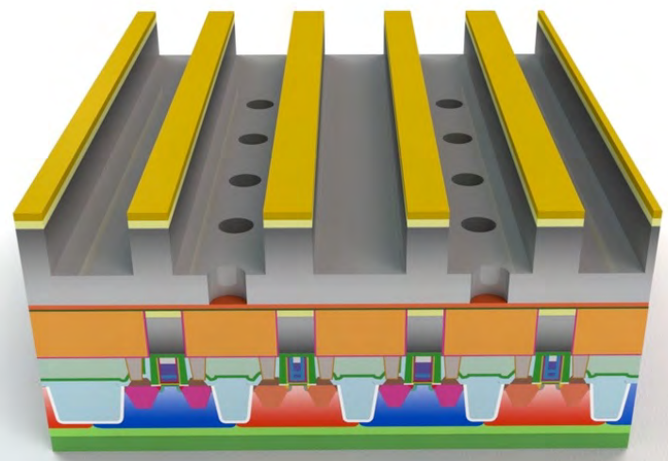
Process Flow for 22nm Planar Logic



© 2018 Threshold Systems Inc.

Presented by

Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Table of Contents:

ThresholdSystems

22nm Planar Process Flow

1. 22nm Logic Fabrication Sequence ..... page 3
2. Shallow Trench Isolation Module ..... page 5
3. Well Module Definition ..... page 23
4. Device module Processing ..... page 33
5. The Contact Module ..... page 79
6. Contact Module: Polished-back Contacts ..... page 85
7. Device Module Gate-Last ..... page 89
8. Back-end Metallization ..... page 103

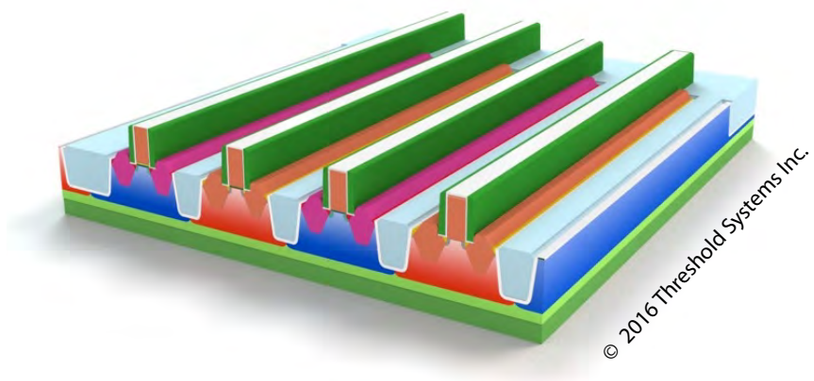
© 2018 Threshold Systems Inc.

All rights are reserved. No part of this publication may not be reproduced, stored in a retrieval system, or transmitted, in any form or by any means , electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

ThresholdSystems

22nm Planar Logic Fabrication Sequence



© 2018 Threshold Systems Inc.

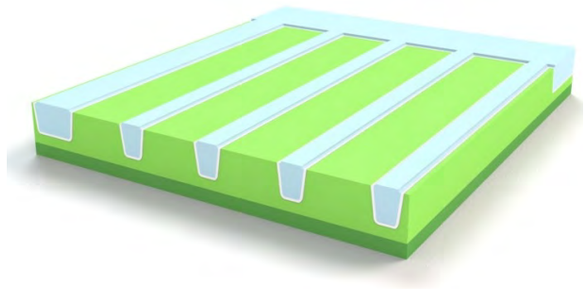
Presented by
Jerry Healey

© 2018 by Threshold Systems Inc. All Rights Reserved.

Processing Modules:

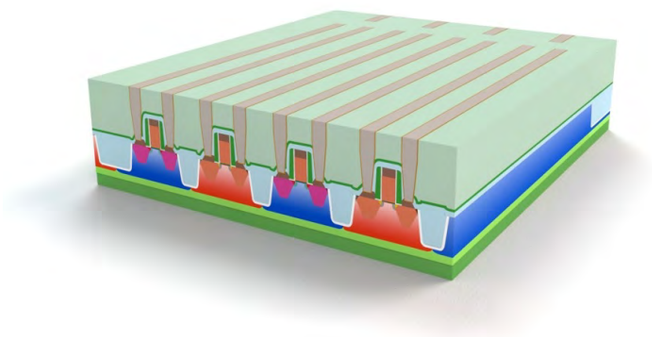
ThresholdSystems

The STI Module



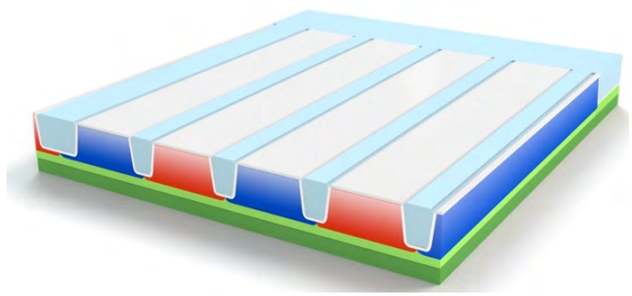
© 2016 Threshold Systems Inc.

The Contact Module



© 2016 Threshold Systems Inc.

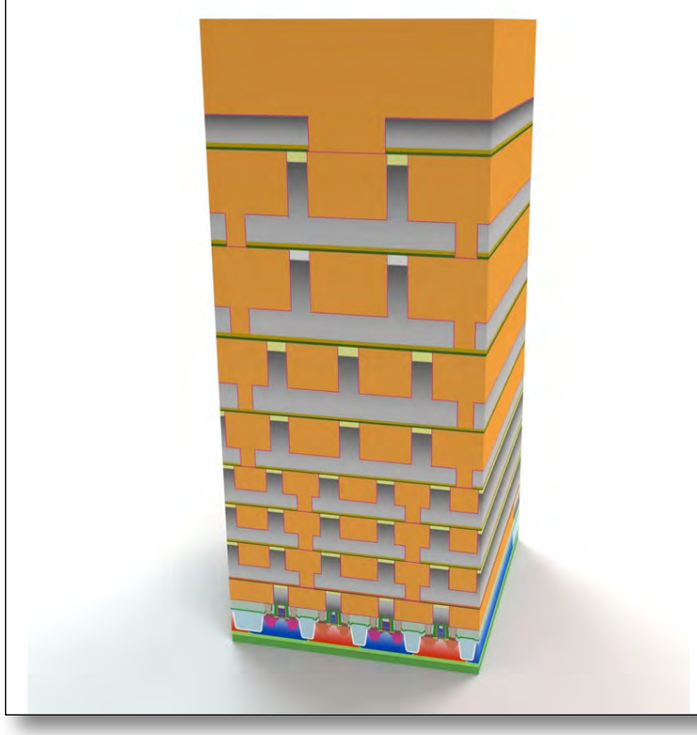
The Well Module



© 2016 Threshold Systems Inc.

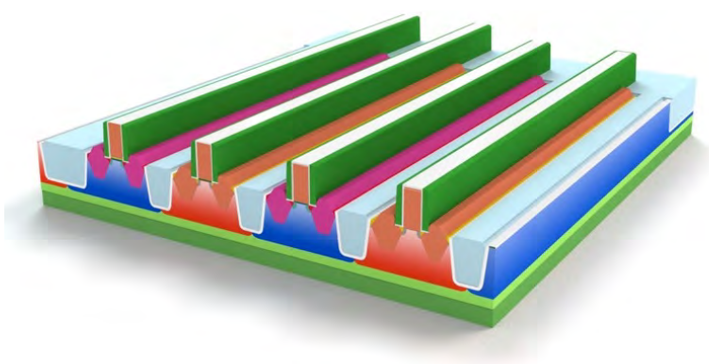
© 2018 Threshold Systems Inc.

Metallization



© 2016 Threshold Systems Inc.

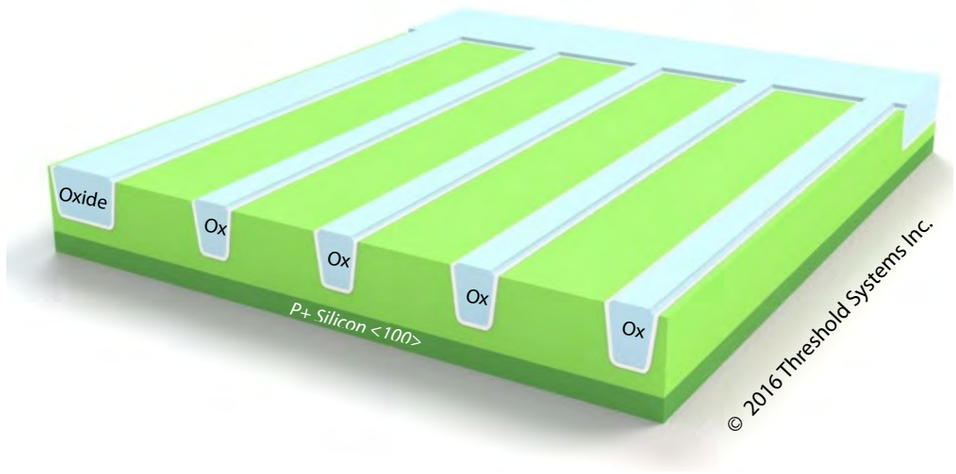
The Device Module



© 2016 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

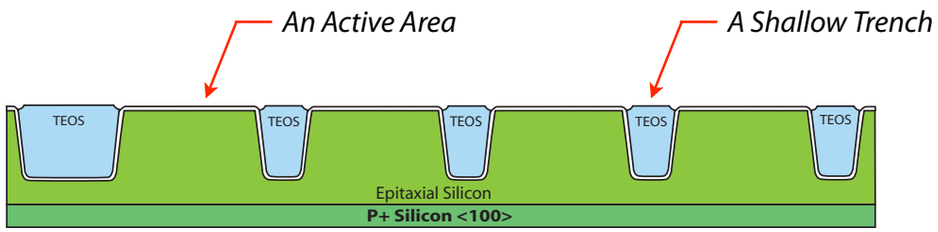
Shallow Trench Isolation (STI):



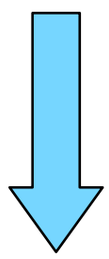
© 2018 Threshold Systems Inc.

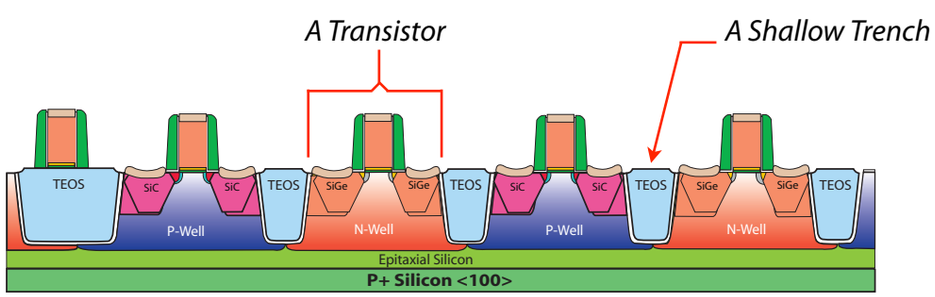
© 2018 by Threshold Systems Inc. All Rights Reserved.

The Purpose of STI:



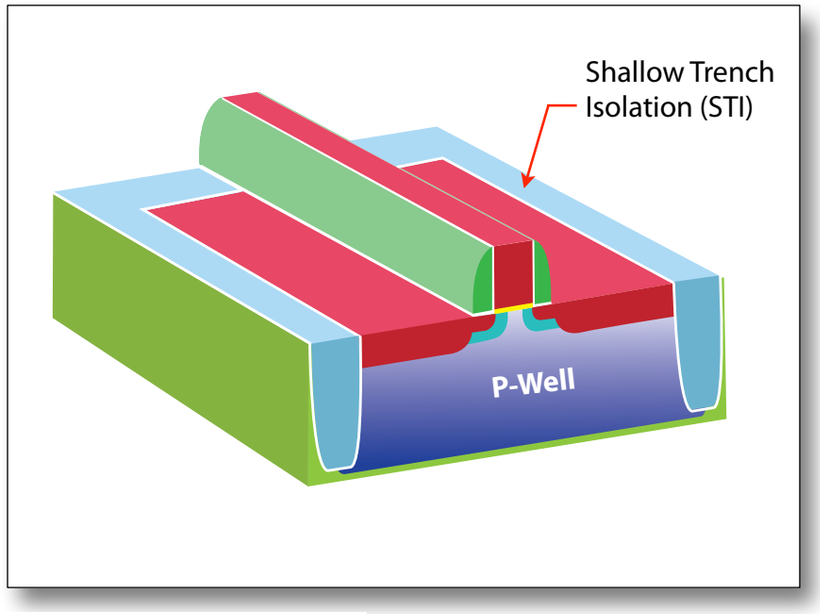
© 2016 Threshold Systems Inc.





© 2016 Threshold Systems Inc.

- *Shallow Trench Isolation consists of a series of trenches that are etched into the silicon surface and which are filed with silicon dioxide (SiO_2) which is simply glass.*
- *Glass (SiO_2) is a superb electrical insulator and by surrounding the active areas (where the transistors will reside) with glass-filled trenches we can ensure that the transistor structures are electrically isolated from each other.*
- *Beginning at the 350 nm node STI replaced LOCOS as the isolation technology of choice.*



© 2016 Threshold Systems Inc.

© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

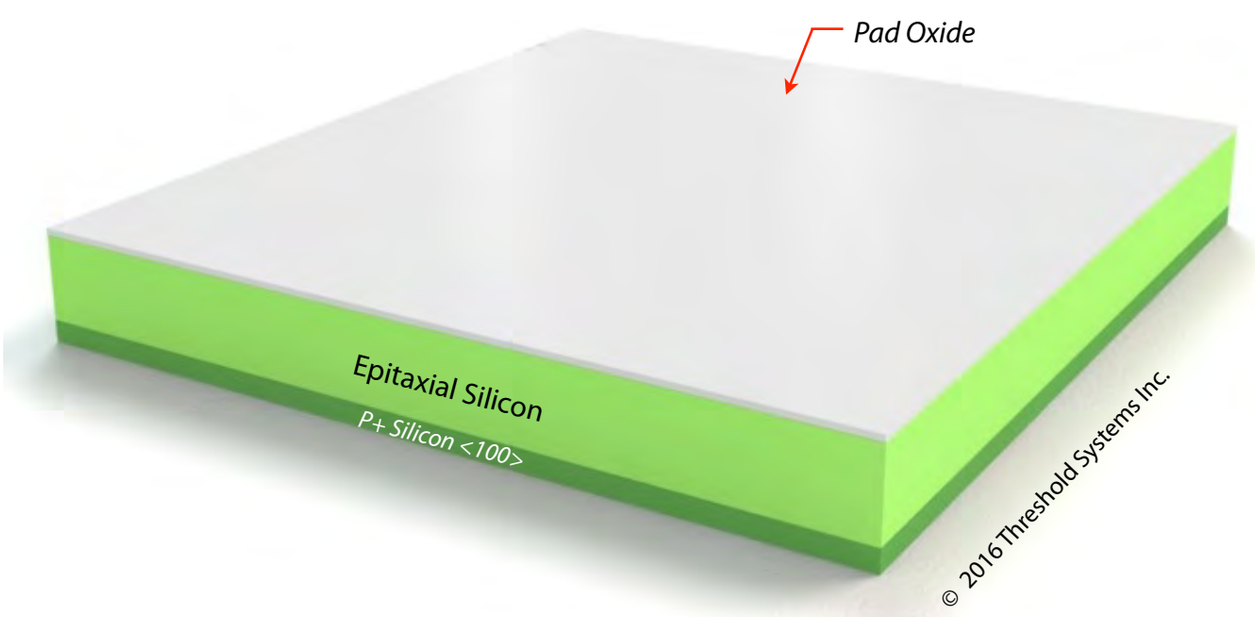
Pad Oxide Growth:

The starting material for this type of process consists of <100> heavily doped P+ Silicon with a typical resistively of 0.01 Ω-cm. The front-side of the wafers are covered in a 2 µm thick layer of Epitaxial Silicon which has a resistively of approximately 14 Ω-cm.

© 2018 Threshold Systems Inc.

Source: Applied Materials

A "wet hood" that contains four cleaning stations.



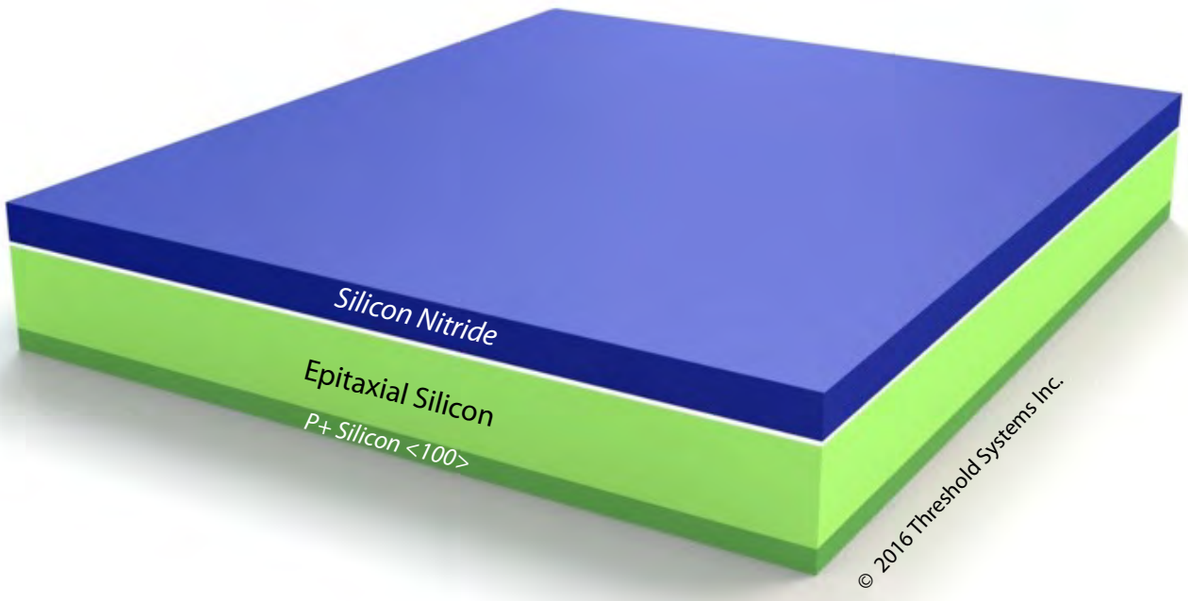
The wafers are cleaned with Piranha + HF + SC1+ SC2 and 100Å of pad oxide is grown across their surface.

The most common cleaning agents are:

- ***Piranha*** *- to remove organics*
- ***SC1*** *- to remove particles*
- ***SC2*** *to remove heavy metals*
- ***Hydrofluoric acid (HF)*** *to remove silicon dioxide*

© 2018 by Threshold Systems Inc. All Rights Reserved.

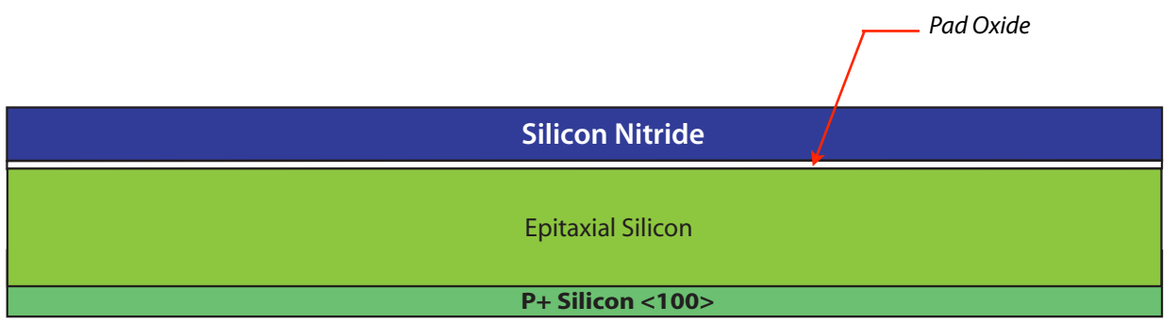
Nitride Barrier Deposition:



Next a 1,200 Å thick layer of Silicon Nitride is deposited via CVD.

The wafers are then cleaned in a Megasonic with SC1.

© 2018 Threshold Systems Inc.



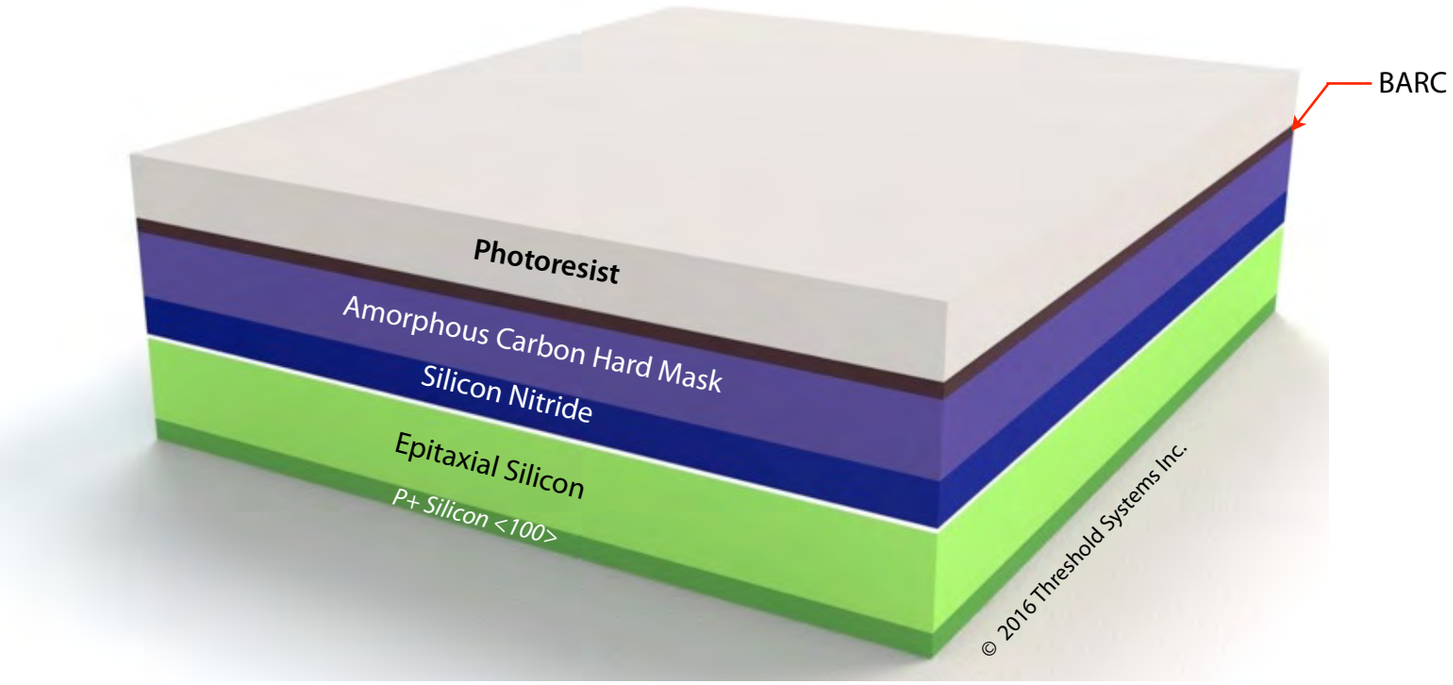
© 2016 Threshold Systems Inc.

Silicon Nitride is generally a very tensile film relative to silicon. They have differing coefficients of thermal expansion and contraction. For this reason nitride is never deposited directly onto a silicon surface as this differential expansion and contraction could induce dislocation faults into the silicon when the wafer is heated and cooled.

As a consequence, when silicon nitride is to be deposited onto a silicon surface, a thin layer of silicon dioxide, know as a "Pad Oxide", is first grown on the silicon. The nitride is then deposited onto the Pad Oxide. The Pad Oxide acts as a stress relief buffer between the silicon nitride and the underlying silicon surface and prevents the silicon nitride from inducing any damage into the silicon.

© 2018 by Threshold Systems Inc. All Rights Reserved.

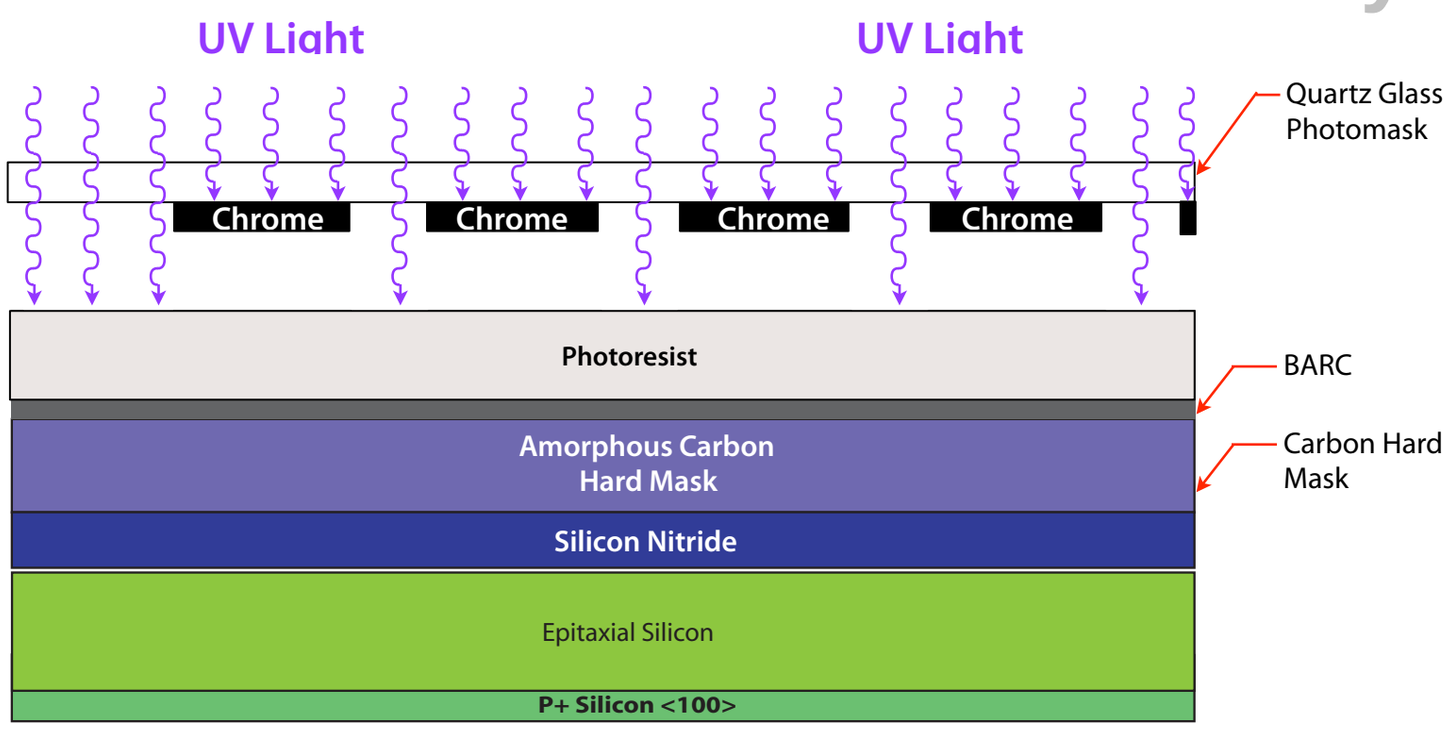
Hard Mask Deposition & Lithography:



2,000 Å thick layer of amorphous Carbon is deposited across the wafers surface using CVD. This is followed by a 500Å layer of Bottom Anti-Reflective Coating (BARC). The amorphous carbon will act as a hard mask.

Next, a ~3,000Å of photoresist is spun onto the wafers and soft baked.

© 2018 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

The wafers are exposed to 193 nm UV light through a photomask that is made of transparent quartz and patterned with chrome that is opaque to UV light.

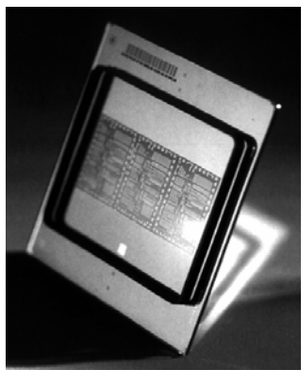
© 2018 by Threshold Systems Inc. All Rights Reserved.

Photolithography:

- *The Lithography tools that print microchip circuits are extremely sophisticated and are the most precise instruments in the world. They are capable of printing images as small as 20 nm and aligning to accuracies of less than 1 nm.*
- *They are called "scanners" because the illuminating light scans across a chip printing one microchip at a time. (They used to be called "steppers".)*
- *Photolithography scanners employ excimer laser light sources at wavelengths of 193nm and super-sophisticated optical optical pathways that terminate in a column of precisely calibrated lenses.*
- *The lens column shrinks the precise patten on the reticle (mask) and prints it into the photoresist located on the wafer's surface.*
- *Through the use of Resolution Enhancement Techniques (RET), Optical Proximity Correction (OPC) and other optical processing, modern scanners are able to print images smaller than the wavelength of the illumination light (193nm) being used.*
- *A state-of-the-art scanners can cost as much as 70 million dollars and can process up to 200-250 wafers per hour. A modern semiconductor fab may require 20-to-30 scanners.*
- *All process flows at the 20nm node employ 193nm immersion scanners which are denoted as "193i".*

© 2018 Threshold Systems Inc.

A Photomask



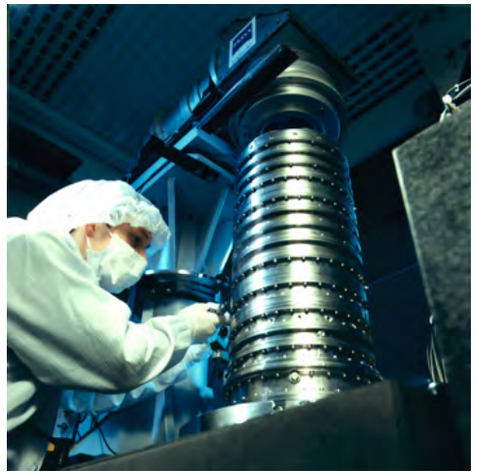
A Lithography Lens

© Courtesy of Zeiss

Lithography Equipment

© 2010 ASML

A 193i immersion scanner.



© Carl Zeiss

The Lens column in a scanner.

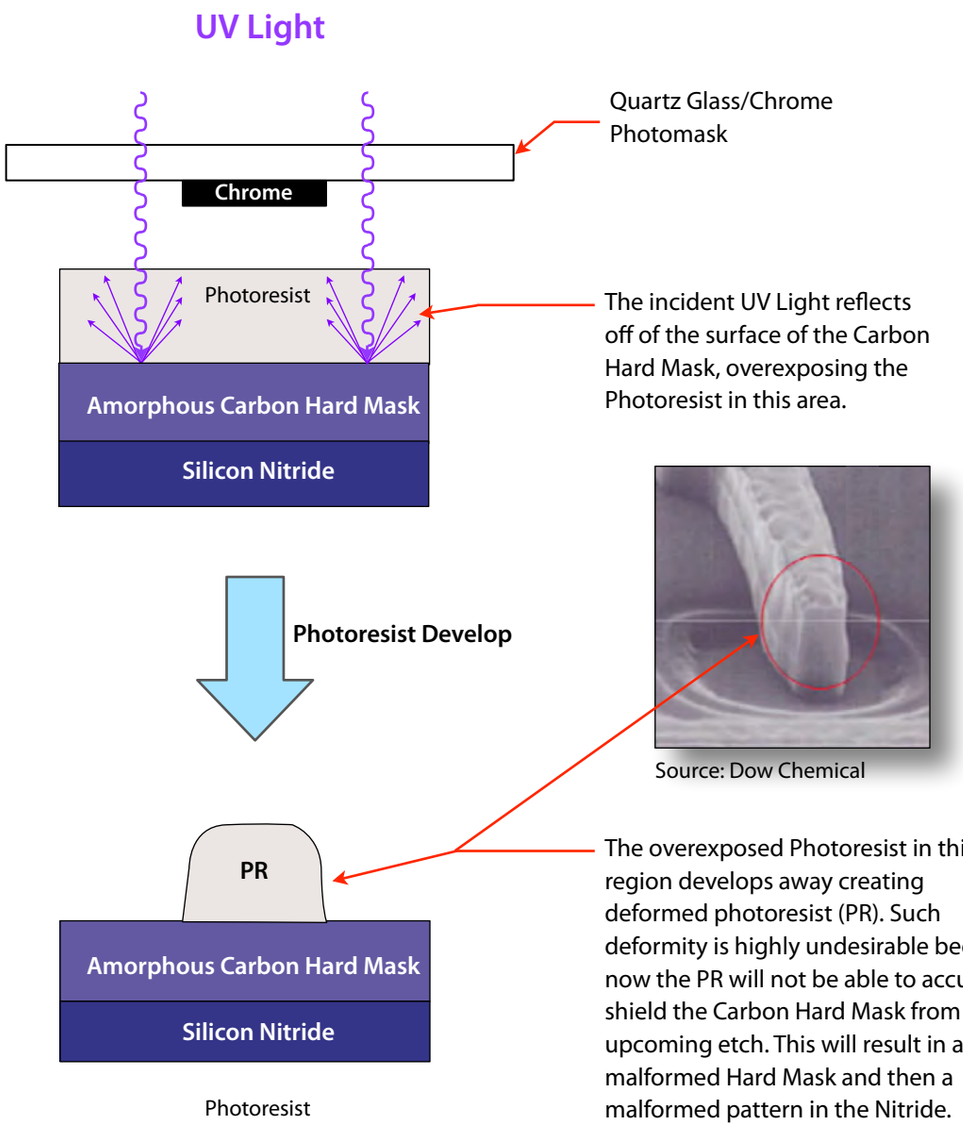
In immersion lithography a tiny drop of water is present between between the lens and the wafer.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Bottom Anti-Reflective Coatings (BARC):

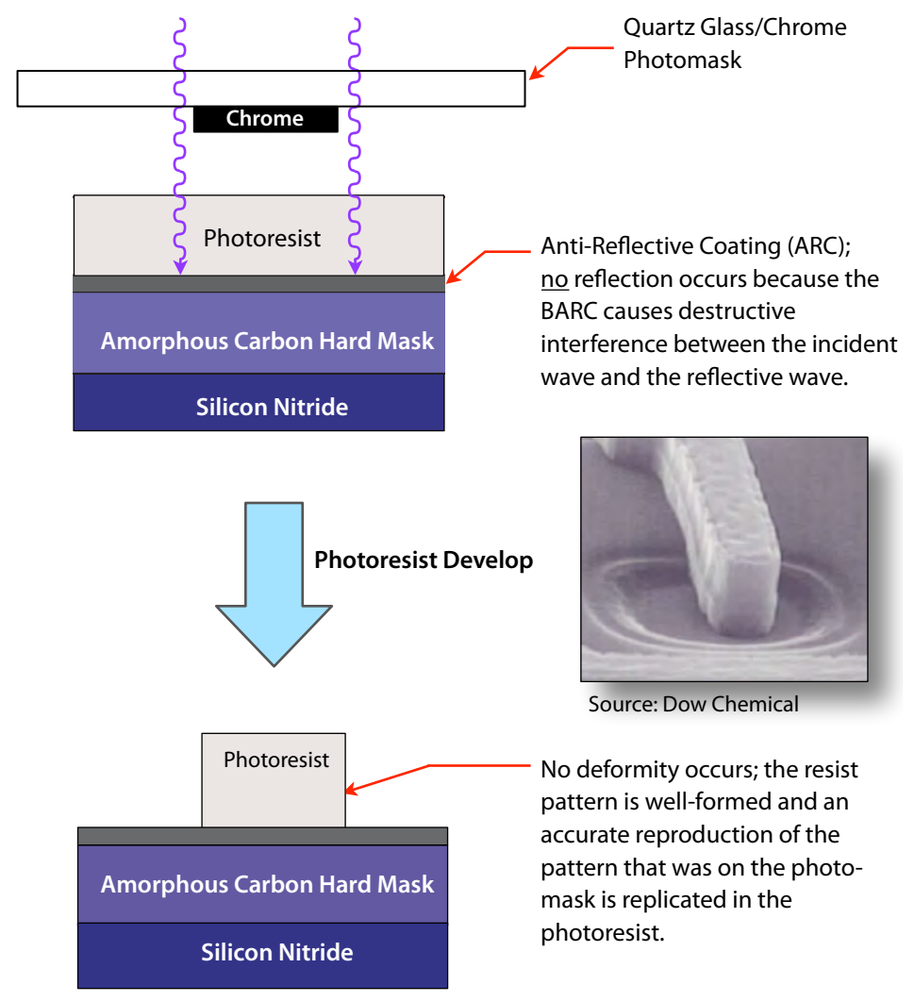
ThresholdSystems

Without BARC



© 2016 Threshold Systems Inc.

UV Light With BARC



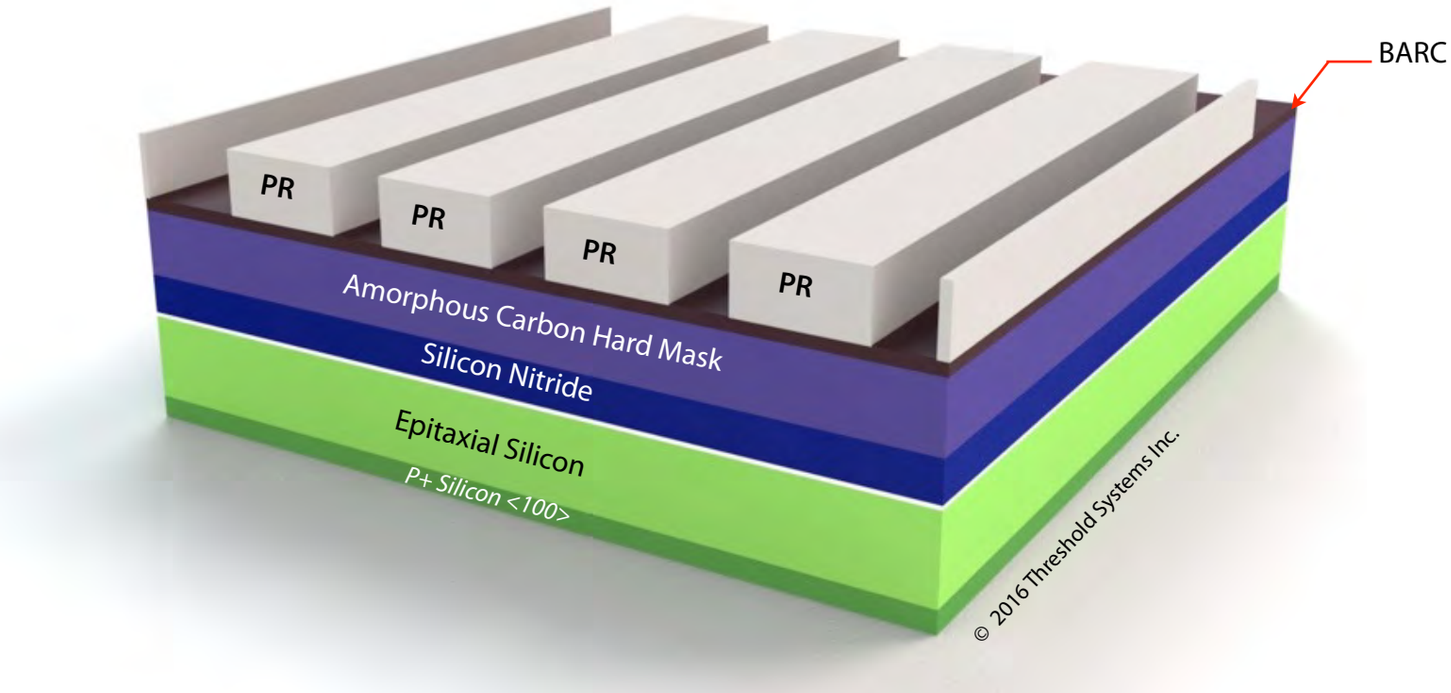
© 2016 Threshold Systems Inc.

© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

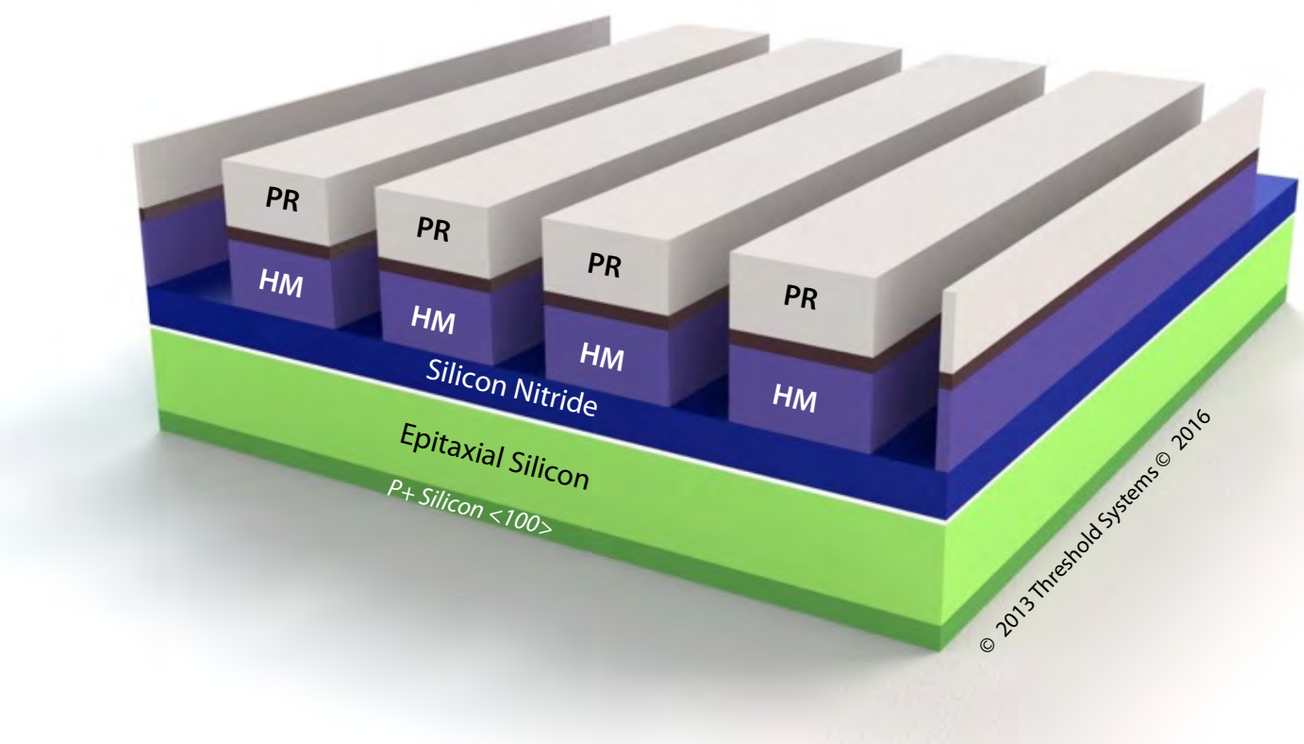
Hard Mask Definition:

ThresholdSystems



After patterning the wafers are soft-baked to give the resist a measure of structural integrity and then the resist is exposed and developed. The resulting photoresist pattern then undergoes a Post Exposure UV Bake (PEB) to finalize the photoresist pattern and to turn the photoresist into a cross-linked imidized mass that will resist etching.

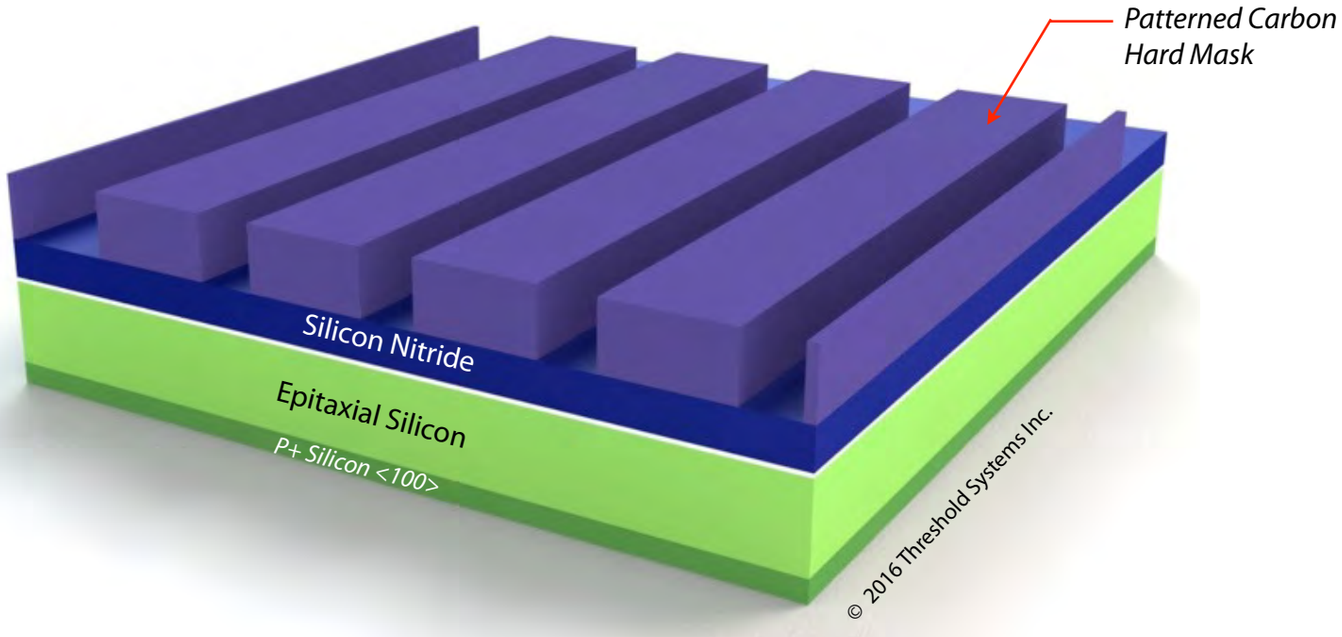
© 2018 Threshold Systems Inc.



A highly anisotropic etch cuts down through the BARC and Hard Mask layers. The photoresist is largely unaffected by the etch process and protect the material beneath it.

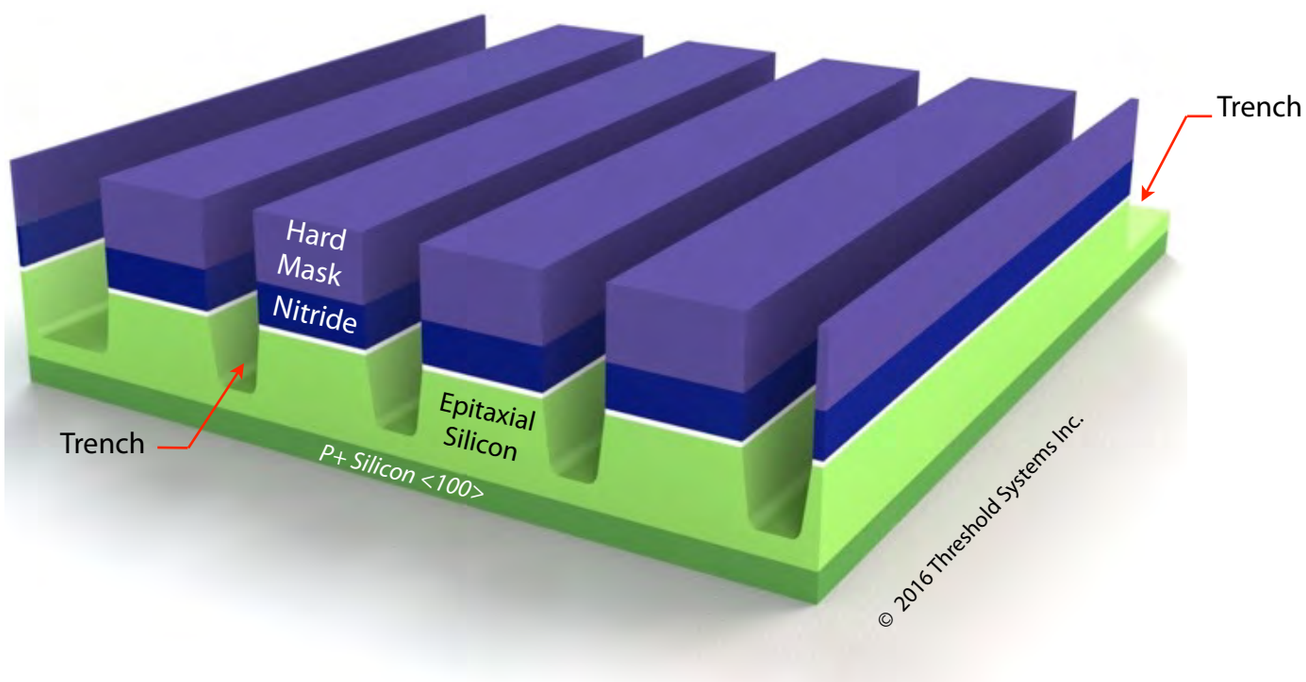
© 2018 by Threshold Systems Inc. All Rights Reserved.

Photoresist Strip & Trench Etch:



The photoresist and the BARC are removed leaving behind the patterned amorphous carbon hard mask. The wafers are cleaned in Piranha to remove any residual photoresist.

© 2018 Threshold Systems Inc.

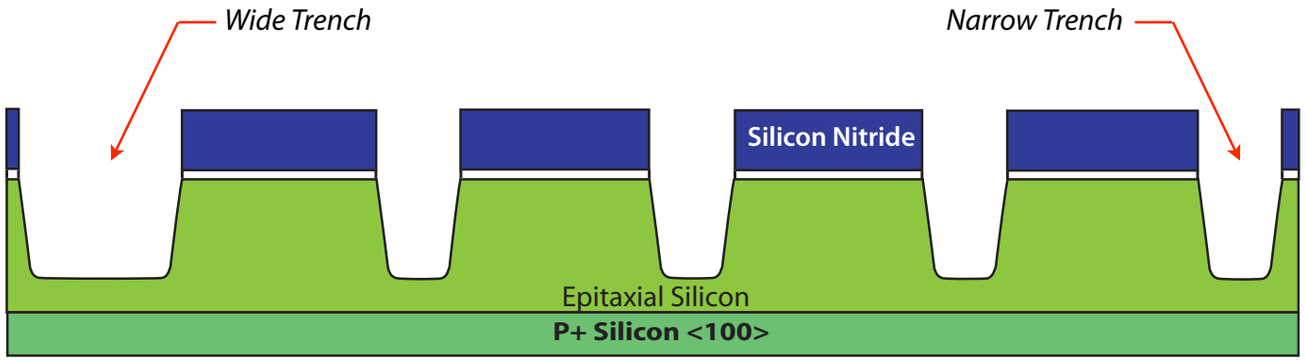


Using the amorphous carbon layer as a hard mask, an anisotropic etch cuts down through the silicon nitride and the pad oxide and into the silicon forming a shallow trench. Once the silicon surface is reached the etcher pressure and gases are changed to make the etch process less anisotropic. This results in trench side-walls that are sloped at an angle from 5 to 15 degrees.

The amorphous carbon hard mask has a selectivity of 20:1. It is also transparent to the 633 nm light used for alignment purposes, & therefore does not interfere with this critical operation.

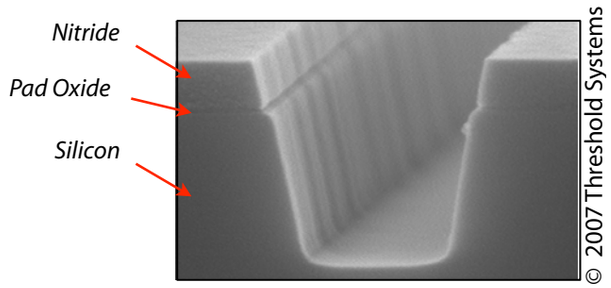
© 2018 by Threshold Systems Inc. All Rights Reserved.

Trench Etch:

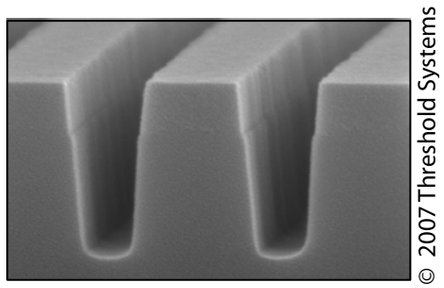


© 2016 Threshold Systems Inc.

The amorphous carbon hard mask is selectively stripped away and the wafers are cleaned in Piranha.



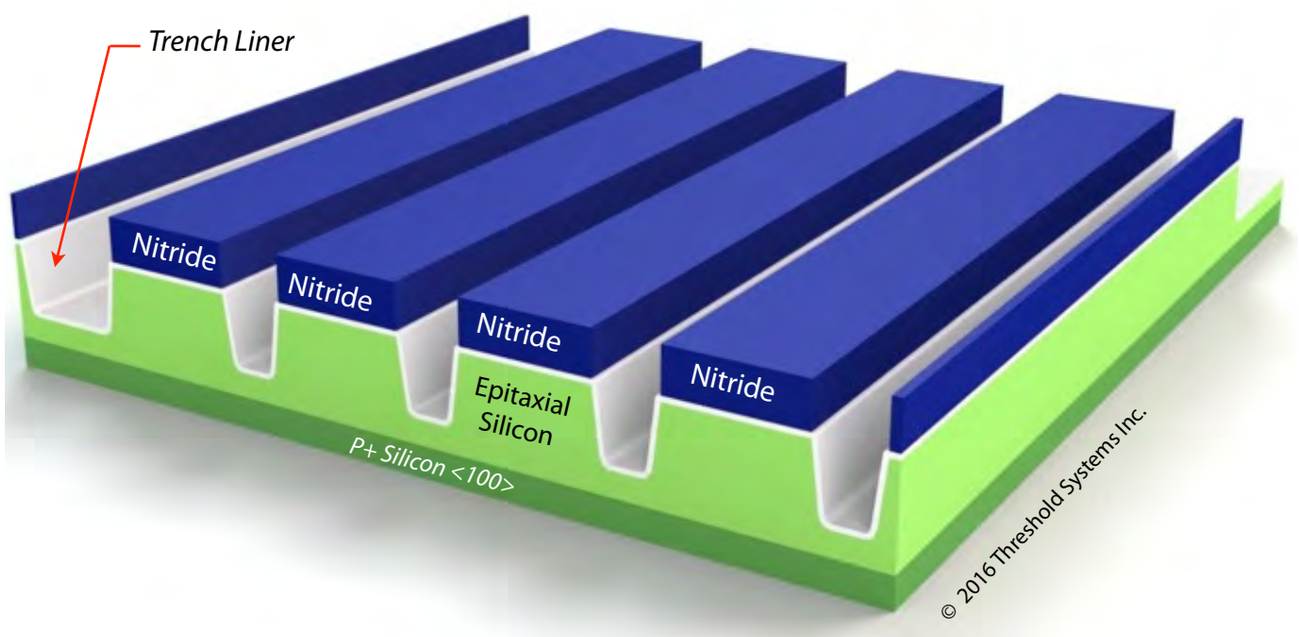
A wide Shallow Trench



Narrow Shallow Trenches

It is important to realize that not all trenches are of the same size. Since any region on the chip that is NOT an active area (I.e. a Well) will most likely be Shallow Trench Isolation (STI), the size and shape of the trenches will vary widely across the chip.

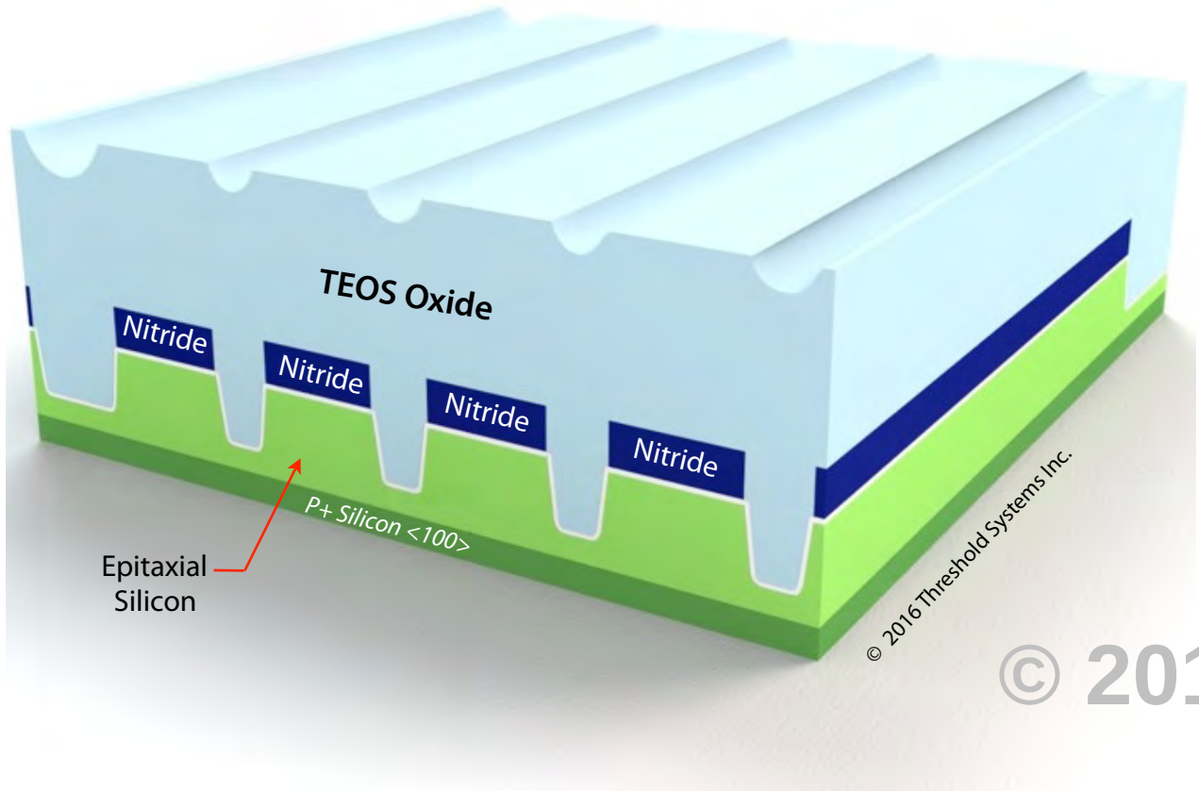
© 2018 Threshold Systems Inc.



The wafers are again cleaned (P/HF/SC1/SC2) and then a 100 Å thick layer of silicon dioxide is grown in the trenches. This oxide layer is known as a "trench liner" and is helps to relieve the stress in the silicon around the upper and lower corners of the trench. The silicon nitride acts as a diffusion barrier and prevents any silicon dioxide from growing beneath it.

© 2018 by Threshold Systems Inc. All Rights Reserved.

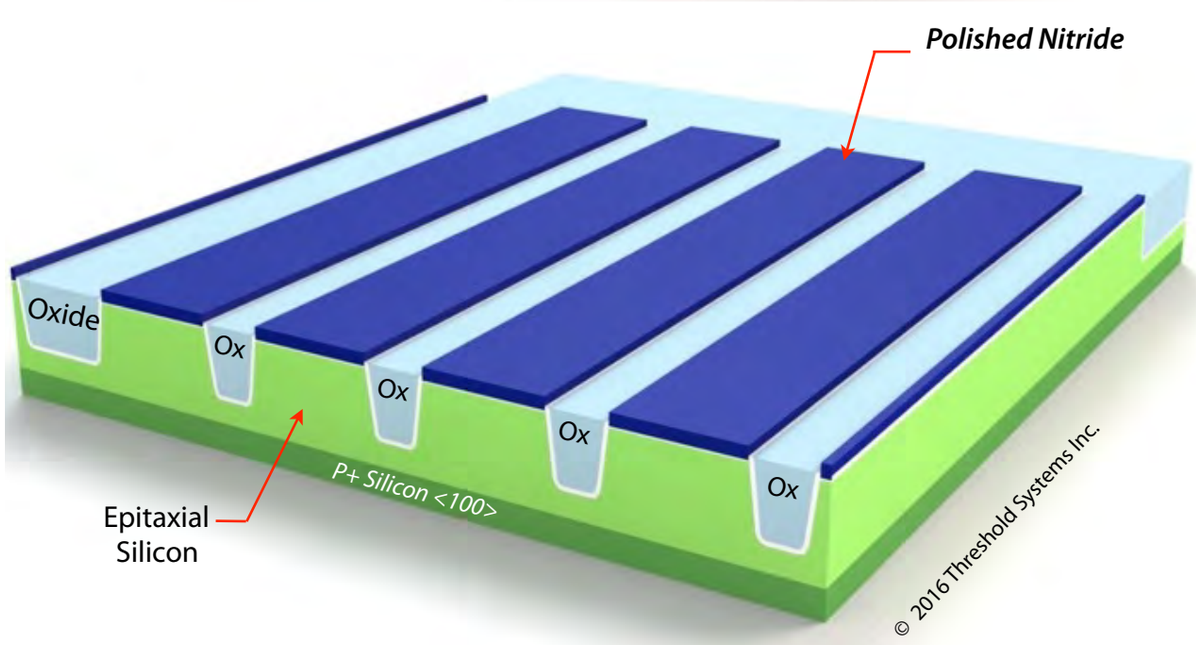
Trench Fill & Oxide Polish:



The surface of the oxide is then cleaned in P/SC1/SC2 and 4,000 Å of TEOS (Tetra-Ethyl-Ortho-Silicate) oxide is deposited. TEOS is an oxide deposited by CVD, as opposed to a grown (thermal) oxide. Its key advantage is that it allows thick layers of oxide to be deposited quickly. It main disadvantage is that a deposited TEOS oxide is not as dense as a grown oxide.

The wafers are then heated at 1,000 ˚C for 20 minutes to densify the TEOS and make it more resistant to wet etches.

© 2018 Threshold Systems Inc.



The wafers are polished back using CMP. The very hard silicon nitride layer acts as a CMP stop. After the polish step the wafers are cleaned in P/SC1.

After the oxide has been polished, the thin layer of nitride on the surface that acted as a CMP stop must be removed.

© 2018 by Threshold Systems Inc. All Rights Reserved.

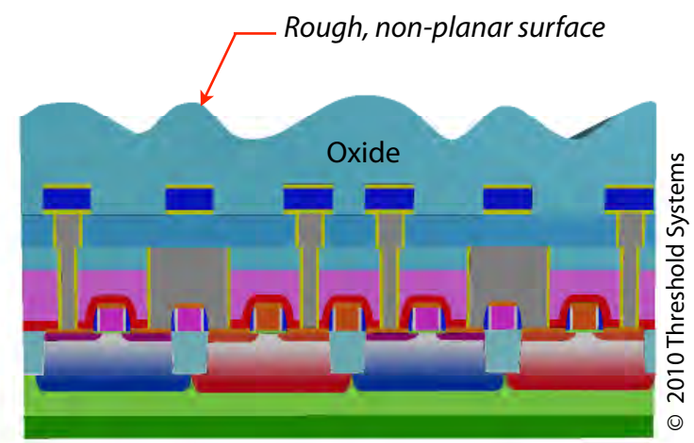
Chemical & Mechanical Polishing(CMP):

ThresholdSystems

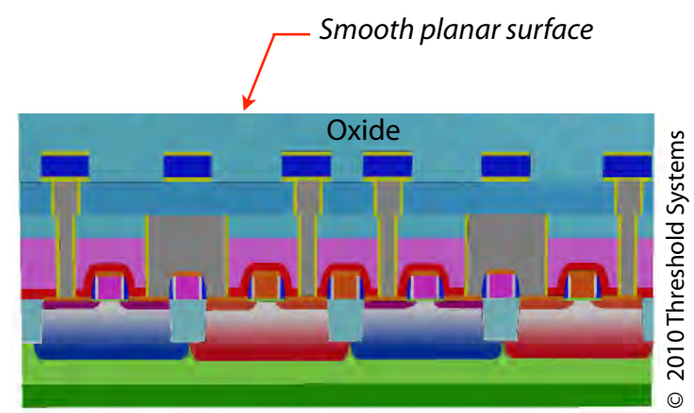
- *There are a number of operations throughout the semiconductor manufacturing process where the surface of the wafers needs to be smoothed or "planarized."*
- *Planarization is required because the many oxide and Copper deposition operations create a rough upper surface on the wafer which is referred to as "topographical variation".*
- *Because microchip manufacturing consists of depositing and patterning a series of structures all stacked on top of each other (known as the "planar" process) it is important that each layer be polished after it is deposited so that its upper surface is smooth and flat.*
- *Chemical Mechanical Planarization (CMP) smoothes out the rough topography of the as-deposited oxide or Copper films by polishing away the film's high spots and producing a smooth planar surface.*
- *CMP is used several times during the semiconductor manufacturing process and is a key enabling technology without which advanced IC's could not be produced.*

© 2018 Threshold Systems Inc.

CMP Planarization



Cross-sectional close-up of a tiny portion of a microchip after an oxide deposition. Note the undulating upper surface.



The same cross-sectional close-up of microchip after it has been polished by CMP. Note how smooth the upper surface is after this polishing.

© 2010 Threshold Systems

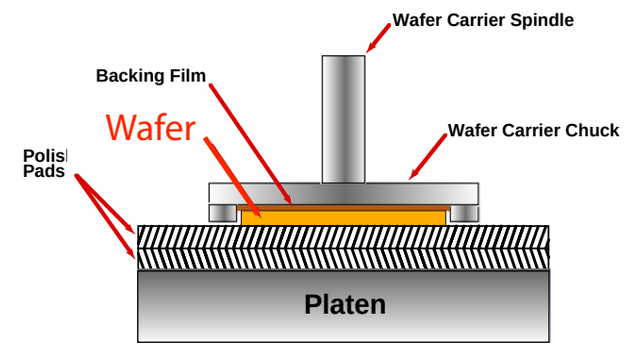
© 2018 by Threshold Systems Inc. All Rights Reserved.

CMP Technology:

- *CMP works by polishing the wafer in an abrasive liquid slurry while applying a small downward force.*
- *The wafer is loaded into a wafer carrier face-down. The wafer carrier is then pressed against a rotating pad called a platen.*
- *Both the platen and the wafer carrier rotate while a liquid abrasive known as a slurry is dispensed onto the platen by a slurry applicator.*
- *As the wafer rotates in the carrier its upper surface is slowly and carefully polished. This results in a smooth, flat upper surface on the wafer.*
- *A device known as an end effector keeps the platen pad clean and maintains a uniform pad surface.*
- *After being polished the wafer is rinsed and cleaned to remove all of the slurry.*

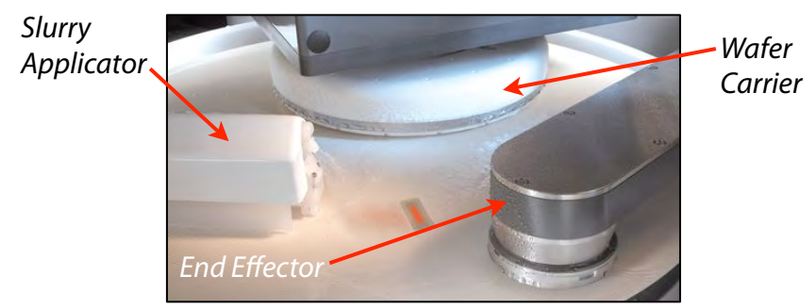
© 2018 Threshold Systems Inc.

CMP Configuration



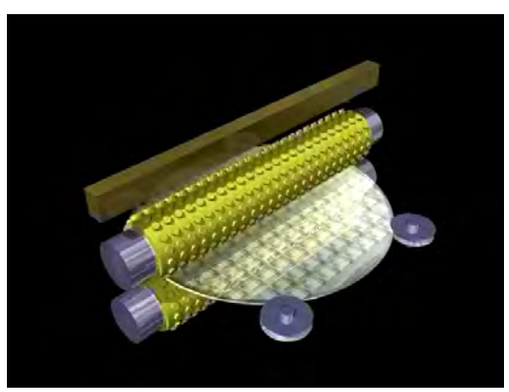
© 2007 Threshold Systems

The wafer is placed face-down in the carrier and pressed against the polishing pads



© Applied Materials

The central components of a CMP System

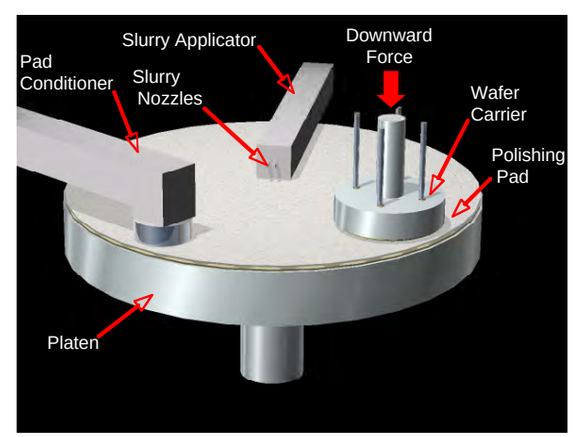


© 2007 Threshold Systems

The wafers are cleaned after being polished.

© 2007 Threshold Systems

A top-down view of the polishing table.



© 2007 Threshold Systems

A side of the polishing table.

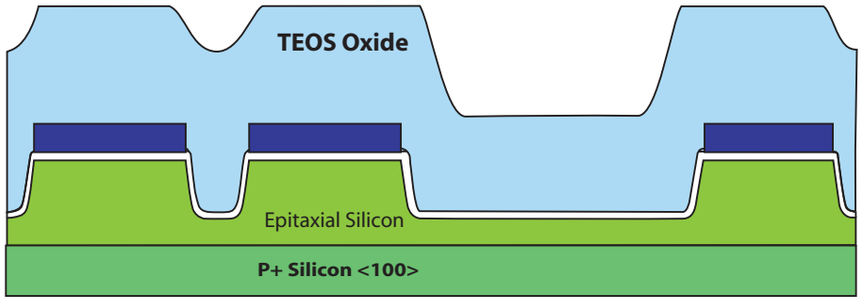
© 2018 by Threshold Systems Inc. All Rights Reserved.

Tiling Technique:

ThresholdSystems

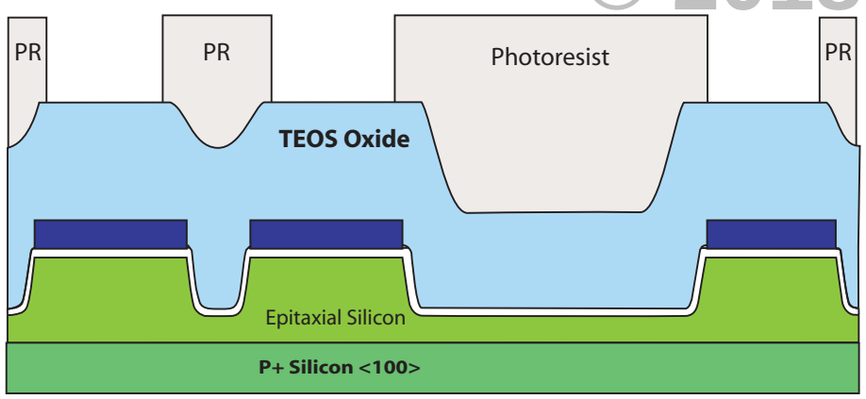
Method A

1.



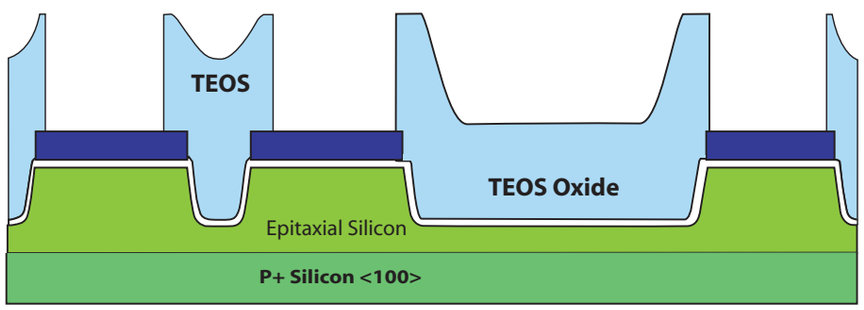
© 2009 Threshold Systems

2.



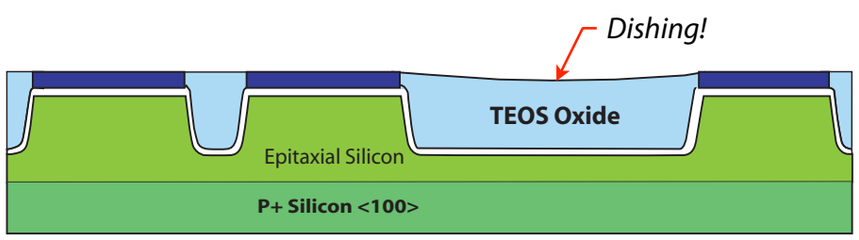
© 2009 Threshold Systems

3.



© 2009 Threshold Systems

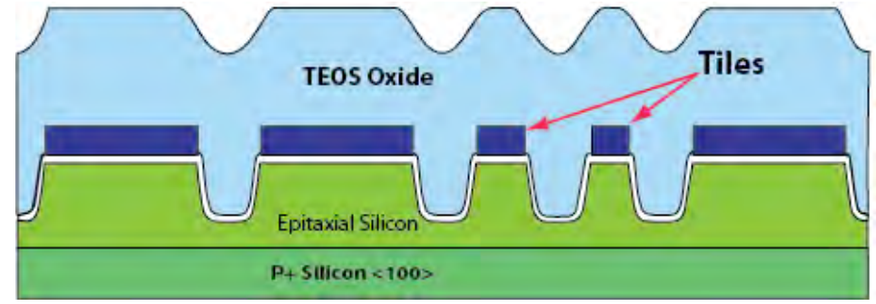
4.



© 2009 Threshold Systems

Method B

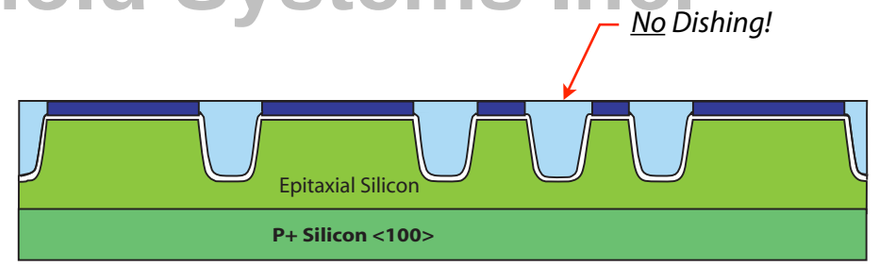
5.



© 2009 Threshold Systems

© 2018 Threshold Systems Inc.

6.



© 2009 Threshold Systems

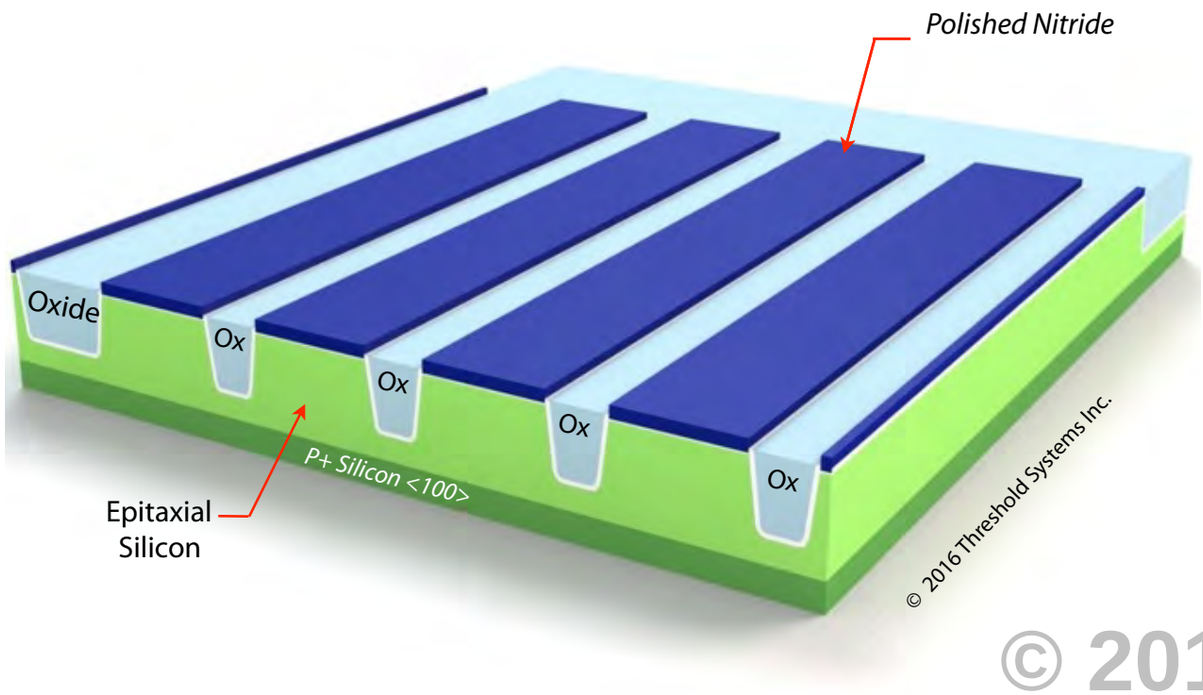
Not all trenches are uniform in size. Some trenches are very large and this poses special problems for CMP planarity. Illustration #1 displays a very large trench that creates a large variation in the TEOS oxide density across the surface of the wafer. This is highly undesirable for CMP.

In order to avoid this problem the trenches and their adjacent area are patterned with photoresist (figure #2) and the oxide over the active areas is etched away (figure #3). This produces an oxide density that is somewhat more uniform and which should polish more evenly. However the very large open trenches can still end up being somewhat dished after the CMP operation is completed (figure #4).

An alternate approach is illustrated in Method B wherein "tiles" are designed into the middle of the large open structures (figure #5). These tiles are dummy structures that perform no function other than to present a uniform oxide surface to the CMP polish platen. This eliminates the dishing effect (figure #6) and also eliminates the photo and etch operations employed in the method A.

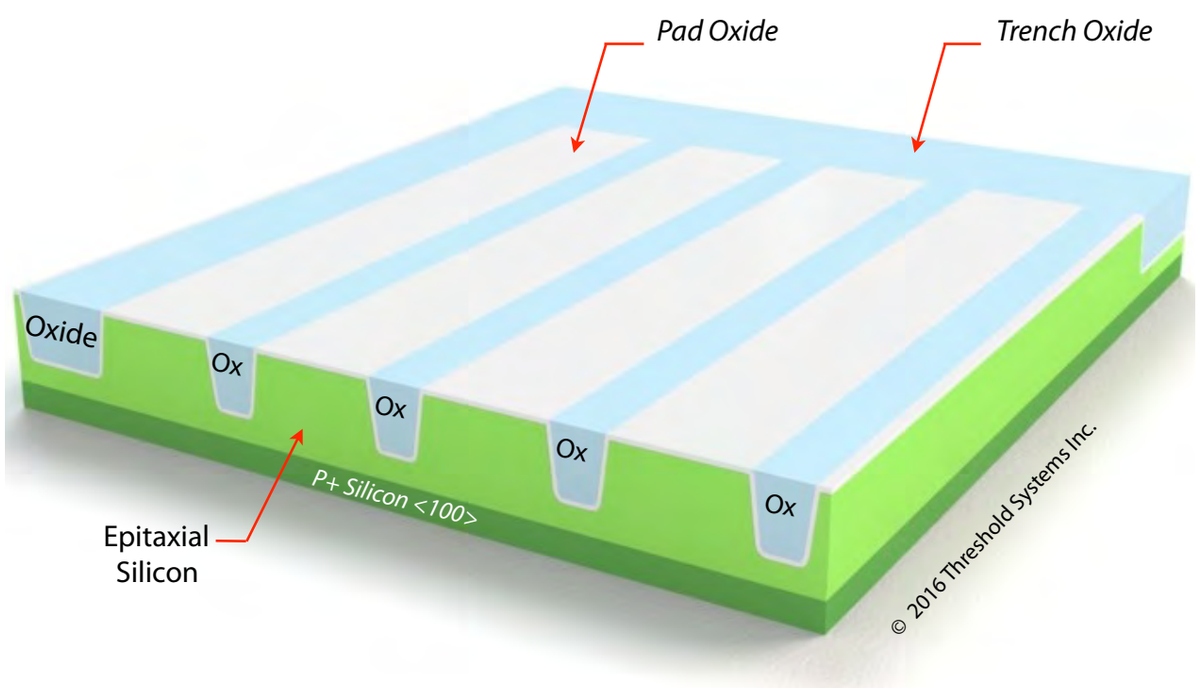
© 2018 by Threshold Systems Inc. All Rights Reserved.

