2004 – The Year of 90-nm: A Review of 90 nm Devices

Dick James
Chipworks Inc.
3685 Richmond Road, Ottawa, Ontario, Canada K2H 5B7
(email: djames@chipworks.com)

Abstract

The year 2004 saw the introduction of the first 90-nm process node devices into the marketplace. This node is notable not only for the expected reduction in feature sizes, but also for the more general adoption of low-k dielectric layers, and the first use of nickel silicide and strained silicon.

Chipworks, as a supplier of competitive intelligence to the semiconductor and electronics industries, monitors the evolution of chip processes as they come into commercial production. Chipworks has obtained parts from leading edge manufacturers, and performed structural analyses to examine the features and manufacturing processes of the devices.

The paper shows how "90nm" has been interpreted by various vendors, details the physical transistor structures we have analyzed, and comments on the introduction of strain into CMOS processing.

INTRODUCTION – WHAT IS 90 NM?

We continually see parts labeled as "130 nm" or "90 nm". To the uninitiated, this must seem to be some form of code, but it is actually just industry jargon for a process generation or 'node'.

The 1994 NTRS (National Technology Roadmap for Semiconductors) predicted a minimum feature size of 130 nm in 2004, with a 90 nm feature size in 2008. It is a reflection of the frenetic pace of process development that we are currently seeing feature sizes of less than 50 nm, and metal pitches ~230 nm.

Table 1 shows a comparison of the metal 1 half-pitch and minimum gate dimensions that we found in the samples analyzed by Chipworks. As can be seen, no manufacturer meets the actual ITRS numbers – Intel is closest in M1 dimensions, and Fujitsu the closest in gate length.

90-NM TRANSISTORS

Sony/Toshiba CXD9797GB from Sony PSX

Toshiba was the first off the mark in January, with the Sony PSX chip, which combined the Sony graphics synthesizer with their Emotion Engine™ on one die. The process used five layers of copper + one aluminium, two low-k layers at metal 2 and metal 3 (M2 and M3), and a 45 nm minimum gate.

Figure 1 shows a SEM image of NMOS transistors stained to show the source/drains. Figure 2 is a TEM image; the structure of the sidewall spacers can be seen. In essence, this seems to be a dimensional shrink from the previous generation.

Table I. Comparison of Published and Found Dimensions with ITRS 2003 Dimensions for the 90nm Node

	Metal 1 Half-Pitch (nm)		Minimum Gate Length (nm)	
	Published	Found	Published	Found
ITRS 2003 (MPU/ASIC)	107		37	
Sony/Toshiba	120	130	65	45
Texas Instruments	165	165	52	47
Intel	110	120	50	45
IBM	158	135	46	45
UMC (Xilinx)	240	240	70	55
TSMC (Altera)	120	125	59	55
Fujitsu (Transmeta)	140	145	40	40
AMD	N/A	130	N/A	49

