Towards Heterogenous 3D-Stacked Reliable Computing with von Neumann Multiplexing

George R. Voicu, Sorin D. Cotofana
Faculty of EE, Mathematics and CS
Delft University of Technology
Delft, The Netherlands
{G.R.Voicu, S.D.Cotofana}@tudelft.nl

***Abstract*—The reliability of near-future nano-meter range CMOS, and novel nano-computing devices is greatly affected by undesired effects of physical phenomena appearing due to continuous technology scaling. The emerging 3D-Stacking Integrated Circuits (3D-SIC) technology allows devices manufactured using different technologies, and thus with different reliability, to be stacked on top of each other and connected with low latency links. In this paper, we propose to take advantage of this new design space dimension, i.e., the individual reliability of devices, when using the von Neumann multiplexing redundancy technique. Our analysis suggests that multiplexing units reliability importance is determined by how high the error rate of individual gates in the system is, i.e., for high error rates the units at the end of the restoration chain are critical, while for low error rates the units at the beginning of the restoration chain are critical. We further introduce and evaluate the first, to the best of our knowledge, heterogeneous 3D-SIC multiplexing arrangements. Our results indicate that assuming that delay and area are doubled for a technology with an order of magnitude higher reliability, a heterogeneous multiplexing scheme with gates having high and medium error rates can achieve a reduction of $1.79\times$ in delay and area, with a 9% loss in the Reliability Improvement Index (RII), over the homogeneous counterpart with only medium reliability gates. For medium and low error rates, a minimum 1% RII loss can be traded for a delay and footprint reduction of $5.66\times$ and $4.25\times$, respectively.**

***Index Terms*—Multiplexing, Fault-Tolerance, 3D-Stacked ICs, Reliability, Error Rate, Through-Silicon Vias (TSVs).**

I. Introduction

With the silicon technology node size scaling rapidly approaching fundamental physical limits, the reliability of CMOS devices is greatly affected [1]. Effects of undesired physical phenomena like quantum behavior of the atoms will make a challenge to manufacture correct functioning devices, tolerant to external noise disturbances, and with a low degree of degradation in time. On the other hand, alternative emerging device technologies, e.g., molecular devices, spintronics, single-electron tunneling, are still in their infancy and lack reliable mass production means [2]. Having a density higher than CMOS devices, they are more susceptible to both transient and permanent errors, at manufacturing or during use. Therefore, history repeats itself, and the pre-solid-state era problem of how to perform dependable computation with unreliable devices comes into play again.

The solution to deal with errors in a circuit is to use a form of redundancy, in either time (e.g., repeat computations), space (e.g., execute the same computation on multiple instances of the same hardware), or information (e.g., encode the data with error correcting codes) domain [3]. Most of the popular spatial redundancy techniques, like Triple Modular Redundancy (TMR) [4], use a voter to decide the correct results from several version of the computation. The foundation for this lies in the seminal work of von Neumann in which he also introduced *multiplexing*, a redundancy technique that harness the power of a high spatial redundancy factor to achieve near-perfect computations with rather faulty logic gates [**?**]. Because of its requirement of an extremely large number of gates, in the order of tens of thousands, it was then considered ahead of its time. However, due to the increase in density of current and foreseeable nano-devices, in the past decade researchers gained interest in using multiplexing to increase the system reliability.

3D Stacking Integrated Circuits (3D-SIC) technology, although originally proposed as an alternative to scaling in improving the performance of a circuit, also opens up new opportunities for reliable computing [5]. The idea behind the 3D-SIC approach is to partition the design in several smaller parts, and to implement each of them on a separate die. The dies are bonded together forming a multi-tier stack, with signals travelling between the tiers in the stack using special vias, i.e., Through Silicon Vias (TSVs). Blocks placed in the planar case far apart from each other and connected by long global wires can now communicate through the short and low-resistance TSVs. Furthermore, there is no restriction on what type of technology is used to fabricate each die in the stack. In view of this, for a given computing architecture, technologies with larger feature sizes, thus more reliable, can now be used for the most critical parts of the architecture to increase the system reliability. The trade-off is mostly an increase in delay and area on the more reliable dies. Even though there is some concern about the reliability of the TSVs, their defects can be detected at manufacturing time, and their large size makes them not as susceptible to degradation, ageing or transient faults as the devices forming the gates in the design. The new dimension introduced by 3D-SIC in the design space, i.e., the individual reliability of devices, does not replace or prohibits the use of redundancy techniques. Instead, by careful mapping the devices on the stack tiers, it can even be used to enhance their reliability improvement potential.

