Project Title: Trigate MOSFET/FinFET or Silicon Nano Wire MOSFET: Structure, Operating Principles and Characteristics

Course: Solid State Devices

Code: ELEC 6221

Department: Electrical and Computer Engineering

Submitted to:

Dr. M. Z. Kabir

By:

Khosa Komal Bir Singh (27135173)

Chunhui Ji (27140452)

2nd December 2015

Contents:

Page No.

List of Figures:

Page No.

Abstract:

The report highlights the features of the **22nm Intel transistor** and its construction and how it came into picture. The starting few section of the report discuss the beginning of the processors by Intel. The next part of the report discusses the 22nm technology which is **3D FinFET** technology with few fabrication details. 22nm technology is codenamed **Ivy Bridge.**

It is necessary to have a look at the competition in the market and hence the report highlights some of the important points about its competitors like **AMD**. Finally in the last sections, the future of the Intel technology is discussed. Some important points about the latest technologies like **Haswell and Knights Landing** have been mentioned. In an attempt to follow the **Moore's Law**, Intel is continuously trying to come up with new techniques of semiconductor fabrication which have helped to reduce the overall size of the chip and to reduce the total weight of the system.

Introduction:

Intel is a momentous name in the field of processor technology. The processors have advanced a lot if we compare the processors of the present day to the processors of the past. One of the earliest processors from Intel was introduced in 1971. The name of the earliest processors by Intel is **4-bit 4004 (1971)** and the list continues with the more advanced processors in the upcoming years and the present day processors include Core i7, Xeon E3 and E5 series (2015).

The processor technology follows the **Moore's law** according to which the number of transistors on a dense IC doubles every 2 years. The period of 2 years is often quoted as 18 months and the Moore's law is not a hard and fast rule.

The Intel processors have advanced from the 32nm technology to the present 14nm technology i.e. from the micron technology to the nanometer technology. The report discusses the Moore's law in some detail in the first section and then the later sections list some of the important concepts and landmarks in the Intel processors, like the atom processor, Pentium processor, 2^{nd} generation and 4^{th} generation processors, thermal sensors in the processors etc.

Moore's Law: [1]

As already stated, the time period of 2 years in Moore's law is often quoted as 18 – 24 months. This is due to the fact that if we want to increase the number of transistors on a given area of the chip, we have to use new techniques to accomplish this since the aspect ratio is decreasing with the new technology. As the transistor technology is reaching below 20nm, the packing of more and more transistors on the smaller area is becoming more difficult.

The node dimensions (20nm, 14nm etc.) now do not give an idea about the actual performance of the transistor. It is because even in the 20nm transistor, there can be few dimensions which are smaller than 20nm and some which are larger than 20nm. Thus the gate length cannot be determined just from the given node name.

The limit on the number of transistors that can be packed on a silicon chip depend on the distance between the transistor gates and the copper wires on the back of the chip.

The extremely small features can be printed on the wafer but this requires very sophisticated lithographic techniques. One of the method used in this case is double patterning. Intel used this technique in the formation of transistors on 22nm chip. In the case of double patterning, the patterning step is divided into two parts and each step can be slightly misaligned from the other.

- In order to achieve the maximum density, double patterning cannot be used. So Intel moved back to single patterning. Also the double patterning method takes more time and cost. Hence it is not a feasible solution for the industry. All these factors have proved that the density of the chips does not doubles with each generation but the ratio is somewhere from 1.6 to 2.
- The patterning technology presently uses the 193nm laser technology. As the demand for smaller chips grows, the problem of wiring comes up. The problem is more pronounced if the size grows goes to below 10nm. It is due to the fact that the copper wires on the chip require a

barrier material to prevent unwanted contacts with the surrounding material. This in turn pushes the size of the copper wires to become smaller and smaller.
This thickness of the wires is a deciding factor to know how closely these wires can be placed. Thus the attempt to make the copper wires thinner and thinner, the resistance and the delays of the circuit increase and hence it adversely affects the performance.

- Thus, in order for the Moore's law to survive, we have to have new technologies to overcome these problems.

Intel Processors: [2]

Intel has developed a number of processors in the past and the list still continues. Starting from the 4-bit processors, the present day processors from Intel are 64-bit. The technology in each of these processors kept on evolving and an overview has been shown in the following figure from Intel:

Fig 1: [2] List of Intel Processors with Other Details Involved

Intel's Tri-Gate MOSFET Technology: [3]

The short channel effects come into picture at the sub-micron level. When the transistor technology reaches the channel length of about 10nm, the multi gate devices are used to avoid the short channel effects. According to the research carried out in this field, it has been studied that the tri-gate MOSFETS are the most efficient type of the multi-gate devices that can reduce the short channel effects in the transistors upto 10nm of the channel length. This is possible due to the capability of the tri-gate MOSFET to have better gate control over the channel and hence presenting a high current flow capability.

At the extremely small channel length of 10nm, the quantum effects have to be taken care of. Due to the quantum effects present, the quantization of energy of the electrons is an important concept. The complete model which describes all the effects in the MOSFET at the sub-micron level is still under study and is very difficult to come up with due to the small distances and the quantum effects involved which render the general laws of physics useless at the sub-micron level.

