

# Enhanced 4T Loadless SRAM Comparison With Selected Volatile Memory Cells

Karol Niewiadomski and Dietmar Tutsch  
 University of Wuppertal  
 Chair of Automation and Computer Science  
 Wuppertal, Germany  
 Email: {niewiadomski, tutsch}@uni-wuppertal.de

**Abstract**—The adaptiveness of Field Programmable Gate Arrays is a key aspect in many mobile applications. Modern vehicles contain up to 100 "Electronic Control Units" in order to implement all necessary functions for autonomous driving. Due to the limited power resources of mobile applications, an appropriate implementation of power reduction measures is crucial for achieving an acceptable amount of power savings. However, effective power reduction mechanisms have to be applied to the backbone of each Field Programmable Gate Array: the look-up table. In this paper, we describe the implementation and comparison of various Static Random Access Memory cells and the related characteristics which are used as a benchmark. All Static Random Access Memory cells have been analyzed in order to evaluate feasible modifications for the sake of lowering leakage currents and modified for minimizing static and dynamic power consumption. The trade-off between low-power use cases, a fast response time and area considerations as well as yield after manufacturing has to be carefully analyzed. Therefore, further aspects like signal to noise ratio and area increase due to additional, required transistors are deliberated about whether the additional efforts delivers desired results. Since speed in terms of a high operating frequency is demanded by many applications, we analyze each design upon its capabilities to run at their maximum speed.

**Keywords**—Field Programmable Gate Array; Static Random Access Memory cell optimization; low-power; signal to noise ratio; max. operating frequency; signal propagation delay.

## I. INTRODUCTION

During the last years, the number of classic desktop computers used in domestic homes has constantly decreased. The reason behind this phenomenon is the rising number of mobile devices such as smartphones and tablets, taking over most of the functionalities provided by desktop computers before [1]. Furthermore, upcoming features like highly automated driving cars or fully autonomous vehicles require a high demand for computing power. Whilst the computing performance of mobile devices is improved constantly to face the challenges of complex applications like video processing for adaptive cruise control on long distance highway drives, the capacities of batteries providing the needed energy resources have not been extended in the same way. A modern, upper-class vehicle contains more than 70 electronic control units (ECUs) to provide all features desired by consumers these days [2]. On-board communication networks like Controller Area Network (CAN), FlexRay and ethernet ensure the communication between these devices, but also introduce a remarkable amount of additional weight of approximately up to 30% (depending on the used technology). In order to counter the limits set by power consumption and overall weight, a significant reduction of the ECU number would be an efficient approach. This could lead to the application of more powerful processors, taking

over many of the functionalities from the large number of slower ECUs used before. The downside of this approach would be a higher power consumption due to higher clock frequencies. A more comprehensive approach focuses on the massive usage of FPGAs in mobile applications. Field Programmable Gate Arrays (FPGAs) offer various advantages compared to processors and Application Specific Integrated Circuits (ASICs). Being fully configurable, FPGAs are well-suited for the execution of various functions which have been spread over several ECUs before, either purely by hardware implementations or software execution running on a softcore processor implemented on the FPGA's fabric. However, FPGAs do not offer similar power saving mechanisms implemented on microprocessors and lack of a substantial power management system. Power consumption saving mechanisms shall be applied to series production passenger cars, which is a cost-sensitive market, hence we choose the Xilinx Spartan-3 low-cost FPGA as a baseline architecture for all further considerations [3]. FPGAs play a major role for the realization of adaptive systems. Partial, dynamic reconfiguration [4], supported by various FPGA designs, offer a vast potential for fast adaption of the implemented functional range within a vehicle, e.g., realizing a requested function by the driver and disengaging a previously implemented vehicle function which is not required any more [5].

In this paper, we evaluate selected Static Random Access Memory (SRAM) cell designs on their suitability for a low-leakage look-up table (LUT) implementation, which are the elementar computational elements. Since the overhead of reconfigurability leads to unused parts within the FPGA, both static and dynamic power consumption are analyzed for each cell design. In Section II, we give an overview about a selection of existing designs and our motivation for improvements. In Section III, we describe a number of leakage reduction techniques and evaluate the feasible adaption on current designs. In Section IV, we investigate the SRAM cell designs on their assets and drawbacks and compare the simulation results. In Section V, circuit improvement methods for standby and active currents reduction are introduced. All investigated SRAM cells are enhanced with these additional improvements and compared again. In Section VI, we use each modified SRAM cell to implement a 4-input LUT reference design and explore the power consumption during the idle and active state. The advantages of reasonable SRAM cell design modifications are presented based upon the simulation results. In Section VII, further timing considerations are described. In Section VIII, all previous discussions are summarized and concluded.

## II. RELATED WORK

Various SRAM cell designs have been under research over the years. Compared to dynamic RAM (DRAM), which is widely used as main memory in many applications, SRAM offers numerous advantages like quick read & write-cycles, cell stability, data retention without refresh cycles, differential outputs and many more. During the pre-Complementary Metal-Oxide-Semiconductor (CMOS) era, the 4T cell [6] was commonly used for cache memories. Considering the additional effort in terms of process variations for implementing the resistor load and weaker signal to noise (SNM) margin, this cell type was replaced by the 6T cell [7]. This design depicts the mostly used approach for combining reliable functionality with a proven in use fabrication process due to its CMOS structure. Being the starting point for benchmarking, cell variations like the 5T SRAM [6] design were developed to eliminate the parasitic capacitance penalties of two bitlines. Further derivations like the 7T cell implementation [8] inherit the characteristics of the reference 6T design and provide power savings by exploiting an effective writing mechanism, putting no further requirements on adaptations to auxiliary circuitry. Features like soft error rate robustness during low-power operation have been explored in a 10T design variation [9]. All of these cell types have been designed during research without applying additional, commonly used power reduction measures. LUT designs have been evaluated and improved on architectural level [10] for power reduction by power gating mechanisms. New FPGA designs were presented and compared to commercial products, by adding structural improvements [11].

Our approach goes one step further and is based on circuit level improvements to a LUT by reasonable selection of a suitable SRAM cell design and substantial modification of the cell circuitry to achieve better leakage reduction and power savings. The improvements achieved on that level are essential for important leakage current suppression and are an inevitable step to be combined with architectural amendments.

## III. LEAKAGE REDUCTION

Three major components of leakage currents can be identified for a Metal-Oxide-Semiconductor (MOS) transistor of gate lengths in nanometer scales:

- Subthreshold leakage
- Direct tunneling gate leakages shown in
- Reverse biased p-n BTBT leakage

Whilst the band-to-band tunneling (BTBT) leakage currents can be neglected for devices exceeding 50nm gate lengths, subthreshold and direct tunneling gate leakage currents come into consideration for our design. Tunneling electrons through gates oxides can be countermeasured by carefully setting an adequate oxide thickness of each transistor. This dependency can be seen in (1):

$$J_{DT} \propto A \left( \frac{V_{ox}}{T_{ox}} \right)^2 \quad (1)$$

$$A = \mu_o C_{ox} \frac{W}{L_{eff}} \left( \frac{kT}{q} \right)^2 e^{1.8}$$

By increasing the oxide thickness  $T_{ox}$ , the direct tunneling current density  $J_{DT}$  can be efficiently lowered to a minimum stage [12]. Increasing the gate length  $L_{eff}$  would have a similar effect, but lead to higher effort in the manufacturing process due to a change in one of the basic technology parameters like the gate length of a transistor. Therefore, this option should be avoided. However, the usage of multi-oxide thicknesses is a technology dependent parameter and requires awareness for the selection of a suitable multi-oxide technology.