In this paper, we address this possibility by investigating the use of the multiplexing technique in the context of 3D-SIC with heterogeneous technologies. In particular we focus on NAND multiplexing, one of the two examples of the logic gates used in the multiplexing scheme by von Neumann. First, we identify the contribution that each unit of the multiplexing architecture has to the system reliability, in order to place important units that have the largest contribution, on tiers manufactured in a more reliable technology. Our analysis suggests that the reliability importance of multiplexing units is determined by how high the error rate of the gates in the system is, i.e., for high error rates the units at the end of the restoration chain are critical, while for low error rates the units at the beginning of the restoration chain are critical. Based upon this insight, we further introduce and evaluate the first, to the best of our knowledge, heterogeneous 3D-SIC von Neumann multiplexing arrangements.

Our results indicate that when assuming that gate delay and area double on a technology having an order of magnitude higher reliability, a heterogeneous multiplexing scheme with gates having high and medium error rates can achieve a reduction of $1.79\times$ in delay and area over the homogeneous counterpart with only medium reliability gates, with a 9% loss in the Reliability Improvement Index (RII), respectively. For medium and low error rates, a minimum 1% RII loss can be traded for delay and footprint reductions of $5.66\times$ and $4.25\times$, respectively.

The remainder of the paper is organized as follows. Section II details the von Neumann multiplexing technique, surveys recent analyses of it found in the literature, and highlights the differences with our work. In Section III we asses the impact of multiplexing units on the overall system reliability. Then, we propose in Section IV novel 3D-SIC von Neumann multiplexing schemes and evaluate their impact on reliability. Finally, Section V concludes the paper.

II. Background

Half a century ago, faced with the problem of building computing blocks with unreliable vacuum tubes, von Neumann devised the multiplexing technique as a neural inspired reliable computing paradigm based on redundancy. An example of a multiplexing architecture for computing NAND is given in Figure 1. The basic processing unit (the NAND gate) is replaced by a multiplexing unit, which contains N replicas of the processing unit working in parallel, with their inputs being a random permutation (U) of the N input ports of the multiplexing unit. The input ports are also replicated from the original ports of the basic processing unit. In the case when the processing units, the interconnection network, and the input multiplication are functioning correctly, all N outputs should have the same value. In the presence of faults, some outputs will have a wrong value. For a defined threshold value Δ, the system is considered faulty if the the number of outputs equal with "1" or "0" are in the range $(\Delta N, (1-\Delta)N)$.

Depending on the processing unit failure probability and on how many inputs have the correct value, the percentage

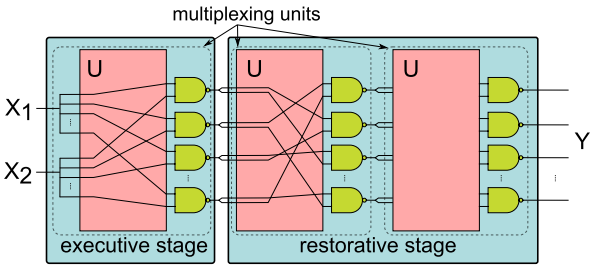


Figure 1: NAND Multiplexing example.

of outputs that have the correct value may be lower than the percentage of inputs having the correct value, i.e., the multiplexing unit can actually decrease the reliability. To compensate for this the primary multiplexing unit, called the *executive* stage, is followed by multiplexing units built on the same principle, with their inputs obtained from the outputs of the previous multiplexing unit. Several *restorative* stages like this can be added to increase the repairing potential.

Von Neumann originally described the scheme using 2 input NAND (NAND-2) gates and 3 input majority (MAJ-3) gates. Nevertheless, the same scheme can be generalized to other Boolean gates, or even considering different degrees of abstraction, from the underlying devices forming the gates, to circuits composed of multiple gates. Care must be taken that the restorative stages produce the same result as the executive stage, thus the derivation of the restorative stages inputs and the number of multiplexing units in a restorative stage have to be specially designed for the processing unit. For example, both inputs of the NAND-2 gate use the output of the previous stage, transforming the NAND-2 into an inverter. Thus, two multiplexing units are needed in one restorative stage, as in Figure 1.

