SEMICONDUCTOR FRONT-END · PART 3-2

Oxidation and Thermal Processing How heat reshapes silicon

Silicon owes much of its success to a material grown from its own surface: silicon dioxide. This lesson follows oxygen through that oxide, explains why steam grows it faster, and shows how heat can activate dopants while also moving them out of place. From batch furnaces to nanosecond laser pulses, the challenge is to obtain the change you need without damaging what is already there.

Semiconductor Front-End Series · Part 3-2 · Intermediate introduction · Helpful background: Lesson 3-1 · Updated 8 November 2026

In this lesson

1. Three ideas to keep in mind
2. Why silicon's oxide changed electronics
3. Growing an oxide: transport first, reaction at the interface
4. Why make the gate oxide thinner?
5. Annealing: activate, repair and preserve
6. Inside the furnace
7. Equipment suppliers and the markets they serve
8. The materials that make the hot zone possible
9. Thermal processing for stacked devices
10. What to remember—and what comes next
11. Check your understanding
12. Glossary
13. References and supporting material
14. Claims and supporting sources

How to read this lesson

Begin with the main text. The shaded physics boxes add optional detail and worked estimates. Figures are conceptual and not to scale. Numbered links lead to the supporting sources.

What you will learn

- Explain why silicon dioxide and its interface with silicon are important, and how a thermal oxide grows.
- Use the Deal–Grove model and a diffusion-length estimate to understand process behaviour and their limits.
- Connect thermal budgets, heating methods and furnace components to the equipment and materials used in fabrication.

1. Three ideas to keep in mind

- **The interface matters as much as the film.** Silicon dioxide is an excellent insulator, and processing can produce a stable, low-defect boundary with silicon. That combination helped make silicon the foundation of integrated circuits.
- **Oxidation consumes the wafer surface.** Oxidant travels through the existing oxide and reacts at the silicon interface. As the film thickens, the longer transport path increasingly slows its growth.
- **Heat is useful—and costly in process terms.** It activates dopants and repairs damage, but it also allows atoms to diffuse. The right temperature, duration and heating depth must be chosen together.

2. Why silicon's oxide changed electronics

A useful insulator with a useful interface

Silicon dioxide, **SiO_2**, is familiar in glass. On a silicon wafer it can serve a much more demanding purpose: separating electrical conductors, protecting the surface or helping a gate control the charge underneath it.

Stanford teaching material identifies SiO_2 and the Si/SiO_2 interface as a central reason for silicon's dominance in integrated circuits. It highlights selective etching, masking of common dopants, a breakdown field around **10^7 V/cm**, and a stable, reproducible interface [1] Reported . Masking performance depends on film thickness and process conditions; an oxide is not an unlimited barrier to every impurity.

The earliest transistor used germanium. Yet a convenient native oxide became an important advantage for silicon. Feigl's 1986 discussion contrasts the quality of silicon's oxide with the water-solubility problem of germanium oxide [2]. Germanium remains useful, but forming and maintaining the desired oxide/interface system presents different challenges.

Calling the oxide silicon's “rust” can be a helpful analogy for oxidation. It should not suggest that the film behaves like flaky iron rust, or that any oxide grown under any conditions is electrically perfect.

From surface stabilisation to MOS devices

Surface defects can trap charge and interfere with electrical control. At Bell Labs, **Mohamed Atalla and Dawon Kahng** overcame a key obstacle by using thermally grown oxide to reduce surface-state effects. The Computer History Museum dates their successful insulated-gate transistor work to **1959**, within its account of the 1960 MOS demonstration [3]. Atalla and colleagues also published their surface-stabilisation work in 1959 [4] Reported .

At Fairchild, **Jean Hoerni's planar process** retained an oxide layer to protect the junctions, rather than treating it only as a temporary processing aid [5]. These developments made oxide part of the device and its manufacturing method.

Over time, silicon oxide has been used for dopant masks, surface passivation, device isolation and gate insulation. Thermal growth can produce particularly clean films and low-defect interfaces [6]. These are process-dependent advantages, not a promise of zero defects.

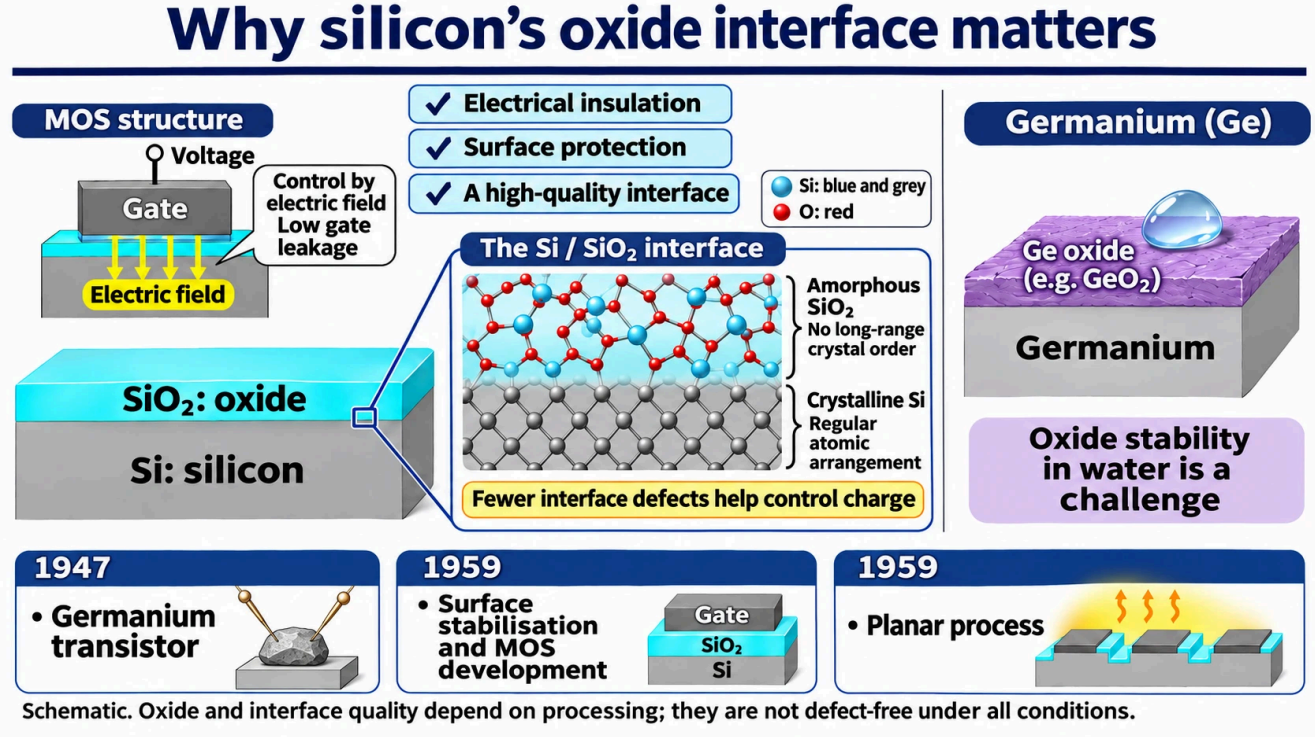


Figure 1 **Why silicon's oxide interface matters** (schematic)

Reading the figure

The oxide is amorphous; the silicon underneath is crystalline. Reducing interface defects helps the gate control charge in the silicon. The historical illustrations are schematic, rather than photographs of the original devices.

3. Growing an oxide: transport first, reaction at the interface

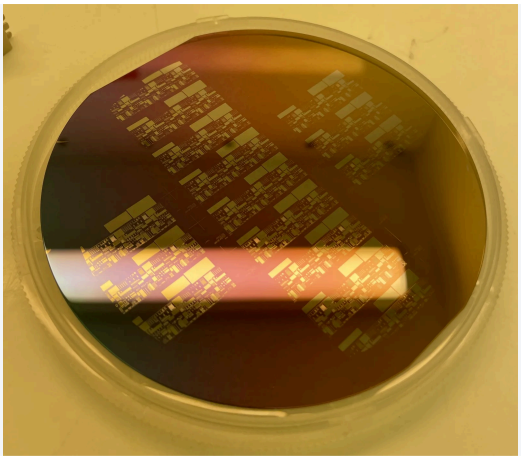
- **Dry oxygen or steam**

Thermal oxidation uses either **dry oxygen** or **water vapour** as the oxidising species [6]. The overall reactions are:

$$\text{Dry oxidation: } Si + O_2 \rightarrow SiO_2$$

$$\text{Wet oxidation: } Si + 2H_2O \rightarrow SiO_2 + 2H_2$$

Teaching examples place conventional oxidation roughly in the **700–1,200°C** range and describe steam oxidation as about ten times faster than dry oxidation [7]. The actual ratio varies with temperature, pressure, thickness and other conditions. Dry oxidation is useful when thin-film quality and thickness control are priorities; wet oxidation is useful when a thicker film must be grown efficiently.