Subthreshold currents can be expressed by the following equation:

$$I_{sub} \propto \frac{W}{L_{eff}} e^{(V_{GS} - V_{t0} - \gamma V_{SB} + \eta V_{DS}) / n V_t} (1 - e^{-\frac{V_{DS}}{V_t}}) \quad (2)$$

Equation (2) shows the parameters which contribute to the overall weak-inversion current, flowing below the threshold voltage  $V_{th}$  of each MOS transistor in the circuit. Several leakage reduction measures can be applied by utilizing these parameters to design a low leakage circuit:

- $W$ : setting the width of a transistor as small as possible leads to a higher resistance of it and therefore to smaller leakage currents
- $V_{gs}$ : Gate biasing is done by applying a  $V_{gs}$  voltage lower than  $Gnd$ , which turns the transistor deeply off
- $V_{sb}$ : Body biasing by tweaking the body voltage of a turned off transistor
- $V_{dd}$ : Lowering the supply voltage mitigates or even completely removes the DIBL (drain-induced barrier lowering) effect, represented by  $\eta$  in (2)

In general, we can distinguish between two classes of leakage reduction techniques [13]. Some can be applied during the design, whereas others can be used during operation time of the circuit. A reasonable extract of these techniques is shown in Table I.

TABLE I. LEAKAGE REDUCTION TECHNIQUES

Design leakage reduction	Static leakage reduction	Active leakage reduction
Dual- $V_{th}$	Stacking	DVS
Multi- $V_{dd}$	Sleep mode	
	VTCMOS	DVTS

Energy efficient circuits should feature multiple supply voltages and at least a dual threshold approach. As shown in Table I, these characteristics need to be added during the development phase. Furthermore, additional techniques working during operation of the circuit can help to continuously reduce the overall power consumption. Dynamic (threshold) voltage scaling (DVS & DVTS), as well as variable threshold CMOS (VTCMOS) circuitry are powerful methods to overcome the side-effects like subthreshold leakage due to progressive scaling to smaller technology nodes.

We analyze the techniques listed in Table I on their careful combination and application to volatile (SRAM) memory cells and therefore automatically to LUTs.

#### IV. SRAM CELL DESIGNS

The backbone of each computational activity within an FPGA is the LUT [14]. Typically an FPGA consists of a sea of tiles which contain the necessary logic in terms of LUTs and interconnection circuitry, shown in Figure 1. Two different groups of logic can be identified: Configurable Logic Block (CLB) and switch matrix. A CLB is used for ensuring the feature of adaptiveness due to the built-in LUTs, therefore it contains the LUTs and additional components, e.g., flip-flops, multiplexers and basic logic gates. On the other hand, the switching matrix is used for providing all necessary interconnections to other tiles / LUTs in case that more complex functions are requested to be implemented and require a combination of multiple CLBs.

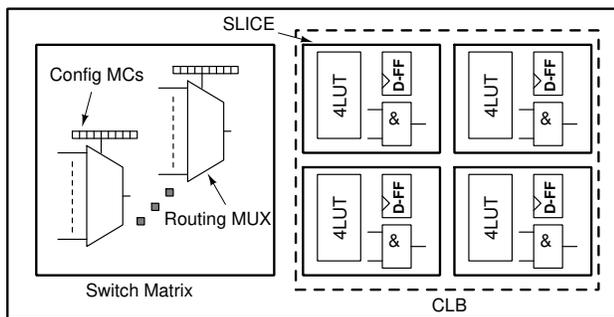


Figure 1. Simplified 'Tile' of an FPGA

By putting a higher focus on the optimization of the configuration RAM cells, these efforts serve not only improving the power balance of the CLBs, but also to decrease the switch matrices standby leakage currents. For communication with peripheral logic General Purpose Input Output (GPIOs) blocks are implemented, which can be used for bidirectional data. However, switch matrices and GPIOs are not subject matter of this paper and will be discussed in later publications.

Depending on the number of the LUT's inputs, a LUT can contain numerous SRAM cells. For example, in case of a 4-input LUT, 16 SRAM cells are necessary for the realization of all possible input value combinations. An exemplary illustration of a LUT is shown in Figure 2.

Since the memory cells are used for configuration, they are also called configuration RAM (CRAM). Once configured during the start-up phase, the content of these memory cells would not be changed until the next reconfiguration cycle. In consequence, the static leakage current reduction is of higher significance for the overall power consumption.

The selection of a low-power SRAM cell design is crucial for an appropriate energy-efficient implementation of integrated circuits. Many memory cell designs have been introduced in the past. The common 6 transistor cell can be found in most FPGAs nowadays [15]. In principle, this memory cell consists of two cross-coupled inverter and two access transistors, connecting the inverters to the bitlines, as shown in Figure 3.

As long as  $M5$  and  $M6$  are in cut-off mode, the cross-coupled inverters are isolated from the bitlines and store the

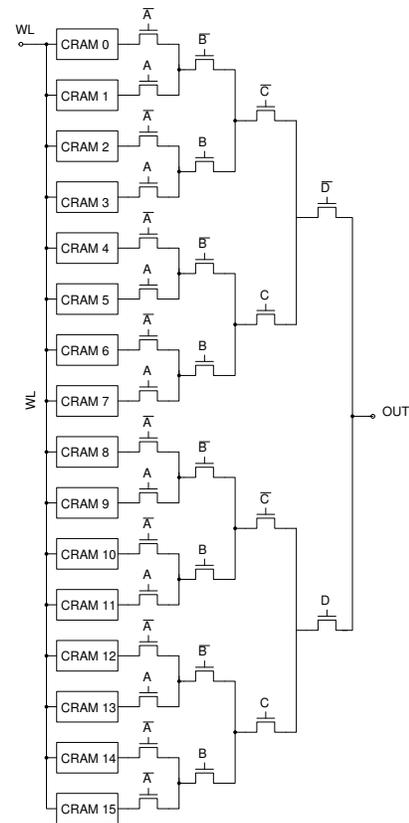


Figure 2. Simplified 4-input LUT

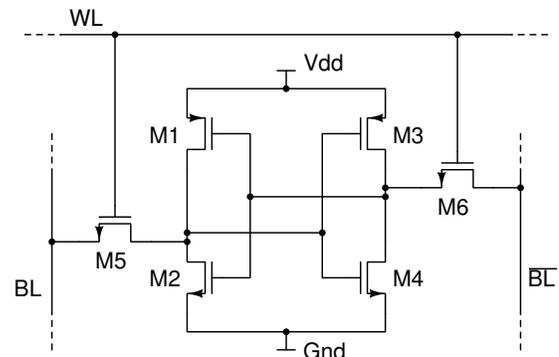


Figure 3. 6T SRAM cell

complementary data value at the output nodes of each inverter. Data retention is ensured as long as a sufficient supply voltage  $V_{dd}$  is applied. Before reading the stored data, both bitlines  $BL$  and  $\overline{BL}$  are precharged to  $V_{dd}$  by a special precharge circuit and the access transistors  $M5$  and  $M6$  are turned on. One of the bitlines will be discharged to  $Gnd$ , whereas the other bitline will remain on  $V_{dd}$ . The voltage drop between  $BL$  and  $\overline{BL}$  will be sensed and evaluated by a sense amplifier. For writing data into the cell, one of the bitlines is kept at  $V_{dd}$ , whereas the other bitline is kept at  $Gnd$ . By turning the access transistors on, the desired value is written. For this purpose, a suitable bitline driver circuit is needed to ensure the proper execution of the writing cycle. Careful transistor sizing is required for avoiding the cell to flip during, e.g., a

read cycle. This cell design is well-elaborated and used for years in integrated circuits. Its stability and reliability is well-known and therefore used in various applications. However, the power consumption of the 6T SRAM cell can be further optimized by some modifications resulting in the SRAM cells described in the following paragraphs:

1) *4T SRAM cell*: A typical implementation of a four transistor SRAM cell is shown in Figure 4. In comparison to the 6T cell, a smaller area of approximately 30% can be achieved [16]. Due to the replacement of all pMOS transistors by polysilicon resistors, only nMOS transistors are used for the pure functionality of this cell. Despite of the space-savings, which could lead to a higher yield after the manufacturing process, the realization of high-resistivity polysilicon resistor adds additional technological steps to the manufacturing process, resulting in higher costs.

