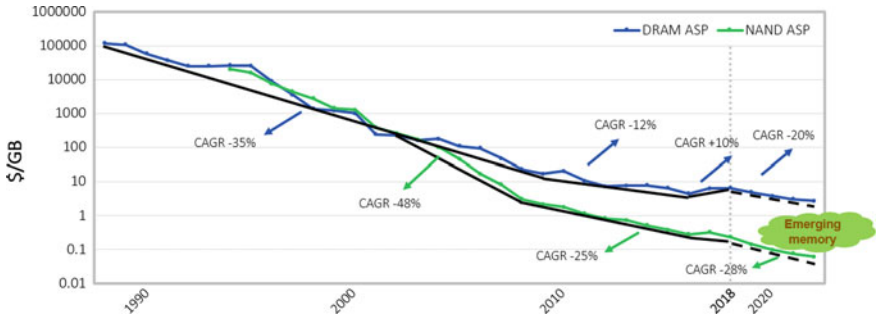

Fig. 10.3 DRAM versus flash ASP (average selling price). *Source* IDC

Early 3D NAND products used 24 layers in the 3D stack, and then the industry released 32, 64, 72, and recently 96 layers to production. This 3D-stack driven roadmap suggests continuing the 3D scaling towards a few hundred layers, thus keeping the scaling of NAND memory products to increase the memory capacity with the corresponding reduction in cost per bit.

Capacitor based DRAM would not allow such 3D scaling and no alternative has been proposed so far to do so for DRAM.

10.2 Alternative Memory Technologies

Over the past decades, a significant R&D effort has been devoted to developing alternative memory technologies. The leading alternative technologies are based on: Phase Change Materials (PCM), Resistive Memory (R-RAM) or Magnetic Memory (M-RAM). These alternative memory technologies have many variations and derivatives with other name branding as well. None of these alternative memory technologies seems to challenge the mainstream technologies—DRAM and NAND. And none of these technologies has been considered as a potential alternative to DRAM.

10.2.1 PCM—3D XPoint

Intel and Micron collaborated in releasing to the market a product named Optane™ as a Storage Class Memory (SCM) to bridge the growing gap between DRAM and 3D NAND. Also, it is considered a 3D memory as it is designed as a cross-point architecture and would not fit the low-cost 3D Scaling in which many memory layers are processed together following the same lithography step. 3D XPoint is not considered as a potential DRAM alternative due to access speed and endurance limitations.

higher than 10^{12} cycles. A similar concept was published by IBM [4]. This work was confirmed and improved on by work such as Fujitsu's [5] and covered in patents filed by Macronix [6] and Micron [7]. It seems that the thin tunneling concept was proposed at a time floating gate rather than charge-trap was the Non-Volatile industry's technology of choice. Moreover, at that time DRAM scaling was in-step with the rest of the industry and a thin tunneling charge-trap did not offer enough of an advantage to be pursued by the memory industry.

In the Flash market for storage applications, NAND architecture became the industry choice as it provides a significantly higher density (lower cost) than a NOR architecture. As illustrated in Fig. 10.4 a NAND architecture with only two diffusion contacts could provide access to a long NAND string, thus reducing the effective size of a memory cell to $4F^2$ [8]. In the NOR architecture, the one diffusion contact per cell increases the cell size to $8F^2$, thus a higher memory cost. The NOR architecture does provide direct access to the selected cell which result in much faster read access time, consequently making it attractive for applications such as program code storage. An alternative architecture shown in Fig. 10.4 as AND architecture provides direct access with a better density than conventional NOR. This architecture often is also called NOR and could be attractive for 3D random access memory structures.

The success of the NAND industry with 3D NAND scaling could now be followed by adopting Charge-Trap for DRAM and changing the memory architecture from a NAND to a NOR (AND) architecture. Such a 3D architecture has been first proposed by Macronix [9], later by MonolithIC 3D Inc. with single crystal channel option [10], and then by Eli Harari [11] and his new company Sunrise Memory Corp (Eli Harari was the founder of SanDisk and won the National Medal of Technology and Innovation from President Barack Obama for his innovations and contributions to flash memory storage solutions). These proposals could be grouped into those with horizontal bit-line orientation and vertical bit-line orientation. In the following, the details of 3D NOR with a vertical bit-line orientation are presented. An important advantage of these structures is the similarity to the common 3D NAND 'Punch and Plug' process and accordingly the advantage in sharing the industry accumulated know-how and manufacturing infrastructure.