The Commons description identifies an approximately 30 nm silicon oxide film, with oxide removed from selected areas. The photograph was taken under yellow room lighting; its colours should not be treated as a film-thickness measurement.
Photo: 90lbNose / Wikimedia Commons / CC0. Resized and converted to WebP for this article.

The interface moves down; the oxide also extends upwards

The oxide does not grow simply by piling new material onto its top. Oxygen or water-derived oxidant travels through the film already present, reaches the **Si/SiO_2 interface**, and reacts with silicon there [6]. The interface advances into the wafer.

For an oxide thickness **d**, approximately **0.44d** of silicon is consumed [6]. The resulting oxide occupies roughly 2.2 times the volume of the consumed silicon [7]. In a simple planar example, a **100 nm** oxide therefore consumes about **44 nm** of silicon and extends about **56 nm** above the original surface **Calculation** .

This geometry explains why the oxide's two surfaces move in different directions. It does not mean that contamination disappears automatically as the interface advances. The pre-clean discussed in Lesson 3-1 and the purity of the oxidation environment remain important.

The Deal–Grove model: three linked fluxes

Deal and Grove's 1965 model connects oxide thickness and time by considering the route from gas to reaction [8]. Three fluxes are useful [6]:

1. **F_1: transport from the gas to the oxide surface.**
2. **F_2: diffusion through the oxide.**
3. **F_3: reaction at the silicon interface.**

Under the model's steady-state approximation, the fluxes are equal [1]. The thickness–time relationship becomes:

$$x^2 + Ax = B(t + \tau)$$

Here x is oxide thickness, t is elapsed oxidation time, and τ accounts for the starting oxide in the model. A and B are process-dependent parameters.

- **In the thin-film limit of the model**, interface reaction dominates: thickness increases approximately in proportion to time. This is the **linear regime**, with rate constant B/A.
- **For thicker oxide**, diffusion through the film increasingly limits growth: thickness increases approximately with the square root of time. This is the **parabolic regime**, characterised by B.

B includes oxidant diffusivity and dissolved concentration; B/A is associated mainly with interface reaction [7]. A thicker film lengthens the path to the reaction site, so adding another increment of oxide takes longer.

A worked example at 1,000°C

City University of Hong Kong teaching material supplies the following parameters [6]:

Oxidation	A (μm)	B (μm²/h)	τ (h)
Dry	0.165	0.0117	0.37
Wet	0.226	0.287	0

For **100 nm = 0.1 μm**, rearrange the model:

$$t = (x^2 + Ax) / B - \tau$$

Case	Substitution	Estimated time
Dry	(0.01 + 0.0165) / 0.0117 − 0.37	About **1.9 hours**
Wet	(0.01 + 0.0226) / 0.287	About **0.11 hours**, or **7 minutes**

With these parameters and starting-film corrections, the wet case takes roughly **one seventeenth** as long Calculation . This is a worked example, not a transferable recipe or a fixed wet/dry speed ratio.



This instrument at LAAS-CNRS, France, measures changes in the polarisation of reflected light. An optical model is then used to infer properties such as film thickness. Measurement checks whether an oxidation process produced the intended result.
Photo: Guillaume Paumier (user:guillom) / Wikimedia Commons / CC BY-SA 3.0. Resized and converted to WebP.

Why steam helps: more oxidant can dissolve in the oxide

A fast reaction at the surface is not the whole explanation. KTH teaching material highlights the much greater solubility of water-derived oxidant in the oxide, giving an approximately **1,000-fold** comparison with oxygen under the example conditions [7]. A higher dissolved concentration means more oxidant is available for transport to the interface.

The model expresses this contribution through:

$$B = 2DC^* / N_1$$

D is oxidant diffusivity in the oxide, C* is its equilibrium concentration at the surface, and N_1 accounts for the oxidant needed per unit oxide volume. The solubility ratio alone is not the growth-rate ratio: diffusivity, reaction kinetics and operating conditions also matter.

The model has limits. Dry oxides thinner than roughly **20–30 nm** can grow faster than the basic model predicts [1][6]. Ultrathin gate-interface layers lie well within the range where corrections and more detailed descriptions are needed.

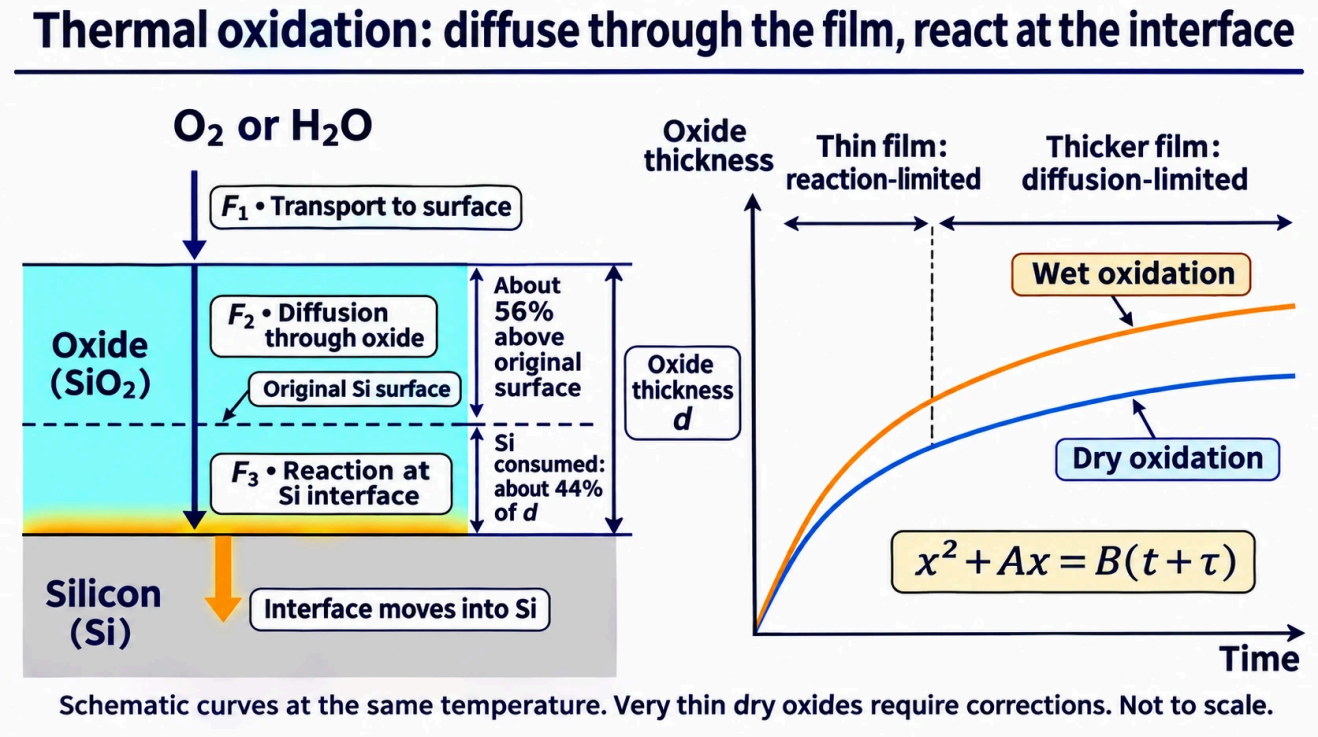


Figure 2 **Oxidant transport, interface motion and growth rate** (schematic)

Reading the figure

The upper part of the oxide is about 56% of its final thickness; the consumed-silicon depth is about 44%. The curves show qualitative behaviour at the same temperature, rather than a calibrated prediction for every oxidation condition.

4. Why make the gate oxide thinner?

- **Start with the goal: smaller, faster transistors**

For decades, the industry improved chips by fitting more transistors into a given area and improving the performance of each device. Moore's law describes the growth in the number of devices on an integrated circuit; it is not, by itself, a rule that guarantees faster switching. In his 2004 review, J. Robertson describes the historical increase in integration density and a roadmap in which gate length and oxide thickness decrease together [39].