Nitride Strip:



After the oxide has been polished back the thin layer of nitride on the surface that acted as a CMP stop must be removed.

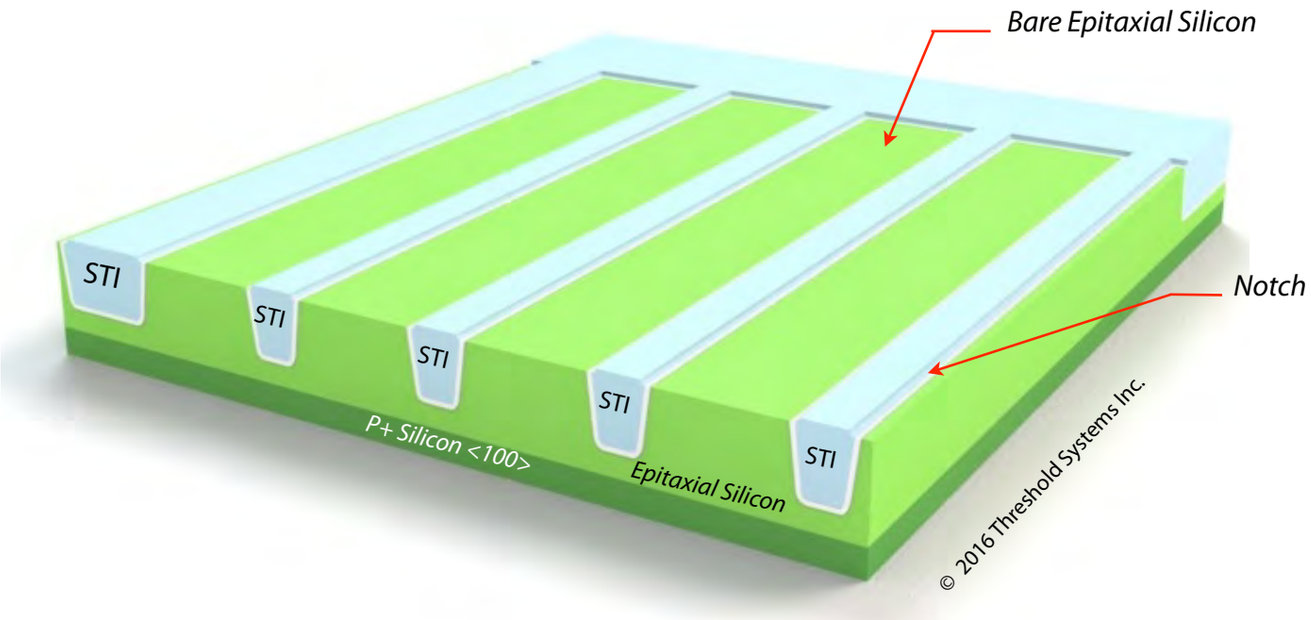
© 2018 Threshold Systems Inc.



The residual silicon nitride is removed by immersing the wafers in hot (140 ºC) Phosphoric acid for 20 minutes. The Phosphoric acid has a high selectivity to nitride over oxide and will not attack the oxide.

© 2018 by Threshold Systems Inc. All Rights Reserved.

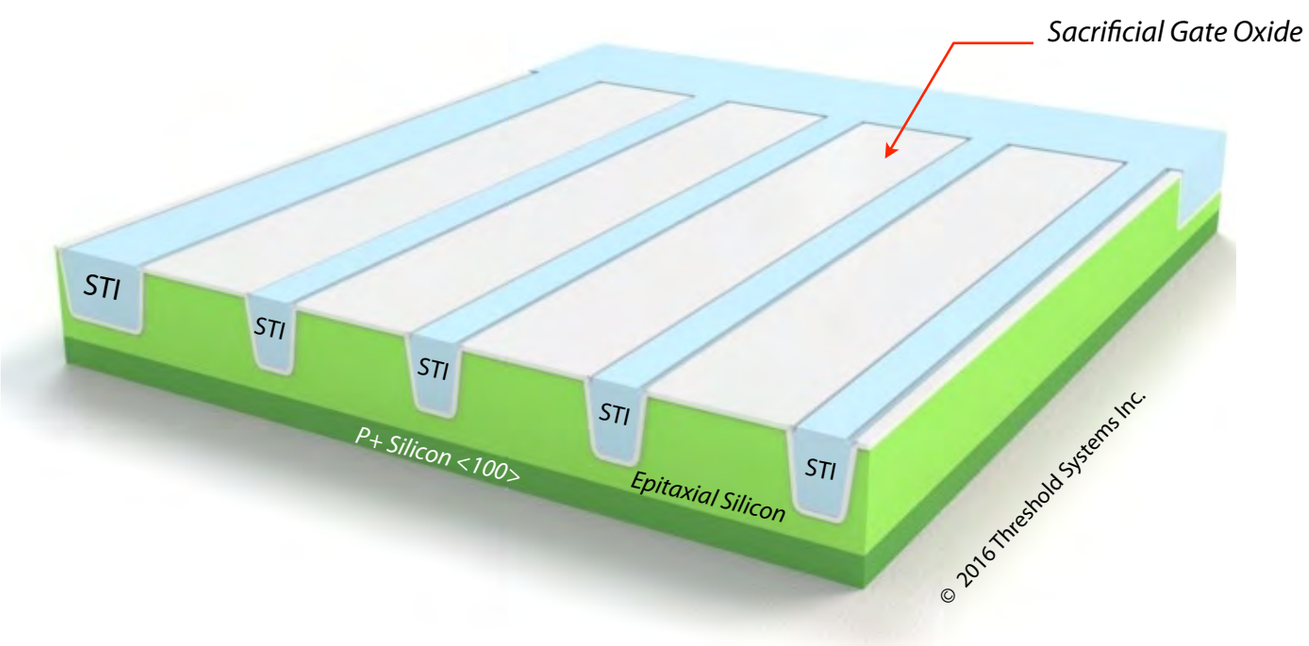
Pad Oxide Strip & Sacrificial Gate Growth:



Next, the Pad Oxide is stripped by immersing the wafers in Hydrofluoric acid.

The Pad Oxide strip unintentionally removes oxide preferentially from the trench corners because these areas are under stress and etch faster. This results in the creation of "notches" along the sides of the trench that can cause problems later in the process.

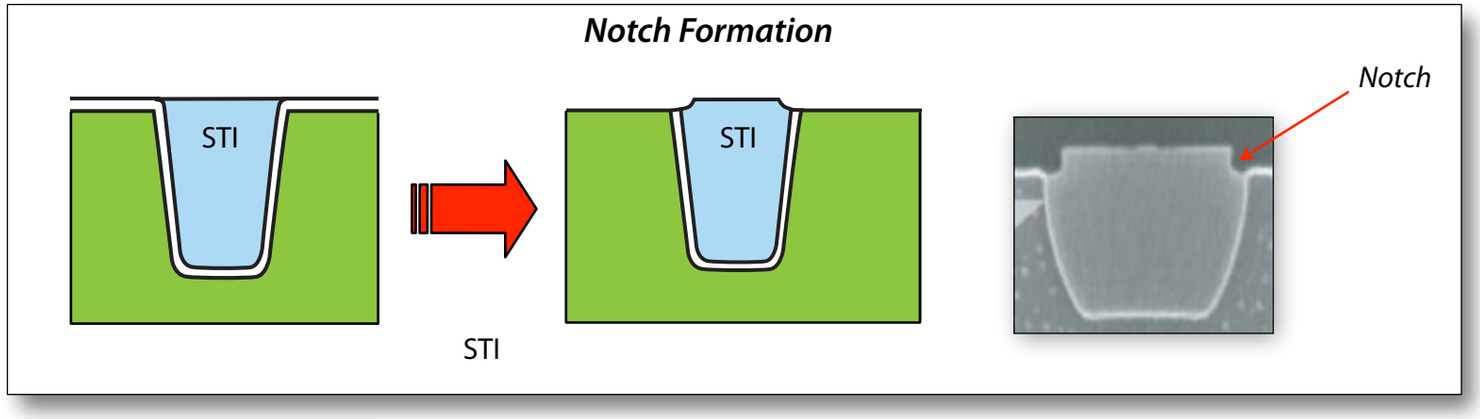
© 2018 Threshold Systems Inc.



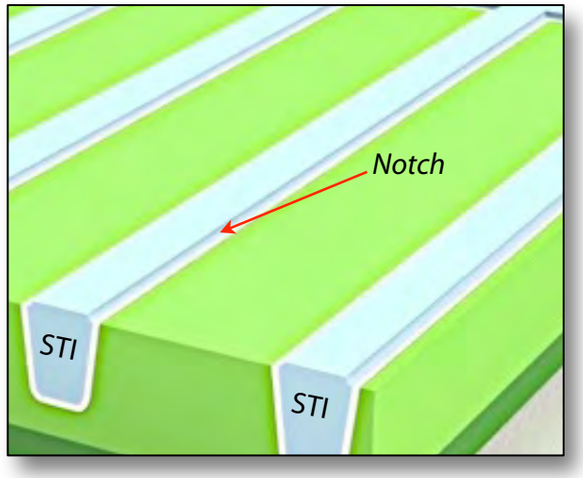
The wafers are cleaned in P/SC1/SC2 and a 50 Å sacrificial gate oxide (Sac Ox) is grown. The sacrificial gate oxide also acts as a screen oxide.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Notch Formation:

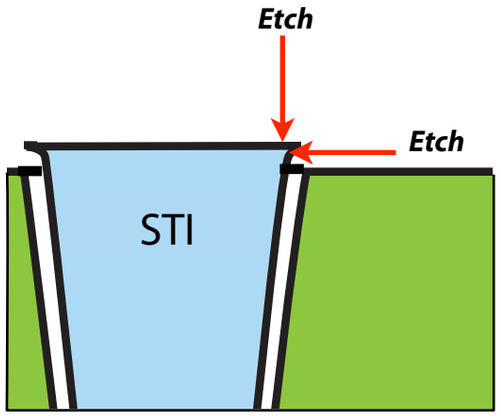


© 2016 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

© 2018 Threshold Systems Inc.



The Pad Oxide strip unintentionally removes oxide preferentially from the trench corners primarily because the STI oxide is a deposited oxide which is much softer than the pad oxide which is a grown thermal oxide. In addition, the corner areas of the STI oxide are under mechanical stress, and microstructures that under stress etch faster.

Finally, the sharp corners of the STI oxide are attacked from two sides during the Pad Oxide etch process and this accelerates the retreat of these corners.

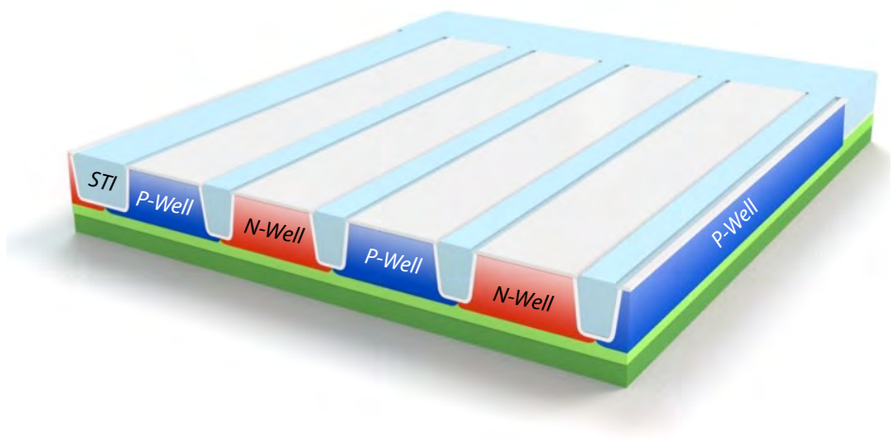
This preferential etching results in the creation of "notches" along the sides of the trench that can cause problems later in the process.

© 2016 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Well Definition Module:



© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

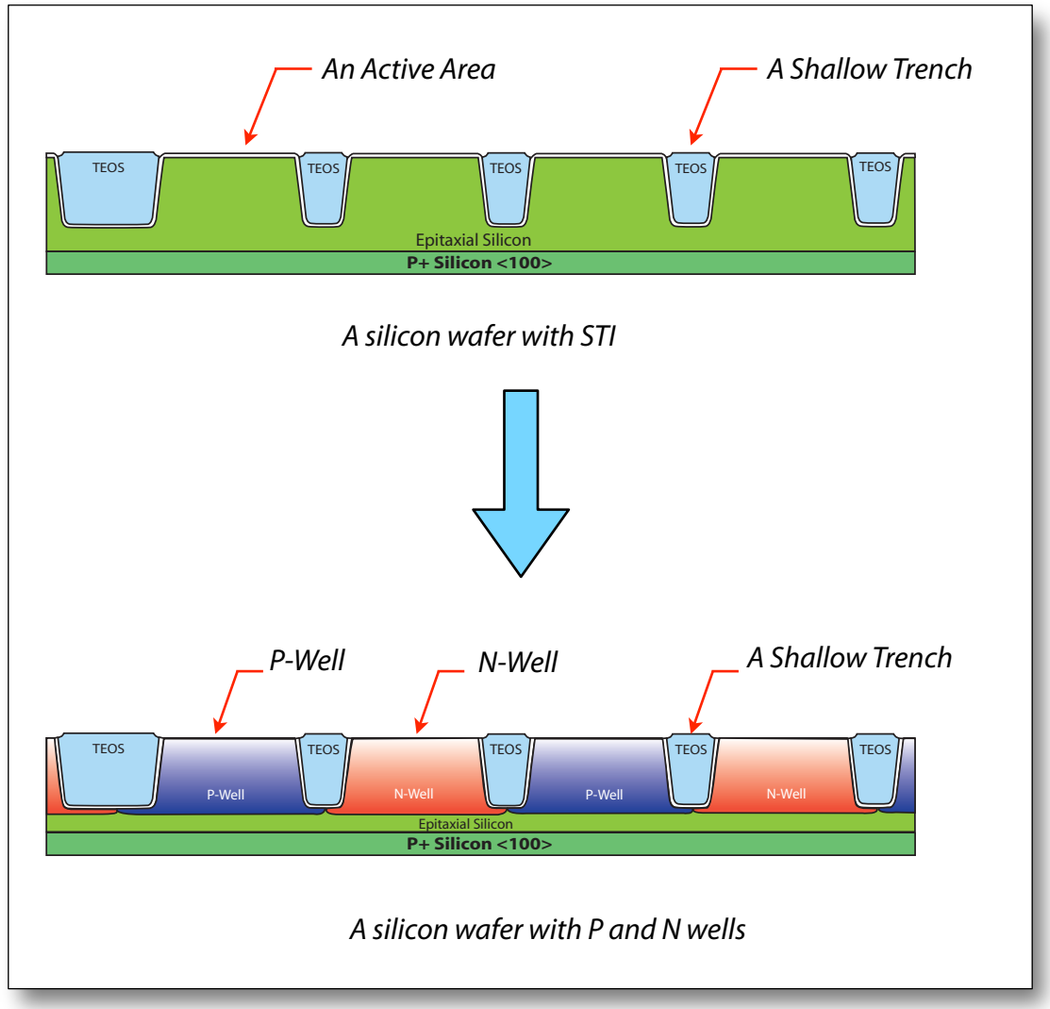
The Well Module:

ThresholdSystems

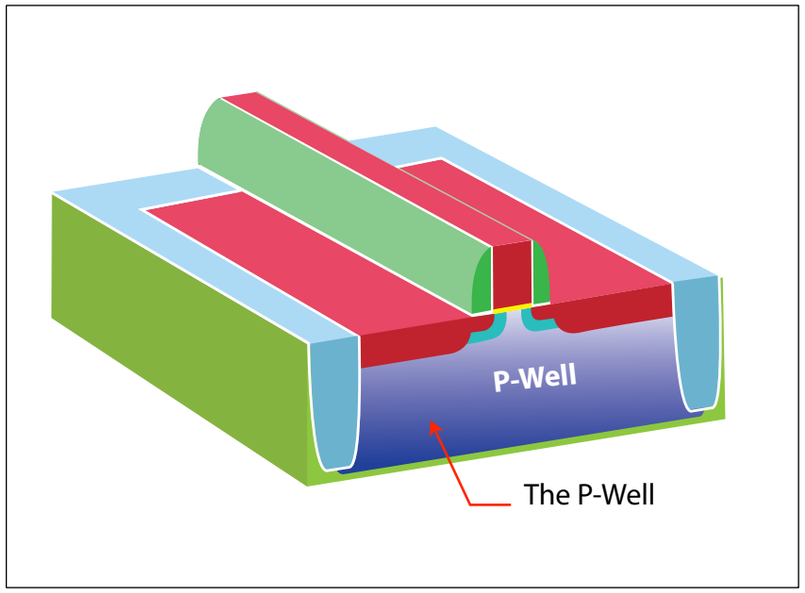
The Well Module:

- *A Well is a localized region of dopant that has been implanted into the silicon surface in order to tailor its electrical characteristics and optimize transistor performance. The Wells also augment the electrical isolation provided by the Shallow Trench Isolation by establishing diode barriers between adjacent PMOS and NMOS devices.*
- *In the well fabrication module the P-Wells and the N-Wells required for the NMOS and PMOS transistors are realized using a series of masking and ion implantation steps.*

© 2018 Threshold Systems Inc.



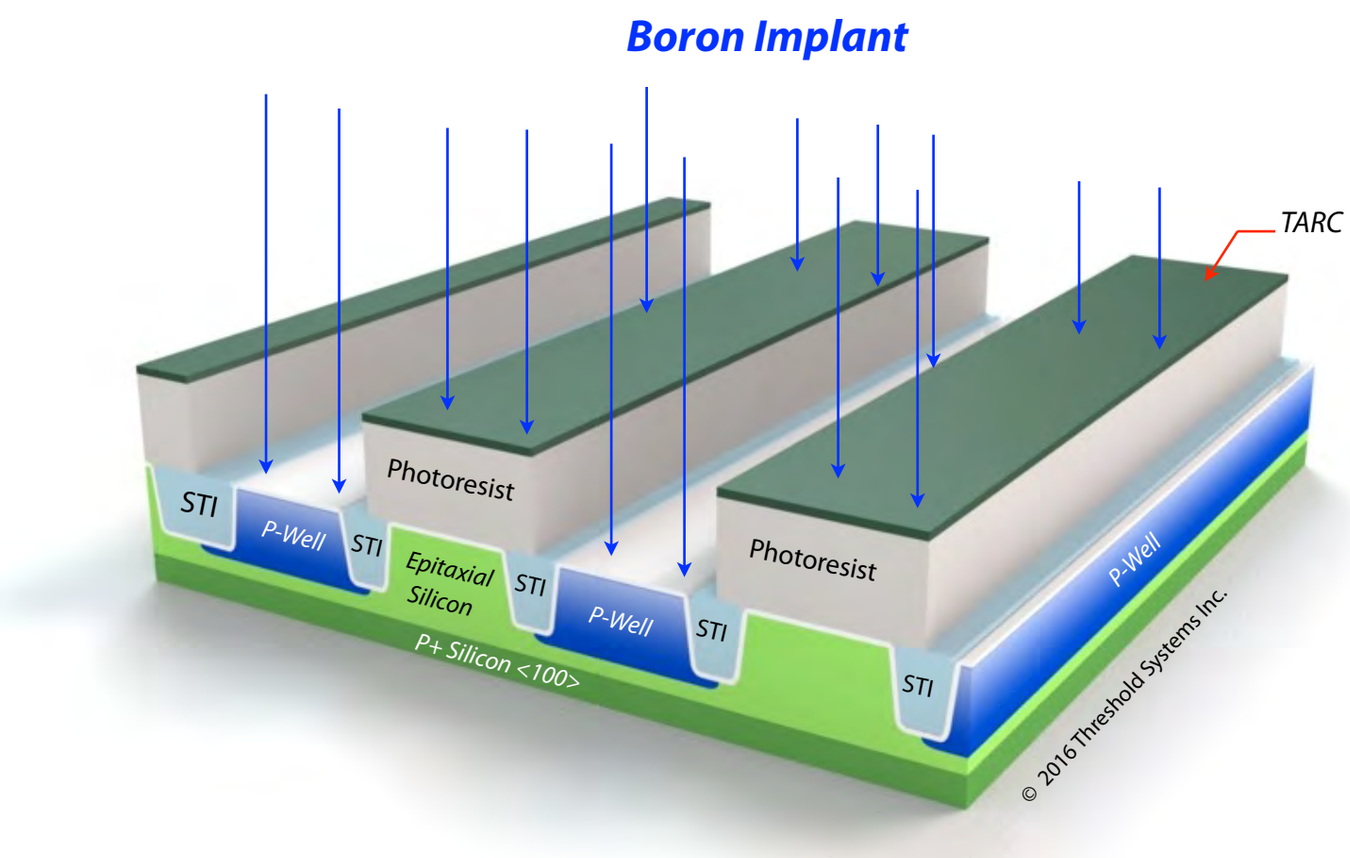
© 2016 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

P-Well Implant:

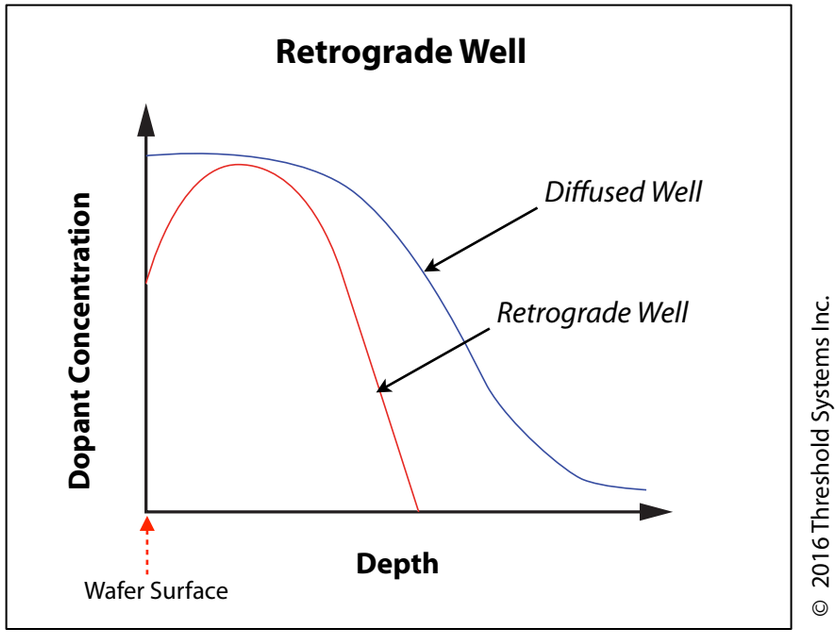


© 2018 Threshold Systems Inc.

The wafers are primed in HMDS and then coated with TARC (Top Anti-Reflective Coating) and photoresist. The photoresist is soft baked to give it a measure of structural integrity. After exposure and development the photoresist has a pattern that covers the region where the N-Wells will be located and which exposes the regions of the wafers where the P-Wells are located.

Next, the wafers are implanted with a chain implant of Boron that establishes the retrograde well P-Well (B, 5E13@200KeV, 7°), a mid-well Field Channel Stop (B, 5E12@50KeV, 7°), and finally a Threshold Voltage adjust (B, 5E11@5KeV 7°).

The P-Well established by the well implant is "Retrograde" in nature. This means that the highest concentration of dopant is not located on the surface of the Well, but below the surface as indicated in the graph below. This arrangement permits high carrier mobility near the surface of the silicon where the transistor channel is located.



© 2018 by Threshold Systems Inc. All Rights Reserved.

Ion Implantation:

- *Ion Implantation is the process of accelerating atoms to a very high energy and then implanting a very exact number of them into the surface of a silicon wafer at a precise depth. These atoms are known as "dopant" atoms.*
- *The kinds of atoms that are implanted are generally Boron, Phosphorous or Arsenic.*
- *These atoms alter the electrical characteristics the silicon and create localized regions that have specific electrical characteristics (like Sources, Drains or Wells) required to form transistors and other electrical structures.*
- *The big advantage of doping with an implanter is that is provides enormous control over the depth, quantity and purity of the dopant atoms that are implanted.*
- *Implants are specified by four parameters: Dopant species, dose, energy and implant angle.*

© 2018 Threshold Systems Inc.

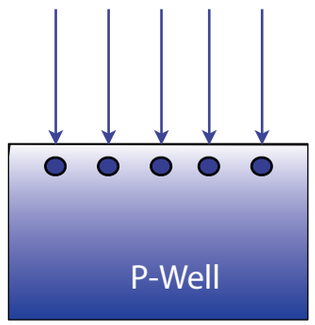
The dopant species. In this case, Boron.

The implant angle (deviation from vertical). In this case zero degrees.

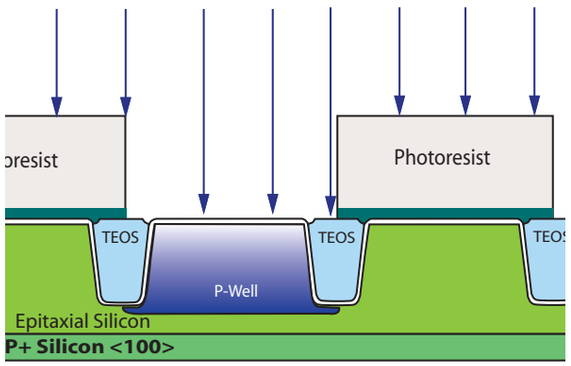
B, 3E15@ 50KeV, 0°

The dose of the implant, expressed in "ions/cm^2". In this case 3 x 10^{15} ions/cm^2.

The implant energy expressed in Kilo-electron Volts (KeV). This determines how deep the ions are implanted.



Ion Implantation is the introduction of precise quantities of atoms into the silicon surface at specific depths. These dopant atoms, once activated, alter the electrical characteristics of the silicon.



The location of where the dopant atoms are placed into the silicon can be controlled by using photoresist. Those areas covered by the photoresist are shielded from the implant. Exposed regions receive the full dose of the implant.

© 2010 Threshold Systems Inc.

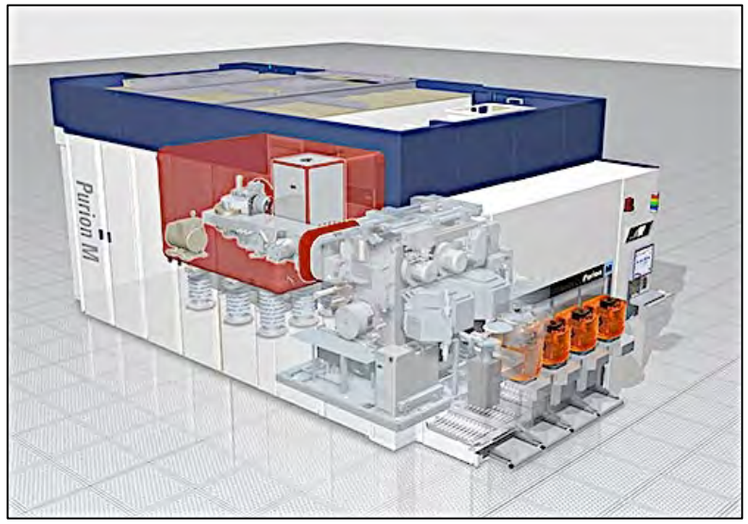
© 2018 by Threshold Systems Inc. All Rights Reserved.

Ion Implantation:

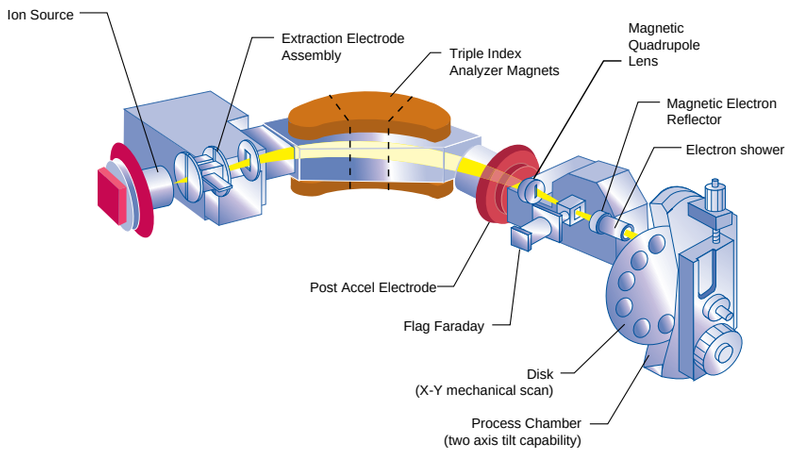
- *Ion Implanters are a special kind of particle accelerator and operate in a similar fashion to standard particle accelerators.*
- *The dopant species to be implanted is first converted into a gas and then ionized so that it has a charge and can be accelerated to a precise energy in an electric field.*
- *The accelerated dopant ions are then filtered in a mass spectrum analyzer that removes any contaminant gas atoms.*
- *The ion beam is then steered by powerful magnets and directed toward the wafer where is rastered across the wafer's surface to implant the dopant species.*
- *The exact dose of the implant is determined by "counting" the dopant atoms by precisely measuring the accumulated charge in the ion beam as it passes a specific point in the implanter.*
- *After dopant atoms have been implanted it is necessary to "activate" them by applying heat in order to move them onto substitutional sites in the silicon lattice where they become electrically active.*
- *Ion Implanters are by far the largest machines in the fab and they cost several million dollars.*

© 2018 Threshold Systems Inc.

Ion Implanters



Source: Axcelis

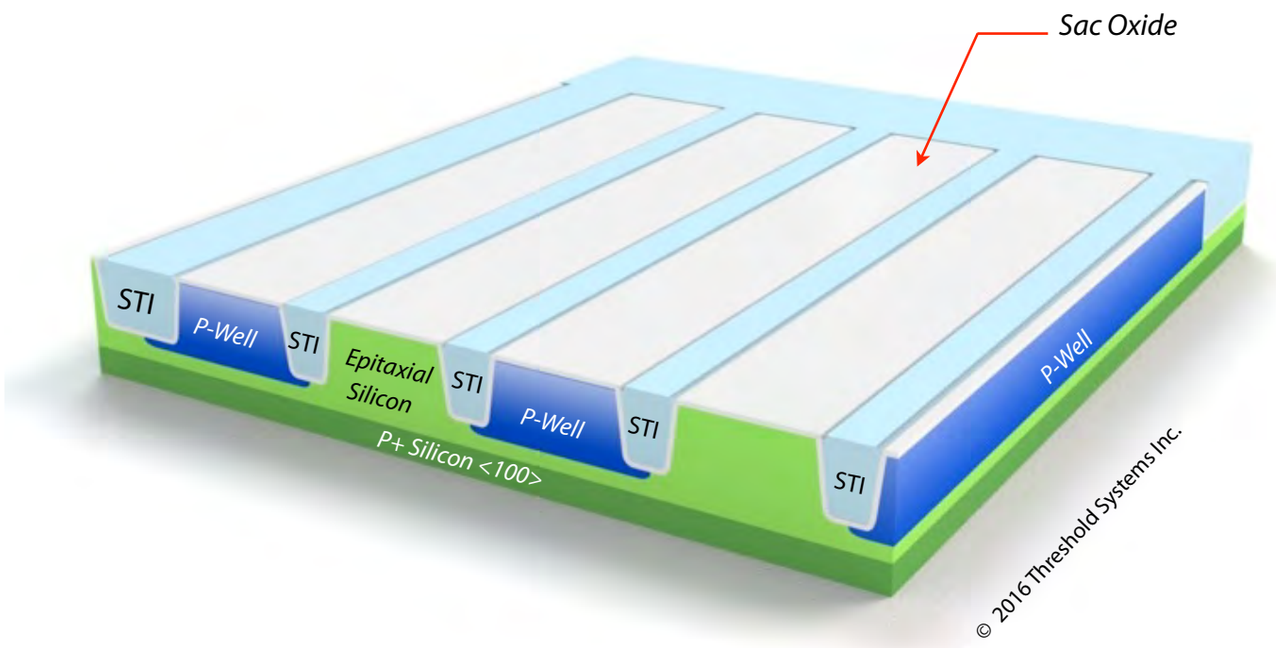


A component of an ion Implanter displaying the Ion source, analyzing magnet and associated components.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Photoresist Strip and N-Well Implant:

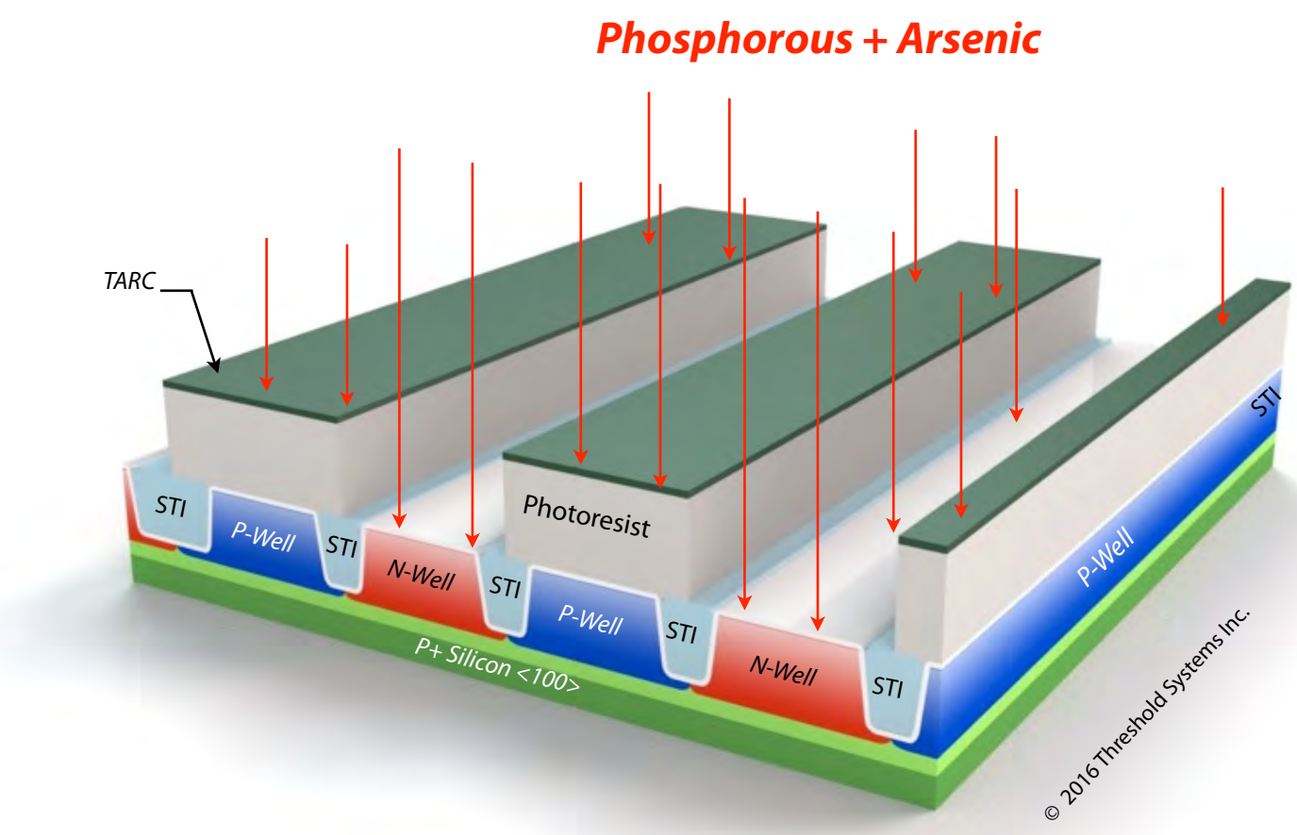
ThresholdSystems



The photoresist from the previous operation is stripped in an ionized oxygen plasma and the wafers are cleaned in Piranha.

The TEOS oxide in the STI trenches is now completely exposed as is the Sacrificial Oxide ("Sac Ox").

© 2018 Threshold Systems Inc.



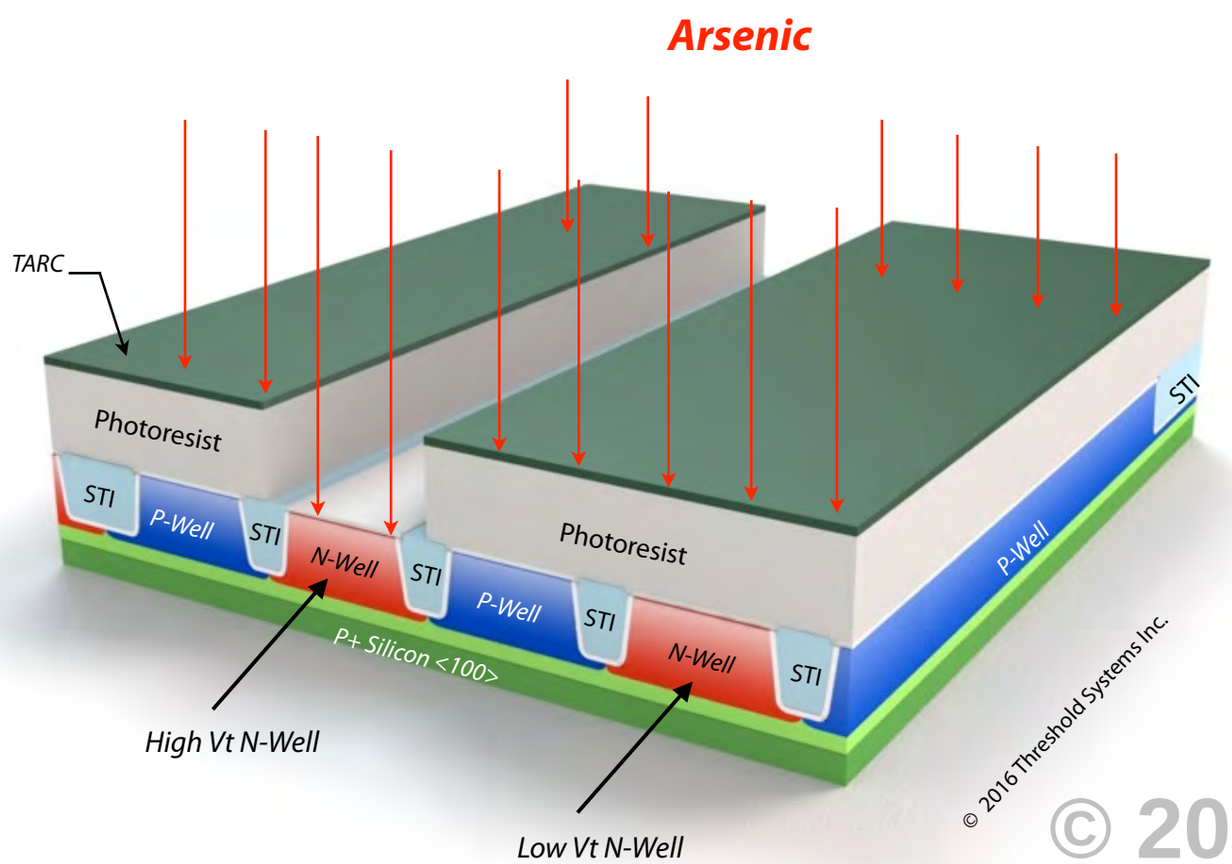
The wafers are primed in HMDS and then coated with TARC, and the photoresist is soft baked to give it a measure of structural integrity. After exposure, development, and hard-bake, the photoresist has a pattern that covers the region where the P-Wells are located and which exposes the N-Well regions.

The wafers then receive a chained implant of Phosphorus and Arsenic that implants the retrograde N-well (P, 5E13 @400KeV, 0º), a mid-well Field Channel Stop (P, 5E12@100KeV, 0º) the Threshold Voltage adjust (As, 5E11@5KeV, 0º).

© 2018 by Threshold Systems Inc. All Rights Reserved.

High Vt N-Well Vt Adjust:

ThresholdSystems

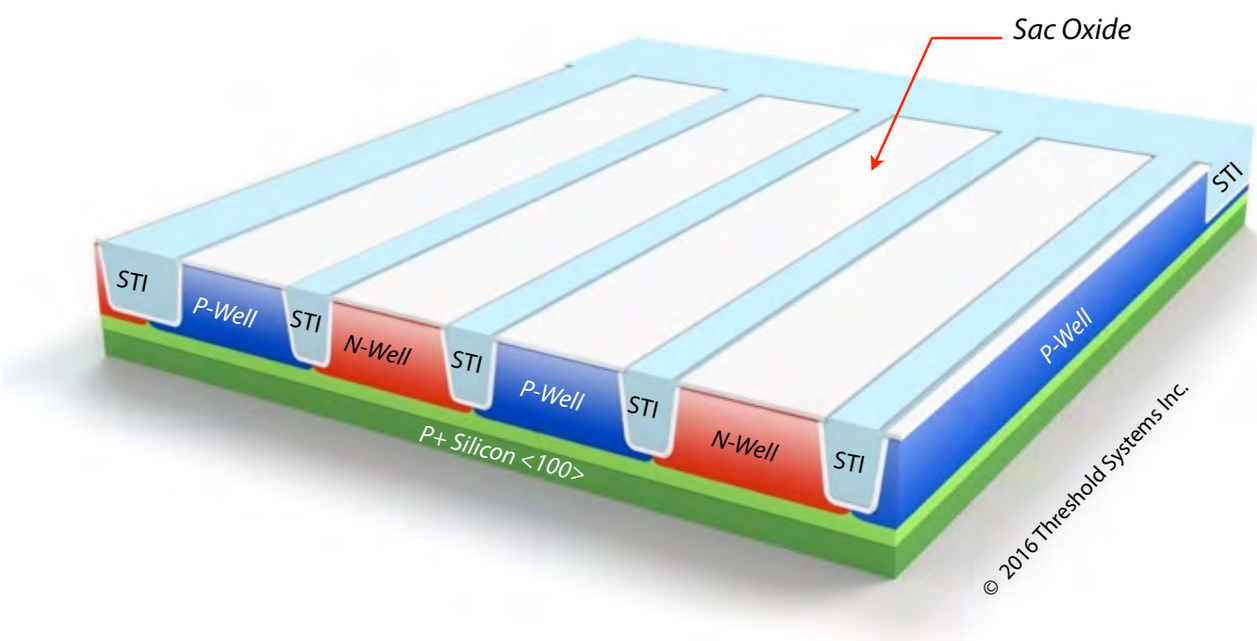


The photoresist from the previous operation is stripped in an ionized oxygen plasma and the wafers are cleaned in Piranha.

The wafers are primed with HMDS, coated in photoresist and TARC, and then patterned. After patterning that portion of the active area that is designated to be the High Vt N-Well processing section is exposed, while all other regions are covered with photoresist.

The High Voltage N-Wells receive an implant of As,3E11@5KeV, 0º to give them a higher Threshold Voltage (Vt) than the Low Voltage N-Wells.

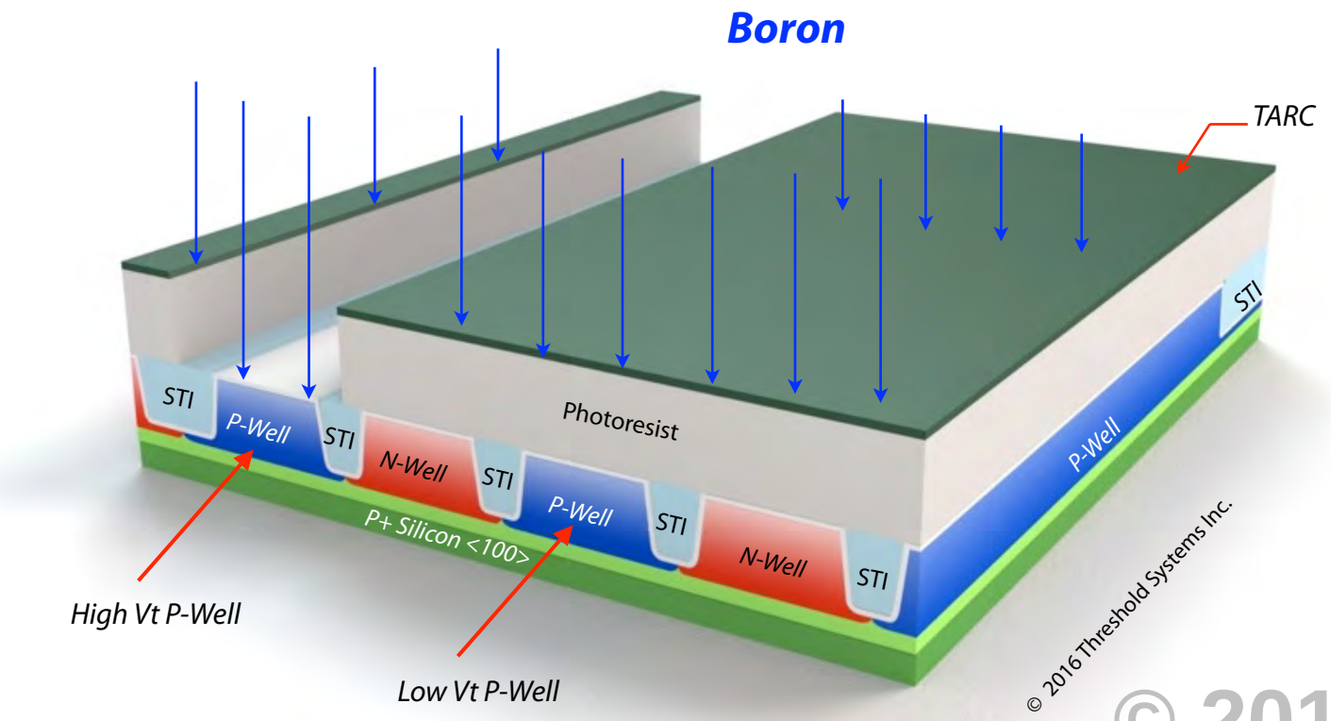
© 2018 Threshold Systems Inc.



The photoresist from the previous operation is stripped and the wafers are cleaned in Piranha to remove any residual photoresist.

© 2018 by Threshold Systems Inc. All Rights Reserved.

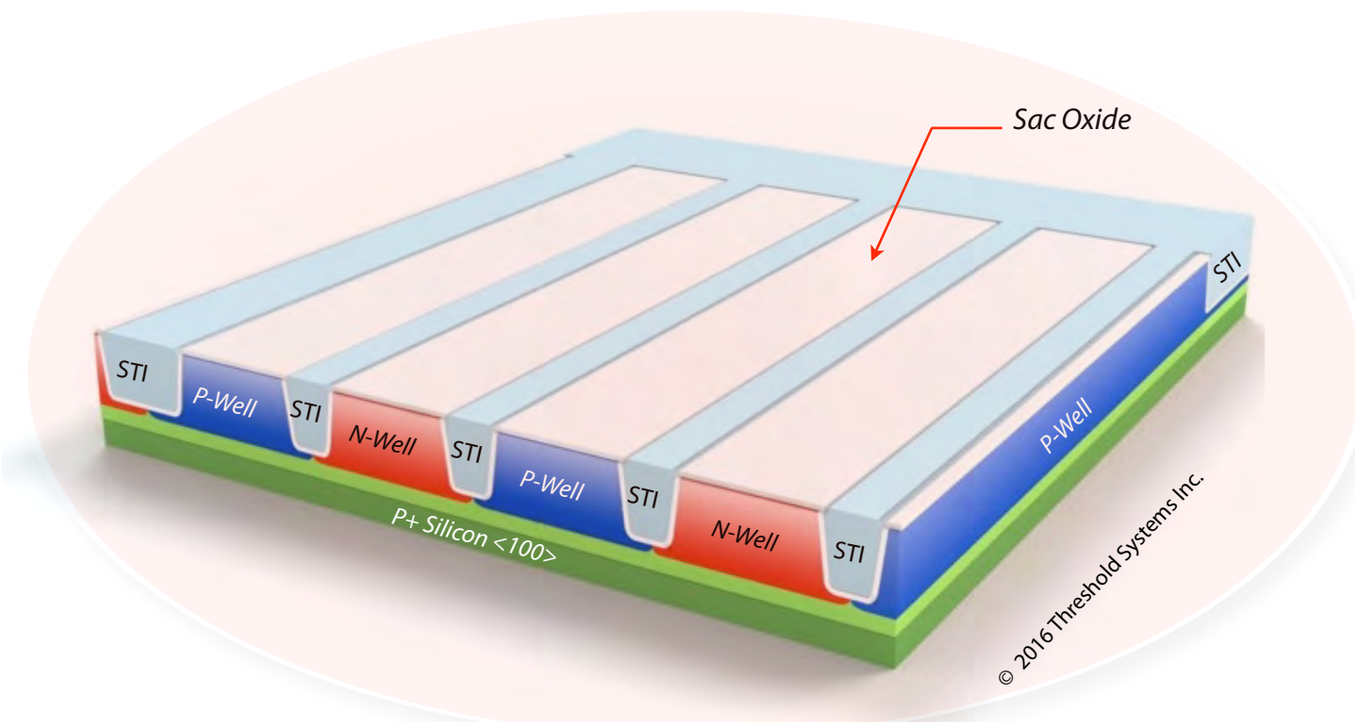
High Vt P-Well Vt Adjust & Well Anneal:



The wafers are primed with HMDS, coated in photoresist and TARC, soft baked and then patterned. After patterning that portion of the active area that is designated to be the High Vt P-Well processing section is exposed, while all other regions are covered with photoresist.

The High Vt P-Wells receive an implant of B, 3E11@5KeV, 0º to give them a higher Threshold Voltage (Vt) than the Low Voltage P-Wells.

© 2018 Threshold Systems Inc.



The photoresist from the previous operation is stripped and the wafers are cleaned in Piranha.

Following the well implants, the wafers undergo a Rapid Thermal Anneal (RTA) at 1,000 ºC for 10 seconds to activate the implanted dopants in the wells.

The anneal makes the dopant atom electrically active and anneals out the damage done to the silicon during the implants.

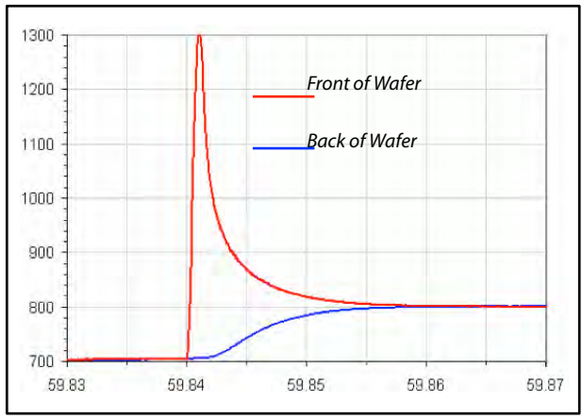
© 2018 by Threshold Systems Inc. All Rights Reserved.

Rapid Thermal Processing (RTP):

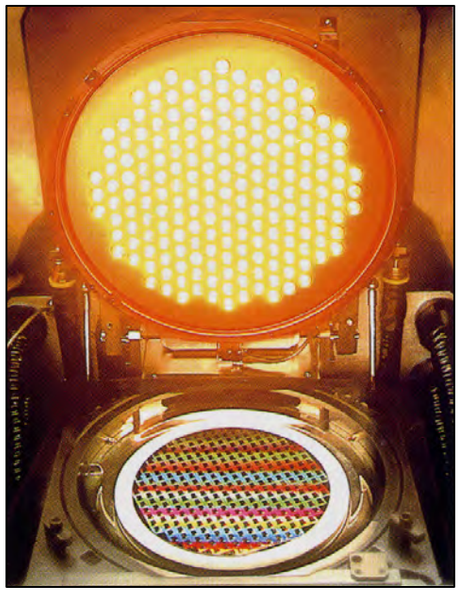
ThresholdSystems

- *Rapid Thermal Processing is a key enabling technology that is required in order to make Ion Implantation work.*
- *It is not enough to simply implant dopant atoms into a silicon wafer. The dopant atoms must be "activated".*
- *Activation is the process of causing the dopant atoms to move onto substitutional sites in the silicon lattice and become electrically active, modifying the Silicon's electrical characteristics.*
- *Activation requires that the wafers be heated to a high temperature, usually from 1,000 - 1,350°C.*
- *However, it is extremely important that the activation process not involve a lot of heat or the heat energy will cause the dopant atoms to diffuse into undesired regions of the wafer.*
- *RTP solves this problem by raising the temperature to the activation point and then rapidly dropping it back down again. It does this so rapidly that the wafers do not have time to heat up.*
- *The RTP equipment consists of an array of high-intensity lamps that heat up very rapidly and impart a short, abrupt flash of thermal energy to the front side of the wafer that does not involve a lot of heat.*

© 2018 Threshold Systems Inc.



RTP abruptly raises the temperature of the front side of the wafer for a fraction of a second without adding a lot of heat.



Traditional RTP chambers consist of a series of very bright lamps that raise the temperature of the wafer's surface for a brief instant.

© Ultra-Tech

Unlike traditional RTP tools, Laser annealing tools activate the implanted dopant while providing virtually zero heat. This is a more sophisticated approach to annealing.

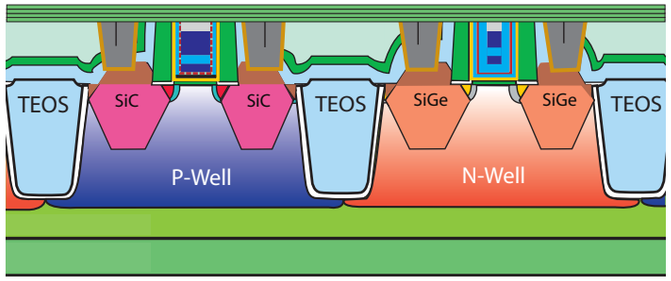
© 2018 by Threshold Systems Inc. All Rights Reserved.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Device Module Gate Last:

ThresholdSystems

Device Module Processing



© 2016 Threshold Systems Inc.

© 2018 Threshold Systems Inc.

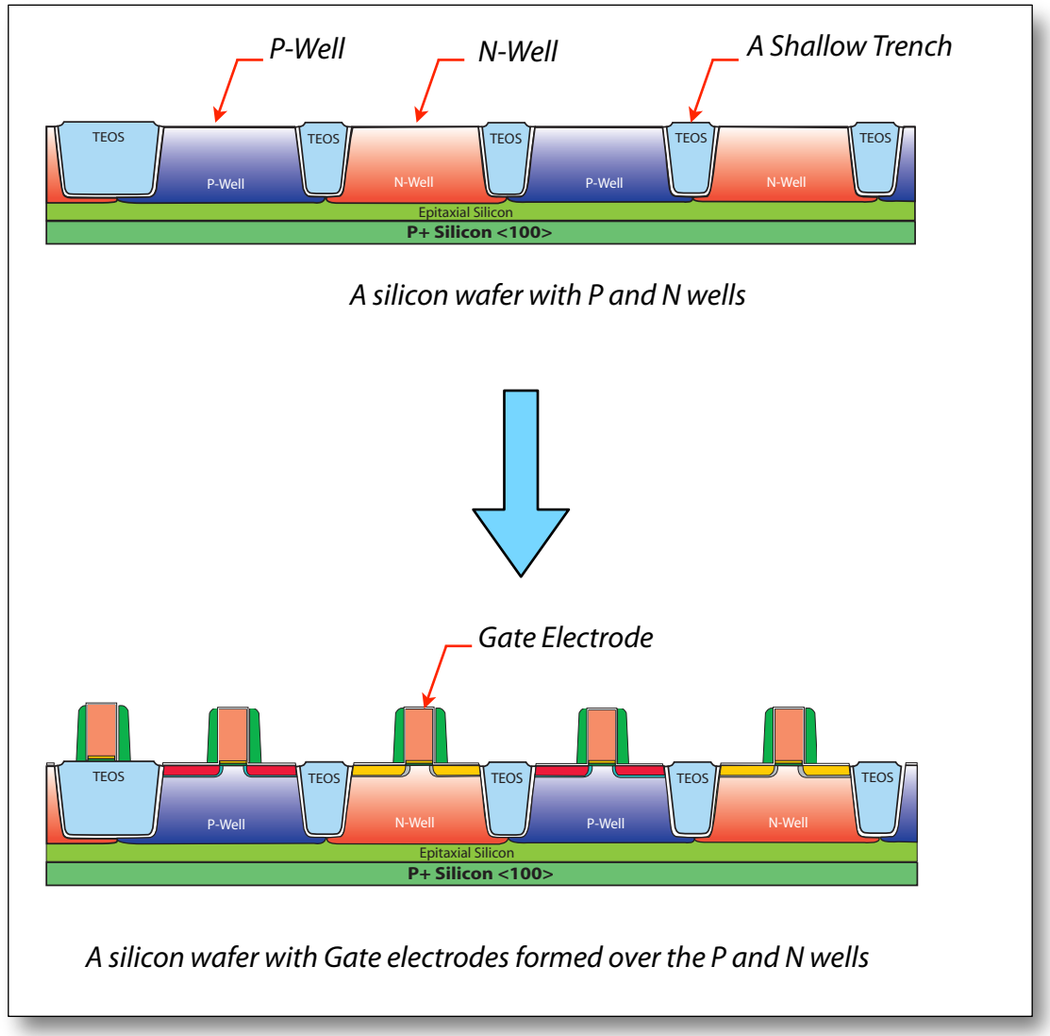
© 2018 by Threshold Systems Inc. All Rights Reserved.

The Gate Module:

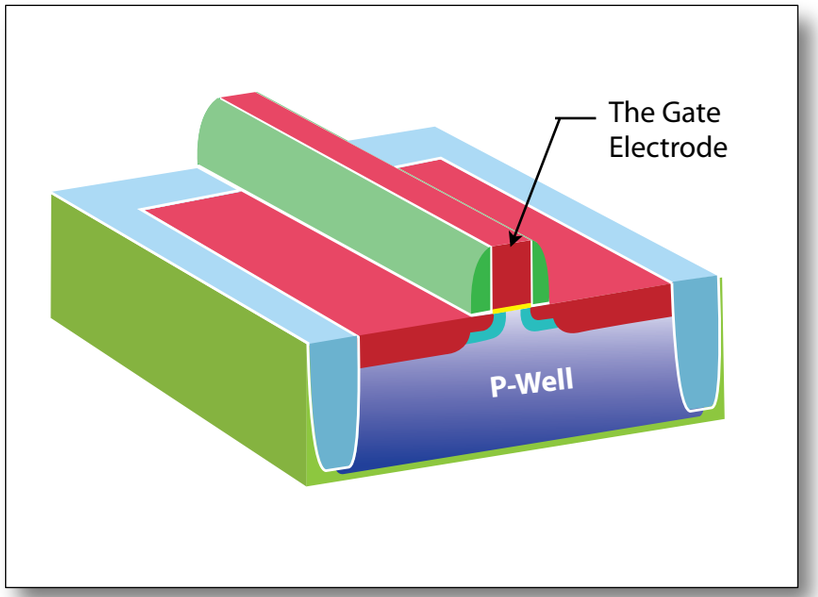
ThresholdSystems

The Gate Module:

- *In the Gate Module the transistor structures are fabricated. They consist of the Source and Drains, the Halo and Extension implants, the Nitride Spacers and the all-important Gate Electrode and gate dielectric.*
- *The Gate Electrode initially consists of a strip of doped amorphous silicon of precise dimensions. In the replacement gate methodology, the amorphous silicon is merely a place-holder that is removed and replaced with a complex series of metals.*
- *Beneath the Gate Electrode is the transistor dielectric.*



© 2016 Threshold Systems Inc.



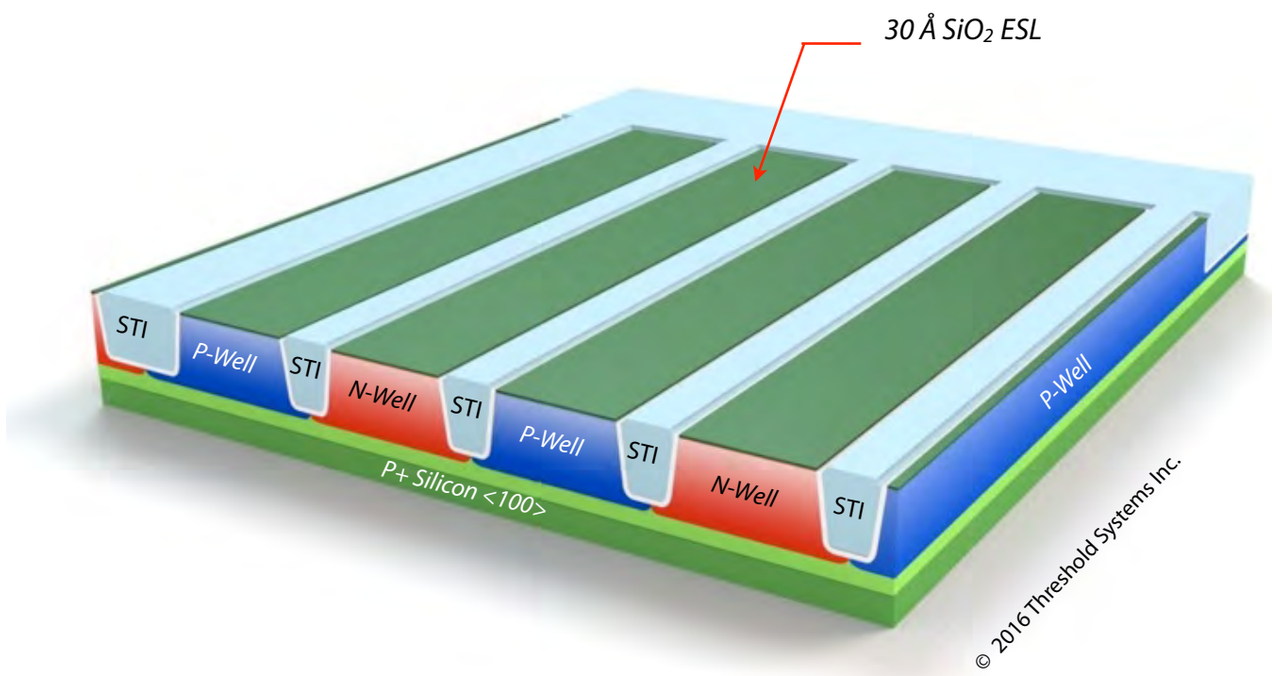
© Threshold Systems 2011

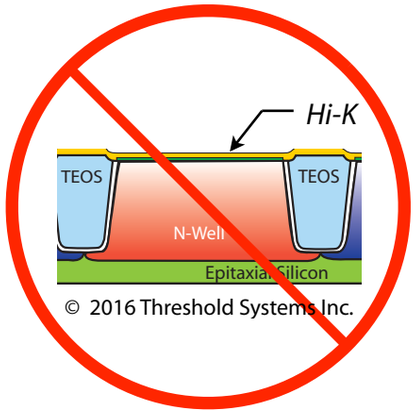
© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Oxide ESL Growth & AmSi Deposition:

ThresholdSystems

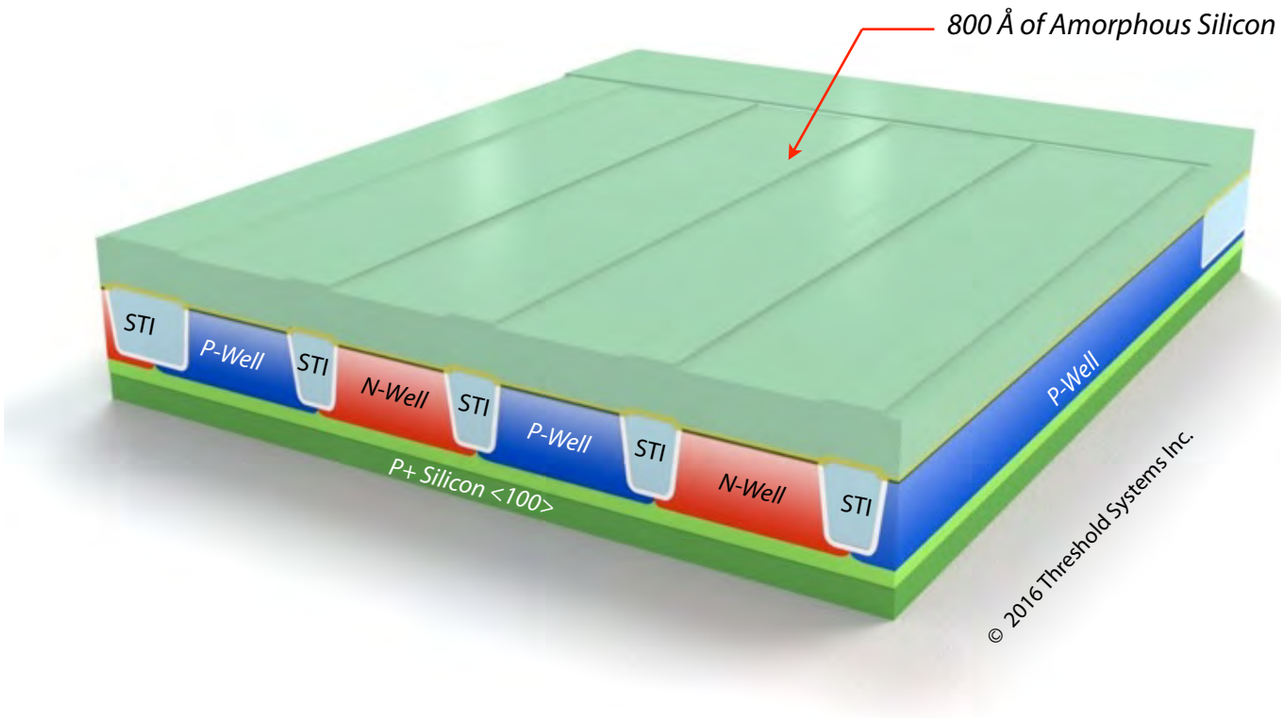


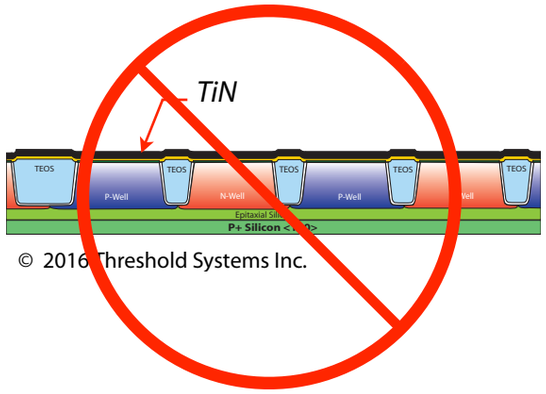


A thin 30Å Etch Stop Layer (ESL) of SiON thermal layer oxide is grown across the surface of the wafer. This oxide layer will act as an etch stop layer to protect the transistor channel when the amorphous silicon gate mandrel (not yet deposited) is etched away.

Note that there is no Hi-k dielectric deposited at this point in the process flow.

© 2018 Threshold Systems Inc.





800 Å of amorphous silicon (AmSi) is deposited using Chemical Vapor Deposition(CVD) technology. The amorphous Silicon is NOT doped.

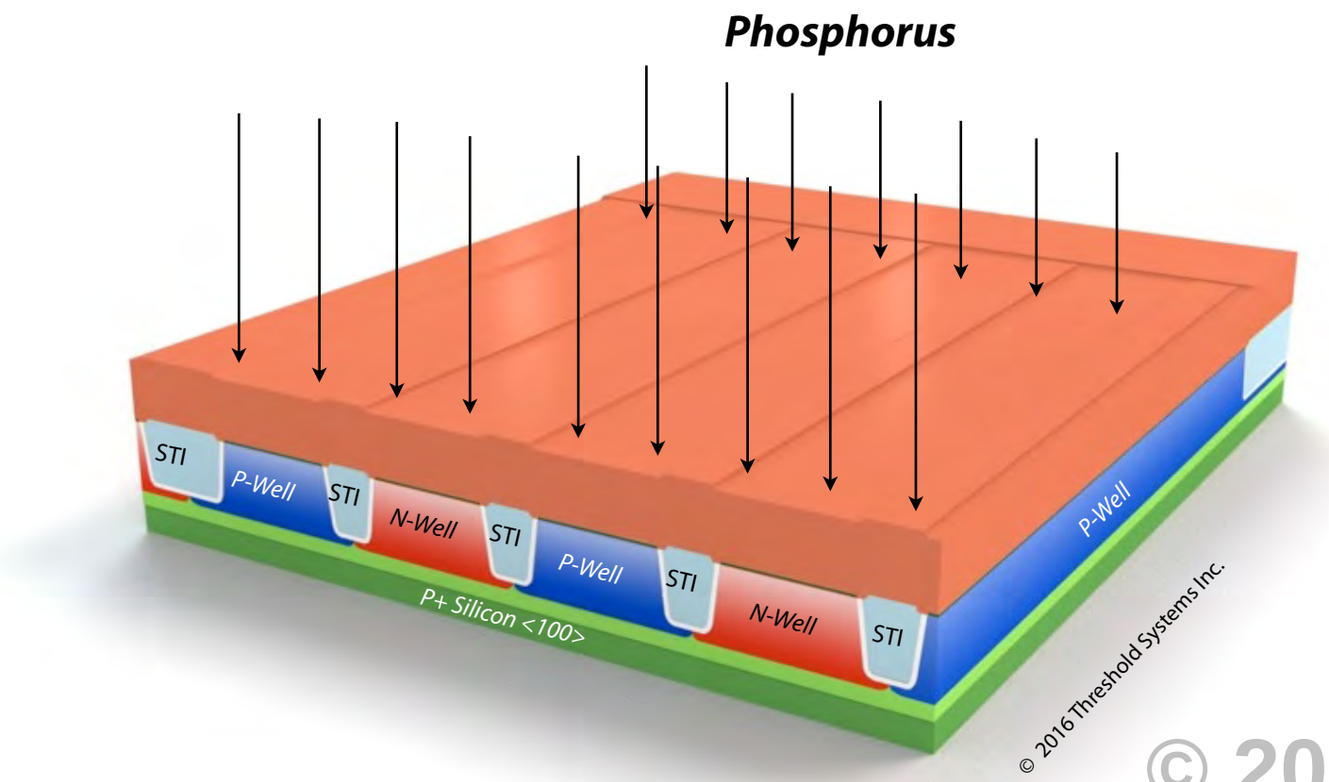
After it has been patterned and etched the AmSi forms the sacrificial gate mandrel.

Note that there is no metal gate beneath the amorphous silicon gate electrode.

© 2018 by Threshold Systems Inc. All Rights Reserved.

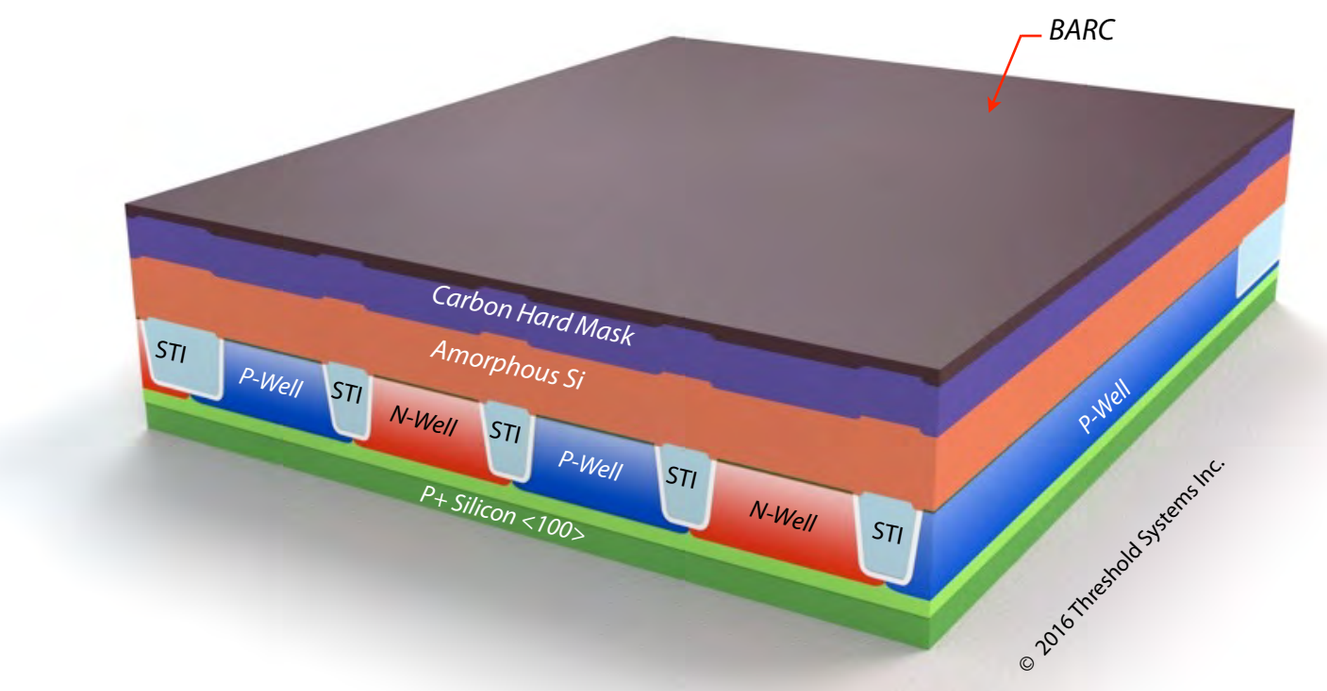
Amorphous Silicon Dope & Hard Mask Dep:

ThresholdSystems



The amorphous silicon is implant doped with Phosphorus (5.0E15@ 5 KeV, 0°) to make it conductive. The gate electrode is sacrificial and will be removed later in the process. Doping the gate electrode aids in its removal but is not necessary.

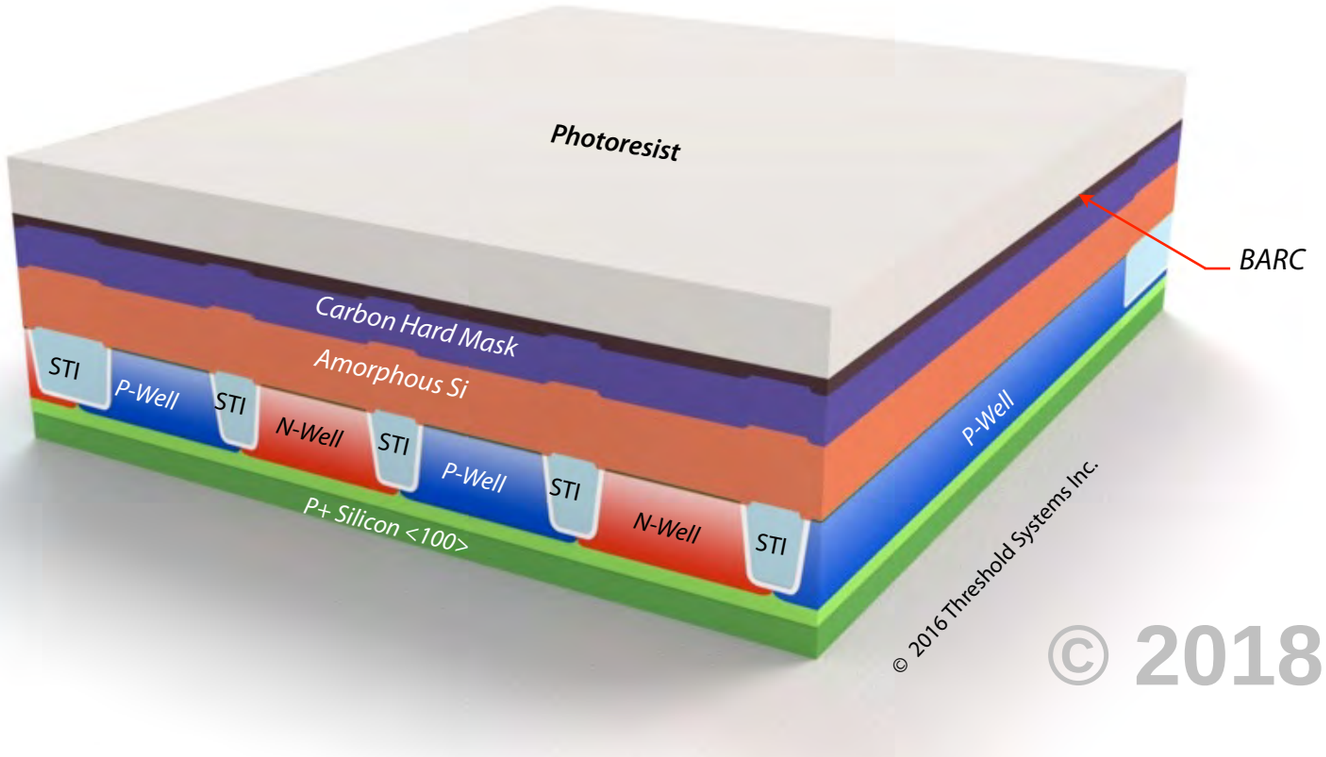
© 2018 Threshold Systems Inc.



A 1,500 Å thick amorphous Carbon hard mask and a 300 Å thick BARC layer are deposited.

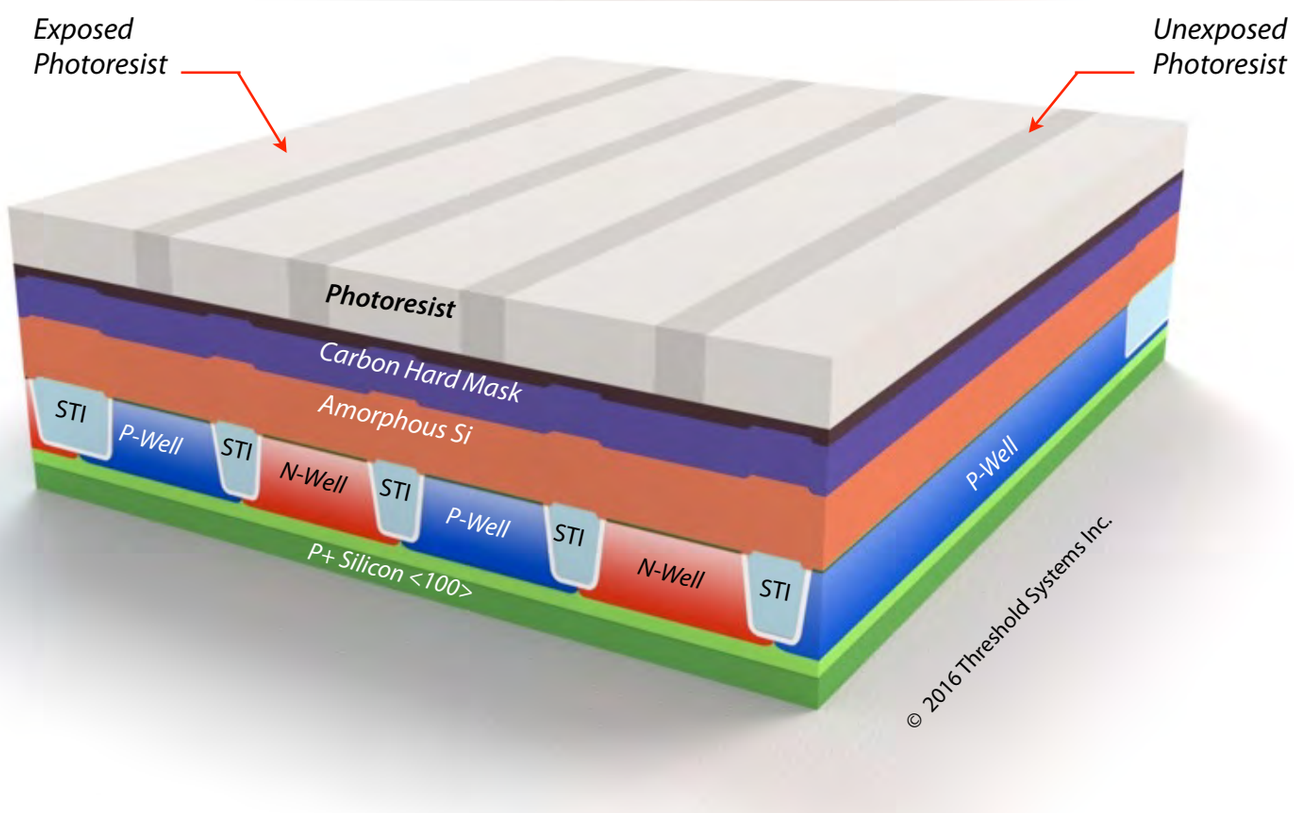
© 2018 by Threshold Systems Inc. All Rights Reserved.