10.4 Charge-Trap 3D NOR (AND)

Just like in 3D NAND, the foundation fabric is a multi-layer fabric such as oxide layers with poly-silicon in-between. The number of poly-silicon layers is a linear relation to the number of memory cells in the 3D memory structure. And just as in 3D NAND the memory process is done for the full multi-layer fabric affecting all the levels together—hence 3D scaling (Fig. 10.5).

Figure 10.6 illustrates a side cut-view of the structure overlaying the structure transistor schematic. It represents an aggressive 3D NOR (AND) structure in which the bit-lines (B0–B4), in blue, serves as Source and Drain to cells on their right side and on their left side. These bit lines could be formed by filling the punch holes with

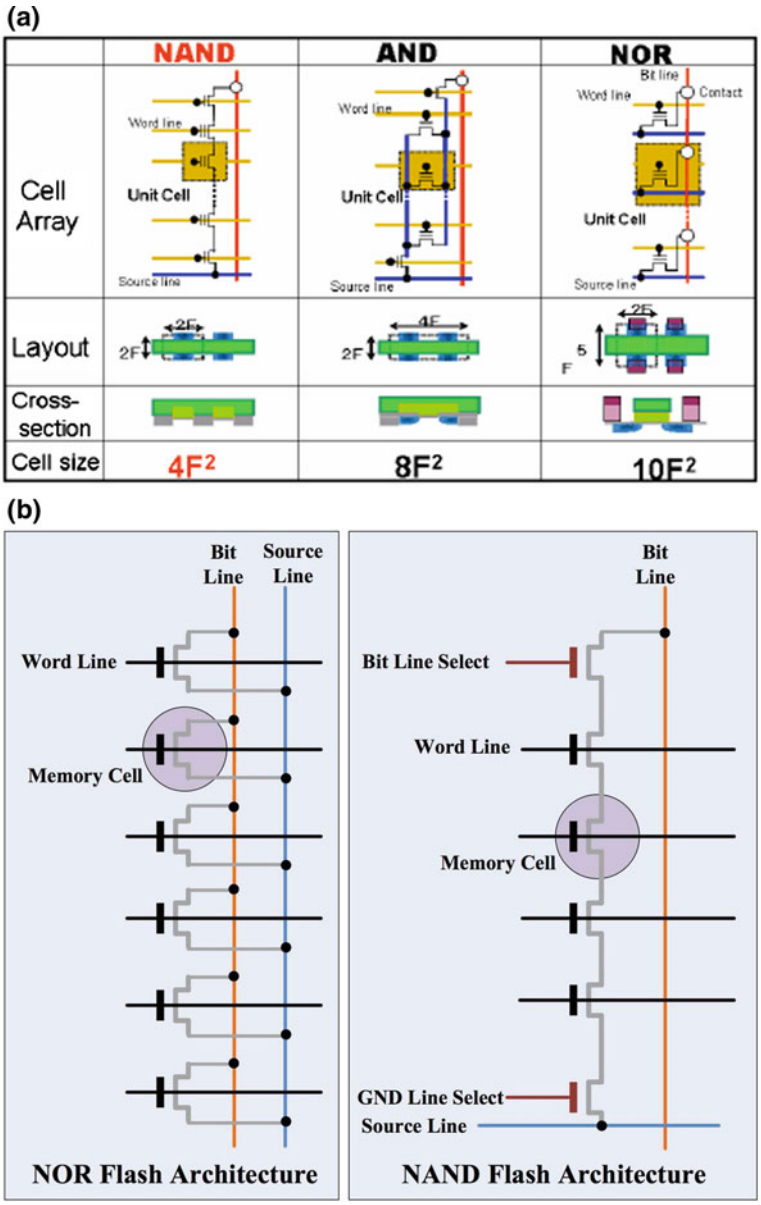

Fig. 10.4 **a** NOR versus NAND flash architecture. **b** NOR versus NAND flash architecture

