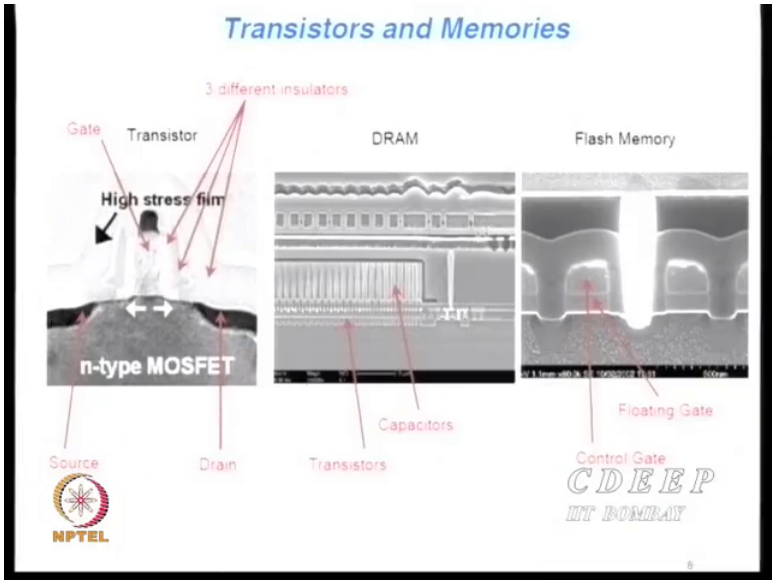
Fabrication of Silicon VLSI Circuits
Professor. AN Chandorkar
Department of Electrical Engineering
IIT Bombay
Lecture No. 2
Introduction "Micro to Nano"
A Journey into Integrated Circuit Technology

(Refer Slide Time: 0:23)



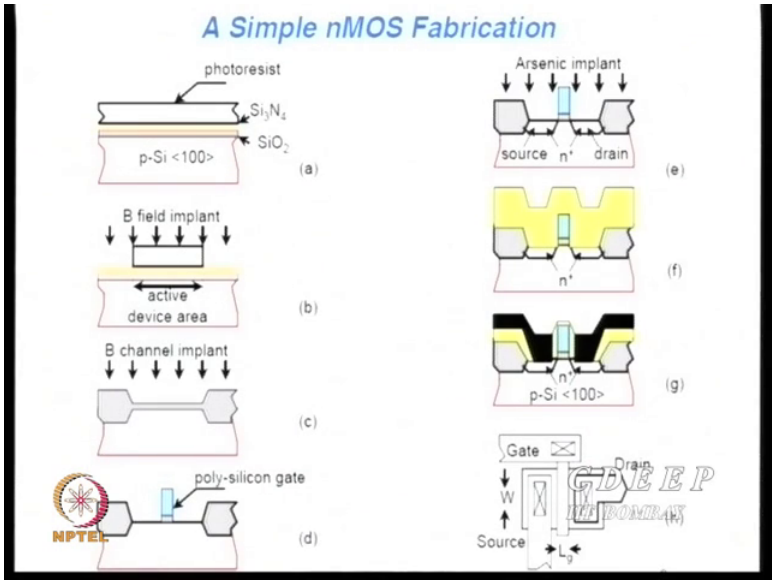
The semiconductor technology has 3 driving forces and these are 3 ones which you will hear, the first of course is the transistor technology, transistors are used in mostly in logic as well as analogue RF systems, from these major transistors you make there are other technologies which is driving semiconductors now, one is of course the memory technology DRAM technology as it is called and all of you know how much DRAMs are now available.

Variety of DRAM come NV DRAM, DD DRAM and many of them they are just RAM, so the dynamic memories are very very strong conductors for changes in normal technology of ICs and that is why there are lot of different research goes on DRAMs or memories in particular and the last but not the least is the current trend for last so many years now is to create a non-volatile memory which is called Flash.

Flash essentially stands for erase by altogether, it is electronically erased but all the bits are erased in one go and one is trying to see that the excess time of these memories are as good as

SRAMs if possible maybe SRAMs will have to go then. These 3 different kinds of technologies are current trends for which the efforts have been tried at many years.

(Refer Slide Time: 1:50)



Just to give an idea what is essentially a fabrication detail which is this course is going to address. So I may just tell you we start with some semiconductor in this case say we are making an (()) (2:01) channel device and because channel device we start with P silicon then we go silicon dioxide by oxidation and then we deposit silicon nitride and over which we have a photoresist.

The word photoresist will be clear to you when I do lithography latter. The idea is like if you have seen a normal photo film there is some kind of emulsion coated there, it is like a resistor essentially, so when you shine light through some object the light does not pass through the object and therefore the rest of the material actually receives light and there are 2 kinds of resistor use one is negative other is positive.

In one case if it is negative the resist which is essentially a polymer actually hardens, hardens means it does not get etch in normal etchants whereas the PPR or positive photoresist essentially are the materials they have also resins (()) (2:59) the resins and the resins whenever they like they actually break their links and therefore etchable for both kinds of patterns are used one with

PPR, one with NPR and we will see this is of course apparently it essays that we are actually using NPR but can be used either PPR NPR depends on the mask we create.

The word mask is very important, in the sense the object in the case of photo equivalently on a glass plate I create a pattern and if I have a darker area light does not pass, the clean area of glass light passes, so wherever the light does not pass this resin below is soft so it is etchable, so I can create a window inside the black region by in the oxide because then I can etch rest of the oxide or small oxide whichever way mask is used and therefore I can create windows in the silicon dioxide that is called masking.

Minimum mask shown here are 5, simplest N mask transistor can be made using 5 masks gallantly as I say CMOS IC is required 32 to 36 mask, so one can imagine the amount typically 1 to 5,000,000 dollar per mask is the cost, okay. So you can say if you add a mask and imagine how much money the industry has to invest. Okay, so the way it is done is that then we actually use the areas which I want something to happen I can protect that and if I do not want that area to be used then I will etch that out.

So in this case I have done using the resist, I have actually doing boron implants and if I do boron, boron cannot go through this resist area and the rest places the boron goes and goes down the way because of energy much below here, these are called P plus areas one can create these are called channel stoppers, so the first masking is done for what we call channel stoppers we will see this in actual technology.

Then he further do additional by etching this area we do further additional boron implants by putting an oxide under before, so there is an oxidation step here. As I say you may not be appreciating immediately but when we deposit polysilicon over it and again etch it. So Polly would have been deposited everywhere but only this part is retained by mass by etching this technique masking technique, so you have a poly gate.

Then I actually implants arsenic to make N plus drain and source resistor is good enough for normal implants, so we do not etch many times many things will show you later on then we actually deposit oxide again take one, okay. Then create windows for connection, metal connection to source and drain into the gate, so we open window here what contact metals etch

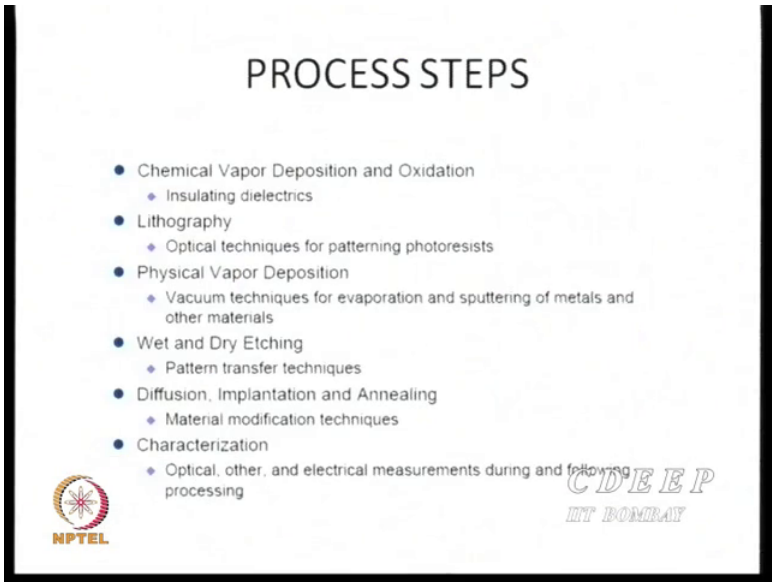
the pattern wherever you want and you get, so this layer portion which I show here are actually composite mass this is the pattern actually you will see on silicon.

But there are number of mass, you can see the first mass which I was doing this area, maybe I do not know whether the big bigger portion of this is the window which I actually opened which you can see this was the area which was blocked, so that is the area which is total area is what I first created, inside then I did some kind of channel implants with this area and then I opened source drain windows, so these are source drain.

I opened a contact gate, so gate was done by this mask, okay and this mask. So each mask actually should get aligned with the earlier ones, so wherever you want that this to come you must align. So you need a process called mask whenever I do is lithography in which I actually see through a microscope allowing the patterns, okay and then only I expose the light there, okay.

So do this, this is essentially very important lithography is a major crux of all processes. So I even done this then you open a contact window for source, drain and gate. So you can see here and then put finally the last mask is the metal mask which connects. Now this is a process step which is as I said 5 mask process and later as I say you 16 mask standard CMOS process but as of now there may be some processes or some chips which may require as high as 32 masks, okay.

(Refer Slide Time: 7:46)

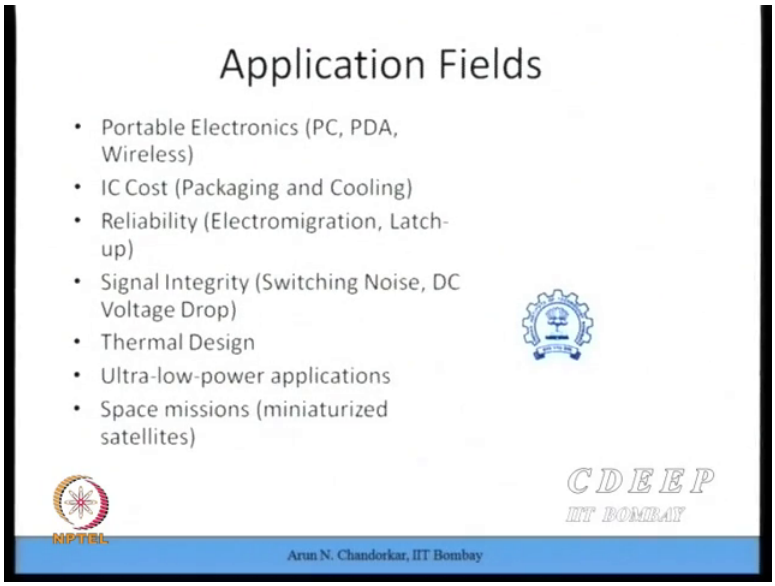


So next is, so I yesterday already said we have all kinds of process steps which are required to deposit material, oxidised, diffuse, implant, incorporation selectively is the basic feature of all IC technology. I want to make N plus certain areas, so all other areas should be blocked and only that should be open, so this is essentially the feature. The smallest feature which I can do any processing is called the node what is number we used to be gave 32 nanometres, it is assumption one of the smallest whatever dimensional I can create a study do but that is not really true these numbers were earlier when we were working on 0.2 5, 0.1 3 microns nowadays the gate oxide is 22 nanometres process is 16 nanometres.

For 16 nanometres it is only 9, okay 11 it may be 5, so we will see that these numbers are not really matching, earlier this was the smallest damage if I say 5 micron process which means the smallest dimension I can print on silicon is 5 by 5 microns, okay. This is our designers sees, designers sees what is the minimum feature it has. So now this note does not say anything about features though we still keep caring because of Moore's law business 0.7 0.7, so we kept on naming the notes like this but in reality these numbers have no direct relationship the feature sizes.

So these are the process steps which is this course is going to look into lithography is, CVDs, PVDs, dry etching, diffusion, implant, Anneal and characterization.

(Refer Slide Time: 9:28)



People always ask where is the use of ICs, so much money is pumped in? Well all of you are working with portable electronics PCs are one, wireless base systems like mobile and others or what is called hand (()) (9:39) PDAs. The basic application is the cost of IC which you create. A typical microprocessor may cost around 160 dollars, okay Pentium 4 for example with the latest version but older version 80 to 85 may be available to160 rupees.

So the question is as technology as technology advances the cast of chip also increases and this is slightly difficult to understand we will say, oh I have made bigger things so I will get smaller money, no it is not because an investment on a process line as I said last maybe 8 billion dollars to 10,000,000,000 dollars per node. So if someone will say shift from 32 nanometre to 22 he is investing 8 billion dollars, okay for the next line.

But if you want to keep the old lines is unnecessarily spending money, so he will actually start charging them also because otherwise that line cannot be sustained, okay. Please remember typical run in any semiconductor industry has at least 250 wafers of 12 in size. In one run 250 wafers actually go into the furnaces and so how much heat cycle it must be having to maintained pressure from 800 to 1200 degrees, so huge thermal budget is required.

So much etching goes on, so much processing goes on roughly industry requires 54,000 gallons of water per day just to maintain your fab line, okay. So why I am telling you cost? Because all

these processes are good or bad when I am teaching a design course I keep saying behind is as good as what customer wants in minimum money, okay that is what the best. Best is nothing called best because you can always improve something but at the cost if the money goes high no one buys it. So the gain is what is best for the customer smallest amount of money is the best design and best technology, okay.

So never say, oh! You have to be there, you do not have to be if the money is this much customers (()) (11:41) only this much use as bad as technology available but it satisfies, okay. Isaac third problem in all applications are reliability and this is major worry because if you are using a chip and every third chip blows then no one will buy your chips. So there is a reliability issue it should stand to at least minimum lifetime maybe let say at least few years 5 years or 4 years.

So the test has to be earlier performed for reliability, reliability is defined as the spects remaining same for given time. So if the spec changes no one buys the chip, okay. So there are issues like reliability which was actually hell of a time when there is also problem in design has to be taken care of a technologist is signal integrity. You may enter some data which may be 3 gigahertz signal but internally it is not reaching 3 gigahertz because of the RC time constant inside.

