Planar fully depleted silicon technology to design competitive SOC at 28nm and beyond

ABSTRACT

This document considers the challenges to obtain competitive silicon technology for the upcoming generation of System-On-Chip ICs. It suggests planar fully depleted technology deserves serious interest. After outlining some implementation choices, a number of circuit-level benchmark results as well as some important design aspects are presented. It is found that this technology combines high performance, power efficiency and cost-effectiveness, which makes it a very attractive candidate to serve the needs of mobile and consumer multimedia SOCs starting at the 28nm node and scalable down to 14nm.

February, 2012

Philippe Flatresse,
Circuit Design and Architecture, STMicroelectronics
Giorgio Cesana,
Technology Marketing, STMicroelectronics
with Xavier Cauchy,
Digital Applications, Soitec

CONTENTS

1. Introduction

2. Motivations

3. Brief Overview of STMicroelectronics’s 28nm planar FD technology

4. Circuit-level benchmarking

5. Design Considerations

6. Perspectives

7. Conclusion

1. Introduction

Dr Martin Cooper, former Motorola VP, placed the first private phone call from a handheld mobile phone in 1973. The DynaTAC weighed 2 pounds (close to 1kg), offered 30 minutes conversation and was costing 3995 USD.
Challenges remain more or less the same today: power, performance and cost.
Except your phone must now playback or encode HD or 3D video, handle 3G+ and 4G data communication, Internet access and all sorts of connectivity, include GPS and camera functionalities, with augmented reality around the corner. Occasionally it is used to place phone calls. All this must fit in a 120g box of affordable price.

To always integrate more complexity and meet the Giga Operation per Second, the semi-conductor industry shrank transistor dimensions dramatically. But nowadays, going to the 22/20nm node – and even already at 28nm – conventional transistors become unable to offer optimal performance without draining your battery or raising the temperature of your smartphone beyond safe limits. To work around this issue and improve performance and power efficiency, multi-gate transistor approaches have been envisioned for a decade. MugFET, TriGate, FinFET, Gate-All-Around, have been proposed to improve transistor behavior and hopefully provide the consumers with great user experience without burning their fingers. However, those novel structures are so complex that they could not be adopted prior to the 22nm node at the earliest, and may not be the best fit for consumer-type System-on-Chip applications.
In this context, STMicroelectronics is implementing a technology that solves those multiple issues at minimum efforts. Planar fully depleted silicon technology will be ready as early as 2012 to compete in the forthcoming 3GHz superphones era and in many other consumer segments.

2. Motivations

2.1. Transistor Technology Requirements

When switching from one CMOS technology generation (or “node”) to the next, chip designers expect reduced transistor dimensions for better circuit density, reduced dynamic and static power per transistor to keep total chip power under control, and improved switching speed for higher performance – all this at a cost per finished die that makes economic sense.
Moreover, with the advent of Internet-connected smart devices offering sophisticated multimedia and running “apps” of all kinds, advanced mobile and consumer applications are engaged in a race which is calling for an exponential growth of the performance of embedded CPU and other IP cores.

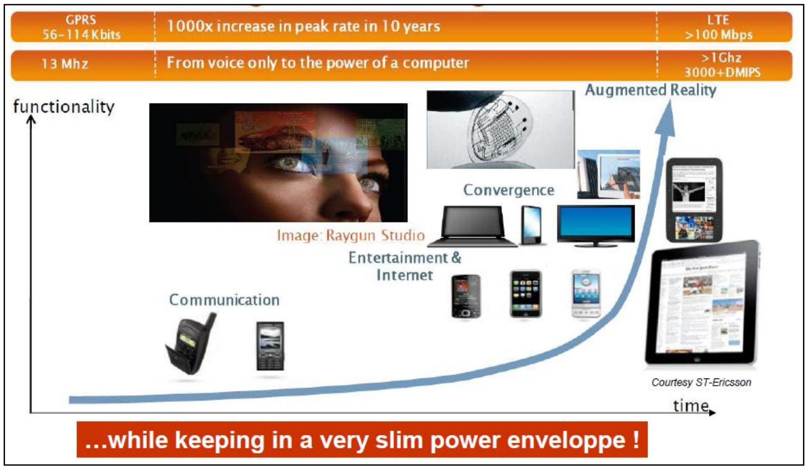


Fig. 1: Exponential increase of SOC performance requirements at constant power consumption

Therefore low-power technology is less than ever low-performance, and there is a convergence between low-power and performance-oriented CMOS technologies. The ideal technology would combine high peak performance, low active power across all use cases – in particular by retaining good performance at reduced power supply (Vdd), low stand-by power and low cost.

2.2. Traditional CMOS technology does not scale well beyond 28nm

Traditional planar CMOS on bulk silicon (or “bulk CMOS”) is facing great challenges to keep up with these simultaneous requirements at the 28nm node, and things will only get worse. The two major detractors to the efficiency of traditional technology at these advanced nodes are transistor variability and electrostatics.

Variability

Transistor variability essentially refers to the fact that each transistor in a chip will exhibit an electrical behavior different from its nominal behavior, with random differences of threshold voltage (VT), ‘on’ current and ‘off’ current. The root cause of variability in bulk CMOS technology is mainly RDF, random dopant fluctuations, i.e. variations in the exact number and position of dopant atoms in the channel of the transistor. At advanced nodes, sensitivity to this effect is high because, as the dimensions shrink, there actually remain few atoms in the channel and statistical averaging is limited.

A wide distribution (in the statistical sense) of VT will affect the predictability of transistor behavior, degrade stability and minimum operating voltage Vmin of SRAM bit cells and logic (thereby limiting the opportunities for dynamic power savings), increase total current leakage at chip level, affect performance and make it more difficult to sign-off designs in corner conditions.

Electrostatics

An ideal CMOS logic transistor behaves as a switch, where the formation of a conduction channel from source to drain is under complete electrostatic control of the gate. In the real world, however, this is not true: the source and the drain also affect to some extent what happens in the channel, and when the channel is very short (as is now the case for advanced nodes, which have shrunk transistors dimensions considerably), these side-effects, sometimes collectively called “short channel effects” (SCE), become extremely pronounced and deleterious.

Two important metrics of the ideality of a transistor are the sub-threshold slope (SS) and the drain-induced barrier lowering (DIBL). The interested reader is referred to e.g. *[1]* for details. Suffice to say here that, at advanced nodes, these parameters reach very poor values and seriously degrade the performance/leakage trade-off of circuits. Obviously, this is a major concern for mobile and consumer multimedia applications, for which lowering Vdd is an important lever to save dynamic power, and cannot afford extensively using extremely low VTs to reach high performance – as other applications such as high-performance computing may afford.

