RTA-Driven Intra-Die Variations in Stage Delay, and Parametric Sensitivities for 65nm Technology

I. Ahsan[1], N. Zamdmer[1], O. Glushchenkov[1], R. Logan[1], E. J. Nowak[2], H. Kimura[3], J. Zimmerman[2], G. Berg[1] ,J.Herman[1], E.Maciejewski[1], A.Chan[1], A.Azuma[4], S.Deshpande[1], B. Dirahoui[1], G.Freeman[1], A. Gabor[1], M. Gribelyuk[1], S.Huang[1], M. Kumar[1], K. Miyamoto[4], D. Mocuta[1], A. Mahorowala[1], E.Leobandung[1], H.Utomo[1], B.Walsh[1]
[1]IBM Systems and Technology Group, 2070 RT. 52, Hopewell Junction, NY 12533 ,USA,
phone: (845)894-7627, fax: (845)892-6264,email: ishtiaq@us.ibm.com
[2]IBM Semiconductor Research and Development Center, 1000 River Road, Essex Junction, Vermont,USA, [3]Sony Electronics Inc., Hopewell Junction, New York, 12533, [4]Toshiba America Electronic Component Inc,Hopewell Junction, New York, 12533

Abstract

We report, for the first time, a detailed study of Intra-Die Variation (IDV) of CMOS inverter delay for the 65nm technology, driven by mm-scale variations of rapid thermal annealing (RTA). We find that variation in V_T and R_{EXT} accounts for most of the IDV in delay and leakage and is modulated by lamp RTA ramp rate. We show a good correlation of inverter delay to mm-scale variation in the predicted reflectivity of the device pattern densities. (Keywords: CMOS, RTA, Variation, and Ramp Rate).

Introduction

Control over intra-die variation has been an important element for the successful manufacture of VLSI circuits. Until recently, focus has been on Across-Chip Line-width Variation (ACLV) [1] as the primary cause of leakage and delay dispersion within die. More recently, attention has been called to the short time scale of so-called 'spike' RTA processing [2, 3], since the length scale, L, over which thermal equilibrium can be reached for a given time, t, is approximately $L\sim(\sigma/c_v*t)^{1/2} \sim 4$mm, (where σ and c_v are the thermal conductivity and specific heat of silicon, respectively), significantly smaller than the size of many VLSI die. Heretofore, focus of RTA-induced variations has been on junction depth and activation, however, we show that minor pattern density variations within a die can result in significant variations in inverter delay, driven primarily by V_T and R_{EXT}, modulation.

Intra-Die Variation Data

A physically compact set of structures was designed to enable characterization of local behavior of thermal anneal in an advanced 65nm-generation CMOS technology. These structures enabled the measurement of doped poly-silicon sheet resistance (R_S), gate length, DC FET currents and inverter delay (via ring oscillator speed). Seven identical copies of this set were placed in specific locations of a test die that had large pattern density differences (Fig. 1); multiple copies of overlap-capacitance structures were also placed. Differences in inverter delays among different location of the die are shown in Fig 1. Large variations in off-current were also found in FETs in these locations. However, Gate-length measurement structures placed in these locations indicated no significant difference among them, as seen in Fig. 2, which shows large inverter delay differences even at the same gate-length between rings located at different locations of the die. NanoSEM measurements were taken on a separate mask-set with similar phenomenon to verify the validity of electrical measurements of the poly length. TEM images were also taken to check for any difference in poly-silicon gate profile between FETs located at different locations of the die. All these results indicated that there were no significant structural differences in the FETs between the different locations.

It was also noted that there were significant variations of both n and p-doped poly-silicon R_S, at these locations (Fig.3). The strong covariance observed indicates that this variation was driven by local activation. A correlation between inverter delay and the sheet resistance is clearly observed (Fig. 4). Similar correlations were found between the sheet resistance and the V_T (Fig 5, 6) for both nFETs and pFETs.

Performance Analysis

In an effort to investigate the cause of the inverter delay variation, detailed analysis of nFET and pFET characteristics were carried out on test transistors adjacent to each inverter ring oscillator. It was found that in addition to large variation in V_Ts, significant co-variation in R_{EXT} occurred. Differences in overlap capacitance were also observed.

Compact model parameters were varied by site to represent the observed behaviors of R_{EXT} and V_T, and inverter delays were simulated for each site. The simulated results (Fig. 7) agreed very well with the observed variation in inverter delay. While it is clear that a more-thorough extraction of variation by site would reveal other contributions to delay variation, the results of Fig. 7 indicate that the observed differences in the ring performance are largely captured by the observed variation in these two parameters, and that others are of secondary importance.