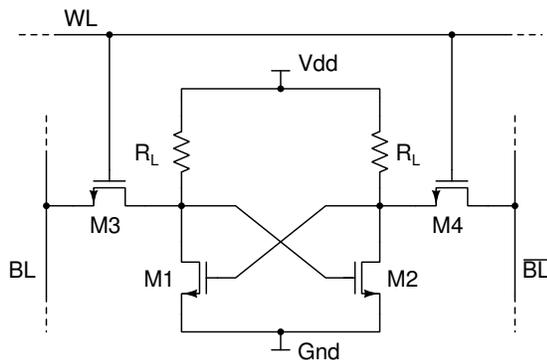


Figure 4. 4T SRAM cell

The 4T (polysilicon) SRAM is a predecessor of all CMOS-based SRAM cells. Lower stability, lower tolerance against soft-errors and a more technically demanding manufacturing process exclude this cell type from further considerations [6].

2) *5T SRAM cell*: The circuitry of a five transistor SRAM cell is shown in Figure 5. The advantage of this cell design compared to the 6T reference cell is the availability of just one access transistor  $M5$  and therefore only one bitline  $BL$  [17]. The connecting bitlines in each slice of an FPGA add undesired parasitic capacitances, which underly the process of charging and discharging during each read- and write-cycle and lead subsequently to higher power consumption. A cell design working with just one access transistor adds space-savings. For a proper and stable functionality of this cell, asymmetric transistor sizing is required, which may complicate the manufacturing process and to modifications of auxiliary circuitry like sense amplifiers, precharge circuits, etc..

3) *7T SRAM cell*: The seven transistor SRAM cell is shown in Figure 6, which enhances the 6T reference cell design by an additional feedback transistor  $M7$  and 2 signal lines  $R$  and  $W$ . The idea behind this design is a write mechanism, which depends only on one of the two bitlines in order to execute a write operation. This can be also expressed in equation (3) [12].

While the activity factor  $\alpha$  equals 1 in conventional memory cells, the 7T SRAM cell reduces this factor to less than 0.5 by exploiting the fact, that most of the bits in memories and caches are zeros [8]. The main asset of this implementation is

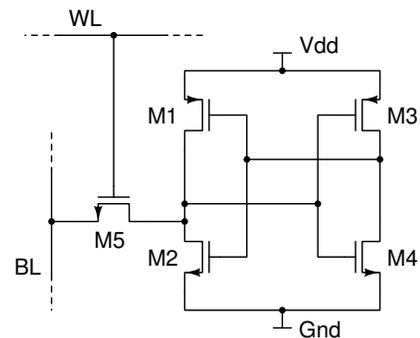


Figure 5. 5T SRAM cell

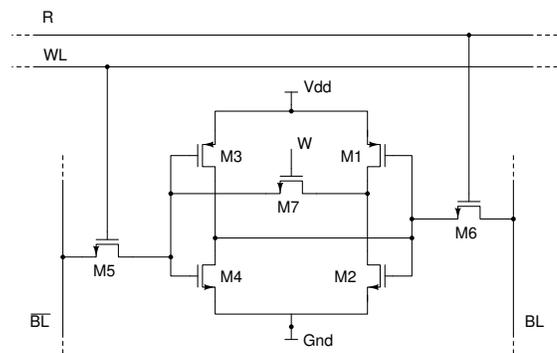


Figure 6. 7T SRAM cell

the reduction of the switching activity and therefore a reduction of charging and discharging cycles of parasitic capacitances. The drawback is the required additional control logic and the loopback transistor, which lead to higher complexity and required space.

$$P = \alpha C_{BL} V^2 F_{write} \quad (3)$$

## V. SRAM CELL DESIGN MODIFICATIONS

The simulation results showed that the choice of a suitable SRAM cell design leads to a significant impact on power consumption of a LUT. In this section we present further improvements on each cell design in order to achieve even better power savings in this essential component. Since Xilinx' Spartan 3(A) is manufactured in a 90nm process and has a recommended internal supply voltage of 1.2V, we choose a 90nm TSMC technology library at an comparable operating voltage of 1.2V.

Coming back to the proposed cell designs in Section IV, we refer to the 4T SRAM cell since its compact design is of interest for further considerations and performance comparison to other design. The major drawback of the 4T SRAM cell is the high-resistive polysilicon resistor, which should be replaced or completely omitted in an improved cell. A possibility how to bypass this drawback is shown in Figure 7.

The previous pull-down network (PDN) consisting of two nMOS transistors is replaced by a pull-up network of two pMOS  $M1$  and  $M2$  transistors [18]. In combination with both nMOS access transistors  $M3$  and  $M4$  a stable and power saving functionality is achieved. Instead of precharging both

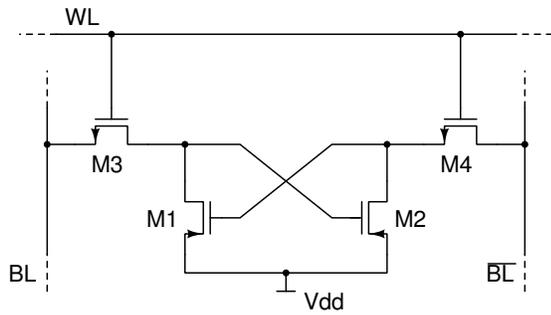


Figure 7. 4T loadless SRAM cell

bitlines to  $V_{dd}$  as a pre-step of the reading-phase, the bitlines are "precharged" to  $Gnd$ , due to the fact that pMOS transistor are used as drivers in this cell. This saves power and ensures compatibility with CMOS logic processes. Nevertheless, minor adaptations to the auxiliary circuitry around the cell have to be done, e.g., modifying the bitline drivers.

#### A. Test results

All SRAM cells have been designed and simulated by usage of the Cadence toolchain and a 90nm technology provided by TSMC at an ambient temperature of 27°C. The main challenge to achieve comparable results was to develop suitable bitline drivers, precharge circuitry and a sense amplifier. Careful design of the bitline drivers is crucial for avoiding the cell to flip during a read cycle. All simulations are performed with a clock frequency of 200MHz and a load of 600aF. Configuration memory cells used in a LUT are not supposed to be written and read at high frequencies, like, e.g., memory arrays in a microprocessor's cache (up to 4GHz). Therefore, we choose a lower frequency, nevertheless all cells have also been successfully tested with a higher clock frequency of 500MHz. All cell designs have been applied to the test circuit in Figure 8.

