# Multi-Bit Non-Volatile Spintronic Flip-Flop

Christopher Münch, Rajendra Bishnoi and Mehdi B. Tahoori

Chair of Dependable Nano Computing (CDNC), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Email: christopher.muench@student.kit.edu, rajendra.bishnoi@kit.edu and mehdi.tahoori@kit.edu

*Abstract*—As leakage increases proportionally with the technology downscaling, it becomes extremely challenging to manage to meet the total power budget. This is because, CMOS-based logic blocks can not be completely power-gated as their flip-flops always require a retention supply to hold the system states. Alternatively, their data can be stored in a separate memory during the standby mode, however, that results in a huge area and energy overhead. *Spin Transfer Torque* (STT) based non-volatile flip-flops can offer normally-off/instant-on computing features to reduce leakage by complete power shut-down without the need to transfer and restore system states separately. The non-volatile component of such flip-flops can be easily shared for the overall design optimizations. In this paper, we design a unique multi-bit non-volatile flip-flop architecture using STT devices to reduce the area and energy costs associated with non-volatile components. This architecture is developed based on the resource sharing principle using a custom design that enables the optimization for the area and energy consumption. Moreover, we have developed a framework in which we have replaced the conventional neighbor flip-flops in the layout with our proposed multi-bit non-volatile designs. Results show that using our multi-bit flip-flop architecture, we improve the system-level area and energy by 26 % and 14 % in average, respectively, compared to the standard single-bit non-volatile flip-flop design.

## I. INTRODUCTION

Technology downscaling and high performance requirements of modern *System-on-Chips* (SoCs) lead to a substantial increase in static power [1, 2]. The static power, that is consumed in the form of leakage when devices are not operating, is nowadays dominating the total power consumption in a chip [3]. Moreover, this trend is increasing with technology advancements. Reduction of such power is extremely important, especially for the battery operated hand held devices. The best way to reduce leakage is by disconnecting the power supply using *power-gating*. However, conventional CMOS based flip-flops, which are the basic building blocks in an SoC, can not be power-gated completely, because they are volatile in nature, always require a supply voltage to retain their value. Furthermore, the conventional CMOS *save and restore* technique [4], in which the flip-flop contents are stored in a separate memory during the power-down and restored back during the wake-up, contributes to severe delay, area and routing overheads. Therefore, researchers are seeking for an alternative technology.

Emerging non-volatile *Spin Transfer Torque* (STT) is a promising technology, that is well-known for its zero-leakage storage capability [5–7]. In addition to the leakage reduction, this technology also has many other beneficial features such as scalability, high endurance, small footprint, CMOS compatibility, fast accessibility and immunity to soft-error due to radiations [8–11]. Recently, STT based *non-volatile* (NV) flip-flop designs have been gaining a lot of attention [12–17], that is primarily due to their *normally-off/instant-on* computing attributes. Such flip-flops are usually designed in the form of a shadow architecture in which the data back-up can be done locally using NV devices for each conventional CMOS flip-flop. This arrangement allows the logic design block to be completely powered-off during the standby mode, thus reduces leakage considerably. During power-up mode, the data is restored back to the conventional flip-flop and the normal operation is presumed. Nevertheless, the shadow component in this NV flip-flop itself occupies significant area and consumes high energy.

In SoCs, typically many flip-flops are placed adjacent (or close) to each other in the layout, meaning their shadow component can be easily shared between such flip-flops in order to gain in terms of area and energy. CMOS-based multi-bit flip-flop architectures share the clock among multiple flip-flops, that improves clock network power and area significantly [18–20]. However, for static power reduction, each of such flip-flops still have to be associated with NV components individually, leading to an overly designed architecture. On the other hand, a mini array type STT-based flip-flop design is proposed in [17] to provide a check-pointing solution for microprocessor applications. This design uses a dedicated reference circuitry, has a complex controlling mechanism and adds considerable area as well as power due to the extra circuitry. Thus, it is not feasible to use this architecture for the design sharing. Overall, a cost effective NV flip-flop architecture that works on the sharing principle to deliver a high area and energy efficiency is still missing.

In this paper, we propose a novel multi-bit shadow latch architecture, which is designed based on the sharing mechanism, that can merge the neighbouring single-bit NV components in the layout. In this design, we have organized the NV storing components around the actual sensing circuitry in such a way that the same sensing scheme can be used to read two bits, leading to a significant energy reduction. The functionality in this design is controlled in a unique way to perform a reliable back-up and restore operation. Moreover, we have developed layouts for both proposed as well as state-of-the-art NV components of the flip-flops to perform realistic area comparison. To demonstrate the system-level benefits, these designs are employed for various benchmark circuits and their evaluation is performed after analyzing the overall SoC physical-design behavior. Our simulation results show that, compared to the single-bit NV flip-flop design, our proposed architecture can reduce the system-level area and read energy by upto 31 % and 17 %, respectively.