- The device dimensions kept on reducing with time and with this reduction, there were other problems that came up. Because of the quantum effects, there is velocity saturation and hence the drive current is limited and also the gate leakage became major concern.
- The velocity saturation has been overcome by straining the lattice in the channel region so that the carrier mobility is increased. Intel used this technique in its 90nm, 65nm, 32nm and 28nm processors.
- The gate leakage, on the other hand, is a more serious problem. It has been tackled to some extent using Hafnium (Hf) based dielectrics. Intel used this technique in its 45nm processors. The leakage in this case is reduced by a factor of 10.

The addressing of these problems has enabled Intel to produce **High K/Metal Gate (HKMG) processors**.

Introduction of Ivy Bridge Structure: [4]

In addition to the above problems, there is still the problem of short channel effects as explained earlier and to overcome these issues, Intel had the option of either using the fully-depleted Silicon-on-Insulator or to use the FinFET technology. Both the technologies have their own merits. Intel has chosen the tri-gate form of FinFET on a bulk substrate to overcome this issue.

Intel introduced its first tri-gate processor in **April 2012** and the channel length was 22nm. The 22nm transistor technology is codenamed by Intel as **"Ivy Bridge"**. One of the general structure to show the inside of such a MOSFET is given in Fig 2. The left side shows the NMOS part with 3 fins and the right side shows the PMOS part with 4 fins.

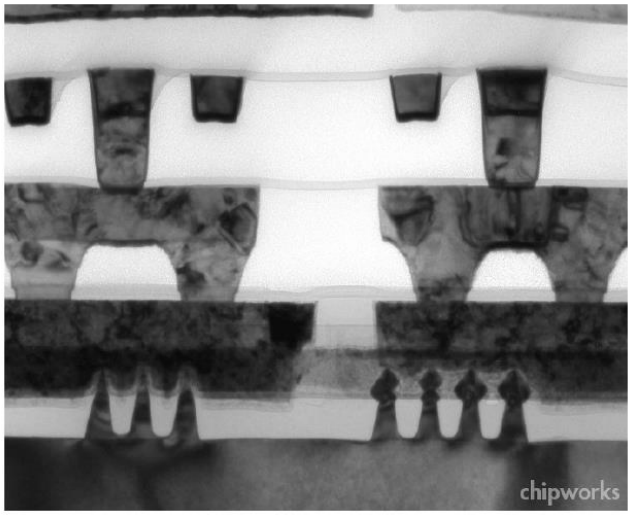

Fig 2: [4] One of the First tri-gate Processors by Intel (NMOS on left (3 fins) and PMOS on right (4 fins))

Intel's Tri-Gate Technology (22nm):

The tri-gate transistor used by Intel in its 22nm technology is far better than its predecessors because of the FinFET technology used it. The structure gives higher conduction capabilities and also better control over the channel by the gate.

22nm Transistor Structure: [5,6]

The **"Fin"** in FinFET is actually the thin Si "fin" from the body of the device, and it covers the channel. The thickness of the fin is measured in the direction from the source to the drain and this gives the effective length of the channel.

The fin height for this technology is 34nm while the width, W_{si} , is 8nm, the gate length is drawn at 34nm and the gate pitch is 90nm. The W_{si} is very narrow in this case and it allows for better gate control of the channel. This gate control also reduces the channel doping density which is required to control the short channel effects. This in turn provides the better control over the metal gate work-function and this reduces the gate oxide field which improves the reliability of the gate oxide. Hence the resulting devices is much better in terms of reliability, performance and power usage.

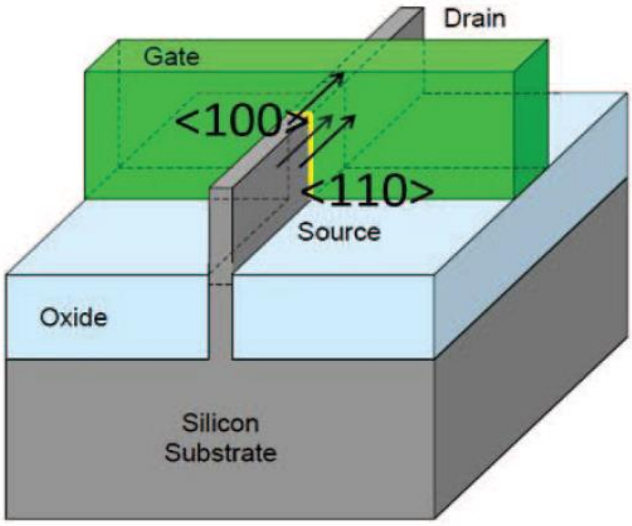

Fig 3: [6] Tri-gate Transistor with <110> Sidewalls and <100> Top Surface

This technology by Intel has various advantages over the previous generation technologies. This section of the report lists some of these features and the reasons why they are important. The Intel processors which use the tri-gate technology fall under the 3rd generation of Intel processors using high K metal gate technology. Some of these features of these processors are:

1. **Fin and Gate Processing:**
   The advantage over the previous generation comes from the fact that this technology provides a better means to control the metal gate work-function so that it can be brought closer to the middle of the bandgap. The benefit of this is that now the field across the oxide is reduced. Also this method provides gate stack optimization and hence this optimization and the use of the fin leads to the area scaling and this shows a relationship between fin and gate processing.