Photoresist Deposition & 1st Patterning:



Next, the wafers are primed in HMDS, coated in BARC (Bottom Anti-Reflective Coating), photoresist and patterned. This is the most critical lithography operation in the entire process flow. It generates a series of parallel lines that will be further defined by a cut mask in a following lithography operation.

© 2018 Threshold Systems Inc.



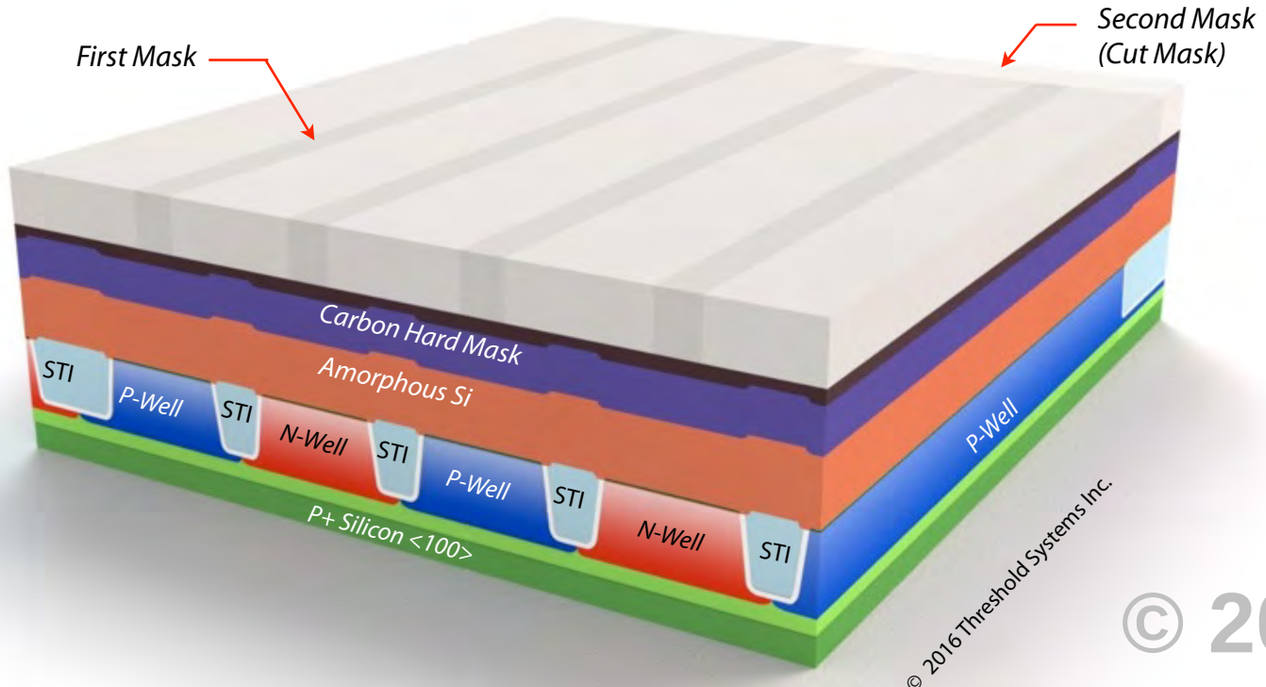
Next, the wafers are exposed using a grating mask. This mask prints straight lines in one direction only.

The photoresist is not developed yet.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Cut Mask 2nd Patterning:

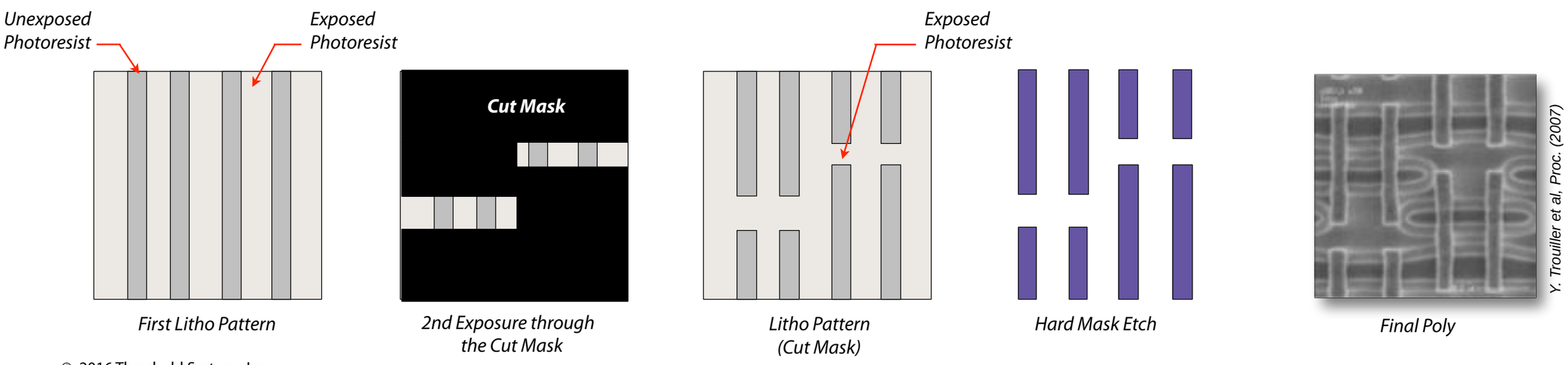
ThresholdSystems



The photoresist is then exposed through a "cut mask" which exposes the undeveloped photoresist in a direction perpendicular to the direction of the first mask. This exposes the photoresist along this axis and allows the continuous grating lines that run all across the chip to be cut.

© 2018 Threshold Systems Inc.

Top-Down View of the Cut Mask Operation

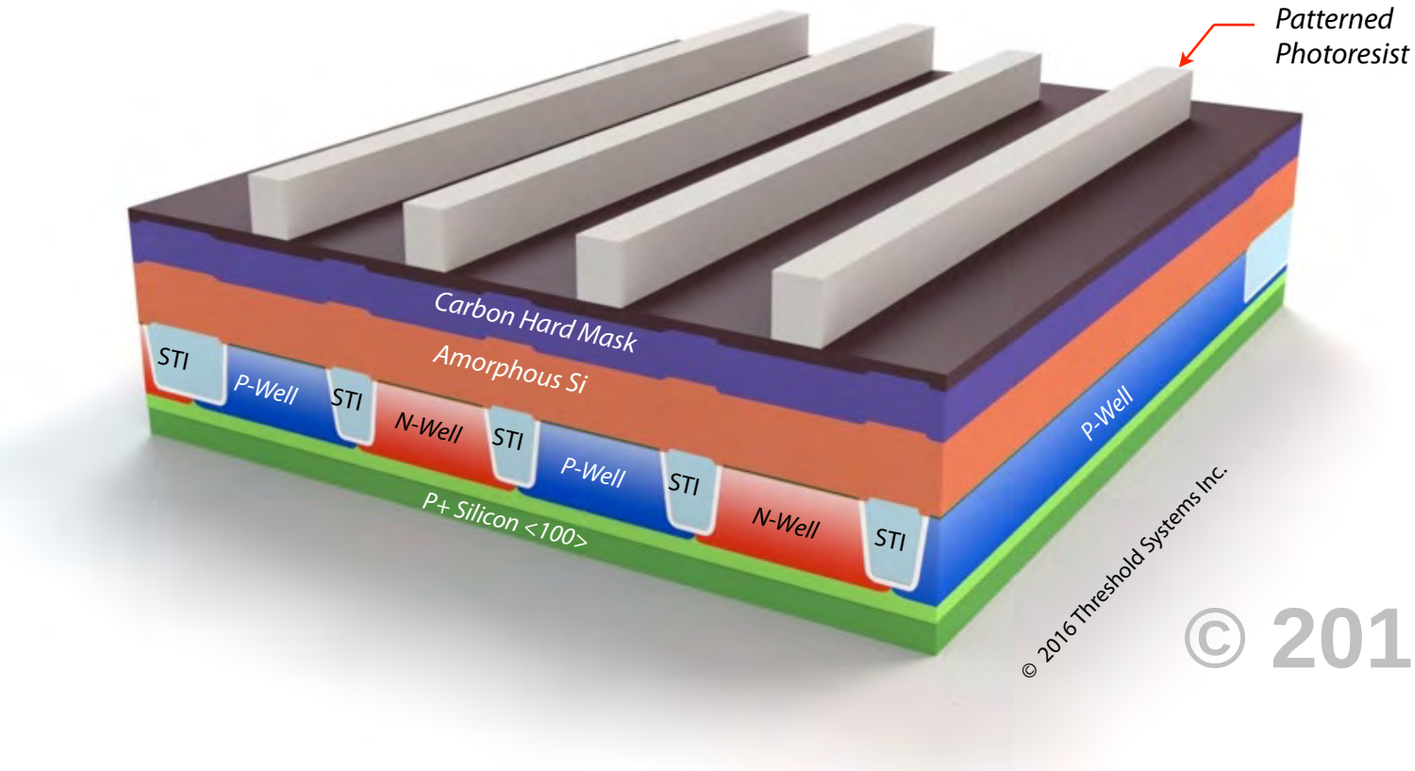


© 2016 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Photoresist Trim Process:

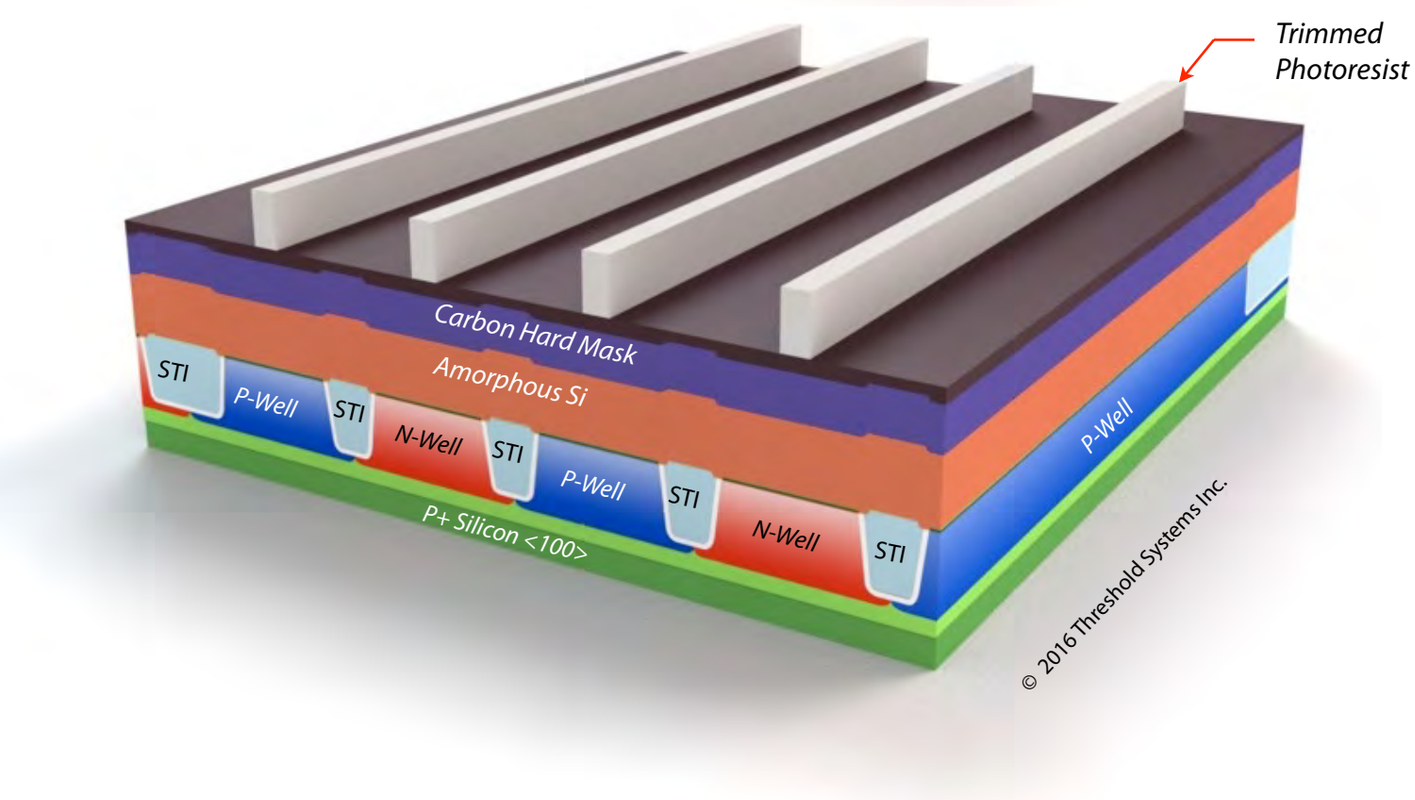
ThresholdSystems



The wafers are then loaded into an etcher and exposed to a controlled oxygen plasma that isotropically etches the patterned photoresist. This causes the photoresist to shrink in size to a targeted dimension.

This process is know as "photoresist trim".

© 2018 Threshold Systems Inc.



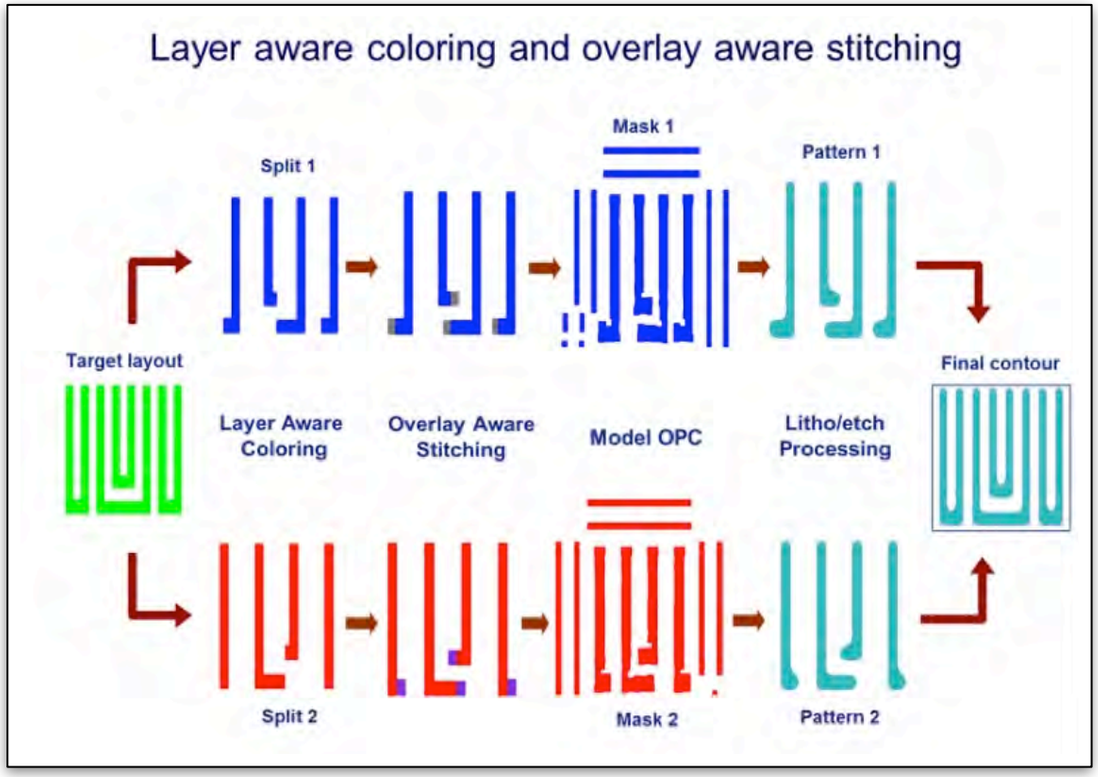
The trimmed photoresist lines are now narrower and can be used to define smaller narrower gate electrodes.

Photoresist trim is a technique that allows the fabrication of very narrow photoresist lines (and thus narrow gate electrode lines) that are smaller that the photoresist lines that are optically printed.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Multiple Mask Patterning:

- *Double exposure is a sequence of two separate exposures of the same photoresist layer using two different photomasks. This technique is commonly used for patterns in the same layer which look very different or have incompatible densities or pitches that cannot be successfully printed with a single mask.*
- *Litho-Etch-Litho-Etch (LELE) is the most basic style of double patterning, where a pattern with unresolvable pitch is decomposed into two mask layouts, each with resolvable features colored differently.*
- *The green colored pattern to the left is the desired pattern created by the chip designer. However the 193i Lithography tool is incapable of resolving such a pattern which has lines that are so densely packed.*
- *The pattern is split into two separate masks (the Red and the Blue masks).*
- *The two fractured patterns are then overlaid to ensure that they will stitch together. Next Optical Proximity (OPC) correction is performed on the pattern and the physical masks are generated.*
- *The patterns printed with mask one and mask two will appear on the wafer after the LELE processing as indicated in the light blue patterns in the diagram to the left. These two patterns are printed separately and overlaid on the wafer to produce the final contour on the wafer.*
- *Multiple patterning is necessary because we don't have a lithography tool that is capable of resolving such fine patterns. It is also adds enormously to the cost of new technology nodes. So much so that the 22/14nm nodes have not presented any of the normal cost reductions associated with new technology nodes.*



Source: ASML

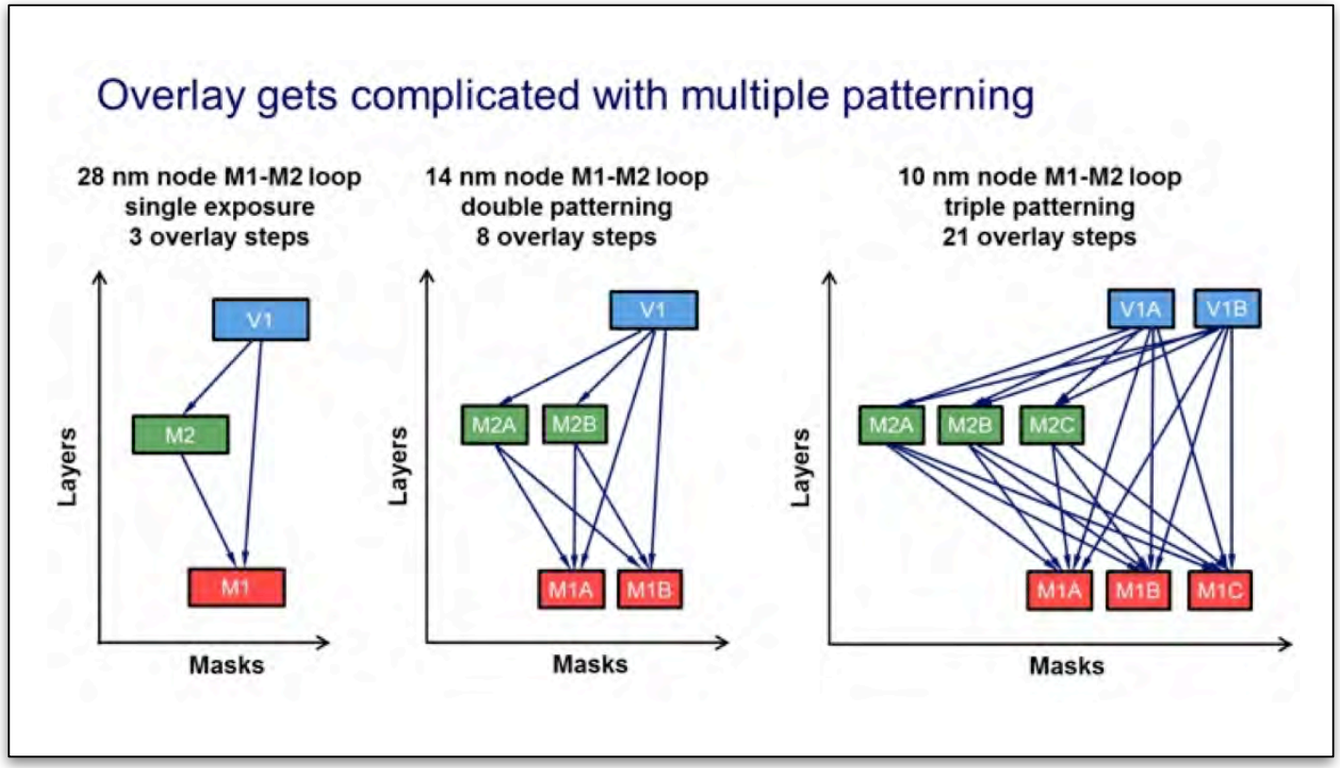
© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Overlay Complexity:

ThresholdSystems

- *The central issue with double and triple patterning is overlay alignment. Aligning multiple patterns is complex, difficult, time consuming and expensive. It doubles or triples the lithography cycle time.*
- *Because we cannot resolve the ultra-fine lines required to pattern sub-20nm node devices, we break the mask patterns up into two or sometimes three masks, print them two or three times, and rely on overlay alignment to stitch the patterns together to form the intended final pattern.*
- *For example, in a traditional single exposure methodology, a via and two associated metal lines aligned together would require 3 overlay operations (refer to the diagram to the right).*
- *For a double patterning of the same structures, each of the metal lines would be split into two different masks. This would require eight separate overlay operations because the via would have to be aligned to the each metal line twice and the metal lines would have to be aligned to each other.*
- *Triple layer patterning (required at the 10nm node), where the metal layers are split onto three masks and the vias into two masks, would require 21 overlay alignments.*
- *The 10nm node may require quadruple layer patterning and the same example would require 40 overlay operations. This would be a very expensive process, perhaps prohibitively expensive.*



Source: ASML

© 2018 Threshold Systems Inc.

Source: ASML

Double and triple layer patterning would not be necessary if EUV were available …

© 2018 by Threshold Systems Inc. All Rights Reserved.

Photoresist Trim:

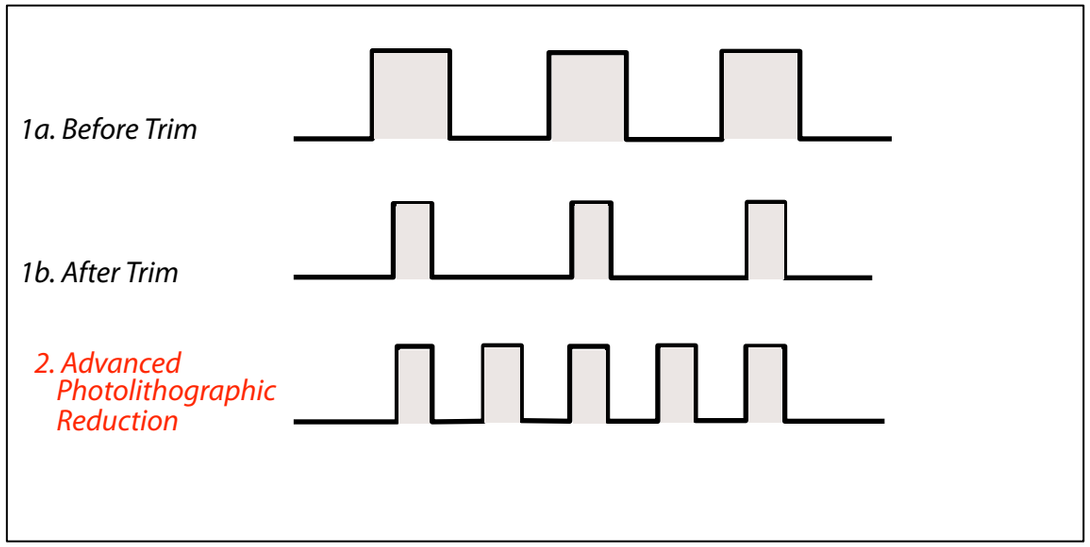
ThresholdSystems

- *Photoresist trimming is a "trick-of-the-trade" used by Etch Engineers to allow them to etch polysilicon to dimensions that are smaller than the photoresist lines printed on the wafer by the lithography process.*
- *It works like this:*

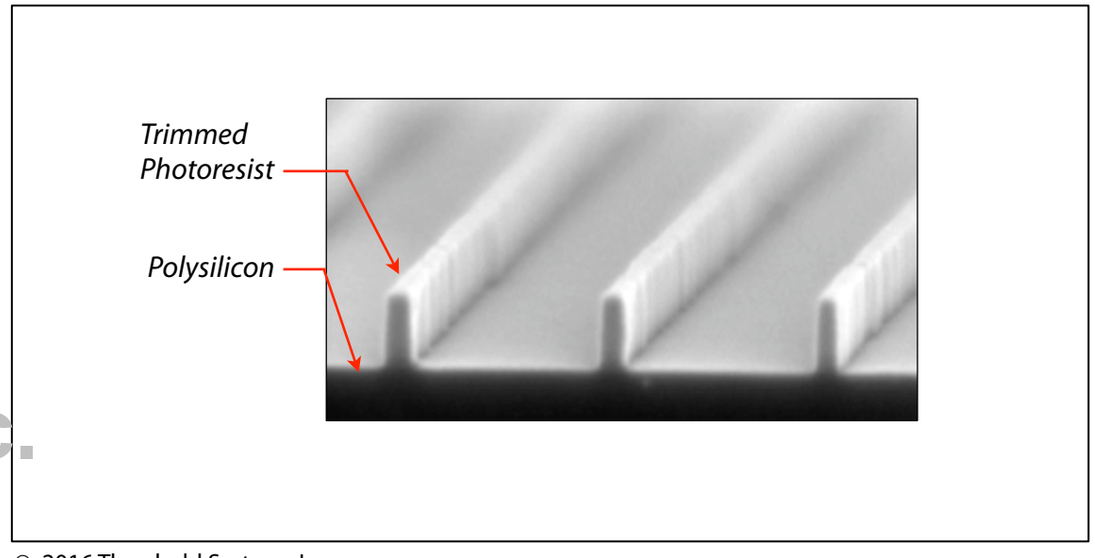
  a) the photoresist lines are printed and developed (1a)

  b) the developed photoresist lines are exposed to an oxygen plasma that uniformly shrinks their length dimension (1b)

  c) these new smaller photoresist lines are used to etch the polysilicon
- *This remarkable trick allows gate electrodes to be etched that are smaller than the smallest dimension that the lithography scanner is capable of printing.*
- *However, there is a limit to how far photoresist can be trimmed (a maximum of ~ 30-40%). Beyond this point the photoresist simply becomes too tall and looses its structural integrity.*
- *It is important to realize that although photoresist trimming reduces the dimension the resist lines, it does not increase their density. Increased density and smaller photoresist lines can only be achieved through advanced photolithographic reduction (refer to #2 above right).*



© 2016 Threshold Systems Inc.



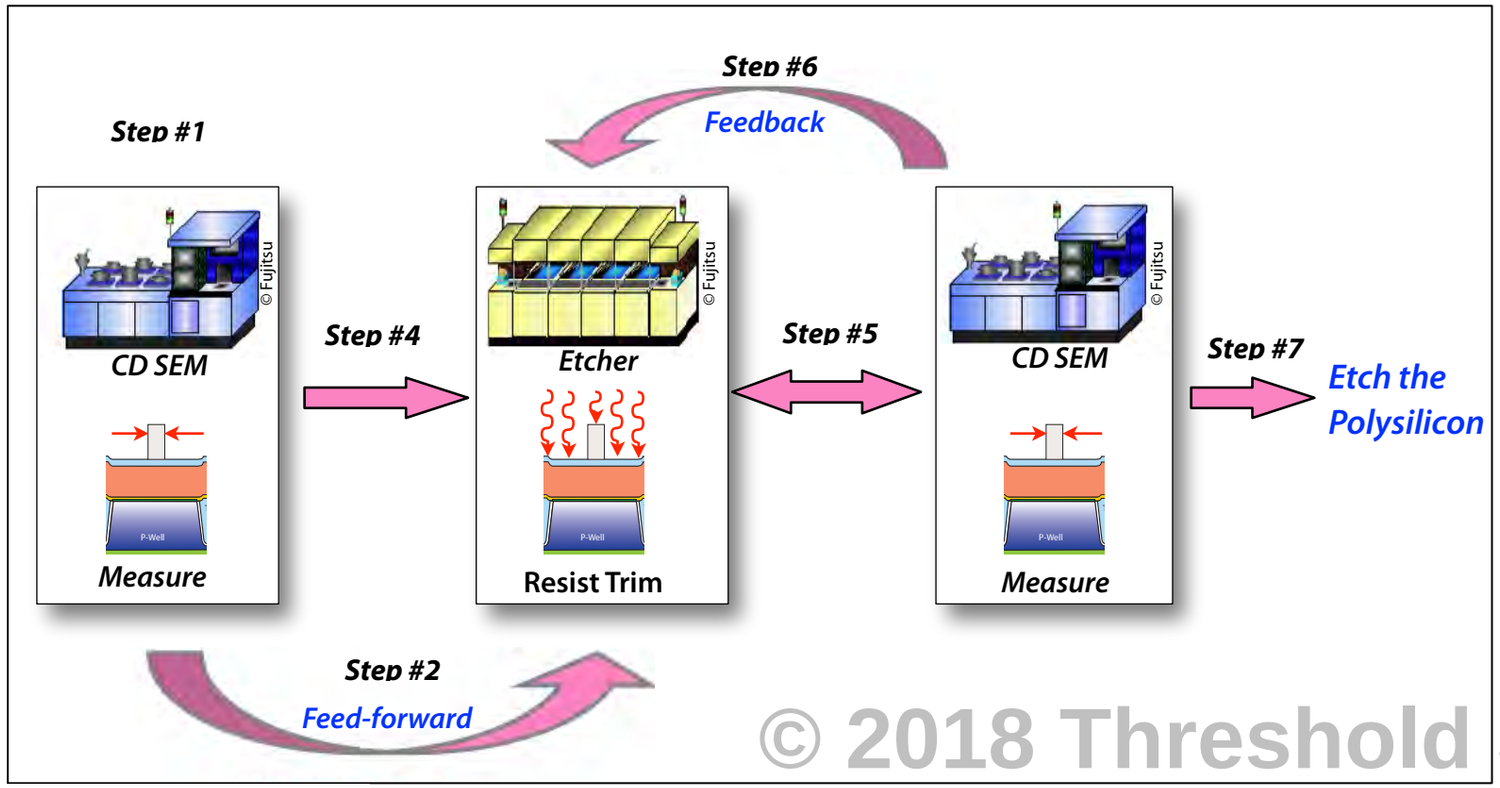
© 2016 Threshold Systems Inc.

© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

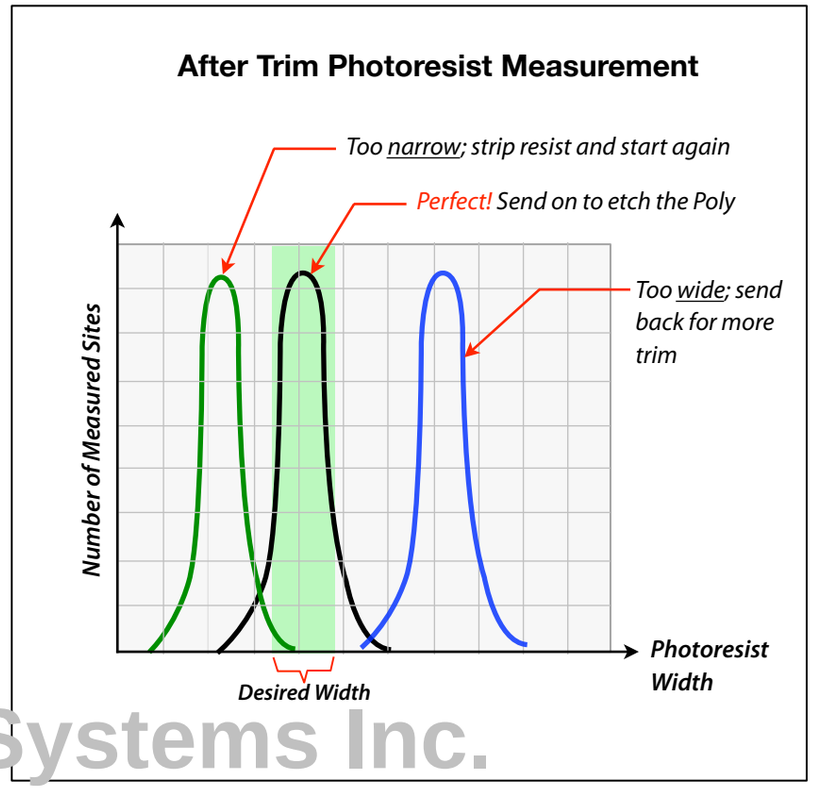
Photoresist Trim Process:

ThresholdSystems



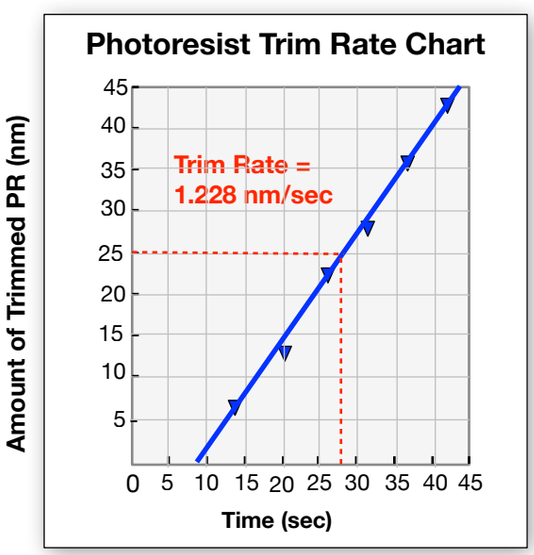
© 2018 Threshold Systems Inc.

© 2016 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

Step #3



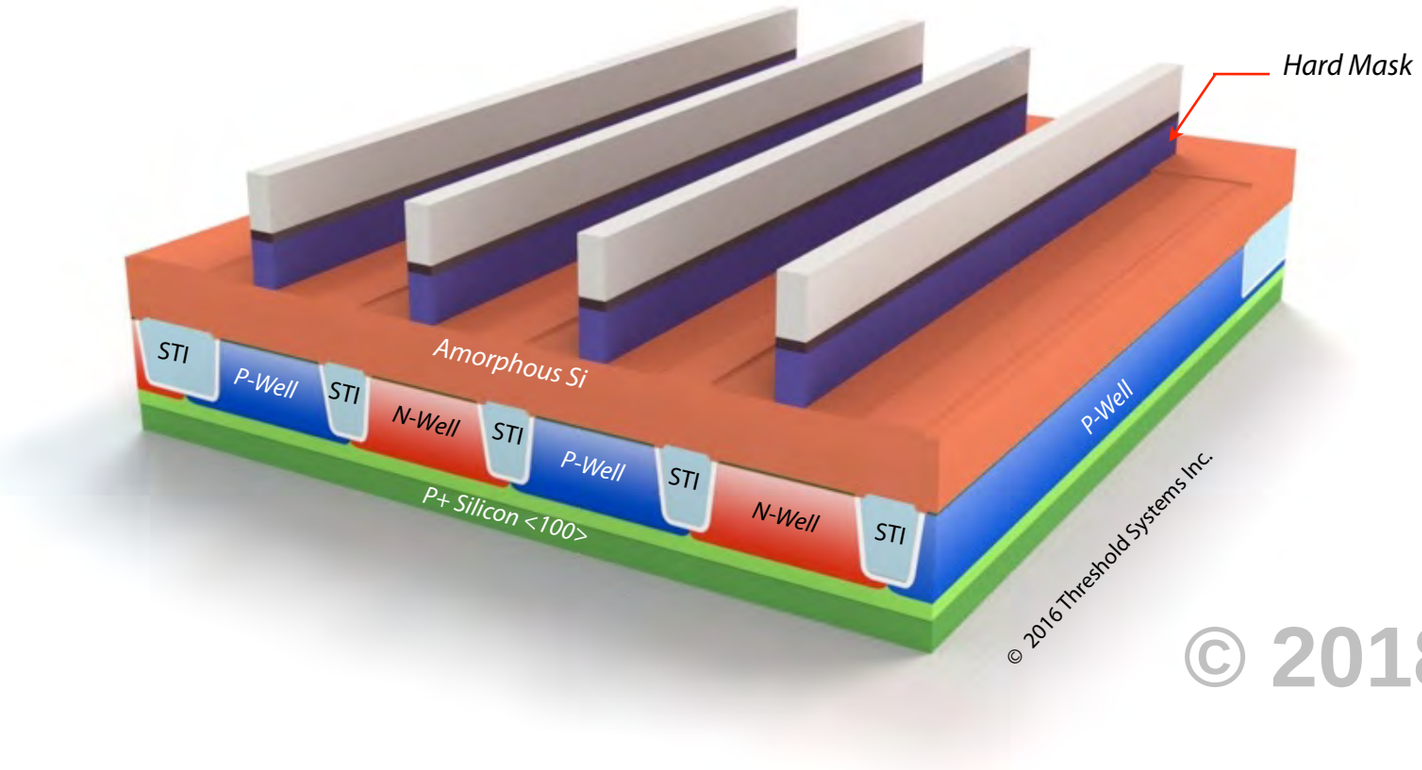
© 2016 Threshold Systems Inc.

- *Trimming photoresist is an iterative process that is performed on every wafer.*
- *After the resist has been printed and developed its width is measured in a SEM (step #1). This measurement information is feed-forward to the etcher that is performing the resist trim (step #2).*
- *A photoresist trim etch rate look-up chart is consulted (step #3) to determine how long to trim the resist for to obtain the desired dimension.*
- *The wafers on then sent on to have the photoresist trimmed in an etcher (step #4) and then the wafers are sent on so that the resist after-trim dimension can be measured in a SEM (step #5).*
- *This information is feedback to the etcher (step #6) and the etch time is adjusted to fine-tune the resist trim process. If the resist is not small enough, the wafers are returned to the etcher to have their resist additionally trimmed.*
- *If the resist is the correct dimension the wafers are sent on to have the polysilicon etched (Step #7). If too much resist has been trimmed, then the entire photoresist pattern will be stripped and the entire photolithographic patterning and resist trim process must be repeated.*
- *This process is completely automated.*

© 2018 by Threshold Systems Inc. All Rights Reserved.

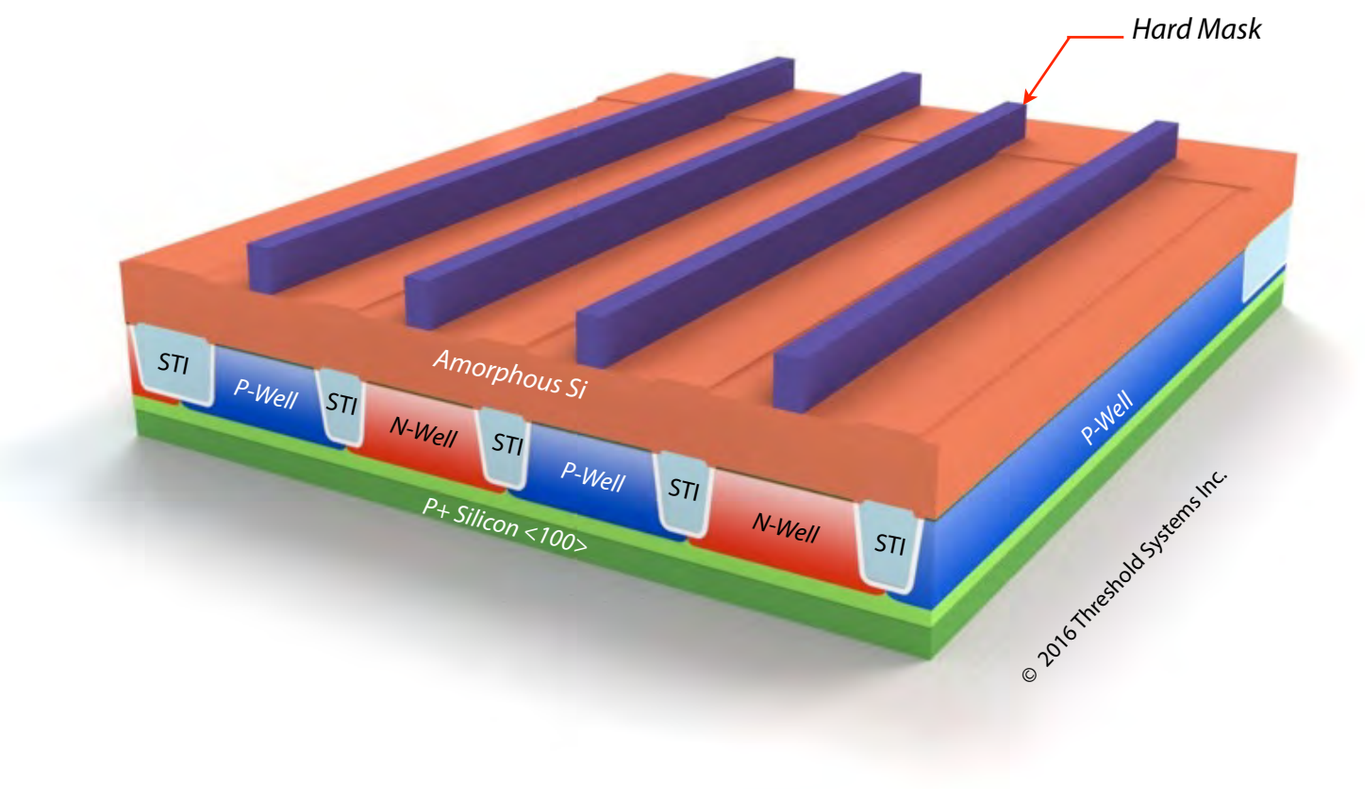
Hard Mask Etch & Photoresist Strip:

ThresholdSystems



The amorphous carbon hard mask is etched in a highly anisotropic etch process that stops on the amorphous silicon.

© 2018 Threshold Systems Inc.

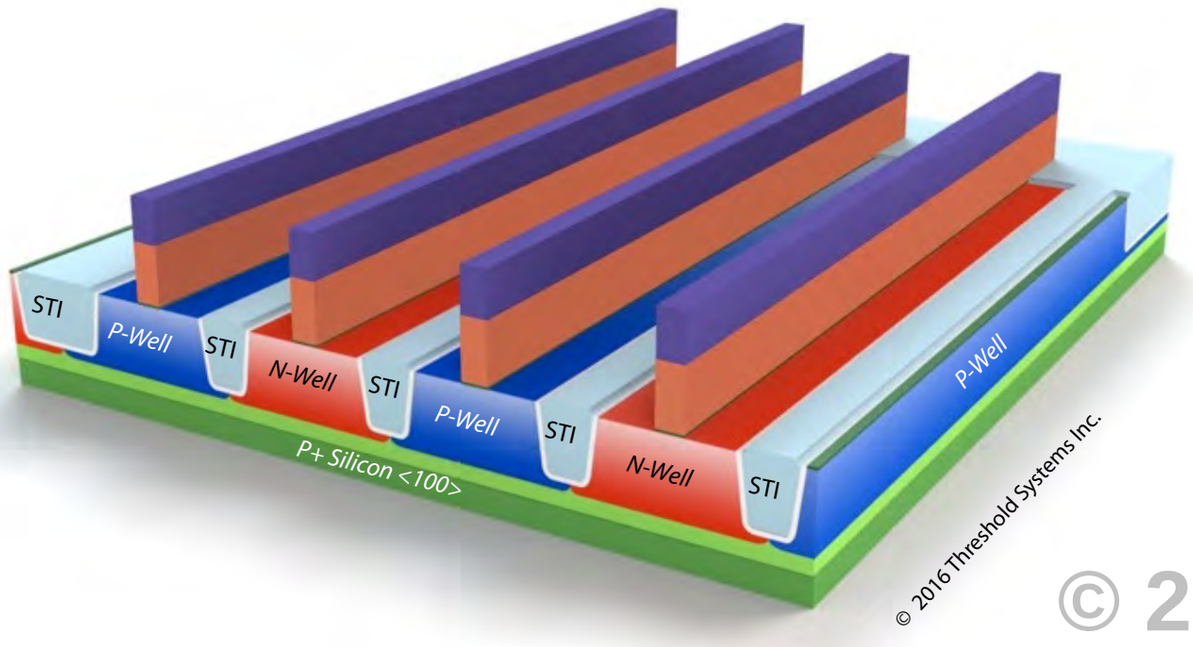


The photoresist is stripped and the wafers are cleaned in Piranha.

© 2018 by Threshold Systems Inc. All Rights Reserved.

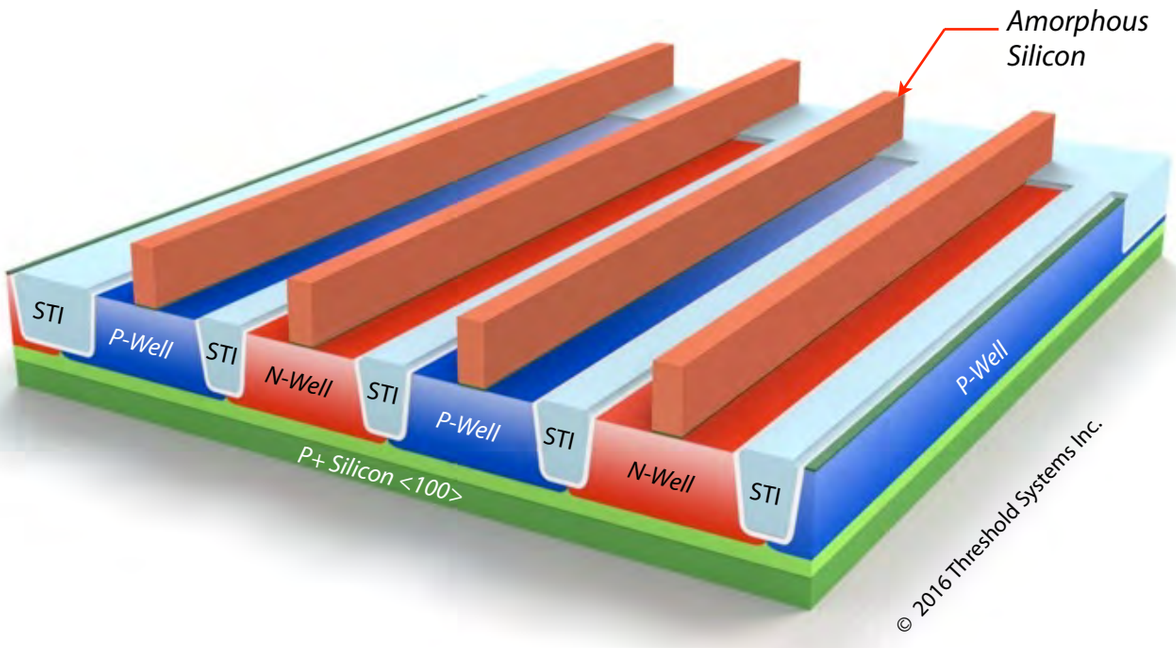
Amorphous Silicon Etch & Hard Mask Strip:

ThresholdSystems



A highly anisotropic etch recipe is then used to transfer the pattern from the hard masks into the amorphous silicon.

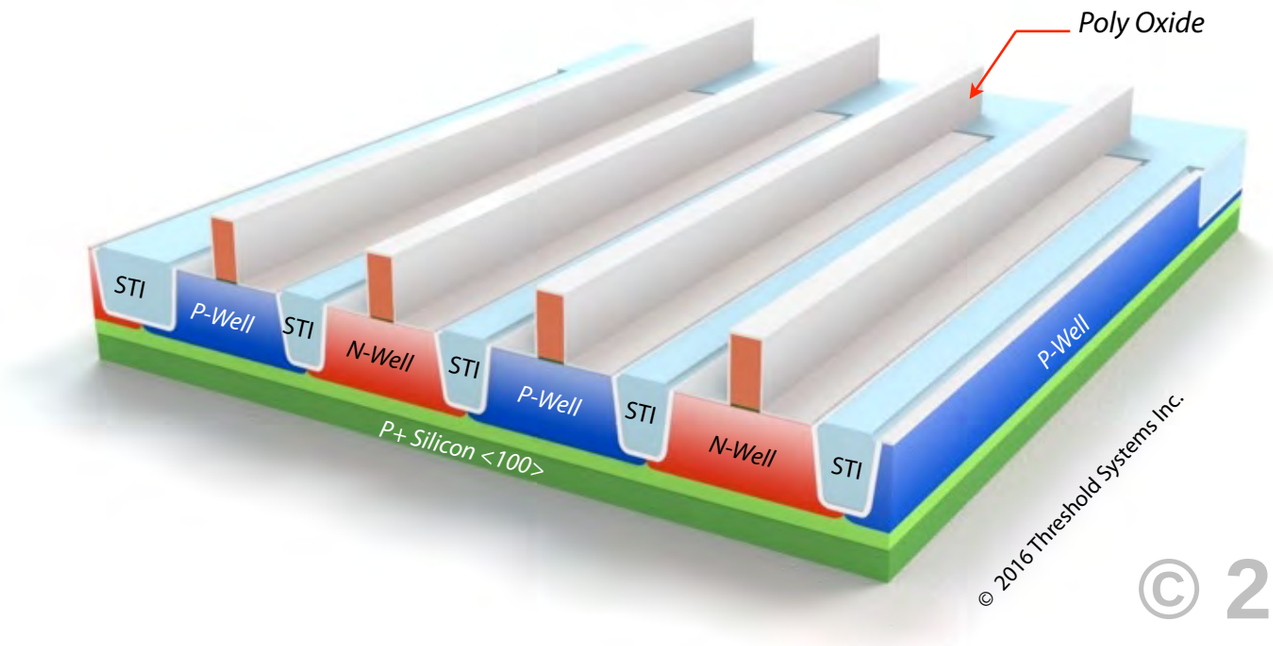
© 2018 Threshold Systems Inc.



The Hard Mask is stripped.

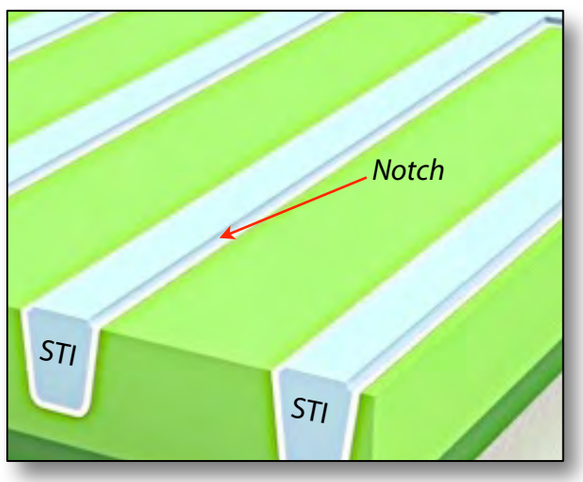
© 2018 by Threshold Systems Inc. All Rights Reserved.

Poly Oxidation & Poly Stringers:

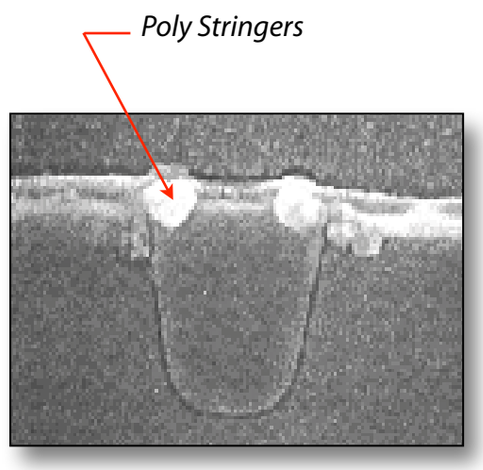


The wafers are cleaned in Piranha and 15Å of thermal poly oxide is grown. This is followed by the deposition of 15Å of CVD oxide. These two oxide layers form what is referred to as the "Offset Spacer" (sometimes also referred to as "Spacer-Zero" or the "Spacer-1"). In previous generation of devices this oxide was much thicker and referred to as "Poly Oxide".

© 2018 Threshold Systems Inc.



© 2016 Threshold Systems Inc.



The original Pad Oxide strip (performed several operations ago after trench fill and nitride removal) removes oxide preferentially from the trench corners resulting in the creation of "notches" all along the sides of the trenches that facilitate the formation of polysilicon stringers.

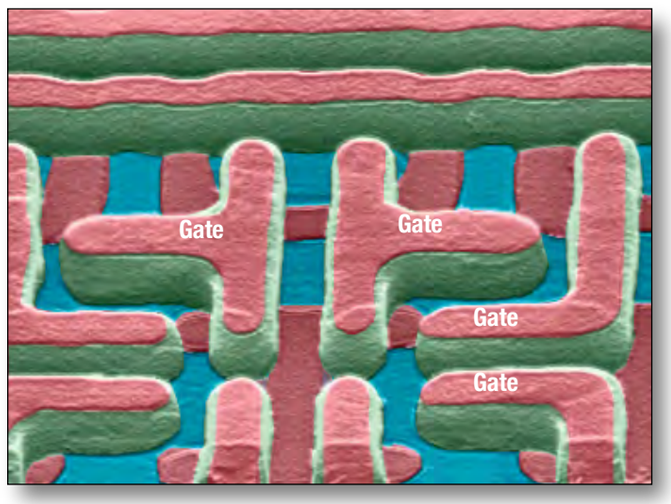
The poly stringers form at the end of the polysilicon etch operation. The final phase of this etch process is highly selective over-etch that is designed to remove any small amounts of polysilicon remaining in the Source/Drain regions and in the notches. If this phase of the etch is not carefully designed, the thick polysilicon stringers will remain in the trench notches and will form unintended conductive pathways that will short-out adjacent transistors.

For this reason it is very important to minimize the size of the notches to facilitate easy stringer removal.

© 2018 by Threshold Systems Inc. All Rights Reserved.

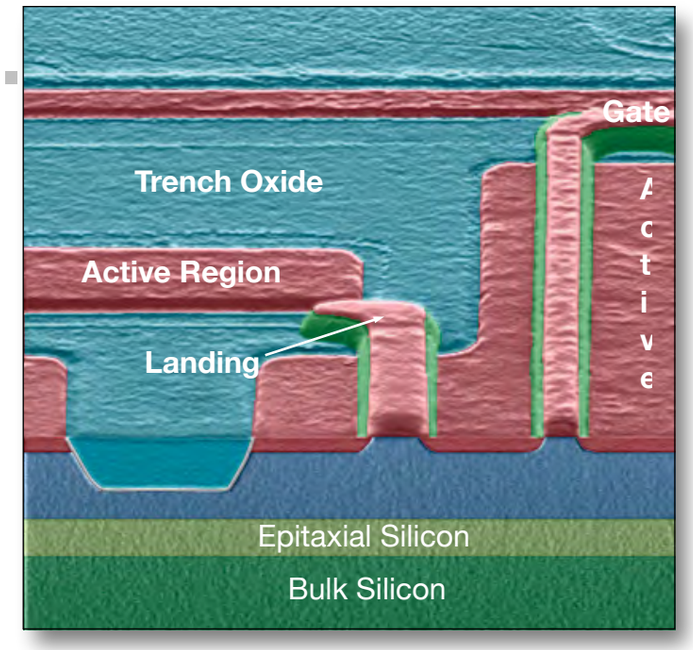
What Do These Structures Really Look Like?

© 2018 Threshold Systems Inc.



© Threshold Systems 2006

An artificially colored section of a memory array displaying 3D angled views of real-world gate electrodes running over active regions and trench oxide.



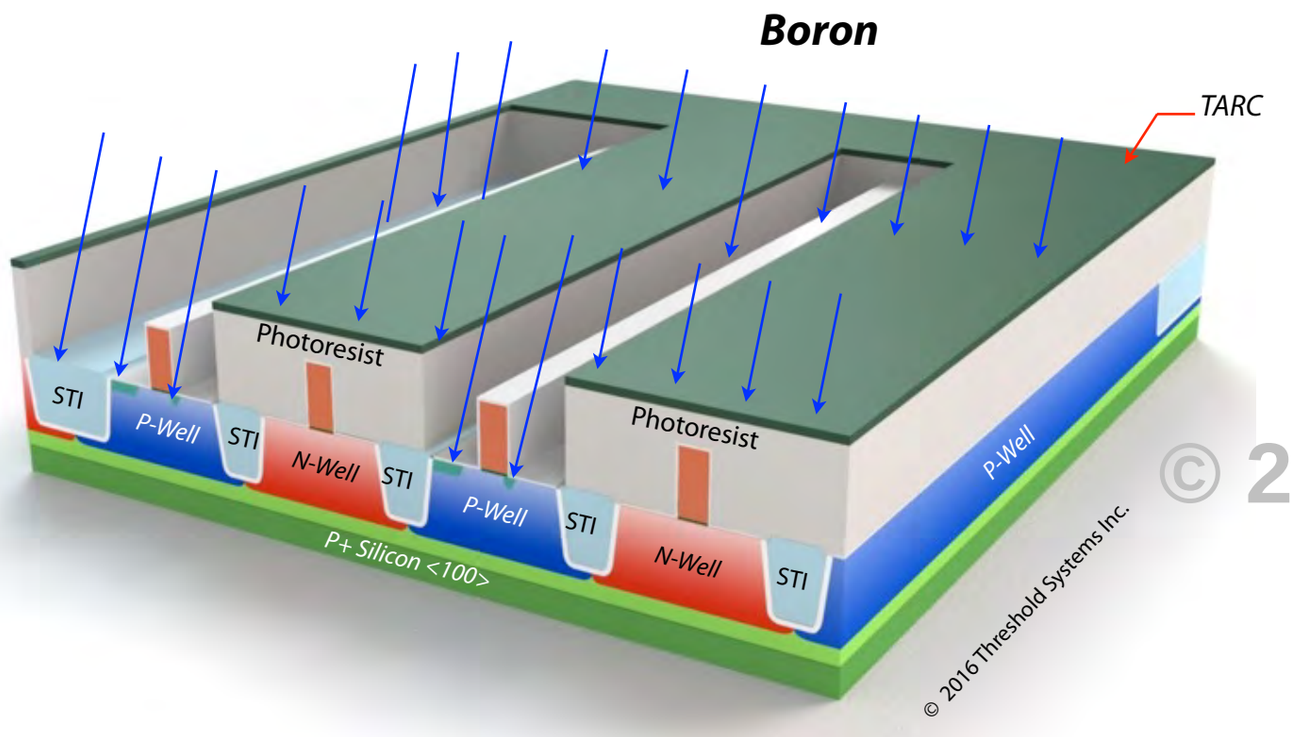
© Threshold Systems 2006

An artificially colored section of a Logic chip displaying an angled view of the trench oxide, active areas and gate electrodes. Note how the trench oxide is not always at minimum dimension and that it assumes a wide range of shapes and sizes.

© 2018 by Threshold Systems Inc. All Rights Reserved.

NMOS Halo Implant:

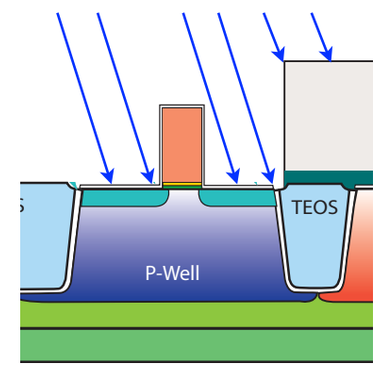
ThresholdSystems

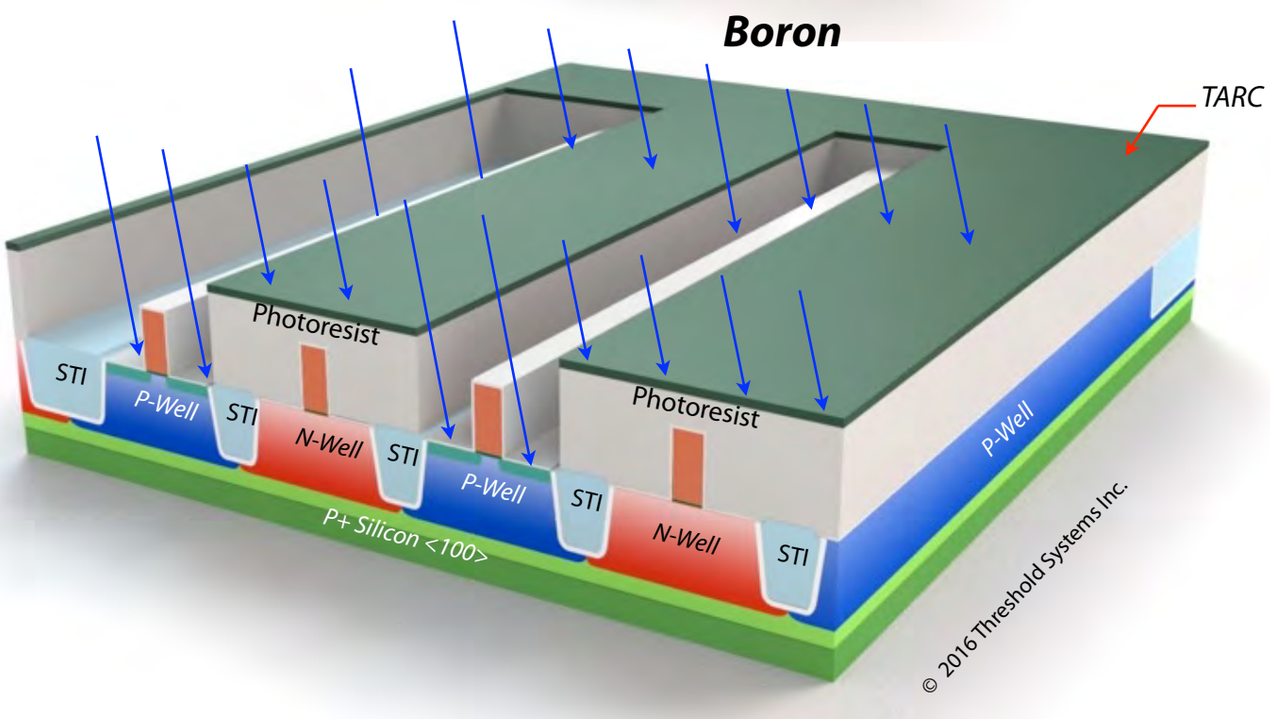


The wafers are coated with photoresist and a Top Anti-Reflective Coating (TARC) and then patterned. After patterning the photoresist covers all of the active areas except the P-Wells.

A Halo implant of Boron (4.4E13@3KeV), at a tilt angle of 30 º, is applied to the High Vt P-Well. This is a quad implant meaning that the wafers are implanted four times and rotated 90 º between implants. This ensures that all sides of the gate electrode have a pocket of Boron implanted beneath the gate edge. The Halo implant is ~ 150 - 200 Å deep.

© 2018 Threshold Systems Inc.



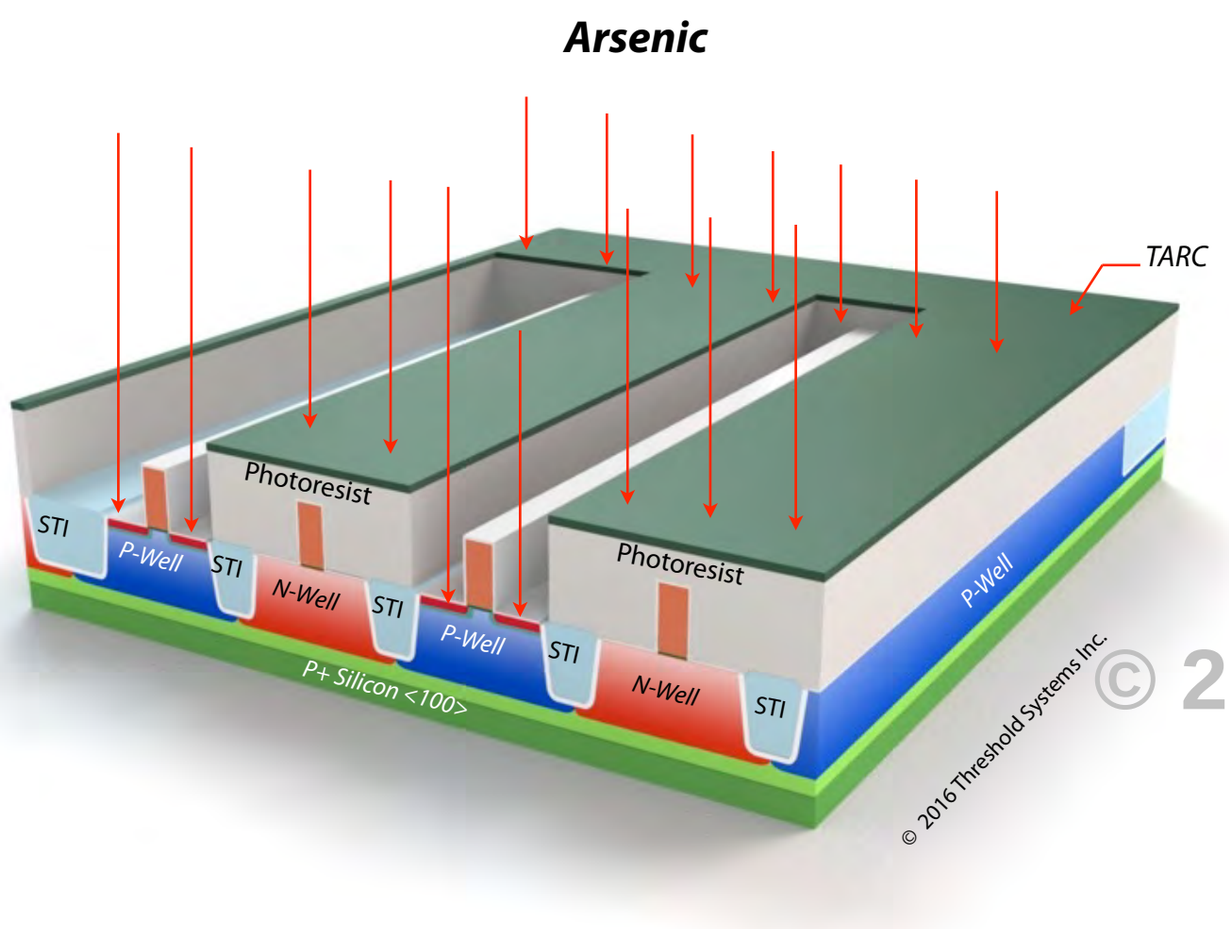


A view of the Boron Halo implant when the wafers have been rotated 180 degrees. This would be the third of the four implants that comprise the quad Halo implant. The others being implanted when the wafer is at angular positions of zero, 90 and 270 degrees.

© 2018 by Threshold Systems Inc. All Rights Reserved.

NMOS Extension Implant:

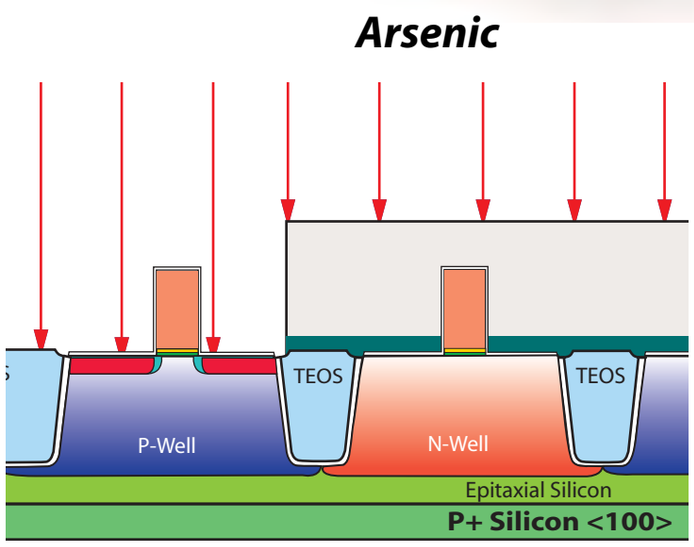
ThresholdSystems



© 2018 Threshold Systems Inc.

Using the same photoresist pattern that was used for the Halo implant, an Arsenic Extension implant (1E15@1 KeV, 0º) is implanted into the P-Wels. This implant establishes the Extension implants of the NMOS transistors.

The extension implants are at approximately the same depth as the Halo implants (150-200 Å for a 20nm device).



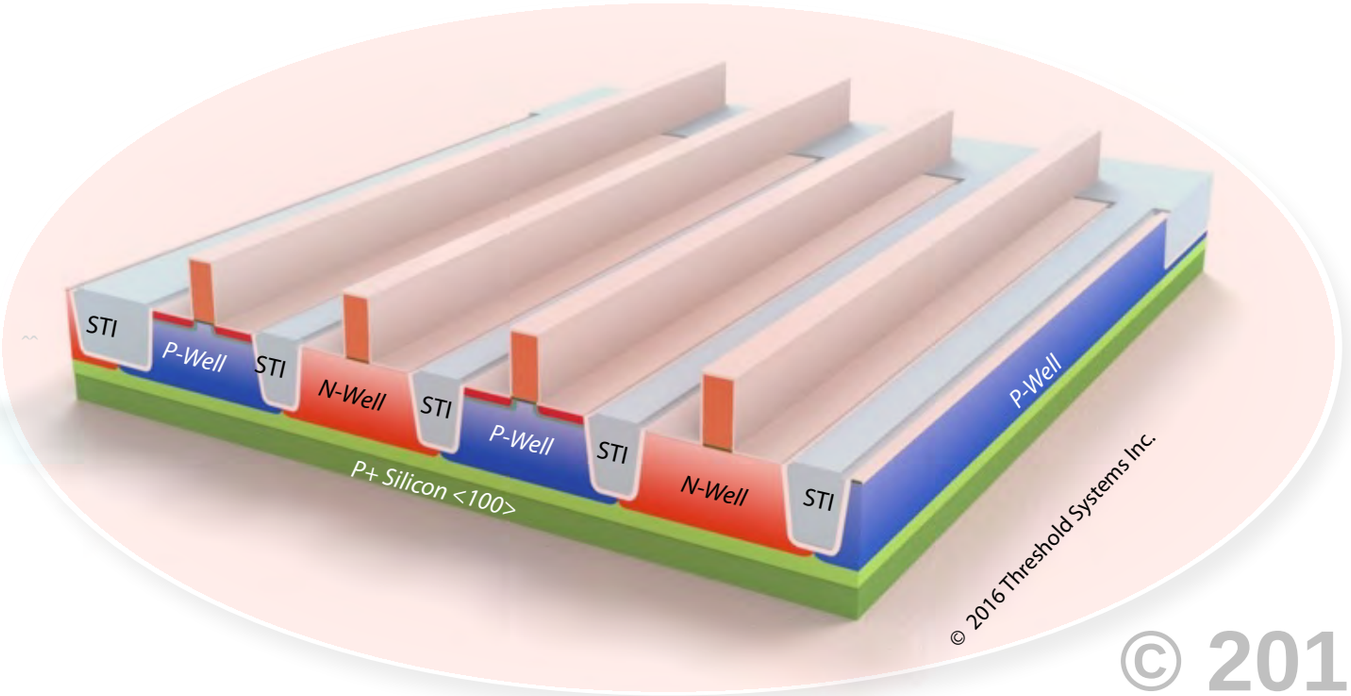
© 2016 Threshold Systems Inc.

The adjacent 2D image clearly displays the Extension implant. This implant is self-aligned to the edge of the gate electrode and this high dose implant overwhelms the Halo implant in the area where the two implants overlap. However, the tip of the Halo implant is tucked under the gate electrode and the electrode shields the tip of the Halo from the Extension implant in this location.

The Extension implant determines the final L-effective (Leff) of the transistor and has a huge effect on Id_{SAT}.

© 2018 by Threshold Systems Inc. All Rights Reserved.

NMOS Anneal & PMOS Halo:

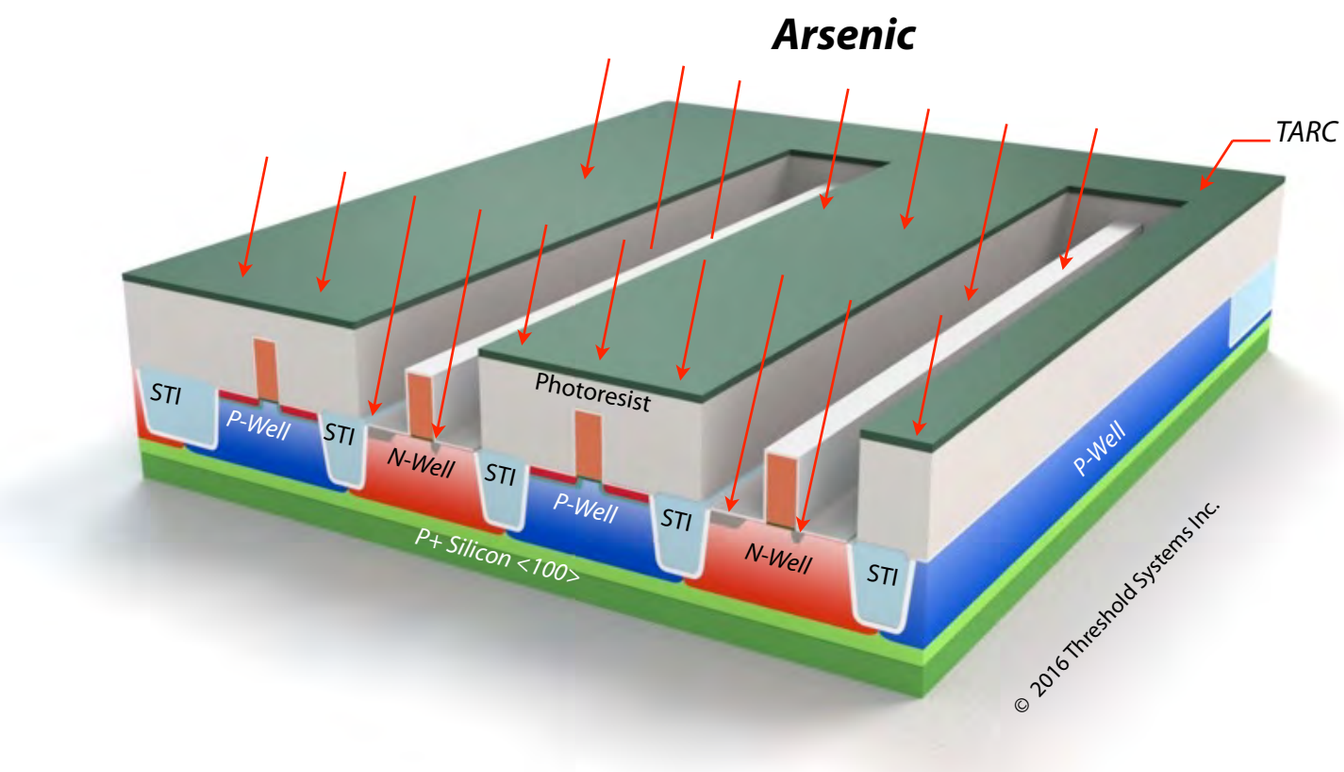


The photoresist from the previous operation is ashed away and the wafers are cleaned in Piranha.

The wafers are then undergo a Rapid Thermal Process (~950 degree C Spike for ~ 1 second followed by a ~1,350 degrees FLASH anneal for 1-3 milliseconds) that activates the NMOS Extension and Halo implants. In some process flows the Flash anneal is replaced by a sub-melt Laser anneal at the same temperature of ~1,350 degrees C. The location of the Extension implant fixes the final Leff of the NMOS transistors.

The anneal makes the dopant atom electrically active and anneals out the damage done to the silicon during the implants.

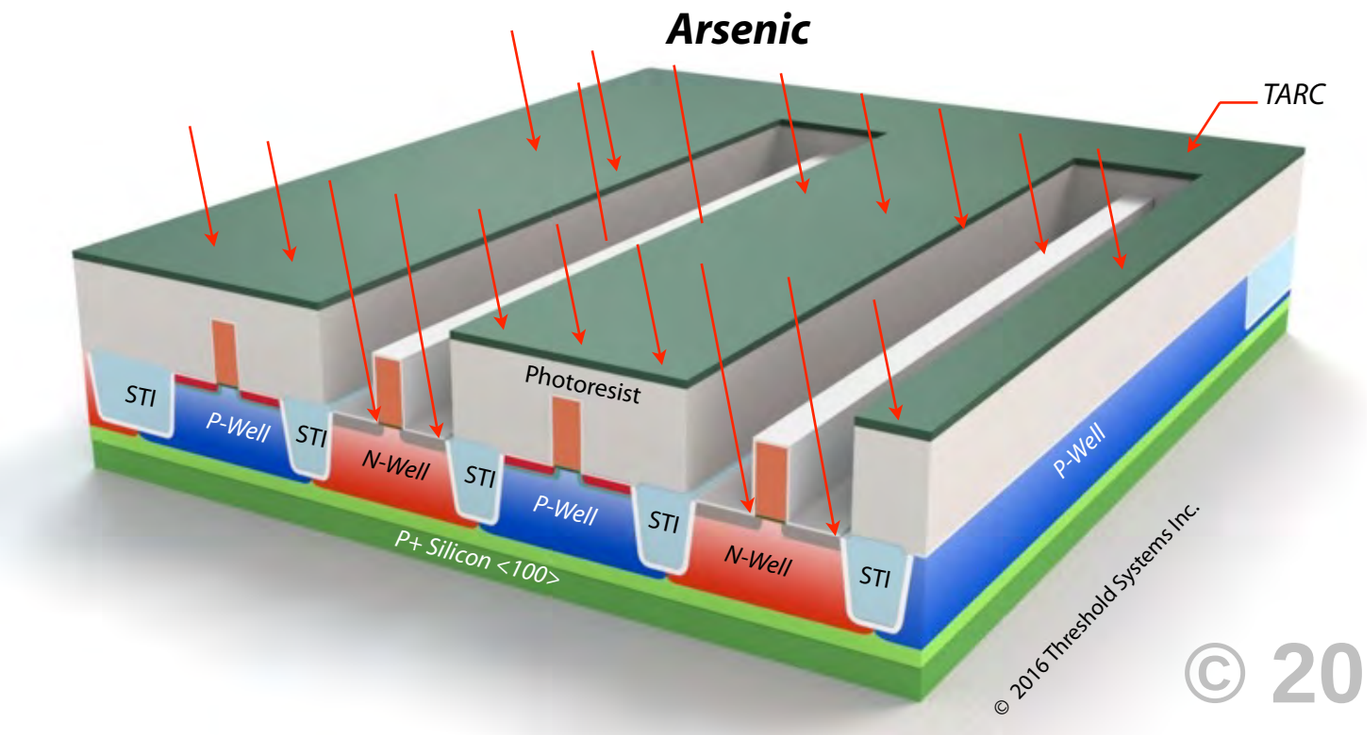
© 2018 Threshold Systems Inc.



Next, the wafers are patterned with photoresist and coated with TARC. After photoresist development only the N-Wells are exposed. A tilted 30 º Arsenic Halo implant (2.2E13@20KeV) is applied to this well. It is a quad implant.

© 2018 by Threshold Systems Inc. All Rights Reserved.

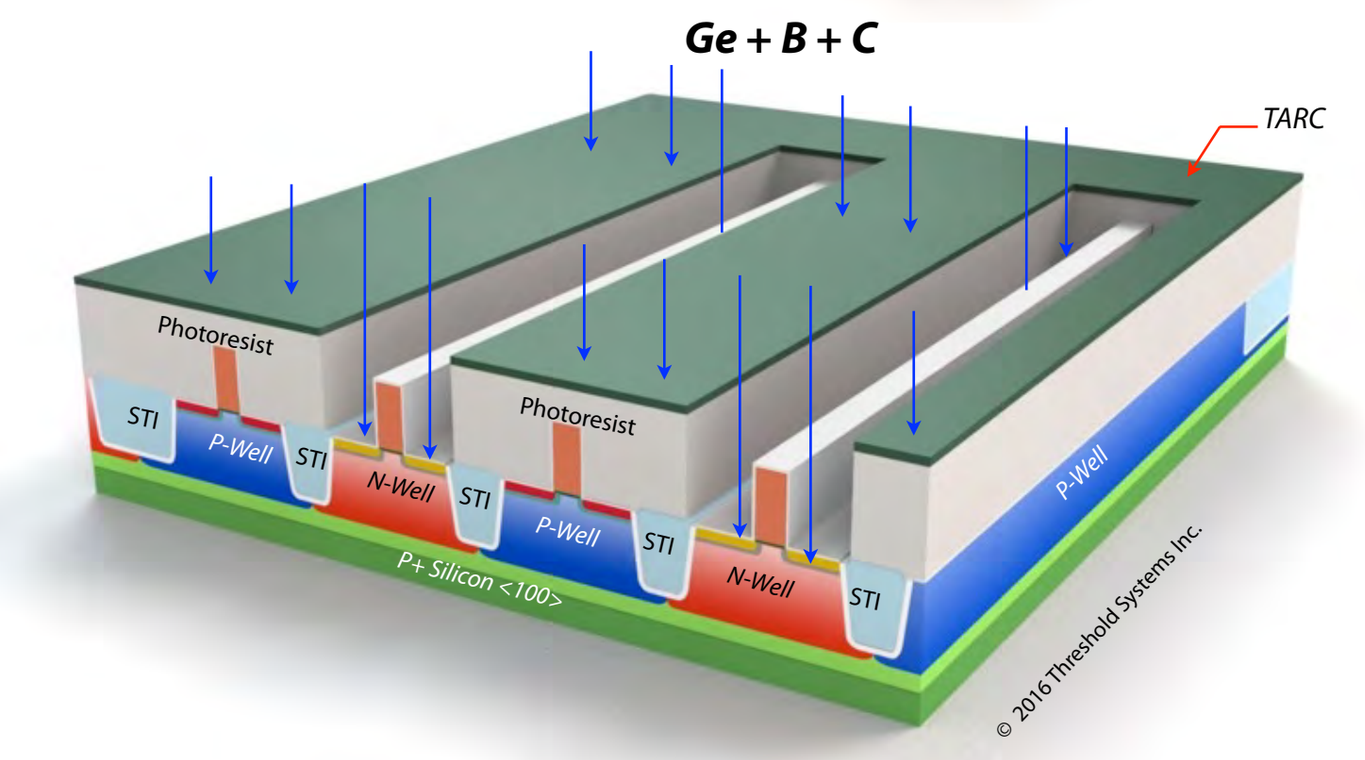
PMOS Halo & Extension:



A view of the Arsenic Halo implant when the wafers have been rotated 180 degrees. This would be the third of the four implants that comprise the quad Halo implant. The others being implanted when the wafer is at angular positions of zero, 90 and 270 degrees.