In [6] it was demonstrated that the technique can achieve error rates less than 10^{-18} for high redundancy factors ($N \geq 20000$), an optimal $\Delta = 0.07$, and statistically independent NAND-2 gate failures with a probability $\epsilon = 0.005$. In fact, for the scheme to work, the NAND-2 gates error rate should be upper bounded by 0.0107, while MAJ-3 gates outperform them and could lead to reliability improvements for $\epsilon < 0.0197$ [6], [7], [8], [9]. However, when taking into account the devices forming the gates, this result can change, e.g., it can reverse in certain cases when using single-electron technology [10]. Reliability analyses were made for NAND- and MAJ-based multiplexing schemes with low and medium redundancy factors (3 to 3000), using various methods to compute or estimate the system reliability: analytical modelling [11],[12],[8],[9],[13], probability transfer matrix [14], [10], probabilistic model checking [15], [16], Monte Carlo simulations [10], [17] or Bayesian networks [18].

All of the above mentioned works consider that gates fail with the same probability, with the exception of [10], where the gates are implemented in Single-Electron Tunnelling (SET) technology, and [17], [18], where manufacturing variability effects on the multiplexing scheme are investigated. Our work also considers gates failing with different probabilities, but from a different perspective, and as an effect of a different

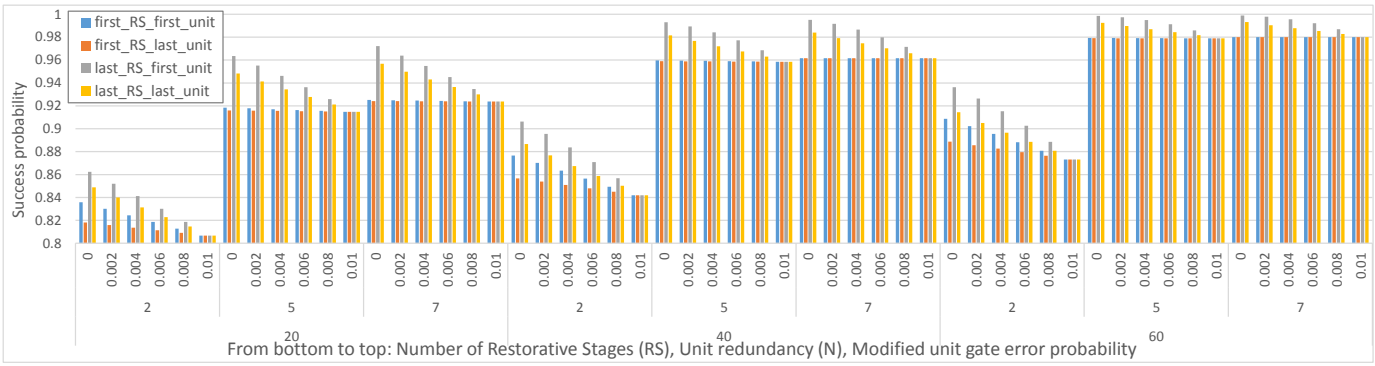


Figure 2: Multiplexing unit reliability contribution in a highly faulty system.

cause. The heterogeneous integration allows for the creation of well-defined clusters of gates, each cluster having the gate error rate in the same reliability range, whereas process variability or SET background charge affects all devices in an uncontrollable way, and only the fault probability range of all gates can be determined. Choosing which gates have what reliability can be used to tweak reliability-performance-cost trade-offs, as presented in the next sections.

III. Reliability in the Multiplexing Chain

According to [12] the probability $P_l(k_l)$ of a multiplexing unit l to have k_l outputs equal with "1" (stimulated) is given by the following recursion:

$$P_l(k_l) = \sum_{k_{l-1}=0}^{N} P_l(k_l|k_{l-1}) P_{l-1}(k_{l-1}), \tag{1}$$

$$P_l(k_l|k_{l-1}) = \binom{N}{k_l} (\overline{z}(k_{l-1}))^{k_l} (1-\overline{z}(k_{l-1}))^{N-k_l}, \tag{2}$$

$$\overline{z}(k_{l-1}) = (1-\epsilon) - (1-2\epsilon)\left(\frac{k_{l-1}}{N}\right)^2, \tag{3}$$