The rest of this paper is organized as follows. In Section II, the basic of STT technology and non-volatile shadow flip-flop architecture is discussed. Section III explains our proposed multi-bit flip-flop architecture, followed by experimental results in Section IV. Finally, Section V concludes the paper.

## II. BACKGROUND

### A. Spin Transfer Torque technology

In STT, a *Magnetic Tunnel Junction* (MTJ) cell is the storing device, in which the value is stored in the form of resistance states. The MTJ cell, as illustrated in Figure 1, comprises of two ferromagnetic layers separated by a thin barrier oxide layer. One of the ferromagnetic layers has always a fixed magnetic orientation, that is known as *Referenced Layer* (RL). Whereas, the magnetic orientation of the other layer can be freely rotated, which is known as *Free Layer* (FL). When the magnetic orientation of the two ferromagnetic layers are parallel to each other ('P' configuration), it exhibits a low
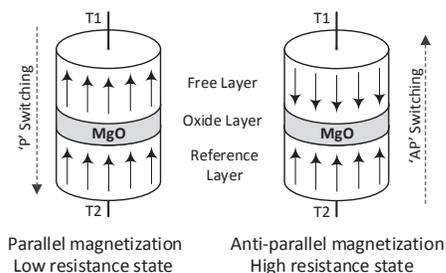
Fig. 1. Spin Transfer Torque based MTJ cell

resistance value. Otherwise, it has a high resistance value, when the magnetic orientation of those two layers are anti-parallel to each other ('AP' configuration). The current flow directions decide the required magnetic switching state as demonstrated in the figure. To read the content of the MTJ, a small read current has to pass through the stack.

### B. Shadow non-volatile flip-flop architecture

The block diagram for a shadow non-volatile flip-flop architecture is shown in Figure 2(a). It consists of total three components, namely, *master-latch*, *slave-latch* and *NV-latch* [12–16]. The conventional CMOS flip-flop is the combination of master and slave latch designs. Whereas, the NV-latch is added to it for the data back-up storage, as shown in the figure. The data from the conventional flip-flop is stored in the NV-latch during the power-down mode, and restored back during the power-up mode of operation. These store and restore operations are controlled by a PD (Power-down) pin, that is activated/deactivated at system-level based on the application. Furthermore, the shadow latch comprises of two MTJs in addition to their read and write components as illustrated in Figure 2(b). In this latch design, the write circuitry is organized in such a way that the two MTJs should always store the opposite magnetization. This is because, this arrangement assists the sensing of the resistance differences during the read process. Overall, using this NV shadow flip-flop architecture, the entire logic core in an SoC can be power-gated after the data back-up unlike for the conventional CMOS-only flip-flop design, leading to a significant static power reduction.

### C. Related work

*Power-gating* is the most efficient technique for the static power reduction, however, this method is complicated for the conventional CMOS-based flip-flops as they always need a supply voltage to retain their data. One solution is to employ *save and restore power-gating* technique in which the content of flip-flops are stored in a memory array during the power-down mode [4]. Nevertheless, this technique is only applicable for very long idle periods and it incurs huge area and latency costs due to the extra storage and data transfer, respectively.
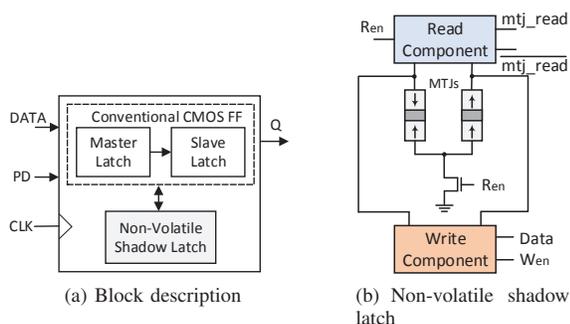
Hence, non-volatile STT-based flip-flop designs are gaining popularity as the entire logic core can be power-gated, and the data back-up storage can be done locally for each flip-flop. This way, the data store and restore during the standby mode become fast and low cost.

Several non-volatile flip-flop architectures have been proposed to reduce the static power consumption. For instance, a conventional MRAM[1] based flip-flop is employed in a 16-stage 8-bit shift register design block [21]. On similar concept, an STT-based flip-flop is demonstrated in [22], where a current sensing scheme for the read and a NOR-based write mechanism is illustrated. Another STT-based retention flip-flop is proposed in [12], in which the read and write circuitry are separated from the storing device. Similarly, several *Spin Orbit Torque* [23, 24] as well as other non-volatile technology based flip-flops are also proposed [25–27]. However, all these architectures need a dedicated read and write mechanism to operate on a single bit. That means, it is not possible to take advantage, by the custom design schemes, for the flip-flops that are placed close to each other.

A NV flip-flop design proposed in [17] demonstrates the multi-bit storage system. In this design, several NV devices are organized in the form of a mini-array to provide a back-up for checkpointing purpose. This scheme requires a special reference cell that is manufactured in a way that its resistance value is tuned to the mid of the two storage resistance values. On the top of that, to access multiple storage devices in this architecture, a decoder logic is required, whose controlling mechanism makes the design more complex. Overall, these additional circuits impose not only extra area but also consume more energy.