Reported

The essential connection is **scaling**. As a transistor gets shorter, the gate still needs strong control over the channel beneath it. Reducing the gate-oxide thickness was one way to preserve that control as other dimensions shrank. A thinner film was a means to a better transistor, rather than the industry's objective in its own right.

- **What does a thinner oxide actually do?**

A gate is an electrode separated from silicon by a very thin insulator: the **gate dielectric**. In an n-channel example, a suitable gate voltage attracts electrons towards the silicon surface, creating a conducting path between source and drain. The field acts through the insulator; useful channel current flows along the silicon, not through the gate oxide.

The gate and channel behave as a capacitor. With the same dielectric material and other conditions held equal, a thinner film gives a higher capacitance per unit area. At a comparable gate bias, that allows more channel charge and potentially more drive current. More drive current can charge or discharge a given circuit load faster. This is a simplified explanation of the performance benefit: actual switching speed also depends on the load, device geometry, mobility and operating conditions. Thin oxide also helps the gate retain electrostatic control as gate length is reduced [39].

The layer being thinned is **the nanometre-scale insulator inside each transistor**. The aim is not to grind the wafer or the whole chip thinner. Wafer thinning is a separate fabrication operation with different purposes.

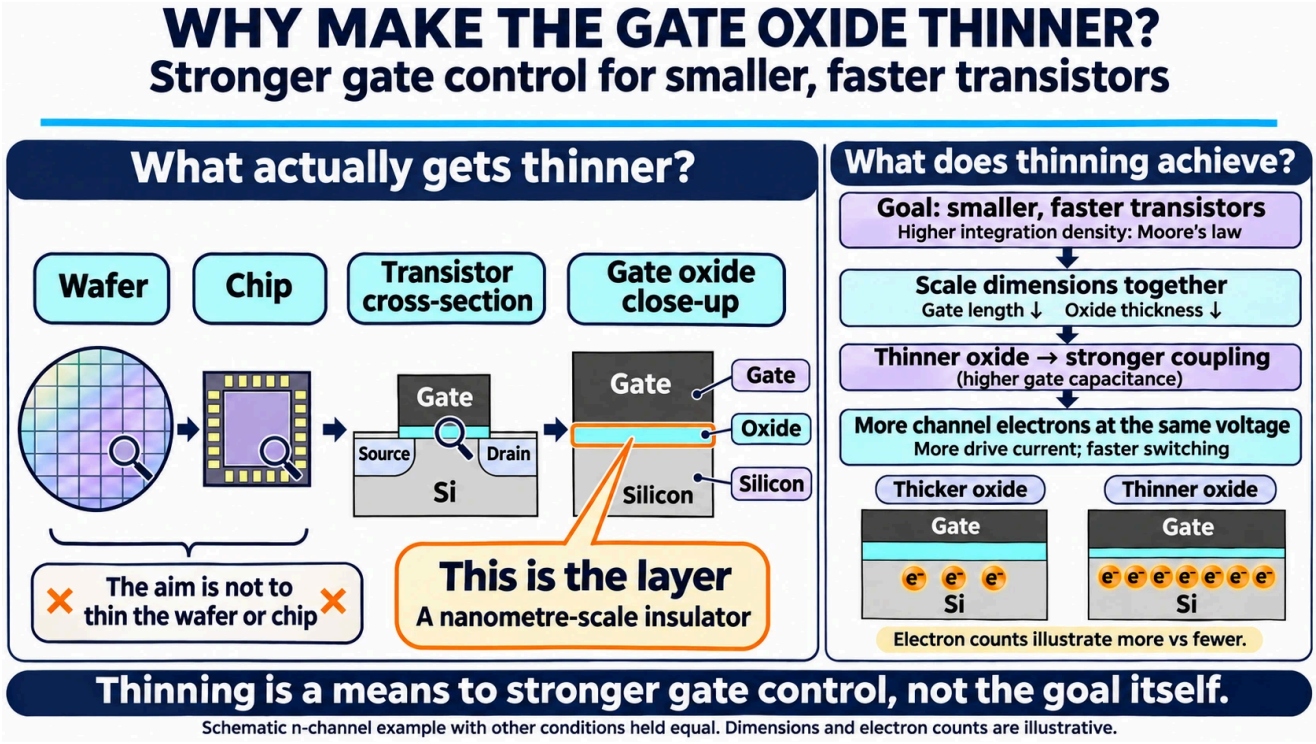


Figure 3 **Gate-oxide thinning: the location and the purpose** (schematic)

Reading the figure

The left-hand sequence identifies the thin insulating layer between gate and silicon. The right-hand sequence connects scaling to gate control and channel charge. The electron sketches are an n-channel example with the same material and comparable voltage, holding other conditions equal; the counts mean more or fewer, not actual electron numbers. Dimensions are not to scale. Stronger drive can help a device switch a given load faster, but the diagram is not a complete circuit-speed model.

Why thinning eventually stopped being enough

Very thin SiO_2 allows electrons to **tunnel through** the barrier, creating unwanted gate leakage. Intel's January 2007 announcement described a **1.2 nm** SiO_2 dielectric in its 65 nm generation, approximately five atomic layers in the company's wording. Further thinning increased leakage. For its 45 nm technology, Intel introduced a hafnium-based high-k dielectric with a metal gate and reported more than a tenfold reduction in gate leakage relative to its earlier technology [9]. Reported

Why replace a thin film with a thicker one? Because the desired result was strong gate control with acceptable leakage. If a different material provides that control at greater physical thickness, it offers another route to the same goal.

Capacitance and EOT: electrically thin, physically thicker

A simple parallel-plate model gives the dielectric capacitance as [39]:

$$C = \varepsilon_0 K A / t$$

Here ε_0 is the permittivity of free space, K is relative permittivity, A is area and t is physical thickness. At fixed area, either reducing t or increasing K raises capacitance. A higher-K material can therefore provide the same capacitance at a greater thickness. The thicker barrier helps suppress direct tunnelling; leakage is not reduced to zero [39].

Robertson's table lists **K = 3.9 for SiO_2** and **25 for HfO_2** [39]. Reported Using those values in an ideal, single-layer calculation, the HfO_2 thickness corresponding to 1.2 nm of SiO_2 is:

$$t = 1.2\ \text{nm} \times 25 / 3.9 \approx 7.7\ \text{nm}$$

That is about **6.4 times thicker** for the same ideal dielectric capacitance. Calculation This is a teaching calculation, not Intel's disclosed high-k film thickness or a universal value for every HfO_2 film.

Equivalent oxide thickness (EOT) expresses the capacitance as the thickness of SiO_2 that would give the same result:

$$EOT = (3.9 / K) \times t$$

The ideal 7.7 nm HfO_2 layer above has an EOT of approximately 1.2 nm. Calculation Physical thickness and electrical thickness answer different questions: how far apart the surfaces are, and how strongly they are capacitively coupled.

Real gate stacks have additional layers and electrical effects. For a high-k layer and an SiO_2 interfacial layer, the dielectric contributions act in series: the SiO_2 layer adds to the stack's EOT. Robertson discusses this limitation [39]. The simple single-layer calculation should not be applied to the entire stack without accounting for its constituents.

The interface layer still matters

High-k did not eliminate silicon oxide. KTH material discusses the transition around the 45 nm generation [7], while imec's 2026 CFET discussion describes HfO_2 protecting an underlying SiO_2 interfacial layer [10]. The layers can perform complementary roles: high permittivity for coupling, and a well-controlled interface with silicon.

- **Making a very thin oxide locally**

In situ steam generation (ISSG) is another oxidation approach. Applied Materials engineers described reacting hydrogen and oxygen near a hot wafer at low pressure, generating reactive oxygen species close to the surface. Their 2000 account discusses films around **1 nm** and a more than threefold improvement in charge-to-breakdown relative to the furnace-grown comparison [11]. These are results for the processes discussed in that account, not universal performance ratios for ISSG.