2.3. The Solution: Fully Depleted Transistor Architectures

There is consensus in the industry that the way to go in the near future is fully depleted transistors. A fully depleted transistor can be planar or tri-dimensional. In the planar flavor, ultra-thin body transistors are fabricated in an ultra-thin layer of silicon sitting over a buried oxide (Fig.2.a). This is done by employing SOI (Silicon-on-Insulator) wafers as starting substrate, in a flavor where the top silicon is extremely thin. In the tri-dimensional flavor, transistors are FinFET or TriGate devices (Fig.2.b): the gate wraps around the sides of a vertical silicon ‘fin’. Ultra-thin body, fully depleted transistors solve the issues faced by conventional bulk CMOS because different physical effects govern their behavior. In particular, owing to the geometry of the body, the gate retains excellent electrostatic control over the channel, dramatically cutting short channel effects and improving sub-threshold slope and DIBL.

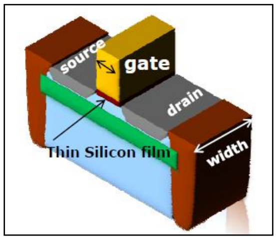


Fig. 2.a (left): 2D (planar) fully depleted transistor

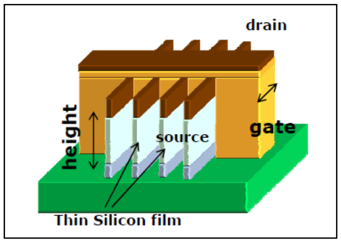


Fig. 2.b (right): FinFET (i.e. 3D fully depleted) transistor

2.4. Motivations for Planar FD

Having identified that conventional planar bulk CMOS would not meet all the requirements of mobile and consumer multimedia System-on-Chip (SOC) ICs in the coming years, STMicroelectronics assessed alternative options. Planar fully depleted CMOS technology was identified as a very effective solution as it combines a low-disruption planar approach with solving the electrostatics and variability issues that plague planar bulk CMOS. Being a comparatively simple evolution from conventional CMOS, it can be developed at reasonable cost and effort. It is therefore possible to propose a 28nm planar FD solution available as second generation shortly after readiness of traditional 28nm on bulk silicon, with better time-to-market than waiting for availability of the 20nm node. Offering a very worthwhile operating frequency boost for embedded CPU, GPU and the likes, plus very low standby and active power at chip level, 28nm planar FD represents a real opportunity for some high-volume applications.
It is also an excellent learning step to prepare a 20nm planar FD process, to cope with the shortcomings of traditional 20nm bulk CMOS, expected to be very serious for some application needs. 20nm planar FD will bridge the gap with an uncertain 14nm node. Keeping a planar architecture, rather than a FinFET-type approach, which would mean a completely new and complex silicon process as well as significant design disruption, it makes possible to offer a 20nm solution in a reasonable timeframe. Our evaluations show that 20nm planar FD has also a very competitive potential performance-wise vs. FinFET for System-on-Chip applications.

3. Brief Overview of STMicroelectronics's 28nm planar FD Technology

For a more in-depth review of planar FD technology, the reader can for example refer to some exhaustive tutorials by T. Skotnicki [2,3] or to technical literature such as [4].

Planar fully depleted silicon-on-insulator ("planar FD", or equivalently "FDSOI") transistors are planar CMOS transistors fabricated in a very thin layer of silicon sitting over a buried oxide (BOX). They are therefore 'ultra-thin body' (UTB) devices: the electrical conduction channel that forms between source and drain is confined into the ultra-thin silicon layer under the gate oxide.

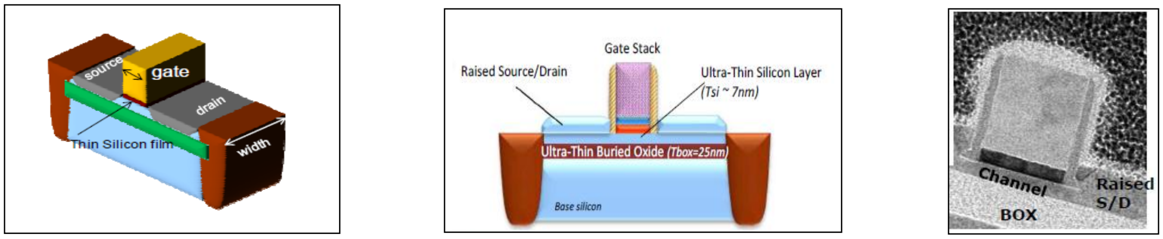


Fig. 3: ST's planar FD device structure (notional perspective, notional cross-section, TEM cross-section)

Immunity to Short Channel Effects and Variability
Having a very thin body ensures all electrical paths between source and drain are very close to the gate, and the latter therefore regains excellent electrostatic control over the channel. As a result, sub-threshold slope, DIBL and other short channel effects exhibit excellent values (usage of a very thin buried oxide also helps, see below).
In addition, the planar FD technology does not demand doping or pocket implants in the channel to control the electrostatic characteristics and tune the threshold voltages. Therefore, the major issue of random dopant fluctuation mostly disappears. The absence of doping also helps reaching good performance since high channel doping induces reduced carrier mobility.

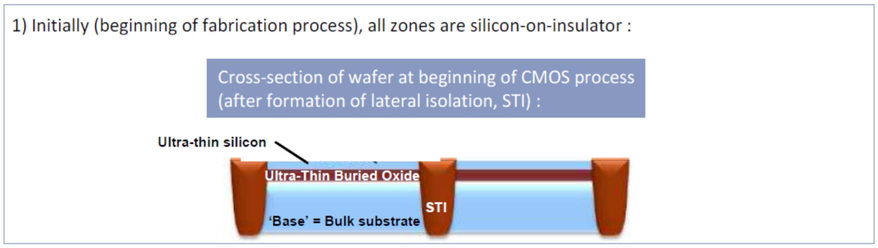
Ultra-Thin Buried Oxide
STMicroelectronics has chosen to implement planar FD on an ultra-thin buried oxide, also called 'Ultra-Thin Body and BOX', UTBB. For the 28nm node, the selected BOX thickness is 25nm. Using an ultra-thin BOX brings several advantages:

- further improves the electrostatic control and relaxes the thinness requirement of the top silicon,
- enables back-biasing through the BOX, as described further down this document,
- enables the implantation, during the fabrication process, of heavily doped "ground planes" or "back-planes" under the BOX, for improved electrostatics and/or VT adjustment and/or best-efficiency of back-bias,

- brings the ability ,during the fabrication process, to locally remove the top silicon and BOX to reach the base bulk silicon and co-integrate a few (non geometry-critical) devices on Bulk with devices on SOI – with a small step height between an SOI zone and a Bulk zone, compatible with lithography tools (Fig.4).

In addition, the BOX offers total dielectric isolation of the very thin active layer and naturally ultra-shallow junctions, leading to lower source/drain capacitance, lower leakage and latch-up immunity.

