REFERENCES

1. M.Bohr, S.U. Ahmed, L. Brigham, R. Chau, R. Gasser, R.Green, W. Hargrove, E. Lee, R. Natter, S. Thompson, K. Weldon and S.Yang. (1994). *IEDM Technical Digest*, p. 273

2. M.Bohr, S.S. Ahmed, S.U. Ahmed, M. Bost, T. Ghani, J.Greason, R. Hainsey, C. Jan, P. Packan, S. Sivakumar, S. Thompson, J. Tsai, and S. Yang. (1996). *IEDM Technical Digest,* p. 847

3. Hemant V.Deshpande, B. Cheng and Jason. C. S. Woo. (2001). *Deep Sub-Micron CMOS device design for Low Power Analog Applications.* Symposium on VLSI Technology Digest of Technical Papers, pp. 87-88

4. Chin, W. K. (2000). *Semiconductor Device and Failure Analysis Using Photon Emission,* John Wiley & Sons Ltd, Inc.USA.

5. Bin Yu, H. W., Concetta Riccobene, Qi Xiang and Ming Ren Lin. (2000). *Limits of Gate Oxide Scaling in Nano-Transistors.* Symposium on VLSI Technology Digest of Technical Papers, pp. 90-91

6. Hiramoto, T. (2004). *Advanced Device Structure for Aggresively Scaled MOSFET.* IEEE International Conference on Integrated Circuit Design and Technology, pp.59-64.

7. James e. Chung, K. N. Q., Charles G. Sodini, Ping-Keung Ko, and Chenming Hu. (2001). *The Effect of Hot Carriers on the Operation of CMOS Active Pixel Sensors, Electron Device Meeting*. IEDM Techical Digest, International, pp. 24.5.1-24.5.4.

8. D.A. Hodges and H.G. Jackson. (1980). *Analysis and Design of Digital Integrated Circuits,* McGraw-Hill Book Co.

9. Wolf, S. (1995). *Silicon Processing For the VLSI Era, Volume 3-The Submicron MOSFET,* Lattice Press, Sunset Beach, California, USA.

10. Ren-Ji.C, Sah C.T.(June 1983). *Two –Dimensional Numerical Analysis of the Narrow Gate Effect in MOSFET,* IEEE Transaction on Electron Devices, Vol.Ed-30, no.6

11. Hess, K. (2000). *Advanced Theory of Semiconductor Devices*, IEEE Press Marketing, 3 Park Avenue, New York, USA.

12. Qiang Chen, K. A. B, Evans M.Harrell and James D.Meindl. (2003). *Double Jeopardy In the Nanoscale Court?.* IEEE Circuits and Devices Magazines. pp.28

13. Shur, M. (1996). *Introduction to Electronic Devices,* John Wiley & Sons Ltd. Inc. USA.

14. James e. Chung, K. N. Q., Charles G. Sodini, Ping-Keung Ko, and Chenming Hu. (2001). *The Effect of Hot Carriers on the Operation of CMOS Active Pixel Sensors, Electron Device Meeting.* IEDM Techical Digest,International, pp.24.5.1-24.5.4.

15. E. P. Vandamme, P. J., and L. Deferm. (August 1997). *Modeling the Subthreshold Swing in MOSFET's.* IEEE Electron Device Letters, 18(8).

16. C. Hu. (1993). *Future CMOS Scaling and Reliability*, Proceedings of the IEEE. 81, p.682

17. Hayes.J (1980).*MOS Scaling,* Proceedings of the IEEE.

18. J. R. Brews, (1979). *IEEE Trans. Electron Dev*, ED-26(11), p1696

19. N. Kotani and S. Kawazu (January 1979) *Solid State Electronics*, 22, 63

20. C.Fiegna et al. (June 1994). *Scaling the MOS Transistor below 0.1µm: Methodology, Device Structures, and Technology Requirements,* IEEE Transactions on Electron Devices, Vol. 41, No.6