2. **Bias Temperature Instability (BTI):**
   BTI in this type of the transistor is improved significantly. BTI refers to the condition where there is the increase in the threshold voltage followed by the decrease in the drain current and hence it affects the transconductance of the MOSFET. Both NMOS as well as PMOS are affected by this. Due to the combined effects of gate stack optimization, Metal Gate (MG) work-function (WF) control and oxide scaling, this problem can be easily solved. The effect is compared in the previous 32nm and the present 22nm technology and it can seen in the curve that for the 22nm technology, NMOS is significantly improved. The tri-gate technology shows a significant improvement over the 32nm planar technology.

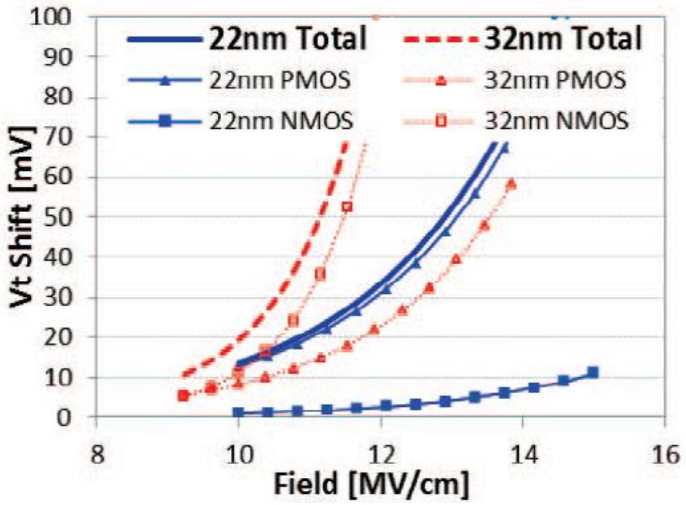

Fig 4: [6] BTI effect compared for 32nm and 22nm Technology

3. Degradation by Hot Carriers:

The degradation due to the hot carriers is much less in the case of 22nm technology than in 32nm technology. This can be attributed to the tri-gate fin architecture and the reduction in the channel length. This degradation can increase or decrease with the fin width.

- With the change on the fin width, the electric field is modified and the effect on the electric field is further enhanced by the doping levels as well as by the junction profiles which are directly affected by the fin width.
- The efficiency is affected by the fact that the gate oxide can capture the scattered carriers. The capturing is also dependent on the fin width. Narrow fin width have a higher probability that the gate oxide will capture a carrier. This process can be seen in the following figure:

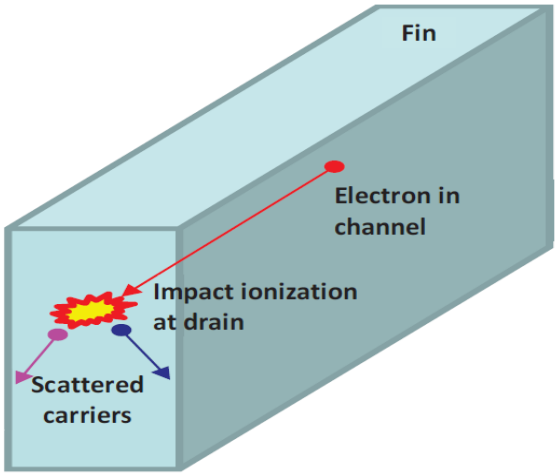

Fig 5: [6] Hot Carriers are more likely to be captured by Gate if FIN Width is Narrow

- Also if the fin width is narrow then the conduction of the heat out of the channel makes the fins hot.

  All these effects and the different values of fin width can lead to different results for the degradation by the hot carriers.

4. **Body Bias Affects:**
   The application of the higher body biases does not have any effect on the functioning of the device because the tri-gate transistors are fully depleted. Also many experiments are conducted for this particular feature and it has been found that as long as the body bias remains below the normal operating voltage, the BTI degradation is very small and the reason for this is that the BTI is dependent on the electric field effects present at the gate oxide interface (this study is conducted for the PMOS BTI).
   Also, as discussed earlier in the report that the gate has a very good control over the channel and hence the hot carrier degradation in the case of NMOS cannot have any adverse effects for the high body biases.

Intel's Present Day Processors (as explained by Intel): [7]

This section of the report will explain the structure of the 22nm transistor which is codenamed "Ivy Bridge" by Intel and finally both the 32nm and 22nm transistor will be shown side by side to show the structural difference.

1. The traditional planar technology of Intel made use of conducting channel on the silicon surface under the gate electrode in the "ON" state.

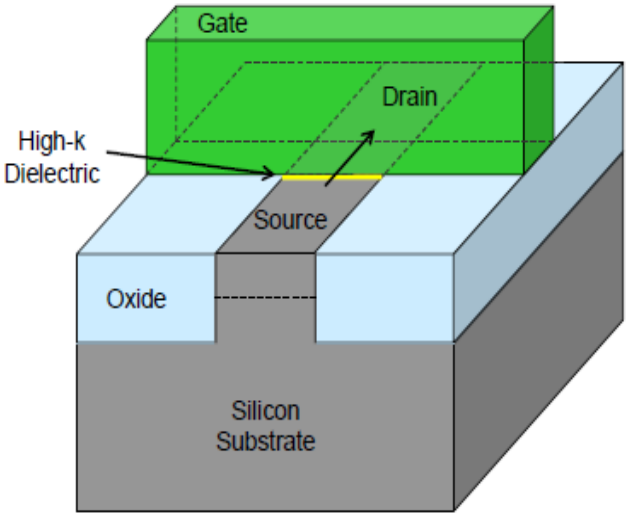

Fig 6: [7] 2-D Planar Technology Transistor by Intel

2. The tri-gate transistors on the other hand, form conducting channels in 3 directions of the vertical fin and hence they provide the **"fully depleted"** operation. Thus with the 22nm technology, Intel entered the 3-dimension technique of transistor manufacturing.