1) Initially (beginning of fabrication process), all zones are silicon-on-insulator :


Cross-section of wafer at beginning of CMOS process (after formation of lateral isolation, STI) :
Ultra-thin silicon
Ultra-Thin Buried Oxide
STI
'Base' = Bulk substrate

2) Selected zones can be turned into bulk silicon zones by locally etching off the top silicon and BOX :

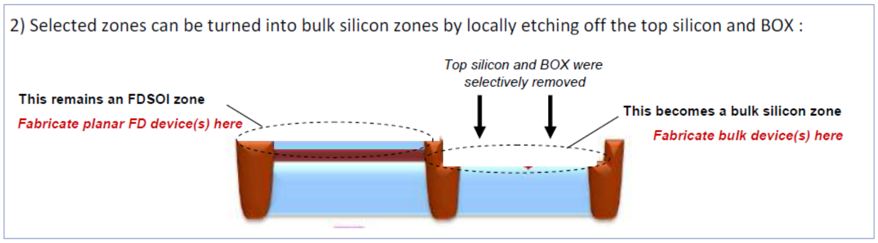

Top silicon and BOX were selectively removed
This remains an FDSOI zone
Fabricate planar FD device(s) here
This becomes a bulk silicon zone
Fabricate bulk device(s) here

Fig. 4: Formation of non-SOI zones for co-integration of bulk and planar FD devices

Multi-VT

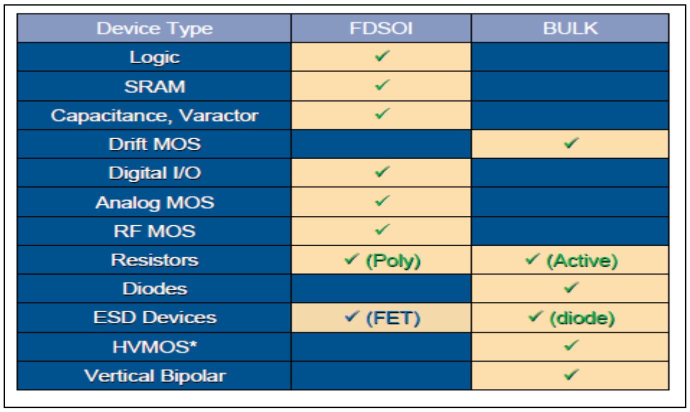
While Multi-VT in planar bulk CMOS technology is achieved by multiple channel implants, this approach is normally not used in planar FD, which is essentially an undoped channel technology – enabling the achievement of record-low variability.

Regardless this limitation, planar FD techology allows several methods for setting the threshold voltage VT, including engineering the gate stack work function, trimming the gate length, counter-doping, and other process engineering techniques.

Thanks to this, STMicroelectronics' 28FDSOI technology is capable of offering 3 VTs (HVT, RVT, LVT), as in traditional bulk CMOS technologies.

Device Menu

We have listed all the devices that can be required in SOCs. A small number of specific devices are more easily implemented through the hybrid Bulk-FDSOI co-integration approach outlined earlier. The table below sums up our choices to offer a full device menu suitable for complex SOC integration, taking care not to increase the overall process cost.



Device Type	FDSOI	BULK
Logic	✓	
SRAM	✓	
Capacitance, Varactor	✓	
Drift MOS		✓
Digital I/O	✓	
Analog MOS	✓	
RF MOS	✓	
Resistors	✓ (Poly)	✓ (Active)
Diodes		✓
ESD Devices	✓ (FET)	✓ (diode)
HVMOS*		✓
Vertical Bipolar		✓

Table 1: Planar FD (FDSOI) device menu

Commonalities with traditional 28nm Low-Power CMOS technology

STMicroelectronics's strategy when developing the 28nm planar FD technology has been to reuse as much as possible the 28nm low-power bulk CMOS process. Figure 5 outlines the process flow; process steps specific to planar FD have been highlighted in blue. The "Back-End of Line", BEOL, part of the process (from contacts up to top metal levels) is a direct copy of the 28nm bulk technology. The "Front-End of Line", FEOL, part of the process (essentially, all transistor fabrication steps) also relies in majority on a direct re-use of equivalent process modules from the bulk technology. Only a few steps have been optimized, added or removed – noting that the excellent performance to meet the demands of high-end mobile applications is reached without having to add complex stressors to the baseline 28nm low-power process approach.

- Lateral isolation (STI)
- N-Well/P-Well/DN-Well implant
- **BP-N/BP-P(back-plane) implant**
- **Local Buried Oxide (BOX) removal**
- High-K
- Metal Gate
- SPACER
- **EPI (raised S/D)**
- LDD implantation
- S/D implantation
- NiPtSi Silicidation
- Standard Back-End of Line

Fig. 5: 28nm planar FD Outline Flow with specificities highlighted

Overall, the **Back-End is 100% identical** to the traditional 28nm bulk low-power CMOS process, and the **Front-End of Line (FEOL) is 80% common** with that same process.

Cost considerations

In addition to sharing many steps with the conventional 28nm Low-Power process and sharing the same fabrication lines, the planar FD process **saves about 10% of the steps** required to fabricate the chips on the wafers. This approximately offsets the cost overhead of the starting wafers. As a result, the 28nm planar FD technology **matches the cost of a conventional low-power technology** while, as will be shown, delivering extremely competitive performance vs. a more complex "G-type" technology (see benchmarking section further down).

4. Circuit-Level Benchmarking

Methodology

To assess how the improved transistor characteristics translate at circuit level, ST has benchmarked a number of representative IP blocks, including an ARM Cortex-A9 CPU core. To that aim, we have extracted logic critical paths with associated RC parasitics from placed-and-routed designs and have re-characterized them by swapping 28nm traditional bulk CMOS transistor SPICE models with 28nm planar FD SPICE models.

Silicon Correlation

The models were first calibrated versus TCAD and extrapolations from early silicon. As the process development advanced, we obtained more silicon test structures and were able to refine the model of each transistor instance – or "model card" in SPICE terminology. With test chips in our 28nm planar FD technology becoming available, we are demonstrating that the models predict well the silicon behavior. We are therefore confident that the benchmarks presented below are reliable and will be matched by SOC implementations.

Results

The following benchmarks will compare the merits at the 28nm node of ST's planar FD technology ("28FD") with a state-of-the-art Low-Power technology ("28LP") and a more performance-oriented, state-of-art General Purpose technology ("28G"). They are all based on evaluation of an ARM Cortex-A9 core. The analysis will mostly focus on the higher end of the range of operating frequencies found in a SOC, since modern mobile and consumer multimedia demand high performance from their master CPU (for example, a Cortex-A9 or the forthcoming A15).
Some of the results presented take into account back-bias. This power management technique is extremely efficient with planar FD on UTBB and will be discussed in some details in the next chapter "Design Considerations". In short, it consists of shifting up or down from its nominal value the voltage applied to the substrate under the BOX of target transistors. This

results in a VT shift. It is therefore possible, on a use case basis, temporally boosting performance (at the expense of increased leakage current) by “forward” back-bias, or cut leakage (at the expense of reduced performance) by “reverse” back-bias.