21. K.E. Kroell, (Sept 1973). *Geometry effects on field effect transistors, European Solid-State Device Res.Conf.(ESSDERC),* Munich,German

22. M.C. Jeng et al. (1987).*Performance and reliability of deep-submicron MOSFETs*, Tech. Dig. IEDM, p.710

23. B.Agrawal,V.K.De and J.D. Meindl, (1993). *Opportunities for Scaling FET's for Gigascale Integration (GSI)*, in Europ.Solid State Device Res. Conf, p. 919

24. R. Yan,A. Ourmazd, and D.Jeon. (Dec 1991). *Scaling the Si metal-oxide-semiconductor field effect transistor into the 0.1µm regime using vertical doping engineering.* Appl. Phys. Lett.,pp. 3315-3317

25. S.Thompson et al. (1998*). MOS Scaling: Transistor Challenges for the 21st Century,* Intel Technology Journal

26. A.Agarwal (2003) *Leakage in Nano-Scale Technologies: Mechanism, Impact and Design Considerations,* West Lafayette, USA

27. Xiao, H. (2001). *Introduction to Semiconductor Manufacturing Technology*. Pearson, Prentice Hall.

APPENDIX

APPENDIX A

1. MOSFET Fabrication Process Technology [27]

1.1 Self aligned gate process by using single Poly Double Metal 0.5um

Polysilicon is the most commonly used material to form gates and local interconnections. This is because polysilicon has high temperature stability, which is necessary for the self-aligned source/drain implantation and post implantation high-temperature anneal process. Single Poly Double Metal 0.5um normally is deposited with the LPCVD process, either SiH_4 or SiH_2Cl_2 can be used as the silicon precursor. Deposition temperature is from 550°C to 750°C, and it can be heavily doped with boron, phosphorus or arsenic, either in situ during the deposition or ex situ by an ion implantation process. **Figure 5.0** illustrates the self-aligned gate process for making a NMOS transistor.

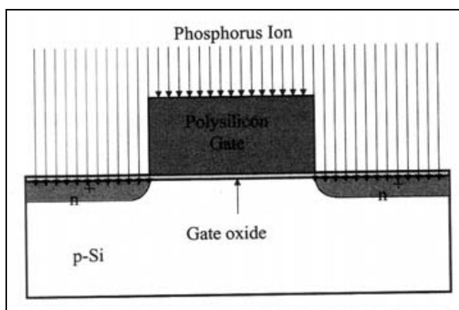

Figure 5.0: Self-aligned gate transistor formation

1.2 Locos isolation

Local oxidation of silicon (LOCOS) has been used in IC production since the 1970s.It has better isolation effect than the blanket field oxide .One advantage is that the silicon dioxide is grown after the channel stop implantation. The field oxide layer is self-aligned with the isolation doping area. By using the channel stop implantation; one can keep the same value of field threshold voltage, V_T while decreasing the field oxide thickness. The thickness of the LOCOS oxide is 5000 to 10, 000 Å. The LOCOS isolation process is listed and illustrated in **Figure 5.1**.