So your whole trip that you are pushing a signal for gigahertz is essentially not inside anywhere for gigahertz there are other problems like there are inductors and they may be lot of transient may occur. So one has to take whatever signal you are processing should remain same throughout that chip area, okay. And that is another issue which is a big job in design which is somewhere related to technology where he say I want at this, this also. So some of the things are actually designers forcing technologist to take care because otherwise you will not be able to get a design which works, okay.

There is always a fight between technology and design. Designers feel that technology is not competent and technologist feel designers do not know what to do, so they just give something and expect that we will do it for them. Earlier times technology was lagging designers but in 2000 maybe ahead now, technology is so advanced, designers does not know now they can give you 1 billion transistors but do not have a system which requires 1 billion transistors they make it 3 billion so what you will do?

Okay, so there was a time in my phase of carrier when technology was lagging people wanted 10,000 transistors you could give only 2000 then they say okay we want this performance you say no I cannot do better than this but then times have changed now designers are at loss to actually create systems which are much more useful cheap as well as and therefore a technology evaluation is possible.

There is another issue in our designs which has to be taken care by technologies is called thermal design. You know if you have 1.5 centimetres by 1.5 centimetres, 2.5 or 2.25 square centimetres and there are let's say 800 million transistors working half of them at least will be turned on hopefully by probability each will be drawing current, okay. So power supply voltage into that current is the power dissipation.

So look at the so many transistor turning on with so much current and voltage products, so the net power which it may consume maybe watts10s of watts to 40 Watts which essentially in a thin silicon layer actually heats the silicon. Now as soon as silicon starts heating device property start changing because all most of the characteristics of semiconductor device is very strong function of temperatures.

So if your temperatures rise too much your circuit will fail irrespective whether it was well good design or not. Even if I had taken care the (()) (15:17) is the same the maximum temperature I expect is 50 but it may reach 65 then what will happen. So the issue starts that your designer now tell thermal design has to be taken care mostly by the technology people that the heat is removed faster, okay.

So there are tricks which allows that it is not that it is impossible and then there are 2 kinds of applications which is driving us, one is of course is the ultra-low-power key applications. Most of the battery handed systems actually require ultra-low-power device. We will show you some of this, how much low is low? And of course only good thing about space is the only want smaller things the rest of the things they allow.

You say one chip costs 10 million dollars NASA at least will not say no, okay because that is what in their some mission they need, so they will say, okay fair enough the performance I want this. So essentially now I am trying to say there are at least digital there are 3 kinds of circuits

which has 3 different technologies, one is high-performance circuits, now this high performance means high-speed. So high-speed circuits at the cost of something it cannot achieve anything otherwise, so it is okay I will allow some more power dissipation I may allow some more area to you but I want performance the circuit must function at 8 gigahertz and that's it. So this is called higher performance circuit, so only feature for them is speed how fast you can do?

The second kind of circuit is low-power circuit, I damn care the speed is very low or it may be one gigahertz or 800 megahertz like GPS this our mobile requirements but it should not consume more than so many watts per square centimetre. Now this is called low-power design in which the easiest way to reduce power is reducing the power supply. So more circuits we are trying is from same events from 5 volts, 3 volts, 2.1 volts, 1.8 volts, 1.5volts, 1.2volts, 0.8 is now the current power supply.

We may prefer to go 0.4 but something else may happen, so we are right now not reaching but maybe someday we will reach. So the power, second circuits are called low-power, so that technology is not same as technology for performance. There are third kinds which is what most mobiles require those standby, when it turns on it should give higher performance, okay.

But when it is not working it should give very low power dissipation, these are called low standby I mean low-power standby these are the 3 different digital circuit technologies which are used in commercial sense. So whenever Apple comes with something or Microsoft or Google, Google of course has sold Motorola now, Lenovo one may say, so they actually figured it out the available in the market what is it doing ?

Because you now want some million applications I do not know how can you use all of them but they have ones on their mobile so many application, each application costs lot of power. So the power dissipation is major very sense nowadays except as I say space they then they have satellite system they may put fence everything cooling, liquid nitrogen, liquid helium they made anything for you but they say performance has to…

Because when something lands on Moon or Mars it should function irrespective and that is where the money is paid for, okay so it may not be silicon, it can be mixture, it can be anything I only want performance, so these are third kind of technologies which are space related which are

generally not known to many because these are normally secret technologies. When 10-20 years when things are already too much advance than they release 20 years ahead technologies. So you realize oh! This was what they were doing 20 years ago, okay but those technologies are very advance but at too much other cost.

So please remember why so many technology studies? Where each application is asking something else and then you want to come up with this it will meet your specs and it also should be low-cost. Now if the cost is high no one is buying it either unless you have monopoly and at least in many countries including yours monopoly is not allowed, you know why AMD is surviving because Intel supports AMD which is the competitive, okay which is very funny, is not it?

Because if AMD goes away Intel will have a problem, so Intel will actually see some market per AMD, so that they can have unrestricted manufacturer, okay. So this is why there is an issue in every sense. So technology development is very costly, very intensive it takes 3 years to go to the next node and investment to 8,000,000,000 dollars roughly these days so much investment if you do in this and if the product does not sell who will put this money, okay.

Of course they do sometimes they fail and then they remove 4500 jobs very or something that is what the quarter 3 this year has did not shown good profits to 2500 are laid off, okay thank you that is the way it is pink sleeve as they call. Why it happens? Because they expected so much sell that did not happen that is the end of it, okay.

So please remember I keep telling why economics matter because only economics is driving force for us. If the money is not generated no one is trying to put money in, okay. This is life and 1 has to accept. I may like to do something great but maybe in my lab, keep it, if I has to growth of industry it should be manageable, sellable only than it will grow the industry. So many of us are so enamoured by electron motion including me that we forget that, okay I have found something interesting but who is going to buy this one, no one because this will take so much cost and buyers will be 2, so why should someone will invest money in you.

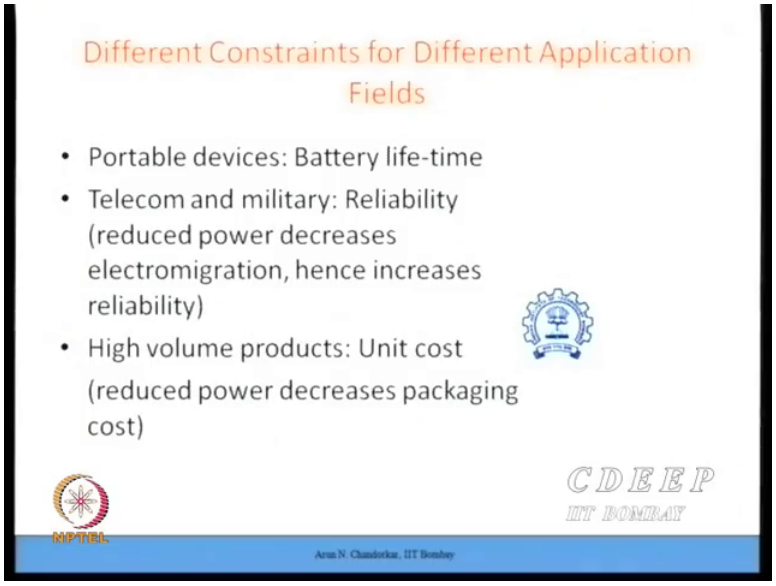
But then I will get papers I may get a paper in nature, I may get some big laurels fair enough but industry does not go by that, industry see when the maturity appears in a process only than it

picks up. So they themselves may be doing research, they also know what they have to do but basically research is normally not done by industry simply because they fear that if they invest there their profits go down. So this is how University survives, they (()) (21:52) in universities you do keep doing research 1 out of thousand I may pick it up, you do it.

But we survive on that, okay, if just one chip works out of 10,000 chips, so we never say 9099 did not work, so old process is cracked we got one chip working that is a process in lab but that is not the process in industry. In the industry if it is 10,000 then at least 9900, okay. So the game has to be understood that why technologies are always killed by everyone is because whatever they create if it is not yield, enough yield it gives that process goes down, okay.

So we may come up with graphical as I show you many other new technologies but it is not going into industry simply because it is not feasible or it is not very great to see that it is financially viable. So read papers and you will figure out why so many technologies have come and gone also.

(Refer Slide Time: 22:53)



There are different constraints one sees for example if you have portable device battery is the major cause, how it should not of course, so one of the things which I suggest you leave microelectronics and look for battery research, okay. If you can create a battery which have large

ampere hours in a smaller area I think you have a mine of gold with you. If that is what you can do let say you have 10 ampere hours or 20 ampere hours.

In a small tablet I suppose many will rush to you only chase you out everywhere, if you create such a technology. So all these IC is not worth that if you can create one battery which can create large ampere hours for you, okay. Some of you all say hydrogen cell may try every other thing like Edison tried 400 experiments or 4000 experiments (()) (23:43), okay but that is something boring, so people say hey there is nothing great in this and therefore no such research happens in IITs where it should actually they are technology people, okay.

The Telecom and military of course as I said they actually look for reliability that is where the money is. Of course they are ready to pay money for it and there are high-volume products which are optional products which are actually cost wise like a DRAM no one wants to sell it at say 5000 rupees, no one will buy even if it is 16 GB or 256 GB, so when I asked to reduce the price and if you reduce the price, price is essential in package.

As I said a set chips are very cheap, okay. It is the package which costs helm. So if I tell you give a product, of course I do not give projects at if I give you some project on that is designed you will say but where is electrical there but that is a problem, okay. Because electrical people think we should only look for at least microelectronics, if there is no electron what is there? You know, there is nothing there, okay but that is the money, money is in packaging actually, think of it, someone amongst you can become an entrepreneur.

Okay, the problem in technology started after 95 ahead then there is a switch, transistor works like a switch they say it for logic, so if it is turned on essentially the switch is on connected, okay. There are few currents which I can say one expect that there is no current from the controlled side into the switch, okay whichever reason like in case of MOS there is no gate current, okay we expect this but it does not happen there is a gate current, okay.

We expect I to be 0 across the gate but it does insulator are not have a good insulator as we thought, okay therefore current flows through them because of the high fields. We expect maximum current as much as current I pushed from the input to go to output believe that no at

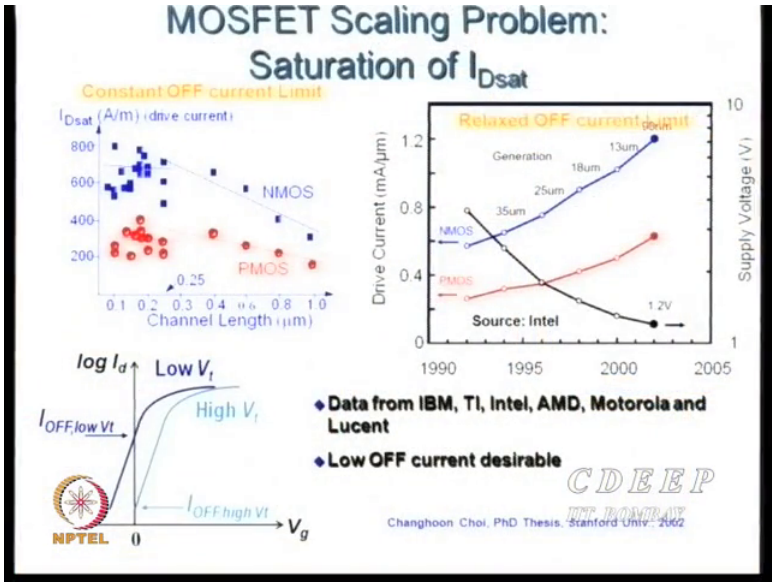
attenuation will be there but that is not true in fact there is a arm resistance which will actually not allow all that current to go, okay.

So another ideal switch we thought that infinite current can go and no control from the control side is only ideal switch. In reality when we started scaling devices (()) (26:11) came we figured out the gate current is not very small even this lateral current which we believe below threshold when the transistor is of should be 0 is not 0 it is in fact very large current these days and because of that ideal switch is only ideal in reality it has too much leakage.

Leakage means what? The power loss, this is unused power loss and that is one reason why most of the low-power research is controlling is leakage paths, okay reduce this leakage current because in standby mode in specific you can see that the currents are flowing even if it is leakage current and if they are larger they will drain the battery. So you keep your mobile, next day morning there is nothing inside, okay because of leakage path unless you power of completely there is no way power shut-off comes, okay.

Now there is an issue, there is a technology as I say low-power technology came out of this because they want this to be as small as possible, so the whole transistors have to be differently thought they give low-powers. Every technology comes from requirement of the user and he says this is what I want.

(Refer Slide Time: 27:36)



As I say as I start scaling, in the case of on currents if a MOS transistor the RV characteristics is thing like this, okay. Now one can see from here, one believes the of current below threshold is very small but if you scale even this current of 0 Vt or small vt the current is very high. So of current, of current means the transistor is switched off officially, input Vg is smaller than Vt but there is a current, okay.

This is whereas worrying us the most right now, how to actually bring them back to then if you reduce Vt this will happen that is a problem we are scaling, we are losing the VDD we are reducing Vt if you reduce Vt the power, low-power actually off power increases. So now here is a case whether you want really low-power or you want off power to be less, so design and design means new technology he says you do this I so why every time we come back and say want this.