The Halo implant is ~ 150 - 200 Å deep for a 20nm node device.

© 2018 Threshold Systems Inc.

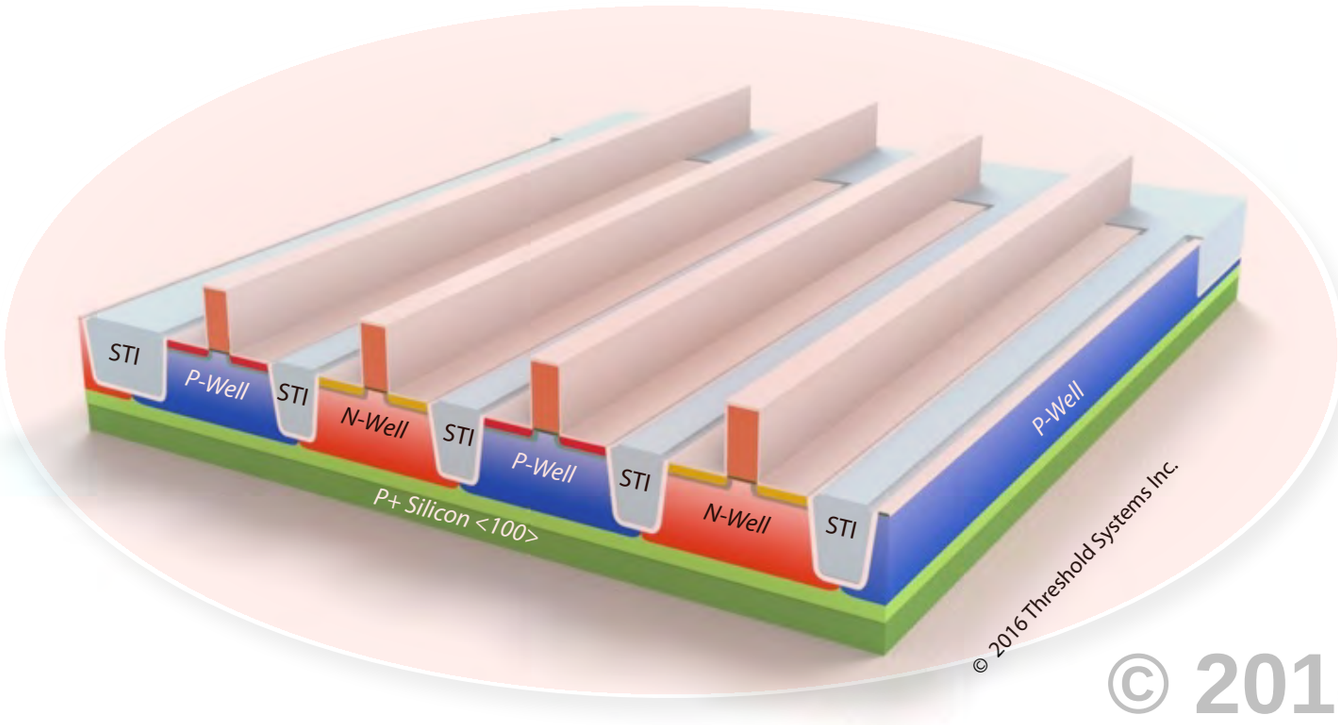


A Pre-Amorphizing Implant (PAI) of Germanium (5E14@4KeV, 0º) is applied to the N-Wells, and then a low energy Extension implant of (B, 8E14@300eV, 0º) is applied. This is followed by a Carbon implant of 5E14@3KeV which facilitates a superior Extension implant profile.

The PAI limits the channeling of the Boron atoms, while the Boron implant establishes the Extension implants of the PMOS transistors.

© 2018 by Threshold Systems Inc. All Rights Reserved.

PMOS Anneal:



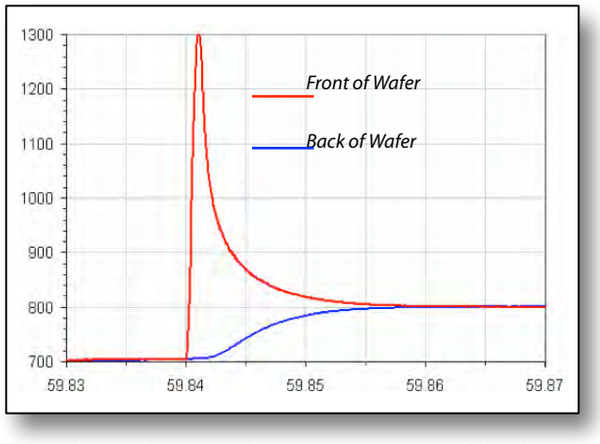
The photoresist from the previous operation is ashed away and the wafers are cleaned in Piranha.

The wafers are then undergo a Rapid Thermal Process (~950 degree C Spike for ~ 1 second followed by a ~1,350 degrees FLASH anneal for 1-3 milliseconds) that activates the PMOS Extension and Halo implants. In some process flows the Flash anneal is replaced by a sub-melt Laser anneal at the same temperature of ~1,350 degrees C.

The location of the Extension implant fixes the final Leff of the PMOS transistors.

The anneal makes the dopant atom electrically active and anneals out the damage done to the silicon during the implants and has a huge effect on Id_{SAT}.

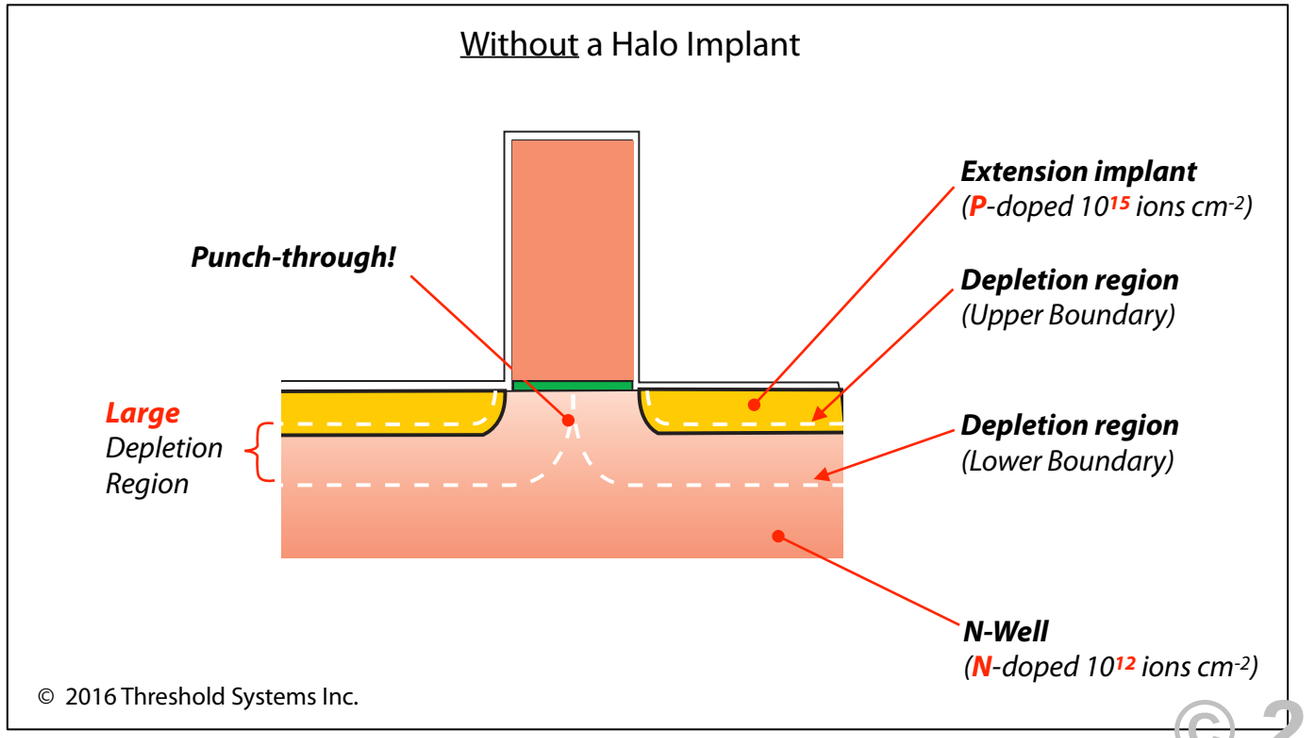
© 2018 Threshold Systems Inc.



RTP abruptly raises the temperature of the front side of the wafer for a fraction of a second without adding a lot of heat.

© 2018 by Threshold Systems Inc. All Rights Reserved.

The Extension and the Halo Implants:

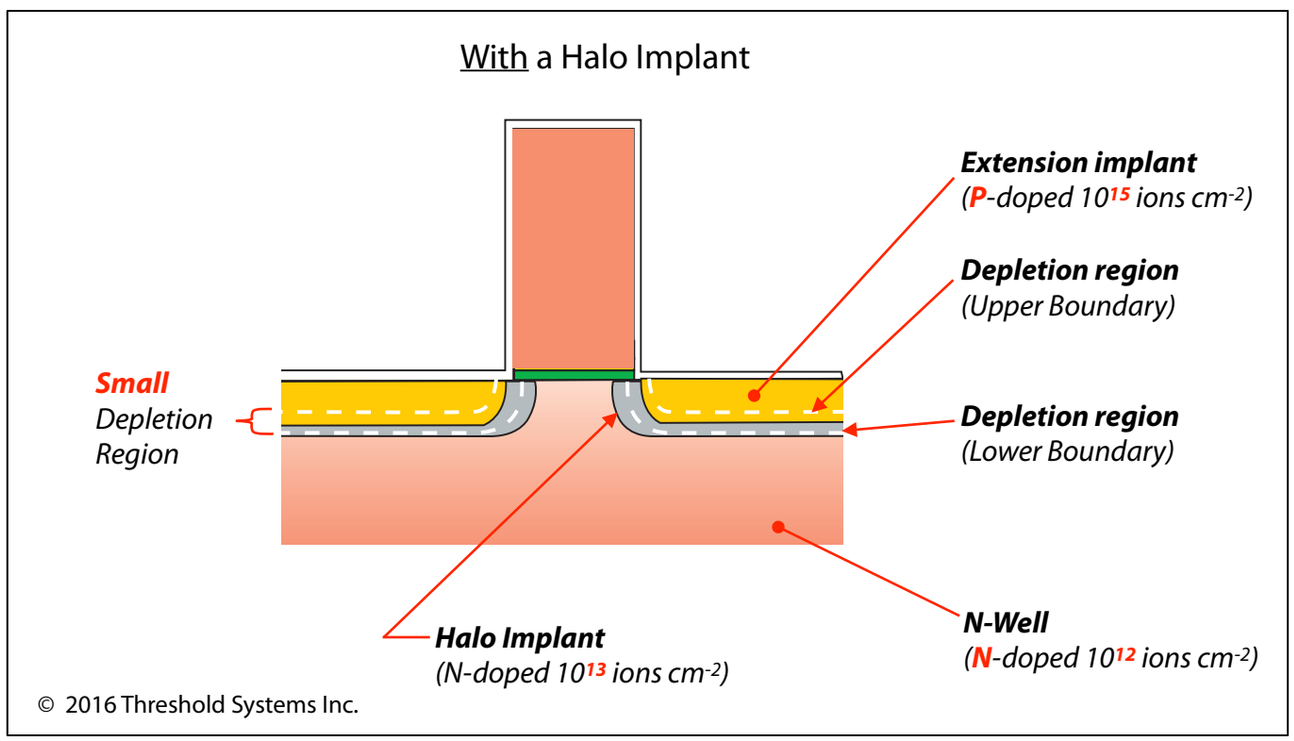


Along any PN junction there is always a "depletion region" which is a zone that is depleted of charged carriers. All of the positive charges along one side of the junction must be equal to all of the negative charges on the opposite side of the junction so that charge neutrality is maintained.

*When the PN junction consist of an Extension implant that is heavily doped (**P**-doped 10^{15} ions cm^{-2}) and a N-Well that is lightly doped (**N**-doped 10^{12} ions cm^{-2}) the depletion region will extend deeply into the lightly doped N-Well so that enough charge in the N-Well can be grabbed to neutralize all of the opposite polarity charge in the Extension implant.*

When the gate electrode is extremely narrow, the lower boundaries of the depletion regions from the Source and Drain of the transistor can touch, resulting in a soft breakdown condition known as "Punch-through". When this happens the transistor will be in the ON state even when it is supposed to be OFF.

© 2018 Threshold Systems Inc.

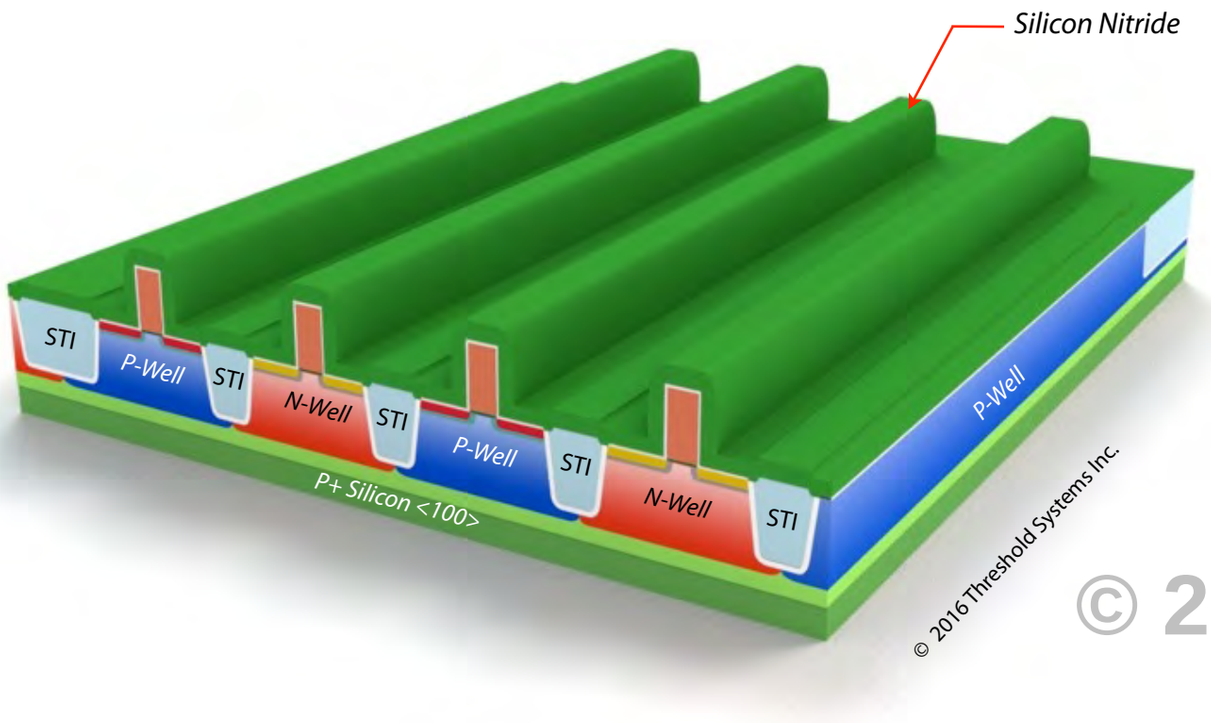


By implanting a N-type Halo implant (N-doped 10^{13} ions cm^{-2}) that is at a higher dose than the N-Well, more charge is supplied on the Well side of the PN junction to balance the opposite charge on the Extension side of the PN junction. This means that the lower boundary of the depletion region does not extend as deeply into the N-Well and the Punch-through condition is avoided.

The Halo implant extends only slightly beneath the edge of the gate electrode and is not located in the vast majority of the channel region beneath the gate electrode, so it does no lower carrier mobility.

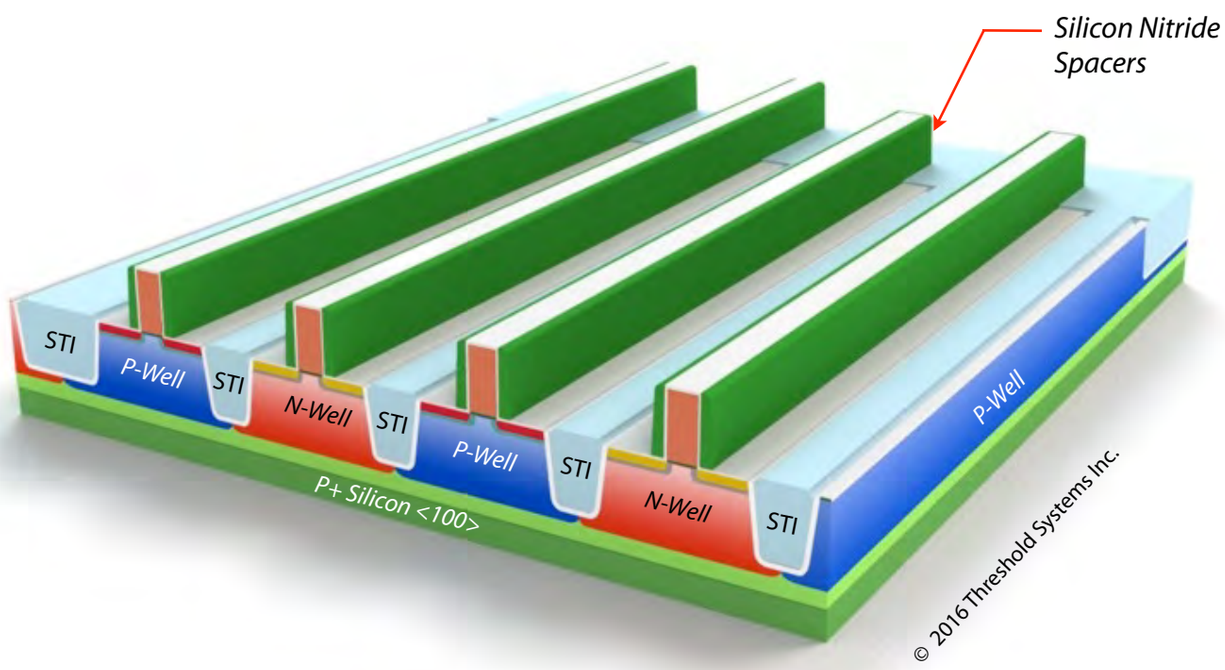
© 2018 by Threshold Systems Inc. All Rights Reserved.

Nitride Spacer Deposition and Etch:



A 150 Å thick blanket layer of silicon nitride is deposited over the surface of the wafers. Because of the conformal nature of silicon nitride, it is thickest along the side-walls of the gate electrodes.

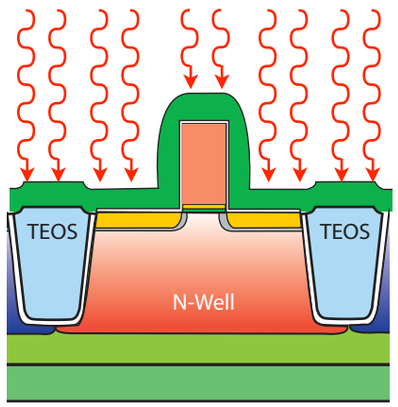
© 2018 Threshold Systems Inc.



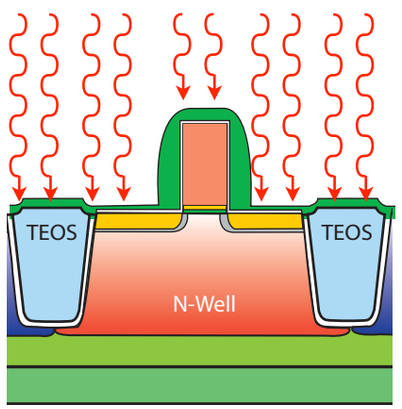
A highly anisotropic etch removes the silicon nitride from all of the horizontal surfaces, and but leaves nitride side-wall spacers along the sides of the gate electrodes where the nitride is thickest. The nitride side-wall spacers act as mini-implant masks during the contact implants that come later in the process.

© 2018 by Threshold Systems Inc. All Rights Reserved.

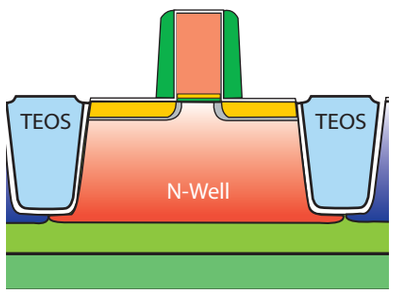
Nitride Spacer Etch:



© 2016 Threshold Systems Inc.



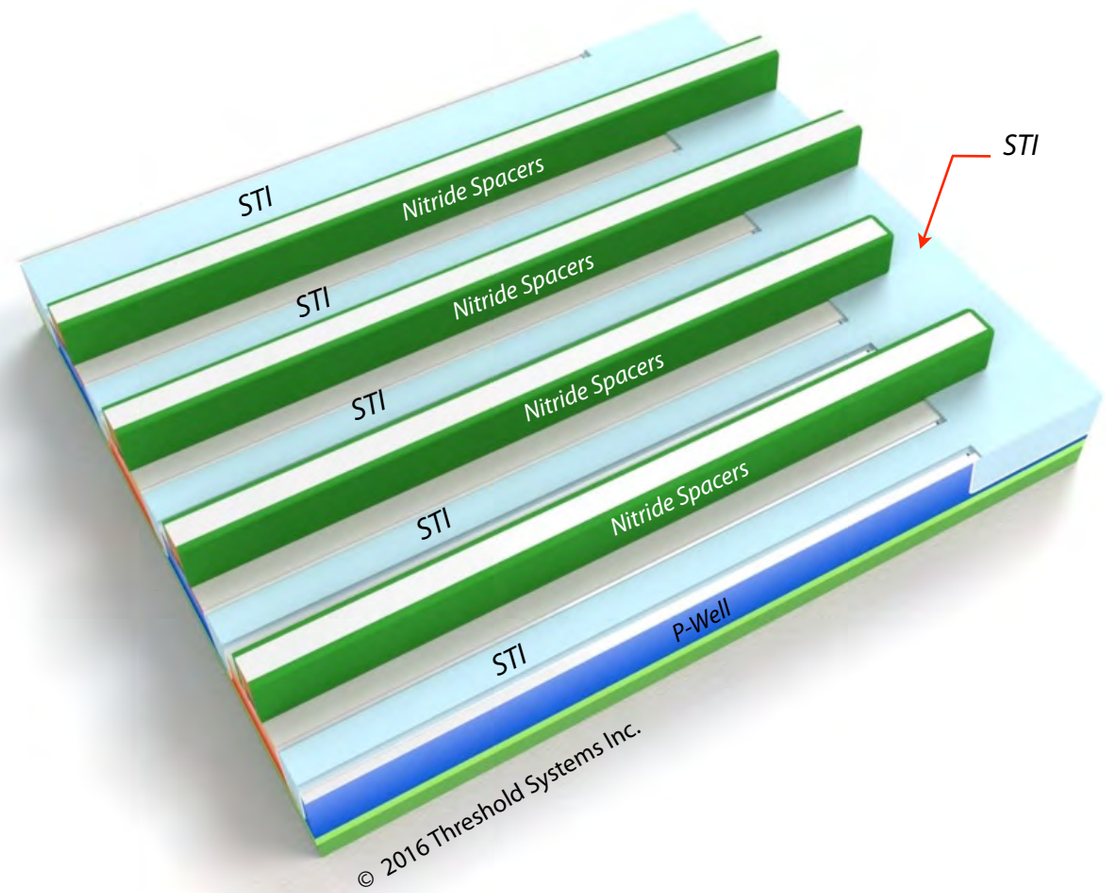
© 2016 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

Because the etch is highly anisotropic (etches in one direction only, i.e. downward) the silicon nitride is cleared everywhere except the on the sidewalls of the gate electrode because this is where the nitride is the thickest.

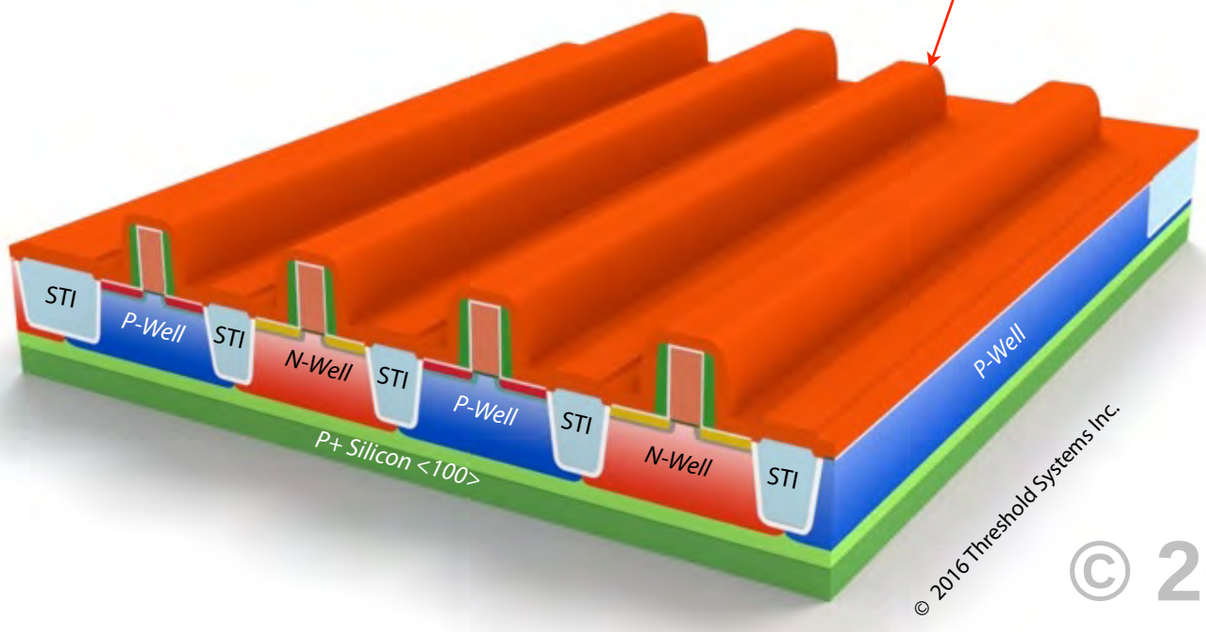
© 2018 Threshold Systems Inc.



A highly anisotropic etch removes the silicon nitride from all of the horizontal surfaces, and but leaves nitride side-wall spacers along the sides of the gate electrodes. The nitride side-wall spacers act as mini-implant masks during the contact implants that come later in the process.

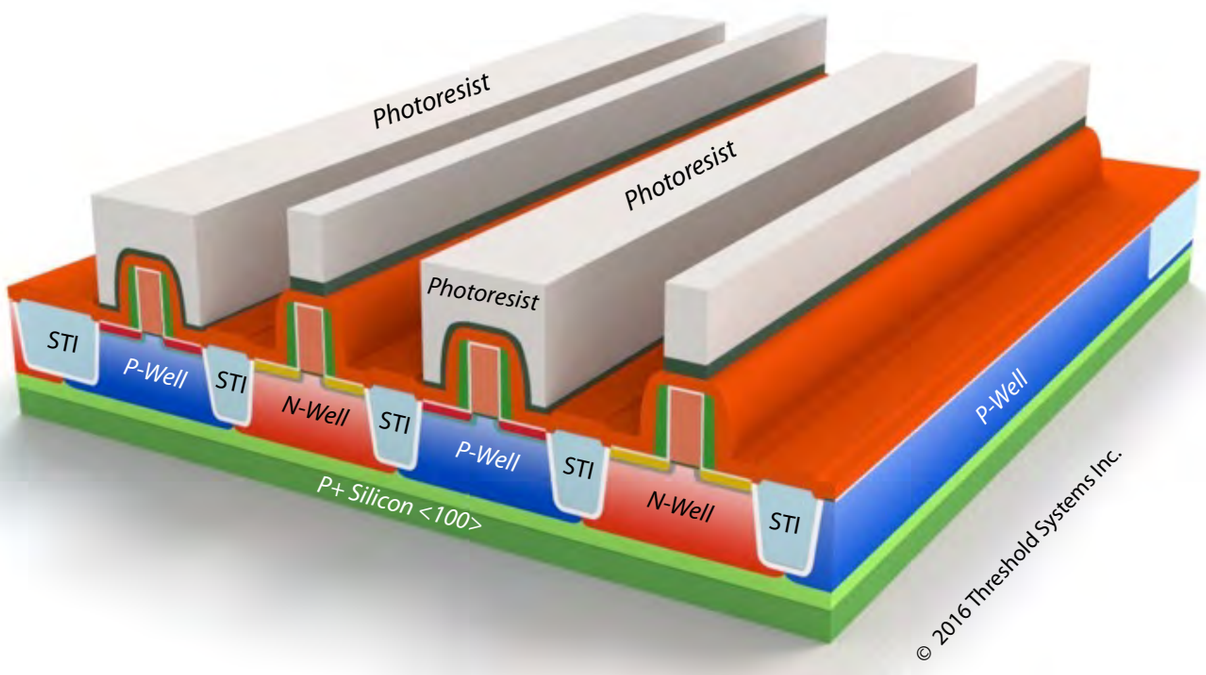
© 2018 by Threshold Systems Inc. All Rights Reserved.

PMOS Hard Mask Deposition & Patterning:



Next, 300Å of SiCN is deposited across the surface of the wafer.

© 2018 Threshold Systems Inc.

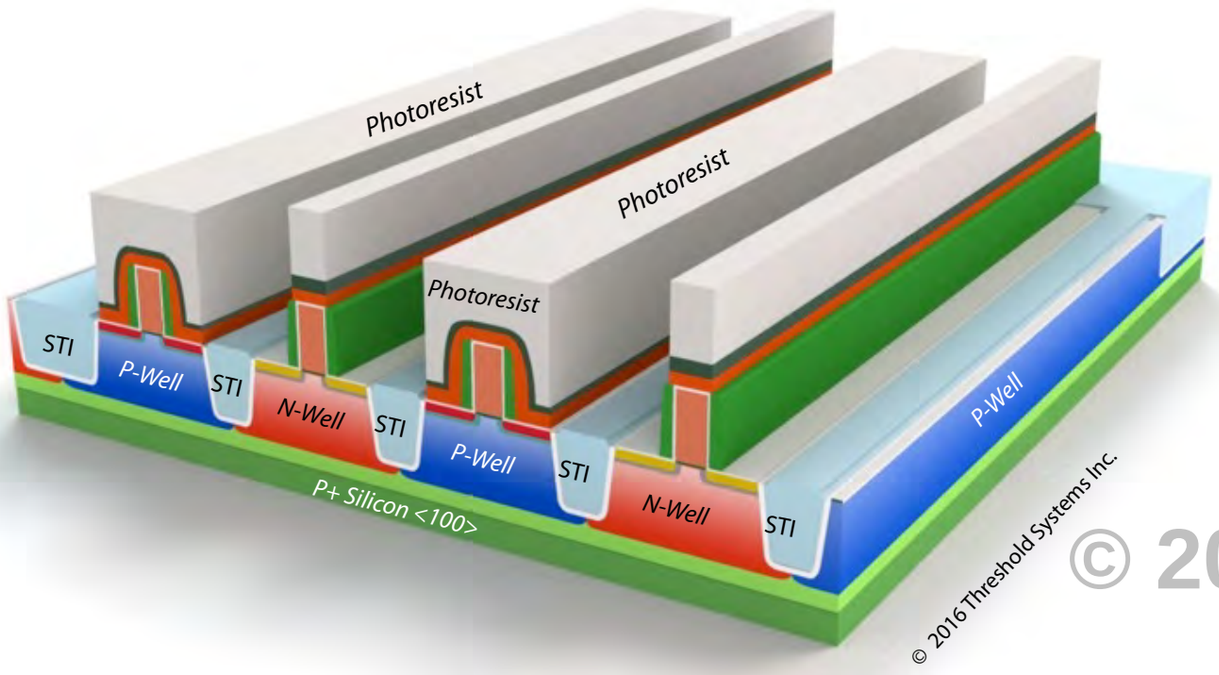


The SiCN is patterned with photoresist.

© 2018 by Threshold Systems Inc. All Rights Reserved.

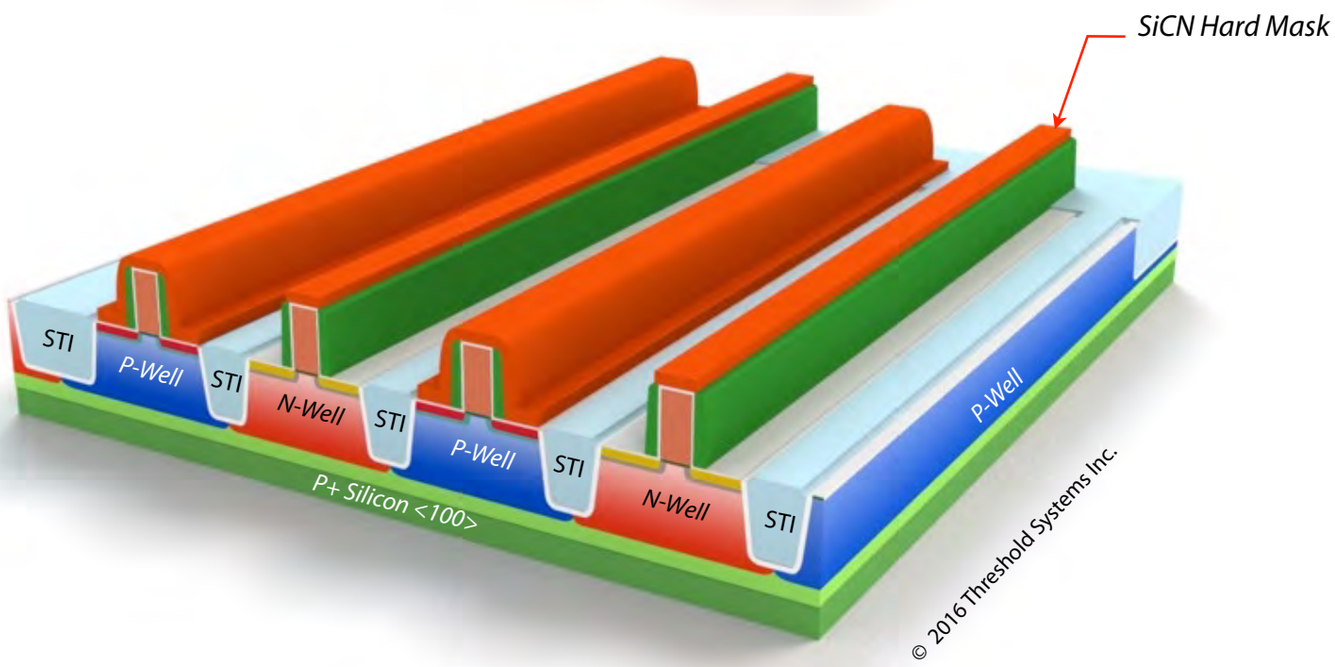
PMOS Hard Mask Etch & Photoresist Strip:

ThresholdSystems



The exposed SiCN is etched away in a highly anisotropic etch.

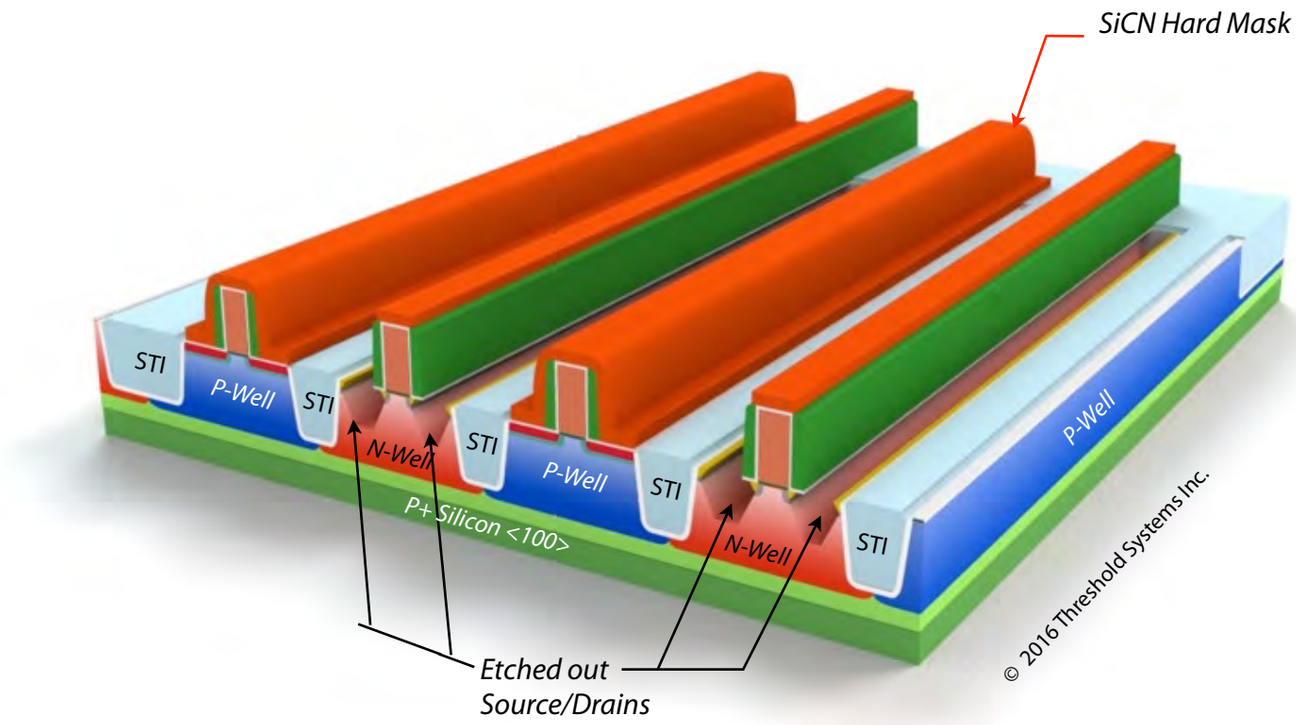
© 2018 Threshold Systems Inc.



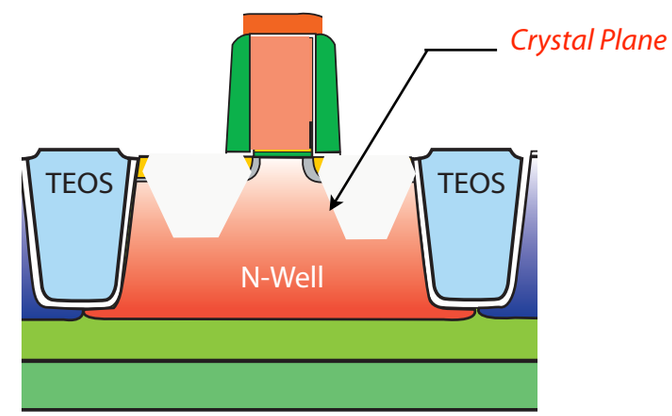
The photoresist is stripped in an ionized oxygen plasma and the wafers are cleaned in Piranha.

© 2018 by Threshold Systems Inc. All Rights Reserved.

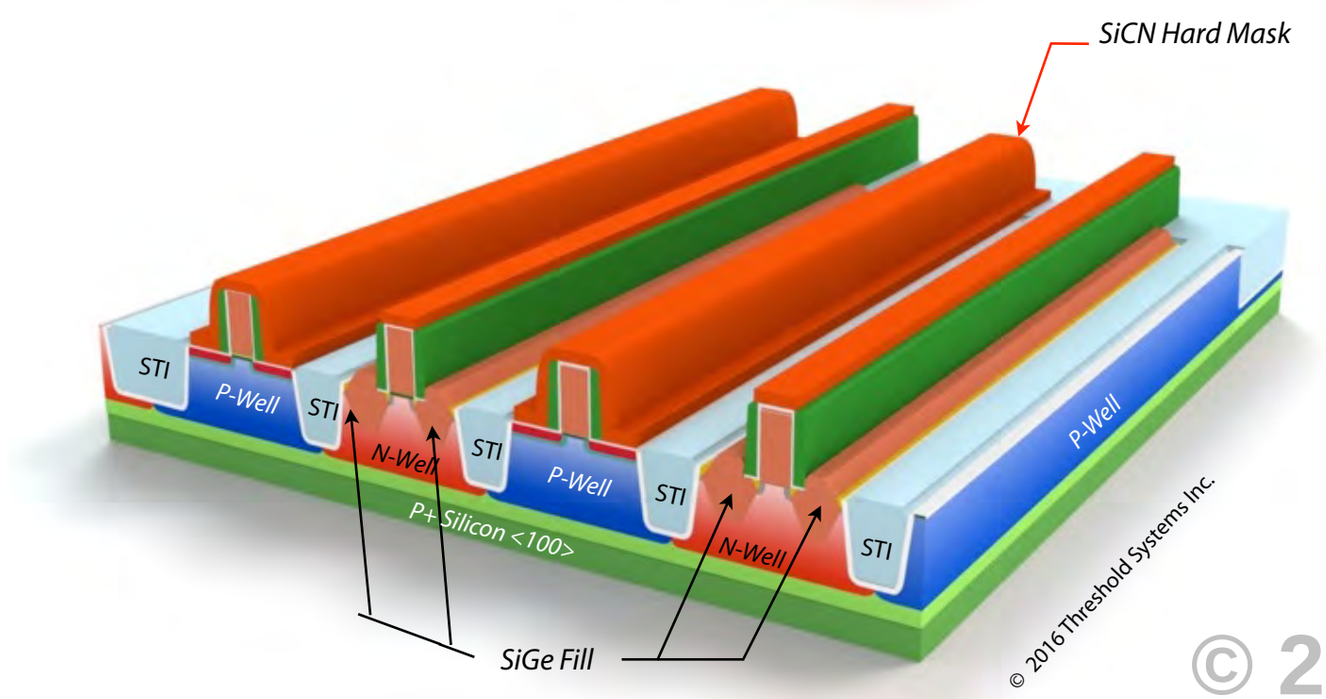
PMOS Source/Drain Etch & SiGe Fill:



The exposed Source/Drains of the PMOS transistor are wet etched away using a NH_4OH solution that etches along crystallographic planes. The P-Wells and the gate electrodes are protected by the SiCN. The etch will not attack the exposed oxide in the STI trenches.



© 2016 Threshold Systems Inc.



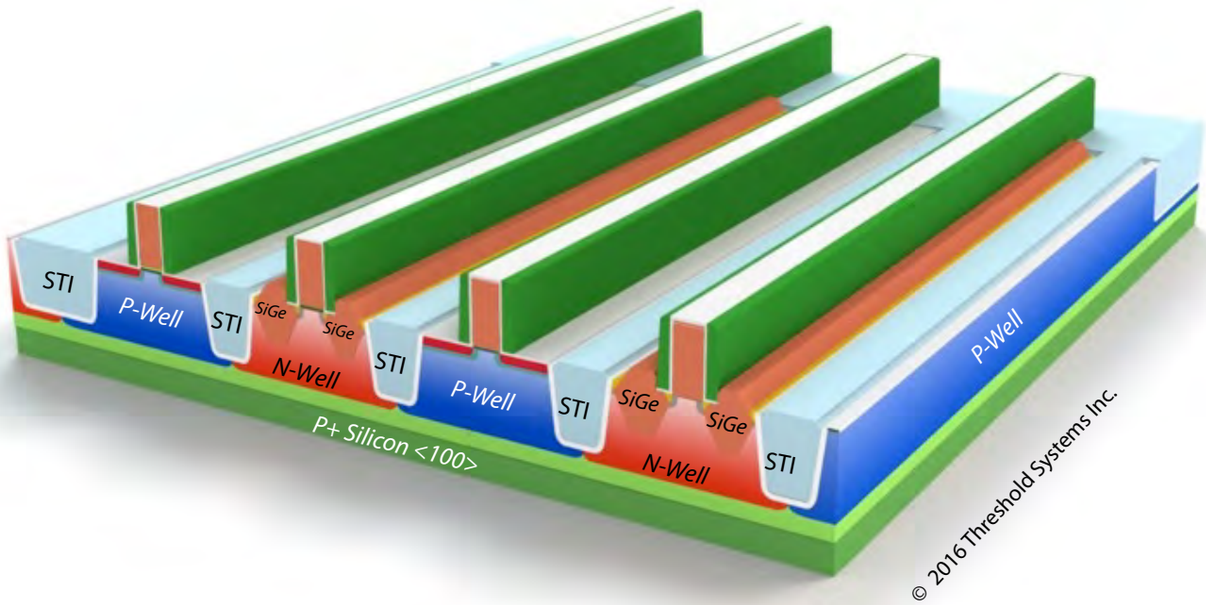
The wafers then undergo a selective epitaxial deposition of SiGe. The SiGe will only nucleate on an exposed silicon surface which means that it will only form in the PMOS Source/Drain regions (every other part of the wafer consists of nitride or oxide, or is covered in SiCN). The SiGe is ~ 40% Ge and it will exert a high level of compressive strain on the PMOS channel, greatly increasing the mobility of the holes in the channel and therefore increase Idsat.

As the Epitaxial SiGe is deposited it is heavily doped with Boron to make it P-Type silicon.

© 2018 Threshold Systems Inc.

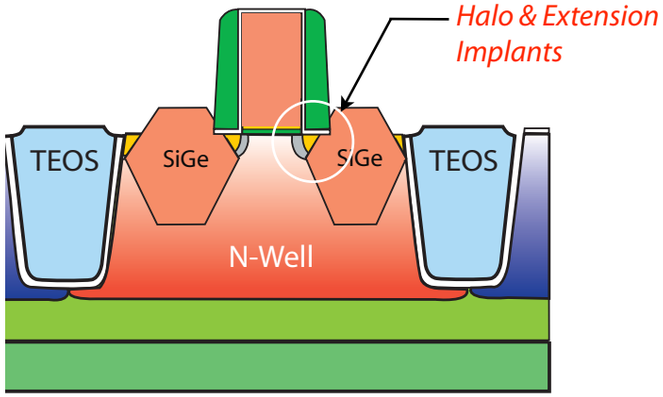
© 2018 by Threshold Systems Inc. All Rights Reserved.

Hard Mask Strip:

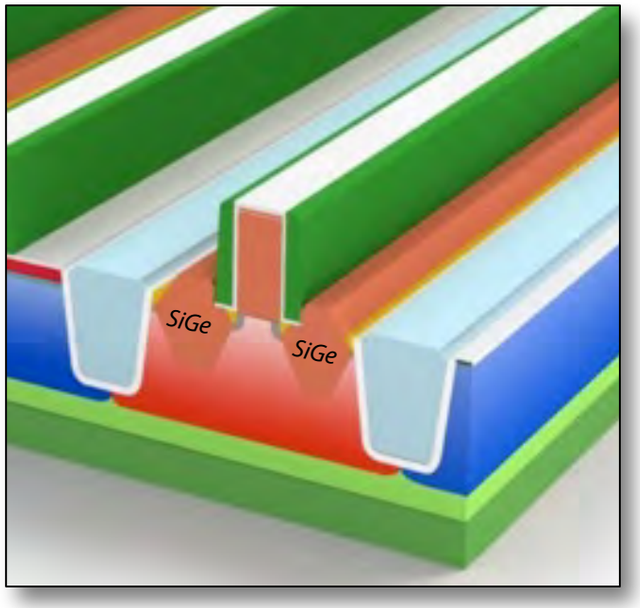


© 2016 Threshold Systems Inc.

The remaining SiCN hard mask is etched away in an isotropic etch.



© 2016 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

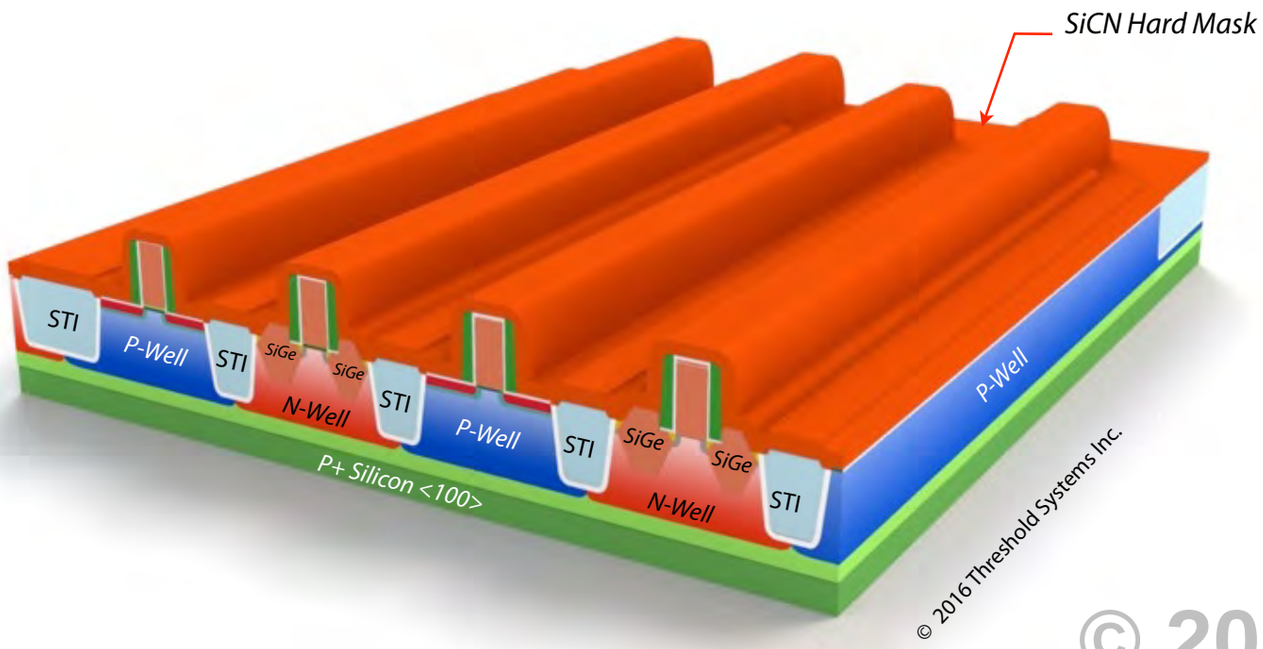
The adjacent image illustrates a close-up view of the SiGe replacement Source Drains for the PMOS devices. Note that even though the Source/ Drains have been replaced the Extension and Halo implants are still present.

© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

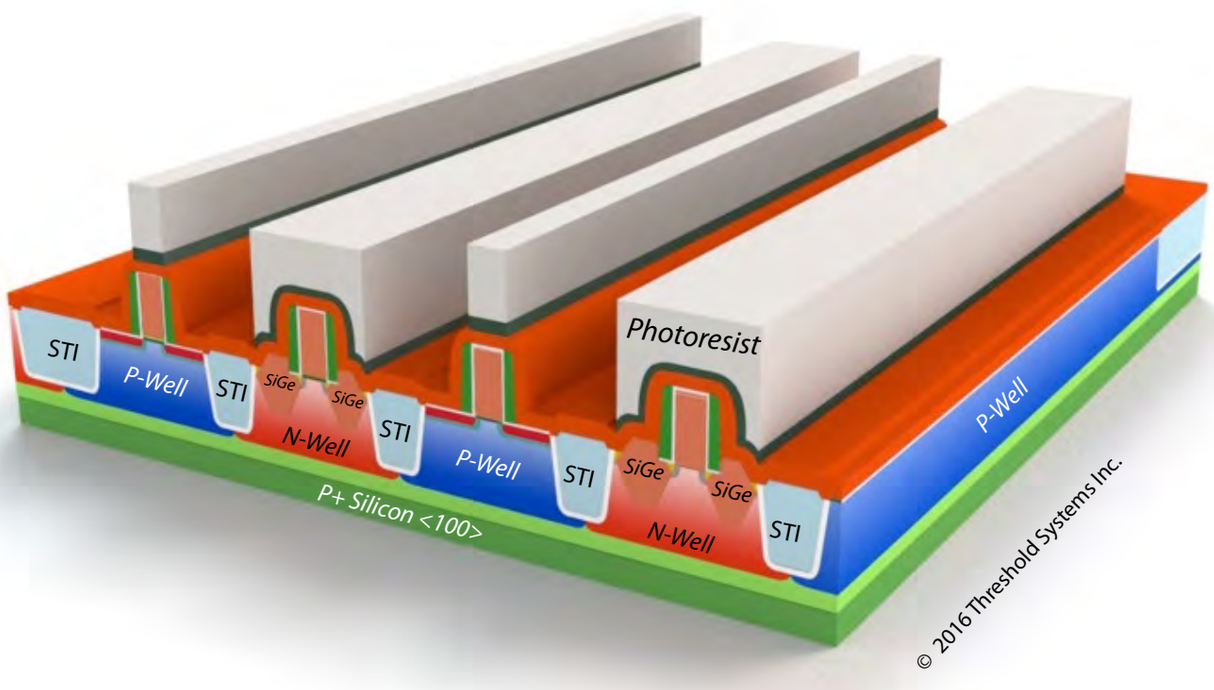
NMOS Hard Mask Deposition & Patterning:

ThresholdSystems



Next, 300Å of SiCN is deposited across the surface of the wafer.

© 2018 Threshold Systems Inc.

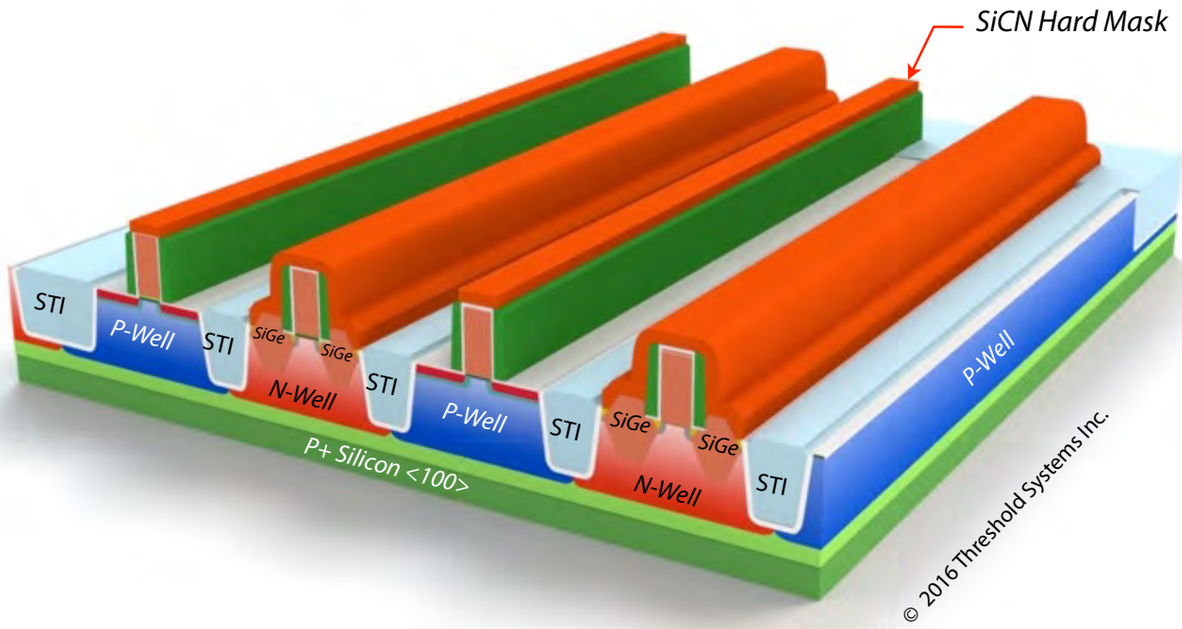


The SiCN is patterned with photoresist.

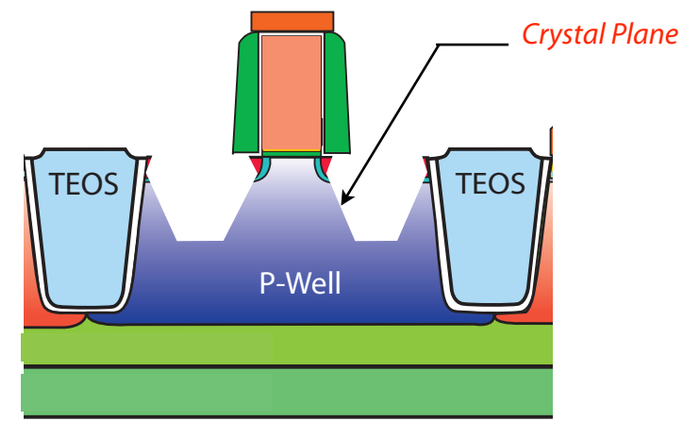
© 2018 by Threshold Systems Inc. All Rights Reserved.

NMOS Hard Mask Etch + Source/Drain Etch:

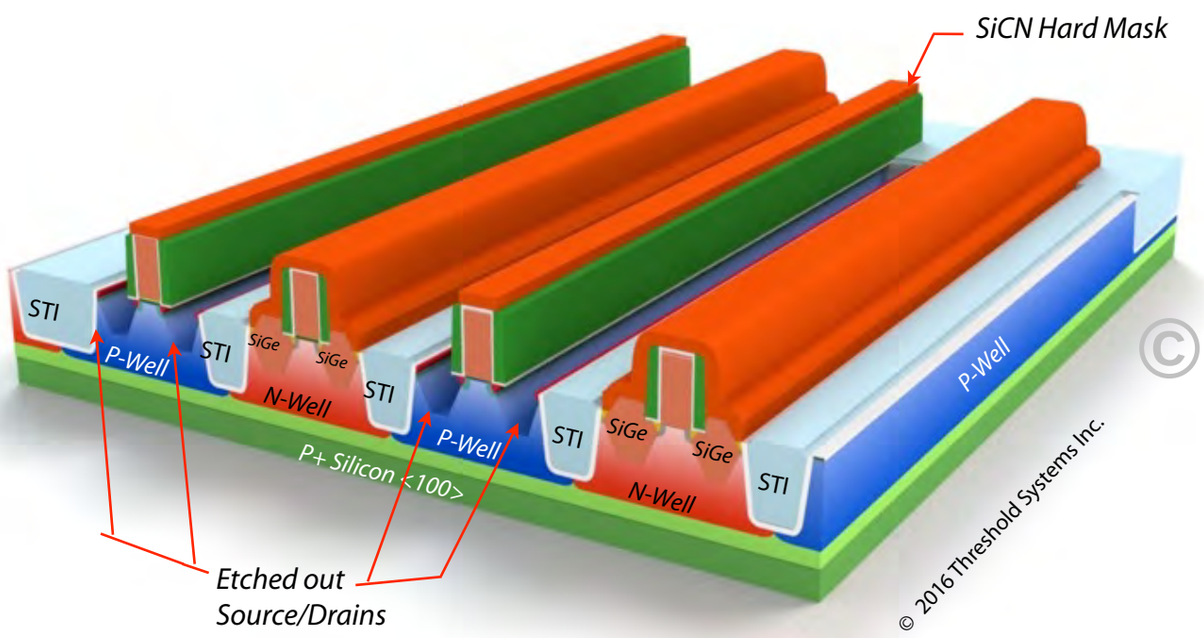
ThresholdSystems



The photoresist is stripped.



© 2016 Threshold Systems Inc.



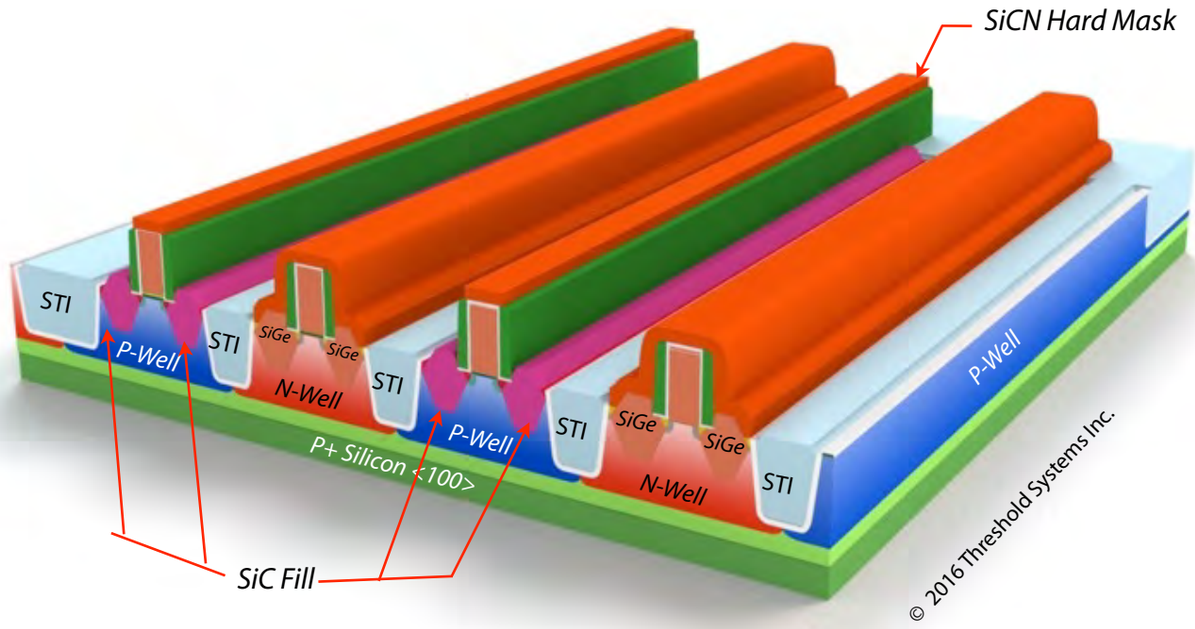
The oxide over the Source/Drains is etched away using a HF wet etch.

The exposed Source/Drains of the PMOS transistor are then wet etched away using a NH_4OH solution that etches along crystallographic planes.

© 2018 Threshold Systems Inc.

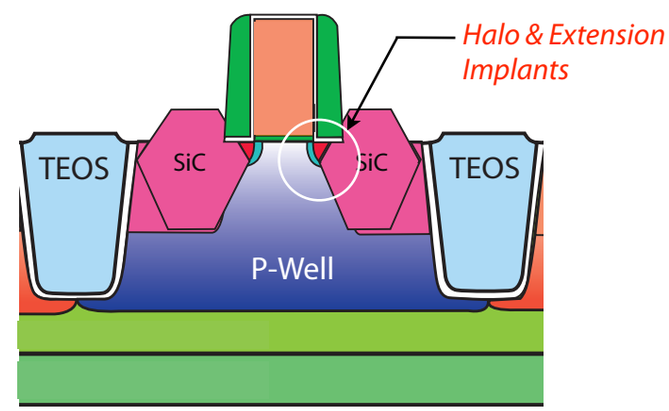
© 2018 by Threshold Systems Inc. All Rights Reserved.

NMOS Source/Drain Selective SiC Dep:



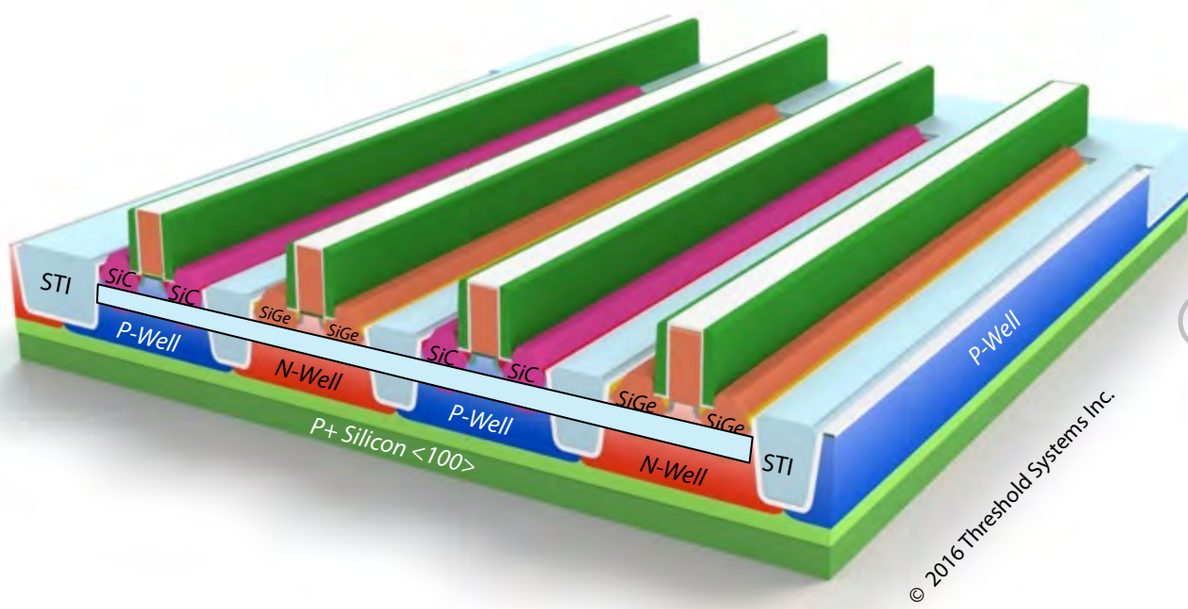
The wafers then undergo a selective epitaxial deposition of SiC. The SiC will only nucleate on an exposed silicon surface which means that it will only form in the PMOS Source/Drain regions (every other part of the wafer consists of nitride or oxide, or is covered in SiCN).

The SiC exerts a very strong tensile strain on the NMOS channel and enhances the mobility of electrons. This increases the Idsat of the NMOS devices. As the SiC is deposited it is doped with Phosphorus to make it N-Type.



© 2016 Threshold Systems Inc.

The remaining SiCN hard mask is etched away in an isotropic etch.

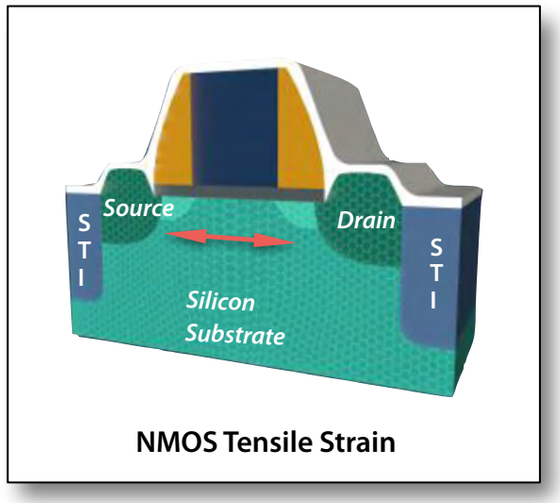


© 2018 Threshold Systems Inc.

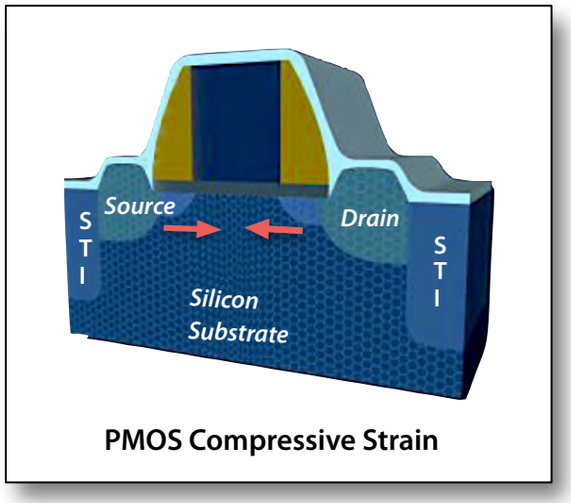
© 2018 by Threshold Systems Inc. All Rights Reserved.

Straining Silicon:

- *Appropriately straining the transistor channel dramatically increase the transistor drive current, and with it, the transistor speed.*
- *The PMOS transistor channel must be placed under compressive strain and the NMOS channel under tensile strain to achieve this effect.*
- *There are several approaches to inducing compressive or tensile strain into a transistor channel, but the most effective way is to replace the transistor source/Drains with a Silicon epitaxy that is doped with atoms that have a larger or smaller lattice constant. This induces compressive or tensile strain, respectively.*

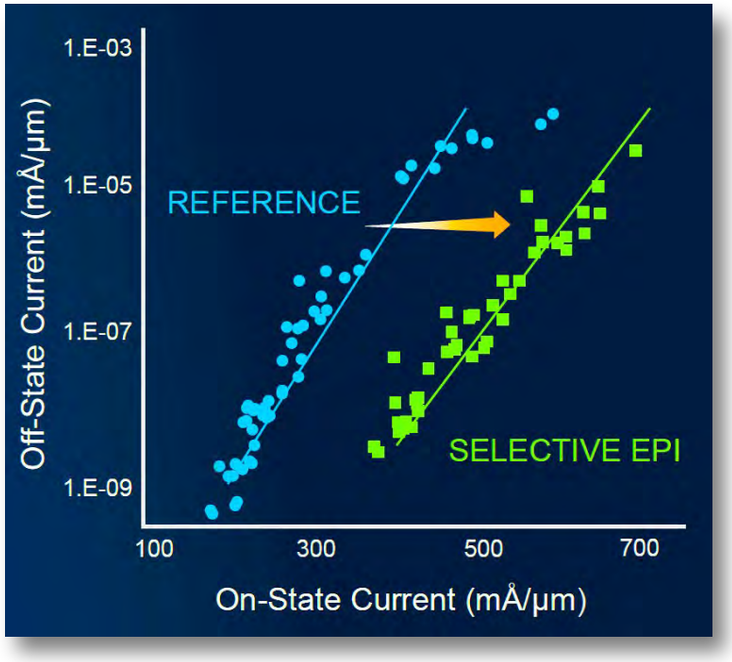


Source: Schubert Chu, Applied Materials 2013



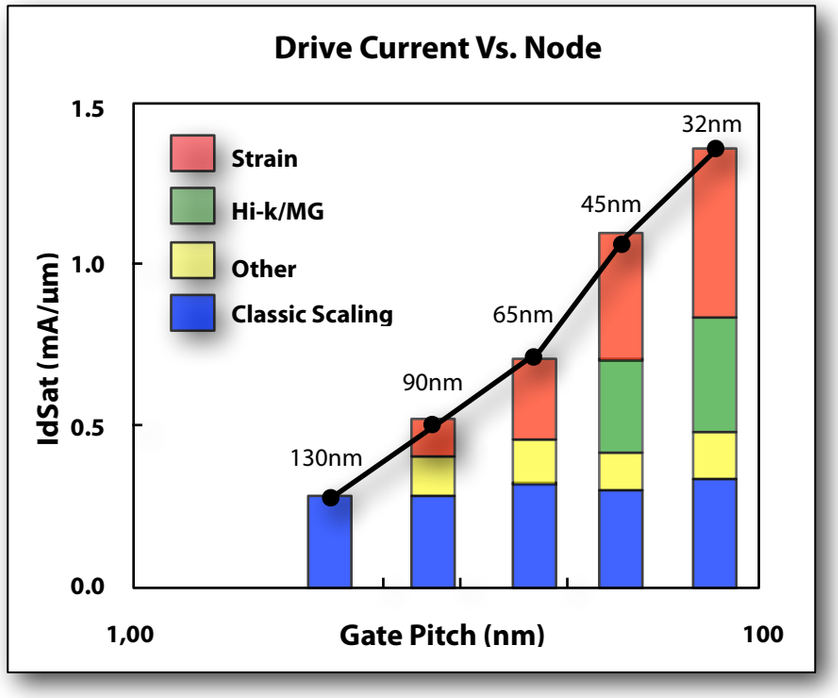
Source: Schubert Chu, Applied Materials 2013

© 2018 Threshold Systems Inc.



Source: Schubert Chu, Applied Materials 2013

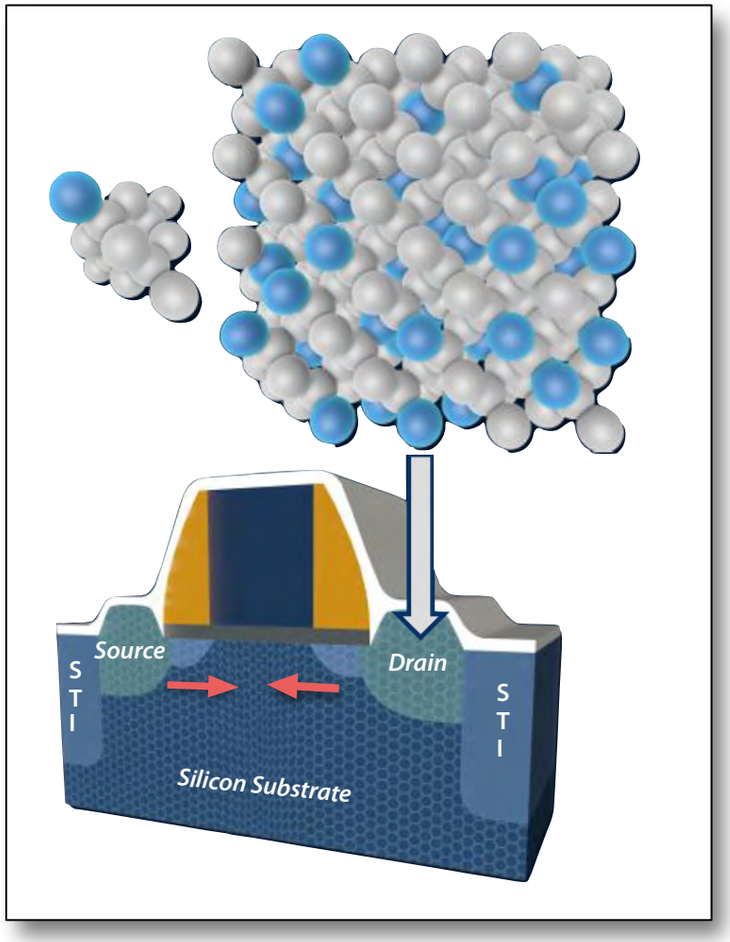
> 60% Improvement in Drive Current!



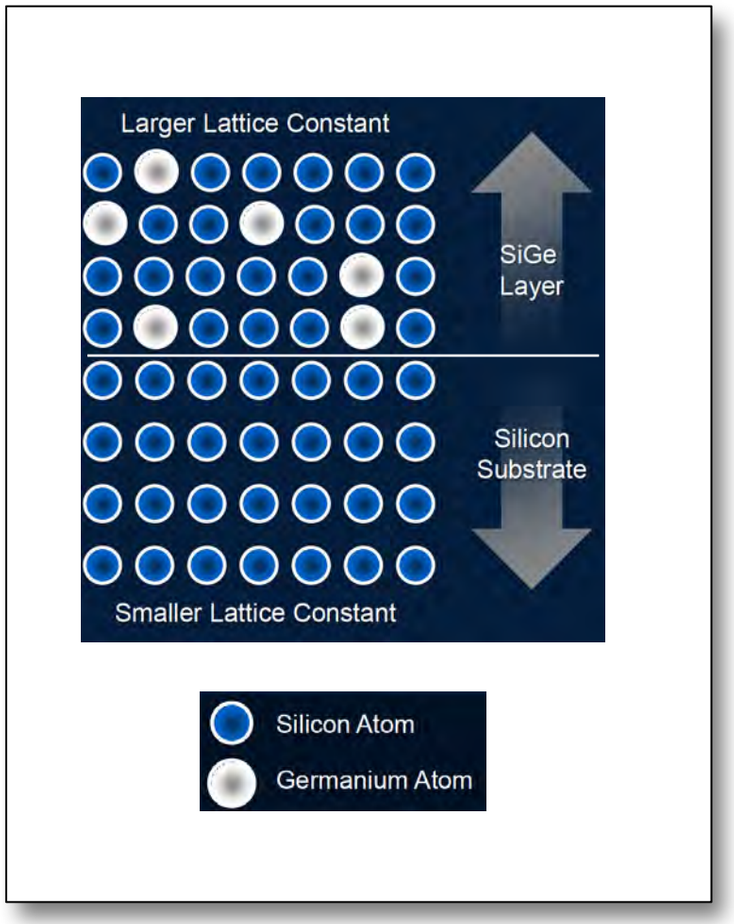
After K.Kahn - Intel

© 2018 by Threshold Systems Inc. All Rights Reserved.

Straining the PMOS Transistor:



Source: Schubert Chu, Applied Materials 2013



Source: Schubert Chu, Applied Materials 2013

- *The Source/Drains are selectively etched away and replaced with Silicon/Germanium. Germanium has a lattice constant that is 4% larger than that of Silicon, and this larger lattice constant will ensure that the transistor channel is compressed, enhancing the transistor drive current.*

© 2018 Threshold Systems Inc.

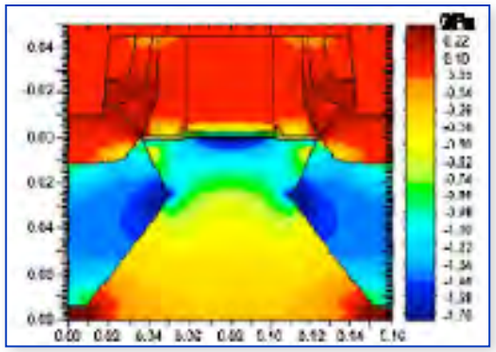
© 2018 by Threshold Systems Inc. All Rights Reserved.

Optimizing PMOS Stress:

ThresholdSystems

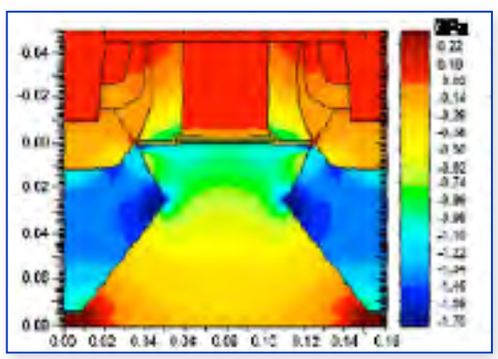
- *As previously mentioned, the amount of Ge in the PMOS SiGe Source/Drains is increased at each new node to maintain the compressive stress on the transistor channel (currently 55% Ge at the 20nm node).*
- *Another technique is to move the pointed edge of the SiGe crystal closer to the channel to induce more stress. This was done at 32nm (refer to the accompanying TEMS).*
- *The replacement gate also offers another method to increase PMOS strain. When the polysilicon gate electrode is removed the spacers and associated structures move inward, compressing the transistor channel and inducing compressive stress.*

45nm PMOS Replacement Gate Stress Map



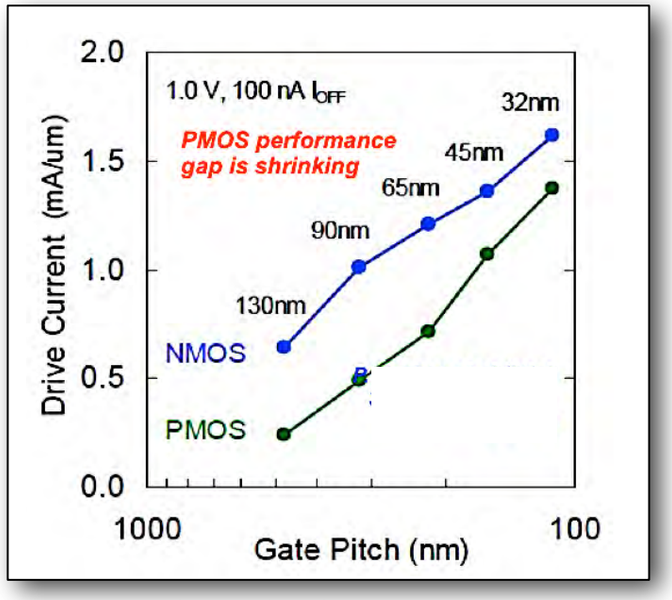
Before Gate Removal
(ambient compressive stress)

Source: C. Auth et al - Intel



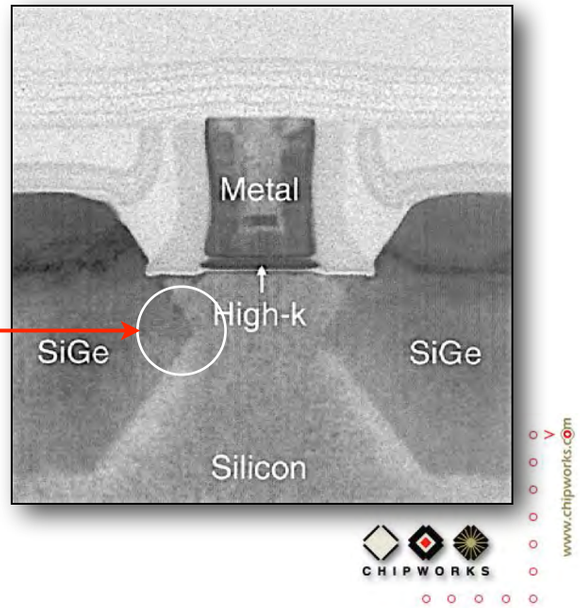
After Gate Removal
(increased compressive stress)

Source: C. Auth et al - Intel



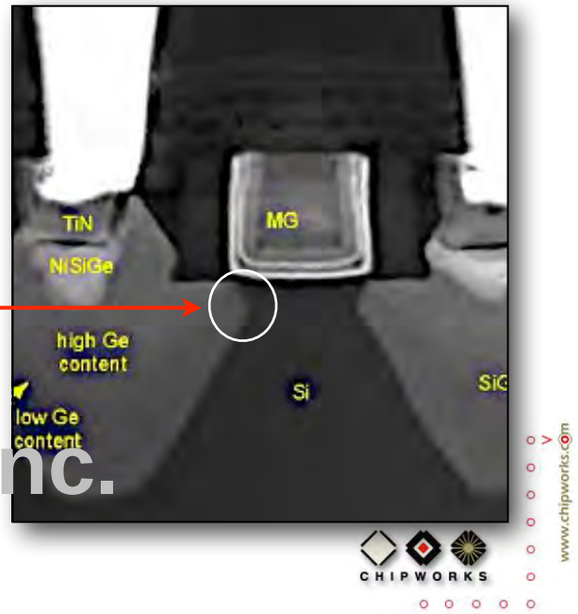
Source: P. Packan et al Intel

45nm PMOS



At 32nm The foot of the SiGe was moved closer to the channel to induce more stress.