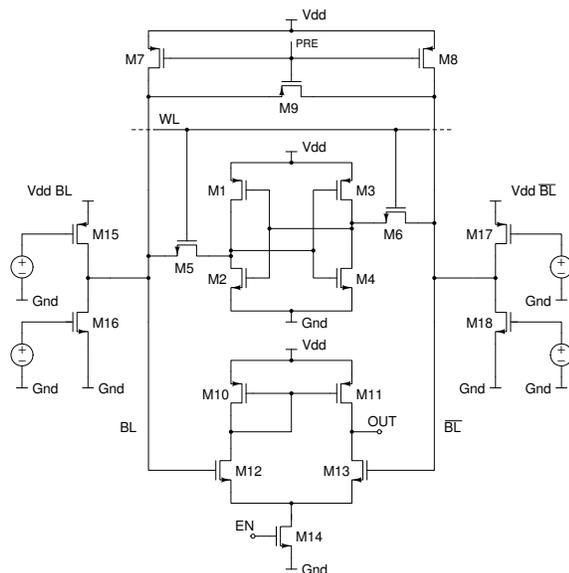
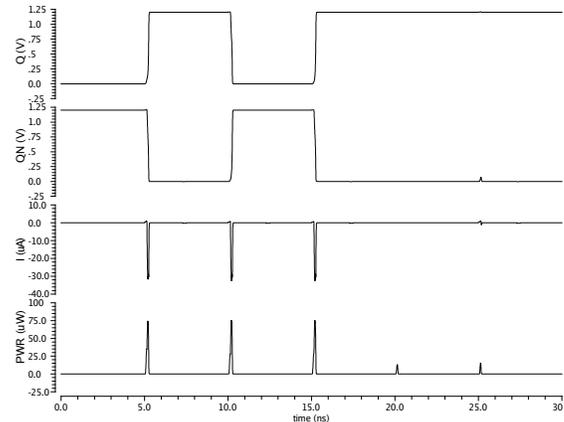


Figure 8. Test circuit

Figure 8 shows a 6T SRAM cell as DUT (design under test), the precharge circuit consisting of transistors  $M7$ ,  $M8$  and an equalizing transistor  $M9$ , two bitline drivers ( $M15$ ,  $M16$  and  $M17$ ,  $M18$ ) a sense amplifier. For the first step, the determination of the best SRAM cell design in terms of power consumption without any further improvements, is done. The simulation results of the 6T cell design are shown in Figure 9.


 Figure 9. Power dissipation and  $I_{Leak}$  of 6T SRAM cell

The average power consumption, the maximum and minimum power consumption during simulation time were traced and summarized in Table II.

TABLE II. SIMULATION RESULTS WITHOUT MODIFICATIONS

SRAM cell	Average Power nW	Max. Power uW	Min. Power pW
4T	334.5	35.07	161.7
5T	587.2	61.26	217.34
6T	927	75.39	250.8
7T	491	49.19	221.7

Compared to the other designs, Table II shows clearly the drawbacks of the reference 6T SRAM cell. Substantial power savings can be achieved by the choice of alternative cell design. For example, the average power consumption of the 6T SRAM reference cell design is 927nW and about 3 times higher than the average power consumption of the 4T loadless SRAM cell, which is only 334.5nW. That results in power savings of approximately 65%.

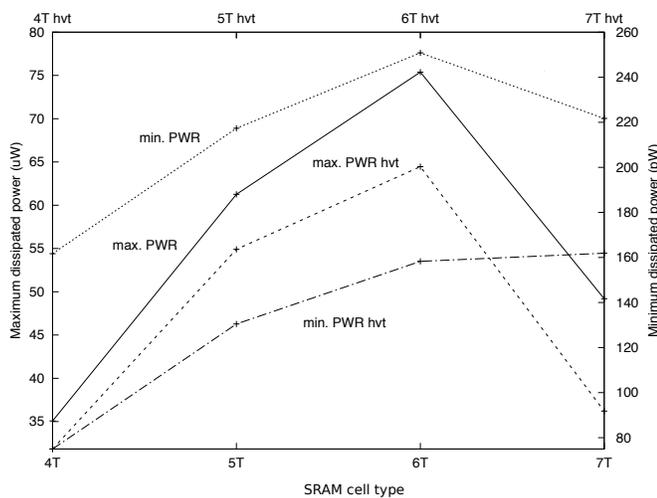
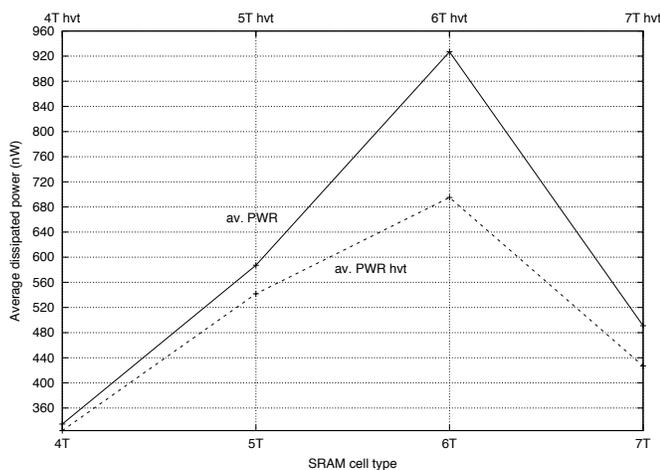
#### B. Dual Threshold CMOS

Further optimizations can be achieved by the introduction of high threshold voltage ( $V_{th}$ ) transistors. High  $V_{th}$  transistors require a higher  $V_{GS}$  voltage at the gate in order to turn the transistor on, which can lead to an increase of the propagation delay within a signal path. Therefore, high  $V_{th}$  should be only used in applications which are not timing-critical. However, the SRAM cells in a LUT are used as configuration RAM (CRAM) and are pertinent for use with high threshold voltage transistors. All cell designs have been modified and the simulations were performed again. These modifications are limited to the core cell only, the precharge circuitry, the sense amplifier and the bitline drivers have not been modified. The results are summed up in Table III.

TABLE III. SIMULATION RESULTS WITH HIGH THRESHOLD VOLTAGE TRANSISTORS (hvt)

SRAM cell	Average Power nW	Max. Power uW	Min. Power pW
4T hvt	324	31.83	74.99
5T hvt	541.78	54.9	130.5
6T hvt	695.1	64.46	158.3
7T hvt	427	36.21	161.9

In comparison to the reference design of the 6T SRAM cell, the introduction of the high  $V_{th}$  transistors adds power savings of about 25%. The performance of the high  $V_{th}$  4T loadless SRAM cell is slightly improved and leads to energy savings of approximately 10nW. In general, we can say that this modification improves both, the maximum and minimum energy consumption of all introduced cells. For illustration purposes, these improvements are shown in Figure 10 and Figure 11.

Figure 10. Comparison of minimum and maximum PWR dissipation between standard and high  $V_{th}$  designsFigure 11. Comparison of average PWR dissipation between standard and high  $V_{th}$  designs

Since the 4T loadless SRAM cells offers an excellent out of the box power balance, the strongest impact of the first

optimizing steps can be primarily noticed on all previously already existing designs, especially the 6T SRAM cell. Independent from the respective implementation, each exploited measurement is positively influenced by the modification done in a step before.

### C. Transistor Stacking

Transistor stacking, shown in Figure 12, which is also known as self-reverse biasing, is a strong technique to reduce subthreshold leakage current by raising the voltage at the source terminal of each transistor. By constantly increasing the source voltage  $V_S$  and keeping the gate voltage  $V_G$  at the same level,  $V_{GS}$  becomes negative at a certain point of time, which leads the transistor into super cut-off mode and turns it deeply off. Subthreshold currents are exponentially reduced.

