

ΟΛΟΚΛΗΡΩΣΗ ΚΑΙ ΕΛΕΓΧΟΣ ΣΕ 2.5 & 3 ΔΙΑΣΤΑΣΕΙΣ

Αλέξανδρος Οικονόμου – alexanco@ece.auth.gr

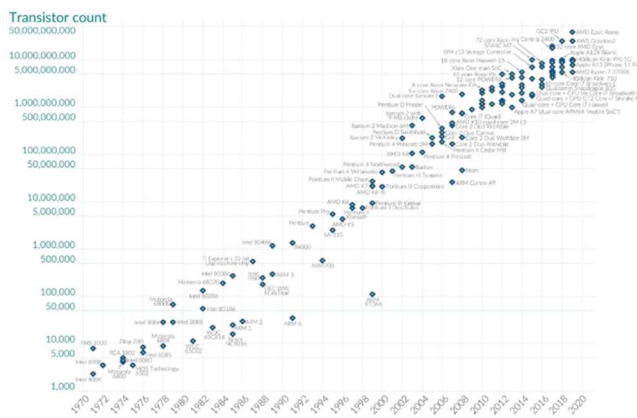
Κωνσταντίνος Χατζής – kachatzis@ece.auth.gr

I ΠΕΡΙΛΗΨΗ

Στη παρούσα αναφορά κάνουμε μια σύντομη περιγραφή στις μεθόδους ολοκλήρωσης των 2.5 και 3 διαστάσεων. Παραθέτουμε ιστορικά στοιχεία σχετικά με την εφεύρεση και τις απαρχές της μη-οριζόντιας ολοκλήρωσης, και τελικά αναφερόμαστε σε ορισμένες από τις γνωστότερες τεχνικές της.

II MORE MOORE

Το 1965 ο Gordon Moore δημοσιεύει την πρόγνωση του: ότι ο αριθμός των εξαρτημάτων ενός ολοκληρωμένου κυκλώματος πρόκειται να διπλασιάζεται κάθε δυο έτη. Μια δεκαετία αργότερα γίνεται η διαπίστωση ότι η πρόγνωση αυτή είναι ορθή, και έτσι καλείται πλέον “ο Νόμος του Moore”. Ο εμπειρικός κανόνας αυτός συνεχίζει να προσδιορίζει την εξέλιξη της τεχνολογίας έκτοτε.



Εικόνα II.1: Εξέλιξη του πλήθους των Transistor σε γνωστούς επεξεργαστές, 1970-2020.

Τα όργανα-οδηγοί στην εξέλιξη αυτή της τεχνολογίας συγκαταλέγονται σε δυο κατηγορίες [1]: την *Push* (ώθηση) και την *Pull* (υστέρηση). Στην πρώτη ανήκουν μέσα τα οποία ωθούν τη τεχνολογία και βοηθούν στην ανάπτυξή της. Σύμφωνα με τον Moore αυτά είναι:

- αύξηση του εμβαδού ολοκλήρωσης,
- μείωση μεγέθους χαρακτηριστικών και
- ευφυία σχεδίασης.

Οι οδηγοί της κατηγορίας *Pull* περιορίζονται σε οικονομικούς παράγοντες. Όπως έγραψε [2] το 1977 ο ο R. Noyce, συνεφευρέτης του Ολοκληρωμένου Κυκλώματος και συνιδρυτής της Intel, «περαιτέρω σμίκρυνση των Transistor είναι λιγότερο πιθανό να περιορίζεται από τους νόμους της φυσικής, αλλά πιθανότερο να περιορίζεται απ' την οικονομία».

Στην *Εικόνα II.1* παρατηρείται ότι αυτή η τροχιά συνεχίζει να υφίσταται στο έτος 2020. Παρ' όλα αυτά, στο τρέχον μέγεθος Transistor που χρησιμοποιείται για την παραγωγή ολοκληρωμένων κυκλωμάτων μεγάλης κλίμακας (~7nm) η φυσική έχει καθοριστικό ρόλο. Χρησιμοποιώντας πυρίτιο είναι αδύνατη η παραγωγή Transistor, η πύλη των οποίων είναι μικρότερη των 7nm (βλ. Quantum Tunneling). Το όριο αυτό θέτει ένα φυσικό περιορισμό στην εξέλιξη του Νόμου του Moore. Έτσι, όπως άλλωστε υποστήριξε και ο ίδιος, νέες μέθοδοι οφείλουν να βρεθούν ώστε να επιτευχθεί ο πολλαπλασιασμός του πλήθους των Transistor ανά μονάδα εμβαδού.

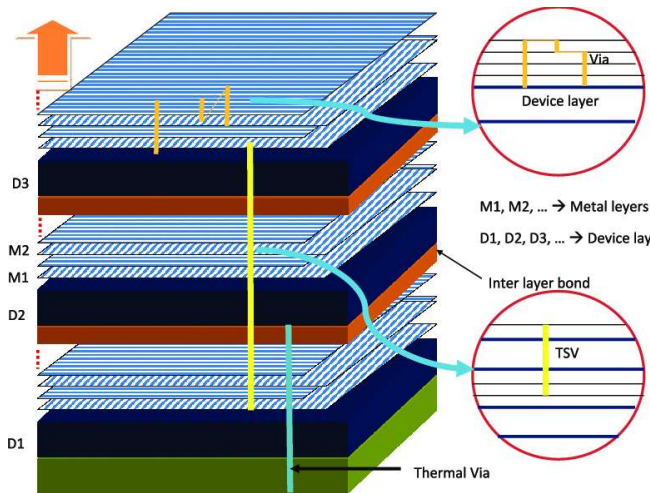
III ΠΕΡΙΣΣΟΤΕΡΕΣ ΔΙΑΣΤΑΣΕΙΣ

Σε αντίθεση με την ολοκλήρωση σε δυο διαστάσεις, όπου τα εξαρτήματα του ολοκληρωμένου τοποθετούνται στο ίδιο επίπεδο πυριτίου, με συνδέσεις να γίνονται σε

επόμενα επίπεδα μετάλλου, η χρήση περισσότερων διαστάσεων έρχεται να τοποθετήσει πολλαπλά επίπεδα εξαρτημάτων κάθετα, με ενδιάμεσα επίπεδα μετάλλου και κάθετες συνδέσεις μεταξύ των μετάλλων αυτών.

Τα δυο κυριότερα πλεονεκτήματα [3] της ολοκλήρωσης σε τρεις διαστάσεις, σε αντίθεση με τις δυο, είναι:

- η σμίκρυνση των συνδέσεων μεταξύ διαφορετικών ολοκληρωμένων (ICs), επομένως και η μείωση του συνολικού εμβαδού ολοκλήρωσης, καθώς και
- η βελτίωση του χρόνου καθυστέρησης στην επικοινωνία ολοκληρωμένων, λόγω του μικρότερου μήκους διασυνδέσεων μεταξύ τους.



