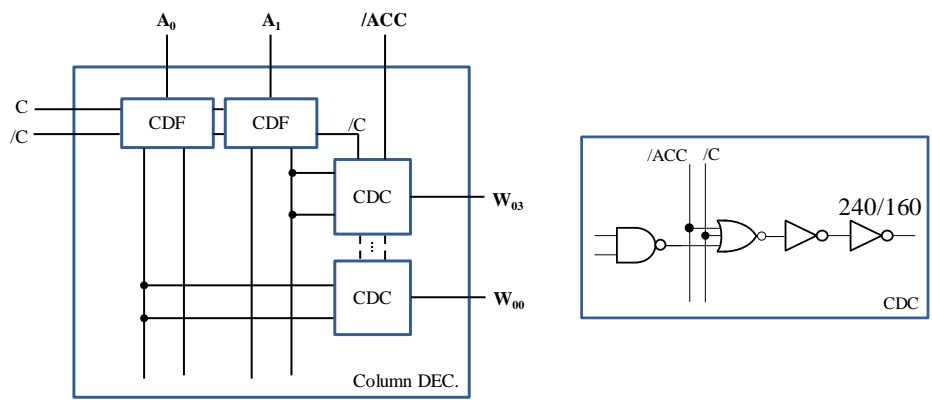


図 2.5.4 カラムデコーダとデコーダセルの構成

2.7 参考文献

- [1] H. B. Bakoglu, “Circuits, Interconnections, and Packaging for VLSI”, Addison-Wesley Publishing Company, Inc., 1990.
- [2] M. Nobakht and R. Niaraki, “A new 7T SRAM cell in sub-threshold region with a high performance and small area with bit interleaving capability”, IET Circuits Devices Syst. **13**, 873 (2019).
- [3] S. Pal, S. Bose, W. H. Ki, and A. Islam, “Characterization of half-select free write assist 9T SRAM cell”, IET Circuits Devices Syst. **13**, 584 (2019).
- [4] VLSI Design and Education Center (VDEC), University of Tokyo, <http://www.vdec.u-tokyo.ac.jp/English/index.html>
- [5] R. J. Baker, “CMOS Circuit Design, Layout, and Simulation,” Hoboken, New Jersey: John Wiley & Sons, Inc., 2010.

第 3 章

不揮発性 SRAM

この章では、はじめに SRAM の基本回路である双安定回路の動作について解説し、不揮発性メモリ素子を用いた双安定回路の不揮発化の原理を述べる。次いで、本論文で不揮発性メモリ素子として用いる強磁性トンネル接合 (Magnetic tunnel junction; MTJ) について述べる。次に、本論文で用いる不揮発性 SRAM (NV-SRAM) の回路構成と動作原理を示し、MTJ の書き込み電流と、NV-SRAM の各種ノイズマージンを指標としたセルの設計方法と、セルのリーク削減アーキテクチャについて述べる。そして、最適設計されたセル・アレイと周辺回路から構成される NV-SRAM の PG 性能について示す。特に、NV-SRAM の電力削減効果と Break-even time (BET) の評価結果について述べる。さらに、不揮発性 SRAM のエネルギー性能をさらに向上することが期待できる階層型ストアフリー電源遮断アーキテクチャ (Hierarchical store-free shutdown architecture; HSF) と無用データのフラッシュを用いた PG アーキテクチャ (Proactive useless-data flush; PUDF) を提案し、その効果も議論する。最後に、試作した Test element group (TEG) から抽出したパラメータを用いて NV-SRAM の電力・エネルギー性能を検証する。

3.1 不揮発性双安定回路の実現方法

第 2 章で述べたように、SRAM セルの基本回路である双安定回路は二つの安定な状態を記憶することができ、そのノイズ耐性はバタフライカーブのローブの開き (SNM) によって決まる。また、バタフライカーブの形状は電源電圧により変化する。図 3.1.1 に十分に大きい電源電圧 (V_{DD}) を供給した場合と、電源遮断を想定して電源電圧を非常に小さい電圧 (V_{DD}) まで絞った場合の、通常のインバータで構成した双安定回路のバタフライカーブを示す。Q、 \bar{Q} はそれぞれ記憶ノードを示す。電源電圧が低下するとバタフライ

カーブのローブは小さくなり、最終的にローブが消失し、双安定状態を実現できなくなる。したがって、通常電圧から電源遮断を行うと、同図(a)に示す安定点 A または B の状態はいずれも非双安定状態の D に遷移して行き、記憶情報は失われる。続いて再び電源を投入しても、同図(b)のように非双安定状態の D を通った状態は、ノイズやばらつきなどによって A または B の状態に、もとの記憶内容とは無関係に遷移する。したがって、電源遮断前に保持されていた状態に復帰するとは限らない。そのため、双安定回路は揮発性記憶回路と呼ばれる。

SRAM セルを不揮発化するには、電源復帰後に電源遮断前と同じ状態に必ず復帰することが必要である。電源遮断前の安定状態への復帰は電源復帰初期時に双安定回路の記憶ノードに、この状態に遷移できるような電位差を強制的に与えることで実現できる。そこで、記憶ノードに不揮発性メモリ素子を接続した不揮発性双安定回路が提案されている[1-17]。図 3.1.2 に不揮発性双安定回路の構成と電源復帰時の記憶ノードの電位を示す。不揮発性メモリ素子は 1 ビットのデータを異なる 2 つの抵抗値などに交換して記憶するため、記憶ノードに接続された不揮発性メモリ素子の状態に応じて記憶ノードの充電速度に差が生じ、記憶ノード間に電位差が発生する。例えば、Q 側の素子が低抵抗状態、 \bar{Q} 側の素子が高抵抗状態の場合、Q の電位が \bar{Q} よりも低くなり、Q が L レベルに \bar{Q} が H レベルに遷移する(図 3.1.2(a))。逆に Q 側の素子が高抵抗状態、 \bar{Q} 側の素子が低抵抗状態の場合、Q の電位が \bar{Q} よりも高くなり、Q が H レベルに \bar{Q} が L レベルに遷移する(図 3.1.2(b))。したがって、双安定回路の記憶ノードの状態に応じて、不揮発性メモリ素子の状態を変化させることができれば、不揮発化が実現できる。

不揮発性双安定回路に用いる不揮発性メモリ素子には、これまでに様々な素子が検討されてきた。原理的には EEPROM のメモリセル[1, 2]や, MRAM[3-7], ReRAM[8-11], PRAM[12,13], FeRAM[14-17]などの不揮発性メモリ素子を用いることができるが、実際にはシステムから要求される要件から選ぶ必要がある。電氣的に書き込みと読み出しができること(読み出しでは大きな抵抗変化率(または静電容量の変化)が得られること)、高密度集積化への適合性や CMOS プロセスとの親和性があることなどは必須の条件であるが、特に動作電圧(書き込み電圧)は重要な判断基準となる。ReRAM, PRAM, FeRAM の不揮発性メモリ素子は書き込みに比較的大きな電圧を有し、この電圧は近年のマイクロプロセッサ(μP)や SoC で用いる電源電圧より高い[18]。一方、MRAM の不揮発性メモリ素子である強磁性トンネル接合(Magnetic tunnel junction; MTJ)は、必要な

書き込み電流が確保できれば、駆動電圧に制約はない。したがって、MTJ は電圧駆動に適した不揮発性メモリ素子であり、不揮発性双安定回路に適した不揮発性メモリ素子であると考えられる。本論文では MTJ を用いて NV-SRAM セル(以後 NV セルと呼ぶ)を構成する。

MTJ は 2 層の強磁性電極でトンネル障壁を挟み込んだ構造を有する不揮発性メモリ素子である(図 3.1.3)。片方の強磁性電極は、磁化されたまま固定され、ピン層(pinned layer)と呼ばれる。ピン層は反強磁性体との磁氣的な交換バイアスによって磁化を固定、または極薄 Ru 層を介した二つの強磁性層による層間反強磁性結合を用いて磁化を固定することで実現できる。MTJ のもう一方の強磁性電極はフリー層(free layer)と呼ばれ、磁化の向きを固定せず磁化反転が可能である。MTJ では強磁性電極の相対的な磁化の向きが平行の場合と反平行の場合とでトンネル抵抗が異なる。強磁性電極の相対的な磁化の向きが平行であればトンネル抵抗は低く、反平行ならば高くなる。これはトンネル磁気抵抗効果(Tunneling magnetoresistance; TMR)と呼ばれる。したがって、MTJ ではこの相対的な磁化の状態として情報を記憶し、このトンネル抵抗から情報を読み出すことが可能となる。この情報の書き込みは、配線に電流を流すことで磁場を発生させてフリー層を磁化反転させることで実現できるが、近年ではスピン注入磁化反転(Current-induced magnetization switching; CIMS)と呼ばれる電流誘起の磁化反転機構を用いることが一般的である。これはスピン・トランスファ・トルクと呼ばれる物理現象に基づくことから、スピン注入磁化反転を用いた MTJ によって構成された MRAM のことを STT-MRAM と呼ぶことが多い。強磁性電極に CoFeB、トンネル障壁に MgO を用いた MTJ では、室温において数 100%もの高いトンネル抵抗の変化率(TMR 比と呼ばれる)が得られ、また室温において 500%を超える TMR 比も報告されている[19]。表 3.1.1 に本研究に用いた MTJ の特性を示す。これらの値には最近の典型的な報告値を用いた[20-22]。また、この MTJ の特性を図 3.1.4 に示す。

3.2 不揮発性 SRAM の回路構成と動作原理

MTJ を用いた不揮発性 SRAM はこれまでに様々なセル構成が提案されている。図 3.2.1(a)-(c)に提案されたセル構成の例を示す[3-5]。これらのセル構成は電源復帰初期時に双安定回路の記憶ノードに、電源遮断前の状態に応じた電位差をつけることで不

揮発化を実現している。しかし、これらは 1.5 節で述べたノーマリオフアーキテクチャ (NOF) を用いることを前提に提案されたもので、通常の SRAM 動作時に MTJ に書き込むように構成されている。MTJ の書き込みでは、書き込みエラーを抑えるために通常は 10ns 程度の MTJ へのストア時間を必要とする。したがって、GHz 動作を前提とした場合、書き込み時のクロックオーバーヘッドが大きく、本論文で目指す不揮発記憶を用いた SRAM 技術には不適である。また、これに加えて、NOF アーキテクチャでは、通常の SRAM 動作時に MTJ に書き込むことから、通常動作時のエネルギー性能が大幅に劣化する。以上から、不揮発記憶だけでなく通常の双安定回路動作も考慮して MTJ を導入する必要がある。

そこで、本論文では図 3.2.2 のように、トランジスタを介して不揮発性メモリ素子を双安定回路に接続したセル構成を用いる[7]。この回路は通常の 6T-SRAM セル (以後 6T セルと呼ぶ) と同様のインバートループからなる双安定回路、この双安定回路の記憶ノード(Q, /Q) と 2 本のビット線(BL, /BL) とを接続するパストランジスタ(M_B , $M_{/B}$)、このパストランジスタを駆動するワード線(WL)、記憶ノードにパストランジスタ(M_1 , M_2) を介して接続された強磁性トンネル接合素子(MTJ₁, MTJ₂)、このパストランジスタに接続された SR 線、MTJ₁ と MTJ₂ に接続された CTRL 線 から構成される。図 3.2.2(a)のセルは 1.3 節で示した仮想電源電圧アーキテクチャにおける構成である。この構成では MTJ のピン層を記憶ノード側に、フリー層を CTRL 線側に接続する。また、 M_1 , M_2 には nMOS を用いる。以上は仮想電源電圧アーキテクチャの構成方法であるが、図 3.2.2(b)に示すように仮想接地アーキテクチャで構成することもできる。この場合、パワースイッチは nMOS、 M_1 , M_2 には pMOS を用いる。また、MTJ はピン層を CTRL 線側に、フリー層を記憶ノード側に接続する。本研究では仮想電源電圧アーキテクチャで構成した NV セルを用いる。

以上のセル構成では、不揮発記憶を用いないときには M_1 と M_2 を用いて不揮発性メモリ素子を双安定回路から電氣的に切り離し、通常の双安定回路と同等の動作を実現することができる。不揮発記憶を用いるとき、すなわち電源遮断を行う直前にこのトランジスタを導通させ、MTJ に記憶ノードのデータを書き込むことで、データを失うことなく電源遮断できる。このセル構成の特徴をまとめると、①通常の SRAM 動作時には双安定回路は不揮発性メモリ素子から分離されるため、MTJ は通常の SRAM 動作に影響しない。②不揮発記憶は電源遮断の前のみ行い、通常の SRAM 動作には用いないため、

不揮発性メモリ素子への書き込みエネルギーを最小限に抑えるとともに、MTJ への書き込みレイテンシは通常の SRAM 動作に影響しない。この2つの特徴により、通常動作の速度性能を劣化させずに、不揮発記憶の機能をパワーゲーティング (PG) に導入することが可能となり、BET (T_{BE}) を最小限に抑えた高効率の PG を実現できる。ロジックシステムに不揮発記憶を用いるアイデアは古くからあるが、これを PG に用いることは、本著者の所属研究室から初めて提案された。また、NOF も提案されているが、これは第 1 章で述べたように PG とは異なる別のアーキテクチャである。

以下、NV-SRAM の動作について述べる。通常の SRAM 動作は 6T-SRAM と同様に読み出しと書き込みの動作を行う。

読み出し動作は BL と /BL をプリチャージ後、BL と /BL を電氣的に切り離す。この状態から WL にバイアスを加えて M_B ($M_{/B}$) を導通させ、L レベルの記憶ノードに接続された BL または /BL の一方が放電される。これをセンスアンプによって検出することで、読み出しが実現できる。書き込み動作は、まず BL および /BL をプリチャージしておき、L レベルにしたい記憶ノードに接続される BL または /BL を放電する。次いで、パストランジスタ M_B , $M_{/B}$ を導通させることで双安定回路を所望の記憶状態 ($(V_Q, V_{\bar{Q}}) = (L, H)$) または (H, L) に強制的に移行させることができる (2.2 節参照)。なお、読み出し/書き込みいずれも安定な動作を確保するために pMOS, nMOS の W/L (チャネル幅/チャネル長) を適切に設計する必要がある。一般に 6T セルの設計では、(インバータの pMOS の W/L) < (nMOS パストランジスタの W/L) \leq (インバータの nMOS の W/L) とすることが多い。

NV-SRAM では不揮発記憶動作 (双安定回路のデータを MTJ に書き込む動作) を行ってから電源遮断を行うことで、電源遮断を行ってもデータを失わず保持し続けることができる。このためには、電源遮断の前に双安定回路から MTJ へデータを書き込み、電源遮断から復帰する場合に MTJ から双安定回路へデータを書き戻す必要がある。これらを以後それぞれストア動作、リストア動作と呼ぶことにする。図 3.2.3 にストア動作、電源遮断、リストア動作における回路の状態を示す。ストア動作は M_B , $M_{/B}$ を閉じた状態で、パストランジスタ M_1 , M_2 を開くだけで、MTJ₁, MTJ₂ にデータを書き込むことができる。まず通常の SRAM 動作終了時に双安定回路の記憶ノードの電位 V_Q , $V_{\bar{Q}}$ がそれぞれ H レベル, L レベルであるとする。ストア動作はこの状態から V_{CTRL} を L レベルにしたまま、 M_1 , M_2 を導通する。この時、2 つの記憶ノードのうち H レベルの記憶ノードから MTJ₁ を介して CTRL 線にストア電流が生じる。ピン層からフリー層に電流が流れ (電子はフリー

層からピン層に流れる), MTJ_1 は反平行磁化で安定化する(はじめに平行磁化であれば反平行磁化に, はじめに反平行磁化であればそのままの磁化状態となる). 以後, この動作を H ストア動作と呼ぶことにする. 次に, M_1 , M_2 を導通した状態で, V_{CTRL} を H レベルに引き上げる. この場合 CTRL 線から L レベルの記憶ノードに向かってストア電流が生じる(電子はピン層からフリー層に流れる)ため, MTJ_2 は平行磁化で安定化する(はじめに反平行磁化であれば平行磁化に, はじめに平行磁化であればそのままの磁化状態となる). 以後, この動作を L ストア動作と呼ぶことにする. H ストア動作と L ストア動作によって, MTJ_1 , MTJ_2 への書き込みが完了すると, セルへの電源の供給を遮断することができる. 電源遮断の状態から再び電源を供給して復帰させる場合にはリストア動作を行う. リストア動作は, V_{CTRL} を L レベルにしたまま M_1 , M_2 を導通した状態でパワースイッチを導通させ, セルの V_{DD} を電源電圧まで引き上げる. H ストア動作時に書き込みを行った MTJ は反平行磁化になっているため記憶ノードから MTJ_1 , MTJ_2 を介して CTRL 線へ流れる電流は小さいが, L ストア動作時に書き込みを行った MTJ は平行磁化になっているため CTRL 線へ流れる電流は大きい. したがって, 反平行磁化の MTJ_1 の接続されている記憶ノードが充電されやすく, 平行磁化の MTJ_2 の接続された記憶ノードよりも高電位となり, インバータループは電源遮断前と同じデータを記憶ノードに復元できる. リストア動作後は M_1 , M_2 を遮断して再び通常の SRAM 動作に移行する.

これまで, 本論文の NV-SRAM を用いた PG は, 本著者の所属研究室で先行研究が進められ, NV-SRAM セルの提案, リーク制御アーキテクチャの開発, NV-SRAM TEG の試作が行われていたが, ばらつきを考慮したセルの設計, アレイレベルの設計・性能評価・アーキテクチャ開発は検討されていない(または検討不十分の)ため, これらの課題を本論文で検討した.

3.3 不揮発性 SRAM の設計

3.3.1 セル設計

図 3.3.1 に NV セルとパワースイッチを示す. 通常電圧の SRAM 動作用 (PS_1) とスリープ用 (PS_2) の 2 種類のパワースイッチを用いて電源配線 (V_{DD}) からセルに電源の供給をする. PS_2 は nMOS でも構成されることがあり, この場合しきい値程度だけ V_{DD} が減少

することにより PS_2 を線形領域で動作させることができるためばらつきの影響を受けにくいという特長を持つ。また、セルを構成するトランジスタのボディは pMOS, nMOS をそれぞれ電源線, 接地線に接続した。これにより V_{DD} の減少にともないトランジスタのしきい値が増加する方向に実効的にボディバイアスが印可されるため、電力削減効果を高めることができる(このようなボディの接続方法は PG で用いられる典型的な構成である [23])。

表 3.3.1 に NV セルとパワースイッチのデバイスパラメータとバイアスの表記をまとめた。セルとパワースイッチの設計は MTJ への書き込み電流(ストア電流)と各動作モードにおける SNM を考慮して決定した。各トランジスタのチャンネル幅 W の設計は、65nm バルク MOSFET の SRAM の設計[24]を参考にして、ロード(LD)トランジスタ: $W_{LD}=100\text{nm}$, ドライバ(DRV)トランジスタ: $W_{DRV}=150\text{nm}$, パス(PSS)トランジスタ (M_B, M_B): $W_B=W_B=100\text{nm}$ と設定した。リファレンスに用いる 6T セルにもこの設計を用いた。また、 M_1, M_2 は M_B, M_B と同じ $W_{M1}=W_{M2}=100\text{nm}$ とした。

図 3.3.2 に NV セルの設計に関するフローチャートを示す。セルの設計では、まず MTJ の書き込み電流を実現することが重要となる。MTJ は臨界電流密度以上の電流を流すことで磁化反転を実現できるが、臨界電流密度は MTJ の断面積に反比例する。そこで、所望の書き込み電流を確保するためには、図 3.3.1 の NV セルで実現できるストア電流が磁化反転の臨界電流より大きくなるように MTJ の直径(ϕ_{MTJ})を決めるか、MTJ の臨界電流を実現できるようにセルサイズを設計することが必要である。いずれも検討したが、本論文では前者の設計方法について説明する。

はじめにストア動作時に MTJ へのストア電流が臨界電流密度(J_c)の 1.2 倍を確保できるよう ϕ_{MTJ} を決定した。MTJ のばらつきによる臨界電流の偏差は 3σ で 20%程度となることが知られているため[25]、本論文でもこのようにストア電流のしきい値を設定した。ここでは PS_1 は十分に大きいサイズ($W_{PS1}=1\mu\text{m}$)で固定した(これは後述の全動作(スリープを除く)の SNM と H ストア時, L ストア時のストア電流が十分に飽和しているとみなせるチャンネル幅である)。続いて、 $W_{PS1}, V_{SR}, V_{CTRL}$ をパラメータとして必要最低限となるストア電流を設計する。ストア電流は大きなエネルギーを消費するため T_{BE} の削減にはこの設計が重要となる。 W_{PS1} が一定値以下、 V_{SR} がある一定値以上の領域では H レベル側の記憶ノードの電位が下がり保持されているデータがフリップ(別の安定点に推移)してしまう。このため、データのフリップに注意して H ストア時のストア電流を W_{PS1}, V_{SR} により最

適化する。ストア電流を確保できなければ所望の電流密度を実現できるまで ϕ_{MTJ} を縮小する。その後、設定した ϕ_{MTJ} , W_{PSI} , V_{SR} のもとで、 V_{CTRL} を適切に選ぶことでLストア時のストア電流を決定した。Lストア時の電流が取れない場合は V_{SR} を増加させて再度Lストア時の書き込み電流を最適化する。その後、全動作についてSNMが確保できるかを確認し、モンテカルロシミュレーションによりローカルばらつきの影響を調べる。SNMが十分に確保できない場合は W_{PSI} を調整する。

図 3.3.3 に ϕ_{MTJ} が 40nm, 50nm, 60nm の MTJ を用いた場合の H ストア時のストア電流密度 ($J_{\text{MTJ}}^{\text{P}\rightarrow\text{AP}}$) の V_{SR} 依存性を示す。点線は MTJ のばらつきを含めた上で必ず書き込むことができるとみなせる J_c の 1.2 倍の値を示している。いずれもデータがフリップしない領域のみプロットしている。 V_{SR} の増加とともにストア電流が増加するが、 ϕ_{MTJ} の増加にともない電流密度が減少し、所望の臨界電流密度が実現できなくなる。このため、すべてのプロセスコーナーで臨界電流密度を確保するため $\phi_{\text{MTJ}}=40\text{nm}$ とした。続いて、ストア電流が必要最低限となるようなバイアス (V_{SR} , V_{CTRL}) とパワースイッチサイズ (W_{PSI}) を設計した。図 3.3.4(a) に H ストア時のストア電流密度 ($J_{\text{MTJ}}^{\text{P}\rightarrow\text{AP}}$) の W_{PSI} 依存性を示す。プロセスコーナーは TT とした。 V_{SR} を 0.6V-0.8V で変化させており、ここでもデータがフリップしない領域のみプロットしている。 V_{SR} の増加によりストア電流は増加し、0.7V 以上で J_c の 1.2 倍を実現できる。また、図 3.3.4(b) に $V_{\text{SR}}=0.7\text{V}$ とし、それぞれのプロセスコーナーでプロットしたものを示す。SS コーナーにおいて最もストア電流が低下するため、このとき所望の電流密度となる $W_{\text{PSI}}=150\text{nm}$ に決定した。次にこの条件のもとで、L ストア時のストア電流を確保する V_{CTRL} を決定した。図 3.3.4(c) に L ストア時のストア電流密度 ($J_{\text{MTJ}}^{\text{AP}\rightarrow\text{P}}$) の V_{CTRL} 依存性を示す。先ほどの条件 ($W_{\text{PSI}}=150\text{nm}$, $V_{\text{SR}}=0.7\text{V}$) のもとで $V_{\text{CTRL}}=0.5\text{V}$ とすることで、L ストア時についても所望の電流密度を確保できる。なお、本論文では簡単のため、平行から反平行への磁化反転と反平行から平行への磁化反転で同一の臨界電流密度を用いた。

次に、図 3.3.5(a) に NV セルの各 SNM (mode=SB, READ, WRITE, H-STR, L-STR, RST) の W_{PSI} 依存性を示す。ここでは第 2 章で述べたように単体セルにパワースイッチを接続した構成を用い、このセルにノイズを入れて SNM を評価した。前述のストア電流を確保する設計 ($W_{\text{PSI}}=150\text{nm}$) において、すべてのプロセスコーナーで 78mV 以上の SNM が実現できている。リストア時にはパワースイッチを完全に導通する前に双安定回路の状態が確定した時点で (V_{DD} が V_{DD} に完全に到達する前に) M_1 , M_2 を遮断するこ

とでエネルギーオーバーヘッドを抑えることができる。本論文ではリストア動作開始から、 PS_1 のゲート電圧(V_{PG1})が0.3Vになった時点で M_1 , M_2 を遮断した(これより速く遮断すると正常に復帰することができない)。SNM(mode=RST)についてはこの時の電圧状態における値を示してある。

PS_2 の設計については $W_{PS2}=100\text{nm}$ とし、ゲートにフルスウィング電圧(1.2V)を印可して複数セルで共有する構成とし、 PS_2 を共有するセル数を N_{SLP} とした。この時、上記のスリープ以外のモードと同様に1セルにノイズを入れることもできるが、ここではワーストケースである、PSを共有するすべてのセルにノイズが入る条件で評価した。図3.3.5(b)にSNM(mode=SLP)の N_{SLP} 依存性を示す。同図にはスリープ時の $V_{DD}(V_{SLP})$ も示している。 N_{SLP} の増加にともない V_{SLP} が減少し、SNMが劣化する。すべてのプロセスコーナーでSNMを確保する最小の N_{SLP} として、 PS_2 は1024セル(64bit×16line)で共有する構成とした。この時 $V_{SLP}\sim 0.8\text{V}$ となる。後で比較のために用いる6Tセルのスリープモードにもこの設計を用いた。図3.3.5(c)に最適設計($W_{PS1}=150\text{nm}/1$ セル, $W_{PS2}=100\text{nm}/1024$ セル)におけるNVセルの各動作時のSNMを示す。いずれも78mV以上の十分に大きなSNMが確保できている。

最後にローカルばらつきを考慮したモンテカルロシミュレーションを用いて冗長セル(リダンダンシ)の評価を行った。最もばらつきの影響を受けるHストア, Lストア時のストア電流, SNMではSNM(mode=READ, SLP, RST)に対して評価を行った。図3.3.6(a)にNVセルのHストア, Lストア時のストア電流, (b)にSNM(mode=READ, SLP, RST)におけるモンテカルロシミュレーションの結果を示す。試行回数は10000回とした。ストア電流とSNMはいずれも、同図に黒の点線で示した下限を十分に上回っており、第2章で示した方法で必要になる冗長セルの割合を評価するといずれの場合も0.1%以下となる。これはチップ全体の一般的な歩留まり(例えば90%程度)に比べて極めて小さいため、従来と同じ割合で冗長セルを配置することで十分に補償できる。表3.3.2にセルの設計結果をまとめる。以後NVセルはこの設計を用いる。

図3.3.7にリファレンスとして用いた6Tセルと最適設計のNVセルのレイアウトを示す。NVセルは6Tセルに対して1.5倍程度の面積オーバーヘッドが生じるが、これは現在用いられているデュアルポートSRAMと同程度であるため、それほど大きな面積増加ではない。

次に、NVセルの読み出し速度について示す。図3.3.8(a), (b)にそれぞれ(V_Q, V_Q)=(0,

V_{DD}), (V_{DD} , 0)の場合における読み出し動作の波形を示す. この波形は第2章で示した評価回路を用いて計算し, クロックは1GHzとした. 実線はNVセル, 点線は6Tセルの波形を示している. 各図の上から, 記憶ノードの電位 (V_Q , $V_{\bar{Q}}$), ビット線の電位 (V_{BL} , $V_{\bar{BL}}$), センスアンプの出力 (V_{SAO} , $V_{\bar{SAO}}$)を示してある. 6TセルとNVセルはいずれもビット線の放電速度にほとんど差がなく, センスアンプの出力もほぼ波形が重なっていることから, 同じ読み出し速度が実現できている. このように6Tセルが十分に機能する(6Tセルで設計された)センスアンプを用いればNVセルで同じクロックで読み出すことができる. すなわち, 速度劣化は生じない. また, 読み出し時の動作時電力は6Tセルが44.73 μ W, NVセルが44.68 μ Wとなり, 同程度となる.

続いてNVセルの書き込み速度と電力について議論する. 図3.3.9(a)に(V_Q , $V_{\bar{Q}}$)=(V_{DD} , 0)から(V_Q , $V_{\bar{Q}}$)=(0, V_{DD})に書き込みを行ったときの波形を示す. クロックは1GHzとし, 上から記憶ノードの電位 (V_Q , $V_{\bar{Q}}$), ビット線の電位 (V_{BL} , $V_{\bar{BL}}$), クロック信号 (V_C , $V_{\bar{C}}$)を示す. NVセルはトランジスタ数が増加するため, セルの容量が増大し6Tセルに比べて若干反転に時間がかかる. 例えば, ビット線電位が $V_{DD}/2$ (=0.6V)に変化する時刻を比べると, Qでは24ps程度の, \bar{Q} では28ps程度の差が生じる. しかし, これはクロック周期(1ns)の2%程度であり, クロックスピードに全く影響は与えない(書き込み時のクロックオーバーヘッドを生じず, 6Tと同じクロックで書ける). 図3.3.9(b)は逆のデータの書き込み時の波形であるがこちらも同様である. よって, 書き込みについても6Tセルで十分に動作させるように設計された周辺回路を用いればNVセルを用いても動作速度上の問題は生じない. また, 動作時電力についても評価した. 6Tセル, NVセルに第2章の評価回路を用いて, セルにおける書き込み時の電力を見積もった. 6Tセルで74.32 μ W, NVセルで75.85 μ Wとなり, NVセルでは約1.6%程度の増加があるが, この程度の増加であれば, アレイを構成したときに周辺回路の影響で隠蔽できる, または周辺回路の工夫で削減することができる.

3.3.2 リーク電力削減アーキテクチャ

NVセルの T_{BE} を小さく抑えエネルギー性能を究極的に高めるには, 電源遮断時のリーク電力の削減が重要となる. また, 6Tセルに比べ, トランジスタ数の増加により発生する余剰なリーク電流も小さく抑えることが重要になる. 本論文ではこのようなリーク電流を, 電源遮断時には1.3節で説明したスーパーカットオフ(SCCMOS), 通常のSRAM動作

時には SR 線のバイアス制御により削減した。

電源遮断時にパワースイッチとセルを流れるリーク電流は SCCMOS によって削減を試みた。図 3.3.10(a)にリーク電流の V_{PG1} 依存性を示す。PS₁ のゲートに V_{DD} (=1.2V) よりも 0.2V 高い 1.4V を印加することで 1.2V を印加した場合に比べて 1 桁程度削減できる。また、 $V_{PG1}=1.4V$ までは顕著な減少がみられるが、それ以上の電圧では削減量が飽和するため 1.4V を採用した。 V_{PG1} は 1.2V から 1.4V まで変化させることで 15%程度増加するが、ゲート絶縁膜の絶縁破壊電界には至らない。また、ゲートリークの増加も、この範囲では PS₁ のサブスレッショルドリークの減少よりもはるかに小さいためリーク電流が削減される (1.4V 以上の領域でリーク電流が微増しているのはゲートリークの増加のためだが、この影響は非常に小さい)。

通常の SRAM 動作時の M₁, M₂ のリーク電流は M₁, M₂ を通常の遮断よりも深く電源遮断することで削減を行った。ここでは SR 線のバイアス制御を試みた。SR 線の電位 (V_{SR}) を負バイアスすることでリーク電流を削減する。図 3.3.10(b)にリーク電流の V_{SR} 依存性を示す。赤線は NV セル、青線は 6T セルの通常電圧のスタンバイ時とスリープ時のリーク電流を示している。 $V_{SR} = -0.2V$ で削減量が飽和し 6T セルと同等のリーク電流を実現できるため、本論文では $V_{SR} = -0.2V$ を用いた。同様の削減効果は V_{CTRL} を若干増加させることでも得られる。こちらを用いることで、負バイアスを用いることなく余剰なリーク電流の削減も可能である。なお、本論文では第 2 章で述べたようにビット線を放電した状態でスリープモードの電流を評価しているが、通常動作時にビット線を V_{DD} に充電した状態で評価した場合、リーク電力削減率はさらに 10%程度小さくなる。

3.3.3 動作波形

図 3.3.11 に最適設計におけるセルの動作波形を示す。各信号は上からセルの PS₁ の制御信号 (V_{PG1})、SR 線と CTRL 線の制御信号 (V_{SR} , V_{CTRL})、MTJ の書き込み電流の絶対値 ($|I_{MTJ}|$)、記憶ノードの電位 (V_Q , $V_{\bar{Q}}$)、セルの消費電力を示している。 $(V_Q, V_{\bar{Q}})$ は初期時に (V_{DD} , 0) とし、1 回目の電源遮断/復帰後に通常電圧の SRAM の読み出し動作、書き込み動作を 1 回ずつ行う。書き込み動作により ($V_Q, V_{\bar{Q}}$) は (0, V_{DD}) となり、その後 2 回目の電源遮断/復帰を行う。MTJ₁/MTJ₂ の磁化の初期状態は、それぞれ平行状態/反平行状態とした。 $|I_{MTJ}|$ は H ストア/L ストア時の途中で、MTJ₁/MTJ₂ の抵抗変化のために電流値が急激に変化しており、正常に双安定回路のデータが MTJ₁/MTJ₂ へ書き込まれ

ている. H ストア時に H レベルの記憶ノードの電位が低下するが, 上述の設計で W_{PS1} と V_{SR} に適切な値を選んだことによりデータのフリップを回避できている. MTJ に書き込まれたデータに基づき, 設計したセルは $(V_Q, V_{\bar{Q}})=(V_{DD}, 0)$, $(0, V_{DD})$ のいずれの場合も電源遮断前の状態に正常に復帰している. 消費電力については, スタンバイ時 ($V_{DD}=V_{DD}$) に比べて電源遮断を行うことで大幅に電力が削減されている. また, スストア時の消費エネルギーが T_{BE}^{PE} (第 2 章参照) を決定するエネルギーオーバーヘッドとなる.

3.4 電力・エネルギー性能の評価

3.3 節ではセルの構成と設計について述べたが, 本節では最適設計したセルを用いた NV-SRAM セル・アレイの電力・エネルギー性能を示す.

まずは最適設計における NV セルのリーク電力を評価する. 図 3.4.1 に通常の 6T セルと NV セルの通常電圧におけるスタンバイ, スリープ, 電源遮断の各モードにおけるリーク電力を示す. NV セルは 3.3.2 節のリーク削減アーキテクチャを導入しない場合のリーク電力を白抜きで示している. 6T セルはスリープモードを用いて, スタンバイ時と比べて 41%程度 (第 2 章の評価回路を用いずに 1 セルで評価した場合も 44%程度であり, ほぼ同程度となる. また, この削減率は文献値との良い一致を示す[26]) の電力を削減できるが, NV セルでは電源遮断によって 99%以上と大幅に電力を削減できる. また, NV セルの通常動作時に余剰なリーク電流によるエネルギーオーバーヘッドは, 前述の通りバイアス制御により 6T セルと同等にまで削減される. これにより NV セルの 6T セルに対する通常電圧のスタンバイ時の余剰リーク電力は 1%以下にまで削減される.

次に, 第 2 章で示した PGP の詳細を以下に示す. NV-SRAM ではストア/リストア動作時に図 3.3.1 の SR 線 (M_1, M_2 のゲートに接続された信号線) を選択するためのデコーダ (以下 SR デコーダ), カラムデコーダ, CTRL 線を駆動するドライバ (以下 CTRL ドライバ) が必要となる. また, これらの制御用の信号には SR デコーダのイネーブル信号とリストア動作の制御信号が必要となる. SR デコーダはワード線の選択用のデコーダと同様の構成だが, ワード線とは異なる別のデコーダで構成する. スストア動作時は SR デコーダとカラムデコーダで 1 ワード (64bit) ごとに選択し, 1 ワードあたり 20ns で MTJ への書き込みを実行する (ここではワーストケースである下層配線の電流許容値を想定して 1 ワードずつ書き込みを行っている. 上層配線を考慮した場合の結果は Appendix に示す). ま

た, SR 線のドライバはリストア動作の制御信号を用いてリストア時に複数のドライバを同時に駆動できる構成にしてある. リストア動作時は2ワードごとにPSを導通して電源復帰する. このとき, パワースイッチを10nsかけて復帰し, M_1, M_2 をパワースイッチの制御開始から7ns経過した時点で遮断した. CTRL線はビット線と並行に設置してドライバもライトドライバを参考に設計した. ただし, 不揮発性メモリ素子への書き込み動作は比較的大きな電流が生じるので余裕を持たせて設計した.

続いて図3.4.1のリーク電力を用いてNV-SRAMアレイの T_{BE} を評価した. 図3.4.2に32kB, 256kB, 2MBのNV-SRAMアレイにおける T_{BE} の τ_{NL} 依存性を示す. 青の点線はMTJに流せる最大電流でストア動作を行った場合, 青の実線は最適化したストア電流を用いた場合の結果を示す. いずれも上述の V_{PG1} による電源遮断時のリーク削減技術を導入してある. 第2章で示したように T_{BE} は $T_{BE}^{EE} + \eta\tau_{NL}$ と書ける. 32kBではストア電流を設計することで T_{BE}^{EE} が50%程度削減でき, 大きな削減効果が見られるが256kB, 2MBと容量が増加するほどその削減効果は小さくなっている. これは第2章で述べたように大容量アレイではストア電流よりもストアを待機しているセルのリーク電流によるエネルギーが支配的になることによる. 図3.4.2の緑の点線は待機セルを通常電圧のスリープモードではなくスタンバイモードとした場合を示しており, 特に大容量ではスリープモードの導入により効果的に T_{BE}^{EE} が削減できることがわかる. しかし, 待機セルのリークにはまだ削減の余地があり, ラストレベルキャッシュ等の大容量アレイにおいてさらに T_{BE}^{EE} を削減するには次節で述べる階層型ストアフリー電源遮断アーキテクチャや無用データのフラッシュアーキテクチャによりスリープモードのセルを電源遮断することが有効である. 一方, η は通常のSRAM動作時の余剰リーク電力に依存する. したがって, 上述した V_{SR} 制御によってこのリーク電力を削減することで大幅に η を削減できる. アレイサイズに関わらず, V_{SR} を用いたバイアス制御により2桁程度削減される.

3.5 ストアエネルギーレイテンシ削減アーキテクチャ

次に, T_{BE}^{EE} のさらなる削減方法として, 階層型ストアフリー電源遮断アーキテクチャ(HSFアーキテクチャ)と無用データのフラッシュアーキテクチャ(PUDFアーキテクチャ)を提案し, その削減効果を検証する.

3.5.1 階層型ストアフリー電源遮断アーキテクチャ

T_{BE} のさらなる削減のためには、 T_{BE}^{EE} と η を削減する必要がある。 η は前節で述べた通常の SRAM 動作時のリーク電力の削減によって相当に小さくできる。一方、 T_{BE}^{EE} はストアに要するエネルギーを削減する必要がある、これにはストア動作をスキップするなどの方法が有効になる。また、ストア動作には、有限の時間 (ストアレイテンシ) が必要となるため、この削減も重要である。

MRAM ではこれまでに様々なストアエネルギーレイテンシの削減技術が提案されている。Zhou らは不揮発性メモリ素子への書き込み動作初期時に専用のコンパレータを用いて MTJ の抵抗状態を読み取り、書き込むデータが MTJ に保持されているデータと一致している場合は書き込み動作を中断する **Early write termination** を提案した[27]。また、MTJ のリテンション時間を犠牲にすることで書き込みのエネルギーを削減し、データの保持には DRAM のように定期的にリフレッシュを用いるアーキテクチャも提案されている[28-30]。そのほか、データの 0/1 の書き込みエネルギーが異なることを利用して、あるいはデータに存在する特定の 0/1 のパターンに着目して、データコーディングにより書き込みエネルギーの小さいデータに変換してから書き込みを行う方法[31,32]や、SRAM と MRAM のハイブリッドキャッシュを用いて、不揮発記憶を用いる必要のあるデータを MRAM に、不揮発記憶を用いる必要のないデータを SRAM へ記憶することで MRAM の書き込み回数を削減する方法もある[33-36]。しかし、これらの技術ではストアエネルギーレイテンシの削減は可能であるが、これらに要するエネルギーとレイテンシのオーバーヘッドは必ずしも小さくないことや、複雑なアーキテクチャを必要になるなどのデメリットが小さくない。

本論文では階層型ストアフリー電源遮断アーキテクチャ (HSF アーキテクチャ) を提案した。本アーキテクチャでは、セルアレイを構成するサブアレイをブロックに分割して、6T セル部と MTJ 部のデータが一致しているブロックに対して、通常の SRAM 動作時に書き込みの有無でストア動作の実行を判定する。このアーキテクチャでは 6T セルと MTJ とのデータの照合や、データのプロセッシングを必要としないため、これに要するエネルギーとレイテンシのオーバーヘッドを最小限に抑えることができる。通常電圧の SRAM 動作時に書き込みアクセスが生じたブロックについて、ストアフリーフラグ (**Store-free flag; SFF**) を専用のラッチにセットする。ラッチはサブアレイ内に設けることもできる[37]。また、各ブロックの 6T セル部のデータの有効/無効を有効データフラグ (**Valid data flag; VDF**)

で判定することもできる。この場合、データ無効であれば、6T セル部と MTJ のデータが不一致であってもデータのストア動作をスキップすることも可能である(ただし、キャッシュの場合、ブロック内のタグの有効ビットをすべて無効にして MTJ にストアするなど工夫が必要になる。また、下層の記憶回路とのコヒーレンシも保てるようにする)。VDF は実装しなくてもよい。本論文では Power management unit (PMU)内に専用のレジスタファイルとして配置する。SFF により一度でも書き込みアクセスがあったブロックを指定できるが、このブロック内では6Tセル部には通常のSRAMとしての書き込みが行われるため、ほとんどの場合でそのデータはMTJのデータと一致しない。電源遮断時にVDFとSFFを参照することで、ストア動作の必要なブロックのみストア動作を実行できる。すなわち、通常のSRAM動作時の書き込みアクセス履歴からストア動作をスキップできるブロックやこれらからなるサブアレイ(以後、ストアフリーブロック、ストアフリーサブアレイと呼ぶことにする)を判定できる。ストアフリーブロック、ストアフリーサブアレイは通常のSRAM動作における書き込みアクセスのアドレスから容易に特定できるため、SFFを生成するハードウェアは容易に実現できる。

図 3.5.1 に HSF アーキテクチャの概念図を示す。簡単のため VDF はすべてのブロックで有効であったとする。HSF アーキテクチャでは、ストアの開始直前に SFF に基づき、まずストアフリーサブアレイを一括して電源遮断し、その後ストアフリーブロックを電源遮断する。最後に残りのブロックを順次スキャンしながらストアを行う。これにより双安定回路と MTJ のデータが一致していない可能性のあるブロックのみストア動作を実施すればよいので、ストアエネルギー、待機セルのリーク電流によるエネルギー、ストアレイテンシをすべて削減することができる。この図はストアフリーサブアレイとストアフリーブロックを分けて制御したものであるが、これらを分けずにすべてストアフリーブロックとして制御してもよい。以下ではこの後者の方法を採用する。

図 3.5.2 に HSF アーキテクチャの実装方法を示す。複数のサブアレイから構成されるメモリアレイと SFF 用のレジスタファイルを内包し、メモリアレイのパワースイッチを制御する PMU から構成される。図にはアドレスやデータを制御する Memory management unit (MMU)も示してある。電源遮断時は、SFF に基づきパワースイッチを制御する。

図 3.5.3 に 32kB, 256kB, 2MB の NV-SRAM アレイにおける T_{BE}^{EE} とストアレイテンシのストアフリー率(全アレイに対するストアフリーフラグで指定されたブロックの割合)依存性を示す。ストアレイテンシはアレイ全体のストア動作に要するレイテンシであり、1 ワー

ド(64bit)のストア動作に要する時間を 20ns としてストアする必要のあるワードをすべて順次書き込みを行った時のレイテンシである。ストアフリー率に応じて待機セル数を減らすことができるため、ストアを待機しているセルの待機時電力を削減できる。したがって、大容量のアレイサイズに対しても効果的に T_{BE}^{EE} とストアレイテンシを削減できる。

3.5.2 積極的無用データフラッシュアーキテクチャ

続いて、 T_{BE}^{EE} の削減のためのもう一つのアーキテクチャについて説明する。ここで提案するアーキテクチャは NV-SRAM と VNR-SRAM の両方に適用できる。NV-SRAM では前節で述べた HSF アーキテクチャと組み合わせることも可能である。以下では、はじめにこのアーキテクチャの NV-SRAM への適用について述べ、次いで、第 4 章で VNR-SRAM への適用について述べる。

これまでの NV-SRAM を用いた PG の議論では、NV-SRAM に書き込まれたデータのすべてを保持することを前提に議論してきた。しかし、NV-SRAM をロジックシステムのキャッシュに応用した場合には、PG 後に不要なデータも存在すると予想される。このような無用なデータをいちいち不揮発記憶するのは不必要なストアエネルギーレイテンシを生じる。したがって、PG 後にも必要となるデータのみを不揮発記憶によって保持し、これ以外は破棄することで、不揮発記憶にともなうストアエネルギーレイテンシは削減できると予想される(ただし、書き換えのあったデータはライトスルーまたはライトバックによって、電源遮断より前に下位の記憶回路に書き戻されているとする)。このようなキャッシュに記憶されている無用な、または重要度の低いデータを本論文では無用データ (Useless data) と呼ぶことにする。本節ではこの無用データを判定し積極的に破棄することで NV-SRAM のエネルギー性能をさらに向上できるアーキテクチャを提案する。これを積極的無用データフラッシュ (Proactive useless-data flush; PUDF) アーキテクチャと呼ぶ。

図 3.5.4 に NV-SRAM を用いた場合の PUDF アーキテクチャの実装方法を示す。PUDF アーキテクチャは前節の HSF アーキテクチャと同様に専用のフラグを保持するレジスタファイルを配置することで実現できる。このレジスタには無用データが記憶されていると判定されたブロックを指定する無用データフラグ (Useless data flag; UDF) を記憶する。この UDF は従来のキャッシュアーキテクチャと類似の方法で生成できる。一般にキャッシュでは、データの書き込み時に書き込み可能な領域に空き領域がない場合、過

去に記憶されたデータを破棄して書き込み領域を確保する(これはキャッシュラインの置き換えや追い出しと呼ばれる)。置き換えデータの判定には, first-in first-out (FIFO) と呼ばれる書き込み時期の古いデータを選択するアーキテクチャ, Least recently used (LRU)と呼ばれる使用頻度の少ないデータを選択するアーキテクチャ, ランダムと呼ばれる無作為に選ぶアーキテクチャなどを用いることができる[38]。これらのアーキテクチャはキャッシュのヒット率に大きく影響するため極めて重要であるが, 一般に LRU, ランダム, FIFO の順にヒット率が低下する。また, ヒット率向上のためこれらを組み合わせた方法も提案されている。このようにキャッシュではヒット率を可能な限り下げない置き換えデータの判定機構が備わっている。本論文で扱う無用データも従来のキャッシュアーキテクチャを用いて判定する。したがって, UDF は既存技術で容易に生成できる。

NV-SRAM の PUDF アーキテクチャでは電源遮断動作を行う直前に, まず UDF に基づいて無用データを記憶しているブロックを一括して電源遮断する。このとき, UDF で指定されたブロックはストア動作を行わない。すなわち, UDF で指定されるブロックのデータを破棄する。したがって, スタア動作にともなうエネルギーとレイテンシのオーバーヘッドを削減できる。残りのブロックについては順次ストア動作と電源遮断を行い, すべてのブロックの電源遮断が完了したところで, 電源遮断への移行を終了する。

NV-SRAM の場合, PUDF アーキテクチャは上述の HSF アーキテクチャと併用することができる。PMU には UDF と SFF を保持するためのレジスタファイルをそれぞれ配置し, スタア動作開始前に UDF に基づき指定されたブロックを SFF の内容に関係なく一括して遮断し, 次に, 残りのブロックの内 SFF で指定されたブロックを一括して遮断する。そして, 最後に残ったブロックを順次ストアする。

HSF アーキテクチャあるいは PUDF アーキテクチャによってストア動作を省略して遮断されるブロックの割合 R_{skip} は, PUDF と HSF を併用する場合, 以下のようになる。

$$R_{\text{skip}} = R_{\text{PUDF}} + r_{\text{HSF}} \cdot (1 - R_{\text{PUDF}}) \quad (3.1)$$

R_{skip} は第 2 章の式(2.25)で定義した ENT 動作を省略するブロックの割合である。 R_{PUDF} は NV-SRAM のメモリアレイに対する UDF によって指定されたブロックの割合, r_{HSF} は UDF で指定されたブロックを除いた残りのメモリアレイに対する SFF によって指定されたブロックの割合である。第 4 章で VNR-SRAM に適用する場合は $r_{\text{HSF}}=0$ とすることで, 式(3.1)で同様に記述できる。

続いて, PUDF アーキテクチャを NV-SRAM に適用した場合の $T_{\text{BE}}^{\text{PE}}$ とストアレイテン

シ(τ_{STR})の削減効果を解析する. 図 3.5.5(a)-(f)に NV-SRAM における T_{BE}^{EE} および τ_{STR} の R_{PUDF} 依存性をそれぞれ示す. これらの図の赤と緑の線は $r_{HSF}=0$ の場合である. NV-SRAM ではストア動作時に UDF で指定されたブロックについてはストア動作をスキップするため, スキップするブロックの分だけストアエネルギーとストアレイテンシを削減できる. したがって, R_{PUDF} の増加にともない, いずれも効果的に削減される. これらの図中の青の線は PUDF と HSF を併用した場合の T_{BE}^{EE} および τ_{STR} である. r_{HSF} は 0%から 100%まで 25%ずつ変化させてある. PUDF と HSF を併用することで NV-SRAM の T_{BE}^{EE} と τ_{STR} をさらに効果的に削減できる. なお, 電源遮断時のリーク電力は R_{PUDF} に依存せず, 99%以上削減される. これは NV-SRAM では最終的にすべてのブロックが電源遮断されるためである.

3.6 試作 TEG による検証

最後に試作した TEG による NV-SRAM の電力・エネルギー性能の検証について述べる. ルネサスエレクトロニクス社の 65nm CMOS プロセス[39]を用いて試作した Test element group (TEG)から抽出した回路性能パラメータを用いてエネルギー性能を評価した. 本論文ではエネルギー性能評価用の TEGを用いた. NV-SRAM について不揮発記憶を含めて動作を検証するためには図 3.6.1(a)のように MTJ を実装する必要がある. しかし, MTJ の CMOS へのエンベディッド技術は近年開発が進んでいるが, 我々の利用可能な汎用のプロセスとしてはまだ提供はない. そこで, エネルギー性能に限って評価を行うことにし, MTJ を省略した図 3.6.1(b)のセル構成の TEG を用いた. この TEG では MTJ は存在しないが, ストア電流を正しく再現できるように設計しており, ストアエネルギーを正しく評価できる(エネルギーの評価には実際に不揮発記憶されたかどうかは関係がないため, エネルギー解析では本構成の TEG で十分である[40, 41]).

図 3.6.2 に試作した NV-SRAM TEG のセルとアレイのレイアウトを示す. 1kb (32bit×32bit)のアレイを試作し, 各種周辺回路を配置した. この TEG はセル, パワースイッチ, ドライバを比較的余裕をもってトランジスタサイズを設計してある. 以下に述べる解析では 3.3 節で述べた最適設計のセルを用いるため, この TEG から抽出した回路性能パラメータの実測値と, TEG と同一構成の回路を用いた HSPICE による計算値を比較して補正值 (Correction factor) をもとめ, この補正值を用いて最適設計での実測値を推

定した. 図 3.6.3 に抽出および補正を行った回路性能パラメータを示す.

図 3.6.4 に実測による回路パラメータを用いて算出した T_{BE} と, HSPICE のシミュレーション結果から求めた T_{BE} の τ_{NL} 依存性を示す. SFF や UDF により指定されたブロックの割合はアプリケーションによるが, ここでは前節の R_{skip} を 50% と仮定した. 実測値とシミュレーションによる結果は概ね一致した. 若干のずれは PDK が開発途上であるため生じたものと考えられる.

図 3.6.5(a) に各アレイサイズにおける T_{BE}^{EE} と η の τ_{NL} 依存性を示す. T_{BE}^{EE} , η は実測値とシミュレーションでよく一致している. 従来のロジックシステムにおける PG の T_{BE} が数 ms 程度[42]であることを考えると, この T_{BE}^{EE} は, これらに比べて 1 桁程度小さい. システムの通常動作時間 (τ_{NL}) が 10ms~100ms 程度以下である場合, 従来よりも大幅に PG の実行頻度を増加させることができる(細粒度 PG を実現できる). また, 図 3.6.4 に示したように τ_{NL} が 100ms 程度以上の領域では T_{BE} は τ_{NL} に比例して増加するが, T_{BE} の値は τ_{NL} よりもおおよそ 2 桁程度短い. したがって, この場合でも従来よりも有効に電源遮断が行えると考えられる. 同図(b)に 6T-SRAM の待機時電力と NV-SRAM の電源遮断時の電力を示す. 電力の削減率は実測値とシミュレーションでほぼ一致した. NV-SRAM は 6T-SRAM と比較して 2 桁以上の電力削減が可能である.

3.7 まとめ

本章では本研究室提案の NV-SRAM に関する各種検討を行った. 以下にその内容をまとめる.

- ・著者の所属研究室から提案された NV-SRAM について, ばらつきの影響を考慮したストア電流と SNM の評価によるセルの設計方法を確立した.
- ・最適設計したセルのリーク電力を評価し, NV-SRAM が 6T-SRAM に比べてはるかに大きい 2 桁以上の電力削減効率を実現できることを示した.
- ・ストア電流の最適化, 各種リーク削減アーキテクチャの導入により NV-SRAM アレイにおける T_{BE}^{EE} を 50% 程度, η を 2 桁程度削減した.
- ・さらなるエネルギー性能の向上に向けて階層型ストアフリー電源遮断アーキテクチャ (HSF アーキテクチャ) と積極的無用データフラッシュアーキテクチャ (PUDF アーキテクチャ) を提案し, これを用いることで T_{BE}^{EE} を効果的に削減できることを示した.

・試作した TEG から抽出した回路パラメータを用いて解析を行い, 解析結果が HSPICE のシミュレーション結果と概ね一致することを示した.

以上から不揮発記憶を用いた PG の実現に向けた NV-SRAM 技術を確立した.

3.8 Appendix

(i) MTJ の書き込み特性が NV-SRAM の T_{BE} に与える影響

NV-SRAM の T_{BE}^{EE} は MTJ の臨界電流密度 (J_c) とストアレイテンシ (τ_{STR}) に大きく依存するため, これらのパラメータ (すなわち MTJ の書き込み特性) が NV-SRAM の T_{BE} に影響を与える. 図 3.8.1(a), (b) はそれぞれ 32kB と 2MB の NV-SRAM アレイの T_{BE} である. 青の実線が 3.3 節で最適設計したセルで構成した NV-SRAM アレイの T_{BE} , 青の点線は MTJ の J_c が 1 桁小さいと仮定した場合の T_{BE} , 赤の点線はさらに τ_{STR} を 1 桁下げた場合の T_{BE} を示す. MTJ の書き込み特性について, J_c と τ_{STR} では T_{BE}^{EE} に与える効果が異なる. NV-SRAM の容量が小さいときは J_c を下げることで T_{BE}^{EE} を削減できるが, 容量が大きくなると, T_{BE}^{EE} は J_c に依存しなくなり, τ_{STR} を短くすることで T_{BE}^{EE} を削減できる. これは第 2 章で説明したように容量が大きくなると, 電源遮断移行動作のエネルギーより待ちサブアレイのリークエネルギーの方が支配的になることによる. 以上から, アレイサイズに依らず MTJ の書き込み特性の改善により NV-SRAM の T_{BE}^{EE} を削減するためには, MTJ の J_c を現状から低減するとともに, 同時に τ_{STR} も削減する必要がある (ただし, この両立は相当に困難である). また, これらを 1 桁程度削減できたとしても次章で示す VNR-SRAM ほどには T_{BE}^{EE} を下げることはできない.