Correlation to Pattern Density and Anneal Ramp Rate

The pattern density (averaged over ~4mm) of the exposed STI layer (i.e. isolation areas not covered by gate poly-silicon) is expected to correlate to local activation since differences in calculated reflectivity are well approximated by this parameter (Table 1). Doped poly-silicon R_S (and hence inverter delay) correlated well with this pattern density (Fig. 8), again strongly indicting differential thermal absorption at the time of anneal.

Experiments with reduced thermal ramp-rates were performed. Intra-die variation of V_T, doped poly-silicon R_S, and inverter delay were all significantly reduced, and in equal proportions (Figs. 4, 5, 6). This clearly proved that the rapid ramp-rate of the lamp thermal anneal was the principle source of the observed intra-die variation of FET parameters and inverter delay.

Conclusion

The need to employ rapid-ramp RTA processing for advanced device structures has introduced new intra-die variation of performance and sub-threshold leakage, which can rival or surpass ACLV in importance. Most of the observed variation can be accounted for by lamp annealing-driven variations in R_{EXT} and V_T: the faster ramp rate RTA results in more variation and the variation itself correlates well with calculated reflectivity for the lamp RTA spectrum and hence depends on the variation of local, mm-scale pattern density. Accurate prediction of chip behavior must account for such pattern dependencies, and future high-speed VLSI processes must carefully consider these issues.

References

[1] D.Chesebro, et al., IBM Journal of Research & Dev., V39, No1/2 1995 pp189-200.
[2] P.J.Timans, et al, 'Challenges for Ulta-Shallow Junction Formation Technologies beyond the 65nm node.', IEEE International Conf. On Advanced Thermal Processing of Semiconductors.-RTP2003.
[3] Th.Feudel et al, Junction Presentation on Scaling for Next Generation Microprocessors, Technical Symposium of Semicon Europe,April 2005

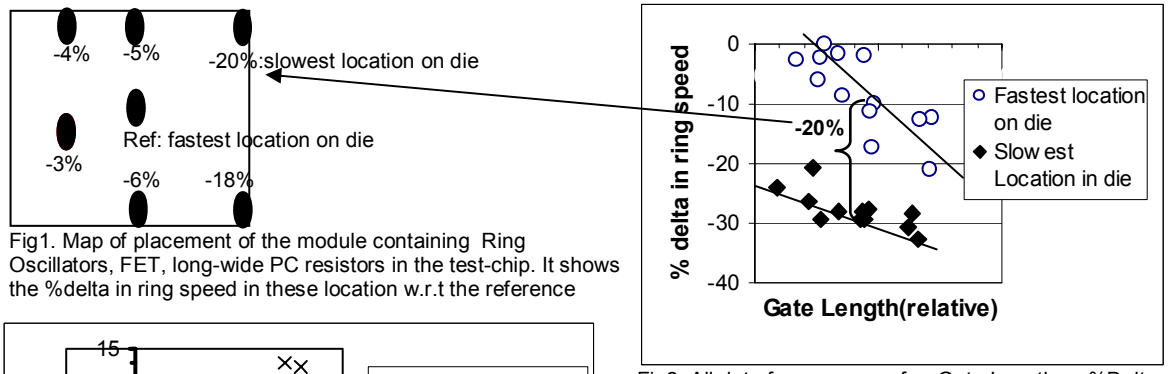

Fig1. Map of placement of the module containing Ring Oscillators, FET, long-wide PC resistors in the test-chip. It shows the %delta in ring speed in these location w.r.t the reference

Fig2. All data from same wafer: Gate Length vs %Delta Ring Speed(delta to fastest ring on wafer) indicating that the large difference in ring speed in different location of the die for the same Gate length

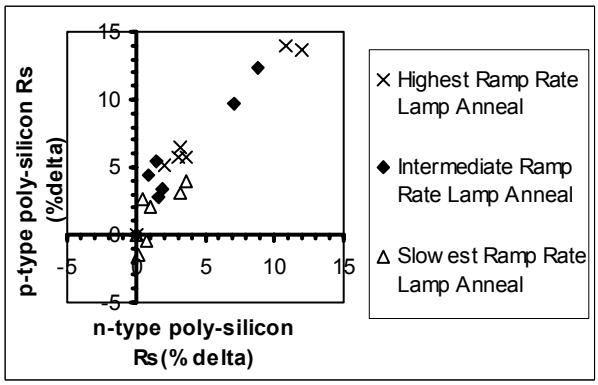

Fig.3 Large delta in unsilicited Poly sheet resistance noted at different locations in the die. Correlation between N and P Sheet resistance indicates *within-die* thermal absorption difference