where $P_l(k_l|k_{l-1})$ is the probability in unit l to have k_l stimulated outputs, when k_l inputs are stimulated. If all the gates in the system fail with the same probability ϵ, then the probability transition matrix with elements $P(k_l|k_{l-1})$ is the same for all multiplexing units. However, if each multiplexing unit can have a different failing rate ϵ_l for its gates, as it is the case when they are implemented with different technologies, each unit output probability conditioned by the unit is given by:

$$P_l(k_l|k_{l-1}) = \binom{N}{k_l} (\overline{z}_l(k_{l-1}))^{k_l} (1-\overline{z}_l(k_{l-1}))^{N-k_l}, \tag{4}$$

$$\overline{z}_l(k_{l-1}) = (1-\epsilon_l) - (1-2\epsilon_l)\left(\frac{k_{l-1}}{N}\right)^2. \tag{5}$$

To qualitatively assess the impact of each multiplexing unit on the scheme reliability from Equations (1), (4) and (5) is a difficult task. Instead, throughout our analyses we employ probabilistic model checking to quantitatively evaluate the reliability of a specified NAND multiplexing configuration. Probabilistic model checking is an automatic procedure for establishing if a desired property holds in a probabilistic system model. A NAND multiplexing system can be modelled as a discrete-time Markov chain, and the property checked is whether the number of correct outputs are within the value range determined by the threshold Δ, i.e., the success probability. We use PRISM [19], a probabilistic model checker that also provides quantitative results, and a modified version of the multiplexing model from [15] that supports specifying individual gate errors in each stage. We keep the same values for the stimulated probability and the output correctness threshold (0.9 and 0.1, respectively), but we also added tests for all three combinations of the input vector: *00*, *11*, and *01* (the same with *10* since the inputs are interchangeable).

The multiplexing architecture consists of one executive stage followed by a serialized chain with an arbitrary number of restorative stages (RS). It is known that the purpose of adding an RS is to obtain a lower error rate at the output than at the input. Thus, it is expected that an RS with more reliable gates, i.e., implemented with gates with a smaller error probability, is more effective in the restoring mechanism. However, it is not clear if a less reliable RS is less effective, neutral, or even detrimental to the restoring mechanism. Furthermore, since the restorative part can be formed from several RSs, it is also of interest to know the difference each of them can make in the overall system reliability.

To find out how much the error rate of a multiplexing unit affects the NAND multiplexing reliability, we decrease, in turns, the error rate of both multiplexing units in the first and the last RS of the restoration chain. The experiments were performed for all input vector combinations, a default error gate of 0.01, different redundancy factors, with the first and the second multiplexing unit from the first RS in the chain, and from the last one, respectively, being more reliable (having a lower error probability). Figure 2 presents the minimum success probability obtained (higher is better) for all vector combination. In every bundle, the last set of four bars (with error probability 0.01, thus unmodified) has the same values, and is the reference success probability for the multiplexing scheme with that redundancy configuration. Furthermore, the first set in each bundle represents the maximum improvement that can be obtained with a perfect multiplexing unit (modified

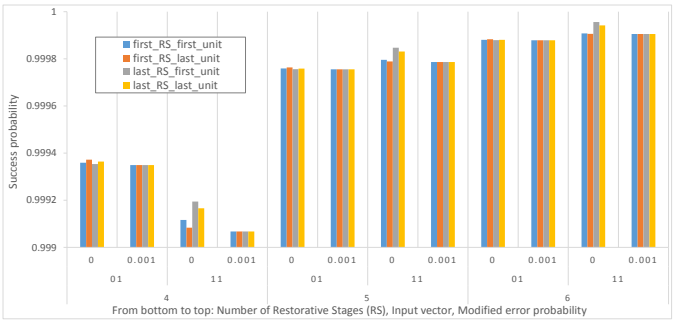


Figure 3: Multiplexing unit reliability contribution for different restorative stages (N=40).

error probability is 0).

It can be observed that for high error rates (in the order of 10^{-2}) it is always better to have a higher reliability in the last RS, since the last two bars in a 4-bars set are always higher than the first two bars. Otherwise, any improvement given by an RS is approximately halved by each subsequent less reliable RS, e.g., when having 7 RSs, any improvement in the first RS has an unnoticeable effect on the system reliability. This means that heterogeneous RSs should be arranged in an ascending order of reliability. Moreover, the reliability improvement when using a more reliable last RS is quite high for low redundancy factors, even exceeding the one obtained by adding more redundancy to the scheme, thus it is possible to replace a large amount of redundancy with just a more reliable multiplexing unit. Inside the RS, the improvement is higher when the first multiplexing unit has a higher reliability than the second one.

