A Die Selection and Matching Method with Two Stages for Yield Enhancement of 3-D Memories

Wooheon Kang¹, Changwook Lee¹,², Keewon Cho¹, and Sungho Kang¹
¹Department of Electrical Engineering Yonsei University, Seoul, Korea
²Probe Test Engineering Team, SK Hynix, Inc., Icheon-si, Gyeonggi-do, Korea
E-mail: sudal@soc.yonsei.ac.kr; changwooklee@soc.yonsei.ac.kr; ckw1505@soc.yonsei.ac.kr; shkang@yonsei.ac.kr

Abstract—Three-dimensional (3-D) memories using through-silicon-vias (TSVs) as vertical buses across memory layers has regarded as one of 3-D integrated circuits (ICs) technology. The memory dies to stack together in a 3-D memory are selected by a die selection method. In order to improve yield of 3-D memories, redundancy sharing between inter-die using TSVs is an effective strategy. With the redundancy sharing strategy, the bad memory dies can become good 3-D memories through matching the good memory dies. To support die selection and matching efficiently, a novel redundancy analysis (RA) algorithm, which considers various repair solutions, is proposed. Because the repair solutions can be various, the proposed die selection and matching is performed with two stages; general die selection and matching in the first stage and re-matched remained memory dies, after the first stage, applying other repair solutions in the second stage. Thus, the proposed die selection and matching algorithm using the proposed RA algorithm can improve yield of 3-D memories. The experimental results show that the proposed die selection and matching method can achieve higher yield of 3-D memories than that of the previous state-of-the-art the die selection and matching methods.

Index Terms—Yield improvement, 3-D integrated circuit (IC), 3-D random access memory (RAM), memory repair, through-silicon-via (TSV)

I. INTRODUCTION

Three-dimensional (3-D) integration technology using through-silicon-vias (TSVs) has regarded as one of the integrated-circuits (ICs) design technologies [1-3]. There are many benefits of the TSV technology including high density, high band-width, low power consumption, and small form factor [3]. However, this technology faces other challenges, such as technological challenges, test and yield challenges, thermal and power challenges, and infrastructure challenges, which should be overcome.

Homogenous memories are good for 3-D IC technology. In [4], for example, DDR3 DRAM using TSV technology is proposed. To produce good 3-D memories, 3-D memory is typically composed of known-good-dies (KGDs), which are repaired with self-contained spare memories. Because the capacity and density of two-dimensional memories have increased, the probability of memory faults has increased [5]. In addition, the yield of 3-D memories is affected by yield of 2-D memories. Therefore, redundancy memories with redundancy analysis (RA) algorithm are used for improving yields. For example, a repairable 2-D memory has spare rows and spare columns. If the 2-D memory has faulty cells, then the redundancies are replaced these faulty cells by RA algorithm. If all the faulty cells can be replaced by the redundancies, then the memory is repaired. Otherwise, the memory is not repairable.

For the repairable memories, the unused redundancies may remain. If the redundancies in a memory die are conventionally designed for 2-D memories, these remained redundancies are wasted. However, these remained redundancies after the memory repair are used to other memories which require additional redundancies to repair themselves. To improve the yield of 3-D memories, many researches on die selection and matching algorithm using inter-die sharing redundancy have been studied [6-10]. With the redundancy sharing strategy using TSVs, a memory that is not repairable using self-contained redundancies can use redundancy resources from its vertical memory die and it can be repaired. Thus, the yield of 3-D memories can be improved because bad memory dies become good 3-D memories.

To support redundancy sharing strategy efficiently, there are two crucial points; an RA algorithm and a die selection and matching algorithm. The RA algorithm is used to find out the number of remained and required redundancies as well as repairable of memories and repair solutions. In [8], the die selection and matching algorithm uses SFCC RA algorithm [11] due to its optimal repair rate. SFCC RA algorithm uses a simple fail count comparison for building a search tree based on line faults. The die selection and matching algorithm in [9-10] uses LRM RA algorithm [12] due to its high repair efficiency and low area overhead. LRM RA algorithm is greedy algorithm that it allocates spare memory to the faulty line has the most faulty cells. However, conventional RA algorithms cannot be suitable for redundancy sharing strategy, because the results of RA can be various. In addition, for the case of the memory is not repairable, it is difficult that the required redundancies cannot be precisely determined. Because the die selection and matching is based on results of RA, the yield of 3-D memories is affected by both the RA algorithms and the die selection and matching strategies.
In this paper, a new die selection and matching method with two stages is proposed. To support the proposed method, a new RA algorithm is also proposed. Since the repair solutions can be various, the proposed RA algorithm holds other repair solutions for two stages of the proposed die selection and matching method. The proposed die selection and matching method consists of two stages. In the first stage, the good memories and bad memories are matched to become good 3-D memories. In the second stage, if there are other repair solutions for the remained good memories and bad memories after the first stage, die selection and matching procedure is again performed with other repair solutions. Thus, the proposed method can enhance the 3-D memory yield.