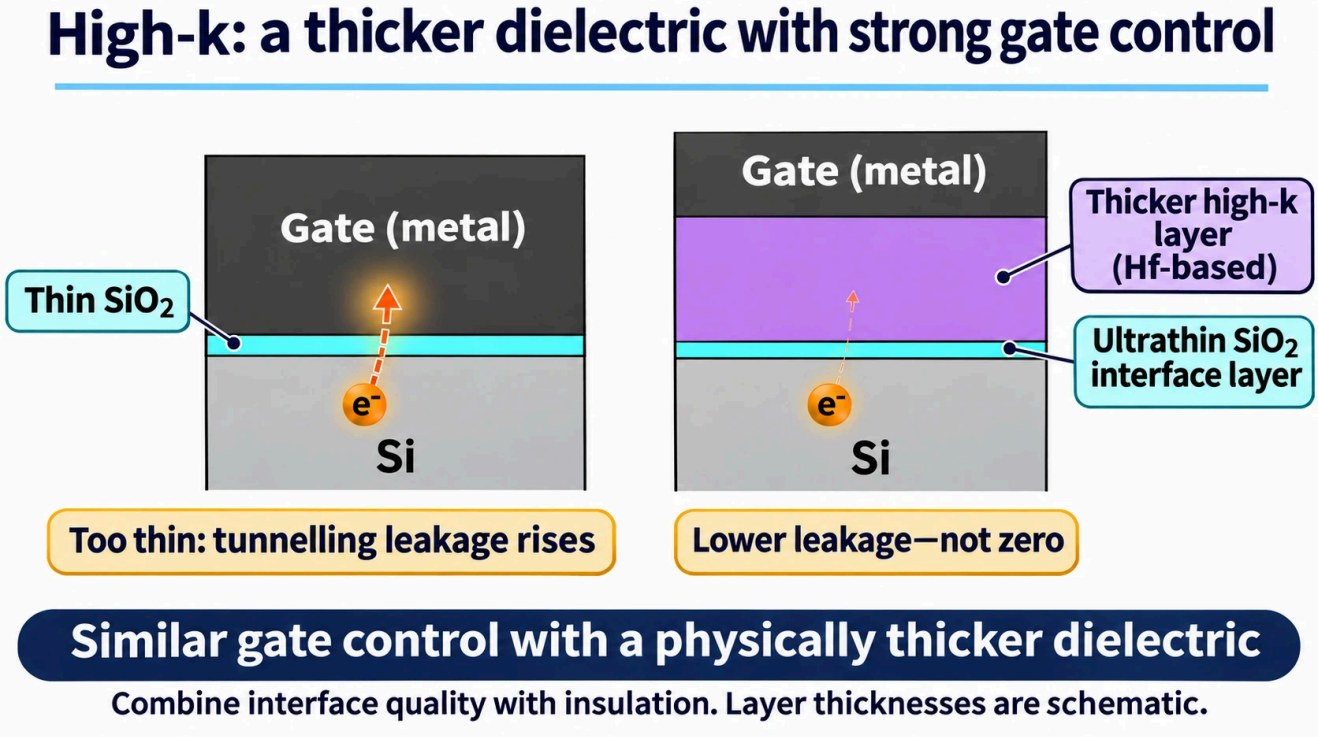


Figure 4 **High-k: greater physical thickness with strong gate control** (schematic)

Reading the figure

The objective is strong gate control with limited leakage, rather than minimum physical thickness for its own sake. A thicker high-k stack can have an EOT comparable to a thinner SiO_2 dielectric; the interface layer contributes to the total EOT. The arrows compare leakage qualitatively, not as a measured current ratio, and do not imply zero leakage. Layer thicknesses are schematic.

5. Annealing: activate, repair and preserve

- **What heat is meant to achieve**

Oxidation deliberately changes silicon into an oxide. **Annealing** uses a chosen thermal cycle to change the state of a material or interface, often without deliberately growing an oxide.

Three common purposes are:

- **Dopant activation.** Implantation introduces atoms such as boron or phosphorus. Thermal processing helps place them in electrically useful configurations, commonly substitutional sites in the silicon lattice, so they can supply holes or electrons.
- **Damage repair.** Implantation disrupts the crystal; annealing can restore order and reduce defects.
- **Interface improvement.** Heating in a hydrogen-containing atmosphere can passivate electrically active defects. KTH examples use roughly **350–500°C**, reducing interface-state densities from around **10^{11}** to **10^{10} cm^{-2} eV^{-1}** [7]. The result depends on the starting interface and treatment conditions.

The same heat also moves atoms

Atoms do not stay fixed while these improvements occur. Dopants diffuse, broadening the carefully positioned distribution. A process that improves activation may therefore make a junction too deep or blur a narrow doped region.

Diffusion length and the thermal budget

A characteristic diffusion scale can be written as [12]:

$$L \approx 2\sqrt{(Dt)}$$

D is the diffusion coefficient and t is time. This expression assumes a constant D; the exact definition and numerical factor depend on the diffusion problem being described.

At the same D, cutting time to one hundredth cuts this length scale to one tenth Calculation . That qualification matters: changing temperature changes D, often very strongly.

A real heating cycle is better understood through its complete temperature history. For diffusion, the contribution of D(T) is accumulated over time. A short high-temperature interval can matter much more than a longer cool interval. A **thermal budget** is therefore not simply added heat energy or temperature multiplied by time.

Stanford scaling discussions connect smaller devices to shallower junctions [13]. The available diffusion allowance shrinks with the structure. Implantation can also cause **transient enhanced diffusion (TED)**, in which excess point defects temporarily accelerate dopant motion during annealing [14].

The engineering task is to balance activation, defect repair and limited redistribution—not merely to maximise temperature or minimise time.

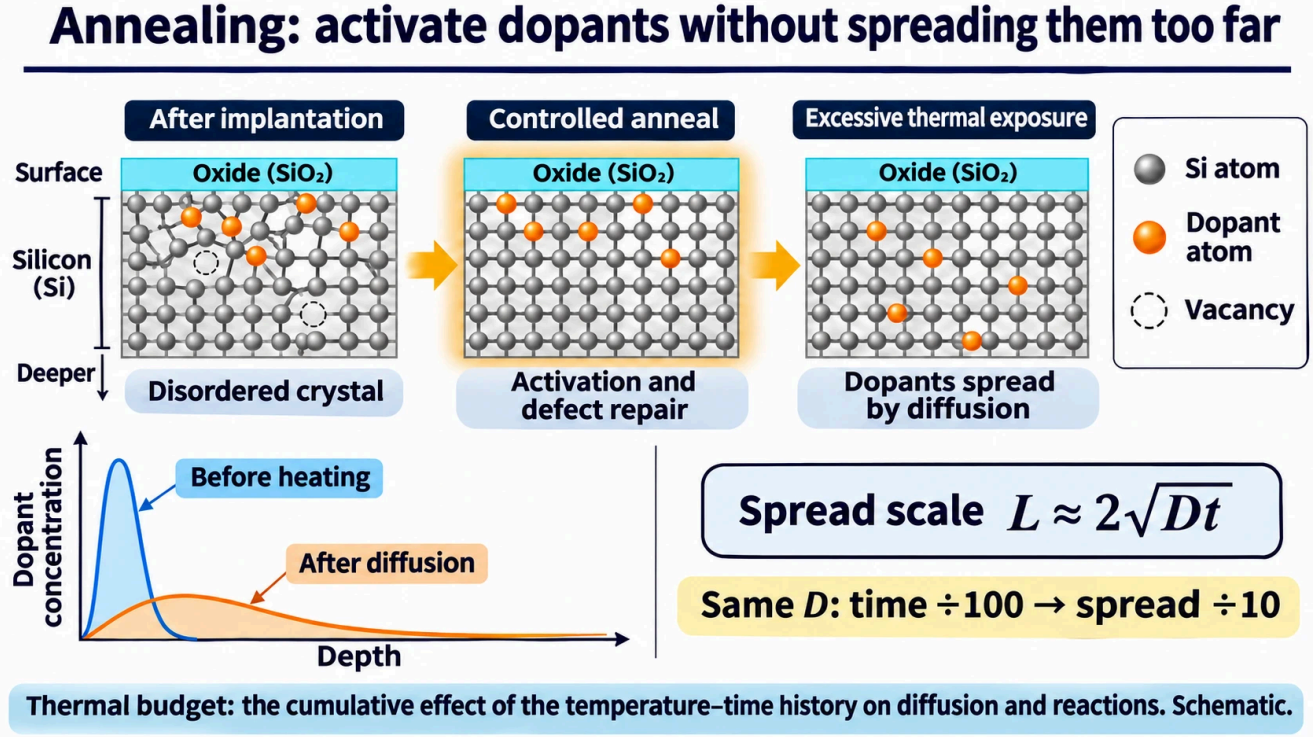


Figure 5 **Activate dopants without spreading them too far** (schematic)

Reading the figure

Orange spheres are dopant atoms, not free electrons. Each panel shows the same illustrative number of dopants. The concentration curves and lattice drawings are qualitative; the time-scaling comparison holds only when D is unchanged.