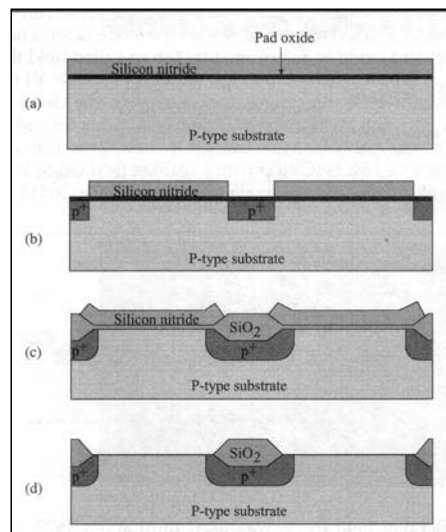

Figure 5.1: LOCOS Isolation Formation

LOCOS process uses a thin layer of oxide (200 to 500Å) as the pad layer to buffer the strong tensile stress of the LPCVD nitride. LPCVD silicon nitride is used as the oxidation mask, which allows only the thick silicon dioxide, called LOCOS, to grow at the designated area. The activation areas where transistors will be built are covered by nitride and do not grow oxide. Pad oxide is needed to relieve the strong tensile stress of the LPCVD nitride. Plasma etch with fluorine chemistry is usually employed for the nitride patterned etch, and hot phosphoric acid is generally used to strip the nitride layer. After the nitride etch, photoresist strip, and wafer clean, a thick layer of oxide (3000 to 5000Å) is grown on the area not covered

by silicon nitride. The silicon nitride is much better barrier layer than the silicon dioxide. Oxygen molecules cannot diffuse across the nitride layer, therefore the silicon underneath the nitride layer does not oxidize. On the area not covered by the nitride, oxygen molecules continuously diffuse across the silicon dioxide layer, where they react with silicon underneath to form more silicon dioxide.

1.3 Twin Well Technology

The p-well and n-well formed by the self-aligned twin well process are not at the same level, which could affect photolithography resolution because of the depth of focus problem. Double photo twin well is common in advanced IC chip manufacturing; the process steps are listed and illustrated in **Figure 5.2**.Both well implantation processes use high energy, low current implanters. Furnaces usually perform the well implantation annealing and driving processes.

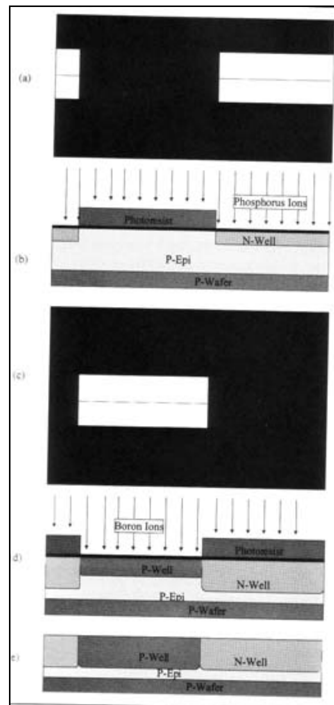

Figure 5.2: Double photo twin-well

1.4 Dry oxidation

Dry oxidation has a lower growth rate that wet oxidation; however, the oxide film quality is better than the wet oxide film. Therefore, thin oxides such as screen oxide, pad oxide and especially gate oxide normally use the dry oxidation process. The parameters for this process are 110Å oxide for the thickness and dry thermal at 980°C.

Usually there are two nitrogen (N_2) sources in an oxidation system, one for the process application, with higher purity and another with lower purity for the chamber purge. Because nitrogen is a stable gas, it does not react with silicon even at 1000°C. For dry oxidation, high purity oxygen gas (O_2) is used to oxidize silicon. Hydrogen chloride, HCL, is also used during the oxidation step to reduce mobile ions in the oxide and minimize the interface state charge. The dry oxidation process normally operates at about 1000°C. In dry oxidation, HCl is commonly used as a getter to remove mobile metallic ions, especially sodium, by forming immobile chloride compounds.

Now the system is ready for oxidation. Turning on the oxygen and anhydrate hydrogen chloride flows and turning off the nitrogen flow, makes oxygen react with silicon to form a thin layer of silicon dioxide on the silicon wafer surface. After the required oxide thickness is reached, the O_2 and HCl flows are terminated and N_2 flow is resumed. The wafers stay at the high temperature for a while to anneal the oxide. This step improves the quality of the silicon dioxide, makes it denser, reduces the interface state, and increase the breakdown voltage. The thin gate oxide can grow in a not be too short to control the process. After the oxide film grows, the film is annealed at > 1000°C in nitrogen ambient to improve oxide quality. After the oxide annealing, the furnace is gradually cooled down to its idle temperature and wafer boats are slowly pulled out from the furnace with a constant nitrogen flow.