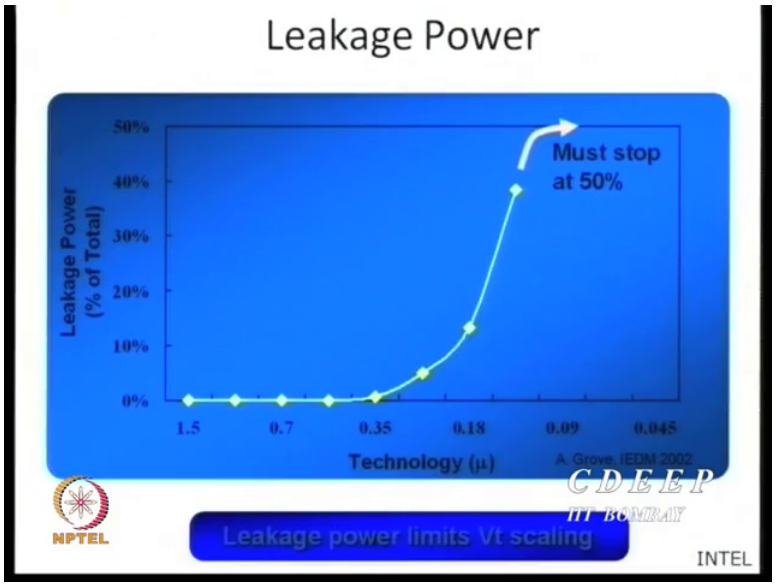
So why every time they come back and say that a new technology has appear because someone else is asking I want this then one has to plan, device people have to think, get models and we use those models to create new technologies which will fit into those and that is why the research and technology is advancing day by day though much of the industry is not changing very much because as I say they want standard products. So when you technology goes 10 years or 5 years ahead afterwards they go into the actual fab, okay.

So if you look for data from Intel, IBM, AMD, Motorola and Lucent many companies everyone has tried and one is trying to, if you reduce the dimensions the current does not keep increasing as scaling (()) (29:22) actually its saturates you can say it saturates which essentially means why scale below 0.1 micron or less than hundred nanometres if on current is that way speed cannot be improved because on current is fixed, okay.

Then if I want to increase speed somehow I must reduce the capacitance it CDV by DD all that I have in hand or I fool the circuit saying that oh! You still can run faster though my current is only this there are methods call paralleling pipelining, these designers will tell you if I am teaching designing for there I will tell in case I am not teaching. So the circuit method now says that technology says at these technologies current cannot be recruit for the on state, okay do whatever it is.

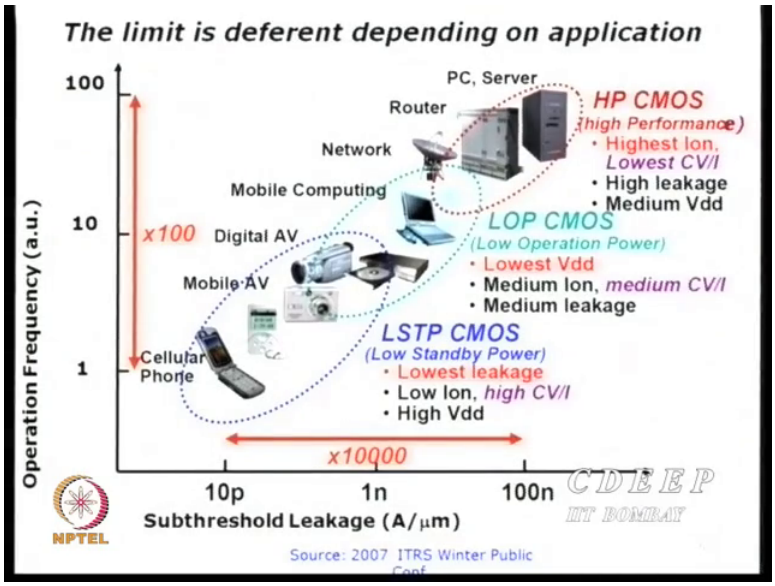
So whenever you do something you have to understand I am giving something at the cost of receiving something. So now I say I scale down everything I have put more devices more 8 million transistors, 800 million transistors I put but your speed is not increasing, okay. Technology you have reduced from 14 to 11 nanometres but speed is not growing, so externally something has to be done to improve your speed, okay.

(Refer Slide Time: 30:54)



By speed I already said videogames nothing else. This is some kind of statistic from Intel many years ago this is leakage current and one can see 50 percent of the, is already reaching as a leakage path, okay. So we are worried that latter as I say in 30 nanometre down technologies off current were 66 percent. So why do a circuit which is losing power without doing anything, okay but that is the game.

(Refer Slide Time: 31:13)

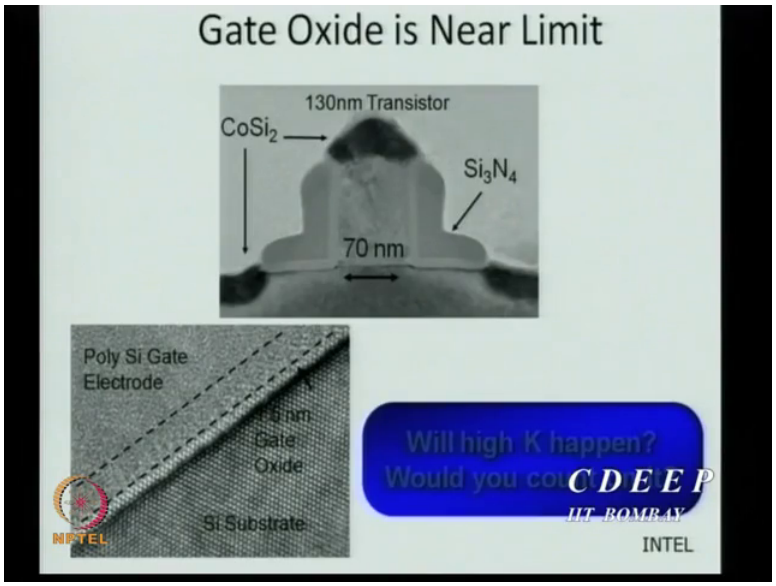


There are also these limits as I say just to show you in different devices systems which are available in market. There is an operational frequency varies from one gigahertz to say hundred gigahertz. Different systems are there mobile, digital AV are computing, network, Routers, servers each has different technology because each has done different requirement. So when I say technologies I am actually for the product, is that point clear to you?

Why are so many technologies? Because product needs something, so technology is catering that kind of there is no universal technology, okay. Each kind of applications because for example like cellular works around 899 megahertz or 1 gigahertz at best, so the kind of thing we can say subthreshold current we want there should be very low because low-power I want, okay.

But frequency I do not want very high, so I may probably able to get something out of that but if someone say I want 100 gigahertz I am low-power am not achieving because then I going to hire leakages, okay. So every product has a different line and different companies cater to different products, not every company enters every market because that is the way they cannot survive, okay or a lot of cost because of that.

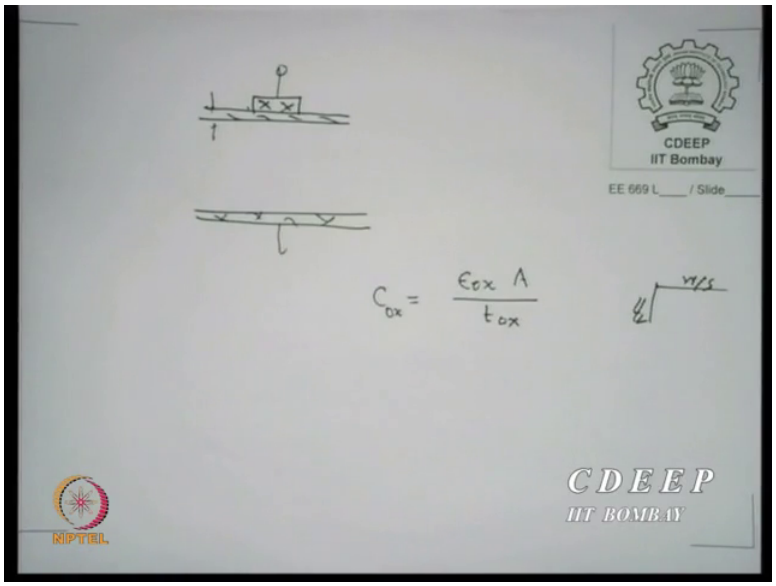
(Refer Slide Time: 32:33)



This is something which in 2000 it right and I will show you the latest one, we were trying to state that if I start scaling the dimensions of the oxide thickness of your mass transistor also scales down. So for example scaling long to say 22 nanometres or even lower the oxide thickness required is less than 5 angstrom, half nanometre. Now if your oxide thickness is half nanometre all over its called thickness of the oxide or insulator.

One atom of silicon dioxide will require 5 silicon layer as 5 angstrom (()) (33:14), okay lattice. So I cannot have less than one monolayer, you want to have reduction lower than that.

(Refer Slide Time: 33:33)



So I made a take back what I did is, I said fine I cannot scale down as then say 10 nanometres out of this one nanometre, so I did something like this, this is my capacitor, this is my insulator, this is my gate, this is a standard capacitor, this is a mass capacitor, the capacitance of this oxide whatever it is Epsilon ox ks Epsilon 0 multiplied by area divided by oxide thickness.

So if you scale you reduce the dimensions W and L area also reduces by square law, oxide thickness goes linearly by the ratio we want to scale down let say S is the ratio, so it t ox by S, A will A by S square because area you reduce that much by each dimensions, you can look at it this dimensional and this dimension if I scale the area reduces by a square, okay. So I reduce area but I reduce oxide thickness by only S, okay.

Now if I start say oxide has to be less than 5 angstrom I can align, no. So I said okay, area I cannot reduce too much because that is decided by the other currents of requirement. So I figured out if I somehow increase my oxide thickness proportional to reduction in oxide I mean directory constant of oxide they use proportion how much I have loose in oxide and let say I have decided 10 angstrom is I can grow minimum or maybe 50 I can grow then as many times I have dielectric material which is that much dielectric constant.

So let say I cannot go less than some 4 times this, so I at least dielectric material which has 4 times dielectric constant of silicon dioxide these are called (()) (35:23) silicon dioxide as a

dielectric constant of 3.9 we are looking for materials which has at least 12, 16, 25 maybe 100, okay. Because if I increase ox this insulator thickness then I can have more reliability, I can create so many atoms, I can put there without worries, okay. The problem started if I have very thin oxide which is what, let say it is less than 10 angstroms this is essentially punched through that means little electric field will allow electrons to go through the thin oxide (()) (35:58).

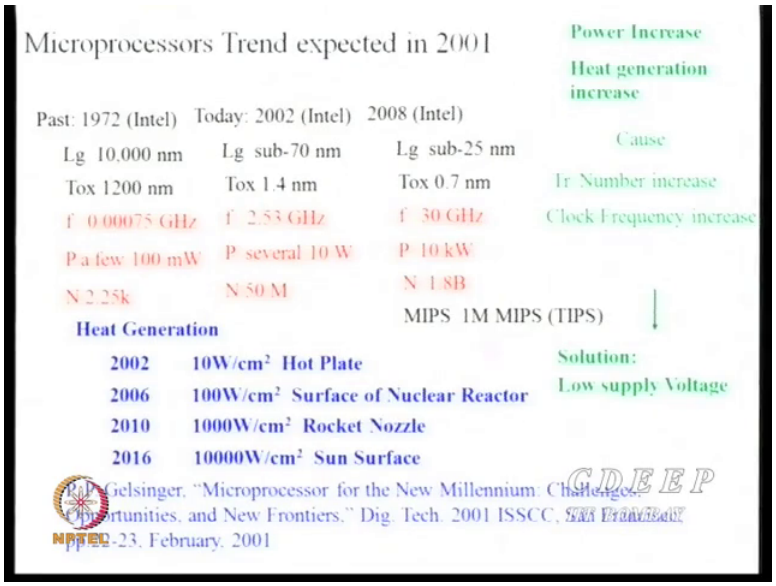
So you cannot have too less thickness of material there because it will end short-circuit the gate which means I cannot really work on very thin oxides anywhere, okay because there are other problems pinholes and many other issues in oxide of thin nature but this is one problem of this, I cannot scale down thickness too much and if I cannot the only possibility was to go for a higher dielectric constant materials. So new technologies are (()) (36:28) dielectric technologies and not silicon dioxide, okay.

This is the only difference which new technology but to create a hike is not very easy because the materials which should stand silicon dioxide is an ideal material in every sense, okay. It is the latest material probably I see people think. It has a very high dielectric constant I mean very high dielectric strength it can go up to 10 to power 7 volts per centimetres fields. It is unetchable in most cases except chlorine atoms it cannot be etched away easily, is very hard, it remains into, its melting point is 1900 or everything what you think the best for an insulator it has a band gap of 9 electron volts, so no electron can easily go through it. All possible goodness was in silicon dioxide.

Much easier to grow a silicon oxidise you have oxide, so everything was in favour of silicon dioxide, we still are using do not think it is not being used but the time has come and we say oh! If you cannot do less than this what do we do? So the new attempt is now to look for what they called hike materials, okay and what are the choices? We will see that, is that point clear?

So why (()) (37:46) have come because as scaling goes oxide thickness reached limits now, okay. Oxide means silicon dioxide, so we are now looking for new materials.

(Refer Slide Time: 37:58)



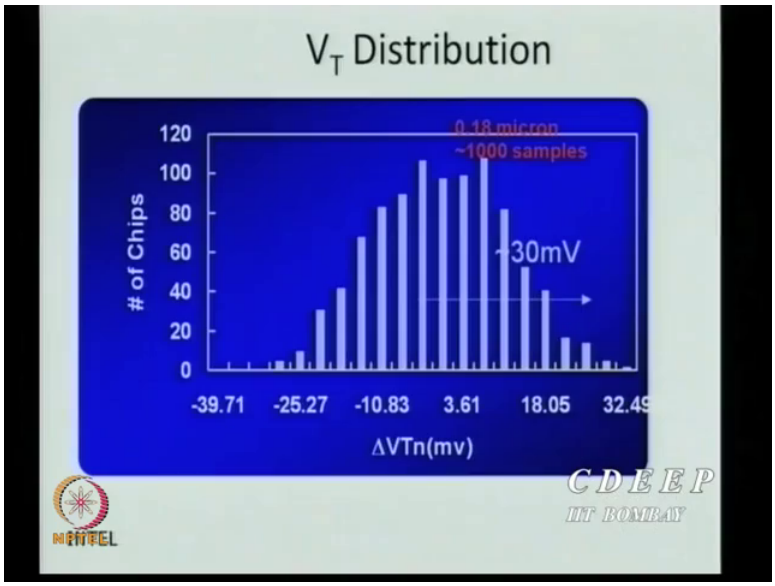
It was predicted as early as in 2001 what will be the microprocessor trend? And this is very interesting to show you. In 1972, 2002, 2008 what Intel thought and by prediction and by doing design they will go for 30 gigahertz frequencies with 24 nanometre gate lengths, okay. Oxide of 70 angstroms what essentially they were looking for? Tera instructions per second that is the ultimate aim everyone is looking for.