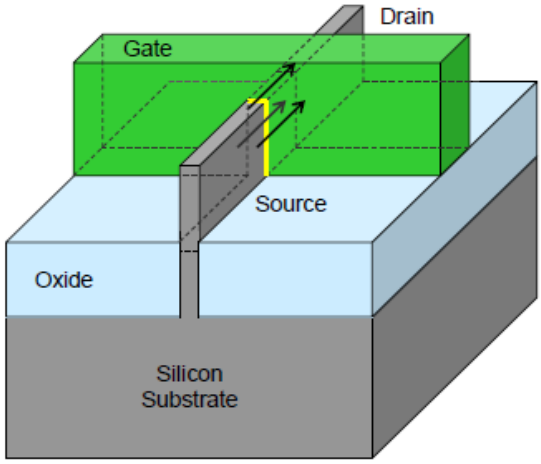

Fig 7: [7] 22nm Tri-Gate Transistor (3-D Tri-Gate Transistors)

3. The tri-gate transistors can have more than one fin and this can improve the overall performance.

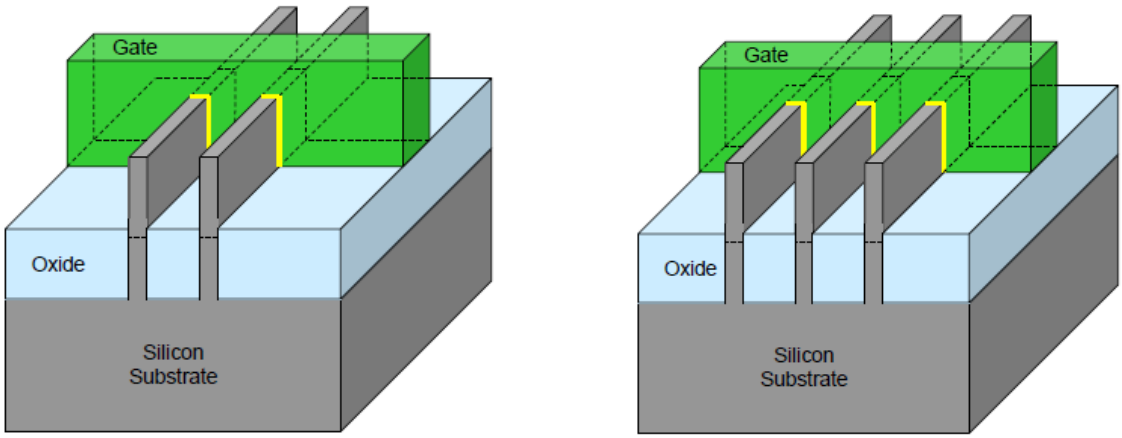

Fig 8: [7] Tri-Gate Transistor with more than One FIN

4. The following figure shows the planar 32nm and 3-D 22nm transistor built by Intel.

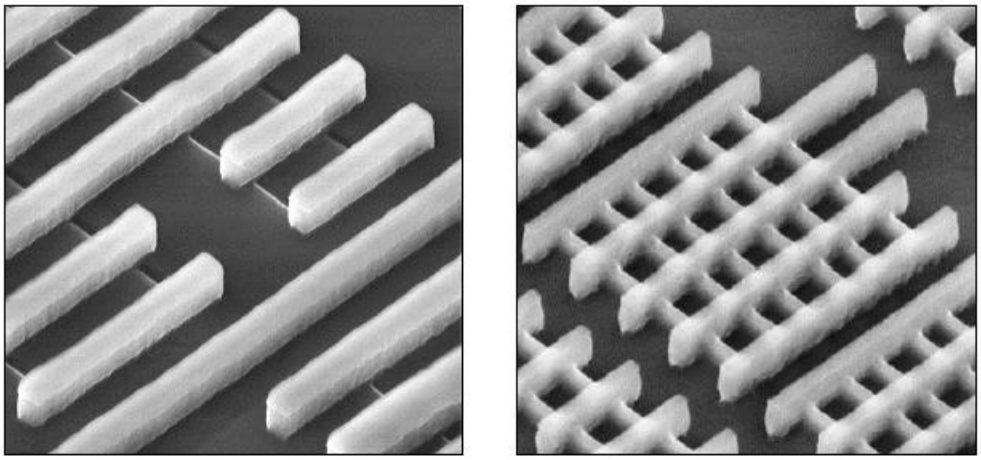

Fig 9: [7] Figure shows the 32nm Planar Transistor (left) and the 22nm Tri-Gate Transistor (right) by Intel

5. The development of the transistors over these years can be seen in the following figure. In the following figure it is visible how the tri-gate transistors came into picture. Intel introduced the strained silicon transistors in 2003 with the channel length of 90nm and now the market has Intel processors which have the transistors with the channel length of 22nm or even less.