32nm PMOS



© 2018 Threshold Systems Inc.

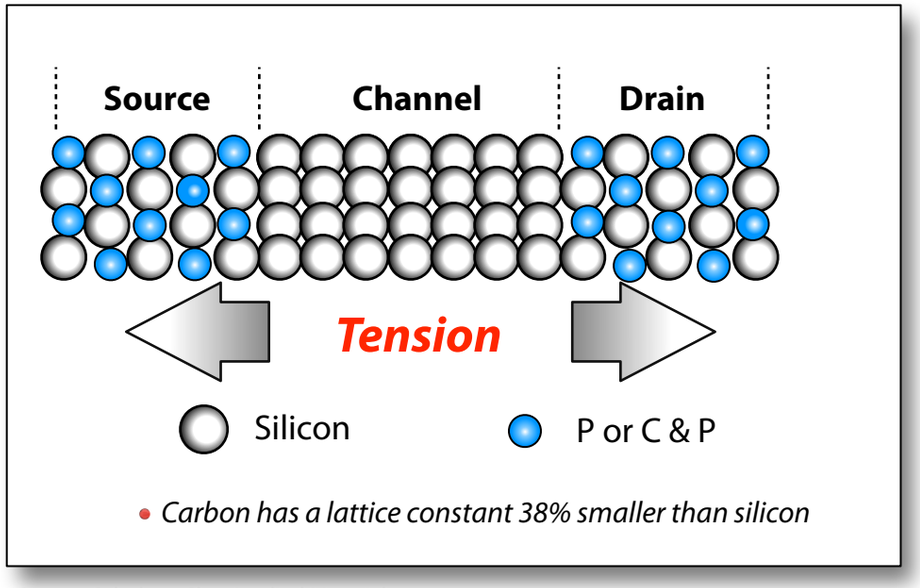
© 2018 by Threshold Systems Inc. All Rights Reserved.

Straining the NMOS Transistor:

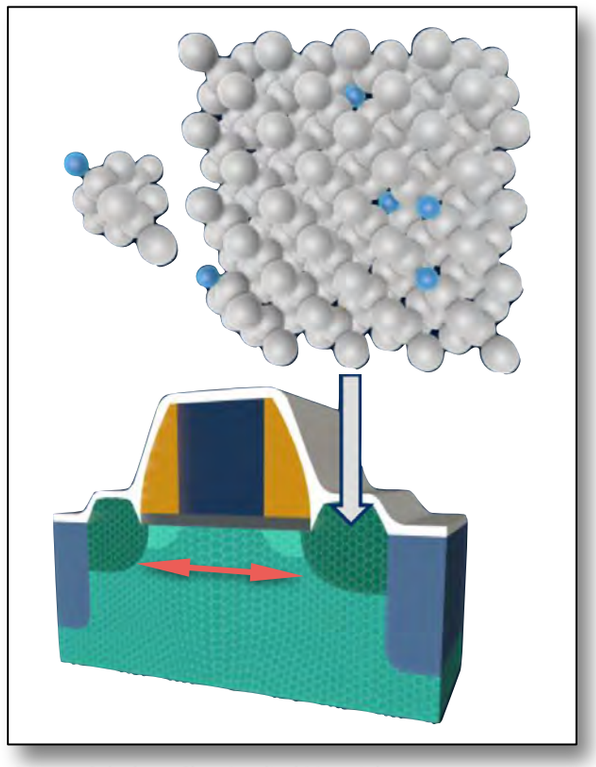
ThresholdSystems

- *Straining the NMOS transistor with the replacement Source/Drain methodology is considerable more problematic due to the difficulty in finding a material with a smaller lattice constant than silicon that also has a high degree of stability. For this reason NMOS tensile strain has typically been implemented using techniques other than replacement Source/Drain.*
- *However, recently a technology has been perfected that incorporates either Phosphorus, or Carbon & Phosphorus, into a Silicon epitaxy. Because Carbon is 38% smaller than Silicon, Source/Drains that are replaced with this epitaxy pull on either side of the channel and create a tensile strain on the channel. This is precisely what is required to enhance the performance of the transistor (up to 20% more drive current).*
- *Either a Phosphorus or a Carbon/Phosphorus epitaxy can be used for this purpose.*

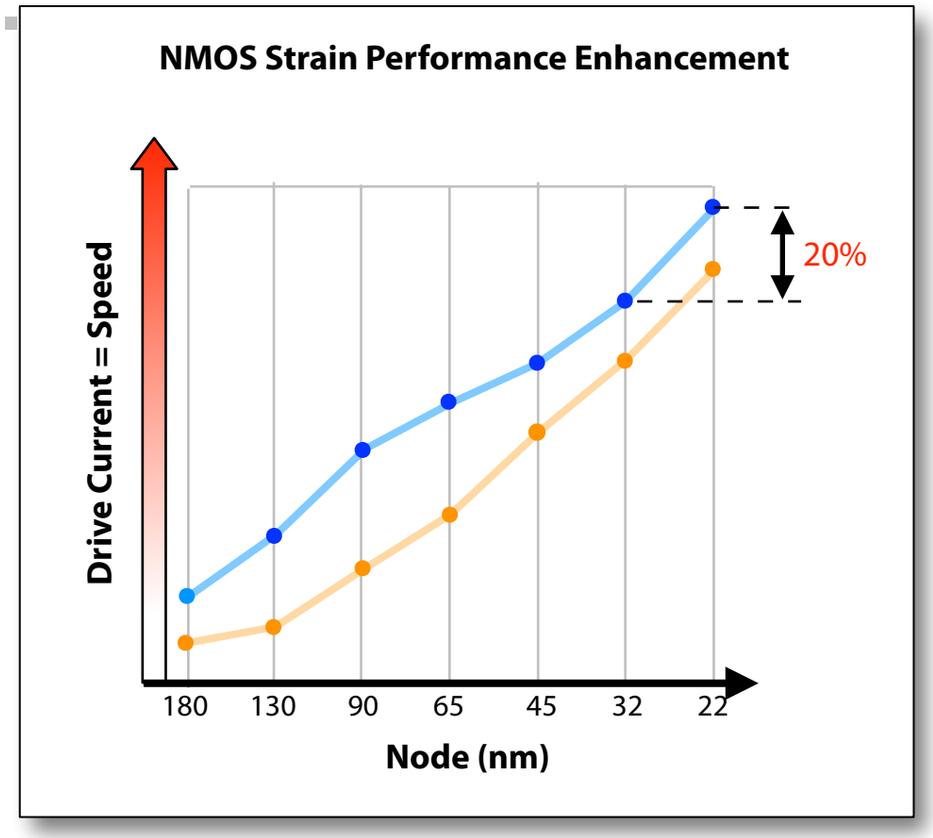
© 2018 Threshold Systems Inc.



Source: Schubert Chu, Applied Materials 2013



Source: Schubert Chu, Applied Materials 2013



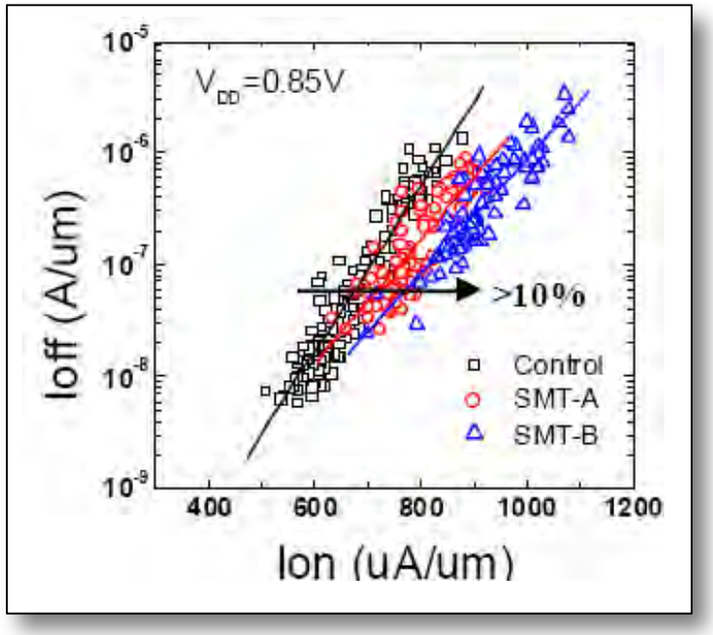
Source: Schubert Chu, Applied Materials 2013

© 2018 by Threshold Systems Inc. All Rights Reserved.

Mask-Edge Dislocation Stress (SMT):

ThresholdSystems

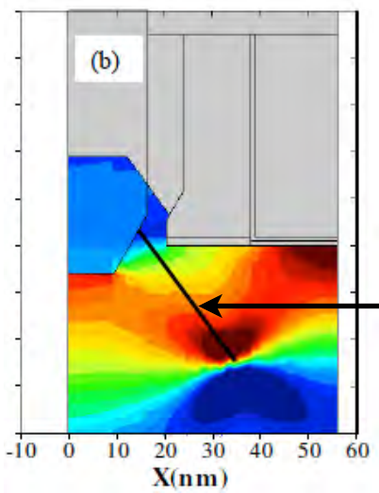
- *A more specific name for Stress Memorization Technique (STM) on the NMOS device is a technique known as "Edge Dislocation Stress". This is a STM technique that generates a mask-edge dislocation inducing dramatic tensile stress in the NMOS channel. It was first deployed at the 45/32nm nodes prior to the advent of P or P&C replacement Source/Drain epitaxy for NMOS transistors.*
- *The process involves a deep amorphization implant followed by the deposition of a stress over-liner of nitride. Following a Rapid Thermal Anneal to induce SMT, the stress liner is removed and the mask-edge dislocations are formed.*
- *The initial pre-amorphization implant (PAI) for SMT creates multiple mask-edge dislocations under the Source/Drain region which enhances short-channel mobility. Overall, a 10% increase in Idsat is achieved.*
- *It is important to realize that different strain techniques are typically used together because their contributions to increased performance is additive.*



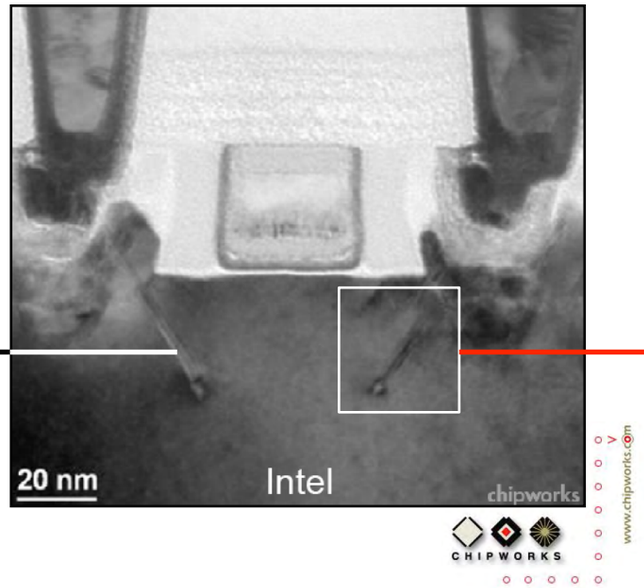
Source: K.Y. Kim et al - Samsung

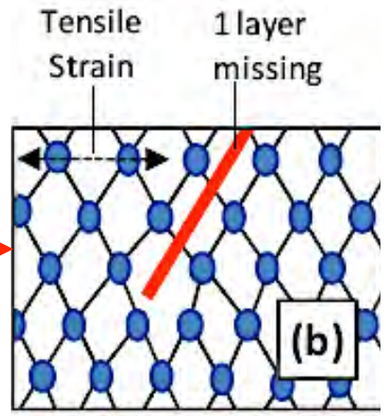
© 2018 Threshold Systems Inc.

Mask Edge Dislocation



Source: C.E. Weber et al Intel

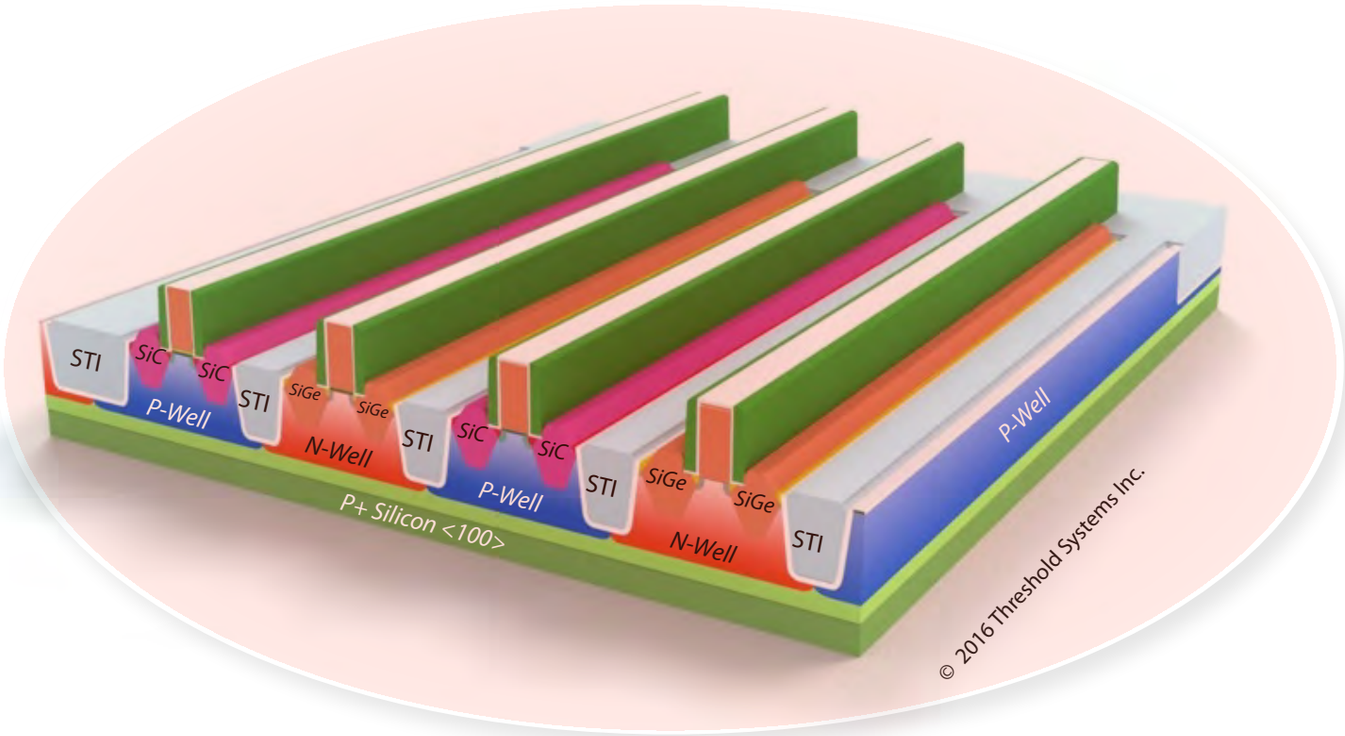




Source: K.Y. Kim et al - Samsung

© 2018 by Threshold Systems Inc. All Rights Reserved.

Source /Drain Anneal:



The wafers are then undergo a Rapid Thermal Process (~950 degree C Spike for ~ 1 second followed by a ~1,350 degrees FLASH anneal for ~1-3 milliseconds) that activates the PMOS and NMOS dopant atoms.

© 2018 Threshold Systems Inc.

In some process flows the Flash anneal is replaced by a sub-melt Laser anneal at the same temperature of ~1,350 degrees C.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Physical Vapor Deposition (PVD):

ThresholdSystems

- *In microelectronic fabrication it is often necessary to deposit thin films of ultra-pure metals or metal alloys such as Nickel, Titanium, Titanium Nitride, Tantalum, Tantalum Nitride and Copper.*
- *This deposition process is accomplished using a technology known as Physical Vapor Deposition (PVD).*
- *This task is accomplished by sputtering metal atoms off of a metal target and allowing them to be sprayed onto the wafer's surface. The atoms of metal adhere on the surface and eventually form a thin film that covers the entire surface of the wafer.*
- *PVD is a valuable technology because it is capable of depositing thin layers of very pure metals, or metal alloys, with good step coverage, in a very short time frame.*

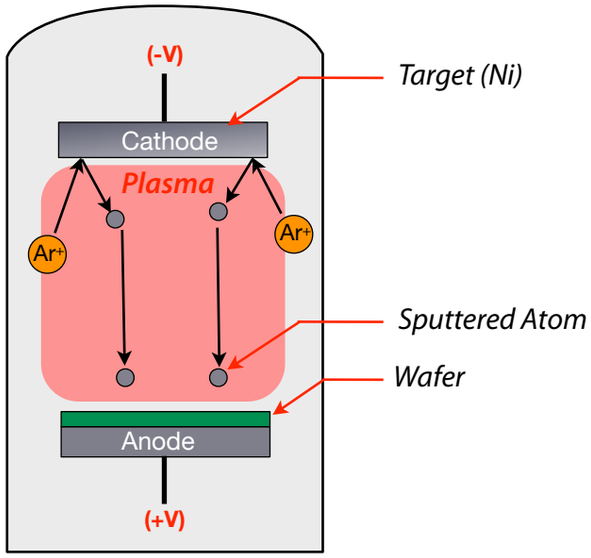
© 2018 Threshold Systems Inc.

© Applied Materials

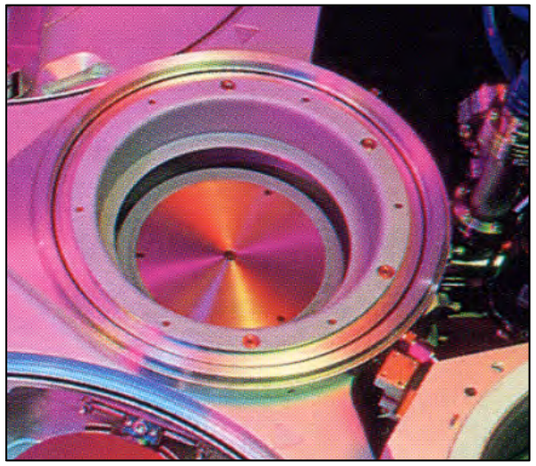
An opened PVD tool showing all of the chambers.

The wafer is placed in an evacuated chamber on top of the anode and a small amount of Argon gas is introduced into the chamber. A plasma is struck and the Argon is ionized. The positively charged Argon ions then bombard the Cathode which consists of the target material, and in so doing, sputter off metal that is deposited onto the wafer.

This thin and very pure coating of metal is rapidly deposited onto the wafer's surface.



© 2016 Threshold Systems Inc.



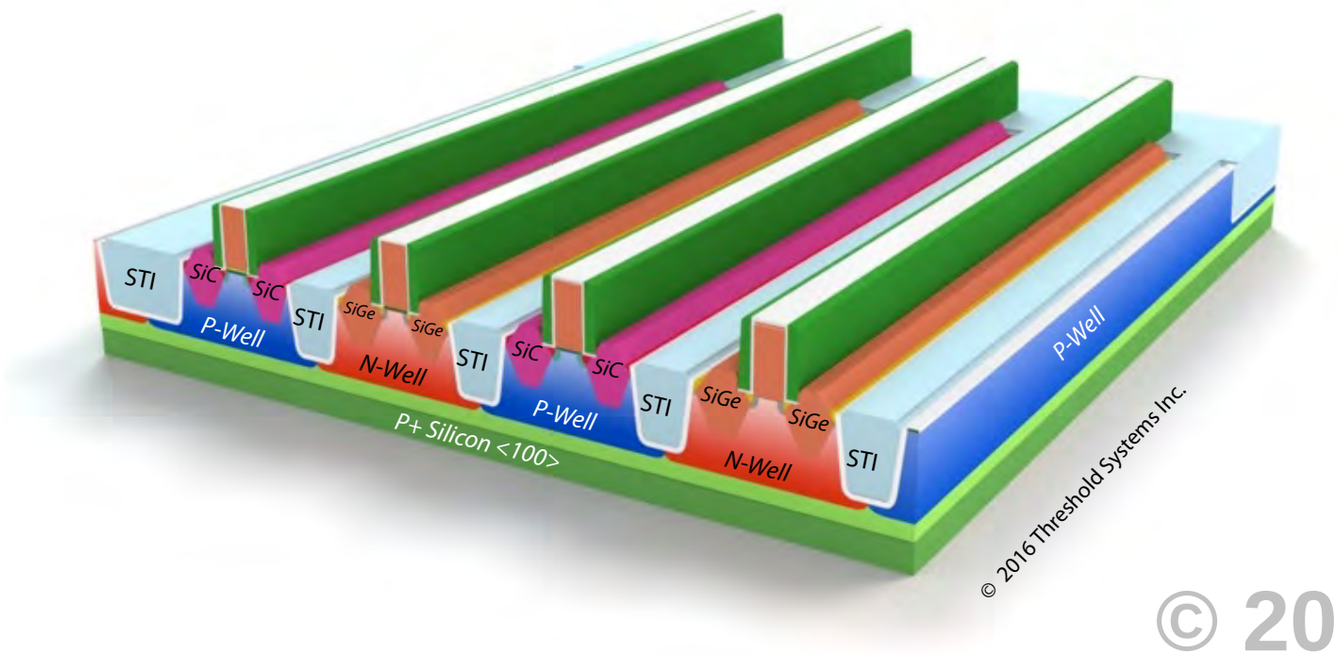
© Applied Materials

An open degas chamber.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Nickel Salicidation Oxide Strip:

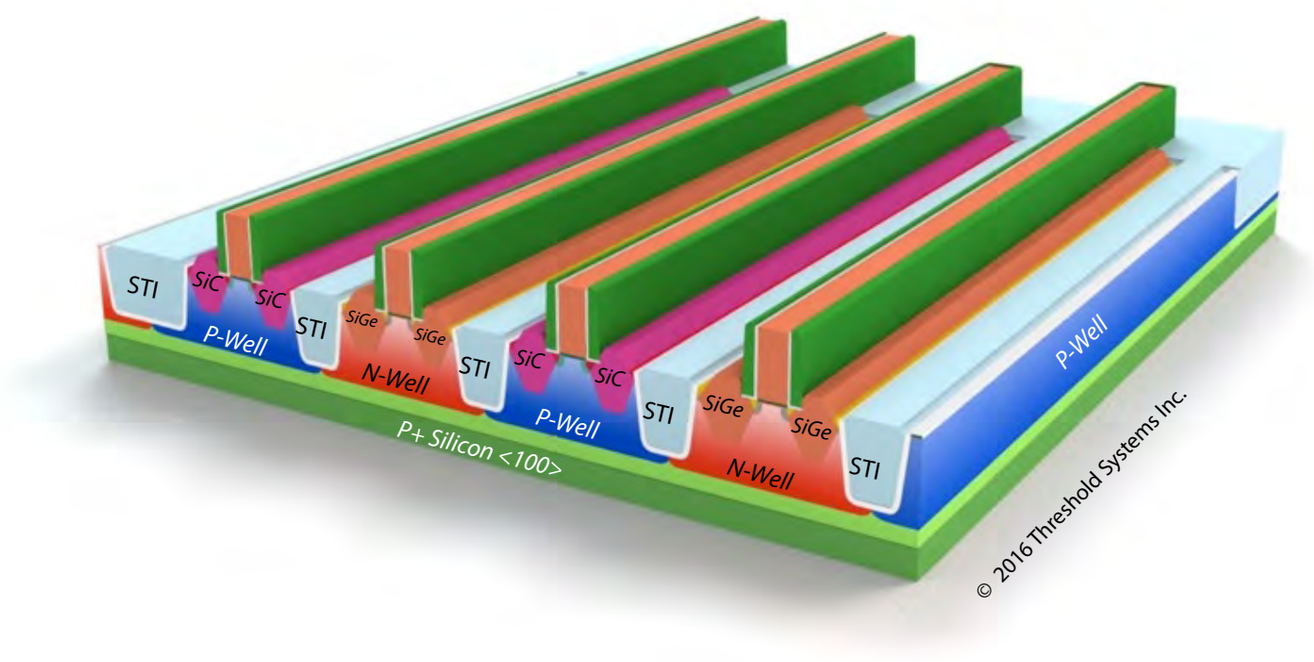
ThresholdSystems



After the anneal the wafers still have a layer of poly oxide on the top of the gate electrodes. There may also be a native oxide over the SiGe and the SiC Source/Drain regions.

Both of these oxides must be removed to expose a bare silicon surface for the salicidation process.

© 2018 Threshold Systems Inc.

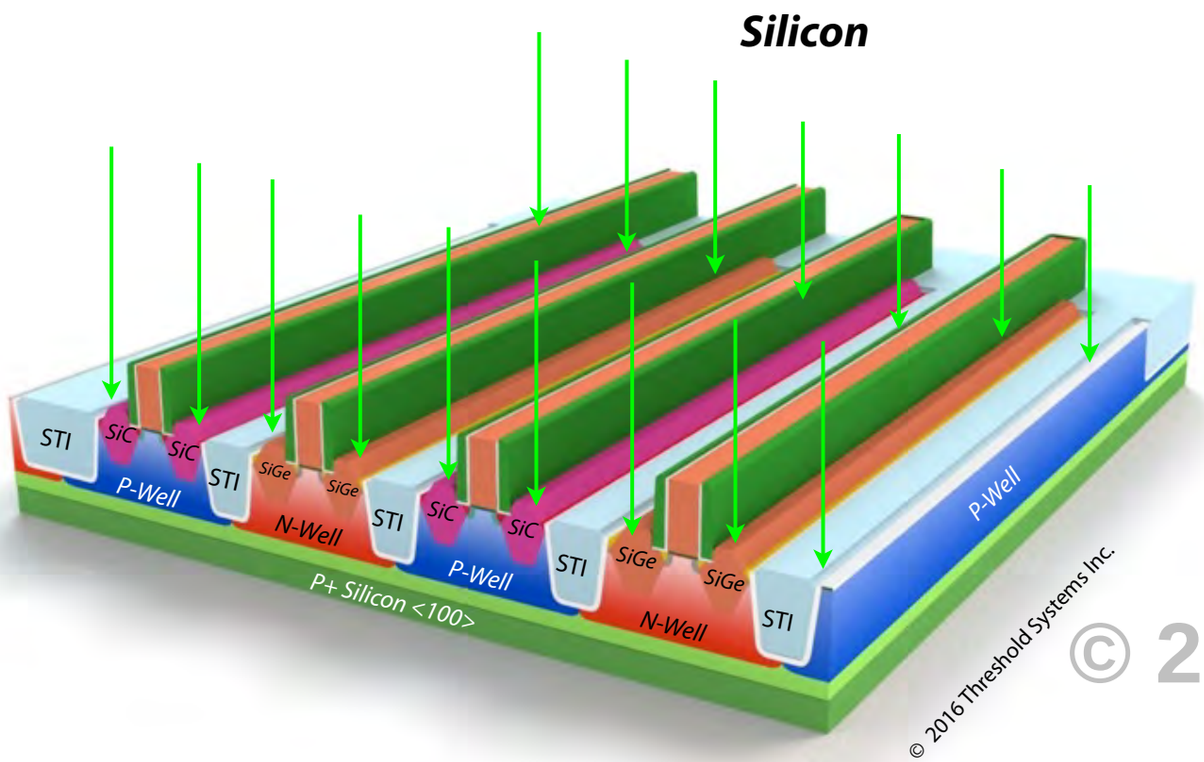


The wafers are then dipped in HF to remove all of the silicon dioxide from tops of the gate electrodes and from the surfaces of the Source/ Drain regions. This operation prepares the wafers for the upcoming salicidation step.

© 2018 by Threshold Systems Inc. All Rights Reserved.

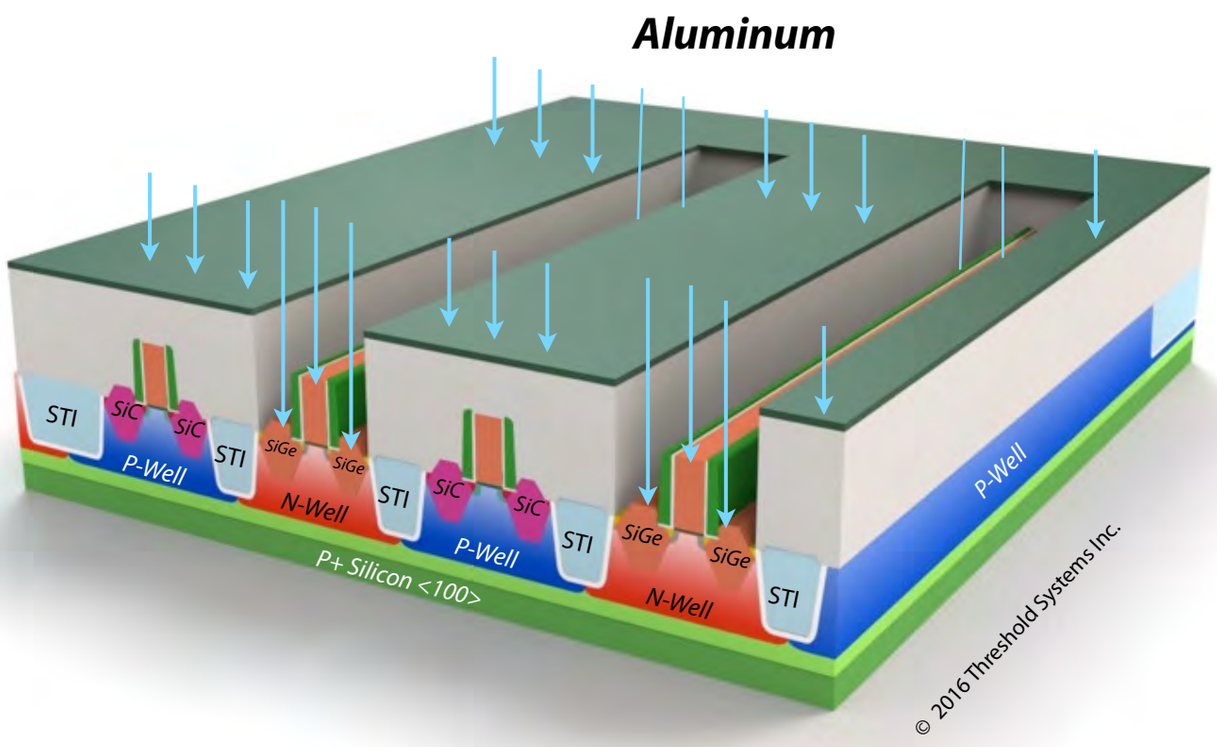
Salicidation Implants:

ThresholdSystems



The wafers are implanted with a Silicon Pre-Amorphization Implant (PAI). (1.0E15@5KeV, 0°) This amorphizes the surface of the silicon and facilitates the formation of a more uniform, lower resistance silicide (formed in the following operations).

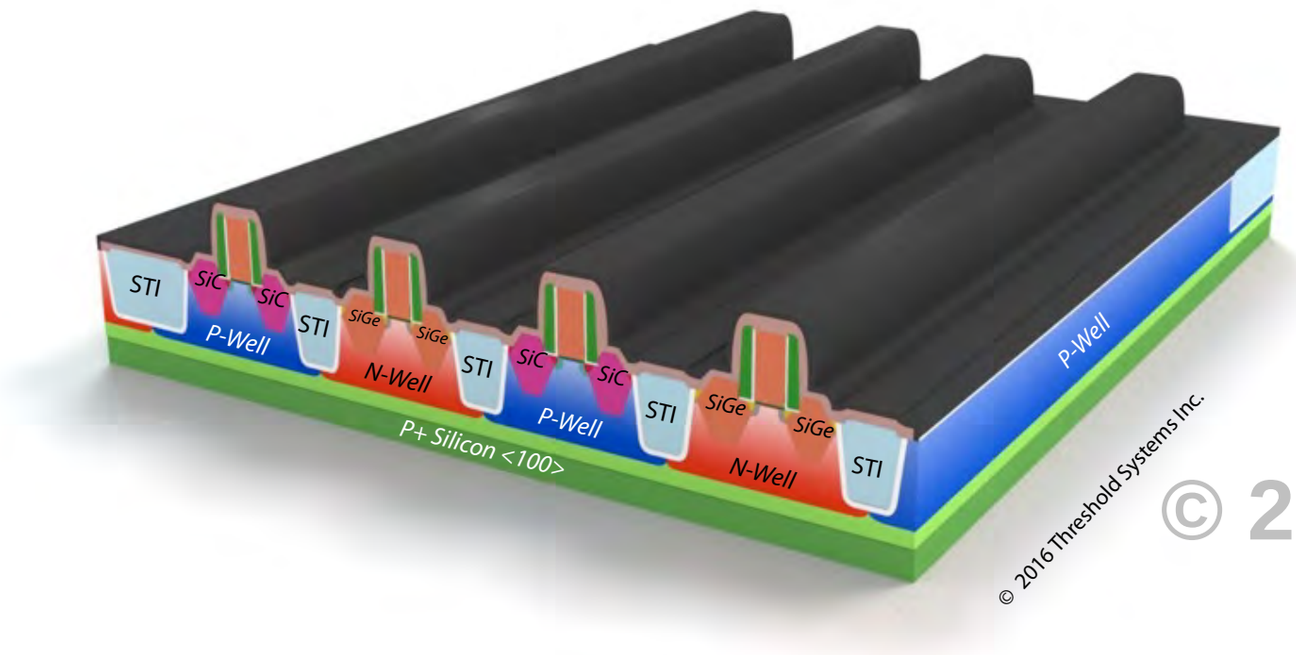
© 2018 Threshold Systems Inc.



The wafers are patterned with photoresist that exposes only the P-Wells. They are then implanted with Aluminum (~2E14@ 5KeV, 0°). The aluminum is positioned so that it will be located at the interface between the SiGe and the yet-to-be deposited Ni/Pt. This implant results in lower contact resistance for the PMOS devices because the Aluminum segregates to the top of the SiGe and being a P-Type dopant it lowers the Schottky barrier height for holes from ~.4eV to ~.12eV and increases drive current by up to 19%.

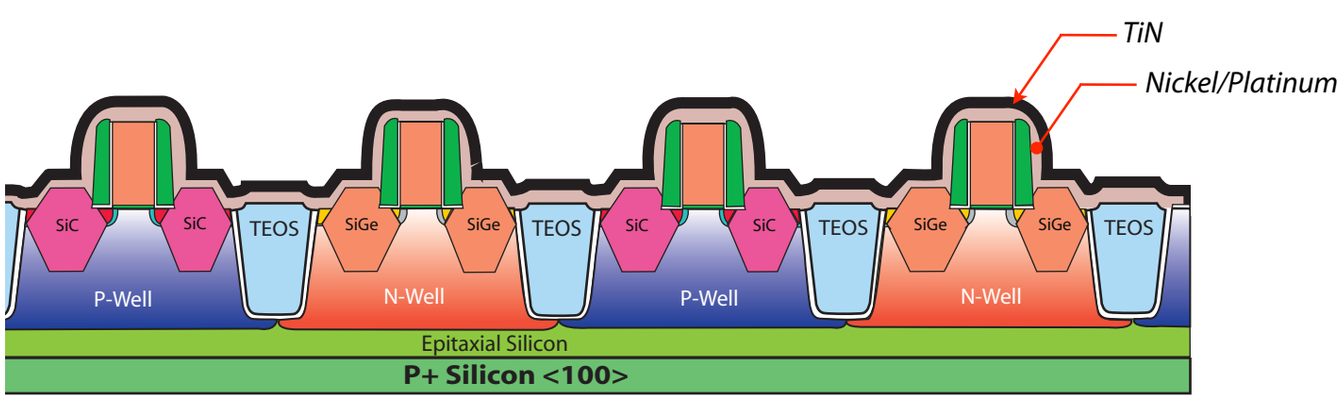
© 2018 by Threshold Systems Inc. All Rights Reserved.

Nickel/Platinum/TiN Deposition:



The wafers are cleaned with P/SC1 and a blanket coating of ~125 Å Nickel/Platinum and ~150 Å of TiN is deposited using Physical Vapor Deposition. The TiN cap on top of the Nickel helps to prevent the Nickel from diffusing away during the high-temperature salicidation process and promotes more uniform NiSi formation. The 5-10% Platinum Platinum helps to lower the barrier at the Ni-Si interface and establish a lower contact resistance. It also postpones the transformation of low resistance NiSi into the higher resistance $NiSi_2$.

© 2018 Threshold Systems Inc.

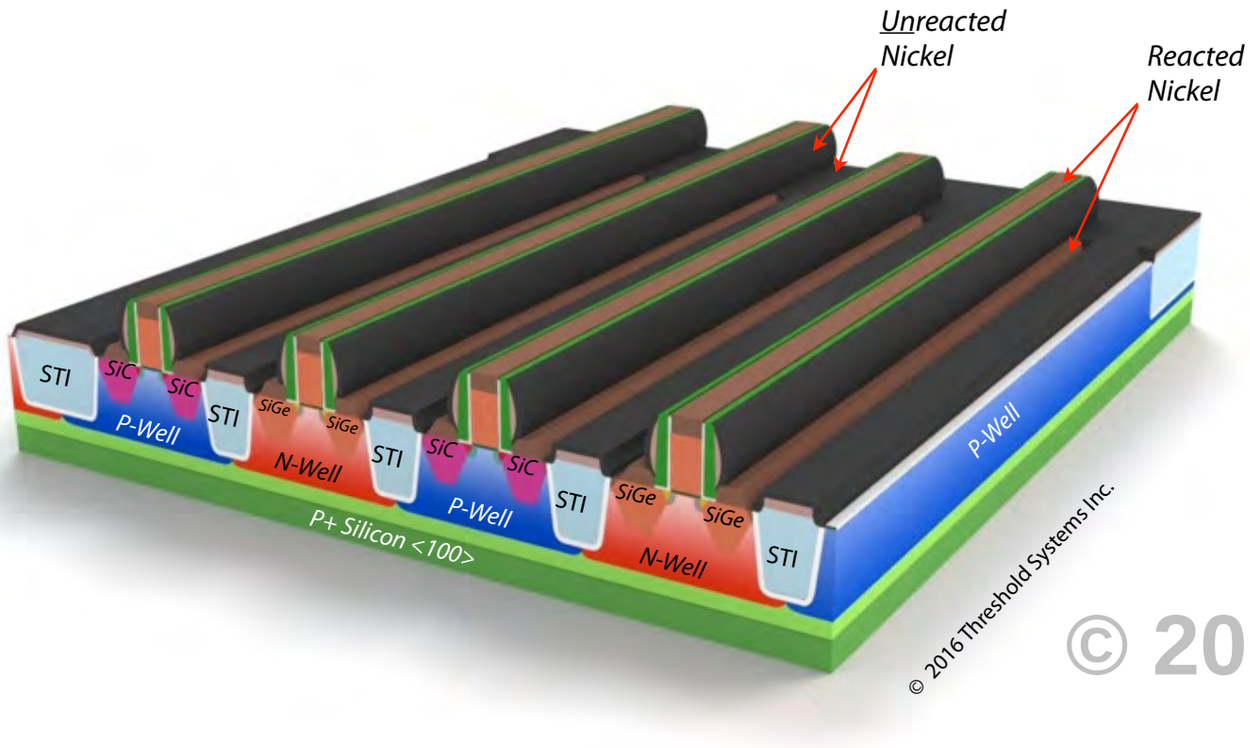


© 2016 Threshold Systems Inc.

The adjacent 2D image illustrates the Nickel/Platinum/TiN stack after initial deposition and prior to the anneal.

© 2018 by Threshold Systems Inc. All Rights Reserved.

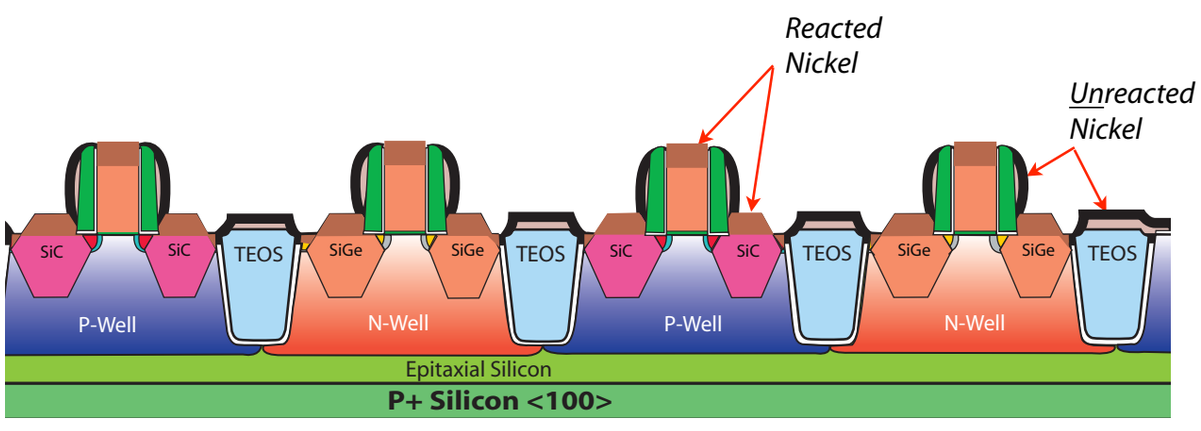
Salicidation Anneal:



The wafers undergo a RTA at ~250 ° C for 30 seconds that converts the nickel-Platinum into the high resistance C49 phase. A second RTA at ~390 ° C for 30 seconds converts the on the wafer's surface into the low resistance (15 µΩ-cm) C54 NiSi phase.

NiSi is necessary because it consumes at least 30% less silicon than the same thickness of $CoSi_2$. This is extremely important because the 20nm node junctions are very shallow. In addition, NiSi has a very smooth surface which minimizes junction leakage.

© 2018 Threshold Systems Inc.

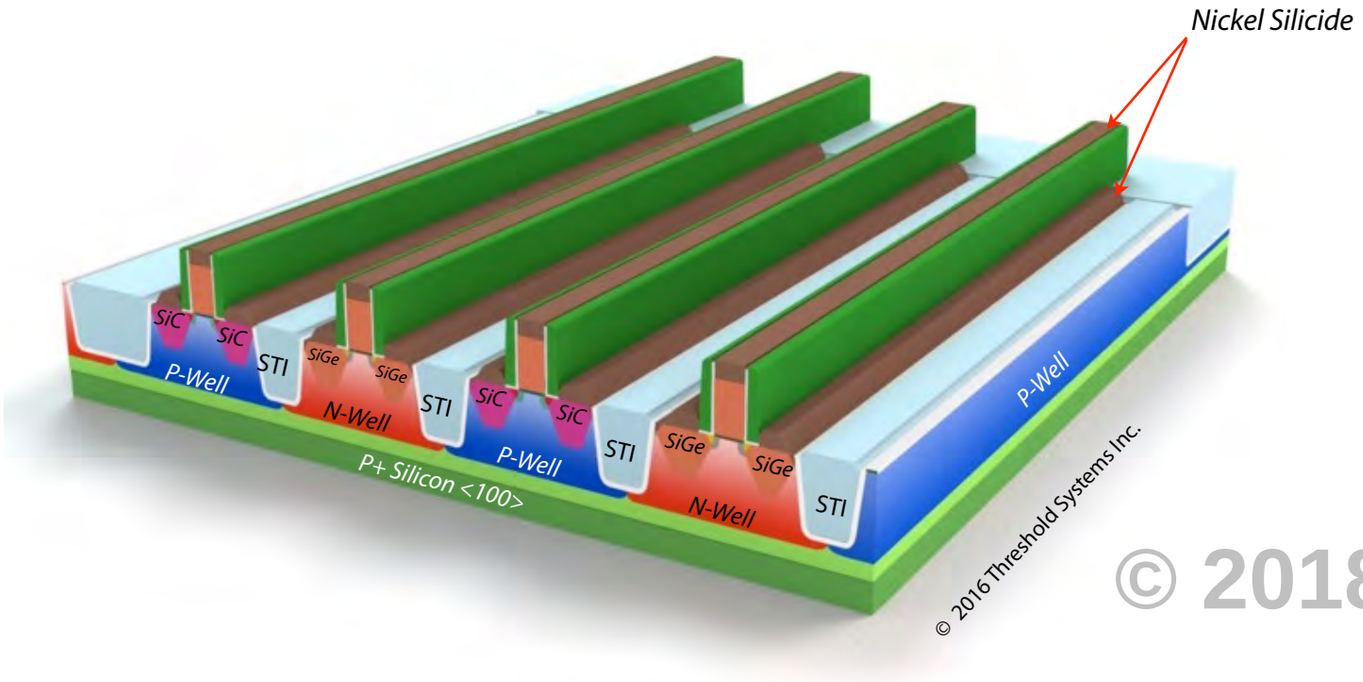


© 2016 Threshold Systems Inc.

The adjacent 2D image illustrates the reacted and unreacted Nickel/ Platinum on the devices. In order to form a Nickel silicide, the Nickel/ Platinum layer must come into contact with an exposed Silicon surface. The only place where exposed Silicon is located is on top of the gate electrodes or on the Source/Drain regions. The Nickel/Platinum on top of the STI and on the Nitride spacers will remain unreacted.

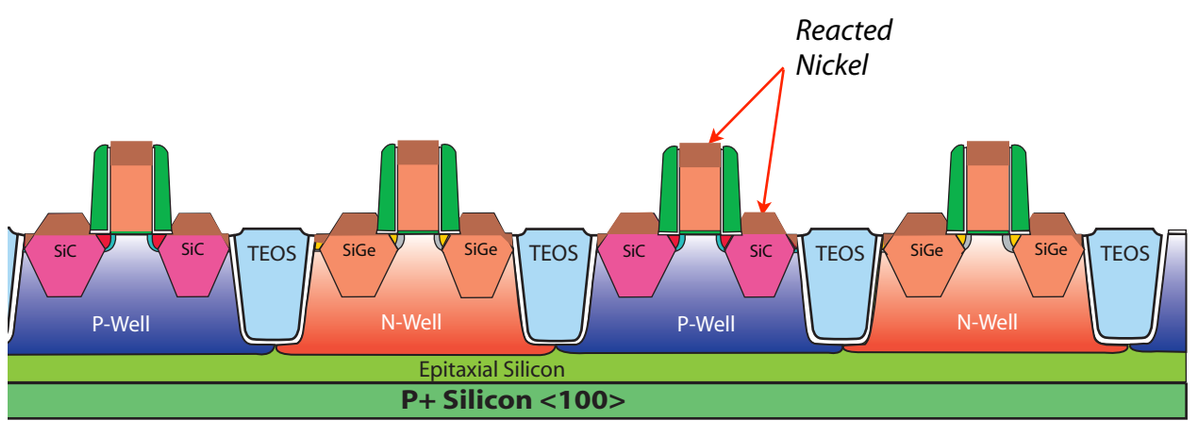
© 2018 by Threshold Systems Inc. All Rights Reserved.

Unreacted Nickel Strip:



The unreacted nickel located on the spacer sidewalls and on top of the STI is etched away with a solution of Piranha and SC1. This etch is highly selective and leaves the reacted salicide on top of the gate electrodes and over the Source/Drains untouched. This etch also removes the TiN capping layer from all surfaces on the wafer.

© 2018 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

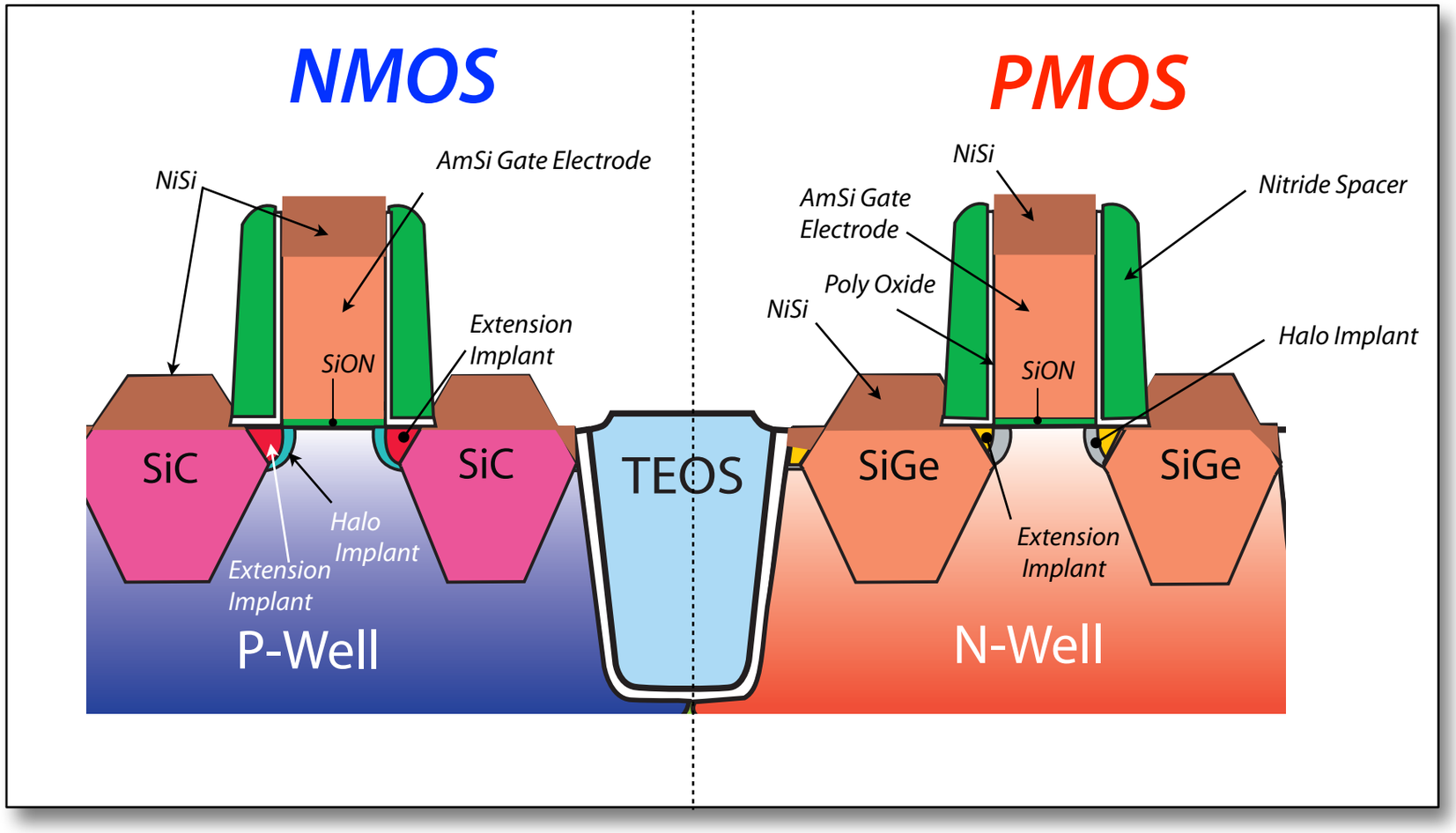
The adjacent 2D image illustrates the fact that the Piranha solvent readily attacked and removed the unreacted Nickel/Platinum from the nitride spacer sidewalls and from the top of the STI.

The Piranha does not attack the reacted NiSi over the Source/Drains or on top of the Gate Electrodes.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Transistor Nomenclature:

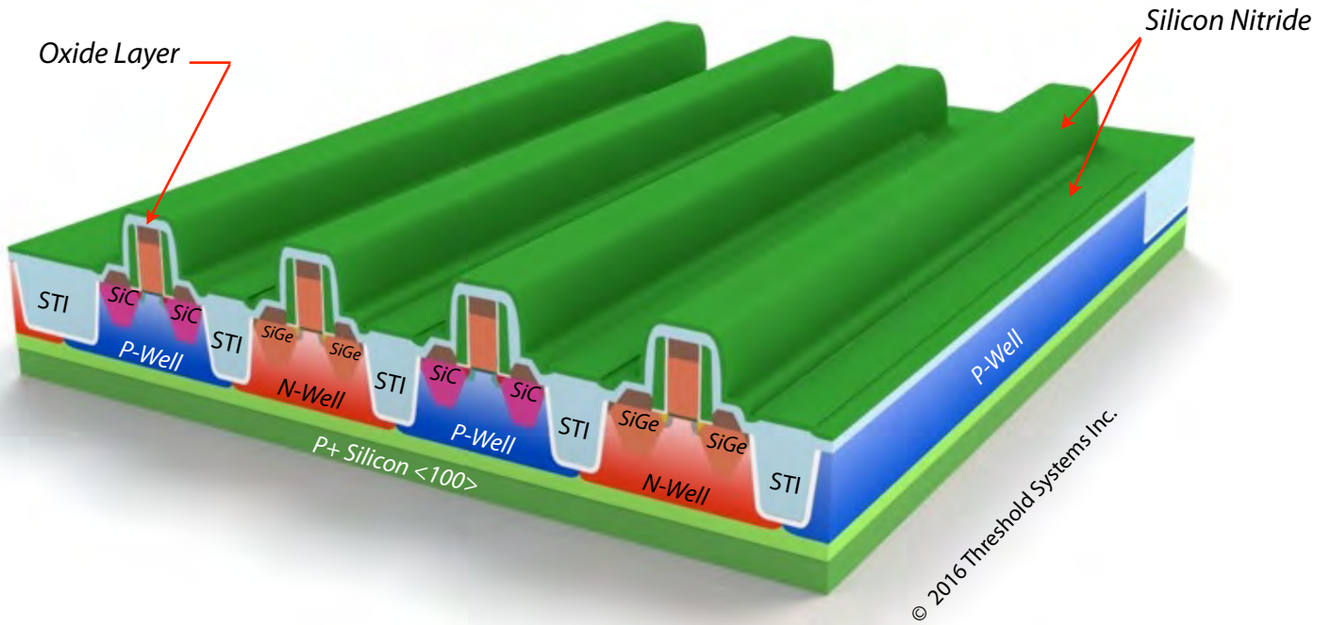
© 2018 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

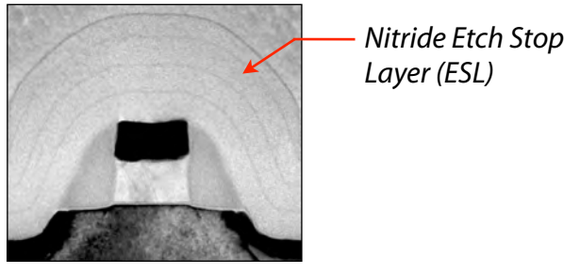
© 2018 by Threshold Systems Inc. All Rights Reserved.

Pre-Metal Dielectric (PMD):

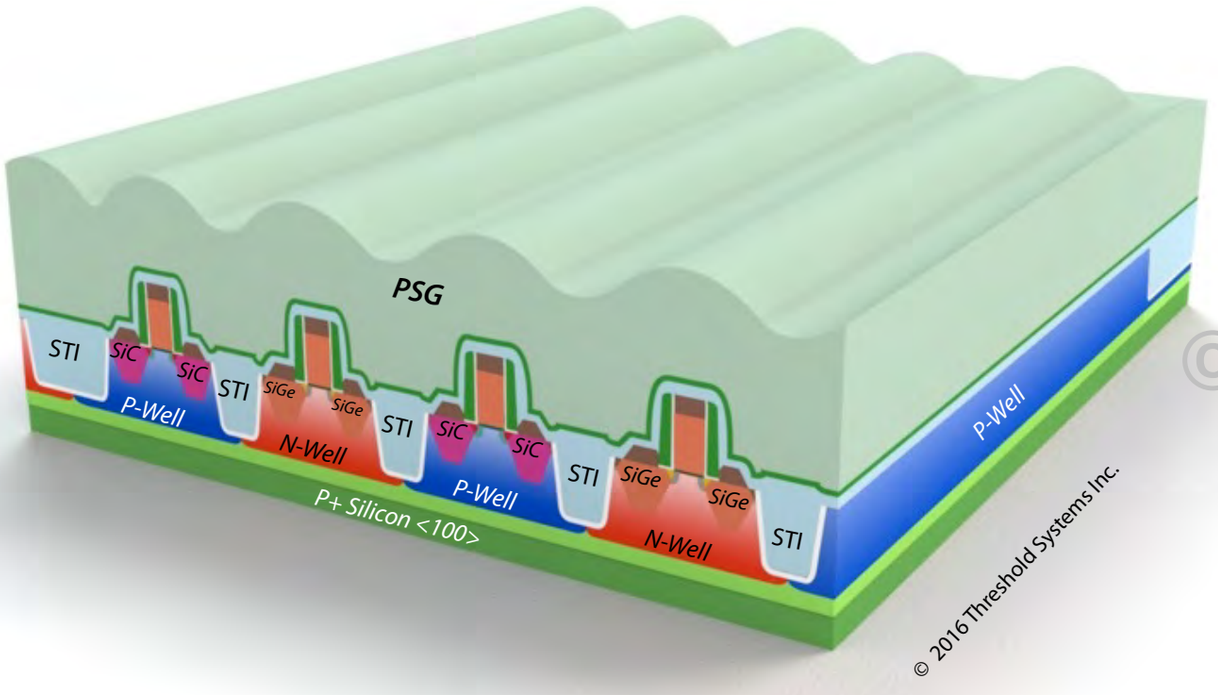


The wafers are cleaned in P/SC1 and thin 75 Å layer of silicon dioxide followed by ~75Å of nitride is deposited. The oxide acts as a diffusion barrier to external dopants that could contaminate the device layer (the "active" Layer).

The silicon nitride will act as an etch-stop layer for the contact etch which will connect the transistor Source/Drains and gate electrodes to the yet-to-be-formed Tungsten trenches.



CH Chan et al, IEDM 2005 3.4



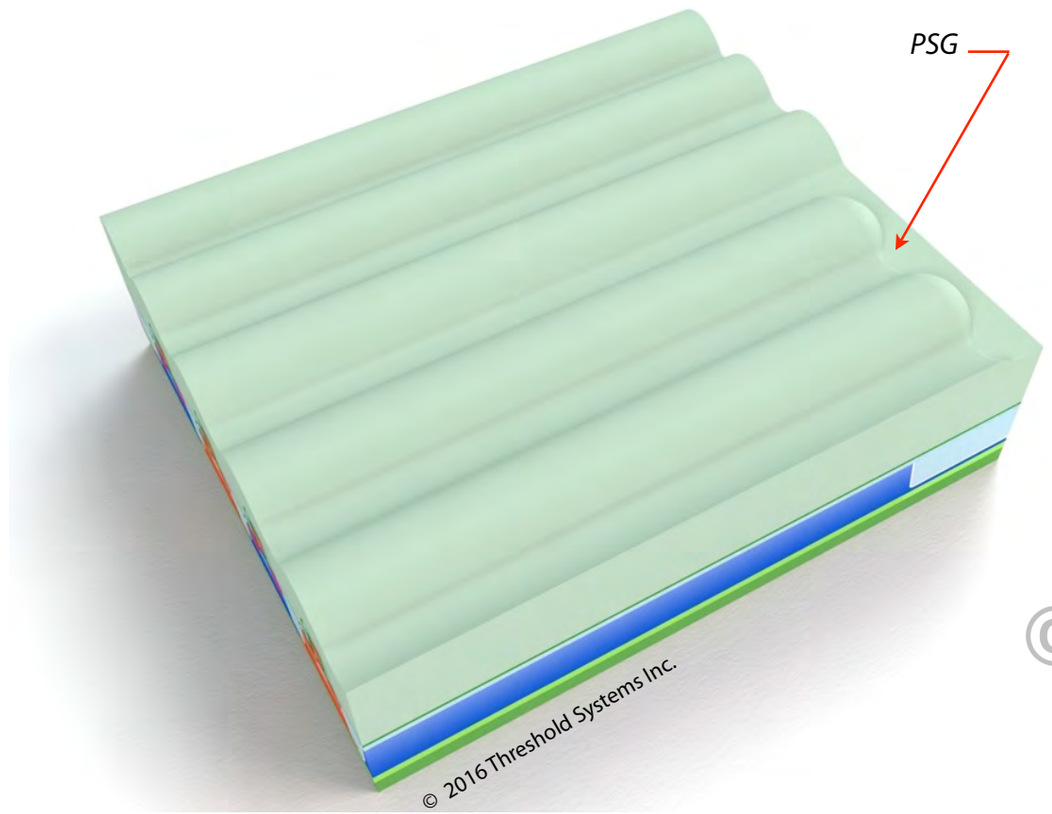
Next, 2,000 Å of a Phosphorus doped glass known as PSG (Phospho-Silicate Glass) is deposited over the wafers using high density CVD. This layer forms the Pre-Metal-Dielectric (PMD).

© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

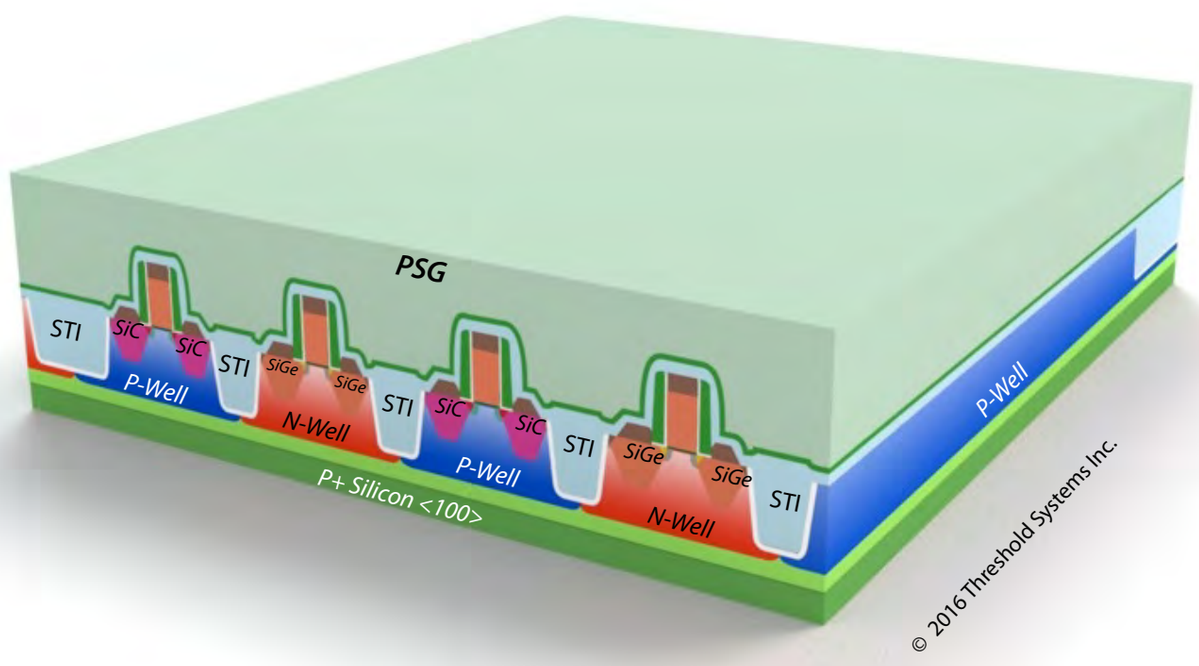
Pre-Metal Dielectric Polish-Back:

ThresholdSystems



A top-down view of the PSG layer prior or polishing. Note the humps in the PSG caused by the underlying topography induced by the height of the gate electrodes.

© 2018 Threshold Systems Inc.

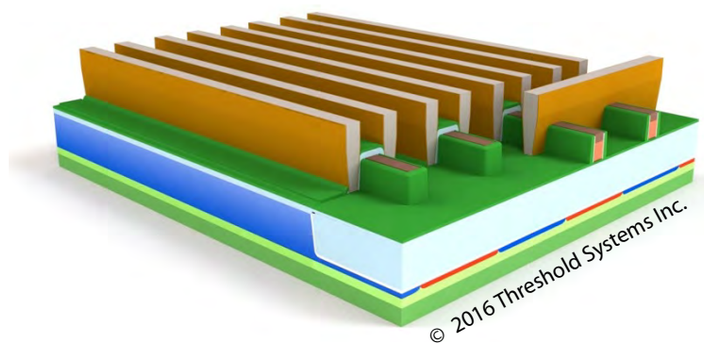


The wafers are polished back using CMP to a thickness of ~1,200 Å.

© 2018 by Threshold Systems Inc. All Rights Reserved.

© 2018 by Threshold Systems Inc. All Rights Reserved.

The Contact Module

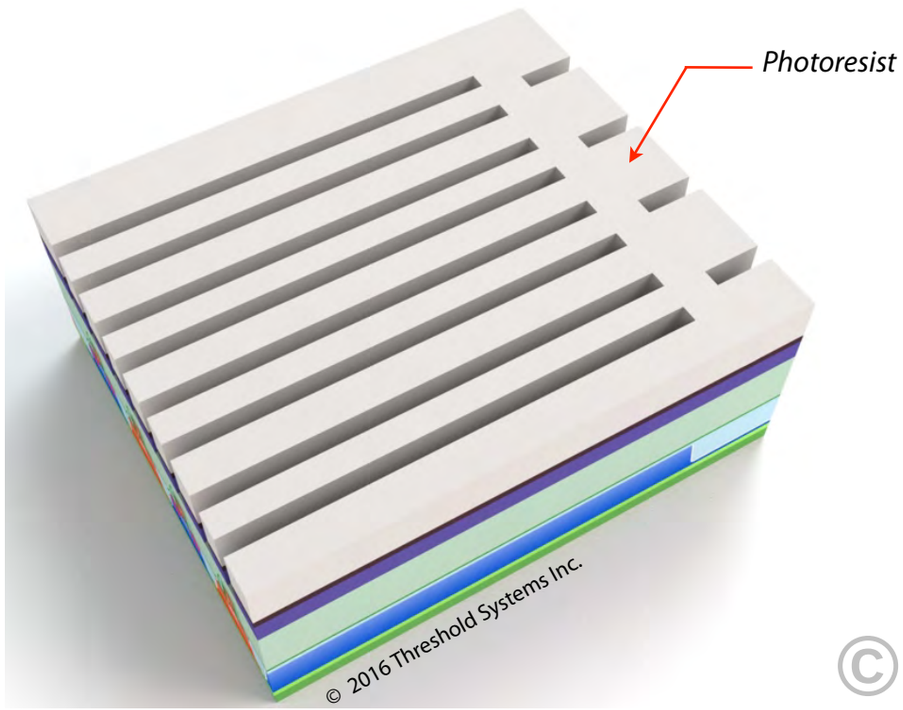


© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Hard Mask Patterning:

ThresholdSystems

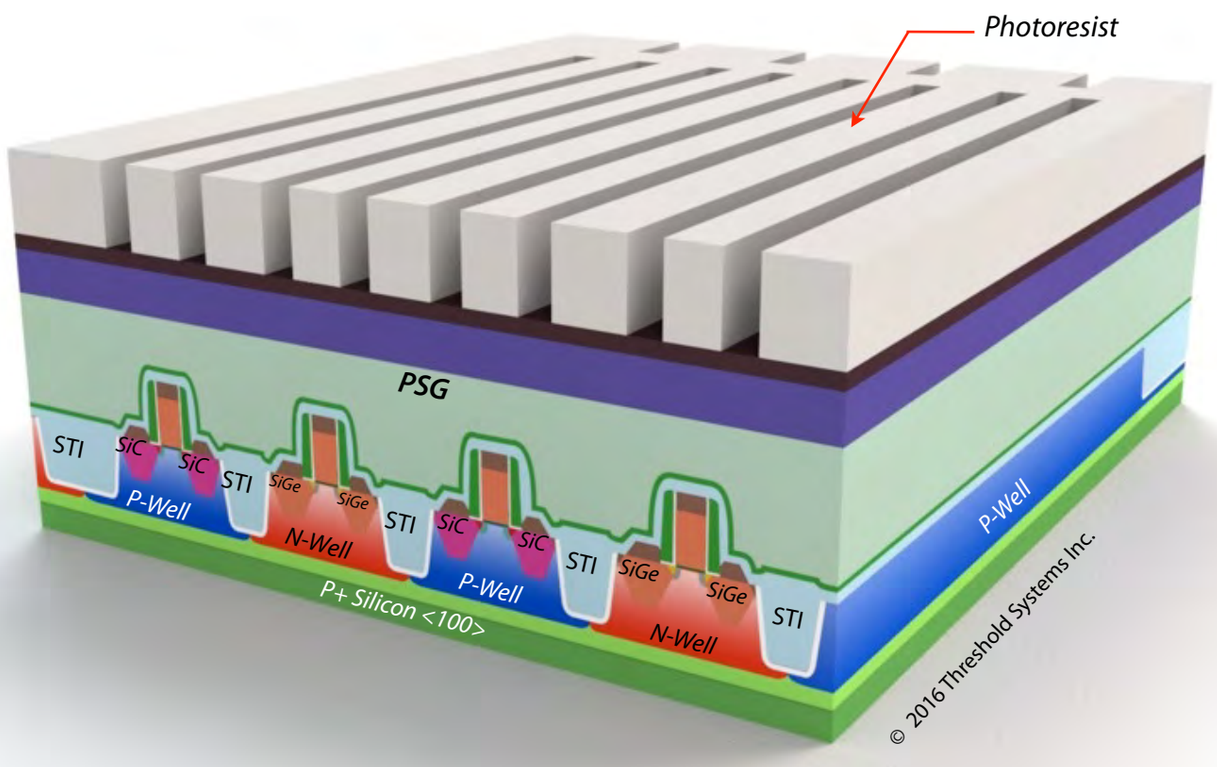


Next, the wafers are primed in HMDS, coated in photoresist and patterned.

The illustration to the left is a top-down view of the trench contact photoresist pattern. Note the gate electrode contact trenches at the right of the image.

This would be a double patterning operation.

© 2018 Threshold Systems Inc.

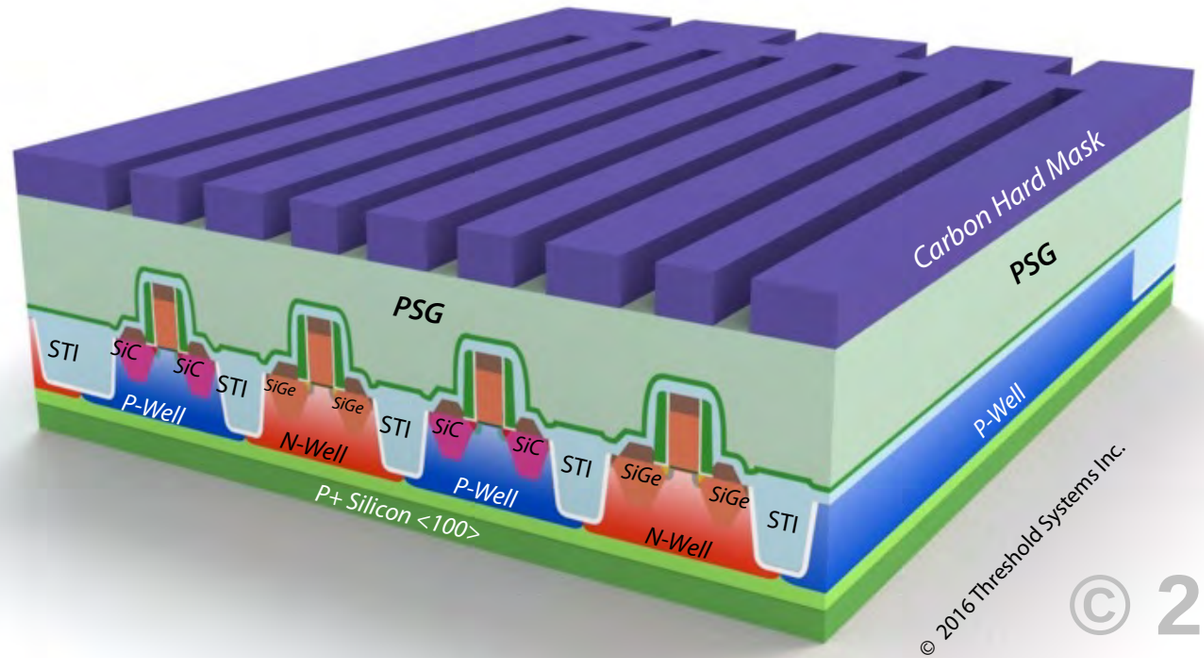


The illustration to the left is a tilted front-end view of the trench contact photoresist pattern. Note the gate electrode contact trenches at the back of the image.

© 2018 by Threshold Systems Inc. All Rights Reserved.

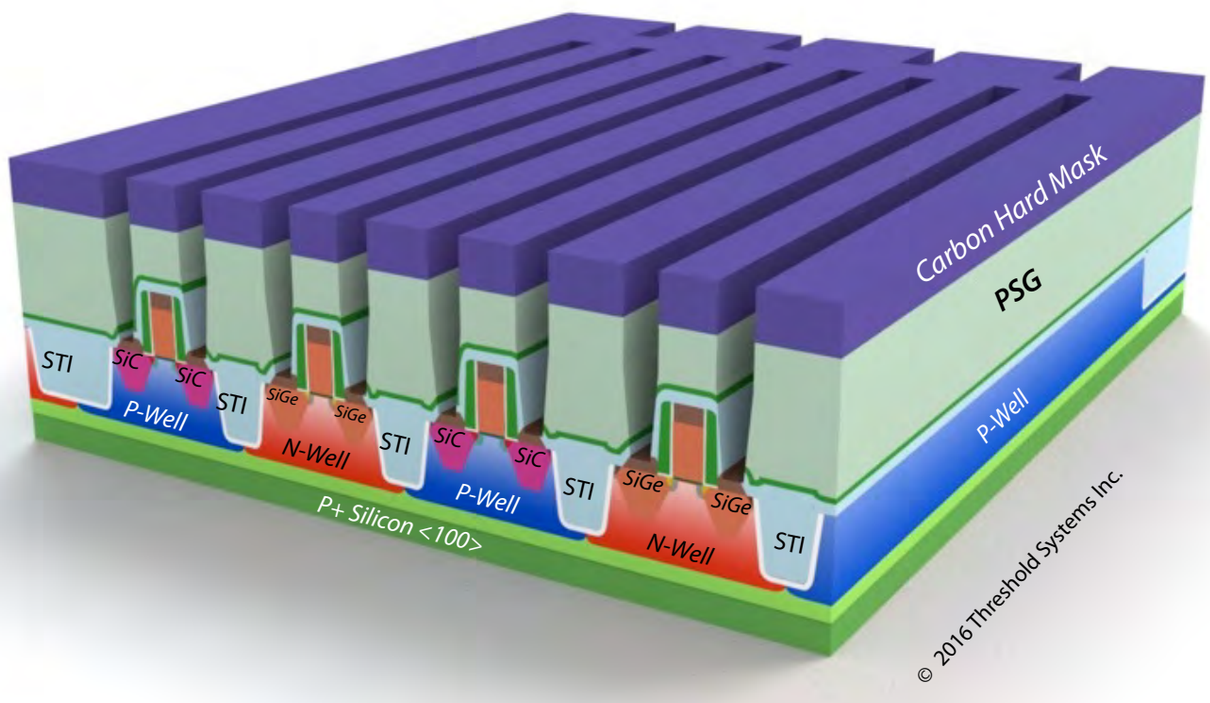
Hard Mask Etch & Photoresist Strip:

ThresholdSystems



The photoresist is ashed away leaving behind the patterned amorphous carbon hard mask. The wafers are then cleaned in Piranha to remove any residual photoresist.

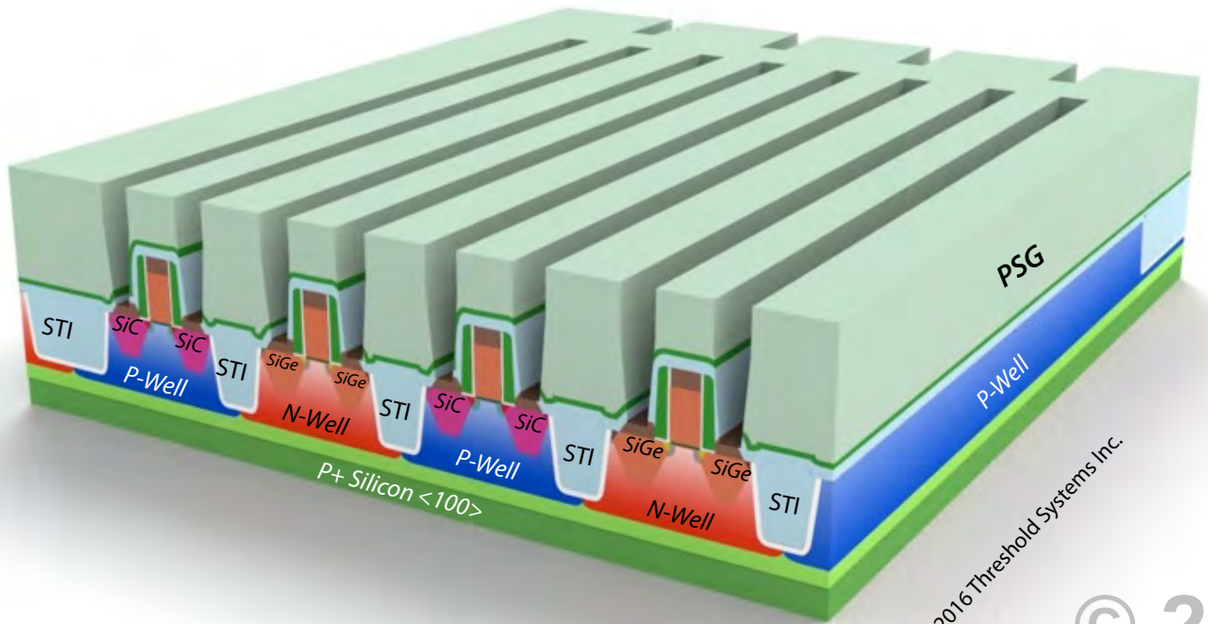
© 2018 Threshold Systems Inc.



Using the Amorphous Carbon Hard Mask as a pattern, a highly anisotropic etch cuts down through the PSG to form the contact holes. The etch stops on the Silicon Nitride located above the Source/Drains and on the top of the gate electrode. The etch chemistry is then changed and the Silicon nitride and underlying oxide is etched away. The etch stops on the Nickel silicide located over the Source/Drains and on top of the Gate Electrode.