(ii) 電流許容値を緩和した場合の T_{BE}

第 3 章では下層配線の電流許容値を仮定して議論したが, 仮に上層配線の電流許容値で NV-SRAM を評価した場合, ストア時に 1 度に 16 ワード分の MTJ への書き込みが, リストア時に 32 ワード分の復帰が可能である. この場合の T_{BE}^{EE} の評価結果を図 3.8.2 に示す. $R_{skip}=50\%$ としている. 複数ワードに同時に MTJ への書き込みを行うことで書き込みレイテンシを大幅に削減できるため, 待機セルのリーク成分の削減が可能となり, 主に大容量のアレイで T_{BE}^{EE} を削減可能である. なお, この場合リストア時のワード数はストア時のワード数に合わせて 16 ワードとする方が実装する上で都合がよい. この時,

リストアの要するレイテンシは若干増加するが、リストア時の消費エネルギーはストア時に比べて非常に小さいのでエネルギーはほぼ増加しない。

(iii) 通常動作時に MTJ を用いた書き込みを行う場合のエネルギー性能の劣化

通常の SRAM 書き込み動作時に常に MTJ に書き込みを行った場合に生じるエネルギー性能の劣化について考察する。本研究室提案の NV セル(図 3.2.2 参照)は上述のように M_1 , M_2 を遮断することで双安定回路と MTJ を電気的に分離することができ、従来の 6T セルと同様の書き込み動作(GHz 動作)を実現できるが、通常の SRAM 書き込み動作時にストア動作も実施し、 M_1 , M_2 を導通して MTJ に書き込みを行うこともできる。このため、通常の読み出し動作時は 6T セルと同様に行い、書き込み動作時のみ MTJ に常に書き込むアーキテクチャで評価した。また、比較のためにこれまでに示された各種不揮発性 SRAM セル[3-5](図 3.2.1 参照)についても、常に MTJ に書き込みを行うアーキテクチャで評価した。これらのセルは NOF アーキテクチャ(スタンバイ時に電源オフ)を前提に提案されたものだが、この場合 GHz 動作は困難であるため、ここではスタンバイ時も電源供給を行うこととした。

続いてセルの設計について述べる。以後、図 3.2.1(a)-(c)のセルはそれぞれ A セル、B セル、C セルと呼ぶことにする。NV セルは 3.3 節と同じ設計とするが、その他のセルは MTJ が双安定回路の内部に挿入されていることにより通常の SRAM 読み出し動作の SNM を劣化させるため、これを十分に確保するようトランジスタサイズと MTJ の抵抗を設計した。図 3.8.3 は各種セルの読み出し動作時の SNM を MTJ の抵抗値 R_p に対してプロットしたものである(ここでは簡単のため TT コーナーのみ検討している)。A セルと B セルは NV セルと同じトランジスタサイズとし、MTJ は抵抗の増大に伴って SNM が劣化するため、それぞれ 78mV を満たす抵抗値である $R_p=4.0k\Omega$, $2.1k\Omega$ とした。C セルはドライバサイズを 150nm とするとプルダウン能力が足りず SNM を十分に確保することができないため、300nm とし、MTJ は $R_p=8.0k\Omega$ とした。NV セルは MTJ の抵抗値に寄らず、6T セルと同じ SNM を保つことができる。なお、NV セルについては SNM が 78mV よりも大きい値となるため、ドライバトランジスタのチャネル幅を小さくすることで他のセルと同じ値に合わせることもできるが、この場合他のセルと比べて通常動作のスタンバイ時におけるリークエネルギーがさらに削減される。

図 3.8.4 に各セルの動作波形を示す。MTJ の書き込みシーケンスを以下に示す。A セル

ルと C セルは文献[3,5]を参考にして MTJ への書き込みを行った。A セルはまずビット線を 0V とした状態でパストランジスタを導通し, Q, /Q のどちらの MTJ にも電流を流して反平行状態にする(フリー層は記憶ノード側に接続する)。その後, 書き込むデータとは逆のデータを双安定回路に書き込んでからビット線を H レベルにした状態でパストランジスタを導通し, L レベル側の MTJ を平行状態にする。MTJ への書き込みがすべて終わった後, 正しいデータを双安定回路に書き込む。C セルは双安定回路にデータを書き込むと V_{DD} 線からビット線に電流が流れることで L レベル側の MTJ が反平行状態になる(ピン層を V_{DD} 側に接続)。その後, パストランジスタを導通した状態で V_{DD} を 0V に落とすことで今度はビット線から V_{DD} 線に電流を流して H レベル側の MTJ を平行状態とする。一方, B セルはもともと磁界を印可することで磁荷反転を行う MTJ を前提に提案されたものであるので, STT-MTJ にも適用可能な書き込みシーケンスを用いた。まず, 双安定回路に逆データを書き込んでからビット線をいずれも V_{DD} に充電した状態でパストランジスタを導通し, L レベル側の MTJ に書き込みを行って反平行状態とする(ピン層を記憶ノード側に接続)。その後, 書き込むべきデータを双安定回路に書き込み, L レベル側のビット線に負バイアス(-0.3V)を印可して, V_{SS} からビット線に電流を流して MTJ を平行状態とする。以上の A, B, C セルの書き込みシーケンスはいずれも逆データの書き込み等のために, 複数回の通常の SRAM の書き込み動作が必要であるが, このエネルギーについては MTJ の書き込みエネルギーに比べて非常に小さいため省略した。また, 後述のバイアス制御時には A セル, B セルはいずれもワード線を 0.65V, C セルは MTJ の書き込み時のみ V_{DD} を 0.2V, ワード線を 0.8V として MTJ への書き込み電流を削減している。ここではセルにプリチャージ回路とライトドライバを接続し, 1 セルの電力を評価した。各セルは MTJ の書き込み時にワード線と電源線の電圧を適宜調整して, MTJ の書き込みエネルギーを可能な限り削減している。通常電圧のスタンバイ時には, 6T セル, NV セル, A セル, B セルについてはほぼ同等のリーク電力となるが, C セル nMOS が導通状態となるために大きな電流が流れる。

図 3.8.5 に各種セルの書き込み動作のエネルギーを示す。MTJ を用いない場合に比べて, MTJ を用いる場合はいずれもエネルギーが 2 桁以上増大する。なお, いずれのセルの場合も MTJ の書き込みエネルギーはバイアス制御を用いることで図 3.8.5 の矢印に示したように削減できるが, MTJ の書き込みを実現する最低限のエネルギー以下には低減できない(この書き込みエネルギーは通常の 6T セルの書き込み動作のエネルギー

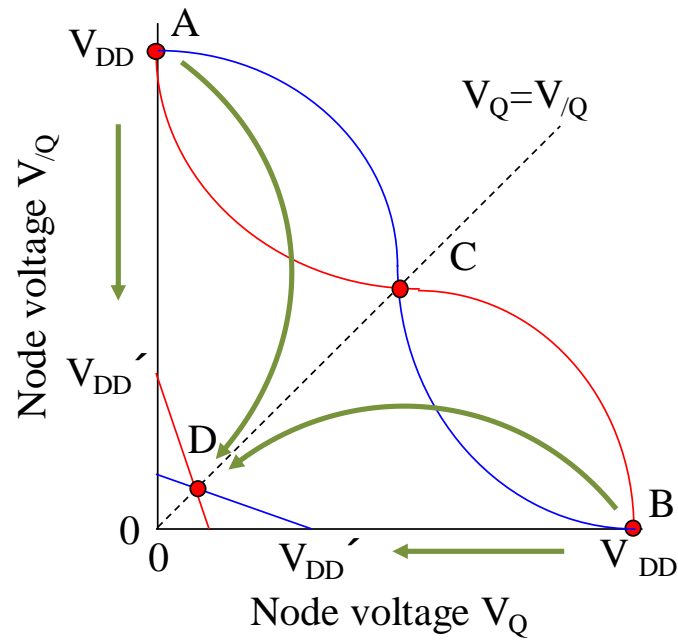
—100 倍程度以上の値である).

続いて, 設計した各種セルのエネルギー性能の比較を行う. 表 3.8.1 に 1 セルの書き込み動作時のエネルギー, スタンバイ時のリーク電力, 読み出し動作時のエネルギーをまとめた. これらの値を用いて 1 サイクルあたりの平均エネルギー ($E_{average}$) は次式のように表すことができる.

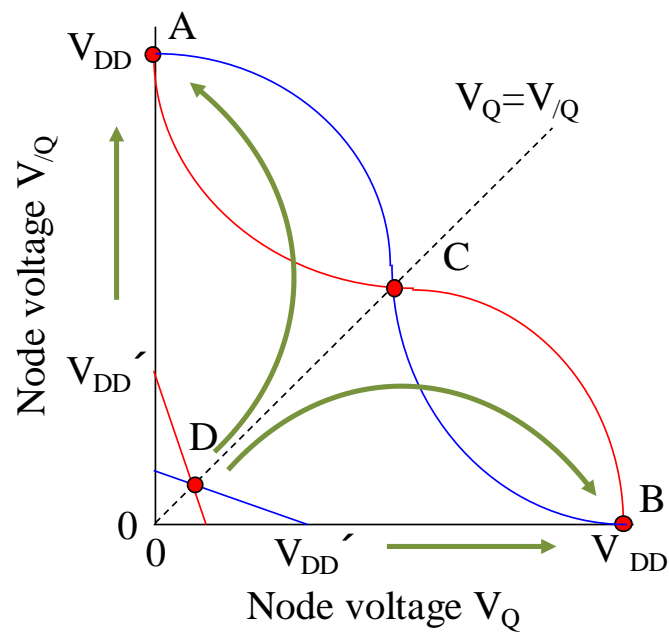
$$E_{average} = \frac{E_{WRITE} \cdot n_{WRITE} + P_{SB} \cdot \tau_{CYC} \cdot n_{SB} + E_{READ} \cdot n_{READ}}{n_{WRITE} + n_{SB} + n_{READ}} \quad (3.2)$$

E_{WRITE} , P_{SB} , τ_{CYC} , E_{READ} はそれぞれ, 各セル構成の書き込み動作のエネルギー, スタンバイ時のリーク電力と 1 サイクルの時間, 読み出し動作のエネルギーを, n_{WRITE} , n_{SB} , n_{READ} は各命令の出現回数を示す. 通常動作時の CPU の命令の出現頻度については文献値を参考にし[43], ストア命令(書き込み動作)を全命令の 10%, ロード命令(読み出し動作)を 27.5%, その他のスタンバイ状態を 62.5%とした. 全サイクル数を 40 サイクルとし, この比率を満たすよう各命令の出現回数はストア命令が 4 サイクル, ロード命令を 11 サイクル, スタンバイ状態を 25 サイクルとした. 1 命令あたりの実行時間 (τ_{CYC}) は 1GHz 動作を想定し, 1ns とした (MTJ の書き込みには H レベル, L レベルで 20ns 必要となるため, クロックオーバーヘッドが生じる). 以上の各動作時の電力・エネルギー, 命令の頻出比率から算出した $E_{average}$ を図 3.8.6 に示す. 同図に示すように, 通常動作時に MTJ に書き込む場合, MTJ を用いずに書き込み動作を行う場合と比べて, 各種バイアス制御を用いたとしても 1 桁以上大きなエネルギーを消費する. 以上から MTJ に常に書き込みを行うような場合にはエネルギー性能が大幅に劣化する. なお, 本論文の対象からは外れるが, GHz 以上の高速動作を必要としないアプリケーションであれば, スタンバイ状態で電源を遮断する NOF アーキテクチャも実現可能であり, 有用となる可能性がある.

3.9 図と表

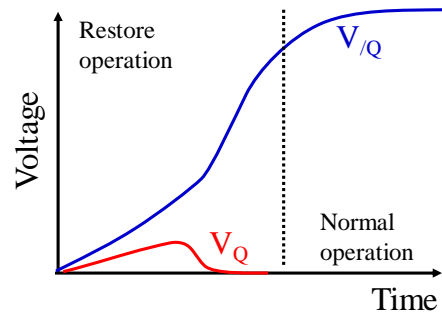
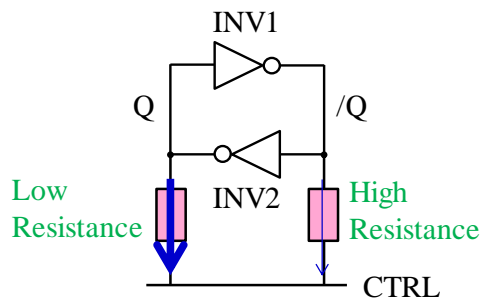


(a)

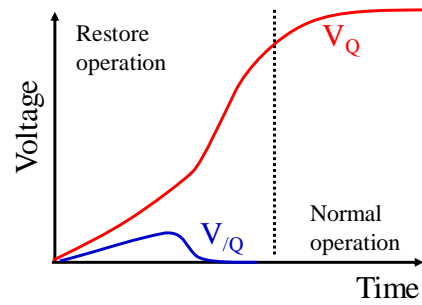
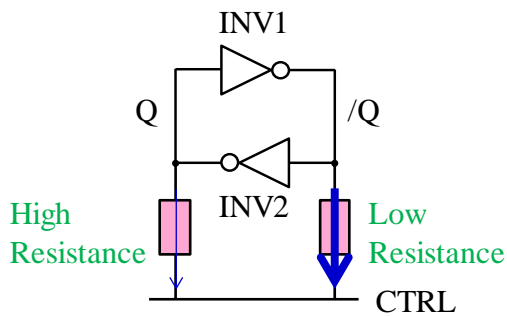


(b)

図 3.1.1 (a)電源遮断時と(b)電源復帰時のバタフライカーブの変化と安定点 A, B の推移

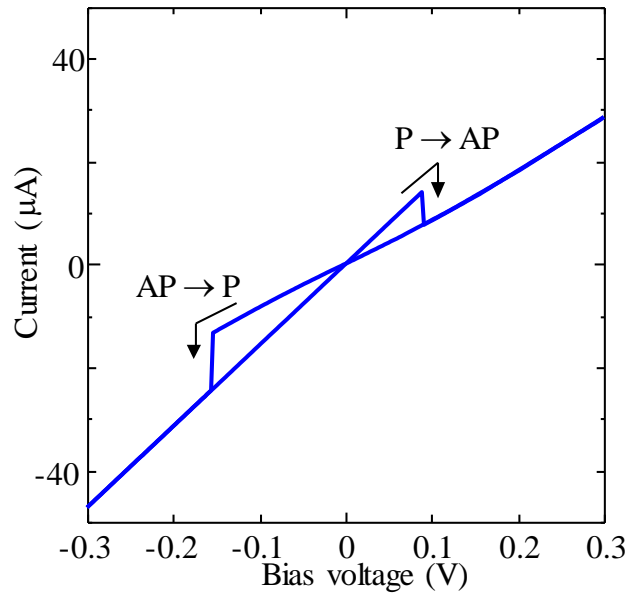


(a)

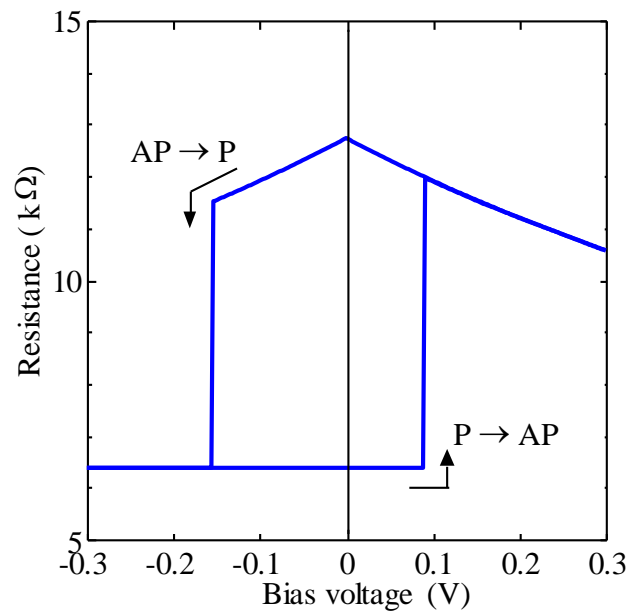


(b)

図 3.1.2 (a)Q 側の MTJ が低抵抗かつ/Q 側の MTJ が高抵抗の場合の電源復帰動作と
(b) Q 側の MTJ が高抵抗かつ/Q 側の MTJ が低抵抗の場合の電源復帰動作



(a)



(b)

図 3.1.4 本論文で用いる MTJ の(a)書き込み電流と(b)抵抗値のバイアス依存性

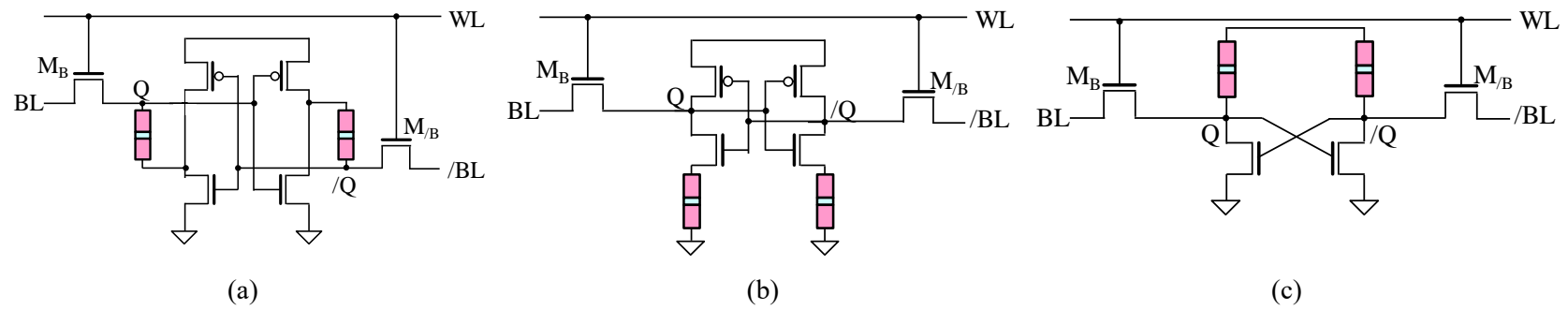
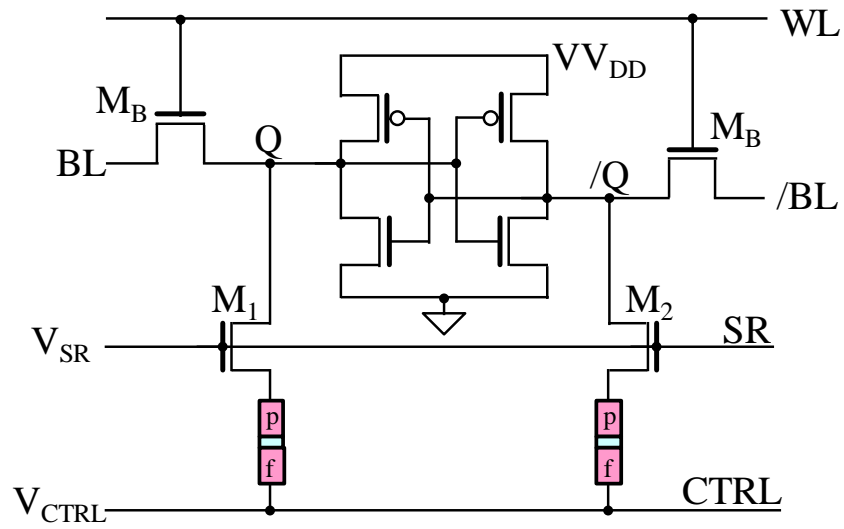
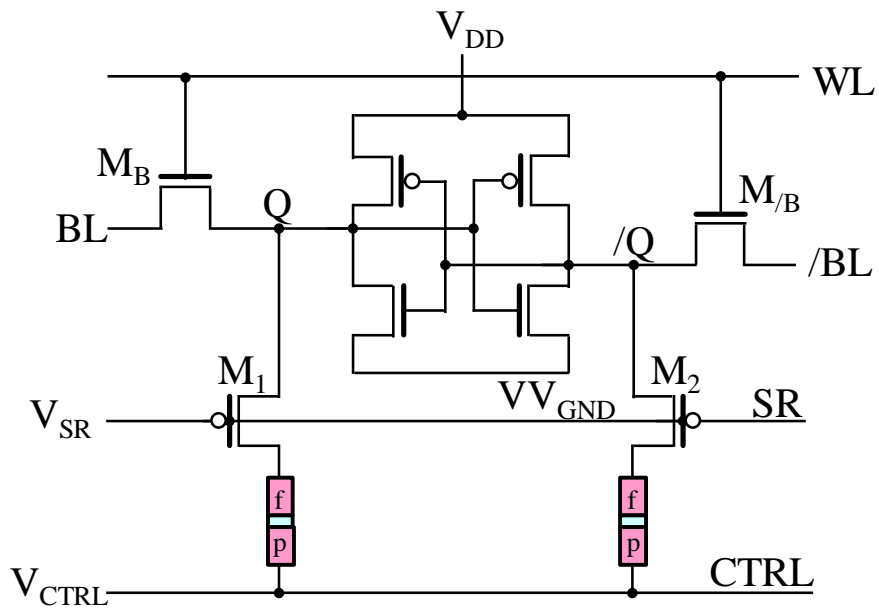


図 3.2.1 これまでに提案されている不揮発性 SRAM のセル構成



(a)



(b)

図 3.2.2 (a)仮想電源電圧アーキテクチャと(b)仮想接地アーキテクチャにおける本論文で検討する NV セル構成

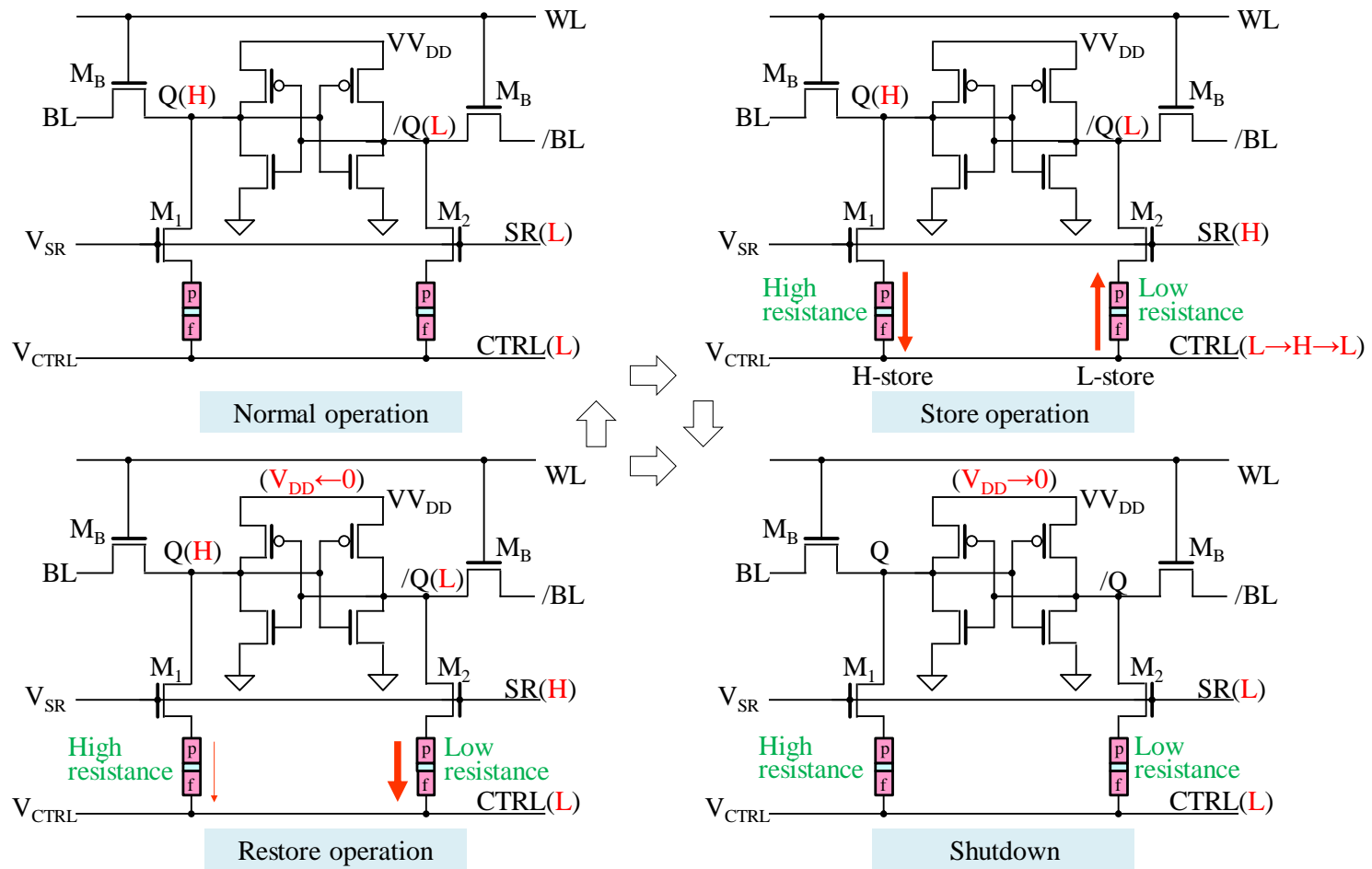


図 3.2.3 NV セルにおけるストア動作, 電源遮断, リストア動作

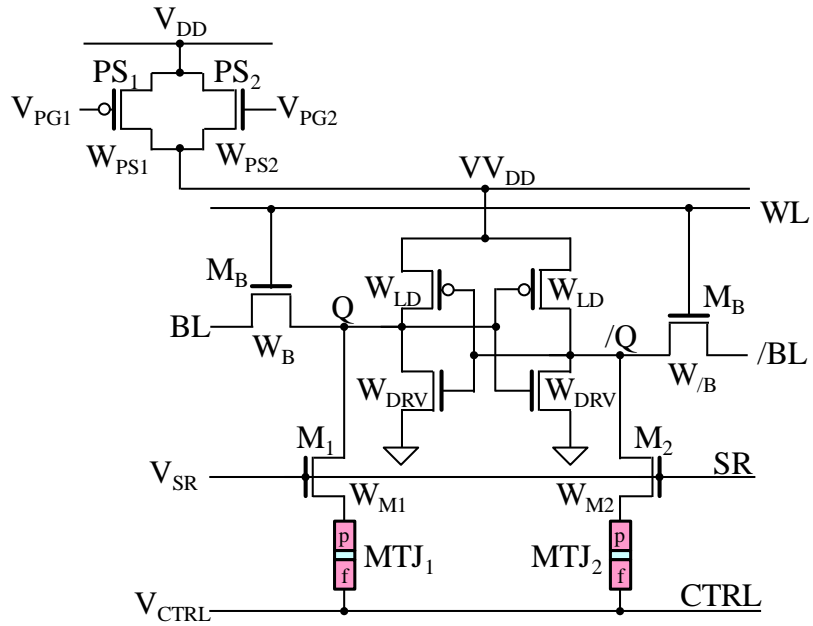


図 3.3.1 NV セルとパワースイッチの構成

表 3.3.1 NV セルとパワースイッチのデバイスパラメータとバイアス

Device parameters and gate biases	
Cell(Channel width)	Power switch (Channel width, gate bias)
LD transistor:	W_{LD} PS ₁ : (W_{PG1} , V_{PG1})
DRV transistor:	W_{DRV} PS ₂ : (W_{PG2} , V_{PG2})
PSS transistor (M_B):	W_B
($M_{/B}$):	$W_{/B}$
pass transistor (M_1):	W_{M1}
(M_2):	W_{M2}

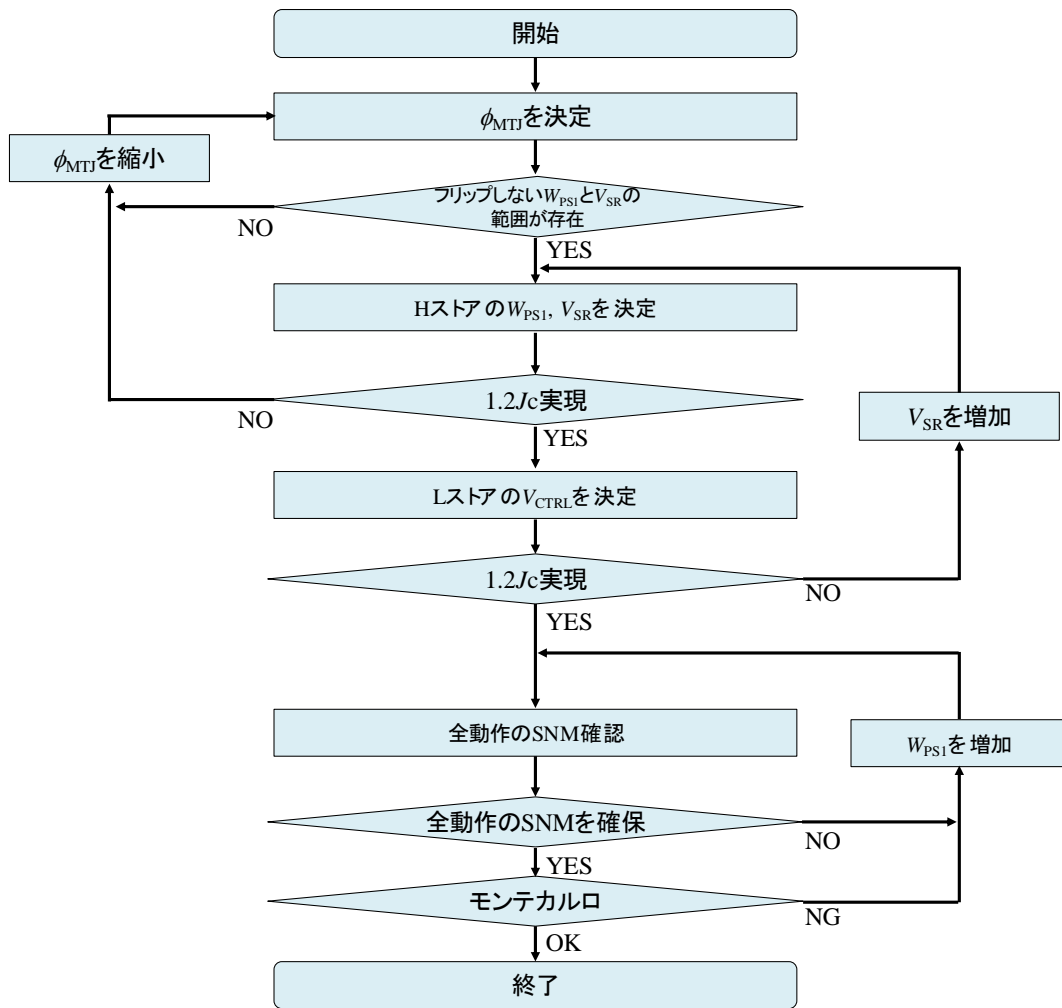


図 3.3.2 NV セルの設計フローチャート

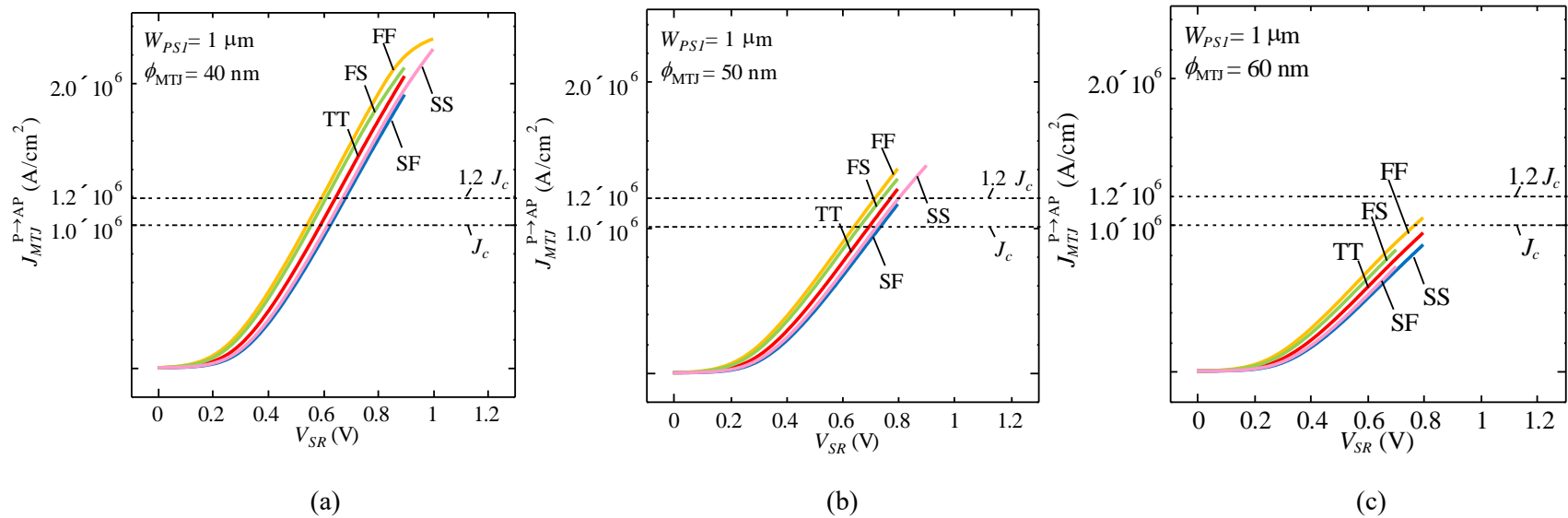


図 3.3.3 MTJ の直径 (ϕ_{MTJ}) が (a)40nm, (b)50nm, (c)60nm の場合の H ストア時のストア電流密度の V_{SR} 依存性

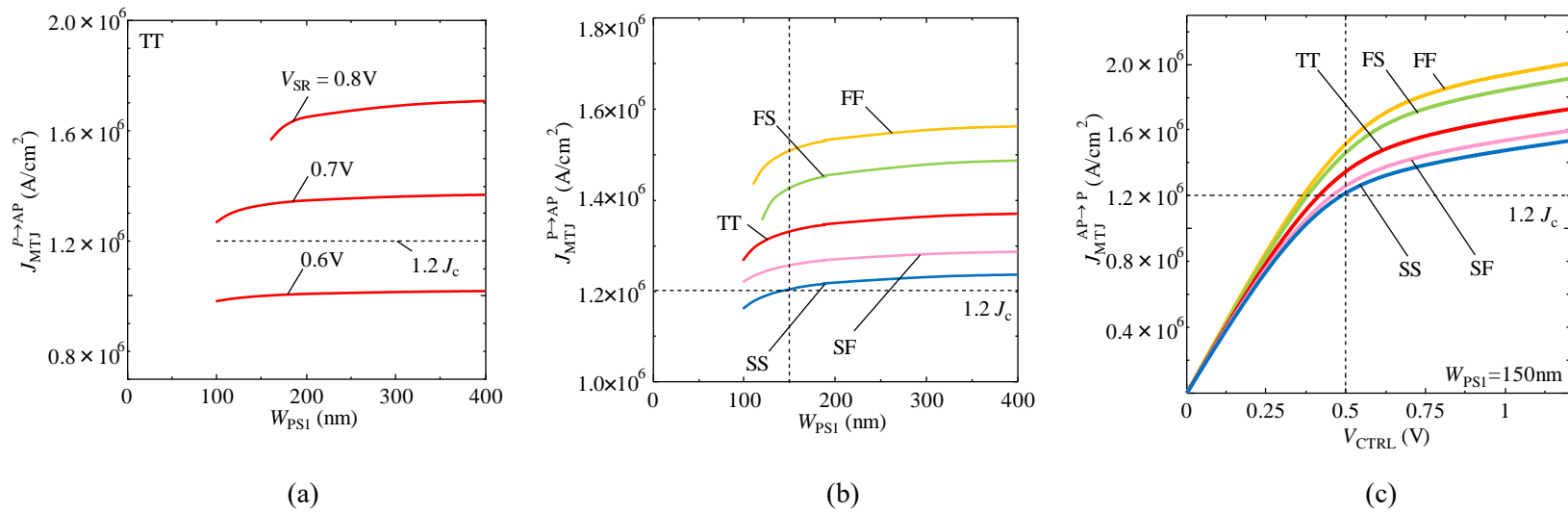


図 3.3.4 (a)H ストア時のストア電流密度の W_{PS1} 依存性, (b) $V_{SR}=0.7V$ での H ストア時のストア電流密度の W_{PS1} 依存性, (c)L ストア時のストア電流密度の V_{CTRL} 依存性

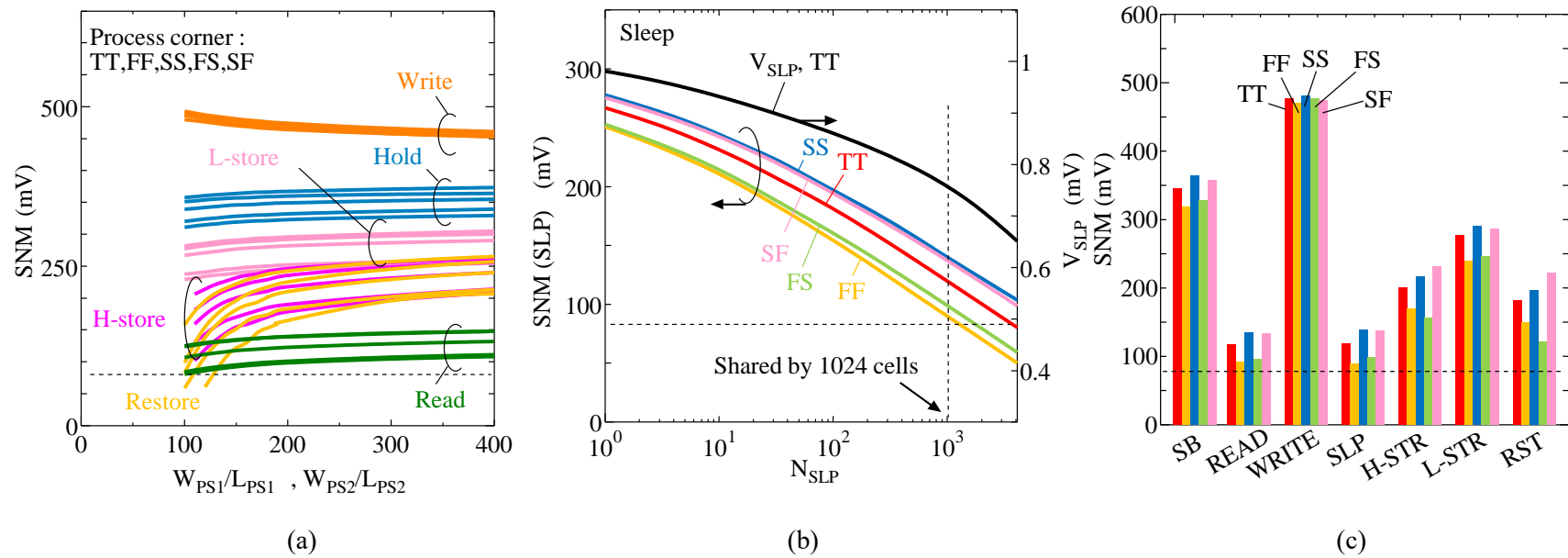
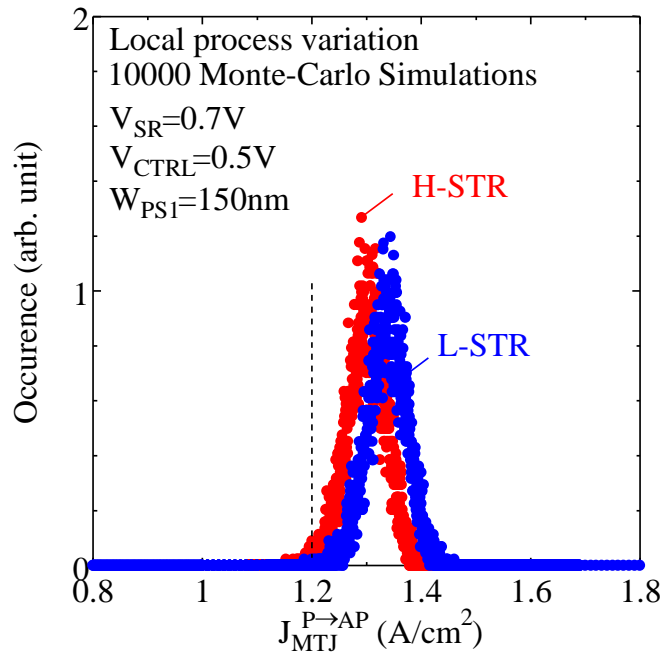
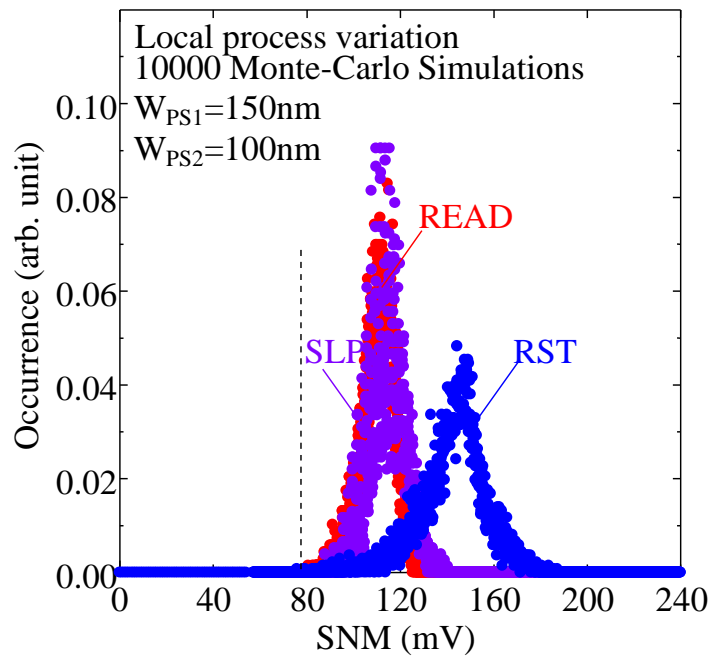


図 3.3.5 (a)各動作モードにおける SNM の W_{PS1} 依存性, (b)スリープ時の SNM の N_{SLP} 依存性, (c)全動作モードにおける SNM のまとめ



(a)

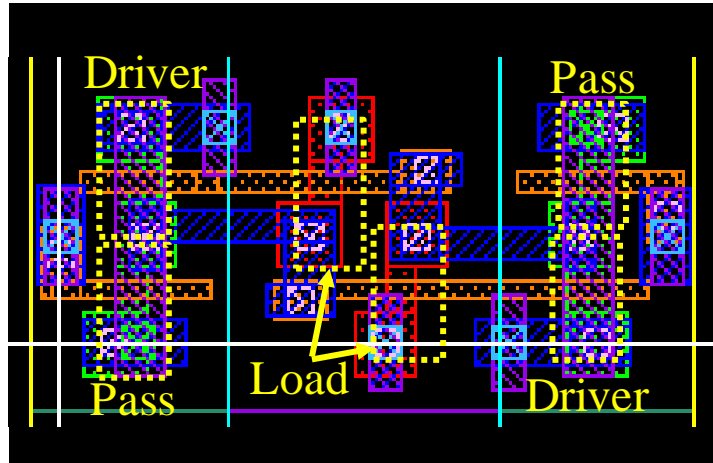


(b)

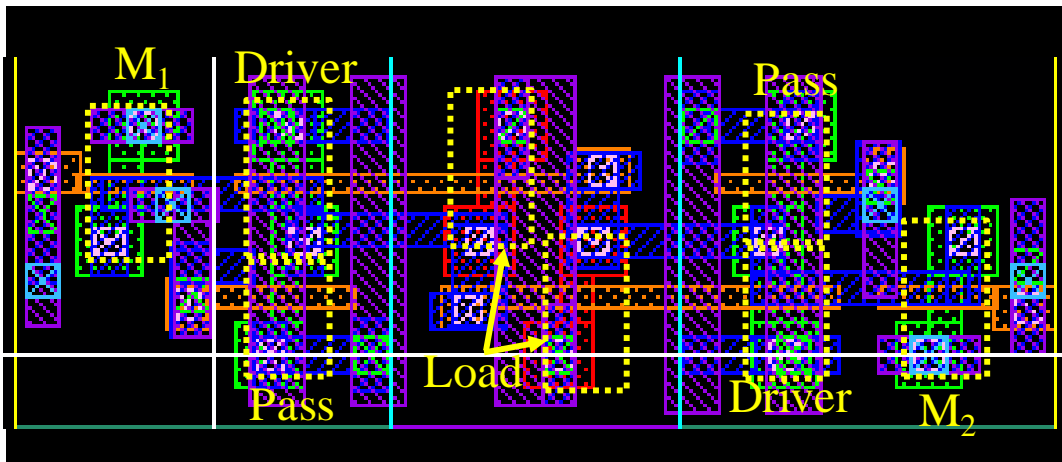
図 3.3.6 (a)H ストア時/L ストア時のストア電流密度と(b)SNM(mode=READ, SLP, RST) についてのローカルばらつきを考慮したモンテカルロシミュレーションの解析結果

表 3.3.2 NV セルの設計結果

Cell and array		Circuit operation and organization	
Channel length L	60 nm	Read/Write speed	1 GHz
Supply voltage V_{DD}	1.2 V		
Channel width W_{cell}	100 nm	Leakage control	
$W_{DRV}:W_{LD}:W_B$ ($W_{/B}$)	$1.5W_{cell}: W_{cell}: W_{cell}$	Cell	V_{SR} -0.2V
W_{M1}, W_{M2}	W_{cell}		V_{PG1} 1.4 V ($V_{DD}=0.02V$)
W_{PS1}	150nm (per 1 cell)	Peripheral	$V_{DD}=0.02V$
W_{PS2}	100nm (per 16 lines)		
Subarray size	8 kB	Sleep control	V_{PG2} 1.2 V ($V_{DD}=0.8V$)
Read/Write bit width	64 bit	Store control	V_{SR} 0.7 V
Store bit width	64 bit		V_{CTRL} 0.5 V
Store-free block size	1 kB	Duration of store operation	20 ns

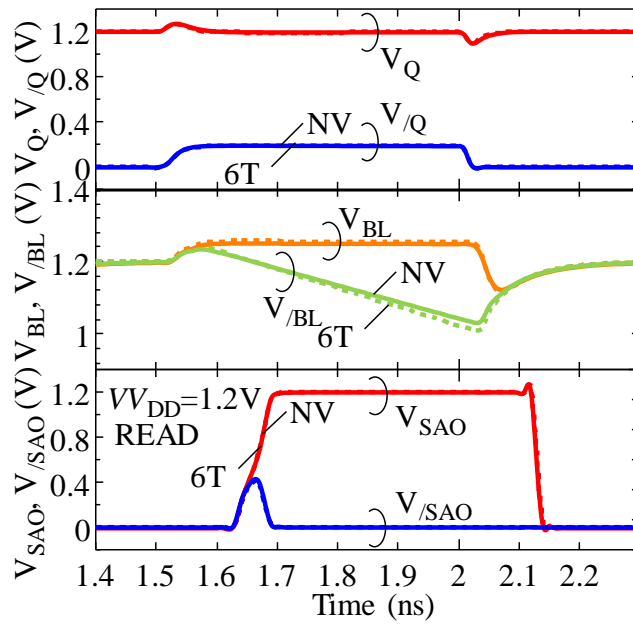


(a)

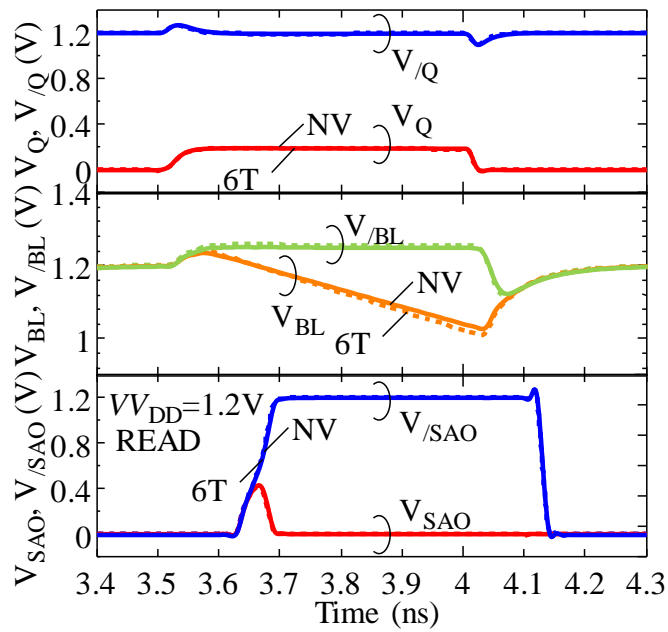


(b)

図 3.3.7 (a)6T セルと(b)NV セルのレイアウト

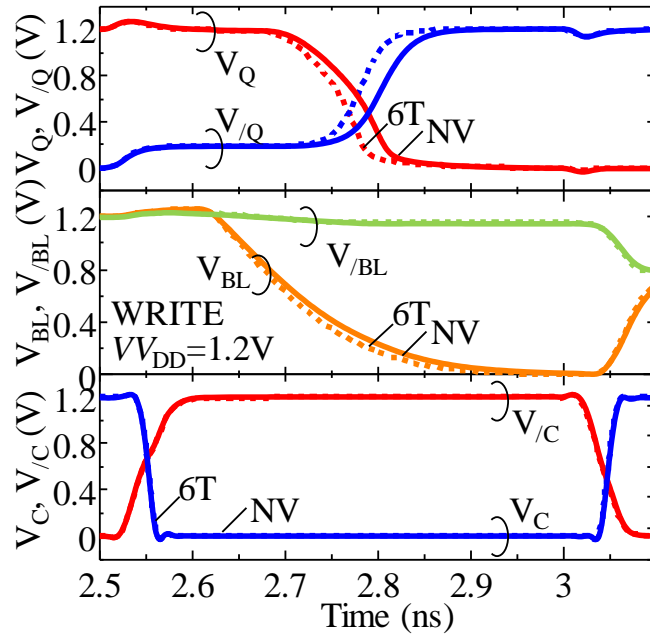


(a)

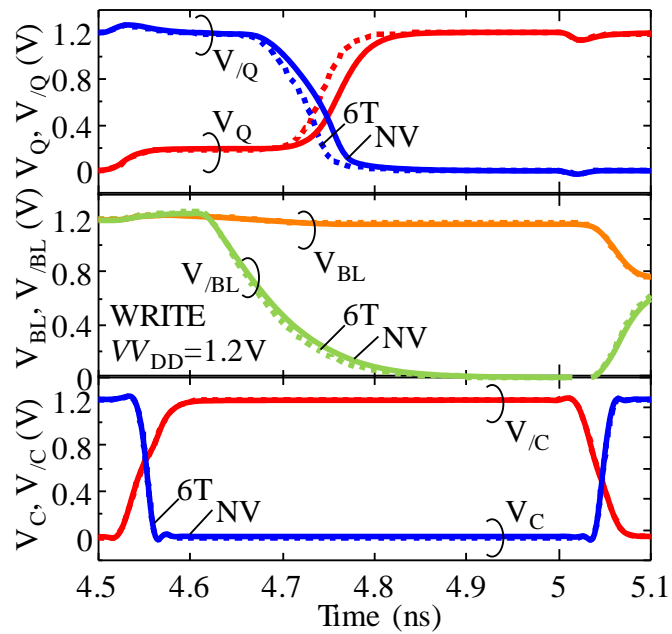


(b)

図 3.3.8 (a) $(V_Q, V_{\bar{Q}}) = (V_{DD}, 0)$, (b) $(V_Q, V_{\bar{Q}}) = (0, V_{DD})$ における通常の SRAM の読み出し動作の波形

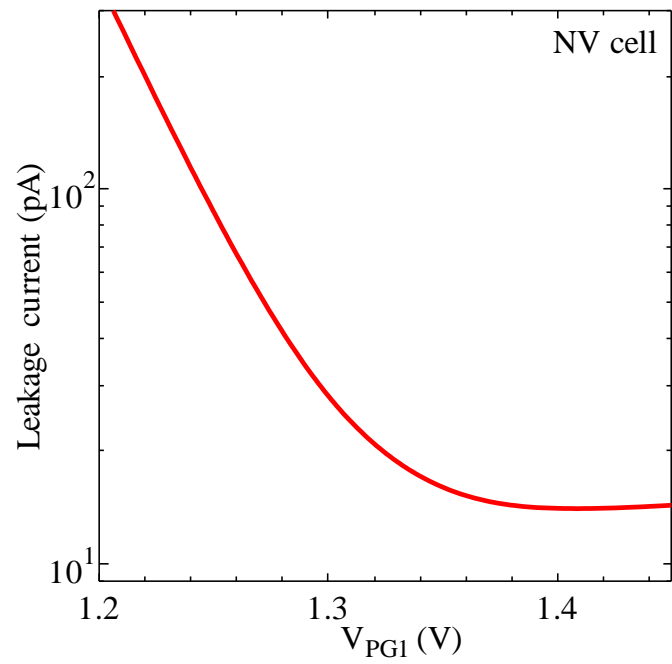


(a)

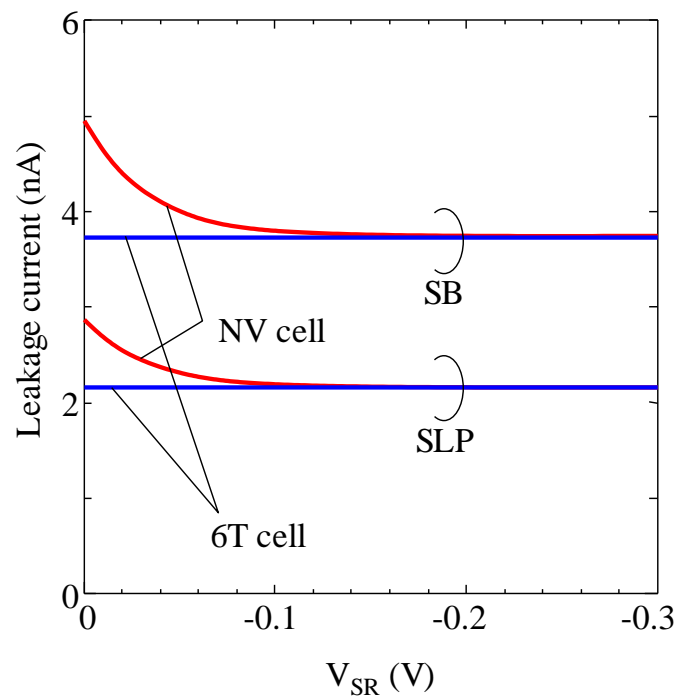


(b)

図 3.3.9 (a) $(V_Q, V_{/Q}) = (V_{DD}, 0)$, (b) $(V_Q, V_{/Q}) = (0, V_{DD})$ における通常の SRAM の書き込み動作の波形



(a)



(b)