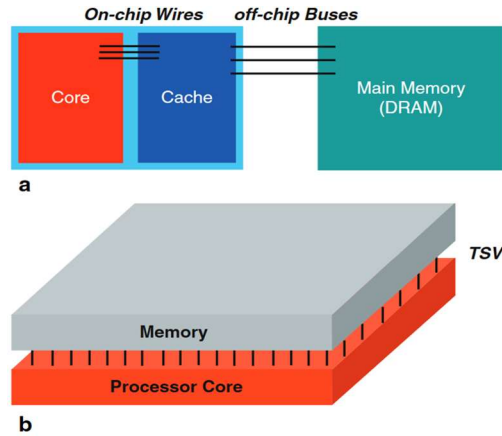
Εικόνα III.1: Δομή τρισδιάστατης μονολιθικής ολοκλήρωσης.

Ως επακόλουθο των τελευταίων, δίνεται πλέον η δυνατότητα σε ένα σήμα να διαμοιράζεται οριζόντια αλλά και κάθετα, επιτυγχάνοντας έτσι καλύτερο παραλληλισμό. Στην Εικόνα III.1 παρουσιάζεται μια τέτοια δομή, όπου εξαρτήματα στα επίπεδα D1, D2 και D3, το καθένα με 4 μέταλλα, ενώνονται σε ένα κοινό θερμικά-αγωγίμο επίπεδο και αποτελούν ένα ενιαίο σύστημα.

Στην Εικόνα III.2 φαίνεται μια πιθανή χρήση ενός τέτοιου ολοκληρωμένου. Σε ένα κλασσικό σύστημα, ο πυρήνας ενός υπολογιστή συνδέεται στην κύρια μνήμη μέσω αγωγών μήκους ορισμένων εκατοστών, περιορίζοντας έτσι σημαντικά την ταχύτητα.

Χρησιμοποιώντας κάθετη ολοκλήρωση, η διασυνδέσεις αυτές μειώνονται σε τάξη χιλιοστών, ως και μικρομέτρων.

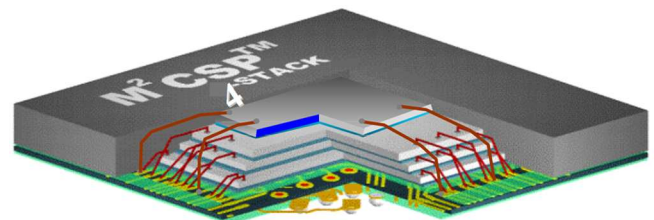
Στις επόμενες παραγράφους συζητούνται οι τρεις γνωστότερες μέθοδοι τρισδιάστατης ολοκλήρωσης.



Εικόνα III.2: Σύνδεση πυρήνα με την κύρια μνήμη, σε 2Δ (a) και 3Δ (b).

IV BONDWIRE STACKING

Η πρώτη μέθοδος είναι αυτή του Bondwire Stacking. Εδώ –συνήθως– αυτοτελή ολοκληρωμένα κυκλώματα τοποθετούνται κάθετα ή και οριζόντια σε μια κοινή ψηφίδα, ενωμένα με τη χρήση λεπτών συρμάτων (bondwires). Οι διασυνδέσεις αυτές συνήθως έχουν διάμετρο 8-500 μm και μήκος ορισμένα εκατοντάδες ή χιλιάδες μικρόμετρα. Για την προστασία των συρμάτων, γίνεται επικάλυψη της ψηφίδας με κάποιο μη-αγώγιμο στερεό υλικό (συνήθως ρητίνη).

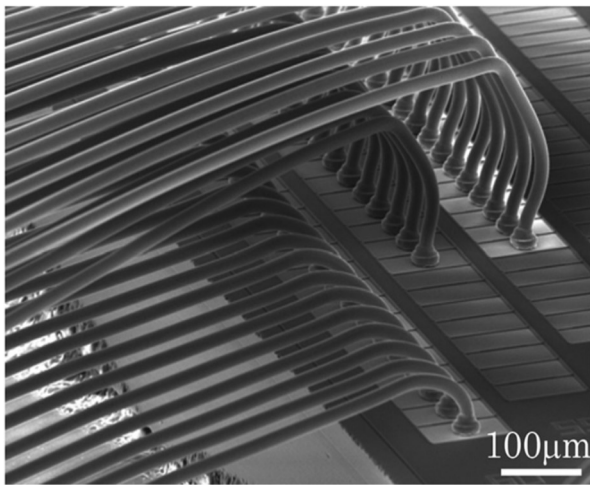


Εικόνα IV.1: Δομή ολοκληρωμένου κυκλώματος με Bondwire Stacking.

Η μέθοδος αυτή χρησιμοποιείται ενεργά από τις αρχές της τρίτης χιλιετίας, κυρίως για τη σύνδεση επεξεργαστών

με δυναμικές μνήμες και NAND Flash σε πακέτα μικρότερου όγκου. Πρακτικά, αυτό είναι και το ένα από τα δυο πλεονέκτημα του Bondwire Stacking, καθώς λόγω της μεγάλης αντίστασης των συρμάτων -άρα και της καταναλισκόμενης ισχύος- δεν κατάφερε να βρεί χρήση σε εφαρμογές εκτός των κινητών υπολογιστών (συμπεριλαμβανομένων κινητών και παιχνιδιομηχανών). Δεύτερο πλεονέκτημα της πρακτικής αυτής είναι το εξαιρετικά χαμηλό κόστος: τα διαφορετικά ολοκληρωμένα κυκλώματα που συμπεριλαμβάνονται σε ένα πακέτο για Bondwiring μπορούν να κατασκευαστούν με διαφορετική τεχνολογία, μέγεθος και υλικά, κάτι το οποίο δεν είναι δυνατό στις υπόλοιπες τεχνικές κάθετης ολοκλήρωσης επόμενων ενοτήτων.

Η προαναφερθείσα, λοιπόν, μέθοδος λέγεται ότι αποτελεί ολοκλήρωση στις 2.5 διαστάσεις, καθώς το κάθε εξάρτημά της (ολοκληρωμένο κύκλωμα) είναι το δικό του αυτοτελές σύστημα. Επομένως η ψηφίδα συνολικά δεν έχει καμία άλλη θέση παρά αυτή του μικροσκοπικού τυπωμένου κυκλώματος (PCB) το οποίο ενώνει μικρότερα κυκλώματα σε περιορισμένο εμβαδό.



Εικόνα IV.2: Διασυνδέσεις ολοκληρωμένων κυκλωμάτων με Bondwiring.

V THROUGH-SILICON VIAS

Αναμφισβήτητα γνωστότερη μέθοδος τρισδιάστατης ολοκλήρωσης είναι το Through-Silicon Via, ή TSV [7].

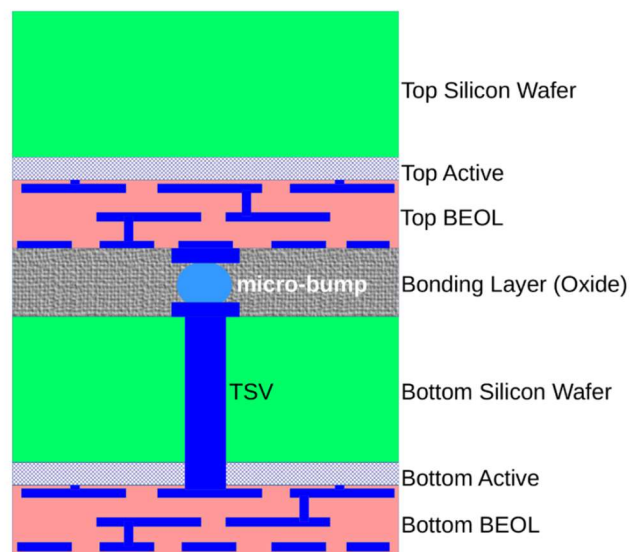
Εδώ εφόσον γίνει κατασκευή δύο (ή περισσότερων) ψηφίδων (dies), αυτές ενώνονται με τη δημιουργία πολλαπλών κάθετων αγωγίμων αυλακιών εντός του πυριτίου, με τη χρήση κάποιου μετάλλου.