They figured out if we do so the heat was a major power, okay. What has happened in that? The power dissipation using the Euler technology 2002 one 10 watt per centimetre square was the temperature I mean wattage it was consuming power density which is like a hot plate, okay. And 2006 when we went to the newer technology it became 100watts per centimetre square which is like a temperature at the surface of a nuclear reactor.

In 2010 this increase to 1000 watts per centimetre square which is like temperature at the rocket nozzle, okay. And in 2016 it may become 10,000 watts per centimetre like a sun surface. So the major worry is actually coming from somewhere else, okay. You may say I improved technology I improve speed but what cost, okay. So one has to understand why technologies get limited even if they think to be very ideal when I do simulations I get everything excellent but in reality nothing works.

The reason is there is a gap between what happens and how much you understood that is why research has to be done what is happening and what I did not understand, is what is major effort is going on, okay. So just to give you this heat is major worry, so that is why I say what for packaging which can remove faster your money, okay.

(Refer Slide Time: 40:09)



The another problem came in technology now is, as I scaled down all the parameters did not remain constant over a wafer or even on a die or even on a transistor, okay. So if you have a 12 inch wafer, each die maybe 2 centimetre by 2 centimetre, let say there may be 10,000, 1 million transistor or million transistor there the parameters process parameters like thickness of everything which I thought should go everywhere is not everywhere same is called variability.

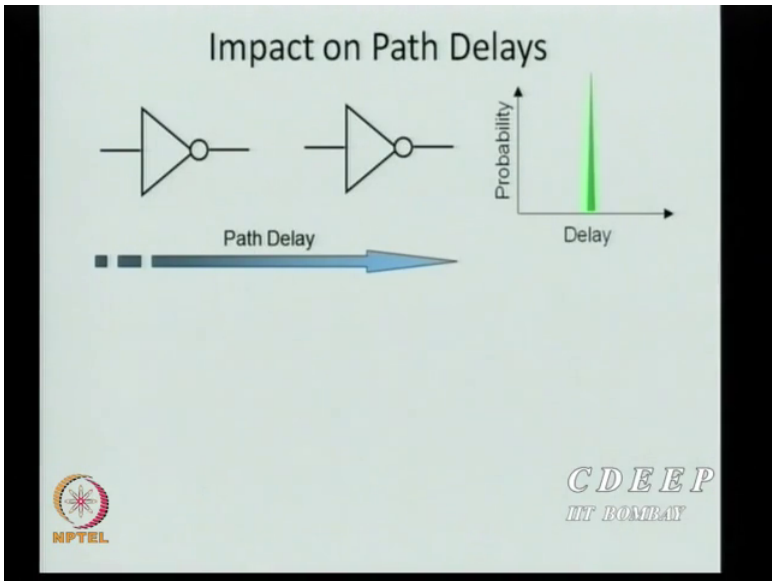
Since there is variability on every parameter I do, the net result if I see and this is an old technology graph from Intel, I have my own instrument I worked recently, so for that 0.1 micron, 0.18 process they have measured thousand samples they figured out there is a threshold shift of 30 millivolts.

In 0.8, 1.2 millivolt does not seem to be very big but in 0.8 and this 10 percent which will go there maybe 60 millivolts to 100 millivolts 10 percent variation is as expected now and if that happens my threshold will be already everywhere, so ships will be working at different speed, different power dissipation sometimes logic may even fail, analogue gains will not be obtained.

Some chips they are working, some chips they are not working my design is ideal for everyone but it is not working. So now the question is how a designer should take care in the process variation which will appear irrespective what I do, okay. So now there is a, earlier designer used to be computer scientists in most cases, electrical entered little late and they always use to look down this E student E people over technologies.

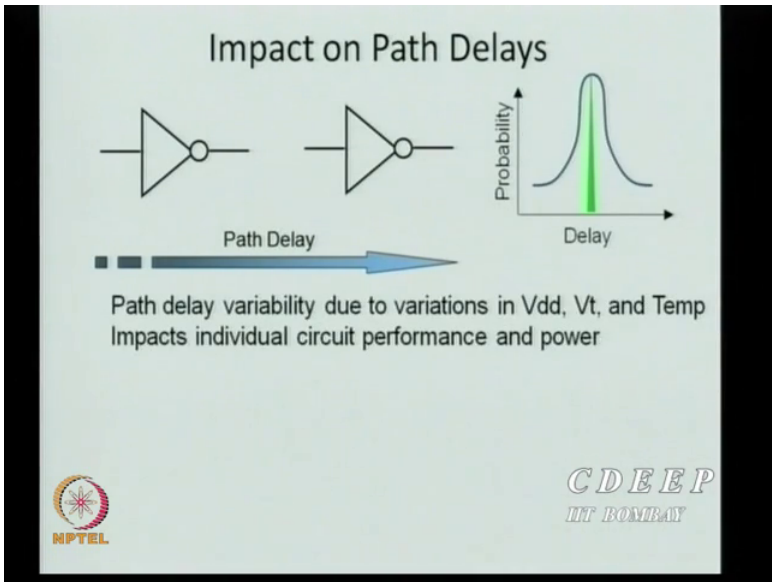
Now they have to come to our place and say please tell us sir what can we do that is the time, okay. Because unless they understand what is happening they cannot design, okay. So now it is a codesign goes with both computer scientist and electrical, chemistry, civil, everyone sit together and actually see what a chip can do best possible for given technology. So this is something new, so that is why we have always been teaching microelectronic independent of everything whether you like or you do not like. We teach devices, we teach technology, we teach systems, we teach design, why? Because we believe from 2005 onwards everyone must know everything, okay.

(Refer Slide Time: 42:43)



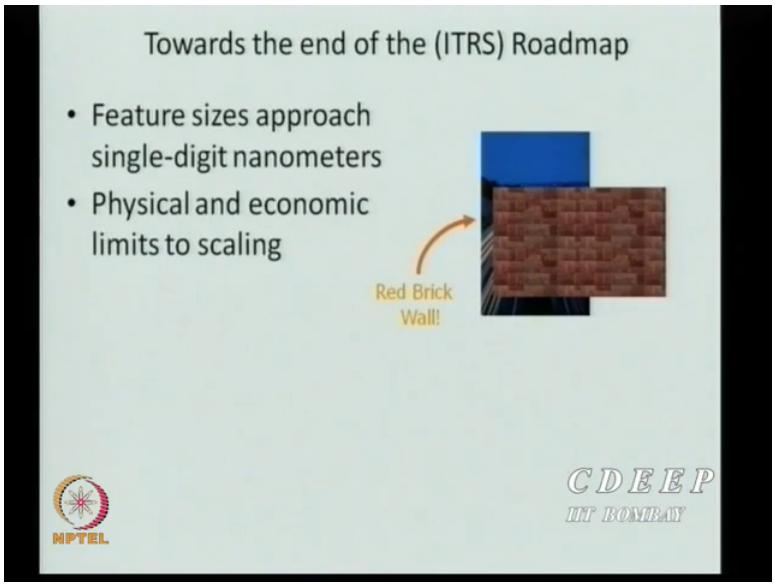
See earlier older technologies if you have an inverter the path delay was known the reason was if you plot the probability of getting one delay, let us say propagation delay TPD it was almost constant, okay at a given value, okay.

(Refer Slide Time: 43:02)



But now is caution, so what Vt I should use? I do not know, what delay I will get? I do not know. So if circuit is performing differently at different point of chip and at this, so what we do now in design is we look for what we call the (()) (43:17) called critical path and actually (()) (43:20) so I know somewhere I am over designing because that is the unfortunate but at least the least one will work better. So lot of money is invested to get critical path designs, okay. Lot of area goes, lot of parallelism has to be done but that is the only way (()) (43:37) can be handled now.

(Refer Slide Time: 43:45)



Okay this is what I said as I scaled down many things I cannot do either by economic or by physics this is called a red brick wall cannot go and as I said humans are very smart and new technologies will appear and had already appeared these are called Nanotech. CAEN is called self assembled chemically electronic nanotechnologies and also in lithography we are working on technology called extreme ultraviolet EUVs.

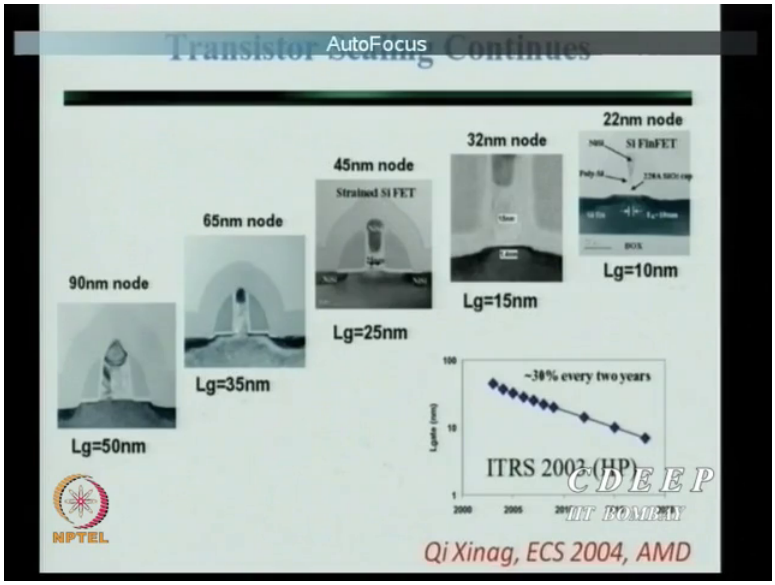
Last 8 years there is an effort all across the world to design a system of UV which we are not able to. One thinks of it what is this EUV? Why it has come? The smallest feature lithography can produce is given by its wavelength, okay. Optic says it is less than wavelength you cannot separate because they are everywhere they fraction around. So you want to separate 2 lines, areas, so we must have wavelength which is of the order but we are made very nice gains, we have the lowest of the low length we are using is 193 nanometre UV, okay which is still 2014 also we are working.

We are trying to see it goes to 10 nanometres then I can make 3 nanometre device, so I am trying to reduce this, so we are trying to push the energy E is equal to CY lambda, so I am trying to see whether energy can be pushed, okay. And if energy is pushed lambda will actually good, so we are looking for optical lithography, okay even now ultraviolet which has a wavelength of the order of 10s of nanometres not 193.

But as humans are ingenious I am ingenious they actually have figured out even with 193 people have made 11 nanometre transistors, of course there are issues, there are costs involved, there are (()) (45:36) issues but this costs multiple water-based new lithography techniques which we have worked. However it costs, so they are working it to 10 years on EUV and I know many universities and few industries I know many of them who are working and they every time they meet they say next year but this next year syndrome is very bad and becomes because when the next year appears it becomes present year, so the next year, okay.

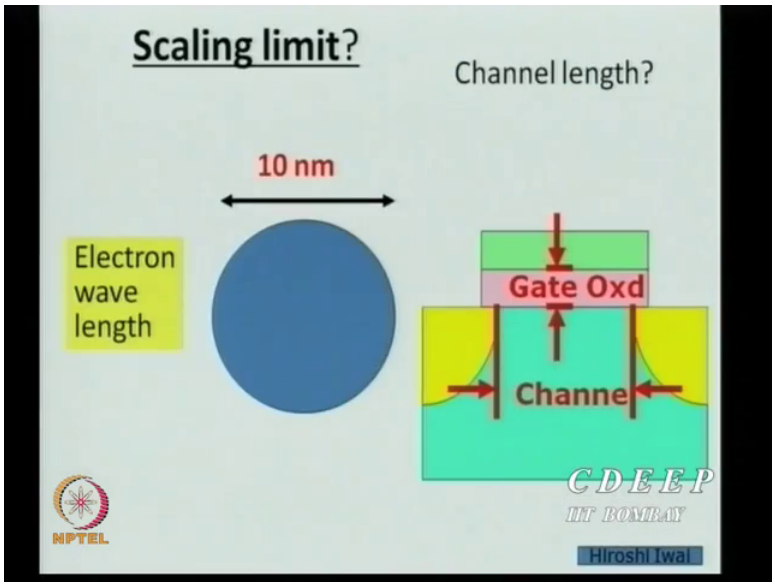
So nano reached next year because next year becomes present, again next year, so last 8 years UV has not come but if it comes these 5 nanometres is achievable very easily.

(Refer Slide Time: 46:30)



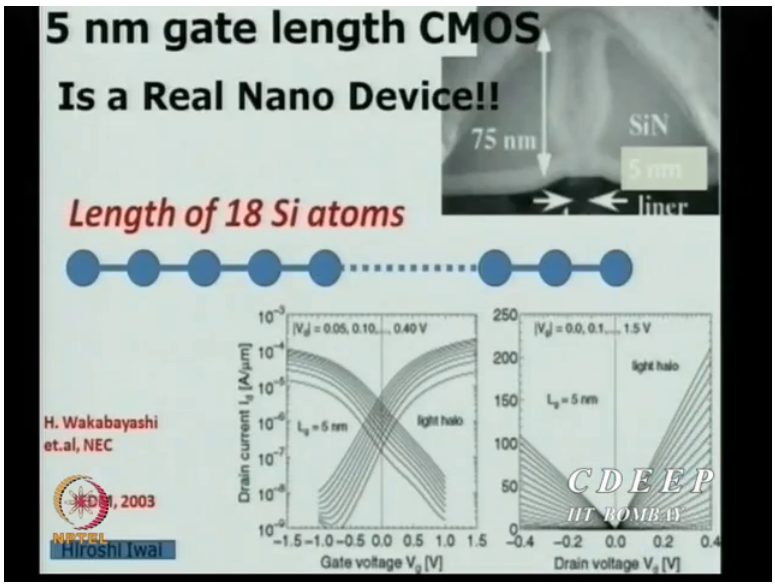
This is a progression we started with 90 nanometre with gate length of 50 and by 2004 we already reached 10 nanometre gate lengths, 22 nanometre node today we are working on 16 nanometre mostly and 11 is of course possible it has already shown its working, 7 also has worked, so I do not know then where it will go 0 probably but this is what is the way progress is done. So for each of them there is a different technology constraints and they have to be solve and that is how the new technology, new node means new technology.

