New Elements and Features in the Process Design Kits for a FinFET Technology

Ning Lu[1], Arnold Baizley[2], Ximeng Guan[3], Jeffrey Johnson[2],
Judy McCullen[1], A. Merve Ozbek[4], Ardasheir Rahman[1], Hailing Wang[1],
Mickey Yu[2], Cole Zemke[2], Werner Rausch[3], and Richard Wachnik[3]

[1]*IBM Systems, Essex Junction, VT 05452 USA, lun@us.ibm.com*
[2]*GlobalFoundries, Essex Junction, VT 05452 USA*
[3]*IBM Systems, Fishkill, NY 12533 USA*
[4]*GlobalFoundries, Fishkill, NY 12533 USA*

ABSTRACT

Parasitic resistance is a primary performance constraint in FinFET technologies. Innovations have been introduced into Process Design Kits for 14nm SOI FinFET technology that enable accurate and efficient modeling and simulations (i) of the effects of different contact schemes to local interconnects (LIs) on LI resistance and (ii) of the effects of shared, unshared, and merged diffusion regions on FinFET drain current. These new elements and features aid circuit layout optimization. Measured 14nm SOI FinFET data demonstrate the accuracy of the new modeling approach.

Keywords: FinFET, process design kit, parasitic resistance, FinFET resistance modeling, ring oscillator delay

1 INTRODUCTION

Over the last decade, the performance evaluation of an advanced semiconductor technology node has evolved from using a simple CV_{dd}/I_{dsat} formula to using a ring oscillator and its effective drain current I_{eff} [1] and further to using a suite of logic circuits that also include NAND and NOR gates. Process Design Kits (PDKs) [2] are now used to evaluate the performance of these logic circuits. In FinFET technologies, MOL/parasitic resistance and capacitance have emerged as bottlenecks to technology performance. Fig. 1 shows measured DC resistance components of 14nm logic pFET [3] in the linear region with both source and drain diffusions being unshared during measurement. Further, when each of source and drain is a shared diffusion (e.g., in the middle of a multi-finger FET), MOL resistance is doubled and its weight in total resistance increases from 28% (Fig. 1) to 44% in the linear region. On the other hand, when each of source and drain is a merged diffusion (i.e., the stacked FET case; e.g., the middle nFET in a NAND3 gate), MOL resistance becomes zero. Large MOL resistance value and its large swing among various layout cases (unshared, shared, and merged diffusions) require careful and physical modeling to accurately predict/reflect circuit performance. Fig. 2 shows the contribution from various resistive components to circuit delay in 14nm SOI FinFET technology [3]. In this paper, we report new and innovative elements and features that have been added into the PDKs for 14nm SOI FinFET technology, which has become an indispensable part of both device and circuit design.

2 ENHANCED LVS/PEX TOOLS

This is the first new element. Layout vs. schematic (LVS) and parasitic extraction (PEX) tools in IBM/GF 14nm SOI FinFET technology [3] have been enhanced to netlist one of several MOL (middle of the line) resistance models based on actual connection scheme to a given local interconnect. For example, for each of LI layout illustrated in Fig. 3, a specific LI resistance model (**`lires_t`**) is netlisted, and for each of LI layout illustrated in Fig. 4, another specific LI resistance model (**`lires_e`**) is netlisted. Related layout dimensions (e.g., the type, number, and locations of contacts V0/V0BAR) are passed to the LI models.

3 PARASITIC RESISTANCE MODELS

This is the second new element. A set of LI resistance models (**`lires_t`** and **`lires_e`**) have been developed [4] and are used in the PEX flow of 14nm SOI FinFET technology. Each model handles one group of layout cases and only uses two nodes (Figs. 5 and 6). Excellent model-hardware correspondence is demonstrated in Fig. 7. Simple two-node resistance models enable efficient circuit simulations.