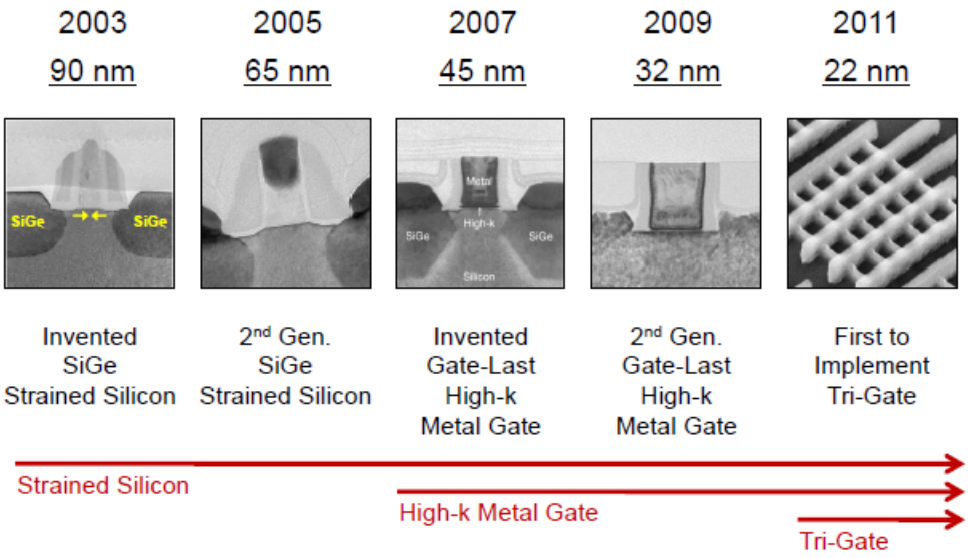

Fig 10: [7] Transistor Channel Length variation over the years by Intel

From Standard Bulk Transistors to Fully Depleted Transistors: [7]

The movement to the fully depleted transistors is a result of the degradation of the turn-off characteristics due to the substrate voltage in the case of the silicon transistor.

1. The degradation is caused due to the inversion layer getting affected by the body voltage.

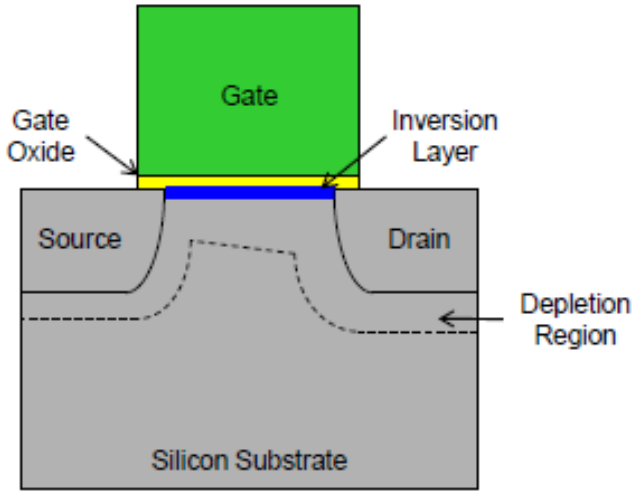

Fig 11: [7] Bulk Transistor, the Inversion Layer is affected by the Substrate Voltage

2. Intel tried using the **partially depleted Silicon-on-Insulator (PD-SOI)** but due to the floating body voltage effects on the inversion layer, this design has also been discarded.

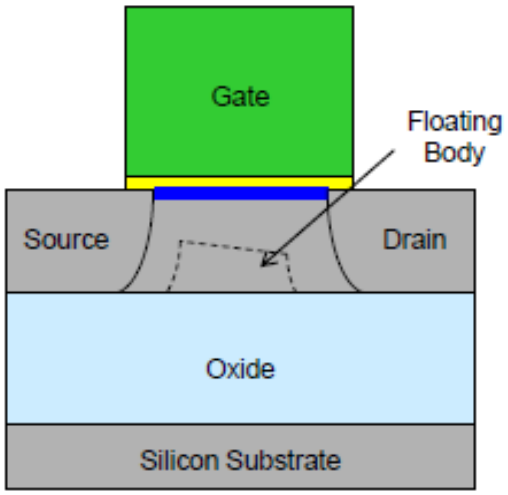

Fig 12: [7] Partially Depleted SOI, Floating Body Voltage influences Inversion Layer

3. The fully depleted transistors involve high cost and hence Intel is using the fully depleted transistors but with the tri-gate formation. In this case the gate can have better control over the channel. The cost is also reduced.

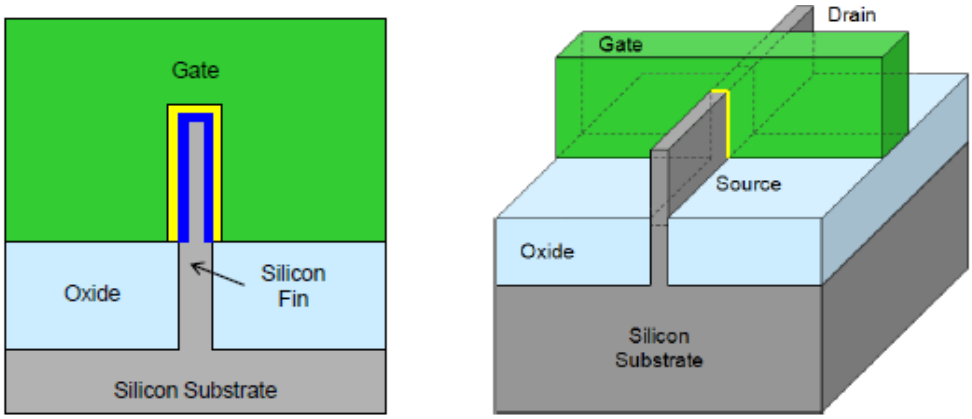

Fig 13: [7] Fully Depleted Tri-Gate Transistor used by Intel

Transistor Characteristics: [7]

Using the **101 transistor**, the transistor characteristics are studied and the following graphs are obtained. These curves are obtained by Intel in order to study the benefits of the fully depleted transistors over other types.