図 3.3.10 (a)電源遮断時のリーク電流の V_{PG1} 依存性, (b)通常の SRAM 動作時のリーク電流の V_{SR} 依存性

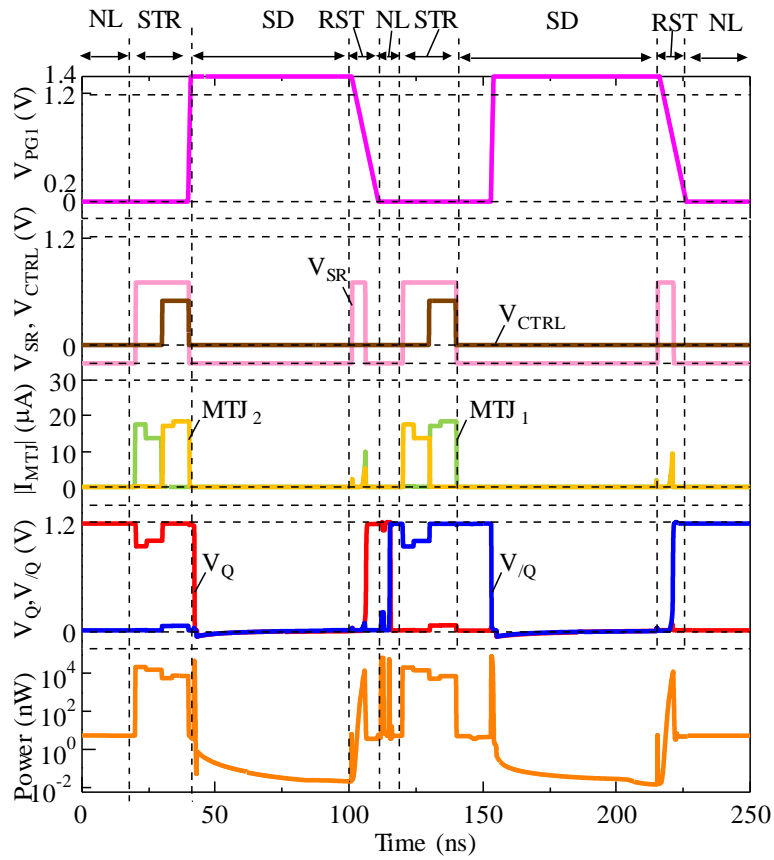


図 3.3.11 最適設計した NV セルの動作波形

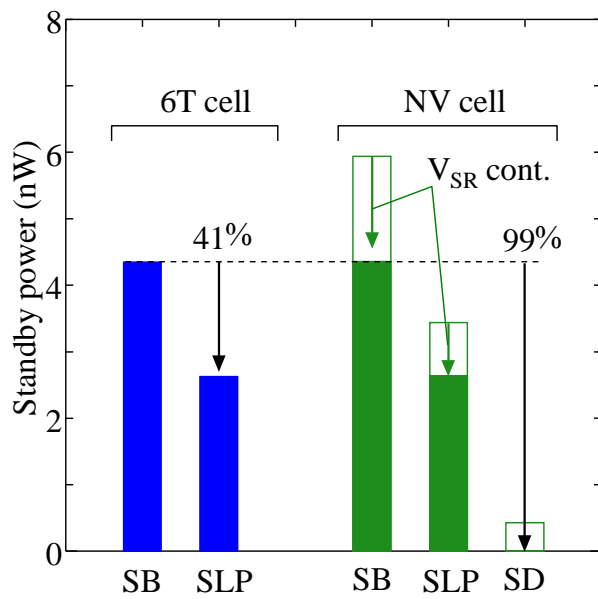


図 3.4.1 6TセルとNVセルの各動作モード(mode=SB, SLP, SD)におけるリーク電力

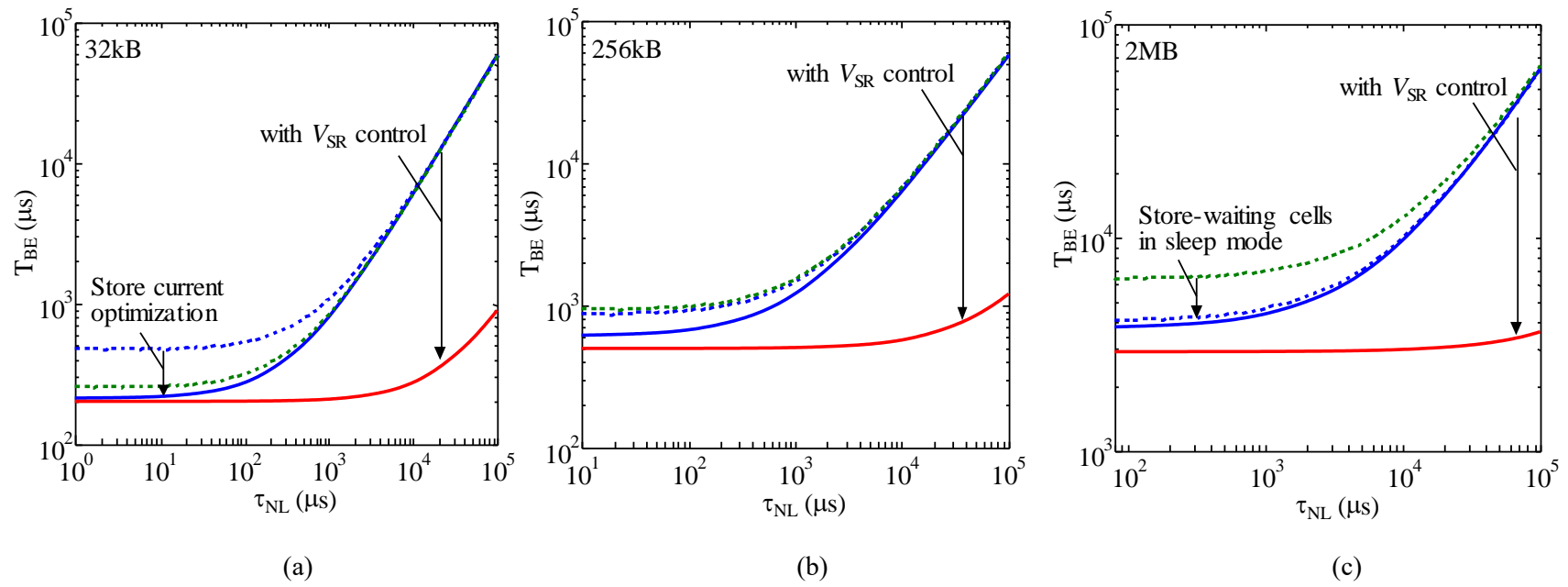


図 3.4.2 (a)32kB, (b)256kB, (c)2MB における T_{BE} の τ_{NL} 依存性

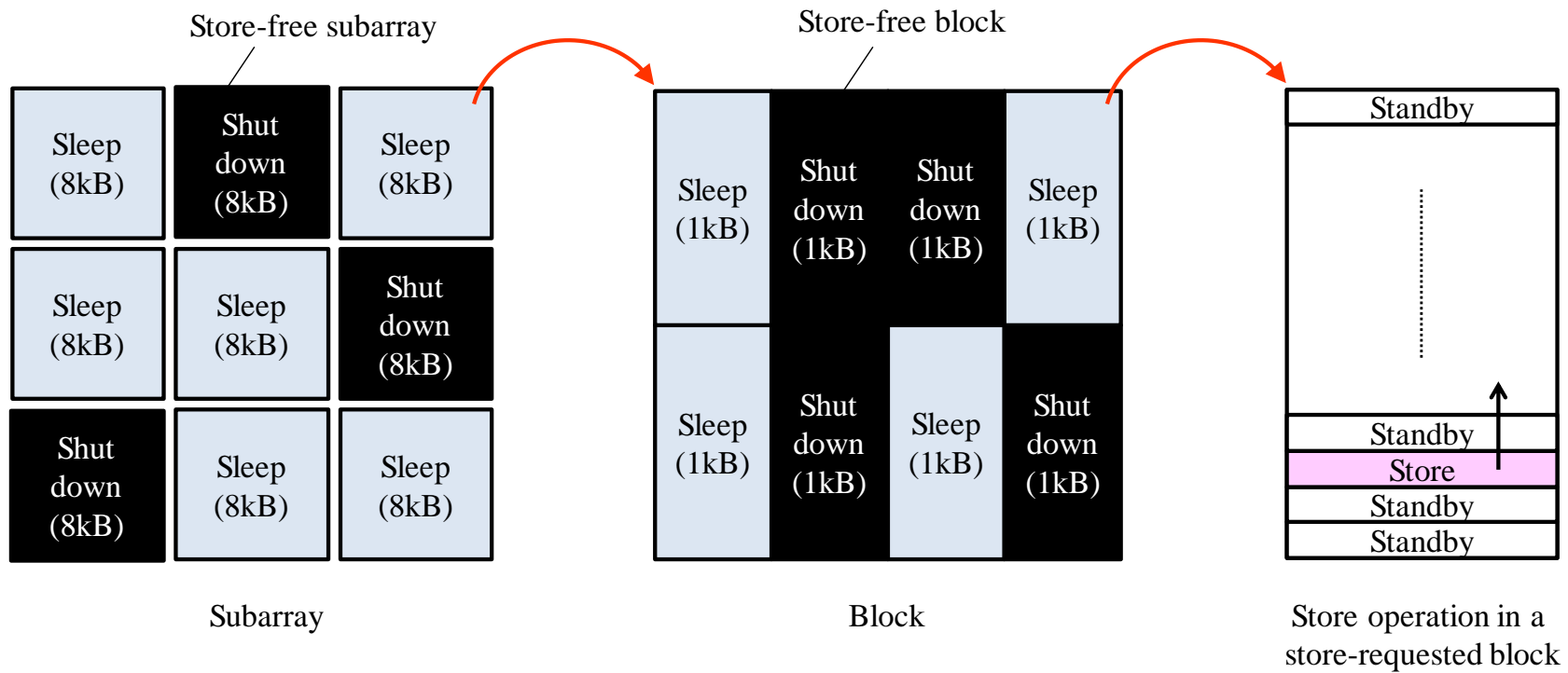


図 3.5.1 HSF アーキテクチャの概念図

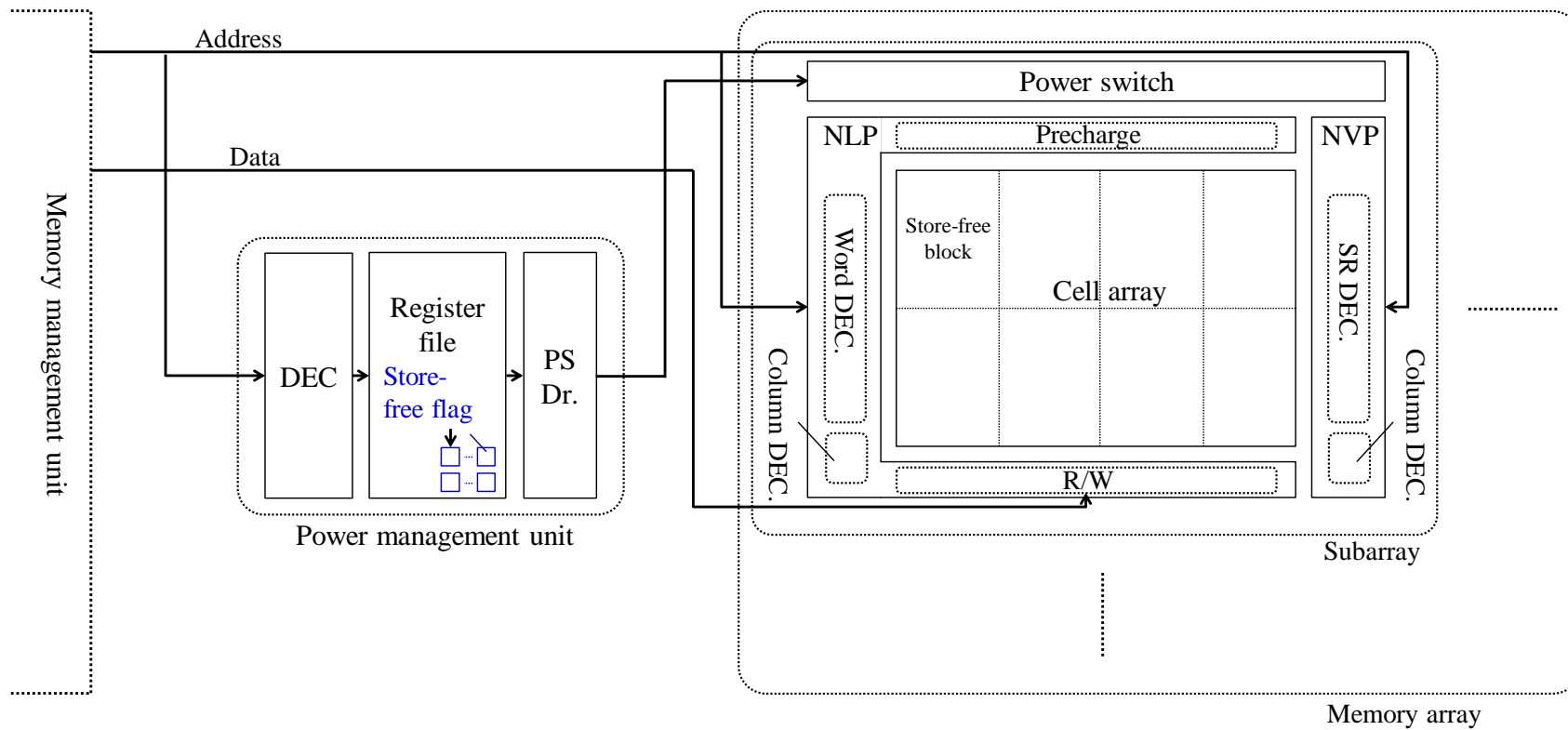


図 3.5.2 HSF アーキテクチャの実装方法

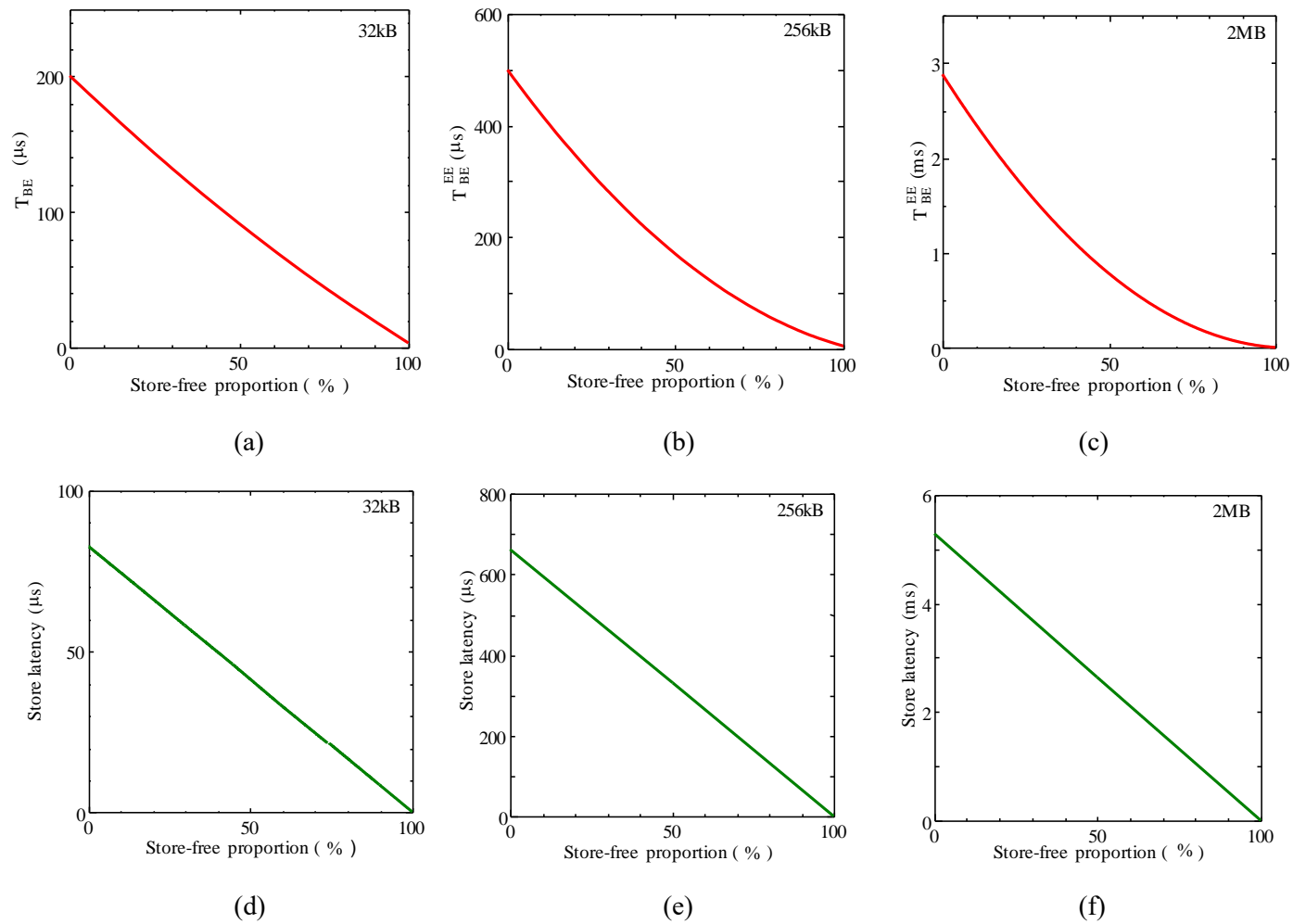


図 3.5.3 32kB, 256kB, 2MB の NV-SRAM における(a)-(c) T_{BE}^{EE} と(d)-(f)ストアレイテンシのストアフリー率依存性

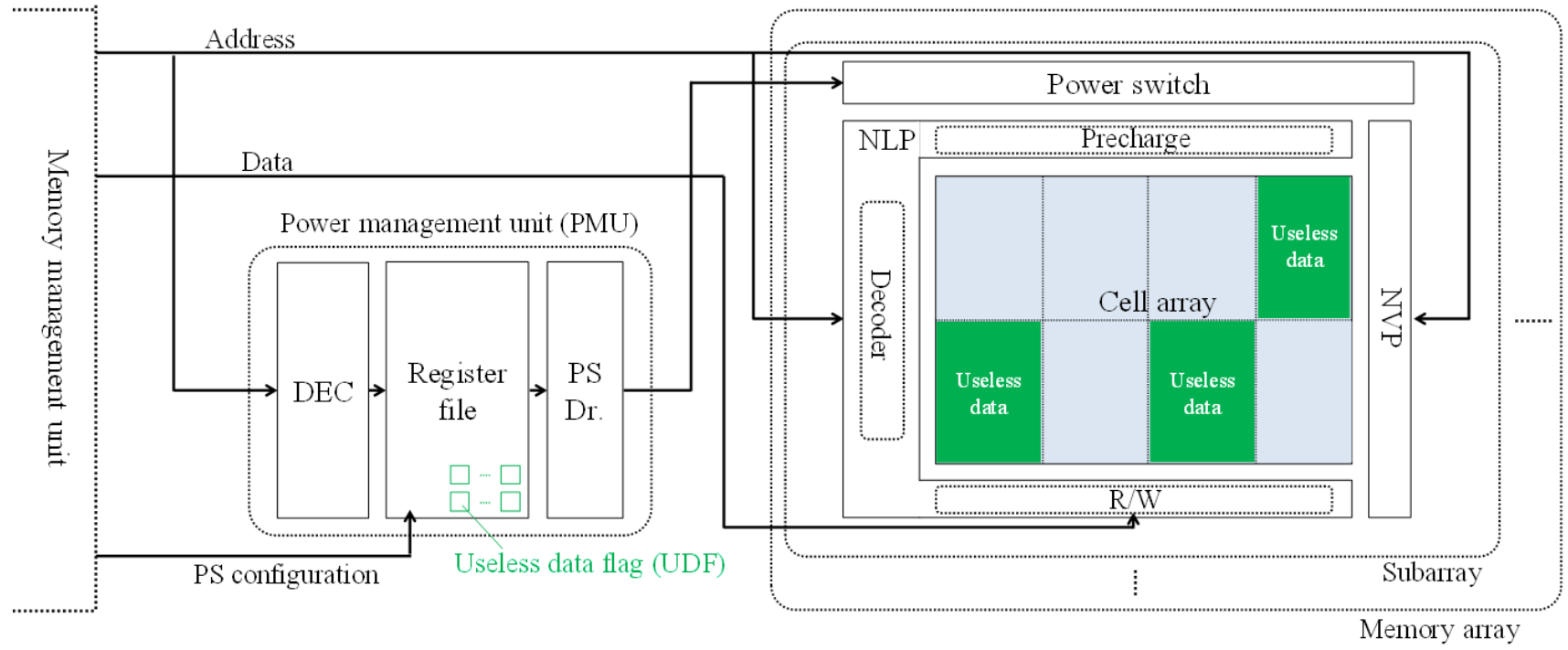


図 3.5.4 PUDF アーキテクチャの実装方法

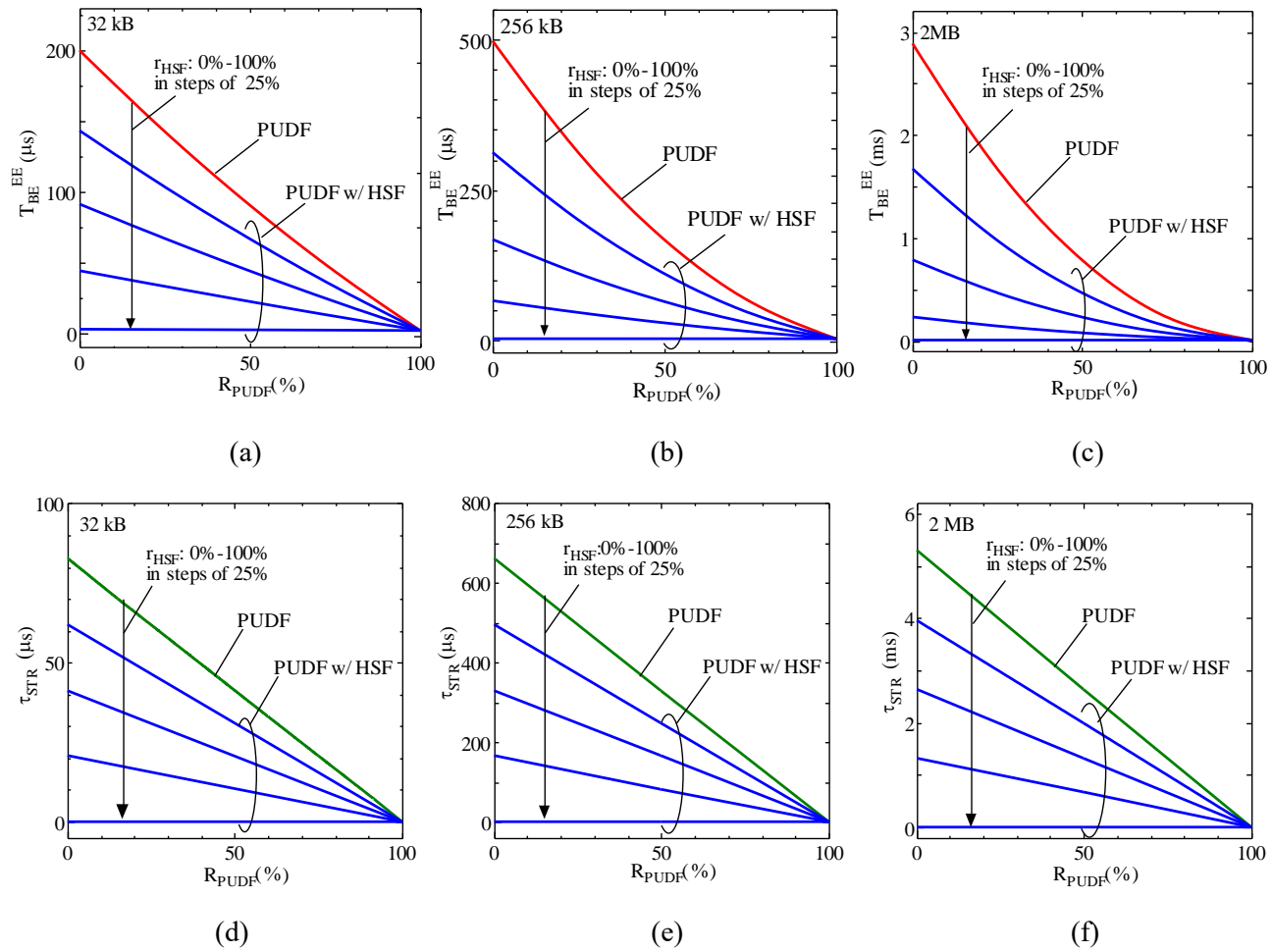
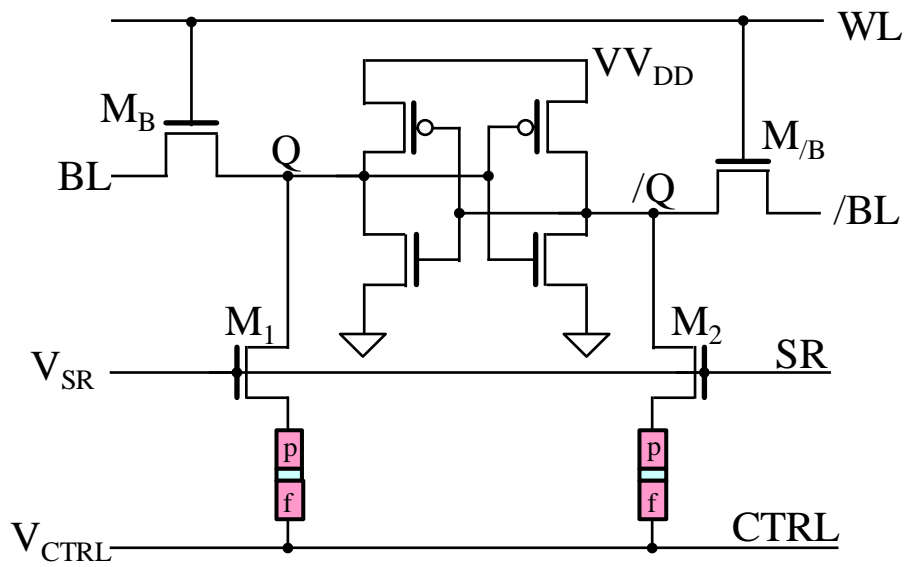
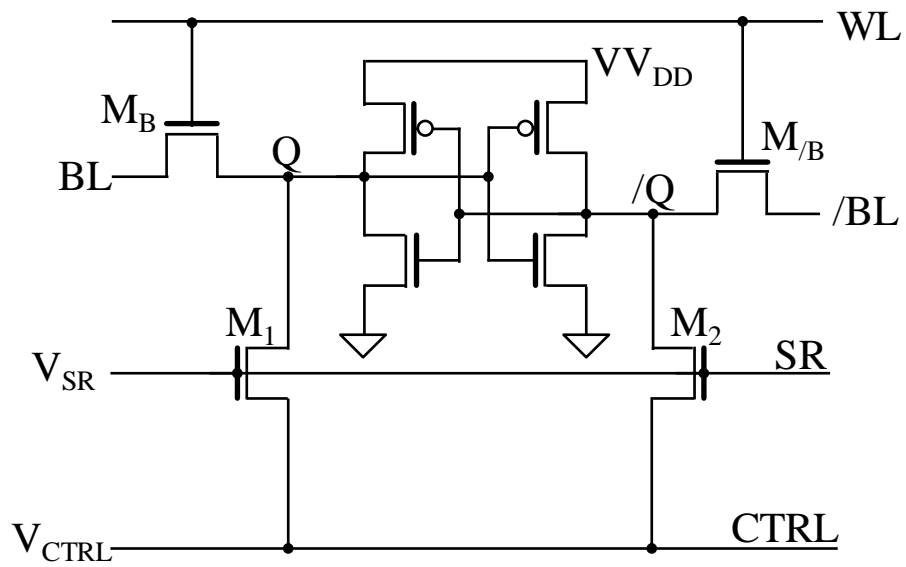


図 3.5.5 (a)32kB, (b)256kB, (c)2MB の NV-SRAM における(a)-(c) T_{BE}^{EE} のと(d)-(f) τ_{STR} の無用データ率(R_{PUDF}) 依存性

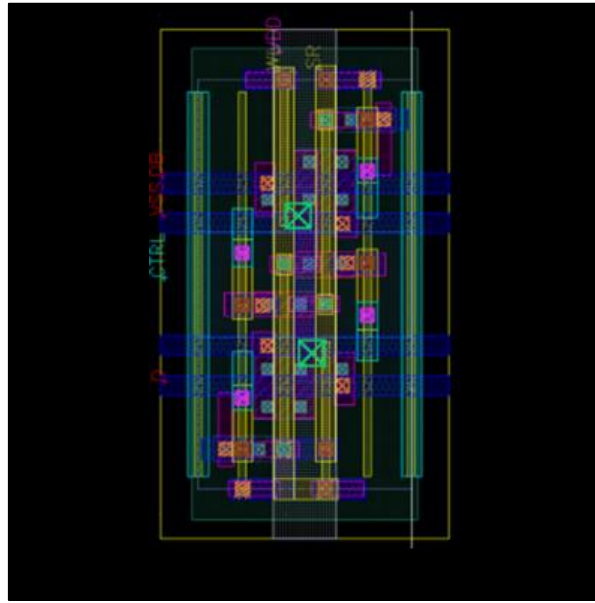


(a)

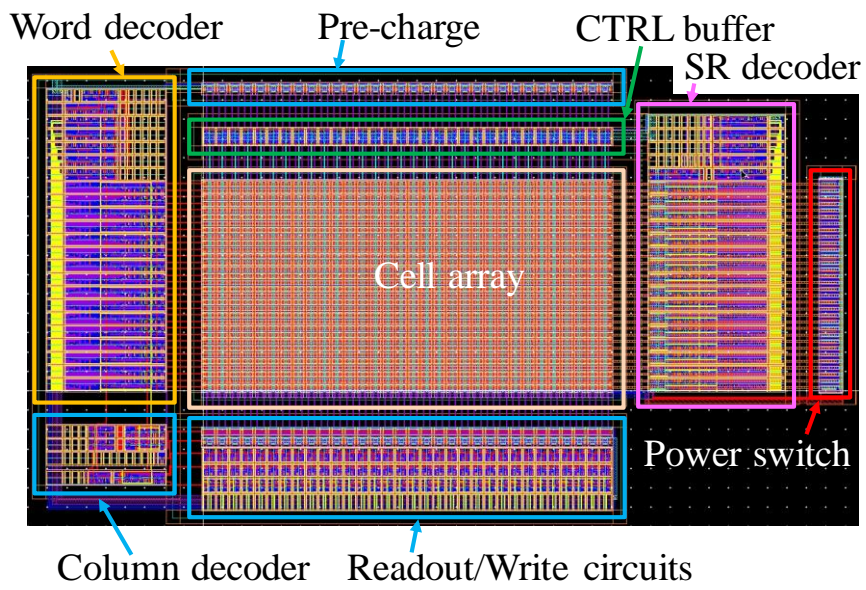


(b)

図 3.6.1 (a)MTJを実装した NV セルと(b)試作した NV セル



(a)



(b)

図 3.6.2 試作した NV-SRAM TEG における(a)セルと(b)アレイのレイアウト

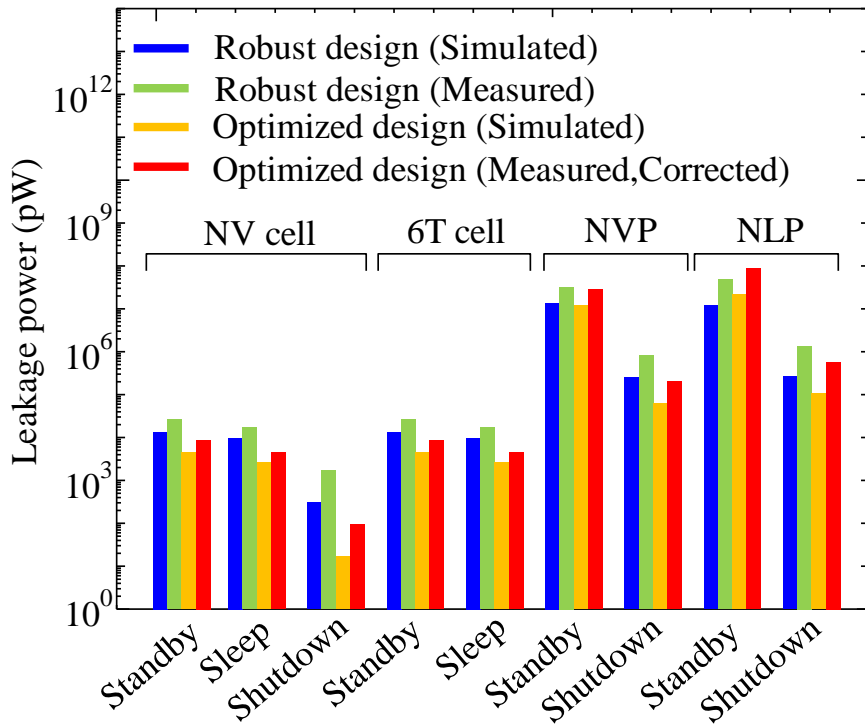


図 3.6.3 試作した TEG から抽出した回路パラメータと TEG と同様の構成における HSPICE シミュレーション結果

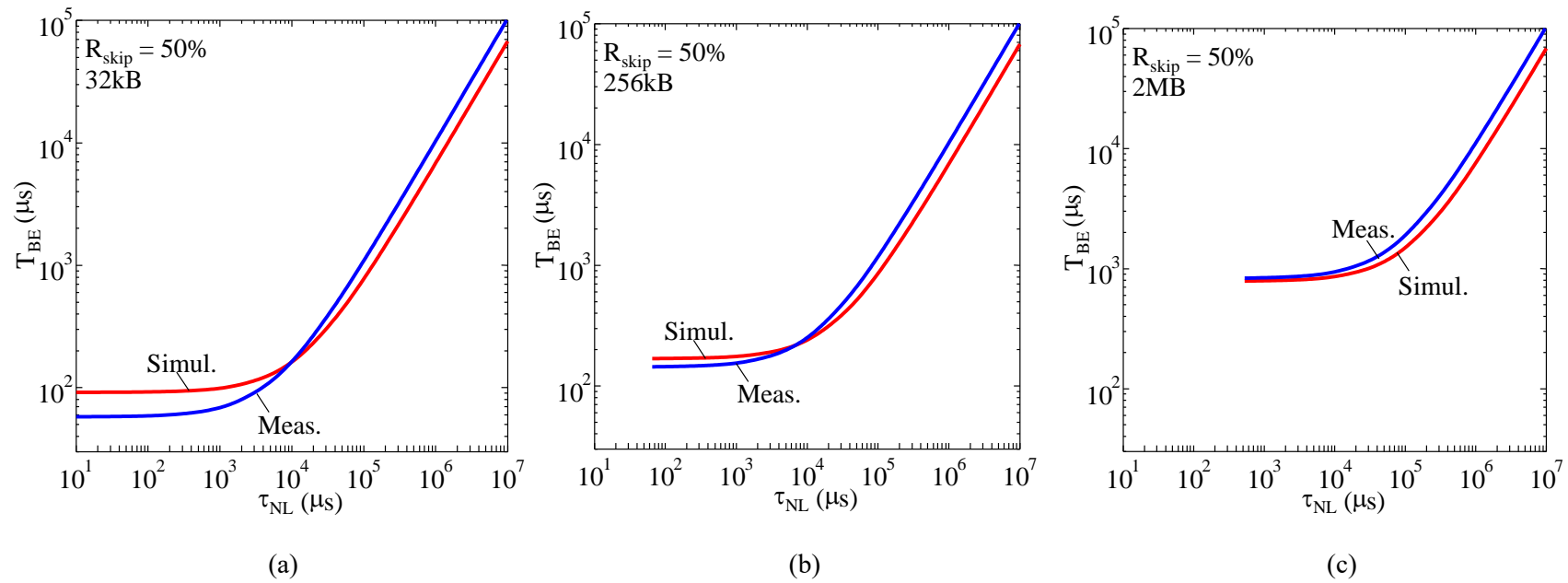
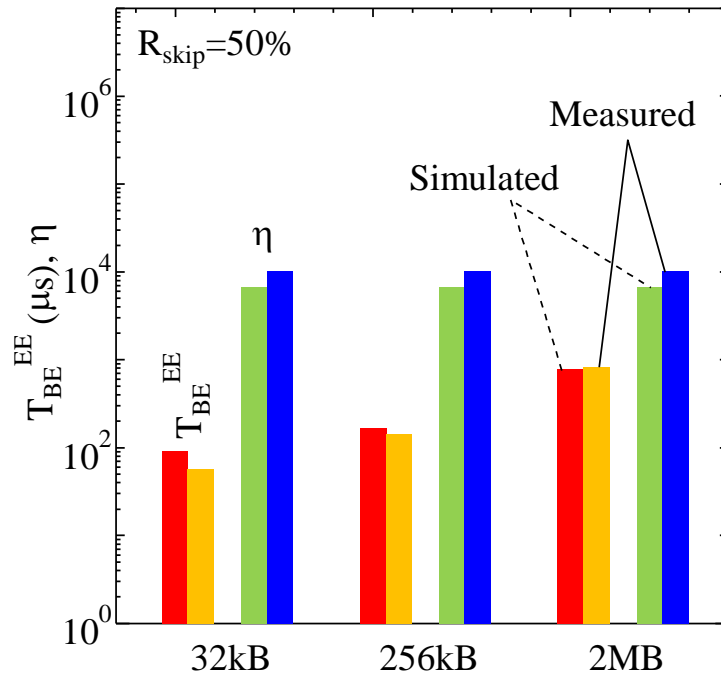
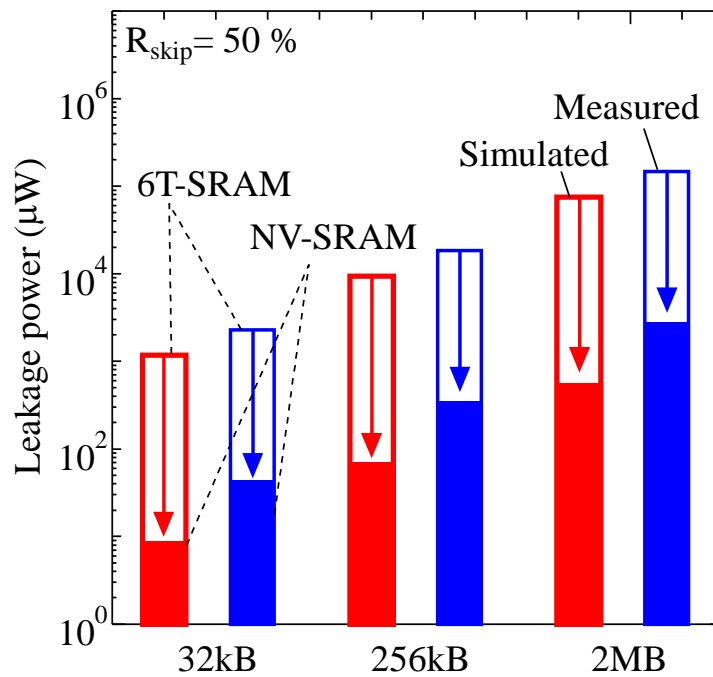


図 3.6.4 (a)32kB, (b)256kB, (c)2MB の NV-SRAM における T_{BE} の τ_{NL} 依存性

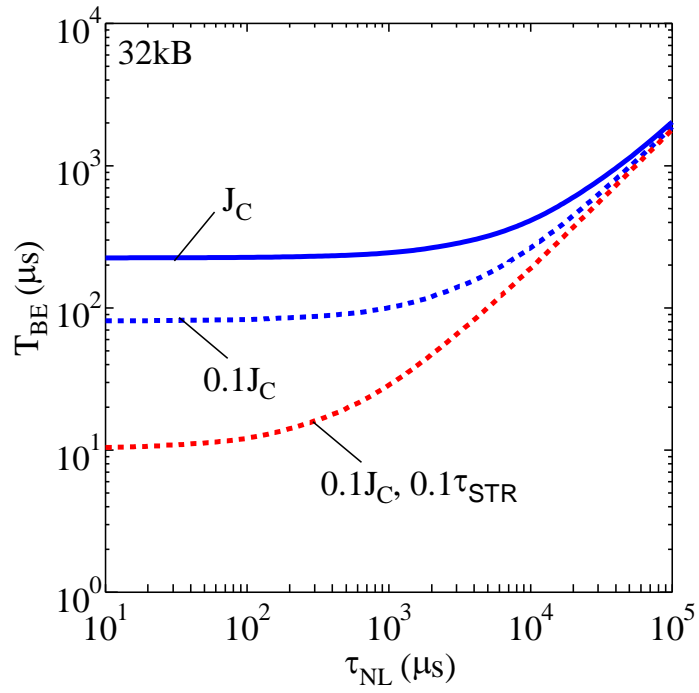


(a)

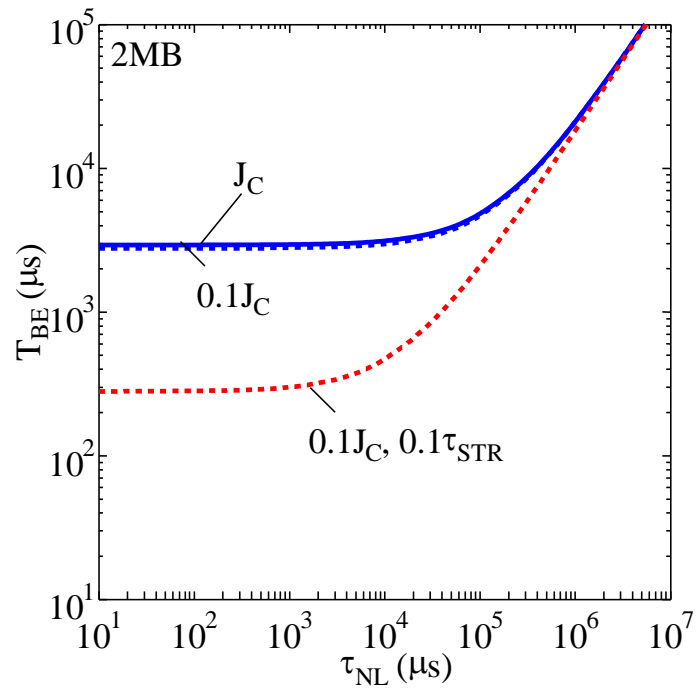


(b)

図 3.6.5 TEG から抽出した回路パラメータを用いて評価した(a) T_{BE}^{EE} と η , (b)待機時電力



(a)



(b)

図 3.8.1 (a)32kB, (b)2MB の NV-SRAM における T_{BE} の τ_{NL} 依存性

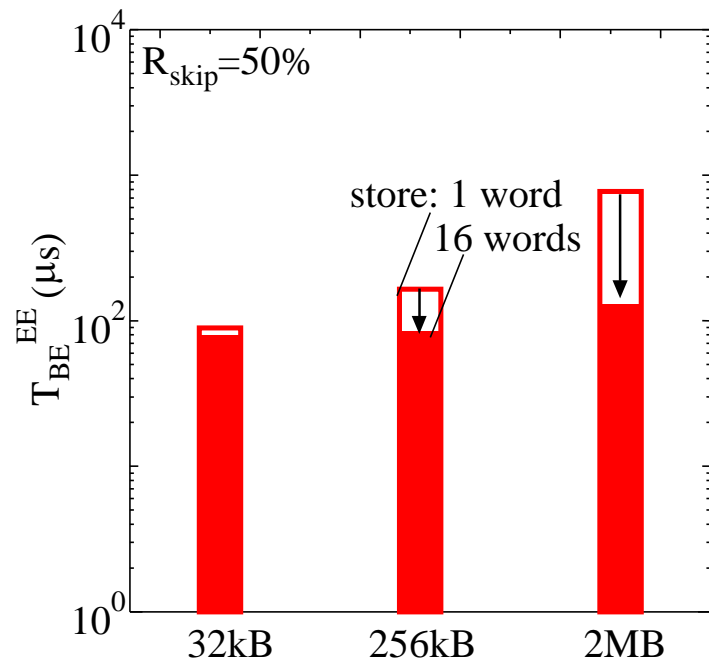


図 3.8.2 2 種類の電流許容値を用いた場合の T_{BE}^{EE}

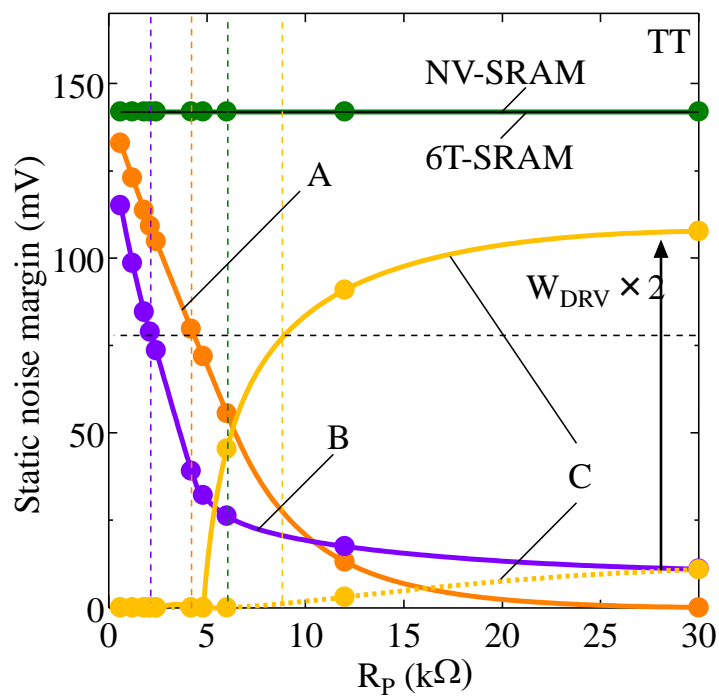


図 3.8.3 各種不揮発性 SRAM セルの SNM の R_p 依存性

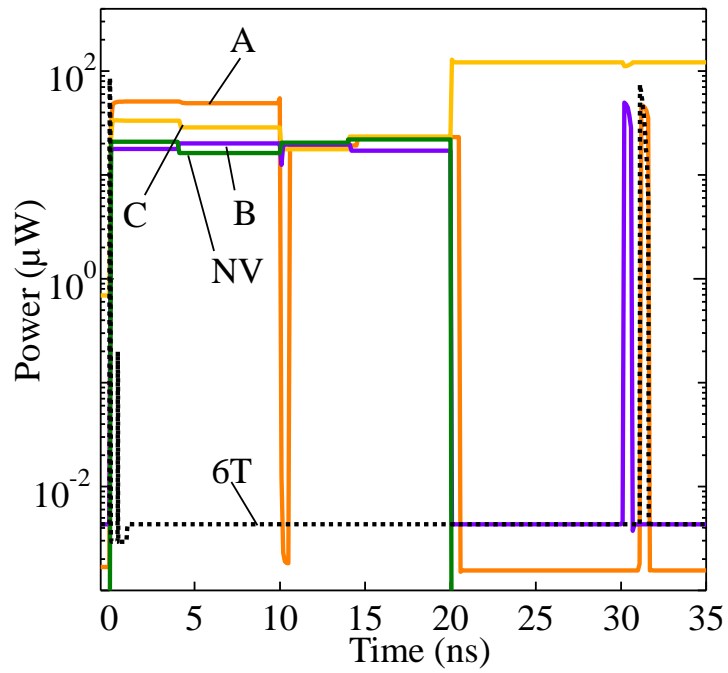


図 3.8.4 各種セルの通常動作の書き込み時, スタンバイ時, 読み出し時の電力の推移

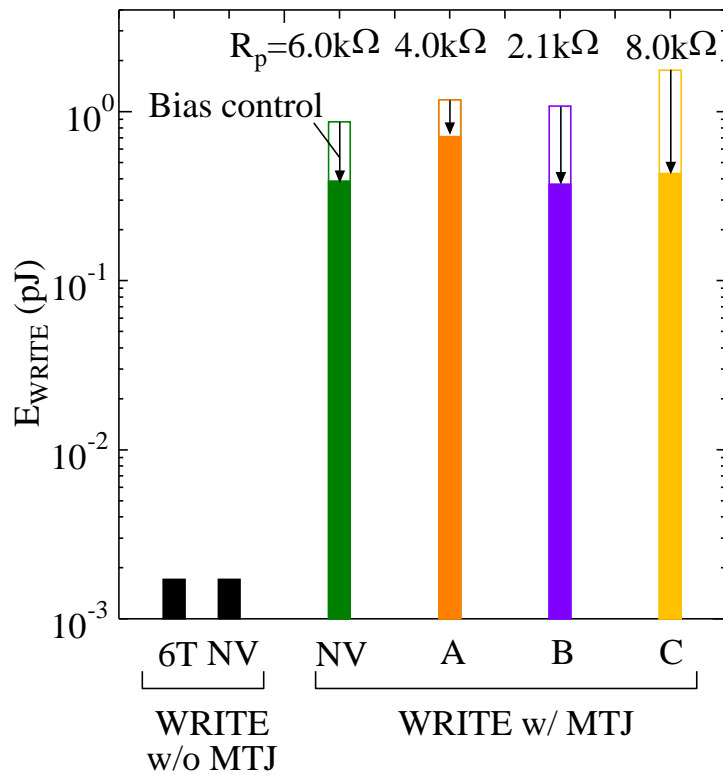


図 3.8.5 各種セルの書き込み動作に要するエネルギー

表 3.8.1 各種セルのエネルギー評価に用いた数値

		E_{WRITE}	P_{SB}	E_{READ}
WRITE w/o MTJ				
6T		3.57 fJ	4.39 nW	17.5 fJ
NV		3.57 fJ	4.42 nW	17.5 fJ
WRITE w/ MTJ				
NV	w/ bias cont.	389 fJ	4.59 nW	17.6 fJ
	w/o bias cont.	905 fJ		
A	w/ bias cont.	714 fJ	4.41 nW	18.0 fJ
	w/o bias cont.	1.26 pJ		
B	w/ bias cont.	372 fJ	4.41 nW	12.0 fJ
	w/o bias cont.	1.09 pJ		
C	w/ bias cont.	430 fJ	150 μ W	134 fJ
	w/o bias cont.	1.85 pJ		

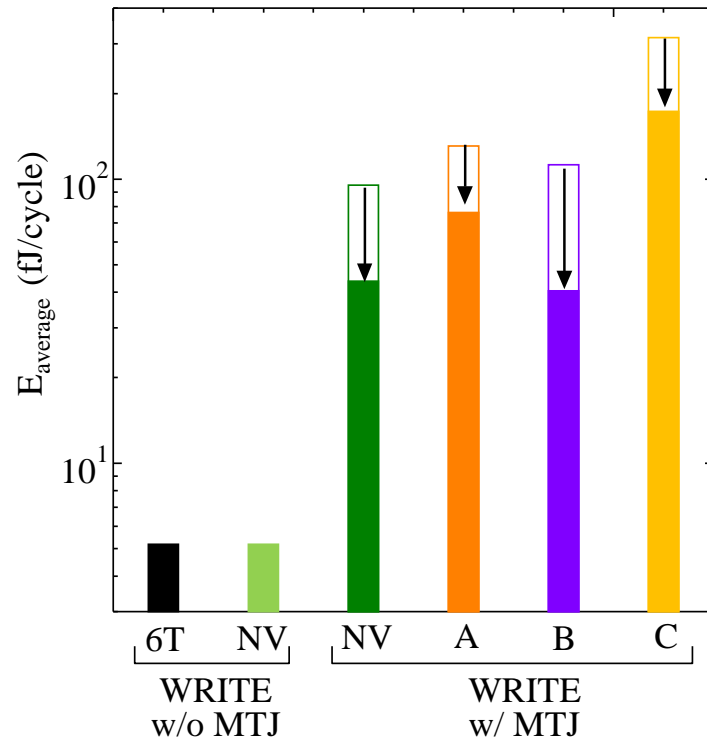


図 3.8.6 各種セルの E_{average} の評価結果

3.10 参考文献

- [1] J. Drori, S. Jewell-Larsen, R. Klein, W. Owen, R. Simko, W. Tchon, M. Darwish, and H. Dill, "A single 5V supply nonvolatile static RAM", IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 148 (1981).
- [2] D. J. Lee, N. J. Becker, A. L. Schlafly, J. A. Skupnjak, and V. K. Dham, "Control logic and cell design for a 4K NVRAM", IEEE Journal of Solid-State Circuits **18**, 525 (1983).
- [3] K. Abe, K. Nomura, S. Ikegawa, T. Kishi, H. Yoda, and S. Fujita, "Hierarchical Nonvolatile Memory with Perpendicular Magnetic Tunnel Junctions for Normally-Off Computing", International Conference on Solid State Devices and Materials (SSDM) 2010, 1144 (2010).
- [4] W. Zhao, C. Chappert, V. Javerliac, and J.-P. Noziere, "High speed, high stability and low power sensing amplifier for MTJ/CMOS hybrid logic circuits", IEEE Trans. Magn. **45**, 3784 (2009).
- [5] T. Ohsawa, F. Iga, S. Ikeda, T. Hanyu, H. Ohno, and T. Endo, "Studies on Static Noise Margin and Scalability for Low-Power and High-Density Nonvolatile SRAM using Spin -Transfer -Torque (STT) MTJs", International Conference on Solid State Devices and Materials (SSDM) 2011, 959 (2011).
- [6] S. Yamamoto and S. Sugahara, "Nonvolatile static random access memory using magnetic tunnel junctions with current-induced magnetization switching architecture", Jpn. J. Appl. Phys. **48**, 043001 (2009).
- [7] Y. Shuto, S. Yamamoto, and S. Sugahara, "Nonvolatile static random access memory based on spin-transistor architecture", J. Appl. Phys. **105**, 07C933 (2009).
- [8] X. Xue, G. Jin, J. Zhang, L. Xu, Y. Ding, Y. Xie, C. Zhao, B. A. Chen, and Y. Lin, "Nonvolatile SRAM cell based on Cu_xO ", IEEE 9th International Conference on Solid-State and Integrated-Circuit Technology 2008, 869 (2008).
- [9] S.-S. Sheu, C.-C. Kuo, M.-F. Chang, P.-L. Tseng, L. C.-Sheng, M.-C. Wang, C.-H. Lin, W.-P. Lin, T.-K. Chien, S.-H. Lee, S.-C. Liu, H.-Y. Lee, P.-S. Chen, Y.-S. Chen, C.-C. Hsu, F. T. Chen, K.-L. Su, T.-K. Ku, M.-J. Tsai, M.-J. Kao, "A ReRAM Integrated 7T2R Non-volatile SRAM for Normally-off Computing Application", 2013 IEEE Asian Solid-State Circuits Conference (ASSCC) 2013, 245, (2013).
- [10] Y. Liu, Z. Wang, A. Lee, F. Su, C.-P. Lo, Z. Yuan, C.-C. Lin, Q. Wei, Y. Wang, Y.-C. King, C.-J. Lin, P. Khalili, K.-L. Wang, M.-F. Chang, H. Yang, "A 65nm

- ReRAM-Enabled Nonvolatile Processor with 6× Reduction in Restore Time and 4× Higher Clock Frequency Using Adaptive Data Retention and Self-Write-Termination Nonvolatile Logic”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 84 (2016).
- [11] S. Yamamoto, Y. Shuto, and S. Sugahara, “Nonvolatile SRAM (NV-SRAM) using functional MOSFET merged with resistive switching devices”, IEEE Custom Integrated Circuits Conference (CICC), 531, (2009).
- [12] M. Takata, K. Nakayama, T. Izumi, T. Shinmura, J. Akita, and A. Kitagawa, “Nonvolatile SRAM based on phase change”, IEEE 21st Non-Volatile Semiconductor Memory Workshop (NVSMW) 2006, 95 (2006).
- [13] J.-M. Choi, C.-M. Jung, and K.-S. Min, “PCRAM flip-flop circuits with sequential sleep-in control scheme and selective write latch”, Journal of Semiconductor Technology and Science **13**, 58 (2013)
- [14] S. S. Eaton, “A ferroelectric nonvolatile memory”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 130 (1988)
- [15] T. Miwa, J. Yamada, H. Koike, H. Toyoshima, K. Amanuma, S. Kobayashi, T. Tatsumi, Y. Maejima, H. Hada, T. Kunio, “NV-SRAM: A nonvolatile SRAM with backup ferroelectric capacitors”, IEEE Journal of Solid-State Circuits **36**, 522 (2001).
- [16] S. Masui, W. Yokozeki, M. Oura, T. Ninomiya, K. Mukaida, Y. Takayama, and T. Teramoto, “Design and applications of ferroelectric nonvolatile SRAM and flip-flop with unlimited read/program cycles and stable recall”, IEEE Custom Integrated Circuits Conference (CICC) 2003, 403 (2003).
- [17] M. Kobayashi, N. Ueyama, K. Jang, and T. Hiramoto, “Experimental Demonstration of a Nonvolatile SRAM With Ferroelectric HfO₂ Capacitor for Normally Off Application”, IEEE Journal of the Electron Devices Society **6**, 280 (2018).
- [18] T. Endoh, H. Koike, S. Ikeda, T. Hanyu, and H. Ohno, “An overview of nonvolatile emerging memories—Spintronics for working memories”, IEEE Journal on Emerging and Selected Topics in Circuits and Systems **6**, 109 (2016).
- [19] S. Ikeda, J. Hayakawa, Y. Ashizawa, Y. M. Lee, K. Miura, H. Hasegawa, M. Tsunoda, F. Matsukura, and H. Ohno, “Tunnel magnetoresistance of 604% at 300K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature”, Appl. Phys. Lett. **93**, 082508 (2008).
- [20] H. Yoda, T. Kishi, T. Nagase, M. Yoshikawa, K. Nishiyama, E. Kitagawa, T. Daibou,

- M. Amano, N. Shimomura, S. Takahashi, T. Kai, M. Nakayama, H. Aikawa, S. Ikegawa, M. Nagamine, J. Ozeki, S. Mizukami, M. Oogane, Y. Ando, S. Yuasa, K. Yakushiji, H. Kubota, Y. Suzuki, Y. Nakatani, T. Miyazaki, and K. Ando, “High efficient spin transfer torque writing on perpendicular magnetic tunnel junctions for high density MRAMs”, *Current Appl. Phys.* **10**, e87 (2010).
- [21] J. H. Park, Y. Kim, W. C. Lim, J. H. Kim, S. H. Park, J. H. Kim., W. Kim, K. W. Kim, J. H. Jeong, K. S. Kim, H. Kim, Y. J. Lee, S. C. Oh, J. E. Lee, S. O. Park, S. Watts, D. Apalkov, V. Nikitin, M. Krounbi, S. Jeong, S. Choi, H. K. Kang, and C. Chung, “Enhancement of data retention and write current scaling for sub-20nm STT-MRAM by utilizing dual interfaces for perpendicular magnetic anisotropy”, *VLSI Symp. 2012*, paper 7.1 (2012).
- [22] M. Gajek, J. J. Nowak, J. Z. Sun, P. L. Trouilloud, E. J. O’Sullivan, D. W. Abraham, M. C. Gaidis, G. Hu, S. Brown, Y. Zhu, R. P. Robertazzi, W. J. Gallagher, and D. C. Worledge, “Spin torque switching of 20 nm magnetic tunnel junctions with perpendicular anisotropy”, *Appl. Phys. Lett.* **100**, 132408 (2012).
- [23] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, “1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS”, *IEEE Journal of Solid-state circuits* **30**, 847 (1995).
- [24] S. Ohbayashi, M. Yabuuchi, K. Nii, Y. Tsukamoto, S. Imaoka, Y. Oda, T. Yoshihara, M. Igarashi, M. Takeuchi, H. Kawashima, Y. Yamaguchi, K. Tsukamoto, M. Inuishi, H. Makino, K. Ishibashi, and H. Shinohara, “A 65-nm SoC embedded 6T-SRAM designed for manufacturability with read and write operation stabilizing circuits”, *IEEE Journal of Solid-State Circuits* **42**, 820 (2007).
- [25] J.-C. Chiu, Y.-J. Tsou, H.-C. Shih, and C. W. Liu, “Write Error Rate Prediction of STT-pMTJ Considering Process Variations and Thermal Fluctuations”, *IEEE International Electron Devices Meeting (IEDM) 2017 MRAM special session*, P-6 (2019).
- [26] S. Rusu, S. Tam, H. Muljono, D. Ayers, J. Chang, “A dual-core multi-threaded Xeon processor with 16MB L3 cache”, *IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, 315 (2006).
- [27] P. Zhou, B. Zhao, J. Yang, Y. Zhang, “Energy reduction for STT-RAM using early write termination”, *IEEE ICCAD 2009*, 264 (2009).
- [28] C. W. Smullen, V. Mohan, A. Nigam, S. Gurusurthi, and M. R. Stan, “Relaxing Non-Volatility for Fast and Energy-Efficient STT-RAM Caches”, *IEEE HPCA 2011*,

50 (2011).

- [29] A. Jog, A. K. Mishra, C. Xu, and Y. Xie, "Cache Revive: Architecting Volatile STT-RAM Caches for Enhanced Performance in CMPs", ACM/IEEE Design Automation Conference 2012, 243 (2012).
- [30] M. H. Samavatian, H. Abbasitabar, M. Arjomand, and H. S. Azad, "An Efficient STT-RAM Last Level Cache Architecture for GPUs", ACM/IEEE Design Automation Conference 2014, 1 (2014).
- [31] A. Nigam C. W. Smullen, V. Mohan, E. Chen, S. Gurumurthi, and M. R. Stan, "Delivering on the Promise of Universal Memory for Spin-Transfer Torque RAM (STT-RAM)", IEEE ISLPED 2011, 121 (2011).
- [32] S. Yazdanshenas, M. P. Pirbasti, M. Fazeli, and A. Patooghy, "Coding Last Level STT-RAM Cache For High Endurance And Low Power", IEEE computer architecture letters **13**, 73 (2013).
- [33] A. Jadidi, M. Arjomand, and H. S. Azad, "High-Endurance and Performance-Efficient Design of Hybrid Cache Architectures through Adaptive Line Replacement", IEEE/ACM ISLPED 2011, 79 (2011).
- [34] J. Ahn, and K. Choi, "Lower-Bits Cache for Low Power STT-RAM Caches", IEEE ISCAS 2012, 480 (2012).
- [35] Z. Wang, D. A. Jimenez, C. Xu, G. Sun, and Y. Xie, "Adaptive Placement and Migration Policy for an STT-RAM-Based Hybrid Cache", IEEE HPCA 2014, 13 (2014).
- [36] W. Cheng, Y. Ciou, and P. Shen, "Architecture and data migration methodology for L1 cache design with hybrid SRAM and volatile STT-RAM configuration", Microprocessors and Microsystems **42**, 191 (2016).
- [37] D. Kitagata, S. Yamamoto, and S. Sugahara, "Design and energy-efficient architectures for nonvolatile static random access memory using magnetic tunnel junctions", Jpn. J. Appl. Phys. **58**, SBBB12 (2019).
- [38] D. A. Patterson, J. L. Hennessy, "Computer organization and design: the hardware/software interface" Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2013, fifth edition.
- [39] VLSI Design and Education Center (VDEC), University of Tokyo, <http://www.vdec.u-tokyo.ac.jp/English/index.html>
- [40] Y. Shuto, S. Yamamoto, and S. Sugahara, "Energy performance of nonvolatile power-gating SRAM using SOTB technology", 46th European Solid-State Device

Research Conference (ESSDERC) 2016, 87 (2016).

- [41] Y. Shuto, S. Yamamoto, and S. Sugahara, "Design and implementation of nonvolatile power-gating SRAM using SOTB technology", International Symposium on Low Power Electronics and Design (ISLPED) 2016, 338 (2016).
- [42] V. George, S. Jahagirdar, C. Tong, K. Smits, S. Damaraju, S. Siers, V. Naydenov, T. Khondker, S. Sarkar, and P. Singh, "Penryn: 45-nm Next Generation Intel® Core™ 2 Processor", IEEE Asian Solid-State Circuits Conference 2007, 14 (2007).
- [43] C. Zarka *et al.*, "Characterization of Alpha AXP performance using TP and SPEC workloads." Proceedings of 21 International Symposium on Computer Architecture. IEEE, 1994.