II. BACKGROUND

A. Stacking Method and Flow of 3-D ICs

There are three integration technologies used in 3-D IC fabrication including 3-D random access memory (RAM): wafer-to-wafer (W2W), die-to-wafer (D2W), and die-to-die (D2D) [7-10, 13-14]. Wafers with the same size of dies are directly bonded together in the W2W integration method. The W2W offers the highest throughput, the thinnest wafers, and high TSV density. However, the W2W can incur a low yield because the bad die can be stacked to another good die, which results in a bad 3-D IC. On the other hand, the D2W and D2D integration methods allow the use of different wafer and die sizes because the dies are bonded after dicing. The D2W and D2D can improve the yield of 3-D ICs due to its high flexibility that the dies are deiced and tested in advance and only the good dies are stacked. However, these methods require a more complex manufacturing process that the dies must be aligned and integrated each other. This results in low throughput and low TSV density. In this paper, the D2D integration method is used to produce 3-D memory stacking.

Fig. 1 shows the stacking flow of 3-D RAMs based on the D2D integration method. The memory dies classified into three types according to the status after pre-bond test and repair: good die, inter-repairable die, and unrepairable die. A good die is fault free die or self-repairable die. The self-repairable die is that the faulty memory die can be repaired using self-contained redundancies. An inter-repairable die is that the faulty memory die cannot be repaired using self-contained redundancies, but can be repaired using sharing inter-die redundancies. An unrepairable die is that the faulty memory die cannot be repaired despite of sharing inter-die redundancies. The dies, which are good dies and inter-repairable dies, are selected to stack 3-D memories. The yield of 3-D memories is influenced by the die selection method. The inter-repairable dies have a chance to form a good 3-D memory because of inter-die sharing redundancy. After stacking, post-bond test and repair is performed to check whether the 3-D memory is good or bad.

B. Inter-die Sharing Redundancy Scheme for 3-D Memory

Fig. 2 shows inter-die sharing redundancy scheme for 3-D memory. In order to improve yield of 3-D memories, the inter-die sharing redundancy architecture is proposed many studies [6-10]. Two separate memory dies can use both the self-contained redundancy memory and inter-die redundancy memory. The interconnections between two memory dies are possible using TSVs. To support this architecture, both programmable decoders and multiplexers are used [6]. The routing overhead support to this strategy is quite low due to the use of short TSVs as routing mechanism. Assume that the one of the spare rows in the memory die 1 is allocated to the memory die 2. In this case, all the spare rows in the memory die 2 are allocated to their layer. The decoder in the memory die 2 inserts high signals to the self-contained spare rows and one of the interconnected to TSV in order to use the spare row in the memory die 1. Thus, the allocated spare row in the memory die 1 to the memory die 2 is controlled by decoder in the memory die 2 through the TSV.

III. PROPOSED YIELD ENHANCEMENT TECHNIQUES

A. Redundancy Analysis for inter-die sharing redundancy

In order to classify the memories under test in the pre-bond test and repair phase, a new RA algorithm for inter-die sharing...
...point of view of column addresses. The repair solution candidates for the column faults are derived in the same manner. Consequently, the repair solution candidates contain all possible cases for a line fault group. After the repair solution candidates for each line fault group are decided, the final repair solution candidates are derived by checking all combinations of repair solution candidates for each repair solution candidates.

Fig. 3 shows an example of faulty memory with two spare rows (Sr) and two spare columns (Sc). There are two line fault groups. For the point of view of row addresses, there are two row line faults and one row single fault in the line fault group 1. The repair solution candidates are 2R for row line faults and 1R row single fault, and 2R for row line faults and 1C row single fault, where R is the spare row and C is the spare column. For the point of view of column addresses, there are two column line faults and one column single fault in the line fault group 1. The repair solution candidates are 2C for column line faults and 1R column single fault, and 2C for column line faults and 1C column single fault. The repair solution candidates for line fault group 2 are {4R, (3R, 1C), (2R, 2C), 3C}. Now, there are 16 final repair candidates for faulty memory in Fig. 3. However, some repair solution candidates, such as 7R, (6R, 1C), (1R, 5C), and 6C, are invalid because the required number of spare rows or columns is greater than the available spare rows or columns despite of inter-die sharing redundancies. In addition, some repair solution candidates, such as (4R, 3C) and (3R, 4C), are repeated, so these repeated repair solution candidates are remained by ones. Thus, the final repair candidates are {(3R, 3C) and (2R, 4C)}. The repair candidates, such as (4R, 3C), and (3R, 4C), are excluded, because these spares do not have the minimum number of spares.