Τα TSV έχουν μεγαλύτερη πυκνότητα διασυνδέσεων ανά τετραγωνικό χιλιοστό σε σχέση με το Bondwire Stacking, όπως επίσης παρουσιάζουν σημαντικές βελτιώσεις την κατανάλωση ισχύος και το εμβαδόν την διασύνδεσης. Η τεχνολογία αυτή συναντάται κυρίως σε περιπτώσεις όπου δυο διαφορετικές ψηφίδες ίδιας τεχνολογίας πρέπει να συνδεθούν κάθετα (π.χ. αισθητήρας εικόνας), ή στην πυκνή διασύνδεση πολλαπλών ίδιων ψηφίδων (π.χ. DRAM). Φυσικά το Bondwire Stacking έχει μακράν χαμηλότερο κόστος έναντι των TSV, γι' αυτό και το τελευταίο καθυστέρησε να βρεί χρήση.

Η δημιουργία των Through Silicon Vias γίνεται με τρεις τρόπους [7]:

- Via-First: το αυλάκι δημιουργείται κατά την έναρξη επεξεργασίας του Wafer,
- Via-Last: το αυλάκι σκάβεται στο ήδη ολοκληρωμένο Wafer, ή
- Via-Middle: το αυλάκι τοποθετείται ταυτόχρονα με τα υπόλοιπα layers του Wafer.

Εκ' των τριών, συνηθέστερα συναντάται το Via-Middle για περιπτώσεις που απαιτούν εξαιρετικά πυκνή ολοκλήρωση, και το Via-Last όταν το κόστος παίζει σημαντικό ρόλο.



Εικόνα V.1: Through-Silicon Via (TSV)

Υπενθυμίζονται, επίσης, δυο έννοιες οι οποίες θα χρησιμοποιηθούν στις επόμενες παραγράφους:

- Front End Of Line (FEOL): το πρώτο τμήμα παραγωγής ενός ολοκληρωμένου, όπου δημιουργούνται τα εξαρτήματά του (Transistor, πυκνωτές, αντιστάσεις κ.α.), και
- Back End Of Line (BEOL): Το επόμενο (ή τελευταίο) βήμα της παραγωγής, όπου γίνεται η διασύνδεση των εξαρτημάτων με επιστρώσεις μετάλλων.

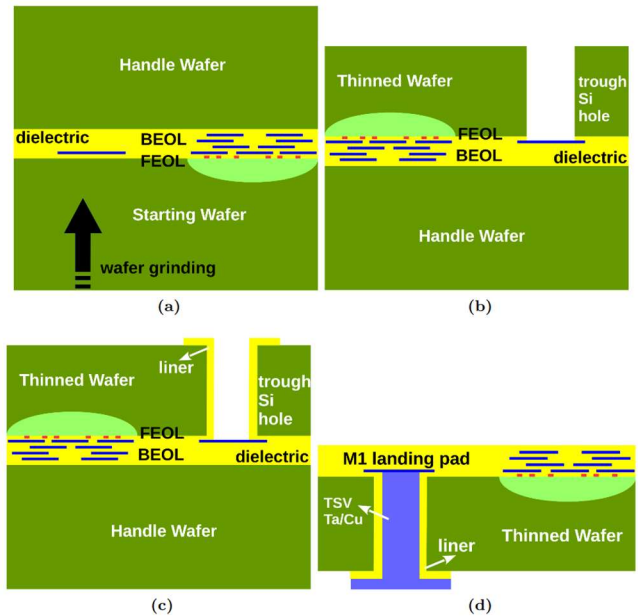
i) VIA-LAST TSV

Στη περίπτωση των Via-Last Vias, η κατασκευή τους γίνεται αμέσως μετά το BEOL έτσι ώστε να παρέχουν ένα μονοπάτι διασύνδεσης μεταξύ του κάτω μέρους του Wafer και του μετάλλου 1. Η διασύνδεση αυτή, όμως, πρέπει να γίνει με τρόπο ο οποίος δεν θα καταστρέψει το ήδη κατασκευασμένο FEOL. Η τεχνική είναι η εξής:

- Το Wafer κολλάται επάνω σε ειδικό Handle Wafer, με σκοπό την ευκολότερη διαχείρισή του, και το πρώτο λειαίνεται με μηχανικό τρόπο.
- Δημιουργούνται αυλάκια εντός του πυριτίου με τη μέθοδο της λιθογραφίας.
- Γίνεται επιστροφή με διηλεκτρικό υλικό (TSV Liner) στο εσωτερικό του αυλακιού, έτσι ώστε το τελευταίο να απομονωθεί από το υπόσωμα πυριτίου.
- Γίνεται έκχυση μετάλλου στο αυλάκι (τυπικά χαλκός).

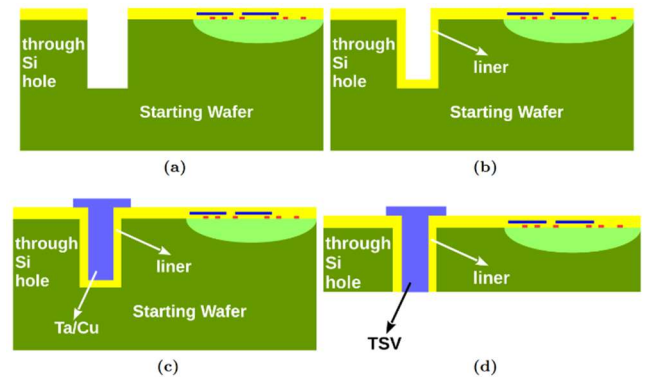
Η τεχνική αυτή είναι επικίνδυνη καθώς το Wafer πρέπει να σκαφτεί λιθογραφικά χωρίς να προκληθεί ζημιά στο FEOL ή το μέταλλο 1, ωστόσο αποτελεί και την φθηνότερη μέθοδο. Το τυπικό μέγεθος ενός Via-Last TSV είναι 5μm σε πάχος και 50μm βάθος, ενώ η αντίσταση που προσδίδει στο κύκλωμα είναι μόλις 60-70mΩ.

Ένα πραγματικό παράδειγμα χρήσης Via-Last στο εμπόριο είναι η DDRAM γνωστών εταιρειών, όπου τέσσερις ψηφίδες των 2GB συνδέονται για να δώσουν συνολικό χώρο 8GB, και ταχύτητα έως και 1.6 Gbit/s.