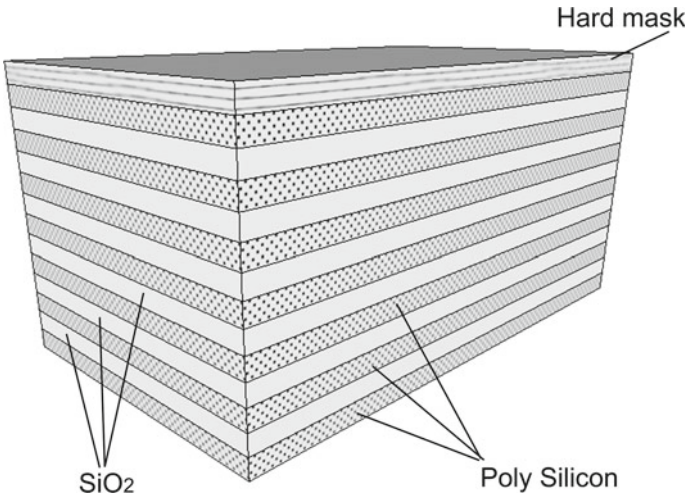

Fig. 10.5 Multilayer fabric as foundation for 3D NAND and 3D NOR

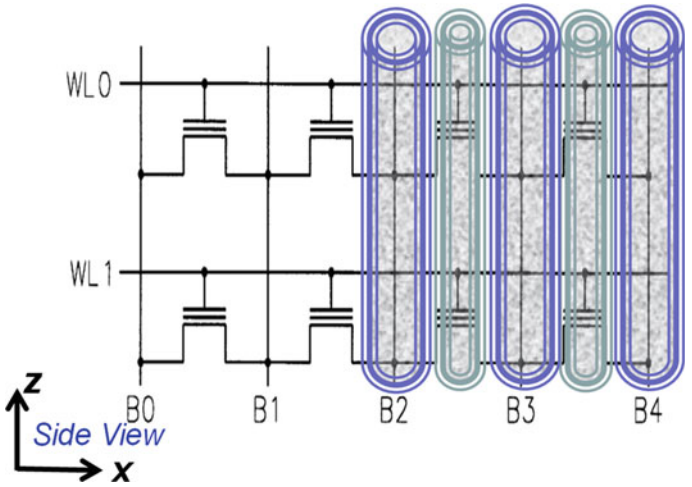

Fig. 10.6 Transistor schematic overlaid by punch and fill source/drain (bit-lines) in the odd holes and channel in the even holes

N+ silicon, or through a combination of N+ layers on the holes' walls and core of metal or even just metal for a Schottky-based structure. The channel holes in between are filled with un-doped polysilicon.

The structure looks like a 3D NAND with N+ holes punched between channel holes. Figure 10.7 illustrates a top view of the structure.

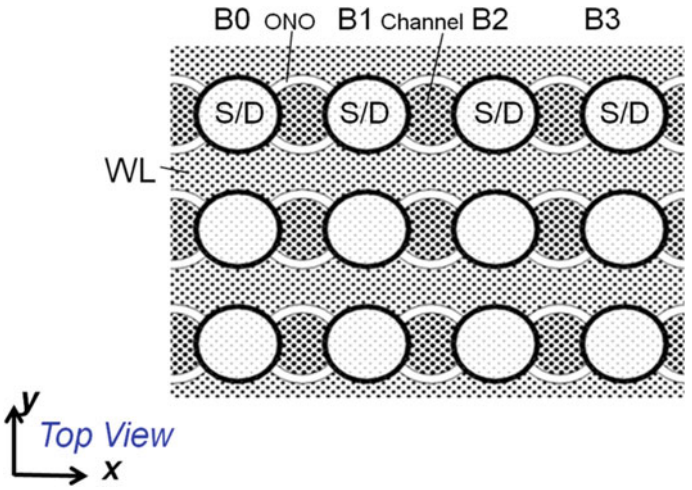

Fig. 10.7 Top-view, source/drain (bit-lines) in the odd holes and channel in the even holes

Figure 10.8 illustrates an alternative for the 3D NOR structure in which no additional holes are 'punched' for the channel but rather forming the channels by use of etch and deposition through the S/D holes.