B. Die selection and matching method

An efficient die selection and matching method is proposed to improve yield of 3-D memories using inter-die sharing redundancy scheme. The proposed die selection and matching method consists of two stages. In order to remain memory dies which have various repair solutions if possible, the memory dies, which have only one repair solution, are preferentially selected and matched. In advance, repairable and inter-repairable memory dies are classified into two groups. The one group consists of memory dies, which have various repair solutions if possible, the memory dies, which have only one repair solution, are preferentially selected and matched. In advance, repairable and inter-repairable memory dies are classified into two groups. The one group consists of memory dies, which have various repair solutions, such as [1, 0, (0)] shown in Fig. 4. This means that there is one spare row is remained. The number zero in the round brackets means that this memory die has only one repair solution. Another group consists of memory dies, which have various repair solutions, such as [0, -1, (1)] shown in Fig. 4. This means that one spare column is required to become a good 3-D memory. The number one in the round brackets means that this memory die has various repair solutions. Because a number of repair solutions can be exist, the number of various repair solutions is limited on three. If the number of repair solutions is bigger, the calculation becomes complex and time consuming. If the repair solution is nR and nC, and 2n is the smallest cost, where R is row and C is column, there can be three repair
solutions; \((n-i)R\) and \((n+i)C\), \(nR\) and \(nC\), and \((n+i)R\) and \((n-i)C\).

If the repair solution is \(nR\) and \(mC\), then there can be two repair solutions; \(nR\) and \(mC\), and \(mR\) and \(nC\). Thus, the maximum number of repair solutions is three.

In the first stage, any conventional die selection and matching method for a stacked 3-D memory can be used. The memory dies, which have only one repair solution, are preferentially selected and matched. And then, the memory dies, which have various repair solutions, are selected and matched. In this paper, the die selection and matching is performed by comparing the number of available spare memories among good dies to the number of required spare memories among inter-repairable dies. If the number of available spare memories of good die is greater than or equal to the number of required spare memories among inter-repairable dies, the good die and the inter-repairable die can be matched. In the second stage, if there are not matched due to their spare memory conditions, the proposed method resets the information on the available spare memories and the required spare memories since other repair solutions can be existed.

Assume that the number of stacked dies is 2 for the simplicity. There are 48 memory dies on two wafers, as shown in Fig. 4. After per-bond test and repair using proposed RA algorithm, the 48 memory dies are classified with status of spare memory as shown in Fig. 5(a). Because the unrepairable dies cannot be stacked, six unrepairable dies are not illustrated. The Die\((n, m)\) is that the memory die has available \(n\) spare rows and \(m\) spare columns, and the Die\((-n, -m)\) is that the memory die requires \(n\) spare rows and \(m\) spare columns to repair this memory die. In the first phase of proposed die selection and matching method, the die selection and matching is performed; the six 3-D RAMs are produced from 13 Die\((0, 0)\) dies, the two Die\((1, 0)\) dies can be stacked the two Die\((-1, 0)\) dies, the two Die\((1, 1)\) dies can be stacked the two Die\((-1, -1)\) dies, the one Die\((2, 0)\) die can be stacked the one Die\((-2, 0)\) die, the four Die\((1, 2)\) dies can be stacked the four Die\((-1, -2)\) dies, and the three Die\((2, 1)\) dies can be stacked the three Die\((-2, -1)\) dies.

After the first phase, one Die\((0, 0)\), one Die\((1, 0)\) die, two Die\((1, 1)\) dies, one Die\((-1, -1)\) die, and one Die\((-2, 0)\) die are remained as shown in Fig. 5(b). Since the Die\((1, 0)\) die and Die\((1, 1)\) dies are good dies, they can matched each other, and 20 3-D RAMs are finally produced if other die selection and matching method is used. However, if the two Die\((1, 1)\) have other repair solution and these are converted to one Die\((2, 0)\) and one Die\((0, 2)\), the remained one Die\((-1, -1)\) die and one Die\((-2, 0)\) are matched and produced as good 3-D memories as shown in Fig. 5(c). Also, the remained one Die\((0, 0)\) die and one Die\((1,0)\) die are matched. Therefore, 21 3-D RAMs are finally produced using the proposed method.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