- **Deliver the heat in a shorter interval**

Method	Typical duration in this comparison	How the heat is delivered
Batch furnace	Tens of minutes to hours	Many wafers are heated inside a process tube
Rapid thermal processing (RTP)	Seconds	Intense radiation rapidly heats a wafer, typically one per chamber
Flash-lamp annealing	Milliseconds	A preheated wafer receives a short, intense light pulse
Laser annealing	Microseconds to nanoseconds	Optical energy is delivered locally or to a shallow surface region

These are broad teaching categories, not limits for every system. Heating depth depends on wavelength, absorption, pulse duration, thermal transport and the process setup.

Mattson's technical presentation discusses surface-focused heating for sub-0.5-second anneals, pulse sources above **10 kW/cm²**, and preheating at around **150 K/s** in its examples [15]. SCREEN's LA-3100 uses xenon flash-lamp pulses lasting a few milliseconds to achieve dopant activation with a limited thermal budget [16].

SCREEN/LASSE material describes **160 ns ultraviolet laser pulses**, including surface-melting approaches to high dopant activation [17]. This is one cited implementation, not the duration or operating mode of all laser annealing.

Short pulses bring their own compromises. The Mattson presentation emphasises the balance between diffusion, activation and defect repair: a millisecond-scale cycle may not fully repair every kind of implantation damage [15]. A shorter process is beneficial only if it also produces the required material state.

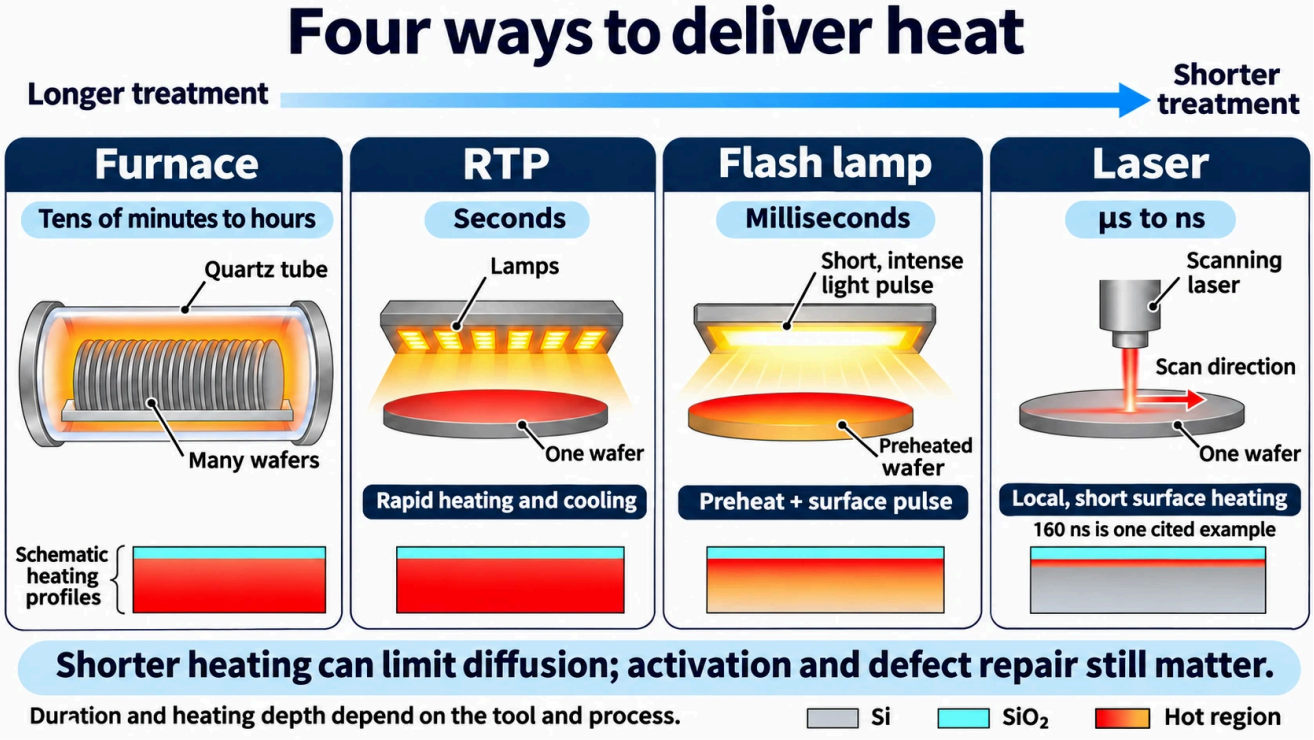


Figure 6 **Four ways to deliver heat** (schematic)

Reading the figure

The furnace at the left is a horizontal example; Figure 7 shows a vertical furnace. The lower heating profiles are schematic. These methods are not interchangeable recipes, and a shallow hot region does not imply that the rest of the structure is completely unaffected.

6. Inside the furnace

- **Horizontal tubes and vertical wafer stacks**

Historically, a horizontal furnace held wafers upright in a long boat inserted into a horizontal quartz tube. The Semiconductor History Museum of Japan describes **100–150-wafer** loads and particle generation associated with contact between the boat and tube. It records joint vertical-furnace development by Hitachi and Kokusai Electric in the mid-1980s, followed by the VERTEX series in **1986**. Boat rotation improved within-wafer uniformity, and the arrangement supported automated handling [18].

In a **vertical furnace**, the tube and boat extend vertically, but each wafer is held **horizontally**. The boat is typically loaded from below. “Vertical” describes the stack and reactor orientation, not an individual wafer standing on edge.



The Commons description identifies centrotherm furnaces used for diffusion and thermal oxidation in Toulouse, France. The horizontal process tubes are arranged at several heights; this is not a photograph of the inside of a vertical furnace.
Photo: Guillaume Paumier (user:guillom) / Wikimedia Commons / CC BY-SA 3.0. Resized and converted to WebP.



Another view of the same installation shows wafers held upright in loading fixtures. It illustrates batch handling for horizontal tubes, rather than the horizontal-wafer stack used in a vertical furnace.
Photo: Guillaume Paumier (user:guillom) / Wikimedia Commons / CC BY-SA 3.0. Resized and converted to WebP.

Batch processing remains useful when many wafers need the same treatment. KOKUSAI ELECTRIC lists a 200 mm vertical-furnace model accommodating **175 wafers** [19]. Stanford material also discusses the use of vertical furnaces, rapid thermal oxidation and fast-ramp furnaces [1]. The choice follows the thermal and productivity requirements of the process.