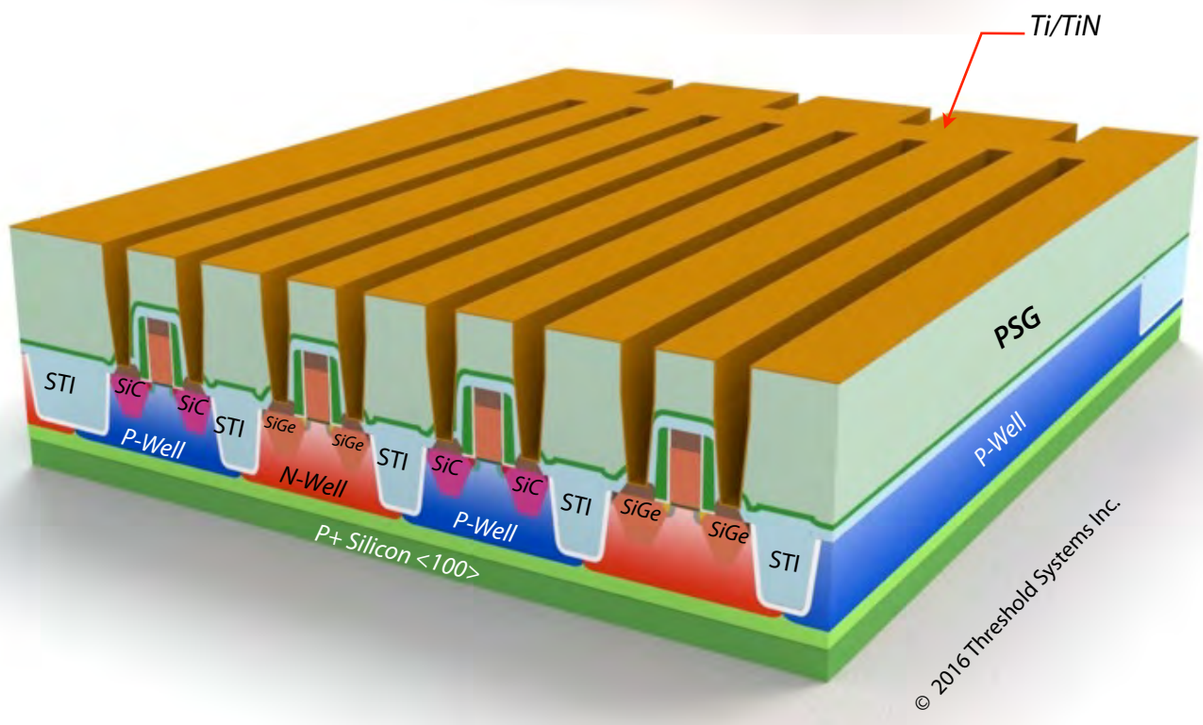
© 2018 by Threshold Systems Inc. All Rights Reserved.

Hard Mask Strip & TiN Deposition:



The hardMask is etched away and the wafers are cleaned in Piranha.

© 2018 Threshold Systems Inc.



The wafers are degassed and then exposed to a soft Argon sputter pre-clean to remove any native oxides present on top of the NiSi at the bottom of the contact holes.

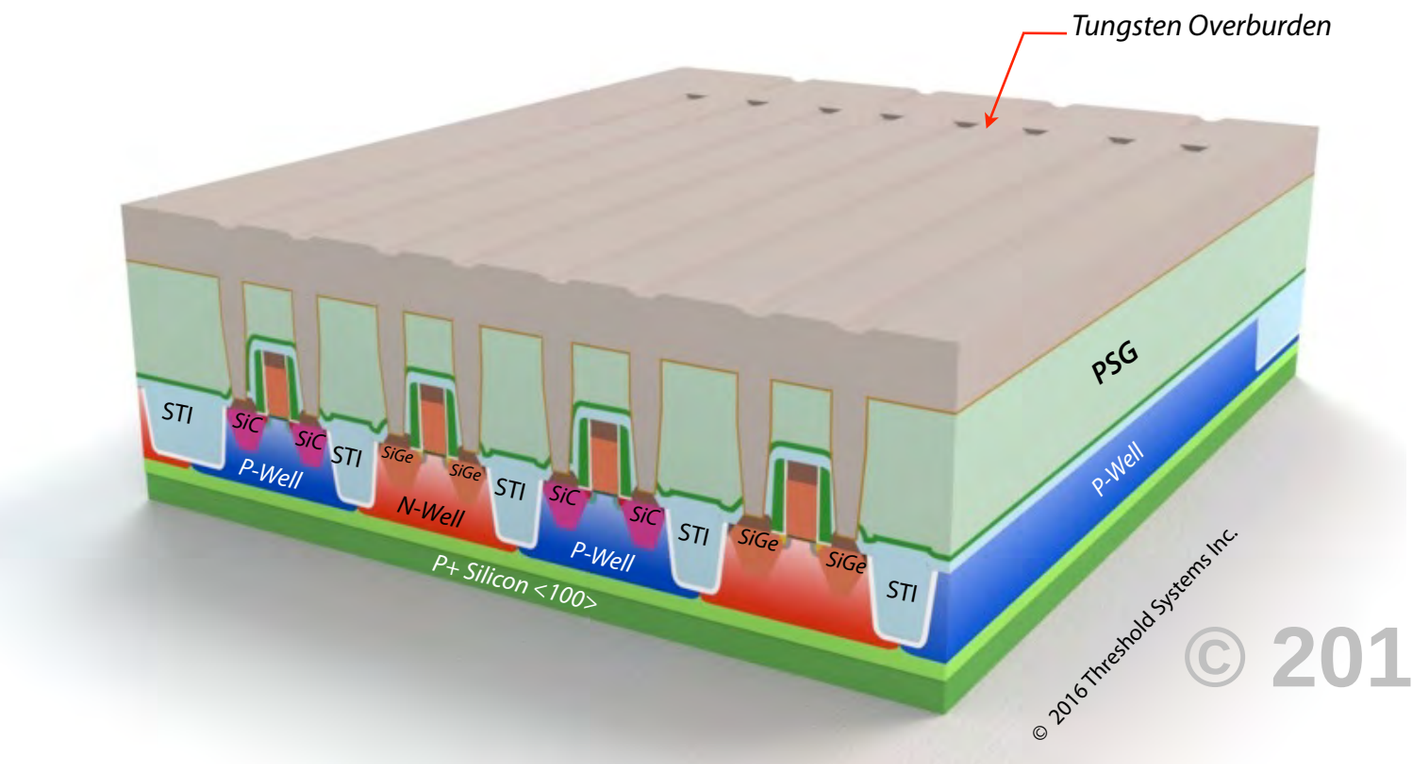
Next, ~60 Å of Titanium (a glue layer) is deposited by Ionized Metal Plasma (IMP) PVD. This is followed by ~30 Å deposition of TiN (a barrier layer) deposited via IMP PVD. The TiN is re-puttered after deposition to enhance sidewall coverage inside the contact.

The wafers are then heated to in an RTA to react the Titanium/Titanium Nitride and set the resistance of the glue layer.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Tungsten Deposition & Polish-Back:

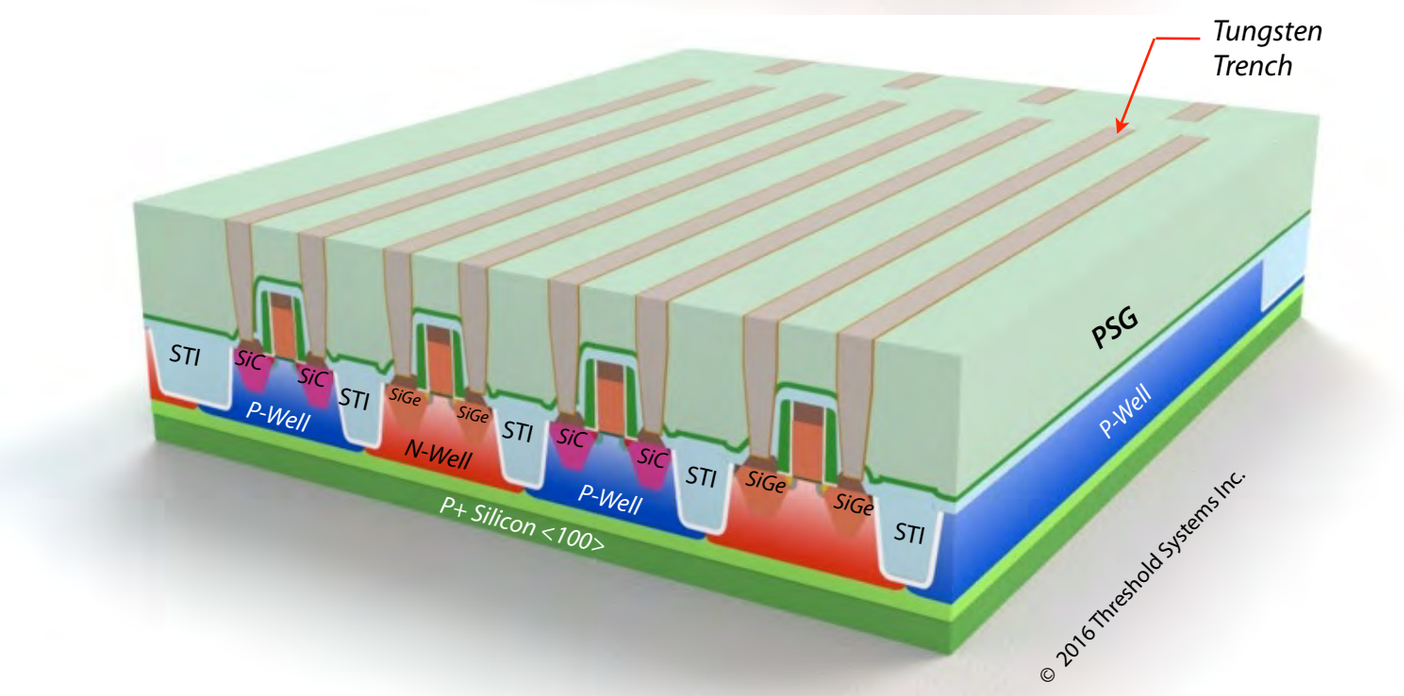
ThresholdSystems



Next, a seed layer of Tungsten is deposited that lines the inside of the contact holes and ensures conformal coverage of the bulk Tungsten deposition.

Finally, ~2,500 Å of Tungsten is deposited via CVD.

© 2018 Threshold Systems Inc.



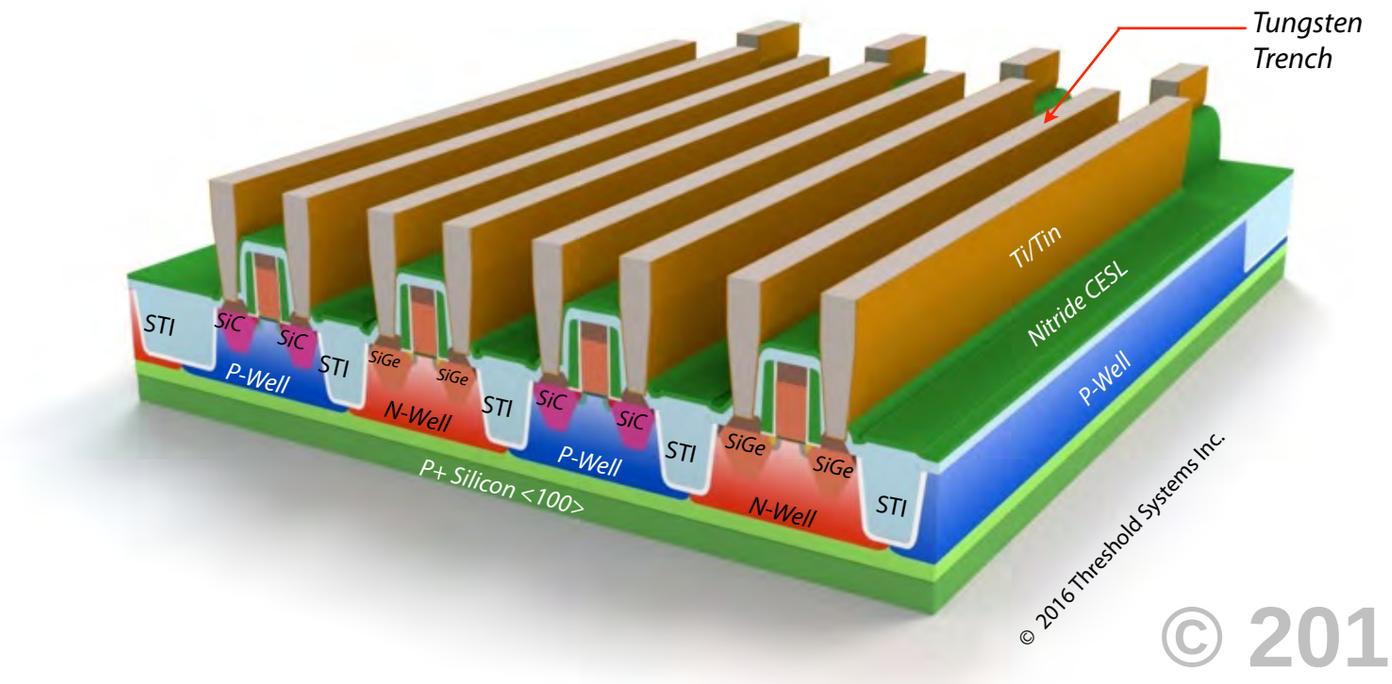
The wafers are again polished using CMP. In this operation the Tungsten overburden is removed and the Tungsten trenches are polished back tops of the PSG.

Next, an oxide buff operation is conducted to smooth out the surface of the PSG oxide and ensure that the surface of the PSG is coplanar with the top of the Tungsten trenches.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Trench Contacts (Alternate Scheme):

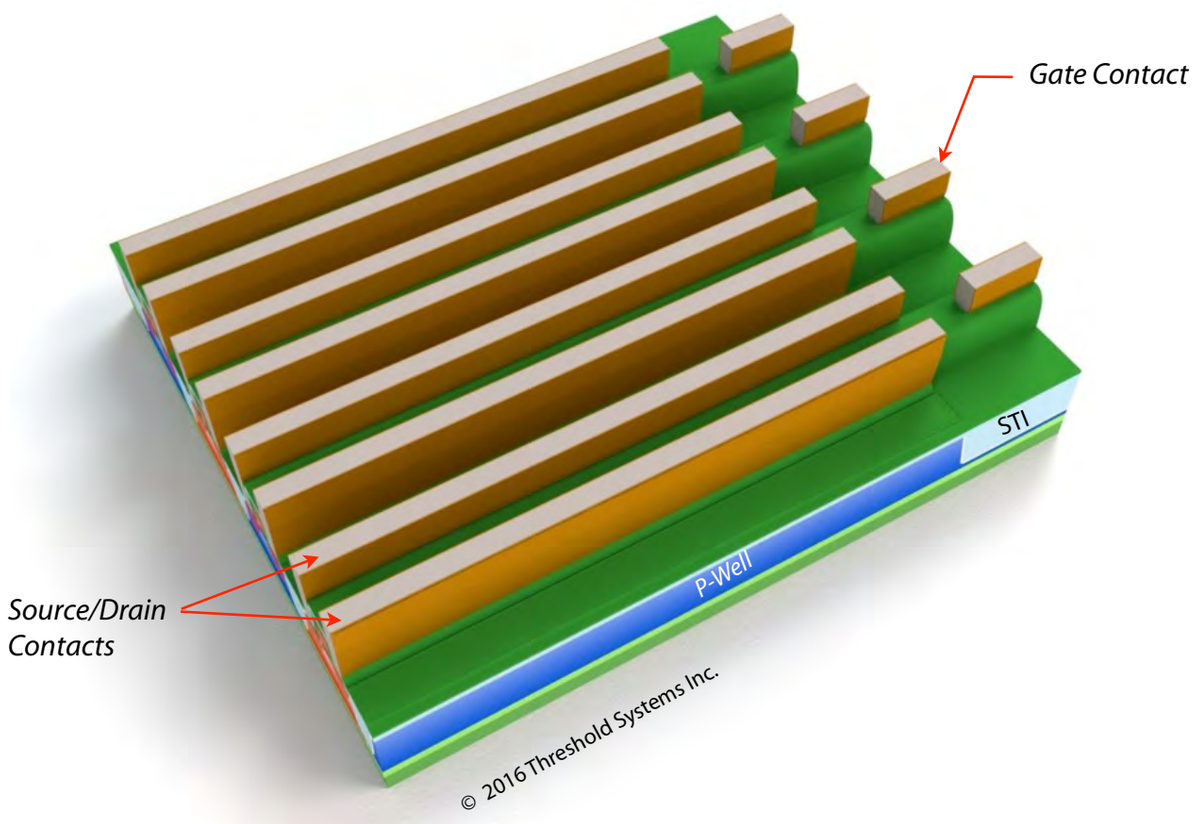
ThresholdSystems

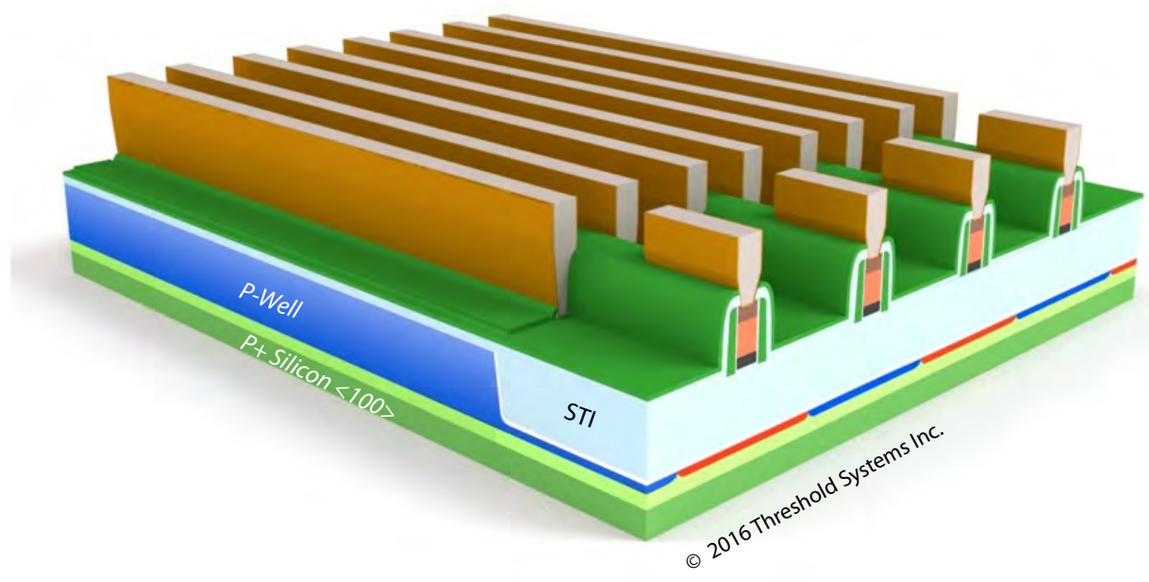


These figures illustrate what the Tungsten trenches would look like if the PSG oxide were removed. Clearly visible are the Tungsten trenches, clad in Ti/TiN, making electrical contact to the Source Drain regions and to the Gate Electrodes.

These views, of course, would not be visible in a production device.

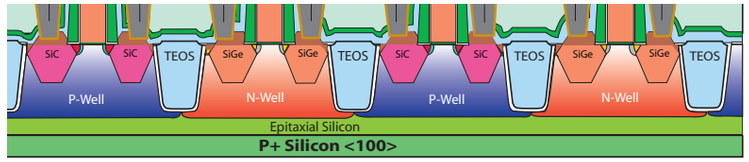
© 2018 Threshold Systems Inc.





© 2018 by Threshold Systems Inc. All Rights Reserved.

Contact Module:
Polished-Back

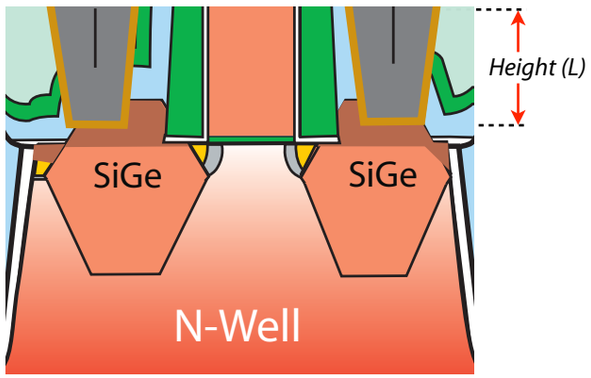


© 2016 Threshold Systems Inc.

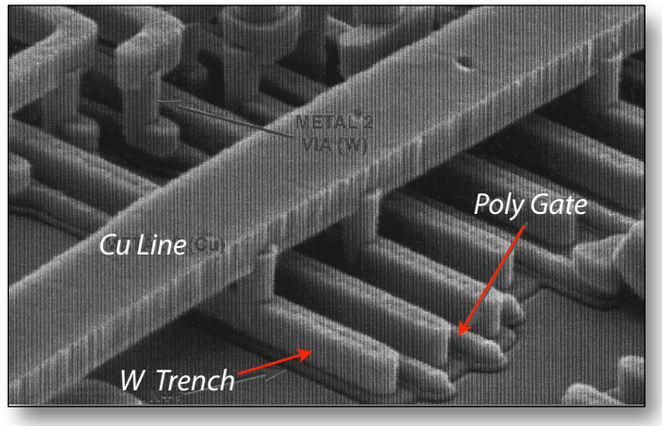
© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Plug Polish Back/Tungsten Trenches:



© 2016 Threshold Systems Inc.



Source: ICE

Polishing the Tungsten plugs down to be coplanar with the tops of the gate electrodes is something new and was first introduced at the 32nm node. Shorter Tungsten plugs have a lower contact resistance. This feature, coupled with the raised Source/Drains for both the NMOS and PMOS transistors, ensure that the contact resistance is kept to acceptable values.

It is difficult to eliminate Tungsten plugs completely and adopt Cu plugs because it is too dangerous to bring the Cu all the way down to the Source/Drains regions.

© 2018 Threshold Systems Inc.

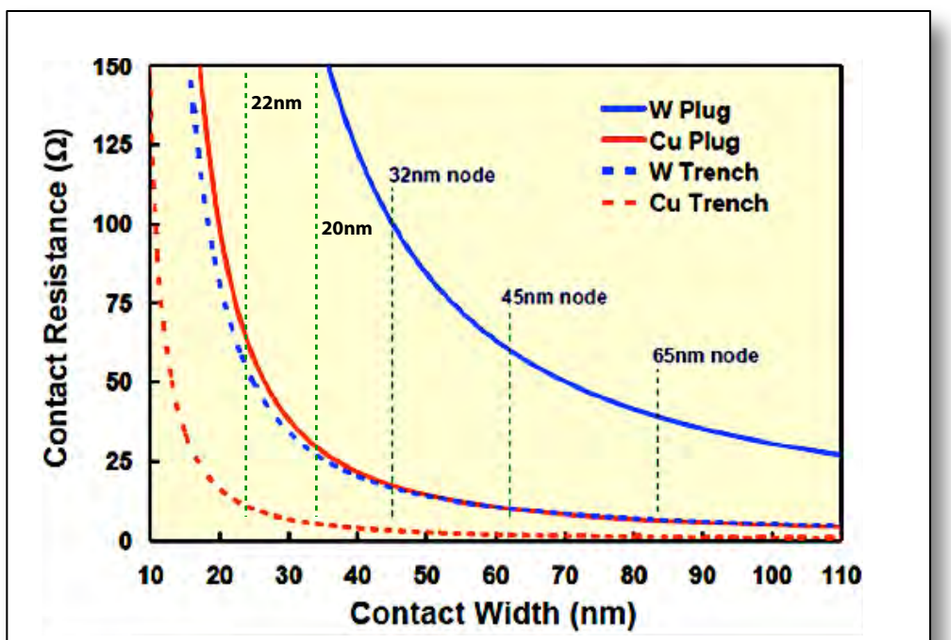
Resistance

Conductor Length

$$R = \rho \left[\frac{L}{Ax} \right]$$

Conductor Cross-sectional Area

Resistivity



Source: Arghavani et al., Semiconductor Fabtech

The data displayed on the graph to the left makes it plain why contact plugs must be polished back to such a short height at the 20nm node. The resistivity of Tungsten is such that at the contact diameters required at the 20nm node the plug resistance soars to unacceptable levels. However, because resistance is a function of length, shortening the length (i.e. the height) of the plug reduces the plugs overall resistance.

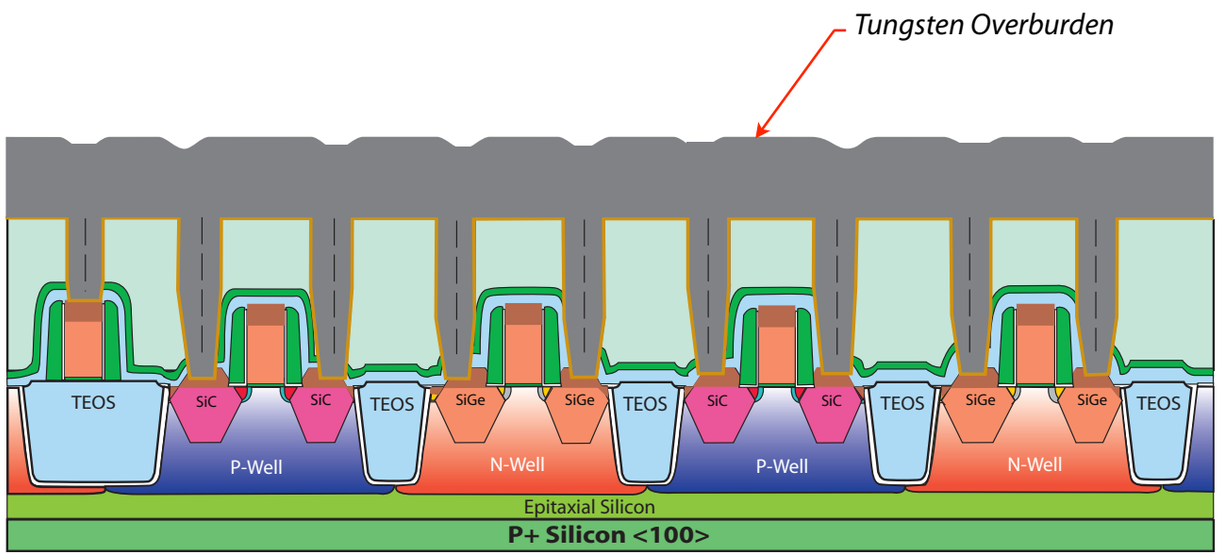
An alternative approach to the plug resistance problem is to keep the plugs narrow so that they fit onto the Source/Drains, but make then very long so that they have a very high contact area, and thus, a low overall resistance. Such contacts form trench-like structures over the Source/Drains and are called "Trenches" because they are formed by etching a trench in the PSG and then filling it with Tungsten.

However, the lower contact resistance of Tungsten trenches is at the cost of higher trench-gate electrode gate capacitance.

The trend of increasing contact resistance and smaller technology nodes appears to make Cu plugs inevitable at 7nm node.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Plug Polish Back:

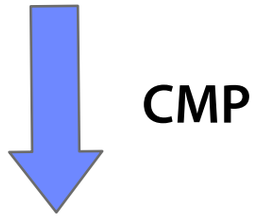


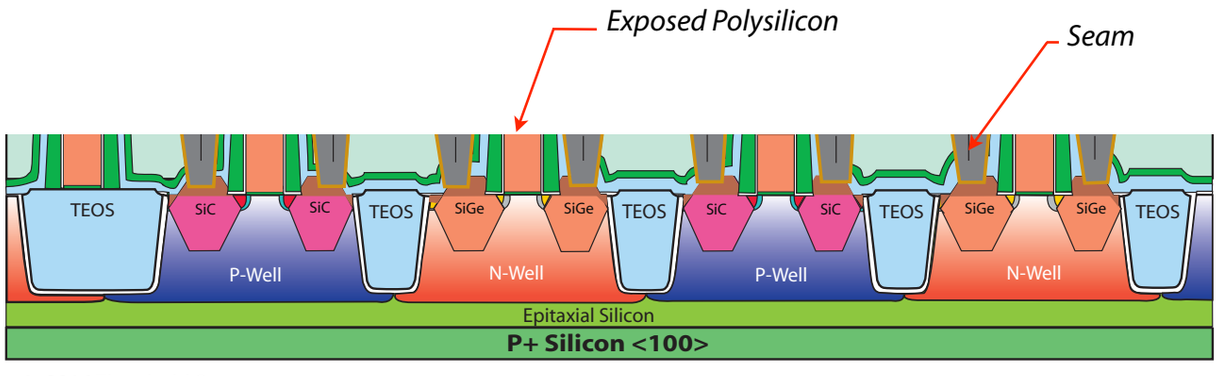
© 2016 Threshold Systems Inc.

Polished-back plugs are initially formed in the same manner as standard plugs. After the Ti/TiN barrier metals are deposited a Tungsten seed layer is deposited followed by the bulk CVD deposition of 2,500Å of Tungsten.

The Tungsten overburden is then polished back until the plugs are so short that they are coplanar with the tops of the gate electrodes.

© 2018 Threshold Systems Inc.





© 2016 Threshold Systems Inc.

Polishing the Tungsten plugs down to be coplanar with the tops of the gate electrodes is something new and was first introduced at the 32nm node. Shorter Tungsten plugs have a lower contact resistance. This feature, coupled with the raised Source/Drains for both the NMOS and PMOS transistors, ensure that the contact resistance is kept to acceptable values.

The Tungsten plugs cannot be eliminated altogether and replaced with Cu plugs because it is too dangerous to bring the Cu all the way down to the Source/Drains.

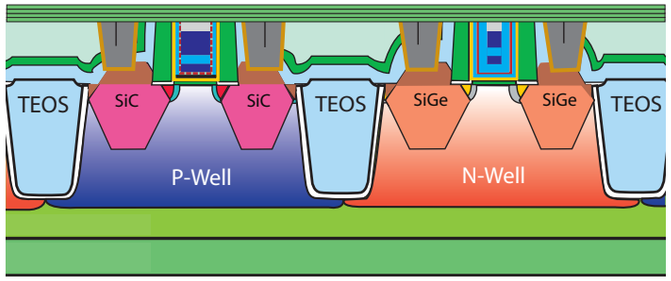
The Gate electrodes have been polished back to a thickness of ~700Å.

© 2018 by Threshold Systems Inc. All Rights Reserved.

© 2018 by Threshold Systems Inc. All Rights Reserved.

ThresholdSystems

Device Module:
Gate Last



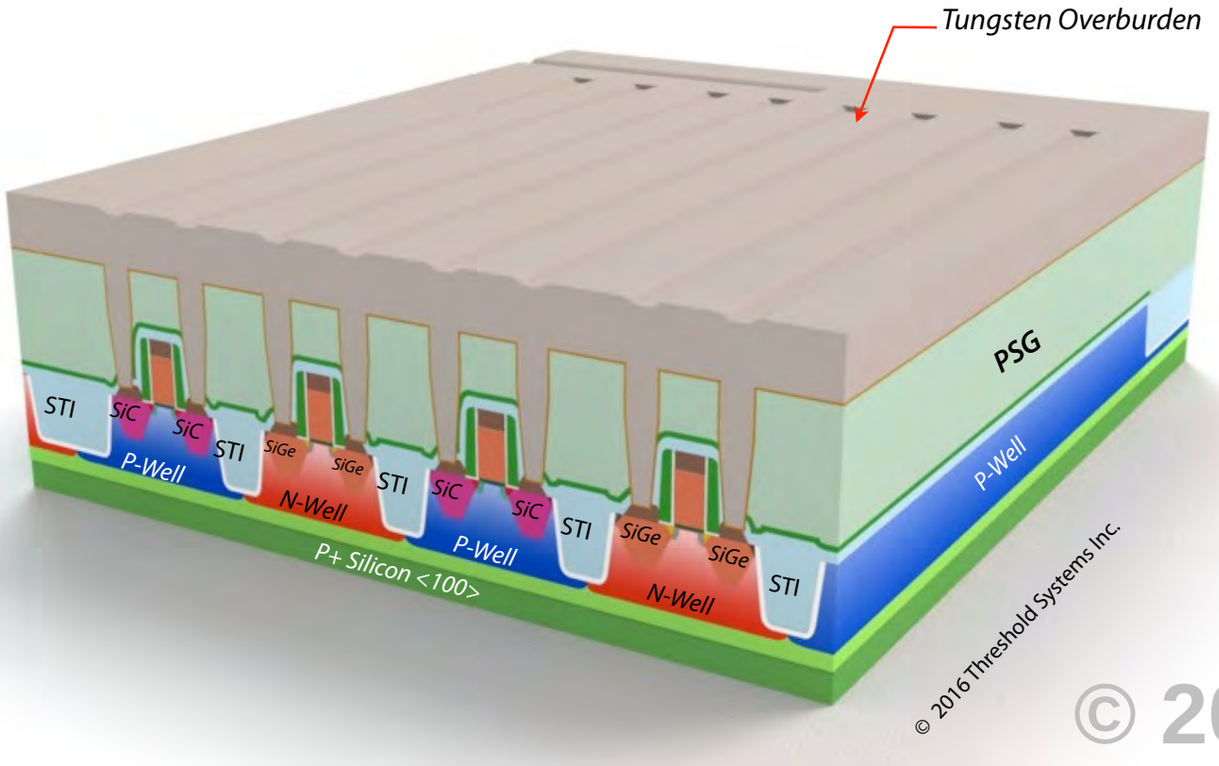
© 2016 Threshold Systems Inc.

© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Polished Back Trench Contacts:

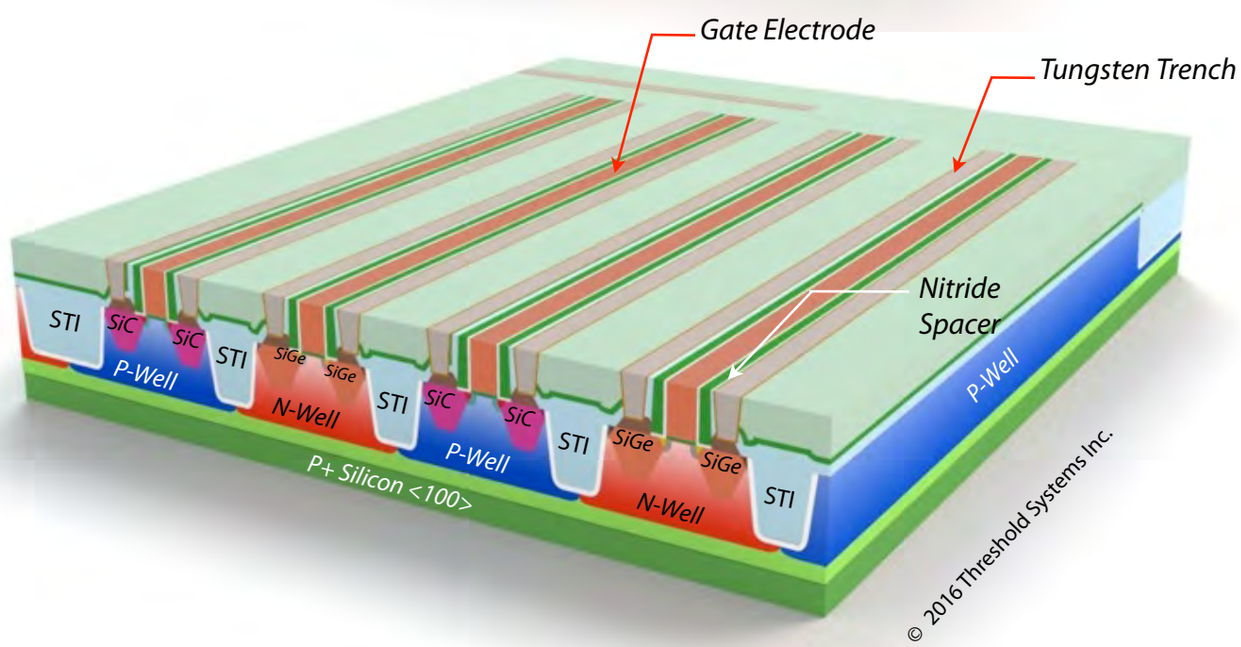
ThresholdSystems



Polished-back trenches are initially formed in the same manner as standard tenches. After the Ti/TiN barrier metals are deposited a Tungsten seed layer is deposited followed by the bulk CVD deposition of 2,500Å of Tungsten.

The Tungsten overburden is then polished back until the trenches are so short that they are coplanar with the tops of the gate electrodes.

© 2018 Threshold Systems Inc.

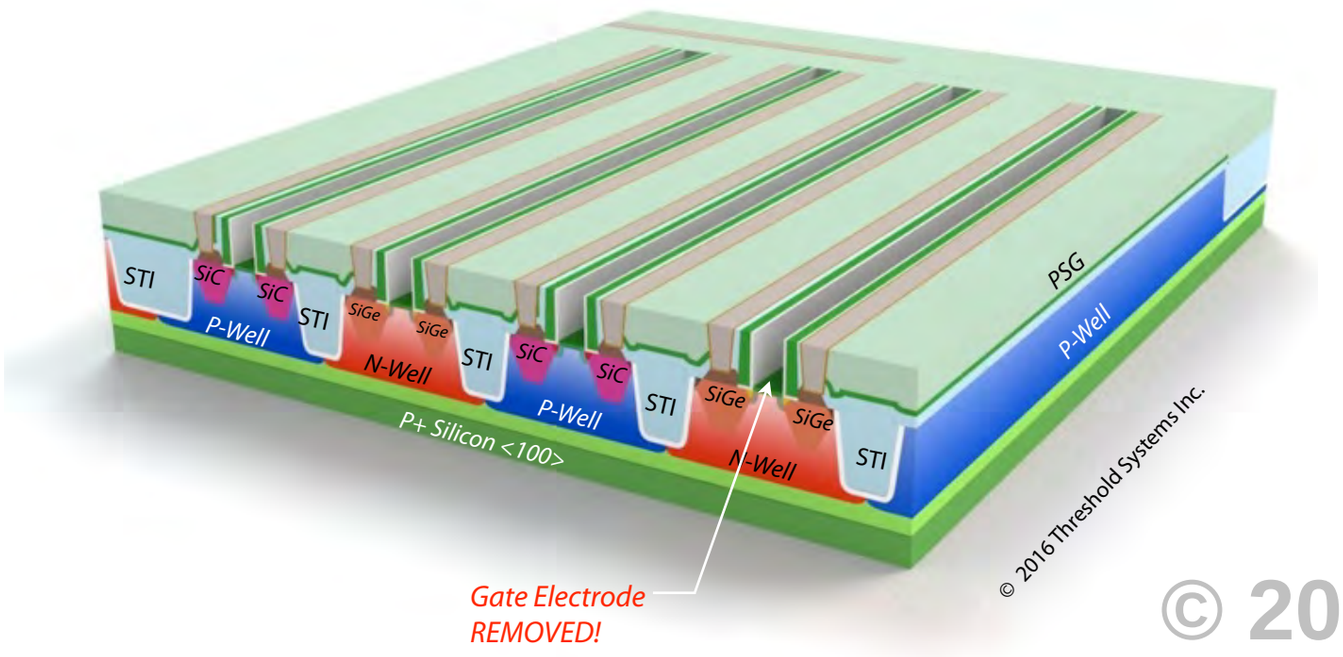


Polishing the Tungsten trenches down to be coplanar with the tops of the gate electrodes is relatively new and was first introduced at the 32nm node. Shorter Tungsten trenches have a lower contact resistance. This feature, coupled with the raised Source/Drains for both the NMOS and PMOS transistors, ensure that the contact resistance is kept to acceptable values.

The Tungsten trenches cannot be eliminated altogether and replaced with Copper plugs because it is too dangerous to bring the Copper lines all the way down to the Source/Drains because Copper is a transistor poison.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Gate Electrode Removal:

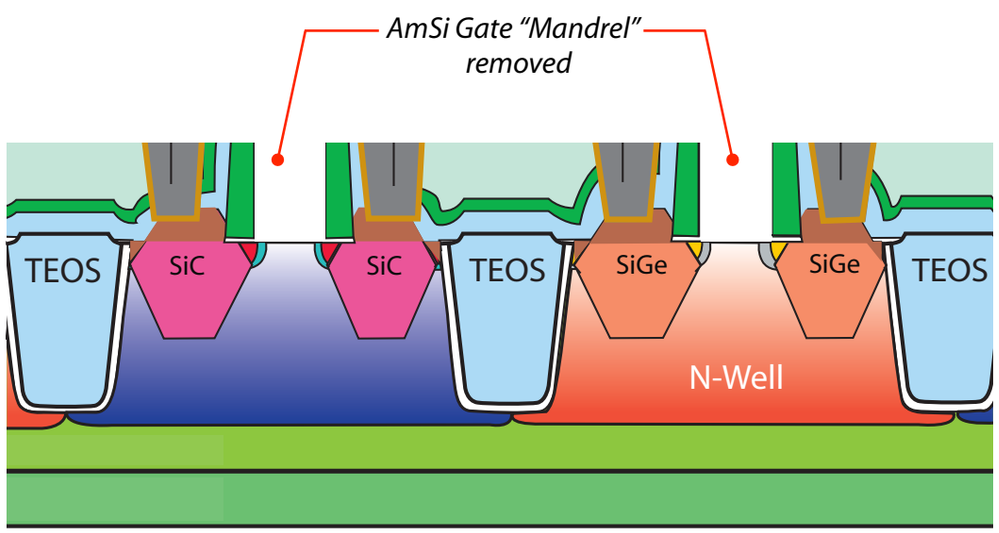


Gate Electrode Removal:

Once the top of the gate electrode has been opened-up, the doped amorphous silicon gate mandrel is etched away. Thus, the gate electrode is said to be "sacrificial". Its only purpose was as a place-holder, commonly referred to as a "Mandrel".

This etch also removes the oxide offset spacer that was grown on the sidewalls of the gate cavity as well as the oxide etch stop layer at the bottom of the gate cavity.

© 2018 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

threshold systems

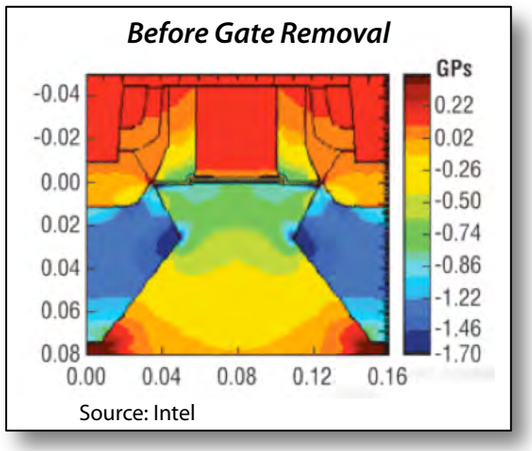
Gate Mandrel Removal:

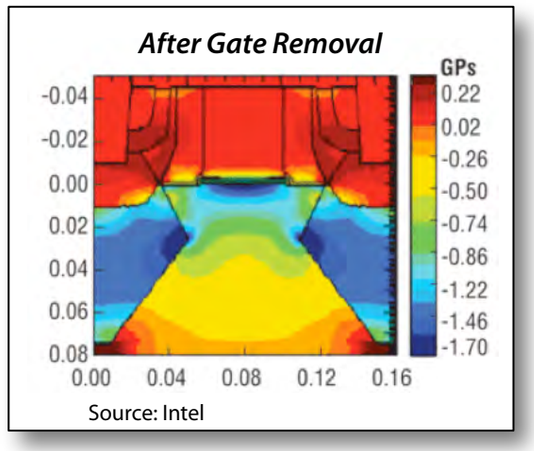
The adjacent illustration displays a cross-sectional view of the CMOS structure after the top of the gate electrode has been opened-up by the CMP polish and the sacrificial amorphous silicon gate mandrel etched away.

This etch also removes the oxide offset spacer that was grown on the sidewalls of the gate cavity as well as the oxide etch stop layer at the bottom of the gate cavity.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Replacement Gate Process:



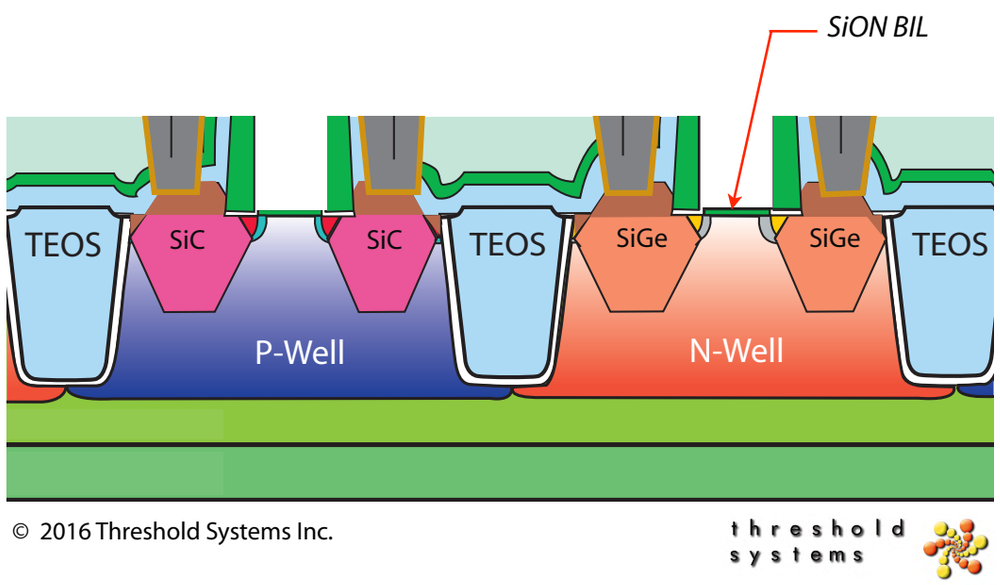


Compressive Strain Enhancement:

The removal of the sacrificial polysilicon gate electrode dramatically increased the amount of compressive strain in the PMOS channel region. This occurs because the channel is already under great compressive strain from the SiGe Source/Drains, and when the polysilicon is removed there is less structural opposition to these compressive forces.

These factors, combined with the fact that the SiGe has been moved closer to the gate edge, increased the PMOS drive current substantially.

© 2018 Threshold Systems Inc.

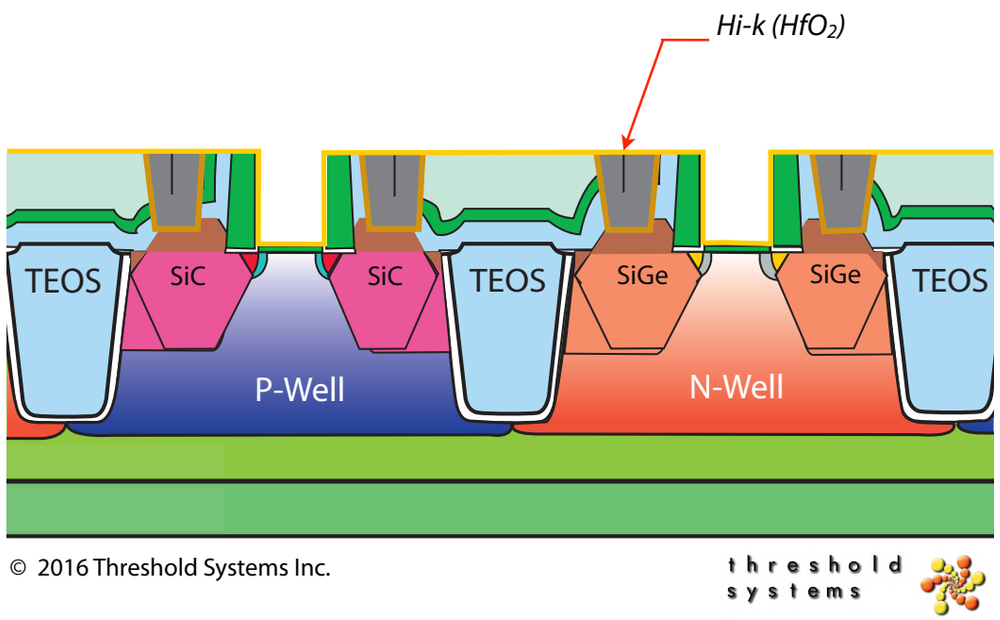


Bottom Interface Layer (BIL) growth:

Next, a layer of silicon dioxide is carefully grown and then nitrated to form a thin ~10Å Bottom Interface Layer (BIL) of Silicon Oxynitride (SiON). The formation of this layer is performed under very strict conditions to ensure that a high mobility interface is maintained between the SiON and the silicon surface.

© 2018 by Threshold Systems Inc. All Rights Reserved.

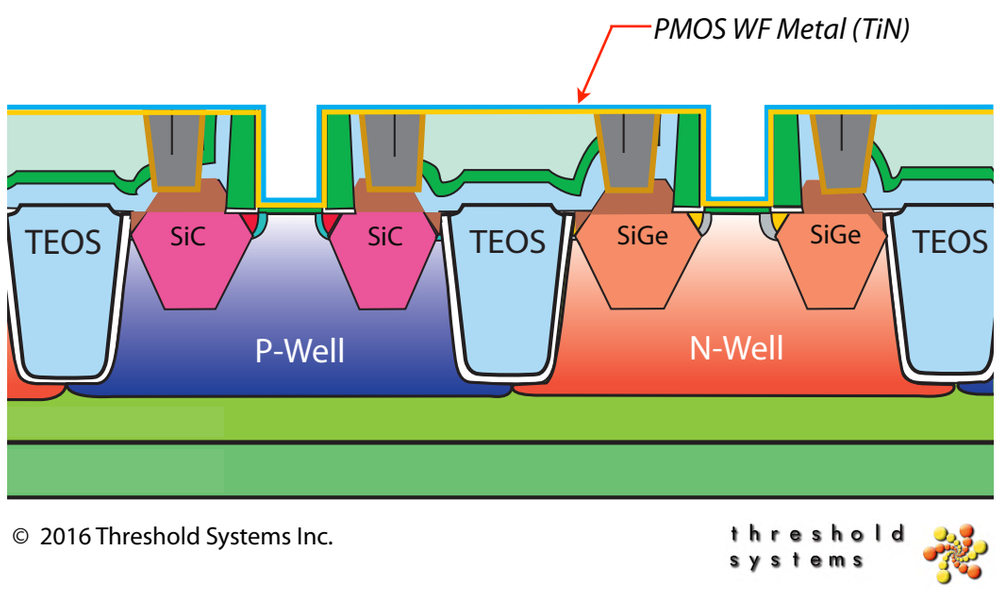
Hi-K Dielectric & PMOS WF Metal Deposition:



Hi-k Deposition:

Next, a thin ~18Å thick layer of Hafnium oxide (HfO_2) is deposited using ALD.

© 2018 Threshold Systems Inc.



PMOS WF Metal Deposition:

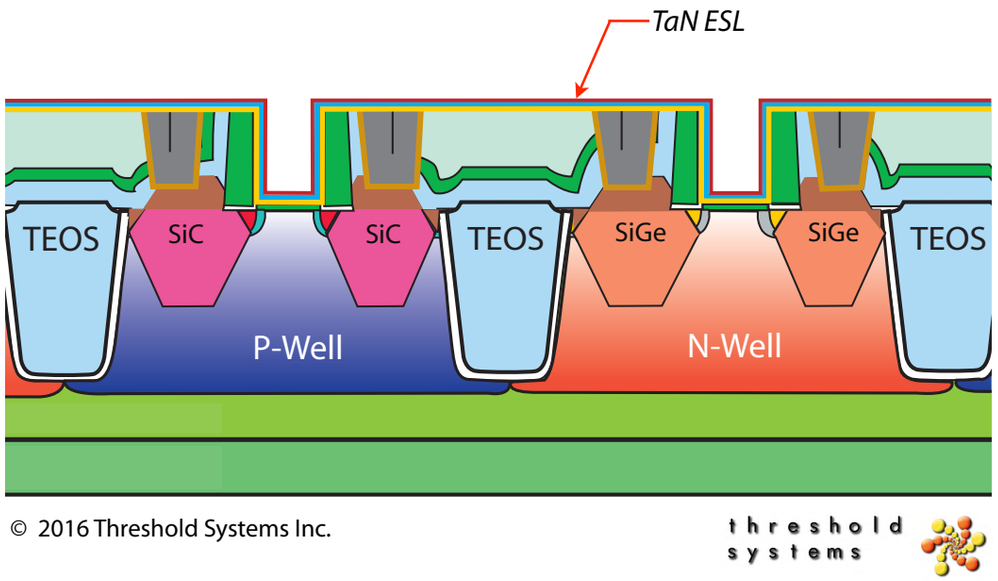
Next, the thin PMOS metal gate is deposited using PVD.

It consists of a thin 2.0 nm layer of TiN is deposited using PVD. This is the PMOS Work Function (WF) metal. It is a conformal layer of TiN that fills both the PMOS and NMOS cavities as well as coats the surface of the wafer.

The TiN is re-spattered after deposition to ensure good sidewall coverage inside of the gate cavity.

© 2018 by Threshold Systems Inc. All Rights Reserved.

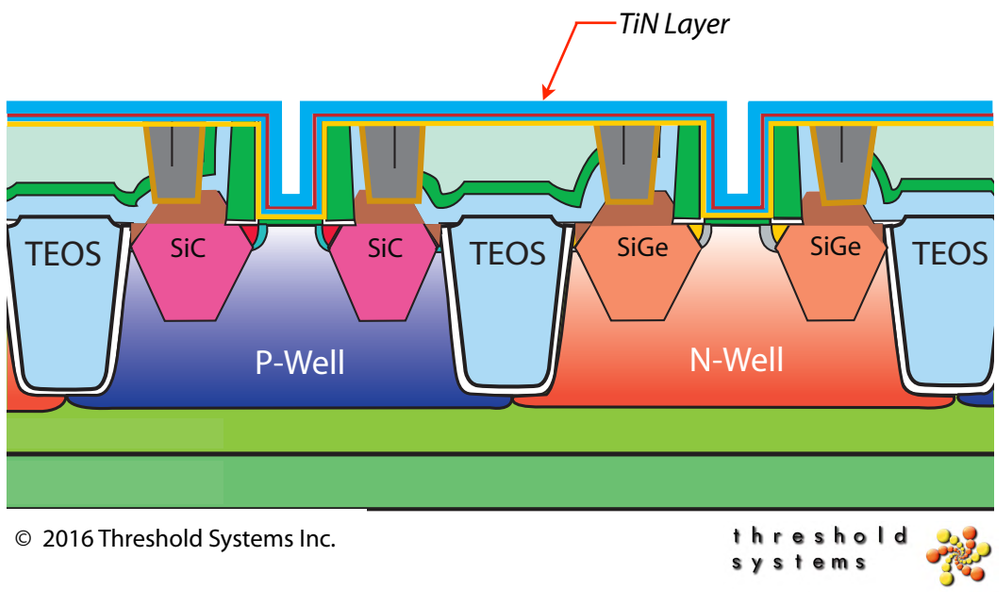
TaN & TiN Metal Deposition:



TaN Etch Stop Layer Deposition:

A ~1.0 nm layer of TaN is deposited using PVD. This layer will serve as an Etch Stop layer (ESL).

© 2018 Threshold Systems Inc.



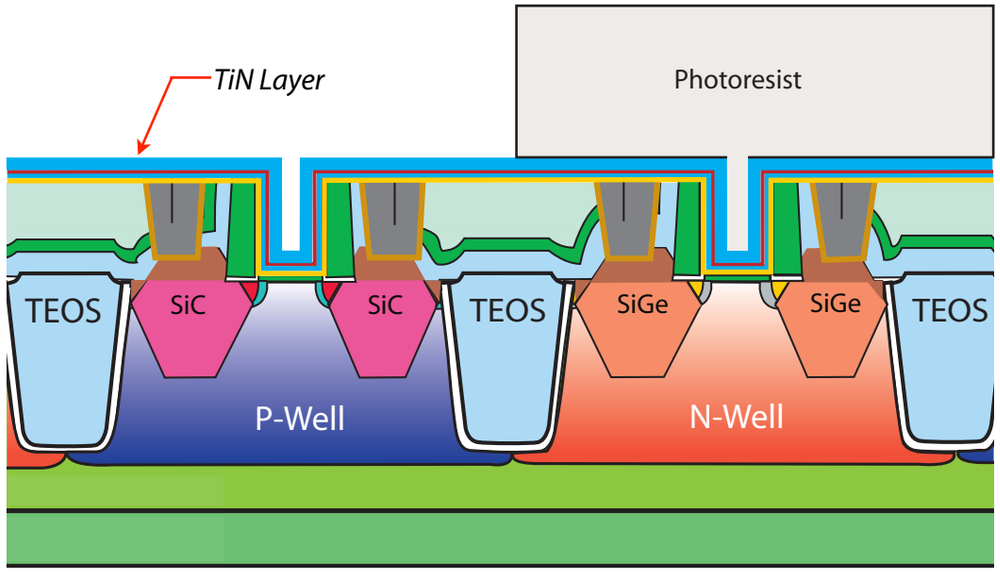
TiN Layer Deposition:

A ~5.0 nm layer of TiN is deposited using PVD. This is the first layer of back-fill metals in the gate stack that will provide strain and electrical contact to the bottom of the gate stack.

The TiN is re-sputtered to ensure good sidewall coverage inside of the gate cavities.

© 2018 by Threshold Systems Inc. All Rights Reserved.

PMOS Metal Patterning & Etch:

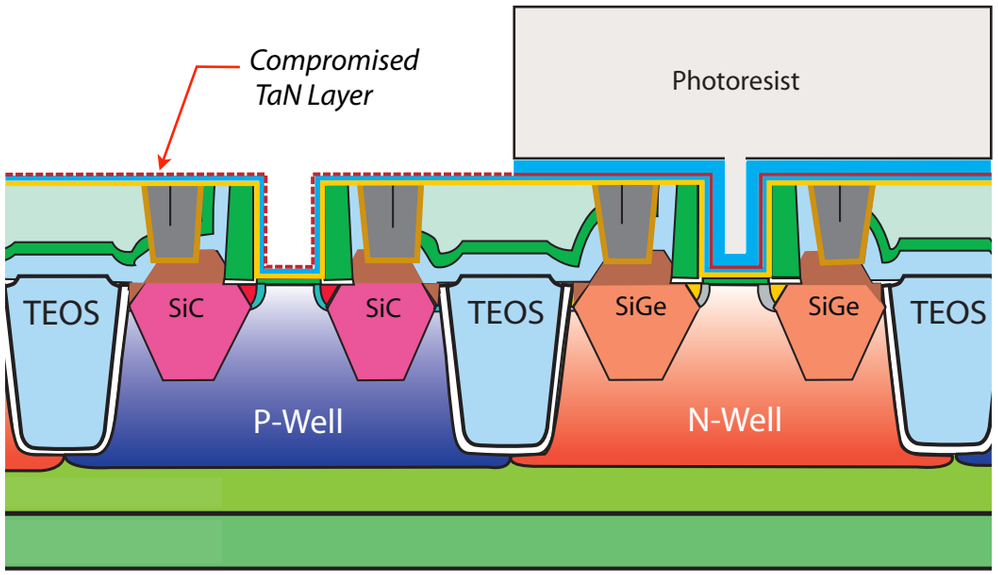


© 2016 Threshold Systems Inc.

TiN Layer Photoresist Patterning:

A layer of photoresist is spun-on and patterned. It protects only the PMOS region and exposes the TiN layer over the NMOS region.

© 2018 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

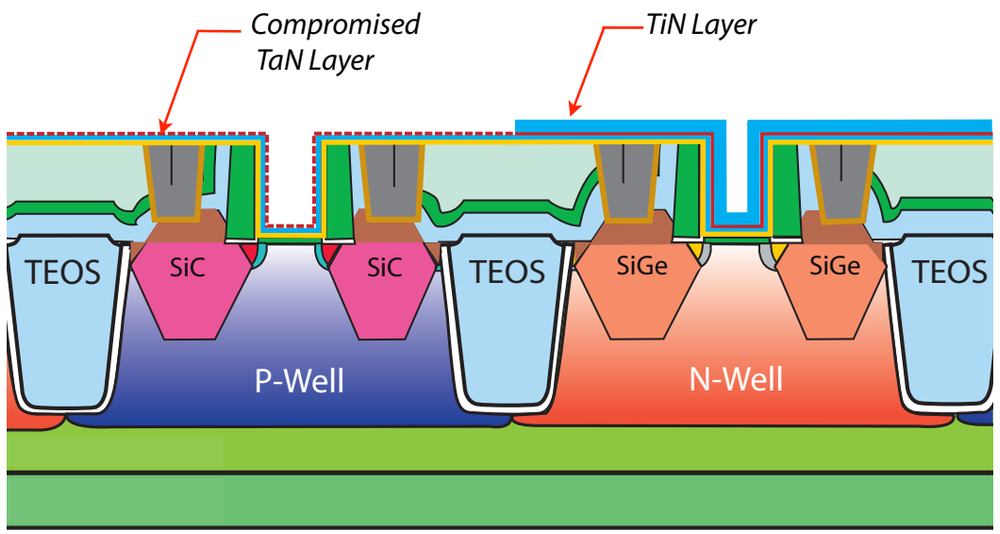
TiN Layer Etch:

The exposed TiN layer is etched away. The TaN acts as an etch-stop layer to this etch. The integrity of the very thin (<10Å) TaN layer is compromised by the TiN etch as indicated.

The fact that the integrity of the TaN layer is compromised is helpful when we diffuse Aluminum into the underlying TiN layer in the NMOS transistor to form the NMOS TiAlN Work Function metal later in the process.

© 2018 by Threshold Systems Inc. All Rights Reserved.

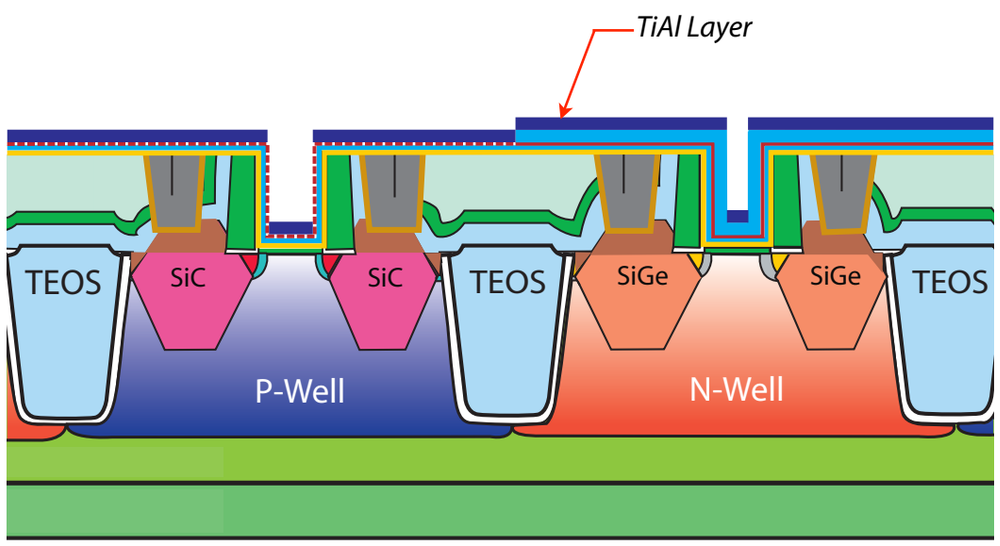
Photoresist Strip & TiAl Deposition:



© 2016 Threshold Systems Inc.

Photoresist Strip:
The photoresist is stripped off and the wafers are cleaned.

© 2018 Threshold Systems Inc.



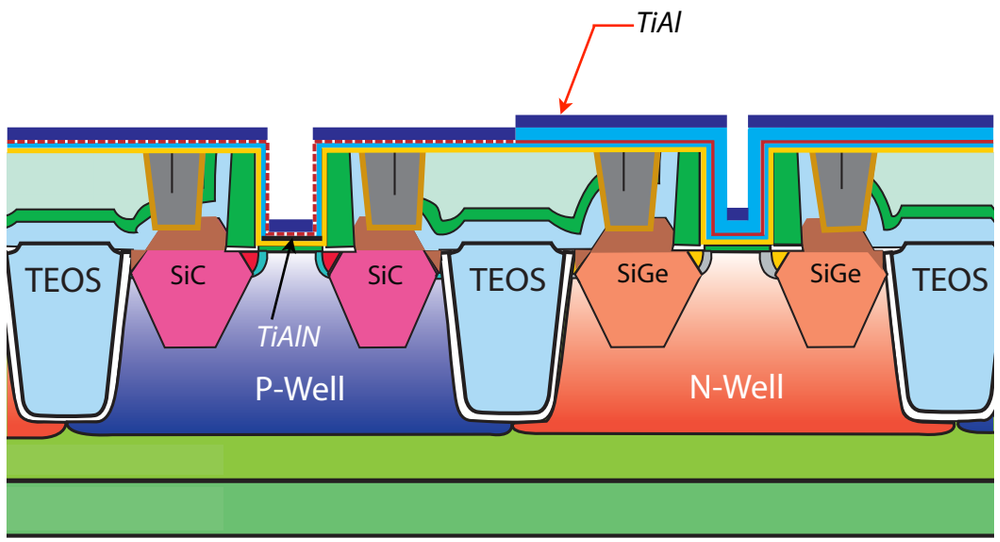
© 2016 Threshold Systems Inc.

TiAl Layer Deposition:

A thin 4.8 nm layer of Titanium Aluminum is deposited using PVD. The PVD process is tuned to deposit only on horizontal surfaces (this is easy to do as you have to really tweak the PVD process to get it to deposit on vertical surfaces).

© 2018 by Threshold Systems Inc. All Rights Reserved.

TiAl Anneal & NMOS WF Metal Formation:



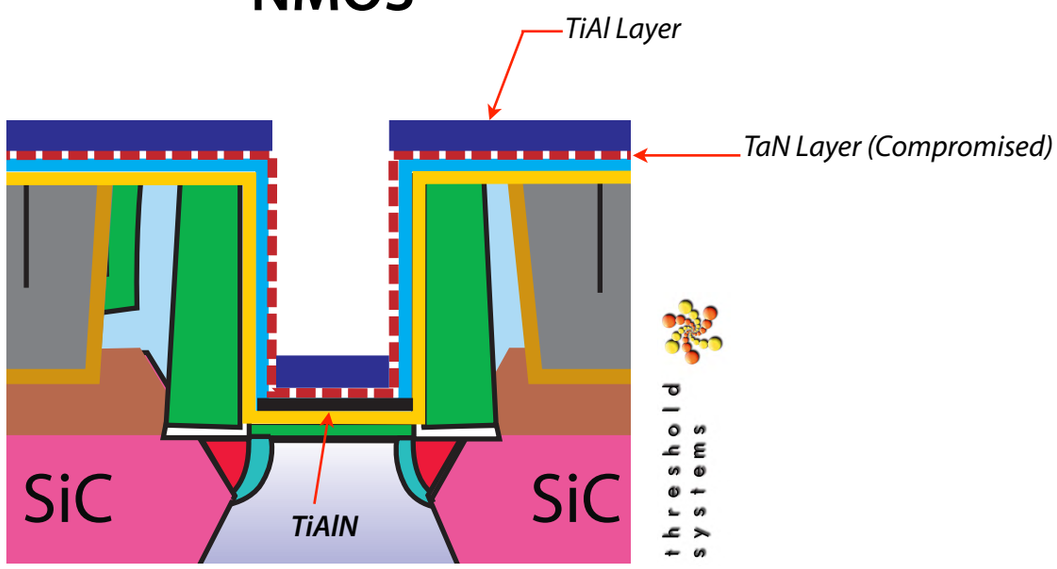
© 2016 Threshold Systems Inc.

Anneal:

An anneal is then performed that causes the Aluminum in the TiAl, in the NMOS region only, to diffuse through the TaN barrier. This creates the TiAlN NMOS Work Function metal located on top of the Hi-k dielectric.

© 2018 Threshold Systems Inc.

NMOS



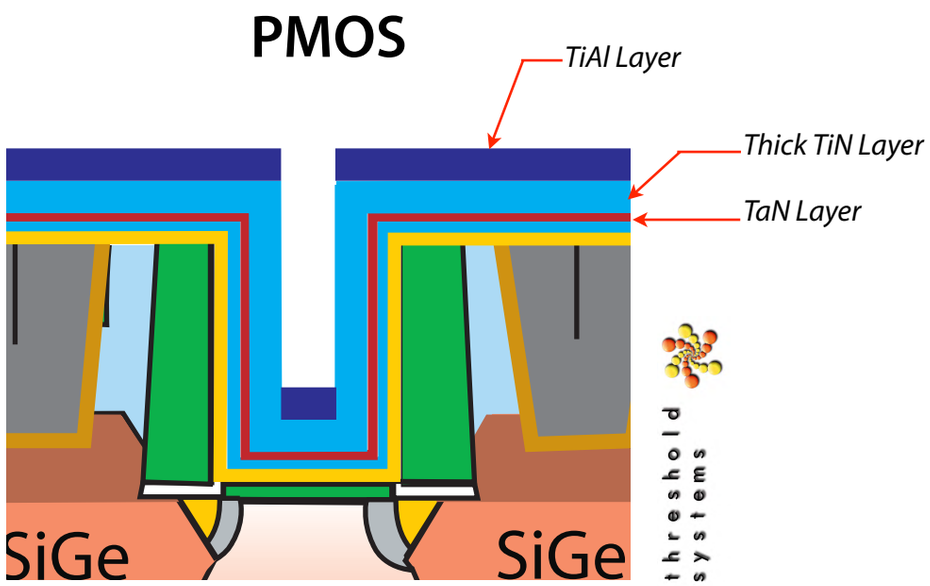
© 2016 Threshold Systems Inc.

Anneal:

The anneal that causes the formation of the TiAlN Work Function metal is performed at a temperature below 400 ° C. Aluminum is very diffusive at this temperature and since the TaN which separates it from the underlying TiN in the NMOS device was previously compromises (it was used as an etch stop layer) the Aluminum freely diffuses into the TiN and forms TiAlN which is the desired NMOS Work Function metal.

© 2018 by Threshold Systems Inc. All Rights Reserved.

TiAl Backfill Deposition:



© 2016 Threshold Systems Inc.

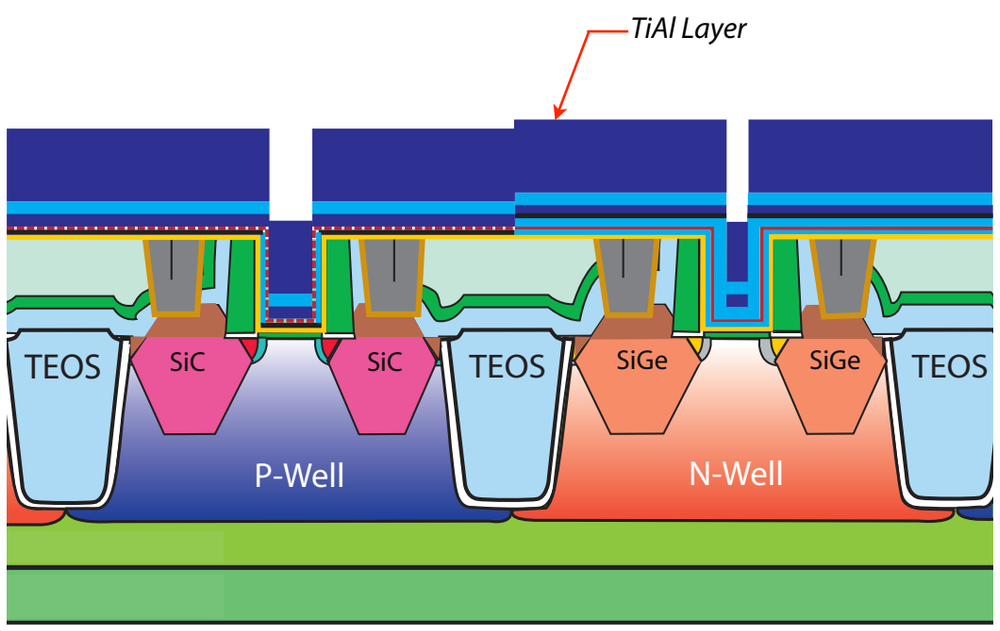
Anneal:

During the anneal that forms the NMOS Work Function metal the thick overlaying TiN layer (which is not present in the NMOS transistor) and the uncompromised TaN layers block the diffusion of Aluminum into the TiN PMOS work function metal. Thus, the TiN PMOS Work Function metal does not become TiAlN. It remains TiN, which is what is desired.

NMOS WF Metal: **TiAlN**

PMOS WF Metal: **TiN**

© 2018 Threshold Systems Inc.



© 2016 Threshold Systems Inc.

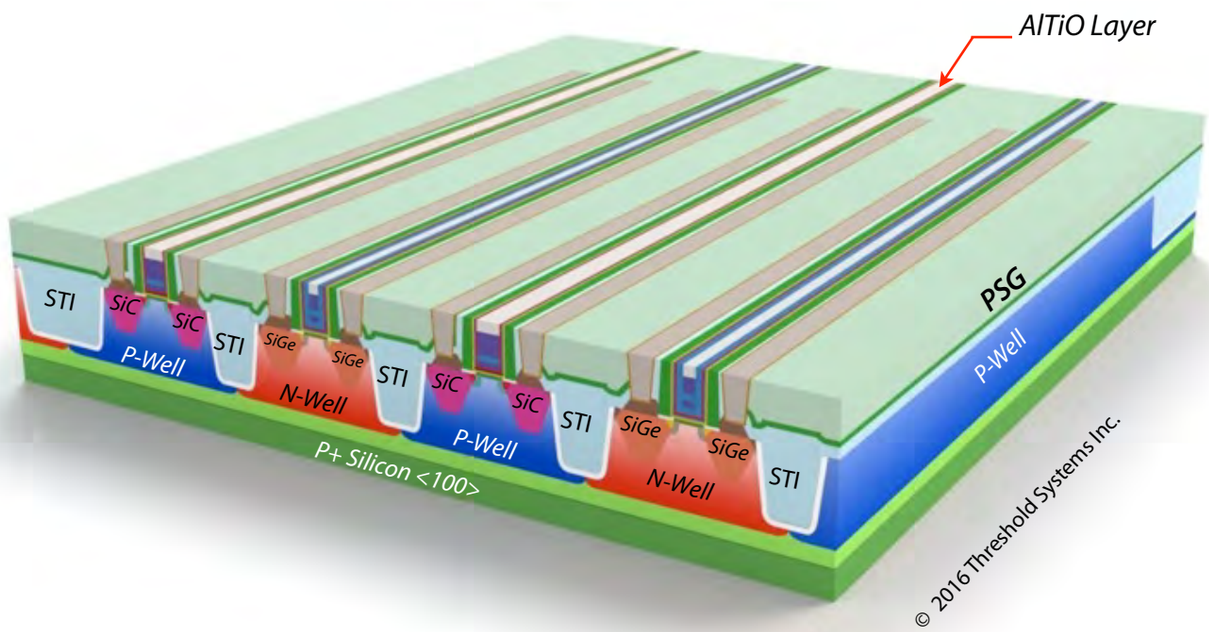
AlTi Layers Deposition:

Two thick layers (11.0nm + 7.0 nm) of Aluminum Titanium are deposited using PVD. The PVD process is tuned to deposit only on horizontal surfaces.

This metal acts as an electrical contact from the top of the gate electrode to the bottom.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Metal Polish-Back:

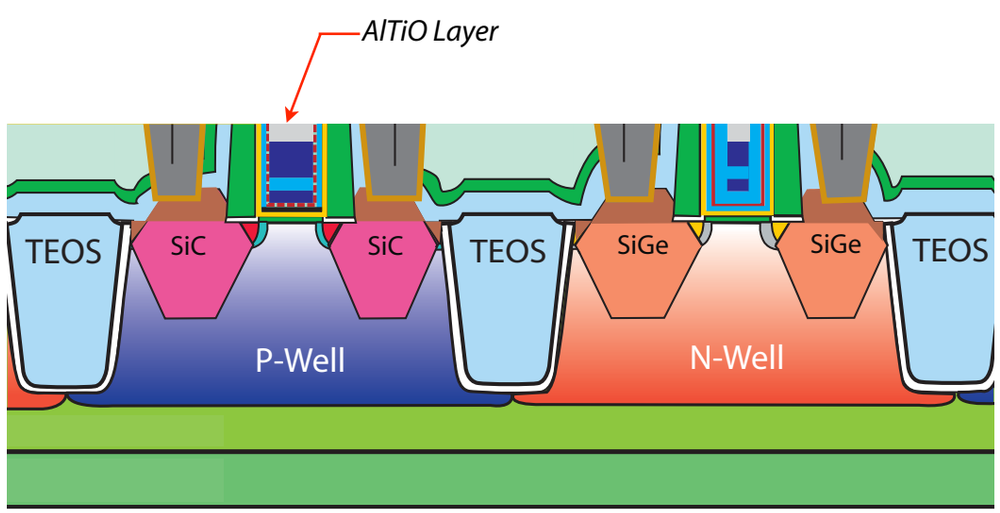


© 2018 Threshold Systems Inc.

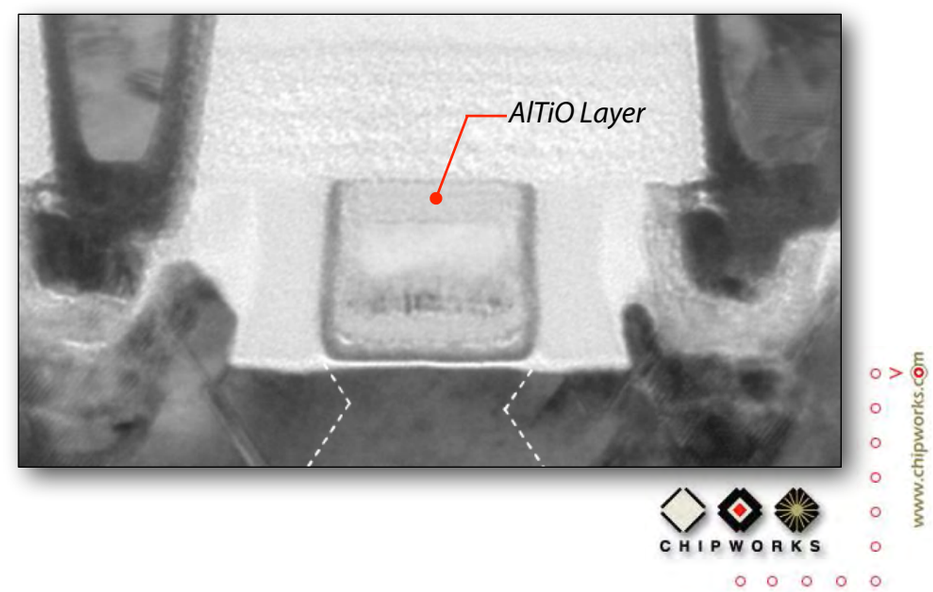
CMP Polish:

The metallization stack is polished back (using CMP) so that it is coplanar with the top of the transistor gate electrodes.

As part of the CMP process the uppermost layer of AlTi is oxidized to form AlTiO.



© 2016 Threshold Systems Inc.

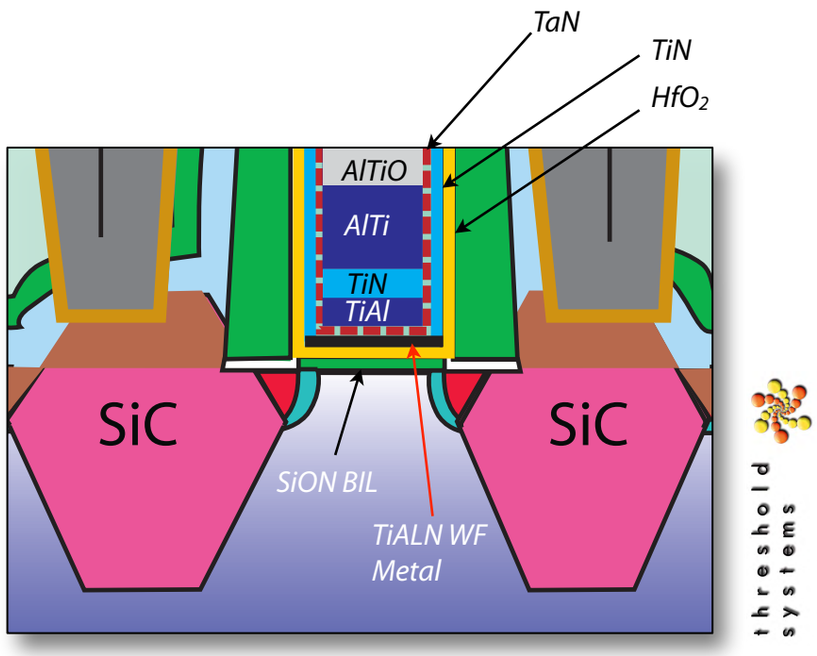


© 2018 by Threshold Systems Inc. All Rights Reserved.