(Refer Slide Time: 46:56)



After all electron wavelengths will decide how best features you can create that is what I just now said. 10 nanometres was initially possible we are looking for 1 nanometre now, let us see whether we can.

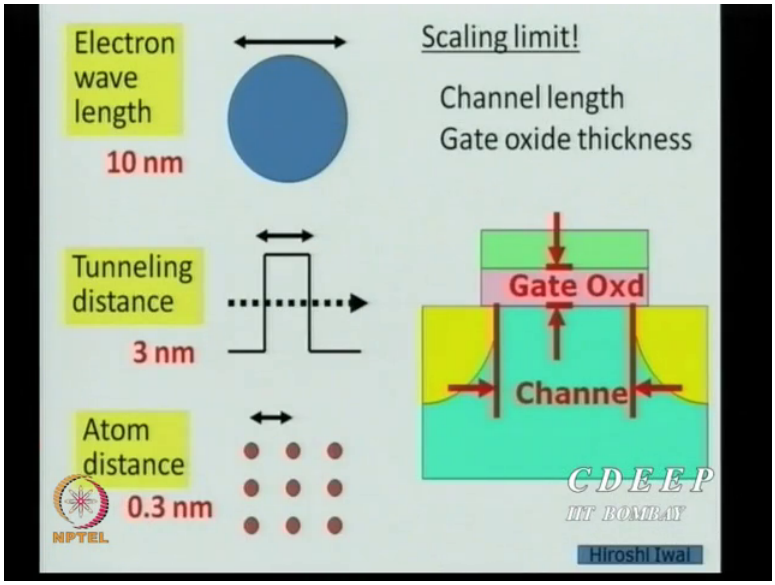
(Refer Slide Time: 47:07)



Just look at this 5 nanometre gate length CMOS is a real nano device, okay. It has a length of 18 silicon atoms, okay. You can see around and in 2003 itself this device was made, so it is not that one is not able to make lower technology node but if I make millions of them only to3 may work

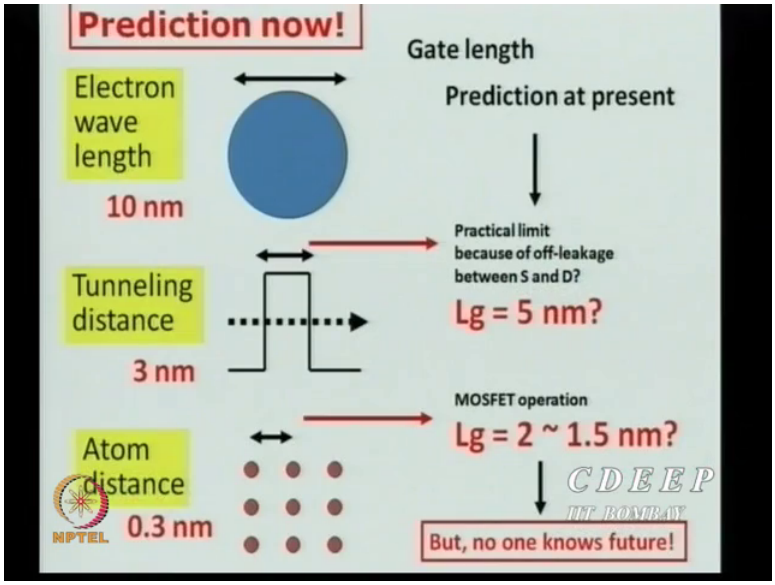
maybe good in IIT but no industry buys it actually, so not the technology is not known but reliability is the major crux which is not still achievable, okay. Ideas have worked everything look to be a good switching technologies.

(Refer Slide Time: 47:51)



There is another problem which is called, okay maybe this.

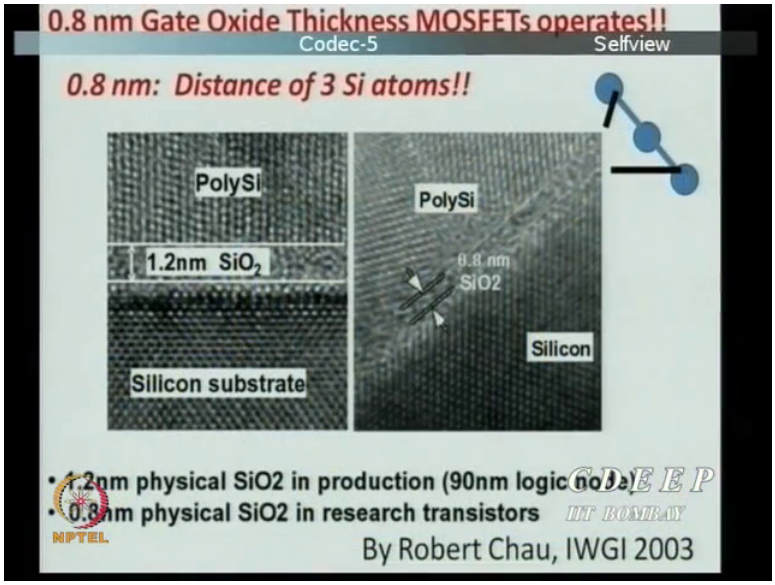
(Refer Slide Time: 47:55)



As you scale down the thickness of channel I am in this, sorry gate thickness one can see that we always thought electron does not have energy enough to cross but there is a possible lady of tunnel, okay. Because of variety of effects, 3 nanometres you do it will just cross, okay. Then there is atomic distance 3 angstroms, what else? 2.6 1 is the minimum (()) (48:23) that's it, so you cannot put less than an atom otherwise what you will do? However all said and done but no one knows the future.

Today, well I am saying but in 1988- 1990s when I was teaching I said something which is proved absolutely wrong by now, so today if I predict something in another 10 years 20 years will be proved absolutely wrong, so I do not want to predict the future but there is something I see fundamental limits.

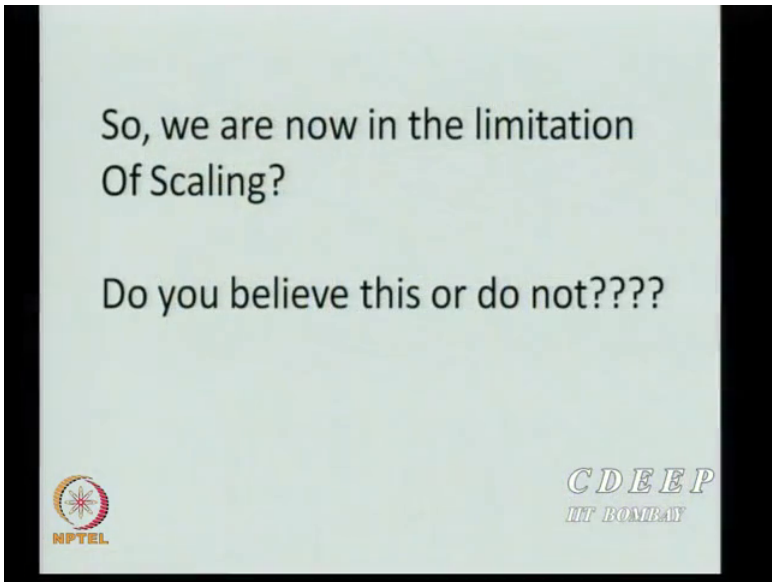
(Refer Slide Time: 49:00)



0.8 for example has only 3 atoms distance, 0.8 nanometres means 3 silicon atoms, so we are really looking for numbers which are accountable. In quantum mechanics we have large number and therefore statistics was allowed. If I have 1 to 3, how will I say this electron moves with this philosophy, so what is quantum? You can say the effect is because of quantization but actually classical mechanics can be applied I1 electron this is the mass, this is the field it should go with this.

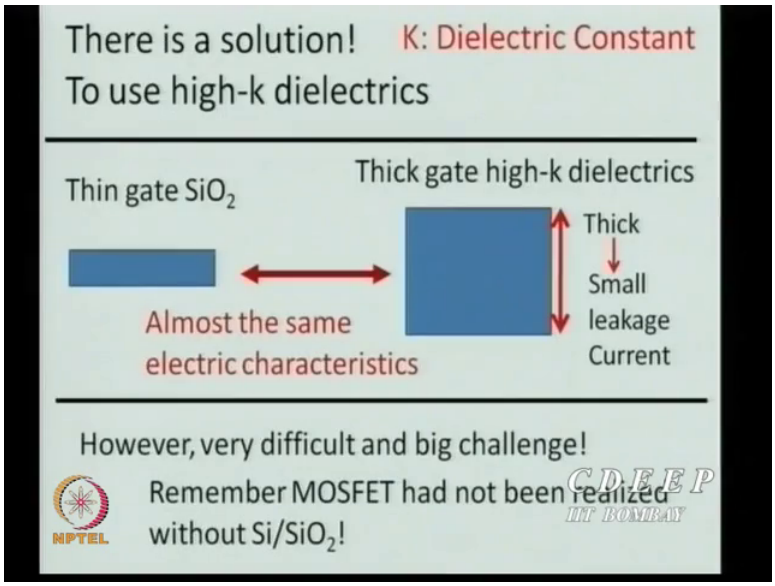
So now the issue as has started even their physics is okay if I go down to this all device people talk to big is that valid now? Your knowledge maybe good but whether it is valid. So that is why issue coming on here in physics now whether it is good enough to have models which will work for less than 10 nanometres.

(Refer Slide Time: 49:54)



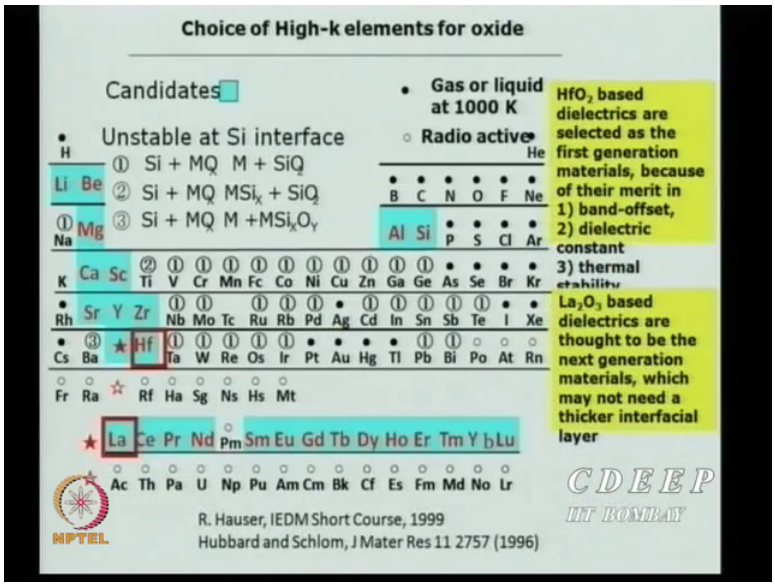
So we are now in the limitation you believe this or do not your choice.

(Refer Slide Time: 49:59)



They then say there is one solution is higher dielectric constant.

(Refer Slide Time: 50:06)



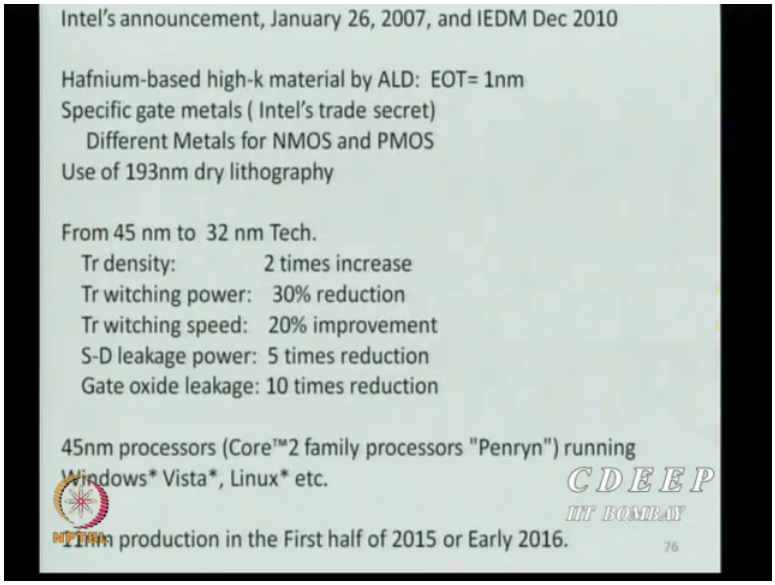
How I do I make high K? I figured out there are certain materials which are higher dielectric constant and also the actually fit into silicon technology that is most the compatibility, they should have band offsets; they should have good dielectric constant, very thermal in stability. So we figured out hafnium oxide is ideal candidate for many years even now most of the silicon chips coming from Intel ti are using hafnium oxide, maybe little nitrogen there called hafnium oxide nitride but that is what the new technologies they have.

Me and my colleagues in Tokyo Institute they have been working earlier than me just go 2 months touch hand there I am also a part. We were working on new material then of course that also new third material has appeared lanthanum oxide. Now lanthanum has a high dielectric constant then hafnium but it has compatibility silicon process is not very good but we do not use much thicker in interfacial layer between silicon and silicon dioxide I mean lanthanum oxide otherwise you need a buffer in between.

So most of the research in the last 5 to 7 years I was involved in Japan for a month, their group has also worked on your Europium, Gadolinium and einsteinium all kinds of high materials which are known to us from the periodic table only when you see, oh! You see these are also materials. So we started looking many of those their oxides, first there should be good dielectric, okay.