第 4 章

擬似不揮発性 SRAM

本章では、はじめに擬似不揮発性 SRAM (VNR-SRAM) セルの構成方法と超低電圧 (Ultralow voltage; ULV) によるデータ保持 (ULV リテンション) の課題について述べる。これまで低電圧 SRAM セルとして検討されてきたシュミットリグインバータで構成した SRAM セルの特徴と問題点を述べ、これらの問題点を解消できるデュアルモードインバータを用いた VNR-SRAM セルの回路構成と動作原理を述べる。続いて、本論文で提案する VNR-SRAM セルについて 2 種類の回路構成を示す。次に、VNR-SRAM セルの最適設計方法を示し、この方法で設計された VNR-SRAM セルの速度性能、Break-even time (BET) および電力削減効果の解析を行う。さらに、電力削減率のさらなる向上を目指し、VNR-SRAM の電力削減効率を究極的に高めるパワースイッチのアーキテクチャと、第 3 章で示した無用データフラッシュアーキテクチャ (Proactive useless-data flush; PUDF) を導入し、その効果を示す。また、最適設計したセルで構成した VNR-SRAM マクロの設計を行い、ポストレイアウト解析により VNR-SRAM の性能を検証する。最後に、先端 CMOS を用いて VNR-SRAM を構成した場合の性能評価を行う。

4.1 擬似不揮発性 SRAM の実現方法

4.1.1 擬似不揮発記憶を用いた PG

第 3 章では NV-SRAM を用いることで、各種アーキテクチャの導入により数 100 μ s 程度の BET (T_{BE}) を実現できること示した。これは従来の PG と比べて、細粒度の PG を実現できることを意味するが、さらなる細粒度化への可能性はある。一般に CMOS ロジックシステムの PG を行う場合、第 2 章で述べたように電源遮断/復帰移行時のラッシュカレントを抑えるため、数 μ s 程度の電源遮断/復帰レイテンシが必要となる[1]。このレイテン

シより短い時間の電源遮断は実行できないため、 T_{BE} はこの最小値より小さくしても意味がない。すなわち、 T_{BE} の実質的な最小値はこの値程度となる。よって、NV-SRAM で実現した T_{BE} には、まだ削減の余地がある。NV-SRAM の T_{BE} を限界まで削減できないのは不揮発性メモリ素子のストアエネルギーが原因であるが、このエネルギーは非常に大きく、第3章で示したように各種削減アーキテクチャを導入してもその削減には限界がある。また、NV-SRAM では CMOS ロジックへの不揮発性メモリ素子のエンベディッド技術が必要となるが、これは技術およびコストの面から容易ではない。

以上から、不揮発性メモリ素子を用いずに CMOS のみで実質的に不揮発記憶を実現できれば、上述の不揮発性メモリ素子に付随する問題を解消できると考えられる。PG における不揮発記憶の意義は電源遮断による待機時電力の大幅な削減である。非常に小さな電圧 (Ultralow voltage; ULV) を用いれば、電源遮断を用いなくても待機時電力を電源遮断時と同等にまで下げることができる可能性がある。したがって、CMOS のみから構成される記憶回路で、このように ULV でデータを安定に保持 (ULV リテンション) できれば、PG においては不揮発記憶と同等の電力削減効果を実現できると考えられる。

図 4.1.1(a), (b) に 6T セルにおけるスタンバイ電力の V_{DD} 依存性を模式的に示す。セルの V_{DD} の低下とともにスタンバイ電力も削減される。一般に、SRAM のスリープモードは、専用のパワースイッチを用いて、これを完全に導通させ、このパワースイッチに共有させるセル数から所望のノイズマージンを実現できる V_{DD} が決定される[2]。通常、この V_{DD} は電源電圧 (V_{DD}) の 7 割程度を用いる。同図の青色の領域にこのスリープ動作で用いるおおよその電圧領域を示す。この領域でも待機時電力は削減できるが、通常動作時の約 50% 程度までしか削減できない。しかし、同図の黄色で示した領域 (V_{DD} の 15% 程度) まで V_{DD} を低減できれば、電力の削減効果は飛躍的に向上し、90% 程度以上の待機時電力が削減できるようになる。以後、特にこのようなことが可能となる電圧のことを超低電圧 (ULV)、ULV でのデータの保持を ULV リテンションと呼ぶことにする。このような ULV は PG の電源遮断時の V_{DD} と同程度であり (通常のロジックシステムでは動作時の V_{DD} をできるだけ高くとるため遮断時においても V_{DD} を完全に 0 にすることはできず、 V_{DD} の数 10% 程度の電圧が生じる[2])、待機時電力削減効果も不揮発記憶を用いた PG に匹敵する。したがって、ULV リテンションは物理的には不揮発記憶ではないが、PG の電源遮断時にデータを保持できるので、PG の範囲内では実質的な不

揮発記憶とみなすことができる。本論文ではこのような ULV リテンション可能な SRAM のことを擬似不揮発性 SRAM (VNR-SRAM) と呼ぶ。

4.1.2 擬似不揮発性双安定記憶回路の構成方法

次に、不揮発性メモリ素子を用いずに CMOS のみで実質的に不揮発記憶を実現する方法について考察する。ULV リテンションを実現するためには、この動作モードにおいて十分なノイズマージンを確保することが重要になる。図 4.1.2(a), (b) に通常のインバータを用いて双安定回路を構成したときのバタフライカーブとスタティックノイズマージン (SNM) を模式的に示す (第 2 章で述べたように SNM はローブに内包される正方形の一边で定義される)。これは 6T-SRAM, 8T-SRAM, 10T-SRAM 等 [3, 4] のリテンションモードの SNM に該当する。通常の双安定回路では同図 (b) のように V_{DD} の低下によりバタフライカーブのローブの広がりや縮小し、SNM が減少してしまう。これに対して、ニアスレッショルド電圧やサブスレッショルド電圧と呼ばれる低電圧領域 (トランジスタのしきい値近傍やそれ以下の電圧) においても十分な SNM を確保できるシュミットリガインバータ (ST インバータ) を用いた双安定回路が提案されている [5]。図 4.1.3(a), (b) に ST インバータの回路構成と、ST インバータで構成した双安定回路のバタフライカーブの一例を示す。ST インバータはメインインバータとフィードバックトランジスタ (FB トランジスタ) で構成され、FB トランジスタの効果によりヒステリシスを有する角形で急峻な伝達特性が得られる。この特性によって、ST インバータで構成した双安定回路では ULV においても広いローブの (SNM の大きな) バタフライカーブが得られる。図 4.1.3(b) は A 点 ($(V_Q, V_Q) = (0, V_{DD})$) の状態で表されるデータが保持されている場合を示しており、データが保持されている安定側 (動作側) のローブが大きく開いている。一方、B 点が動作点となる場合では B 点側のローブが大きく開く。したがって、いずれの場合でも SNM が飛躍的に増加する。ST インバータで構成した双安定回路は、その高いノイズ耐性から低電圧動作の SRAM などへの応用が検討されてきた [5-18]。以下、ST インバータで構成した SRAM を ST-SRAM、そのセルを ST セルと呼ぶことにする。ST インバータは低電圧動作時の安定性が改善されるが、これを通常電圧で動作させた場合にはリーク電力が大幅に増大してしまうという問題が生じる。したがって、ST-SRAM は低電圧動作専用の SRAM としては有用であるが、通常電圧の SRAM 動作と ULV リテンションを用いて PG を実現する場合には性能上の問題を引き起こす。すなわち、ST-SRAM は擬似不揮発

記憶に利用できるが、通常動作時の回路性能を劣化させるため、このままでは PG に活用できない。

この問題は ST セルの通常電圧の SRAM 動作時に FB トランジスタのバイアスを変化させることで解消できる。本論文では通常電圧の SRAM 動作時は従来のインバータと同等に動作し、ULV リテンション時のみ ST インバータとして動作するデュアルモードインバータ (DM インバータ) を提案した。図 4.1.4(a) に DM インバータの構成方法を示す。図中の CTRL ドライバによって、FB トランジスタのバイアスを切り替えることができる。ST インバータとして用いる場合は同図(b)のように、FB トランジスタのバイアスは pMOS 側が V_{GND} 、nMOS 側が V_{DD} に接続されており、従来の ST インバータと同様の動作を実現できる。これをシュミットトリガモード (ST mode) と呼ぶことにする。これに対して、同図(c)では FB トランジスタのバイアスを切り替え ST モードと逆転したバイアス構成となる (FB トランジスタのバイアスは pMOS 側が V_{DD} 、nMOS 側が V_{GND} に接続される)。この場合、pMOS の FB トランジスタがメインインバータのプルアップトランジスタ、nMOS の FB トランジスタがメインインバータのプルダウントランジスタとして動作し、通常のインバータと同様に動作する (FB トランジスタによるフィードバック効果が消失する)。この場合では、ST モードに比べてリーク電流を小さく抑えることができる。以下、この動作をブーステッドインバータモード (BI mode) と呼ぶことにする。以上のように DM インバータでは CTRL ドライバによって FB トランジスタのバイアスを切り替えることで、通常電圧の SRAM 動作と ULV リテンションのそれぞれに適した動作モードに切り替えることができる。したがって、DM インバータを用いて双安定回路を構成することで、PG に適した VNR-SRAM セルを実現できると考えられる。以下では、DM インバータを用いた VNR-SRAM セルとその課題を述べ、その解決策として新たな VNR-SRAM セルを提案する。

4.2 擬似不揮発性 SRAM の回路構成と動作原理

以下、VNR-SRAM セルを VNR セルと略記する。図 4.2.1 に図 4.1.4 に示した DM インバータを用いて構成した VNR セルを示す。上述のようにこの VNR セルでは通常の SRAM 動作 (通常電圧での SRAM 動作とこれより低い (ULV ではない) 電圧でのスリープ動作) は BI モード、ULV リテンションは ST モードで行う。この構成では CTRL ドライバの電源側の電圧を V_{FBH} (例えば 1.2V)、接地側の電圧を V_{FBL} (例えば 0V) とする。こ

のセルは 14 トランジスタで構成され、モード切替に CTRL ドライバが 2 つ必要になるため面積オーバーヘッドが非常に大きい。面積オーバーヘッドを最小限に抑えるためには、メインインバータのプルアップ側またはプルダウン側のどちらか一方にのみ FB トランジスタによるフィードバックを用いればよい。図 4.2.2 にメインインバータのプルダウン側のみ FB トランジスタを用いた DM インバータと、これを用いて構成した VNR セルを示す。以下では、この構成の VNR セルを nFBPD-VNR セル (VNR cell using nMOS feedback transistors for the pull-down side of its main inverter) と呼ぶことにする。また、図には示していないが、同様にメインインバータのプルアップ側のみ FB トランジスタを用いた構成を pFBPU-VNR セルと呼ぶことにする。nFBPD-VNR セルと pFBPU-VNR セルの FB トランジスタは、これらのもととなっているシュミットトリガインバータの構成から、それぞれ nMOS と pMOS である。以上の VNR セルは従来の ST-SRAM セル (ST セル) [5] に CTRL ドライバを付加した構成である。nFBPD-VNR セルおよび pFBPU-VNR セルでは FB トランジスタをメインインバータのプルダウン側またはプルアップ側のみとするためにフィードバック効果がバタフライカーブの一方の伝達特性にしか発現しないが、これによる SNM の減少分はトランジスタサイズを適切に設計することで解消することができる [19]。

VNR セルは ULV リテンション時に 6T セルに比べてノイズマージンを大きくすることができる。しかし、ULV の ST モードではメインインバータのプルダウン側では nMOS の FB トランジスタで接続ノード (図 4.2.2(b) の N) をプルアップ、メインインバータのプルアップ側では pMOS の FB トランジスタで接続ノードをプルダウンしているために、ULV ではフィードバック効果が弱められている可能性がある (これらの FB トランジスタ利用形態ではそれぞれのしきい値が接続ノードの充電/放電に影響する)。これを改善できれば ULV リテンション時のノイズマージンをさらに大きくできると考えられる。また、nFBPD-VNR セルおよび pFBPU-VNR セルではモードの切り替えに、CTRL ドライバを用いた制御と、CTRL ドライバによるセル面積とリーク電流のオーバーヘッドを生じている。CTRL ドライバを用いることなく、 V_{DD} の大きさのみでモード切替ができれば、これらの問題も解消できる。

図 4.2.3(a) に図 4.2.2(b) に示した nFBPD-VNR セルをループ型表示したものを示す。上述のように、この構成では FB トランジスタが nMOS で構成され、ULV リテンション時 (ST モード時) のフィードバック効果が弱い。また、モード切替に CTRL ドライバを用いて

いるため、制御性、リーク電流とセル面積のオーバーヘッドの課題がある。そこで、図 4.2.3(b)に示すように FB トランジスタを ULV であってもプルアップ特性の優れた pMOS で構成し、さらに FB トランジスタに常に一定のバイアスを加える構成を提案した。以下では、このセルを pFBPD-VNR セル (VNR cell using pMOS feedback transistors for the pull-down side of its main inverter) と呼ぶことにする。この pFBPD-VNR セルでは FB トランジスタを pMOS で構成し、そのゲート端子をもう一方の DM インバータの出力に接続する。また、 V_{FB} は ULV リテンション時の V_{DD} 程度 (通常電圧の SRAM 動作時の V_{DD} と比べて十分に低い電圧) に固定する。通常電圧の SRAM 動作時は $V_{DD} \approx V_{DD} - V_{FB}$ となるため、BI モードで動作し、ULV リテンション時は $V_{DD} \approx V_{FB}$ となるため自動的に ST モードに移行する。したがって、CTRL ドライバを用いることなく V_{DD} の変化のみによって、自動的にモード切替が可能となる。同様に、メインインバータのプルアップ側に nMOS の FB トランジスタを用いた構成を nFBPU-VNR セルと呼ぶ。近年のマイクロプロセッサ (μP) では、内部に複数の電源電圧を供給できる DC-DC コンバータが内蔵されており、その電圧値の種類は 10 種類以上にも及ぶ。したがって、これらのセルでは、 V_{FB} のバイアスをこのような電源から供給すれば、CTRL ドライバが不要であることから PGP は不要となる。

図 4.2.4(a)に通常電圧のリテンション時、ULV リテンション時における pFBPD-VNR セル、nFBPD-VNR セルを構成する DM インバータの伝達特性を示す。簡単のため、以後これらをそれぞれ pFB セル、nFB セルとさらに略記する。各セルはそれぞれ後述の最適設計を用い、通常電圧の SRAM 動作時の V_{DD} を 1.2V、ULV リテンション時の V_{DD} を 0.2V としている。また、リファレンスとして通常のインバータ (後述のリファレンス用の 6T セルを構成するインバータ) の伝達特性も示す。nFB セルは通常 SRAM 動作時に BI モードで動作させることで通常のインバータ (青の破線) とほぼ同等の (ヒステリシスの消失した) 伝達特性を示し、ST モードに切り替えることで急峻でヒステリシスのある伝達特性が発現する。この伝達特性は ULV リテンション時にも実現できるが、そのヒステリシス幅は狭くなる。これは FB トランジスタの接続ノードのプルアップに nMOS を用いており、ULV ではフィードバック効果が弱まるために生じる。また、この nFB セルのモード切替は、CTRL ドライバを用いて行っている。

pFB セルを構成する DM インバータでは V_{FB} を 0.2V に固定してある。このため $V_{DD}=1.2V$ のとき、FB トランジスタのバイアスは相対的に接地状態に近く、この DM イン

バータは BI モードで動作する。したがって、通常のインバータと同等の伝達特性を示す。 V_{DD} を 0.2V まで低下させると、 $V_{DD}=V_{FB}$ となることから、自動的に ST モードに移行し、急峻でヒステリシスのある伝達特性を示す。このヒステリシス幅は nFB セルの DM インバータの伝達特性よりもはるかに広い。これは FB トランジスタの接続ノードのプルアップに pMOS を用いているため、ULV でも十分なフィードバックの効果を生じるためである。図 4.2.4(b) に 6T セル、nFB セル、pFB セルの ULV リテンションにおけるバタフライカーブを示す。各セルは同図(a) に用いた各インバータの設計を用いて構成した。pFB セルは 6T セル、nFB セルに比べて動作点側のローブが大きく開き、大きな SNM が得られる。定量的な解析は次節で行う。

4.3 擬似不揮発性 SRAM の設計

4.3.1 パワースイッチの構成

本論文では主にヘッダパワースイッチアーキテクチャを用いるが、デュアルパワースイッチアーキテクチャについても 4.5 節で検討する。図 4.3.1 にヘッダパワースイッチアーキテクチャを用いた場合の pFB セルとパワースイッチの構成を示す。通常 SRAM 動作時の V_{DD} は V_{DDH} ($=1.2V$) の電源線から PS_1 , PS_2 を用いて生成した。 PS_1 は通常電圧における SRAM 動作に、 PS_2 はスリープに用いる。ULV リテンションの V_{DD} は V_{DDL} の電源線から PS_3 を用いて供給した。ただし、 PS_3 のボディの接続には注意が必要である。 V_{DDL} の電源線に接続すると通常 SRAM 動作時にしきい値を下げる方向にボディバイアスが生じてリーク電流が増加する。一方、 PS_3 のボディを V_{DDH} の電源線に接続すると ULV リテンション時にしきい値が大きく上昇し、SNM が減少してしまう。そこで、通常 SRAM 動作および ULV リテンションのそれぞれの場合で適切にボディバイアスが印可されるように PS_3 のボディを仮想電源線に接続した。これら PS_1 , PS_2 , PS_3 によって生成される V_{DD} は仮想電源線を通して各セルに供給される。セル内の各トランジスタの基板バイアスは、ULV リテンション時の安定性を保つため、pMOS については仮想電源線に接続し、nMOS については接地線に接続する(後述するようにデュアルパワースイッチアーキテクチャでは pMOS のボディを V_{DDH} の電源線に接続することができる)。表 4.3.1 にセルとパワースイッチのデバイスパラメータとバイアスの定義をまとめた。

4.3.2 セルの設計方法

本論文では, VNR-SRAM を検討するが, 上述のように, より優れた ULV リテンション特性の期待できる pFB セルを中心に議論し, 比較のため nFB セルや 6T セルも用いる. セルとパワースイッチは各動作モードにおける SNM を指標として設計する. 第 2 章で述べたようにグローバルばらつきの影響を考慮して設計を行い(各プロセスコーナーの SNM を指標にする), ローカルばらつきによる SNM の分布からリダダンシを評価する. セルの設計では ULV リテンションの SNM と通常電圧の SRAM 動作時の読み出し動作の SNM がワーストケースになり得るが, 読み出し動作の SNM はワードラインアンダードライブ[20]などのバイアスアシストにより容易に改善できるため設計の自由度が高い. そこで, ULV リテンションの SNM を指標としてトランジスタサイズを決定した(ただし, このように設計を行ってもバイアスアシストが必ず必要になるとは限らない).

図 4.3.2 に pFB セルの設計フローチャートを示す. はじめに, V_{DDL} を暫定値に設定し, パワースイッチとセルの設計を行う. 各パワースイッチ PS_1 , PS_2 , PS_3 のチャンネル幅 W_{PS1} , W_{PS2} , W_{PS3} は十分に大きなサイズにしておき, セルの各トランジスタのサイズは以下のように初期値を設定する. 65nm バルク MOSFET の SRAM の設計[21]を参考にして各トランジスタのチャンネル幅を, ロード(LD)トランジスタ: $W_{LD}=100\text{nm}$, ドライバ(DRV)トランジスタ: $W_{DRV}=150\text{ nm}$, パス(PSS)トランジスタ(M_{PSS}): $W_{PSS}=100\text{nm}$ と設定し, 面積オーバーヘッド削減のため FB トランジスタは最小サイズとした($W_{FB}=100\text{nm}$). 100nm は本論文で用いた PDK における最小チャンネル幅である. pFB セルは 6T セルに比べて DRV トランジスタが縦積みであるためその駆動力が落ちるが, この初期値設定では同じサイズとした(これも最適設計のセル面積をできるだけ小さくするための方策である. 最終的にすべての動作における SNM を満足するように調整するためこのように初期値を選んでも問題はない). そして, パワースイッチサイズを小さくしていき, この状態を保持できるパワースイッチのサイズを決定する. 次に, ULV リテンション時の各プロセスコーナーにおける SNM を設計指標として, DRV トランジスタと FB トランジスタのチャンネル幅をデバイスパラメータ α_{DRV} , α_{FB} ($W_{DRV}=150\alpha_{DRV}$, $W_{FB}=100\alpha_{FB}$)を用いて微調整することで, この段階における最適設計を得る(SNM を最大化できる α_{DRV} と α_{FB} を求める). セルのサイズを変更するとパワースイッチサイズの最適値が変動する可能性があるため, 以上のセルとパワースイッチの設計プロセスをセルフコンシステントになるまで繰り返す. ここで得られた最適設計における ULV リテンションの SNM が目標値に対して余裕があれば, V_{DDL}

を削減して再度、以上の最適化を行う。このようにして V_{DDL} とセルの最適設定の暫定値が決定される。次にこの最適設計の暫定値を用いて、各動作(スリープを除く)における SNM の評価を行う。すべての SNM が目標値を満足していれば、この設計が最適設計となる。満足していなければ、設計を変更する前に、問題となっている動作モードにバイアスアシスト等を用いて SNM の改善を行う(ULV リテンションを除くと読み出し動作がワーストケースになることから、バイアスアシストは有効である)。バイアスアシストを用いない場合は、ULV リテンションの SNM に余裕が生じるように V_{DDL} を増加させ、ULV リテンションの SNM の目標値を満たす範囲内で α_{DRV} と α_{FB} が問題の動作モードの SNM を満たすように再調整する。最後に、SNM のローカルばらつきによる影響を評価して、必要なリダンダンシの割合を求め、これが許容範囲内にあれば設計終了となる(プロセスによる歩留まりに比べて十分に小さければよい)。許容範囲内になれば、グローバルばらつきによる SNM にさらに余裕を持たせ、ローカルばらつきによるリダンダンシが許容範囲内に収まるまで以上の設計を繰り返す。スリープについては、最適設計されたセルを用いて、第3章で述べた6Tセルと同様にして、パワースイッチ、セル数、スリープ電圧を決定した。以下で議論するように pFB セルでは $V_{DDL}=0.2V$ で上記条件を満たす。nFB セルは比較のため $V_{DDL}=0.2V$ での最適設計を用いる(SNM が許容範囲内に収まるかどうかにかかわらず $V_{DDL}=0.2V$ での設計値を用いる)。また、リファレンスに用いる6Tセルについては第3章で用いた構成とした。以下、第2章で述べたように、各モードにおける SNM を SNM(mode) で表すことがある(mode=READ, WRITE, SB, SLP, ULVR)。

pFBセルは次段の出力のフィードバックを利用するため、SNMの評価には図4.3.3に示す回路を使用した(ゲート電圧のスweepには通常のDC解析ではなく、非常に長い時間をかけてスweepするTransient解析を用いた)。同図において、初段の入力 V_{in} を順スweepと逆スweepで入力し、出力 V_{out} を評価してバタフライカーブを作成する。この時、次段のDMインバータはフィードバックトランジスタのゲートをハイレベルに固定した状態で常に通常のインバータとして動作させる。これにより、入力 V_{in} の順スweep時には初段にフィードバックが正常にかかり、逆スweep時にはフィードバックトランジスタが遮断されるので、次段の影響は無視でき、ノイズマージンを正確に評価できる。

4.3.3 設計結果

以下ではpFBセルを中心に、nFBセルをリファレンスの1つとして議論する。上述のよ

うに、パワースイッチとセルはセルフコンシステントになるように最適設計される。以下では最終的な設計を用いて各設計パラメータの依存性から最適設計の様子を示す。pFBセルの最適設計の結果を表 4.3.2 に示す。SNM の目標値は第 2 章で述べたように 78mV である。

図 4.3.4(a)に pFB セルにおける各 SNM(mode)の W_{PS1} , W_{PS3} 依存性を示す。各モードは mode=SB, READ, WRITE, ULVR である。 W_{PS1} , W_{PS3} は 1 セルに対するパワースイッチのサイズである。また、評価に用いたプロセスコーナーは TT, FF, SS, FS, SF である(第 2 章参照)。セルは上述のように最適設計されたものを用いてある。ULVリテンションの電圧は $V_{DDL}=0.2V$ とした。通常電圧の SRAM 動作では、いくつかの動作モードで SNM は W_{PS1} に依存するが、十分なチャネル幅をとることで、どの SNM も概ね飽和する。ULVリテンションでは、SNM の W_{PS3} 依存性は小さい。SNM の飽和値からの変動の許容範囲を 10%程度としてパワースイッチサイズを決定すると PS₁ と PS₃ では $W_{PS1}=300nm$, $W_{PS3}=100nm$ となる。ただし、パワースイッチの最小設計サイズは 100nm である。

スリープモードでは第 3 章でも述べたようにパワースイッチを完全に導通し、1 つのパワースイッチで共有するセル数を調整することで目標とする SNM を満たすスリープ電圧 V_{SLP} を実現する。図 4.3.4(b)に pFB セルにおける SNM(SLP)と V_{SLP} のセル数 N_{SLP} 依存性を示す(W_{PS2} は 100nm としている)。pFB セルでは FB トランジスタに常に一定のバイアスを印加しているため、スリープモードにおいても ST モードのようなフィードバックの効果が弱く生じている。この影響によって、6T セルに比べて、より低電圧でスリープモードを実現できる可能性があるが、共有セル数が膨大になりすぎるため配線による副次的効果を考慮して、6T セルと同じセル共有数を用いてスリープモードを実現することとした。この時、 V_{SLP} も 6T セルと同じ 0.8V 程度となる。

図 4.3.5 に pFB セルにおける SNM(ULVR)の α_{DRV} , α_{FB} 依存性を示す。ULVリテンションの SNM は、 $\alpha_{DRV}<\sim 1.13$, $\alpha_{FB}>\sim 1.00$ で SF コーナーがワーストケースとなり、 $\alpha_{DRV}>\sim 1.13$, $\alpha_{FB}<\sim 1.00$ で FS コーナーがワーストケースとなる。そこで、両方の SNM が一致しワーストケースの-marginが最大となる $\alpha_{DRV}\sim 1.13$, $\alpha_{FB}\sim 1.00$ で設計した。この図では上述のようにセルフコンシステントに決定したセルとパワースイッチの構成を用いていることから、これらが最適設計値である。この時、第 2 章で述べたように ULVリテンション時の SNM 改善のために、ビット線は 0.2V に充電している。この改善効果の詳細は Appendix に記した。

図 4.3.6 に 6T セル, nFB セル, pFB セルの SNM(mode)の値をまとめた. 同図(a)は mode=SB, READ, WRITE, SLP, (b)は mode=ULVR の結果である. nFB セルは pFB セルと同様に設計し, $\alpha_{DRV}=0.93$, $\alpha_{FB}=1.00$ とした. この設計は $V_{DDL}=0.2V$ での最適設計の結果であり, pFB セルと同様にリードアシストを用いずに通常の読み出し動作時の SNMを確保できる(SNMの目標値を満足できなくても V_{DDL} の最適化は行っていない). 6Tセルについては, 設計の初期値に用いたリファレンスセルの構成をそのまま用いてある. スリープを含む通常の SRAM 動作では, すべてのセルで 78mV 以上の SNM を確保できている. ULV リテンションについては, 6T セルでは TT コーナーで 20mV 程度, ワorstケースとなる FS コーナーではほとんど SNM をとれず, 目標値を全く満足できない(ULV リテンション時, 6T セルは既存のスリープアーキテクチャにおいて, V_{PS2} のゲート電圧を下げることで 0.2V のデータ保持を行っている). nFBセルでは TT コーナーでは 78mV を実現できるものの FS および SF コーナーでは満足できない. 一方, pFB セルでは ULV リテンションでもすべてのプロセスコーナーで 78mV 以上の SNM を確保できる. これは pFB セルでは ULV においてもより強いフィードバック効果の ST モードを実現できるためである.

図 4.3.7 に $V_{DDL}=0.15V$, 0.2V を用いて最適設計された pFB セルの ULV リテンションにおける SNM(ULVR)を示す. $V_{DDL}=0.2V$ のとき, pFB セルは SNM の目標値を満たすが, これより V_{DDL} が下がると設計の最適化を行っても目標値を満足できない. したがって, V_{DDL} の最小値は 0.2V となる.

上述の設計において, 6T セル, nFB セル, pFB セルにおける SNM のローカルばらつきの影響について調べた. 図 4.3.8 に各セルにおける SNM(ULVR)のローカルばらつきによるモンテカルロシミュレーションの結果を示す. 同図には SNM(ULVR)のしきい値も点線で示している. 各トランジスタにおけるしきい値のローカルばらつきは第 2 章で述べた分布を用いた. 試行回数は 10000 回である. それぞれのセルにおける SNM の分布の中心は TT コーナーにおける SNM の値とほぼ一致している. nFB セルおよび pFB セルは ST モードにおけるフィードバック効果によって 6T セルに比べてばらつきの影響が抑制され, 分布の広がりが小さくなる. pFB セルが nFB セルに比べて分散が大きいのは, FB トランジスタと DRV トランジスタの極性が異なるため, ばらつきの影響がより大きく出るためである. 以上の結果から, 第 2 章で示した方法を用いて必要となるリダンダンシを評価すると, pFB セルでは 0.1%程度以下となり一般的なチップの歩留まりに比べて十

分小さく、通常の冗長セルの範囲内で実現が可能となる。一方で、nFB セルでは 100% 以上、6T セルではさらに多量の冗長セルが必要となる。以上から、 $V_{DDL}=0.2V$ の ULV リテンションは、pFB セルのみで実現可能となる。ただし、本論文で目標値とした 78mV は厳しい条件であるため、目標値を $kt(=26mV)$ 、 $2kt(=52mV)$ と設定した場合、nFB セルでも十分にデータ保持できる可能性がある。

図 4.3.9(a)に pFB セル、(b)に nFB セルの最適設計におけるレイアウトを示す。ともに $V_{ULV}=0.2V$ における最適設計を用いた。pFB セルでは LD トランジスタと FB トランジスタを同じウェル (n-well) 上に平行に配置することができるため集積密度が高い。一方、nFB セルでは FB トランジスタが nMOS であるため LD トランジスタとウェルを切り離す必要があり集積密度が低下する。このため、nFB セルでは面積オーバーヘッドが pFB セルに比べて増加する (nFB セルではセル面積削減のためトランジスタの配置を工夫しているが、どのように配置しても pFB セルよりも増大する)。また、CTRL ドライバを用いる場合はさらに面積オーバーヘッドが増える。6T セルに対する面積オーバーヘッドは pFB セルでは 1.77 倍、nFB セルでは 1.87 倍である。このオーバーヘッドはデュアルポート SRAM と比べると少し大きい。従来技術では到達できない低消費電力化が実現できるためメリットが大きい。また、SRAM の面積削減技術には、6T セルの LD トランジスタを薄膜トランジスタ (TFT) としてセル上部の配線層につくりこむ技術がすでに提案されている [23]。この技術を適用できれば、FB トランジスタを同様の TFT とすることで、pFB セルおよび nFB セルのセル面積は、いずれもデュアルポート SRAM 程度 (6T セルの 1.5 倍程度) まで削減できる。一方、6T セルの面積を pFB セルや nFB セルと同じ面積にまで拡大した場合には、面積の増加によりばらつきが低減され、SNM の分散を抑えられるため、スリープ時の V_{DD} を多少下げることができる。しかし、この場合でも SNM の分散の中心値 (TT コーナーにおける SNM 値) はインバータの特性から決まってしまう、セル面積にほとんど依存しないことから、6T セルの面積を pFB セル程度に増加させてもスリープ時の V_{DD} が約 0.1V 低減する程度の効果しかなく、pFB セルのように 0.2V のような超低電圧でのリテンションは実現できない。SNM の分散における中心値を増大するためには上述の ST モードが必須となる。

次に、pFB セルの読み出し速度について示す。図 4.3.10(a)、(b)にそれぞれ (V_Q, V_{DD})、($V_{DD}, 0$) の場合における読み出し動作時の V_Q 、および V_{Q0} の波形を示す。クロックは 1GHz である。実線が pFB セル、点線が 6T セルの波形を示している。各図の上

から、記憶ノードの電位 (V_Q , $V_{\bar{Q}}$), ビット線の電位 (V_{BL} , $V_{\bar{BL}}$), センスアンプの出力 (V_{SAO} , $V_{\bar{SAO}}$)を示してある. pFB セルは 6T セルに比べて DRV トランジスタが縦積みになっているため、ビット線の放電速度には差が生じる. しかし、この影響は小さく、例えばリードイネーブル信号の入力時点 (図 4.3.10(a)の 1.62ns)でのビット線電位の差分は電源電圧の 1%程度である. そのため、センスアンプの出力でみると同じ読み出し速度が実現できている (図 4.3.10(b)も同様). このように 6T セルが十分に機能する (6T セルで設計された) センスアンプを用いれば、pFB セルも同じクロック周波数で読み出すことができる (通常センスアンプの出力変化はクロックより十分に早く応答するように設計されるため). すなわち、pFB セルの動作速度の劣化は GHz 動作に影響しない. 読み出し動作時電力については、図 4.3.10(a)の場合では、6T セルで 44.7 μ W, pFB セル 40.0 μ W となる.

次に、pFB セルの書き込み速度と電力について議論する. 図 4.3.11(a)に (V_Q , $V_{\bar{Q}}$)=(V_{DD} , 0)から (V_Q , $V_{\bar{Q}}$)=(0, V_{DD})に書き込みを行ったときの波形を示す. クロックは 1GHz である. 各図の上から、記憶ノードの電位 (V_Q , $V_{\bar{Q}}$), ビット線の電位 (V_{BL} , $V_{\bar{BL}}$), クロック信号 (V_C , $V_{\bar{C}}$)を示している. pFB セルはトランジスタ数が増加するため、セルの容量が増大し 6T セルに比べて若干反転に時間がかかる. 例えば、ビット線電位が $V_{DD}/2$ (=0.6V)に変化する時刻を比べると、Q では 32ps の、 \bar{Q} では 41ps の差が生じる. しかし、これはクロック周期 (1ns) の 4%程度であり、クロックスピードに全く影響は与えない (書き込み時のクロックオーバーヘッドを生じず、6T と同じクロックで書ける). 図 4.3.11(b)にクロックを 2GHz に増加させたときの波形を示す. このとき、この高速動作を実現するため、書き込みドライバを上述の 1GHz の場合から倍のサイズとしている. こちらも(a)と同様にビット線電位が $V_{DD}/2$ (=0.6V)に変化するところで比較すると、Q では 13ps の、 \bar{Q} では 19ps の差が生じる. しかし、この場合でもクロック周期 (500ps) に影響はない. 最近の μ P のクロック周波数は数 GHz であって、これは μ P におけるクロックの事実上の上限であることから (これ以上のクロックの向上は望めない), 性能の向上はマルチコア化することで実現している. よって、数 GHz 程度でクロックオーバーヘッドを生じることなく書き込みができれば応用上の速度劣化はない. 以上から書き込みについても 6T セルで十分に動作させるように設計された周辺回路を用いれば、pFB セルを用いても動作速度上の問題は生じない. 書き込み動作時電力については、6T セルで 74.3 μ W, pFB セルで 75.5 μ W となり、pFB セルでは約 1.6%程度の増加がある. しかし、この程度の増加であれ

ば、アレイを構成したときに周辺回路の影響に隠蔽される、または周辺回路の工夫で削減することができる。

図 4.3.12 に pFB セルの動作波形を示す。2.4 節で示した周辺回路構成を用いてシミュレーションを行った。パワースイッチの制御信号 (V_{PG1} , V_{PG3}) に従って V_{DD} が変化する。通常電圧の SRAM 動作では V_{DD} が V_{FB} (図中の点線) より十分大きく、ULV リテンションでは V_{DD} と同電位となるため、自動的にモード切り替えを実現している。動作モードの切り替えにより記憶ノードの電位 (V_Q , V_Q) はどちらの安定点のデータであっても正常に電源遮断/復帰が実現できていることが確認できる。ULV リテンションにおける待機時消費電力は、通常の SRAM 動作時に比べて大幅に削減される。電源遮断/復帰動作時にはラッシュカレントが生じ、これが pFB セルにおけるエネルギーオーバーヘッドとなる。

4.4 電力・エネルギー性能の評価

図 4.4.1(a) に pFB セルおよび nFB セルの最適設計におけるリーク電力を示す。nFB セルは通常電圧の SRAM 動作時に ST モードを用いた場合大きな電力を消費するが、BI モードに切り替えることで電力を半分程度まで削減することができる。pFB セルでは通常電圧のスタンバイ時には自動的に BI モードとして動作し、さらに、 V_{FB} 線が定電圧でバイアスされていること、および基板バイアス効果によって、FB トランジスタを流れるリーク電流が削減され、nFB セルに対してさらに半分程度に電力が削減されている。また、縦積みされた DRV トランジスタもリーク削減に効果がある。これにより pFB セルの通常電圧のスタンバイ時の余剰リーク電力は 6T セルのリーク電力以下まで削減される。

図 4.4.1(b) に 6T セルの通常電圧のスタンバイ電力を基準とした 6T セルのスリープ、nFB セルおよび pFB セルの ULV リテンション、NV セルの電源遮断におけるリーク電力削減率 (第 2 章の r_{PG}) を示す。6T セル、NV セルは第 3 章で示したように、41%程度、99%程度の削減率となる。nFB セルは ULV リテンションによって 91%程度、pFB セルは 94%程度まで電力を削減でき、NV セルに匹敵するリーク電力の削減率となる。以上の議論における pFB セルは 78mV を SNM の下限として最適設計を行ったセルの評価結果であるが、第 2 章で示した基準のうち 52mV (=2kt) を下限として設計すると、pFB セルはさらに低電圧 (0.15V 程度) でも ULV リテンションが可能となり、96%まで待機時電力を

削減できる。

次に、上述の $V_{DDL}=0.2V$ において最適設計された pFB セルおよび nFB セルで構成した VNR-SRAM アレイの T_{BE} を検証する。図 4.4.2(a), (b)にアレイの電源遮断移行 (ENT) モードおよび電源復帰移行 (EXT) モードにおける各サブアレイの状態を示す。図中の各正方形はメモリアレイ内のサブアレイを示している。また、各サブアレイには動作モード (ENT, EXT, ULVR, SB) も付記してある。VNR-SRAM ではサブアレイごとに電源遮断/復帰の制御を行う (パワースイッチはサブアレイ内の適切なサイズのブロックごとに配置されるが、同一のサブアレイ内ではこれらのパワースイッチは同時に制御される (各ブロックからなるサブアレイがパワードメインとなる))。pFB セルは ENT/EXT 動作時にパワースイッチの制御のみ行うが、nFB セルはパワースイッチの制御に加えて、動作モード (BI モード, ST モード) の切り替えを行い、これにかかるエネルギーオーバーヘッドも考慮している。ここではサブアレイの電源遮断/復帰 (つまりパワースイッチの導通/遮断) によって生じる突入電流 (ラッシュカレント) が、第 2 章で述べた上層配線の許容電流値を超えないようにパワースイッチの制御シーケンスを決定した。図 4.4.3(a)に ENT/EXT 動作時における VNR-SRAM のパワースイッチの制御シーケンスを示す。ENT 動作では図 4.3.1 の PS_1 を遮断してから PS_3 を導通することで、パワードメイン (各ブロック) の放電にともなうラッシュカレントが V_{DDL} 線に流れる。このとき、 PS_1 から PS_3 へのダイレクトな電流パスが生じないように PS_1 を完全に遮断してから PS_3 を導通する。一方、EXT 動作では PS_3 を遮断してから PS_1 を導通し、パワードメインの充電にともなうラッシュカレントが V_{DDH} 線から流れる。これらのラッシュカレントを抑えるようにパワースイッチの制御を適切に行う必要がある。 V_{PG1} と V_{PG3} のスイープ時間を制御することで、ラッシュカレントを制御する。これらのスイープ時間をそれぞれ τ_{PG1} および τ_{PG3} とする。通常の PG では ENT 動作時にはラッシュカレントは生じないため、EXT 動作時のパワードメインの充電にともなう突入電流をラッシュカレントと呼ぶ。しかし、本論文の VNR-SRAM は PS_3 を使用して ULV リテンション時の V_{DD} を生成したことで、ENT 動作時にも電流が流れる。そこで、本論文では ENT/EXT 動作時の充放電により生じる電流をいずれもラッシュカレントと呼ぶ。図 4.4.3(b)に pFB セルを用いた VNR-SRAM の ENT/EXT 動作におけるラッシュカレントの最大値 (I_{rush}^{max}) および ENT/EXT 動作にかかるレイテンシである τ_{ENT1} , τ_{EXT1} のパワースイッチのスイープ時間 (τ_{PG1} , τ_{PG3}) 依存性を示す (τ_{ENT1} , τ_{EXT1} はパワースイッチの制御開始からラッシュカレントが十分に減衰するまでとして定義した)。いずれも

第 2 章の解析モデルで評価し、サブアレイやメモリアレイのラッシュカレント、レイテンシはこの解析モデルをアレイサイズに応じて拡張し評価した。パワースイッチのスweep時間を短くすると急激なパワードメインの充放電のために大きなラッシュカレントが生じるが、スweep時間を長くすることでラッシュカレントを低く抑えることができる。しかし、長すぎるとレイテンシが大きく増大してしまう。このため、pFB セルの ENT/EXT 動作におけるスweep時間は I_{rush}^{max} が上層配線の許容値を満たす最短の値である~20n および~100ns とした(nFB セルも同様に動作させることで十分にラッシュカレントを抑えることができる)。なお、ENT 動作時のラッシュカレントは EXT 動作時よりも小さく抑えられている。これはパワードメインの放電時に PS₃ のゲート-ソース間電圧 ($V_{PG3}-V_{DD}$) が減少し、自動的に PS₃ の駆動力(すなわち放電のスピード)が落ちるためである。一方、EXT 動作時は PS₁ のゲート-ソース間電圧 ($V_{PG1}-V_{DDH}$) が V_{DD} に依存せず一定であるため充電が速く、パワースイッチのスweep時間をより長くとする必要がある。

図 4.4.4(a)に 32kB, 256kB, 2MB の VNR-SRAM アレイにおける T_{BE}^{EE} の評価結果を示す(pFB セルは比較用の 6T セルよりもリーク電力が小さいため、 η は負の値をとる)。比較のために第 3 章で評価した NV-SRAM の T_{BE}^{EE} も示している(ここでは、NV-SRAM は $R_{skip}=0\%$ とした)。nFB セルで構成した VNR-SRAM では、仮想電源線の充放電時のレイテンシは上述の pFB セルで構成した VNR-SRAM と同様に評価し、モード切替のレイテンシは CTRL ドライバの出力の充放電時間で決定し、BI モードから ST モード、ST モードから BI モードの切り替えをいずれも 15ns とした。pFB セルおよび nFB セルで構成した VNR-SRAM はいずれも不揮発性メモリ素子を用いた書き込み動作を必要としないため、NV-SRAM と比較して 2 桁程度以上短い T_{BE}^{EE} を実現できている。また、nFB セルで構成した VNR-SRAM は pFB セルで構成した VNR-SRAM に比べて T_{BE}^{EE} の若干の増加がみられるが、これはセルのリーク電力が pFB セルに比べて増加していることとモード切替時の CTRL ドライバの切り替えの余剰エネルギーのためである。図 4.4.4(b)に VNR-SRAM アレイの電力を示す。白抜きの棒グラフは 6T-SRAM アレイにおける通常電圧のスタンバイ電力を示している。6T-SRAM に比べて VNR-SRAM はいずれも電源遮断と同程度の削減率を実現できる。

4.5 リーク電力削減アーキテクチャ

VNR-SRAM のエネルギー性能をさらに高めるデュアルパワースイッチを用いた低消費電力技術について検討する。また、第 3 章で説明した無用データのフラッシュアーキテクチャは VNR-SRAM にも適用可能であるので、その効果についても検証する。

4.5.1 デュアルパワースイッチを用いた低消費電力化技術

これまでヘッダパワースイッチを用いた pFB セルについて議論を行った。本節ではデュアルパワースイッチを用いて ULV リテンション時のリーク電力をさらに削減する方法を述べる。ここでは基板バイアス効果を積極的に活用する。

図 4.5.1(a)は図 4.3.1 に示したヘッダパワースイッチ構成の pFB セルについて、各トランジスタのボディの接続を明記したものである。セルの pMOS と nMOS のボディはそれぞれ V_{DD} と V_{GND} に接続されている。したがって、pMOS に基板バイアス効果は生じない。一般のヘッダパワースイッチ構成の PG では、pMOS のボディを V_{DD} ではなく V_{DD} 側に接続して、基板バイアス効果によって電源遮断時のリークを削減することもよく用いられている。しかし、ヘッダパワースイッチ構成の pFB セルでは上述のように pMOS のボディは V_{DD} に接続している。これは以下の理由による。pFB セルの pMOS のボディを V_{DDH} 線に接続すると、ULV リテンション時に pMOS に実効的に大きな基板バイアスが印可され、しきい値が増大するため電力削減効果は増大する。しかし、このとき pMOS と nMOS の電流駆動能力のバランスが崩れ、ULV リテンション時のバタフライカーブが大きく歪むことから、SNM が劣化してしまう(図 4.5.1(b))。したがって、ヘッダパワースイッチ構成の pFB セルでは、既存の電源線を利用して基板バイアス効果によって電力削減効率を高めることが難しい(基板バイアスを用いるためには外部からバイアスを別途用意しなければならない)。しかし、pMOS と nMOS の電流駆動能力のバランスを確保しながら、基板バイアスを印可できれば、SNM を十分に保ちながら、電力削減効果をさらに向上させることができると予想される。そこで、図 4.5.2(a)に示すヘッダパワースイッチとフッタパワースイッチを両方配置したデュアルパワースイッチ構成を検討した。この構成では pMOS のボディは V_{DDH} 、nMOS のボディは V_{SSL} に接続している。ULV リテンション時に V_{DD} と V_{GND} に供給される V_{DDL} と V_{SSH} を適切に選ぶことによって、ULV リテンションでも SNM を確保し、しかもどちらのトランジスタにも有効に基板バイアス効果を発現さ

せることができる(図 4.5.2(b)). 図 4.5.3 に pFB セルにおける SNM(ULVR)の V_{GNDH} 依存性を示す. この図ではセルに 0.2V の電圧が加わるように, V_{DDL} は $V_{\text{DDL}} = V_{\text{GNDH}} + 0.2\text{V}$ とした. また, セルはこの図の SNM の極大値近傍で SNM が最大化するように最適化してある ($\alpha_{\text{DRV}}=1.03$, $\alpha_{\text{FB}}=1$). また, 同図には ULV リテンション時のリーク電流も示してある. 同図から $V_{\text{GNDH}}=0.5\text{V}$ のとき, SNM(ULVR) が最大付近, リーク電流を最小に抑えることができる. また, 同図に点線で示したヘッダパワースイッチ構成の pFB セルの SNM(ULVR) とリーク電流と比べると, デュアルパワースイッチ構成ではヘッダパワースイッチ構成と同程度の SNM(ULVR) を確保し, リーク電流はさらに削減できていることがわかる. これは, pMOS と nMOS の両方に効果的に基板バイアスがかかることによる.

図 4.5.4 に 6T セルのスリープ時, ヘッダパワースイッチ構成の pFB セルの ULV リテンション時, デュアルパワースイッチ構成の pFB セルの ULV リテンション時における γ_{PG} を示す. デュアルパワースイッチ構成ではヘッダパワースイッチ構成の 94.4% から 99.1% 程度まで削減率が向上している. デュアルパワースイッチでは, パワースイッチが増加することによる面積オーバーヘッドがあるが, 後述の図 4.6.3 の 8kB の pFB セルで構成した VNR-SRAM マクロにおけるパワースイッチの占有面積は全体の 4.3% 程度であることから, デュアルパワースイッチ構成ではこの 2 倍程度の 8.6% 程度となることが予想でき, それほど大きなオーバーヘッドは生じない. 図 4.5.5 にヘッダパワースイッチ構成とデュアルパワースイッチ構成の pFB セルで構成した 32kB, 256kB, 2MB の VNR-SRAM アレイの $T_{\text{BE}}^{\text{EE}}$ を示す. デュアルパワースイッチ構成の場合でも, ヘッダパワースイッチ構成の場合と同程度の $T_{\text{BE}}^{\text{EE}}$ が実現できる.

4.5.2 無用データのフラッシュアーキテクチャ

第 3 章で示した PUDF アーキテクチャは VNR-SRAM にも応用できる. 無用データと判定されたブロックを積極的に電源遮断して, 残りを ULV リテンションすればよい. NV-SRAM の場合と同様に PMU に UDF 用のレジスタファイルを設けることで実装できる.

VNR-SRAM では図 4.3.1 に示した PS_1 と PS_3 の両方を遮断することで, 完全な電源遮断も実現できる ($\text{mode}=\text{SD}\neq\text{ULVR}$). 第 2 章で示した設計では電源遮断時は $V_{\text{DD}}=0.02\text{V} (\ll 0.2\text{V})$ となる. この場合, VNR-SRAM アレイの電源遮断時のリーク電力

は ULV リテンション時と比べてさらに削減される。ULV リテンション時に、UDF で指定されたブロックについては完全に電源遮断できるためリーク電力をより多く削減できる。また、電源遮断する場合、ENT 動作時に PS_3 を導通する必要がないため、ラッシュカレントが生じない。その結果、ENT 動作時のエネルギーとレイテンシも削減できる。ただし、EXT 動作時については、ULV リテンションした場合と電源遮断した場合ではいずれも PS_1 を導通するため同程度のエネルギーが消費される。

図 4.5.6(a)-(c)に 32kB, 256kB, 2MB の VNR-SRAM アレイにおける T_{BE}^{EE} の R_{PUDF} 依存性を示す。最適設計されたヘッダパワースイッチ構成の pFB セルで構成した VNR-SRAM を用いた。32kB では ULVR から SB へ復帰するよりも SD から SB へ復帰する場合に若干大きいエネルギーを生じるために微増し、2MB では T_{BE}^{EE} は R_{PUDF} の増加にともない ENT 動作時のエネルギーとレイテンシが削減されるため若干の減少がみられるものの、すべてのアレイサイズにおいてほぼ一定となる。いずれのアレイサイズにおいても、極めて短い T_{BE}^{EE} を実現できている。また、図 4.5.6(d)-(f)に各 VNR-SRAM における ENT/EXT 動作に要するレイテンシを示すが、こちらも T_{BE}^{EE} と同様におおよそ一定となる。一方、VNR-SRAM アレイのリーク電力はアレイサイズによらず、 R_{PUDF} 増加に伴い大きく向上する。図 4.5.7 に γ_{PG} の R_{PUDF} 依存性を示す。VNR-SRAM のメモリアレイをすべて ULV リテンションした場合の電力 ($R_{PUDF}=0\%$ の場合) に比べて R_{PUDF} が増加すると、完全に電源遮断できるセルが増加するため電力が効果的に削減される。例えば R_{PUDF} が 50% の場合、6T-SRAM のスタンバイ電力からの電力削減率を 98% 程度まで向上させることができる。

4.6 擬似不揮発性 SRAM マクロの電力・エネルギー性能

本論文では第 2 章で示した解析モデルでの評価からアレイのエネルギー性能を算出したが、この方法ではアレイ内のすべての寄生抵抗・寄生容量成分を評価できないため、アレイ全体をレイアウトして評価した結果とは異なる可能性がある。そこで、VNR-SRAM の 8kB マクロの設計およびレイアウトを行い、そこから寄生抵抗・寄生容量を抽出して、高速 SPICE による大規模シミュレーションにより性能の評価・検証を行った。図 4.6.1 にこのマクロの設計フローと使用した設計ツールを示す。セルアレイは上述の最適設計された pFB セルで構成した。周辺回路はプリチャージ、センスアンプ、ライトドライバ、セレ

クタは HSPICE を用いてトランジスタレベルで設計レイアウトし、デコーダと制御部は Verilog で RTL を記述し、スタンダードセルを用いて論理合成した。その後、各回路ブロックを自動配置配線した。最後に作成したマクロの機能検証を行い、正常に動作しない場合は周辺回路を正常に動作するようになるまで再設計した。

図 4.6.2 に設計した 8kB マクロのフロアプランを示す。セルアレイは 128bit×128bit からなるブロック 4 つで構成し、周辺回路は図のように配置した。ワードデコーダは図の左右のブロックで共有し、センスアンプとライトドライバは図の上下のブロックで共有している。図 4.6.3 に設計後のレイアウトを示す。リファレンスに用いる 6T-SRAM は簡単のため、上述の VNR-SRAM マクロのセルを 6T セルに置き換えることで構成した。したがって、周辺回路は VNR-SRAM と同じ構成である。

図 4.6.4 に上述の VNR-SRAM マクロの動作波形を示す。シミュレーションには高速 SPICE (Synopsys 社の CustomSim) を用いた[24]。上からそれぞれパワースイッチの制御信号 (V_{PG1} , V_{PG3}), V_{DD} , 記憶ノードの電位 (V_Q , $V_{\bar{Q}}$), 消費電力を示している。いずれも ULV リテンション後に ULV リテンション前の状態を正常に復元することができている。

図 4.6.5 に上述の VNR-SRAM マクロを解析した結果を示す。また、第 2 章で示した解析モデルでの評価結果も示す。図 4.6.5(a)は抽出した単体セルのリーク電力、(b)と(c)はそれぞれ 32kB, 256kB, 2MB のアレイの T_{BE}^{EE} とリーク電力の評価結果を示す。(c)については、比較用の 6T-SRAM のスタンバイ時のリーク電力を白抜きの棒グラフで示している。図 4.6.5 から、解析モデルによる評価はマクロのポストレイアウト解析と比べて若干の違いはあるが、よく一致している。

以上から VNR-SRAM の動作が、マクロレベルでも正常に行われることを確認した。また、マクロレベルのエネルギー性能評価結果が、第 2 章で示した解析モデルを用いた評価結果とよく一致することを示した。

4.7 先端 CMOS を用いた性能評価

本節では、上述の pFB セルを 45nm CMOS, 15nm FinFET で構成した場合の性能評価を行う。PDK は NANGATE の 45nm CMOS プロセスと 15nm FinFET プロセスを用いた[25, 26]。図 4.7.1(a)-(c)に 65nm CMOS, 45nm CMOS, 15nm FinFET における、pFB

セルの ULV リテンションの SNM とリーク電力削減率(γ_{PG})の評価結果を示す。65nm については先ほどの評価結果を示しており、4.3 節で示した設計方法を用いて最適化した後、レイアウトして寄生抵抗・寄生容量を抽出した pFB セルと 6T セルのネットリストを用いて評価した。一方 45nm, 15nm の評価結果については、pFB セルは最適な設計ではあるが、寄生抵抗・寄生容量の抽出は行わずに評価した。6T セルについては 45nm では 65nm のロード、ドライバ、パストランジスタのチャンネル幅比を用いて、ロードトランジスタとパストランジスタを最小チャンネル幅で構成した。パワースイッチのチャンネル幅は 90nm (最小チャンネル幅)とした。15nm については、Fin 数をロードトランジスタ、パストランジスタは 1 本で、ドライバトランジスタは読み出しの SNM を確保するため 2 本で構成した(比較用の 6T セルも同様の構成)。パワースイッチは 1 本とした。また、65nm におけるデバイスの 3σ のしきい値ばらつきは 10%程度、45nm では 15%程度であり、15nm は PDK にパラメータセットが用意されていないため、報告値を参考にして 15%を仮定した[27]。図 4.7.1(a)-(c)はいずれも 6T セル 0.2V で 78mV 程度以上の SNM を実現しており、 γ_{PG} は 45nm で 95%程度、15nm で 93%程度が実現できる。以上から先端 CMOS においても擬似不揮発性 SRAM 技術は有効である。特に 15nm は駆動能力が高く、各 SNM が 65nm, 45nm に比べて大きく向上している。以上の結果から、VNR-SRAM は最先端 CMOS デバイスでも有効であると結論できる。また、しきい値の 3σ ばらつきが 15%程度以下なら非シリコンチャンネルの先端 CMOS でも同様に有効である可能性がある。したがって、本技術は CMOS デバイスの数種に依らず、将来にわたって活用できると考えられる。

4.8 まとめ

本章の内容を以下にまとめる。

- VNR-SRAM を提案し、この設計方法や電力・エネルギー性能を初めて明らかにした。
- nFB セルを超低電圧における安定性を大幅に改善した pFB セルを開発した。
- pFB セルの設計方法を確立し、最適設計した pFB セルが $V_{DD}=0.2V$ において高い SNM と電源遮断に匹敵する 94%程度の電力削減率を実現できることを示した。
- VNR-SRAM アレイの T_{BE} と電力削減率を評価し、NV-SRAM に比べて T_{BE} が 2 桁程度削減され、同程度の電力削減率を実現できることを示した。

- VNR-SRAMセルの電力削減率のさらなる向上のため、デュアルパワースイッチ構成を提案し、99%程度まで電力削減率を高めることに成功した。
 - VNR-SRAMにおける無用データのフラッシュアーキテクチャの効果を検証し、電力削減率をさらに向上できることを示した。
 - 8kBのVNR-SRAMマクロを設計・レイアウトし、そのポストレイアウト解析により事実上の最小値である短い数 μs 程度の T_{BE} と90%以上の電力削減率を実現できることを示した。また、この結果が第2章の解析モデルの結果とよく一致することを確認した。
 - 先端CMOSを用いた場合のpFBセルの評価を行い、擬似不揮発性SRAM技術が有用となる可能性があることを示した。
- 以上から、擬似不揮発記憶を用いたSRAM技術を確立した。

4.9 Appendix

ULVリテンション時のビット線の充電によるSNMの改善

本章で述べたpFBセルおよびnFBセルはULVリテンション時にビット線を $V_{\text{DDL}} (=0.2\text{V})$ に充電することでSNMを改善させている。ここでは、この改善効果について述べる。図4.9.1にビット線電位(V_{BL})を0Vと0.2Vにした場合のバタフライカーブをそれぞれ示す。それぞれ本章の設計方法における最適設計を用いている。点線の $V_{\text{BL}}=0\text{V}$ の場合には V_{in} を増加させるにつれ、ハイレベルノードの電位が劣化しバタフライカーブの急峻性が失われるためにSNMが損なわれている。これは、双安定回路のロードトランジスタの駆動力が低下し、パストランジスタのリークによりドライバトランジスタが導通して放電される前に記憶ノードが放電され始めてしまうことによる(通常電圧ではロードトランジスタのオン電流がリーク電流に比べて非常に大きいためこのような現象は生じない)。そのため、超低電圧の安定したデータ保持のためにはハイレベルノードのパストランジスタのリーク電流を下げるのが重要となる。 $V_{\text{BL}}=0.2\text{V}$ とした場合、パストランジスタのソース/ドレインの電圧はいずれも0.2Vとなり、 V_{GS} は負バイアスとなる。その結果、リーク電流を大幅に削減することができ、バタフライカーブの急峻性が改善し、SNMが増加する。

4.10 図と表

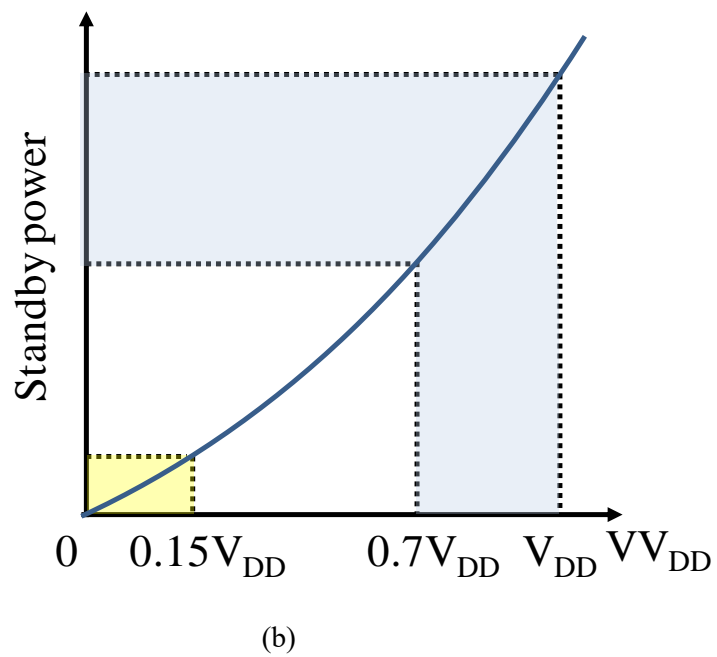
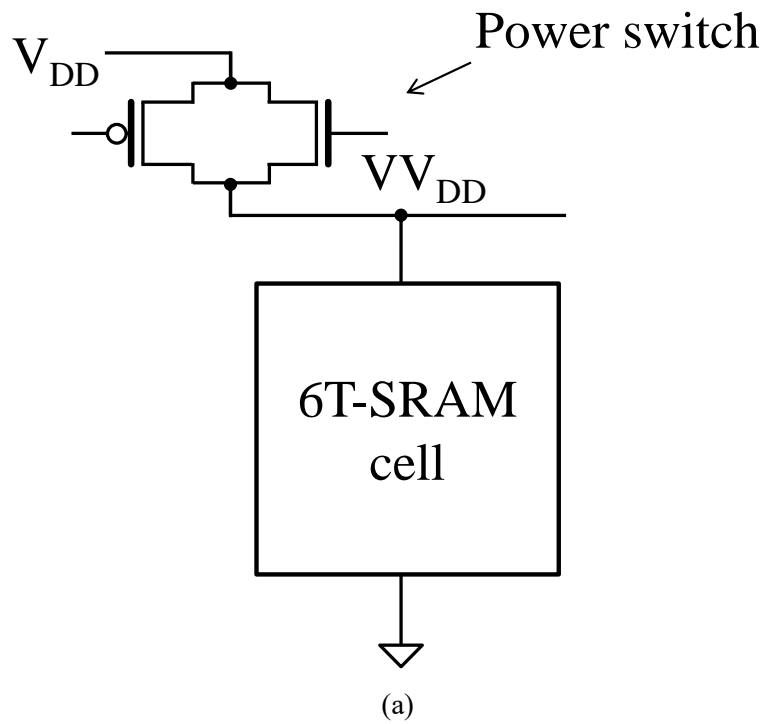
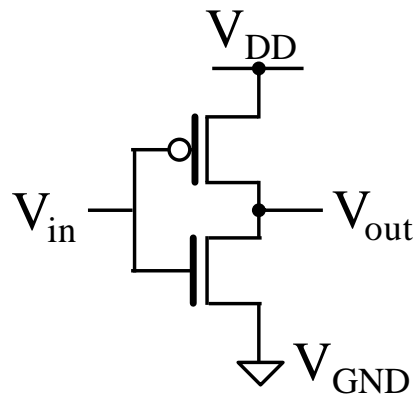
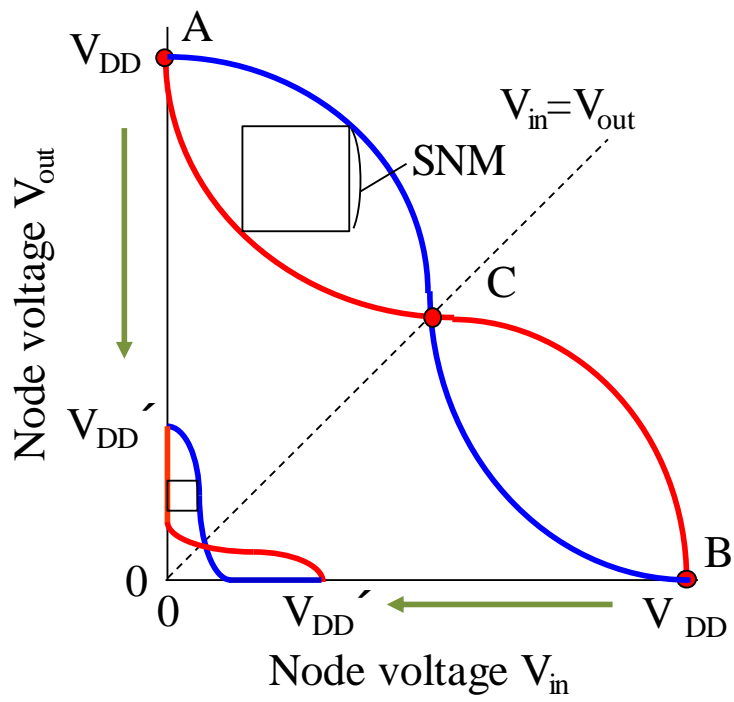