1.5 Tungsten silicide process

Metal CVD is widely used to deposit metal in IC processing. CVD metal films have very good step coverage and gap-fill capability and can fill tiny contact holes to make the connections between metal layers. CVD metal thin films normally have poorer quality and higher resistivity than those of PVD metal thin films. Therefore, they are mainly used for plug and local interconnection and not applied for the global interconnections. The most commonly used metal CVD processes are tungsten, titanium, titanium nitride and for this fabrication, the tungsten silicide was selected. Most metal depositions are thermal processes; external heat or from the heating elements provides the free energy needed for the chemical reaction. In some cases, remote plasma sources generate free radicals and increase the chemical reaction rate.

Tungsten silicide deposition normally is used for the gate and local interconnection applications. Both SiH4 and SiH2Cl2 (DCS) are employed as silicon sources gases, and WF6 is the tungsten precursor.SiH4/WF6 chemistry requires lower process temperature (typically at 400°C), whereas DCS/WF6 chemistry requires a temperature of about 550 to 575°C.For SiH4/WF6 chemistry:

WF6 + 2 SiH4 → WSi2 + 6HF + H2………………………………………… (5)

This process is similar to the nucleation step of the tungsten CVD process. The main difference is the ratio of the SiH_4/WF_6 flow rate; when the ratio is lower than 3, the chemical replacement reaction deposits silicon-rich tungsten instead of tungsten silicide. To ensure tungsten silicide deposition, the SiH_4/WF_6 flow ratio must be larger than 10.The DCS/WF_6 chemistry can be expressed as:

$2\,WF_6 + 7SiH_2Cl_2 \rightarrow 2WSi_2 + 3SiF_4 + 14HCL$………………………………… (6)

DCS/WF_6 based process requires higher deposition temperature. It has a higher tungsten silicide deposition rate and better film step coverage than SiH_4/WF_6 based process. It also has much lower fluorine concentration in the film and fewer film peeling and cracking problems due to its lower tensile stress. DCS/WF_6 silicide process is gradually replacing the silane-based process. The advantages of WSi_x over $TiSi_2$ are fewer process steps and its easier integration with polysilicon deposition in one process tool. However, it has higher resistivity than titanium silicide process can from titanium silicide on the gate and source/drain at the same time.

1.6 P-LDD/N-LDD CMOS process with anti-punch-through implant

1.6.1 Lightly Doped Drain (LDD)

When the gate width is smaller than 2 microns, the vertical component of the electric field induced by the bias voltage between source and drain may be high enough to accelerate electrons tunneling through the thin gate oxide layer. This is called the hot electron effect, which can affect transistor performance from the gate leakage and cause reliability problems for the IC chips because of the trapping of electrons in the gate oxide. **Figure 5.3** illustrates the hot electron effect of the MOS transistor.

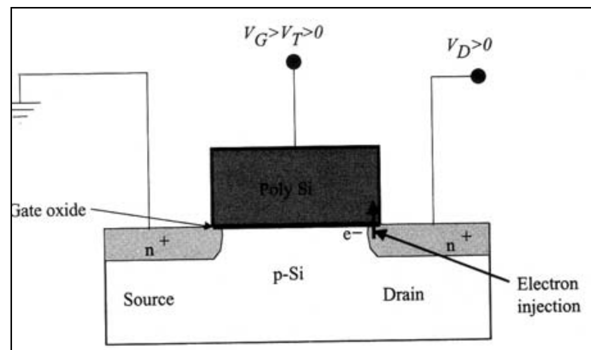

Figure 5.3: Hot electron effect of the MOS transistor

The most widely used method to suppress hot electron effect is called lightly doped drain (LDD), as shown in **Figure 5.4**.