A pure PEX approach represents LI resistance using a resistive network (with fixed resistance values) in a SPICE netlist (Fig. 8b). In such an approach, LI resistance always decreases with increasing fin number N_{fin}, even when the fin number N_{fin} is large. The inaccuracy of the pure PEX approach is illustrated in Fig. 8a, where LI resistance values simulated using full FinFET netlists are also provided. Besides the accuracy issue, the pure PEX approach also leads to a much longer SPICE time. A product design team has seen its SPICE simulation time increased from ten minutes to forty minutes, from one day to four days, and from four days to two weeks after switching from using 14nm SOI FinFET technology (where the model approach for LI resistance is used) to using a 14nm bulk FinFET technology (where a pure PEX approach for LI resistance is used).

Since fixed resistance values are used in the pure PEX approach, the impact of LI resistance corners on circuit performance cannot be evaluated easily. (It needs separate corner PEX decks and re-extraction of netlist.) In addition, the change directions of horizontal (r_{se}) and vertical (r_p) LI resistance components with respect to a LI size variation can be either the same or the opposite. For example, a wider LI reduces both the horizontal and vertical resistance components in the LI. On the other hand, a taller LI decreases the horizontal resistance component but increases the vertical resistance component in the LI. Distinguishing and correctly enabling various LI resistance corners in a pure PEX approach will be a challenge task. The LI resistance models in 14nm SOI FinFET technology support various LI resistance corners, plus providing a Monte Carlo simulation capability.

During the development of an advanced semiconductor technology, one often needs to know the contribution from each resistive component to the total delay of a logic circuit (Fig. 2). This is accomplished by SPICE simulations together with setting a particular parasitic resistive element to zero in a SPICE netlist or in the LI models (sensitivity analysis). When a set of LI models are used, setting silicon-silicide interfacial resistance (the largest parasitic resistance component in Fig. 2) to zero in a SPICE netlist is very simply: Just override a related model parameter in the LI models. Should the pure PEX approach were used, setting silicon-silicide interfacial resistance to zero (the dominant component of r_p in Fig. 8b) in a SPICE netlist would be both tedious and error prone. As illustrated in Fig. 8a, when the ratio of r_{se} /r_p (the ratio of per-fin horizontal to vertical resistances) is changed from a small value to a large value, the inaccuracy of the pure PEX approach is also changed from being small to being large.

4 SCHEMATIC FINFET MODEL

This is the third new element. For schematic simulations (e.g., using a single FinFET instance to represent a multi-finger FinFET), IBM/GF 14nm FinFET PDK provides schematic FinFET models that also include gate resistance, parasitic MOL and M1 resistance, and parasitic MOL capacitance elements associated with the multi-finger FinFET [5]. Ring-oscillator (RO) switching delay can be easily simulated using a compact netlist that contains only pairs of FinFET instances to obtain RO delay as a function of fin number N_{fin} and/or finger number N_f. This enables efficient exploration of design choices by designers. Figure 9 plots RO delay as a function of fin number N_{fin}, and illustrates that there is an optimal fin number for RO design.

The RO behavior shown in Fig. 9 can be explained by the behavior of the MOL resistance and the behavior of parasitic/MOL capacitance in FinFET. (i) When both MOL resistance and capacitance are turned off in the schematic FinFET models ("No R, No C"), RO delay is insensitive to N_{fin}. This behavior is anticipated.

(ii) When only C_{MOL} is turned on ("C only"), RO delay decreases with N_{fin}. This is due to the existence of an extra amount of fringe capacitance (~$c_{\mathrm{end\_fringe}}$/2) at each end of gate (PC) (see Fig. 10). This leads to a decrease in per-fin capacitance when fin number N_{fin} increases,

$$C_{\mathrm{MOL}} / N_{fin} = c_{p,\mathrm{1fin}} + c_{\mathrm{end\_fringe}} / N_{fin} , \quad (1)$$

$$C_{tot} / N_{fin} = c_{\mathrm{fet,1fin}} + C_{\mathrm{MOL}} / N_{fin} , \quad (2)$$

where $c_{\mathrm{fet,1fin}}$ represents single-fin's capacitance contained in an intrinsic FinFET model and C_{MOL} is MOL/parasitic capacitance contained inside the schematic FinFET model but outside the intrinsic FinFET model. Each of $c_{\mathrm{fet,1fin}}$, $c_{p,\mathrm{1fin}}$, and $c_{\mathrm{end\_fringe}}$ is positive.

