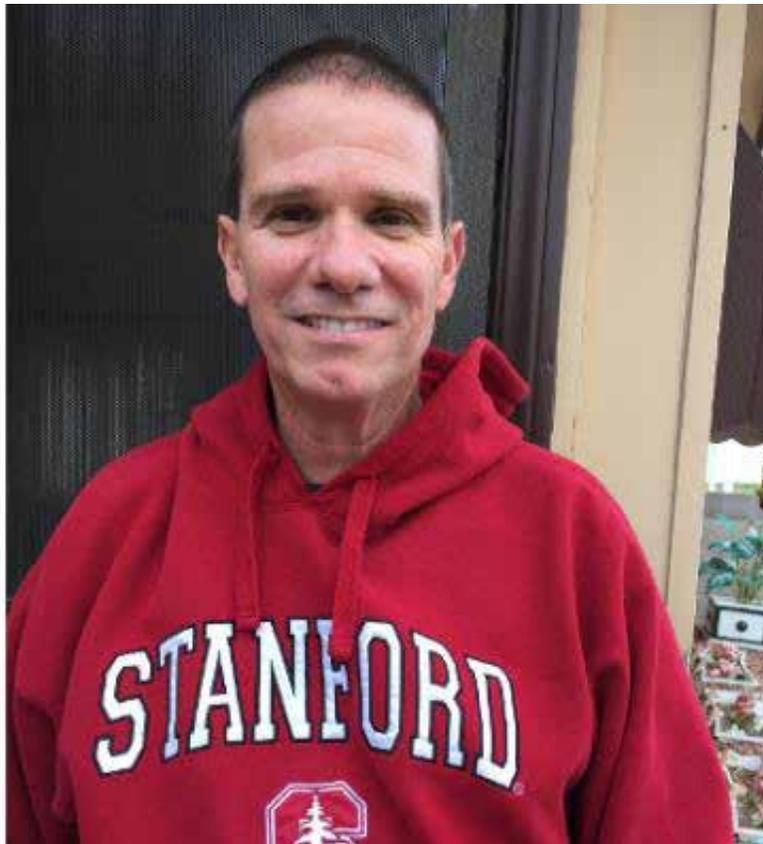


CEO interview: Silicon Valley is still the place where the magic happens

December 15, 2020 //By Peter Clarke



We interview Mark Santoro, CEO of Micro Magic Inc. (Sunnyvale, Calif.), a company that has produced a a 64bit RISC-V processor that can clock at 5GHz but also operate at near-threshold voltage.

Santoro pointed out to us that while the high clock frequency is exceptional (see EDA company claims world's fastest 64bit RISC-V core) the more significant achievement may be the RISC-V processor's low-power performance. Santoro told

us dropping the operating voltage to near the threshold voltage at 350mV decreases performance by a factor of five but increases the power efficiency of computation by a factor of more than 9.

And getting such power consumption efficiency is likely to be the big win that will get licensees to sign up to use this core – if that is the direction Santoro chooses to take the company.

Santoro provided some background to his company to put the performance – up to 13,333 CoreMarks and power consumption as low as 10mW – into context.

He pointed out that the company has a history that stretches back to Sun Microsystems and before. The company is a relatively small group of engineers – less than 50 according to LinkedIn – who specialize in datapath design and optimization and high-speed memory design. The company was originally founded in January 1995 but was acquired for \$260 million in December 2000 by Juniper Networks Inc. The founders of Micro Magic restarted the company in 2004.

Although the company's founders were interested in leading-edge processor implementation they had to develop their own EDA tool suite to do that. Since its reformation in 2004, Micro Magic has been a vendor of those EDA tools, and a design services firm, using its tools to improve customers' ASICs.

One of the benefits of those EDA tools is the ability to place and route circuits in response to timing requirements, thereby making design for performance and timing closure easier, Santoro said.

CoreMarks per watt

“CoreMarks seemed a reasonable benchmark but we designed for power efficiency and not just performance,” Santoro said. CoreMarks is a performance benchmark suite produced by EEMBC, the Embedded Microprocessor Benchmark Consortium, a non-profit, member-funded organization.

“But CoreMarks per MHz is not significant, as CoreMarks do not directly scale with clock frequency. The more important measure is CoreMarks per watt,” said Santoro. He then provides the key benchmark figure with power consumption attached. This is the power consumption for the RISC-V core and first-level caches.

As previously reported, at 1.1V the processor clocked at 5.14GHz and achieved 13,333 CoreMarks while consuming about 500mW. That equates to nearly 27k CoreMarks/watt.

We also previously reported that at 0.8V the processor runs at 4.3GHz and achieves 11,111 CoreMarks while consuming 200mW. That is 55.5k CoreMarks/watt.

At 0.6V the processor runs at 3.1GHz achieving 8,461 CoreMarks and consumes about 70mW. That's roughly 121k CoreMarks/watt.

But such was the teams' attention to design-for-power, that the core – and this is a real chip, not a design – continues to operate down to 350mV and achieves a clock frequency of 1GHz and 2500 CoreMarks while consuming 10mW. That's 250k CoreMarks/watt.

Dropping the voltage to one third improves the power consumption efficiency by more than a factor of 9 roughly in line with the V^2R power rule. The processor was benchmarked without 'binning' meaning the same chip would be able to achieve both the highest clock frequency performance and the highest computational power efficiency.

Unfortunately, it is not always easy to compare processors because of different size caches and how much peripheral logic is included.