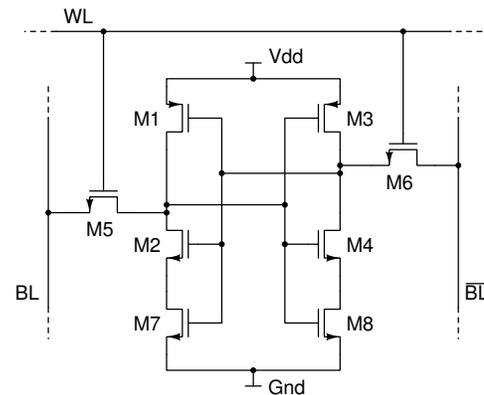


Figure 12. 6T SRAM cell with stacking

At the same time, the body to source potential  $V_{SB}$  also becomes negative, since the body terminal of a nMOS transistor is usually kept at  $Gnd$ . In consequence, the body effect is intensified, thus  $V_{th}$  is tuned by that effect to a higher level. This fact can be further exploited by continuing stacking transistors in series, but the effect of subthreshold current reduction becomes diminished with a rising number of transistors. This technique implies a trade-off between power savings and size ratio of the chip. Despite the gradual technology shrink up to 16nm FinFET, on-chip space is not an unlimited resource and should be used carefully. Therefore, we choose to add two stacking transistors only in order to have a reasonable compromise between leakage current reduction and size-ratio of the cells. The simulation results are shown in Table IV and Table V.

TABLE IV. SIMULATION RESULTS WITH STANDARD TRANSISTORS AND STACKING

SRAM cell	Average Power nW	Max. Power uW	Min. Power pW
4T	346.8	35.31	137.6
5T	327.4	25.1	189.4
6T	826.6	72.05	274
7T	540.4	31.64	168.3

If the used manufacturing process does not support dual-threshold CMOS technology, Table IV shows that a noteworthy reduction of leakage currents within the 4T SRAM cell is achieved by approximately 90%. Even the standard 6T SRAM

TABLE V. SIMULATION RESULTS WITH *hvt* TRANSISTORS AND STACKING

SRAM cell	Average Power nW	Max. Power $\mu$ W	Min. Power pW
4T hvt	336.6	32.79	70.42
5T hvt	327.4	25.1	189.4
6T hvt	672.4	61.28	167.4
7T hvt	461.8	30.84	523.9

cell features important amendments in terms of power savings ( $\approx 12\%$ ) and leakage currents.

The combination of both techniques, dual-threshold CMOS and transistor stacking, puts additional improvements to the overall power savings parameters. Since most of the currently available technologies feature dual-threshold CMOS, the feasibility of this combination is high.

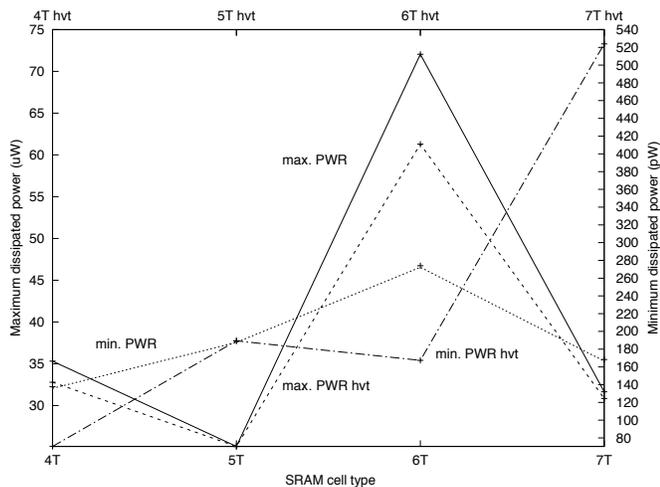
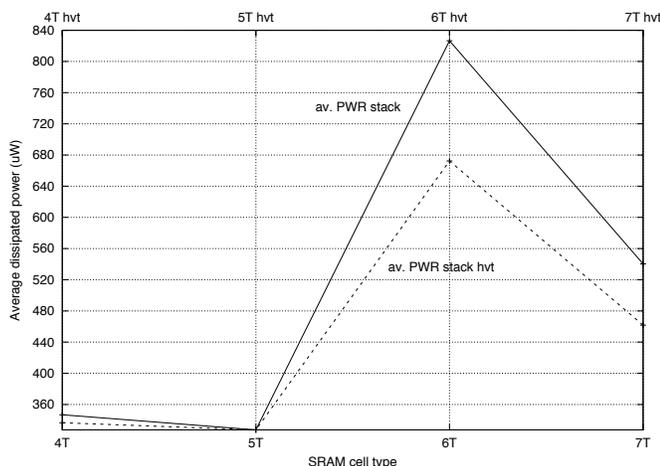
Figure 13. Comparison of minimum and maximum PWR dissipation between standard and high  $V_{th}$  designs with stackingFigure 14. Comparison of average PWR dissipation between standard and high  $V_{th}$  designs with stacking

Figure 13 and Figure 14 display different characteristics of the consumed power by referring to the values of Table IV and V. The 5T SRAM offers a slight advantage compared to the 4T loadless SRAM design due higher active and standby

intrinsic power consumption of less transistors when applying stacking to a logic design. Nevertheless, the 4T loadless SRAM cell still performs better than the remaining memory circuits. Another interesting aspect to be considered is the signal to noise ratio, which gives a benchmark about the margin between the transferred signal or stored data inside a memory cell and the influence of background noise on the signal lines, which can not be neglected. This factor is even of higher significance when volatile memory cells are equipped with high  $V_{th}$  transistors, replacing their standard  $V_{th}$  counterparts. To investigate potential undesirable side effects on the intended functionality, Figure 15 shows the butterfly plots of two 6T SRAM implementations, each realized with high and standard  $V_{th}$  transistors. It can be seen that this modification has a small impact on the signal to noise margin, but this is still acceptable due to predominant benefits in terms of power savings.

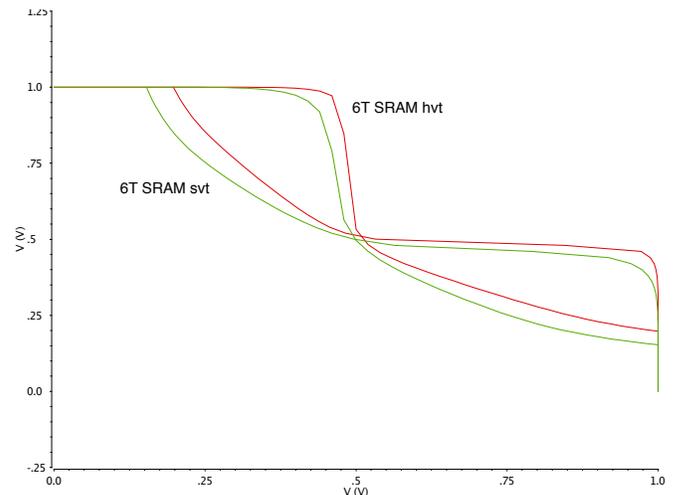


Figure 15. Butterfly plots of 2 different 6T SRAM implementations

#### D. Dynamic Voltage Scaling

The higher the supply voltage is, the faster the operation of the integrated circuit will be, since high  $V_{dd}$  allows fast charging and discharging of parasitic capacitances. In case of low demand on performance such as for CRAMs, the supply voltage can be lowered while still ensuring data retention within the cell. Dynamic voltage scaling (DVS) depends usually at least on an operating system and a regulation loop to recognize the circuit speed and to cover a wide range of operating voltages. Our approach simplifies this principle by introducing two additional transistors, shown in Figure 16.