Additional details for the 3D NOR structures and alternative process flows to form them could be found in the referenced patents and applications [9–12].

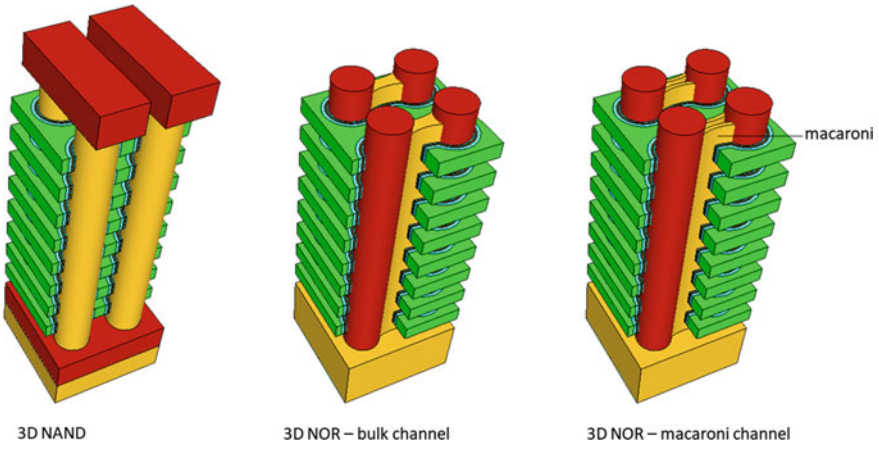

Fig. 10.8 Some alternatives for 3D memory structures

10.5 Schottky Barrier and Dopant Segregated Schottky Barrier ("DSSB")

In flash devices there are a few writing mechanisms that are frequently used. One is Fowler–Nordheim (FN) tunneling commonly used in NAND flash devices and another is Hot Carrier Injection (HCI), also called Hot Electron, often used in NOR flash devices. Flash cell writing using FN tunneling is orders of magnitudes more efficient than HCI as in FN most of the current is the tunneling current while in HCI only a small fraction of the current through the channel is actually the hot carriers being driven over the quantum barrier thus to be trapped.

In a paper [13] titled "Performance breakthrough in NOR flash memory with dopant-segregated Schottky-barrier (DSSB) SONOS devices" a few orders of magnitude improvements were reported by the use of Schottky Barrier devices, as is illustrated in Fig. 10.9 (Fig. 3 of the paper [13]).

This improvement in hot carrier write time and efficiency was reported in other papers including devices without dopant segregation, and devices utilizing poly silicon channels [14–16]. Using metalized Source/Drain lines in the 3D NOR device improves the bit-line conductivity and thus enhances the device P/E efficiency and speed.

Comparing such a 3D NOR technology to Stacked Capacitor DRAM suggests many advantages such as: higher density, 3D scaling, lower power, reduced rate of refresh, non-destructive read. Yet Charge-Trap 3D NOR is expected to have much longer erase time. Proper design of a 3D NOR device could support a full segment erase scheme, which combined with proper system design and support software, could compensate for the erase time deficiency.

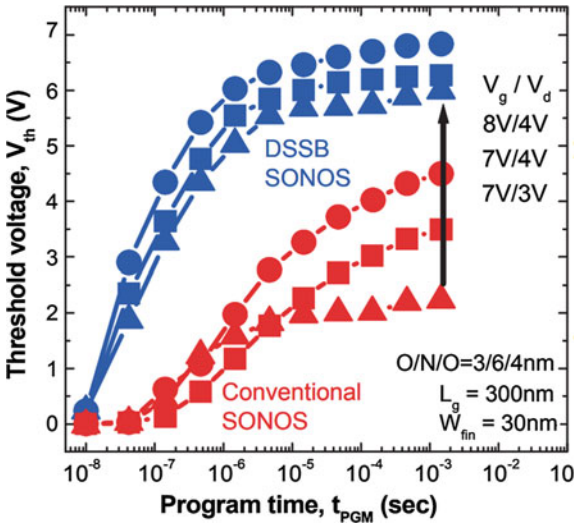

Fig. 10.9 Program/erase characteristics for NOR flash memory cell (double gate), DSSB and conventional SONOS devices