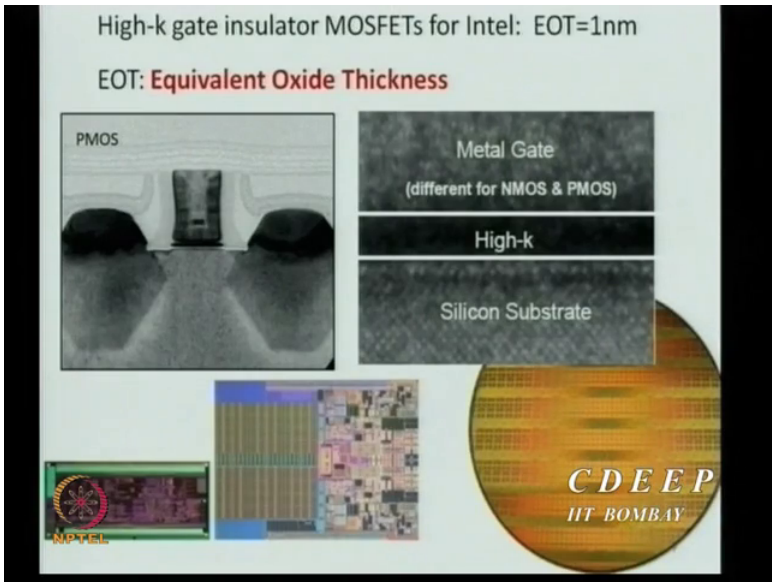
So good dielectric should have good dielectric strength reproduce able technology wise and compatible to CMOS silicon process that is major. You have very limited research but we are still working. Americium was also tried, failed it is very leaky.

(Refer Slide Time: 52:04)



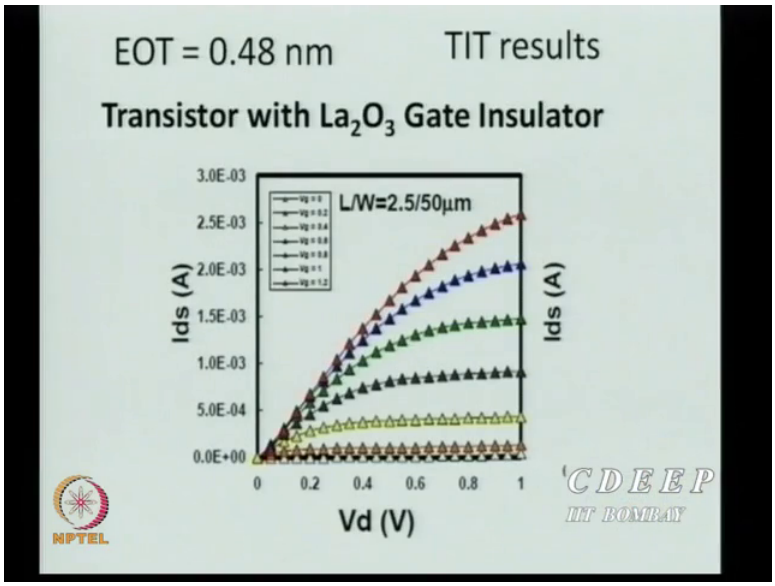
As I said there is a word we use called EOT equivalent oxide thickness silicon dioxide is major thickness. So 1 nanometre EOT is for silicon dioxide equivalent to this multiplied by directly constant and that is equivalent oxide. So even if directed constant is higher divide by equal to silicon dioxide and find what is equivalent thickness? We started with 1 nanometre and current this is around 0.2 nanometres also can be possible for EOTs, okay. 11 nanometre process first half of 2015 or 2016 would be actually on production we hope so and I only can predict.

(Refer Slide Time: 52:53)



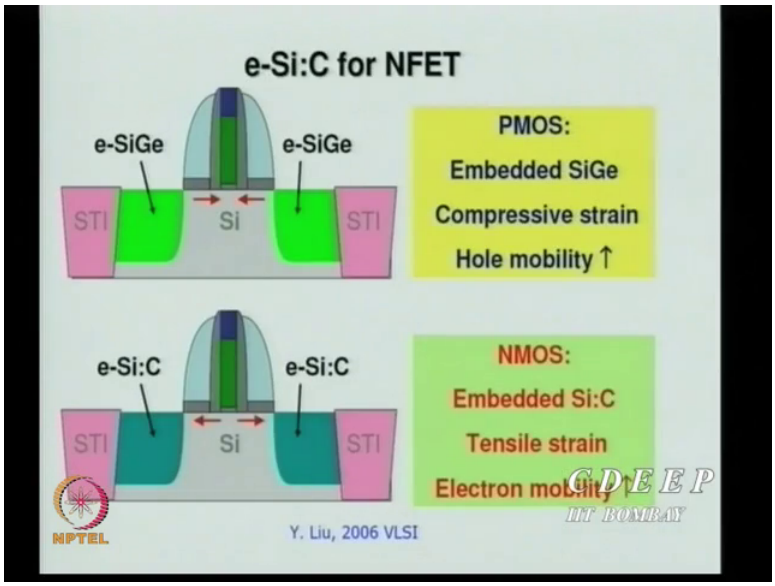
This is what high K looks like actual devise which was made in TIT and we have a metal gate instead of silicon gate which also is another change in technology and this is a high K based design, it has a small 4 bit microprocessor which has a bank of SRAMs, okay. This has an EOT of 1 nanometre.

(Refer Slide Time: 53:18)

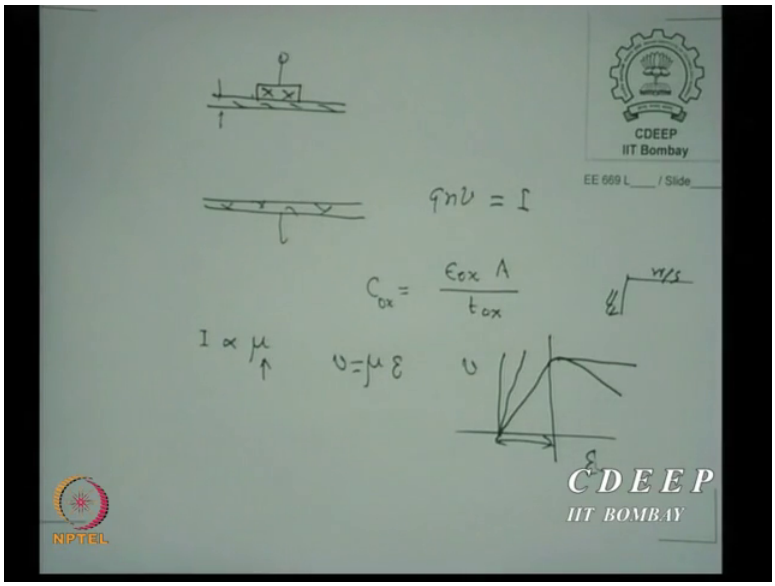


We have worked on 0.4 8 nanometres and it gives good iv characteristics that is what the importance is.

(Refer Slide Time: 53:27)



(Refer Slide Time: 53:45)



We have the (()) (53:28) Intel works, Intel has been working on process which is called one of the major interest in high-speed was if you look at it one of the most important current is I is proportional to mu in MOS transistors Ids is proportional to mu c ox, okay. So the speed is limited by available mobility. So in silicon are given high fields it may be around typically 600 centimetres square per volt second is highest ability one can get in a MOS structure.

Normally you may get even 400 to 450 best device maybe 600, 1300 is its actual bulk material but at the surface this is only possibility because of scattering. I want to improve this I must somehow see scattering were reduced and if scattering is reduced or I have a different material silicon instead of silicon which has higher mobility like gallium arsenic 8500, so I have at least 6 times higher, okay.

But then the issue started coming something like this, velocity is normally mobility times the electric field. So if I plot this we believe it is something only linear but in reality most semiconductors actually show something like this or if not falling saturates, okay. Which means after certain electric field irrespective what you do the velocity is fixed, current is q into v into n number of carriers, qnv is the current, so if v saturates that is the end of it, okay.

So how can I increase the current by increasing electric field? Because I get saturate, so I am not somehow have technology or somehow have a material you still remain here, so I increase the

field, I must get higher crunch, okay our mobility itself should be high enough (()) (55:29) something like this, okay. It but happens that I have high-speed circuits, okay. So what Intel is trying or have been doing last 10 years.

(()) (55:42) mobility varies with strains, okay. You know your mechanical engineer first-year some way we have learned Young's modulus and many other simple thing there we say that sphere is proportional to strain and if it exceeds that certain value which is called from there, permanent strain appears, okay. So we figured out that if I have a strain material in one direction the mobility wave carriers going in that direction can improve, okay.

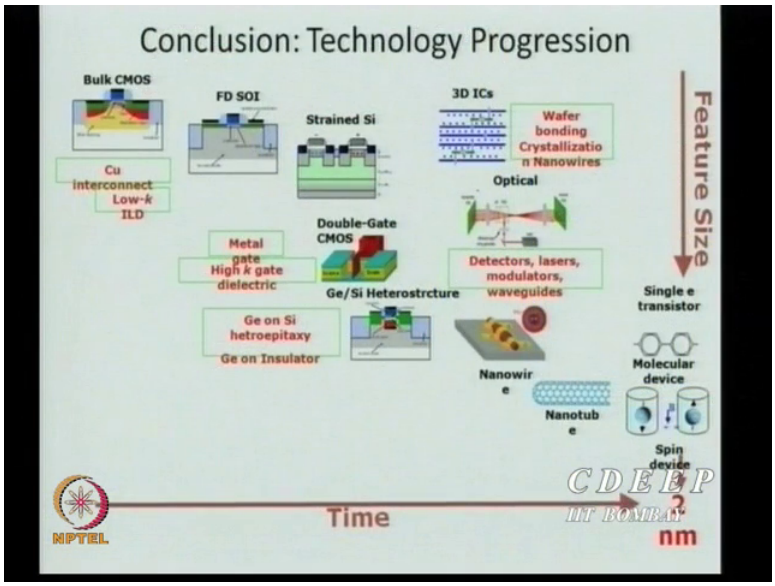
So if I put instead of normal silicon the source then made of silicon germanium than it gives a permanent strain because silicon germanium lattices are not same lattice constants. So there is a strain and they give larger mobility along the lateral directions, the. This is essentially called compressive strengths. So you have large y mobility in the channel device. For N channel what they did instead of normal silicon dioxide covers they actually put nitride and they give the (()) (56:48) the material which actually include for N channel device the mobility in lateral engines.

So now the new technology they have is not (()) (56:57) silicon germanium is here and silicon nitride at the top they are actually making mobility's much higher compared to what the limit was 600 we can go at least 800 to 900 now or even 1000. So did I look these are only technology which we figure out we can do this, okay.

However if you want higher than for other materials look for silicon carbide, they look for silicon nitride, silicon nitride-based transistors, carbon nano tips, refined based tubes all kinds of newer processes will be required if I want much higher mobility's then what you are now getting and the whole research rest on this. So there is a search for new material for gate, new high k material, new channel materials it working every year new technology appears.

New and different for source drain, so even if you say that so much is known still there is so much to do. So do not think that oh! It is all over, nothing is over till it is over therefore he is saying nothing is over till it is over. So till it is over is too far. Okay now few other slides I will show you in the last. What is going to happen technology progression?

(Refer Slide Time: 58:17)



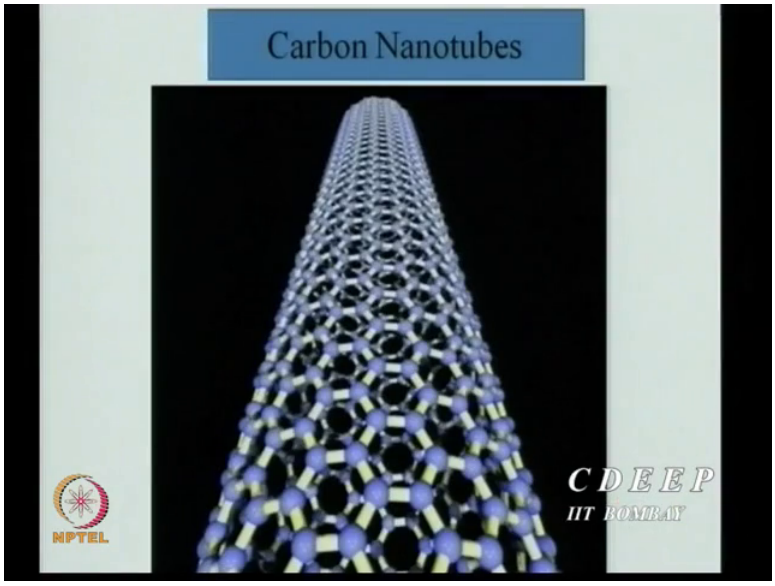
We went from bulk CMOS, we went for fully depleted silicon SOI silicon on insulator devices, strained silicon, we are looking for high k which is called double gate FET which is very first better MOS transistor created, we can make few more gates will it is called trigates, triple fin gate or multi-fin gates which is what Intel is working, so these will give better performance compact to because there is a problem which I never said.

When I scaled down there is the effect called short channel effect because of that mobility is actually reduced. So whatever you are trying to achieve actually it is lost simply because and there is a possibly be of tunnelling in the oxide because of short channels. Now because of this problem of scaling we are looking for newer technologies in which these effects can be minimized even if I scaled down and the structures which I am showing you essentially take care of that.

Okay, this is bulk CMOS, this is fully depleted silicon on insulator, strained silicon, double gate, triple gate, silicon germanium hetero structures. Now we are also looking for germanium transistor that to what we were earlier in 1940s we are looking germanium again, okay does more advantage but germanium on silicon not only pure germanium. They are looking for nano wires, we are looking for nano tubes, we are looking for single transistor devices and we are looking for spin devices which are magnetically controlled.