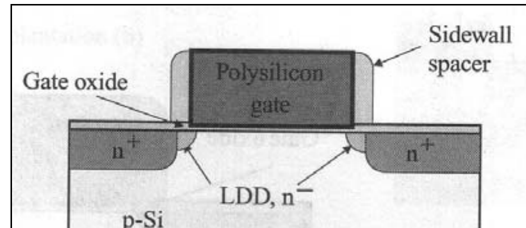

Figure 5.4: LDD of MOS transistor

An LDD junction can be formed by using low energy, low current implantation. It is a shallow junction with very low dopant concentration, extended just underneath the gate. After depositing and etching back the dielectric layers, sidewall spacers are formed on both sides of the polysilicon gate. High current, low energy ion implantation forms the heavily doped source/drain junctions, which are kept apart from the gate by the sidewall spacers. This reduces the vertical component of the source/ drain bias induced electric field and reduces the available electrons for tunneling, thus suppressing the hot electron effect. Transistors with LDD can be made in process steps listed and illustrated in **Figure 5.5**.

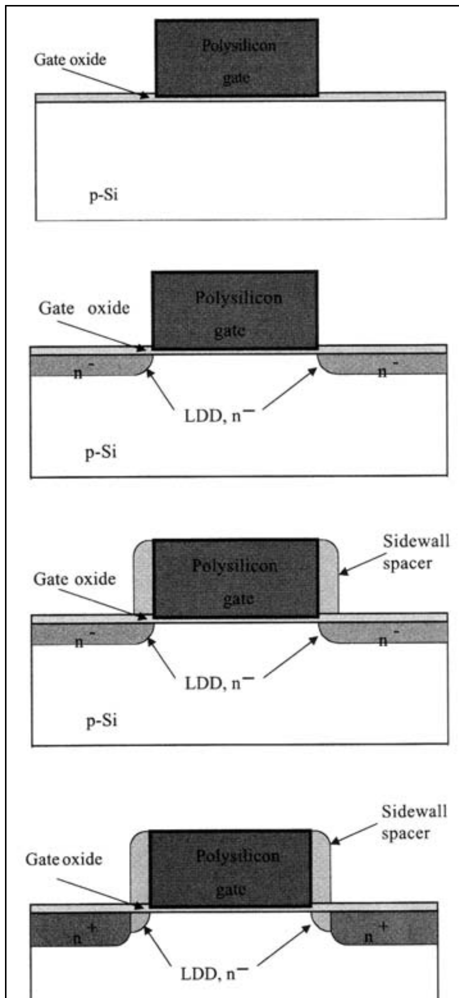

Figure 5.5: Process step to make a MOS transistor with LDD

When transistor feature size reduces to sub 0.18 micron, and the power supply voltage drops to 1.5 volt, the hot electron effect may not be so important anymore. The LDD implantation process probably is no longer needed. However, the sidewall spacers are still needed to provide a diffusion buffer for the dopant in the source/drain junction. Otherwise, the dopants atoms in the source and drain could diffuse too close to each other during the post-implantation annealing process because of the small gate dimension, as shown in **Figure 5.6**.