図 4.1.1 (a)6T セルとパワースイッチの構成と(b)待機時電力の V_{DD} 依存性の模式図

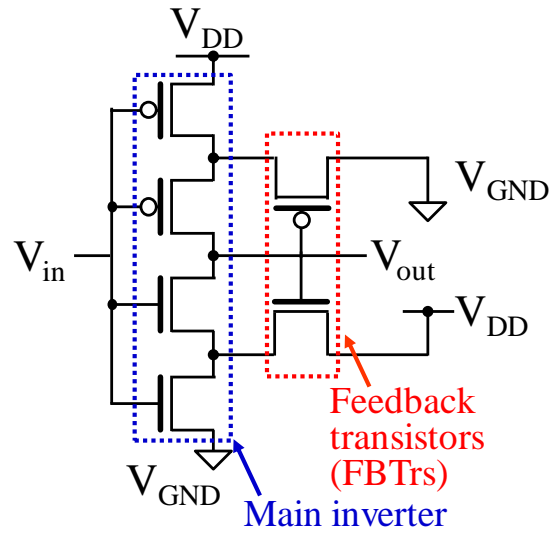


(a)

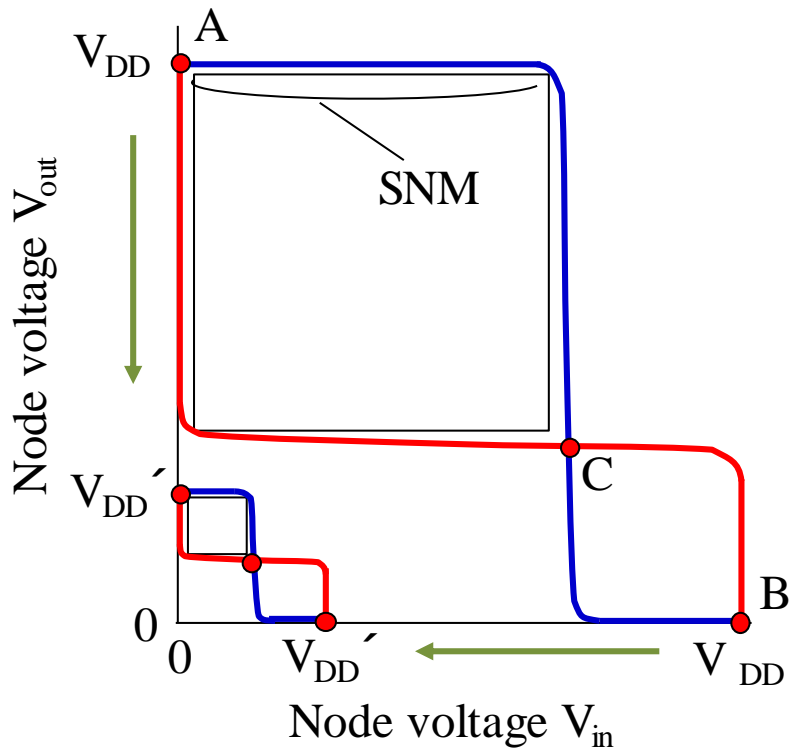


(b)

図 4.1.2 (a)通常のインバータの構成と(b)双安定回路におけるバタフライカーブ



(a)



(b)

図 4.1.3 (a)ST インバータの構成と(b) 双安定回路におけるバタフライカーブ

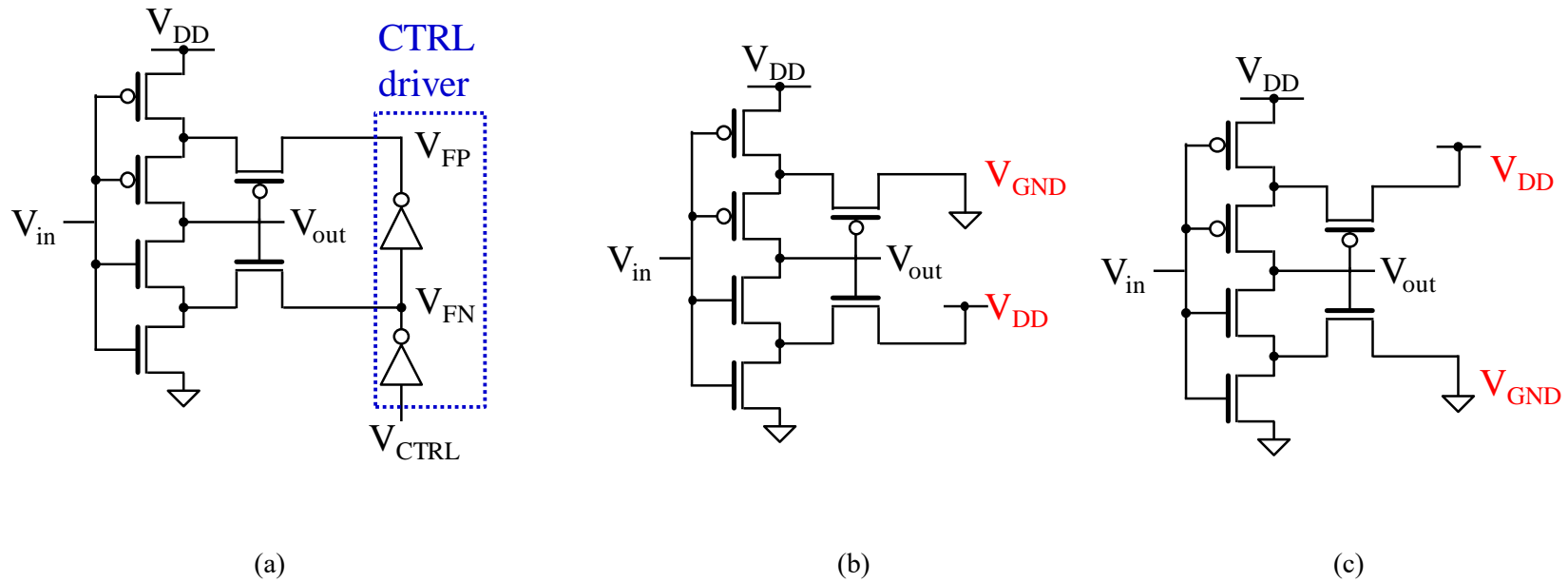
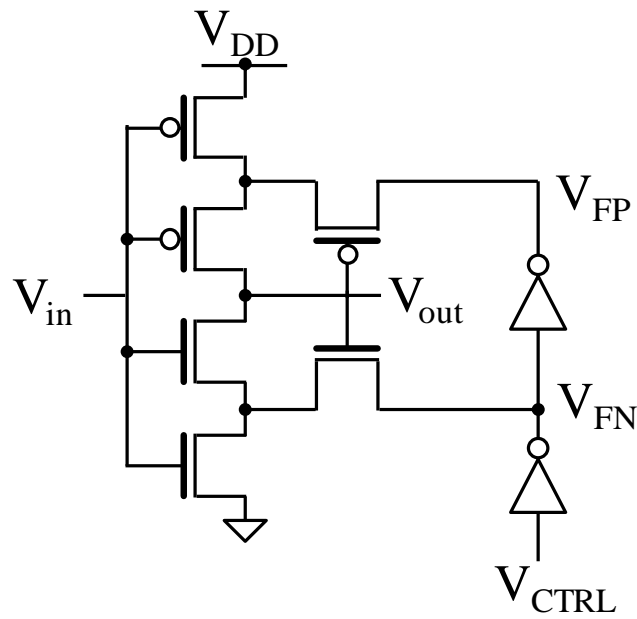
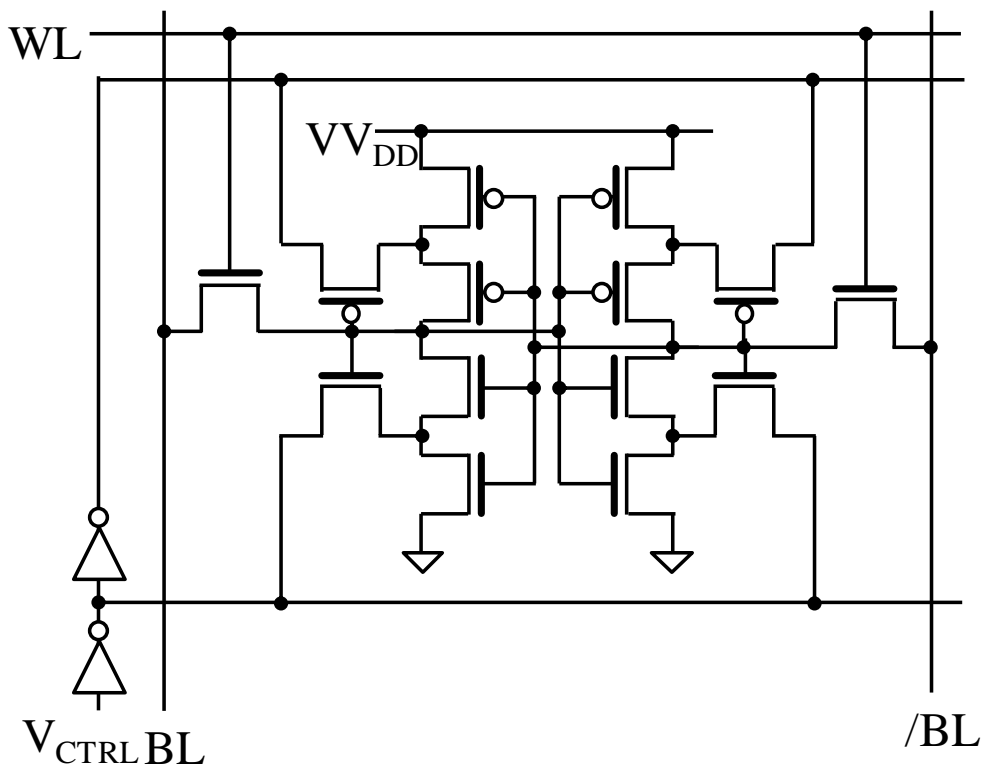


図 4.1.4 (a)DM インバータの回路構成と(b)ST モード時および(c)BI モード時のフィードバックトランジスタのバイアス構成

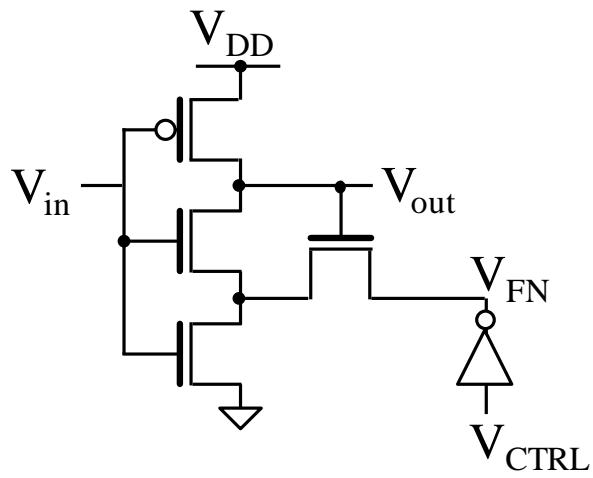


(a)

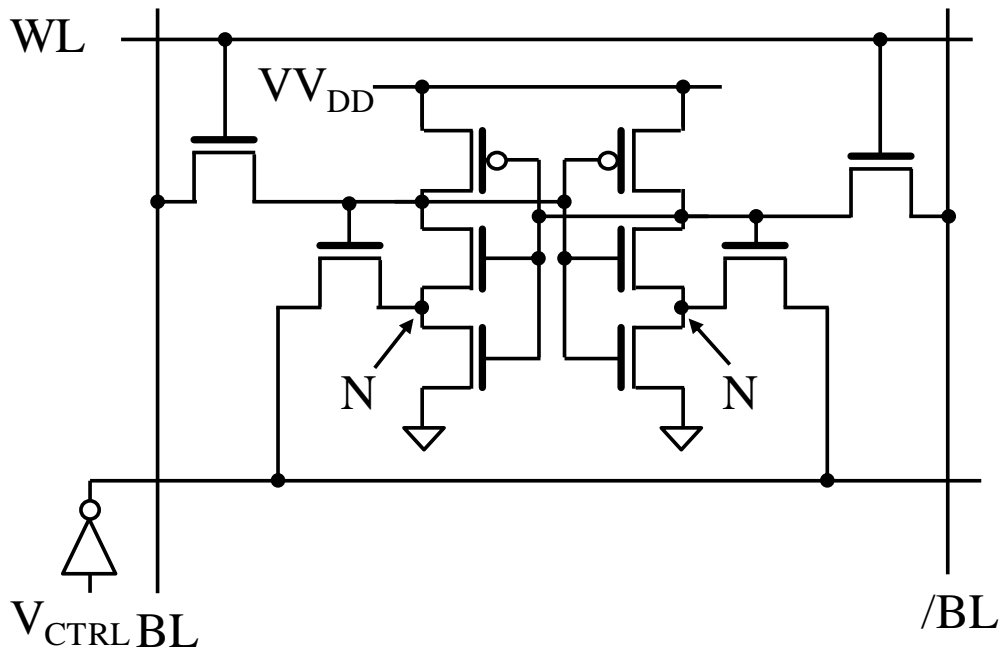


(b)

図 4.2.1 (a)プルアップ側とプルダウン側の両方にFBトランジスタを接続した構成のDMインバータと(b)これを用いて構成したVNRセル



(a)



(b)

図 4.2.2 (a)プルダウン側のみに FB トランジスタを接続した構成の DM インバータと(b) これを用いて構成した VNR セル

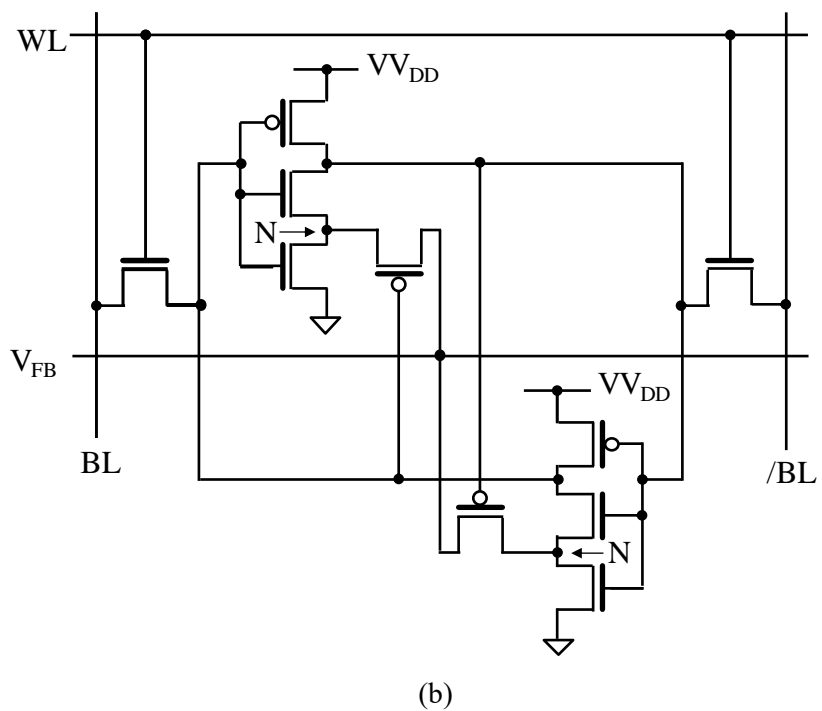
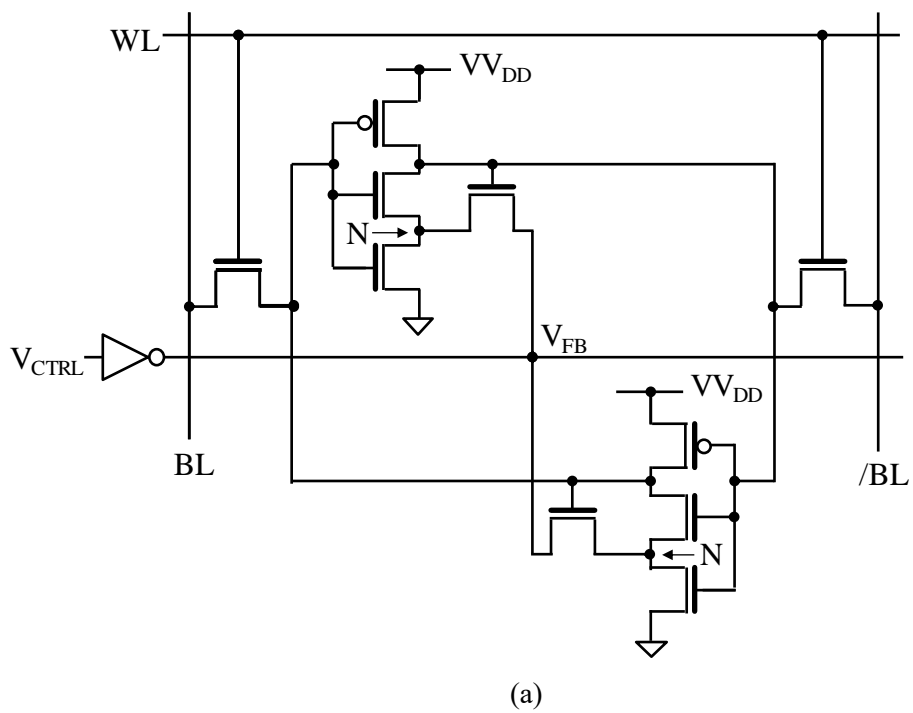
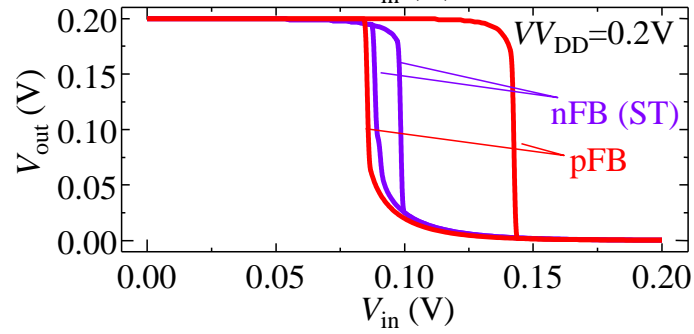
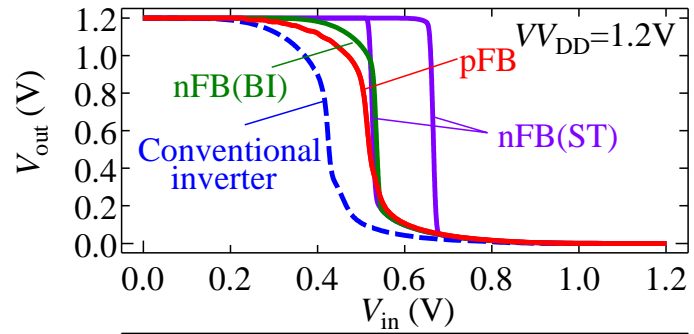
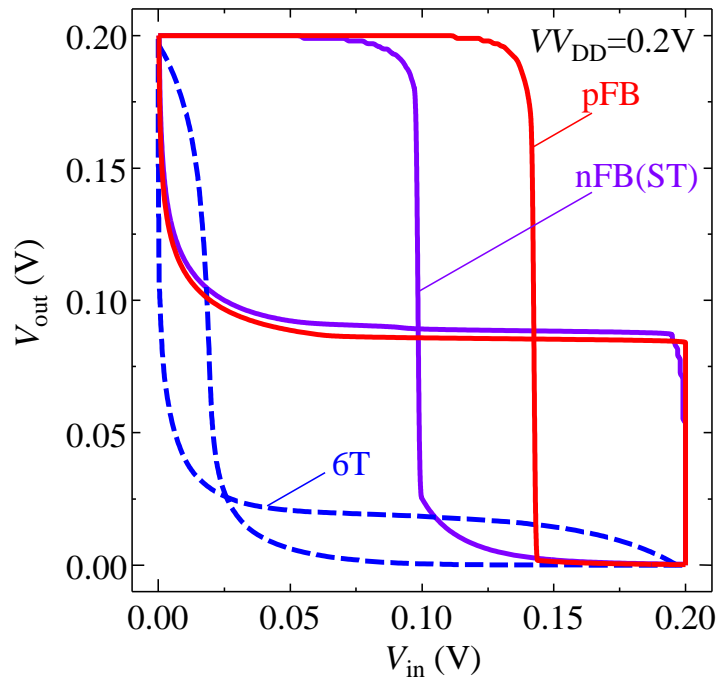


図 4.2.3 (a)nFBPD-VNR セルと(b)pFBPD-VNR セルのループ型表示



(a)



(b)

図 4.2.4 (a) $V_{DD}=1.2V$ と $0.2V$ におけるインバータ特性と(b) $V_{DD}=0.2V$ におけるバタフライカーブ

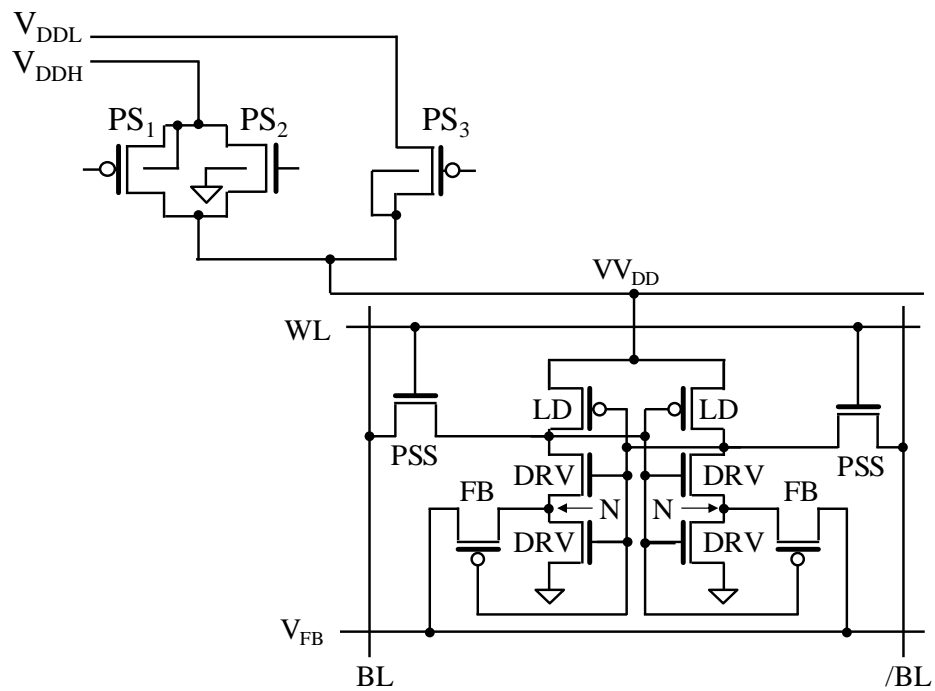


図 4.3.1 pFB セルにおけるセルとパワースイッチの構成

表 4.3.1 pFB セルとパワースイッチのデバイスパラメータとバイアスの名称

Device parameters and biases		Power switch (Channel width, gate bias)	
Cell (Channel width)			
LD transistor:	W_{LD}	PS ₁ :	(W_{PG1}, V_{PG1})
DRV: transistor:	W_{DRV}	PS ₂ :	(W_{PG2}, V_{PG2})
PSS transistor:	W_{PSS}	PS ₃ :	(W_{PG3}, V_{PG3})
FB transistor:	W_{FB}		

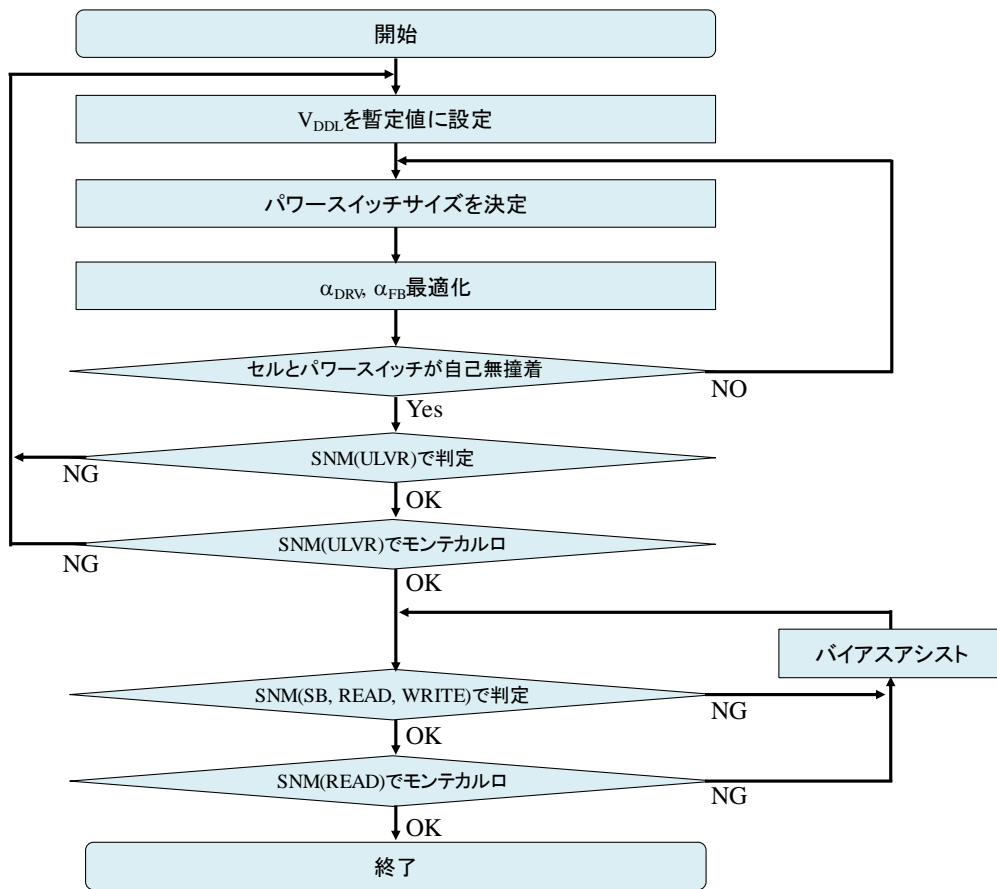


図 4.3.2 pFB セルの設計フローチャート

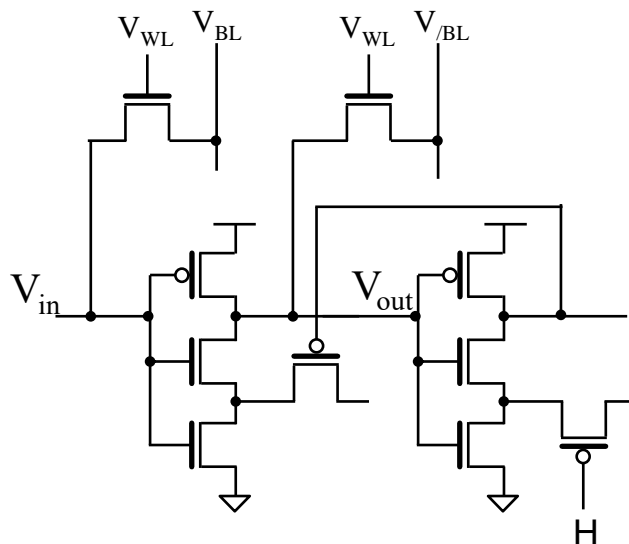
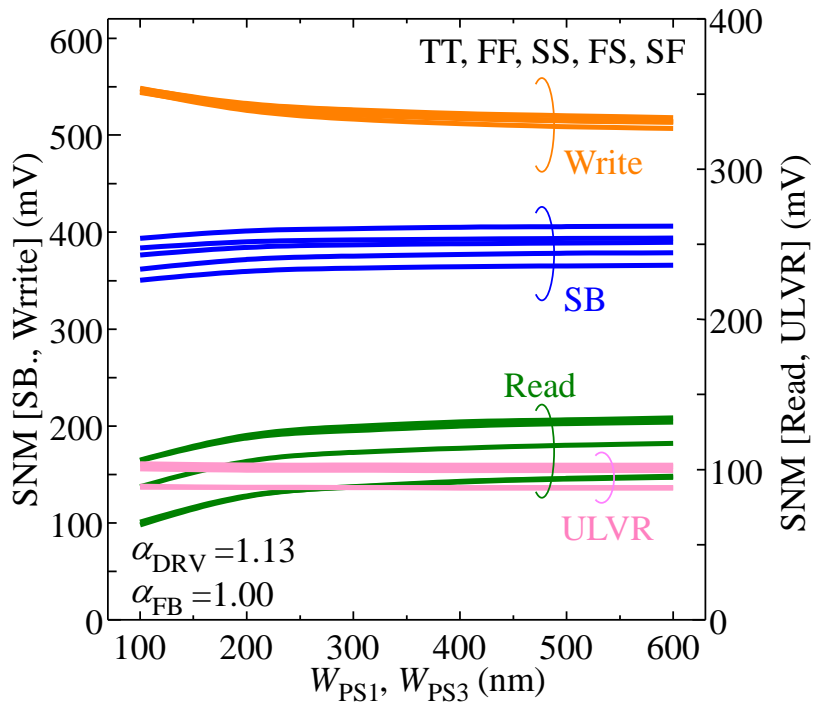


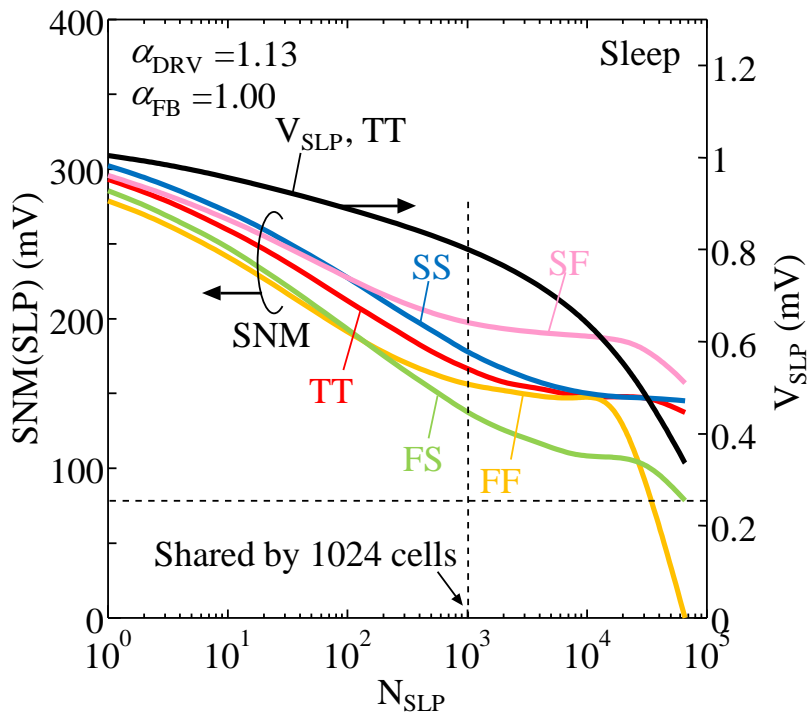
図 4.3.3 SNM の評価回路

表 4.3.2 pFB セルの最適設計におけるデバイスパラメータとバイアス

Cell design		Circuit operation and organization	
Channel length L	60 nm	Subarray size	8 kB
Supply voltage V_{DD}	1.2 V	Read/Write bit width	64 bit
Channel width W_{cell}	100 nm	Read/Write speed	1 GHz
W_{LD}	100 nm	Leakage control	
W_{PSS}	100 nm		
W_{DRV}	170 nm	Cell	V_{FB} 0.2V
W_{FB}	100 nm		V_{PG1}, V_{PG3} 1.4 V ($V_{DD}=0.02V$)
W_{PS1}	300nm (per 1 cell)	Peripheral	$V_{DD}=0.02V$
W_{PS2}	100nm (per 16 lines)		
W_{PS3}	100nm (per 1 cell)	Sleep control	V_{PG2} 1.2 V ($V_{DD}=0.8V$)

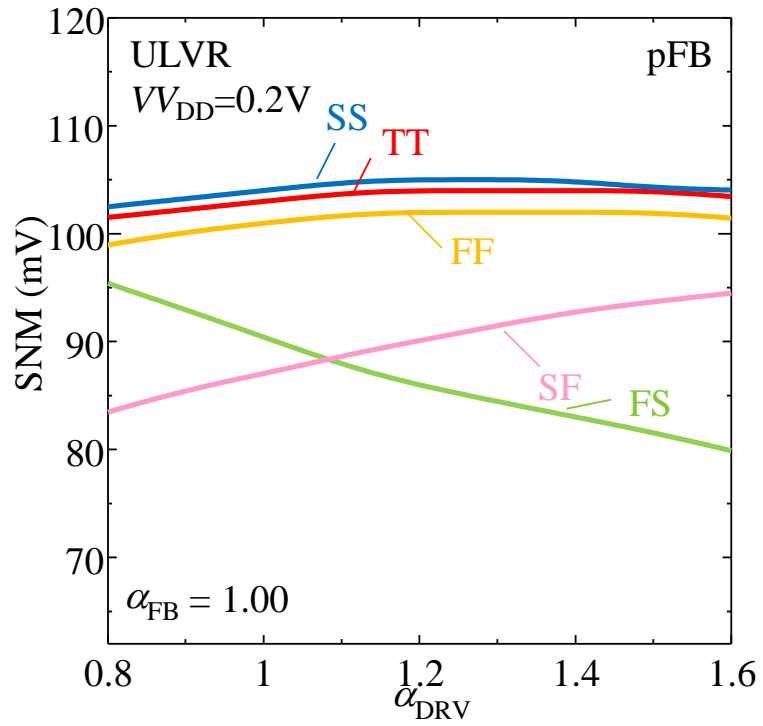


(a)

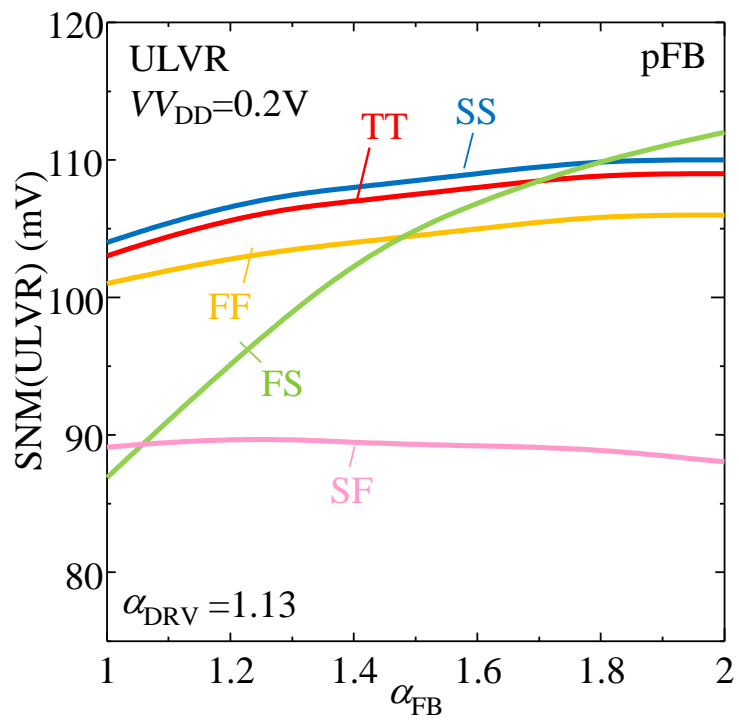


(b)

図 4.3.4 (a)各動作モード(mode=SB, READ, WRITE, ULVR)における SNM の W_{PS1} , W_{PS3} 依存性と(b)スリープモード(mode=SLP)における SNM と V_{SLP} の N_{SLP} 依存性

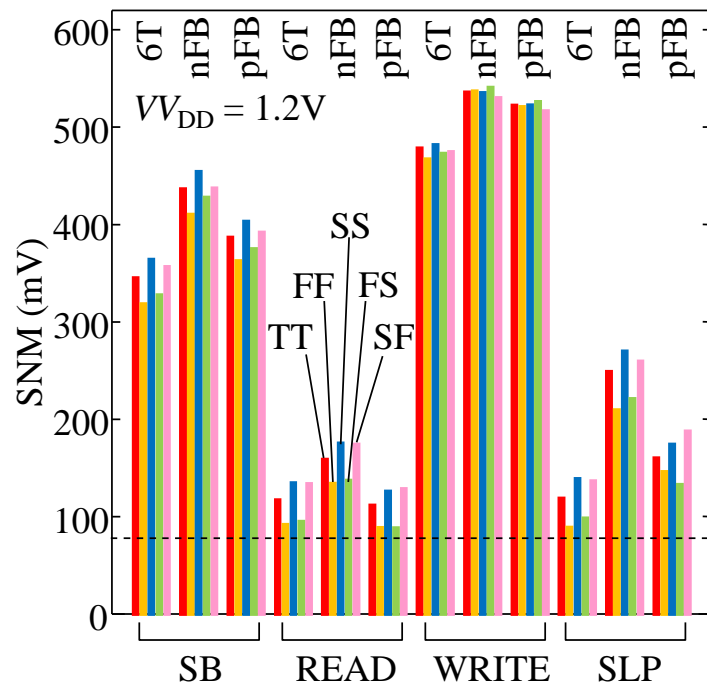


(a)

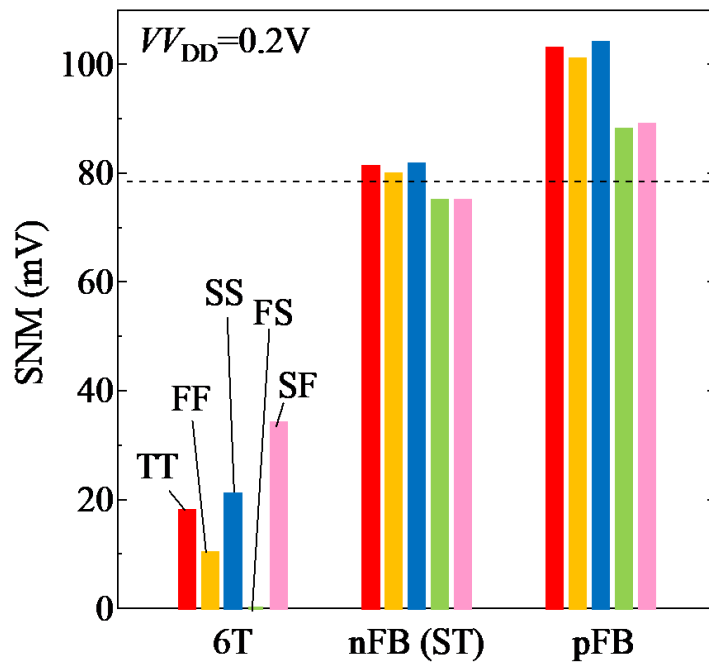


(b)

図 4.3.5 ULV リテンションにおける SNM の(a) α_{DRV} 依存性と(b) α_{FB} 依存性



(a)



(b)

図 4.3.6 (a)通常の SRAM 動作時(mode=SB, READ, WRITE, SLP)の SNM と(b)ULV リテンション(mode=ULVR)の SNM

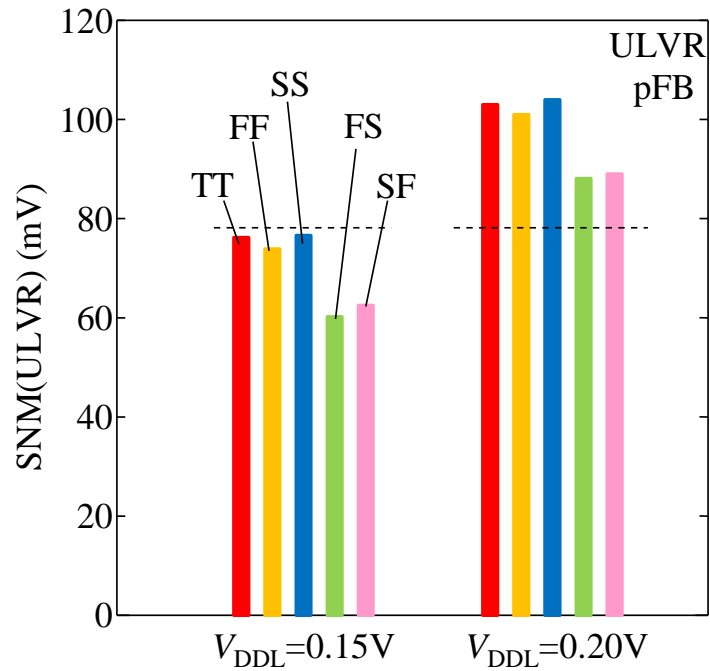


図 4.3.7 $V_{DDL}=0.15V, 0.20V$ における ULV リテンションの SNM

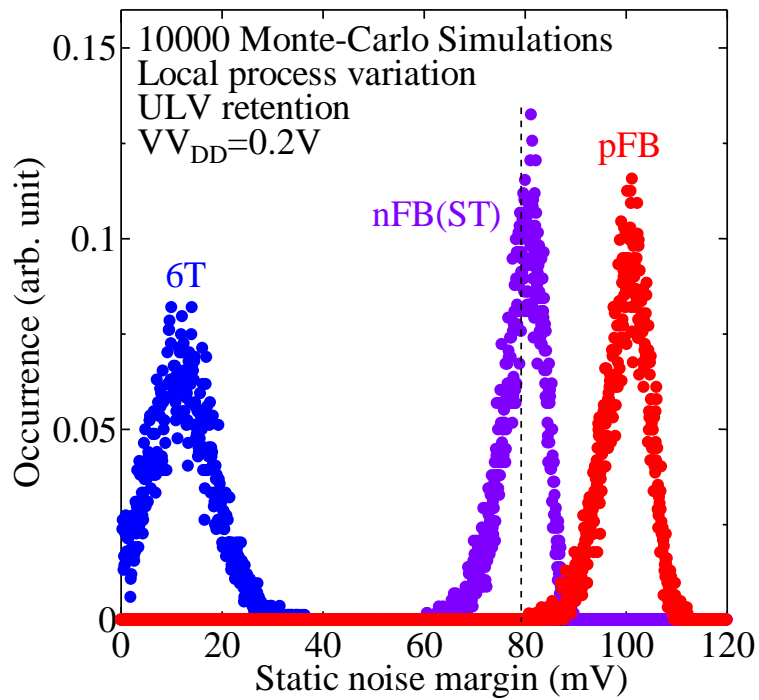
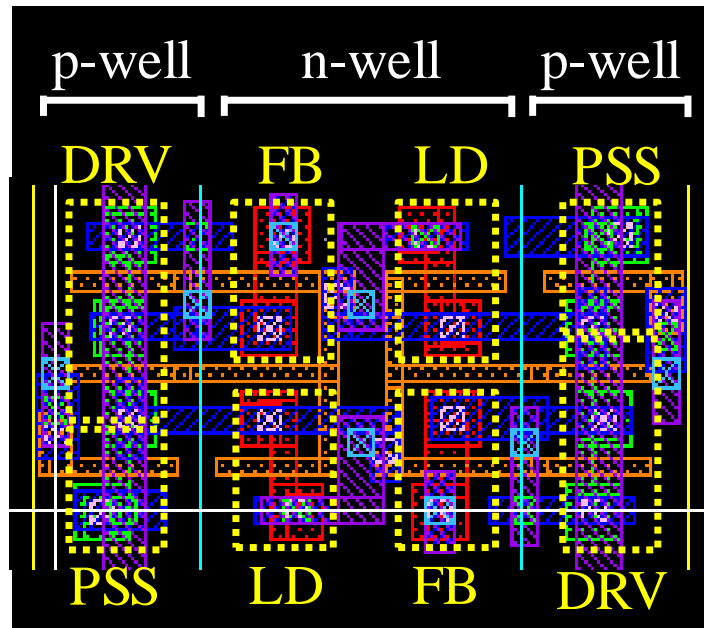
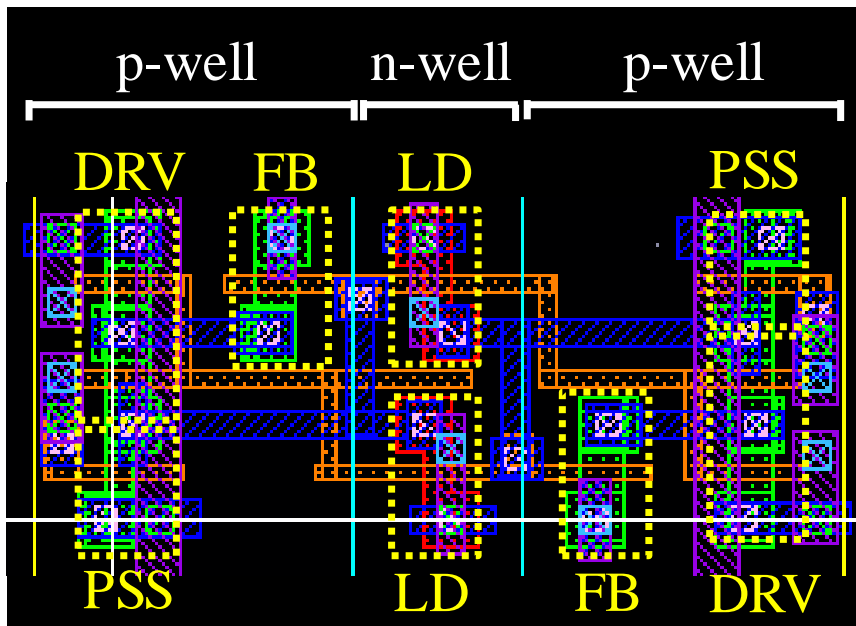


図 4.3.8 ULV リテンションの SNM におけるローカルばらつきを考慮したモンテカルロシミュレーションの解析結果

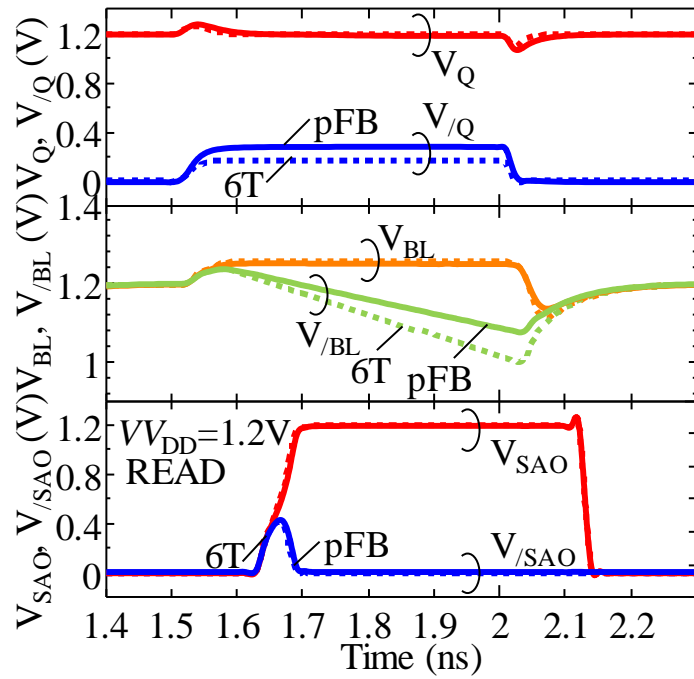


(a)

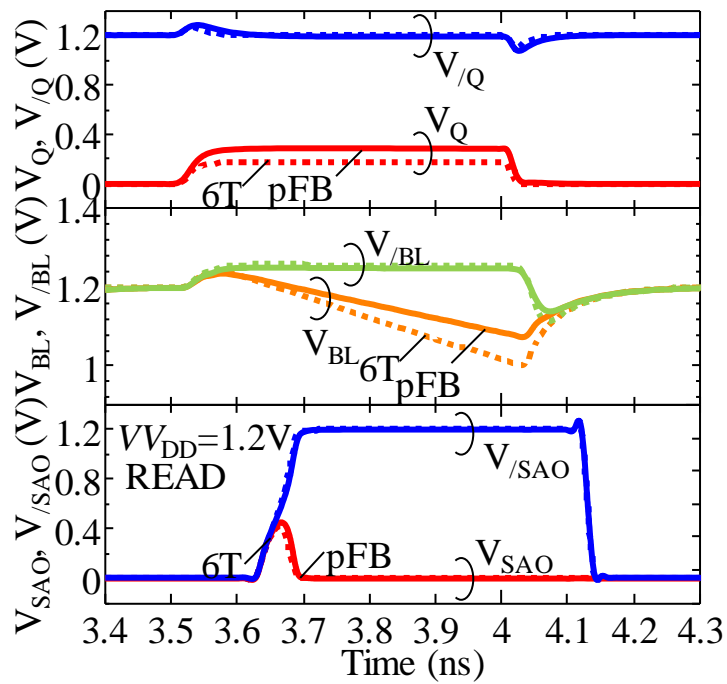


(b)

図 4.3.9 $V_{DDL}=0.2V$ で最適設計された(a)pFB セルと(b)nFB セルのレイアウト

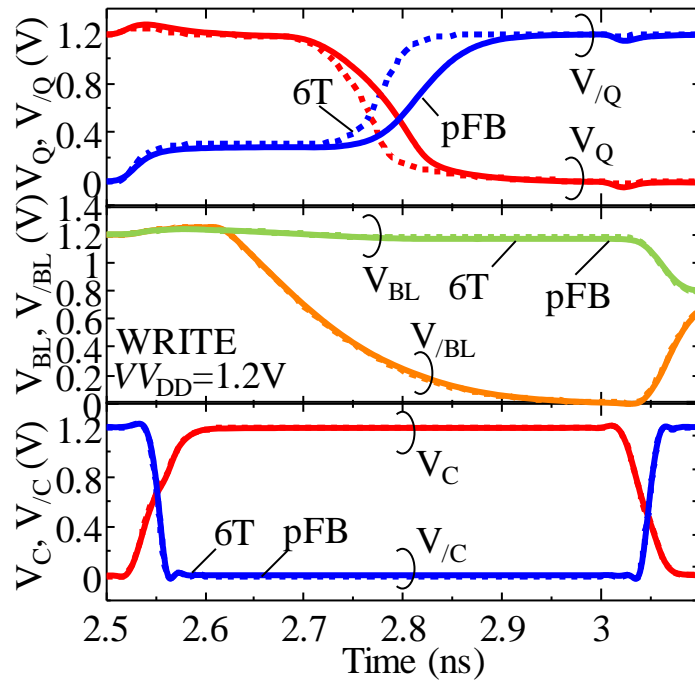


(a)

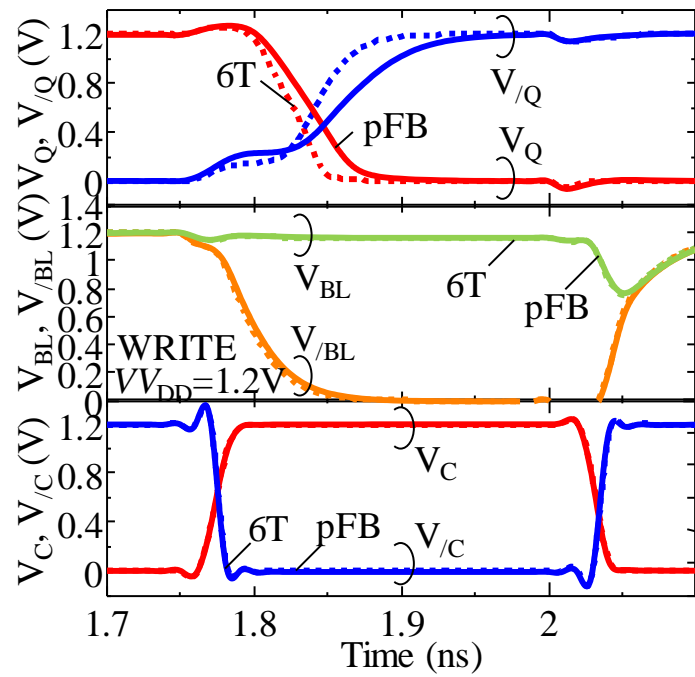


(b)

図 4.3.10 (a) $(V_Q, V/Q) = (V_{DD}, 0)$, (b) $(V_Q, V/Q) = (0, V_{DD})$ における通常の SRAM の読み出し動作の波形



(a)



(b)

図 4.3.11 (a)1GHz 動作と(b)2GHz 動作における書き込み動作の波形

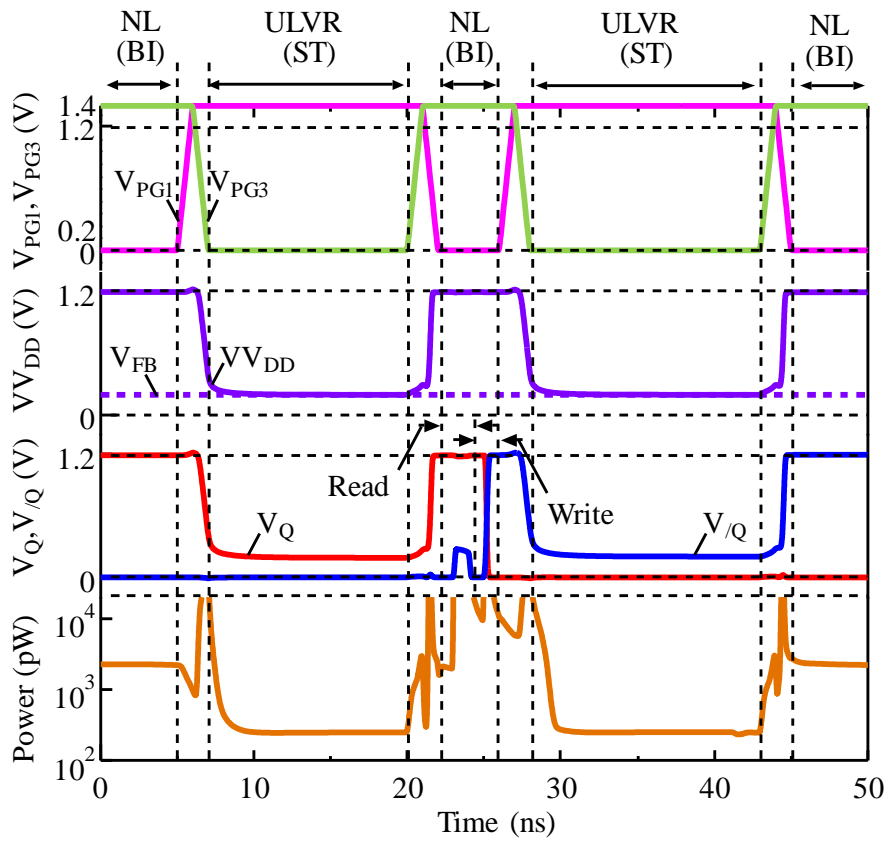
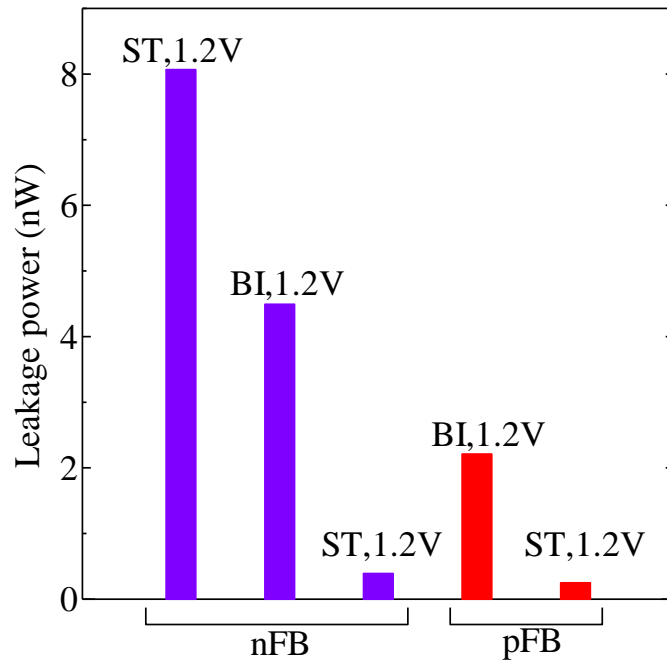
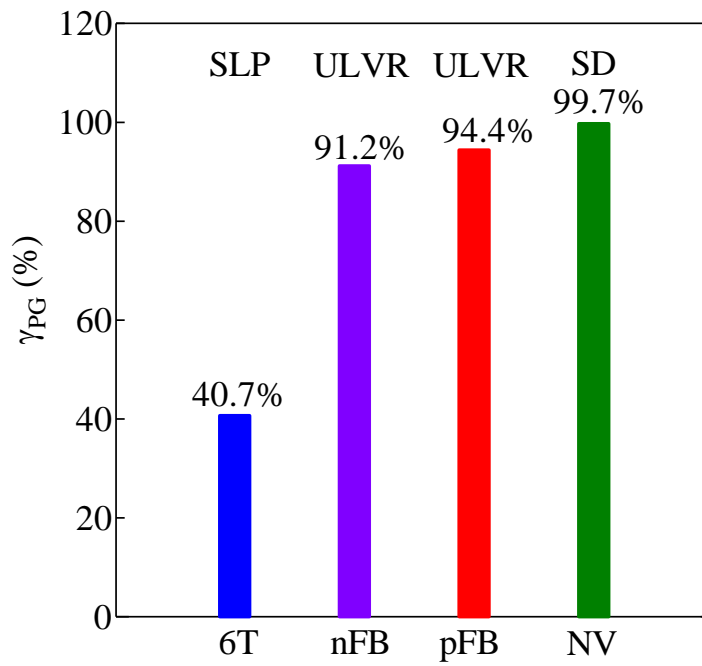


