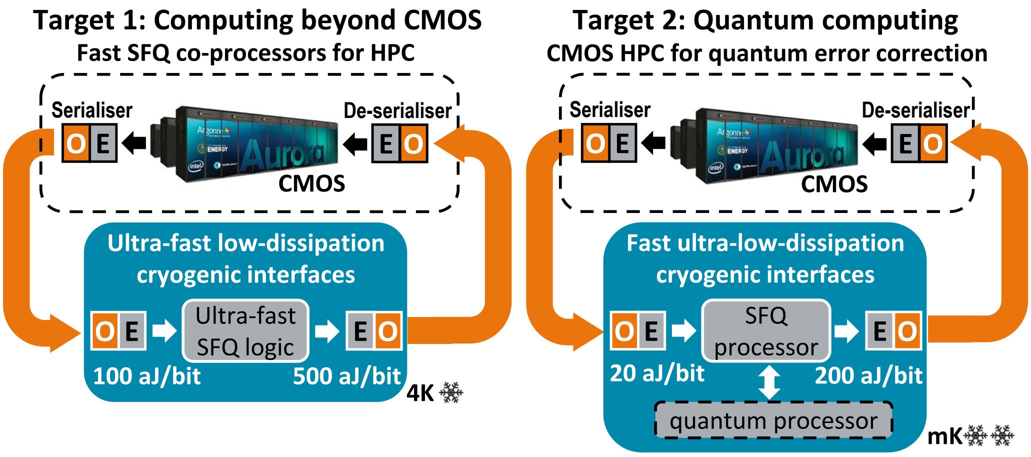
A cold touch to supercomputers

At the beginning of October 2020, we have kicked-off the EU project aCryComm (Horizon 2020 FET Open, grant agreement ID: 899558). The main goal of the project is to develop suitable interfaces to let traditional supercomputers communicate with special processors operated at very low temperatures that can help boosting computing power and efficiency, especially when used for those specific computing tasks where they stand out.



To better understand what we are talking about, it is probably a good idea to recall what happened about 20 years ago to personal computers (PC) and, more in general, to computer gaming. PCs till year 2000 included a central processing unit (CPU) to process data and a graphic card to help displaying 3D images and animations as fast as possible. Then, Graphic Processing Units (GPU) came to the market, i.e. dedicated processor chips to efficiently handle the heavily parallel tasks required by 3D rendering: it was (literally) a game changer. In this computing architecture, the CPU takes care of the computing tasks, whereas the GPU (to be more exact the so-called General-Purpose GPU) co-processor steps in whenever 3D rendering comes into play. Today this approach has become a commonplace, found not only in almost any laptop but also in most smartphones.

Beyond consumer electronics, suitable combinations of CPUs and GPUs are nowadays at the core of fastest supercomputers. In fact, the ability of GPUs to handle 3D rendering, makes them particularly suitable for vector and matrix calculations, that are at the core of several scientific calculations. CPUs take care of serial calculations whereas GPUs take care of highly parallel calculations.

One major limitation of supercomputers is power consumption, often exceeding 10 MW, which would cover the needs of several thousands of homes. The energy dissipation is due to CPU and GPU processors being made of silicon, where electrical signals generate heat due to the relatively high resistance of the integrated circuits. In turn, this requires additional power to cool down the electronics and the room hosting the machine, to operate it in the optimal temperature range. Combining CPU and GPU has already helped reducing the energy consumption to up to a factor of three, but resorting to superconducting technology would help to reduce it even further, up to several orders of magnitude for some calculations. In fact, our vision is that the power-hungry silicon-based GPUs could be replaced by superconducting logic, that can be at the same very fast (up to two orders of magnitude faster than silicon chips) and energy efficient (several orders of magnitude, even when taking into account the power consumption needed to cool the chip down).

Contact Matteo Cherchi, matteo.cherchi@vtt.fi

We are reviving here a relatively old technology called single flux quantum (SFQ), that had been proposed about thirty years ago as a possible alternative to silicon technology (despite the name, we are talking about a technology for fully classical processors). Aiming at full replacement of silicon chips with superconducting chips doesn't make much sense. In fact, ubiquitous silicon chips have reached impressing maturity and achievements and they work at room temperature, unlike superconducting devices that require temperatures close to absolute zero (-273°C). On the other hand, we believe that developing GPU co-processors based on SFQ technology could have a major impact both on performance and power consumption of future supercomputers.

Another important addition to future supercomputers will be the quantum processing unit (QPU), i.e. an universal quantum computer able to solve certain types of problems exponentially faster than classical computers, thanks to the high parallelism ensured by the unique properties of quantum bits, called qubits. The most mature technology for such processing units is based on superconducting qubits, which recently allowed Google to demonstrate for the first time so-called quantum supremacy, using a machine with 53 qubits. Even though this was a very important milestone in the history of computation, it was just a proof of concept based on a problem created *ad hoc* for such demonstration, but with no practical interest.

It is generally agreed that, in order to achieve practical quantum computers, the number of qubits must be scaled up to about one million, which is a major leap compared to the state of the art. In our vision, SFQ logic can play an important role also to reach this ambitious goal, by providing a classical co-processor operating in the same cryostat of the QPU with low energy dissipation.

In summary, we believe that SFQ logic can play an important role in the future of supercomputers, both as a possible replacement of silicon technology for GPUs and as a classical co-processor to interface future QPUs with the classical processors. Adding SFQ technology to supercomputers, necessarily means bringing cryostats in the supercomputer room, but also developing suitable links to let them communicate with the processors at room temperature.

We have conceived our project aCryComm exactly to lay the groundwork for the development of this communication links, and in particular the interfaces to take the data to and from the SFQ logic inside the cryostat. Our natural choice for the physical communication channel is the optical fibre, because it ensures at the same time incomparable speed, bandwidth, and energy efficiency, as well as negligible thermal conduction. Our main challenge is the development of the interfaces to be placed in the cryostat. On the input side, we need convert optical light signals coming from the room-temperature processors into electrical signals that can drive the SFQ logic. On the output side, we need to convert the electric signals coming from the SFQ into optical signals going back to the room-temperature processors. In both cases, the operation at cryogenic temperatures poses uncommon challenges compared to standard optical communication technologies, but it also offers unique opportunities that we are trying to identify and exploit.

The project is coordinated by VTT, where we have conceived the general vision in the last couple of years, thanks to discussions among colleagues with very diverse backgrounds and competencies including superconducting devices, integrated photonics, packaging, and metrology. Based on this vision, we have then attracted some of the best experts at European level addressing all the key aspects of the envisaged development.

Contact Matteo Cherchi, matteo.cherchi@vtt.fi