

NAND 3D - Grossissement et pas rétrécissement !

par Axel Stoermann, Directeur Général Marketing Technique et Ingénierie Applications, chez Toshiba Electronics Europe

Les cris pour réclamer plus de capacité mémoire à des prix inférieurs sont devenus presque assourdissants. Jusqu'à récemment, le rétrécissement lithographique était la seule manière d'augmenter la densité de mémoire, c'est-à-dire le nombre de bits pouvant être stockés par unité de silicium.

Cependant, des progrès récents dans les techniques de production de semiconducteurs, ont permis aux ingénieurs de produire des puces contenant des structures cellulaires empilées en 3D. Ces structures non seulement répondent à tous ceux qui réclament une densité de bits supérieure, mais surmontent aussi plusieurs des défis posés par le rétrécissement lithographique des puces.

Avec l'arrivée des structures NAND 3D en production, les lecteurs SSD (Solid State Drive, ou lecteur semiconducteur) d'une capacité de plus de 15 To sont désormais à portée de main.

Piloter la loi de Moore

Depuis l'introduction des NAND par Toshiba en 1984, la technologie flash NAND est à la pointe de la miniaturisation lithographique dans l'industrie électronique. Par conséquent, les dispositifs NAND sont parmi les CI les plus denses produits en série. Au cours des 30 dernières années, la taille du noeud de traitement a chuté de 350 nm à 15 nm.

Cela plus l'introduction de nouvelles technologies de cellules multi-niveaux permettant de gérer davantage de bits par cellule, a permis d'augmenter la densité de bits des puces NAND d'un facteur de plus de 2.000. Parallèlement à cette augmentation de densité, on a pu observer une chute impressionnante du prix par Go, qui a baissé encore plus vite que la densité de bits n'augmentait.

L'un des grands défis pour ceux qui veulent utiliser les toutes dernières NAND brutes dans leurs appareils, est que ces nouvelles technologies NAND ont en général besoin d'une correction d'erreurs (ECC pour Error Code Correction) beaucoup plus puissante, intégrée aux contrôleurs. Un autre défi majeur est qu'alors que la lithographie rétrécit et que les bits par cellule augmentent, la durabilité des cellules diminue.

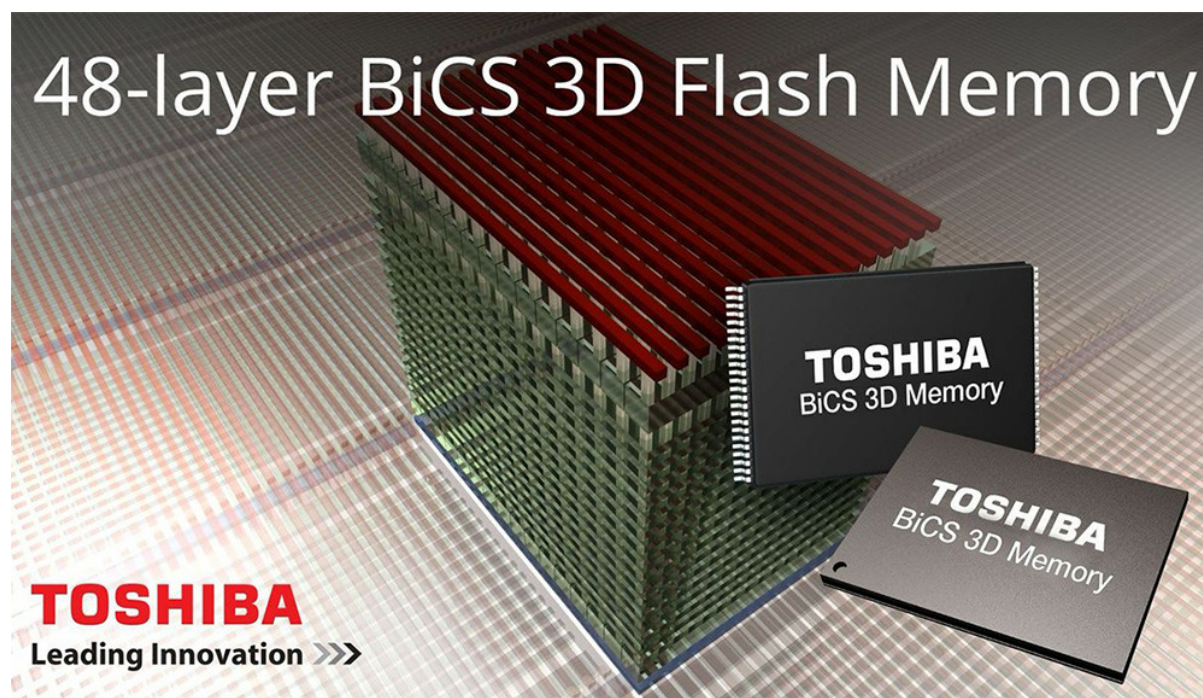
Le développement de contrôleurs hôte NAND est donc devenu un élément déterminant du secteur des mémoires, et d'une façon générale, les contrôleurs sont souvent en retard par rapport au rétrécissement des puces.