1. It is found that the use of the tri-gate technology with the fully depleted transistors can help reduce the leakage current and an improved subthreshold slope which means a better gate control.

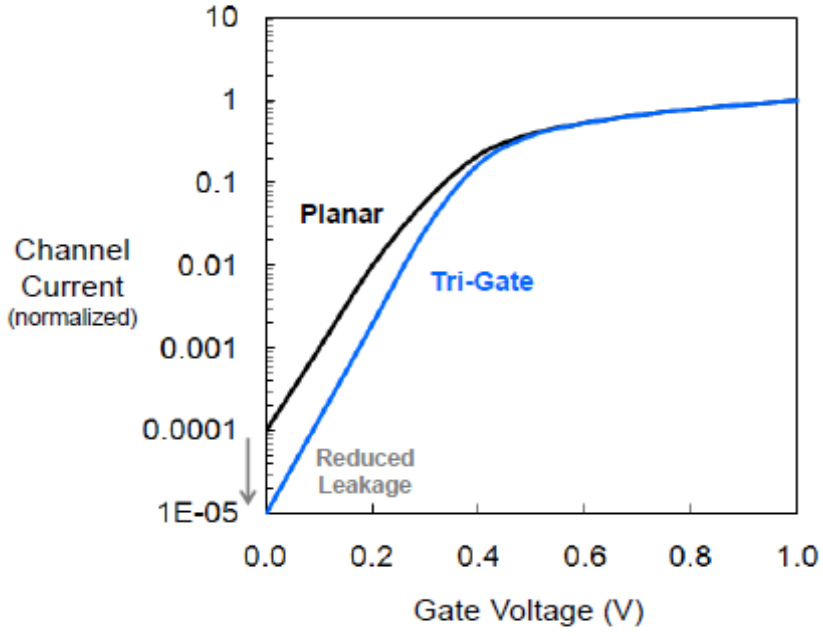

Fig 14: [7] Fully-Depleted Tri-Gate Transistor Characteristics

2. Another important point that is to be noted about these kind of transistors is that the improved sub-threshold slope can result in lower threshold voltage and hence it allows the transistor to operate at lower voltage and hence improving the speed and reducing the power consumption.

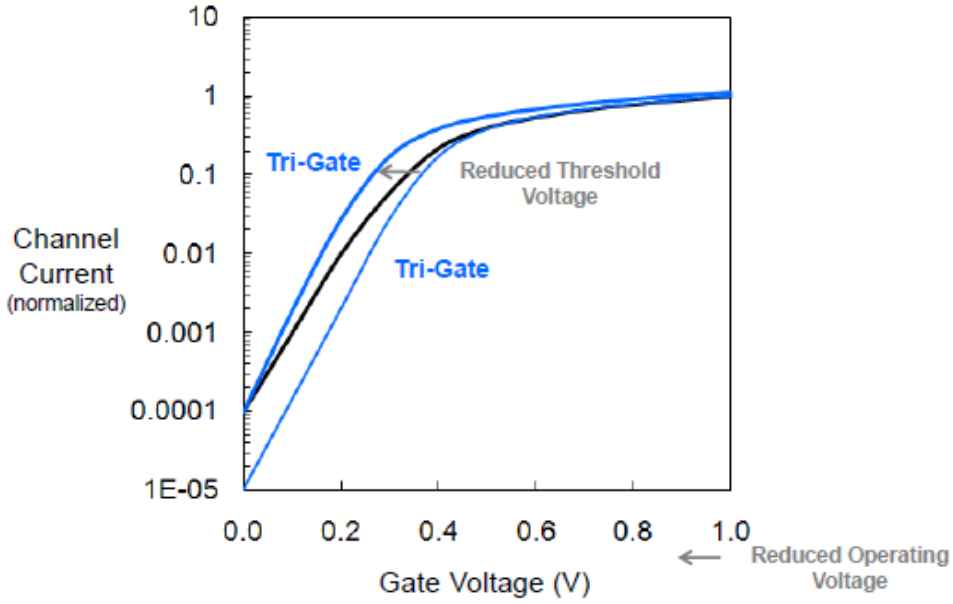

Fig 15: [7] Curve showing how Fully-Depleted Tri-Gate Transistor can improve Performance

3. The gate delay has been found to improve in the case of 22nm transistor due to the lower working voltage but the performance still needs to be improved further. This is so because the performance is better at lower voltage as compared to that at higher voltage as shown in the figure below (right). Thus at the lower voltage, the active power is reduced by almost 50%.