In this work, we propose a multi-bit non-volatile latch architecture that is designed based on the resource sharing principle and with its unique custom-design strategy, the overall design area and energy are optimized. Please note that, multi-bit CMOS-only based flip-flop designs, that are highly adopted by industries for the reduction of clock network power and area, are working on the similar concept of the resource sharing [19, 20]. However, for the static power reduction, these designs also require a back-up using a non-volatile storing component. Therefore, our proposed multi-bit non-volatile component can easily be integrated in such designs, that can further enhance the overall efficiency of an SoC in terms of both static and dynamic energy consumption as well as area.

### III. PROPOSED NON-VOLATILE MULTI-BIT LATCH

In this section, we first describe the overview of the idea and then a block description of the proposed design. Afterwards, implementation details of the proposed design and their simplified controlling mechanism are explained. In the end, the integration with the multi-bit CMOS flip-flop is described, followed by the explanation of the design scalability for our proposed architecture.

### A. Overview

In general, in an SoC, the placement of standard-cells is done in an automated way using EDA (*Electronic Design Automation*) tools. During this process, regardless of the timing constraints, a large number of flip-flops are placed very close to each other (see Figure 9). The NV shadow components for such flip-flops are also placed close by. This means that some of the circuit components that are common to these shadow part can be merged. A block diagram that illustrates this sharing principle is shown in Figure 3. As shown, the multi-bit NV shadow component is shared between the two conventional CMOS flip-flops. Here, values from these two conventional CMOS flip-flops are stored in the shadow latch during the power-down mode



(a) Block description      (b) Non-volatile shadow latch

Fig. 2. Illustration of shadow flip-flop architecture

---

[1]Here MRAM is *Magnetic Random Access Memory* technology where field induced magnetic switching scheme is used for the storage.
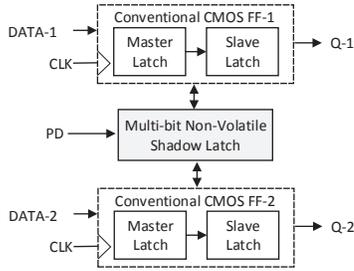
Fig. 3. Overview of multi-bit shadow flip-flop architecture

before disconnecting the power supply. During power-up, values from the shadow latch are restored back to their respective flip-flops, so that the normal operation can be resumed. Similar to the standard single-bit NV flip-flop, these store and restored operations are also controlled using a common PD pin. In this work, we propose this multi-bit shadow latch architecture that enables the optimization for the area as well as energy.

*B. Block-level implementation*

Figure 4(a) shows another way of the implementation of the shadow latch, as opposed to the conventional design (demonstrated in Figure 2(b)). Unlike the standard NV shadow latch, here the two MTJs are connected above the read component and the read operation is enabled using a PMOS transistor based on the $R_{en}$ signal, as shown in the figure. The two output signals, mtj_read and $\overline{\text{mtj\_read}}$ are generated based on the resistance states of those two MTJs. Similar to the standard NV design, these two MTJs are also connected in such a way that they should always store opposite values during the write operation.

We propose an architecture in which the above mentioned latch architecture (as shown in Figure 4(a)) and the standard NV latch design (as shown in Figure 2(b)) can be combined to store two bit values. The block description of this combined architecture is illustrated in Figure 4(b). As shown, in this way, the common circuit elements such as read component and other controlling circuits, can be merged. On the other hand, we intentionally do not modify the write components to maintain the error-free storing. This is due to the fact that the MTJ store operation is very sensitive to the current value and its duration of flow, and any minor disturbance in the current level during the write operation can easily lead to an error. Additionally, these write components can be easily overlap with the master/slave circuitry of the conventional CMOS flip-flop design [24]. Therefore,



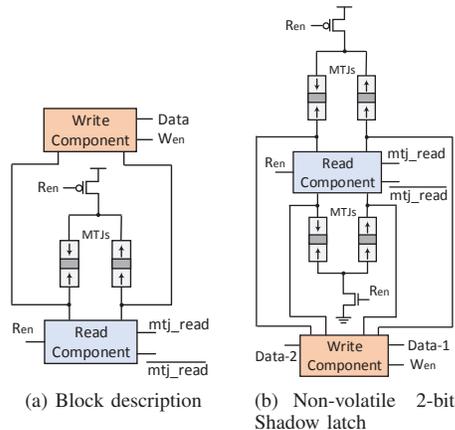(a) Block description   (b) Non-volatile 2-bit Shadow latch

