Xilinx responds to Altera's FPGA benchmarks


It can't be Friday already! I spent yesterday (Thursday) celebrating the 30th anniversary of my 21st birthday. I'm not finished celebrating yet; I don't want to get back to work... Oh well... perhaps not surprisingly, I've been receiving feedback with regard to Altera's recent "How To" on performing FPGA benchmarks (article #208400643). One reader wrote:

Hi Max, there was a consortium known as PREP that existed about 15 years back. It included all PLD makers at the time (a boatload!) and consisted of a set of agreed upon benchmarks and methodology. PREP was led by Stan Baker of EE Times in those days.

I think our friends at Altera are treating the OpenCores designs as if they are something without "bias", which I don't agree with. Anyway, new benchmarks simply become fertile ground for "new spin", as far as I'm concerned!

This reader also pointed me at a proposal for benchmarks relating to High Performance Computing, which goes beyond PREP and beyond FPGA benchmarks per se, but is nonetheless very interesting:


Another response was a tad more strident:

Hey Max, regarding Altera's last "unbiased" analysis... Don't believe everything you read. It is my understanding that Altera used a "not-yet-released" version of Quartus and may have (accidentally?) used some less than optimal settings on the Xilinx tools (as there are about a billion settings).

From what I hear, when these "settings" in the Xilinx tools (10.1SP1 which is currently available) were corrected and run against the currently available Quartus tools, the results favored Xilinx by more than 2:1 on the same set of OpenCores designs.

I've been using programmable logic in one form or another since the early 80's. My first experience with Xilinx was really, really bad (1987) and when the Altera FAE walked in and showed me schematic capture, I was blown away and used Altera for a decade. I have horror stories about how Altera manipulates and deceives their customers. I then became an FAE for one of the Xilinx distributors and got to see what was really happening. I won't say that Xilinx isn't without its issues, however, as a designer and as someone who had to clean up "messes" that the sales people left, I am FAR happier.
Regarding your recent benchmark article, the article weighs heavily for Altera's FPGAs vs. Xilinx's, but then it was written by Altera.

Unfortunately, in marketing products, as well as in politics, attacks must be responded to - if there is any bit of truth in them whatsoever.

What does Xilinx have to say in defense?

Well, that's a very good question. In fact, as soon as I'd posted the article, as a matter of courtesy I emailed Bruce Fienberg (the Communications Manager at Xilinx) to let him know it was there and invite his response. Bruce replied as follows:

Hi Max, we're not going to provide a competing article because we don't want to get into a tit-for-tat thing with Altera over these benchmarks and we don't think that would serve the best interests of your readers. These benchmarks we're run by Altera and even though they used OpenCore designs, the results are still biased in their favor no matter how you cut it.

I'm not sure how critical you want to be about all this, but I'll share the following with you anyway, which is based on our communications to the field:

Note that this is not the first time Altera has published benchmarks based on OpenCore designs. A couple of years back they did a DSP performance comparison between Stratix II and Virtex-4 based on 6 OpenCores designs. Interestingly enough, none of the 6 designs they picked back then are represented in the current list of 7.

It turns out that we've been aware of Altera's benchmarking strategy for some time and so we've been performing the same tests, with the same designs, ourselves. Based on our findings, here are four reasons why Altera's benchmarks misrepresent ISE and Virtex-5 capabilities:

1. The most obvious flaw in Altera's benchmarks is that they are not using comparable settings in the backend tools. They use the weakest effort for ISE
which is "standard" (-ol std). For Quartus, they use "Auto Fit" but that effort is not the weakest effort, it's "Fast Fit".

2. They use ISE 9.2i when our current version is 10.1_sp1. All Xilinx internal benchmarks have shown that 10.1 is a great improvement over 9.2i.

