Sunday, April 8. 2007

Is the race to 32nm a race to a dead end?

Whenever you talk about advancements in processor technology it boils down to process technology. The race between AMD and Intel gave us jumps from 130nm to 90 nm to 65 nm to 45 nm and 32nm looms at the the horizon. At most you can contribute this race to Intel, as for a long time they needed one shrink more to be competitive with the respektice products from AMD. The things changed a little bit since the invention of the Core architecture. But this isn´t the point i want to make.
When you look at all this process technology step you should look at the other part of the equation. With every shrink, the fabs to produce this chips got more and more expensive, as you need the newest and most developed equipment to produce this processors. The price for such fabs almost reach the edge to be unamortisable within the lifetime of a process, at least when you use it to produce the CPU for your daugthers and sons gaming PC (Besides of HPC the only market with an ever elastic demand for computing power).

More and more companies starts to step back to think about going the way to 32 nm. TI, our venerable manufacturer of the SPARC chips. Obviously the media makes a "TI is toast, and all of it´s partners, too ." of of it (besides some exceptions, like).

Well, i have a little more differentiated view to it. The nanometer race is nothing more than the reverberation of the megahertz race some years ago. It´s an easy way to get more performance, but will hit a wall sooner or later. But the wall at the end of the nanometer dead end is a little bit more complicated. The wall is made by the financial department. Even when the technology is able to go farther ... you can´t go the path, when the fab doesn´t pay of withing 4 to 5 years or so.

So, here we sit in front of the next wall. Crash imminent. What to do? At least for Sun i have an idea. I´m not aware of any developments this far in the future, but many things lie open to the public. Do you remember the annoucement of proximity computing? Do you remember the whitepapers about the internal structure of our processors with their internal crossbars. Now mash them up. Think about a Niagara V that couples to other Niagara V in a chip package via proximity computing. Imagine a Niagara V Cube augumented with dices full of cache in a single chip package. Something Intel wanted to sell as an innovation with their 80core processor ... but what in principle is an old hat at Sun. You don´t have to invent more expensive processes (I think, with this development, 12 nm will costs more than the GDP of a middel sized economy) ... you have only to invent to glue dices together in a cheap and efficient way.

I don´t think that the future of processors lies in smaller structure sizes ... i think it will lie in vastly faster interconnects to bridge between seperate dices manufactured with cost effective processes. It´s a more intelligent way to build faster computers and leads us to the next crisis in computing: The fall and rise of software development techniques in really massive parallel environments and their introduction to enterprise computing. But that´s a different story.

Posted by Joerg Moellenkamp in English at 09:28

IIRC 12nm would be beyond the physical limits of the current materials so that's pretty much out without inventing a whole new set of materials and figuring out what and how expensive that will be is probably a bit too early at this point.
The obvious wins at going to lower nms is using less power and being able to fit more processors on a silicon wafer.
In the end it seems pretty simple economics (with the possible exception of doing good for the environment).
  Anonymous on Apr 8 2007, 14:28

IMHO, a complete new set of materials would be incredibly expensive as all the R&D that has to pay off in a few fiscal quarters.
  Anonymous on Apr 8 2007, 15:05

The obvious win of smaller semiconductor features is switching speeds, this being the result, primarily, of less charge accumulating between the legs of a transistor, due to their smaller size and the shorter distances that charges need to move in order to effect a switch. Smaller features, with current silicon technology, leads to more current leaks, so higher power consumtion, not lower.
  Anonymous on Apr 9 2007, 00:19