All the results in Figure 2 correspond to input vector *11*. However, for lower error rates, as the number of stages is increasing the potential to correct inputs *11* is higher than the one to correct inputs *10*, as can be observed in Figure 3, when going from 4 restorative stages to 5. Input vector *01* becomes the worst-case combination, and thus it dictates also the potential contribution of any multiplexing unit in the scheme.

A detailed view on the individual contributions of multiplexing units, including the executive stage, for all input vector combinations in the case of a default gate error of 0.001, is presented in Figure 4. First, we observe that the system reliability values for input combination *00* are two orders of magnitude higher than for the other combination. The reason for this is that vector *00* is the easiest to correct: (i) if one of the NAND gate input is faulty and has a 1 instead of a 0, a correct NAND gate still outputs the correct 1 value, (ii) the same phenomenon happens also for the *01* input vector, but the probability to have gates with both inputs incorrect is considerably lower in the case of *00* input. In conclusion, in contrast with the other input combination, the majority of faulty inputs in the input vector *00* are always corrected by the executive stage, and thus it is not limiting the system reliability.

The second thing that we can observe in Figure 4 is the inversion in reliability contribution of multiplexing units between input vectors *01* and *11*, in the position of the RS in which the multiplexing unit is, as well as in the position of the multiplexing unit inside the RS. This is a consequence of the inversion of the correct value that needs to output from 1 to 0. Thus, contrary to the case for the *11* input vector, heterogeneous RSs should be arranged in a descending reliability order, starting with the executive stage. The behaviour of the system reliability in case of a more reliable executive unit simply reinforces the aforementioned conclusions: for *11* input vector the improvement in the executive stage is masked by the subsequent faultier RSs, while for the *01* input vector the improvement is accentuated by the earliest position of the executive stage in the multiplexing chain.

IV. 3D-Stacked Heterogeneous Technology NAND Multiplexing

The previous section indicated that the position of a more reliable multiplexing unit in a heterogeneous reliability multiplexing chain is of importance in the system reliability. To evaluate the impact that different 3D-SIC arrangements can have on the system reliability we use the well-known Reliability Improvement Index (RII), defined as [20]:

$$RII = \frac{\log(\text{gate reliability})}{\log(\text{system reliability})}. \quad (6)$$

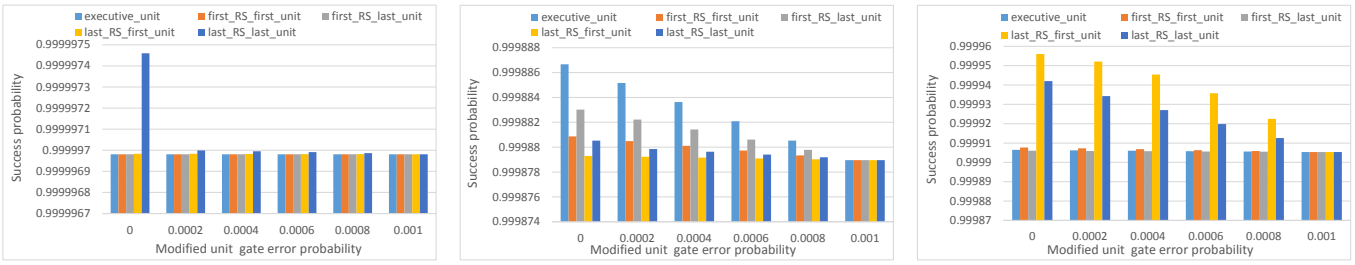
When the system is composed of gates with different reliability, the multiplexing system reliability should be higher than the reliability of the best gate in order to make sense to apply multiplexing in the first place, otherwise we can simply use one instance of the more reliable gate. Thus, the RII in a reliability heterogeneous system gives a measure of the improvement potential relative to the best gate in the system:

$$RII = \frac{\log(1 - \text{minimum gate error probability})}{\log(\text{success probability})}. \quad (7)$$

A positive reliability improvement is obtained only when $RII > 1$. Having sub-unitary RII values means that the system reliability is worse than the most reliable part in the system.