Note- VT flavors in the following are denominated: HVT for high-VT, RVT or SVT for regular-VT or standard-VT, LVT for low-VT and uLVT for ultra-low-VT.

4.1. Performance at nominal Vdd : best speed/leakage trade-off

At nominal voltage (0.85V for the G-type technology, 0.9V for 28FD and 1V for the LP-type technology), for comparable leakage power, 28FD consistently outperforms both 28LP and 28G. In addition, applying forward back-bias (FBB) to the planar FD technology enables further pushing the performance, obviously at the expense of increased leakage, but without degrading the performance/leakage ratio – making this a possible solution for bursts of activity at ultra-high speed without calling for ultra-low VT.

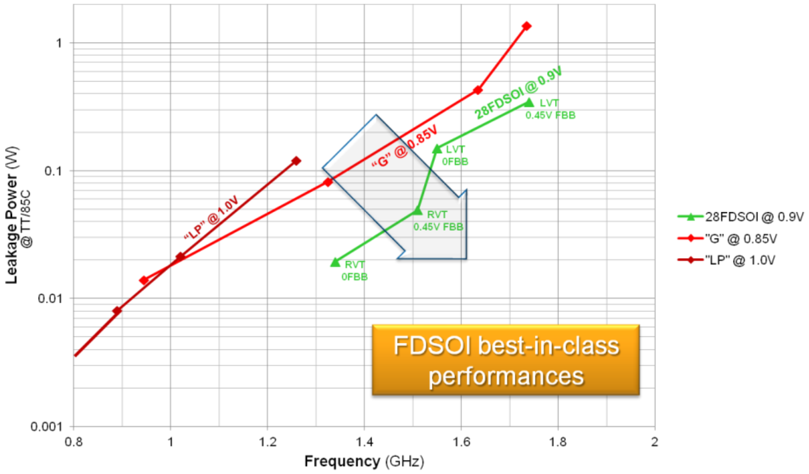


Fig. 6: Best operating frequency for any class of leakage (TT process, 85C)

4.2. Competitive speed/leakage trade-off enables operation at reduced Vdd

Figure 7 shows that operating 28FD with 100mV underdrive (Vdd=0.8V instead of 0.9V nominal) still provides remarkable performance. The performance/leakage ratio is comparable to what conventional bulk CMOS technologies achieve at nominal Vdd (no underdrive); in addition, with adequate FBB it is still possible to reach 2GHz operation in typical process conditions. Of course reducing Vdd is a very good way to save dynamic power. It is therefore realistic to envisage building 28FD chips that match 28G or 28LP performance at a fraction of the power consumption.

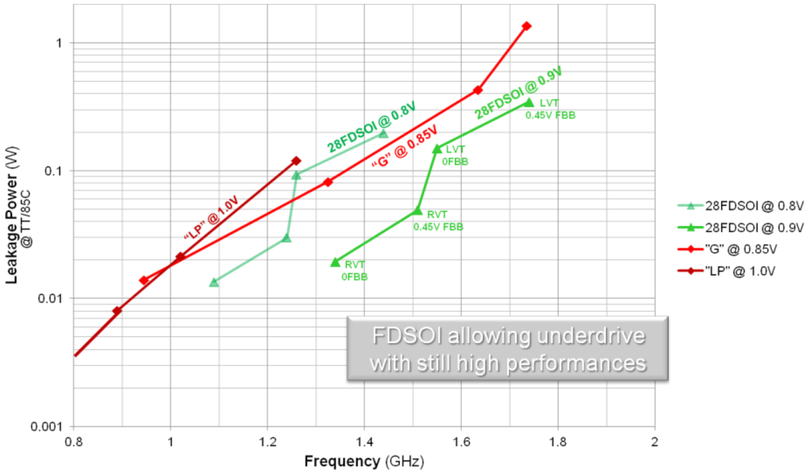


Fig.7: Remarkable performance in underdrive mode (TT process, 85C)

4.3. Leading-edge performance across the full Vdd range

Fig. 8 compares performance results vs. Vdd of the different technology flavors under worst-case process corner ("slow-slow") and worst-case temperature conditions, based on the evaluation of a particular ARM Cortex-A9 CPU core implementation. Different VT options are considered and comparison here is no longer at matched leakage.

Worst-case process and temperature conditions, cross-Vdd analysis

It appears that only the LVT and uLVT flavors of the 28G technology can compare, performance-wise, with the LVT flavor of 28FD. Note however that both 28G-LVT and 28G-uLVT are considerably leakier than 28FD-LVT, as seen in Fig.9. In addition, 28FD has two additional weapons to further push performance: its effective FBB capability and its larger Vdd overdrive capability – overdriving a "G"-type technology is typically limited, in practical applications, by time-dependent dielectric breakdown (TDDB): repeatedly overdriving the gate voltage by a significant extent would permanently alter the gate dielectric and cause premature aging of the chip and serious reliability issues.

Performance at worst-case Vdd

One interesting perspective is to look at the sign-off performance in worst case conditions, including Vdd: to account for all sources of voltage drops, the **worst-case Vdd** is typically taken as 0.1V below the nominal Vdd value. The exercise can be done with and without Vdd overdrive.

Without using overdrive, at worst case supply (Vnom-10%) 28FD technology allows similar performances as 28G technology, with similar dynamic power consumption but lower leakage figures (exact values varies depending the 28G Vt choice and 28FD FBB usage, for reference see fig. 11).

About overdrive, we estimate 28G processes having a small room for increasing the power supply (exact values depends on product mission profile) whereas 28FD technology allows for much larger room, enabled by the ticker (LP) oxide used. This translates in far higher top speed performances capabilities (ref. Fig. 8).