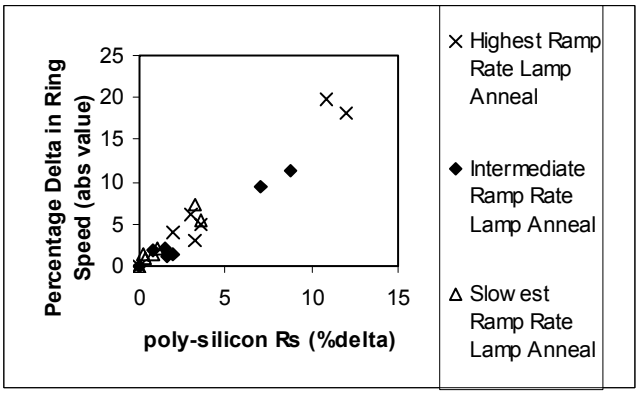

Fig4. Percentage delta in ring speed vs delta in unsilicided long-wide poly sheet resistance at different location in die indicating correlation between delta in *within die* thermal absorption and ring speed

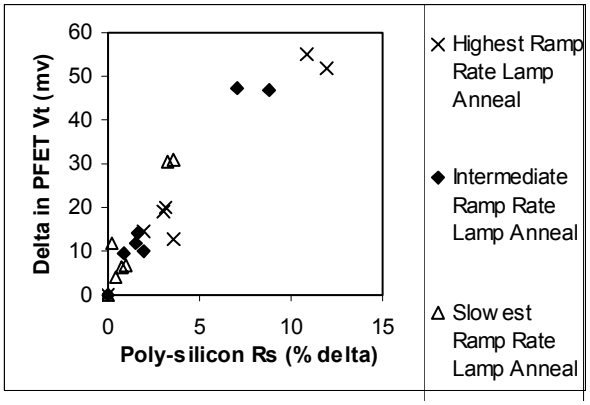

Fig5. *Within die* Percent delta in unsilicided Poly resistance vs *Within die* Delta Vt for PFET (abs. value) indicating Vt difference due to thermal absorption.

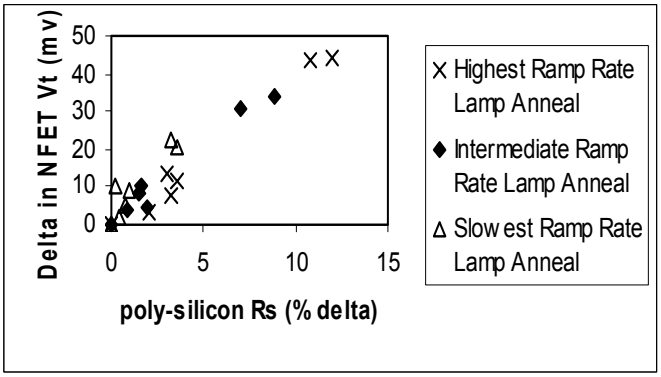

Fig6. *Within die* Percent delta in unsilicided Poly resistance vs *Within die* Delta Vt for NFET.. Similar correlation like PFET delta Vt

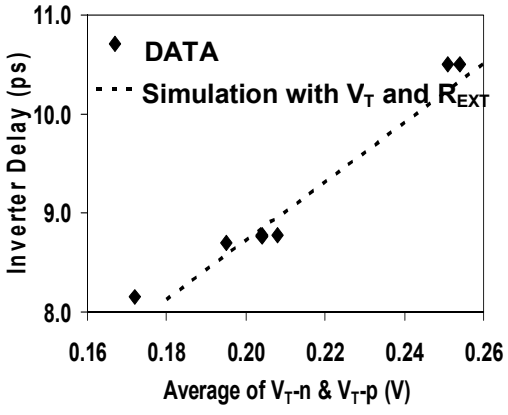

Fig7, Hardware vs Model correlation with model override for Rext and Vt for FETs indicating delta in ring speed is primarily coming from thermal absorption induced difference in Vt and Rext in the FETs.

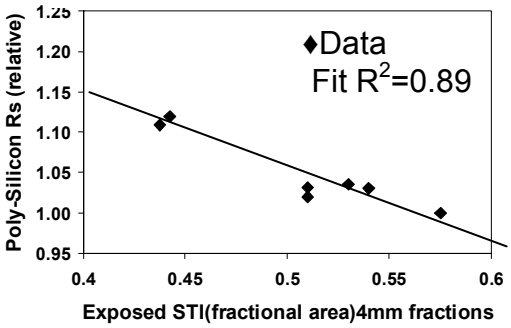

Fig8. unsilicided poly sheet Rs vs exposed STI density of the chip indicating strong dependence of thermal absorption on pattern density

Table 1: Calculated Reflectivity of RTA Radiation.

Region	Reflection Coefficient
N+ source/drain	0.57
P+ source/drain	0.57
Gate over isolation	0.54
Gate over Transisto	0.45
Trench Isolation	0.20