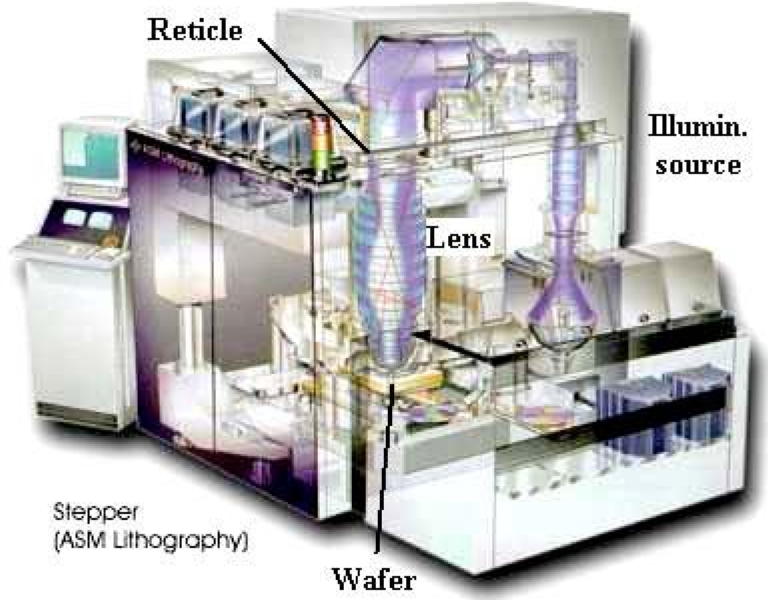
Monday, February 16, 2009

Lithography for dummies (well, us EDA guys)

EDN



You probably already know that designs are transferred onto chips using a photographic process. The wafer is coated in a solution called photoresist and then exposed to light passed through a "mask" which alters its chemical composition. The exposed (or sometimes unexposed, depending on type) photoresist is then removed with a powerful acid and some semiconductor process takes place through the gaps created: diffusion of impurities, implantation of ions, etching of metal and so forth.

Originally a mask was the size of an entire wafer and all the die (technically the plural of die is dice but it looks so Las Vegas I'll stick with die) were exposed to light at the same time through a mask the same size as the wafer (which was 2", 3" or 4" back then; now they are 12" with 18" in planning). For about the last twenty years, though, each die has been exposed individually through a reticle, a smaller mask that is stepped across the chip one die at a time by an expensive piece of equipment called a stepper. Usually the reticle is a multiple of the actual die size and the stepper has reduction optics rather like a photographic enlarger in reverse.

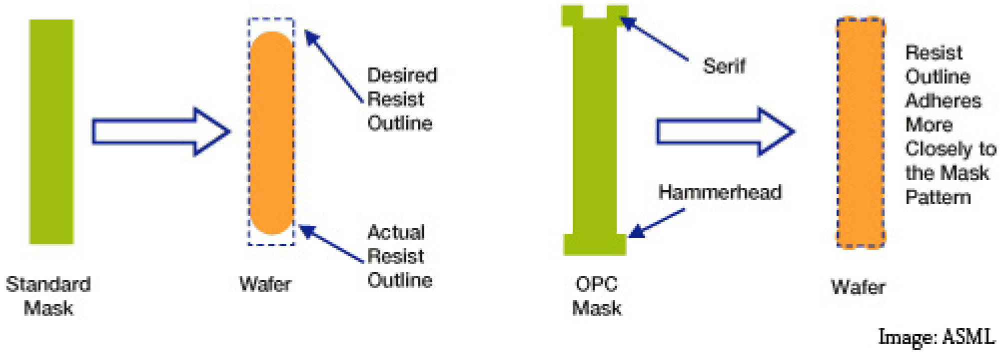
Originally we used 436nm mercury lamps which was a much shorter wavelength than the 1um or so feature size we were trying to achieve on the die so we didn't have to worry about all those strange things in optics that you may remember from high-school or college physics: Young's slits, diffraction gratings, wave interference. What was drawn on the layout designer's screen, what was put on the mask and what ended up on the silicon were pretty much the same thing.

As feature sizes got smaller, we reduced the wavelength of light, first to 248nm and then to 193nm. We are still at 193nm for two reasons. We had developed technology for DUV (deep-ultra-violet) at 157nm but it was really expensive and unattractive. We also discovered immersion lithography where the gap between the lens and the wafer is filled with water not air, which improves things enough that we can continue to use 193nm for the time being.

The basic problem is that as the wavelength gets shorter and shorter, we are moving out of the part of the electromagnetic spectrum where we can focus light with lenses, and into the part where we essentially have X-rays that go straight through the lens and through pretty much anything else too. The next step looks like it will have to be e-beam lithography, where a beam of electrons is steered in the same way as in an old TV. This is well-understood technically but it has a very slow write speed which, so far, makes the whole process uneconomical for mass production.

But being stuck at 193nm means we have a new problem. We have feature sizes on chips that are much less than 193nm (which is around 0.18um which was many process nodes ago). All sorts of optical effects happen due to wave interference of light and we needed to put very different patterns on the mask from the original layout, in order to get the eventual feature on the die to match what we first thought of. It became anything but WYSIWYG.

There is a whole gamut of techniques that have come to be known as RET, for resolution enhancement technologies. Optical proximity correction (OPC) changes the shape of what is on the mask so that what ends up on the wafer is what is required. For example, corners have extra lumps added so that they don't get etched away. Phase shift masking (PSM) etches the reticle by fractions of a wavelength so that the interference that results is desirable. The generic name for putting these extra features onto the mask is known as RET decoration. Since this might multiply the billion or so shapes on a layer by a factor of ten it is computationally very expensive.



A whole subsegment of EDA grew up when this first became important, under the generic name of **DFM**, **design for manufacturability**. Many companies were started in the segment and it is instructive to look at this since it is the most recent example of an area of technology where the basic cycle from foundation to exit is pretty much complete. That can wait until another day.