(iii) When only parasitic resistance is turned on ("R only"), RO delay increases with N_{fin}. One reason for this behavior is the existence of a horizontal resistance component ($r_h N_{fin}$) in source/drain (S/D) local interconnect as well as a fixed amount of S/D resistance component (r_0) that does not decrease with increasing fin number N_{fin} (e.g., discrete contact/V0 resistance) [4]. This leads to an increase in "per-fin" DC resistance (i.e., the inverse of per-fin DC conductance) when fin number N_{fin} increases:

$$N_{fin} R_{S/D} = r_v + r_0 N_{fin} + r_h N_{fin}^2 , \quad (3)$$

$$N_{fin} R_{tot,DC} = r_{\mathrm{fet,1fin}} + N_{fin} R_{S/D} , \quad (4)$$

where $r_{\mathrm{fet,1fin}}$ represents single-fin's DC (voltage dependent) resistance contained in the intrinsic FinFET model and $R_{S/D}$ is the source/drain resistance contained inside the schematic FinFET model but outside the intrinsic FinFET model. Each of $r_{\mathrm{fet,1fin}}$, r_v, r_0, and r_h is positive. Another reason for this behavior is the existence of a horizontal component ($r_{g,h} N_{fin}$) in gate resistance as well as a fixed amount of gate resistance component (r_{g0}) that does not decrease with increasing fin number N_{fin} (e.g., metal gate resistance outside diffusion region and the resistance in the local interconnect that connects to the metal gate) [5]. This also leads to an increase in "per-fin" gate resistance when fin number N_{fin} increases:

$$N_{fin} R_G = r_{g,v} + r_{g0} N_{fin} + r_{g,h} N_{fin}^2 , \quad (5)$$

where each of $r_{g,v}$, r_{g0}, and $r_{g,h}$ is positive.

(iv) After including the effects of both MOL resistance and capacitance, it follows from Eqs. (1)–(4) that the product of total DC resistance and total capacitance has the following N_{fin} scaling relation:

$$R_{tot,DC} C_{tot} = (\tau_{-1} / N_{fin}) + \tau_0 + \tau_1 N_{fin} + \tau_2 N_{fin}^2 , \quad (6)$$

with

$$\tau_{-1} = (r_{\mathrm{fet,1fin}} + r_v) c_{\mathrm{end\_fringe}} ,$$

$$\tau_0 = (r_{\mathrm{fet,1fin}} + r_v)(c_{\mathrm{fet,1fin}} + c_{p,\mathrm{1fin}}) + r_0 c_{\mathrm{end\_fringe}} ,$$

$$\tau_1 = r_0 (c_{\mathrm{fet,1fin}} + c_{p,\mathrm{1fin}}) + r_h c_{\mathrm{end\_fringe}} ,$$

$$\tau_2 = r_h (c_{\mathrm{fet,1fin}} + c_{p,\mathrm{1fin}}) .$$

RO delay vs. N_{fin} scaling relation (which includes an additional delay component from gate resistance R_G) is similar to Eq. (6) but with slightly different τ_{-1}, τ_0, τ_1, and τ_2

values. Note that each of τ_{-1}, τ_0, τ_1, and τ_2 is positive. Thus, we have explained that (a) RO delay decreases with increasing fin number N_{fin} when the fin number N_{fin} is small (due to the reduction of per-fin capacitance), (b) RO delay increases with increasing fin number N_{fin} when the fin number N_{fin} is large (due to the increase of "per-fin" resistance), and (c) there is an optimal fin number at which RO delay is minimized. Measured data in Fig. 11 support that (N_{fin}·DCR_{eff}) and (N_{fin}·R_{on}) increase with N_{fin} at a constant nFET-pFET mean saturated threshold voltage.