To evaluate the performance of the proposed die selection and matching method, we implemented a simulator in C-language. The simulator performs memory test and repair using RA algorithms, such as LRM, SFCC, and the proposed RA algorithm, and estimates the final yield of 3-D memories according to the die selection and matching algorithms. The simulations was performed for 1,024×1,024 bit memory dies with nine redundancy configurations, which are combinations of one to three spare rows and one to three spare columns. To evaluate the die selection and matching methods in realistic situations, Polya-Eggenberger distribution is used. Polya-Eggenberger distribution is suitable for modeling integrated circuit yields [12, 15-17]. Fig. 6 shows the two
Because this method shares only spare columns, this method
yields for the without inter-die redundancy. For evenly
distributed faults, the 3-D memory yield and yield
improvement for clustered faults, the 3-D memory yields
for evenly distributed faults is larger than the
memory dies with the low number of faults for
increases compared to the without inter-die redundancy.

Selection and matching method. In Table I,
spare rows and three spare rows and one spare column are
shared, the yield improvement increases
when the number of spare columns is large. For evenly
distributed faults, the 3-D memory yield and yield
improvement are remarkable, when one spare row and three
spare columns are used comparing to two spare rows and two
spare columns and three spare rows and one spare column are
used, although the total number of spares is same.

In [8], the die selection and matching method uses SFCC
algorithm which has optimal repair rate. Because SFCC
analyzes single faults after analyzing line faults, the method in
[8] can consider single faults for the die selection and matching.
For 2-D redundancy sharing scheme without considering single
faulst, once the repair solutions are fixed, there can be bad cases
compared with sharing 1-D redundancy. If a self-repairable
memory die remains one spare row and an inter-repairable
memory die requires one spare column, then these memory dies
cannot become a good 3-D memory due to their fixed repair
solutions. However, if these faulty memories contain single faults,
these memory dies can become a good 3-D memory.

The proposed method can achieve the highest 3-D memory
yields due to its second stage. In the first stage, there are many
unmatched both the self-repairable memory dies and the
inter-repairable memory dies 2-D redundancy sharing because

Table I summarizes the simulation results of 3-D memory
yields without and with inter-die redundancies of the previous
dies selection and matching methods and the proposed dies
selection and matching method. In Table I, \( nR/mC \) denotes a
memory die with \( n \) spare rows and \( m \) spare columns. The yield
of 3-D memories with inter-die redundancy significantly
increases compared to the without inter-die redundancy. Because
the memory dies with the low number of faults for
evenly distributed faults is larger than the memory dies with the
clustered faults, the 3-D memory yields for evenly distributed
faults are higher than the 3-D memory yields for clustered
faults. The single faults are also more flexible than line faults for
repair solutions because a single fault is repaired by either a
spare row or a spare column. Thus, a 3-D memory yield
improvement for evenly distributed faults is greater than the
yield improvement for clustered faults.

Since the die selection and matching method in [10] uses
LRM algorithm, the repair rate is not optimal and this leads the
lowest 3-D memory yields for the without inter-die redundancy.
Because this method shares only spare columns, this method
adopts row first strategy RA. After the spare rows are exhausted,
the spare columns are used to repair faulty cells. Since only
spare columns are shared, the yield improvement increases
when the number of spare columns is large. For evenly
distributed faults, the 3-D memory yield and yield
improvement are remarkable, when one spare row and three
spare columns are used comparing to two spare rows and two
spare columns and three spare rows and one spare column are
used, although the total number of spares is same.

In [8], the die selection and matching method uses SFCC
algorithm which has optimal repair rate. Because SFCC
analyzes single faults after analyzing line faults, the method in
[8] can consider single faults for the die selection and matching.
For 2-D redundancy sharing scheme without considering single
faulst, once the repair solutions are fixed, there can be bad cases
compared with sharing 1-D redundancy. If a self-repairable
memory die remains one spare row and an inter-repairable
memory die requires one spare column, then these memory dies
cannot become a good 3-D memory due to their fixed repair
solutions. However, if these faulty memories contain single faults,
these memory dies can become a good 3-D memory.