< > BACK
TSMC Exhibit 1025, Page 1 of 5

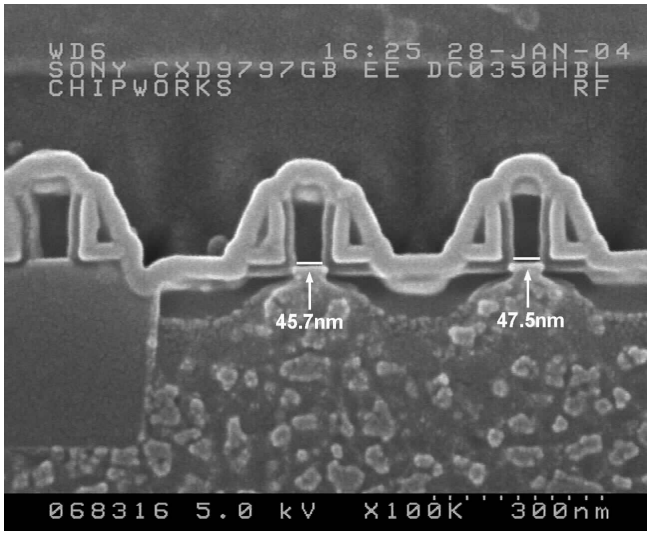


Figure 1. SEM Cross-Section of NMOS Transistors in Sony PSX

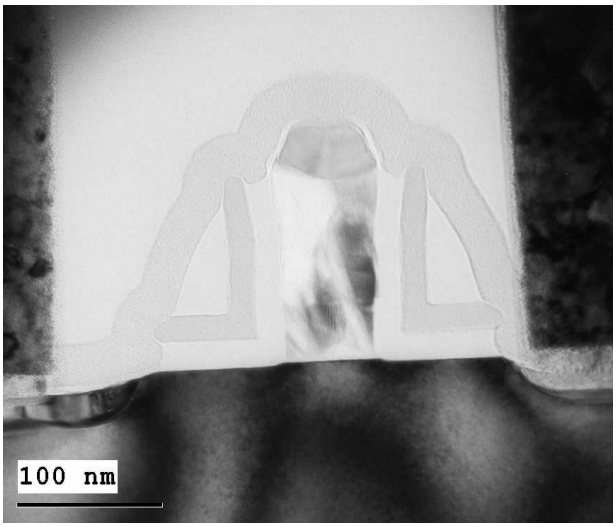


Figure 2. TEM Cross-Section of Sony PSX Transistor

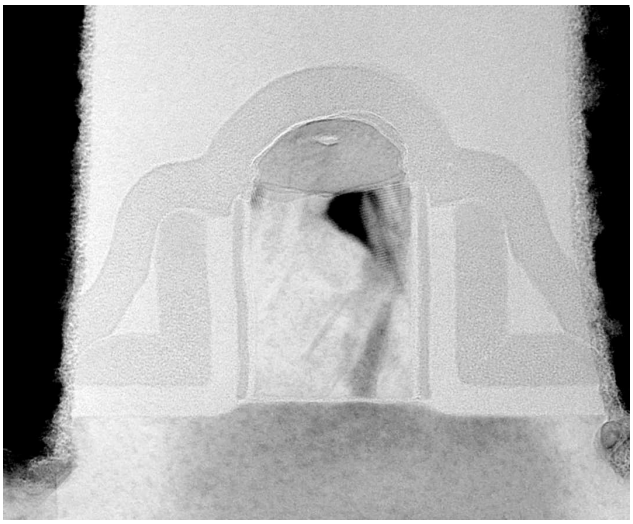


Figure 3. Transistor in TI DSP

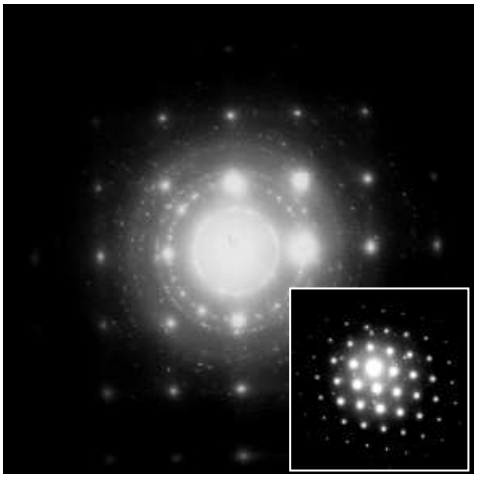


Figure 4. TEM Electron Diffraction Image of TI Transistor Channel Region (Inset – Intel Prescott, <110> Channel Direction)

Texas Instruments TMX320C6416T 1 GHz DSP

The TI DSP is a 7-metal part, with low-k at M1 – M4. The transistor again appears to be a shrink down to a 47 nm gate length, with a nitride liner on the side of the gate electrode. Figure 3 is a TEM image of the transistor cross-section.

The distinguishing feature of the TI process revealed itself in an electron diffraction image (Figure 4). This showed that the channel was oriented in the <100> direction, compared with the more normal <110> direction. This orientation gives higher PMOS mobility.

Intel Prescott

Intel first announced their 90-nm P1262 process in March 2002. Process details such as the use of strained silicon and nickel silicide were given later in the year and at the 2002 IEDM [1], and at IEDM 2003 Intel showed the different strain techniques for NMOS and PMOS transistors [2]. The Prescott die analysed by Chipworks had seven metal levels, with five low-k layers at M2 – M7.

An NMOS transistor is imaged in Figure 5. Localized tensile stress is applied to the channels of NMOS devices by a highly stressed nitride sealant layer, in undercut regions under the edges of the sidewall spacers. This increases electron mobility by ~10%. The gate oxide thickness is ~1.3 nm.