Εικόνα V.2: Τεχνική επεξεργασίας Via-Last για Through-Silicon Via

ii) VIA-MIDDLE TSV



Εικόνα V.3: Τεχνική επεξεργασίας Via-Middle για Through-Silicon Via

Αντίστοιχα, στα Via-Middle η κατασκευή γίνεται μετά το FEOL, αλλά πριν το BEOL. Η τεχνική είναι η εξής:

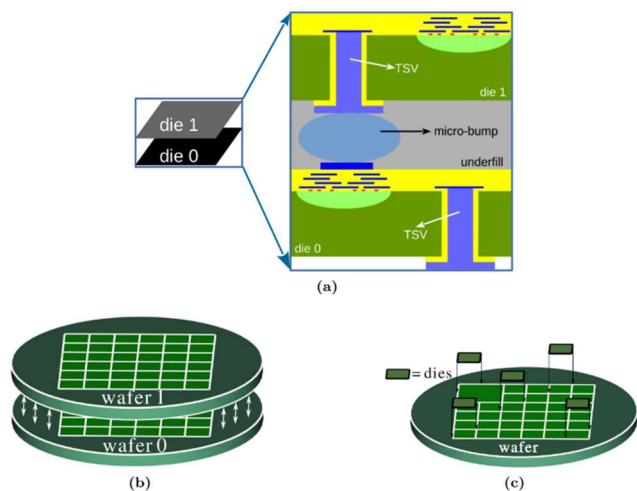
- Δημιουργείται αυλάκι στην άνω πλευρά του Wafer με τη μέθοδο της λιθογραφίας. Το αυλάκι αυτό δεν φτάνει ως την κάτω πλευρά του Wafer.
- Στη συνέχεια γίνεται επιστροφή του αυλακιού με διηλεκτρικό υλικό. Εδώ το κάτω μέρος του

αυλακιού δεν χρειάζεται να αφαιρεθεί, όπως συμβαίνει στα Via-Last.

- Γίνεται έγχυση του αυλακιού με μέταλλο.
- Λειαίνεται και λεπτύνεται η κάτω πλευρά του Wafer για να εμφανιστεί το Wafer στην κάτω πλευρά.

Το τυπικό μέγεθος και η αντίσταση των Via-Middle TSV είναι ίδιο με αυτό των Via-Last. Η πυκνότητα κάθετης ολοκλήρωσής τους είναι μεγαλύτερη, καθώς επίσης έχουν και καλύτερη οικονομική απόδοση (yield) σε σχέση με τα Via-Last. Κοστίζουν, ωστόσο, περισσότερο λόγω του τελευταίου βήματός τους, δηλαδή τη λέπτυνση του Wafer ως την εμφάνιση του μετάλλου, όπου μόρια χαλκού μπορεί να εισχωρήσουν στο πυρίτιο και να καταστρέψουν τη συσκευή. Επομένως απαιτείται μια ειδική διαχείριση για τον επιτυχή καθαρισμό του περισσίου χαλκού.

Ανεξαρτήτως της μεθόδου, αποδεικνύεται πρακτικά ότι το Landing Pad (βλ. *Εικόνα V.5*) στην κορυφή του TSV δεν μπορεί να είναι μικρότερο από 7.5μm.



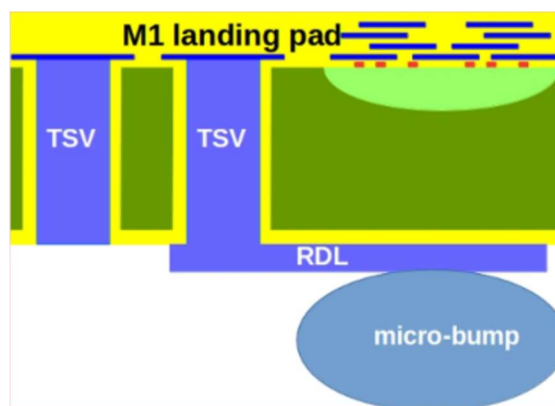
Εικόνα V.4: Ένωση (a) Die-to-Die, (b) Wafer-to-Wafer και (c) Die-to-Wafer.

Εφόσον δημιουργηθούν τα TSV, ανεξαρτήτως μεθόδου, ακολουθεί η ένωση των δυο ολοκληρωμένων κάθετα. Η ένωση αυτή μπορεί να επιτευχθεί με δυο τρόπους (βλ *Εικόνα V.4*):

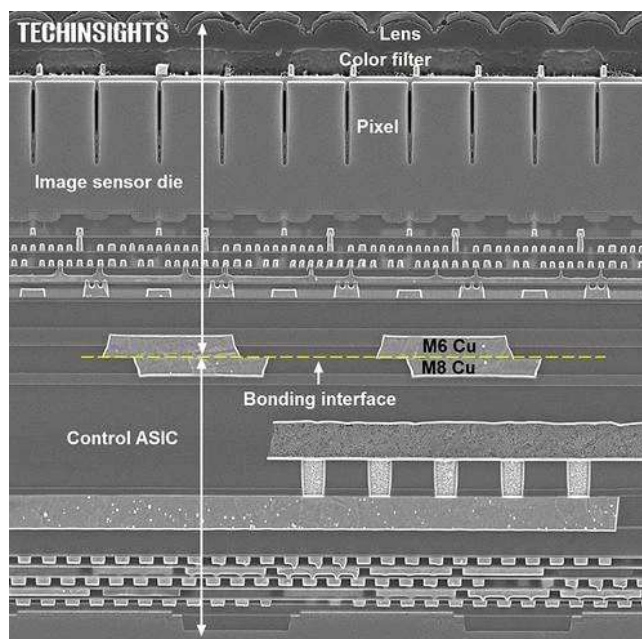
- Die-to-Die (D2D) Bonding: όπου συνδέονται κάθετα δυο ψηφίδες με TSV. Η ένωση γίνεται

με Micro-Bumps (μικρο-εξογκώματα) μεγέθους περίπου 40μm. Τα Micro-Bumps είναι προφανώς μεγαλύτερα των TSV, επομένως περιορίζουν σημαντικά την πυκνότητα οριζόντιων διασυνδέσεων.

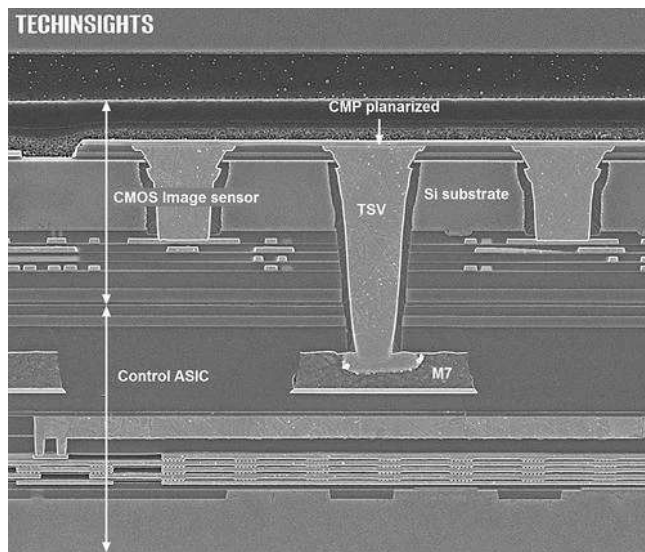
- Wafer-to-Wafer (W2W) Bonding: όπου για να αποφευχθούν τα Micro-Bumps γίνεται η κόλληση ολόκληρων Wafer, χρησιμοποιώντας χημικές μεθόδους ή ακτινοβολία. Ένα παράδειγμα της τεχνικής αυτής είναι η κόλληση Χαλκού-Χαλκού (Cu-Cu Bonding), όπου δυο Wafers τοποθετούνται σε θερμοκρασία 300-400C και υδραυλική πίεση των 200kPa έως ότου οι δυο επιφάνειες μετάλλου ενωθούν.