Fig. 4. Block diagram description of shadow flip-flop architecture
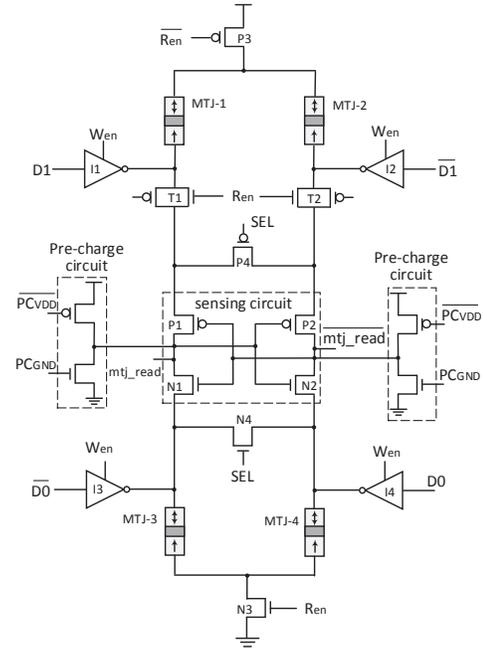


Fig. 5. Schematic of the proposed 2-bit shadow latch architecture

in this work, our main focus is to optimize the read components. The detail of these read/write components and their simplified controlling mechanism is discussed in the next subsection.

*C. Circuit-level implementation*

The transistor-level schematic of our proposed 2-bit non-volatile latch design is shown in Figure 5. As shown, the proposed circuit is composed of a sense amplifier, pre-charge circuitry, four MTJs and write circuitry. The sensing circuitry is designed based on the pre-charged current based sense amplifier [28]. The purpose of the pre-charge circuit is to maintain the two outputs at the equal potential, which is necessary before the actual every read operation. Furthermore, as described earlier, out of four MTJs, two are connected above the sensing circuit whereas the other two are connected below the sensing circuit. Additionally, a PMOS (P4) and an NMOS (N4) transistors are connected to stabilize the outputs during the read operation. On the other hand, the write components are designed to drive the write current (enabled only during the standby mode) through the MTJs. As described earlier, these write components can be overlapped with the master/slave inverters with a transmission gate. Nevertheless, for the sake of completeness, we demonstrate these write components using tristate inverters as shown in the figure.

Our proposed 2-bit shadow latch design operates in two phases, i.e. *store* and *restore* phases. In reality, these two phases are controlled using the clock and the power-down (PD) signal, which are the inputs to this design. The working of store and restore phases is described as follows:

*1) Store phase:* During the store phase, when the design is in the standby mode, the two bit data have to be written in the NV component of the latch. To write the value, a definite amount of current has to pass through the MTJs for a constant amount of time. The tristate inverters facilitate this write current to flow and the direction is decided based on the input values. Both pairs of MTJs have two tristate inverters each, which have always complementary inputs. For instance, to write the D0 value from the FF0, the D0 is applied to I4 inverter and its complimentary value ($\overline{D0}$) is applied to
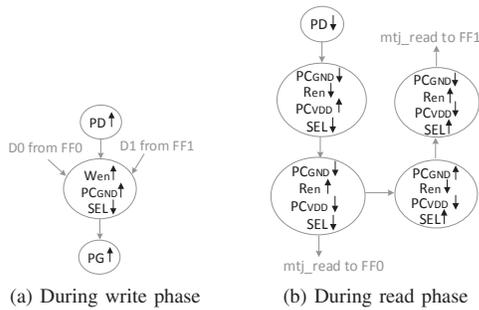
(a) During write phase    (b) During read phase

Fig. 6. Working sequence of the proposed multi-bit latch. Here, ↑ and ↓ indicate signals are enabled and disabled, respectively. PD is for *power-down* and PG is for *power-gating* signal.



(a) Pre-charge circuit    (b) working details

Fig. 7. Optimized pre-charge mechanism to improve the controlling scheme of the proposed design.

I3. In case, D0 is low, the current flows from I4 to MTJ-4, MTJ-3 and sinks at I3. The write current flow directions lead to store the opposite values for MTJ-3 and MTJ-4. Similarly, the data D1 from FF1 can be written at upper pair of the MTJs. Since these two pair of MTJs have independent paths, the data can be written in parallel. The working of the store phase is also illustrated in Figure 6(a). Please note that during the store phase, transistors P3, N3, transmission gates (T1 and T2), P4 and N4 are in the off state. Hence, at a time, only one write current path is developed and a reliable store operation to NV component is performed. Furthermore, the mtj_read and $\overline{\text{mtj\_read}}$ nodes are required to be at GND potential during the write operation, so that a proper write current path is established for MTJ3 and MTJ4.

*2) Restore phase:* In the restore phase, the content of the two MTJ pairs are read, that is required during the wake-up mode. This is performed by a sensing mechanism using two back-to-back connected inverters and the output is obtained at mtj_read and $\overline{\text{mtj\_read}}$, as shown in Figure 5. Here the two MTJ pairs are read individually. The restore phase is further divided into two parts, i.e., (1) pre-charge, and (2) evaluation. During the pre-charge stage, the two output nodes are equalized at the same potential value, and the sensing operation is actually performed during the evaluation stage. Based on the resistance values of MTJs, one of the output nodes stabilizes at the logic low and the other one at the logic high level. The working sequence of the restore phase is described in Figure 6(b). As mentioned, the value from the lower pair of MTJs read first, followed by reading the upper pair. These read values are propagated to their respective active flip-flops.