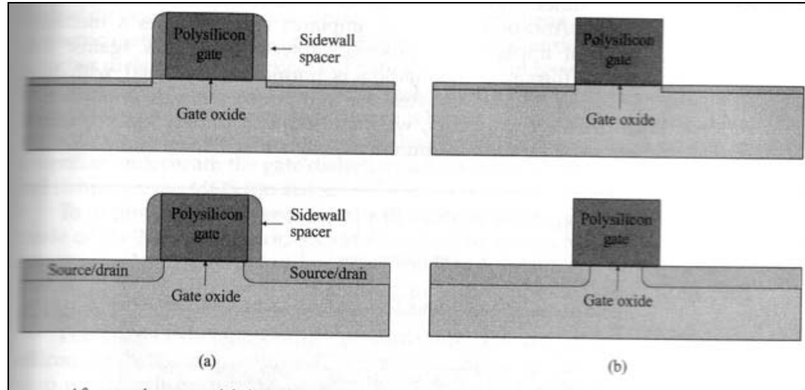

Figure 5.6: Source-drain annealing (a) with and (b) without sidewall spacer

1.6.2 Anti-punch-through process

A punch-through effect occurs when the depletion regions of the source and drain short each other under the influence of both gate-substrate bias and source-drain bias. Anti-punch-through implantation, which is a medium-energy, low-current implantation process, protects transistors against this effect. Anti-punch-through implantation is normally performed with well implantation. **Figure 5.7** illustrates the anti-punch-through implantation process. With the MeV implanter, well, anti-punch-through, and V_T adjust implantations can be performed simultaneously after STI formation.

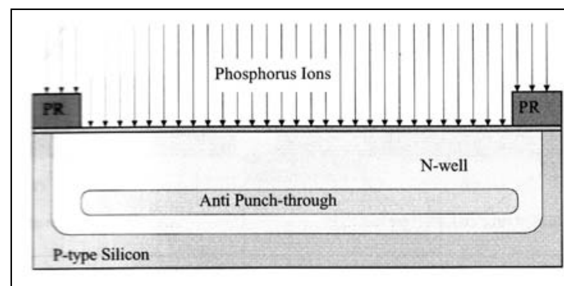

Figure 5.7: Anti-punch-through implantation process

1.7 p+ Halo implant

Another process commonly used to prevent punch-through effect is halo implantation. It is a low-energy, low current implantation, process with 45° incident angle. The halo junction formed in this implantation process can help to suppress punch-through effect. **Figure 5.8** illustrates a halo implantation process.

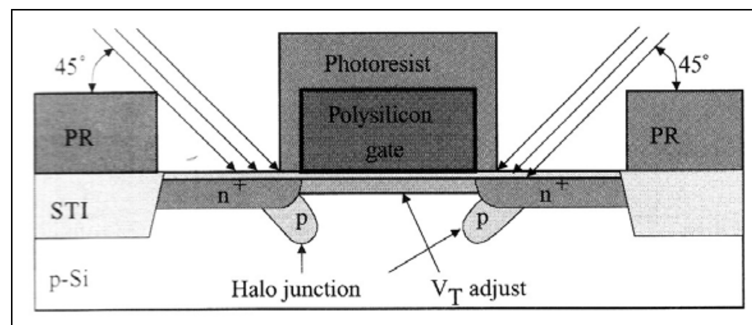

Figure 5.8: Threshold adjustment implantation process

1.8 Back End, Al-Si=Cu Multilevel Interconnect with W-plug Process

When the number of transistors increases, one layer of metal interconnection cannot route all the transistors on the chip. Multilayer metal interconnection started to be used at this point. The earlier interconnection process always left a rough surface, which caused problems in photolithography and metal deposition. As device feature size reduces, there is no room for wide, tapered connection, which allows aluminum PVD to cover the bottom of the contact holes.

Tungsten CVD process has been introduced to fill the narrow contact and via holes. The basic interconnection process steps are dielectric CVD, dielectric planarization, dielectric etch; tungsten PVD, bulk tungsten removal, metal stack PVD, and metal etch. The most commonly used dielectric is silicate glass, both doped such as PSG or BPSG for PMD, and undoped (USG) for IMD. Dielectric

planarization can be achieved with thermal flow (only for PMD), etch back, and CMP.CMP became more popular in dielectric planarization and bulk tungsten removal applications in the late 1990s.Dielectric etch forms contact and via holes. The most commonly used metal stack is titanium as the welding layer, aluminum-copper alloy as the main conductor, and titanium nitride as the antireflective coating (ARC). Metal etch defines the metal interconnection lines. **Figure 5.9** describes the interconnection between metal 1 and metal 2.The interconnection process steps for metal 3, metal 4, and up to metal 7 are almost identical to these process steps.