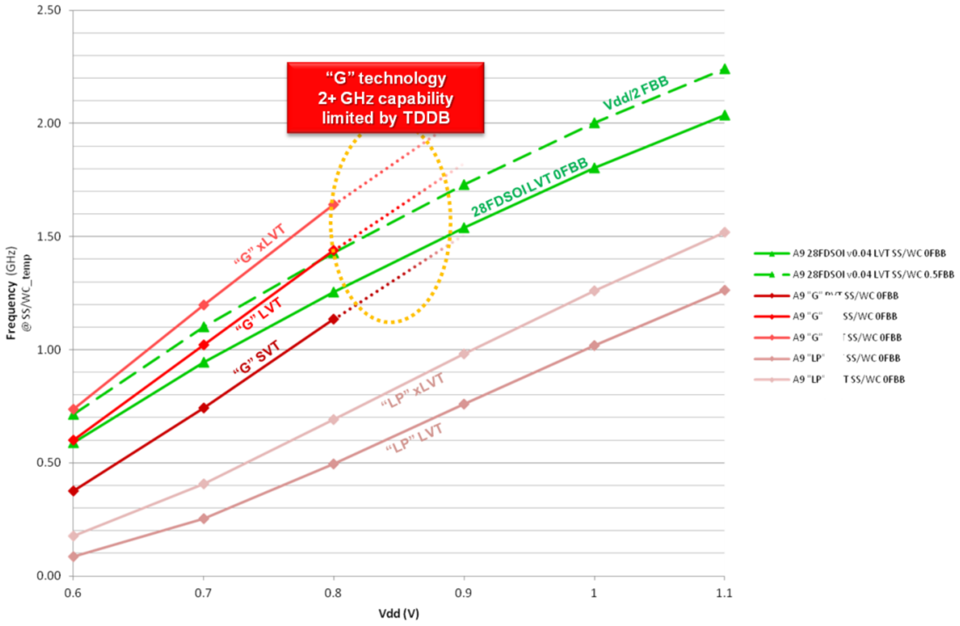


Fig.8: Compared Performance of G, LP and planar FD technologies (SS process, WC temp.)

Planar FD in very-low Vdd mode: Vdd=0.6V / 0.7V

As expected from a fully depleted technology, the relative performance loss of 28FD when decreasing Vdd is much lower than with the bulk CMOS technologies. For example, at 1.0V, a 28FD LVT critical path runs +45% faster than the same critical path in 28LP uLVT, but at 0.7V it is +130% faster.

One additional observation is that 28FD saves about 200mV on Vdd for same performance as 28LP, at comparable leakage. Therefore, a chip or physical block not requiring leading-edge performances will save up to 50% power when implemented in 28FD vs. a 28LP implementation.

4.4. Best Power Efficiency Across Use Cases

Besides getting the best possible performance for a selected class of leakage, it is important to have access to the best possible total power consumption (dynamic power plus leakage power) across a wide range of operating frequencies.
Looking at Fig. 9 figures, we can observe that:
The 28G technology can reach very high target frequencies in cores when using extra-low VT devices and high Vdd, but ultra-low Vt is very power-inefficient when operated at lower Vdd; at the same time, SVT allows better power efficiency but far reduced speed range
The 28LP technology is penalized by high dynamic power consumption (Vdd is higher) which negatively affects total power figures regardless the very poor leakage figures;
In contrast, **the 28FD technology** is **power-efficient across the full Vdd and target frequency range**. A logic path based on, for example, low-VT and operated at maximum Vdd (with or without FBB) to reach high target frequencies has comparable or better power efficiency as the best 28G alternative in this zone of operation; and the same logic path when operated at lower Vdd (typically, when the use case requires a lower target frequency) still has better efficiency than the most power-efficient alternative in this other zone of operation.

One consequence is that a low-power technique based on voltage scaling, such as DVFS (where operating frequency is reduced in use cases of limited workload and supply voltage is scaled down accordingly), will bring more benefits to a design based on the FD technology than on the "G" technology. A logic path made of a given mix of VT cells will operate at a reduced target operating frequency at lower total power.

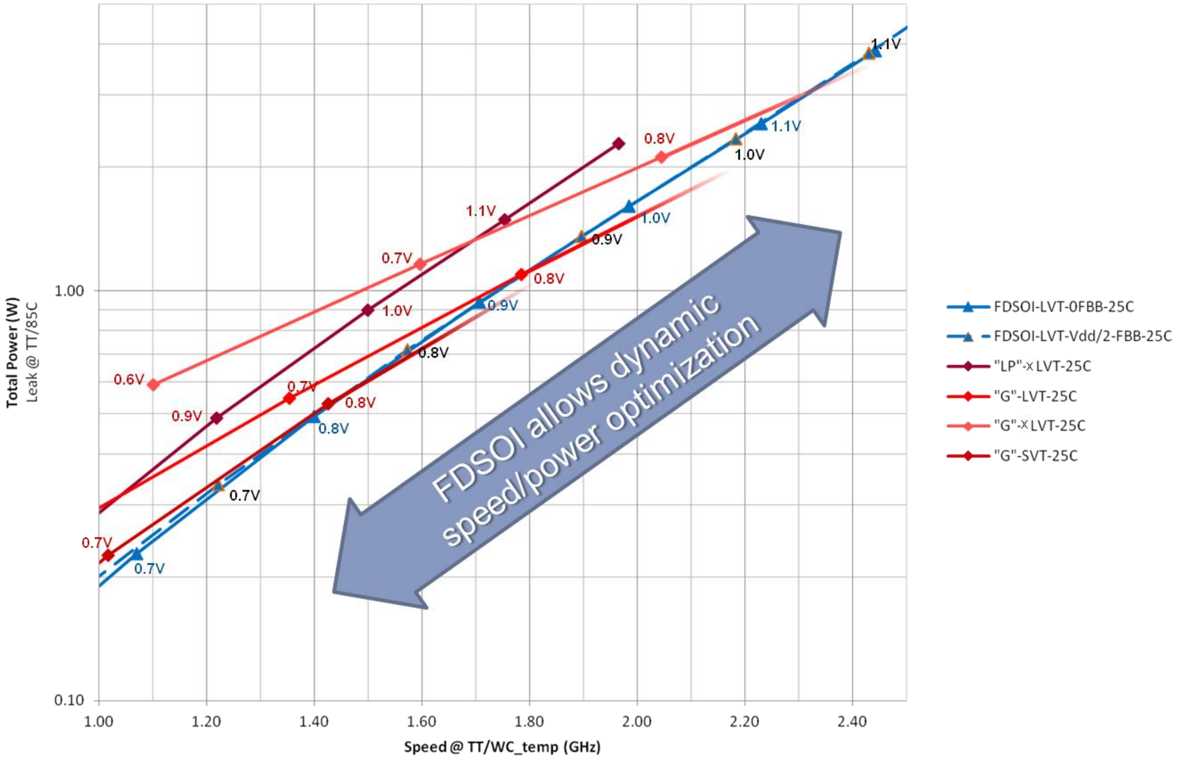


Fig.9: Power efficiency across all use cases (TT process, WC temp)

4.5. Focus on SRAM

All data presented here is for 'Typical-Typical' process corner, at 25deg.C, nominal bitcell Vdd (1V for 28LP, 0.85V for 28G, 0.9V for 28FD), unless specified otherwise.

Leakage and performance at nominal Vdd

The bitcells proposed in 28FD technology have very competitive cell current (Icell) vs. standby current ratio, which is representative of the performance/leakage power trade-off for SRAM arrays. This is true for all bit cells flavors: high-density and low-leakage oriented, or high-speed oriented. The footprint of the 4 bitcells proposed in 28FD is the same as that of the 4 bitcells proposed in 28LP.