Both transistors  $M9$  and  $M10$  are used to connect the SRAM cell to two different supply voltages,  $V_{dd}$  and  $V_{ddL}$ , whereas  $V_{dd}$  equals the primary 1.2V. On the one hand, the prerequisite of this method is a dual- $V_{dd}$  setup, representing a simple alternative to the mentioned operating system driven regulation loop, and on the other hand, a modified power gating approach is implemented. Since the 4T SRAM cell has no connection to  $Gnd$  in its core, power gating is achieved by the possibility to fully cut-off the supply voltage, if needed. However, power gating should be introduced at a coarse-grain level, e.g., by powering or switching off groups of cells at a higher abstraction layer. By lowering the supply voltage to  $V_{ddL}$ , which equals 1V, we can further reduce leakage

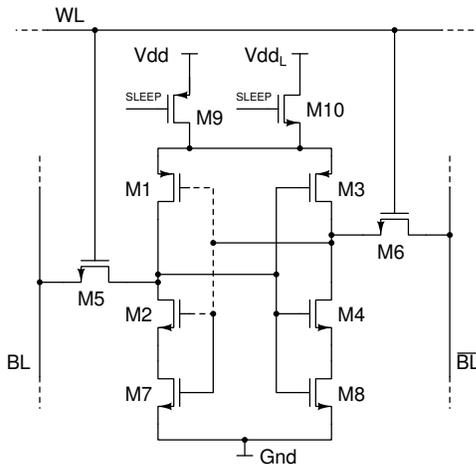


Figure 16. 6T SRAM cell with *hvt* transistors, stacking and DVS

power consumption. Experimental results have shown that data retention will still be ensured at supply voltages down to 400mV. A combination of all three power saving mechanisms in a 6T SRAM cell is shown in Figure 16.

TABLE VI. SIMULATION RESULTS WITH *hvt* TRANSISTORS, STACKING AND DVS

SRAM cell	Average Power nW	Max. Power uW	Min. Power pW
4T <i>hvt</i>	232.9	21.27	49.59
5T <i>hvt</i>	327.4	25.1	189.4
6T <i>hvt</i>	458.7	44.67	166.1
7T <i>hvt</i>	368.3	26.53	167

In order to achieve an average power consumption of 232.9nW at a clock frequency of 200MHz and full data retention like shown in Table VI, we combined all three power saving methods introduced in the chapters before with careful transistor sizing of an efficient memory cell design. We present the modified, loadless 4T SRAM cell in Figure 17.

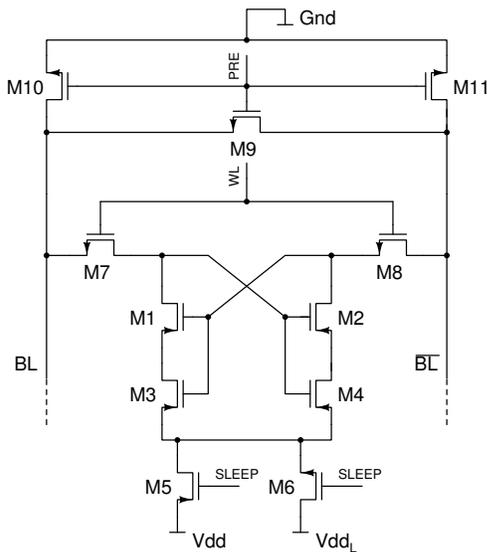


Figure 17. Modified 4T SRAM cell

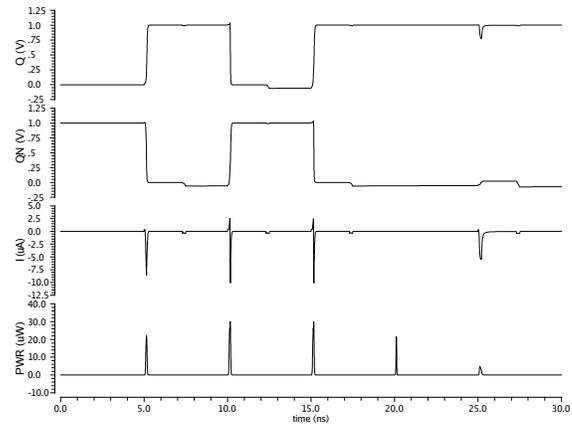


Figure 18. Power dissipation and  $I_{Leak}$  of a modified 4T SRAM cell

The simulation was done by injecting a 1 → 0 → 1 sequence and one read cycle at the end of the simulation time, which can be seen in Figure 18. By comparing the results of Figure 18 with the outputs shown in Figure 9, we see a reduction in both, power and current spikes. Looking back on the continuous improvements added to each cell type, we see the benefits in reduction of average power consumption in Figure 19.

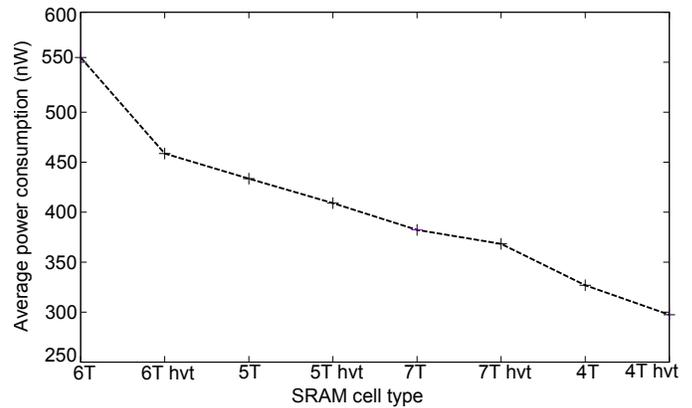


Figure 19. Power dissipation reduction

Analogue to the previous sections and in terms of a good overview, all results from Table VI are displayed in Figure 20. For highlighting the effects of all applied modifications, the simulation plots of all original designs are added to the same Figure.

Figure 20 and Figure 21 underline the numbers in Table VI. Both illustrations, especially Figure 21, clearly depict the advantages of combining an SRAM cell design with inherent power efficiency and appropriate modifications for even better energy savings in applications with limited resources.

## VI. LUT SIMULATIONS

The LUT was implemented with each cell type investigated in the previous chapters. In order to achieve an equal distribution of bits, all memory cells have alternating bits stored and are not connected to the bitlines by switching off all access transistors. As a matter of lucidity, we present a comparison

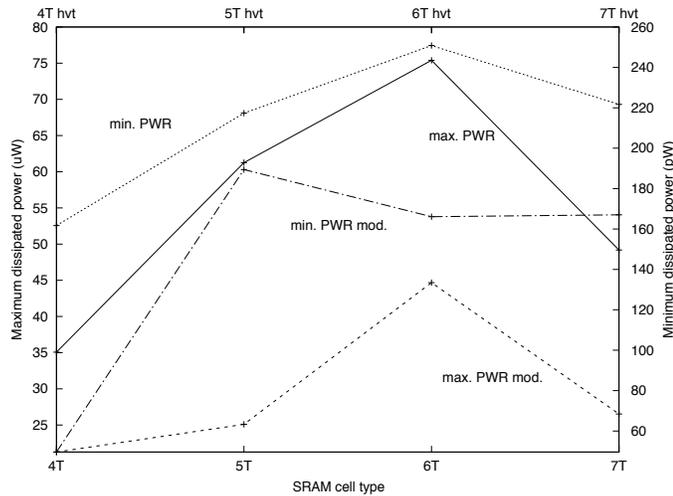


Figure 20. Comparison of minimum and maximum power dissipation between original and modified cell designs