図 4.3.12 最適設計された pFB セルの動作波形

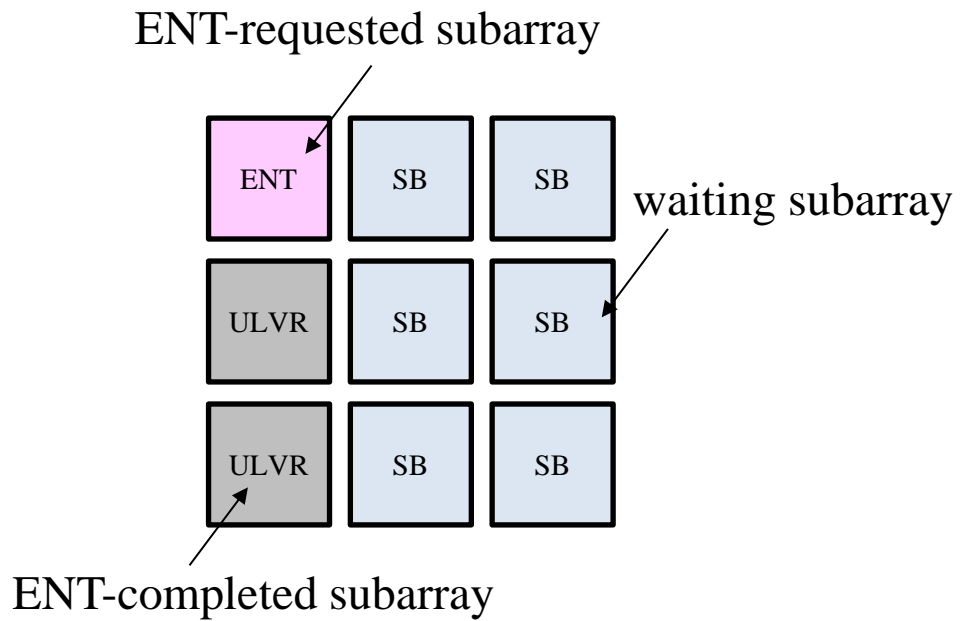


(a)

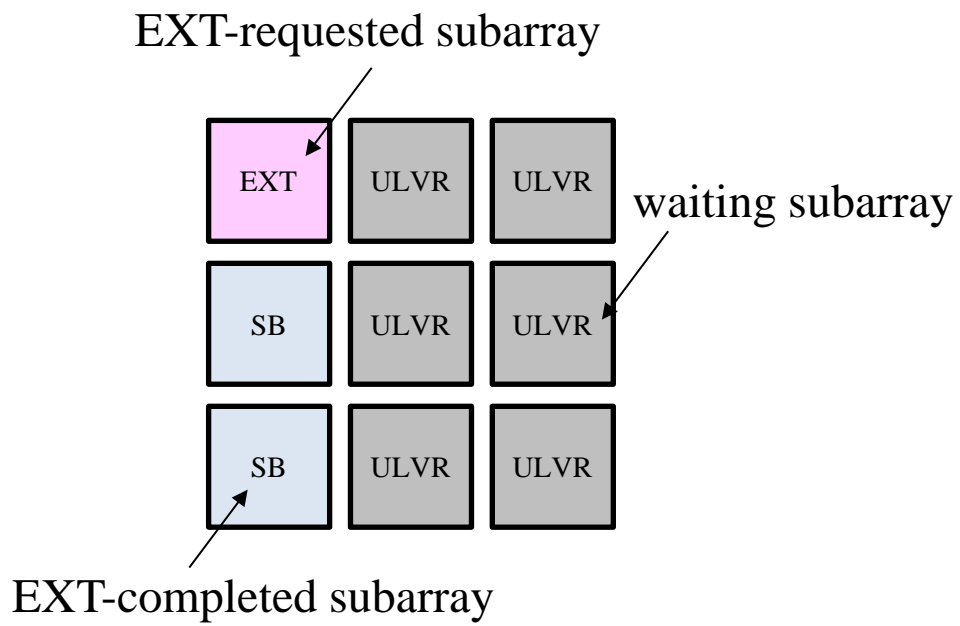


(b)

図 4.4.1 (a)nFB セルと pFB セルの各動作時(mode=SB, ULVR)におけるスタンバイ電力と(b)各種セルの γ_{PG}

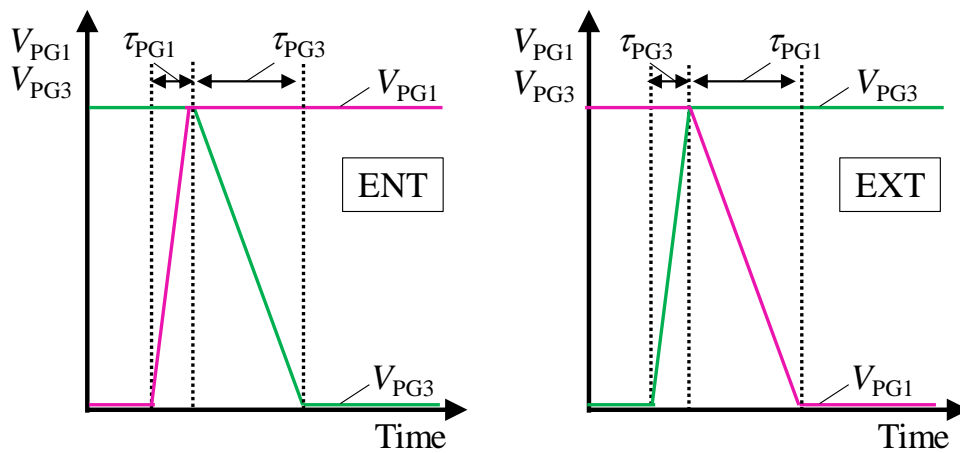


(a)

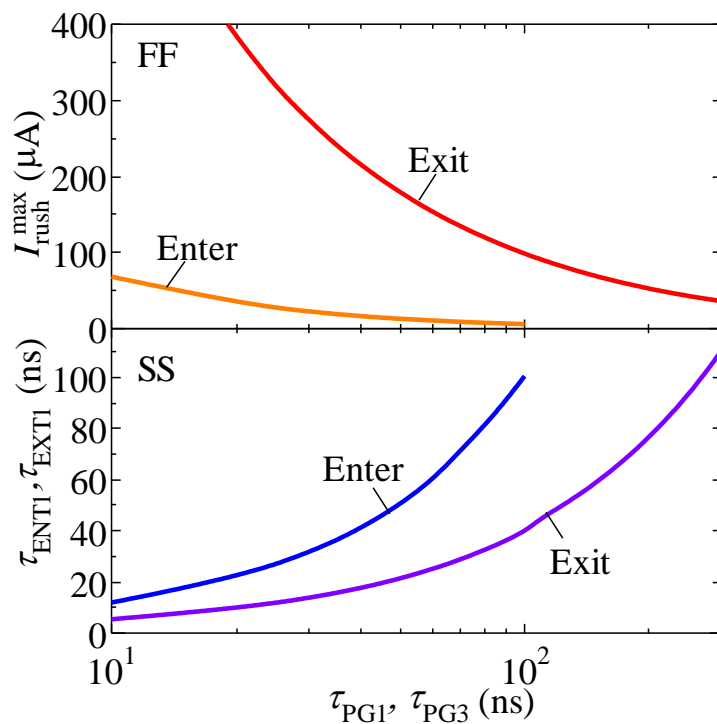


(b)

図 4.4.2 (a) ENT モードと(b)EXT モードにおける各サブアレイの回路状態

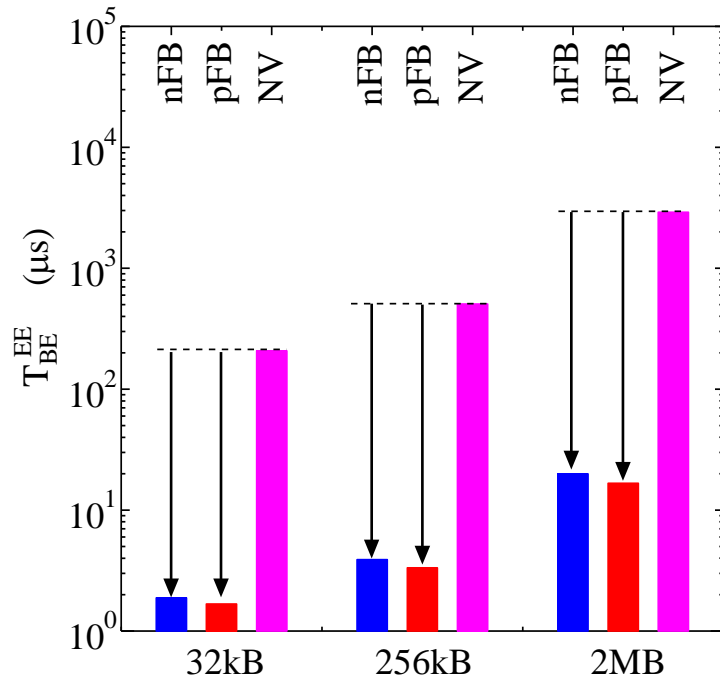


(a)

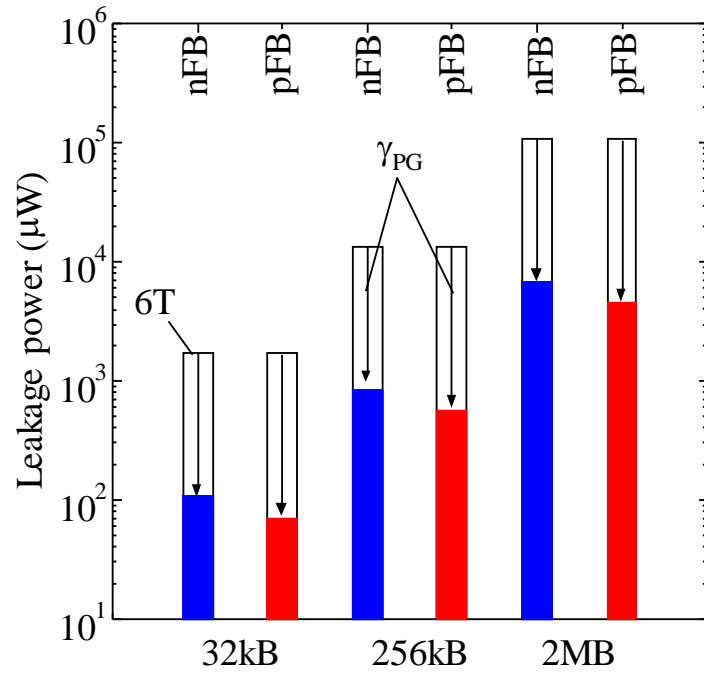


(b)

図 4.4.3 (a)ENT 動作と EXT 動作における PS_1 と PS_3 の制御シーケンスと(b)ENT/EXT 動作における I_{rush}^{max} および τ_{ENT1} , τ_{EXT1} の τ_{PG1} , τ_{PG3} 依存性

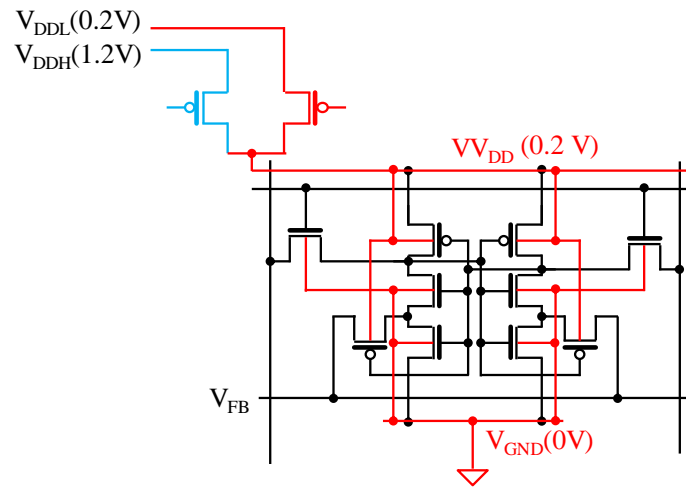


(a)

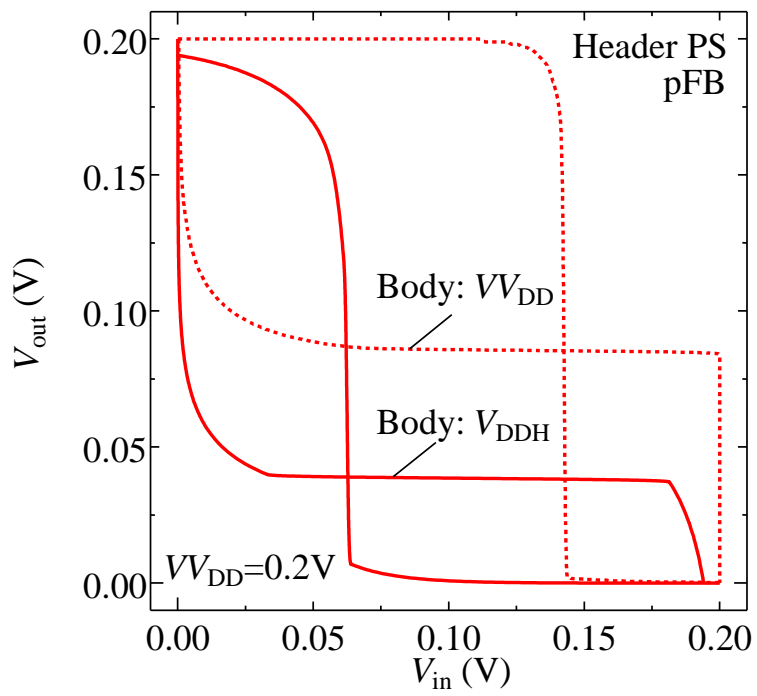


(b)

図 4.4.4 32kB, 256kB, 2MB のアレイにおける(a) T_{BE}^{EE} と(b)リーク電力の解析結果

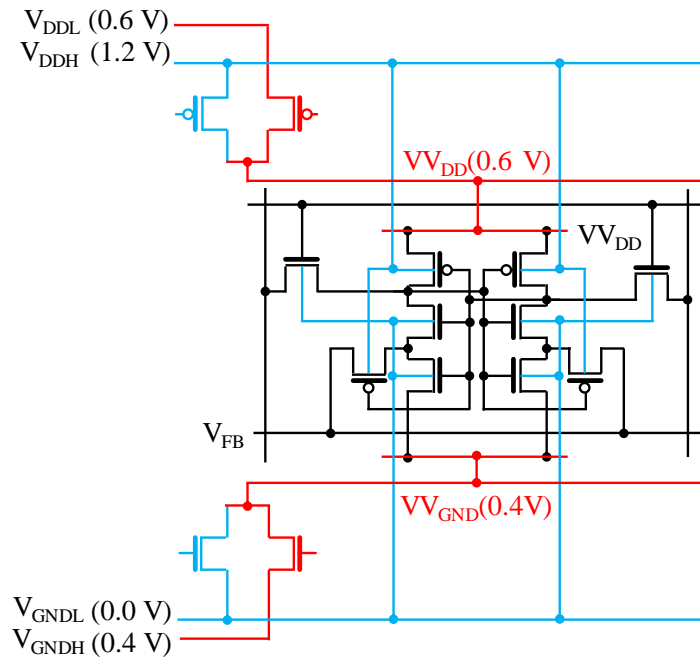


(a)

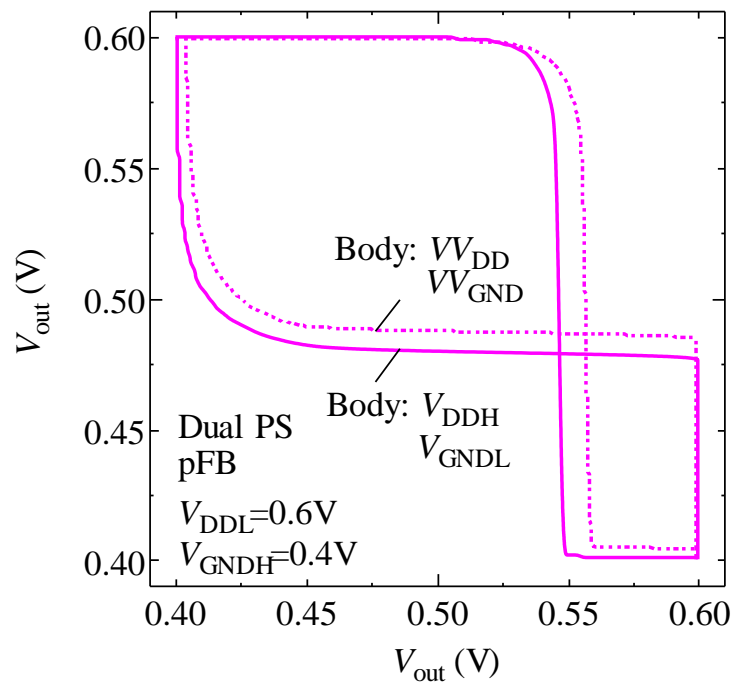


(b)

図 4.5.1 (a) ヘッダパワースイッチ構成の pFB セルの回路構成と(b) ULV リテンションにおけるバタフライカーブ



(a)



(b)

図 4.5.2 (a) デュアルパワースイッチ構成の pFB セルの回路構成と(b) ULV リテンションにおけるバタフライカーブ

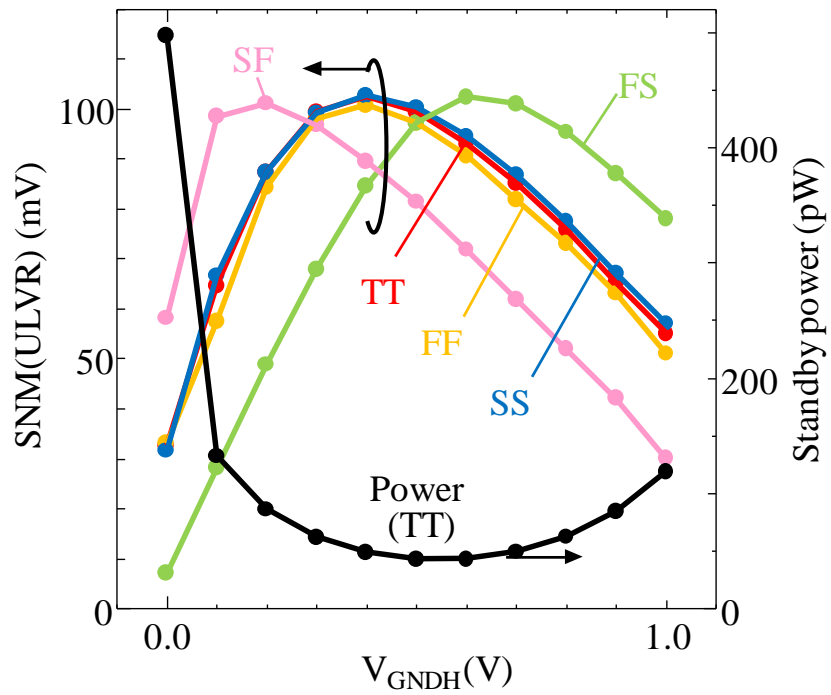


図 4.5.3 デュアルパワースイッチを用いた pFB セルにおける SNM(ULVR) とリーク電力の V_{GNDH} 依存性

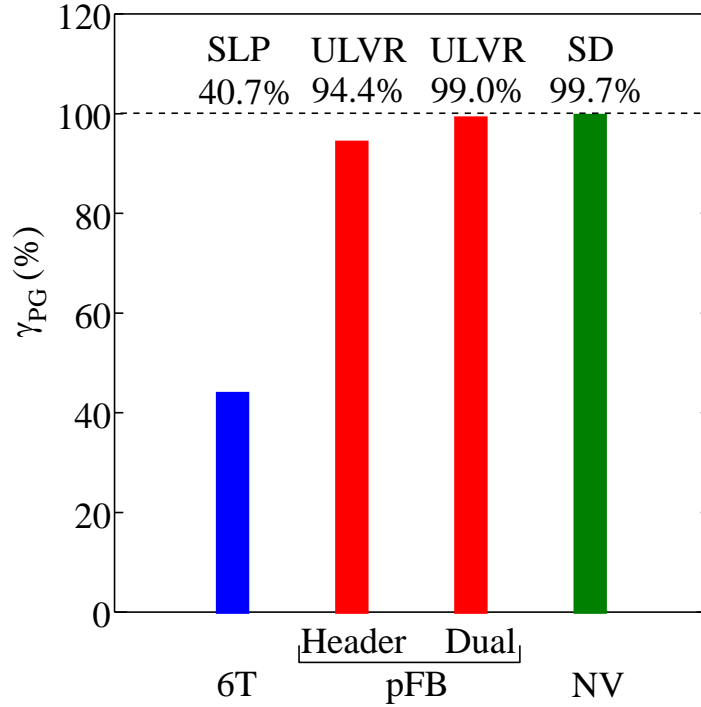


図 4.5.4 各種セルの γ_{PG}

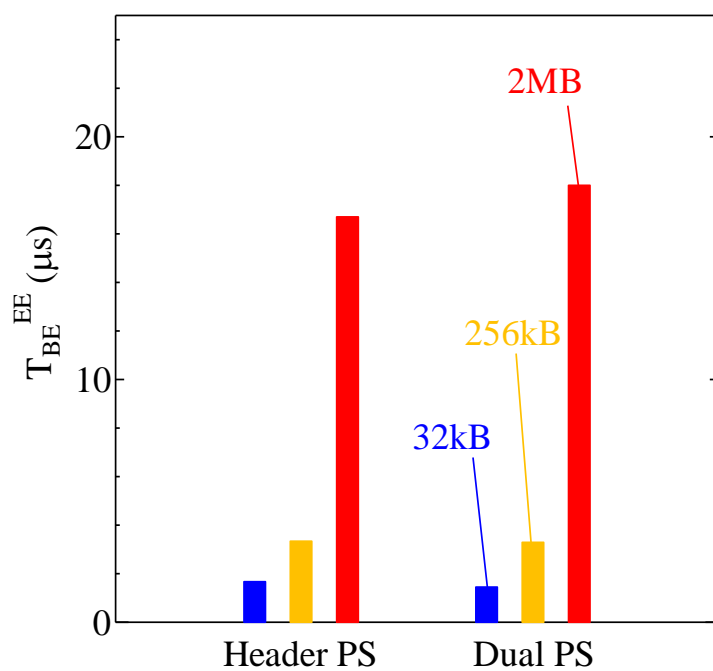


図 4.5.5 ヘッダパワースイッチ構成とデュアルパワースイッチ構成の pFB セルの T_{BE}^{EE}

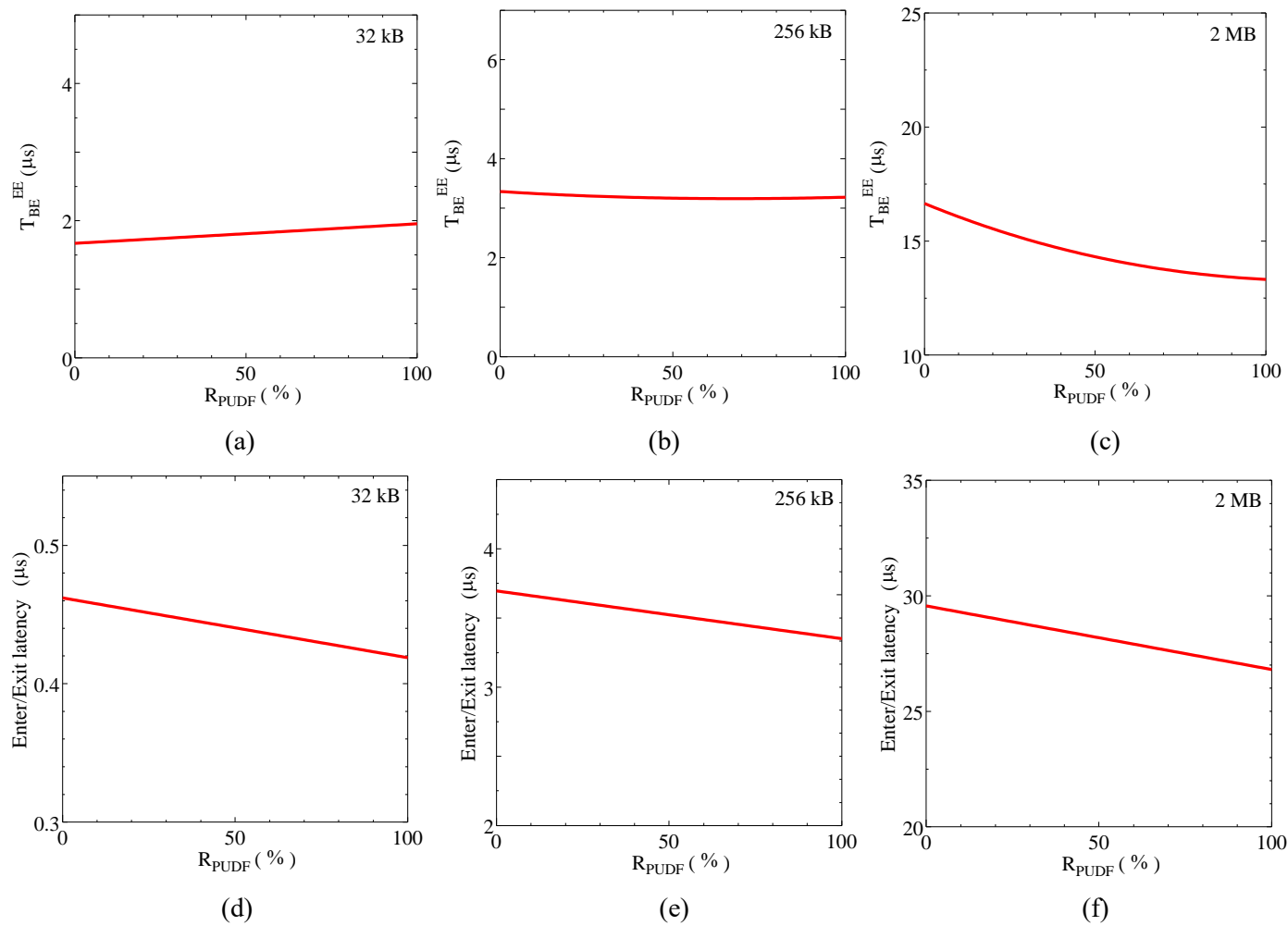
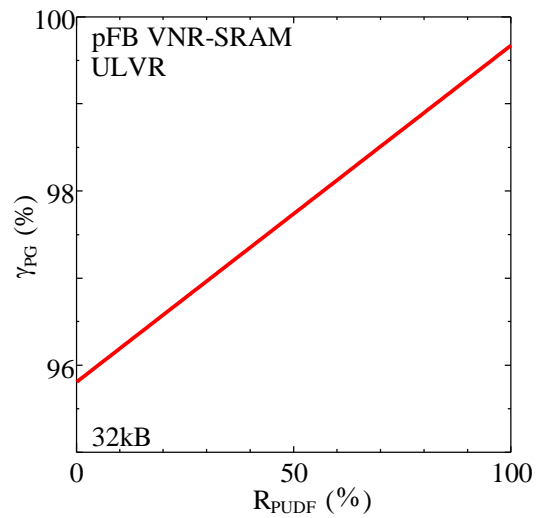
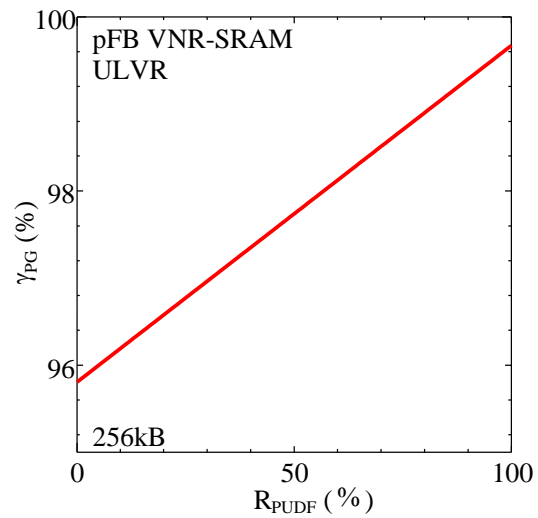


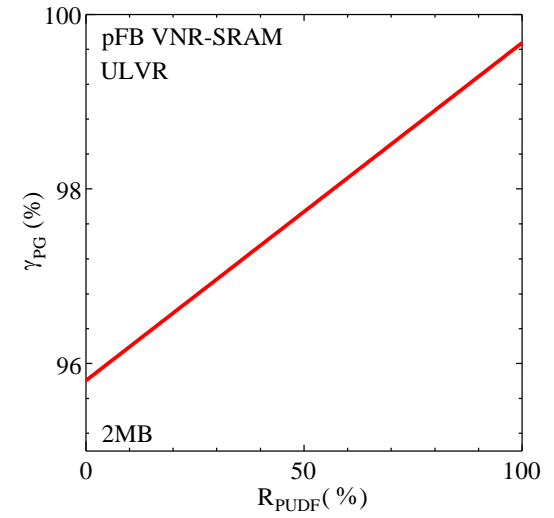
図 4.5.6 PUDF アーキテクチャを用いた場合の(a)-(c) T_{BE}^{EE} と(d)-(f)ENT/EXT レイテンシの R_{PUDF} 依存性



(a)



(b)



(c)

図 4.5.7 (a)32kB, (b)256kB, (c)2MB の VNR-SRAM における γ_{PG} の R_{PUDF} 依存性



図 4.6.1 VNR-SRAM マクロの設計フローと使用した設計ツール

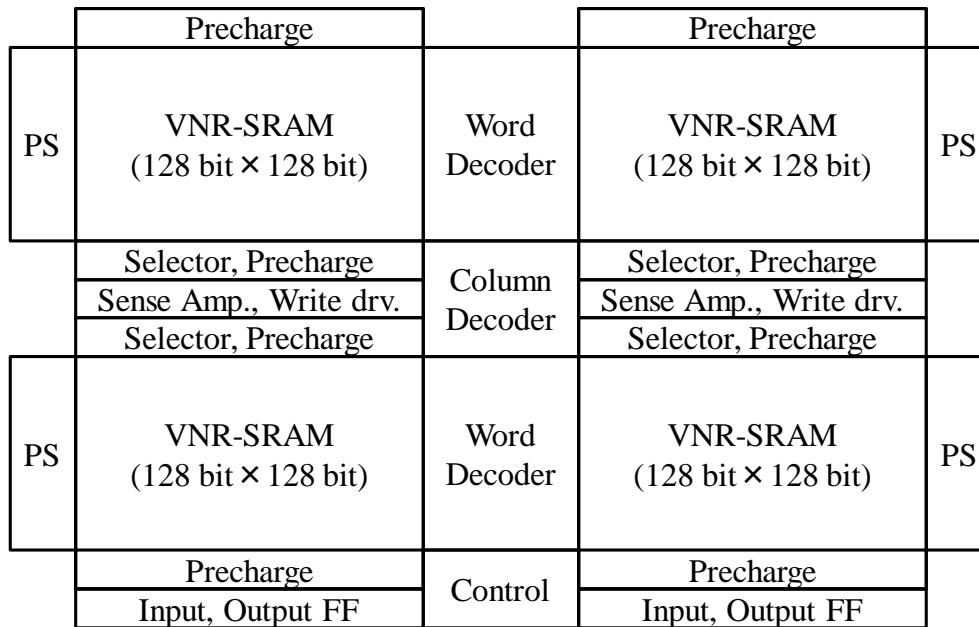


図 4.6.2 8kB サブアレイのフロアプラン

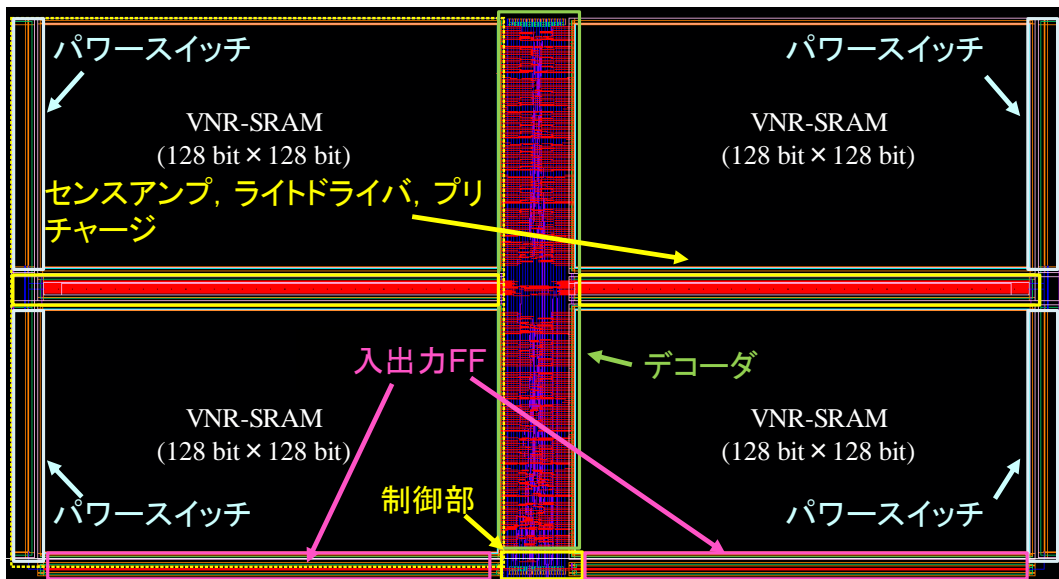


図 4.6.3 8kB サブアレイのレイアウト

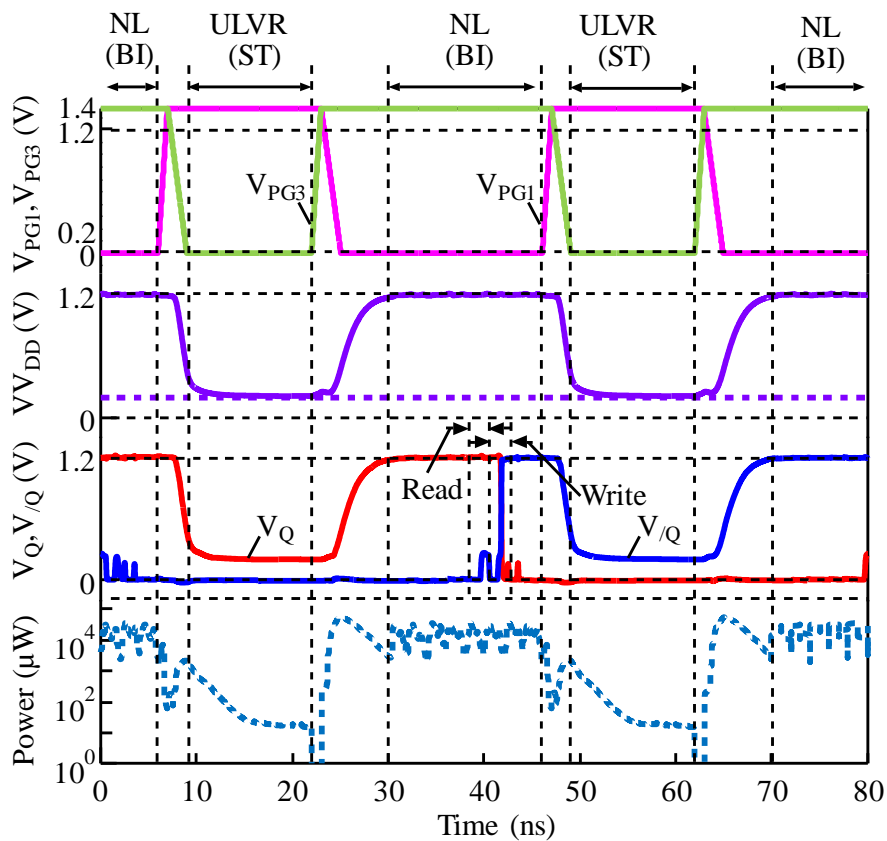


図 4.6.4 サブアレイの動作波形

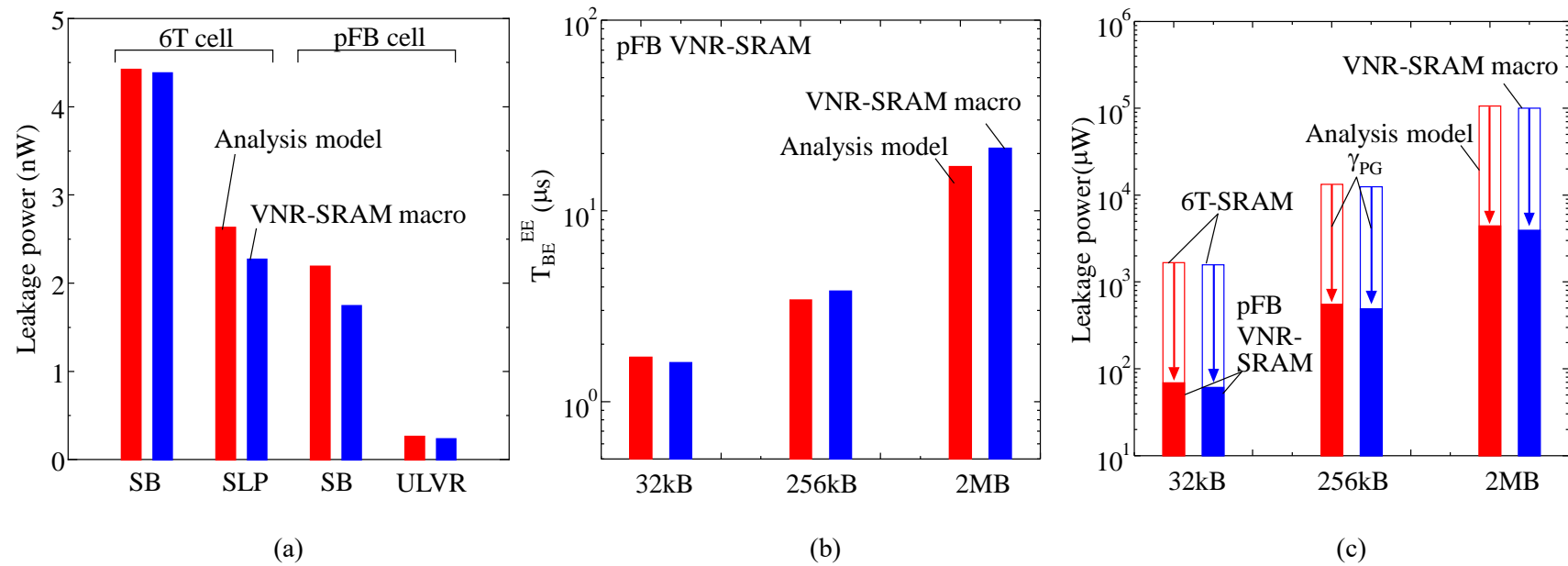


図 4.6.5 第 2 章の性能解析モデル(赤の棒グラフ)と設計した VNR-SRAM マクロ(青の棒グラフ)から抽出した回路パラメータを用いた評価結果の比較. (a)抽出した単体セルのリーク電力の評価結果と 32kb, 256kb, 2MB のアレイの(b) T_{BE}^{EE} と(c)リーク電力の評価結果(比較のために 6T-SRAM のスタンバイ時のリーク電力も示している)

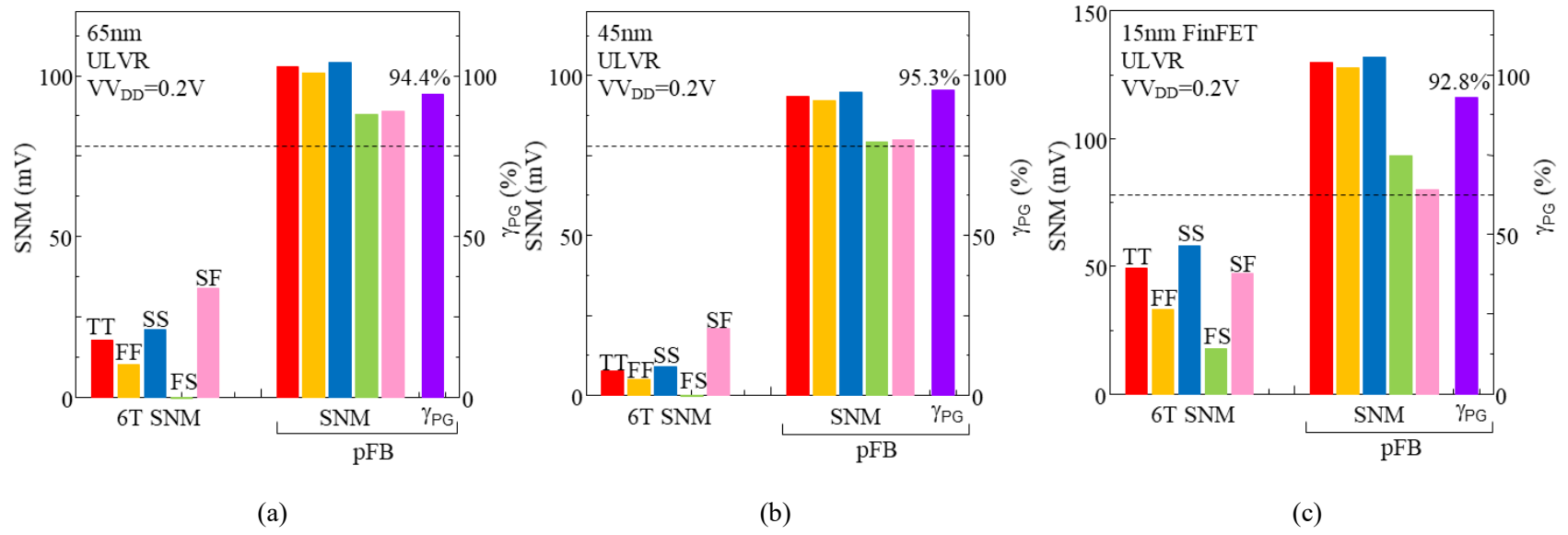


図 4.7.1 (a)65nm, (b)45nm, (c)15nm プロセスにおける 6T セルと pFB セルの SNM と γ_{PG}

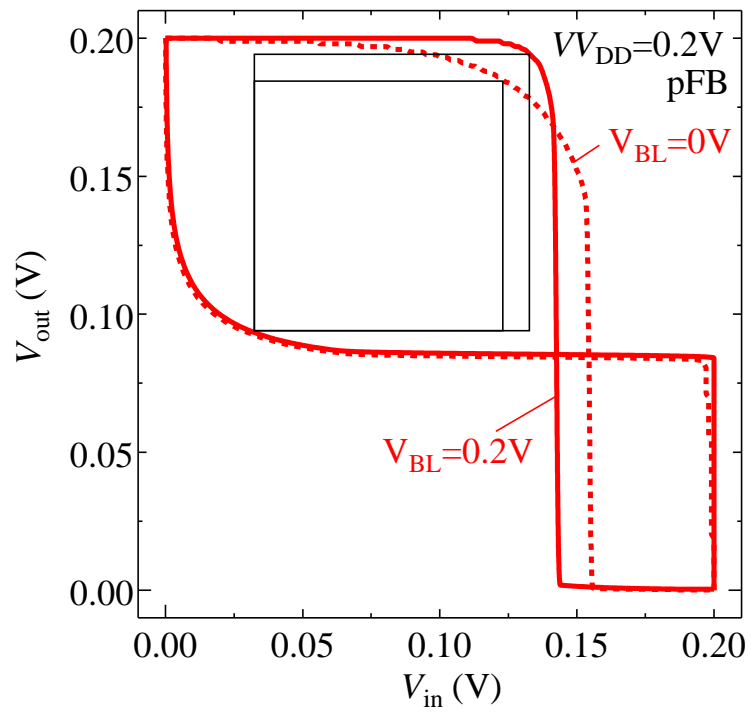


図 4.9.1 $V_{BL} = 0V, 0.2V$ における pFB セルのバタフライカーブ

4.11 参考文献

- [1] Y. Kanno, H. Mizuno, N. Oodaira, Y. Yasu, and K. Yanagisawa, “ μ I/O architecture for 0.13- μ m wide-voltage-range system-on-a-package (SoP) designs”, Symp. on VLSI Circuits Dig. Tech. Pap., 168 (2002).
- [2] S. Rusu, S. Tam, H. Muljono, D. Ayers, J. Chang, “A dual-core multi-threaded Xeon processor with 16MB L3 cache”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 315 (2006).
- [3] L. Chang, D. M. Fried, J. Hergenrother, J. W. Sleight, R. H. Dennard, R. K. Montoye, L. Sekaric, S. J. McNab, A. W. Topol, C. D. Adams, K. W. Guarini, and W. Haensch, “Stable SRAM cell design for the 32 nm node and beyond”, IEEE Symposium on VLSI Technology 2005, 128 (2005).
- [4] A. Islam, and M. Hasan, “Leakage characterization of 10T SRAM cell”, IEEE Trans. on Electron Devices **59**, 631 (2012).
- [5] J. P. Kulkarni, K. Kim, and K. Roy, “A 160 mV robust Schmitt trigger based subthreshold SRAM”, IEEE Journal of Solid-State Circuits **42**, 2303 (2007).
- [6] J. P. Kulkarni, K. Kim, S. P. Park, and K. Roy, “Process variation tolerant SRAM array for ultra low voltage applications”, 45th annual Design Automation Conference, 108 (2008).
- [7] K. Roy, J. P. Kulkarni, and M. Hwang, “Process variation tolerant SRAM array for ultra low voltage applications”, IEEE SENSORS, 185 (2009).
- [8] Y. N. Chen, C. Y. Hsieh, M. L. Fan, V. P. H. Hu, P. Su, and C. T. Chuang, “Impacts of Intrinsic Device Variations on the Stability of FinFET Subthreshold SRAMs”, IEEE International Conference on IC Design and Technology, 1 (2011).
- [9] J. P. Kulkarni, and K. Roy, “Ultralow-Voltage Process-Variation-Tolerant Schmitt-Trigger-Based SRAM Design”, IEEE Trans. on VLSI Syst. **20**, 319 (2012).
- [10] C. Y. Hsieh, M. L. Fan, V. P. H. Hu, P. Su, and C. T. Chuang, “Independently-Controlled-Gate FinFET Schmitt Trigger Sub-Threshold SRAMs”, IEEE Trans. on VLSI Syst. **20**, 1201 (2012).
- [11] N. Yadav, S. Jain, M. Pattanaik, and G. K. Sharma, “A novel stability and process sensitivity driven model for optimal sized FinFET based SRAM”, Microelectronics Reliability **55**, 1131 (2015).
- [12] S. Ahmad, M. K. Gupta, N. Alam, and M. Hasan, “Single-Ended Schmitt-Trigger-Based Robust Low-Power SRAM Cell”, IEEE Trans. on VLSI Syst.

- 24, 2634 (2016).
- [13] J. Cai, J. Yuan, L. Chen, and Y. Hei, “High noise margin 12T subthreshold SRAM cell with enhanced read speed and eliminated half-selected problem”, IEEE International Conference on Solid-State and Integrated Circuit Technology, 582 (2016).
- [14] D. Sreenivasan, D. Purushothaman, K. S. Pande, and N. S. Murty, “Dual-threshold single-ended Schmitt-Trigger based SRAM cell”, IEEE International Conference on Computational Intelligence and Computing Research, 1 (2016).
- [15] T. V. Reddy, B. and K. Madavi, “Designing of schmitt trigger-based architecture 8T SRAM of 256bit cells under 45 NM technology for low power applications”, International Conference on Intelligent Computing and Control, 1 (2017).
- [16] A. S. Rajput, M. Pattanaik, and R. K. Tiwari, “Design and Analysis of Schmitt Trigger Based 10T SRAM in 32 nm Technology”, IEEE International Symposium on Nanoelectronic and Information Systems, 234 (2017).
- [17] M. Shrivastava and V. K. Yadav, “Low power schmitt trigger based sram using 32nm finfet devices”, Materials Today: Proceedings **5**, 1578 (2018) 1578.
- [18] K. Sanapala, R. Sakthivel, and S. S. Yeo, “Schmitt trigger-based single-ended 7T SRAM cell for Internet of Things (IoT) applications”, The Journal of Supercomputing **74**, 4613 (2018).
- [19] D. Kitagata, H. Yoshida, S. Yamamoto, and S. Sugahara, “Virtually Nonvolatile Retention SRAM cell Using Dual-Mode Inverters”, SOI-3D-Subthreshold Microelectronics Technology Unified Conference, 1 (2018).
- [20] E. Karl, Y. Wang, Y.-G. Ng, Z. Guo, F. Hamzaoglu, U. Bhattacharya, K. Zhang, K. Mistry, and M. Bohr, “A 4.6 GHz 162Mb SRAM design in 22nm tri-gate CMOS technology with integrated active V MIN-enhancing assist circuitry”, ISSCC Dig. Tech. Papers, 230 (2012)
- [21] S. Ohbayashi, M. Yabuuchi, K. Nii, Y. Tsukamoto, S. Imaoka, Y. Oda, T. Yoshihara, M. Igarashi, M. Takeuchi, H. Kawashima, Y. Yamaguchi, K. Tsukamoto, M. Inuishi, H. Makino, K. Ishibashi, and H. Shinohara, “A 65-nm SoC embedded 6T-SRAM designed for manufacturability with read and write operation stabilizing circuits”, IEEE Journal of Solid-State Circuits **42**, 820 (2007).
- [22] J.P. Kulkarni, K. Kim, and K. Roy, “A 160 mV fully differential, robust Schmitt trigger based sub-threshold SRAM”, IEEE International Symposium on Low Power Electronics (ISLPED) 2007, 171 (2007).

- [23] K. Ishibashi, T. Yamanaka, and K. Shimohigashi, "An alpha-immune, 2-V supply voltage SRAM using a polysilicon PMOS load cell", *IEEE J. Solid-State-Circuits* **25**, 55 (1990).
- [24] D. A. Hodges, H. G. Jackson, and R. A. Saleh, "Analysis and Design of Digital Integrated Circuits In Deep Submicron Technology" McGraw-Hill Companies Inc., 1221 Avenue of the Americas, NY, USA, 2004, third edition.
- [25] J. Knudsen, Nangate 45 nm Open Cell Library. EMEA: CDNLive, 2008.
- [26] M. Martins *et al.*, Open cell library in 15nm freepdk technology, ISPD, 2015.
- [27] M. Jin *et al.*, "Reliability Characterization of 10nm FinFET Technology with multi-VT gate stack for Low Power and High Performance", *IEEE International Electron Devices Meeting (IEDM) 2016*, 15-1 (2016).

第5章

CMOS ロジックシステムへの応用

本章では第3章、第4章で確立した不揮発/擬似不揮発性 SRAM (NV-/VNR-SRAM) 技術をマイクロプロセッサ (μP) のキャッシュに応用した場合の効果について議論を行う。はじめに、不揮発/擬似不揮発記憶を用いたパワーゲーティング (PG) を評価するためのシステムモデルを構築し、その電力削減率と面積増加率から不揮発/擬似不揮発記憶の導入による効果を評価する。また、システムの電源遮断頻度のモデリングを行い、リーク電力の削減に対する損益分岐時間 (BET) の効果を議論する。以上から、不揮発/擬似不揮発記憶を用いた PG の意義を明らかにする。そして、これまで本論文で述べた NV-SRAM および VNR-SRAM の性能・特徴から、これらを総合評価する。

5.1 不揮発/擬似不揮発記憶を用いた PG の性能評価

不揮発/擬似不揮発記憶を用いた PG の検証には、ロジックシステムの具体的なベンチマークが必要になる。本論文の対象は第1章で述べたようにサーバや PC に用いるマイクロプロセッサ (μP) であることから、 μP におけるキャッシュに不揮発/擬似不揮発記憶を導入した場合のベンチマークが重要になる。一般に μP におけるキャッシュのベンチマークの解析では、まず、様々なアプリケーションを想定したキャッシュの利用状況を専用のプログラムを用いて解析し、次に、この得られた利用状況を用いて、キャッシュの素過程の電力、エネルギー、レイテンシなどを用いて、具体的な性能を検証する(これらは CACTI など一般的なエミュレータから見積もることができる)。不揮発/擬似不揮発記憶を用いた場合では、さらに BET (T_{BE}) と PG による電力削減率が重要なパラメータとなる。キャッシュの PG についてはこれを検証できるベンチマークプログラムはまだ存在しないことから、これらのパラメータを用いて、不揮発/擬似不揮発記憶 (NV-/VNR-SRAM) を用いた PG を検証する。

はじめに、不揮発/擬似不揮発記憶を用いた SRAM 技術を適用したレジスタ(後述の NV-/VNR-FF で構成), キャッシュ(NV/VNR-SRAM で構成)を用いた μP のモデリングを行い, PG によるリーク電力削減率($\gamma_{\text{PG}}^{\text{system}}$)の定式化を行う. また, 従来の μP に対する, 不揮発/擬似不揮発記憶を用いた μP の面積増加率も定式化する. 次いで, これらの指標を用いて, 不揮発/擬似不揮発記憶を用いた PG を実行する場合の電力削減効果と面積増加率について評価する. なお, 本章で用いる回路構成(各種 SRAM, FF, 周辺回路等)におけるリーク電力, エネルギーオーバーヘッド, レイテンシ, 面積等のパラメータを表 5.1.1 にまとめた. 以降ではこの表の数値を使用する.

5.1.1 CMOS ロジックシステムのリーク電力削減率と面積増加率

μP におけるリーク電力とその削減率はキャッシュ, レジスタ, ロジックなどといったシステムの構成に依存する. 図 5.1.1(a)に本論文で用いる従来の μP のシステムモデルを示す. このシステムはコアとラストレベルキャッシュ(LLC)で構成され, コアはファーストレベルキャッシュ(FLC), ステイト・リテンション用のレジスタ(REG), ロジックゲート(LG)からなる. 本章では簡単のため, コア内の高階層キャッシュは FLC のみで構成されるものとした.

ここでは図 5.1.1(a)で示す従来のシステムにおいて, ステイト・リテンション用のレジスタを後述のデュアルパワーレールを用いたリテンション FF[1](これを以後 DP リテンション FF と呼ぶことにする)で構成し, LLC, FLC を 6T-SRAM で構成したシステムとする. LG については, NAND ゲートがロジック領域のトランジスタ密度を評価するときの指標として用いられることがあるため[2], これを用いてリーク電力を評価した. 本論文の対象はサーバ, PC に用いる μP であることから, システム構成は実際に報告された μP を参考にして決定した. LLC, FLC のチップの全面積に対するキャッシュの占有率は Intel 社のプロセッサにおける典型的な報告値を参考に, LLC の占有率を 40%から 70%, FLC の占有率を 10%とする[3-9]. ただし, LLC の占有率を固定する必要がある場合には 50%とした. また, REG を構成する FF については, 占有率のコアの 20%とした.

次に従来のシステムのリーク電力の定式化について述べる. 上述のシステムモデルにおけるリーク電力(P_{total})は次式で表せる.

$$P_{\text{total}} = \frac{S_{\text{LLC}}}{S_{\text{LLC}}^{\text{C}}} P_{\text{LLC}}^{\text{C}} + \frac{S_{\text{FLC}}}{S_{\text{FLC}}^{\text{C}}} P_{\text{FLC}}^{\text{C}} + \frac{S_{\text{REG}}}{S_{\text{REG}}^{\text{C}}} P_{\text{REG}}^{\text{C}} + \frac{S_{\text{LG}}}{S_{\text{LG}}^{\text{C}}} P_{\text{LG}}^{\text{C}} \quad (5.1)$$

$P_{\text{Circ}}, S_{\text{Circ}}$ (Circ=LLC, FLC, REG, LG)はそれぞれ LLC, FLC, REG, LG の総電力と総面積を示す. 上付きの添え字の C はシステムの各回路を構成する単位セルの電力と面積を表し, $S_{\text{Circ}}/S_{\text{Circ}}^{\text{C}}$ は単位セルの個数を表す. P_{total} はそれぞれの単位セルの個数と単位セルのリーク電力の積を足し合わせたものであり, システム全体のリーク電力となる. 式(5.1)はシステムの全面積と LLC, FLC, REG, LG の占有率を用いて次式のように書き換えられる.

$$P_{\text{total}} = \left\{ \frac{R_{\text{LLC}}}{S_{\text{LLC}}^{\text{C}}} P_{\text{LLC}}^{\text{C}} + (1 - R_{\text{LLC}}) \cdot \left(\frac{r_{\text{FLC}}}{S_{\text{FLC}}^{\text{C}}} P_{\text{FLC}}^{\text{C}} + \frac{r_{\text{REG}}}{S_{\text{REG}}^{\text{C}}} P_{\text{REG}}^{\text{C}} + \frac{r_{\text{LG}}}{S_{\text{LG}}^{\text{C}}} P_{\text{LG}}^{\text{C}} \right) \right\} \quad (5.2)$$

R_{LLC} はシステムの全面積 (S_{total}) に対する LLC の占有率 ($=S_{\text{LLC}}/S_{\text{total}}$), $r_{\text{FLC}}, r_{\text{REG}}, r_{\text{LG}}$ はコアの全面積 ($S_{\text{core}}=(1-R_{\text{LLC}}) \cdot S_{\text{total}}$) に対する FLC, REG, LG の占有率 ($=S_{\text{FLC}}/S_{\text{core}}, S_{\text{REG}}/S_{\text{core}}, S_{\text{LG}}/S_{\text{core}}$) とした. 式(5.2)を用いて, 従来のシステムにおけるリーク電力削減率 $\gamma_{\text{PG}}^{\text{system}}$ は, 式(5.2)を用いて次式で表せる.

$$\gamma_{\text{PG}}^{\text{system}} = \frac{P_{\text{total}}^{\text{SB}} - P_{\text{total}}^{\text{SDULVR}}}{P_{\text{total}}^{\text{SB}}} = 1 - \frac{\frac{R_{\text{LLC}}}{S_{\text{LLC}}^{\text{C}}} P_{\text{LLC}}^{\text{C,mode}} + (1 - R_{\text{LLC}}) \cdot \left(\frac{r_{\text{FLC}}}{S_{\text{FLC}}^{\text{C}}} P_{\text{FLC}}^{\text{C,mode}} + \frac{r_{\text{REG}}}{S_{\text{REG}}^{\text{C}}} P_{\text{REG}}^{\text{C,mode}} + \frac{r_{\text{LG}}}{S_{\text{LG}}^{\text{C}}} P_{\text{LG}}^{\text{C,mode}} \right)}{\frac{R_{\text{LLC}}}{S_{\text{LLC}}^{\text{C}}} P_{\text{LLC}}^{\text{C,SB}} + (1 - R_{\text{LLC}}) \cdot \left(\frac{r_{\text{FLC}}}{S_{\text{FLC}}^{\text{C}}} P_{\text{FLC}}^{\text{C,SB}} + \frac{r_{\text{REG}}}{S_{\text{REG}}^{\text{C}}} P_{\text{REG}}^{\text{C,SB}} + \frac{r_{\text{LG}}}{S_{\text{LG}}^{\text{C}}} P_{\text{LG}}^{\text{C,SB}} \right)} \quad (5.3)$$

上付きの添え字には, 前述の C に加えて回路状態 (mode) を記載している. 本論文では電力削減率の基準を従来システムの通常電圧におけるスタンバイ時 (SB) のリーク電力としているため, 分母は常にこのリーク電力を用いる (上付きの添え字の SB は通常電圧のスタンバイを示す). 一方, 分子は評価する回路状態 (mode=SB, SLP, SD) におけるシステムの各回路のパラメータを用いる (表 5.1.1 参照). 従来のシステムでは, システムの PG 時に LLC はスリープモード, FLC は電源遮断 (第 1 章で述べたフラッシュ), REG は通常電圧のスタンバイモードあるいはスリープモード (本来 FF に SRAM のスリープ技術を適用することはないが, 比較のため, あえてこれを導入した場合の電力も検討した), LG は電源遮断とした. 式(5.3)に表 5.1.1 から対応する値を選び代入することで $\gamma_{\text{PG}}^{\text{system}}$ を算出できる.

