SWCP Portal
February 2021

Apple Aims at Datacenters

By Mark Costlow, SWCP President

As mentioned in our January issue, Apple has recently introduced a new range of laptops using their own CPU chips called **Apple Silicon**. The **CPU** (Central Processing Unit) which runs the Operating System is the essential heart of any computer, and what Apple is doing is basically putting it all on a single integrated chip called, logically enough, a **System on a Chip**, or **SoC**. Their first is named the "M1".

Since this sort of replacement affects everything, hardware and software must be updated to match. This is difficult enough to do for a personal computer, laptop, or smartphone, but what about the internet datacenters upon which they all heavily rely? Here is another market, one where Apple Silicon chips could potentially make an even bigger splash.

ARM vs. x86

Every computer needs a set of instructions to run its microprocessors. There are two basic architectures dominant in the market today: "**ARM**", found in most smartphones and tablets, and an older one called "**x86**", still used in the majority of laptops and PCs. While x86 is good for intensive computing, ARM uses less juice.

This is very important because energy costs are a major factor in datacenter operation. The most significant advantage of ARM-driven processors are their increased efficiency, which means both reduced electric consumption and less heat production.

For every dollar you spend to power a computer in a datacenter, you spend at least another dollar to remove the heat that it generates. Reducing power consumption can be quite a big deal because every microwatt of electricity fed into a CPU to do a calculation is converted into heat. Datacenters must then use more energy for air conditioning to dump that waste heat somewhere else.

Some firms have experimented with putting datacenters in frigid environments to lower operating costs. Microsoft has one on the floor of the North Sea, chilled by the deep's constant low temperature.

The new Macbooks don't need internal fans to stay cool, and even massive M1 server farms should show similar savings. Since anything that reduces the amount of waste heat put into the environment saves energy, money, and carbon, savings in this era of climate change are more than just economic.

Only a few massive companies buy the largest share of the world's server computers. Amazon, Google, and Microsoft are the big three cloud providers, and Intel, which invented the x86 architecture, provides almost all of the CPUs that they use. But at least two of these international corporations are already working on ARM-based SoCs, not so different from what Apple has done. In fact, Amazon already has them in production. You can buy AWS services and specify you want "graviton2" CPUs instead of Intel's x86s.

Microsoft isn't quite as far along, but they might arguably have a bigger impact. They certainly own the best position of any company to make sure the ARM-based servers will run Microsoft Windows well. And of course they can tweak Windows to make it run better in that environment as they see fit.

Apple doesn't have as many datacenters as the others, but they do have a large and growing cloud presence and are actively scaling up their datacenter holdings. One presumes they will leverage the M1 and its successors to run those datacenters at lower cost. If they can get the same computing power with only *20%* of the electricity consumption, that will make a huge difference in datacenter economics.

Amazon and Microsoft might never sell these servers on the open market, but it seems clear that this trend could push Intel out of the datacenter, a market they have universally dominated for decades.

With Apple leading the way in putting non-Intel CPUs on people's desks and laps, Windows users are likely to start demanding the same cool-to-the-touch long-running machines from Dell, Lenovo, and HP. It's not a stretch to think they could license CPU technology from Microsoft or Amazon to give these same features to their customers.

Another potentially revolutionary aspect is that the computer industry has been lulled into sticking with x86 backwards-looking compatibility for decades. It's been a very long time since anyone has been willing to put significant effort into a different framework. Apple's success at doing it is going to feed others' willingness to take the plunge.

Some people will strongly resist it because Compatibility is King. For decades, home and office computing has benefited from "PC Compatible" x86

computer architectures. They all have the same type of CPU and a modular design that lets you add and remove individual parts at will. Here's a short list of parts you can swap in and out of a standard PC but you that CANNOT change in a new M1 Mac:

- The CPU
- RAM (main memory)
- Disk drive
- GPU (Graphics controller)
- Network adaptor

All of these are integrated parts of the M1 SoC. Many people, including this writer, lament the fact that Apple computers are all but unrepairable. If something goes wrong, there are no authorized third-party repair shops, only Apple. And they are not likely to "fix" your computer, just replace it, at possibly a high cost both to you and to the environment.

On the other hand, the history of computing is a series of coalescing black boxes. As a specific technology matures, it is made physically smaller, and eventually subsumed within the larger black box of the "system" as a whole, no longer a distinct, replaceable, modifiable appendage. This has already happened to Math Co-processors, I/O Ports, WiFi, Ethernet, Memory Controllers, and Graphics Controllers.

When this happens, innovation on that device slows to a crawl. The only people working to improve it are the ones now responsible for assimilating it into the larger system. This reduction in flexibility is traded off for increased speed and lower cost.

But the consumer is left with no choices, and any improvements must wait for an entire system replacement. Generally this coincides with the natural slowing of innovation for that device. For example, nobody has consciously considered the Math Co-processor in their PC since 1995.

In a very few, special areas there still remains a market for separate devices. The most obvious example left is with Graphics Processors. Several niche markets (gaming, crypto-currency, machine-learning) have fueled the continued development of high-end graphics cards that can be swapped in and out of a personal computer independent of other system components. People who don't need that extra capability use the included on-board graphics, while those that do need it have the flexibility to add it.

Apple's M1 SoC includes all of these things within a tightly closed system. You can't add or alter any one of them. Once you buy the computer, its configuration is frozen into place. To add memory or disk space, get better graphics, or replace something broken, you will have to buy a whole new machine.

The truth is that this model works just fine for the majority of users. Even many of those who dislike the idea of it will admit that they don't actually make any upgrades to their laptop over the course of its natural life. But it can still irk the inner geek to have to give up the ABILITY to upgrade it if so desired.

This trade-off will play out in the marketplace over the next few years. Is the extra speed and efficiency of a highly integrated SoC worth it to give up the freedom to add the flashiest graphics card every year?

Along with inflexibility is the inability to repair. Together these might make other players more like Apple. Their devices may also become proprietary black boxes that will not work with anything outside their ecosystem. And thus, more expensive for users.

Regarding the Internet of Things, I don't think this will have too much effect. IoT devices already use super low-power (and are not very computationally-powerful either), because that's all they require. After all, they just need to be able to talk to WiFi and turn a few bulbs on and off – just enough to be useful to hackers.

It remains to be seen how all this will affect security. One hopes that ARM devices will not be subject to the same kinds of flaws that Intel has had exposed in the past couple of years, but no one knows yet. A flowering of different manufacturers comes with the double-edged sword of a more diverse supply chain: consumers (in this case, the engineers that fill datacenters) will have to decide whether a given product is robust enough or if the maker is not yet mature. But by the same token, if those Intel security flaws had been worse, the present monoculture could lead to a spectacular meltdown.

Conclusion

The blow that Apple just dealt Intel may not be fatal, but it could be deadly. Apple is not the first to move away from Intel, but they are the most visible. They've never had a huge market-share of the home and office PC market, but they have an outsized mind-share. Success with the M1 SoC will inspire copycats at home, in the office and datacenters, which may add up to a thundering herd running away from Intel towards faster, cooler, and cheaper alternatives.

Then again, Intel has been on the ropes before. Perhaps this will sharpen their focus and we'll see some amazing innovations from them in the next couple of years. Let's wish them luck!

Keep Chrome up to date

Google just released a Chrome update to fix a zero-day flaw already being exploited in the wild. That means the bad guys knew about this one before the good guys, so update your Chrome browser ASAP.