The PMOS transistors (Figure 6) use epitaxial SiGe source/drains to create compressive stress – this improves hole mobility by as much as 50%. The substrate is etched out in the source/drain regions, and local epi-SiGe (~17% Ge) is deposited into the cavities. Crystalline facets form raised source/drains, reducing the series resistance. The tensile stress from the nitride layer is counteracted by the compressive effect.

The strain has induced dislocations at the SiGe/Si interface. The SiGe was likely *in-situ* boron-doped, and also implanted with the polysilicon gate implant. The boron

TSMC Exhibit 1025, Page 2 of 5

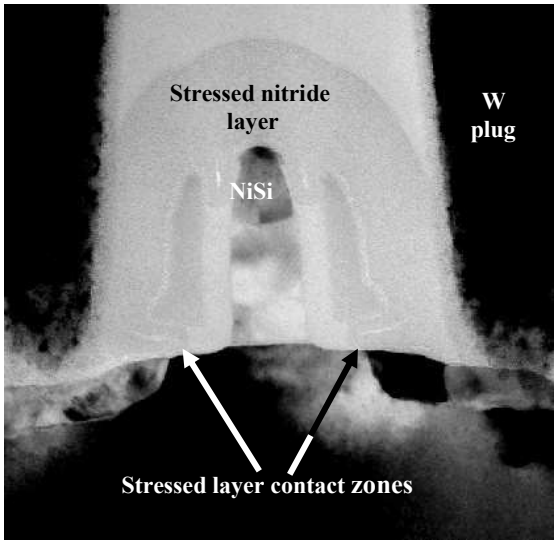


Figure 5. NMOS Transistor in Intel Prescott

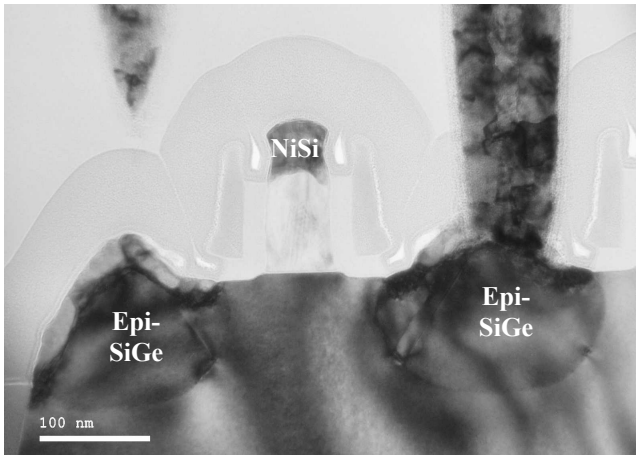


Figure 6. Intel PMOS Transistor

out-diffuses from the SiGe into the silicon to neutralize the potential leakage caused by the dislocations. Nickel silicide was chosen because it consumes less silicon and forms at a lower temperature, making it more compatible with the SiGe.

Figure 7 is a longitudinal section of a PMOS source/drain, clearly showing a stacking fault extending from the epitaxial interface to the silicide layer.

IBM PowerPC PC970FX

The IBM PowerPC is built on a 45 nm-thick SOI layer, with a 45 nm gate length with a 1.5 nm gate oxide thickness. The part has ten metal levels with a hybrid FSG/USG dielectric structure.

IBM had stated that the part would use strain technology, and press reports in the spring of 2004 speculated that the "Strained Silicon Directly on Insulator" (SSDOI) technique, which was announced at IEDM 2003 [3], would be used.