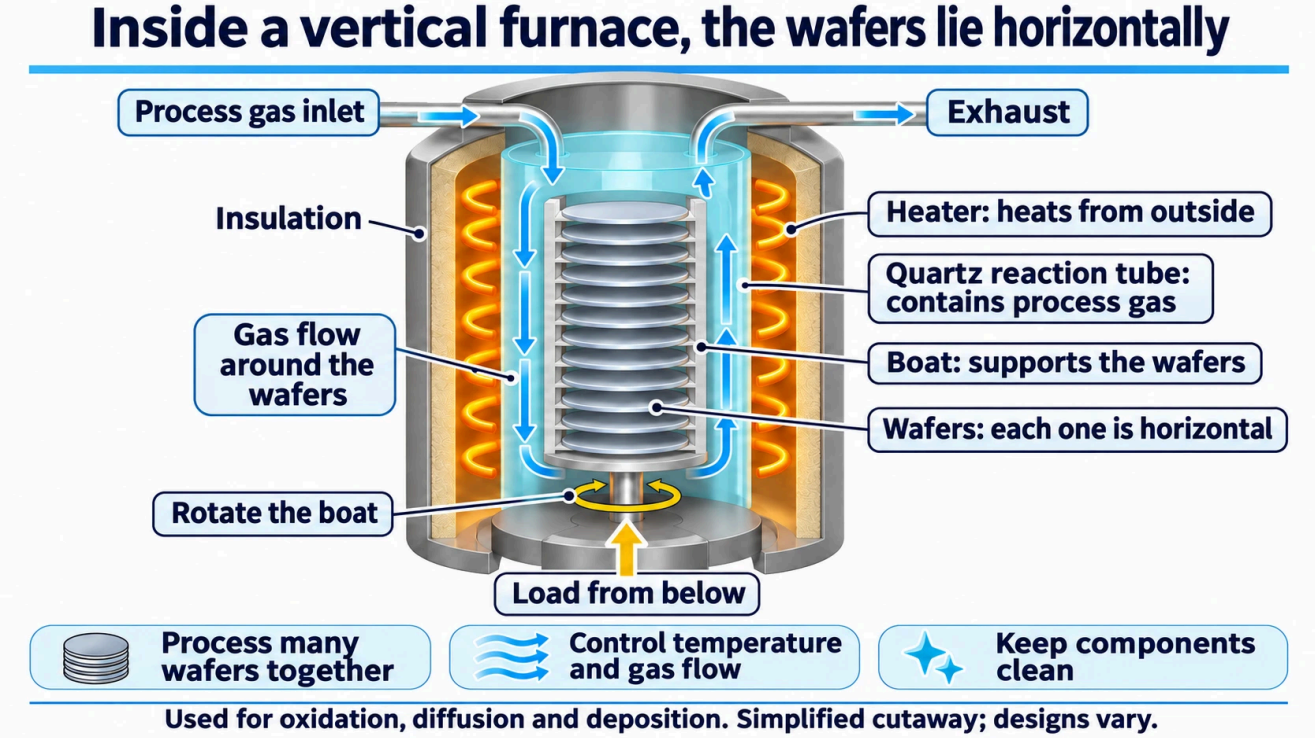


Figure 7 **A vertical furnace with horizontal wafers** (schematic)

Reading the figure

Heaters sit outside the reaction tube. The tube contains process gas, while the boat supports separated horizontal wafers. The gas path and mechanical arrangement vary between actual tools; the drawing is a simplified teaching example.

- **Knowing the temperature is part of controlling the process**

Rapid heating is useful only when wafer temperature can be measured and controlled. Optical thermometry avoids touching the wafer, but must account for how the wafer and surroundings emit, reflect and transmit radiation.

A **2006 NIST paper** discusses traceable non-contact temperature measurement and reports an uncertainty around **2.3°C** for a calibration example involving a thermocouple-equipped wafer [20]. That figure describes the study's measurement arrangement, not a universal error for present-day RTP tools.

7. Equipment suppliers and the markets they serve

The following companies illustrate different parts of thermal processing. The table is alphabetical. A furnace used for deposition and an RTP tool used for annealing do not belong to exactly the same market category.

Supplier	Headquarters / context	Examples and disclosed scope
Applied Materials	United States	RTP, including Vantage/Radiance; annealing and oxide/oxynitride processing in the cited filing [24]
ASM International	Netherlands	SONORA and A400 vertical furnaces [23]
KOKUSAI ELECTRIC	Japan	Batch deposition, oxidation, diffusion and annealing; reports 49.2% of the 2025 batch-deposition category [22]
Mattson Technology	United States; acquired by a China-based investor consortium	RTP and millisecond annealing [15][26]
NAURA	China	Vertical furnaces and RTP; thermal-equipment sales above CNY1 billion in the first half of 2025 [25]
SCREEN	Japan	Flash-lamp annealing, including LA-3100 [16]
Tokyo Electron	Japan	Vertical furnaces, including TELINDY PLUS; estimates a leading diffusion-furnace position [21]

Read the market definition before the ranking

Japan-based suppliers have a substantial presence in batch furnaces, but their published leadership claims do not measure identical categories.

Tokyo Electron describes itself as number one in diffusion furnaces, including thermal deposition within its classification, based on its own estimate. Its furnace business began with a Thermco distribution agreement in **1964**, followed by domestic production in **1968** [21].

KOKUSAI ELECTRIC, whose predecessor business was part of Hitachi Kokusai Electric, listed on the Tokyo Stock Exchange Prime Market in **2023**. Its 2026 report cites Gartner for a **49.2%** share of 2025 batch deposition, up eight percentage points, and gives its own estimate above **80%** for batch-ALD-capable equipment. It reports revenue of **JPY235.1 billion** for the year ended March 2026; regional shares include China **39%**, South Korea **24%**, Taiwan **18%** and Japan **10%** [22].

Tokyo Electron's diffusion-furnace category and KOKUSAI's Gartner-derived Tube CVD category should not be compared as if they were one league table. In a separate treatment category combining RTP with oxidation/diffusion, KOKUSAI's report cites a **US$2.8 billion** 2025 market and an **11.8%** company share [22].

NAURA reports thermal-equipment sales above **CNY1 billion** in the first half of 2025 and cumulative vertical-furnace shipments reaching **1,000 units**. Its report estimates a 2024 global thermal-equipment market around **CNY23 billion** [25]. The dates and definitions matter when placing these figures alongside other disclosures.

For short thermal cycles, the source set includes Applied Materials' RTP portfolio [24], SCREEN's flash systems [16] and Mattson's RTP/millisecond work [15]. Mattson's **2016** acquisition by an E-Town Dragon-led consortium provides relevant ownership context [26], rather than a measure of technical market share.

A proposed acquisition that did not close

Applied Materials agreed in July **2019** to acquire Kokusai Electric from KKR for **US$2.2 billion** [27]. The agreement ended in March **2021** after Chinese regulatory approval was not obtained by the deadline; Applied Materials paid a **US$154 million** termination fee [28].

This historical transaction illustrates how consolidation in equipment supply can depend on decisions across several jurisdictions. It is not evidence that the suppliers' different product categories are interchangeable.

8. The materials that make the hot zone possible

- **Clean surfaces at high temperature**

A reaction tube or wafer boat must survive repeated heating without adding unacceptable contamination, distortion or particles. Fused silica—often called **quartz glass** in equipment literature—is widely used because it combines useful purity, thermal and processing characteristics. Shin-Etsu Quartz describes it as an unusually well-suited material for these demands [29].

It is not the only option. Larger or more demanding structures may use **silicon carbide (SiC)** for its strength and stiffness; CoorStek discusses this contrast with conventional quartz tubes [34]. The right material depends on temperature, chemistry, mechanical load and the contamination requirement.

Supplier	Headquarters / context	Examples
CoorStek	United States	SiC tubes and boats [34]
Ferrotec	Japan	Quartz components, SiC boats and tubes, silicon components [32][38]
Heraeus Covantics	Germany	Fused silica for semiconductor processing [33]
Shin-Etsu Quartz	Japan; shareholders include Shin-Etsu Chemical and Heraeus	Fabricated quartz products and quartz crucibles for crystal growth [30]
Tosoh Quartz	Japan	Oxidation/diffusion tubes and boats, including tubes over 1,000 mm in diameter [31]

Tosoh Quartz traces its company history to a Japanese quartz-glass manufacturer founded in **1936** [31]. Ferrotec reported approximately **JPY34.4 billion** of quartz-product sales for the year ended March 2026 and described growing ceramic/quartz demand associated with generative-AI-related semiconductor equipment [32]. These disclosures show particular businesses and capabilities, not a combined global market share.

- **Upstream raw materials also matter**

One route to high-purity quartz products starts with exceptionally pure natural quartz. At Spruce Pine, North Carolina, **Sibelco** announced an approximately **US$200 million** expansion in 2023, intended to double high-purity-quartz capacity [35].

This is one supply route. Not all quartz glass comes from that location, and synthetic fused silica provides another manufacturing route. A mine, a glass producer and a precision component fabricator perform different jobs in the chain.

9. Thermal processing for stacked devices

The next challenge is to apply the required heat where it helps, while preserving structures already built underneath or nearby.

- **Sequential 3D integration.** In imec and Soitec's **2018** demonstration, the upper-layer process stayed below **525°C** to manage the thermal budget [36]. This is a constraint for that integration example, not a universal temperature limit for every stacked device. IRDS also discusses the need for lower-temperature steps after forming DRAM peripheral circuitry [37].
- **New transistor stacks.** imec's 2026 CFET discussion identifies the interaction between high-temperature annealing and lower-temperature gate modules as an integration challenge [10].
- **Local surface treatment.** SCREEN/LASSE material presents short laser treatments as a way to limit heat reaching buried structures [17]. Wavelength, pulse duration and stack design still determine the effect; a nanosecond pulse is not automatically harmless to everything below it.
- **Film quality with less thermal exposure.** KOKUSAI describes increasing demand for lower-temperature deposition [22].
- **Energy and emissions.** Tokyo Electron reports about **25% less CO_2 per wafer** for EVAROS compared with its stated conventional-tool comparison [21]. The result belongs to that comparison and its assumptions, rather than every furnace operation.