Replacement Gate Anatomy:

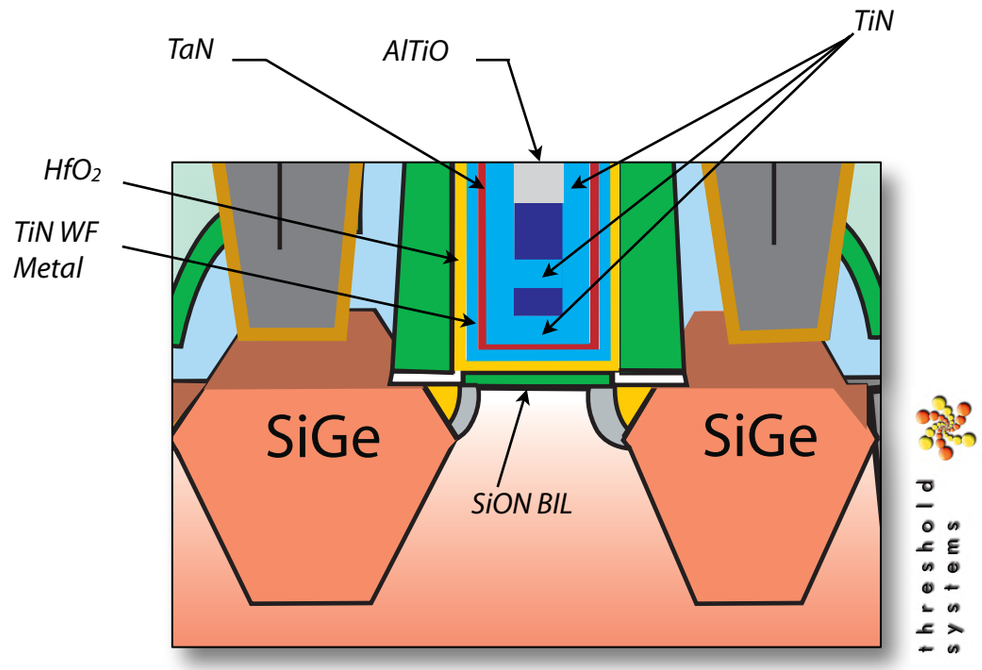
ThresholdSystems

NMOS



© 2016 Threshold Systems Inc.

PMOS



© 2016 Threshold Systems Inc.

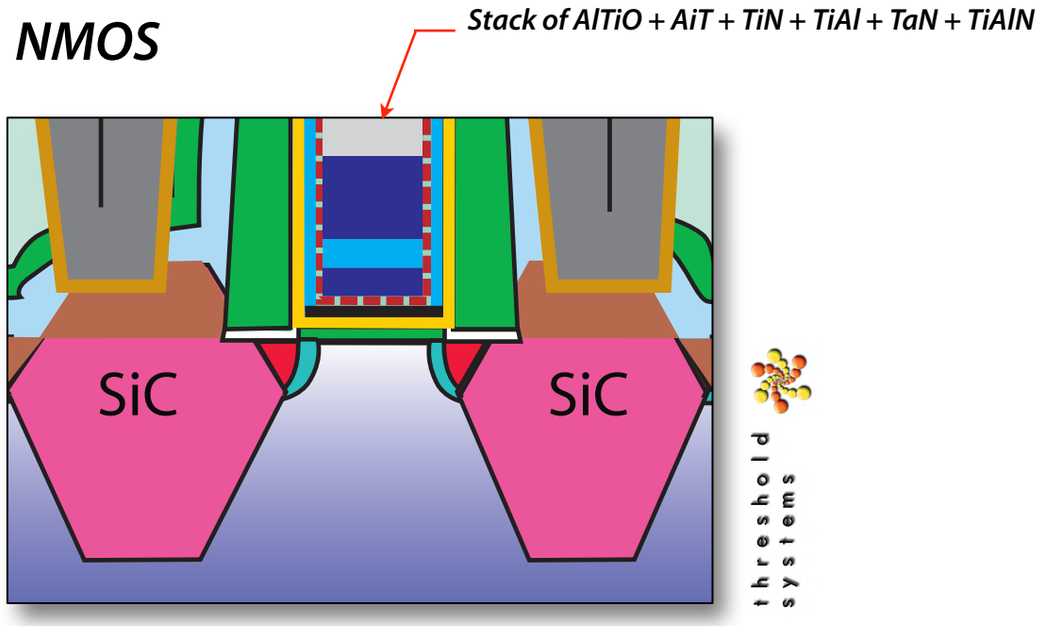
© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Strain Enhancement:

ThresholdSystems

NMOS



© 2016 Threshold Systems Inc.

© 2018 Threshold Systems Inc.

AlTi Layer Deposition:

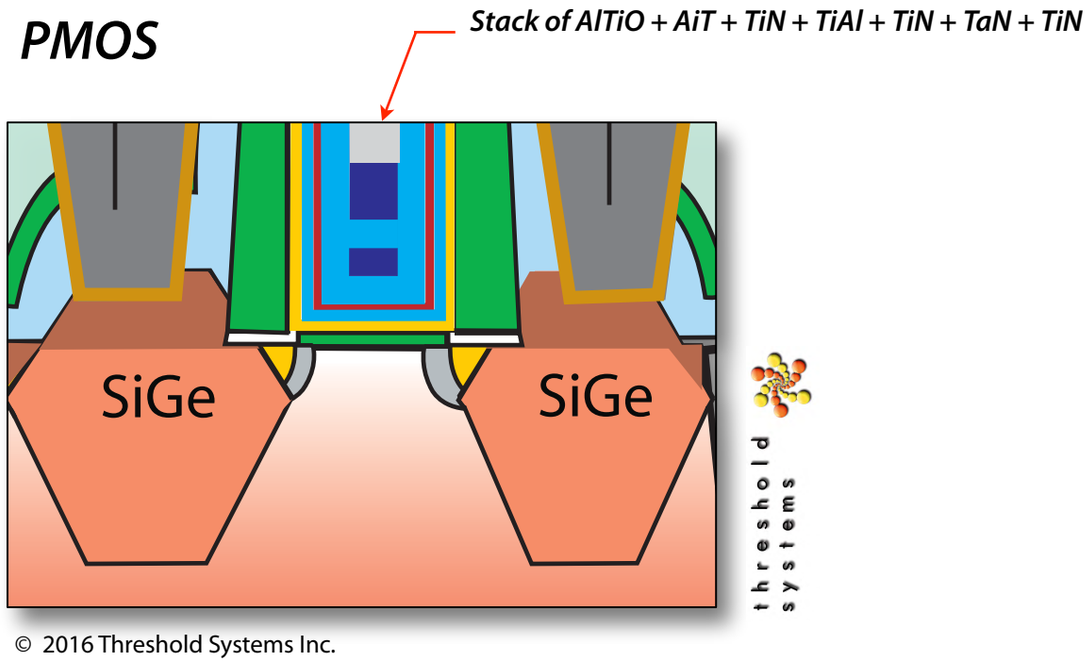
Two thick layers (11.0nm + 7.0 nm) of Aluminum Titanium are deposited using PVD. The PVD process is tuned to deposit only on horizontal surfaces. The composition is 65% Titanium and 35% Aluminum.

NMOS Strain Enhancement:

The metal stack in both the gate cavities is very compressive in the Z-direction. It pushes down on the transistor channel and introduces tensile strain which enhances the drive current of the NMOS device.

This tensile strain in the PMOS device degrades the drive current of the PMOS device somewhat, but the induced tensile force is overwhelmed by the compressive force introduced by the SiGe located in the Source/ Drains.

PMOS

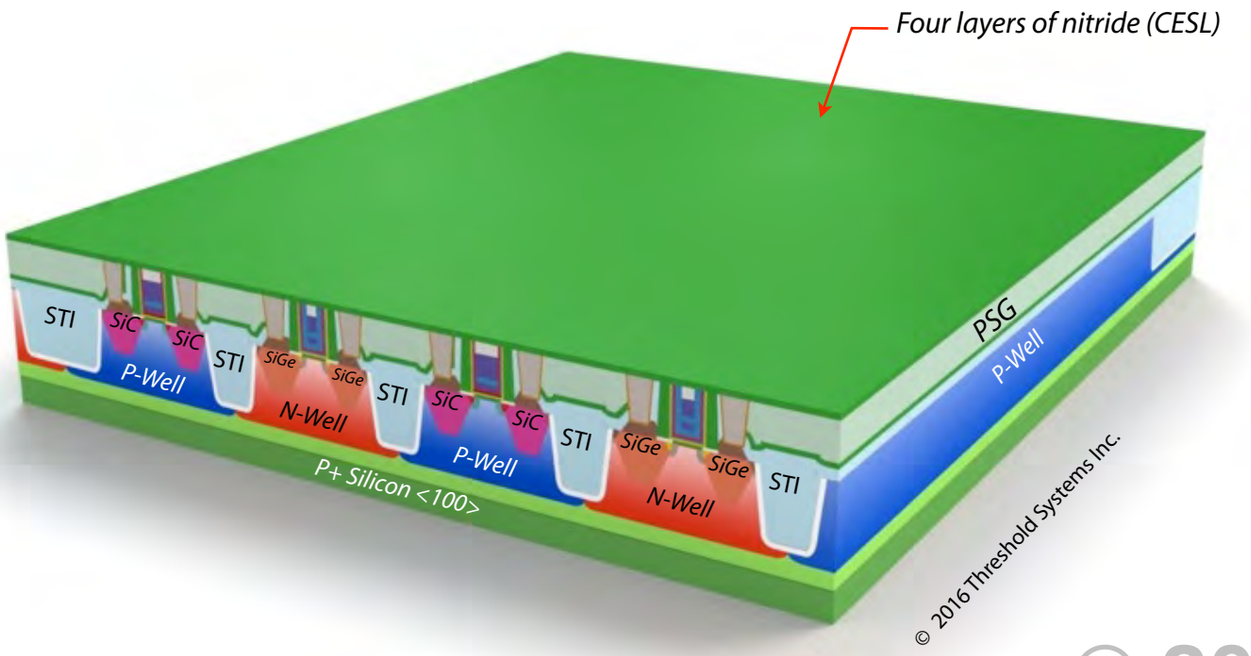


© 2016 Threshold Systems Inc.

This complex metallization structure is what makes the Replacement Gate so difficult!

© 2018 by Threshold Systems Inc. All Rights Reserved.

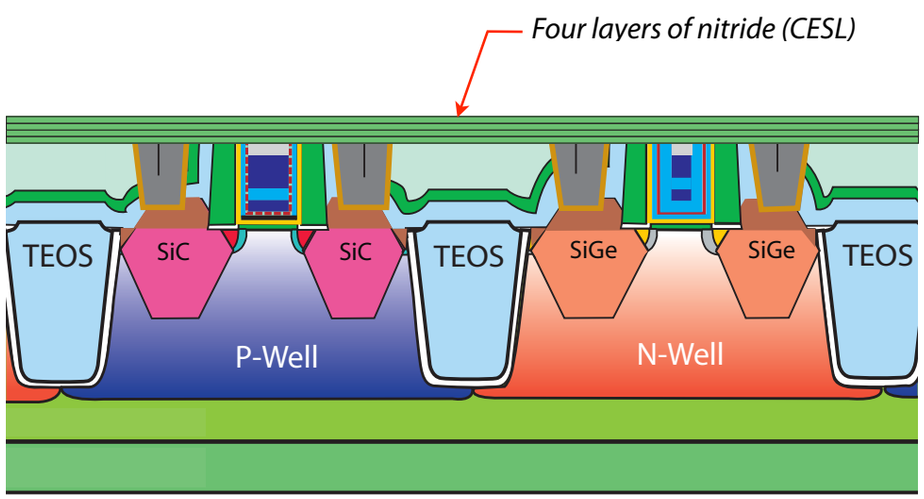
Nitride CESL Deposition:



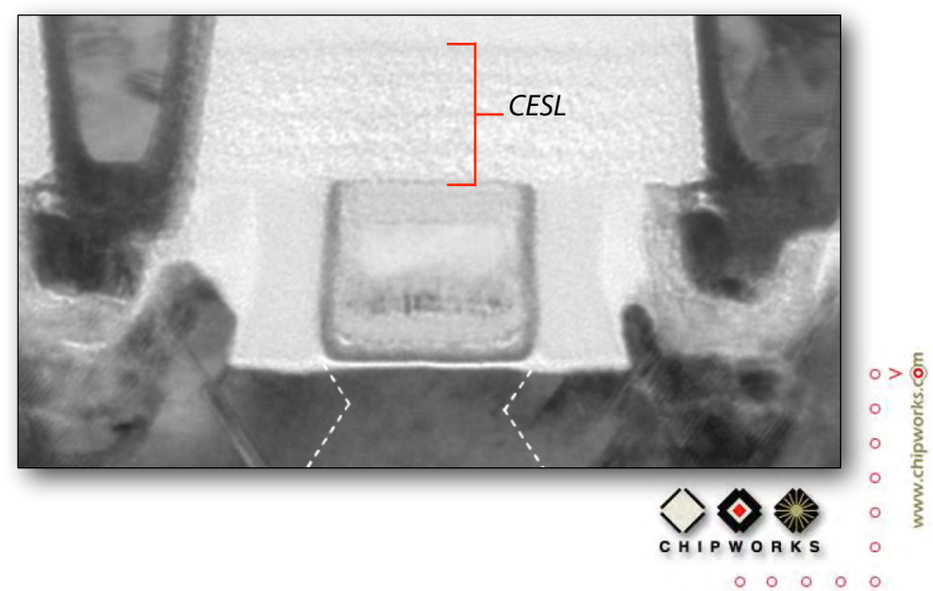
CONTACT ETCH STOP LAYER (CESL):

Next, four layers of Silicon Nitride is deposited over the surface of the wafer. This seals the tops of the transistors and will acts as an etch stop layer for the contact trenches that have yet to be fabricated.

© 2018 Threshold Systems Inc.



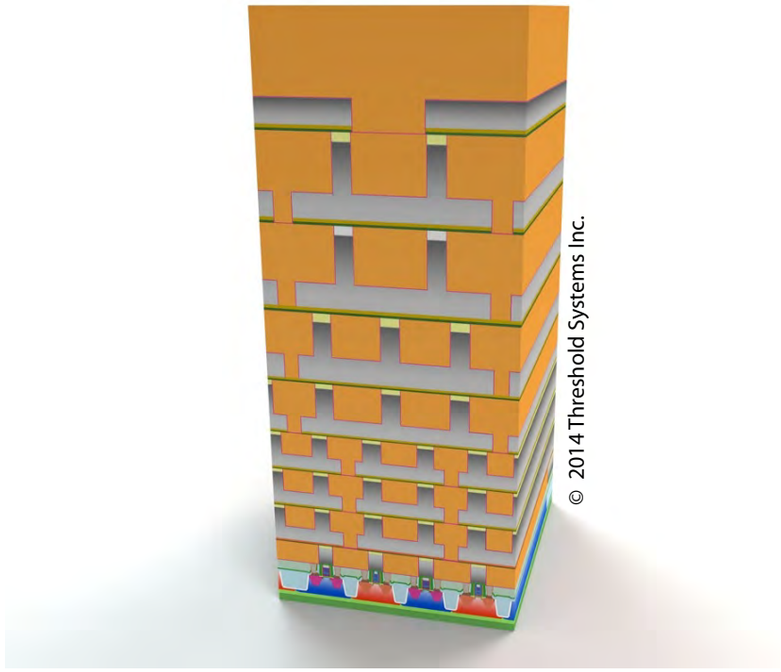
© 2016 Threshold Systems Inc.



© 2018 by Threshold Systems Inc. All Rights Reserved.

ThresholdSystems

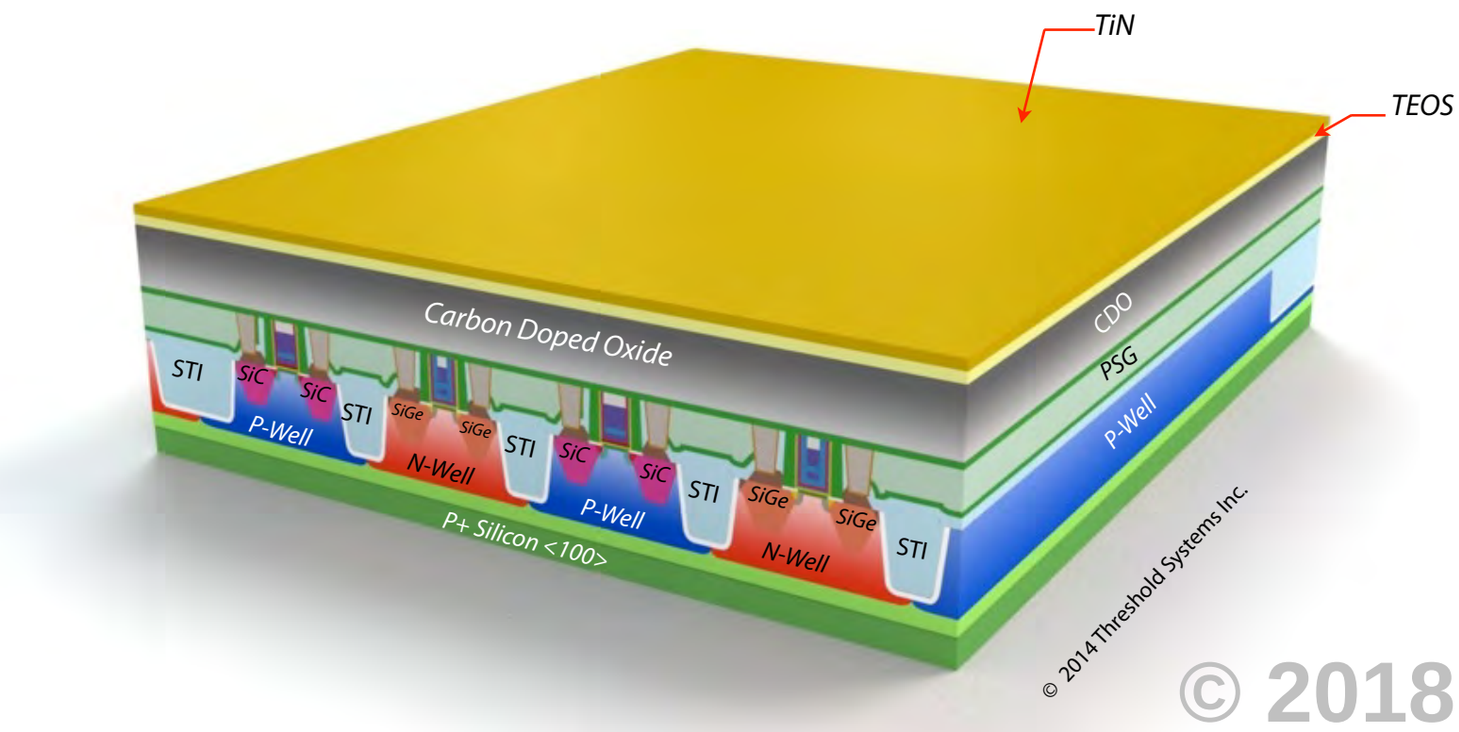
Wafer Fab Backend Metallization



© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

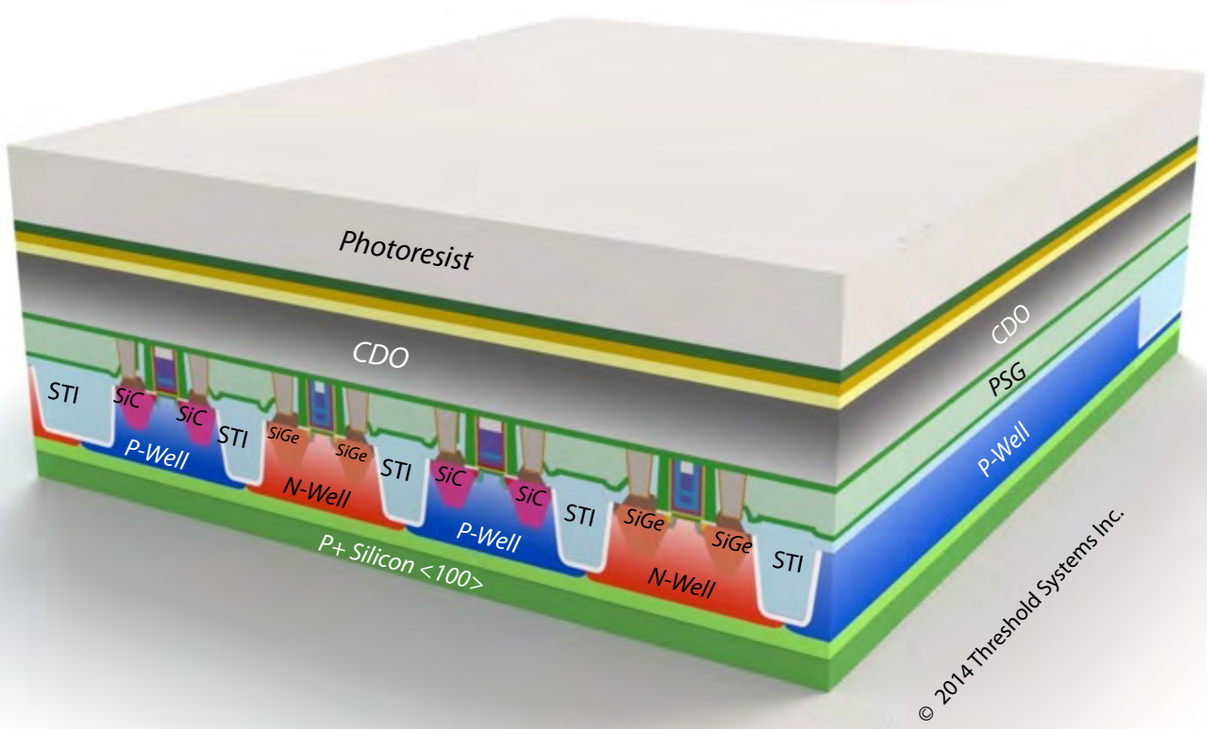
CDO/TEOS/ TiN Deposition:



The first Inter-layer Dielectric (ILD) is deposited. It consists of a ~1,000Å layer of Carbon Doped Oxide (CDO) that is UV cured after deposition to bake out porogens and lower its k value.

This is followed by the deposition of a 400Å TEOS CMP stop layer and followed by a 300Å TiN hard mask.

© 2018 Threshold Systems Inc.

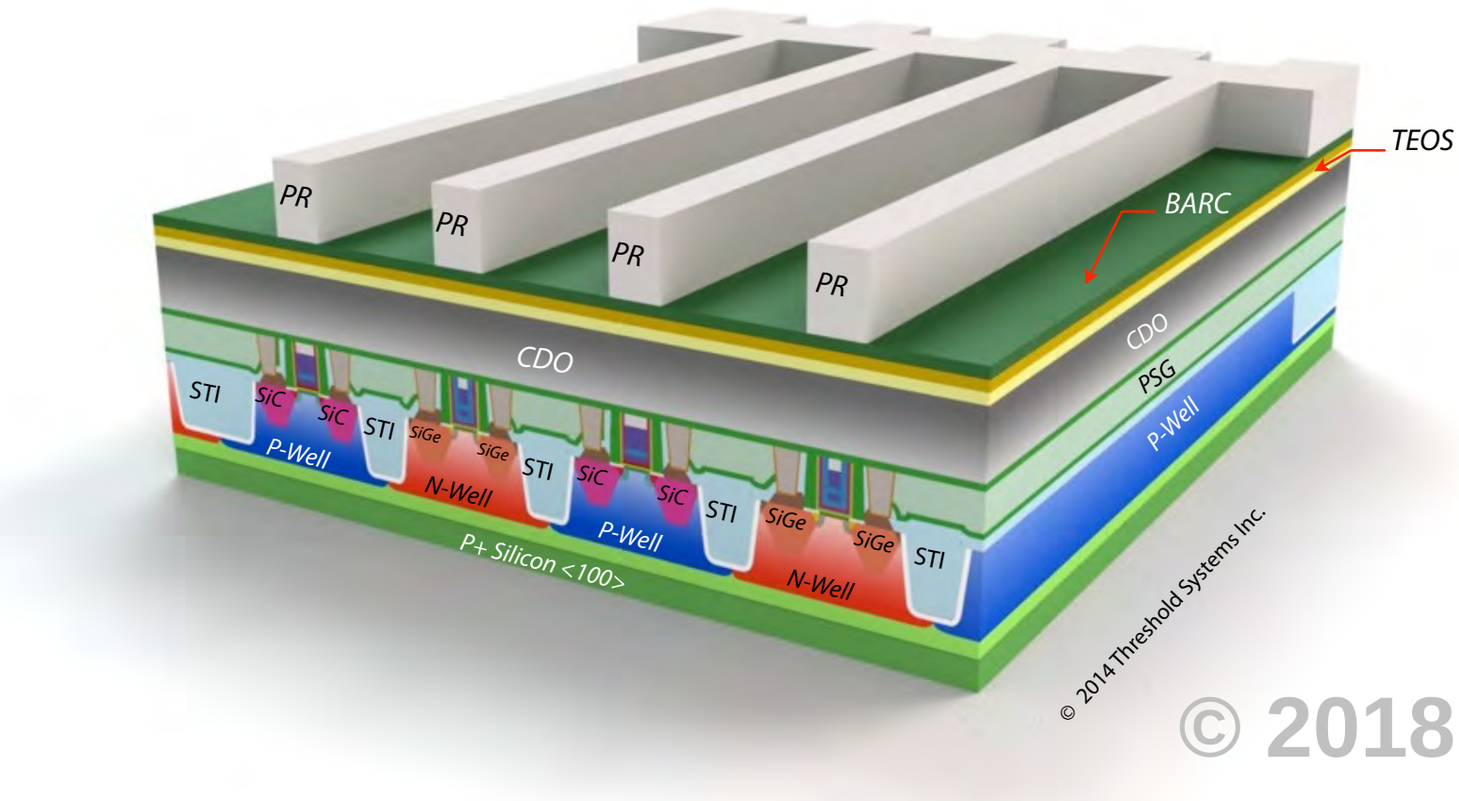


Next, the wafers are primed in HMDS (to promote photoresist adhesion) and coated with Bottom Anti-Reflective Coating (BARC) and a layer of photoresist.

© 2018 by Threshold Systems Inc. All Rights Reserved.

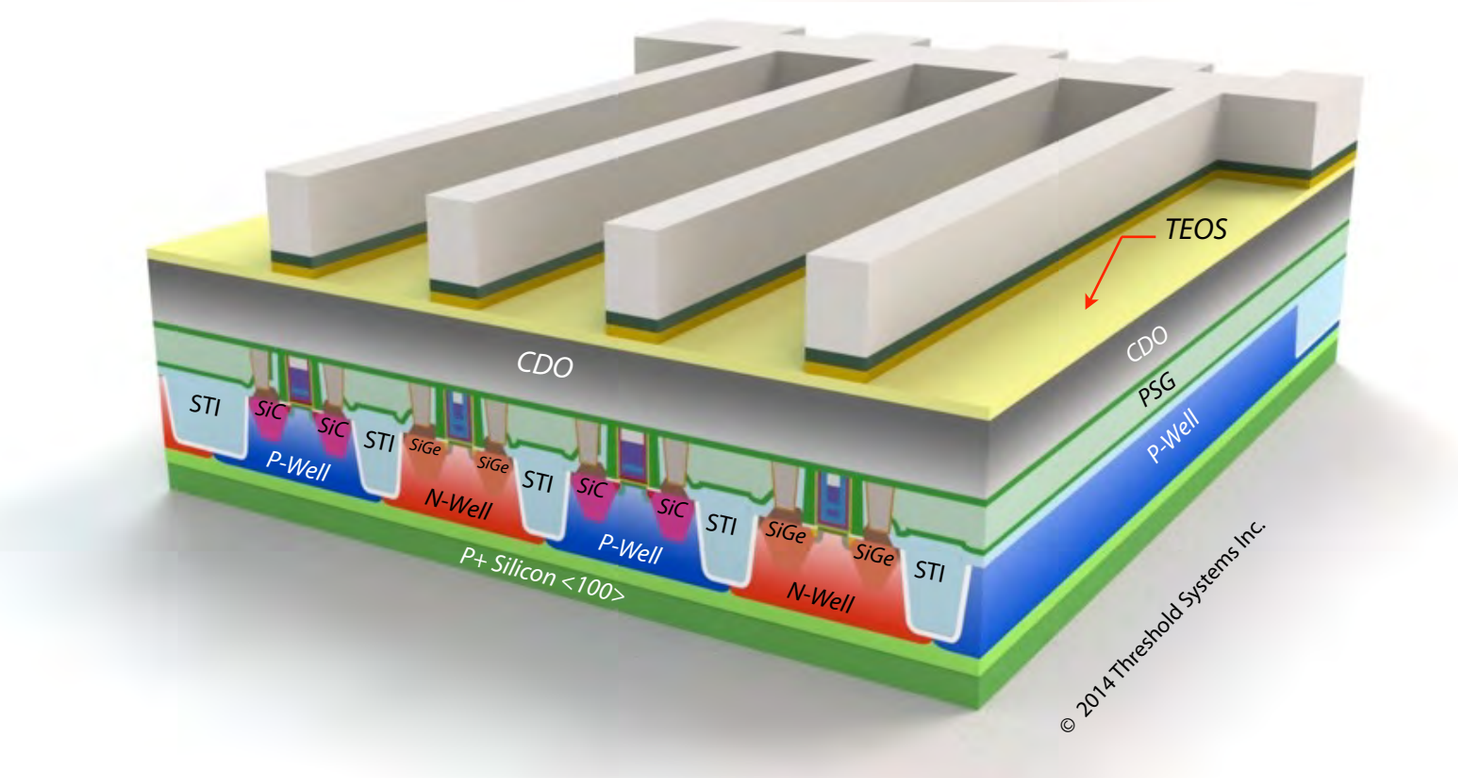
TiN Hard Mask Patterning & Etch:

ThresholdSystems



The wafers are then soft-baked to give the resist a measure of structural integrity, then the resist is exposed and developed. The resulting photoresist pattern then undergoes a Post Exposure UV Bake (PEB) to finalize the photoresist pattern and to turn the photoresist into a cross-linked imidized mass that will resist etching.

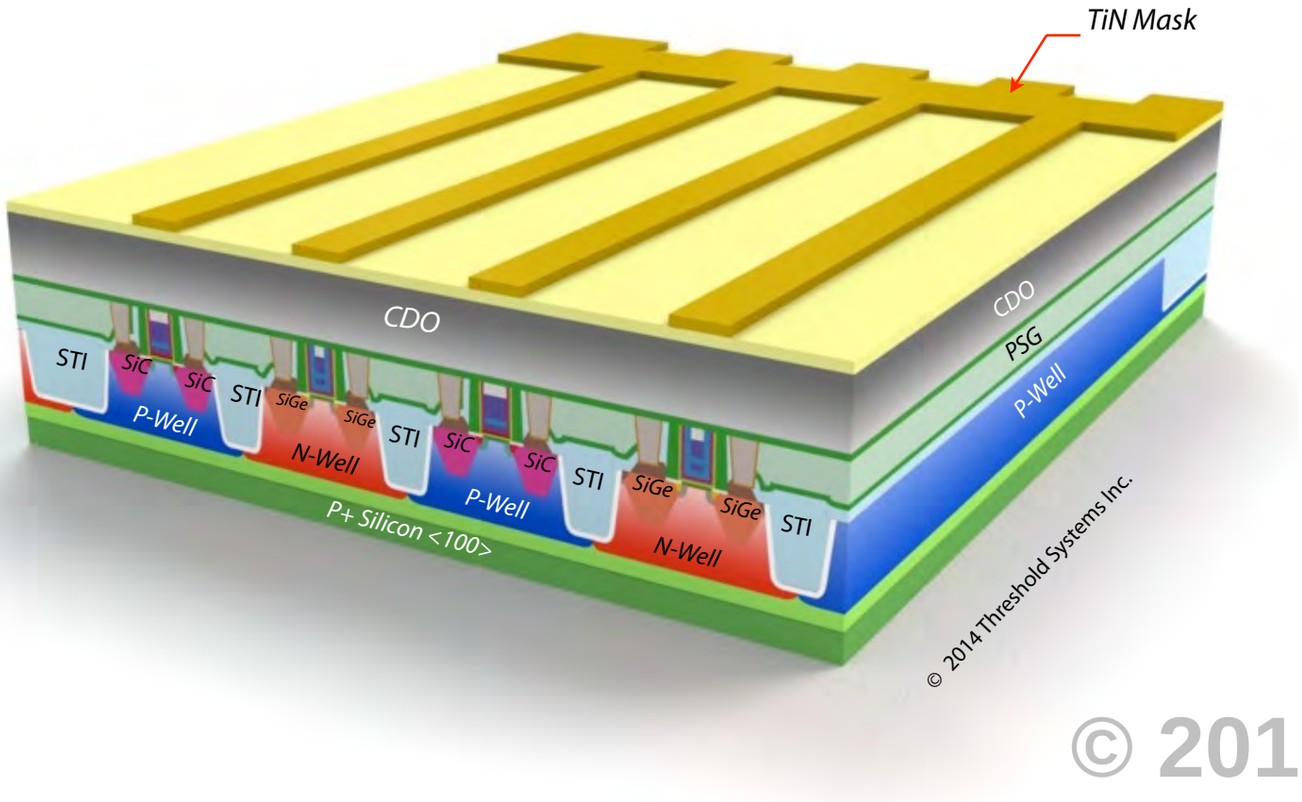
© 2018 Threshold Systems Inc.



The BARC is etched followed by an anisotropic etch of the exposed portion of the TiN hard mask. The etch stops on the TEOS layer.

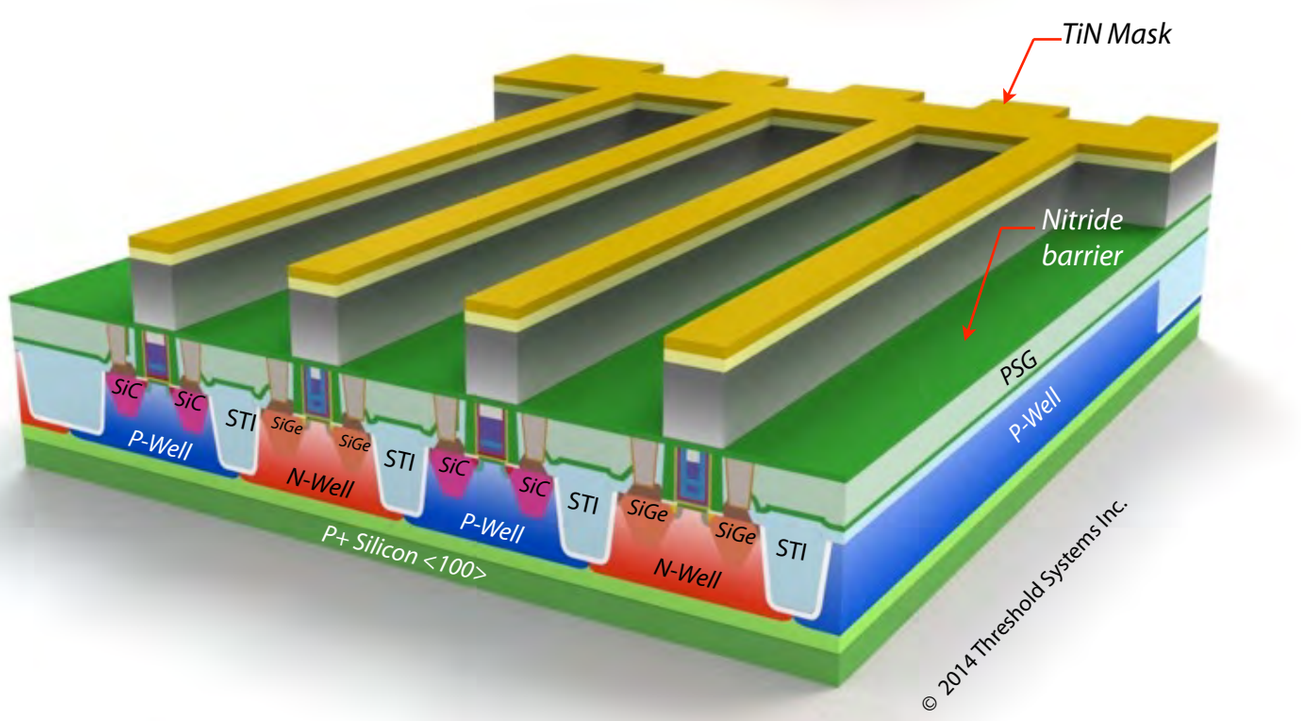
© 2018 by Threshold Systems Inc. All Rights Reserved.

Photoresist Strip & Metal-1 Trench Etch:



© 2018 Threshold Systems Inc.

The photoresist and the BARC are stripped away and the wafers are cleaned. The TiN hard mask has now been fully formed.

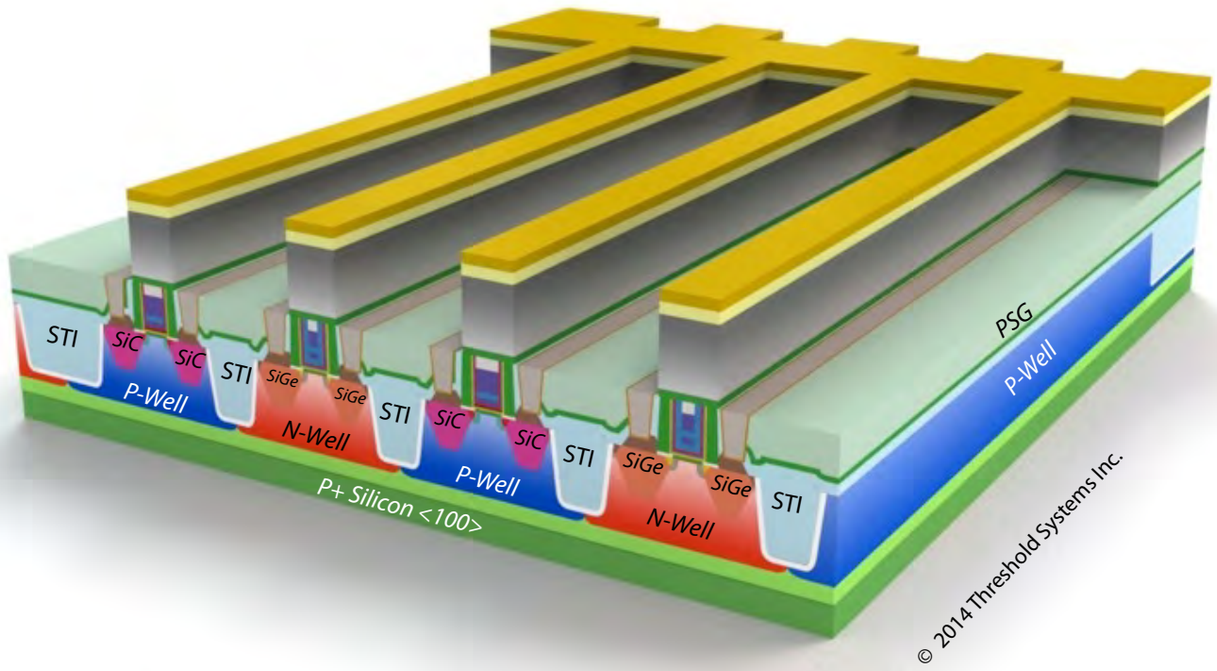


Using the TiN hard mask as a pattern, a highly anisotropic etch cuts through the TEOS and CDO layers to define the metal-1 trenches. The etch stops on the SiN etch-stop layer.

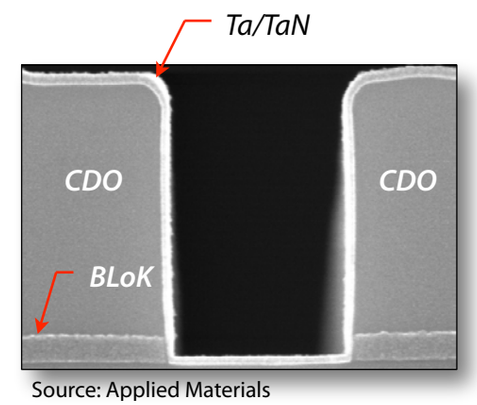
© 2018 by Threshold Systems Inc. All Rights Reserved.

Nitride Barrier Opening TaN/Ta Deposition:

ThresholdSystems

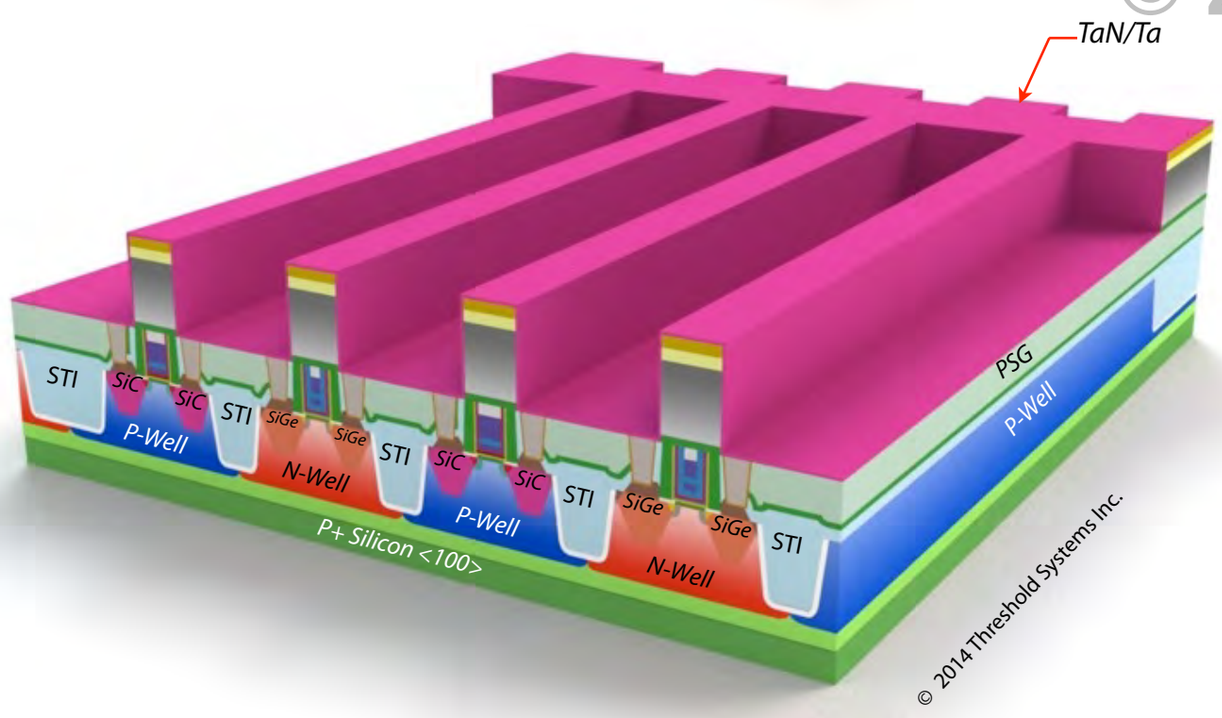


The wafers are cleaned and soft Argon sputter etch opens up the Silicon Nitride etch stop layer at the bottom of the trench.



Source: Applied Materials

© 2018 Threshold Systems Inc.



The wafers are degassed to boil off contaminants in the trenches and then a wet clean, followed by a soft Argon sputter, cleans out any remaining polymeric residue and carbon contaminants from the trenches.

Next, 80 Å of TaN followed by 30 Å of Ta are deposited using IMP PVD. These two metals act as a barrier that will confine the not-yet-deposited Copper to the trenches. A re-sputter step follows both the Ta and TaN operations to ensure good sidewall coverage.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Ta/TaN Barrier Characteristics:

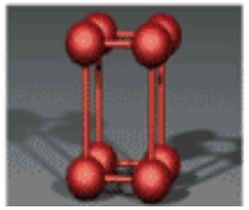
ThresholdSystems

- *The barrier material of choice is TaN to entomb the Copper in the Dual Damascene trenches and vias. Ta is the liner that facilitates the Copper fill of these structures.*
- *However, Ta cannot be directly deposited onto an oxide or it will configure as the high resistance beta-phase. Since the Ta is competing with Cu for volume in the trench, the Beta phase of Ta must be avoided.*
- *TaN adheres relatively well to oxide and is deposited first. TaN is also chemically stable and unreactive with Copper.*
- *Ta is deposited second because when it is deposited on top of TaN it manifests itself in the low resistivity alpha phase which is highly desirable*
- *Copper is a transistor poison and cannot be allowed to contaminate the active layer. Therefore it must be entombed on all sides by diffusion barriers (Ta/TaN on five sides & a SiN seal on the top).*

© 2018 Threshold Systems Inc.

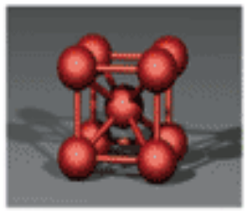
Tantalum has two phases, the high resistance Beta-phase and the low-resistance Alpha phase

β-Ta
ρ≈180 μΩcm



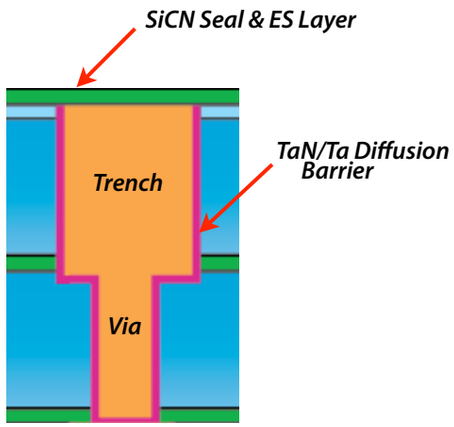
When placed directly onto oxide Ta assumes its high resistance Beta phase. Since it is desirable to have the Ta resistance as low as possible, this phase is highly undesirable.

α-Ta
ρ≈20-30 μΩcm



When placed deposited onto a TaN coated surface, Ta assumes its low resistance Alpha phase.

M. Travig et. Al. Semiconductor International

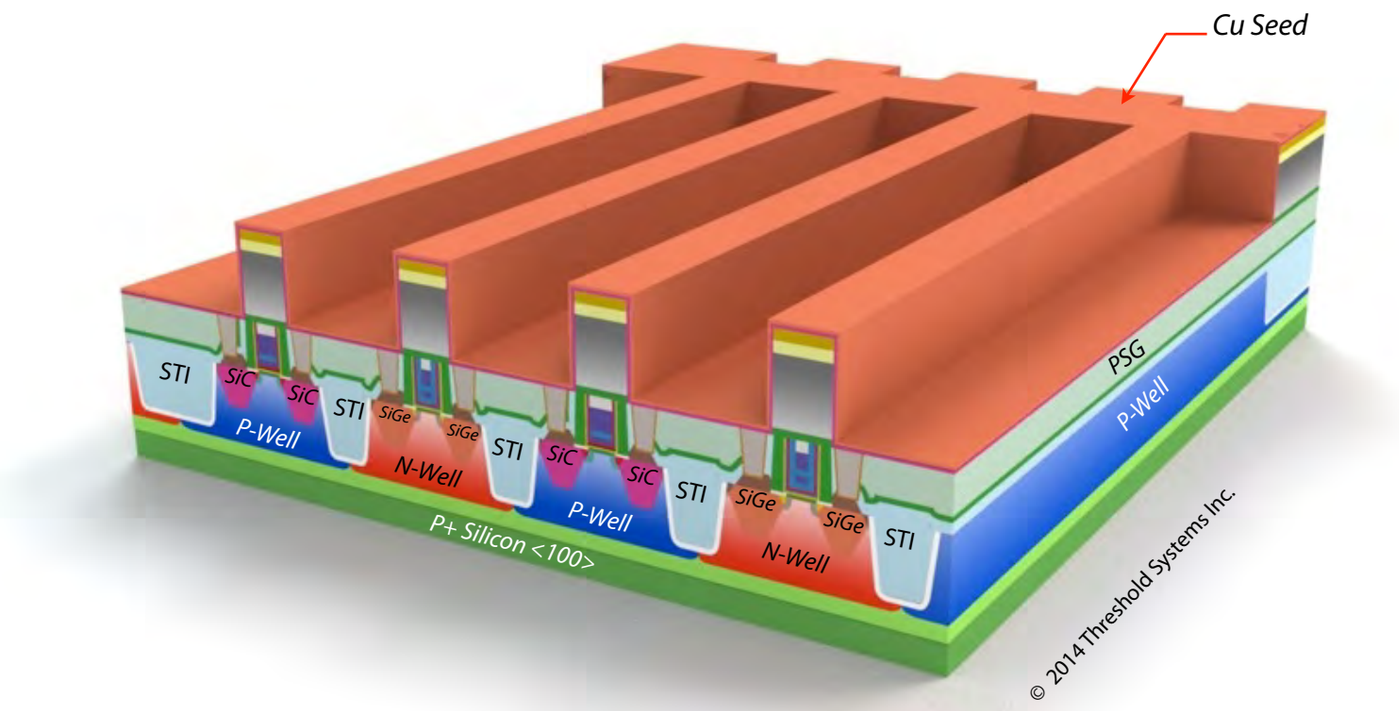


© 2007 Threshold Systems

© 2018 by Threshold Systems Inc. All Rights Reserved.

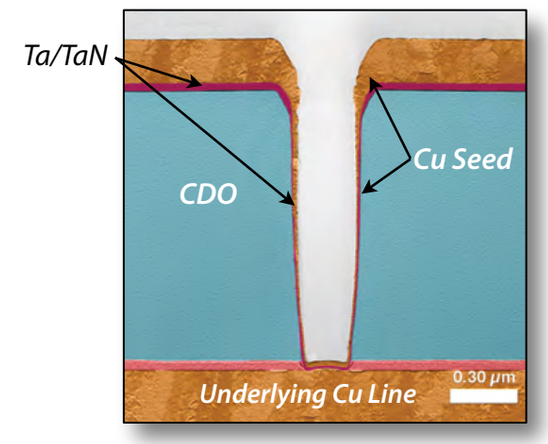
Copper Seed & Copper Bulk Deposition:

ThresholdSystems

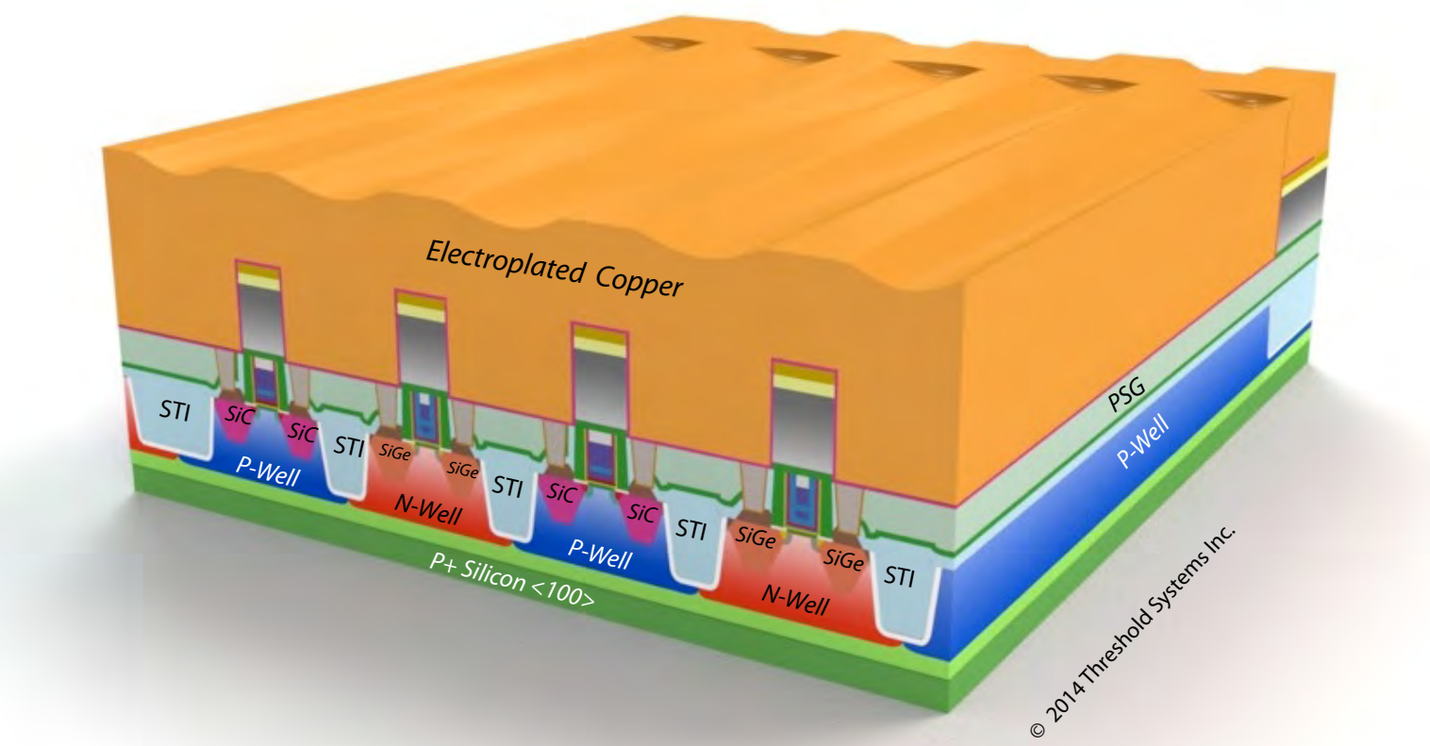


Without breaking vacuum, 300Å of Copper seed is deposited across the wafers which lines the interior of the trenches.

The Copper seed is required to ensure conformal void-free bulk Copper deposition and is a necessary precondition for the electroplating process.

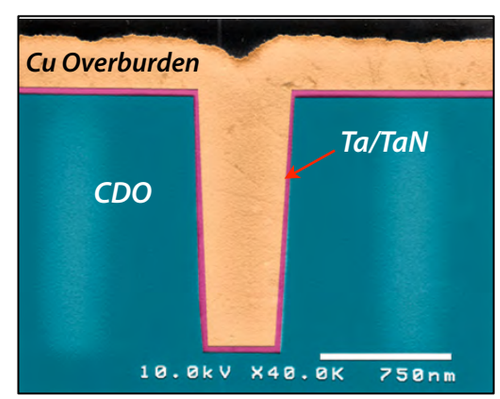


© 2018 Threshold Systems Inc.



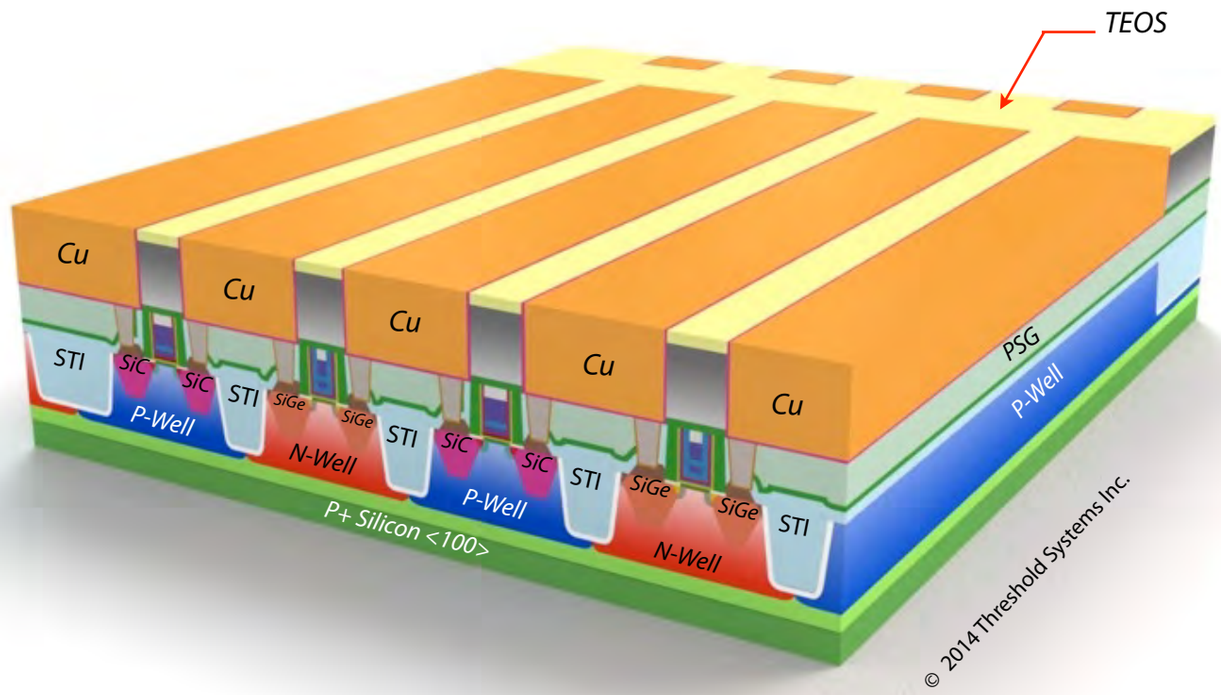
Next, 6,000 Å of bulk copper is then deposited using electrochemical deposition.

After deposition, the Copper is annealed at 300 °C for 90 seconds in forming gas to set its grain structure and reduce its resistance.



© 2018 by Threshold Systems Inc. All Rights Reserved.

Copper-Ta/TaN-TEOS CMP:

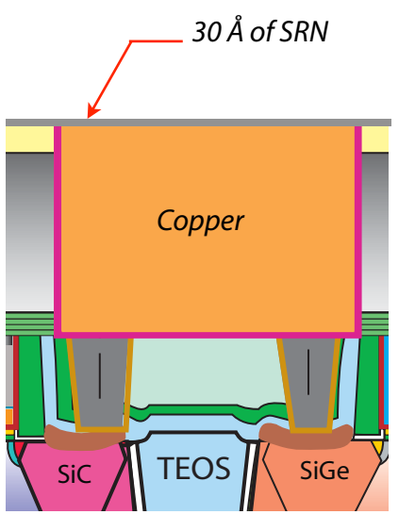


The Copper is polished back using CMP. This polishing also removes the Tantalum that was on the upper surfaces of the TEOS. The softer TEOS acts as a polish stop for the CMP process.

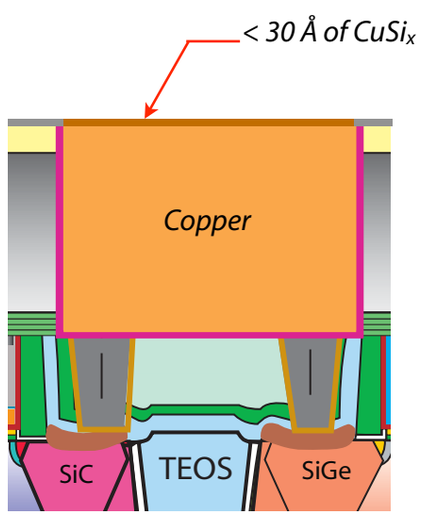
Next, an oxide buff operation is conducted to smooth out the surface of the TEOS oxide and ensure that the surface of the TEOS is coplanar with the top of the Copper lines.

After each CMP operation the wafers are cleaned.

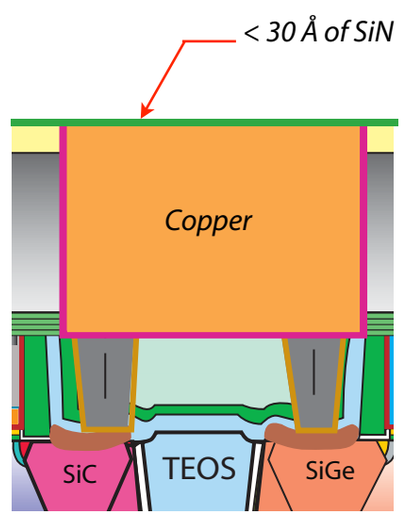
© 2018 Threshold Systems Inc.



© 2011 Threshold Systems



© 2011 Threshold Systems



© 2011 Threshold Systems

The Cu surface is cleaned of organic impurities left behind from the CMP clean. This is accomplished by employing a multiple step plasma assisted reducing chemical that eliminates Cu oxide and other organic impurities.

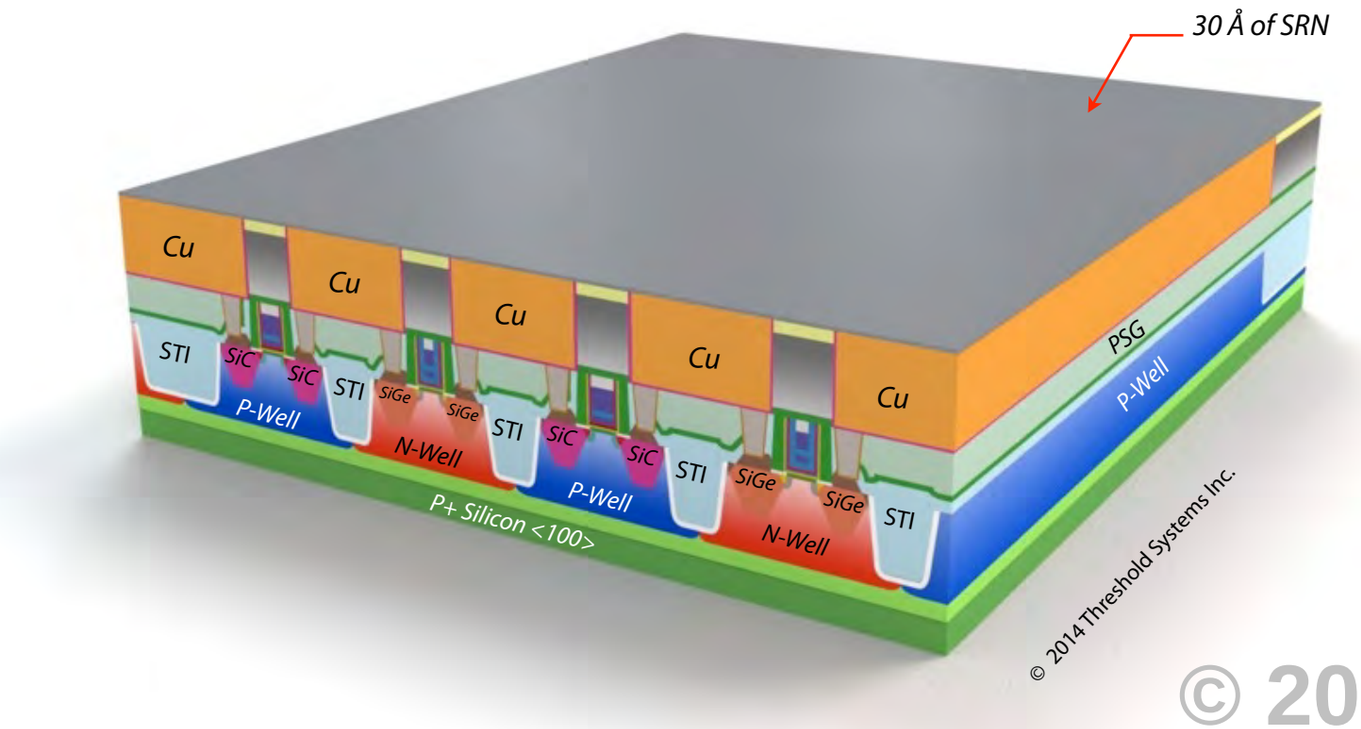
Next, a thin 30Å layer of Silicon Rich Nitride (SRN) is deposited that reacts with the Cu to form a Cu silicide ($CuSi_x$). This is a self-limiting reaction and the $CuSi_x$ limits the Cu diffusion by locking-up the Cu on the upper surface of the trench in the SRN matrix. The $CuSi_x$ is electrically leaky so it is next subjected to a nitridization step to convert it to SiN film.

The SRN treatment is highly effective in sealing in the Cu and enables a thinner SiCN sealant film to be deposited. This results in a lower $K_{effective}$ for the ILD stack.

© 2018 by Threshold Systems Inc. All Rights Reserved.

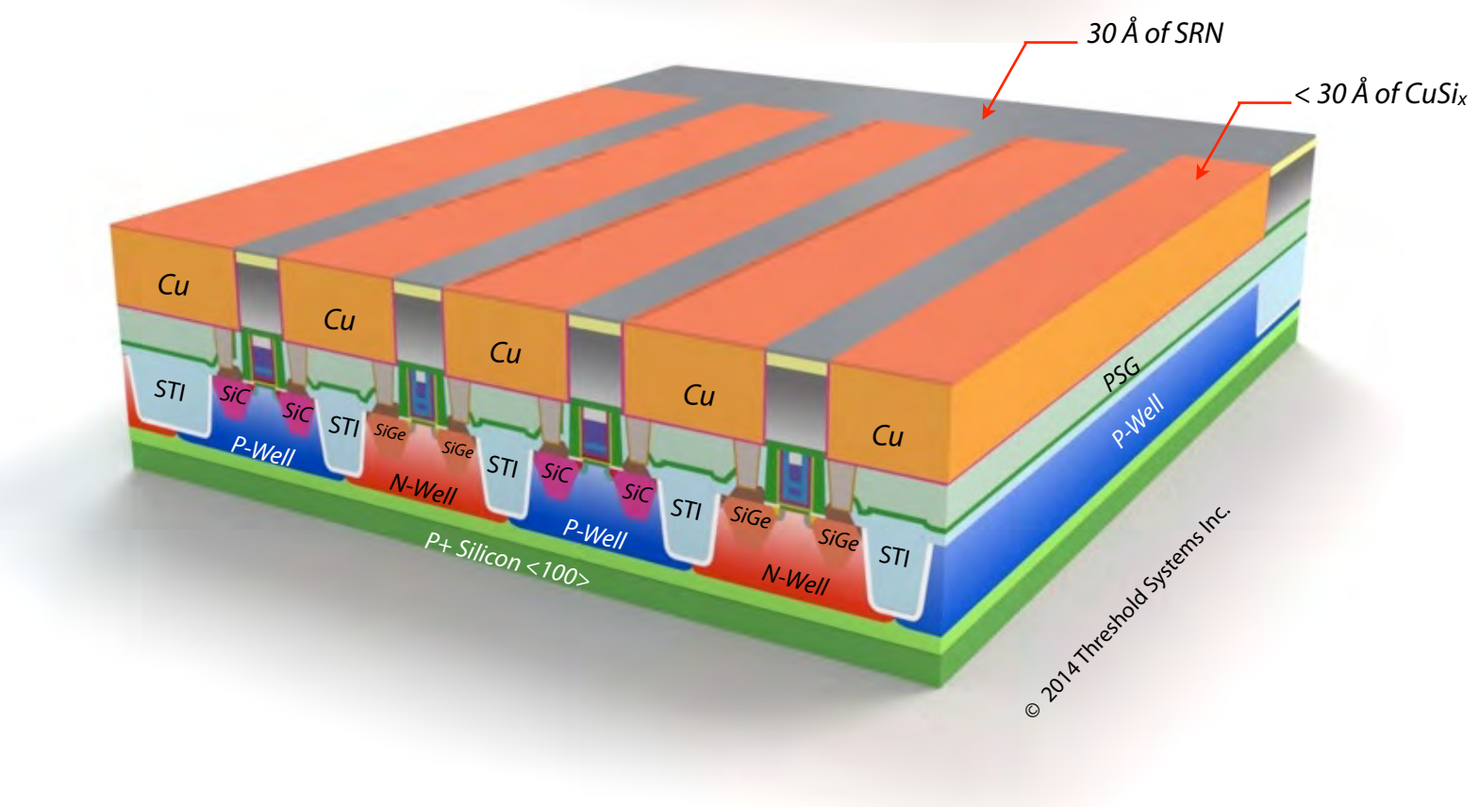
Copper Sealing:

ThresholdSystems



A thin 30Å layer of Silicon Rich Nitride (SRN) is deposited across the wafer.

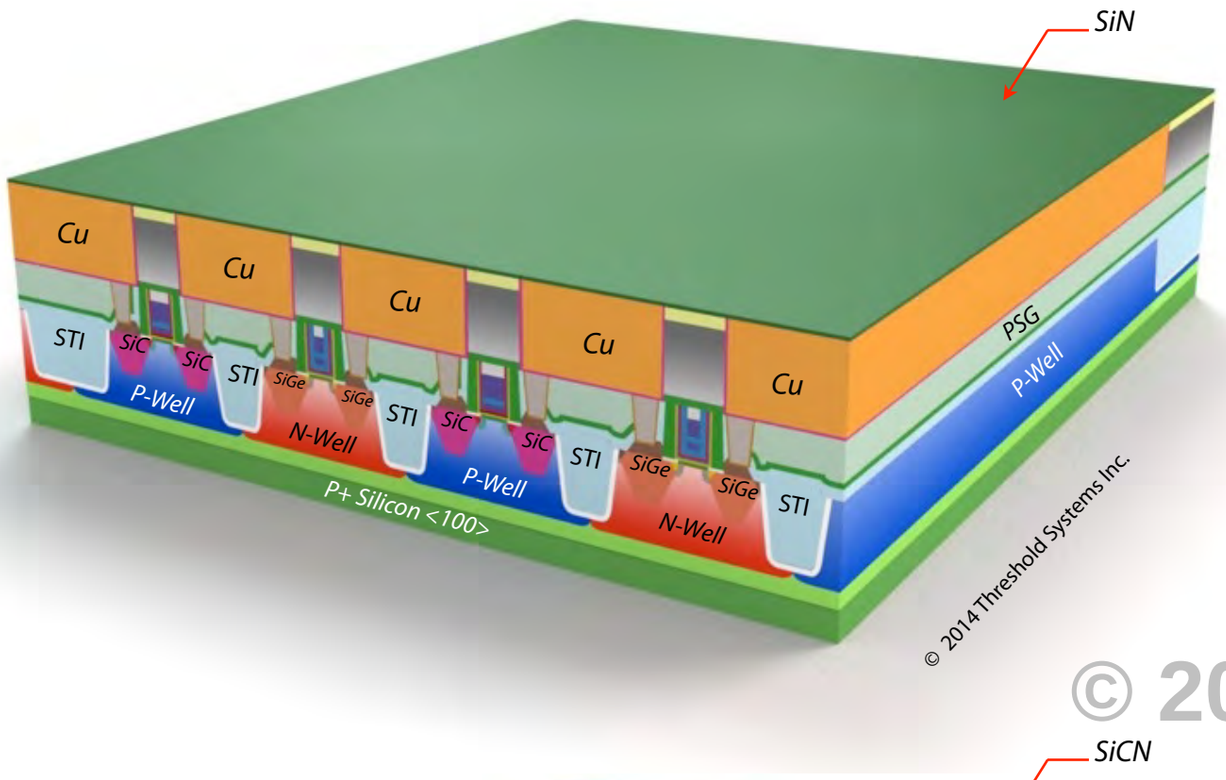
© 2018 Threshold Systems Inc.



A thin 30Å layer of Silicon Rich Nitride (SRN) reacts with the Copper metal to form a Copper silicide ($CuSi_x$).

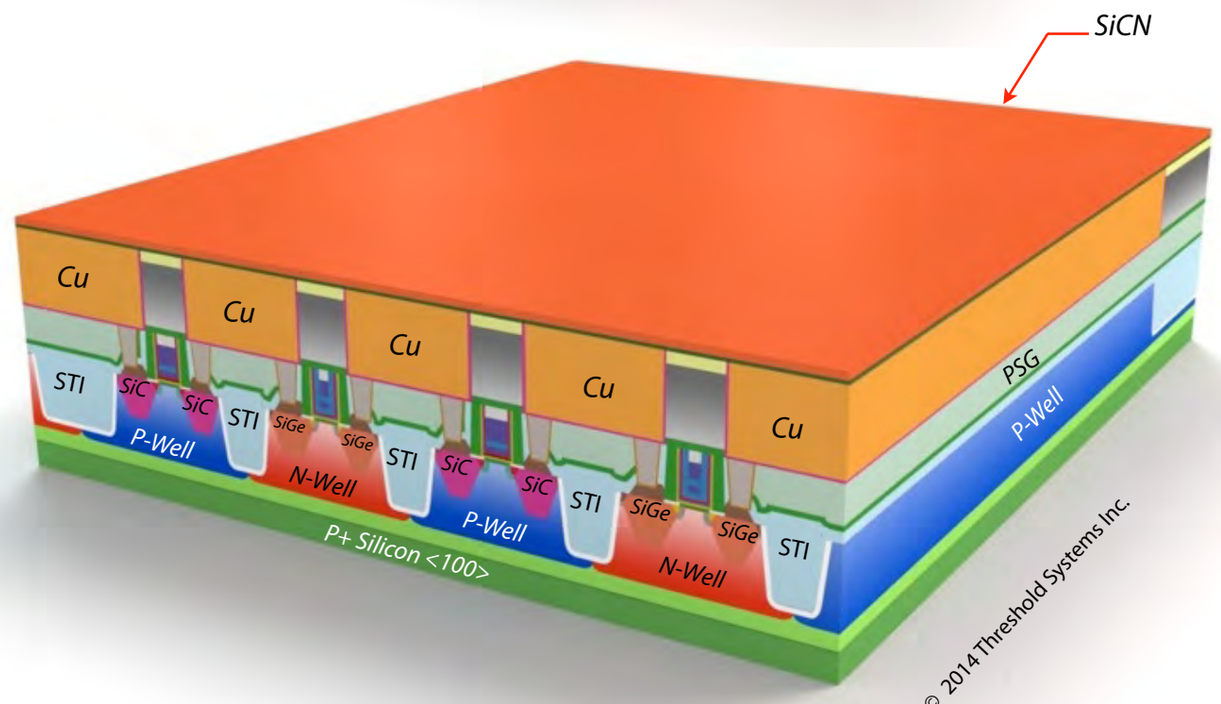
© 2018 by Threshold Systems Inc. All Rights Reserved.

Copper Sealing & Barrier Deposition:



The $CuSi_x$ is then subjected to a nitridization step to convert it to SiN film.

© 2018 Threshold Systems Inc.



A 300Å etch-stop layer of SiCN is deposited to seal off the top of the first layer of metal. The Copper in the first layer of metal is now entombed with Ta/TaN on five sides and SiCN on the top.

© 2018 by Threshold Systems Inc. All Rights Reserved.

The Role of a Local Interconnect:

ThresholdSystems

© 2018 Threshold Systems Inc.

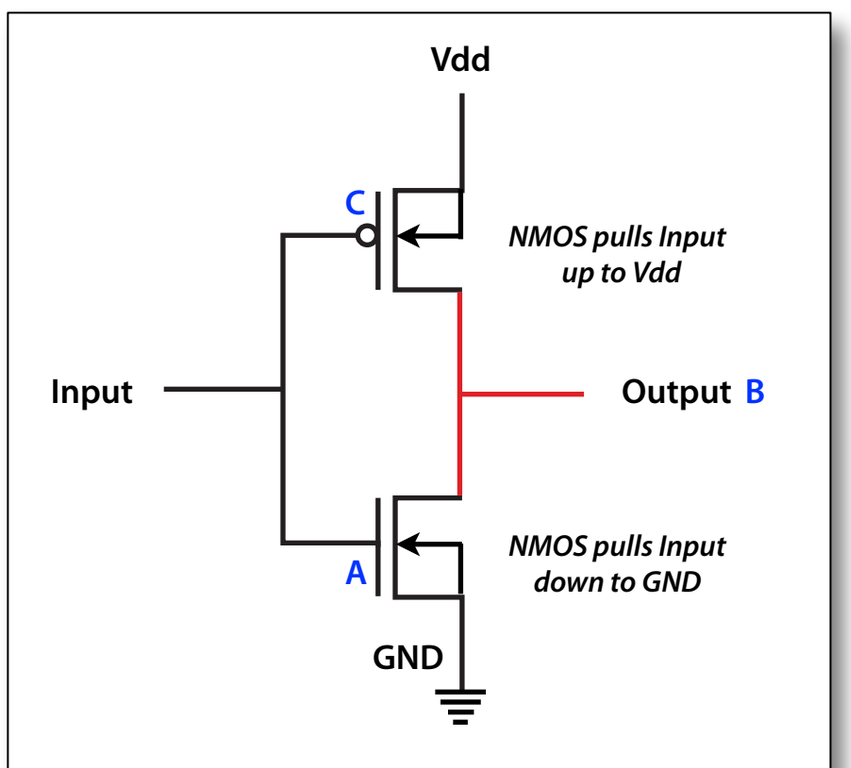
Local Interconnects

Local Interconnects enhance packing density by allowing electrical inter-connections to be made at the device level without having to route signals to several higher levels of metal. They simplify signal routing and are highly desirable for devices with extremely high transistor counts.

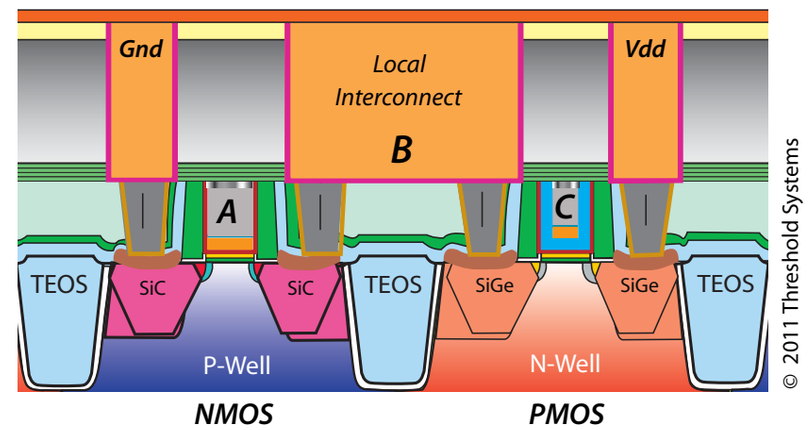
To understand how Local Interconnects are used, consider the example of a simple inverter, the schematic of which is presented below. In this example the Drain of a NMOS device (denoted as "A") and the Source of a PMOS device (denoted as "C") are joined together and pulled out to a common output denoted as "B".

The top-down and cross-sectional view presented to the right illustrate how an inverter might be realized in Silicon. We can see that the Drain of the NMOS device (A) is physically and electrically connected to the Source of the PMOS device (C) by the Cu Local Interconnect (B).

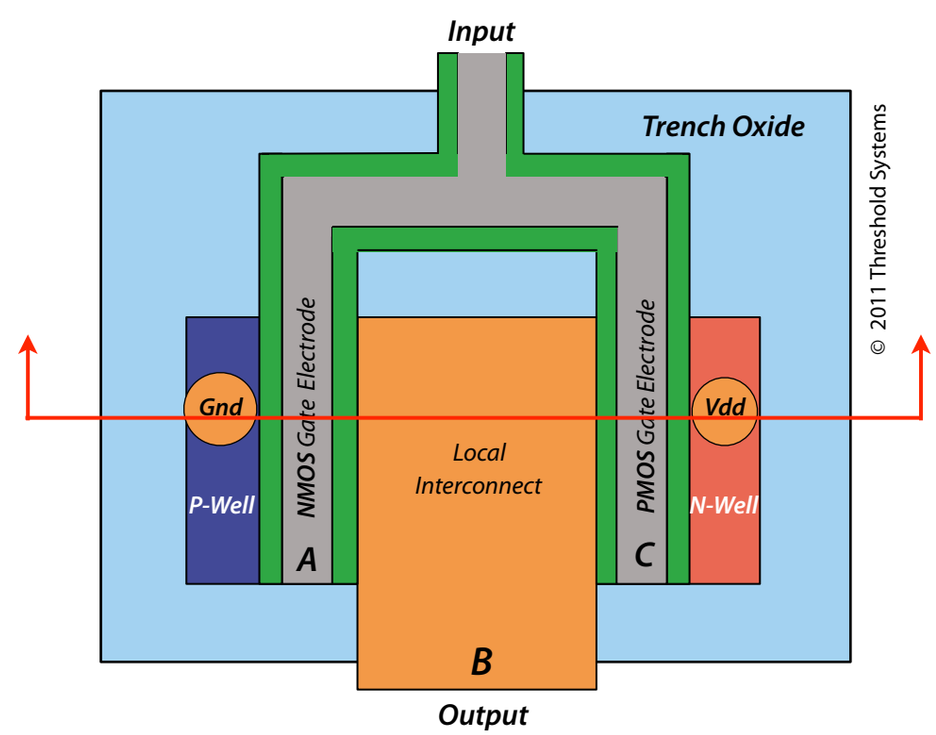
Schematic of an Inverter



Cross-Sectional View

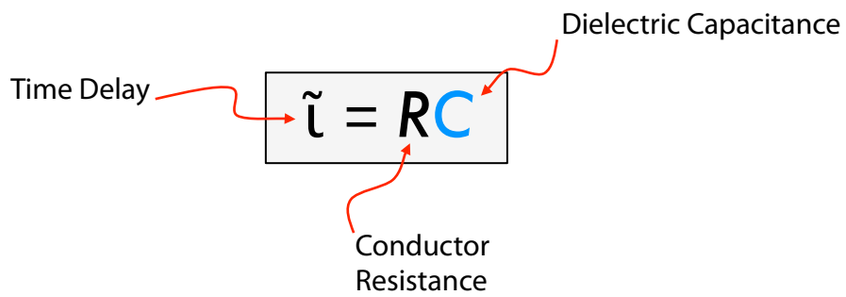


Top-Down View



© 2018 by Threshold Systems Inc. All Rights Reserved.

