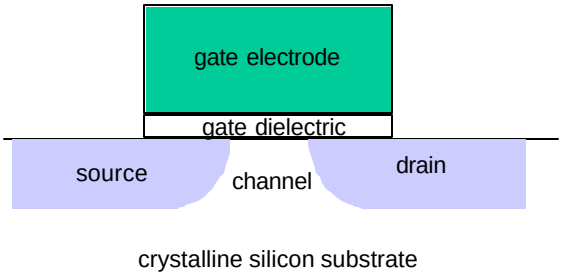
Glossary for high-k/metal gate



Basic CMOS transistor

transistor – A simple on/off switch. Current flow from the source to the drain is determined by whether the gate is at high or low voltage, much as a light switch's state (up or down) controls whether or not current flows to a light bulb. Digital chips such as microprocessors consist of millions of such transistors connected together by copper wires in a specific pattern. As technology advances, the objective is to make these transistors ever smaller, faster, cheaper, and less power-hungry, all of which leads to more powerful chips.

source – The part of the transistor where current flows from. It consists of doped silicon, that is silicon with some impurities, which lower its resistance.

drain – The part of the transistor where current flows to. It is doped with impurities in the same way as the source. A transistor is completely symmetrical, i.e. current can flow from source to drain, or vice versa.

gate – (also known as gate electrode) A region at the top of the transistor whose electrical state determines whether the transistor is on or off. Traditionally, the gate is made of polycrystalline silicon ("polysilicon"), that is, silicon whose atoms are randomly placed, and not in a grid-like structure.

channel – The region between the source and drain, where current flows when the transistor is in the 'on' state. It consists of silicon in the crystalline state, that is, silicon in an orderly grid-like (lattice) structure.

gate dielectric – A thin layer underneath the gate that isolates the gate from the channel. In today's chips, it consists of silicon dioxide.

silicon dioxide – Molecules consisting of one silicon and two oxygen atoms, which form a good insulator (non-conductor of electricity). For a gate dielectric, a thin silicon dioxide layer is desirable for high performance. The problem is that the thinner the layer, the higher the leakage through it; hence the effort to replace it with new materials that preserve its properties, but don't need to be so thin.

high-k material – A material that can replace silicon dioxide as a gate dielectric. It has good insulating properties and also creates high capacitance (hence the term "high-k") between the gate and the channel. Both of these are desirable properties for high performance transistors. "k" (actually the Greek letter kappa) is an engineering term for the ability of a material to hold electric charge. Think of a sponge. It can hold a lot of water. Wood can hold some but not as much. Glass can't hold any at all. Similarly, some materials can store charge better than others, hence have a higher "k" value. Also, because high-k materials can be thicker than silicon dioxide, while retaining the same desirable properties, they greatly reduce leakage.

leakage – Current flowing through the gate dielectric. In an ideal situation, the gate dielectric acts as a perfect insulator. But as it is made ever thinner (in Intel's 90nm process, it is a mere 5 atomic layers thick!), current leaks through it. This results in undesirable results. The transistor doesn't behave as it should, and it consumes more power than it should. Think of a leaky faucet that drips water, hence being very wasteful.

NMOS transistor – (also known as n-type transistor) A transistor which is on when its gate is at high voltage, and off when its gate is at low voltage

PMOS transistor – (also known as p-type transistor) A transistor that is opposite to an NMOS transistor, i.e. it is off when its gate is at high voltage and on when its gate is at low voltage (think of a faucet for which water flows when the handle is turned clockwise).

CMOS – Complementary Metal Oxide Semiconductor. A process technology in which both NMOS and PMOS transistors exist. All modern logic chips such as microprocessors and chipsets use CMOS due to its ability to deliver a combination of high performance and low power, all at low cost.

Threshold voltage – The voltage level between high and low which distinguishes whether a transistor is on or off. For an NMOS transistor, if its gate is above the threshold voltage, it is "on"; below the threshold voltage, it is "off". A PMOS transistor exhibits complementary behavior. Transistors are designed to have a low threshold voltage, as this leads to high performance (think of a racing car with low center of gravity).

Threshold voltage pinning – (also known as Fermi level pinning) One of two undesirable effects when a high-k gate dielectric is combined with a polysilicon gate electrode. Due to some defects that arise at the gate dielectric/gate electrode boundary, it becomes difficult to adjust the threshold voltage to a low value, which is needed for high performance. The problem goes away when the gate electrode is a specific metal, rather than polysilicon. The choice of metal is different for NMOS and for PMOS transistors.

Phonon scattering – The second undesirable effect when a high-k gate dielectric is combined with a polysilicon gate. This phenomenon limits electron mobility and hence degrades performance. The problem goes away when a gate made from a specific metal replaces the polysilicon gate, and the right process recipe is applied.

Related terms

Moore's Law – A prediction (not really a law) made by Intel co-founder Gordon Moore that the number of transistors on a chip double every two years. Intel's microprocessors have followed this 'law' very closely, beginning with the 4004 in 1971, with just over 2000 transistors, and leading up to today's Itanium® 2 processor which has 410 million transistors. In general, transistor density is roughly doubled with each new process generation, which occurs every two years.

Strained silicon – A technique for speeding up transistors. As described above, the silicon atoms in the channel are packed neatly in a grid-like (lattice) structure. It has been known for decades that stretching the grid so the silicon atoms are slightly farther apart than in their natural state makes NMOS transistors switch faster (similarly, compressing the lattice slightly speeds up PMOS transistors). This stretching/compressing is known as straining. Intel uses special techniques to strain its 90nm process NMOS and PMOS transistors to improve their performance.

Tri-gate transistor – A new type of transistor that Intel has designed for possible deployment in the 45nm process in 2007. The transistors described earlier in this document are planar transistors. That is, they have a single flat gate that is parallel to the surface of the silicon substrate. A tri-gate transistor employs a novel three-dimensional structure where the gate wraps around three sides of the silicon channel. A traditional planar transistor could be likened to a highway on top of a mesa (a flat-top mountain with vertical sides) with the electronic signals traveling like cars across the flat surface of the mountaintop. With the new, elevated 3D design, the signals travel not only across the flat top, but along both vertical sidewalls as well; hence, the 'tri-gate' name. Both high-k/metal gate transistors and tri-gate transistors are candidates for Intel's 45nm process. We don't know yet which will be chosen. A hybrid is also possible.

TeraHertz transistor – A class of transistors introduced by Intel in 2001, which operate at over a TeraHertz (that is, they are capable of switching over a trillion times per second) and are scalable through the end of this decade. The TeraHertz transistor is designed to maximize performance while minimizing power dissipation and manufacturing cost. Both high-k/metal gate and tri-gate designs fall within the class of TeraHertz transistors.