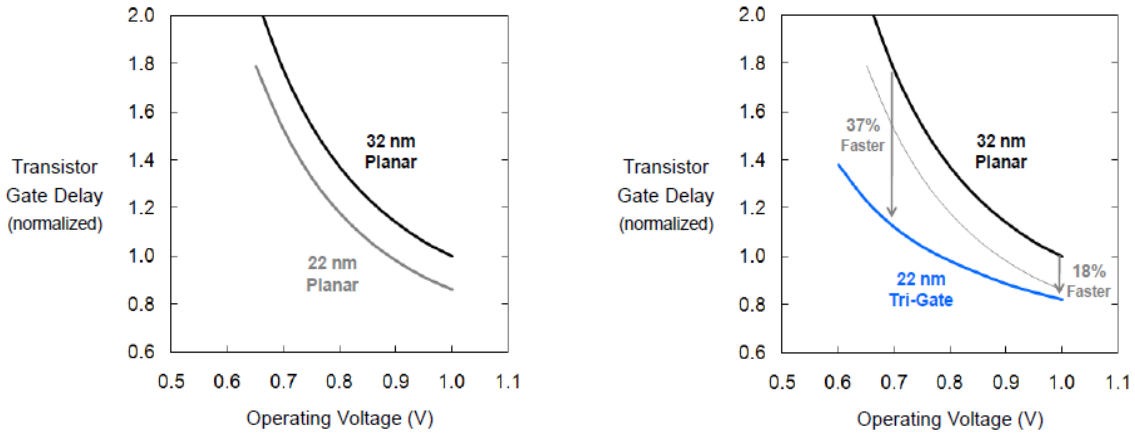

Fig 16: [7] Improved Gate Delay for 22nm as compared to 32nm

Intel vs AMD: [8]

When talking about the processors, there is another name which is quite famous for making good quality processors and that is "AMD". AMD processors are not equally famous as the Intel processors but they are one of the competitors of Intel. In this section of the report, a comparison is made between the two of the biggest processor manufacturing companies: Intel and AMD. The comparison is based on one of the **article on the FORBES website.**

1. **Intel and Advanced Micro Devices (AMD)** have diverged into two different markets of processors. These companies were once pioneers in the field of microprocessors but Intel's introduction of the 3D technology in 22nm transistors has enabled it to surpass AMD.
2. One of the major reasons why there is less competition in the field of the semiconductor fabrication technology is that the total investment has increased from the net value of $2-3 billion to $10-12 billion. Hence, Intel was more capable of remaining in the field as compared to AMD.
3. AMD has remained ahead of Intel until Intel came up with its own 64-bit design. AMD is the first among the two to produce the 64-bit processor, the processor with 1GHz speed, and also the dual core CPU.
4. Thus, AMD went completely out of competition when **Globalfoundries** bought its factories in 2009. Since then AMD has remained out of the fabrication process and has acted more like a vendor of the processors like **Qualcomm, nVidia and ARM holdings.**
5. AMD is now concentrating on other market shares where it can position itself as a giant. One such sphere is the graphics processors like **Carrizo and Carrizo-L.** Also the fast growing market of mobiles and tablets have seen a rapid increase in the AMD processors being used in these devices. Intel is trying to make a position for itself but is unsuccessful to do so.
6. One of the other devices that are using the AMD technology are the gaming consoles and companies like **Sony, Microsoft and Nintendo** use AMD technology.

Thus Intel and AMD will continue to compete each other in various fields but for now it is clear that the two companies are trying to establish themselves in different areas.

Future of Intel Processors:

This report shows how Intel has established itself in the processor market and how it achieved the current technology in transistors. Although Intel is ahead of its competitors for sure but the research continues at Intel and now the research is taking place for the transistors with the channel length of 14nm or even less.

1. **Haswell: The Fourth-Generation Intel Core Processor [9]**

- **Haswell** can be thought of as an improved version of the **Ivy Bridge** which is explained in the previous sections of this report.
- This type of the processor uses the improved tri-gate transistors to reduce the leakage current by almost 2 to 3 times. This is possible because of the **11 metal interconnect layers present in Haswell as compared to 9 in the Ivy Bridge technology.**

- Haswell uses the **System on Chip (SoC)** technology which saves a lot of space and also the voltage regulator is improved which further saves space and also cost.
- The new core design in new technology uses the **Intel TSX (Transactional Synchronization Extensions)** synchronization method in case of parallel workloads. Intel TSX is nothing but the extra addition of the hardware so that the multithreaded applications can be run smoothly.

In doing so, it should be ensured that the threads coordinate properly and hence this is made sure by a protocol. These protocols see that the threads take the right value to avoid any error and they do not access the critical sections. A software construct called **LOCK** prevents this.

2. Intel's Xeon Phi: [10]

Xeon Phi is one of the latest coprocessors by Intel. **Knights Landing** is the second generation coprocessor after **Knights Corner** which was announced in 2011 and used 22nm technology. On the other hand, Knights Landing uses **14nm technology** and it was announced in 2013.

- The coprocessors are actually the processors which help the main processor to perform various tasks and hence they can accelerate the system. But the functioning of the coprocessors is totally dependent on the main processor because the coprocessor cannot function on its own to fetch instructions, manage memory etc.
- **Knights Landing uses 61 cores** and has a memory of **16GB and it runs the Linux OS.** This coprocessor is about four times more power efficient and gives a better parallel computing capability. The speed is usually around **1TFLOPS (1 Tera-Floating point Operations Per Second).**
- Although **Knights Corner** is a coprocessor but Knights Landing can act as a processor.
- The micro architecture is based on the 14nm technology and the **Intel Atom core.**
- One of the latest processor in this line is **Knights Hill.** Knights Hill will be based on **10nm technology.** It is the **third generation MIC (Many Integrated Core)** architecture processor. It is supposed to be available somewhere in the last quarter of 2015.