Specifically,

> Compared to 28G:

- 28G is mostly out of the game as far as bitcell efficiency is concerned, with leakage over an order of magnitude worse than 28FD for the same Icell, and limited performance at any practical leakage value.

> Compared to 28LP:

Taking a performance-centric view:

- A given performance (specifically, cell current Icell) target is obtained from 28FD at significantly better leakage, and the highest-speed bit cell of 28FD is significantly faster than the highest-speed of 28LP, with still lower leakage;

and, taking a leakage-centric view:

- A given leakage target obtained with 28FD delivers a significantly better current Icell, and the lowest-leakage bitcell of 28FD is better than that of 28LP with still significantly better performance.

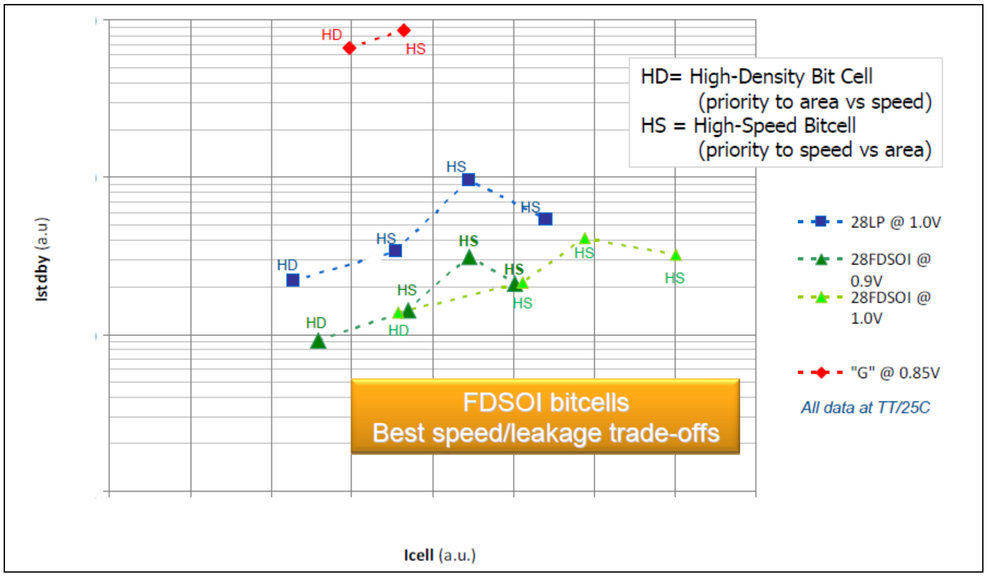


Fig.10: SRAM memory bit cells performance/leakage

Leakage and performance at reduced Vdd

Fig.10 shows that the power supply of 28FD SRAM arrays can be lowered by 100mV from nominal and still match the performance of 28LP SRAM arrays operated at nominal Vdd, while offering a 2x to 5x reduction in leakage power.

Vmin

We have verified on silicon that Vmin – the lowest supply voltage for which all SRAM arrays are still 100% functional – is also improved by over 100mV with 28FD.

5. Design Considerations

Designing on planar FD requires specific extraction deck and SPICE models. Apart from that, the design flows, methodologies and tools do not need any adaptation that would be specific to planar FD.

5.1. SPICE Models

The equations describing the electrical behavior of fully depleted transistors are different from those used for conventional bulk CMOS. Therefore, specific SPICE compact models have been developed for accurately representing planar FD transistors. At ST, for 28nm we have selected a model developed by LETI, called “UTSOI Spice Model”. Alternative models for planar FD transistors are also available or under development, for example BSIM-IMG from UC Berkeley. The model we use is now integrated in all major commercially available simulators, such as Mentor’s ELDO, Synopsys’s HSPICE and XA or Cadence’s SPECTRE. A model card has been extracted for transistor or other device available in our 28nm planar FD technology.

5.2. Flow and Design Platform

With adequate SPICE models integrated in the PDK, the design flow is identical to that used with conventional 28nm Bulk CMOS technology. We have developed a full design platform for SOC, re-using work done for 28nm Bulk. It consists of standard cell libraries (multi-channel and multi-VT) with power management elements (power switches, level shifters etc.), embedded memories, analog foundation IP (such as PLLs and the likes) and specialty IP (Antifuse etc.).

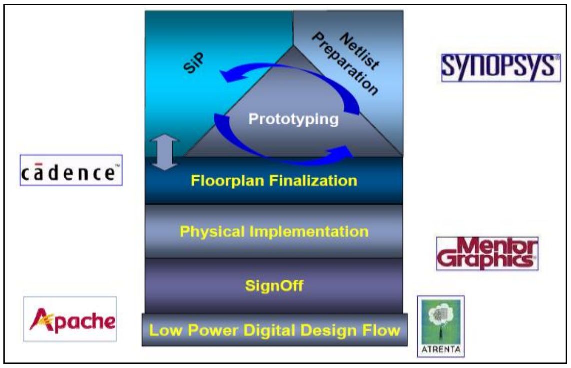
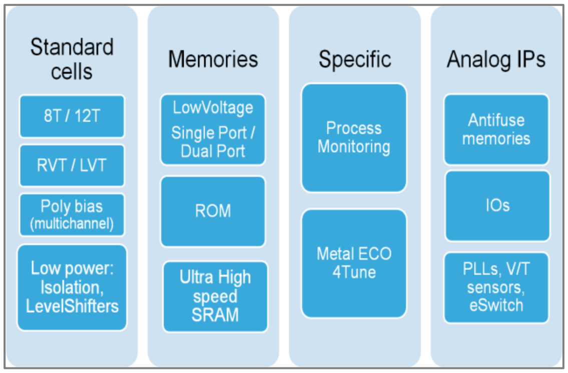

SiP
Netlist Preparation
Prototyping
Floorplan Finalization
Physical Implementation
SignOff
Low Power Digital Design Flow
SYNOPSYS
cadence
Mentor Graphics
Apache
ATRENTA

Fig.11: ST’s SOC implementation flow outline



Standard cells	Memories	Specific	Analog IPs
8T / 12T	LowVoltage Single Port / Dual Port	Process Monitoring	Antifuse memories
RVT / LVT	ROM	Metal ECO 4Tune	IOs
Poly bias (multichannel)	Ultra High speed SRAM		PLLs, V/T sensors, eSwitch
Low power: Isolation, LevelShifters			

Fig.12: 28nm planar FD design platform

5.3. Migration from Bulk (IP & SOC) at moderate effort

A design platform developed for bulk CMOS technology can be ported to planar FD by re-characterization using planar FD SPICE models, which we have done for a variety of back-biasing conditions. Only a limited number of critical IPs need to be tuned or redesigned: Analog IP, IOs, Fuse.