Where materials specialists contribute

Opportunities include cleaner fused-silica and SiC components, heaters and coatings that tolerate repeated cycling, precursor chemistry that works at lower temperatures, and dielectric/interface systems that achieve the required electrical properties within a restricted thermal budget. Glass, ceramics, metals, chemistry and metrology all contribute to what can initially look like a simple heating step.

10. What to remember—and what comes next

Topic	Main point
Silicon's advantage	A useful oxide and a high-quality interface helped enable reliable devices.
Thermal growth	Oxidant crosses the oxide; the reaction consumes Si and moves the interface into the wafer.
Growth rate	The Deal–Grove model moves from an approximately linear to an approximately parabolic regime.
Gate-oxide scaling	The aim is stronger electrostatic control in smaller transistors, not a thinner wafer or chip.
High-k and EOT	Higher permittivity allows greater physical thickness at comparable capacitance. EOT compares that capacitance with SiO_2; an interface layer adds to the stack EOT.
Annealing	Activation and repair must be balanced against dopant redistribution.
Short heating	RTP, flash and laser approaches control different temperature–time–depth histories.
Industry	Supplier rankings depend on whether the category is diffusion, deposition or broader treatment.
Future integration	Local heating, lower-temperature processes and robust components help protect stacked structures.

The next lesson, **3-3: Deposition by CVD and ALD**, changes how the film is made. Thermal oxidation converts part of the silicon substrate into oxide; deposition adds material supplied from outside. CVD and ALD provide ways to build many different films, including the batch-ALD processes discussed in the equipment section.

11. Check your understanding

1. Why was silicon dioxide important to silicon's success in integrated circuits?

   ▼ Show answer

   It provides insulation and protection, can mask selected processing steps, and can form a stable, low-defect interface with silicon. The useful combination is film quality plus interface quality, not simply the existence of an oxide.

2. A planar thermal oxide is 100 nm thick. Where is it relative to the original silicon surface?

   ▼ Show answer

   About 44 nm of silicon has been consumed below the original surface, while the oxide extends about 56 nm above it. The Si/SiO_2 interface moves into the wafer; the whole film does not grow only downwards.

3. What changes between the linear and parabolic regimes of the Deal–Grove model?

   ▼ Show answer

   In the thin-film model limit, interface reaction mainly controls growth and thickness is approximately proportional to time. In thicker oxide, diffusion through the existing film mainly controls growth and thickness scales approximately with the square root of time.

4. Why can wet oxidation be faster than dry oxidation?

   ▼ Show answer

   Much more water-derived oxidant can dissolve in the oxide, increasing the concentration available for transport to the reaction site. The cited solubility comparison is not a direct growth-rate multiplier; diffusivity, reaction kinetics and operating conditions also matter.

5. Why was gate oxide made thinner? Does this mean thinning the wafer or chip?

   ▼ Show answer

   The thin insulator inside each transistor was scaled to strengthen gate coupling and maintain channel control as gate length shrank. With other conditions held equal, higher capacitance supports more channel charge and drive current. The objective was smaller, higher-performing transistors, not a thinner wafer or chip.

6. What problem did high-k address in Intel's 2007 announcement?

   ▼ Show answer

   Very thin SiO_2 allowed excessive tunnelling leakage. Because $C = \varepsilon_0 KA/t$, raising K permits a thicker dielectric at comparable capacitance; EOT expresses the result as an equivalent SiO_2 thickness. Intel reported a more than tenfold gate-leakage reduction in its stated comparison, not zero leakage. The 7.7 nm HfO_2 example in this lesson is an ideal calculation, not Intel's stated film thickness.

7. At unchanged D, what happens to $L \approx 2\sqrt{(Dt)}$ when time is divided by 100?

   ▼ Show answer

   The characteristic length is divided by 10. If temperature and therefore D also change, time alone is insufficient to calculate the result.

8. Why not always use the shortest possible anneal?

   ▼ Show answer

   The process must activate dopants and repair defects as well as limit diffusion. A very short cycle may not achieve all the required material changes, and heating depth and uniformity also matter.

9. What does “vertical” mean in a vertical furnace, and what were some advantages of the arrangement?

   ▼ Show answer

   The tube and wafer stack are vertical; each wafer is horizontal. The historical account links the arrangement to improved handling, reduced boat–tube rubbing and improved uniformity with boat rotation.

12. Glossary

Thermal oxidation
Growing oxide by reacting silicon with an oxidant at elevated temperature.

Dry oxidation
Oxidation using dry oxygen.

Wet oxidation
Oxidation using water vapour.

Si/SiO_2 interface
The boundary between silicon and its oxide.

Interface state
An electronic state near the interface that can trap charge.

Passivation
Reducing unwanted chemical or electrical activity at a surface.

Planar process
Fabrication that retains oxide to protect and define device regions.

Deal–Grove model
A model relating thermal oxide growth to oxidant transport and interface reaction.

Linear regime
Approximate growth proportional to time within the reaction-limited model range.

Parabolic regime
Approximate growth proportional to the square root of time when diffusion dominates.

Diffusion coefficient
A parameter describing how rapidly species spread under specified conditions.

Solubility
The equilibrium amount of a species that can dissolve in a material under given conditions.

Tunnelling
Quantum-mechanical transmission through a barrier.

Gate dielectric
The thin insulating layer between the gate electrode and semiconductor channel; it is distinct from wafer or chip thickness.

Relative permittivity (K)
A material property describing its dielectric response relative to free space; in the simple model it determines capacitance together with area and thickness.

Equivalent oxide thickness (EOT)
The thickness of SiO_2 that would give the same dielectric capacitance per unit area as the layer or stack being compared.

High-k dielectric
An insulator with higher relative permittivity than SiO_2, allowing greater physical thickness for comparable capacitance. Interfacial layers also contribute to the stack EOT.

ISSG
In situ steam generation; reactive oxidising species are generated near a hot wafer.

Annealing
A controlled thermal treatment used to change material or interface properties.

Dopant activation
Making introduced dopant atoms electrically effective in the semiconductor.

Diffusion length
A characteristic spreading scale; $2\sqrt{(Dt)}$ is used here for a constant-D example.

Thermal budget
The cumulative effect of a temperature–time history on processes such as diffusion and reaction.

Transient enhanced diffusion
Temporary acceleration of dopant diffusion associated with excess defects after implantation.

RTP
Rapid thermal processing, commonly using radiation to heat a wafer quickly.

Flash-lamp annealing
Short, intense lamp-pulse heating, often after preheating.

Laser annealing
Thermal processing using absorbed laser energy, with heating controlled in space and time.

Vertical furnace
A vertical process tube holding a stack of horizontal wafers.

Batch ALD
Atomic layer deposition applied to multiple wafers in one batch.

Fused silica / quartz glass
High-purity amorphous SiO_2 used in optical and process components.

Sequential 3D integration
Building a further active-device layer above an existing one, with thermal constraints on the later processing.