続いて図 5.1.1(b)で示す不揮発/擬似不揮発記憶を用いたシステムの構成とリー

ク電力削減率について述べる. このシステムは, 従来のシステムから, LLC, FLC, REG をそれぞれ, NV-/VNR-SRAM, NV-/VNR-FF にセル数を一定に保ちながら置換して構成する. システムの各回路のセル数は一定であるため, 従来のシステムの SRAM, FF の個数と置換後のシステムの SRAM, FF の個数は等しい. このため, 式(5.3)の分子の $P_{\text{Circ}}^{C,\text{mode}}$ にのみ, 不揮発/擬似不揮発記憶を用いたシステムの各回路を構成するセルのリーク電力を代入すればよい. システムの電源遮断/ULV リテンション時には NV-SRAM/FF は電源遮断, VNR-SRAM/FF は ULV リテンションとし, LG はいずれの場合も電源遮断する. なお, 次節の REG の検討では従来技術の一つとしてバルーン FF を評価するが, この場合も式(5.3)の DP リテンション FF をバルーン FF に置換することで評価している.

続いて, システムの面積の増加率 (ΔS_{total}) について定式化する. 不揮発/擬似不揮発記憶を用いた SRAM, FF (および従来技術であるバルーン FF) は従来の 6T-SRAM, DP リテンション FF に比べてトランジスタ数の増加により面積が増大する (表 5.1.1 参照). このため, 従来のシステムにおける 6T-SRAM, DP リテンション FF をその他の SRAM, FF で置き換えた場合にシステム全体の面積が増大する. このとき, システムの面積増加率 (ΔS_{total}) は次式のように表せる.

$$\begin{aligned} \Delta S_{\text{total}} &= \frac{S'_{\text{LLC}} + S'_{\text{FLC}} + S'_{\text{REG}} + S'_{\text{LG}}}{S_{\text{LLC}} + S_{\text{FLC}} + S_{\text{REG}} + S_{\text{LG}}} - 1 \\ &= \frac{\gamma_{\text{LLC}}^{\text{area}} R_{\text{LLC}} + (1 - R_{\text{LLC}})(\gamma_{\text{FLC}}^{\text{area}} r_{\text{FLC}} + \gamma_{\text{REG}}^{\text{area}} r_{\text{REG}} + \gamma_{\text{LG}}^{\text{area}} r_{\text{LG}})}{R_{\text{LLC}} + (1 - R_{\text{LLC}})(r_{\text{FLC}} + r_{\text{REG}} + r_{\text{LG}})} - 1 \end{aligned} \quad (5.4)$$

$\gamma_{\text{LLC}}^{\text{area}}$, $\gamma_{\text{FLC}}^{\text{area}}$, $\gamma_{\text{REG}}^{\text{area}}$ は従来の 6T-SRAM, DP リテンション FF に対する置換後のセル面積の増加率である.

以上の式(5.3), (5.4)を用いて, 次節以降はシステム内の記憶回路がシステムのリーク電力削減率 ($\gamma_{\text{PG}}^{\text{system}}$) と面積増加率 (ΔS_{total}) に与える影響を解析する.

5.1.2 不揮発/擬似不揮発性 FF の導入による効果

(i) システムの構成と各種リテンション FF の回路構成と動作原理

まず, コア内の REG が $\gamma_{\text{PG}}^{\text{system}}$ に与える影響を検証する. REG を構成する FF についても, 本論文で開発した不揮発/擬似不揮発性 SRAM 技術を応用することで電力削減率をさらに改善できる可能性がある. REG を構成する FF は DP リテンション FF, バルーン FF[10], NV-FF[11], VNR-FF[12]を検討する. LLC とコア内の FLC は通常の

6T-SRAM で構成し、LG は NAND ゲートを用いてリーク電力を評価した。

次に上述の 4 種類のリテンション FF の回路構成と動作について述べる。図 5.1.2 は従来技術である DP リテンション FF の構成を示す。DP リテンション FF は通常の Delay Flip Flop (DFF) をデュアルパワーレールにして構成する。コアの電源遮断時に、DP リテンション FF には独立の電源線を用いてパワースイッチを介さずに電源を供給するため、電源遮断時にもデータを保持することができるが、このデータの保持に大きな電力を消費するためコアの PG のリーク電力削減率が劣化する。図 5.1.3 は従来技術の別のリテンション FF であるバルーン FF の構成を示す。これは、通常の FF 部のスレーブラッチの記憶ノードにパスゲートを介して、バックアップラッチを接続した構成である。バックアップラッチには、コアや通常の FF 部とは別の電源線を用いて電源供給するため、電源遮断時にもデータを保持できる。また、この場合上述の DP リテンション FF とは異なり、バックアップラッチは通常の FF 動作における速度性能にはほぼ影響を与えないため、高しきい値かつ最小サイズのトランジスタで構成することができる。その結果、パワードメインの電源遮断時のリーク電力を大きく削減できる。しかし、電源遮断前後の通常の DFF 部とバックアップラッチとのデータ転送のタイミングに関する制御が複雑になる上、信号の切り替えにドライバが必要となるため面積が増加するという課題がある。このように、従来のリテンション FF は一長一短がある。

上述の DP リテンション FF、バルーン FF における課題は、NV-/VNR-SRAM 技術を FF に応用した NV-FF、VNR-FF を用いることで解決できる可能性がある。図 5.1.4 に NV-FF の回路構成を示す。NV-FF は NV-SRAM と同様にスレーブラッチの記憶ノードにパストラジスタを介して Magnetic tunnel junctions (MTJ) を接続した構成であり、電源遮断前にスレーブラッチのデータを MTJ に書き込み、復帰時に MTJ からスレーブラッチにデータを書き戻す。完全に電源遮断できるため、4 種類のリテンション FF の中で最もリーク電力を削減できるが、MTJ のエンベディッド技術が必要となる。図 5.1.5 は VNR-FF の回路構成を示している。スレーブラッチを pFBPD 型のデュアルモード (DM) インバータ(第 4 章参照)で構成することにより ULV リテンションを実現できる。VNR-FF は ULV リテンションにより、データを保持しながら電源遮断と同等の電力削減率を実現できる。また、CMOS のみで構成でき、制御も極めて簡単である。図 5.1.1 から図 5.1.5 の数値は各トランジスタのチャネル幅を示す。各リテンション FF の設計は Appendix に示す。

(ii)各種リテンション FF を用いたシステムのリーク電力削減率と面積増加率

図 5.1.6(a)に上述の 4 種類のリテンション FF を用いた場合の γ_{PG}^{system} の r_{REG} 依存性を示す。 r_{REG} は従来のシステム (6T-SRAM, DP retention FF, NAND) の値を用いてプロットしている。PG を実行する場合、DP リテンション FF は通常電圧のスタンバイモード (黒の実線) と $V_{DD}=0.8V$ まで電源電圧を落としたスリープモード (黒の点線) を検討した。バルーン FF では別の電源線に接続されたバックアップラッチによるデータ保持, NV-FF は電源遮断, VNR-FF は ULV リテンションとした。LLC/FLC は 6T-SRAM で構成してスリープ/フラッシュとし、ロジックゲートは NAND を用いて電源遮断とした。LLC/FLC の占有率はそれぞれ 50%/10%とした。図 5.1.6(a)から DP リテンション FF を用いた場合、リーク電力が大きいため、 r_{REG} が増加すると γ_{PG}^{system} が大きく劣化する (スリープモードを導入してもこの劣化は回避できない)。これに対してバルーン FF, NV-FF, VNR-FF を用いた場合は、いずれもほぼ同等の削減率を実現できている。バルーン FF はバックアップラッチを高いしきい値の最小トランジスタで構成するため電源遮断を行わずとも、電源遮断可能な NV-FF, VNR-FF と同程度にリーク電力を最小限におさえることができる。

次に図 5.1.6(b)に ΔS_{total} の r_{REG} 依存性を示す。バルーン FF は通常の FF 部とバックアップラッチとの接続の切り替えに 4 つのドライバが (図 5.1.3 参照), NV-FF はスレーブラッチと MTJ の接続の切り替えに 2 つのドライバが必要になり (図 5.1.4 参照), いずれもパワースイッチの信号制御以外に別途信号が必要となるために DP リテンション FF に比べてセルの面積増加率が大きい。一方、VNR-FF は第 4 章で述べたように V_{DD} の変化のみでモード切替が可能であり、FB トランジスタに一定バイアスで固定するため (バイアスの切り替えが必要ないので) ドライバが不要となる。したがって、VNR-FF はバルーン FF, NV-FF に比べて面積増加率を小さく抑えることが可能である。

以上の結果から、各リテンション FF の特徴をまとめる。DP リテンション FF は最も容易に構成でき面積も小さく抑えられるが、リーク電力が大きく、PG 時のリーク電力削減率が劣化する。また、バルーン FF を用いた場合、リーク電力削減率のみを考慮すると十分な電力削減が可能になるが、パワースイッチ制御の他に複数の制御信号を用いることや、バックアップラッチとスレーブラッチのデータの転送・復帰に精密な制御信号のタイミング設計を必要とするという課題がある。NV-FF もバルーン FF と同様に、信号の切り替えにドライバが必要になるために面積が増加する (ただし、完全な電源遮断が実現できる

ため、バルーンFFに比べてリーク電力削減率はさらに削減される)。この点、VNR-FFはパワースイッチ以外の制御信号が不要となるので制御が容易であり、上述のバルーンFF、NV-FF に比べて面積が削減され、ULV リテンションにより大幅に電力を削減できる。

5.1.3 不揮発/擬似不揮発性 SRAM の導入による効果

続いて、キャッシュ(特に LLC)が γ_{PG}^{system} に与える影響を解析した。図 5.1.7(a)はステイト・リテンション用のレジスタを VNR-FF、ロジックゲートを NAND とし、LLC、FLC を 6T-SRAM、VNR-SRAM、NV-SRAM を用いて構成したシステムの γ_{PG}^{system} の R_{LLC} 依存性を示している。各種 SRAM はいずれも第 3 章、第 4 章で示した 6T セル、NV セル、ヘッダパワースイッチ構成の pFB セルを用いて構成した。まず、黒の実線は従来の 6T-SRAM で LLC と FLC を構成し、LLC を通常電圧のスタンバイモードとしてコアのみ電源遮断した場合の γ_{PG}^{system} を示している。この場合、 R_{LLC} が 40%から 70%の範囲(典型的なシステムに対する占有率)ではおおよそ 40%から 70%程度の電力削減となる。また、ここからさらに LLC をスリープモードとした場合の γ_{PG}^{system} を黒の点線で示す。スリープによりさらに削減効率が向上するが、スリープを導入したとしても電力削減率を究極的に高めることはできず、まだ削減の余地がある。一方、ピンクと緑の実線はそれぞれキャッシュを VNR-SRAM、NV-SRAM で構成したシステムで PG を実行した場合の γ_{PG}^{system} を示している。この場合、コア内の FLC は VNR-SRAM では ULV リテンション、NV-SRAM では電源遮断としており、(6T-SRAM のフラッシュとは異なり)データを保持できるため PG を頻繁に行うことが可能である。さらに、LLC を ULV リテンションあるいは電源遮断することにより究極的に電力を削減することができる。図 5.1.7(b)は ΔS_{total} の R_{LLC} 依存性を示す。VNR-SRAM は NV-SRAM よりもセル面積が増大するため、チップ面積も NV-SRAM に比べて 10%程度増加する。しかし、第 4 章で述べた TFT でフィードバックトランジスタを構成する技術[13]を適用すると、同図の点線のようにチップ面積の増加率を大きく削減できる可能性がある(ただしこの場合、チップコストは増大する)。

5.2 PG の平均リーク電力に対する BET の効果

続いて、システムの T_{BE} について議論する。本章では簡単のためシングルコアのシステムを仮定して議論する。まずはシステムの非実行時間についてモデリングし、その後、不揮発/擬似不揮発記憶を用いた PG の平均リーク電力に対する T_{BE} の効果について述べる。

5.2.1 非実行時間のモデリング

図 5.2.1 はシステムの通常動作時、スタンバイ時、電源遮断あるいは ULV リテンション時の電力の推移の模式図である。EXE は通常動作の実行時間を、SB はスタンバイ、SD は電源遮断あるいは ULV リテンションを表す。以後、ULV リテンションする場合も含めて電源遮断と呼ぶことにする。これまでは通常電圧のスタンバイ時(SB)のリーク電力に対する電源遮断時(SD)のリーク電力の削減率を議論してきた。この削減率は図 5.2.1 の γ_{PG}^{system} に相当する。しかし、システムの動作には様々な非実行時間(処理を行わないため PG を実行できる可能性のある時間)が存在し、すべての非実行時間で電源遮断するわけではなく、その長さが T_{BE} より長いときのみ電源遮断が有効となる。したがって、電源遮断せずにスタンバイモードとする場合のリーク電力と電源遮断する場合のリーク電力を平均したシステム全体の平均リーク電力 ($P_{L,ave}^{system}$) も重要となる。 $P_{L,ave}^{system}$ を評価するには、システムの動作を解析し、その非実行時間がどのように分布しているかを調べる必要がある。

そこで、まずはシステムの非実行時間のモデリングを行う。各非実行時間(図 5.2.1 の SB と SD) をスタンバイ・電源遮断に関わらず τ_{SBi} ($i=1, 2, 3, \dots$) とし、各非実行時間の出現回数を N_{SBi} とする。このとき、非実行時間の総和 τ_{SB}^{total} は次式で表せる。

$$\tau_{SB}^{total} = \sum \tau_{SBi} \cdot N_{SBi} \quad (5.5)$$

τ_{SBi} が T_{BE} より長ければ電源遮断し、 τ_{SBi} が T_{BE} より短ければ電源遮断せずに通常電圧のスタンバイモードとする。平均のリーク電力 $P_{L,ave}^{system}$ は電源遮断時と通常電圧のスタンバイ時のリーク電力を評価し、それぞれの割合を用いることで求めることができる。この各非実行時間の割合を評価するため、本論文ではその出現回数を与える密度分布 (n_{SB}) は簡単のため、正規分布でモデリングした。本論文では正規分布を用いたが、分布の中心値と広がり記述できる他の密度分布を用いても同様の結果が得られ、定性

的な結果は変わらない。また、任意の分布は正規分布の重ね合わせを用いて表現できるため、結果の一般性は保たれていると考えられる。したがって、正規分布を用いてモデリングを行った。以上から、式(5.5)は次式のように表せる。

$$\tau_{SB}^{\text{total}} = \int \tau_{SB} \cdot n_{SB}(\tau_{SB}) d\tau_{SB} \quad (5.6)$$

図 5.2.2 に $n_{SB}(\tau_{SB})$ の概形を示す。分布の中心は μ 、分布の広がりには σ とした。また、以下の解析ではすべて、 τ_{SB}^{total} を 1s とした。この正規分布と T_{BE} の関係で、次節で述べるように PG を実行するかどうか判断できる。

5.2.2 キャッシュの平均リーク電力

まずはキャッシュのみを考慮した場合の平均リーク電力を解析する。図 5.2.3 は非実行時間の分布を、 μ を変化させて模式的にプロットしたものである(この図は σ が十分小さい場合を示している)。キャッシュの T_{BE} (T_{BE}^{cache}) も点線で示している。それぞれ、正規分布の μ が T_{BE}^{cache} よりも十分小さい場合 ($\mu = \mu_1$)、 T_{BE}^{cache} と同程度の場合 ($\mu = \mu_2$)、 T_{BE}^{cache} よりも十分大きい場合 ($\mu = \mu_3$) を示している。 μ が増加するにつれて、正規分布の高さが低くなるのは、 τ_{SB}^{total} を 1s で一定としたためである(式(5.6)参照)。すなわち、非実行時間 τ_{SB} が短い場合非実行時間の出現回数が多く、非実行時間 τ_{SB} が長い場合出現回数が少なくなる。この正規分布と T_{BE}^{cache} の位置関係により、PG を実行するかどうかを判断する。正規分布が T_{BE}^{cache} よりも十分に短い領域に存在する場合にはすべてスタンバイモード、 T_{BE}^{cache} 付近の領域では一部をスタンバイモード、残りを電源遮断、 T_{BE}^{cache} より十分に長い領域ではすべて電源遮断する。したがって、 T_{BE}^{cache} に対する μ の位置が重要となり、さらに、後述のように分布の広がりを表す σ も重要な役割をもつ。 μ や σ はアプリケーションやプロセスにより異なるため様々な値をとる可能性があるが、正規分布でモデリングすることで、この 2 変数をパラメータにして様々なアプリケーションを想定して解析することができる。

以上のモデリングを用いるとキャッシュの平均リーク電力 ($P_{L,\text{ave}}^{\text{cache}}$) は次式のように表せる。

$$P_{L,\text{ave}}^{\text{cache}} = \frac{E_{SB} + E_{SD}}{\tau_{SB}^{\text{total}} + \tau_{LTC}^{\text{total}}} =$$

$$\frac{\int_0^{T_{BE}^{cache}} n_{SB}(\tau_{SB}) \cdot P_{L,SB}^{cache} \cdot \tau_{SB} d\tau_{SB} + \int_{T_{BE}^{cache}}^{\infty} n_{SB}(\tau_{SB}) \cdot (\Delta E_{EE} + P_{L,SD}^{cache} \cdot \tau_{SB}) d\tau_{SB}}{\int_0^{T_{BE}^{cache}} n_{SB}(\tau_{SB}) \cdot \tau_{SB} d\tau_{SB} + \int_{T_{BE}^{cache}}^{\infty} n_{SB}(\tau_{SB}) \cdot \tau_{LTC}^{cache} d\tau_{SB}} \quad (5.7)$$

E_{SB} はスタンバイ時 (T_{BE}^{cache} よりも短い非実行時間) のリークエネルギーの総和を, E_{SD} は PG 実行時 (T_{BE}^{cache} よりも長い非実行時間) のリークエネルギーと, PG の実行に要するエネルギーオーバーヘッド (ΔE_{EE}) の総和を表す. これらのエネルギーの, PG の実行に要する総時間 (待機時間と ENT/EXT 動作時のレイテンシの総和) における平均値を $P_{L,ave}^{cache}$ としている.

図 5.2.4(a)-(c)に, 上述のモデリングを用いて計算した 2MB の NV-/VNR-SRAM の $P_{L,ave}^{cache}$ の μ 依存性を示す (それぞれ自身の通常電圧のスタンバイ電力で規格化してある). キャッシュの電力 ($P_{L,SB}^{cache}$, $P_{L,SD}^{cache}$), ΔE_{EE} , PG の実行に要するレイテンシ (τ_{LTC}^{cache}) については, 第 3 章, 第 4 章で示した 2MB アレイの計算結果を用いた (ここでは $R_{skip}=0\%$ を仮定しているが, $R_{skip}=50\%$ としても以降の結果の傾向は変わらない). 赤線と緑線はそれぞれ VNR-SRAM と NV-SRAM の $P_{L,ave}^{cache}$ の解析結果を示す. また, 赤の点線で VNR-SRAM の, 緑の点線で NV-SRAM の T_{BE}^{cache} を示している (本章では T_{BE}^{cache} には T_{BE}^{EE} の値を用いた). まず, 図 5.2.4(a) は分布の広がり σ を $0.1\mu s$ (VNR-SRAM の T_{BE}^{cache} の $1/100$ 程度) としたときの μ 依存性である. このとき, VNR-SRAM, NV-SRAM はいずれも, 正規分布が T_{BE}^{cache} を超えると電源遮断できるようになるため, T_{BE}^{cache} 付近で急激に平均電力は削減される. 次に図 5.2.4(b) に σ を $10\mu s$ (VNR-SRAM の T_{BE}^{cache} 程度) まで増加させたときの μ 依存性を示す. VNR-SRAM では, σ が T_{BE}^{cache} 程度になると, μ が T_{BE}^{cache} を横切る前から平均電力が削減でき, μ が十分小さくても電力を削減できる. 一方で, この σ は NV-SRAM の T_{BE}^{cache} の $1/100$ 程度であり, T_{BE}^{cache} より十分小さいため, NV-SRAM では先ほどと同じような挙動を示す. 図 5.2.4(c) に, さらに σ を増加させ $2ms$ (NV-SRAM の T_{BE}^{cache} 程度) としたときの μ 依存性を示す. この場合, NV-SRAM でも μ が T_{BE}^{cache} を横切る前から平均電力が削減でき, μ が十分小さくても電力を削減できるようになる. しかし, この時, この σ は VNR-SRAM の T_{BE}^{cache} の 100 倍程度であるため, VNR-SRAM はさらに電力を削減でき, μ によらず十分に削減できる. T_{BE}^{cache} より μ が小さい領域で平均リーク電力が削減されているのは, 図 5.2.5 のように σ が増加することにより正規分布の裾が T_{BE}^{cache} を超え, 電源遮断が有用となる非実行

時間の割合が増加するためである。つまり、 T_{BE}^{cache} が短いと、分布の中心が T_{BE}^{cache} を超えていなくても、分布の広がりによっては十分に電力を削減できることがわかる。

5.2.3 システムの平均リーク電力

上述のキャッシュのみを考慮した場合と同様の議論をコアと LLC からなるシステムでも試みた。以下の解析では 5.1 節で示したシステムモデルを用いる。ただし、マルチコアプロセッサの場合には各コアの遮断状況に依存し、複雑なアーキテクチャが必要となるため、ここでは簡単のためシングルコアとした。VNR-SRAM で構成したコア、NV-SRAM で構成したコアは、いずれもコア内のステイト・リテンション用のレジスタは VNR-FF で構成し、ロジックゲートは NAND を用いた。また、LLC と FLC についてはそれぞれ NV-SRAM と VNR-SRAM で構成し、コアの T_{BE} (T_{BE}^{core}) については、FLC, REG, LG を構成する 1 セルのエネルギーオーバーヘッド、リーク電力、占有率から評価し、VNR-SRAM, NV-SRAM を用いたコアでそれぞれ $0.8\mu s$ 程度、 $8.8\mu s$ とした。また、LLC の T_{BE} (T_{BE}^{LLC}) については、前節の値を用いた。

このシステムモデルで PG を実行する場合、まずコアにおいて PG を実行し、その後可能であれば LLC においても PG を実行する。図 5.2.6 は上述のシステムで PG を行う場合について非実行時間の分布を、 μ を変化させて模式的にプロットしたものである (図 5.2.3 と同様に σ が十分小さい場合を模式的に示している)。 μ が T_{BE}^{core} よりも短い領域 ($\mu=\mu_1$) ではコアと LLC はいずれもスタンバイモードとする。 μ が T_{BE}^{core} よりも長く、 T_{BE}^{LLC} よりも短い領域 ($\mu=\mu_2$) にある場合はコアのみ電源遮断できる。 T_{BE}^{LLC} よりも長い領域にある場合 ($\mu=\mu_3$) はコアと LLC をいずれも電源遮断する。この場合、キャッシュのみを考慮した場合と異なり、電源遮断する判断の基準となる T_{BE} の値が 2 種類存在し、コアと LLC をいずれもスタンバイとする場合、コアのみ電源遮断する場合、コアと LLC をいずれも電源遮断する場合の 3 つの状態が存在する。以後それぞれ (SB,SB), (SD,SB), (SD,SD) と呼ぶことにする。システムの平均リーク電力は次式のように表せる。

$$P_{L,ave}^{system} = \frac{E_{(SB,SB)} + E_{(SD,SB)} + E_{(SD,SD)}}{\tau_{SB}^{total} + \tau_{LTC}^{total}} =$$

$$\begin{aligned}
& \int_0^{T_{BE}^{core}} n_{SB}(\tau_{SB}) \cdot (P_{SB}^{core} + P_{SB}^{LLC}) \cdot \tau_{SB} d\tau_{SB} + \\
& \int_{T_{BE}^{core}}^{T_{BE}^{LLC}} n_{SB}(\tau_{SB}) \cdot (P_{SD}^{core} \cdot \tau_{SB} + \Delta E_{EE}^{core} + P_{SD}^{LLC} \cdot \tau_{SB}) d\tau_{SB} + \\
= & \frac{\int_{T_{BE}^{LLC}}^{\infty} n_{SB}(\tau_{SB}) \cdot (P_{SD}^{core} \cdot \tau_{SB} + \Delta E_{EE}^{core} + P_{SD}^{LLC} \cdot \tau_{SB} + \Delta E_{EE}^{LLC}) d\tau_{SB}}{\int_0^{T_{BE}^{core}} \tau_{SB} \cdot n_{SB}(\tau_{SB}) d\tau_{SB} +} \quad (5.8) \\
& \int_{T_{BE}^{core}}^{T_{BE}^{LLC}} \tau_{LTC}^{core} \cdot n_{SB}(\tau_{SB}) d\tau_{SB} + \int_{T_{BE}^{LLC}}^{\infty} (\tau_{LTC}^{core} + \tau_{LTC}^{LLC}) \cdot n_{SB}(\tau_{SB}) d\tau_{SB}
\end{aligned}$$

なお、上式はコアが電源遮断/復帰中の LLC のリークエネルギーと LLC の電源遮断/復帰中のコアのリークエネルギーを含まないが、これらはいずれも小さいため省略した。

図 5.2.7(a)-(c)に NV-/VNR-SRAM を用いたシステムの平均リーク電力 $P_{L,ave}^{system}$ の μ 依存性を示す。赤が VNR-SRAM を用いたシステムの、緑が NV-SRAM を用いたシステムの平均リーク電力を示している。また、赤の点線が VNR-SRAM を用いた場合の T_{BE}^{core} と T_{BE}^{LLC} 、緑の点線が NV-SRAM を用いた場合の T_{BE}^{core} と T_{BE}^{LLC} を示している。ここでは、 ΔE_{EE}^{core} は簡単のため 0 とし、 τ_{LTC}^{core} は $1\mu s$ とした。図 5.2.7(a)は NV-/VNR-SRAM の T_{BE} よりも σ が十分小さい場合の μ 依存性を示しており、前節のキャッシュのみの場合と同様に μ が T_{BE} を横切ると平均リーク電力が削減できる(若干のディップが生じているのはレイテンシが増加するためである)。また、図 5.2.7 の(b), (c)は同図(a)に比べて σ を増加させているが、こちらも前節と同様に σ が増加するにつれて NV-SRAM, VNR-SRAM はいずれも μ が T_{BE} を横切る前から電力の削減が見られ、 T_{BE} のより短い VNR-SRAM は NV-SRAM よりも電力の削減効果が大きい。したがって、システム全体での電力に関しても、 T_{BE} を短くすることで細粒度パワーゲーティングが実現できるだけでなく、 σ の広がりによっては平均リーク電力を効果的に削減することができる。ここで、NV-SRAM の T_{BE} については電流の基準値として下層配線の電流許容値を用いた場合の値を示しているが、仮に上層配線を用いた場合の結果(第 3 章の Appendix)を使用したとしても VNR-SRAM と同程度まで T_{BE} を削減することはできない。つまり、NV-SRAM の T_{BE}^{cache} はより短くなるが、VNR-SRAM の方が電力削減効果は大きく、図 5.2.7 の結論は変わらない。

5.3 不揮発/擬似不揮発性 SRAM の総合比較

最後に不揮発/擬似不揮発記憶を用いた PG の総合比較を行う。はじめに本論文における前提と比較を行うセルについて述べる。本論文では第 1 章で述べたように対象をサーバ、PC に用いる μP に設定した。このため、キャッシュに使える NV-/VNR-SRAM としては、GHz 動作が前提となる。したがって、本論文の第 3 章、第 4 章では、採用するアーキテクチャにはリファレンスとする 6T-SRAM を含めて、このような動作が可能なものに限定した。本節でもこの前提で議論をする。

SRAM は他の用途も多く、例えばしきい値を高く設定してリークをできるだけ抑制した構成などもあるが、これらの動作周波数は高くても数百 MHz 程度で、本論文の目的には適合しないので除外した。また、8T セルや 10T セルなどように低電圧動作に特化して提案された SRAM セルもあるが、これは PG に用いることは想定されておらず、セルのリテンションのメカニズムは 6T セルと同じである (8T, 10T などの低電圧セルは読み出し動作に特化して構成されている)。これらはパワースイッチを用いないので、リテンションの電圧は少し低めに設計できる。これにパワースイッチを加えて、PG に用いることは可能ではあるが、リテンションのメカニズムは 6T セルと同じため、低電圧でのリテンションの特性は本論文で示した 6T セルと同じか、またはトランジスタが増えているためリークは増えてしまう。したがって、このような低電圧セルも比較から除いた。NV-SRAM については、本論文のセル以外は NOF アーキテクチャ用に提案されたもので、PG に利用できるものは、本論文で用いたもののみである。NOF では第 3 章で述べたように、通常動作に不揮発記憶を用いることから、 μP に応用できるような高速動作が難しく、また、電源遮断時の電力削減率は同程度ではあるが、通常動作時の電力・エネルギーが極めて大きいことから、比較の対象とはしない。また、これまでに提案されたシュミットトリガインバータを用いた SRAM では、すべての動作を低電圧で行うことを前提としている。これを通常電圧下で使うと待機時電力が大きく増大するため、これを PG に用いることは意味がない。したがって、他のシュミットトリガインバータを用いた SRAM との比較は行わない。以上から本節では本論文で用いた構成の 6T-SRAM, NV-SRAM, VNR-SRAM について比較を行う。

表 5.3.1 に従来の 6T-SRAM, NV-SRAM, ヘッドパワースイッチ構成の pFB セルで構成した VNR-SRAM, デュアルパワースイッチ構成の pFB セルで構成した

VNR-SRAM の特徴を列記した。

プロセスについては NV-SRAM では MTJ のエンベディッド技術が必要になるが、VNR-SRAM では CMOS のみで構成可能である。セル面積は NV-SRAM では 6T-SRAM の 1.5 倍程度となる(これは現在用いられているデュアルポート SRAM と同程度である)。VNR-SRAM については 1.77 倍とやや増加する。システムの面積は NV-SRAM を用いた場合 1.3 程度、VNR-SRAM を用いた場合 1.4 倍程度となる。VNR-SRAM についてはフィードバックトランジスタを TFT で配線層に作りこむ技術[13] を用いることができればセル面積を 1.3 倍程度まで、システムの面積を 1.2 倍程度まで削減できる(ただし、コストが増大する)。NV-SRAM は MTJ のプロセスに係るマスクの増加により非常に高いコストが必要となるが、VNR-SRAM (Header PS, Dual PS) では面積の増加を考慮しても CMOS のみで構成できるという点で中～低コストである。

パワースイッチの構成は 6T-SRAM, NV-SRAM, VNR-SRAM はいずれもヘッダパワースイッチ構成とした。VNR-SRAM については、ヘッダパワースイッチとフッタパワースイッチを両方用いるデュアルパワースイッチ構成とすることで、リーク電力削減能力をさらに向上でき、これに伴うパワースイッチの面積オーバーヘッドは 10% 程度以下である。

セルの書き込み電力については、NV-/VNR-SRAM における増加率は 1.6% 程度となるが、第 3 章、第 4 章で述べた通り、この程度の増加であれば、アレイを構成したときに周辺回路の影響に隠蔽される、または周辺回路の工夫で容易に削減することができる。

電力削減率はセルの場合、6T-SRAM の 41% 程度に比べて、NV-/VNR-SRAM の電力削減率は 94%-99% 程度まで向上している。本章では 65nm CMOS を用いて、NV-/VNR-SRAM の比較を行ったが、さらにスケールリングされた CMOS や FinFET などの先端 CMOS でも、第 4 章で述べたように同様の結果が得られる。また、本論文の対象からは外れるが、最近では、トランジスタのしきい値が高く、よりリークの低い LSTP (Low standby leakage) デバイスもある。これは、速度より低リークが重要となるマイクロコントローラ (MCU) や一部の SoC で検討されている(本論文の対象としているようなサーバや PC に用いるマイクロプロセッサには応用は難しい)。65nm CMOS の LSTP デバイスを用いた NV-SRAM, VNR-SRAM の検討も行ったが、この場合、本論文で用いた LP デバイスと比べて電力の絶対値は大幅に減少するが、削減率という意味では同様の結果が得られる。したがって、しきい値の異なるデバイスについても同様の結果が得られる。ヘッダパワースイッチ構成の VNR-SRAM の電力削減率は NV-SRAM にやや劣るものの、

デュアルパワースイッチ構成とすることで NV-SRAM と同等まで電力削減率を向上できる。システムの電力削減率(γ_{PG}^{system})については、6T-SRAM が 73%程度の削減率にとどまるのに対して、NV-/VNR-SRAM では 97%-99%程度まで削減可能である。

NV-/VNR-SRAM では、第 2 章で述べたようにアレイサイズ(32kB-2MB)に依存するため、 T_{BE}^{EE} については 32kB-2MB の値の範囲を示した。また、6T-SRAM の値については T_{BE} の報告値を示している[7]。NV-SRAM の T_{BE}^{EE} は HSF, PUDF をいずれも用いた場合、 $R_{skip}=50\%$ と仮定すると 89.0 μ s-740 μ s 程度である。それに対し、VNR-SRAM では 1.67 μ s-16.7 μ s 程度であり、NV-SRAM に比べて 1 桁から 2 桁程度の削減が実現できる。VNR-SRAM では、究極的に短い T_{BE}^{EE} の実現により、5.2 節で示したように細粒度の PG が実現できる上、PG の平均リーク電力も効果的に削減することが可能となる。

以上の総合比較から、VNR-SRAM はセル(およびチップ)面積が増加するというデメリットがあるが、従来の 6T-SRAM と比べると従来のシステムでは到達できない程度まで低消費電力化を実現できるというはるかに大きいメリット(利点)がある。このことから、面積よりも低消費電力性が重視される μ P などの場合には特に有用である。

5.4 まとめ

本章のまとめを以下に示す。

- ・マイクロプロセッサにおける PG の効果を調べるため、システムにおけるリーク電力削減率と面積増加率の定式化を行った。
- ・コアのステイト・リテンション用のレジスタについて、従来技術である DP リテンション FF とバルーン FF を、本論文の SRAM 技術を応用した NV-/VNR-FF と比較し、VNR-FF が最も有用であることを示した。
- ・LLC に NV-/VNR-SRAM を用いることで CMOS ロジックシステム全体の電力を 97%程度以上削減できることを示した。
- ・不揮発/擬似不揮発記憶を用いたロジックシステムの非実行時間をモデリングし、平均リーク電力に対する BET の効果を明らかにした。
- ・NV-/VNR-SRAM の総合比較から今後のプロセッサ応用には VNR-SRAM が適していることを示した。

以上から CMOS ロジックシステムにおける不揮発/擬似不揮発記憶を用いた PG の性能を評価し、不揮発/擬似不揮発記憶を用いた PG の有用性・有効性を示した。

5.5 Appendix

(i)各種リテンション FF の設計

5.1.節で説明した各種リテンション FF について以下に設計方法と設計結果を示す。まず、設計の基準となる通常の DFF セルはスタンダードセルを参考に簡略化した設計を用いた(図 5.1.2)。これを標準セルと呼ぶことにする。このセルを基準に 4 種類の FF を設計した。また、以下の 4 種類のリテンション FF ではいずれも、パワースイッチ(PS)は速度性能・ V_{DD} が劣化しないよう十分に大きいサイズとした(いずれの PS も $W/L=30$: チャネル幅 $1.8\mu\text{m}$)。

DP リテンション FF については標準セルと同じトランジスタサイズとした、パワーレイルはデュアルパワーレイルで構成し、コアの電源遮断時はリテンション状態(通常電圧(1.2V)でデータ保持)あるいはスリープ状態(通常電圧の 7 割程度の電圧(0.8V)でデータ保持)とした。コア内の FF 部以外は PS を用いて電源遮断する。DP リテンション FF のデータ保持については、FF のスレーブラッチのみデータ保持する方法も考えられるが、ここでは簡単のために FF 全体に電源を供給してデータを保持する構成とした。

バルーン FF は標準セルのスレーブラッチの記憶ノードにパスゲートを介してバックアップラッチを接続し、データ転送の制御を行うためのドライバを設けた(図 5.1.3)。DP リテンション FF と同様に、パワーレイルはデュアルパワーレイルで構成し、バックアップラッチにはパワースイッチを介さずに独立に電源供給し、データ保持を行う。通常の FF 部は標準セルと同じ設計を用い、バックアップラッチおよび、ドライバについてはリーク電力を抑える必要があるため、最小サイズかつ高しきい値のトランジスタで構成した。

NV-FF は FF 部を含むコアを完全に電源遮断できるため、コア内のすべての領域を同じパワースイッチを用いて電源供給を行うことができる。このため、1 本の電源線と 1 つのパワースイッチを用いて電源供給を行う構成とした。こちらも通常の FF 部は標準セルの設計を用いた(図 5.1.4 参照)。MTJ への書き込みに用いる CTRL ドライバは十分な大きさに設計してある。NV-FF でも NV-SRAM と同様に MTJ の書き込み電流を精密に設計する必要があるため、MTJ へのパストランジスタのチャネル幅とチャネル長 (W_{PSS}/L_{PSS})で調整した。図 5.5.1 に H ストア時、L ストア時における MTJ への書き込み電流の W_{PSS}/L_{PSS} 依存性を示す(ワーストケースである SS コーナーのみ示してある)。同図は W_{PSS} を最小サイズの 100nm で固定し、 L_{PSS} を変化させている。パストランジスタのチャネル長を伸ばすことで MTJ への書き込み電流を最小限に抑えることができる。緑の実

線は MTJ の書き込み電流の下限となる臨界電流 (J_c) の 1.2 倍を示し、ここでは H ストアに必要な電流を満たすため、パストランジスタは $W_{PSS}/L_{PSS}=100\text{nm}/144\text{nm}=0.61$ とした。なお、ここでは MTJ の直径は $\phi_{\text{MTJ}}=60\text{nm}$ とした。

VNR-FF はスレーブラッチ以外を標準セルと同じ設計とし、スレーブラッチはノイズマージンを考慮してフィードバックパスをフォワードパスと同サイズとして DM インバータに置換した(図 5.1.5)。第 4 章の VNR-SRAM の設計方法と同様に、まず ULV リテンションにおけるスレーブラッチの安定性が重要となるため標準セルの設計を基準としてそこからスレーブラッチを構成するトランジスタのチャンネル幅を微調整して SNM が最大になるように設計した。図 5.5.2(a)-(c)に ULV リテンションの SNM の $\alpha_1, \alpha_2, \alpha_3$ 依存性を示す。図 5.1.5 における各チャンネル幅は調整パラメータ $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ を用いて $W_n=240a_1, W_p=260a_2, W_{FB}=240a_3, W_{p4}=260a_4, W_{n4}=240a_4, W_{p5}=260a_5, W_{n5}=240a_5$ とした。同図には比較用の標準セル(紫線と緑線)と VNR-FF (赤線と青線)についてワーストケースとなりうる FS コーナーと SF コーナーの SNM を示している。FF のスレーブラッチは SRAM と異なり、左右の記憶ノードに接続された回路構成が非対称であるため、 $(V_{NS}, V_{NS})=(L, H)$ の場合と $(V_{NS}, V_{NS})=(H, L)$ の場合でバタフライカーブのローブが異なる。前者の場合を実線で、後者を点線で示してある。このとき、マスタラッチ-スレーブラッチ間のパスゲートを介するリーク電流が ULV リテンションの SNM を劣化させるため、これは高しきい値のトランジスタで構成している。図 5.5.2(a)-(c)から $\alpha_1=1.6, \alpha_2=0.5, \alpha_3=0.6$ でワーストケースの SNM を最大化することができ、いずれも目標値である 78mV 以上の SNM を実現できる。続いて、マスタラッチ-スレーブラッチ間のパスゲートとマスタラッチのフォワードパスのインバータのチャンネル幅(設計パラメータ α_4)を微調整することで、リーク電力とクロック-出力遅延を調整した。また、図 5.5.3(a), (b)にリーク電力とクロック-出力遅延の α_4 依存性を示す。いずれも、 V_{FB} を 1.2V とし、通常電圧での動作時に ST モードで駆動した場合の結果も点線で示してある。ST モードで駆動した場合リーク電力が大幅に増大するが、BI モードで動作させることによりリーク電力を大きく削減できる。また、遅延は出力 Q が H レベルに遷移する場合の遅延については BI モードを用いることで半分程度に削減することができる(L レベルに遷移する場合はフィードバックトランジスタの効果が影響しないので、BI モードと ST モードの遅延はほぼ等しい)。また、 $\alpha_4 > 1.5$ の領域ではトランジスタのチャンネル幅が長く、トランジスタを折り返して並列にレイアウトする必要があり、配線遅延が増大してしまう。以上から、図 5.5.3(a), (b)から $\alpha_4=1.5$ とすることで、遅延

を最小に、リーク電力を標準セルと同程度とすることができる。最後に入力ドライバのチャネル幅(設計パラメータ α_5)を調整してセットアップタイム(τ_{setup})とホールドタイム(τ_{hold})を調整した。図 5.5.3(c)に τ_{setup} , τ_{hold} の依存性を α_5 示す。 τ_{hold} は負の値をとるため絶対値を示している。負の値の τ_{hold} は設計上問題がないため、 τ_{setup} を考慮して $\alpha_5=1.0$ とした。以上から、ULV リテンションの SNM, リーク電力, クロック-出力遅延, セットアップ/ホールドタイムをいずれも最適化した VNR-FF を設計した。PS は、スレーブラッチについては VNR-SRAM と同様に構成し、1.2V と 0.2V の 2 本の電源線から 2 つのパワースイッチを用いて V_{DD1} を生成し、コアの電源遮断時は ULV リテンションとした。スレーブラッチ以外の回路は別のパワースイッチを用いて 1.2V の電源線から電源を供給して V_{DD2} を生成し、コアの電源遮断時に電源遮断とした。

(ii) 上層配線の電流許容値を用いた場合の平均リーク電力の評価結果

上述の議論では NV-SRAM の T_{BE} とレイテンシについて、下層配線の電流許容値で設計したときの値を用いた場合の結果を示したが、ここでは VNR-SRAM と同様に上層配線の電流許容値を用いた場合の結果について述べる。上層配線の許容値は下層配線よりも緩和されるため、ラッシュカレントを抑えることができ、 $T_{\text{BE}}^{\text{cache}}$ は 2MB アレイで $\sim 120\mu\text{s}$ 、レイテンシは $\sim 390\mu\text{s}$ となる。図 5.5.4 にこの場合のキャッシュとシステムの平均リーク電力の評価結果を示す。この場合も図 5.2.4, 図 5.2.7 と同じ傾向を示し、同様の結論となる。

(iii) 表 5.3.2 の値の導出について

表 5.3.1 に示したシステムの面積, $\gamma_{\text{PG}}^{\text{system}}$, $T_{\text{BE}}^{\text{EE}}$ の計算に必要なパラメータを表 5.5.1 から表 5.5.3 にまとめた。いずれの表も左の列はそれぞれのセルに共通のパラメータであり、右の列は計算に使用するパラメータである。

5.6 図と表

表 5.1.1 シミュレーション条件と各種 SRAM, FF に関するデータ

<u>HSPICE simulation</u>				<u>Energy overhead (per cell)</u>				<u>Peripheral leakage (per gate)</u>		
65nm CMOS device (LP)				NV-SRAM (Store)	388.9	fJ		Standby	Shutdown	
Room temperature				(Restore)	5.611	fJ	Inverter	13.58	nW	43.88
				VNR-SRAM (ENT)	1.645	fJ	NAND (2input)	23.41	nW	194.8
				(Header, pFB) (EXT)	1.902	fJ	NOR (2input)	23.56	nW	51.62
				VNR-SRAM (ENT)	1.577	fJ	NAND (3input)	35.44	nW	478.4
				(Dual, pFB) (EXT)	1.660	fJ	NOR (3input)	35.89	nW	75.52
				VNR-FF(pFB)	20.3	fJ	sense amplifier	89.80	nW	100.1
				NAND	0.186	fJ	FF (Small Drv.)	141.7	nW	627.1
							FF (Large Drv.)	438.0	nW	2.911
<u>Cell leakage power (SRAM)</u>				<u>ENT/EXT latency (per subarray)</u>				<u>Area occupancy in a logic system</u>		
6T (SB)	4.429	nW		NV	τ_{ENT1}	20.68	μ s	Last-level cache	40%-70% (in a logic system)	
6T (SLP)	2.628	nW		NV	τ_{EXT1}	198.4	ns	First-level cache	10% (in a core)	
6T (SD)	10.35	pW		pFB (Header)	τ_{ENT1}	21.46	ns	Register	20% (in a core)	
NV (SB)	4.436	nW		pFB (Header)	τ_{EXT1}	94.03	ns			
NV (SLP)	2.633	nW		pFB (Dual)	τ_{ENT1}	30.0	ns			
NV (SD)	11.70	pW		pFB (Dual)	τ_{EXT1}	60.0	ns			
pFB(Header) (SB)	2.191	nW								
pFB(Header) (ULVR)	249.1	pW								
pFB(Header) (SD)	10.42	pW								
pFB(Dual) (SB)	2.77	nW								
pFB(Dual) (ULVR)	45.2	pW								
				<u>Active power (per cell)</u>				<u>Cell area</u>		
				6T	Read	44.73	μ W	6T	1.306	μ m ²
					Write	74.32	μ W	NV	1.999	μ m ²
				NV	Read	44.68	μ W	pFB	2.305	μ m ²
					Write	75.85	μ W	Logic gate (NAND)	1.404	μ m ²
				pFB (Header)	Read	39.99	μ W	DP retention FF	8.973	μ m ²
					Write	75.53	μ W	VNR FF	10.73	μ m ²
				pFB (Dual)	Read	40.25	μ W	NV FF	14.15	μ m ²
					Write	75.44	μ W	Balloon FF	17.86	μ m ²
				<u>Cell array</u>						
				arraysize		32	kB-2 MB			
				Read/Write speed		1	GHz			
				Read/Write bit width		64	bit			
				Store bit width		64	bit			
<u>Cell leakage power (FF, NAND)</u>										
DP retention FF (SB)	56.40	nW								
DP retention FF (SLP)	25.79	nW								
VNR FF (pFB) (SB)	56.19	nW								
VNR FF (pFB) (ULVR)	1.28	nW								
NV FF (SB)	65.81	nW								
NV FF (SD)	431.8	pW								
Balloon FF (SB)	58.14	nW								
Balloon FF (BFFR)	2.390	nW								
NAND (SB)	7.099	nW								
NAND (SD)	3.145	pW								

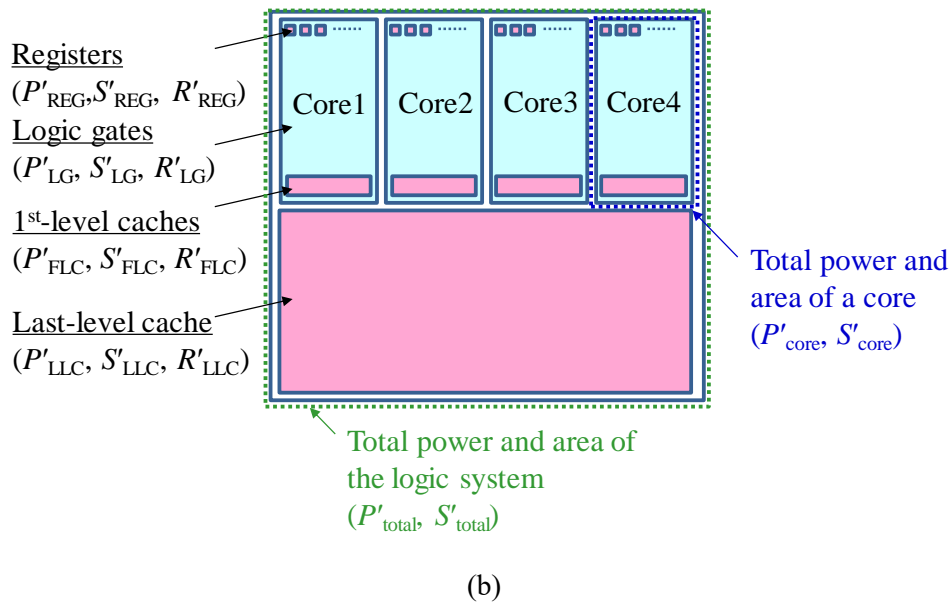
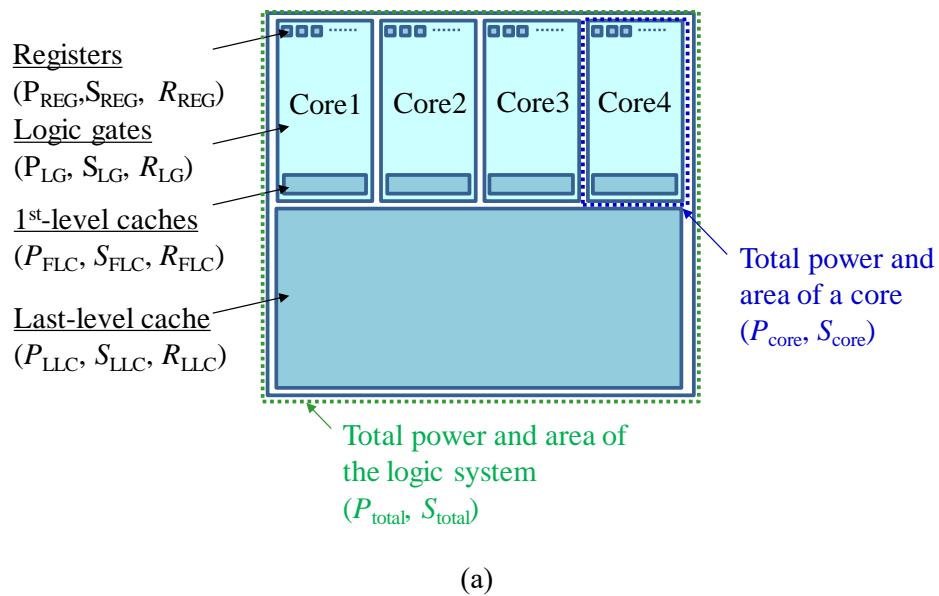


図 5.1.1 (a)従来の CMOS ロジックシステムと(b)不揮発/擬似不揮発記憶を用いた CMOS ロジックシステムの構成と各回路の電力, 面積, 占有率

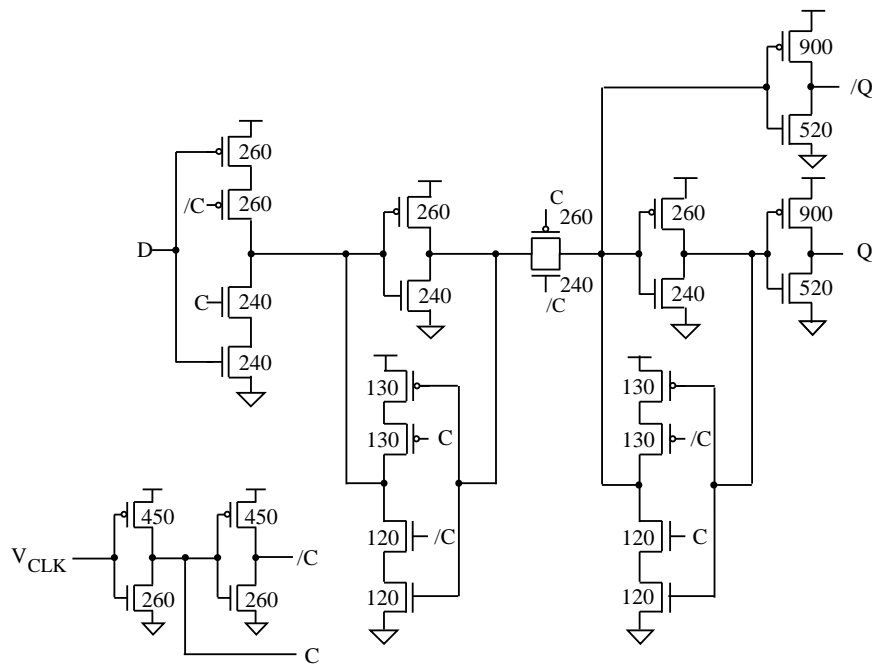


図 5.1.2 DP リテンション FF の回路構成

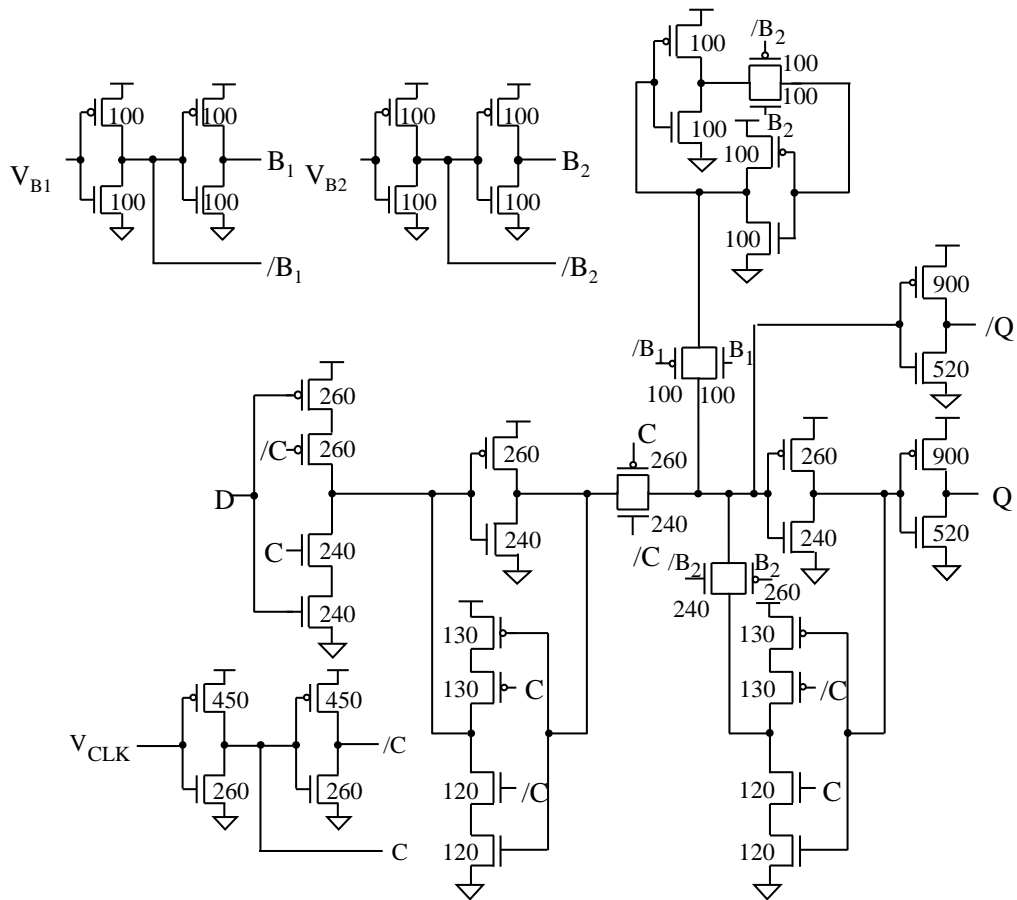


図 5.1.3 バルーン FF の回路構成

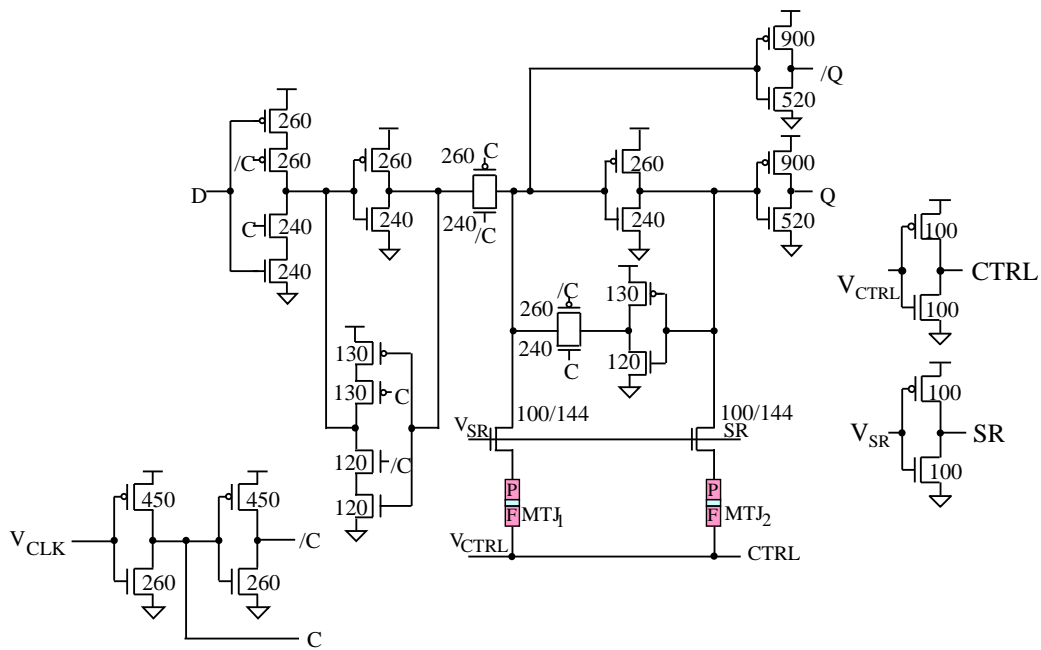


図 5.1.4 不揮発性 FF の回路構成

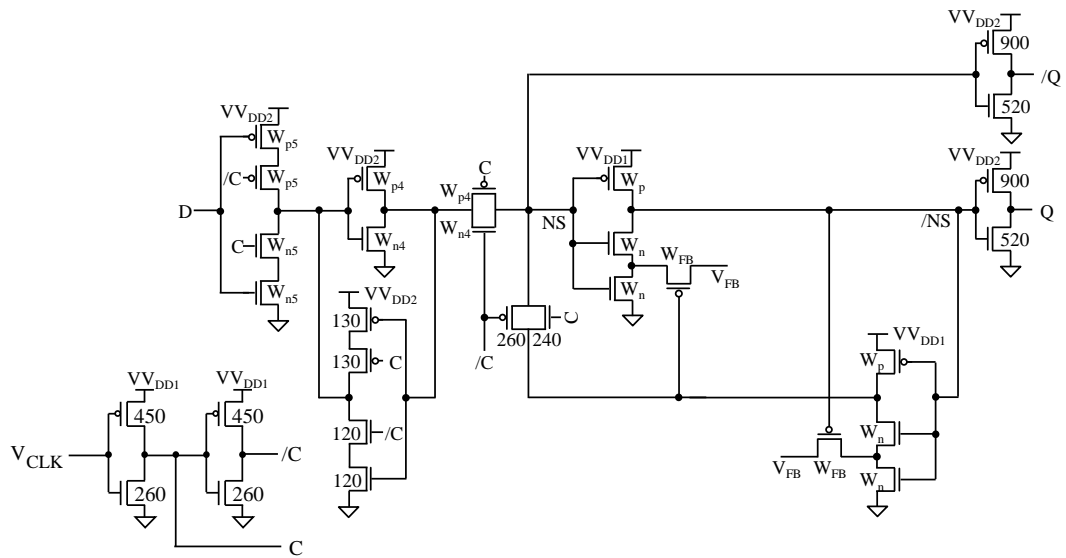
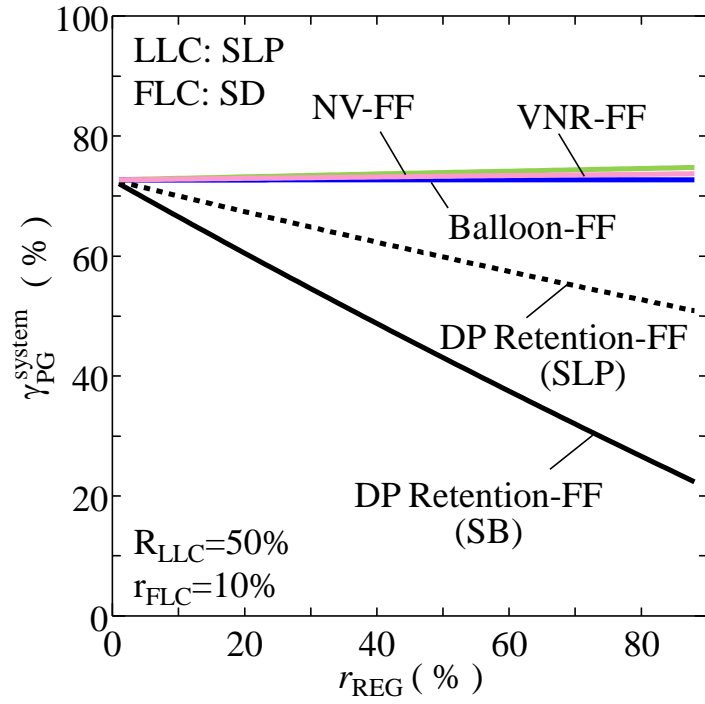
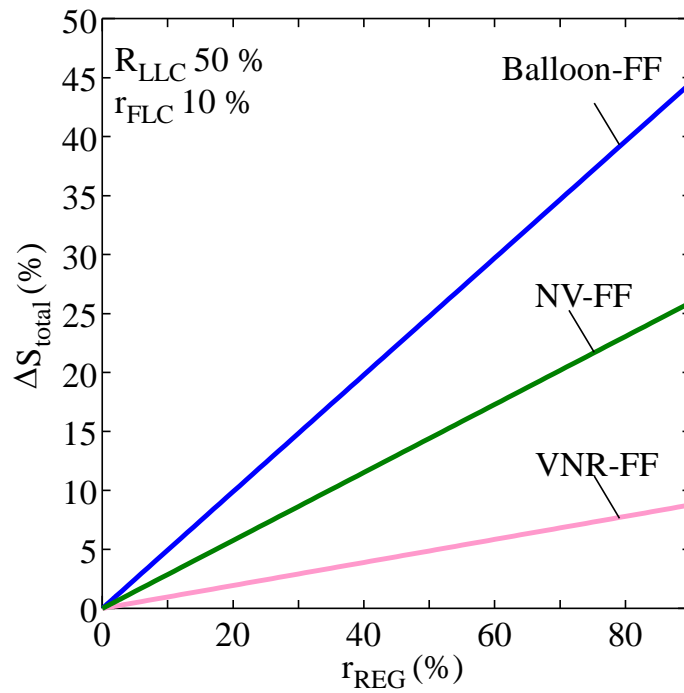