Εικόνα V.5: Through-Silicon Via Landing Pad



Εικόνα V.6: Αισθητήρας Εικόνας Sony IMX260 με διασύνδεση Χαλκού-Χαλκού



Εικόνα V.7: Αισθητήρας Εικόνας Sony IMX240 με διασύνδεση Via-Last TSV

Στις εικόνες V.6 και V.7 παρουσιάζεται η τομή δυο ολοκληρωμένων κυκλωμάτων που λειτουργούν ως αισθητήρες εικόνας. Στην πρώτη η σύνδεση των δυο ψηφίδων έχει γίνει με τη μέθοδο Cu-Cu (Χαλκού-Χαλκού). Στο πάνω μέρος του ολοκληρωμένου φαίνεται ο φακός του αισθητήρα, ο οποίος ενώνεται με τα το σύστημα αίσθησης. Τα Cu-Cu ενώνουν την άνω αυτή ψηφίδα, με την κάτω η οποία λειτουργεί ως ελεγκτής του συστήματος. Αντίστοιχη λειτουργία έχει και το σύστημα της εικόνας V.7, με τη διαφορά ότι η συνένωση των δυο γίνεται με τη χρήση Via-Last Through-Silicon Via.

VI ΜΟΝΟΛΙΘΙΚΗ ΟΛΟΚΛΗΡΩΣΗ

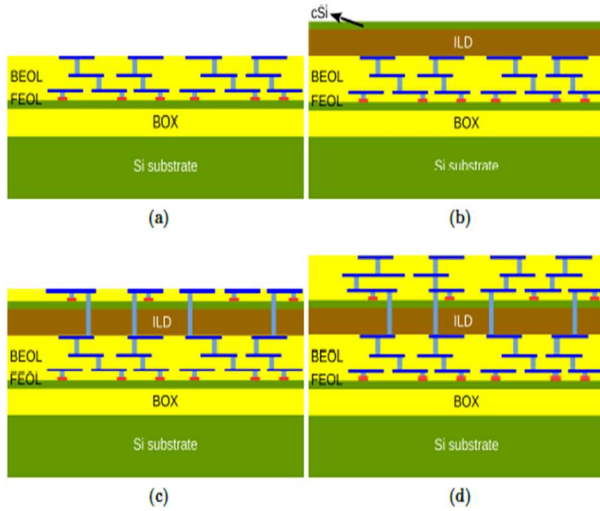
Αν και οι 3D τεχνολογίες βάσει TSV έχουν σημαντικά προνόμια, στην απόδοση του συστήματος και την κατανάλωση ενέργειας, σε σχέση με τις τυπικές 2D ολοκληρώσεις, πάσχουν από αισθητές ασυμμετρίες μεταξύ των διαστάσεων των τρανζίστορ και των διαστάσεων των TSV. Το μήκος του καναλιού σε μοντέρνα τρανζίστορ έχει φτάσει πλέον διαστάσεις κάτω των 10 νανομέτρων ενώ

η διάμετρος μοντέρνων TSV βρίσκεται στην κλίμακα αρκετών μικρομέτρων. Αυτή η τεράστια διαφορά στις διαστάσεις είναι ο βασικός λόγος που οι ολοκληρώσεις με TSV περιορίζονται σε χαμηλή πυκνότητα ολοκληρωμένων. Επιπλέον, το φορτίο χωρητικότητας ενός τυπικού TSV αντιστοιχεί σε φορτίο αρκετών χιλιάδων πυλών. Με αποτέλεσμα να καταναλώνει σημαντική δυναμική ενέργεια και να προκαλεί RC delay. Τα μονολιθικά 3D περιορίζουν αυτά τα προβλήματα μειώνοντας τις διαστάσεις των κάθετων διασυνδέσεων σε νανόμετρα. Οι διασυνδέσεις αυτές ονομάζονται monolithic inter tier vias η αλλιώς MIVs. Έτσι επιτρέπονται πρωτόγνωρα επίπεδα σε πυκνότητα και βαθμό λεπτομέρειας.

Η MONO 3D τεχνολογία αντιπροσωπεύει μια ριζική αποχώρηση από παραδοσιακές κατασκευαστικές τεχνολογίες. Τα εξαρτήματα πρώτα σχεδιάζονται ξεχωριστά, η διαδικασία αυτή ονομάζεται front end of line (FEOL) και στη συνέχεια ενώνονται και συνδέονται μεταξύ τους, η διαδικασία αυτή ονομάζεται back end of line (BEOL). Η τεχνολογία αυτή επιτρέπει την παροχή υπερβολικά πυκνών ολοκληρωμένων κυκλωμάτων, αφού δίνει την δυνατότητα να κατασκευαστούν πολλαπλά layers από τρανζίστορ πάνω σε ένα μόνο υπόστρωμα. Η αύξηση της πυκνότητας στις 3D διασυνδέσεις όμως περιορίζει το thermal budget κατά την σχεδίαση των ανώτερων layers, δηλαδή την επιτρεπτή θερμοκρασία που μπορεί να φτάσει το κύκλωμα χωρίς επιπτώσεις. Το thermal budget των ανώτερων layers επηρεάζει το FEOL και το BEOL των κατώτερων layers. Συγκεκριμένα, για το FEOL, θερμοκρασίες πάνω από 500oC και για περισσότερο από δύο ώρες επηρεάζουν τα τρανζίστορ, τις τάσεις κατωφλίου και τις περιοχές του source και του drain. Για το BEOL, υψηλές θερμοκρασίες μπορούν να προκαλέσουν κενά στις χάλκινες διασυνδέσεις, αλλά και διάχυση του χαλκού στο γύρω διηλεκτρικό με αποτέλεσμα να δημιουργούνται προβληματισμοί σχετικά με την αξιοπιστία των διασυνδέσεων.

i) Si over Si

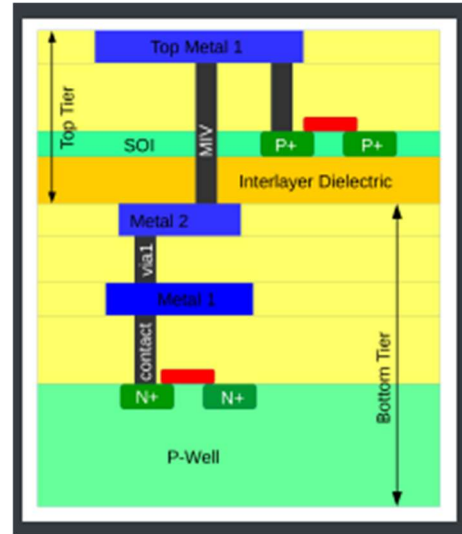
Η κύρια πρόκληση του σχηματικού ολοκλήρωσης SI over Si είναι η απόκτηση χαμηλής θερμοκρασίας στα FEOL των άνω επιπέδων. Η διαδικασία της Si-over-Si Mono3D ολοκλήρωσης είναι η ακόλουθη:



Εικόνα VI.1: Διαδικασία ολοκλήρωσης CMOS over CMOS

Για το πρώτο επίπεδο μπορεί να χρησιμοποιηθεί ένα ήδη σχεδιασμένο wafer. Έπειτα για το FEOL μπορούν να χρησιμοποιηθούν όλες οι γεωμετρίες τρανζίστορ, όπως τα FinFET ή τα επίπεδα CMOS και οποιοδήποτε υπόστρωμα από πυρίτιο (SOI / bulk). Για το BEOL προτείνεται η χρήση ενός μετάλλου το βολφράμιο (tungsten) για την αποφυγή προβλημάτων σταθερότητας στις διασυνδέσεις ή μολύνσεων που μπορεί να προκύψουν κατά την επεξεργασία των άνω επιπέδων. Έπειτα, τοποθετείται πάνω από το wafer ένα στρώμα διηλεκτρικού (inter layer dielectric ILD), το οποίο χωρίζει τα δυο επίπεδα. Στη συνέχεια τοποθετείται ένα λεπτό στρώμα πυριτίου σε κρυσταλλινή μορφή πάνω από το διηλεκτρικό, τα οποία ενώνονται με wafer-to-wafer bonding. Τα

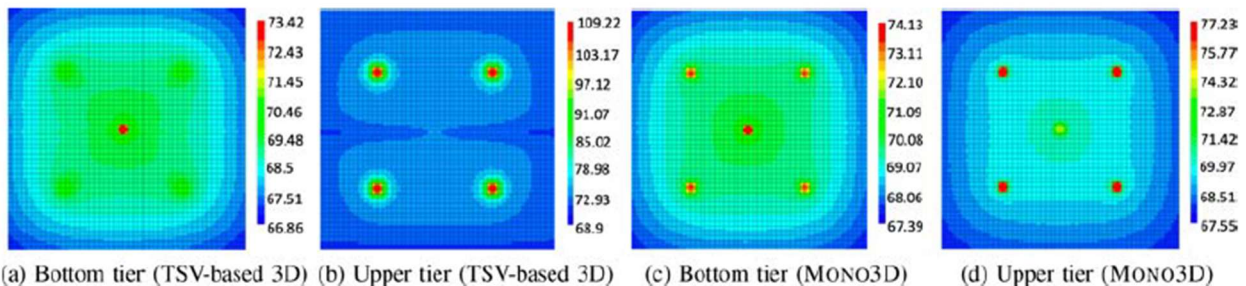
τρανζίστορ μπορούν στη συνέχεια να σχεδιαστούν πάνω σε αυτό το λεπτό στρώμα πυριτίου. Σχηματίζονται τα MIVs τα οποία ενώνουν το στρώμα αυτό με το στρώμα πυριτίου του από κάτω επιπέδου. Τέλος, σχεδιάζονται οι διασυνδέσεις του πάνω επιπέδου, το BEOL δηλαδή. Άλλοι τρόποι σχεδίασης αντί της χρήσης πυριτίου για τα πάνω στρώματα είναι η χρήση carbon nano-tubes και γερμανίου.



Εικόνα VI.3: Μονολιθική 3D ολοκλήρωση

ii) Θερμική Ακραιότητα

Ένας μεγάλος περιορισμός στα 3D ολοκληρωμένα είναι η αποτελεσματική διασπορά της θερμότητας από επίπεδα μακριά από την ψήκτρα. Τα Mono3D αντιμετωπίζουν επιπλέον προβλήματα λόγω της πυκνής ενσωμάτωσης των εξαρτημάτων. Υπάρχουν διάφορες τεχνικές για την μείωση των θερμικών προβλημάτων. Η προσεκτική σχεδίαση από power



Εικόνα VI.2: Θερμικό mapping TSV και Mono3D ολοκληρωμένων

delivery networks (PDN) μπορεί να μειώσει τις θερμοκρασίες στα μικροκυκλώματα ως και 5%.

Γενικά, το πυρίτιο και άλλα μέταλλα δρουν ως αγωγοί θερμότητας, ενώ τα διηλεκτρικά ως μονωτές. Η θερμότητα σε ένα 3D ολοκληρωμένο ρέει ως προς δυο κατευθύνσεις, οριζόντια στα στρώματα πυριτίου και κάθετα προς τις ψήκτρες όπου και χάνεται. Ο πρώτος μηχανισμός βοηθάει στη μείωση της έντασης από hot spots, ενώ ο δεύτερος μειώνει τη μέση θερμοκρασία στη ψηφίδα (die). Στα Mono3D παρόλο που το πολύ λεπτό στρώμα πυριτίου στο πάνω επίπεδο προκαλεί χαμηλή θερμική αγωγιμότητα, ο συνδυασμός από πολύ λεπτό διηλεκτρικό και πολύ μικρή απόσταση μεταξύ των MIVs αποτελεί αποτελεσματική κάθετη θερμική διάχυση. Σε finite element προσομοιώσεις έχει παρατηρηθεί 53% μείωση σε θερμοκρασία, τοποθετώντας νανο-κολώνες κατά την σχεδίαση σε transistor-level για τη διασπορά της θερμότητας από επιλεγμένες hot spot περιοχές.

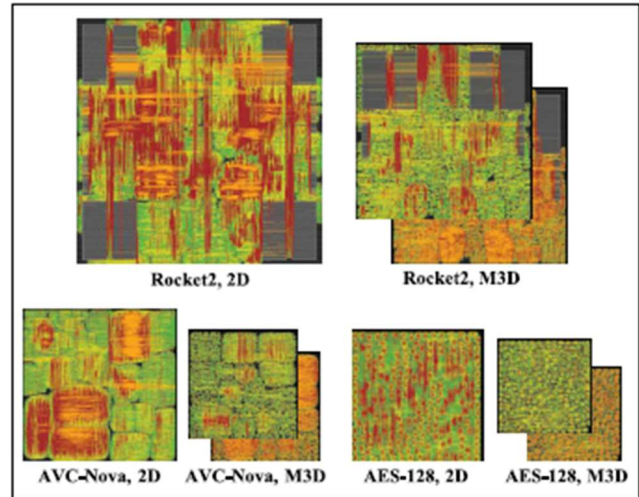
Για να αποδειχθεί ότι τα μονολιθικά 3D ολοκληρωμένα έχουν καλύτερη διασπορά θερμότητας από τα αντίστοιχα TSV-based ολοκληρωμένα έχουν πραγματοποιηθεί διάφορες προσομοιώσεις. Mono3D και TSV-based 3D ολοκληρωμένα δύο επιπέδων μοντελοποιούνται μαζί με τα στρώματα μετάλλων με τη χρήση ενός architecture level thermal simulator (HotSpot-6.0) [4]. Τα μοντέλα προσομοιώνονται τοποθετώντας ένα σημείο hot spot στο κέντρο του κάτω επιπέδου, το οποίο είναι κοντινότερα στη ψήκτρα και τέσσερα σημεία hot spots στο άνω επίπεδο μακριά από την ψήκτρα. Τα αποτελέσματα φαίνονται στο σχήμα. Τα TSV-based ολοκληρωμένα έχουν υψηλότερες θερμοκρασίες από αντίστοιχα Mono3D. Ο λόγος που