And Micro Magic still isn't revealing which foundry produced the silicon. All we are being told is that it is a FinFET manufacturing process and that the company has compared three foundries PDKs shooting for the broadest compatibility. That indicates that the selected process is 28nm or below and for reasons of cost and comparability is not down below 10nm. The list of suspects is fairly narrow and includes Globalfoundries, Samsung and TSMC, with Intel and SMIC as outside possibilities.

No silver bullet

But when we asked Santoro how his team had achieved 2,500 CoreMarks at 10mW he said: "There's no silver bullet. There's not one thing for low power. It is the combination of a lot of things we have to pay attention to." Santoro pointed out that back at Sun Labs he and fellow engineers had achieved the highest performing SRAM with a 900 picosecond access time.

"Also when you design for high performance you HAVE to pay attention to power. If you don't design for low power you end up melting wires," he added. Santoro re-iterates that when he originally founded Micro Magic the standard industry tools were not suitable for this style of design. "We had to write tools specific for building and analysing memories. We had to create tools for timing-aware placement and routing."

He continued: "You may remember SiByte caused a stir when it announced a MIPS processor capable of operating at 2GHz. That was placed using Micro Magic software."

SiByte was startup founded by Dan Dobberpuhl. The company disclosed its networking processor in 2000 and was acquired by Broadcom in November 2000 in a deal worth more than \$2 billion in stock.

One of the traditional ways to achieve high clock speed is by creating a fine-grained pipeline in the ALU so that smaller chunks of processing are done, with each stage requiring less logic, and therefore able to be performed at a higher clock frequency. This also means that many instructions are being acted on in the pipeline in parallel. However, this benefit comes with costs.

Santoro pointed out that the deeper the pipeline the more processing has to be discarded when there is a stall in the pipeline. This occasional but substantial wasted effort acts against power efficiency. Such stalls can come for a variety of reasons including interrupts and failing to find an instruction or data entry in the cache, requiring the loading of a new page.

In addition, a deeper pipeline requires more control logic, which also hurts power efficiency and there is a temptation to increase sophistication with out-of-order and speculative execution and it all ends up requiring more flip-flops and more power.

The KISS principle

Santoro does not admit to how many pipeline stages the Micro Magic design has and in fact such a question often depends on how a processing stage is defined. "It is small and simple. Even within the design team people argue over how many pipeline stages it has. Our CoreMarks may be high but we are not designed for CoreMarks." He continued: "The architecture does matter but the tools matter a lot."

Santoro explained that by designing the datapath for high-speed access Micro Magic can get a pipeline stage done quickly while keeping the ALU simple and thereby promoting lower power.

Nor will Santoro reveal the size of the L1 caches designed into the working silicon. "You always want bigger but we are somewhere in the conventional size. 16kbyte or 32kbyte would be difficult to operate at 5GHz. This suggests 4kbyte or 8kbyte L1 cache sizes.

Although the working silicon produced by Micro Magic is a single core, Santoro acknowledged that licensees of the IP might well want to deploy the core in a quad- or octocore configuration. He said the Micro Magic design could easily go to a multicore implementation and there are some hooks present that would help with such a design.

When asked if Micro Magic would consider an implementation on a fully-depleted silicon-on-insulator process Santoro said: “FDSOI is extremely interesting. There is nothing in our design that precludes the use of FDSOI.”

However, Santoro pointed out that the company had chosen a FinFET manufacturing process to try and maximize ease of availability for the industry. FinFET also offers the obvious route down to leading-edge manufacturing processes – now at 7nm and 5nm. FDSOI may not make it down to such geometries.

“There are some differences between FinFET processes and FDSOI but we designed our tools to be adaptable so it is fairly easy to retarget a design to a different process.” He said. Santoro did throw in the caveat. “If you want to make use of the back-bias capabilities of FDSOI it does get a little more involved.”

Santoro said it is notable that applications for leading-edge FinFET processes at below 10nm are tending to focus on two poles of capability addressing to application sectors: smartphones and high-performance computing. One is limited by power consumption while still needing significant performance. The other requires high performance but must still seek power efficiency. Santoro leaves his chip’s performance at 5GHz and 1GHz to speak for itself.

In parting Santoro points out that in terms of tool sets Micro Magic has had three-dimensional lay-out tools for more than 15 years. However, because the mainstream EDA vendors did not support 3D design it failed to come in as quickly as Santoro had thought it would. “If you are too far ahead of your time you don’t get taken up,” he said wistfully.

Now, with chiplet-style packaging starting to come into mainstream manufacturing Micro Magic is well placed to enable such designs (see Google, AMD tipped as early adopters of TSMC chiplet manufacturing).

Business choices

It would also appear that Micro Magic is in a good place to migrate from design services – which it has successfully performed for many years but which do not scale well as a business – into a more scalable, product-oriented IP licensing business. Alternatively, it can continue with tool licensing as an EDA company, also a scalable business model.

The risk with being an EDA tool vendor is that the company can get sucked into design services. Licensing out cores as products also has challenges. It takes will power to stick to the product plan and turn down requests for additional bells and whistles

on cores. The IP vendor that succumbs to that also that also gets sucked into what is effectively custom design.

“We built the RISC-V design to show what we can do, and what our tools can do,” said Santoro. “We’re getting a lot of interest now. The plan is to license out cores, although things have become more complicated. We don’t want to license out cores if we are about to be acquired.”

Given the premium value attained by Micro Magic in its first sale and by SiByte back in 2000, it can be seen why Santoro is considering the company’s options.