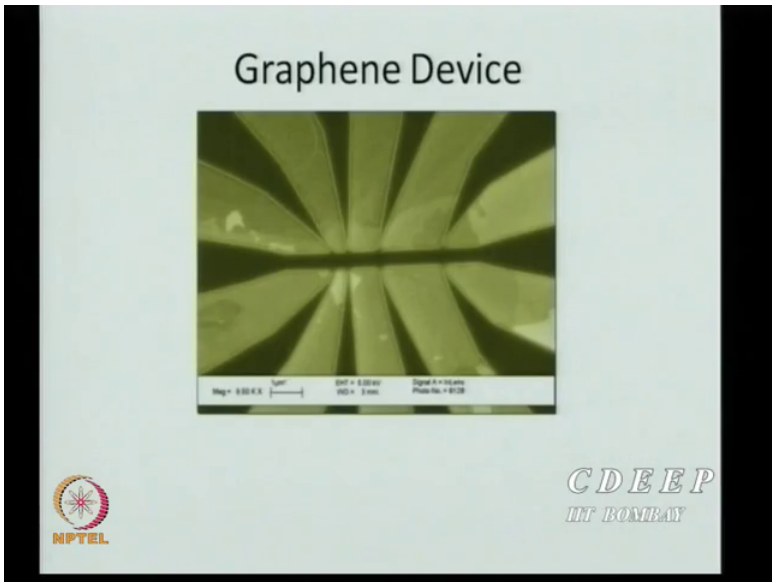
So there are number of devices under which may have 2 nanometre kind of structures and may actually do very high-speed comments, very low-power performance but why we are not reached so fast here, they also can have optical devices made out of that. So the problem there is all these cost and rather if you can make systems which are highly reliable low-cost only then this may go into fab of a company but effort is on and I think there is no reason why we will not be able to reach and the most optimistic. There are 2 kinds of people I think many people know, if the glass is half filled the optimist say it is half full, pessimist says it is half empty it is a matter of your way, okay.

(Refer Slide Time: 60:38)



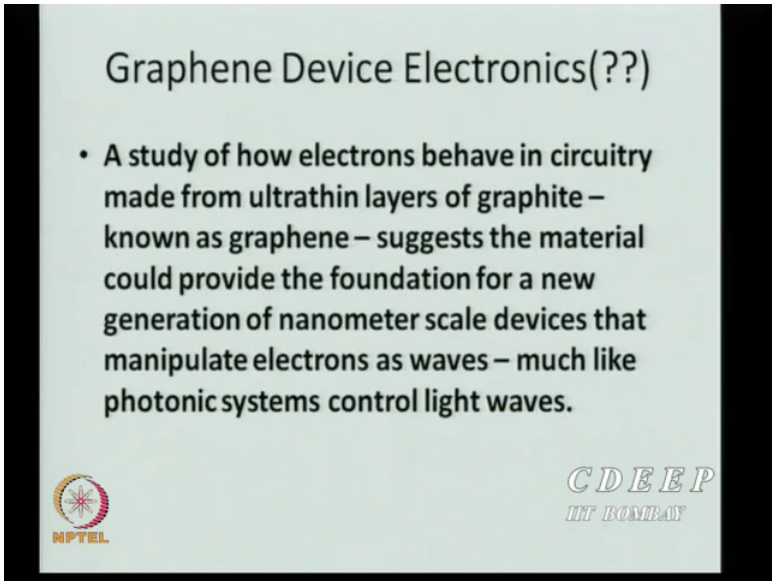
This is a carbon nano tube which is what it looks like, it is hexagonal structure.

(Refer Slide Time: 60:46)



And there is a new material a new nano wire has appeared which is based on Graphene, Graphene is one form of a carbon which has a hexagonal diamond lattice, okay. It is one of the hardest material right now available on Earth, okay. So one is looking for Graphene as a new material between either interconnect or instead of normal carbon nano tube you are trying to put Graphene films there, okay.

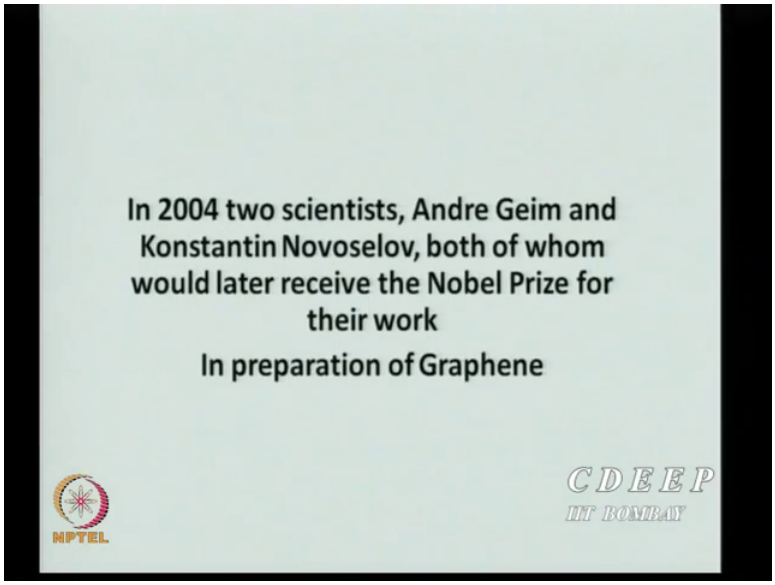
(Refer Slide Time: 61:16)



It is suggested electron behave in circuitry made of ultra-thin layers of graphite known as Graphene it suggest the material could provide foundation for a new generation of nanometre delivers is that manipulate electrons as waves and that is most important. Electrons are not no more particles they will act like EM wave, okay and that is what we are looking for and therefore we can make photonic Systems out of this material, okay because we need photons, okay.

So that is something what is coming this is the major feature of people who are looking in devices. The band structure of Graphene is very funny there is no band gap like silicon semiconductors there is no (()) (61:58) like metals, okay but conduction balance band meet at one point at one k, okay. So you can control the current motion only when you reach that k. So that if you are away from that no current if you are at that base moment you get currents on and off, okay. So Graphene can now be also looked as a switch device which is very recent and people are looking for it. Look at it, normal band gap for silicon is, there is a band gap of 1.1ev between conduction band and a balance band there is only one point where it has common point.

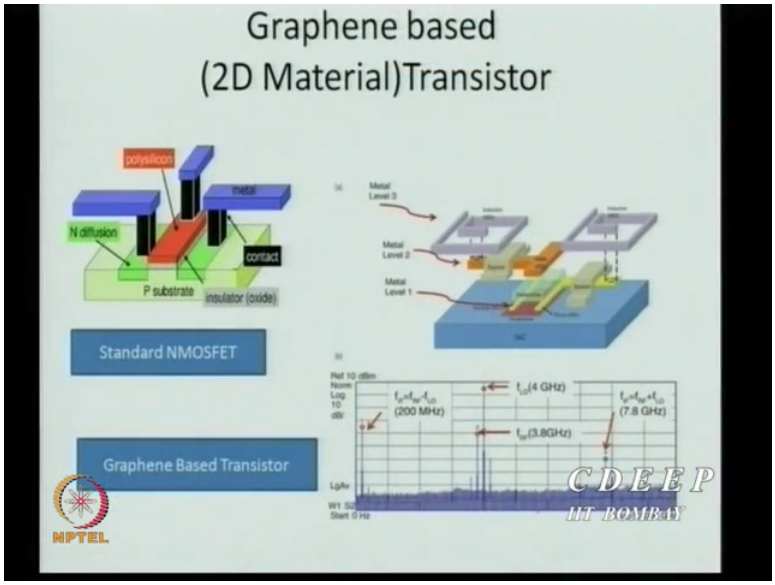
(Refer Slide Time: 62:39)



In 2004, 2 scientists Andre Geim and Konstantin Novoselov both from Manchester University won the Nobel Prize for their Graphene research. Before that carbon is known for 200 years earlier, coal is known many many years, devices people also know as early as 30s or 20s microphones were used with carbon films. So it is not that carbon was unknown but it took almost another 80 years before Graphene-based devices were thought, okay.

So someone has to even whatever people say Newton could see that Apple falling, apples are still falling something else is also falling but we do not care but Newton cared that market. So same way things are still happening your still chance to win the Noble prize this is what it is.

(Refer Slide Time: 63:31)



This is a normal mass transistor and this is Graphene has a small film below, so there is electron moves through the Graphene area their conductivity is very very high typically it could be infinite also in ideal state, so huge currents can actually flow no on resistance, okay. So it can be very high speed device this is what is being tried. So we are working for a, I do not know whether you are.

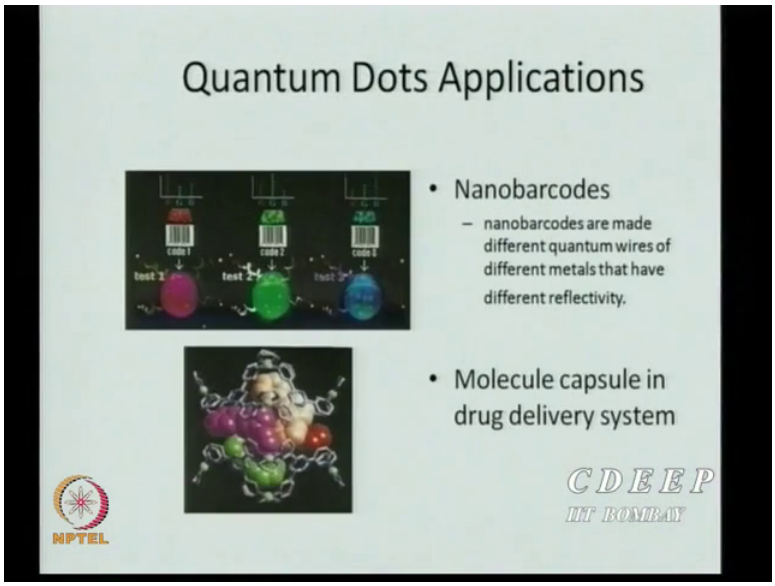
So far now there is a Micro band which was around 100 gigahertz microwaves millimetre waves and we have our optical range which was around 1000 microns and above, infrared and above there is a in between band which is called terahertz band and it is not being used by anyone and this is the band which Graphene may allow that is why this material is now ideal material for photonics, ideal material for semiconductor is research and ideal material for EM waves theories. So this material probably may actually revolutionise our thinking, so new technologies are coming.

Last but not least compatible Graphene there is another example we are trying, what were my problem when the carrier go from source to drain they scatter and therefore mobility goes down.

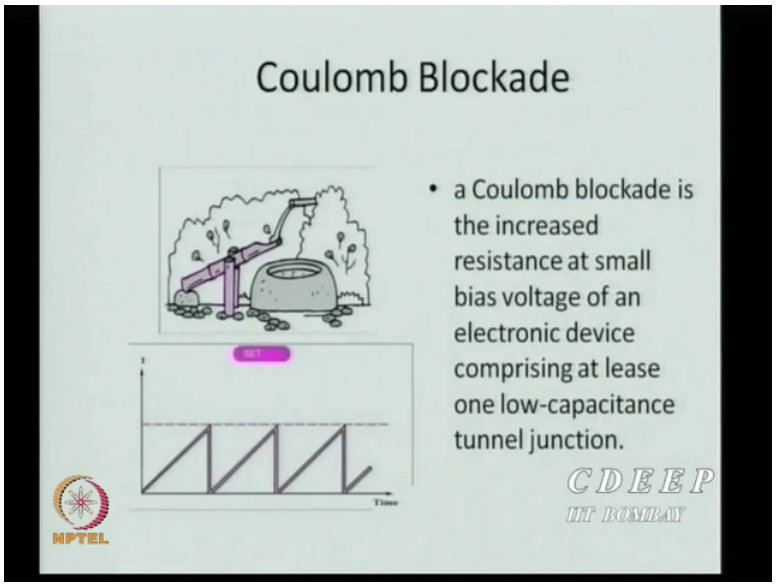
If I keep the vacuum no interactions mean (()) (64:45) is infinite, so all carriers will go. If I do this than the vacuum tubes are 1906 I am bringing back, okay. That is what they did, okay.

I am now trying to create a vacuum tube in semiconductor, okay. How to create this vacuum here is a game which is not so easy but possibly yes, okay and therefore tomorrow you may say a solid-state vacuum tube, okay which is the …

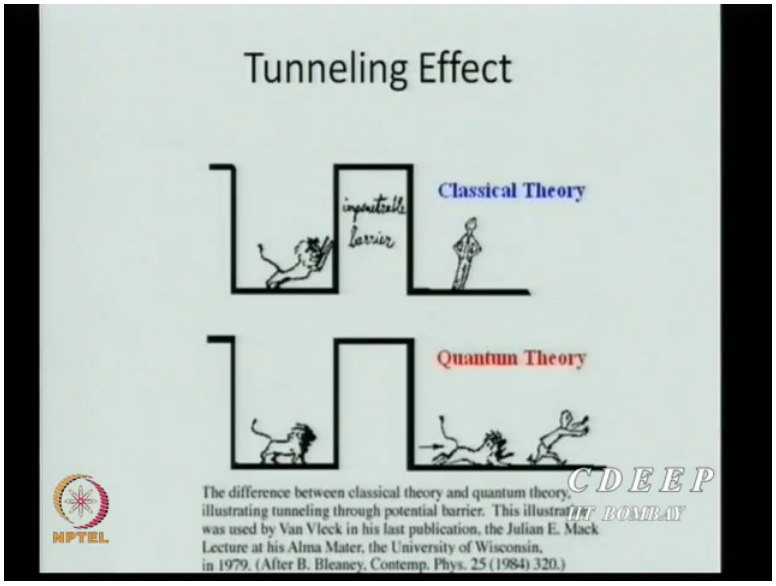
(Refer Slide Time: 65:18)



(Refer Slide Time: 65:22)



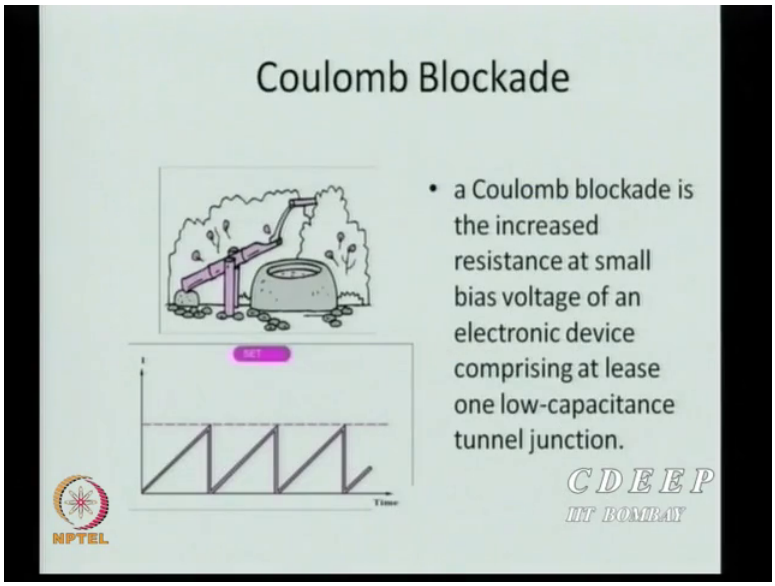
(Refer Slide Time: 65:23)



Okay, we will move further are many plot devices, many applications, this is interesting many of these devices which Coulomb these are called Coulomb blockade devices they work on tunnelling, there is a lion and there is a human being, of course now we can say Leopard, there is a barrier and maybe we have made a fencing which was very small or probably I (()) (65:40) many places whichever reason.