5 DISTINGUISH BETWEEN SHARED & UNSHARED LOCAL INTERCONNECTS

This was fourth new element. A local interconnect in a diffusion region can be "unshared LI" or "shared LI and simultaneous switching (SS)" (Fig. 12). An example of a "shared LI and simultaneous switching" is each inner LI in a multi-finger FinFET, where two active channels next to an inner LI are always simultaneously switched. Such a layout or switching difference has an impact on the actual resistance experienced by the drain current in the diffusion region and/or in the LI. 14nm LVS tool analyzes a layout and passes the information on shared vs. unshared to a netlisted **`lires_t`** or **`lires_e`** model instance (through a switch **`is_shared`**). Our early LI resistance models returned a large resistance value for a shared LI.

6 HANDLE ALL LAYOUT-PLUS-SWITCHING CAES

This is the fifth new element. Besides "unshared LI" and "shared LI and simultaneous switching", there are two other layout-plus-switching cases. One is merged diffusion (i.e., stacked FET) case (Fig. 13a), which occurs between two nFETs in a NAND gate or between two pFETs in a NOR gate. The other is "shared LI but non-simultaneous switching (nonSS)" (Fig. 13b), which could occur between two pFETs in a NAND gate or between two nFETs in a NOR gate. Figures 14 and 15(a) explain that there are three different resistance values for a given diffusion region, and the smallest value is in the stacked FET case. When a FinFET model contains all diffusion resistance (including silicon-silicide contact resistance) (typical TCAD simulated I-V curves do include them), two cases in Fig. 12 can be handled correctly but not the two cases in Fig. 13. Both theory and TCAD simulation (Fig. 16) have been used to find a relation among the three different diffusion resistance values. Using the resistance networks in Fig. 14, it can be shown that ΔR_1 and ΔR_2 in Fig. 15 are equal. We have thus changed FinFET models to stacked FET based (subtract out ΔR_2), and add ΔR_2 into the LI models (Fig. 15b). Stacked FET based FET and LI resistance models correctly count the diffusion resistance values in all four layout-plus-switching cases (Fig. 17). Unshared LI based FET and LI resistance models over-estimated the diffusion resistances for the two cases in Fig. 13. The resulted under-estimation on FET's drain current in the liner region is not negligible (Fig. 18).

7 SUMMARY

We have developed several innovative elements for the PDKs of FinFET technologies. Available hardware data have been used for verification. These new elements enabled accurate performance evaluation of 14nm SOI FinFET technology. These new elements also enabled both layout optimization and efficient FinFET circuit simulations for circuit designers.

REFERENCES

[1] M.H. Na, E. J. Nowak, W. Haensch, and J. Cai, IEDM, 2002, p. 121.

[2] G. Yeric, B. Cline, S. Sinha, D. Pietromonaco, V. Chandra, and R. Aitken, CICC, 2013.

[3] C.-H. Lin *et al*., IEDM, 2014, p. 74.

[4] N. Lu, P. M, Kotecha, R. A. Wachnik, CICC, 2014. N. Lu and R. A. Wachnik, IEEE Trans. Circuits and Systems–I, Vol. **62**, 2015, p. 1899.

[5] N. Lu, T. B. Hook, J. B. Johnson, C. Wermer, C. Putnam, and R. A. Wachnik, IEEE Electron Device Lett., Vol. **34**, 2013, p. 1100.

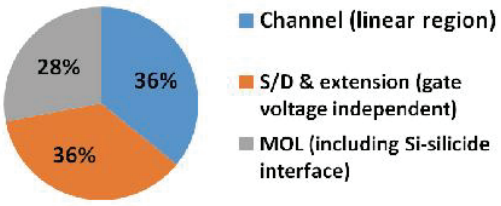

Figure 1. Experimental partition of DC resistance components for logic pFET in 14nm SOI FinFET technology in the linear region. A Kelvin FET structure was used and both source and drain diffusions are unshared during measurement. (MOL stands for "middle of the line".)

Component	Contribution	
Intrinsic FET	78.7%	
Si-silicide Interface	9.5%	Sum of MOL and BEOL parasitics = 21.3%
Lower LI	1.9%	
Upper LI	3.4%	
Gate (PC)	1.8%	
Gate (LI)	2.2%	
V0 (S/D and gate)	0.7%	
M1 (S/D and gate)	0.1%	
V1 (S/D and gate)	1.0%	
M2 (S/D and gate)	0.8%	

Figure 2. A partition of simulated contribution from resistive components to the AC performance of critical-path-like of logic circuits in 14nm SOI FinFET technology.