Aller au delà des limites de la lithographie

Afin de surmonter les défis posés par le rétrécissement des puces tout en augmentant la densité de bits, les fabricants de semiconducteurs ont développé le moyen d'empiler des cellules l'une sur l'autre, pour former des structures NAND 3D.

L'idée de base des NAND 3D est d'empiler les cellules de mémoire Flash verticalement sur le substrat en silicium, ce qui augmente considérablement la densité de bits par rapport à une mémoire Flash NAND plane, où les cellules sont entassées côte à côte sur la puce.

Toshiba a développé une mémoire Flash à structure cellulaire tridimensionnelle à 48 couches, appelée BiCS (Bit Column Stacked, ou colonne de bits empilée), qui surpasse la capacité des mémoires Flash NAND bidimensionnelles traditionnelles, tout en améliorant la fiabilité d'écriture-effacement, l'endurance et aussi les vitesses d'écriture.



L'amélioration de fiabilité d'écriture-effacement est possible parce que la nouvelle technologie BiCS permet le recours à des procédés lithographiques plus grossiers, tout en améliorant malgré tout la densité de bits. En fait, l'amélioration de fiabilité d'écriture-effacement et d'endurance est si spectaculaire, que non seulement des cellules double-niveau (MLC : 2 bits par cellule) et triple-niveau (TLC : 3 bits par cellule) sont déjà produites en série, mais des mémoires à 4 bits par cellule sont également envisagées.

Dans un proche avenir, Toshiba prévoit que les NAND 2D et les NAND 3D de type BiCS vont coexister, car il n'y a pas de superposition au niveau des capacités offertes par ces dispositifs. Les NAND 2D plafonnent à une capacité de 128 Gbits par puce, alors que les BiCS ciblent des capacités au delà de 128 Gbits par puce. Toshiba prévoit que les mémoires Flash BiCS trouveront leurs

premières applications dans les disques SSD d'entreprise, conçus pour les grands centres de stockage et de traitement de données "cloud".

Une différence significative dans la structure des dispositifs est que les architectures BiCS font appel à des cellules "charge trap" (piège de charge) qui stockent les électrons dans une couche de nitrure de silicium, plutôt que dans le silicium polycristallin dopé, typique des cellules à grille flottante; utilisées dans les structures NAND 2D. Ces nouvelles cellules mémoire à piège de charge sont bien plus durables que les cellules à grille flottante utilisées traditionnellement.

Les cellules à grille flottante peuvent être comparées à un verre rempli d'eau qui fuirait dès qu'une infime zone du verre est abîmée, ce qui remettrait en cause la durabilité et la longévité du dispositif. En comparaison, les couches mémoire composées de cellules à pièges de charge sont comme une éponge qui stockerait un certain volume d'eau même si une partie de l'éponge est cassée, augmentant ainsi la durabilité et la durée de vie.

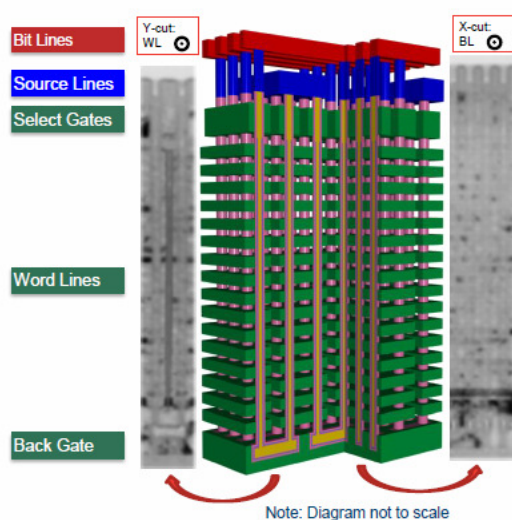


Figure 1. NAND BiCS à cellules pièges de charge

La forme en U des BiCS permet un rendement matriciel maximum, et permet la fabrication sur des lignes de fabrication (légèrement modifiées) actuelles. En premier lieu les dispositifs BiCS seront fabriqués grâce aux techniques actuelles de lithographie en immersion. Cependant, la lithographie en ultra-violet extrême (EUVL) ou la lithographie nanoimprint (NIL) permettraient à des motifs bien plus complexes d'être transférés, ce qui augmenterait encore la souplesse de conception.