συμβαίνει αυτό είναι επειδή έχουν μεγαλύτερη κάθετη αντίσταση, αφού είναι μακρύτερη η διαδρομή της ροής θερμότητας (μέσω των μεγαλύτερων TSV) και έχουν και πιο χοντρό bonding layer ανάμεσα στα διάφορα επίπεδα.

iii) Design Comparison

Η σύγκριση των Mono3D ολοκληρωμένων με τα αντίστοιχα 2D γίνεται με τη βοήθεια ορισμένων μετρικών [5] που υπολογίζουν την κατανάλωση ενέργειας, την απόδοση και την συνολική επιφάνεια. Το φυσικό σχέδιο γίνεται με τη χρήση εξαρτημάτων 28nm. Το Mono3D design έχει 50% μικρότερο αποτύπωμα σε σχέση με το 2D design και παρουσιάζει μια αξιοσημείωτη μείωση στο μήκος καλωδίων, αφού οι διασυνδέσεις των στρωμάτων γίνονται μέσω των MIVs. Η μείωση των καλωδίων έχει ως άμεση συνέπεια τη μείωση της συνολικής ισχύος.

Τα Mono3D κατά τη ροή της σχεδίασης εκτελούν το final routing ξεχωριστά για κάθε επίπεδο.



Εικόνα VI.4: GDS Design Layouts

	Rocket2, 800MHz			AVC-Nova, 730MHz			AES-128, 3.44GHz		
	2D	M3D	Δ(%)	2D	M3D	Δ(%)	2D	M3D	Δ(%)
Footprint (mm^2)	0.746	0.373	-50.01	0.286	0.142	-50.21	0.210	0.105	-50.2
Std. cell count	279983	278622	-0.49	132725	132299	-0.32	109878	109634	-0.22
Std. cell area (mm^2)	0.363	0.362	-0.30	0.190	0.189	-0.79	0.124	0.121	-2.71
Wirelength (mm)	4712.46	4045.65	-14.15	2336.97	1954.92	-16.35	1494.13	1196.88	-19.89
MIV count	0	111395		0	56989		0	44782	
Total Pwr. (mW)	371.90	366.25	-1.52	122.50	120.31	-1.79	269.59	258.96	-3.95
Net Switching Pwr. (mW)	82.70	78.00	-5.68	23.55	22.32	-5.20	67.65	61.57	-8.99
Cell Internal Pwr. (mW)	224.84	224.26	-0.26	63.46	63.20	-0.41	174.66	171.86	-1.60
Leakage Pwr. (mW)	64.36	63.99	-0.58	35.50	34.79	-1.99	27.29	25.52	-6.48
WNS (ps)	-54.27	-49.47		-60.48	-135.57		-14.88	-26.87	

Πίνακας VI.1: Σύγκριση αποτελεσμάτων μεταξύ 2D και 3D ολοκληρωμένων

Αυτό έχει ως αποτέλεσμα να μην μπορεί να γίνει ακριβής βελτιστοποίηση του συγχρονισμού για το συνολικό σχέδιο. Έτσι ο τελικός συγχρονισμός είναι ατελής με αποτέλεσμα την ύπαρξη ενός συνόλου από worst case σενάρια στα οποία υπάρχει χρονική καθυστέρηση, παρόλο του αποτελεσματικού σχεδιασμού. Αυτή η χρονική καθυστέρηση είναι στο ίδιο εύρος με το 2D design και είναι λιγότερη από 10% της επιθυμητής περιόδου του ρολογιού.

Επιπλέον, κατά την σχεδίαση του άνω επιπέδου, το κάτω επίπεδο μπορεί να παραμορφωθεί με αποτέλεσμα να αλλοιωθούν τα τρανζίστορ του και τα RC να έχουν παράσιτα. Υπάρχουν κάποιες τεχνικές που αντιμετωπίζουν αυτό το πρόβλημα μεταφέροντας κάποια critical path cells στο γρηγορότερο άνω επίπεδο.

iv) Design for test

Τα Mono3D παρουσιάζουν μοναδικές προκλήσεις στο τεστάρισμα, επειδή έχουν πολλά layers από τρανζίστορ και έναν υψηλό αριθμό MIVs. Η έντονη κλιμάκωση του πάχους του διηλεκτρικού στρώματος και η αυξημένος αριθμός των MIVs (30 εκατομμύρια ανά mm²) είναι οι κύριες αιτίες για τα λειτουργικά και χρονικά σφάλματα στα συστήματα αυτά.

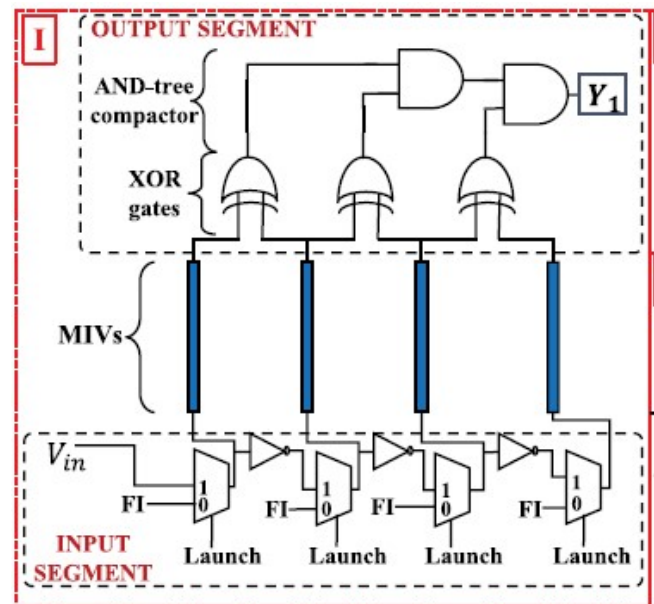
Τα τυπικά μοντέλα σφαλμάτων είναι παρόμοια με αυτά των μοντέλων σφαλμάτων των διασυνδέσεων σε 2D ολοκληρωμένα. Σε γενικές γραμμές είναι βραχυκλώματα, ανοιχτοκυκλώματα και stuck-at faults (SAFs). Τα SAFs προκύπτουν όταν κάποια σήματα ή κάποιοι ακροδέκτες κολλάνε στο λογικό 1, 0 ή X. Βραχυκλώματα συμβαίνουν συνήθως όταν ένα σωματίδιο μολύνει τον χώρο ανάμεσα σε 2 MIVs. Και τα SAFs προκύπτουν όταν υπάρχει βραχυκύκλωμα μεταξύ MIV και τροφοδοσία ή MIV και γείωση.