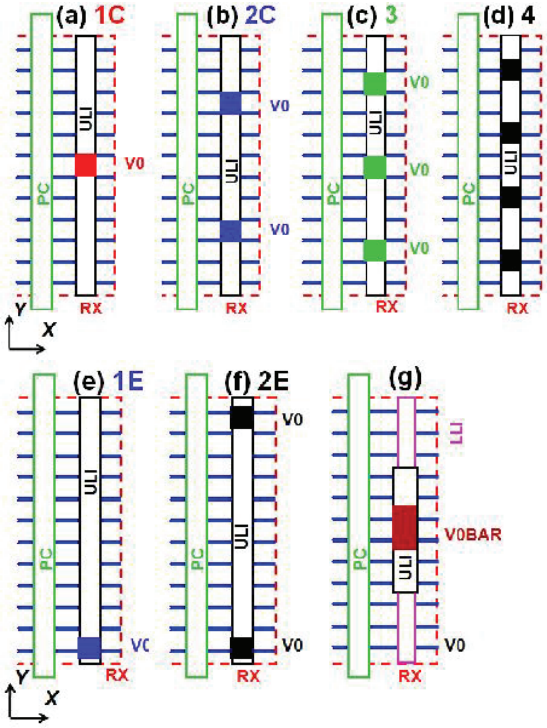

Figure 3. (a) One group of common (e.g., drain side) contact cases for local interconnect (LI). ULI (LLI) stands for upper (lower) local interconnect.

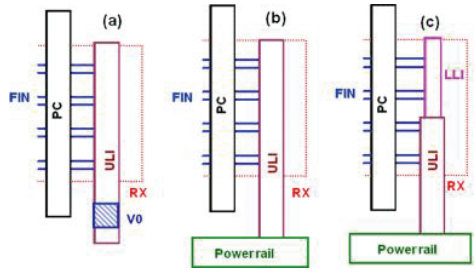

Figure 4. Another group of common (e.g., source side) contact cases for local interconnect.

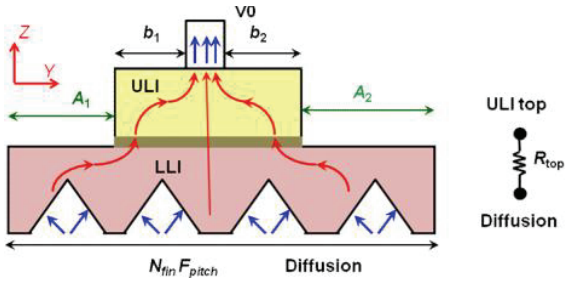

Figure 5. A two-node LI resistance model **`lires_t`** is netlisted to handle each layout case in Fig. 3. Some model parameters are shown. The current flow in Fig. 3(g) is shown.

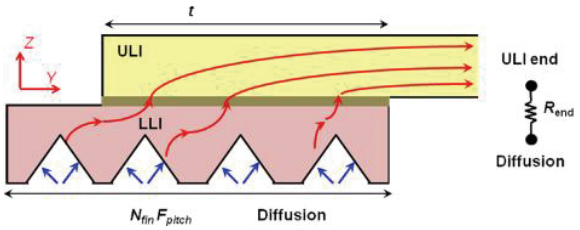

Figure 6. Another two-node LI resistance model **`lires_e`** is netlisted to handle each layout case in Fig. 4. Model parameters are illustrated. The flow of drain current in Fig. 4(c) is depicted.