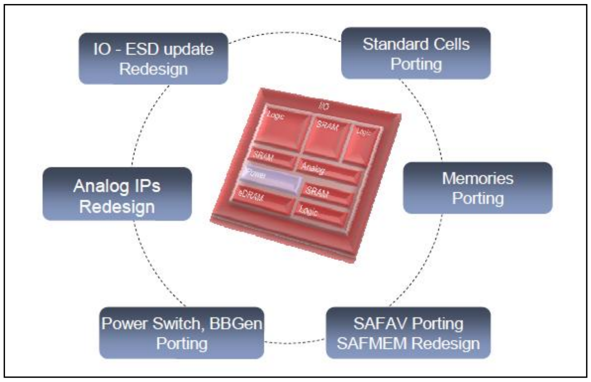

IO - ESD update Redesign
Standard Cells Porting
Analog IPs Redesign
Memories Porting
Power Switch, BBGen Porting
SAFAV Porting SAFMEM Redesign
IO
Logic
SRAM
SRAM
Analog
Fuse
eDRAM
SRAM
Logic

Fig.13: Bulk CMOS to planar FD design migration

At SOC level, migrating an existing design from bulk to planar FD represents an effort comparable to half-node migration, for example from 45nm to 40nm. In other words, it brings very worthwhile benefits at reasonable efforts. A typical approach could be:

- CPU and GPU: the main objective is maximum peak performance and the design is re-worked, making the most of FBB;
- Other SOC blocks: the main objective is power savings, by reaching the target operating frequencies at lower Vdd; there is no change to block design, Timing Analysis is re-run and ECO (Engineering Change Order) is performed to fix violations if needed.
- Other IP such as IOs and PHY blocks are swapped for their planar FD counterpart.

5.4. Power Management and Design Techniques

All techniques used in low-power designs are applicable to planar FD. Fig. 14 lists the main techniques. Those that can be enhanced with planar FD are highlighted in green – keeping in mind that we have also seen earlier that voltage scaling is particularly efficient with FD.

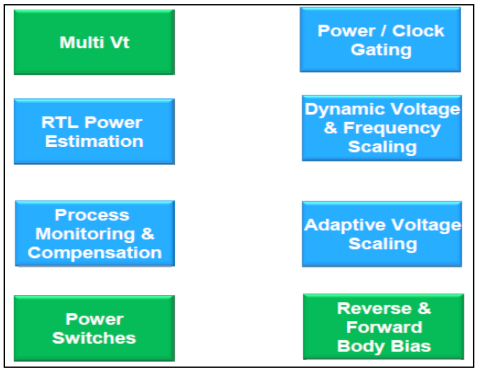

Multi Vt
Power / Clock Gating
RTL Power Estimation
Dynamic Voltage & Frequency Scaling
Process Monitoring & Compensation
Adaptive Voltage Scaling
Power Switches
Reverse & Forward Body Bias

Fig.14: Power Management Techniques

Multi-VT:

Although VT are set differently at process level (as outlined in chapter 3), this is transparent for designers. However, they get the additional possibility, if wished, to “program” VT by back-biasing: a continuum of VT values around the nominal value is available by modifying the back-bias voltage around its nominal value.

Power Switches:

Power switches are used to cut the power supply to a particular block in a SOC when the functionalities it brings are not needed. This is an efficient leakage reduction technique, for low standby power consumption.

Planar FD can exploit the back-bias capability in a special way to create extremely efficient power switches. By implementing a hard gate-to-well connection, a 25% reduction of Ron is obtained. This means lower voltage drop (or “IR drop”), which can be exploited in two ways:

- for the same number of power switches around a given block, this block will see a higher effective supply voltage and run with better performance,
- or, for the same voltage drop, fewer power switches need to be inserted, leading to area savings and a reduction of the leakage currents -- we estimate that stand-by power can be reduced by over 25%.

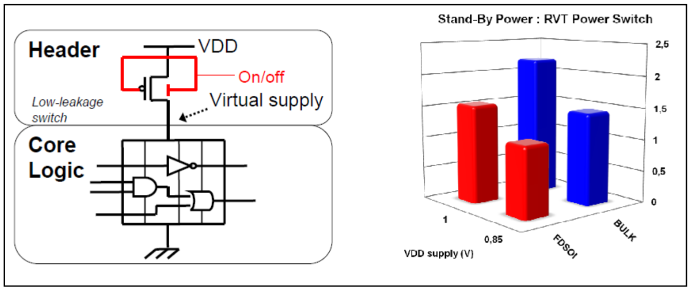


Fig.15: Planar FD optimized power switch

Back-Bias:
Back-biasing consists of applying a voltage just under the BOX of target transistors. Doing so changes the electrostatic control of the transistors and shifts their threshold voltage VT, to either get more drive current (hence higher performance) at the expense of increased leakage current (forward back-bias, FBB) or cut leakage current at the expense of reduced performance. Figure 16 illustrates the concept. For NMOS (top curve of the VT vs. Vbs diagram), forward bias means Vbs>0, that is bringing gnds above gnd – and reverse bias is the opposite. Conversely for PMOS (bottom curve), forward bias means Vbs<0, that is bringing Vdds below Vdd.
Back-biasing can be utilized in a dynamic way, on a block-by-block basis. It can be used to boost performance during the limited periods of time when maximum peak performance is required from that block (at these times, the leakage current is a second order concern). It can also be used to cut leakage during the periods of time when limited performance is not an issue (one obvious example being when the block is not being used i.e. can be put in stand-by). In other words, back-bias offers a new and efficient knob on the speed/power trade-off.

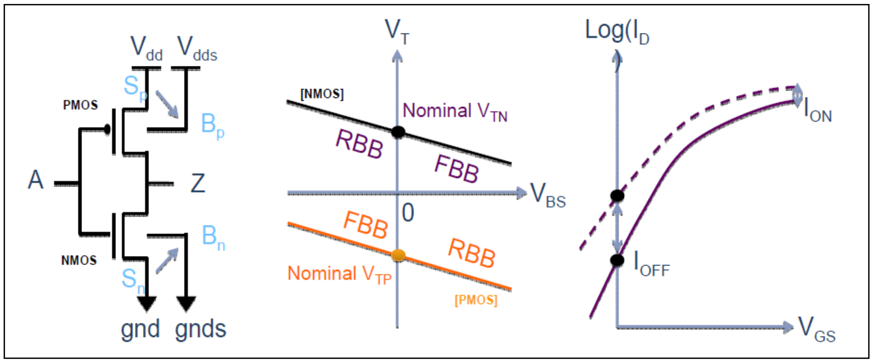


Fig.16: back-biasing concept

At transistor level, the back-biasing voltage supply is contacted to the substrate underlying the BOX through the top silicon and the BOX. There is no area penalty to insert these substrate ties in the layout compared to implementing the substrate ties required in standard bulk CMOS technology.