3. Technologies for Extreme Performance:

Intel's Haswell model is termed as **'tick'** by Intel and the next model they introduced in this line is Broadwell which is termed as **'tock'.** Thus both of these models comprise the tick-tock design model.

a. Broadwell: [11]

- Broadwell is based on **14nm technology.** It is the advanced version of Haswell which shrinks the die size and hence reducing the overall size of the chip.
- Broadwell can be incorporated into the server models like Xeon hence giving a high performance speed of **3.8 GHz for Core i7 5775R.** Broadwell based processors were introduced in September 2014 and since then laptop companies like Lenovo are already using it.
- Broadwell is said to be the **fifth generation** in the processors.
- Broadwell is capable of supporting high end graphics and video.

b. Skylake: [12]

- It is the next model in the Intel processors line after Broadwell. It serves as **'tock'** in the tick-tock design scheme.
- It is also based on the **14nm technology** as Broadwell.
- Skylake is termed as the **sixth generation** of the processors. It was launched recently in August 2015.
- One of the major difference in Skylake and Broadwell is that Skylake processors use less power and give higher performance for both GPU and CPU.
- Maximum speed available using Skylake can reach upto **4.0 GHz for Core i7 6700K.**

c. [13] Intel has already come up with the plans to introduce successor to Sklake processors. It would be termed as **Kaby Lake.** Originally, Skylake was supposed to be succeeded by a **10nm technology** named **Cannonlake acting as 'tick',** but the plan has been postponed.
The upcoming processors will provide a **better processing of the 4K videos** and better **3D graphics** in addition to providing **USB 3.1 support** on the chip and support for both **DDR3L and DDR4 SDRAM.**

Conclusion:

- Starting from the **1st generation** of processors in desktops codenamed **"Lynnfield" (45nm)** and **"Clarkdale" (32nm)**, Intel came up with **2nd generation** of the processors codenamed **"Sandy Bridge" (32nm)** and with the introduction of the **3rd generation** of the processors codenamed **"Ivy Bridge" (22nm)**, Intel entered the 3D level of the transistors i.e. **FinFETs**.
- The report explained the various techniques used by Intel in its processors. It is clear that Intel has been trying to follow the Moore's Law as much as possible. It has always come up with a new technique that could meet the growing needs of the industry. Apart from all the competitions in the market, it has established itself as one of the leading semiconductor fabrication company.
- But coming up with new and innovative ideas is not always easy. Because of the decreasing size of the electronics, the components are getting smaller and smaller and hence the usual manufacturing techniques don't work. Also because of the reducing channel length, the semiconductor physics changes a lot.
- Channel length has already reached upto 22nm in the form of the 3D technology by Intel and Intel is working on 14nm technology and may be by the end of 2015, 10nm technology will hit the market. At this very small channel length, nano effects or the quantum effects into picture and the fabrication becomes a very complex process.
- Various other technology competitors like AMD and ARM are there to challenge Intel but the fields in which they excel are different. If Intel is good in computer processors then AMD and ARM are better for the mobile and tablet processors.
- Thus the future technology by Intel will only compete with itself trying to surpass the previous technology. But in doing so care must be taken about other factors too like the lithographic techniques required for the small surface of the silicon chips, power efficiency, cooling of CPUs etc.

Acknowledgement:

This report has been submitted as a requirement for the course **"Solid State Devices" (ELEC 6221)**, under the guidance of **Associate Professor Dr. M. Z. Kabir**, **Concordia University**, **Montreal, Canada**. The authors would like to thank the professor for his help and guidance which led to the successful completion of this report.

References:

[1] Rachel Courtland, *"The End of the Shrink",* SPECTRUM.IEEE.ORG, Nov 2013

[2] http://www.intel.com/content/www/us/en/history/history-intel-chips-timeline-poster.html

[3] P. Vimala et al, *"New Analytical Model for Nanoscale Tri-Gate SOI MOSFETs Including Quantum Effects"*, IEEE Journal of the Electronic Devices Society, Vol. 2, No. 1, January 2014

[4] Dick James, *"Intel Ivy Bridge Unveiled – The First Commercial Tri-Gate, High-k, Metal-Gate CPU",* IEEE 2012

[5] https://en.wikipedia.org/wiki/Multigate_device

[6] S.Ramey et al, *"Intrinsic Transistor Reliability Improvements From 22nm Tri-Gate Technology",* IEEE 2013

[7] http://download.intel.com/newsroom/kits/22nm/pdfs/22nm-details_presentation.pdf

[8] http://www.forbes.com/sites/rogerkay/2014/11/25/intel-and-amd-the-juggernaut-vs-the-squid/

[9] Per Hammarlund et al, *"Haswell: The Fourth-Generation Intel Core Processor",* IEEE Computer Society, IEEE 2014

[10] https://software.intel.com/sites/default/files/managed/e9/b5/Knights-Corner-is-your-path-to-Knights-Landing.pdf

[11] https://en.wikipedia.org/wiki/Broadwell_(microarchitecture)

[12] https://en.wikipedia.org/wiki/Skylake_(microarchitecture)

[13] https://en.wikipedia.org/wiki/Kaby_Lake