In order to read the lower pair of MTJs, the two output nodes are initially pre-charged at VDD potential by enabling $PC_{VDD}$. Afterwards, the pre-charge circuit is disabled and the sensing process is activated by turning on the N3 transistor. Please note that the two transmission gates T1 and T2 are also turned on with the activation of the N3 transistor as both of them are controlled by the same input signal (i.e. $R_{en}$). That means, these two transmission gates are in the off state when the pre-charge is enabled. Additionally, during this process, the transistor P4 is required to be in the on state as well. This is to equalize the source terminals of the P1 and P2 transistors, to make sure that the resistance states of the upper MTJ pair should not affect the reading process of the lower MTJs. Whereas, the transistor N4 remains in the off state.

To read the upper pair of MTJs, the two output nodes are pre-charged at GND potential by enabling $PC_{GND}$ (see Figure 6(b) for working sequence). In this part of the operation, the sensing mechanism is activated by turning on P3 transistor and at the same time, the two transmission gates T1 and T2 are also turned on, like in the previous case. During the evaluation, based on the resistance states
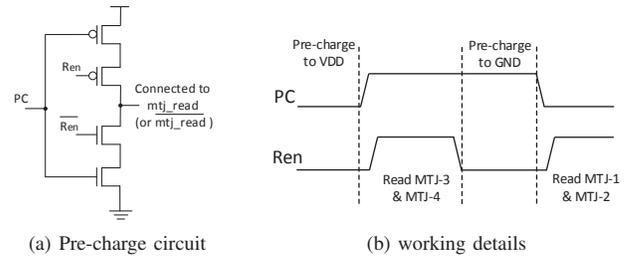
of upper pair of MTJs, the two outputs stabilize at complimentary values. Similar to previous case, here N4 is activated, so that the reading process of the upper pair of MTJs should not be disturbed by the resistance states of the lower pair of MTJs.

*D. Simplification of controlling mechanism*

The main purpose of any NV latch is to reduce the leakage using the power-gating scheme when the design is in the standby mode. During the standby mode, in our proposed design, the data back-up (NV store) operation is enabled based on the PD (Power Down) signal (as described in Figure 6). This PD signal is controlled globally for the entire logic block similar to the standard NV latch design. The wake-up signal is also activated at system-level based on the same PD signal. As discussed earlier, the pre-charge of the proposed architecture has the dependency on two signals (i.e. $PC_{GND}$ and $PC_{VDD}$), and the SEL signal is necessary for the stable outputs during the evaluation stage. The dependencies on these three signals can be optimized further to reduce it to just one signal (i.e. PC), as illustrated in Figure 7. Note that, both P4 and N4 transistors are controlled using $\overline{PC}$ signal. As shown in Figure 7(b), depending on the $Ren$ signal and based on the value of PC, the output nodes are first pre-charged to VDD and the lower pair of MTJs are read, followed by the output nodes are pre-charged to GND and upper pair of MTJs are read. Furthermore, the two output nodes can be pre-charged to GND during the write operations (which is desirable as described earlier) as $R_{en}$ would be low during that period. The only difference from the standard latch design is that in our proposed design, the MTJ read is performed sequentially for the two bits. This is due to the fact that the pre-charge operation, followed by the sense operation are performed one by one to read two bits. This sequential read has almost no delay penalty because the read evaluation is so fast that both reads can be finished within a typical cycle time duration (see Section IV-B for more details).

IV. EXPERIMENTAL SETUP AND RESULTS

In order to evaluate the efficacy of the proposed NV multi-bit latch architecture, a detailed circuit-level analysis is performed in this section. Based on that, a system-level analysis is performed and the resulting area as well as energy consumption are evaluated for various benchmark designs.

*A. Experimental setup*

The circuit-level simulations were performed using *Cadence Spectre Simulator* tool. For that, we have employed the MTJ model as proposed in [29] and for CMOS components, we have used TSMC 40 nm low-power SPICE models. The details of the design parameters of the simulations are depicted in Table I. For process corner analysis, we have considered $\pm 3\sigma$ variations for the product of *Resistance-Area* (RA), *Tunnelling Magneto Resistance* (TMR) value and switching current. Layouts were developed using *Cadence*

TABLE I.    CIRCUIT-LEVEL SETUP

| Parameter | Value |
|---|---|
| VDD and Temperature | 1.1V and 27°C |
| MTJ radius | 20 nm |
| Free/Oxide layer thickness | 1.84/1.48 nm |
| RA | 1.26 $\Omega\mu m^2$ |
| TMR @ 0V | 123% |
| Critical current | 37 uA |
| Switching current | 70 uA |
| 'AP'/'P' resistance | 11 k$\Omega$/5 k$\Omega$ |

*Virtuoso* tool. For the system-level analysis, we have used *Synopsys Design compiler* for synthesis and *Cadence Encounter* tool for the physical design activities such as floorplan, placement and routing.