図 5.1.5 VNR-FF の回路構成 (最適設計では $W_p=130$ nm, $W_n=385$ nm, $W_{FB}=145$ nm, $W_{p4}=390$ nm, $W_{n4}=360$ nm, $W_{p5}=260$ nm, $W_{n5}=240$ nm)

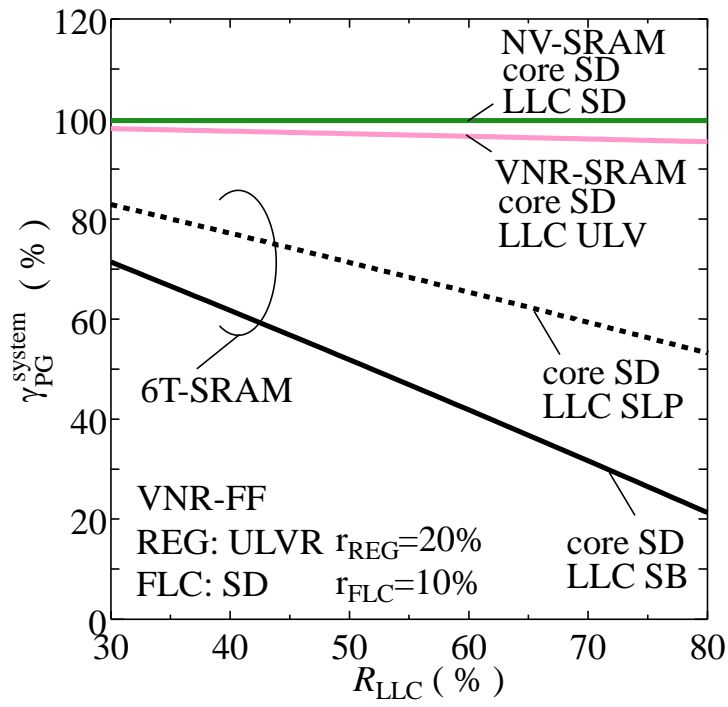


(a)

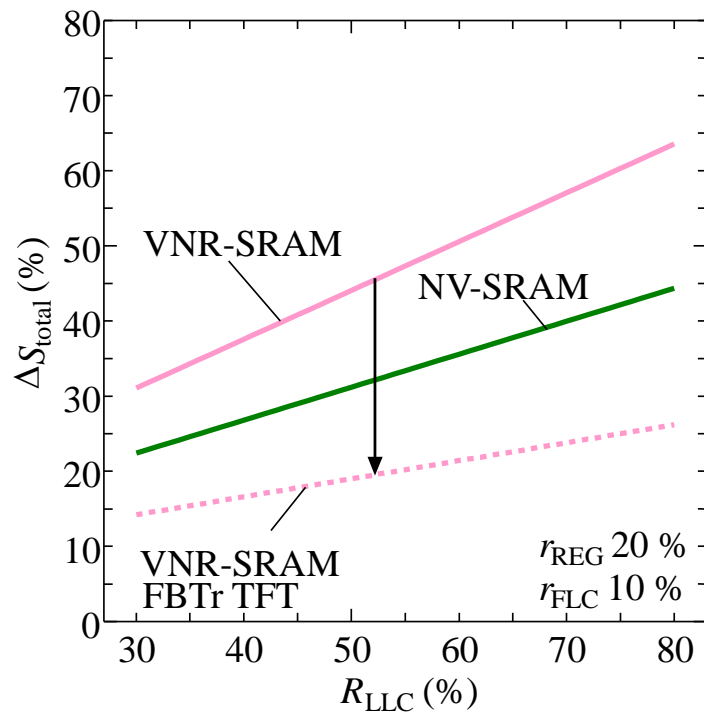


(b)

図 5.1.6 各種 FF で構成したレジスタを用いた場合の(a) γ_{PG}^{system} と(b) ΔS_{total} の r_{REG} 依存性



(a)



(b)

図5.1.7 各種SRAMで構成したLLC, FLCを用いた場合の(a) γ_{PG}^{system} と(b) ΔS_{total} の R_{LLC} 依存性

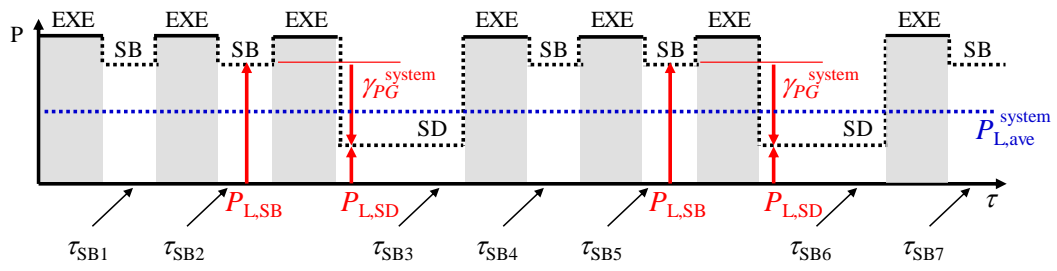


図 5.2.1 ロジックシステムの動作の概念図

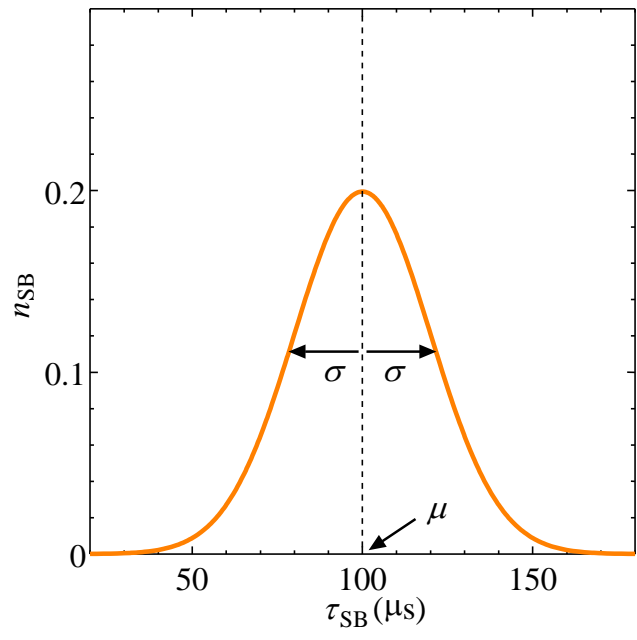


図 5.2.2 非実行時間の分布関数

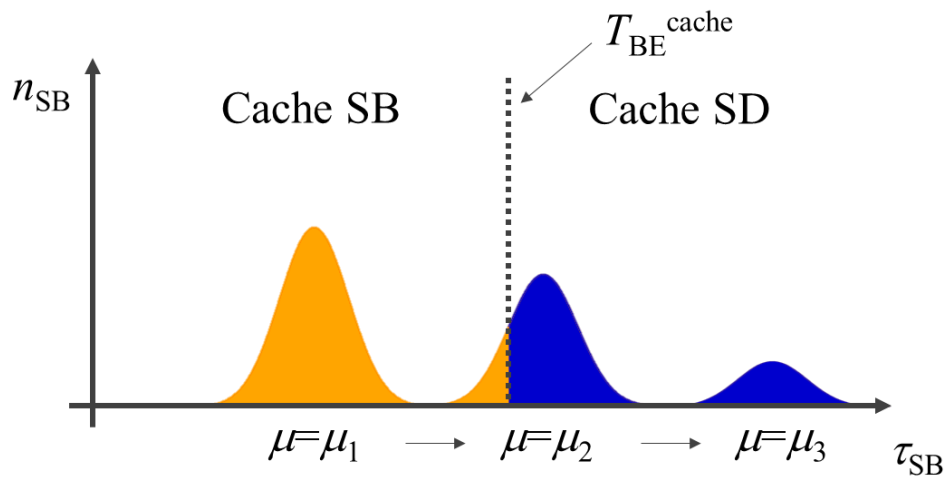


図 5.2.3 μ を変化させたときの非実行時間の正規分布とキャッシュの電源供給状態

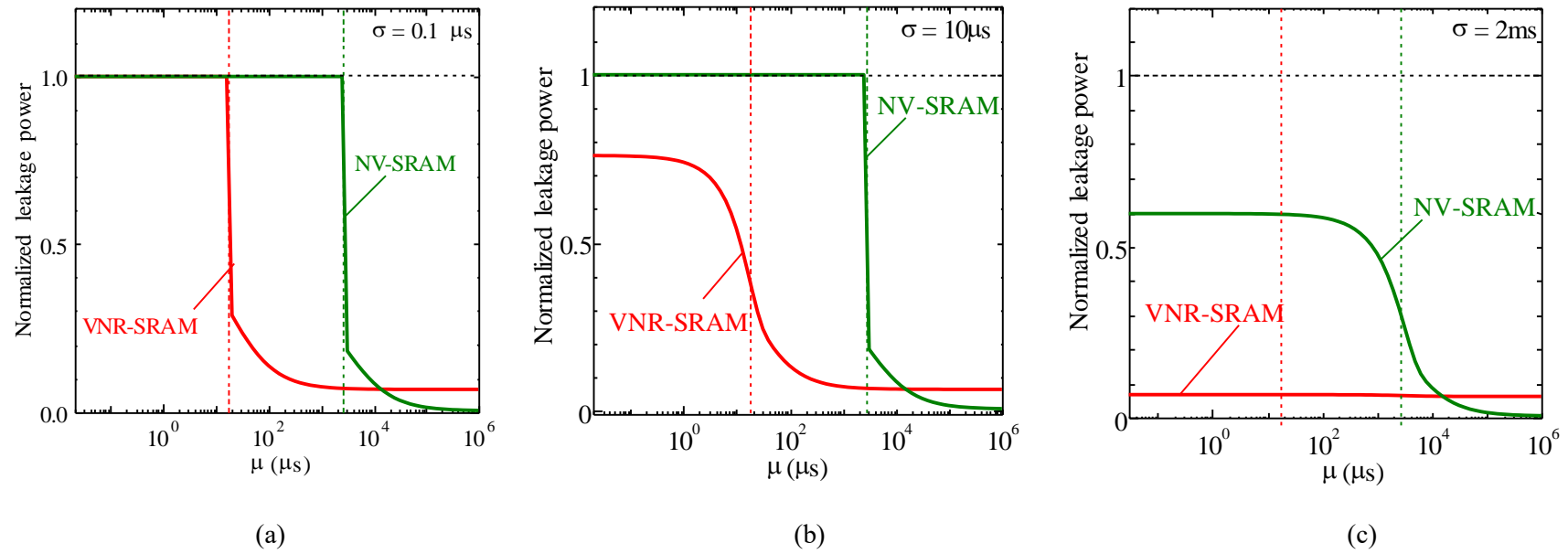


図 5.2.4 (a) $\sigma = 0.1 \mu\text{s}$, (b) $\sigma = 10 \mu\text{s}$, (c) $\sigma = 2 \text{ms}$ における $P_{L, \text{ave}}^{\text{cache}}$ の μ 依存性

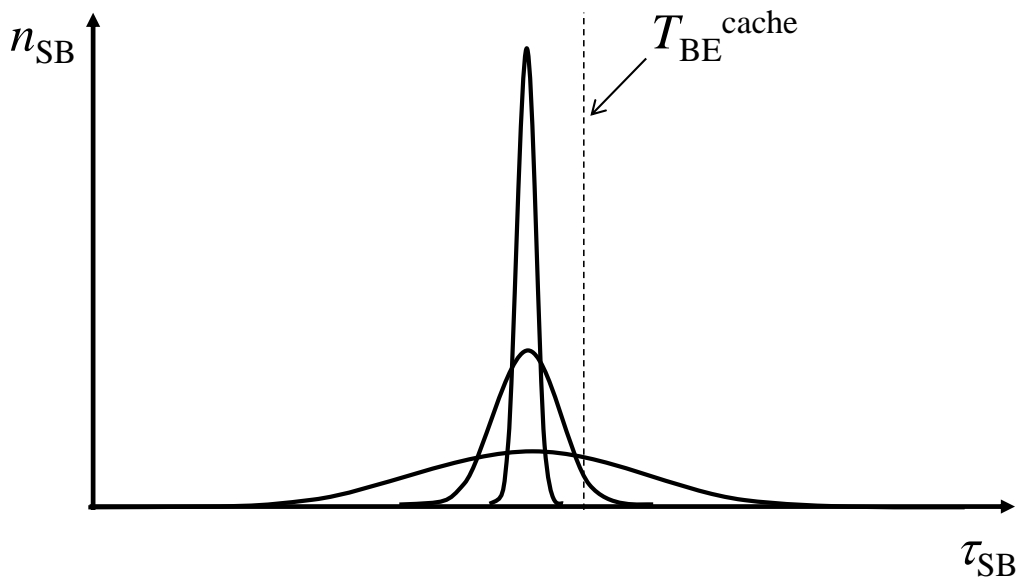


図 5.2.5 σ を変化させたときの正規分布と T_{BE}^{cache} の関係の模式図

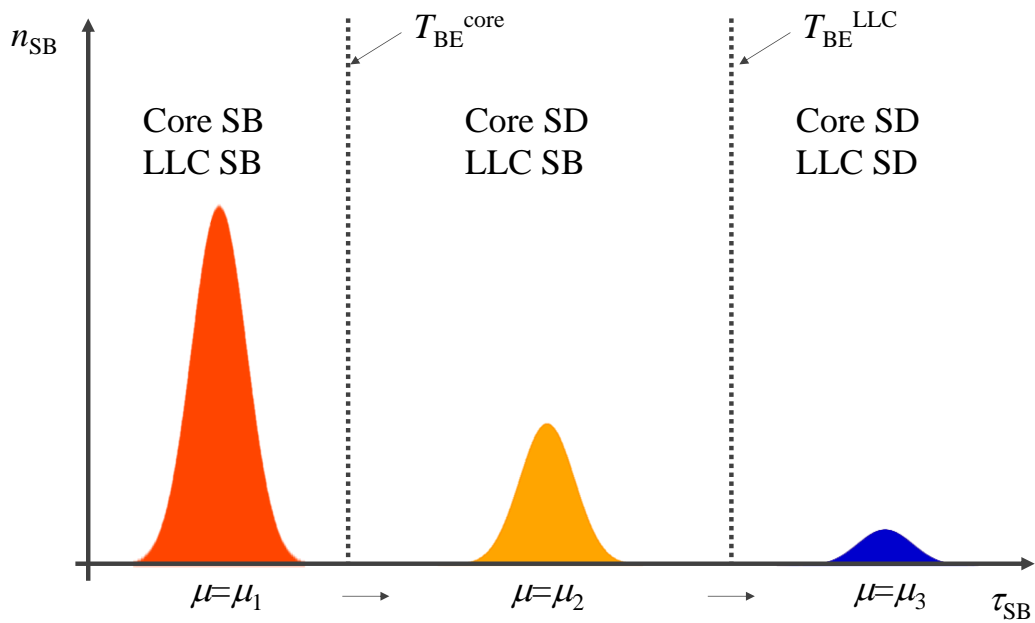


図 5.2.6 μ を変化させたときの非実行時間の正規分布とシステムの電源供給状態

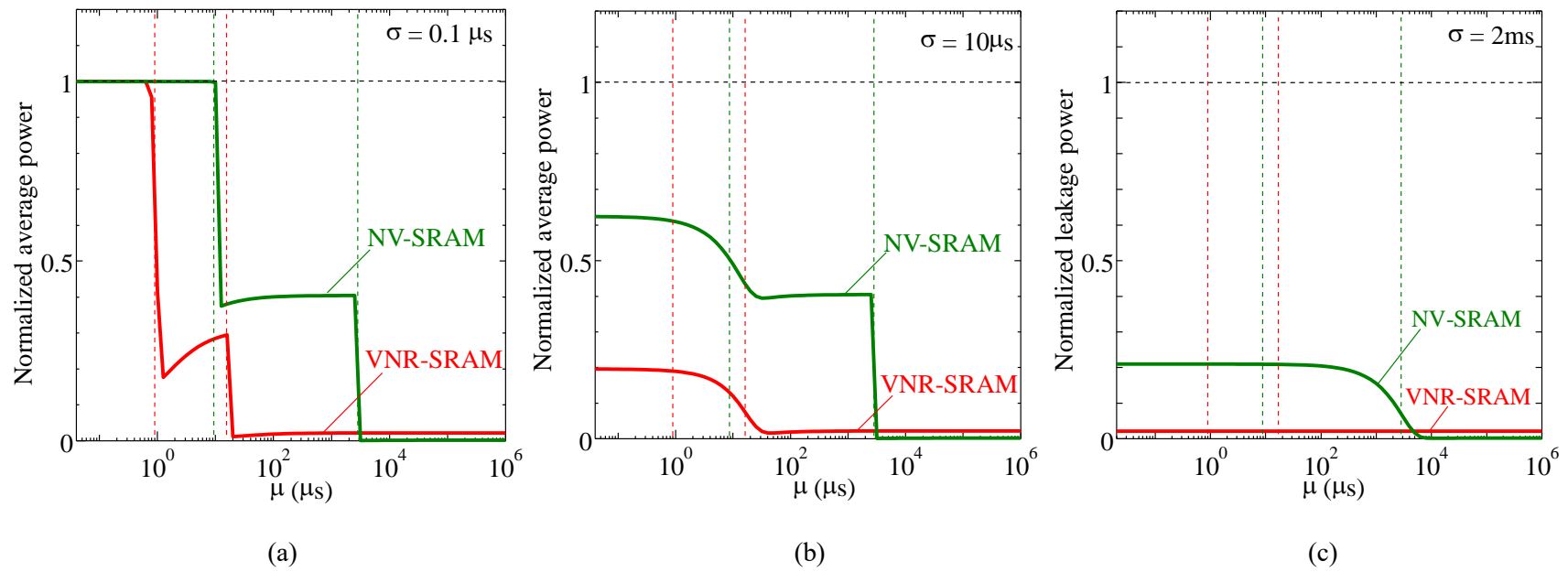


図 5.2.7 (a) $\sigma=0.1\mu\text{s}$, (b) $\sigma=10\mu\text{s}$, (c) $\sigma=2\text{ms}$ におけるシステムの $P_{L,\text{ave}}^{\text{cache}}$ の μ 依存性

表 5.3.1 各種 SRAM の性能比較

		6T-SRAM	NV-SRAM	VNR-SRAM (Header PS)	VNR-SRAM (Header PS, TFT)	VNR-SRAM (Dual PS)
プロセス		CMOSのみ	MTJエンベデッド 技術	CMOSのみ	CMOS+TFT	CMOSのみ
面積	セル (6Tで規格化)	1	1.53	1.77	1.31	1.77
	システム (従来システムで規格化)	1	1.3	1.4	1.2	1.4
コスト**		1	5	2-3	4	2-3
PSの構成		ヘッダ	ヘッダ	ヘッダ	ヘッダ	ヘッダ & フッタ
セルの書き込み電力(規格化)		1	1.016	1.016	-	1.015
電力削減率	セル(γ_{PG})	40.7%	99.7%	94.4%	-	99.0%
	システム(γ_{PG}^{system})	72.9%	99.6%	97.1%	-	99.2%
T_{BE}^{EE}		3ms ***	89.0 μ s-740 μ s	1.67 μ s-16.7 μ s	-	1.45 μ s-18.0 μ s

*CMOS 65nm LP device, $R_{LLC}=50\%$, $r_{FLC}=10\%$, $r_{REG}=20\%$

**chip cost: 1(Low cost) – 5(High cost)

***reported values

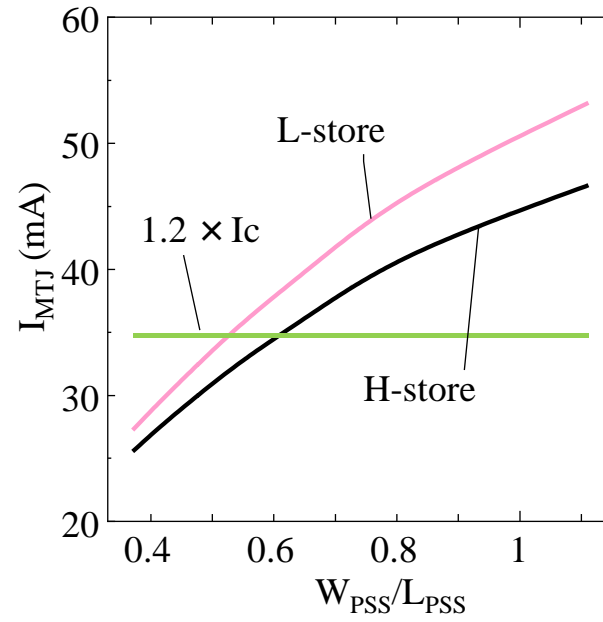


図 5.5.1 H ストア時, L ストア時における MTJ の書き込み電流の W_{PSS}/L_{PSS} 依存性

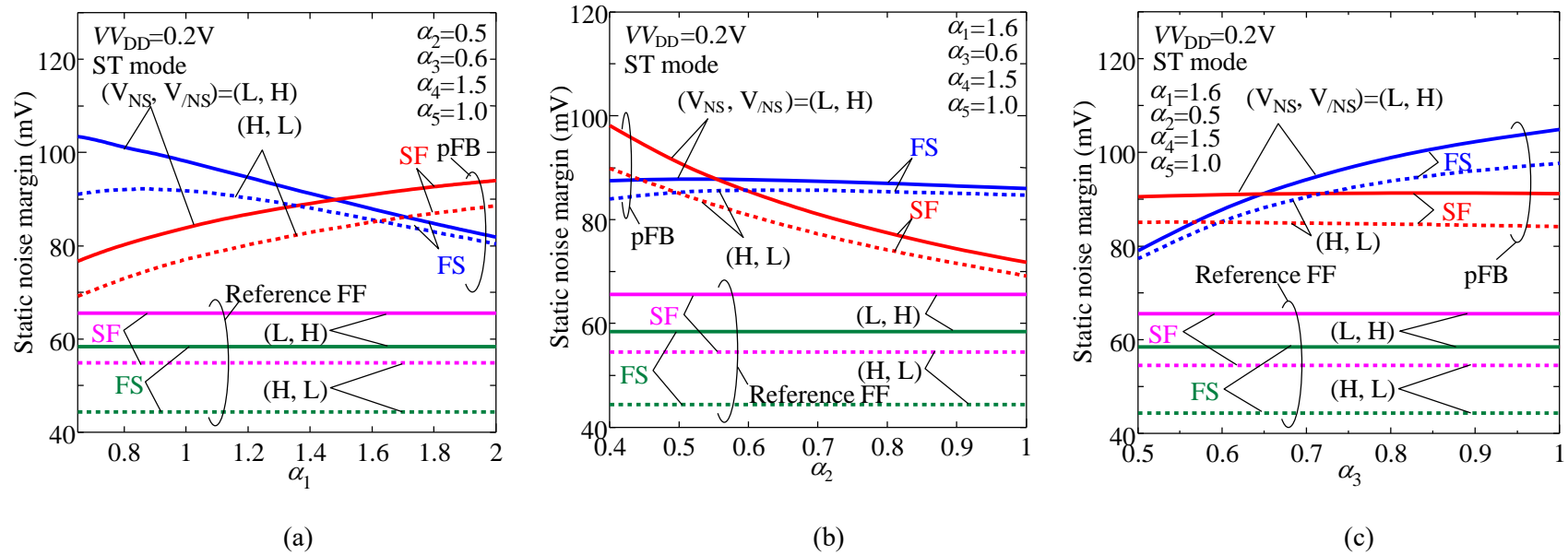


図 5.5.2 ULV リテンションの SNM の(a) α_1 , (b) α_2 , (c) α_3 依存性(それぞれスレーブラッチの記憶ノードのデータが $(V_{NS}, V_{NS})=(H,L)$, (L,H) の場合を示している)

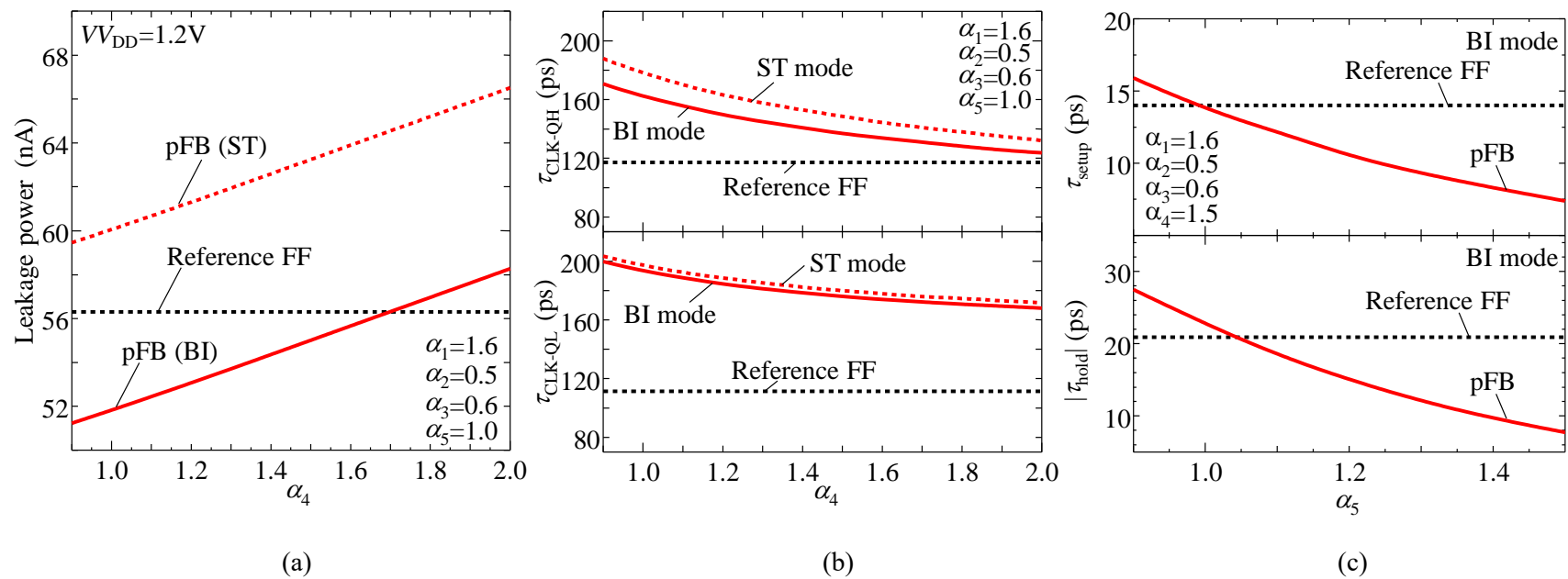


図 5.5.3 通常電圧での動作時における(a)リーク電力と(b)クロック-出力遅延の α_4 依存性と(c)セットアップタイム/ホールドタイムの α_5 依存性

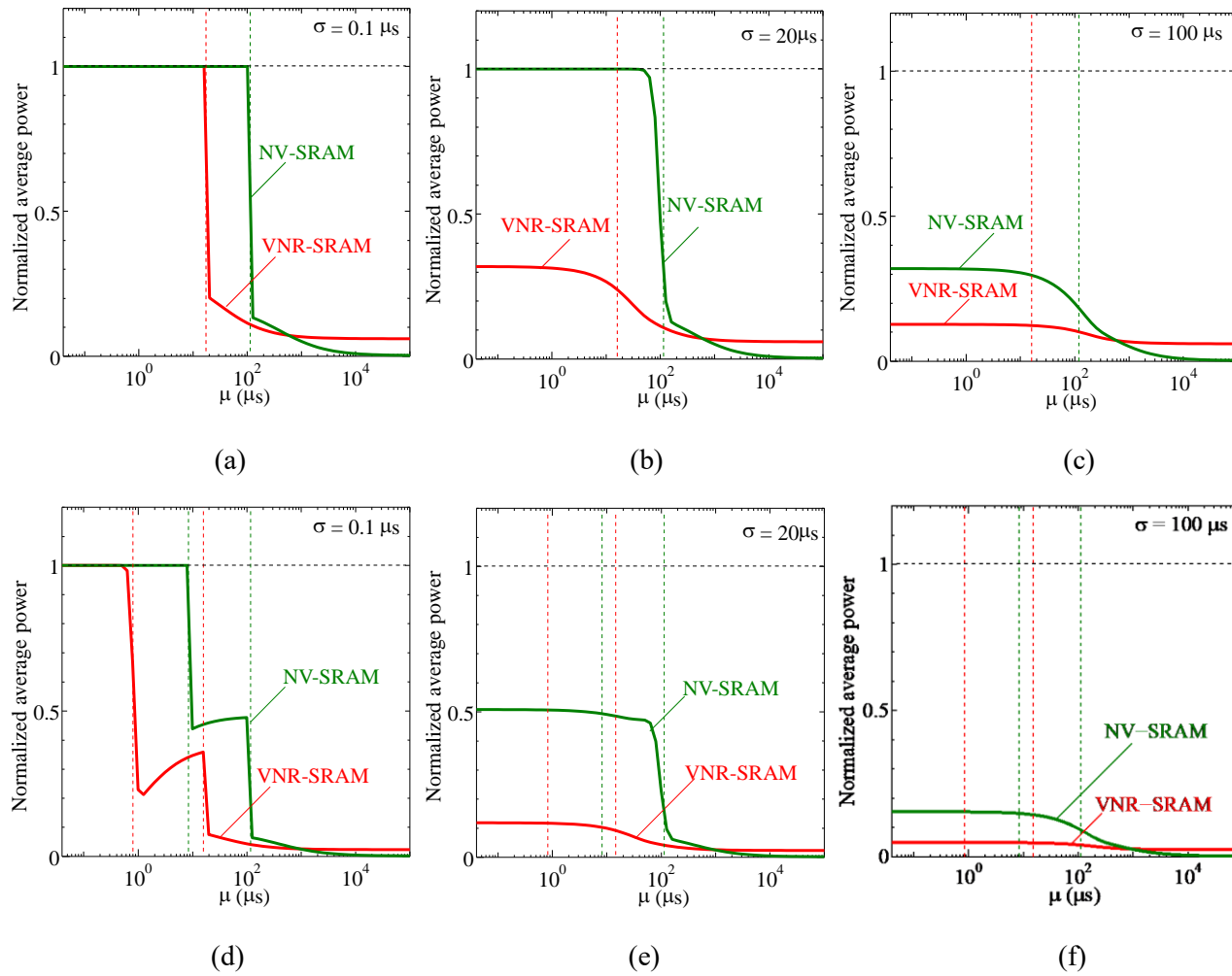


図 5.5.4 下層配線の電流許容値を用いた場合の(a)-(c)キャッシュと(d)-(f)システムの平均リーク電力の μ 依存性

表 5.5.1 システムの面積の計算に用いたパラメータ

$R_{\text{LLC}} = 50\%$	<u>NV-SRAM</u>
$r_{\text{FLC}} = 10\%$	$\gamma_{\text{LLC}} = \gamma_{\text{FLC}} = 1.53$
$r_{\text{REG}} = 20\%$	<u>pFB (Header, Dual)</u>
$r_{\text{LG}} = 70\%$	$\gamma_{\text{LLC}} = \gamma_{\text{FLC}} = 1.77$
$\gamma_{\text{REG}} = 1.2$ (VNR FF)	<u>pFB (TFT)</u>
$\gamma_{\text{LG}} = 1$ (NAND)	$\gamma_{\text{LLC}} = \gamma_{\text{FLC}} = 1.31$

表 5.5.2 $\gamma_{\text{PG}}^{\text{system}}$ の導出に用いたパラメータ

$R_{\text{LLC}} = 50\%$	<u>6T-SRAM</u>
$r_{\text{FLC}} = 10\%$	$P_{\text{LLC}}^{\text{C,SLP}} = 2.628 \text{ nW}$
$r_{\text{REG}} = 20\%$	$P_{\text{FLC}}^{\text{C,SD}} = 10.35 \text{ pW}$
$r_{\text{LG}} = 70\%$	<u>NV-SRAM</u>
$S_{\text{LLC}}^{\text{C}} = S_{\text{FLC}}^{\text{C}} = 1.306 \mu\text{m}^2$ (6T)	$P_{\text{LLC}}^{\text{C,SD}} = 11.70 \text{ pW}$
$S_{\text{REG}}^{\text{C}} = 8.973 \mu\text{m}^2$ (DP retention FF)	$P_{\text{FLC}}^{\text{C,SD}} = 11.70 \text{ pW}$
$S_{\text{LG}}^{\text{C}} = 1.404 \mu\text{m}^2$ (NAND)	<u>pFB (Header)</u>
$P_{\text{LLC}}^{\text{C,SB}} = 4.429 \text{ nW}$ (6T)	$P_{\text{LLC}}^{\text{C,ULVR}} = 249.1 \text{ pW}$
$P_{\text{FLC}}^{\text{C,SB}} = 4.429 \text{ nW}$ (6T)	$P_{\text{FLC}}^{\text{C,ULVR}} = 249.1 \text{ pW}$
$P_{\text{REG}}^{\text{C,SB}} = 56.19 \text{ nW}$ (VNR FF)	<u>pFB (Dual)</u>
$P_{\text{REG}}^{\text{C,ULVR}} = 1.28 \text{ nW}$ (VNR FF)	$P_{\text{LLC}}^{\text{C,ULVR}} = 38.80 \text{ pW}$
$P_{\text{LG}}^{\text{C,SB}} = 7.099 \text{ nW}$ (NAND)	$P_{\text{FLC}}^{\text{C,ULVR}} = 38.80 \text{ pW}$
$P_{\text{LG}}^{\text{C,SD}} = 3.145 \text{ pW}$ (NAND)	

表 5.5.3 T_{BE}^{EE} の導出に用いたパラメータ

NV-SRAM		VNR-SRAM (Header)		VNR-SRAM (Dual)	
(ENT/EXT mode)	(32kB)	(ENT/EXT mode)	(32kB)	(ENT/EXT mode)	(32kB)
$\Delta P_{ENT1}^{Array} = 1.232 \text{ mW}$	$N = 4$	$\Delta P_{ENT1}^{Array} = 5.027 \text{ mW}$	$N = 4$	$\Delta P_{ENT1}^{Array} = 6.248 \text{ mW}$	$N = 4$
$\Delta P_{ENT1}^{PGP} = 131.4 \text{ } \mu\text{W}$	$\Delta E_{EE} = 114.4 \text{ nJ}$	$\Delta P_{ENT1}^{PGP} = 0 \text{ W}$	$\Delta E_{EE} = 930.1 \text{ pJ}$	$\Delta P_{ENT1}^{PGP} = 0 \text{ W}$	$\Delta E_{EE} = 1.137 \text{ nJ}$
$\Delta P_{EXT1}^{Array} = 1.915 \text{ mW}$	$\Delta E_{leakage} = 20.77 \text{ nJ}$	$\Delta P_{EXT1}^{Array} = 1.326 \text{ mW}$	$\Delta E_{leakage} = 112.5 \text{ pJ}$	$\Delta P_{EXT1}^{Array} = 1.597 \text{ mW}$	$\Delta E_{leakage} = 113.2 \text{ pJ}$
$\Delta P_{EXT1}^{PGP} = 131.4 \text{ } \mu\text{W}$		$\Delta P_{EXT1}^{PGP} = 0 \text{ W}$		$\Delta P_{EXT1}^{PGP} = 0 \text{ W}$	
$\Delta P_{L,SLP1}^{Array} = 162.0 \text{ } \mu\text{W}$	(256kB)	$\Delta P_{L,SB1}^{Array} = 143.6 \text{ } \mu\text{W}$	(256kB)	$\Delta P_{L,SB1}^{Array} = 159.1 \text{ } \mu\text{W}$	(256kB)
$\Delta P_{L,SD/ULVR1}^{Array} = 952.1 \text{ nW}$	$N = 32$	$\Delta P_{L,SD/ULVR1}^{Array} = 17.04 \text{ } \mu\text{W}$	$N = 32$	$\Delta P_{L,SD/ULVR1}^{Array} = 2.648 \text{ } \mu\text{W}$	$N = 32$
$\Delta P_{L,OFF1}^{NLP} = 619.4 \text{ nW}$	$\Delta E_{EE} = 915.6 \text{ nJ}$	$\Delta P_{L,OFF1}^{NLP} = 619.4 \text{ nW}$	$\Delta E_{EE} = 7.441 \text{ nJ}$	$\Delta P_{L,OFF1}^{NLP} = 619.4 \text{ nW}$	$\Delta E_{EE} = 9.094 \text{ nJ}$
$\Delta P_{L,OFF/ON1}^{PGP} = 597.1 \text{ nW}$	$\Delta E_{leakage} = 1.713 \text{ } \mu\text{J}$	$\Delta P_{L,OFF/ON1}^{PGP} = 0 \text{ W}$	$\Delta E_{leakage} = 9.275 \text{ } \mu\text{J}$	$\Delta P_{L,OFF/ON1}^{PGP} = 0 \text{ W}$	$\Delta E_{leakage} = 9.336 \text{ nJ}$
(SD/ULVR mode)		(SD/ULVR mode)		(SD/ULVR mode)	
$\Delta P_{L,SD/ULVR1}^{Array} = 160.8 \text{ } \mu\text{W}$	(2MB)	$\Delta P_{L,SD/ULVR1}^{Array} = 144.7 \text{ } \mu\text{W}$	(2MB)	$\Delta P_{L,SD/ULVR1}^{Array} = 159.1 \text{ } \mu\text{W}$	(2MB)
$\Delta P_{L,OFF/ON1}^{PGP} = -597.1 \text{ nW}$	$N = 256$	$\Delta P_{L,OFF/ON1}^{PGP} = 0 \text{ W}$	$N = 256$	$\Delta P_{L,OFF/ON1}^{PGP} = 0 \text{ W}$	$N = 256$
$\tau_{ENT1} = 20.68 \text{ } \mu\text{s}$	$\Delta E_{EE} = 7.325 \text{ } \mu\text{J}$	$\tau_{ENT1} = 21.46 \text{ ns}$	$\Delta E_{EE} = 59.53 \text{ nJ}$	$\tau_{ENT1} = 46.55 \text{ ns}$	$\Delta E_{EE} = 72.75 \text{ nJ}$
$\tau_{EXT1} = 198.4 \text{ ns}$	$\Delta E_{leakage} = 112.7 \text{ } \mu\text{J}$	$\tau_{EXT1} = 94.03 \text{ ns}$	$\Delta E_{leakage} = 610.2 \text{ nJ}$	$\tau_{EXT1} = 92.00 \text{ ns}$	$\Delta E_{leakage} = 614.2 \text{ nJ}$

5.7 参考文献

- [1] V. Zyuban and S. V. Kosonocky, “Low power integrated scan-retention mechanism”, IEEE International Symposium on Low Power Electronics and Design (ISLPED) 2002, 98, (2002).
- [2] C. Auth *et al.*, “A 10nm high performance and low-power CMOS technology featuring 3rd generation FinFET transistors, Self-Aligned Quad Patterning, contact over active gate and cobalt local interconnects”, IEEE International Electron Devices Meeting (IEDM) 2017, 29 (2017).
- [3] M. Yuffe *et al.*, “A Fully Integrated Multi-CPU, GPU and Memory Controller 32nm Processor”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 264 (2011).
- [4] N. A. Kurd *et al.*, “Westmere: A Family of 32nm IA Processors”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 96 (2010).
- [5] C. Park *et al.*, “A 1.2 TB/s On-Chip Ring Interconnect for 45nm 8-Core Enterprise Xeon® Processor”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 180 (2010).
- [6] R. Kumar *et al.*, “A Family of 45nm IA Processors”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 58 (2009).
- [7] V. George, S. Jahagirdar, C. Tong, K. Smits, S. Damaraju, S. Siers, V. Naydenov, T. Khondker, S. Sarkar, and P. Singh, “Penryn: 45-nm next generation Intel® core™ 2 processor”, IEEE Asian Solid-State Circuits Conference (ASSCC) 2007, 14 (2007).
- [8] N. Sakran *et al.*, “The Implementation of the 65nm Dual-Core 64b Merom Processor”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 106 (2007).
- [9] S. Rusu *et al.*, “A Dual-Core Multi-Threaded Xeon® Processor with 16MB L3 Cache”, IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, 118 (2006).
- [10] S. Shigematsu, S. I. Mutoh, Y. Matsuya, Y. Tanabe, and J. Yamada, “A 1-V high-speed MTCMOS circuit scheme for power-down application circuits”, IEEE J. Solid-State Circuits **32**, 861 (1997).
- [11] S. Yamamoto, Y. Shuto, and S. Sugahara, “Nonvolatile delay flip-flop using spin-transistor architecture with spin transfer torque MTJs for power-gating systems”, Electron. Lett. **47**, 1027 (2011).
- [12] D. Kitagata, S. Yamamoto, and S. Sugahara, “Design and Performance of Virtually

Nonvolatile Retention Flip-Flop Using Dual-Mode Inverters”, IEEE New Generation of CAS (NGCAS) 2018, 186 (2018).

- [13] K. Ishibashi, T. Yamanaka, and K. Shimohigashi, “An alpha-immune, 2-V supply voltage SRAM using a polysilicon PMOS load cell”, IEEE J. Solid-State-Circuits **25**, 55 (1990).

第 6 章

結論

本論文では不揮発/擬似不揮発記憶を用いた SRAM 技術を開発し、この技術を用いて高効率に待機時エネルギーを削減できるパワーゲーティング (PG) 技術を開発した。そして、マイクロプロセッサ (μP) などの CMOS ロジックシステムにおける不揮発/擬似不揮発記憶を用いた PG の有用性・有効性を示した。

不揮発/擬似不揮発性 SRAM (NV-/VNR-SRAM) の性能を評価するためのシステムを構築し、評価指標である損益分岐時間 (Break-even time; BET) と待機時電力の削減率の定式化を行った。また、このシステムを用いて NV-/VNR-SRAM アレイの性能解析モデルを構築し、ばらつきを考慮した NV-/VNR-SRAM セルの設計指標を示した。

本研究室で提案された NV-SRAM についてばらつきを考慮した詳細なセルの設計方法を確立し、最適設計されたセル・アレイと周辺回路から構成される NV-SRAM の PG 性能を解析し、NV-SRAM による PG の効果を明らかにした。また、ストアエネルギー・レイテンシを効果的に削減する新技術として、階層型ストアフリー電源遮断アーキテクチャ (HSF アーキテクチャ) と積極的無用データフラッシュアーキテクチャ (PUDF アーキテクチャ) を提案し、その削減効果を示し、その有用性を明らかにした。さらに、試作した NV-SRAM の Test element group (TEG) から抽出した回路パラメータを用いて NV-SRAM の電力・エネルギー性能を検証し、本論文で述べた結果の妥当性と具現性を示した。

超低電圧 (ULV) によるデータ保持 (ULV リテンション) を用いた新たな PG を提案し、これを実現することができる VNR-SRAM セルを提案した。ULV リテンションにおける SNM を最大化する VNR-SRAM のセルの設計方法を確立し、VNR-SRAM を用いれば ULV リテンションを用いて大幅に待機時電力を削減できる PG を実現できることを示した。リーク電力削減率のさらなる高効率化を実現するデュアルパワースイッチアーキテクチャと PUDF アーキテクチャの VNR-SRAM への適用効果を評価し、その有用性を示した。最適設計した pFB セルで構成した 8kB の VNR-SRAM マクロを設計し、大規模シミュレ

ーションによるポストレイアウト解析から、本章で示した VNR-SRAM の性能の妥当性を明らかにした。さらに、FinFET などの先端 CMOS を用いても VNR-SRAM は十分に機能し、この技術は CMOS ロジック分野の将来にわたって活用できることを示した。

不揮発/擬似不揮発記憶を用いた PG を評価するための CMOS ロジックのシステムモデルを構築し、その電力削減率の解析から、この新しい PG によって μP のリーク電力を従来技術と比べて大幅に削減できることを示した。さらに、システムの待機状態分布のモデリングを行い、リーク電力削減に対する BET の効果を明らかにして、短い BET は単に時間的に細粒度の PG を実施できるだけでなく、待機状態の分布によっては平均のリーク電力を大幅に削減できること明らかにした。最後に、NV-/VNR-SRAM の総合比較から VNR-SRAM が今後のプロセッサ応用に適した SRAM 技術であることを明らかにした。

以上の結果から、本論文では不揮発/擬似不揮発記憶を用いた PG の有用性・有効性をはじめ明らかにし、これを用いることで、CMOS ロジックの低消費電力化が実現できることを示した。

今後、不揮発/擬似不揮発記憶技術を用いた PG を μP に応用していくためには、まず、様々なアプリケーションによって実際にどの程度待機時エネルギーの削減効果があるのかを詳細に検証していく必要がある。このためには NV-SRAM や VNR-SRAM を用いてキャッシュを構成し、実際の応用に適した回路・アーキテクチャを開発して実装する必要がある。特にキャッシュの PG に関する(ハードウェアや OS を含めた)制御アーキテクチャが重要になる。このためには、デバイス、回路、アーキテクチャ、OS の観点から、これらの連携によるさらにソフisstケートされた技術群の開発が必要になる。

本研究に関する発表

(a) 英文学術論文 (筆頭著者 2 件)

1. D. Kitagata, S. Yamamoto, and S. Sugahara, “Design and energy-efficient architectures for nonvolatile static random access memory using magnetic tunnel junctions”, Jpn. J. Appl. Phys., vol. 58, no. SB, 2019, pp. SBBB12/1-10.
2. D. Kitagata, S. Yamamoto, and S. Sugahara, “Proactive useless data flush architecture for nonvolatile SRAM using magnetic tunnel junctions”, IEICE Electronics Express, 17-20200032.

(b) 特許 (4 件)

1. 特許登録 発明者:菅原聡, 北形大樹, 山本修一郎 出願日:平成 29 年 11 月 30 日 出願番号:特願 2017-230447 公開番号:特開 2019-102111
2. 特許出願, 発明者:菅原聡, 北形大樹, 山本修一郎 出願日:令和 1 年 5 月 30 日 出願番号:特願 2019-101720
3. 特許出願, 発明者:菅原聡, 北形大樹, 山本修一郎 出願日:令和 1 年 10 月 9 日 出願番号:特願 2019-186042
4. 特許出願, 発明者:菅原聡, 北形大樹, 山本修一郎 出願日:令和 2 年 3 月 18 日 出願番号:PCT/JP2020/012099

(c) 招待講演 (筆頭著者 1 件)

1. 北形大樹, 周藤悠介, 山本修一郎, 菅原聡, “不揮発性 SRAM のアーキテクチャとエネルギー性能”, 電気情報通信学会集積回路研究会, 港区, 東京, 2017 年 4 月 21 日

(d) 国際会議 (筆頭著者 7 件)

1. D. Kitagata, S. Yamamoto, and S. Sugahara, “A New Store Energy and Latency Reduction Architecture for Nonvolatile SRAM Using STT-MTJs: Proactive Useless Data Flush Architecture”, IEEE International Electron Devices Meeting (IEDM) MRAM special session 2019, San Francisco, USA, December 7-11,

- 2019.
2. D. Kitagata, S. Yamamoto, and S. Sugahara, “Design and Performance of Virtually Nonvolatile Retention Flip-Flop Using Dual-Mode Inverters”, 2nd New Generation of Circuits & Systems Conference (NGCAS2018), Valletta, Malta, November 20, 2018, pp. 182-185.
 3. D. Kitagata, H. Yoshida, S. Yamamoto, and S. Sugahara, “Virtually Nonvolatile Retention SRAM cell Using Dual-Mode Inverters”, 2018 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S Conference 2018), San Francisco, USA, October 15-18, 2018, paper 13.5.
 4. D. Kitagata, S. Yamamoto, and S. Sugahara, “A New Architecture of Store Energy and Latency Reduction for Nonvolatile SRAM Based on Spintronics/CMOS-Hybrid Technology”, 2018 International Conference on Solid State Device and Materials (SSDM2018), Tokyo, Japan, September 9-13, 2018, paper B-4-03, pp. 119-120.
 5. D. Kitagata, S. Yamamoto, and S. Sugahara, “Virtually Nonvolatile Retention Flip-Flop Using FinFET Technology”, 2018 IEEE Silicon Nanoelectronics Workshop (SNW 2018), Honolulu, USA, June 17-18, 2018, paper P2-17.
 6. D. Kitagata, S. Yamamoto, and S. Sugahara, “Hierarchical Store-Free Architecture for Nonvolatile SRAM Using STT-MTJs”, IEEE International Electron Devices Meeting (IEDM) MRAM special session 2017, San Francisco, USA, December 2-6, 2017, paper P-23.
 7. D. Kitagata, Y. Shuto, S. Yamamoto, and S. Sugahara, “Analysis of Break-Even Time for Nonvolatile SRAM with SOTB Technology”, Electron Device Technology and Manufacturing Conference, Toyama, Japan, February 28-March 2, 2017, paper 4B-5.

(e) 国内会議 (筆頭著者 6 件, 共著 5 件)

1. 原拓実, 吉田隼, 北形大樹, 山本修一郎, 菅原聡, “各種リテンション技術を用いた SRAM のパワーゲーティング性能”, 第 80 回応用物理学会秋季学術講演会, 札幌市, 北海道, 2019 年 9 月 18 日-21 日, paper 18a-B11-10.
2. 瀧口憲一郎, 北形大樹, 松崎翼, 山本修一郎, 菅原聡, “不揮発/擬似不揮発性 FF を用いたパワーゲーティングの性能評価”, 第 80 回応用物理学会秋季学術講演会, 札幌市, 北海道, 2019 年 9 月 18 日-21 日, paper 18a-B11-9.

3. 吉田隼, 北形大樹, 山本修一郎, 菅原聡, “不揮発/擬似不揮発記憶を用いた SRAM のパワーゲーティング性能”, 第 80 回応用物理学会秋季学術講演会, 札幌市, 北海道, 2019 年 9 月 18 日-21 日, paper 18a-B11-8.
4. 北形大樹, 山本修一郎, 菅原聡, “NV-SRAM を用いた Useless data の積極的破棄による不揮発性パワーゲーティング”, 第 80 回応用物理学会秋季学術講演会, 札幌市, 北海道, 2019 年 9 月 18 日-21 日, paper 18a-B11-11.
5. 吉田隼, 北形大樹, 山本修一郎, 菅原聡, “デュアルパワースイッチを用いた擬似不揮発性 SRAM の設計と解析”, 第 66 回応用物理学会春季学術講演会, 目黒区, 東京, 2019 年 3 月 9 日-12 日, paper 10p-S221-6.
6. 北形大樹, 松崎翼, 山本修一郎, 菅原聡, “擬似不揮発性 FF の速度性能優先設計とその回路性能”第 66 回応用物理学会春季学術講演会, 目黒区, 東京, 2019 年 3 月 9 日-12 日, paper 10p-S221-5.
7. 吉田隼, 北形大樹, 山本修一郎, 菅原聡, “デュアルモードインバータを用いた擬似不揮発性 SRAM の設計と解析”, 第 79 回応用物理学会秋季学術講演会, 名古屋市, 愛知, 2018 年 9 月 18 日-21 日, paper 20a-CE-6.
8. 北形大樹, 山本修一郎, 菅原聡, “デュアルモードインバータを用いた擬似不揮発性 FF の設計と解析”, 第 79 回応用物理学会秋季学術講演会, 名古屋市, 愛知, 2018 年 9 月 18 日-21 日, paper 20a-CE-5.
9. 北形大樹, 山本修一郎, 菅原聡, “階層型ストアフリー電源遮断を用いた不揮発性 SRAM のエネルギー性能”, 第 65 回応用物理学会春季学術講演会, 新宿区, 東京, 2018 年 3 月 17 日-20 日, paper 18p-G203-3.
10. 北形大樹, 山本修一郎, 菅原聡, “強磁性トンネル接合を用いた不揮発性 SRAM の待機時電力削減能力”, 第 78 回応用物理学会秋季学術講演会, 福岡市, 福岡, 2017 年 9 月 5 日-7 日, paper 8p-C18-8.
11. 北形大樹, 周藤悠介, 山本修一郎, 菅原聡, “不揮発性 SRAM の設計とエネルギー性能の解析”, 第 64 回応用物理学会春季学術講演会, 横浜市, 神奈川, 2017 年 3 月 14 日-17 日, paper 16a-412-7.

謝辞

本論文の最後に、このような研究機会を与えて頂いた菅原聡准教授に深く感謝いたします。学部の研究室配属以来、大変すばらしい研究環境の中で充実した研究生活を送ることができました。日々の議論を通じて研究活動のみならず人格形成の面でも大変勉強になりました。常にどのような質問にも丁寧に答えて下さり、知識を深めることはもちろんのこと、学術研究の面白さを教えていただきました。また、何件もの国際会議に参加させていただき、海外での経験から様々なことを学ぶ機会を与えていただきました。研究活動以外の面でも雑学、ディベートなどを通じて暖かくご指導いただき、感謝しております。

本著者が修士2年に進学するまで隔週で行っていた宗片・菅原研究室の合同朝ゼミでの研究紹介や安全衛生リスクマネジメントでご指導いただき、半期ごとの合同ゼミで貴重なご意見をいただきました。宗片比呂夫教授に深く感謝いたします。

審査員としてご指導いただきました中川茂樹教授、宮本恭幸教授、岡田健一教授、杉野暢彦教授、高木信一教授、前島英雄名誉教授に深く感謝いたします。

博士進学時の博士エンカレッジ奨学金に関する業務や修士課程の論文審査員を引き受けていただき、大変お世話になりました若林整教授に深く感謝いたします。

修士論文および博士論文発表会において、ご指導を賜りました筒井一生教授、伊藤浩之准教授、飯野裕明准教授に深く感謝いたします。

周藤悠介博士(現ソニー株式会社)には修士1年では実験器具のトラブルを即座に解決していただき、修士2年では毎週勤務後に大学に立ち寄っていただき、研究内容について相談に乗っていただきました。研究環境、研究内容の両面で大変お世話になり、深く感謝いたします。

博士進学についての心得をご教授いただき、研究に関する相談にも乗っていただきました中川研究室、高村陽太博士に深く感謝いたします。

学部4年の研究室配属から3年間、スピン輸送デバイスの評価についてのプログラムの書き方やデバッグのコツなどをご教授いただき、また、実験装置の使用方法やデバイス作製方法、研究発表等のスライド作成についても多くのことを学ばせていただいた恵七泰樹氏(現 TDK 株式会社)、川目悠氏(現株式会社日立製作所)、生瀬裕之氏(現

日本電気株式会社), 同期として様々なことについて議論し切磋琢磨した近藤剛氏(現三菱電機株式会社), 岩崎大和氏(現パナソニック株式会社)に深く感謝いたします。

博士課程の3年間でも回路研究に邁進した菅原研究室回路班の吉川智英氏(現株式会社キーエンス), 塩津勇作氏(修士2年), 吉田隼氏(修士1年), 松崎翼氏(修士1年), 瀧口憲一朗氏(修士1年), 原拓実氏(学部4年)には, 議論や指導を行う中で新たな発見が得られることも多く, 作業の効率化や分担など多くのご協力をいただきました。本論文の内容の一部は回路班の皆様の協力によるものです。深く感謝いたします。

また, 学生生活をともに過ごした菅原研究室の菊池靖彦氏(現キャノン株式会社), 金子将士氏(現ソニーLSIデザイン株式会社), 清野稔仁氏(現株式会社NTTドコモ), 大久保岳氏(現株式会社東京精密), 千脇那菜氏(現ソニー株式会社), 山下涼音氏(現株式会社NTTドコモ), 熊谷颯人氏(修士1年), 佐藤晟安氏(修士1年), 遠藤弘之氏(学部4年)に深く感謝いたします。

研究生活の中で大変お世話になりました宗片研究室の西林一彦博士(現東京大学), 西沢望博士, 松田喬博士(現三菱電機株式会社), Ronel Roca 博士(現豊田工業大学博士研究員), 青山政樹氏(現三菱UFJモルガン・スタンレー証券), 高橋克典氏(現株式会社東洋テクニカ), 堀江慧氏, 飯塚謙太氏(現株式会社フジクラ), 新保理志氏(現富士通株式会社), 道廣健斗氏(現株式会社ベイカレントコンサルティング), 小川竣太氏(現Lam research), 小野寺蓮太氏, 高橋一真氏(修士2年), 森田有美子氏(修士2年), 谷井智樹氏(現東京大学大学院修士2年), 浜田享志(現キャノン株式会社), 伊藤涼平氏(修士1年), 辰巳堯氏(修士1年), 河島眞弥氏(学部4年), 舩橋昇太郎氏(学部4年)に深く感謝いたします。

事務手続きや施設管理等で長年わたり研究室をバックアップしていただいた石井和雄氏, 徳岡留美氏に深く感謝いたします。

学務手続きでお世話になりました物理電子システム創造専攻事務の岩松瑞恵氏, 電気電子系事務の清水恵子氏に深く感謝いたします。

最後に, 学業に専念できるよう支えてくれた家族に心より感謝いたします。