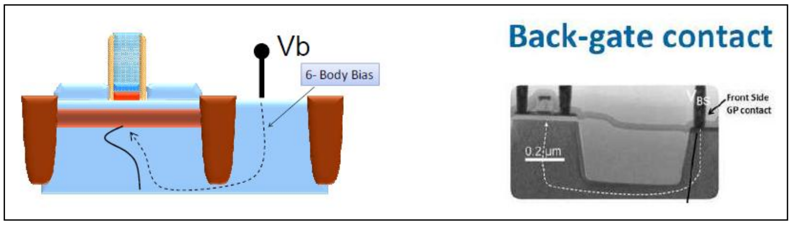


Fig.17: back-biasing implementation at silicon technology level

While back-bias in planar FD is somewhat similar to body-bias that can be implemented in bulk CMOS technology, it offers a number of key advantages. First, body bias is rapidly losing efficiency in bulk technology as transistor dimensions are reduced for advanced nodes; while back-bias in planar FD remains very effective. Second, the maximum amplitude of the bias, and therefore its impact, is rather limited with bulk CMOS: the conventional bulk transistor structure does not allow reaching outside a +-300mV body bias range, otherwise leakage currents become unacceptable. With planar FD, because the buried oxide provides complete dielectric isolation of the source and drain, the possible back-bias range is much wider and large performance boost factors can be obtained. The limiting factor is now the p-n junction between wells, which must not be forward biased e.g., by making the p-well bias significantly higher than the n-well bias.

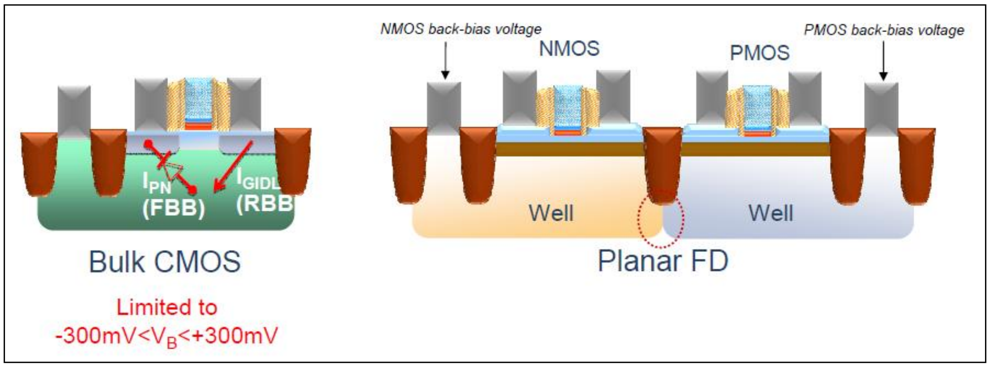


Fig.18: Notional cross-section of a back-biased structure

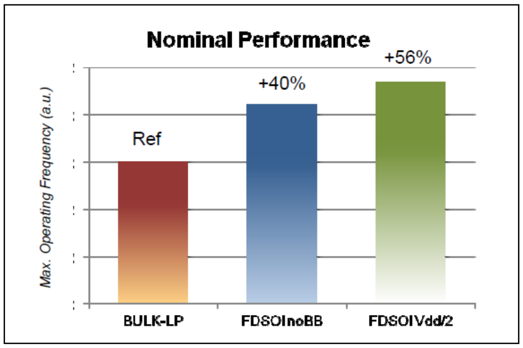


Fig.19.a: Performance boost with FBB.

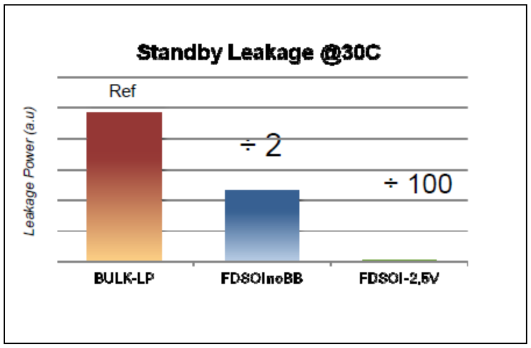


Fig.19.b: Leakage reduction with RBB

6. Perspectives

6.1. 28nm

With the 28nm planar FD technology, on top of preparing the work for 20nm where the kind of power/performance trade-off enabled by planar FD will be key, we are already able to demonstrate very attractive results. We expect to sign-off designs breaking the 2GHz barrier under worst-case conditions, in a power-efficient and cost-efficient way. For lower performance targets, there is also the opportunity to design ultra-low-power chips that can fulfill their functional specifications using a very low Vdd, for example in the 0.6-0.8V range.
The Process Design Kit (PDK) is available, targeting the technology to be open for risk production by mid-2012.

6.2. 20nm

We intend to scale our planar FD technology to 20nm, introducing a number of improvements to continue pushing the performance and retain a low power consumption. The objective is to bring up a solution that will improve on what mobile-optimized planar bulk CMOS will achieve, and will be extremely competitive vs. potential FinFET-based approaches for SOC – while keeping a simple and cost-efficient approach. The design rules will be compatible with 20nm bulk CMOS. This technology will bridge the gap to 14nm and provide an interesting alternative to the cost and complexity of introducing Extreme-UV and FinFET structures.
Evaluation SPICE models are available, and full PDK is scheduled by end of 2012, with risk production for 13Q3.

6.3. 14nm

Based on the assessments we have performed, we are confident that the planar FD technology is shrinkable to 14nm. Silicon and buried oxide thickness will need to be reduced to within limits that wafer manufacturers and CMOS process technology can handle.

7. CONCLUSION

The findings exposed in this document indicate planar FD is a promising technology for modern mobile and consumer multimedia chips. It combines high performance and low power consumption, complemented by an excellent responsiveness to power management design techniques. The fabrication process is comparatively simple and is a low-risk evolution from conventional planar bulk CMOS – and there is little disruption at design level, too.

At 28nm, we find that planar FD more than matches the peak performance of “G”-type technology, at the cost and complexity of a low-power type technology, with better power efficiency across use cases than any of the conventional bulk CMOS flavor.

Looking further, for 20nm and 14nm, we believe planar FD will be extremely competitive with respect to alternative approaches in terms of performance and power, while being both simpler and more suited to low-power design techniques. In short, a better choice for the type of SOC we offer.

REFERENCES

[1] T Skotnicki et al.: « Innovative Materials, Devices, and CMOS Technologies for Low-Power Mobile Multimedia”, IEEE Transactions on electron devices, January 2008

[2] T Skotnicki, IEDM Conference Short Course: “Low Power Logic and Mixed-Signal Technologies”, December 2009

[3] T Skotnicki, IEDM Conference Short Course: “CMOS Technologies: Trends, Scaling and Issues”, December 20010

[4] O. Faynot et al.: ”Planar Fully Depleted SOI Technology: a powerful architecture for the 20nm node and beyond”, IEDM Conference, December 2010