### B. Circuit-level results

For circuit-level analysis, we have developed netlists for our proposed 2-bit latch schematic (Figure 5) as well as for the standard design (Figure 2(b)). For the standard latch design, the read component was designed based on the pre-charge sensing circuit [28], tristate inverters used as write component and a set of transmission gates were employed to isolate the read components during the write operation. Using SPICE simulations, the design parameters such as latency and energy consumption for both store and restore operations were extracted. The design parameters for our proposed 2-bit latch as well as for the two standard 1-bit latch design for three corner cases are shown in Table II. For fair comparison, we have considered equal number of storage bits for both designs. Therefore, we have multiplied all single bit standard latch results by a factor of two, except for the layout area (explained later). As shown in the table, our proposed design has better read active energy efficiency (around 19%) than that of the standard design. This is due to fact that we have designed the read controlling mechanism as well as the pre-charge circuitry in such a way that it leads to a fewer number of transitions.

The read latency is significantly different for our proposed and standard designs as shown in Table II. The single bit read for our proposed design is almost similar to that of the standard design. But since in our proposed circuitry, we read sequentially each bit, its read latency is approximately twice of the standard design. Nevertheless, the read can be easily completed within a clock cycle time duration as read operation is exceptionally fast, as opposed to the write operations. Additionally, it is demonstrated that the wake-up time for a STT based embedded microcontroller can be as high as 120 ns [30]. This is mainly due to the stabling period of their power signal. The total read delay of our proposed design is significantly lower than the system wake-up delay. On the other hand, both designs have employed the same writing methodology, so they have similar write energy and latency values, that is around 104 fJ and 2 ns for the worst case, respectively. As described earlier, we maintain our write mechanism reliable so that the circuit organization of our proposed design does not lead to any sneaky current flow.

Another important design metric is area, where our proposed design have shown considerable improvement over the conventional single bit NV design as the read component is shared between the 2-bits. The number of transistors (excluding write components here) used for our proposed 2-bit design is only 16 in comparison to 22 for the standard design. That means, compared to a single bit standard design, we just need to add five additional transistors to design our proposed multi-bit latch. For accurate area evaluations, we have drawn layouts for both proposed design as well as the standard design. The layout of the proposed design with 12 tracks, which uses upto M2 (metal-line 2), is demonstrated in Figure 8. For the area calculation
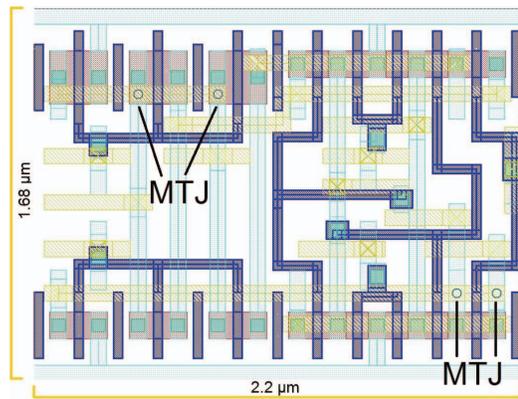


Fig. 8.    Layout of the proposed 2-bit non-volatile latch architecture

TABLE II.    COMPARISON RESULTS FOR TWO STANDARD 1-BIT LATCH AND PROPOSED 2-BIT LATCH DESIGN.

| | Two standard 1-bit Latch | | | Proposed 2-bit latch | | |
|---|---|---|---|---|---|---|
| | worst | typical | best | worst | typical | best |
| Read energy[fJ] | 6.348 | 5.650 | 4.916 | 4.799 | 4.587 | 4.327 |
| Read delay[ps] | 310 | 187 | 127 | 600 | 360 | 228 |
| Leakage[pW] | 4998 | 1565 | 424 | 4960 | 1528 | 394 |
| # of transistors | 22 | | | 16 | | |
| Area[$\mu m^2$] | 5.635 | | | 3.696 | | |

of the standard design, we have also added the minimum spacing margin in addition to the twice of the width of the actual layout block. At cell-level, the area of our proposed design is improved by around 34 % compared to the 2 bits of standard design. It is worth to mention that, in general the conventional CMOS multi-bit layout covers two rows during the placement and the similar layout can be easily generated for our proposed design as well.