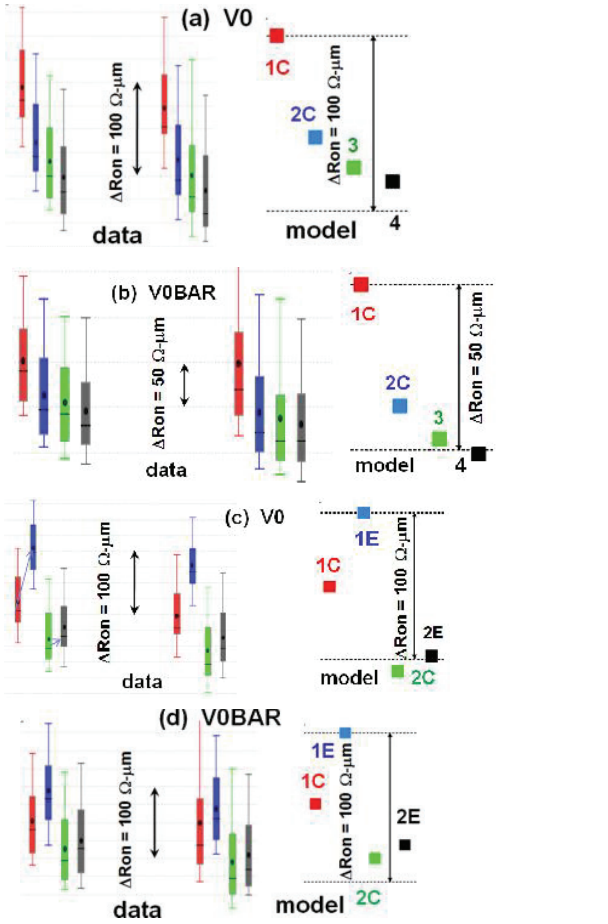

Figure 7. Silicon verification of 14nm LI resistance models. Excellent MHC was obtained on the changes to R_{on} caused by the changes in the number, locations, and types of contact. The left column shows measured 14nm R_{on} distribution, and the right column gives nominal R_{on} from 14nm LI models, with a same color indicating a same layout (see Fig. 3). N_f = 1, N_{fin} = 12, and source and drain layout are the same.

TechConnect Briefs 2016, TechConnect.org, ISBN 978-0-9975-1173-4

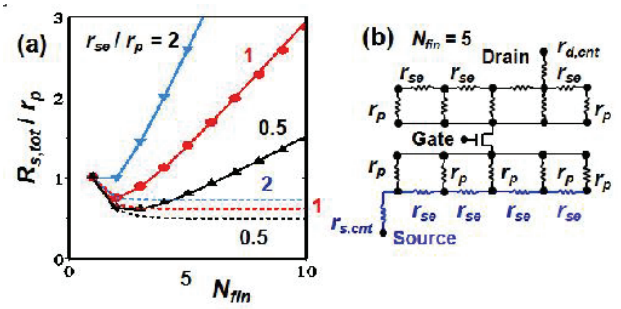

Figure 8. (a) Source-LI resistance values among 3 different approaches: full netlist (symbols; $R_{on,1fin} = 200r_{se}$), LI model (solid curves), and a pure PEX approach (dashed curves) shown in (b).

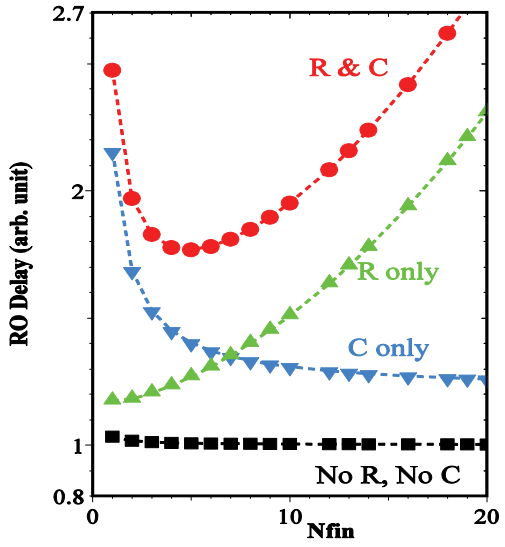

Figure 9. FinFET technology RO delay vs. fin number (N_{fin}) has a minimum values (red curve, "R & C") (N_f = 1, source LI to power raid, 1 or more V0s above drain LI).