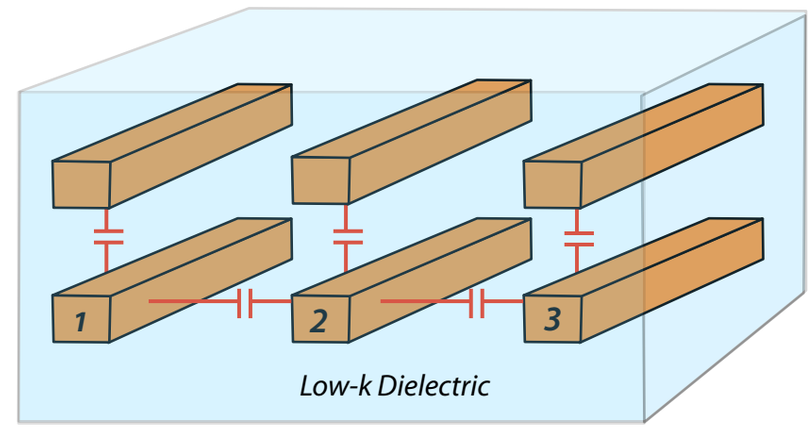
Why Do We Need Low-k Dielectrics?



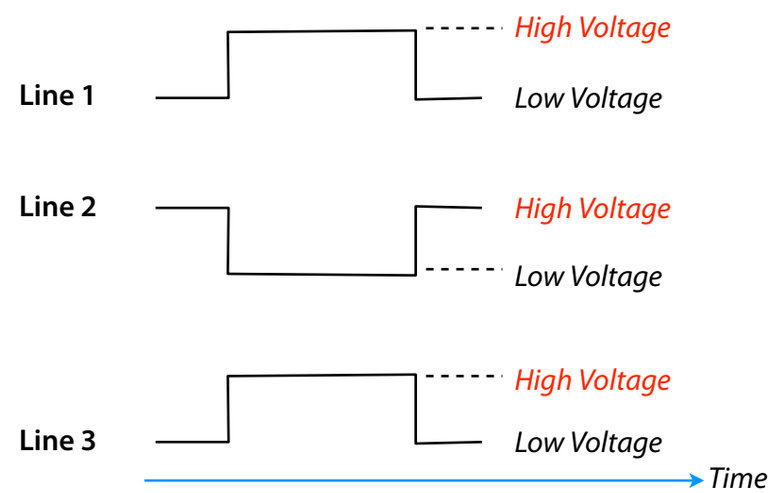
© 2018 Threshold Systems Inc.

- *Interconnect delay can be nearly 20 times greater than gate delay at the 22nm node.*
- *Copper alone is insufficient to minimize interconnect delay. Low-k dielectrics must be used between the Cu lines to minimize capacitive coupling between adjacent metal lines, and to a lesser extent, between stacked metal lines.*
- *The use of Cu metal lines and Low-k dielectrics dramatically reduces the RC time constant enhancing chip speed.*
- *Low-k materials will not easily support an electric field, and therefore inhibit adjacent metal lines from capacitively coupling. This minimizes parasitic capacitances and lowers the probability of cross-talk between adjacent metal lines, a mechanism that can result in data corruption.*

Cross-Talk



© 2011 Threshold Systems Inc.



Adjacent metal lines embedded in a dielectric will capacitively couple with each other. At very small geometries, and especially at the 22nm node, this capacitive coupling will lead to cross-talk between the lines. This results in signal distortion. A worst case scenario occurs when one metal line is transitioning into a low voltage state, when lines on either side of it are transitioning to a high voltage state. The line that is attempting to enter the low-voltage state will experience strong cross talk as a result of capacitive coupling, and will distort the digital signal that it is carrying. A low-k dielectric is extremely helpful in avoiding this kind of problem because it will not readily support the electric field that facilitates cross-talk.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Kinds of Low-k Dielectrics:

ThresholdSystems

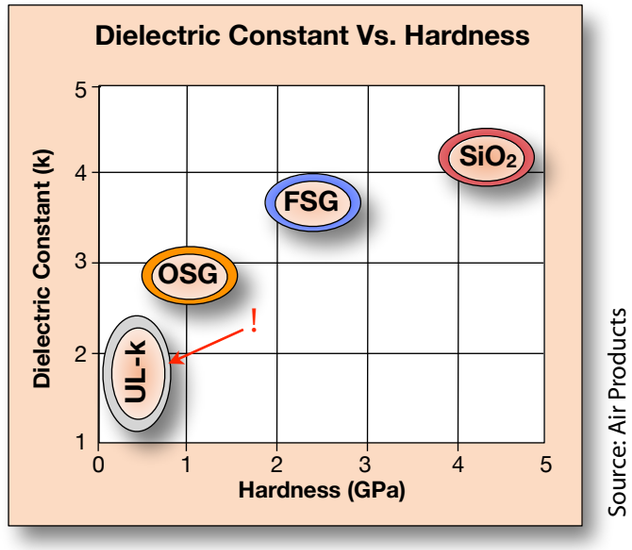
- *The baseline generation of backend dielectrics was SiO_2 which served the industry well for several generations. It was hard, easy to etch, had good coefficients of thermal expansion and conduction, a high breakdown voltage and was an excellent insulator. However, with a K-value of 3.9 - 4.2 it was an unacceptable dielectric for advanced technology nodes.*

- *The first generation of low-k dielectric to achieve widespread adoption was Fluorinated Silicon Glass (FSG) which was introduced at the 130 nm node and used extensively at the 90 nm node. It is simply SiO_2 that has been doped with Fluorine to make it less polar and reduce its k-value to ~ 3.5*

- *The second generation of low-k dielectrics was carbon doped oxide (CDO), which are also known as Organo-Silicate Glass (OSG), or more specifically as SiCOH (marketed as "Black Diamond"). This material has a k-value of ~ 3.0 and saw widespread use at the 90 nm and 65 nm nodes. The k-value can be lowered to 2.7 by increasing the Carbon content of the film, and a porous version of this dielectric is used at the 45/32/28nm nodes having a k-value of ~2.5*

- *A third generation of low-k dielectrics consisting of porous polymers with k-values <2.0 has proven to be elusive.*

Low-K Dielectric Generations

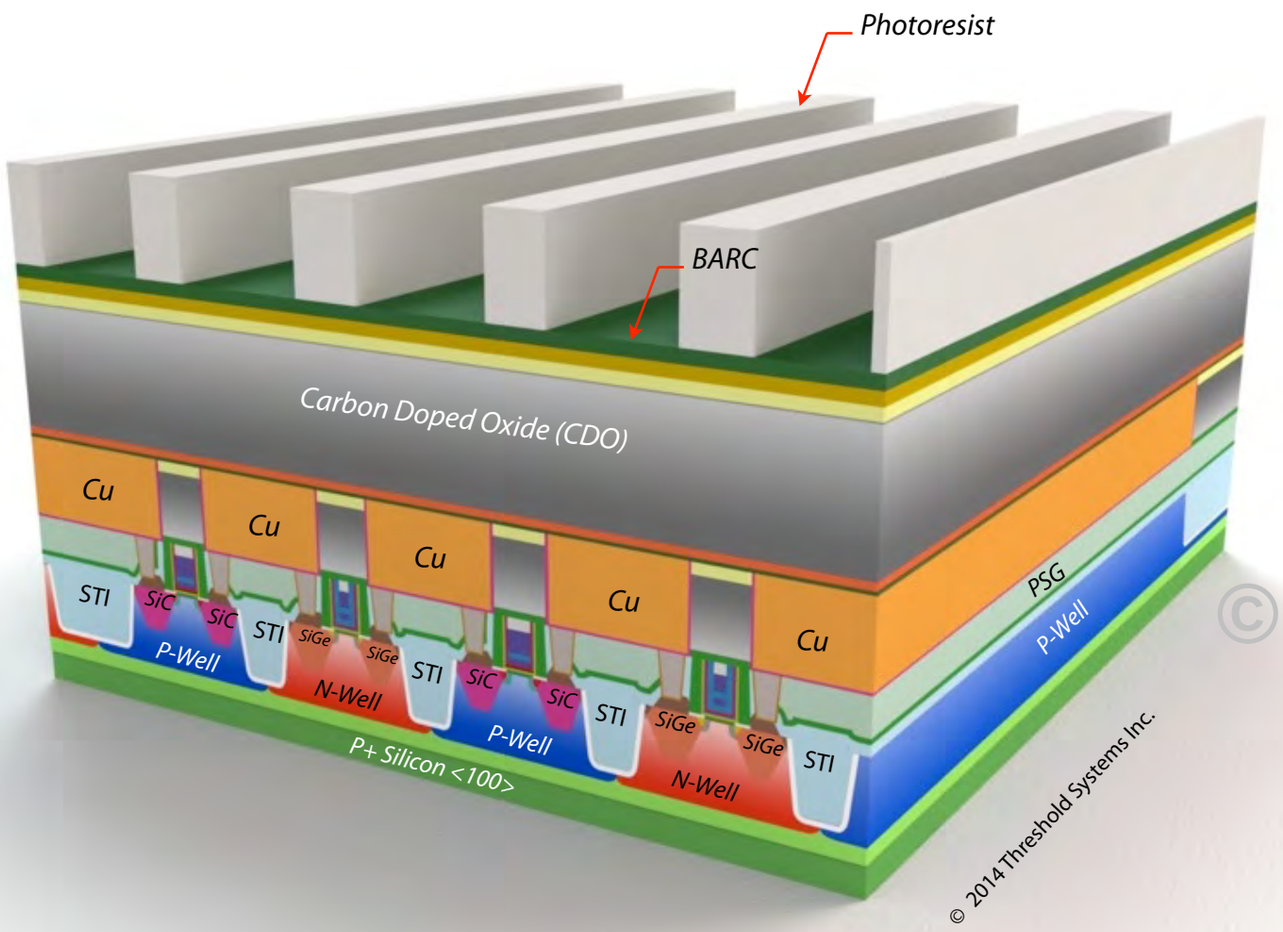


As the k-value of a dielectric film declines so does its hardness. This reduction in mechanical strength introduces a range of serious processing issues that has become one of the biggest challenges for sub-32nm node technology.

© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

ILD-2 Deposition:



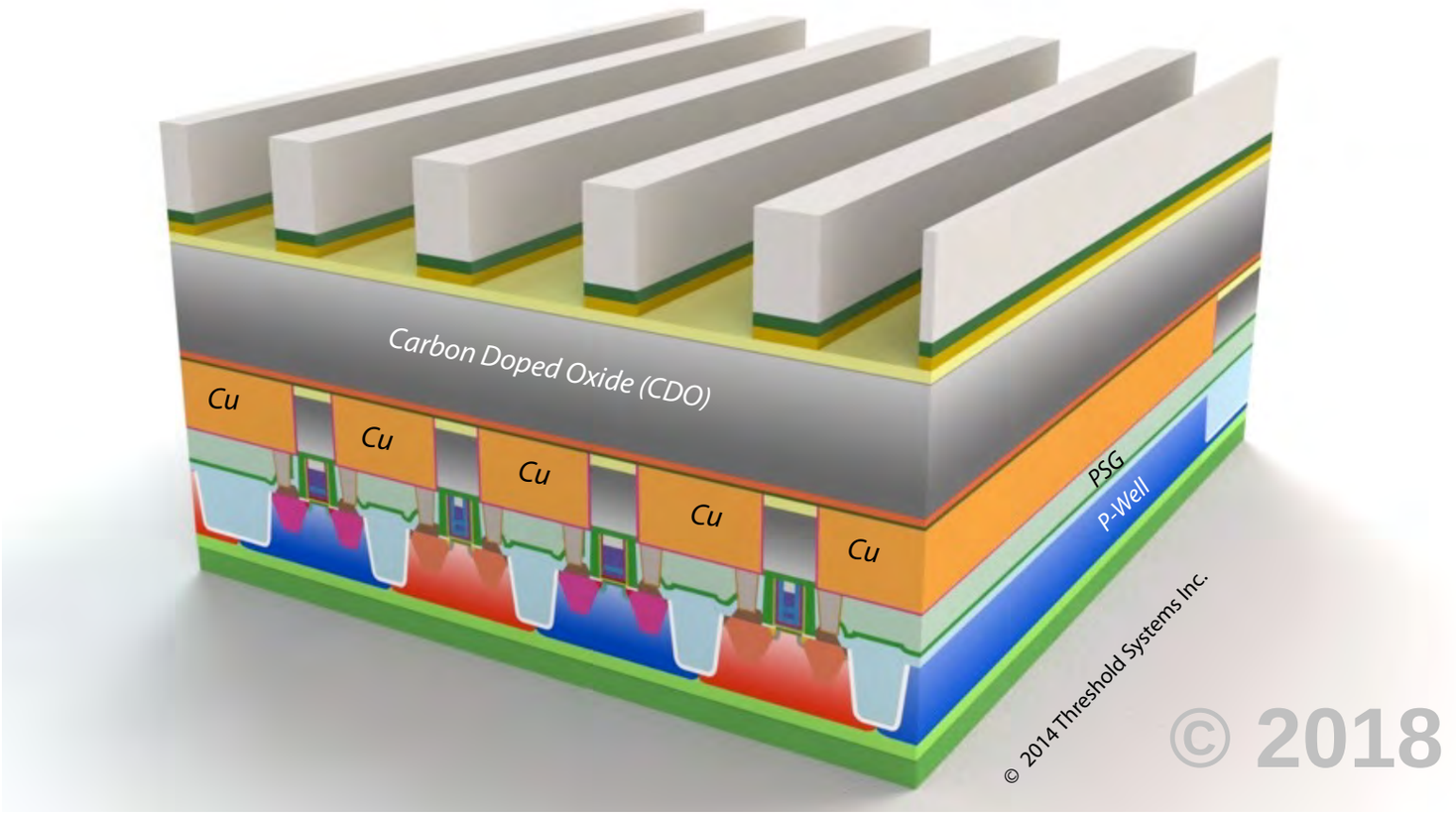
Next, 7,000 Å of a Low-k SiCOH is deposited and then UV baked to remove the porogens and lower the k-value of the dielectric. This is followed by 400 Å of TEOS (a polish stop), followed by a 300Å layer of TiN hard mask and 300Å of BARC and a layer of photoresist.

The photoresist is then patterned.

© 2018 Threshold Systems Inc.

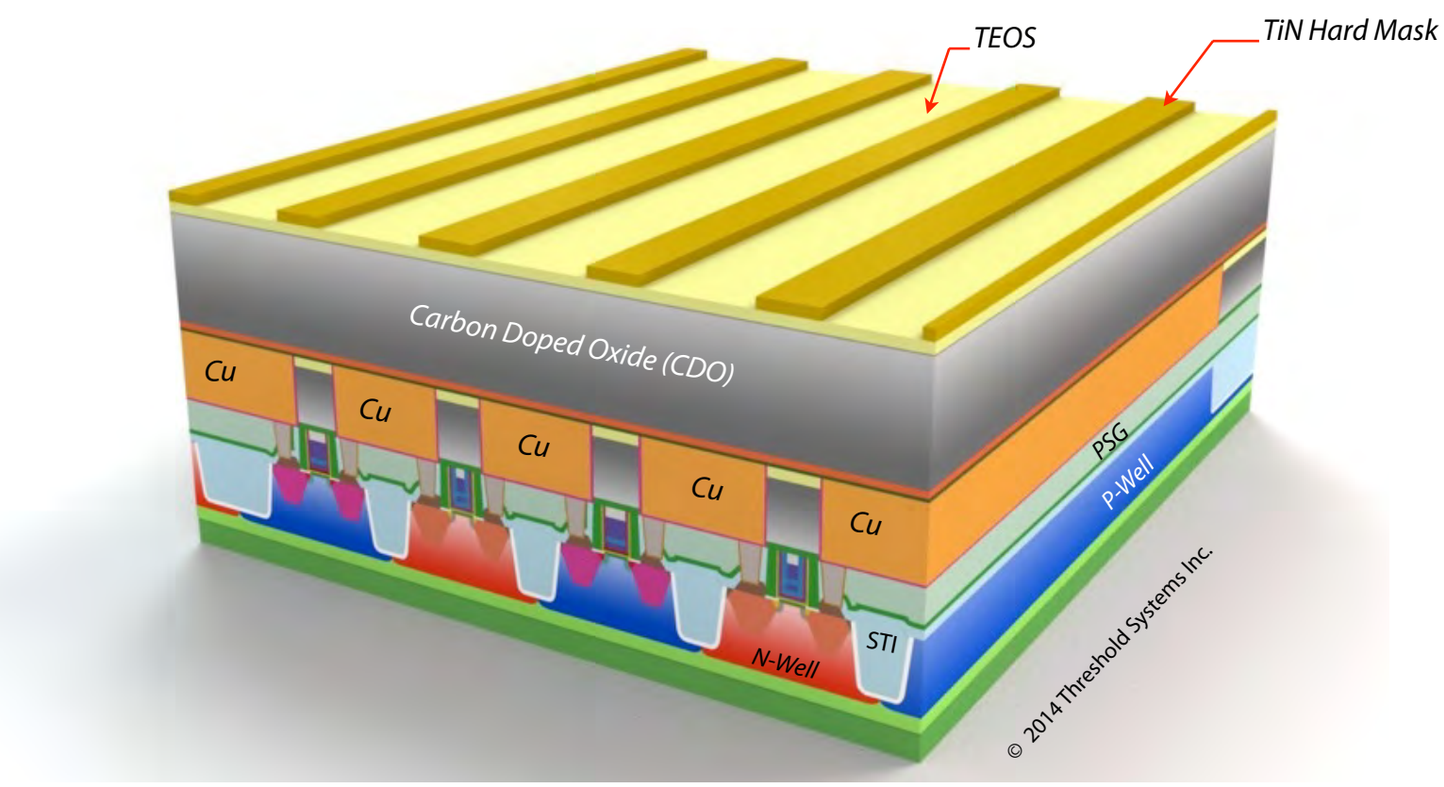
© 2018 by Threshold Systems Inc. All Rights Reserved.

TiN Hard Mask Etch & PR Strip:



The BARC is etched followed by an anisotropic etch of the TiN hard mask.

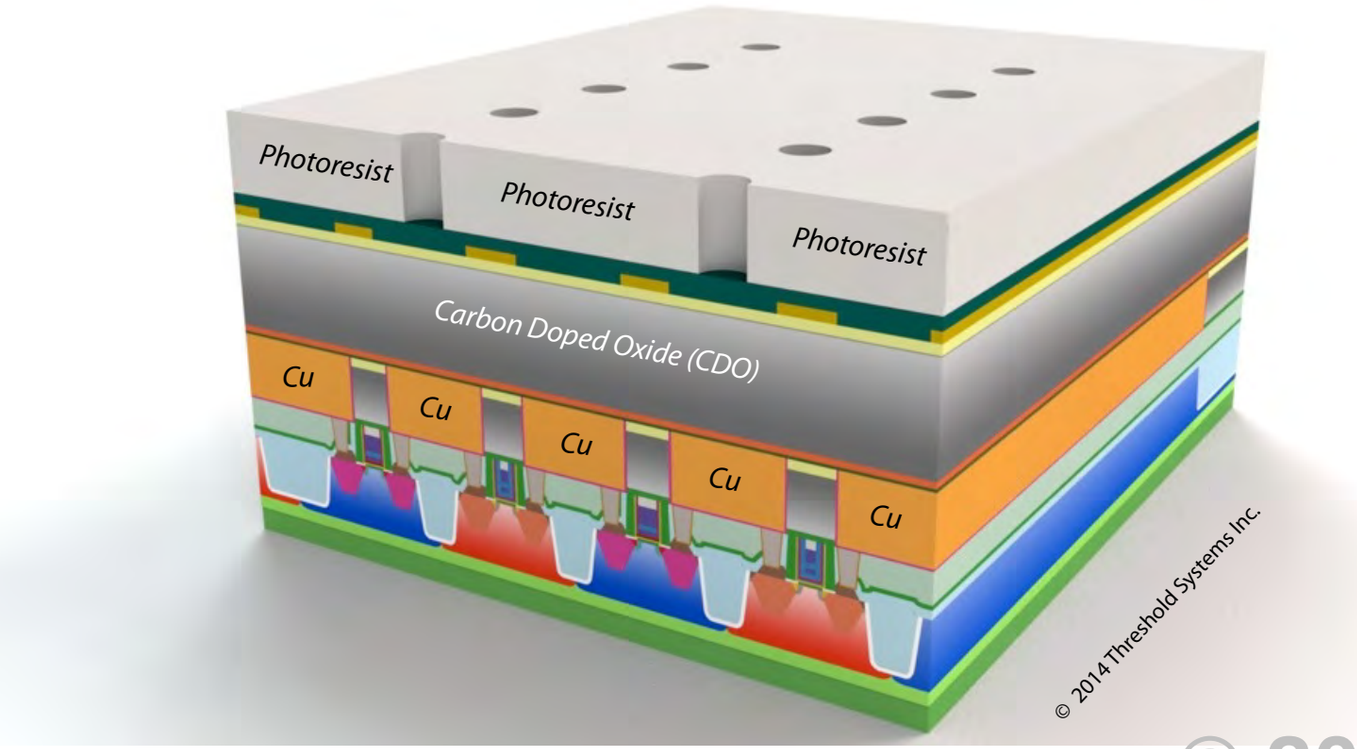
© 2018 Threshold Systems Inc.



The BARC is etched followed by an anisotropic etch of the TiN hard mask. The photoresist is then stripped and the wafers cleaned. The TiN hard mask remains behind.

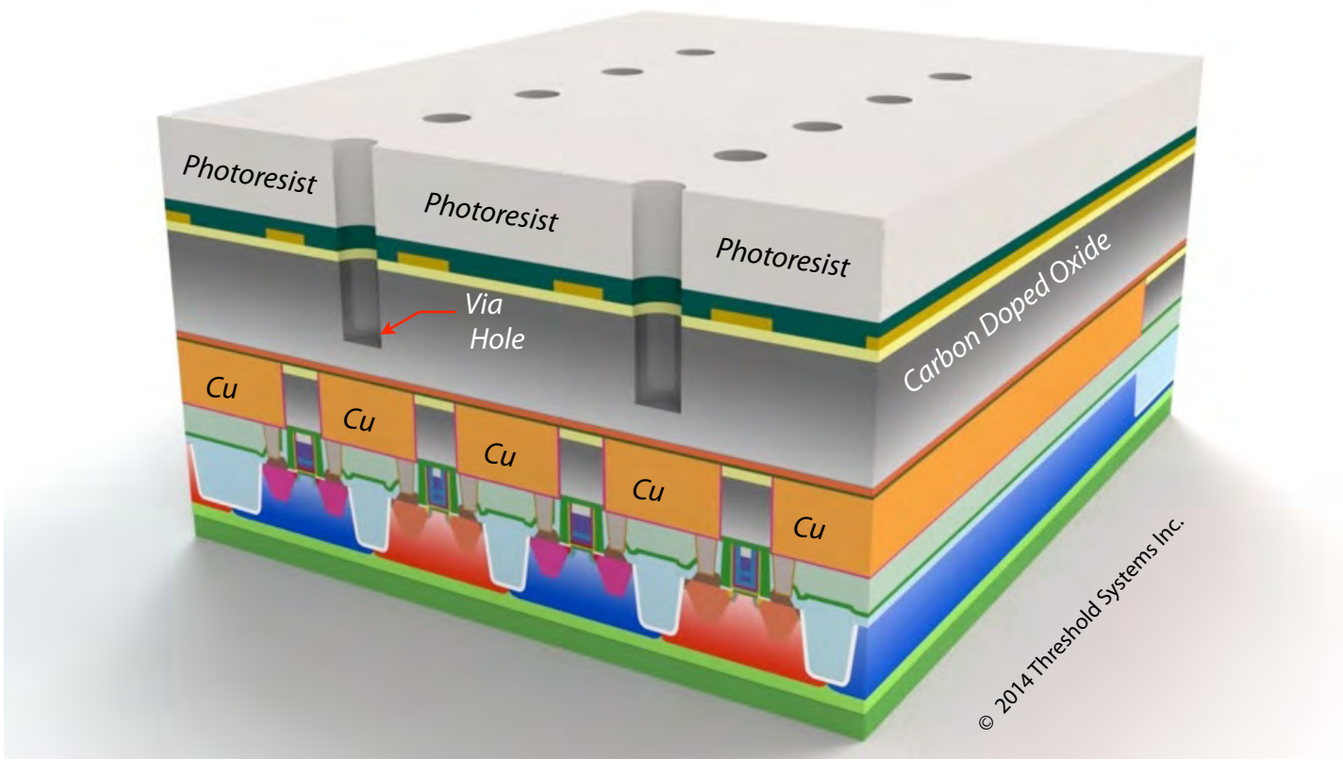
© 2018 by Threshold Systems Inc. All Rights Reserved.

Via Photo & Via Partial Etch:



The wafers are then coated in BARC and Photoresist. The photoresist is then patterned.

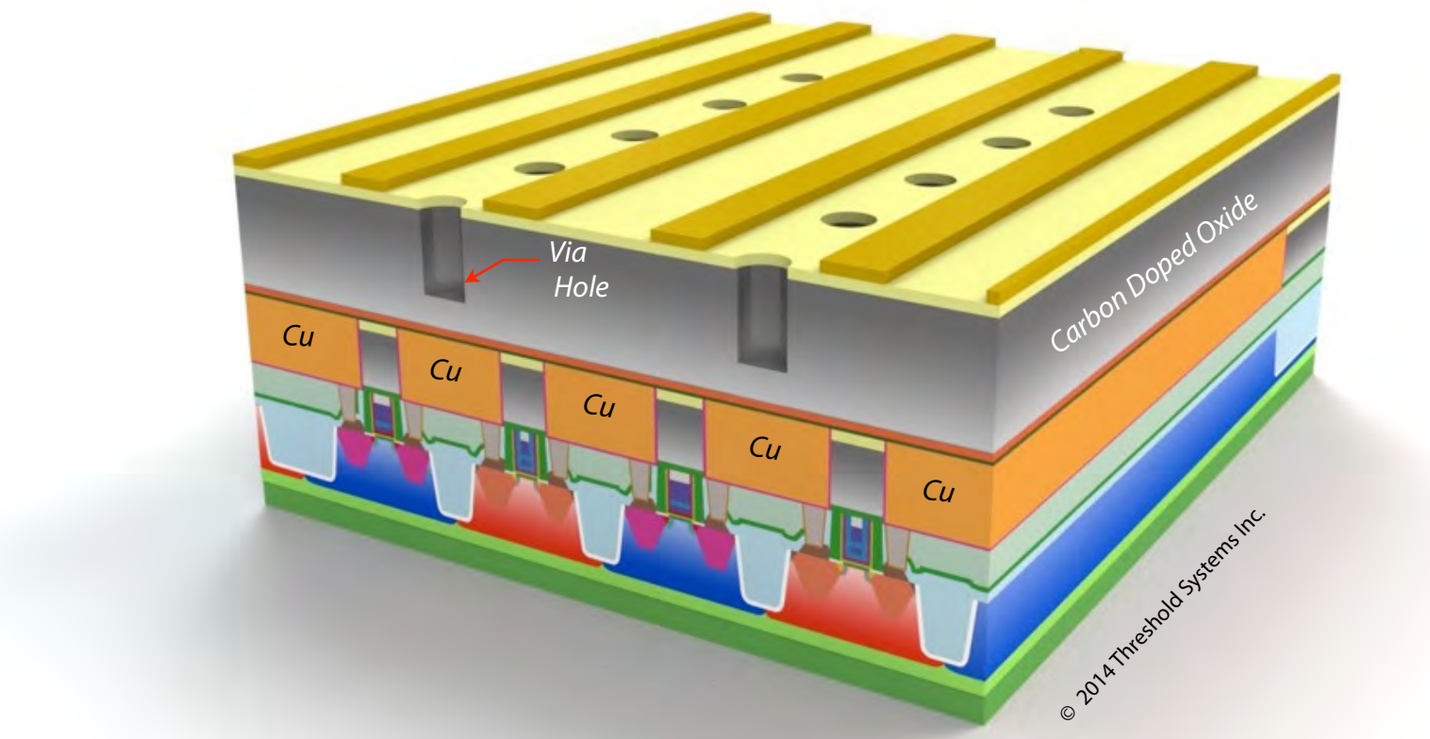
© 2018 Threshold Systems Inc.



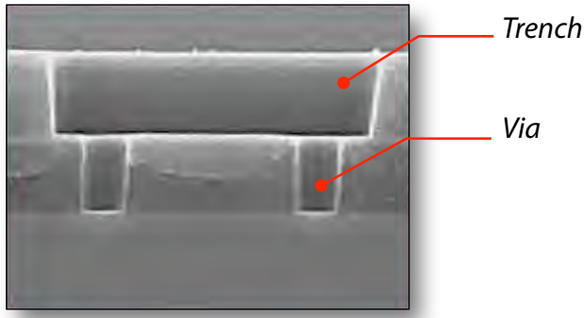
The vias are partially etched in a highly anisotropic etch recipe. They are not etched all the way down to the SiCN etch-stop layer.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Photoresist Strip & Trench Etch:

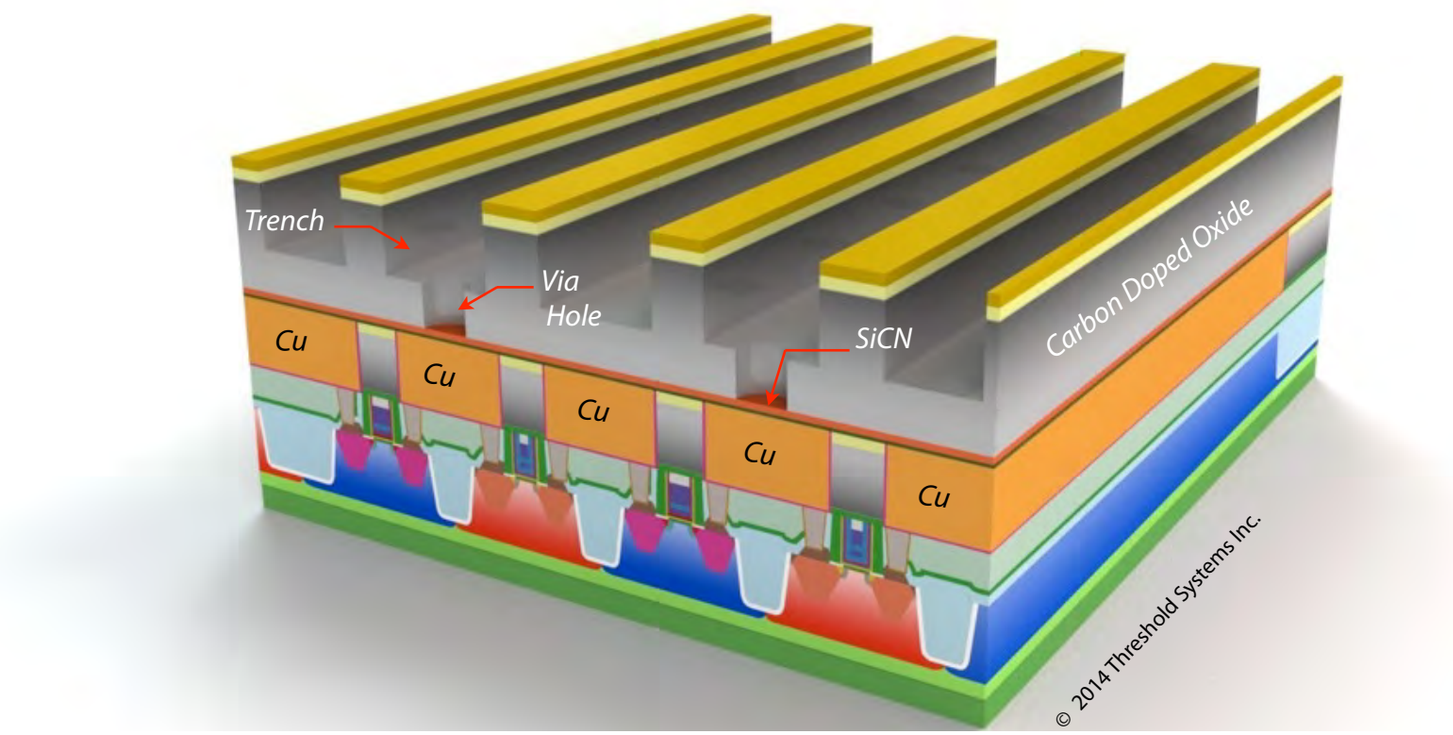


The BARC and the photoresist are stripped and the wafers are cleaned.

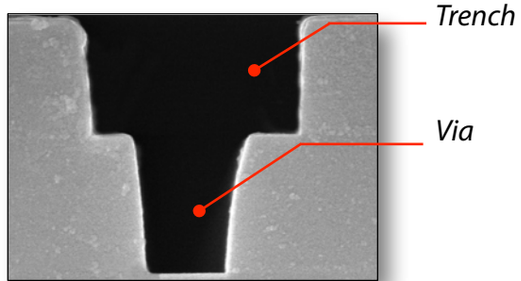


Source: Lam Research

© 2018 Threshold Systems Inc.



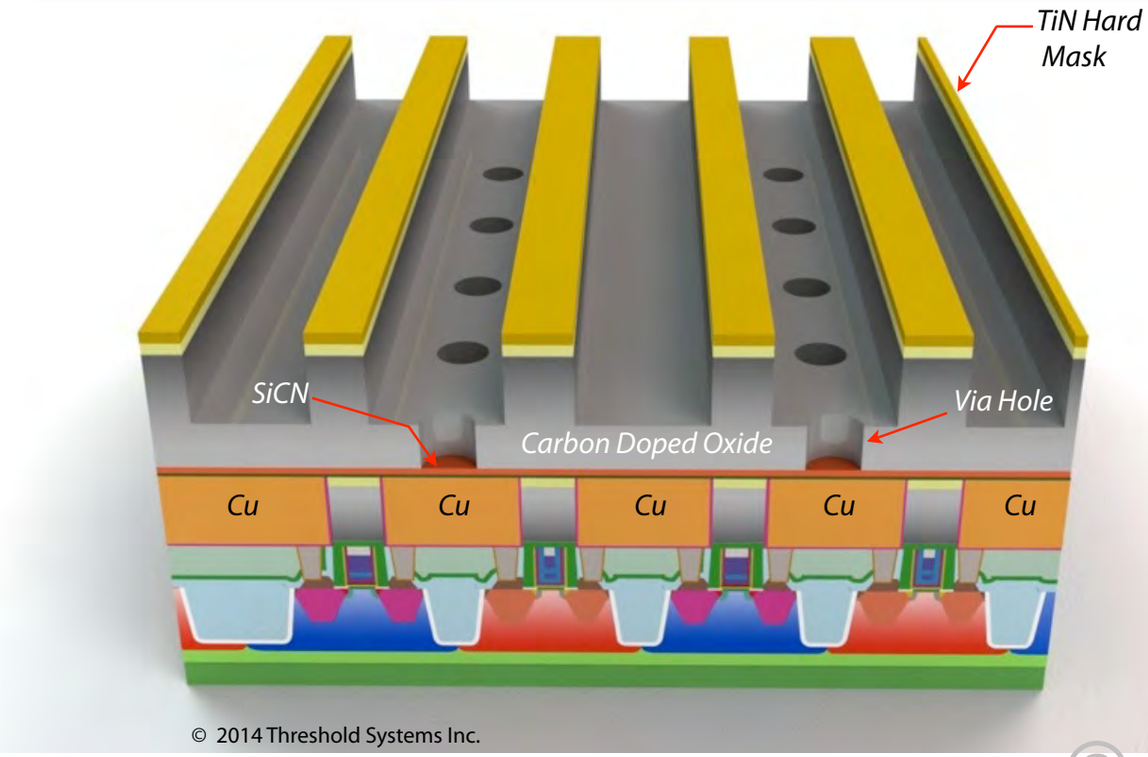
Using the TiN hard mask pattern as a template, the Metal-2 trenches are etched. During this operation the vias are fully etched. The via etch stops on the SiCN etch stop layer and may etch slightly into this layer.



© 2010 Threshold Systems

© 2018 by Threshold Systems Inc. All Rights Reserved.

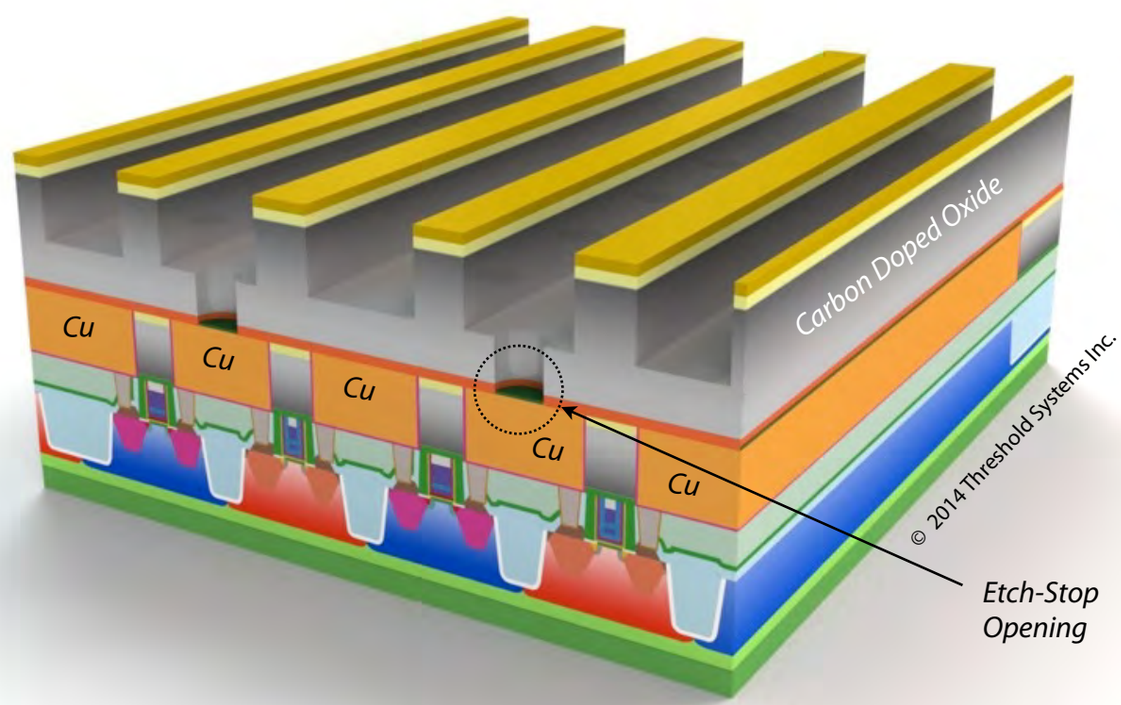
SiCN Barrier Opening:



A head-on view of the Dual Damascene structure displaying the trench and numerous via holes.

Note that the SiCN barrier at the bottom of the vias has not yet been punctured.

© 2018 Threshold Systems Inc.



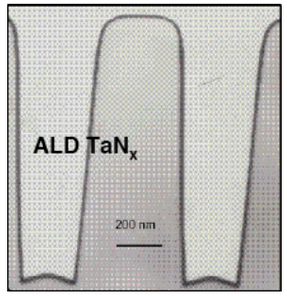
A soft Argon sputter etch is used to open the SiCN BLOk layer at the bottom of the vias and expose the underlying Copper line. This is a very delicate step. To aggressive a sputter etch and Copper will be sprayed up into the unshielded Dual Damascene structure. To mild an etch and the SiCN barrier will not be completely removed resulting in high via resistance.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Barrier Deposition Methodology:

ThresholdSystems

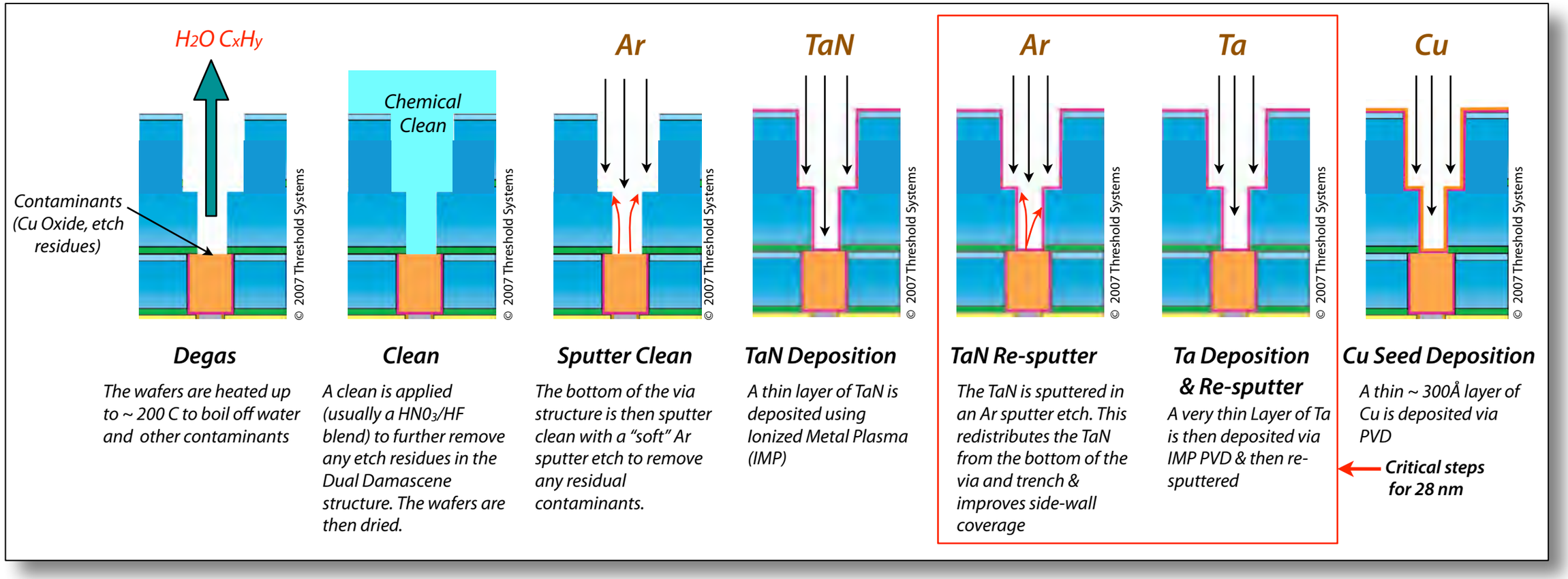
- *The re-sputtering of TaN is essential for 32/28 nm technology because it ensures more conformal coverage of the inside of the Dual Damascene structure.*
- *PVD will continue to be the predominant technology for thin barrier deposition at the 22nm and 14nm nodes and probably for the foreseeable future.*
- *Although ALD technology offers vastly superior conformal coverage and can can be scaled to extremely small dimensions, it's deposition rates are very slow and the carbon precursor used with ALD degrades adhesion. In addition, film continuity and "steric hinderance" are problematic with ALD. It appears that PVD technology will be used for barrier deposition for the foreseeable future.*



O. Van der Straten et al., MRS Proc., vol 716, 2002, pg 471

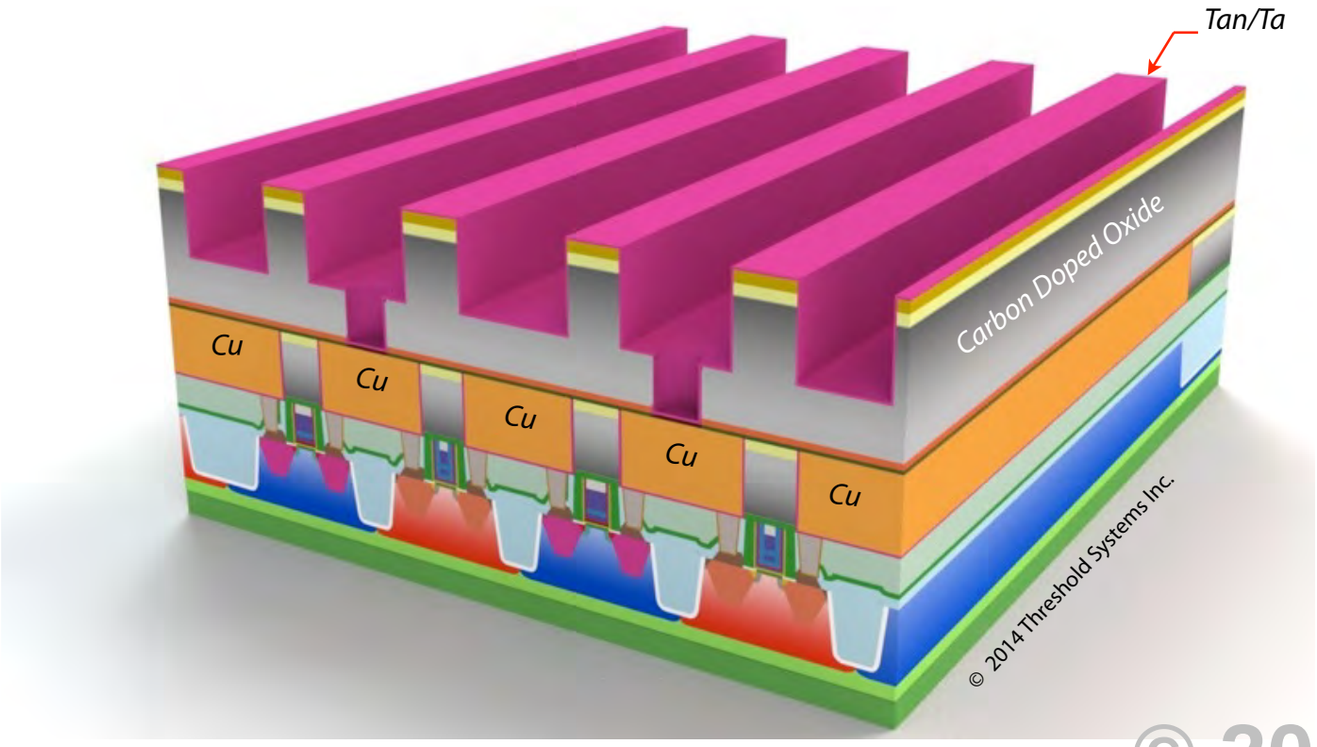
ALD Barriers are extremely thin and conformal, but this technology will probably not be used until the 10 nm node or below.

© 2018 Threshold Systems Inc.



© 2018 by Threshold Systems Inc. All Rights Reserved.

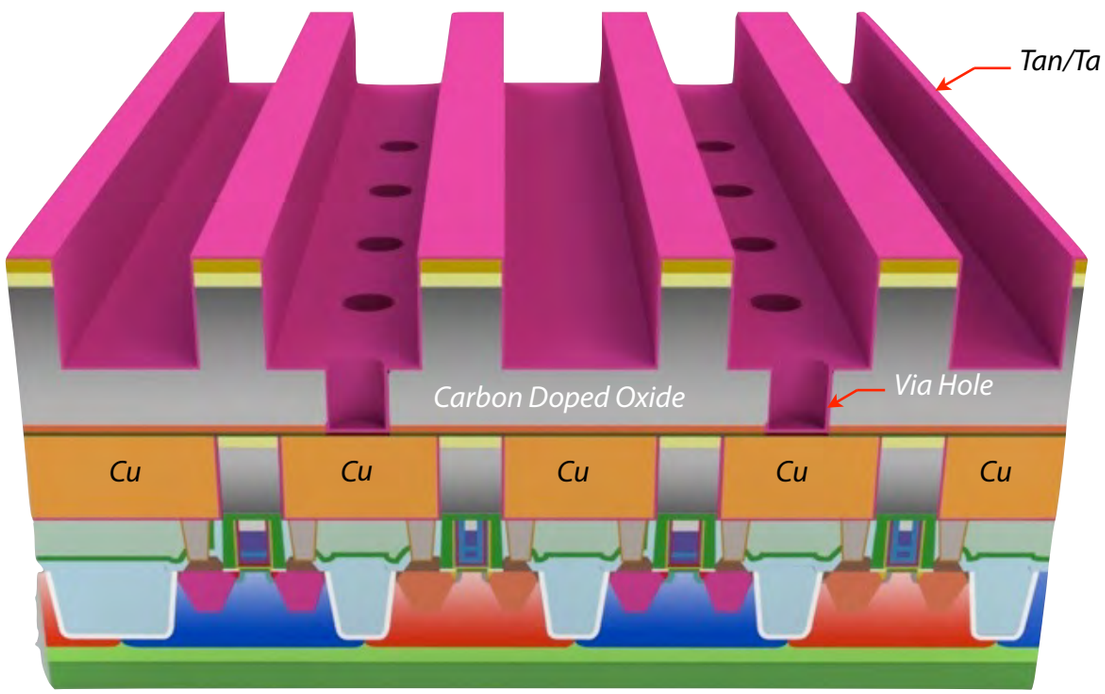
TaN/Ta Barrier :



The wafers are degassed to boil off contaminants in the trenches and then a wet clean followed by a soft Argon sputter to clean out any remaining polymeric residue, CuO, and carbon contaminants.

Next, ~80 Å of TaN followed by ~30 Å of Ta are deposited using IMP PVD. These two metals act as a barrier that will confine the not-yet-deposited Copper to the trenches. A re-sputter step follows both the Ta and TaN operations.

© 2018 Threshold Systems Inc.



© 2014 Threshold Systems Inc.

A head-on view of the Dual Damascene structure displaying the trench and numerous via holes coated in Tan/Ta.

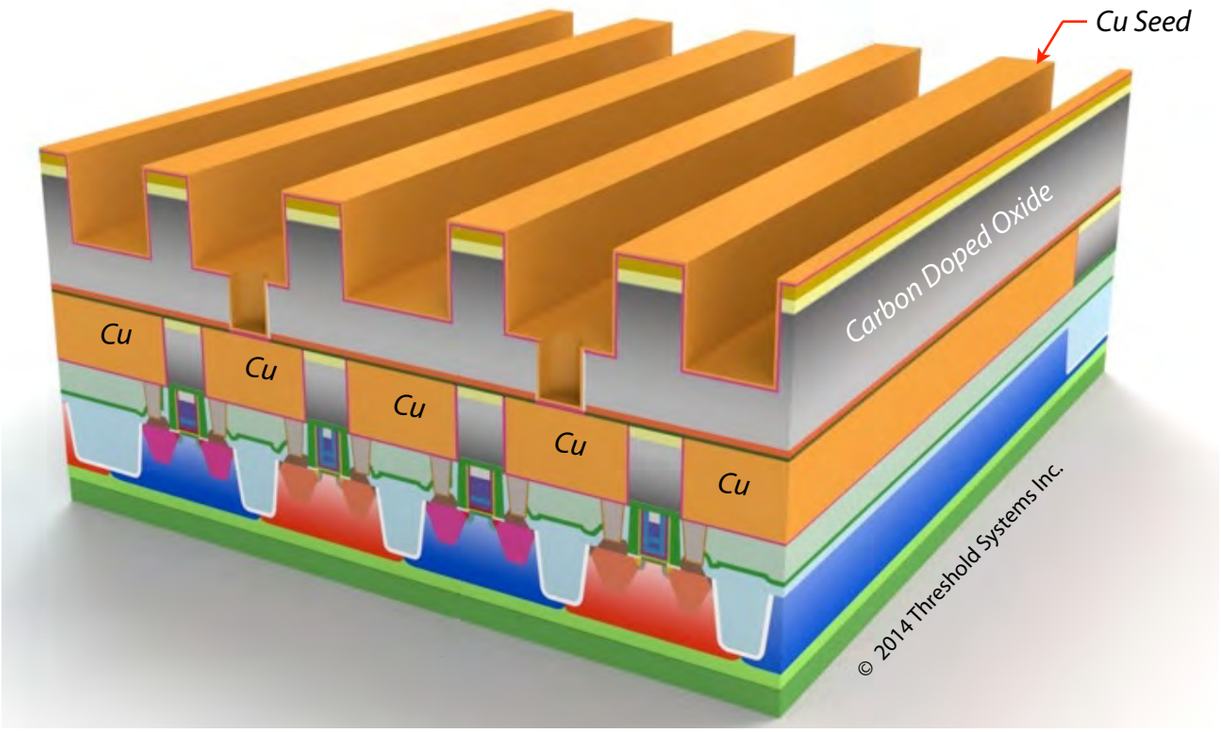
Like the contact holes, via holes are very tiny and difficult to fill with barrier metal. However, this problem is especially important for via holes since it is critical that they be continuously lined with Ta/TaN in order to entomb the copper that will later fill them. Any discontinuity in the Ta/TaN barrier metal will allow copper atoms to leak out of the via and will result in transistor poisoning.

Dedicated Ta/TaN ionized PVD tools combined with a sputter operation ensure that proper barrier coverage is achieved.

© 2018 by Threshold Systems Inc. All Rights Reserved.

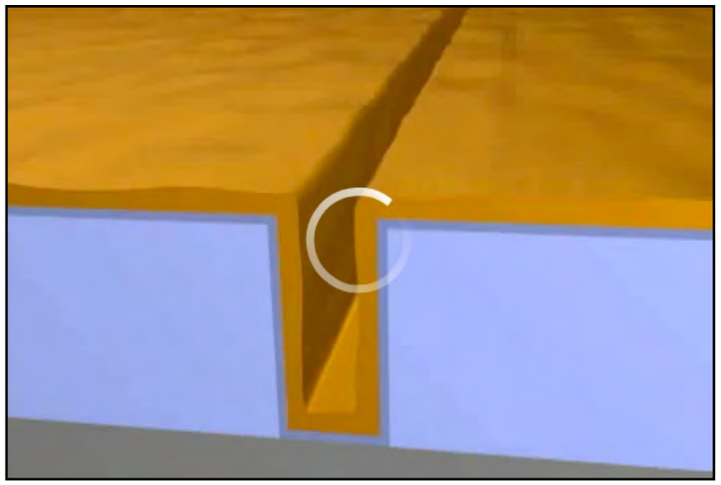
Cu Seed Deposition & Electroplated Copper:

ThresholdSystems



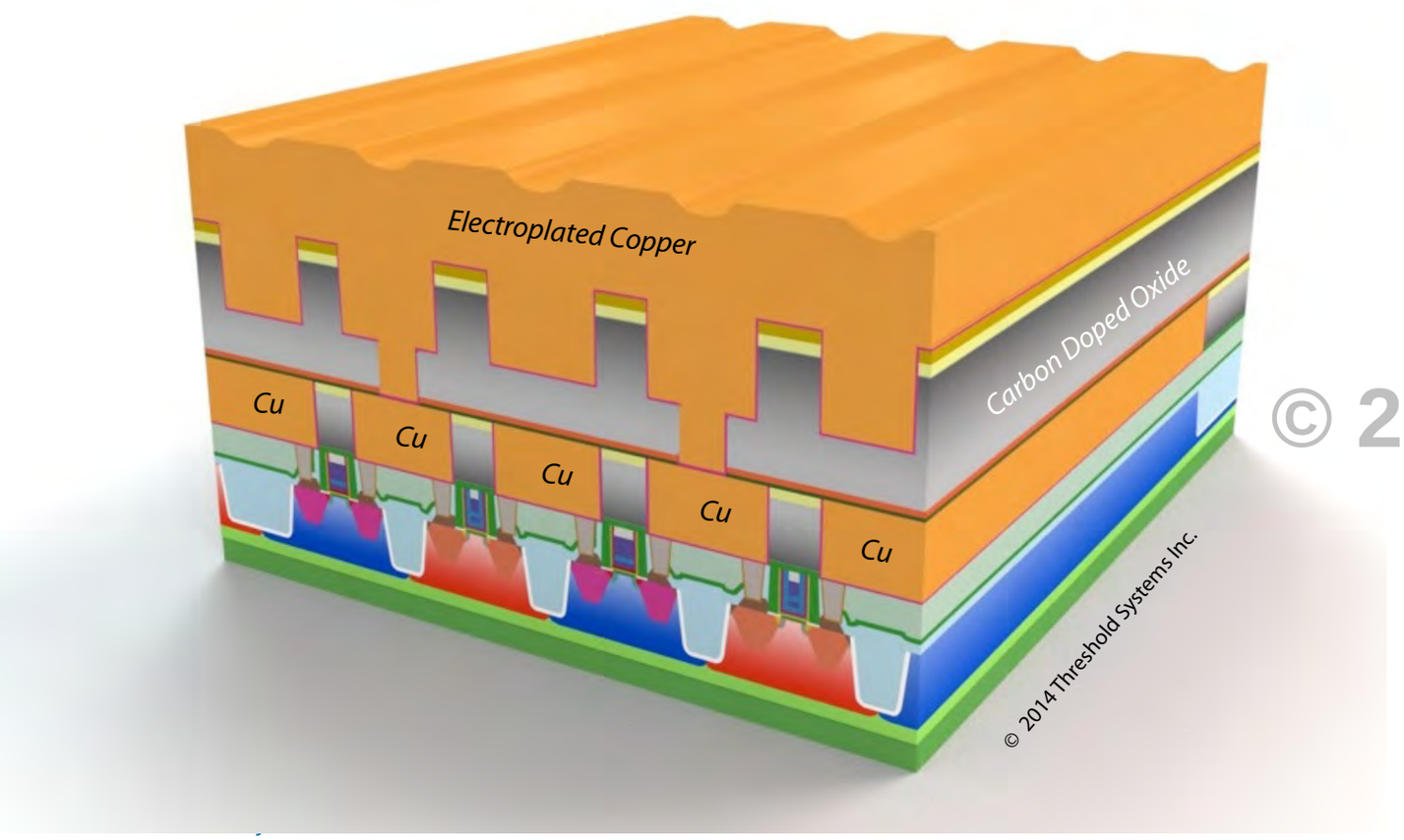
Without breaking vacuum, 300Å of Copper seed is deposited across the wafers which coats the interior of the trenches.

The Copper seed is required to ensure conformal void-free bulk Copper deposition (the following operation) and is a necessary precondition for the electroplating process.



Next, 6,000 Å of bulk copper is deposited using electrochemical deposition.

After deposition, the Copper is annealed at 300 °C for 90 seconds in forming gas to set the grain size.

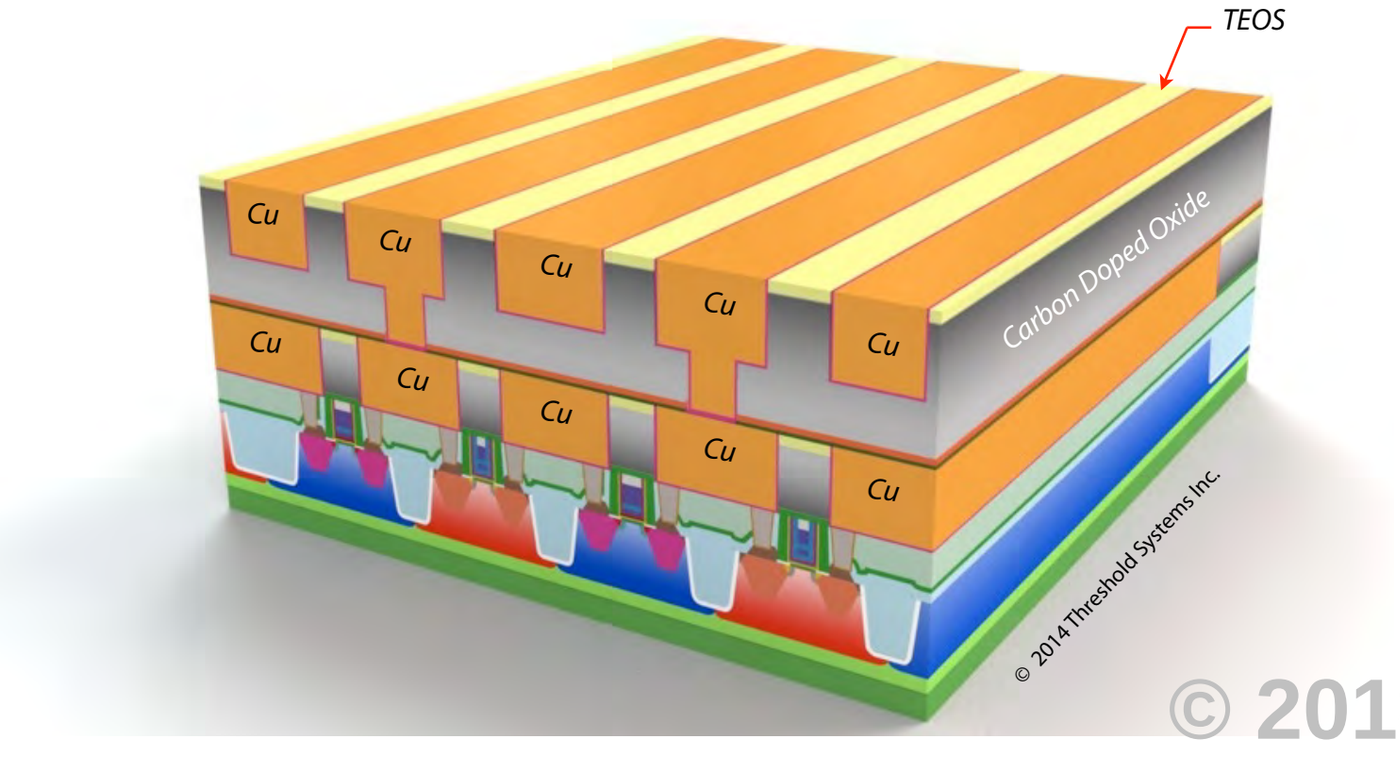


© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Copper CMP Polish & Final Seal:

ThresholdSystems

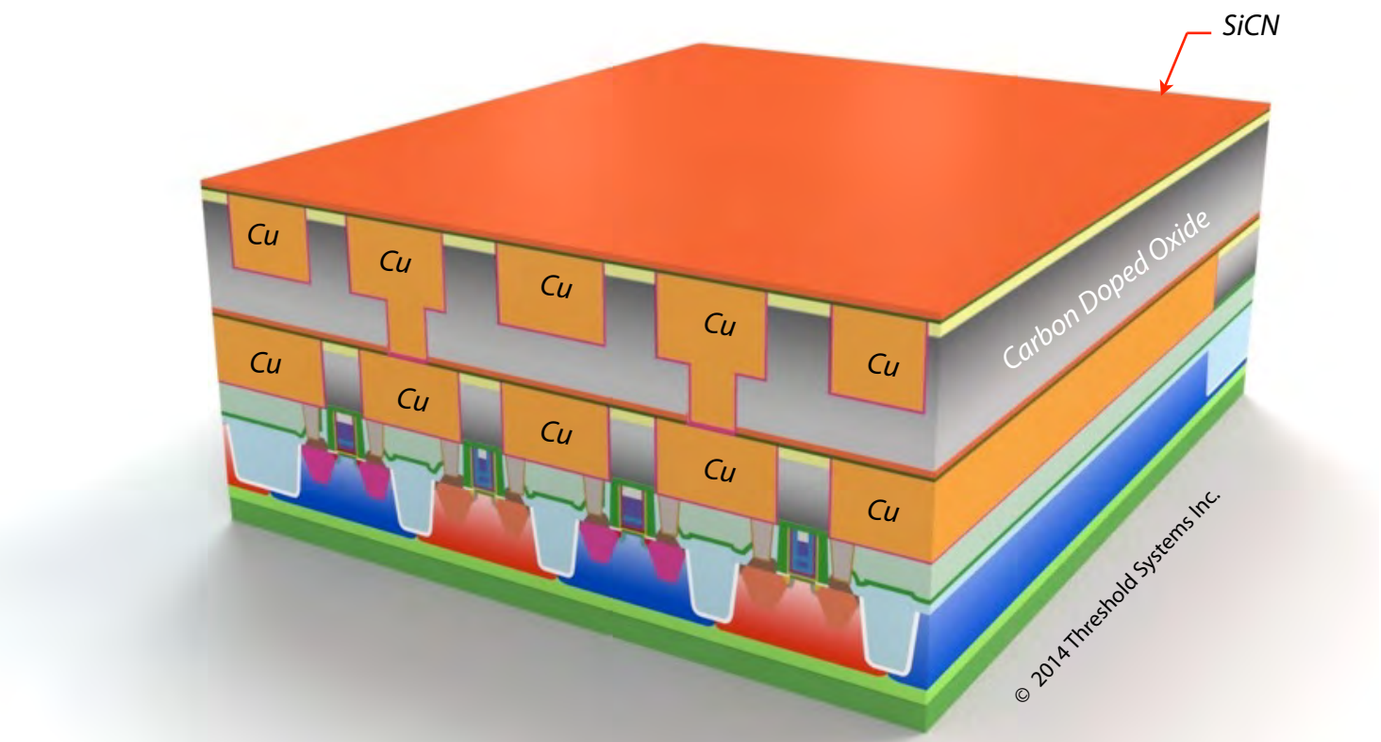


The Copper is polished back using CMP. This polishing also removes the Tantalum that was on the upper surfaces of the TEOS. The TEOS acts as a polish stop for the CMP process.

Next, an oxide buff operation is conducted to smooth out the surface of the TEOS oxide and ensure that the surface of the TEOS is coplanar with the top of the Copper lines.

After each CMP step the wafers are cleaned.

© 2018 Threshold Systems Inc.



After the same SRN layer is deposited, converted to CuSi and then to SiN, a 300Å layer of SiCN is deposited to seal off the top of the second layer of metal. The Copper in the second layer of metal (and its via) is now entombed with TaN/Ta on five sides and SiCN on the top.

Note that in this Dual Damascene integration methodology the photoresist never touched the low-k dielectric.

Seven further layers of metal trenches and vias are defined in a similar fashion.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Backend Metallization:

Resistance

Conductor Length

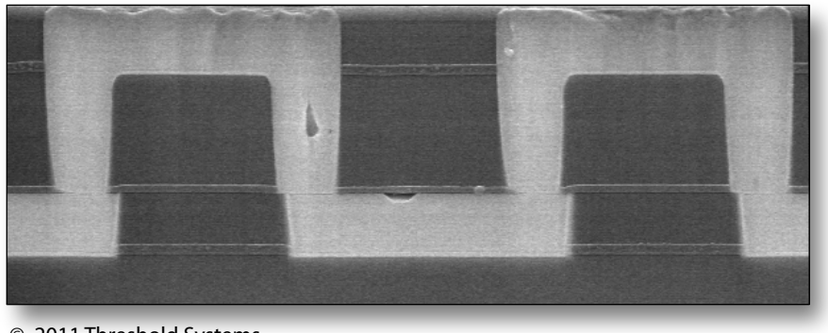
$$R = \rho \left[\frac{L}{A_x} \right]$$

Conductor Cross-sectional Area

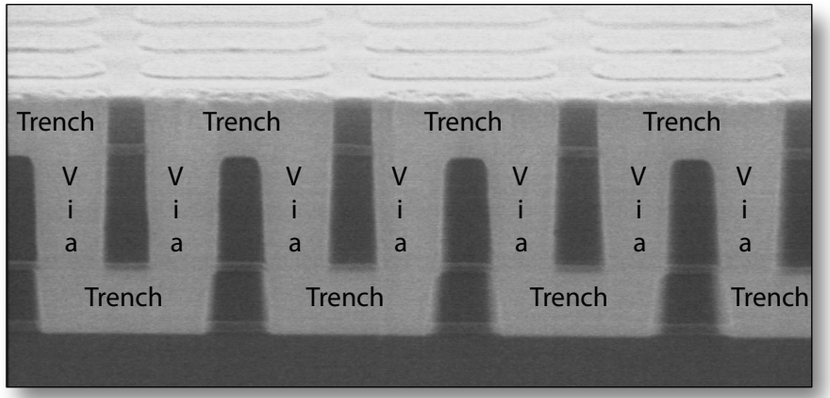
Resistivity

We want R to be LOW!

$$\tilde{\iota} = RC$$



© 2011 Threshold Systems



© 2011 Threshold Systems

© 2018 Threshold Systems Inc.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Metal Hierarchy:

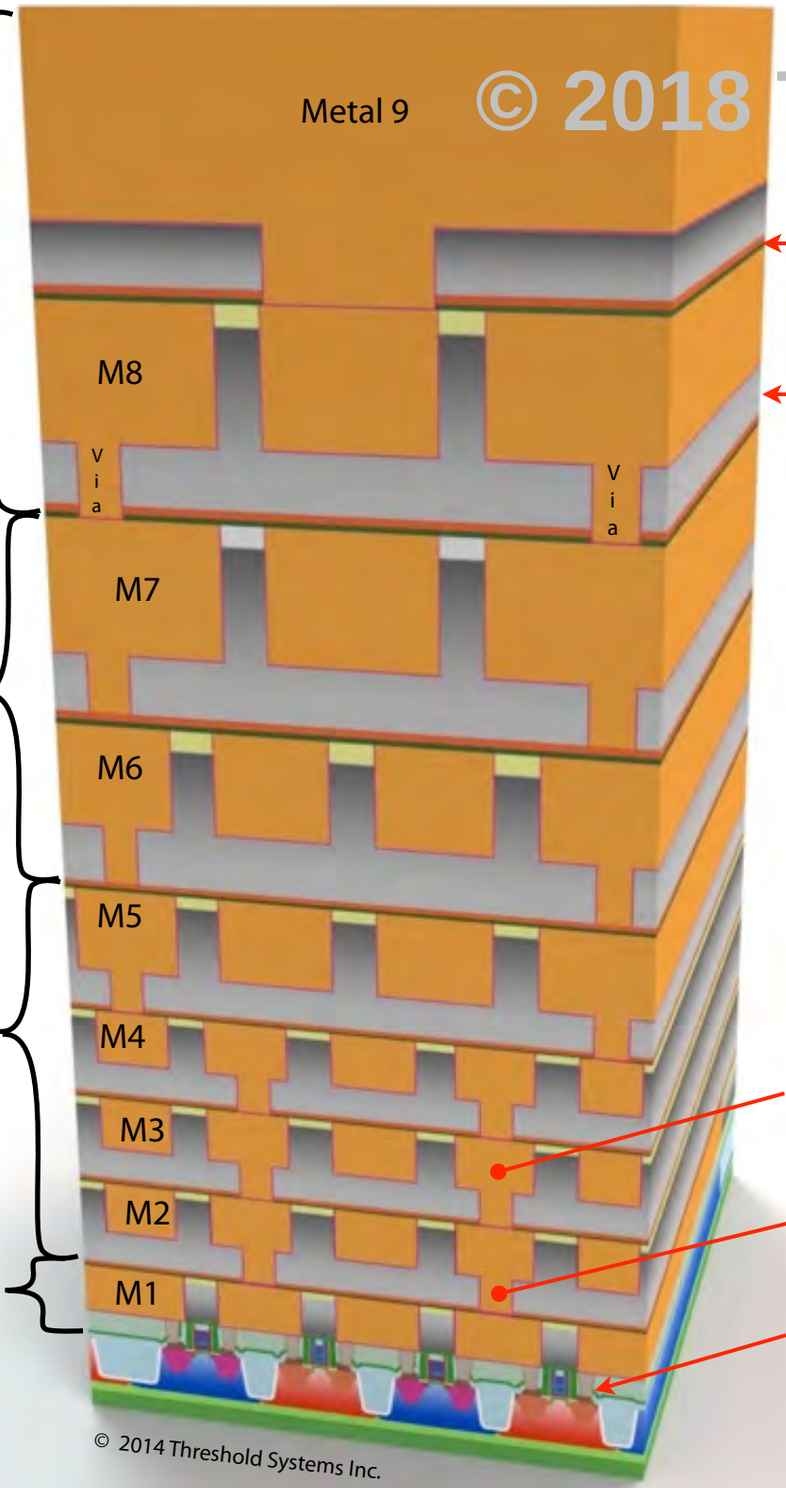
© 2018 Threshold Systems Inc.

Global Wires (up to 2 layers) usually 5-10 mm in length; often in FSG or oxide. They carry power, ground and clock signals

Semi-Global Wires (up to 4 layers) typically 0.5-5 mm in length; usually in CDO.

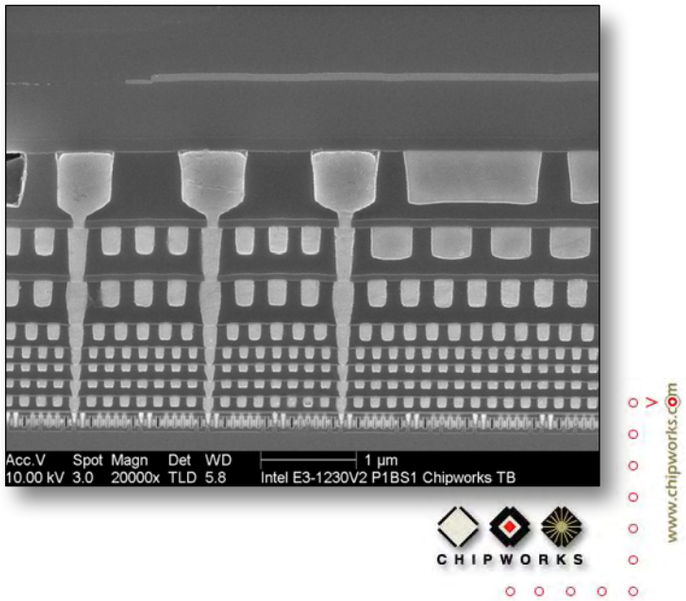
Intermediate Wires (up to 5 layers) typically <100 µm in length; usually in CDO

M1 & contacts (one layer) typically <50µm in length; usually in CDO



© 2014 Threshold Systems Inc.

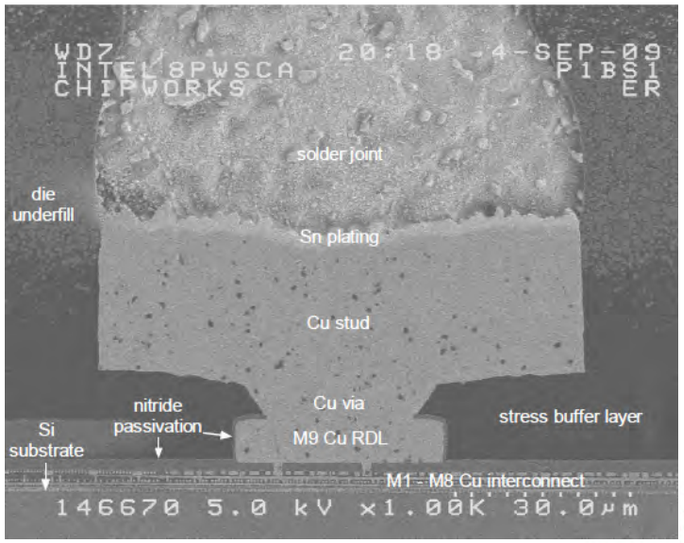
Wiring Hierarchy



© 2018 by Threshold Systems Inc. All Rights Reserved.

C4 Connections to the External World:

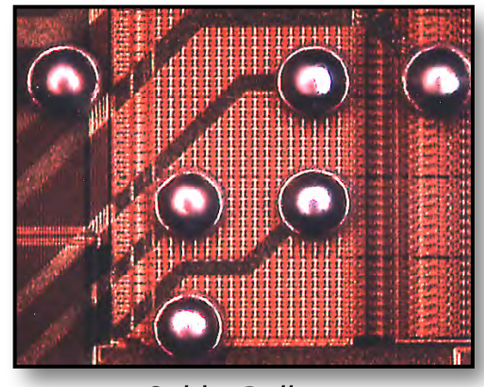
ThresholdSystems



© 2010 Chipworks - Courtesy of Chipworks

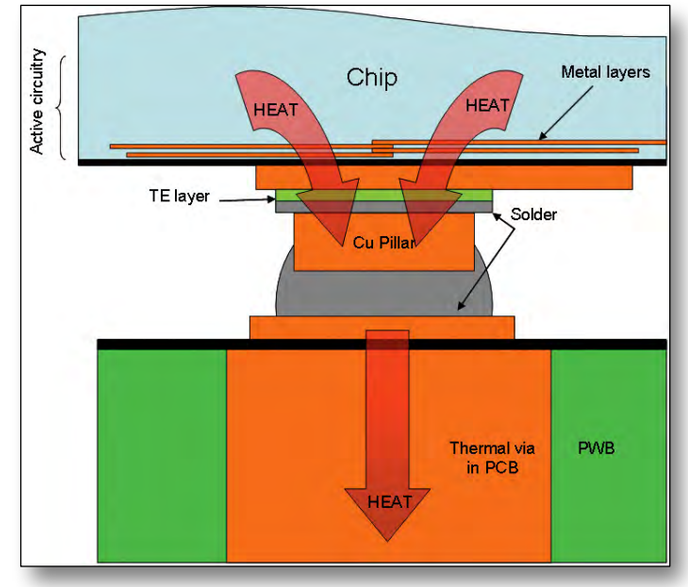
© 2010 Chipworks - Courtesy of Chipworks

C4 - "Controlled Collapse Chip Connection"



Solder Balls

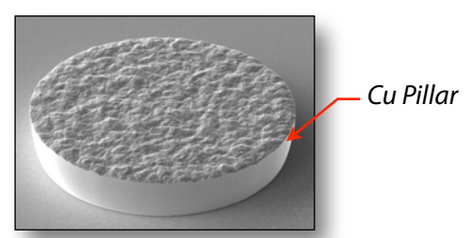
© 2011 Threshold Systems



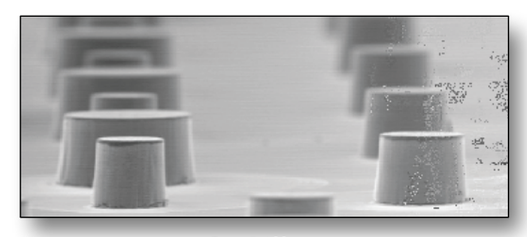
Copper Pillars

© 2018 Threshold Systems Inc.

Copper Pillars



Cu Pillars



Cu Pillars

The conventional solder balls have been replaced at the 45/32 nm node with "Copper Pillars" to comply with RoHS (Restriction of Hazardous Substances) and to act as a Peltier effect cooling conduit to drain heat away from the chip.

© 2018 by Threshold Systems Inc. All Rights Reserved.

Characteristics of the 20nm Generation:

© 2018 Threshold Systems Inc.

Advantages

- *Incorporated 3rd generation Metal Gate High-k technology with a more aggressively scaled dielectric (EOT reduced from 10Å to 9Å).*
- *Strained silicon continued to provided a dramatic speed boost and this technology node employed 4th generation strain technology that incorporating 40% Ge in the SiGe of the PMOS and the use of 1st generation SiC in the NMOS Source/Drains.*
- *The lithography process moved from 193 dry litho to 193i immersion lithography.*
- *NMOS S/D leakage has been reduced 5X; PMOS S/D leakage has been reduced 10X.*
- *The use of porous CDO (Carbon Doped Oxide) significantly lowered the k value of the inter-layer dielectric stack, which in turn permitted higher back-end speeds.*
- *TiN metal hard masks substantially improved trench morphology and enabled very narrow trenches to be cleanly defined.*
- *Amorphous carbon hard masks greatly enhanced the definition of STI, gate electrodes and narrow contact holes.*
- *The use of double patterning dramatically improved gate electrode scaling.*
- *Transistor counts at the 20 nm node reached 1.7+ billion.*

Disadvantages

- *This generation of microchips demanded the extensive use of 193 immersion lithography and extremely sophisticated and expensive reticles sets.*
- *The requirement for damascene-like replacement gates and their associated complex internal metallization morphologies dramatically increased processing complexity.*
- *Ultra-thin barriers for copper metallization faced new deposition challenges at the 20nm nodes as the metal lines scaled to extremely small dimensions.*
- *The extremely narrow gate electrodes present at the 20nm node posed special etch uniformity problems that required innovative forward/feedback mechanisms to control.*
- *The ultra-shallow Source/Drain regions required complex and careful implant schemes and millisecond spike anneal technology.*
- *The cost of a manufacturing facility dedicated to the production of 20 nm devices easily approached or exceeded seven billion+ dollars.*
- *The high level of processing complexity associated with the replacement metal gate significantly increased overall cost.*
- *There are significant questions regarding the scalability of the replacement gate methodology, particularly at the 7 nm node.*

© 2018 by Threshold Systems Inc. All Rights Reserved.