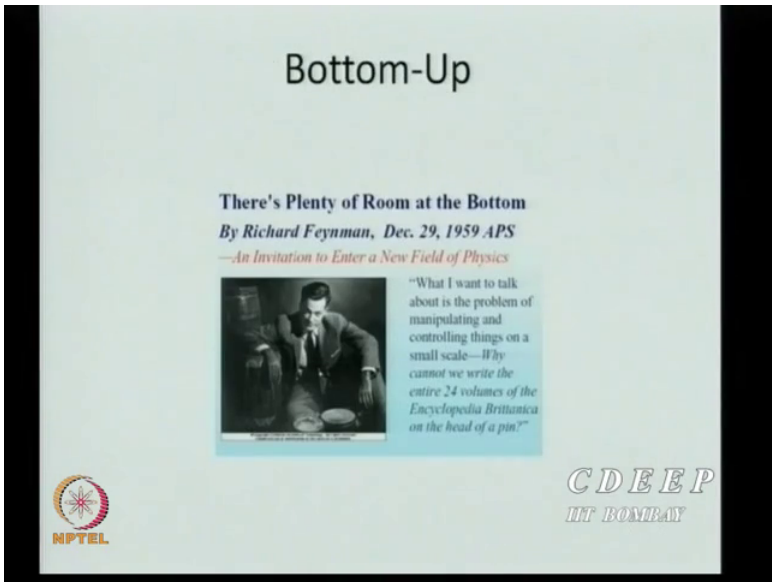
So Leopard could not climb for some years, 1 or 2 years Leopard did not come to our side because the barrier was too much for them, what they did? Some human helps them and cut the wires make a ditch, so now they can cross across this barrier without climbing and they can chase you, okay this is tunnelling.

(Refer Slide Time: 66:08)



Based on this the new devices may come which are Coulomb blockade the advantage is when they cannot go of stake then they can go on stake switch very fast, tunnelling is very high-speed phenomena, so one expects very high-speed transistor is coming out of Coulomb blockade devices.

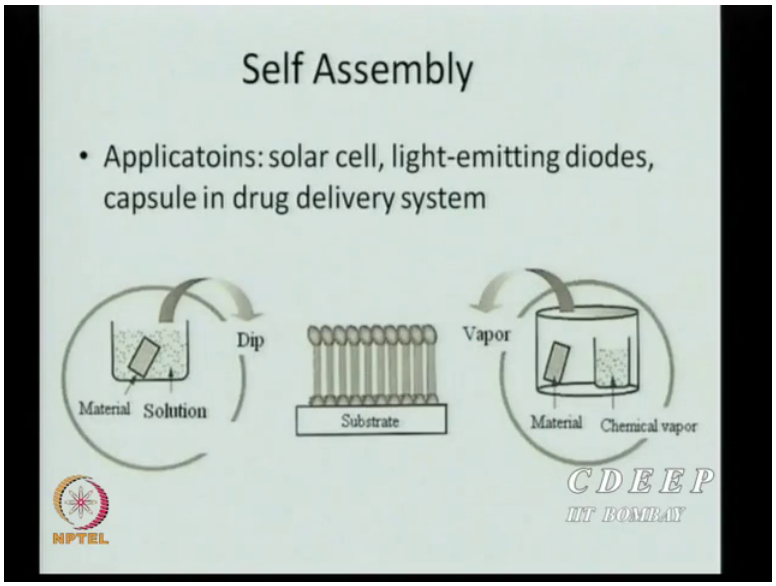
(Refer Slide Time: 66:24)



Last one Feynman is a very famous physicist, he said what I want to talk about is the problem of manipulating and controlling things on a small scale because they used to have macro scale; physics always use to work on macro scales. So why cannot we write and that 24 volumes of Encyclopaedia Britannica on the head of a pin? Why cannot we think like this?

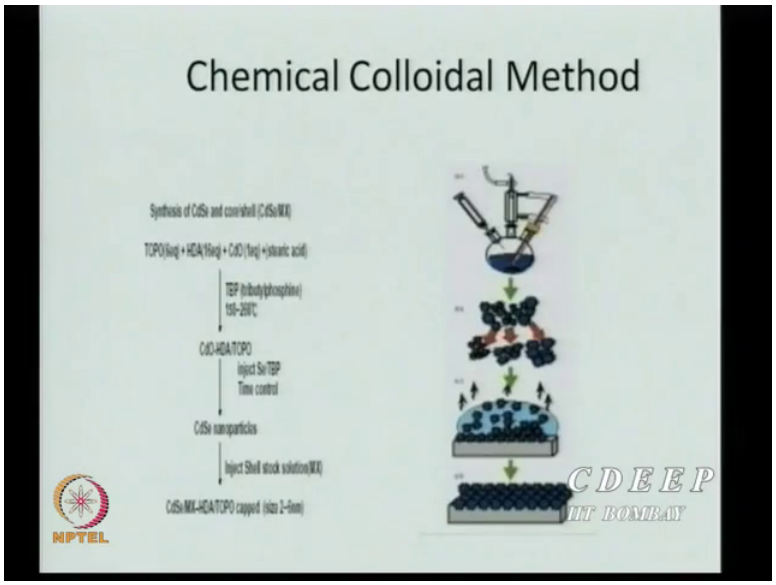
Okay that is Micro or nanostructures, this is Feynman idea in 1959 and we are still working for him. By the way he was a famous Novell (()) (66:57) and very funny person and very interesting person they say.

(Refer Slide Time: 67:02)



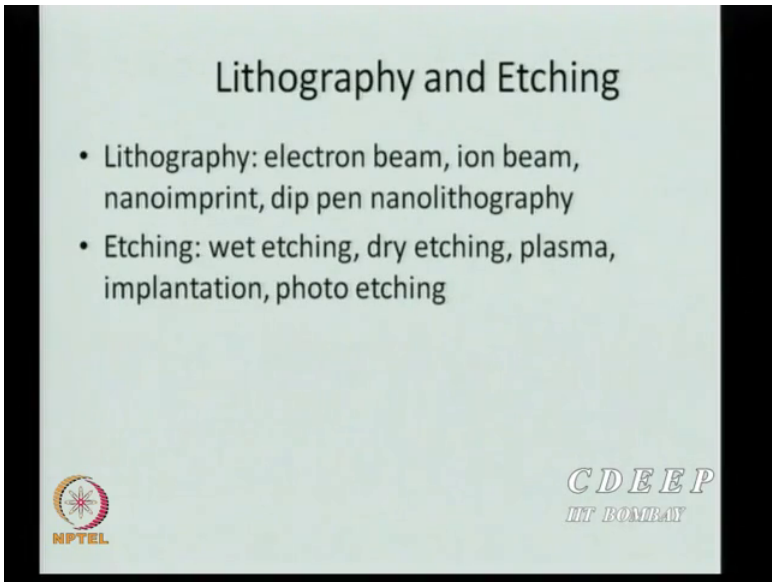
Last part is that there are some molecules which can self align in solutions or by applying light these are called self align lectures.

(Refer Slide Time: 67:12)



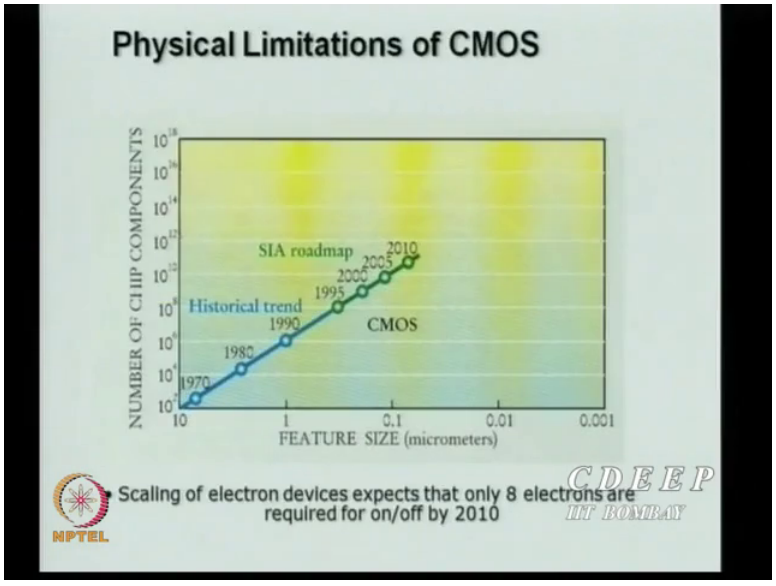
So you can see there are some materials you can inject the atoms will come and self align no lithography, you decide there they will sit there, okay. So something may appear of course success but again large numbers is not a success.

(Refer Slide Time: 67:27)



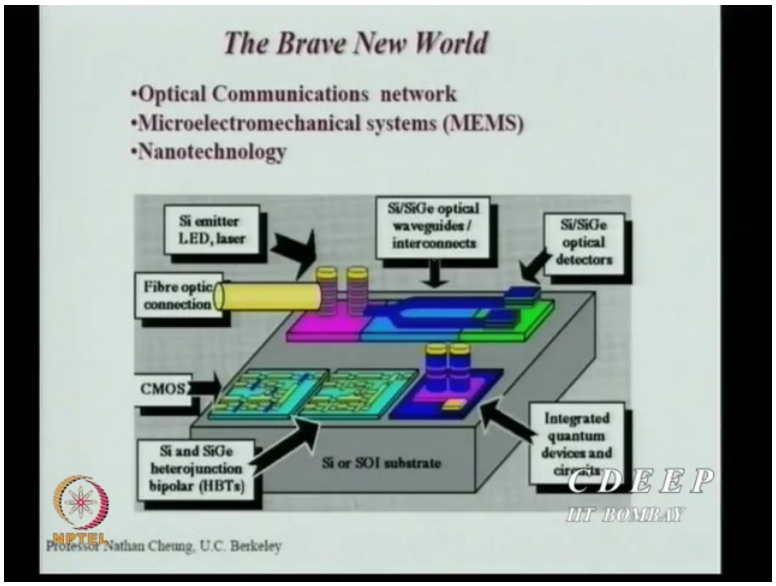
The problem in lithography can be solved or not solved is a major crux. Anything should be dry or wet.

(Refer Slide Time: 67:37)



Okay, so last but not the least limitation of CMOS in 10 years, we are looking for 8 electrons per bit, Manufacturing cost 50 billion dollars per fab line, I said 8 now it may become this, Moore's law scaling to an end? We do not know.

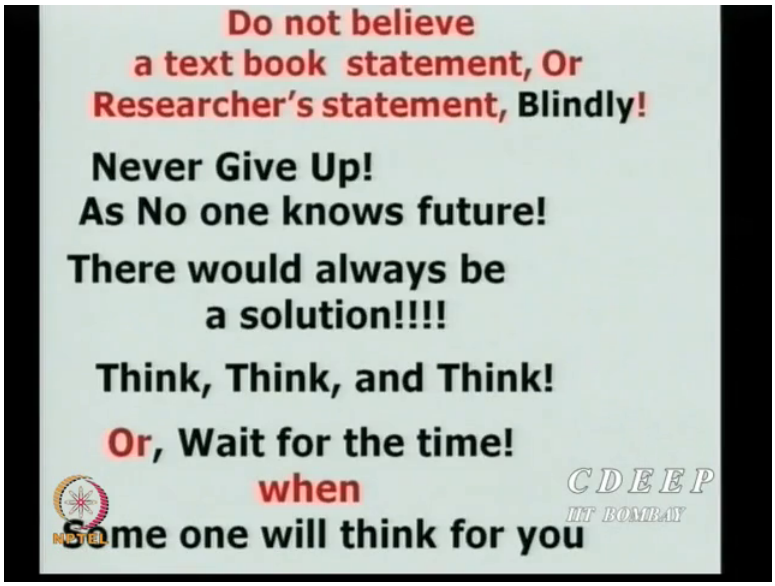
(Refer Slide Time: 67:55)



However humans are great we are now looking for a Brave New World which has come from Professor Nathan from Berkeley. Can see every kind of system can be built on silicon substrate, okay. You have CMOS devices, we have silicon germanium heterojunction devices, we have fibre optics connection, we have silicon LEDs, we have silicon germanium optical waveguides, we have silicon germanium optical detectors and we have quantum dot-based devices, okay.

All in one, okay. All in one, so any kind of system you think probably is too able if only you apply your mind and time. Pretty interesting this will be shown on your website.

(Refer Slide Time: 68:42)



So one last word for you do not believe a textbook statement or a researchers statement blindly including me. Never give up as no one knows future. There would always be a solution think, think and think or wait for the time when someone else will think for you. That is what most of us do. So as much as you can think probably there is hope many slides were taken from various people, various sites, so I acknowledge all of them.

Iwai Hiroshi is my close friend he gave me many slides, Shekhar Borkar from Intel gave me slides, Professor (()) (69:26) they also gave their slides. So these slides are as per permissions and they are not shown without their permissions, thank you..