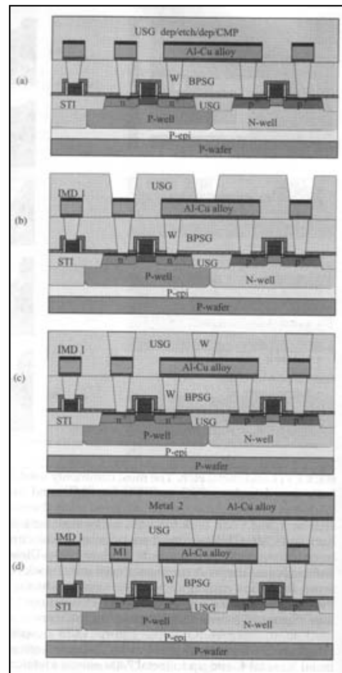

Figure 5.9: “Traditional” interconnection process step

APPENDIX B

2.1 Source/Drain Engineering

Source drain junction depth is another important consideration. The substitution of arsenic for phosphorus as the dopant element has provided considerable junction depth scalability. The 0.12µm depths required by factor of 10 scaling are well within the capability of a high current arsenic implant process particularly when complemented by laser or electron beam annealing techniques.

Scaling the thermal oxide will make low temperature, high pressure oxidation economically feasible; eliminating the oxidation induced stacking faults and continental drift effects associated with high temperature oxide formation. Scaling both the intermediate passivation and the gate electrode materials should also be straightforward, thus helping better meet the associated pattern definition and dimension control requirements. Aluminum alloys providing the necessary electromigation properties will probably evolve, as will barrier metals avoiding "alloy spiking" through shallow,
scaled, source/drain junctions.The undesireable effects of simply reducing MOS channel length can be overcome by scaling down the vertical and horizontal dimensions while at the same time proportionally decreasing voltages and increasing the substrate impurity level.

Several additional problems arise from the need to scale vertical dimensions. The cross section of part of a simple MOS integrated circuit as shown in **Figure 8** indicates the six basic vertical dimensions that require scaling. There are polysilicon, gate oxide, junction, metal, intermediate passivation and thermal field oxide.

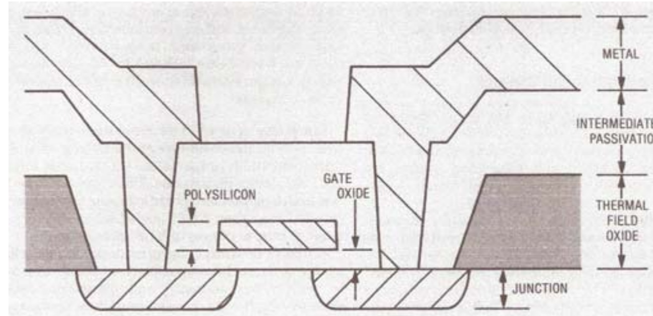

Figure 8: MOS IC cross-section, showing the six vertical dimensions which require scaling [17]

The source/drain extension (SDE) depth and gate overlap for MOSFETs of 1μm and below is shown in **Figure 9**. The SDE is the shallow diffusion that connects the channel with the deep source and drain. Junction depth always refers to the SDE junction depth. The deep source/drain junction depth is held constant. Overlap is defined as the distance the SDE extends under the gate. The metallurgical spacing is the distance between the source and drain SDE.

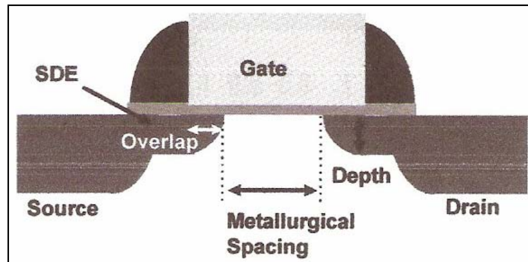

Figure 9: Terminology of source drain junction [25]