Based on the quantitative analysis of the impact on system reliability of individual multiplexing units from Section III, we propose the following generic two-tier 3D-SIC heterogeneous technology arrangements:

1) *3D-NM-first*: the executive stage and a number of consecutive multiplexing units from the beginning of the restorative chain are placed on the high reliability tier, and the rest of the units are placed on a low reliability tier,
2) *3D-NM-last*: the executive stage and a number of consecutive multiplexing units from the end of the restorative chain are placed on the high reliability tier, and the rest of the units are placed on a low reliability tier,
3) *3D-NM-first-last*: the executive stage and an equal number of consecutive multiplexing units from the beginning and the end of the restorative chain are placed on the high



(a) Input vector *00*. **(b) Input vector *01* or *10*.** **(c) Input vector *11*.**

Figure 4: Multiplexing unit reliability contribution for different input vectors (N=40, 6 restorative stages).

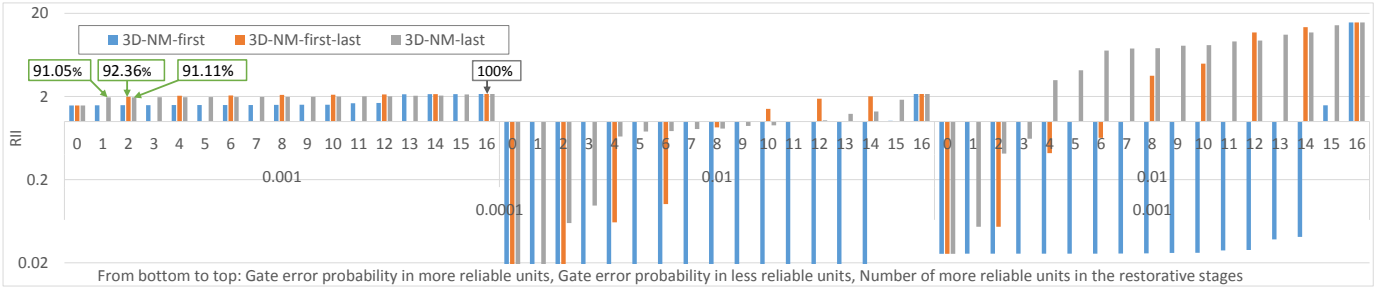


Figure 5: Reliability improvement of 3D-Stacked NAND Multiplexing for high error rates (N=40, 8 restorative stages).

reliability tier, and the rest of the units are placed on a low reliability tier.

Even though it makes sense to also take advantage of the difference in reliability impact between multiplexing units inside a RS and place them on separate tiers, we do not consider such an approach as it increases the number of inter-tier connection. Given the fact that the minimum TSVs dimensions are several orders of magnitude higher than the normal inter-metal vias, we keep the TSV density reduced by keeping together on the same tier consecutive multiplexing units, and make use only of the difference in reliability impact between RSs.

To perform a fair comparison we plot the RII of the three schemes varying the number of multiplexing units placed on the most reliable die, with the stages equally divided between the start and the end of the restorative chain for *3D-NM-first-last*. Figure 5 presents the results for high gate error rates combinations (10^{-3} with 10^{-2}, 10^{-4} with 10^{-2}, and 10^{-4} with 10^{-3}). When the worse gate error rate is close to the theoretical limit (10^{-2} in the middle and the right part of the figure), their effect of destroying any improvement previously obtained with more reliable units is devastating. This explains the very low RII when only the executive stage is implemented on the more reliable tier (0 more reliable units in the RSs), and the small RII increase for the *3D-NM-first*. Even in the arrangements that have higher reliability gates at the end of the restorative chain, i.e., *3D-NM-first-last* and *3D-NM-last*, too many multiplexing units need to be of high reliability in order to achieve an improvement comparable with the maximum RII, corresponding to a planar homogeneous system (all 16 units being more reliable).