The device appears to use a stressed nitride layer to

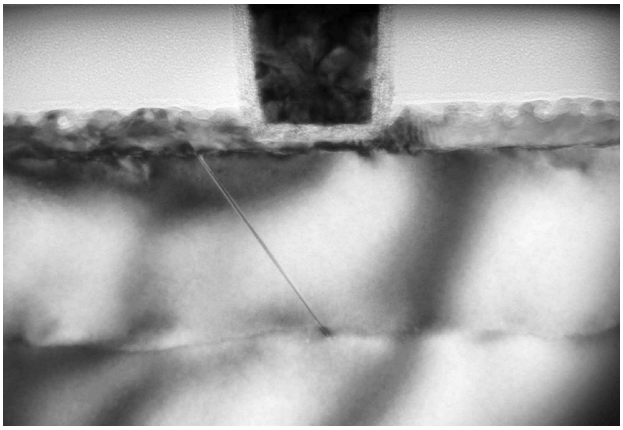


Figure 7. Longitudinal Section of Intel PMOS Source/Drain

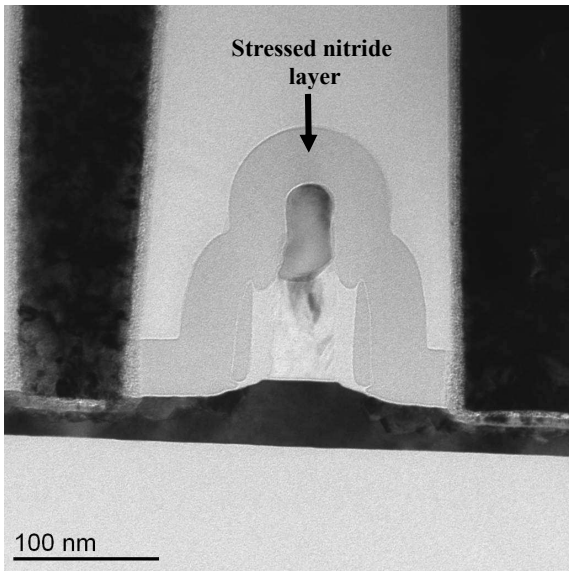
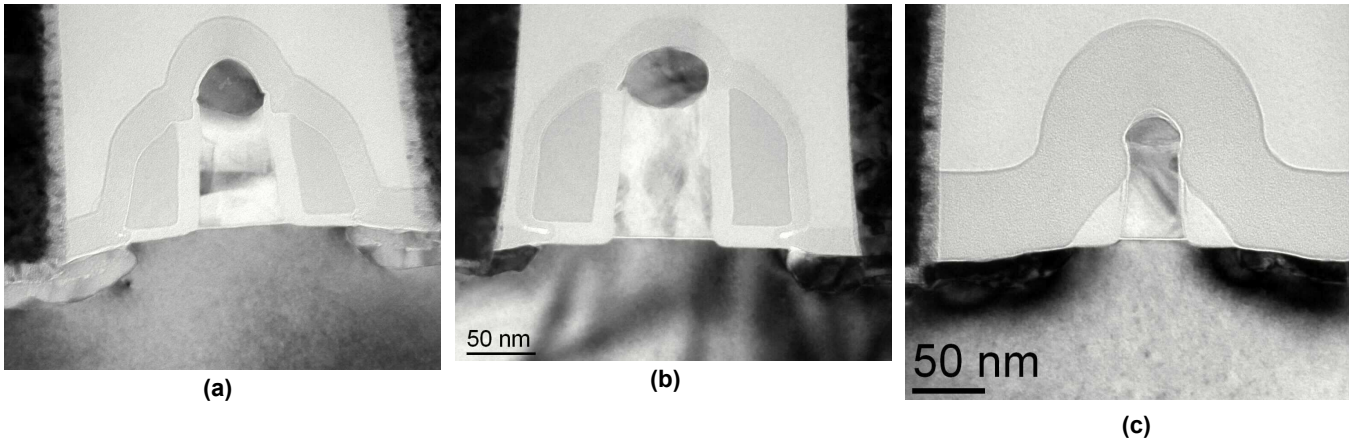


Figure 8. NMOS Transistor in IBM PPC970FX

Figure 9. PMOS Transistor in IBM PPC970FX

enhance the NMOS transistors. Figure 8 is a TEM image of a NMOS transistor, and it can be seen that the sidewall spacers are very slim, and it is only ~20 nm from the gate edge to the cobalt silicide of the source/drain. Comparison with a PMOS transistor (Figure 9) illustrates that the

TSMC Exhibit 1025, Page 3 of 5



(a) (b) (c)

Figure 10. Transistors in (a) Xilinx Spartan-3, (b) Altera Stratix-II, (c) Transmeta Efficeon-2

NMOS spacers have indeed been etched back in order to reduce series resistance, and place the stressed nitride layer closer to the gate edge. Electron diffraction studies on the SOI body confirmed <110>-channel orientation, and there was no indication of a SSDOI substrate. The IEDM SSDOI paper indicated that ~1.5% lattice distortion was required to give useful mobility enhancement.