Οι παραδοσιακές μεθοδολογίες τεσταρίσματος δεν είναι εφικτές λόγω του υψηλού αριθμού ύπαρξης των MIVs και της μη δυνατής χρήσης των μεθόδων σε πολλαπλά layers, αφού για να εντοπιστεί ένα σφάλμα πρέπει να τεσταριστούν όλα τα στρώματα. Αυτό αυξάνει την πολυπλοκότητα και δεν εντοπίζεται η ακριβής περιοχή

του σφάλματος. Επιπλέον, σφάλματα σε πολλαπλά MIVs μπορούν πιθανώς να εξουδετερωθούν μεταξύ τους.

a) MIVs XOR-BIST architecture

Στην αρχιτεκτονική MIV XOR-BIST τοποθετούνται XOR πύλες δύο εισόδων μεταξύ των εξόδων γειτονικών MIVs στο κομμάτι της εξόδου της μονάδας του BIST. Επομένως, για N MIVs τοποθετούνται N-1 πύλες XOR. Οι εξοδοί των πύλων XOR περνάνε από έναν ισορροπημένο AND-tree compactor με N-1 εισόδους και με απόκριση Y1 ενός bit (pass/ fail). Ένα σημείο παρατήρησης (scan flop) είναι αρκετό για την παρακολούθηση της υγείας πολλαπλών MIVs. Ένα σφάλμα στο σετ των MIVs που μελετάται μπορεί να εντοπιστεί παρατηρώντας το Y1. Στο κομμάτι της εισόδου της μονάδας του BIST, μια τάση εισόδου παρέχει συμπληρωματικές εισόδους στα γειτονικά MIVs κατά το τεστάρισμα μέσω μιας αλυσίδας από inverter. Ένας 2:1 πολυπλέκτης δίνει την δυνατότητα μετάβασης από την κατάσταση λειτουργικότητας στην κατάσταση του τεσταρίσματος.



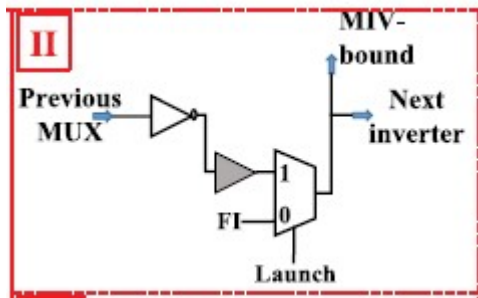
Εικόνα VI.5: MIVs XOR-BIST architecture

Δύο κύκλοι ρολογιού χρησιμοποιούνται για το τεστάρισμα σφαλμάτων των MIVs. Θέτοντας την τάση εισόδου σε 0 και 1 για διαδοχικούς κύκλους. Το τελικό αποτέλεσμα είναι "101..." για τον πρώτο κύκλο και

"010..." για τον δεύτερο κύκλο. Ένα σετ από MIVs δεν παρουσιάζει σφάλματα αν το Y1 είναι μονάδα και στους δύο κύκλους ρολογιού.

b) MIVs dual-BIST architecture

Η XOR-BIST αρχιτεκτονική που αναλύθηκε προηγουμένως σε περίπτωση που προκύψουν SAFs, τα πιθανά σφάλματα των MIVs μπορούν να καλυφθούν και να μην εντοπίζονται. Για την μείωση της πιθανότητας να συμβεί αυτό, τοποθετείται παράλληλα μια δεύτερη μονάδα εξόδου που συνδέεται και αυτή στις εξόδους των MIVs. Έτσι, σχηματίζεται η μονάδα BIST-A και η μονάδα BIST-B με εξόδους Y1 και Y2 αντίστοιχα. Στην BIST-B, όμως, αντί των πυλών XOR και AND τοποθετούνται πύλες XNOR και OR αντίστοιχα. Με αποτέλεσμα στο συνολικό dual-BIST να μην μπορεί να καλυφθεί ένα MIV σφάλμα.



Εικόνα VI.6: Enhanced dual-BIST for resistive fault detection

c) Enhanced dual-BIST for resistive fault detection

Το μέγεθος βραχυκυκλώματος ή ανοιχτοκυκλώματος καθορίζει το μέγεθος της αντίστασης του. Όσο πιο μεγάλο το βραχυκύκλωμα τόσο πιο μικρή και η αντίστασή του και αντίστοιχα όσο μεγαλύτερο το ανοιχτοκύκλωμα τόσο μεγαλύτερη και η αντίστασή του. Για να εντοπιστεί το μέγεθος ενός σφάλματος, προστίθεται ένα προγραμματιζόμενο delay element. Ένας N-stage inverter, όπου N ο αριθμός των single-stage inverters. Η ένωση γίνεται στις εισόδους των MIVs στην κατάσταση τεσταρίσματος.

VII ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] C. A. Mack, "Fifty Years of Moore's Law," in IEEE Transactions on Semiconductor Manufacturing, vol. 24, no. 2, pp. 202-207, May 2011, doi: 10.1109/TSM.2010.2096437.
- [2] Noyce, Robert N. "Microelectronics." *Scientific American*, vol. 237, no. 3, Scientific American, a division of Nature America, Inc., 1977, pp. 62-69, <http://www.jstor.org/stable/24920319>.
- [3] S. Maeda, "Activities of the Research and Development Association for Future Electron Devices," in Proceedings of the IEEE, vol. 77, no. 9, pp. 1420-1429, Sept. 1989, doi: 10.1109/5.35246.
- [4] K. Dhananjay, P. Shukla, V. F. Pavlidis, A. Coskun and E. Salman, "Monolithic 3D Integrated Circuits: Recent Trends and Future Prospects," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 3, pp. 837-843, March 2021, doi: 10.1109/TCSII.2021.3051250.
- [5] A. Chaudhuri et al., "Advances in Design and Test of Monolithic 3-D ICs," in IEEE Design & Test, vol. 37, no. 4, pp. 92-100, Aug. 2020, doi: 10.1109/MDAT.2020.2988657.
- [6] Abbas Sheibanyrad, Frédéric Pétrot, Axel Jantsch, "3D integration for NOC-based SOC Architectures", 2010, Integrated Circuits and Systems, doi: 10.1007/978-1-4419-7618-5.
- [7] Chaourani Panagiotis, "Sequential 3D Integration - Design Methodologies and Circuit Techniques", KTH Royal Institute of Technology, 2019, ISBN: 978-91-7873-228-9.
- [8] M. Kawano, X. -Y. Wang and Q. Ren, "Trench Isolation Technology for Cost-effective Wafer-level 3D Integration with One-step TSV," 2020 IEEE 70th Electronic Components and Technology Conference (ECTC), 2020, pp. 1161-1166, doi: 10.1109/ECTC32862.2020.00186.
- [9] A. Chaudhuri et al., "Advances in Design and Test of Monolithic 3-D ICs," in IEEE Design & Test, vol. 37, no. 4, pp. 92-100, Aug. 2020, doi: 10.1109/MDAT.2020.2988657.
- [10] K. Dhananjay, P. Shukla, V. F. Pavlidis, A. Coskun and E. Salman, "Monolithic 3D Integrated Circuits: Recent Trends and Future Prospects," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 3, pp. 837-843, March 2021, doi: 10.1109/TCSII.2021.3051250.
- [11] D. B. Ingerly et al., "Foveros: 3D Integration and the use of Face-to-Face Chip Stacking for Logic Devices," 2019 IEEE International Electron Devices Meeting (IEDM), 2019, pp. 19.6.1-19.6.4, doi: 10.1109/IEDM19573.2019.8993637.