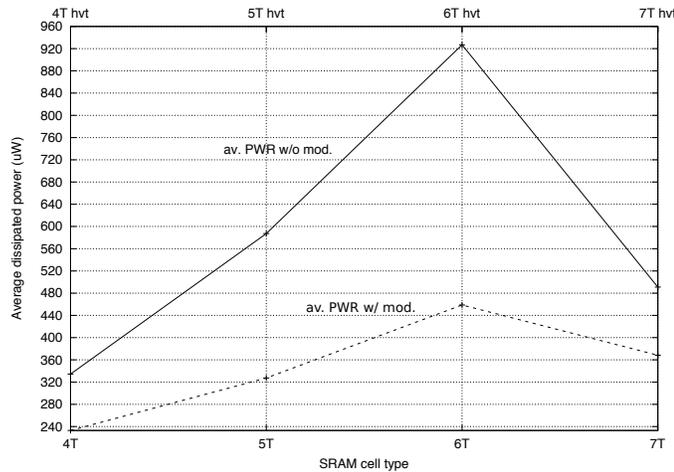


Figure 21. Comparison of average power dissipation between original and modified cell designs

between the 6T SRAM- and 4T SRAM LUT implementation. As expected, the 4T SRAM cell design shows a better performance in terms of power savings and leakage current reduction than the 6T SRAM cell design does. By comparing a LUT implementation with a standard 6T SRAM cell and our modified 4T SRAM design, Table VII summarizes the results and highlights the improvements in power dissipation, which equals power savings of approximately 16%. Figure 22 shows the related leakage current of the 4T SRAM based LUT.

TABLE VII. LUT COMPARISON

SRAM cell	Average PWR nW	Max. PWR uW	Min. PWR nW	Energy aJ
4T hvt	424.2	40.94	0.24	127
6T	500	42.99	2.8	150

It should be mentioned that either the precharge circuit nor the sense amplifier have been optimized for power efficiency. Optimizing these parts will lead to even better results and raise the duration of a battery charge, independent of the target application. Further optimization can be achieved by

coarse-grain power gating of CRAM blocks within the LUT architecture. Unused CRAMs should be completely powered off by adding additional, thick-oxide transistors, cutting off the cell from  $V_{dd}$  and  $G_{nd}$ .

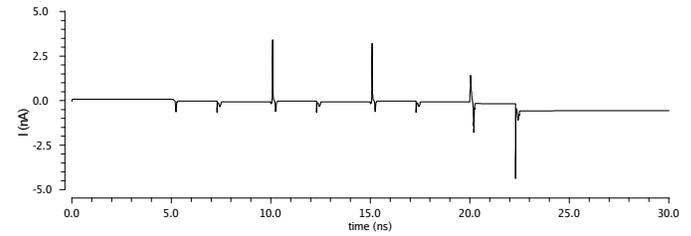


Figure 22. Leakage current of an improved 4T SRAM based LUT

The modified 4T memory cell design introduced in Figure 17 is superior in terms of low power aspects compared to all other investigated cell designs. However, this solution requires additional space, since it requires at least four additional transistors to achieve its intended power-efficient functionality.

## VII. TIMING CONSIDERATIONS

Despite the fact that timing aspects play a minor role for volatile memory cells used for configuring LUTs, a closer investigation of, e.g., the maximum operating frequency  $f_{max}$  is helpful to sound the limits of these circuits for their intended usage. In special cases like critical real time calculations, fast reconfigurability of a programmable logic device may be a inevitable requirement. This maximum operating frequency can be determined by (4):

$$f_{max} = \frac{1}{t_{HL} + t_{LH}} \quad (4)$$

The summands  $t_{HL}$  and  $t_{LH}$  display the time necessary for a *HIGH*  $\rightarrow$  *LOW* and *LOW*  $\rightarrow$  *HIGH* transition respectively. Figure 23 shows a typical 1  $\rightarrow$  0 switching event with additional measurement marks at 90% and 10% of the supply voltage  $V_{dd}$ , including the  $\Delta$  of time, which is considered to be  $t_{HL}$ . All considered SRAM cells have been investigated upon these characteristics and compared against each other. The correspondent results are summarized in Table VIII.

TABLE VIII. TRANSITION TIMES AND MAXIMUM OPERATION FREQUENCY

SRAM cell	time LH ps	time HL ps	Max. freq. GHz
4T hvt	40.22	38.74	12.66
5T hvt	58.24	132.41	5.25
6T hvt	102.53	60.7	6.12
7T hvt	62.47	380.87	2.25

The results reveal the superior performance of the 4T loadless SRAM in terms of elapsed time for both  $t_{HL}$  and  $t_{LH}$ . In direct comparison to the reference 6T SRAM cell, we achieve an improvement of  $\approx 60\%$  for the *LOW*  $\rightarrow$  *HIGH* transition. The improvement for the complementary operation *HIGH*  $\rightarrow$  *LOW* is less, but energy savings of  $\approx 36\%$  are still noteworthy.

For having a better overview about the different slew rates and maximum operating frequencies, all results were visualized in Figure 24. The 4T loadless SRAM cell outperforms the

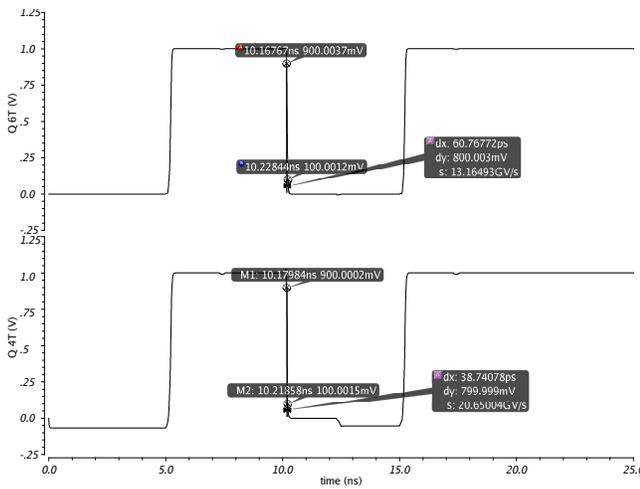
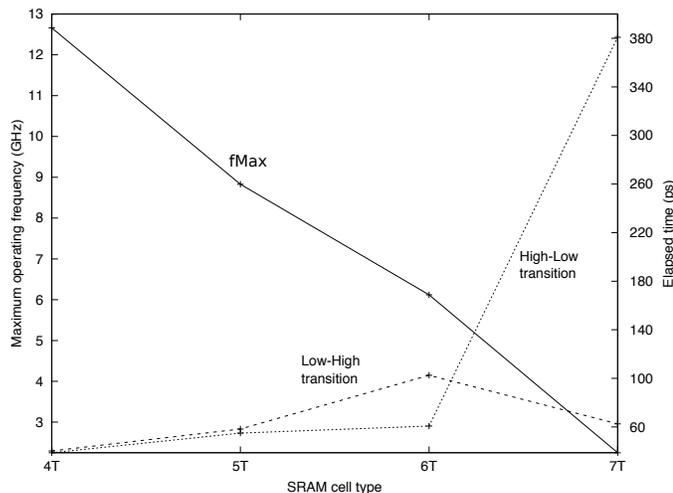


Figure 23. 4T &amp; 6T SRAM HIGH LOW transition

other cells in each considered aspects. The maximum operating frequency gives an impression about the capabilities of this newly developed cell to be used for calculations in critical real time environments. It should be mentioned, that this designed was not optimized for short channel effect suppression. The recent proceedings in process technology lead to a continuous design shrink, which come along with a significantly higher yield in manufacturing. The downside of these achievements are undesirable physical effects, e.g., the short channel effect. The smaller the channel length becomes, the higher leakage currents in standby mode will be, due to tunneling effects of electrons from drain to source even without establishing a steady voltage  $V_{GS} > V_{th}$  at the gate of a transistor.