It is interesting to note that at IEDM 2004, IBM and AMD presented a joint paper [4] on "dual stress liner" technology, which uses stressed nitride layers optimized to apply tensile stress to NMOS transistors, and compressive stress to PMOS devices.

This technique is to be integrated into 90-nm product in the first half of 2005. There was no mention of the use of SSDOI substrates.

UMC (Xilinx Spartan-3)

The Xilinx part was the first 90-nm part from a pure foundry analyzed by Chipworks. The transistor structure was very similar to the 130-nm node device, with the gate shrunk to 55 nm (Figure 10(a)). The gate length was smaller that the UMC-published value of 70 nm, but in other respects the part was manufactured using a more conservative process; FSG was used as the inter-metal dielectric, with seven copper + one aluminium metal levels (although UMC states that low-k dielectrics are available as an option).

There does not appear to be any use of strain to enhance mobility – the nitride layer is thinner that those used in strained devices, and there seems to be no structural changes to move the nitride contact zone closer to the gate edge. However, the gate edge/silicide distance is a respectably small 45 –50 nm.

TSMC (Altera Stratix™-II)

The other major foundry to launch a 90nm process in 2004 was TSMC. The product analyzed by Chipworks was the Altera Stratix™-II FPGA. This displayed a gate length of 55 nm, similar to UMC, although closer to the published 59 nm physical gate length of the Nexsys™ process. This is another ten-metal part (9 Cu + 1 Al), using low-k at M1 _ M6.

The transistor structure (Figure 10(b)) is not a direct shrink from the 130nm node; it has been modified to change the sidewall spacer structure from the typical L-shaped spacer of previous process generations. Again, the nitride layer appears to be too thin to be used as a stress layer for mobility enhancement. Gate/silicide spacing is ~55 nm.

The other major feature of the TSMC 90-nm process is its evolution towards a second-generation low-k dielectric process – the trench and via etching are noticeably cleaner than at the 130-nm node, and the cap layer has changed to a SiCO (rather than SiCN) composition.

Fujitsu (Transmeta Efficeon-2)

Transmeta announced shipment of its Efficeon-2 in September, fabricated by Fujitsu at their Akiruno Advanced Technology Center. This part also seems to use a strained nitride layer. Figure 10(c) displays what is becoming a characteristically thick nitride sealant layer, which combined with the heavily etched sidewall spacers, leads us to conclude that nitride stress is being used.

Fujitsu had announced the use of nitride strain for the 65-nm node at VLSI 2003 and 2004 [5][6], and they seem to have used the technique in their nominal 90-nm process. Given that the gate length is an impressively small 40 nm, and the gate length for the 65-nm process is quoted at 35 nm, this part could be regarded as an interim step to 65-nm. (They also published a dual-stress process at IEDM '04 [7])

In addition, electron diffraction shows that the channels are oriented in the <100> direction, so Fujitsu is the first manufacturer to combine channel re-orientation and strain to optimise mobility.

The dielectric layers are also advanced, using Novellus Coral low-k (not SiLK!) with SiCO copper-cap and etch-stop layers at M1 – M5. It is an eight-metal part, again with a top aluminium layer.

AMD Athlon 64 Mobile

AMD continued their use of "double spacer" technology at the 90-nm node, with a gate shrink to ~50 nm. The NMOS device (Figure 11) consistently has dislocations under the inner spacer, presumably caused by the drain extension implant. These do not appear in the PMOS transistors (Figure 12), which likely have a different amorphisation implant to fix the drain extension.

There is no evidence of either nitride strain or substrate strain in this part; the nitride is relatively thick, and the double spacer places the contact zone further from the gate. The joint AMD/IBM "dual stress liner" paper [4] showed an Athlon fabricated with the technique, and announced production in 2005.

AMD have also evolved their interconnect structure for the 90-nm node, compared with the 130 nm; the metallisation (9 Cu + 1 Al) is now predominantly dual-damascene, although whilst the dielectric structure is still a hybrid low-k/FSG.