When having available technologies with error rates of 10^{-4} and 10^{-3} (left part of the figure), it is enough to have two reliable multiplexing units (one in the executive stage and the last one in the restorative chain) out of the total 17 multiplexing units, in order to obtain 91.05% from the maximum RII. Thus, the trade-off for *3D-NM-last* can be viewed as follows: lose 9% in RII, but gain in speed and area since there are $15 \cdot 100/17 = 88.23\%$ multiplexing units faster and smaller. The actual gains depend on the exact speed and area differences between the two technologies. For example, if we assume that the delay and area values of a gate double for an order of magnitude more reliable technology, the gain in speed and area is $1.79\times$. The loss in RII is actually small when compared with the 90% RII reduction corresponding to the case when the entire system is using gates with 10^{-4} error rate. Even though the last units in the restorative chain offers the highest reliability gain, the *3D-NM-first-last* solution proves to be even better, achieving 92.36% from the maximum RII, for one reliable unit at the beginning and one at the end. However, since now only $13 \cdot 100/17 = 76.47\%$ multiplexing units are faster and smaller, the potential to reduce cost and area under the same assumption is only $1.61\times$.

The previously observed trend holds true also in Figure 6, when the more reliable gates have a low error rate, in the order of 10^{-7}. Furthermore, as the errors in the presence of input vector *01* become the hardest to correct, the RII improvement offered by *3D-NM-first-last* is also growing (the right part of the figure). When the reliability of the faultiest gates in the system is increased to an error rate in the order of 10^{-4}, having only the executive stage in a more reliable technology achieves 99.0153% of the maximum RII. Essentially, this means that for NAND multiplexing with gates having low error rates it is possible to achieve almost the same degree of

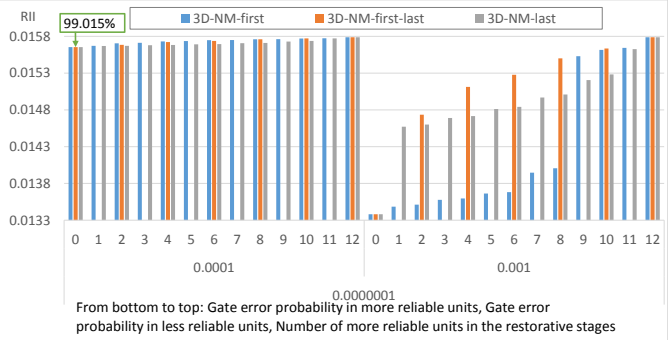


Figure 6: Reliability improvement for low error rates (N=60, 6 restorative stages).

reliability if all the RSs are implemented using only medium reliability gates. Considering the same assumption that the gate delay and area values double for a technology with an order of magnitude more reliable gates, the delay and area of such a 3D-SIC multiplexing architecture are reduced with a factor of $5.66\times$. For 3D-SIC designs, the footprint is a more relevant metric than area, thus the footprint improvement of the considered heterogeneous 2-tier 3D-SIC multiplexing over the homogeneous 2-tier stack with almost the same redundancy is $4.25\times$.

Note that the RII is sub-unitary, so the redundancy has to be extremely high to increase the system reliability over the reliability of gates with an error rate of 10^{-7}. Unfortunately, the simulation of a scheme with an enough redundancy factor is memory and time prohibited. However, simulations with different redundancy factors like the ones in Figure 2 confirm that the trend does not change when increasing redundancy, hence the potential to greatly enhance the reliability-performance-cost trade-off is still present. If needed, the RII can further be pushed higher by using the *3D-NM-first* arrangement. The *3D-NM-first-last* is a compromise between technologies with high and low reliability, and for example it can be used to account for ageing effects that increase the fault rate of devices over time.

V. Conclusions

In this paper, we proposed to take advantage of a new dimension introduced by 3D-Stacking Integrated Circuits (3D-SIC) technology in the design space, i.e., the individual reliability of devices, to improve the reliability-delay-cost trade-offs of the NAND Multiplexing redundancy technique. We first studied the contribution of each unit of the multiplexing architecture to the total system reliability, in order to place important units that have the largest reliability contribution on tiers manufactured in a more fault-tolerant technology. Our analysis determined that multiplexing units importance is determined by how high the error rate of individual gates in the system is, i.e., for high error rates the units at the end of the restoration chain are more important, while for low error rates the units at the beginning of the restoration chain are critical. Based on this, we introduced the first heterogeneous 3D-SIC multiplexing arrangements. Evaluation results indicated that if we consider that delay and area are doubled for a technology with an order of magnitude higher reliability, a $1.79\times$ delay and area reduction can be obtained with only 9% loss in the Reliability Improvement Index (RII). For medium and low error rates, a minimum 1% RII loss in can be traded for delay and footprint reductions of $5.66\times$ and $4.25\times$, respectively.