Figure 24. Slew rates and  $f_{max}$ 

## VIII. CONCLUSION

We analyzed a typical LUT structure of an FPGA in terms of power dissipation and leakage current. Our approach was to integrate power savings mechanisms at the basic circuit level before heading for further optimizations on architectural level. Different SRAM cell structures have been investigated on their power characteristics in order to evaluate the best design

for implementing a LUT, which features inherent low-power characteristics. Simulations have shown that the 4T loadless SRAM cell features the required properties. We applied various low-power techniques and enhanced this cell for standby leakage current mitigation. Hence, we presented a modified 4T loadless SRAM cell design. By combining dedicated techniques during design time and during operating time, we achieved a reduction of the average power consumption within the LUT of 16% during simulation time. Subsequently, this leads to overall energy savings of 127aJ compared to the origin 150aJ of a 6T SRAM cell based LUT implementation. The leakage current  $I_{leak}$  is reduced dramatically from 1.741nA to approximately 0.2nA, showing the strong impact of leakage reduction methods on power-critical circuitry. On the other hand, the 4T loadless SRAM design offers very fast reconfiguration capabilities due to its remarkably high top operating frequency  $f_{max}$  above 12GHz, which outperforms the alternative designs significantly. This goes back to a good slew rate for each of both transitions during signal processing. The overall performance of this design could be further enhanced by careful adaption of the gate length to catch up disadvantages in terms of short channel effects. It can be predicted that this measure would further decrease standby leakage current  $I_{leak}$ , but it would automatically lead to increasing the width of the transistors to avoid any penalties in speed. This would be acceptable if area considerations do not play a role, which is rather unlikely due to a preferable high yield at the end of a semiconductor manufacturing process. All identified and investigated pros come along at the cost of certain adaptations to peripheral circuitry, e.g., the sense amplifier, which needs to be redesigned for the usage with this newly developed memory design. Additional wiring for the DVS transistors is required which is synonymous with more parasitic capacitances and area consumption. These challenges have to be faced during layout phase after synthesis and strongly depend on the chosen process node. Since registers and GPIOs occupy a large amount of area in FPGAs, further power savings can be achieved by adapting the architecture of these circuits. This will be addressed in future publications.

FPGAs support adaptiveness of whole systems by re-configuration abilities on demand of the respective application. The presented low-power cell design reduces power consumption significantly during the charging and discharging cycles of re-configuration tasks within an FPGA and delivers an overall good performance leading to an appropriate suitability for mobile low power applications.

## ACKNOWLEDGMENT

The authors thank Amit Majumdar, from Xilinx, for his support and interesting discussions on FPGA architectures. Also, we want to give credit to Ray Chiang, from TSMC, for his explanations and suggestions on the used technology. We are grateful to Andreas Ullrich, from University of Wuppertal, for his restless dedication in PDK compilation and tool maintenance.

## REFERENCES

- [1] K. Niewiadomski, C. Gremzow, and D. Tutsch, "4t loadless srams for low power fpga lut optimization," in Proceedings of the 9th International Conference on Adaptive and Self-Adaptive Systems and Applications (ADAPTIVE 2017), February 2017, pp. 1–7.
- [2] S. Fürst, "Challenges in the design of automotive software," in Proceedings of the Conference on Design, Automation and Test in Europe, ser. DATE '10. 3001 Leuven, Belgium, Belgium: European Design and Automation Association, 2010, pp. 256–258, last accessed on 2017-11-13. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1870926.1870987>
- [3] XA Spartan-3A Automotive FPGA Family Data Sheet, Xilinx, 04 2011, rev. 2.0.
- [4] M. Ullmann, M. Hübner, B. Grimm, and J. Becker, "An fpga run-time system for dynamical on-demand reconfiguration," in Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International. IEEE, 2004, p. 135.
- [5] R. Anthony, A. Rettberg, D. Chen, I. Jahnich, G. de Boer, and C. Ekelin, "Towards a dynamically reconfigurable automotive control system architecture," in Embedded System Design: Topics, Techniques and Trends. Springer, 2007, pp. 71–84.
- [6] A. S. Pavlov, "Design and Test of Embedded SRAMs," Ph.D. dissertation, University of Waterloo, Ontario, May 2005.
- [7] J. P. Uyemura, CMOS Logic Circuit Design. Norwell, MA, USA: Kluwer Academic Publishers, 1999.
- [8] R. E. Aly, M. I. Faisal, and M. A. Bayoumi, "Novel 7t sram cell for low power cache design," in Proceedings 2005 IEEE International SOC Conference, Sept 2005, pp. 171–174.
- [9] S. M. Jahinuzzaman, D. J. Rennie, and M. Sachdev, "A soft error tolerant 10t sram bit-cell with differential read capability," IEEE Transactions on Nuclear Science, vol. 56, no. 6, Dec 2009, pp. 3768–3773.
- [10] A. Lodi, L. Ciccarelli, D. Loparco, R. Canegallo, and R. Guerrieri, "Low leakage design of lut-based fpgas," in Proceedings of the 31st European Solid-State Circuits Conference, 2005. ESSCIRC 2005., Sept 2005, pp. 153–156.
- [11] T. Tuan, S. Kao, A. Rahman, S. Das, and S. Trimberger, "A 90nm low-power fpga for battery-powered applications," in Proceedings of the 2006 ACM/SIGDA 14th international symposium on Field programmable gate arrays. ACM, 2006, pp. 3–11.
- [12] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, Digital integrated circuits- A design perspective, 2nd ed. Prentice Hall, 2004.
- [13] C. Piguet, Low-power processors and systems on chips. CRC Press, 2005.
- [14] C. Maxfield, The Design Warrior's Guide to FPGAs: Devices, Tools and Flows, 1st ed. Newton, MA, USA: Newnes, 2004.
- [15] K. Itoh, VLSI Memory Chip Design, ser. Springer Series in Advanced Microelectronics. Springer Berlin Heidelberg, 2001, last accessed on 2017-11-13. [Online]. Available: <https://books.google.de/books?id=p2FsQgAACAAJ>
- [16] A. Bellaouar and M. I. Elmasry, Low-Power Digital VLSI Design Circuits and Systems, 1st ed., J. Allen, Ed. Norwell, MA, USA: Kluwer Academic Publishers, 1995.
- [17] I. Carlson, S. Andersson, S. Natarajan, and A. Alvandpour, "A high density, low leakage, 5t sram for embedded caches," in Solid-State Circuits Conference, 2004. ESSCIRC 2004. Proceeding of the 30th European, Sept 2004, pp. 215–218.
- [18] J. Yang and L. Chen, "A new loadless 4-transistor sram cell with a 0.18 m cmos technology," in 2007 Canadian Conference on Electrical and Computer Engineering, April 2007, pp. 538–541.