### C. System-level results

For system-level evaluation, we have considered several benchmark circuits and performed *synthesis*, *floorplan* and *placement* on their RTL netlist. In this process, we have adopted mostly the default mode of option for the design constraints. Moreover, we assumed that each sequential element in the benchmark circuit require a back-up and therefore, the existing flip-flops were replaced with the shadow flip-flop architecture. After the placement, we observed the flip-flops that are very close to each other which can be replaced by our proposed 2-bit NV flip-flop design. The limit of closeness of two flip-flop is decided in such a way that there should not be any timing penalties. Therefore, we only considered the cases where the two flip-flops are apart less than the twice of the width of the NV component of the standard single-bit design (that is, $\leq 3.35 \mu$m). To illustrate such flip-flop cases, a floorplan design for s344 circuit is shown in Figure 9. The identification of such neighbor flip-flops in the layout is done using a script, that is executed over the DEF (*Design Exchange Format*) file. Note that, in this analysis, those flip-flops which can not be merged were replaced with the standard single-bit NV flip-flop.

The system-level analysis results for various benchmark circuits are shown in Table III. Here, the improvement is with respect to the case where all flip-flops are replaced with the standard single bit NV flip-flop design. As shown, both area as well as read energy improvements are increasing with the increase in the number of 2-bit NV flip-flop designs. The average area and read energy improvements for all benchmark circuits are 26% and 14%, respectively.

| Benchmark | Number of total flip-flops | Number of 2-bit NV flip-flops | Back-up using 1-bit NV latch | | Back-up using proposed 2-bit NV latch | | Improvement | |
|---|---|---|---|---|---|---|---|---|
| | | | Area [in $\mu m^2$] | Read energy [in fJ] | Area [in $\mu m^2$] | Read energy [in fJ] | Area (%) | Read energy (in %) |
| s344 | 15 | 5 | 42.255 | 42.375 | 32.565 | 37.06 | 22.93% | 12.54% |
| s838 | 32 | 12 | 90.144 | 90.4 | 66.888 | 77.644 | 25.80% | 14.11% |
| s1423 | 74 | 23 | 208.458 | 209.05 | 163.884 | 184.601 | 21.38% | 11.70% |
| s5378 | 176 | 64 | 495.792 | 497.2 | 371.76 | 429.168 | 25.02% | 13.68% |
| s13207 | 627 | 259 | 1766.259 | 1771.275 | 1264.317 | 1495.958 | 28.42% | 15.54% |
| s38584 | 1424 | 473 | 4011.408 | 4022.8 | 3094.734 | 3520.001 | 22.85% | 12.50% |
| s35932 | 1728 | 472 | 4867.776 | 4881.6 | 3953.04 | 4379.864 | 18.79% | 10.28% |
| b14 | 215 | 90 | 605.655 | 607.375 | 431.235 | 511.705 | 28.80% | 15.75% |
| b15 | 416 | 189 | 1171.872 | 1175.2 | 805.59 | 974.293 | 31.26% | 17.10% |
| b17 | 1317 | 542 | 3709.989 | 3720.525 | 2659.593 | 3144.379 | 28.31% | 15.49% |
| b18 | 3020 | 1260 | 8507.34 | 8531.5 | 6065.46 | 7192.12 | 28.70% | 15.70% |
| b19 | 6042 | 2530 | 17020.314 | 17068.65 | 12117.174 | 14379.26 | 28.81% | 15.76% |
| or1200 | 2887 | 1269 | 8132.679 | 8155.775 | 5673.357 | 6806.828 | 30.24% | 16.54% |
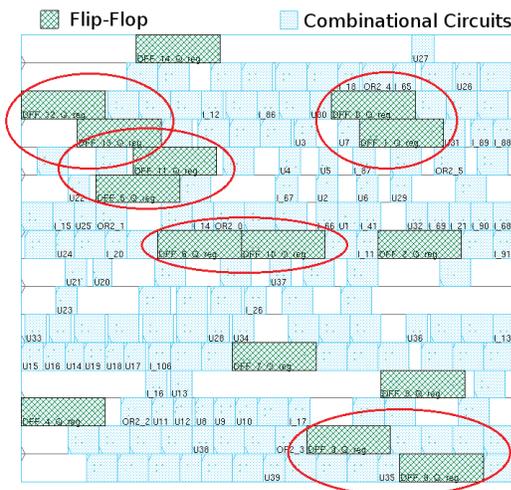


Fig. 9. Demonstration of the floorplan of standard cells for s344 design. Flip-flops that can be merged are encircled.

## V. Conclusions

*Spin Transfer Torque* based non-volatile flip-flop designs are promising for the static power reduction, that is mainly due to their normally-off/instant-on computing features. In this paper, we exploited the fact that many flip-flops are generally placed adjacent to each other in the layout, and shared their shadow components. We proposed a multi-bit non-volatile flip-flop architecture to reduce the area and energy costs associated with the non-volatile components using resource sharing concept. We have developed the layout of the proposed design and modified the physical design flow to replace the adjacent single bit storing cells with the proposed multi-bit cells. Results show that, compared to the single bit non-volatile flip-flop design, our proposed multi-bit architecture improves the system-level area and energy by 26 % and 14 % in average, respectively.

## VI. Acknowledgement

## References

[1] K.-S. Yeo and K. Roy. *Low Voltage, Low Power VLSI Subsystems*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 2005.