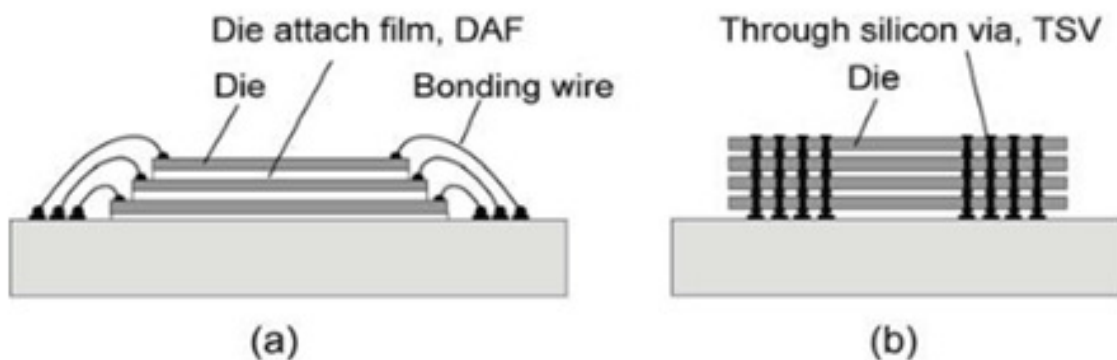
Etant donné que l'espace entre les cellules est agrandi, le bruit et les parasites de cellule à cellule est moindre. Ceci permet des vitesses d'écriture séquentielle plus élevées pour les dispositifs BiCS, que pour les dispositifs NAND TLC 2D. Un dispositif BiCS peut atteindre des vitesses d'écriture séquentielle supérieures à 30-40 Mo/s, soit sensiblement plus que les 20-30 Mo/s des dispositifs TLC 15 nm. Ces vitesses se traduisent par un débit d'environ 533 Mbits/s pour les dispositifs BiCS.

Traversées silicium et au-delà

Pour augmenter encore ces débits, de nouvelles technologies doivent être mises en oeuvre pour connecter les cellules entre elles. Dans les premiers dispositifs BiCS "standard", chaque couche est connectée grâce aux techniques "wirebonding" (liaisons à fils) traditionnelles.

Toshiba a collaboré avec la New Energy and Industrial Technology Development Organisation (Organisation pour le développement de la technologie industrielle et des énergies nouvelles) pour intégrer des connexion à la fois plus rapides et offrant un meilleur rendement énergétique. La société a récemment annoncé avoir développé les premières mémoires Flash NAND à 16 puces au monde, à l'aide de la technologie TSV (Through Silicon Via, ou traversée à travers le silicium) pour connecter les différentes puces d'une structure BiCS.

A l'inverse du wirebonding, la technologie TSV utilise des électrodes verticales et des vias (trous métallisés de traversée) pour passer à travers les puces silicium et assurer les connexions. Ceci permet d'obtenir des entrées-sorties très rapides, tout en réduisant la consommation.



La technologie TSV Toshiba permet d'obtenir un débit E/S de plus de 1 Gbits/s, soit plus que n'importe quelle autre mémoire Flash NAND alimentée en basse tension : 1.8V pour le coeur et au 1.2V pour les circuits E/S. Cela permet de réduire de 50% la puissance requise pour effectuer les opérations d'écriture, de lecture et d'E/S, par rapport à la technologie wirebonding.

Naturellement, le futur ne s'arrête pas à la 3D. Les mémoires ReRAM (Resistive RAM, ou RAM résistive) et les mémoires PCM (Phase Change Memory, ou mémoire à changement de phase) commencent à apparaître au delà de l'horizon NAND 3D, et pourraient se substituer aux BiCS à long terme.

Mais la mémoire la plus disruptive à venir est la mémoire MRAM (Magneto-resistive RAM, ou RAM magnéto-résistive). La MRAM est un autre type de mémoire non-volatile, qui atteint des vitesses proches de celles des DRAM, mais à l'inverse des NAND ces mémoires ont une endurance à l'écriture-effacement, quasi-illimitée.

La MRAM a le potentiel de modifier la façon dont tous les systèmes électroniques accèdent aux données. Aujourd'hui, les gens utilisent la Flash NAND comme un disque dur ; ils copient les données à la mise sous tension de la Flash NAND vers la DRAM, puis exécutent le code en DRAM ; c'est ce que l'on appelle le "shadowing".

Avec la MRAM, vous n'avez plus besoin d'avoir recours au shadowing. Les données sont déjà là quand vous mettez sous tension, et cela offre l'avantage supplémentaire de ne plus devoir compter autant sur l'endurance de la Flash NAND, qui peut encore servir à l'archivage à plus long terme.

La révolution mémoire arrive sur nous

Nous vivons une époque formidable ; la révolution des mémoires à semiconducteurs arrive sur nous. Le rétrécissement lithographique a déjà vu les puces rétrécir d'un facteur 2.000, et les structures mémoire 3D les verront rétrécir encore davantage. Ces évolutions pousseront encore plus la mémoire à semiconducteurs sous le feu des projecteurs, et nous verrons un pourcentage toujours plus grand de données stockées dans de telles mémoires.

Ces mémoires vont non seulement permettre un accès aux données plus rapide, mais elles vont aussi révolutionner la conception des appareils, en permettant à des volumes de données encore plus grands d'être stockés, consultés et analysés, plus vite que jamais.

Mars 2016

Réf : 6699A2