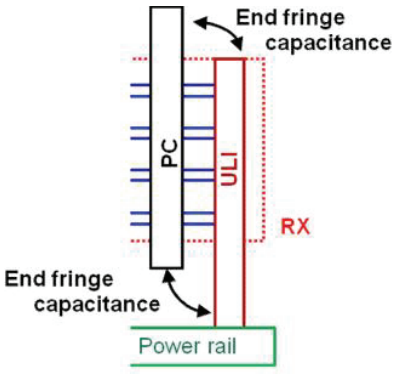

Figure 10. Illustration of end capacitance c_{end_fringe}.

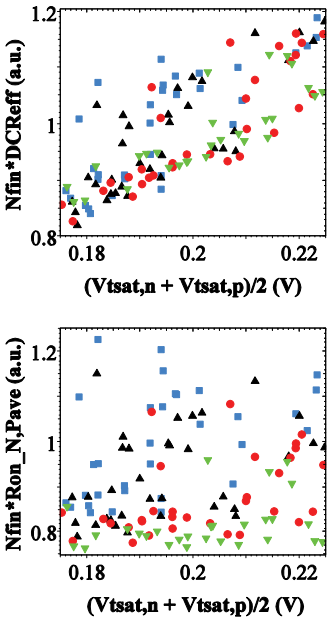

Figure 11. Measured 14nm SOI FinFET data on N_{fin}·DCR_{eff} and $N_{fin}(R_{on,n} + R_{on,p})/2$ vs. the average of nFET-pFET Vt,sat. DCR_{eff} = $(V_{dd}/I_{eff,n} + V_{dd}/I_{eff,p})/2$. Symbols: 2 fins, 3 fins, 4 fins, and 5 fins.

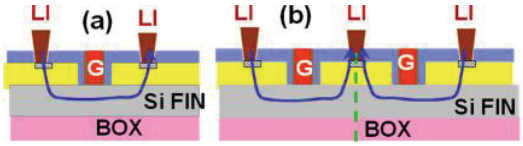

Figure 12. (a) Each LI is an unshared LI. (b) The LI in the middle is shared and simultaneous switching. The two drain current flow arrows in (b) indicate simultaneous switching (SS). The silicon-silicide interfacial resistance experienced by the drain current of a single finger when passing through a shared LI is twice the resistance when passing through an unshared LI, since the passage area for the drain current of a single finger when passing through a shared LI is half the area when passing through an unshared LI. When the silicon-silicide interfacial resistance is included in an intrinsic FinFET model, this interfacial resistance difference cannot be handled by the intrinsic FinFET model.

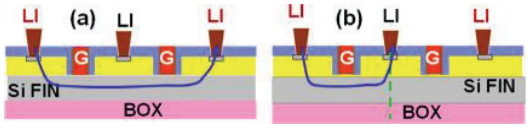

Figure 13. (a) The LI in the middle belongs to merged diffusion (i.e., stacked FET) case. (b) The LI in the middle is shared but non-simultaneous switching.