[2] N. Kim, et al. Leakage current: Moore's law meets static power. *Computer*, 36(12):68–75, Dec 2003.

[3] C. Singh and R. Tangirala. As nodes advance, so must power analysis. Available: http://semiengineering.com/as-nodes-advance-so-must-power-analysis/, 2014.

[4] M. Padhye and D. Gross. Freescale: Wireless Low-Power Design and Verification with CPF. Available: https://www.si2.org/?page=1061.

[5] M.-T. Chang, et al. Technology comparison for large last-level caches (L-3 Cs): Low-leakage SRAM, low write-energy STT-RAM, and refresh-optimized eDRAM. In *HPCA*, pages 143–154, 2013.

[6] M. Gajek, et al. Spin torque switching of 20 nm magnetic tunnel junctions with perpendicular anisotropy. *Applied Physics Letters*, 100(13):132408, 2012.

[7] International Technology Roadmap for Semiconductors. http://www.itrs.net, 2015.

[8] S. A. Wolf, et al. The promise of nanomagnetics and spintronics for future logic and universal memory. *Proceedings of the IEEE*, 98(12):2155–2168, 2010.

[9] A. D. Kent and D. C. Worledge. A new spin on magnetic memories. *Nature nanotechnology*, 10(3):187–191, 2015.

[10] R. Bishnoi, et al. Self-timed read and write operations in STT-MRAM. *TVLSI*, 24(5):1783–1793, 2016.

[11] R. Bishnoi, et al. Improving write performance for STT-MRAM. *TMAG*, 52(8):1–11, 2016.

[12] K. Ryu, et al. A magnetic tunnel junction based zero standby leakage current retention flip-flop. *TVLSI*, 20(11):2044–2053, 2012.

[13] S. Yamamoto, et al. Nonvolatile flip-flop using pseudo-spin-transistor architecture and its power-gating applications. In *ISCDG*, pages 17–20, 2012.

[14] Y. Lakys, et al. Low power, high reliability magnetic flip-flop. *EL*, 46(22):1493–1494, 2010.

[15] S. Yamamoto, et al. Nonvolatile delay flip-flop using spin-transistor architecture with spin transfer torque MTJs for power-gating systems. *EL*, 2011.

[16] R. Bishnoi, et al. Design of defect and fault-tolerant nonvolatile spintronic flip-flops. *TVLSI*, 25(4):1421–1432, 2017.

[17] D. Chabi, et al. Ultra low power magnetic flip-flop based on checkpointing/power gating and self-enable mechanisms. *TCS-I*, 61(6):1755–1765, 2014.

[18] Y. T. Shyu, et al. Effective and efficient approach for power reduction by using multi-bit flip-flops. *TVLSI*, 21(4):624–635, 2013.

[19] G. P. Singh, et al. High speed multiple-bit flip-flop, July 16 2002. US Patent 6,420,903.

[20] K. Gourav, et al. Using multi-bit flip-flop custom cells to achieve better SoC design efficiency. http://www.embedded.com/design/mcus-processors-and-socs/4433619/Using-multi-bit-flip-flop-custom-cells-to-achieve-better-SoC-design-efficiency, 2014.

[21] N. Sakimura, et al. Nonvolatile magnetic flip-flop for standby-power-free socs. *IEEE Journal of Solid-State Circuits*, 44(8):2244–2250, 2009.

[22] W. Zhao, et al. Spin-mtj based non-volatile flip-flop. In *NANO*, 2007.

[23] K. Jabeur, et al. Spin orbit torque non-volatile flip-flop for high speed and low energy applications. *EDL*, 2014.

[24] R. Bishnoi, et al. Non-volatile non-shadow flip-flop using spin orbit torque for efficient normally-off computing. In *ASP-DAC*, pages 769–774, 2016.

[25] I. Kazi, et al. A ReRAM-based non-volatile flip-flop with sub-V T read and CMOS voltage-compatible write. In *NEWCAS*, pages 1–4, 2013.

[26] J.-M. Choi, et al. PCRAM flip-flop circuits with sequential sleep-in control scheme and selective write latch. *JSTS*, 13(1):58–64, 2013.

[27] S. Khanna, et al. An FRAM-Based Nonvolatile Logic MCU SoC Exhibiting 100State Retention at VDD= 0 V Achieving Zero Leakage With < 400-ns Wakeup Time for ULP Applications. *Journal of Solid-State Circuits*, 49(1):95–106, 2014.

[28] W. Zhao, et al. Design considerations and strategies for high-reliable STT-MRAM. *Microelectronics Reliability*, 51(9):1454–1458, 2011.

[29] A. Mejdoubi, et al. A compact model of precessional spin-transfer switching for MTJ with a perpendicular polarizer. In *MIEL*, 2012.

[30] N. Sakimura, et al. A 90nm 20mhz fully nonvolatile microcontroller for standby-power-critical applications. In *ISSCC*, pages 184–185, 2014.