3. Some of the designs picked by Altera (3 out of 7 or 43%) are register-limited for Virtex-5. The LX330 offers 207,360 registers when the EP3S340 offers 270,400 registers. That's 30.4% more! Obviously these kind of register abundant designs are more likely to fit better in Stratix III. The question now is how often are designs register-limited when targeting Virtex-5 devices? The selection from Altera would lead that it happens often (3 out of 7). We checked our customer database of submitted designs and found that only 34% of designs are register limited in Virtex-5. Therefore, it is safe to say that Altera biased their sample in their favor by using more register-limited designs.

4. Altera uses a bottom-up approach, they compile one instance and then instantiate it several times as a black-box with a wrapper. The problem with this approach is that Synplify Pro cannot do "resource management". "Resource management" is the capability of the synthesis tool to allocate resource depending on what's available on the device. This affects the design called oc_aquarius which requires several BRAMs. After a few instances of oc_aquarius, all the BRAMs are used up and the design does not fit, but when the block is synthesized flat (outside the black-box flow), Synplify Pro allocates distributed RAM in lieu of BRAM and this allows the design to fit.

We also looked into the designs to study the coding style. The design oc_aquarius is using more BlockRAMs for Virtex-5 than needed due to an inefficiently coded 4kx32 RAM in HDL. The RAM has been truncated in 512x8 segments and logic has been added to manage write enable and output muxing. Recoding the HDL should allow for entire 4kx32 memory to be mapped onto just one block, thus enabling Virtex-5 to fit more "stamps" for this design. The critical path in oc_or1k is a 32x32 signed multiplier with input and output register. Since these registers are coded with asynchronous registers, the
tools can't MAP them onto the DSP48E. Using synchronous reset will allow for better performance and less fabric registers being used.

We took the designs and recreated the test cases and scripts. We aligned the effort level in the backend tools. We used "high" effort in ISE because it's the most often used effort by our customers and it also make sense to use a high effort since we want software to maintain performance as the device fills up. In Quartus we used the effort above what Altera used in their tests with "Standard Fit3". Since that particular effort (even though branded as "high") does not set the Quartus Placer and Router to their high effort, we did set the Placer and Router multiplier4 efforts "3" (it can take any value from "1" to "4").

We used the slowest speed grade for our tests because it's used more often than the fastest speed grade.

We ran Synplify Pro 9.2 (latest publicly available release) iteratively until negative slack was reported. Constraining the tool until negative slack is measured is a good indication that all synthesis optimizations have been applied and that the logic has been optimized for performance as much as it can be.

Then, we iteratively increased the constraint for the backend tools also using all the latest publicly available tools. For ISE we used 10.1_sp1 and for Quartus II we used 7.2_sp3. Note that we don't have access to Quartus II 8.0 which is what Altera used for their tests.

Overall, the Xilinx software solution with Synplify Pro and ISE 10.1 dominate. Despite the bias in the samples chosen by Altera, all software metrics (performance, utilization, ability to maintain performance) show great with ISE and Synplify compared to Quartus.

Our results show that:

- ISE maintains the performance better than Quartus when the devices are filled up, delivering up to 60% faster performance.

- The runtime for the Xilinx flow (Synplify Pro + ISE) is 16% faster on average than the Altera flow
The number of instances that can be stamped is 3% better for the Virtex-5 device even though Altera carefully biased their sample to get more register-rich designs.

Also, you must keep in mind the Stratix III availability issue and the fact that, in today's design world, it's no longer just about the silicon, but also the tools, resources and other solutions that go around it.

Now my head hurts (although maybe this is due to too much celebrating yesterday). Wouldn't it be nice if there was some independent group who could create a suite of FPGA benchmarks that were scalable across devices (small to large) and evaluated everything (capacity, utilization, performance, power consumption, ...). Actually, if someone were prepared to fund the effort, I'd be more than happy to head-up such a project myself (I know people who know people who know how to do this sort of thing). Maybe one day...

Questions? Comments? Feel free to email me – Clive "Max" Maxfield – at max@techbites.com. And, of course, if you haven't already done so, don't forget to Sign Up for our weekly Programmable Logic DesignLine Newsletter.