The proposed method can achieve the highest 3-D memory
yields due to its second stage. In the first stage, there are many
unmatched both the self-repairable memory dies and the
inter-repairable memory dies 2-D redundancy sharing because

\[
\text{TABLE I} \quad \lambda=14 \text{ and } \alpha=2.382
\]

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
R/C & \text{without} & \text{with} & \text{without} & \text{with} & \text{without} \\
\hline
1R/1C & 2.92% & 5.84% & 3.02% & 6.04% & 3.02% & 6.04% \\
1R/2C & 6.68% & 13.36% & 6.70% & 13.40% & 6.70% & 13.40% \\
1R/3C & 12.80% & 25.60% & 12.80% & 25.62% & 12.80% & 25.62% \\
2R/1C & 6.16% & 12.34% & 6.70% & 13.40% & 6.70% & 13.40% \\
2R/2C & 15.34% & 24.12% & 16.00% & 24.34% & 16.00% & 25.32% \\
2R/3C & 24.26% & 40.20% & 25.04% & 40.76% & 25.04% & 41.96% \\
3R/1C & 11.38% & 22.70% & 12.72% & 24.04% & 12.72% & 25.44% \\
3R/2C & 18.84% & 37.68% & 20.84% & 41.70% & 20.84% & 41.70% \\
3R/3C & 32.86% & 51.30% & 36.16% & 59.52% & 36.16% & 61.30% \\
\hline
\end{array}
\]

\[
\text{TABLE II} \quad \lambda=14 \text{ and } \alpha=0.6232
\]

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
R/C & \text{without} & \text{with} & \text{without} & \text{with} & \text{without} \\
\hline
1R/1C & 19.86% & 32.08% & 20.00% & 32.24% & 20.00% & 32.24% \\
1R/2C & 28.32% & 44.22% & 28.32% & 44.52% & 28.32% & 44.64% \\
1R/3C & 35.42% & 55.04% & 35.42% & 55.50% & 35.42% & 55.50% \\
2R/1C & 27.84% & 38.32% & 28.32% & 44.54% & 28.32% & 44.54% \\
2R/2C & 35.18% & 49.96% & 35.64% & 53.62% & 35.64% & 54.32% \\
2R/3C & 41.54% & 59.42% & 42.02% & 70.66% & 42.02% & 71.62% \\
3R/1C & 34.74% & 43.94% & 35.64% & 55.18% & 35.64% & 55.18% \\
3R/2C & 40.78% & 54.68% & 41.84% & 71.12% & 41.84% & 71.20% \\
3R/3C & 46.46% & 62.44% & 47.58% & 84.02% & 47.58% & 84.12% \\
\hline
\end{array}
\]
of fixed repair solutions. In the second stage, these remained memory dies are matched again by applying another repair solution. This includes the considering single faults. In addition, this stage considers cases that there are more than two repair solutions without single faults. For the clustered faults, the yield of proposed method is 61.30% and yield of the method in [8] is 59.52% when three spare rows and three spare columns are used with inter-die redundancy. The proposed method improves on average a 1.75% and 0.35% of yield improvement for clustered and evenly distributed in comparison to the method in [8], respectively. Since the method in [8] allocates single faults to spare rows and/or spare columns, the repair solutions can be various. However, if there is no single fault, the repair solutions are fixed, the yield improvement can be low. On the other hand, the proposed method can apply various repair solutions even though there is no single fault. Therefore, there exists the yield improvement comparing to previous methods.

Table II shows the simulation results of 3-D memory yields without and with inter-die redundancies of the proposed die selection and matching method. There are three result categories with the inter-die redundancies; after the first stage, without the second stage, and after the second stage. If the second stage is not applied, then the final yields of 3-D memories are calculated after the first stage. After the first stage some yields are smaller than yields without inter-die redundancy, such as 30.34% when two spare rows and two spare columns for evenly distributed faults. The reason is that repairable memory dies are still remained to apply the second stage. For the without the second stage, such as conventional methods, the remained repairable memories are matched and produced 3-D memories if there are repairable memories after the first stage. On the other hand, for the proposed method, the remained repairable memories are examined that they have other repair solutions, and they are matched with other repair solutions in the second stage.

The proposed method improves on average a 20.83% and 5.06% yield improvement for clustered and evenly distributed in comparison to die selection and matching methods using fixed repair solutions, respectively. As shown in simulation results, therefore, the final yields of 3-D memories significantly increase using the proposed method.

V. CONCLUSION

An outstanding die selection and matching method with two stages for 3-D memory using inter-die sharing redundancy scheme is proposed. To support a new die selection and matching method with two stages, a new RA algorithm is also proposed. When the repair solution is fixed using 2-D redundancy sharing, the yield improvement can be low because the types of remained redundancies and required redundancies cannot be matched. On the other hand, since various repair solutions are required in the second stage, the proposed RA algorithm can find all repair solutions for each memory die without RA time penalty. The simulation results show that the proposed die selection and matching method can further improve 3-D memory yields. Therefore, the proposed die selection and matching method can achieve higher yield of 3-D memories than that of the previous state-of-the-art the die selection and matching methods.

VI. ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012R1A2A1A03006255).

REFERENCES