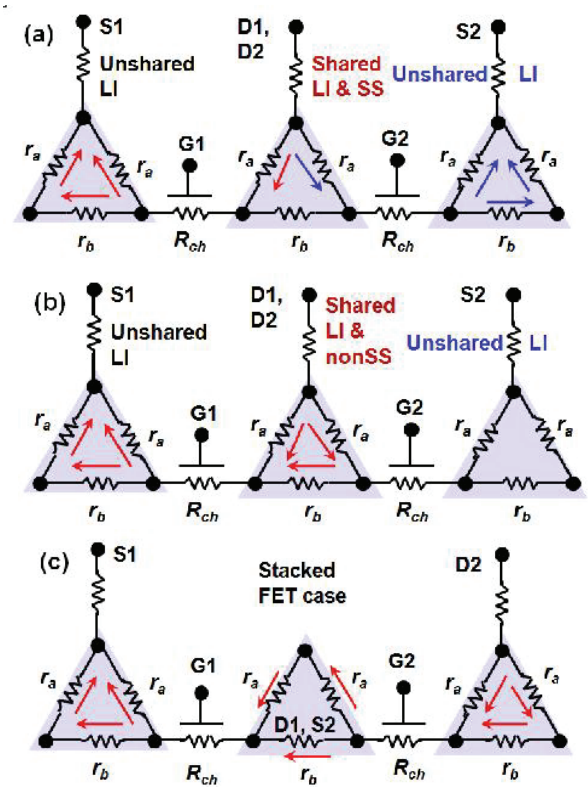

Figure 14. Resistance networks to illustrate different current paths (and thus different resistance values) in the diffusion (triangle) region among 4 cases: "unshared LI" in (a) and (b), "shared LI and simultaneous switching" in (a), "shared LI but non-simultaneous switching" in (b), and stacked FET case in (c). The challenge is to correctly and compactly represent all four cases in a set of FET and LI resistance models. The resistance value in the "shared LI but non-simultaneous switching" case is the same as that in the "unshared LI" case.

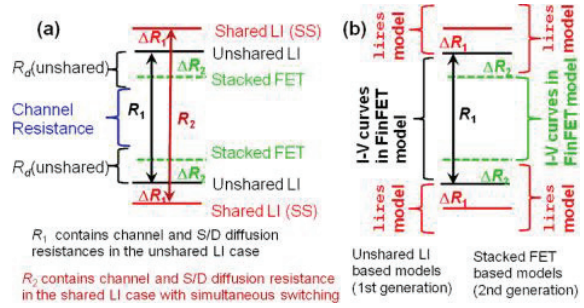

Figure 15. (a) A schematic representation of the amount of resistances in unshared LI, shared LI and simultaneous switching (SS), and stacked FET (merged diffusion) cases. (b) Re-division of resistance boundary lines between intrinsic FinFET and LI resistance models.

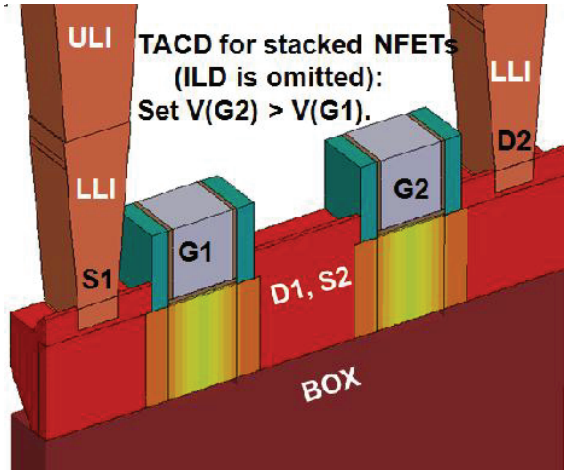

Figure 16. A carefully constructed TCAD simulation on 14nm SOI FinFET gave $\Delta R_2/\Delta R_1 = 1.03$. Half fin structure is depicted here. Routine TCAD simulation for two stacked FETs is a challenge.

Figure 17. A same resistance value r_2 is used for all 4 layout-plus-switching cases in Figs. 12 and 13. No instance parameter on layout is needed. Use only 1 node (minimum node number) in each LI resistance model for the diffusion region. The value of r_2 is $(R_{LI} + \Delta R_1)$. FinFET model is stacked FET based. A Kelvin FET structure actually measures the I-V curves in the stacked FET based FinFET models.

Case	Resistance	Current
S or D is a merged diffusion	10.0%	–9.1%
Both S and D are merged diffusions	23.8%	–19.2%
Shared LI nonSS in S or D	8.6%	–7.9%
Shared LI nonSS in S and D	17.2%	–14.6%

Figure 18. The amount of over-estimation for total resistance and that of under-estimation for drain current in the linear region when a set of unshared LI based FET and LI resistance models were used. Resistance components in Figs. 1 and 2 are used. The sum of FET channel and S/D extension resistances in the liner region is 72%. Each of source- and drain-side MOL resistance is 14%. Within MOL resistance, 19/31 is attributed to silicon-silicide interface resistance, and the rest 12/31 is attributed to the sum of lower LI, upper LI, and V0 resistances.