

## The CPU is dead. Long live the CPU!



Scene from Gladiator when he is decided on the end of somebody's life.

Look at the computers and laptops sold at your local computer shop. There are just few systems with a separate GPU, neither as PCI-device nor integrated on the motherboard. The graphics are handled by the CPU now. The Central Processing Unit as we knew it is dying.

To be clear I will refer to an old CPU as "GPU-less CPU", and name the new CPU (with GPU included) as plain "CPU" or "hybrid Processor". There are many names for the new CPU with all their own history, which I will discuss in this article.

The focus is on X86. The follow-up article is on whether the king X86 will be replaced by king ARM.

Know that all is based on my own observations; please comment if you have nice information.

AMD: Fusion and the APU



AMD saw it coming that heterogeneous computing was the future, and bought graphics card manufacturer ATI in 2006. Simple goal: build a single chip which is an hybrid of an AMD CPU and ATI GPU. Management en marketing loved it and in 2006 they bragged about their plans a lot. But the merger of the two companies took more time than anticipated, and the heating-problem was not well-solved back then to get a good performing embedded GPU. The company needed to work on the product until 2011 to be able to come up with a production-ready hybrid processor.

Last year the GPU was comparable to high-end GPUs of then 5 years ago, now the difference is 4 years. Meaning that in 2014 the architecture could be as fast as today's high-end GPUs. This 5 years delay has also a lot to do with the semiconductor device fabrication process - the 2006's 65nm seemed not to be enough.

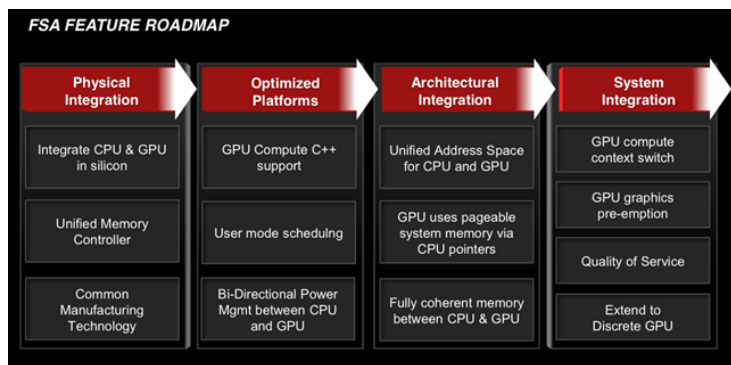
The most interesting part of this new CPU is the much lower total power-usage: starting at an incredible 4.5 till a 100 Watt TDP. The [Wikipedia-article for Fusion](#) has a good overview of Watt-usage.

AMD still develops high-end CPUs and GPUs to let the market decide when they're ready to drop these lines. For instance they

target the professional graphics industry with high-end FirePro and multi-screen graphics cards. Extreme gamers, who always have been the drive for next-generation graphics, are more and more fine with their 2 year old GPU or own a game-console for their games.

In 2011 they introduced the new processor with a marketing-campaign "Fusion" (see image), which I think is a very good name for the new architecture. Now they just use "[APU](#)", which stands for Accelerated Processing Unit. From the start all APUs supported OpenCL.

The level of integration is not full at the moment. Below you see the road-map of AMD for integrating the two parts. FSA stands for Fusion System Architecture.



## Intel: Sandy Bridge and Ivy Bridge



Like AMD Intel has CPUs with embedded GPU since 2011. Their GPU-technology is licenced from Nvidia. They managed to get a good-performing embedded GPU, but only supported OpenCL for the Sandy Bridge CPU. On Ivy the GPU supports OpenCL.

Just like AMD Intel needed to solve the heating-problem, so earlier than 2011 seemed impossible. Intel had merger-problems with their GPU-partner on their own way - it ended with [paying 1.5 Billion USD to NVidia](#) beginning of 2011. From the then the road was open and also the competition with AMD.

If you check the [Wikipedia-page of Ivy Bridge](#) on Watt-usage of the various processors, you see that AMD has the advantage in the FLOPs/Watt. More about that later.

Intel promotes its processors a little different than AMD. Where AMD has very clear APUs, GPUs and CPUs, Intel has Ivy Bridge

processors with and without GPUs, and motherboards with GPUs. So you can buy an Ivy Bridge and a motherboard, ending up between 0 and 2 GPUs.

## Intel MIC, Phi (and many other names)

Intel has a 50-core accelerator housed on a PCI-card. So it looks like a discrete GPU, but actually is a CPU. It supports X86, but not really as it still needs to be recompiled and ported. Or as some others say, that it actually/eventually will be fully portable and will be a better fit for NVidia's acceleration-device Tesla. If you read all the articles written on this line of products, it seems that Intel demanded a lot from themselves and it was not an easy road. What I find unexpected is that Intel chose for a PCI-device - one of the problems the hybrid CPUs solve is the too high transfer-time the PCI-bus causes.

Anyhow, there are many thoughts on this device. There have been so many names and even different devices with the same name, that I don't want to get into this architecture too much. Will the future CPU by Intel be based on this processor?

## Nvidia: Tegra



Foxconn Tabelt with Nvidia-branding (2010)

Nvidia makes GPUs, not CPUs. At least not X86 - in that space only AMD and Intel rule. Licensing GPU-IP to Intel is not enough to secure their future, so they chose to focus on the high-end High-Performance Computing and ARM-processor based tablets/phones. As a company that is being kicked out their main business (the dedicated graphics cards), they currently make more risky decisions - that makes it more interesting to watch NVidia's decisions than Intel's - Intel is very protective of their market and does not easily switch technologies now the X86-market is still very profitable.

Let's focus on Nvidia's ARM-processors. ARM-processors have been hybrid processors for a long time, labelled "System on a Chip" (SoC). A GPU-designer (as there are many for ARM-processors) licenses a processor from ARM and then builds their GPU around it. Nvidia has the advantage of being a well-known brand in the mass-market, meaning that a tablet could be sold to consumers more easily by unknown brands (see photo with Nvidia-brand). They also chose competitive prices (\$20-\$25), which increased interest from tablet-builders even more. Read [here](#) what Forbes thinks about it.

## The new kingdom of X86

So it seems that AMD and Intel will rule the kingdom of X86, but Nvidia cannot be ignored. Actually not much changed, if you look at the brands only.

But a front-line is building up, now the borders of ARM-land are getting closer. Will ARM break open the 2-party CPU-market? Using OpenCL a 2-5 GFLOPS ARM-processor can pump out up to 50 GFLOPS when the GPU is used - 100 GFLOPS is spoken of already. The X86-GPUs are up to 5000 GFLOPS, but those GPUs uses hundreds of Watts, while the ARM-devices use far less than 5 Watt. More about this next time.

And no, replacing the name CPU with HPU, APU or xPU will not happen. People know the name CPU or "the processor" and will

not follow technology if it comes to naming. If in the future memory can be integrated into the CPU, it will still be called a CPU (or "fast CPU") and not mCPU. Therefore I chose the famous sentence "Le Roi est mort, Vivre le Roi" as title for this post - the king is completely different from the previous one, but his name will remain King.

Here at StreamHPC we're happy to see so much influence from OpenCL, the technology that makes the new processors run fast. We offer [coding/consultancy](#) and [training](#).