SUMMARY AND DISCUSSION

We have seen eight examples of 90-nm processes in this review, with transistor gate lengths ranging from 40 – 55 nm. Intel, IBM, and Fujitsu have introduced strain for mobility enhancement, and notably Intel has taken the extra steps of embedded SiGe strain for PMOS, and nickel silicide. Texas Instruments and Fujitsu have also used re-orientation of the substrates to improve mobility.

It appears that 90-nm is the node at which strained silicon is coming into production, although not by using epitaxially strained substrates. The strain method of choice has been localised uni-axial strain, as opposed to the bi-axial strain created by strained SiGe layers.

It is perhaps ironic that stress in silicon nitride layers, a problem in previous process generations, has now been harnessed to apply strain. So far, it has been used only to provide tensile stress to enhance NMOS performance (at least in the three examples we have shown), but the dual-stress approach announced by IBM/AMD and Fujitsu indicates that the layer can also be tuned to provide compressive stress to improve hole mobility for PMOS devices. This technique has the noteworthy convenience that it can be accomplished by process tuning, and involves no extra cost for either capital equipment or special substrates.

It remains to be seen if the embedded SiGe approach used by Intel will become the dominant method of applying compressive stress for PMOS. Texas Instruments has indicated that it will be using local SiGe strain and nickel silicide in its 65-nm process [8], so Intel will not be alone.

All in all, 2004 was an important year in CMOS process evolution, with significant developments in transistor engineering, and further improvement in low-k dielectric usage. We will see what 2005 brings!

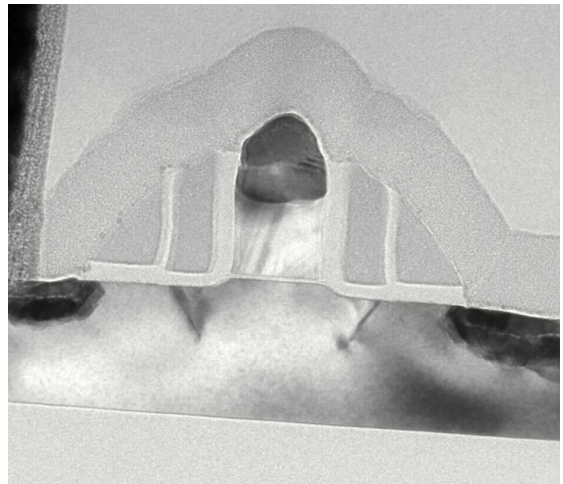


Figure 11. NMOS Transistor in AMD Athlon

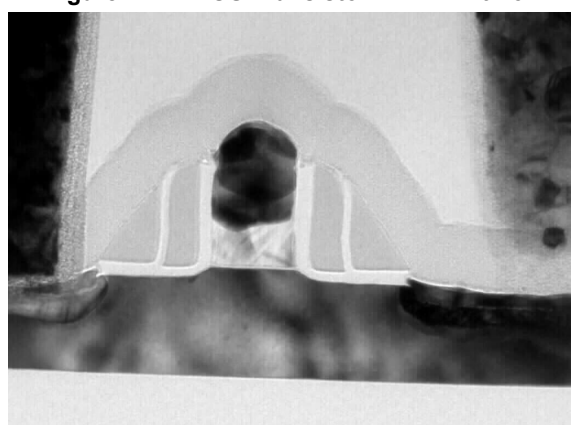


Figure 12. PMOS Transistor in AMD Athlon

ACKNOWLEDGEMENTS

I would like to thank Chipworks' laboratory staff and process analysts, and our contractors, who actually did all the hard work of analyzing these complex devices. They did a great job!

REFERENCES

[1] S. Thompson, et al., *IEDM Tech. Digest,* pp. 61-64, (2002)

[2] T. Ghani, et al., *IEDM Tech. Digest,* pp. 978-980, (2003)

[3] K. Rim, et al., *IEDM Tech. Digest,* pp. 49-52, (2003)

[4] H.S. Yang, et al., *IEDM Tech. Digest,* pp. 1075-1077, (2004)

[5] K. Goto, et al., *Symp. VLSI Techn. Digest,* pp. 49-50, (2003)

[6] S. Pidin, et al., *Symp. VLSI Tech. Digest,* pp. 54-55, (2004)

[7] S.Pidin, et al., *IEDM Tech. Digest,* pp. 213-216, (2004)

[8] P.R. Chidambaram, et al., *Symp. VLSI Tech. Digest,* pp. 48-49, (2004)