References

[1] "ITRS - Process Integration, Devices & Structures, 2012 Update," International Technology Roadmap for Semiconductors, Tech. Rep., 2012. [Online]. Available: http://www.itrs.net

[2] "ITRS - Emerging Research Devices, 2012 Update," International Technology Roadmap for Semiconductors, Tech. Rep., 2012. [Online]. Available: http://www.itrs.net

[3] W. H. Pierce, *Failure-tolerant computer design*. New York: Academic Press, 1965.

[4] R. E. Lyons and W. Vanderkulk, "The use of triple-modular redundancy to improve computer reliability," *IBM Journal of Research and Development*, vol. 6, no. 2, pp. 200–209, 1962.

[5] S. Safiruddin, D. Borodin, M. Lefter, G. Voicu, and S. D. Cotofana, "Is 3D integration the way to future dependable computing platforms?" in *13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*, 2012, pp. 1233–1242.

[6] J. von Neumann, "Lecture notes on: Probabilistic logics and synthesis of reliable organisms from unreliable components," in *Automata Studies*, C. Shannon and J. McCarthy, Eds. Caltech - Pasadena, CA, Jan. 4-15, 1952: Princeton University Press, 1956, pp. 43–98.

[7] W. S. Evans and L. J. Schulman, "On the maximum tolerable noise of k-input gates for reliable computation by formulas," *IEEE Transactions on Information Theory*, vol. 49, no. 11, pp. 3094–3098, 2003.

[8] A. S. Sadek, K. Nikoli, and M. Forshaw, "Parallel information and computation with restitution for noise-tolerant nanoscale logic networks," *Nanotechnology*, vol. 15, no. 1, pp. 192–210, Jan. 2004.

[9] S. Roy and V. Beiu, "Majority multiplexing - economical redundant fault-tolerant designs for nanoarchitectures," *IEEE Transactions On Nanotechnology*, vol. 4, no. 4, pp. 441–451, Jul. 2005.

[10] V. Beiu, W. Ibrahim, Y. A. Alkhawwar, and M. H. Sulieman, "Gate failures effectively shape multiplexing," in *21st IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems, DFT'06.*, 2006, pp. 29–40.

[11] K. Nikolic, A. Sadek, and M. Forshaw, "Fault-tolerant techniques for nanocomputers," *Nanotechnology*, vol. 13, no. 3, p. 357, 2002.

[12] J. Han and P. Jonker, "A system architecture solution for unreliable nanoelectronic devices," *IEEE Transactions On Nanotechnology*, vol. 1, no. 4, pp. 201–208, Dec. 2002.

[13] G. Roelke, R. Baldwin, and D. Bulutoglu, "Analytical models for the performance of von Neumann multiplexing," *IEEE Transactions On Nanotechnology*, vol. 6, no. 1, pp. 75–89, Jan. 2007.

[14] S. Krishnaswamy, G. F. Viamontes, I. L. Markov, and J. P. Hayes, "Accurate reliability evaluation and enhancement via probabilistic transfer matrices," in *Design, Automation and Test in Europe, 2005. Proceedings*, 2005, pp. 282–287.

[15] G. Norman, D. Parker, M. Kwiatkowska, and S. Shukla, "Evaluating the reliability of NAND multiplexing with PRISM," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 24, no. 10, pp. 1629–1637, 2005.

[16] D. Bhaduri, S. Shukla, P. Graham, and M. Gokhale, "Comparing reliability-redundancy tradeoffs for two von Neumann multiplexing architectures," *IEEE Transactions On Nanotechnology*, vol. 6, no. 3, pp. 265–279, May 2007.

[17] F. Martorell, S. D. Cotofana, and A. Rubio, "Fault tolerant structures for nanoscale gates," in *7th IEEE Conference on Nanotechnology, IEEE-NANO*, 2007, pp. 605–610.

[18] V. Beiu and W. Ibrahim, "Devices and input vectors are shaping von Neumann multiplexing," *IEEE Transactions on Nanotechnology*, vol. 10, no. 3, pp. 606–616, May 2011.

[19] M. Kwiatkowska, G. Norman, and D. Parker, "PRISM 4.0: Verification of probabilistic real-time systems," in *Computer Aided Verification*, 2011, pp. 585–591.

[20] T. Klaschka, "Method for redundancy scheme performance assessment," *IEEE Transactions on Computers*, vol. C-20, no. 11, pp. 1371–1376, Nov. 1971.