13. References and supporting material

1. **Stanford EE212, Plummer / Deal / Griffin** — Chapter 6: Thermal Oxidation, 1999 handout, hosted by University of Washington EE528 — Read source
2. **F. J. Feigl** — VLSI Technology and Dielectric Film Science, Physics Today 39(10), 47, 1986 — Read source
3. **Computer History Museum** — 1960: Metal Oxide Semiconductor (MOS) Transistor Demonstrated — Read source
4. **M. M. Atalla, E. Tannenbaum and E. J. Scheibner** — Stabilization of Silicon Surfaces by Thermally Grown Oxides, Bell System Technical Journal 38(3), 749, 1959 — Read source
5. **Computer History Museum** — 1959: Invention of the Planar Manufacturing Process — Read source
6. **P. K. Chu, City University of Hong Kong** — AP6120 Chapter 4: Oxidation — Read source
7. **C. Henkel and M. Östling, KTH** — IH2655 Lecture 5: Oxidation, 2012 — Read source
8. **B. E. Deal and A. S. Grove** — General Relationship for the Thermal Oxidation of Silicon, Journal of Applied Physics 36, 3770, 1965 — Read source
9. **Intel** — Transistor technology breakthrough: high-k and metal gates, 27 January 2007 — Read source
10. **imec** — Advancing the CFET-based device roadmap, 7 September 2026 — Read source
11. **S. Kuppurao, J. Joo and G. Miner, Applied Materials** — In situ steam generation, Solid State Technology, July 2000 — Read source
12. **MIT OpenCourseWare 6.152J** — Lecture 4: Diffusion, 2005 — Read source
13. **Stanford EE212** — Chapter 7: Dopant Diffusion — Read source
14. **Stanford EE212** — Chapter 8: Ion Implantation — Read source
15. **P. Timans et al., Mattson Technology** — Flat-Top Flash Annealing, AVS JTG presentation, 18 July 2012 — Read source
16. **SCREEN Semiconductor Solutions** — LA-3100 flash-lamp annealing system — Read source
17. **SCREEN / LASSE** — Laser annealing presentation, CEA-Leti Workshop at SEMICON West, 2018 — Read source
18. **Semiconductor History Museum of Japan** — 1986: Vertical furnaces — Read source
19. **KOKUSAI ELECTRIC** — Products: vertical-furnace systems — Read source
20. **B. K. Tsai, NIST** — A Summary of Lightpipe Radiation Thermometry Research at NIST, Journal of Research of NIST 111, 9, 2006 — Read source
21. **Tokyo Electron** — Integrated Report 2026: history, market positions and EVAROS (Japanese) — Read source
22. **KOKUSAI ELECTRIC** — Integrated Report 2026: company, markets and results (Japanese) — Read source
23. **ASM International** — Vertical-furnace products — Read source
24. **Applied Materials** — Form 10-K for fiscal 2022: RTP and oxide/oxynitride processing — Read source
25. **NAURA** — 2025 interim report: thermal processing, pp.11 and 14 (Chinese) — Read source
26. **Mattson Technology** — Acquisition by E-Town Dragon completed: SEC exhibit, 11 May 2016 — Read source
27. **Applied Materials** — Agreement to acquire Kokusai Electric: SEC exhibit, 1 July 2019 — Read source
28. **Applied Materials** — Termination of the Kokusai Electric acquisition agreement: SEC exhibit, 29 March 2021 — Read source
29. **Shin-Etsu Quartz** — Semiconductor applications (Japanese) — Read source
30. **Shin-Etsu Quartz** — Company profile (Japanese) — Read source
31. **Tosoh Quartz** — Products and company background — Read source
32. **Ferrotec Holdings** — Results presentation, 14 August 2026, pp.2 and 7 (Japanese) — Read source
33. **Heraeus Covantics** — Fused silica for semiconductor applications — Read source
34. **CoorStek** — Diffusion and LPCVD processing — Read source
35. **Sibelco** — Expansion of Spruce Pine high-purity-quartz operations, 21 April 2023 — Read source
36. **imec / Soitec** — Sequential 3D planar device demonstration at low temperature, 9 July 2018 — Read source
37. **IRDS** — 2023 Edition: More Moore — Read source
38. **Ferrotec Holdings** — Corporate brochure 2025: quartz and SiC components (Japanese) — Read source
39. **J. Robertson, University of Cambridge** — High dielectric constant oxides, European Physical Journal Applied Physics 28, 265-291 (2004) — Read source

14. Claims and supporting sources

Claim or topic	Supporting material and limits	Type
SiO_2 properties, interface quality, steady-state fluxes and model limits	Stanford Chapter 6 [1]	Reported
Silicon oxide and the germanium-oxide comparison	Feigl, 1986 [2]; historical discussion	Reported
1959 insulated-gate development within the 1960 MOS demonstration account	Computer History Museum [3]	Reported
Thermal-oxide surface stabilisation in 1959	Atalla et al. [4]	Reported
Planar fabrication and protective oxide	Computer History Museum [5]	Reported
Oxidation reactions, interface motion, 0.44d Si consumption and 1,000°C growth parameters	City University teaching material [6]; example conditions	Reported
700–1,200°C range; wet/dry and solubility examples; rate constants; interface annealing	KTH lecture [7]; conditions retained	Reported
Linear–parabolic growth equation	Deal and Grove, 1965 [8]	Reported
1.2 nm SiO_2 in 65 nm technology; high-k and metal gates in 45 nm; leakage comparison	Intel, 2007 [9]; company-reported comparison	Reported
Historical scaling, gate capacitance C = ε0KA/t, EOT and example K values of 3.9 / 25	Robertson, 2004 [39]; ideal dielectric model distinguished from a complete gate stack	Reported
HfO_2 / SiO_2 interface stack and CFET thermal-budget integration challenge	imec, 2026 [10]	Reported
ISSG mechanism, approximately 1 nm films and charge-to-breakdown comparison	Applied Materials technical account, 2000 [11]	Reported
Characteristic diffusion scale 2√(Dt)	MIT lecture [12]; constant-D teaching estimate	Reported
Shallower junctions with device scaling	Stanford Chapter 7 [13]	Reported
Transient enhanced diffusion after implantation	Stanford Chapter 8 [14]	Reported
Sub-second surface-focused heating, pulse sources, preheating and activation/repair balance	Mattson technical presentation [15]; example conditions	Reported
Millisecond xenon flash-lamp system	SCREEN LA-3100 [16]	Reported
160 ns UV-laser example and local heating of stacked structures	SCREEN / LASSE presentation [17]; implementation-specific	Reported
Horizontal-furnace loading, vertical-furnace development and 1986 VERTEX history	Semiconductor History Museum of Japan [18]	Reported
A 200 mm furnace model accommodating 175 wafers	KOKUSAI products [19]	Reported
Optical thermometry and approximately 2.3°C calibration uncertainty	NIST, 2006 [20]; study-specific uncertainty	Reported
Diffusion-furnace position, Thermco history and approximately 25% CO_2 comparison	Tokyo Electron report [21]; own category and stated comparison	Reported
KOKUSAI history, 2023 listing, 49.2% batch-deposition share, over 80% batch-ALD estimate, revenue and regional mix	KOKUSAI report [22]; Gartner-derived and company figures distinguished	Reported
US$2.8bn treatment category and KOKUSAI share of 11.8%	KOKUSAI report [22]; distinct category	Reported
SONORA / A400 and RTP / oxidation product scopes	ASM [23]; Applied Materials filing [24]	Reported
NAURA thermal sales, 1,000 cumulative furnaces and its market estimate	NAURA interim report [25]; periods retained	Reported
Mattson acquisition in 2016	Mattson SEC exhibit [26]	Reported
US$2.2bn acquisition agreement; termination after regulatory deadline; US$154m fee	Applied Materials SEC exhibits [27][28]; historical events	Reported

Claim or topic	Supporting material and limits	Type
Quartz applications, shareholders and fabricated products	Shin-Etsu Quartz [29][30]	Reported
Tosoh products and company history	Tosoh Quartz product and company information [31]	Reported
Ferrotec quartz sales and AI-related demand; SiC components	Ferrotec presentation [32] and brochure [38]	Reported
Semiconductor fused silica and quartz/SiC material trade-offs	Heraeus Covantics [33]; CoorStek [34]	Reported
US$200m Spruce Pine expansion announced in 2023	Sibelco [35]; announced plan, not a completion claim	Reported
Upper-layer processing below 525°C in the 2018 sequential-3D example	imec / Soitec [36]; not a universal temperature limit	Reported
Lower-temperature processes after DRAM peripheral circuitry	IRDS 2023 [37]	Reported
44 / 56 nm geometry, approximately 1.9 h and 7 min growth times, time ratio, diffusion-length ratio and ideal 1.2 nm SiO_2 / 7.7 nm HfO_2 capacitance comparison	Arithmetic in the text with stated parameters and assumptions	Calculation
Intuitive explanations, layer roles, thermal-budget interpretation and illustrative drawings	General teaching explanation; schematics are not calibrated process simulations	Explanation

Updated 8 November 2026. Technical explanations draw on papers, university lectures, roadmaps and equipment literature. Company figures retain their stated periods and market definitions; rankings are attributed to the companies or surveys reporting them. Calculations and interpretation are identified separately. All seven diagrams are teaching schematics, not process simulations. Photographs retain their individual credits and licences.

© 2026 Troy Technical. Original text and illustrations: all rights reserved, subject to applicable legal exceptions. The credited photographs retain their stated licences. Links to this article are welcome.
Teaching diagrams localised with the built-in Image Gen tool for this English edition.