Essay

My research topic is "Computational affinity estimation of protein-protein complexes." The main idea is to develop a tool for protein-protein complexes affinity prediction (stable / not stable) from the SKEMPI database (https://life.bsc.es/pid/skempi2/) based on their structural descriptors. The database contains much information about the binding constant of a complex, kinetics, and thermodynamics upon protein mutation. At first glance, HPC does not seem to apply to such a problem, but here we should go a little bit deeper. How much information we have to keep for one protein? Standard "PDB" file from the mentioned database contains, at average, 2000 residues, and each residue contains up to 20-30 atoms. Considering the necessity of processing every pair of atoms in neighboring residues and a large number of proteins (each protein also has several mutants) at last, we have to deal with a large amount of data. Even to properly read all the files and extract the basic properties of the molecule, it takes much time. The atoms processing part is the so-called embarrassingly parallel problem, so here HPC obviously can help. For now, I use a small part of the database only for algorithm design. But when the amount of data increases, knowledge about MPI, for example, definitely helps to increase the overall efficiency of protein data analysis.

For sure, I know plenty of HPC applications in the pharma/bio industry. So in this essay, I want to write about something different: Spaceborne, the supercomputer that was sent to the International Space Station (ISS), and various applications in IoT:

1. This supercomputer, called the Spaceborne computer, was part of a year-long experiment conducted by NASA to run a high-performance commercial off-the-shelf (so-called COTS) computer system in space, which has never been done before. The goal was for the system to operate seamlessly in the harsh conditions of space for one year – roughly the amount of time it should take to travel to Mars. So the global aim is to prepare humanity for Mars mission because for such mission onboard supercomputer can significantly enhance the effectiveness of various space research projects. Today most of the calculations needed for experiments on ISS are still done on Earth due to the limited computing capabilities in space, which creates an additional challenge when transmitting data from and to space. A future mission to Mars will require sophisticated onboard computing resources that are capable of extended periods of uptime. To meet these requirements, modern scientists need to improve technology's viability in space to better ensure mission success. Spaceborne was successfully launched in 2017 by the SpaceX rocket (famous Elon Musk company), and the year after was also successfully returned on Earth. So the first experiment on International Space Station was done well, and as far as I know, NASA wants to use a similar supercomputer in further Mars mission development.

2. The IoT market grows extremely fast these days, so the amount of data to be processed in real-time increases. So again, we have a similar situation - HPC could help to process all that data. But how can IoT and HPC become close friends? The beautiful synergy can be seen in a particular example [El Baz, Didier. (2014). IoT and the Need for High-Performance Computing. 10.1109/IIKI.2014.8.]. The ADREAM building in Toulouse, France, features 6000 various sensors (temperature sensors, light sensors, motion sensors, cameras, e.t.c). The building also features a mobile grid in order to fix lights, sensors, Motion Capture (MoCap) cameras with IP addresses. Besides solar panels, ADREAM building features devices like a geothermal exchanger with very low energy and energy storage batteries. Basically, a lot of IoT stuff. Real-time management of such a smart building gives rise to many problems like management of data from the many sensors and optimal scheduling of tasks in relationship with heating/air conditioning and light management, which is a very difficult optimization problem whose solution demands intensive computation. Optimal scheduling of tasks that consume energy like light and air conditioning is even an NP-complete problem. HPC solutions can take great benefit from new devices like GPUs that have been reported to reduce dramatically computing time by factors from 50 up to 150!

So there are still plenty of possible applications of HPC in various spheres of our life - from smart houses to space missions. The only questions - who will implement HPC? Perhaps Skoltech students?

Appendix

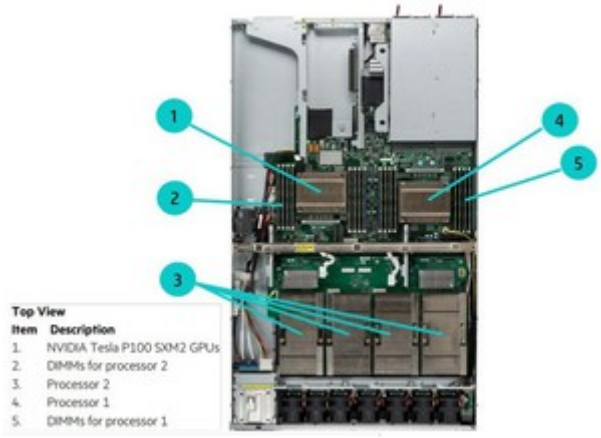

Figure 1. Spaceborne supercomputer (what was inside)

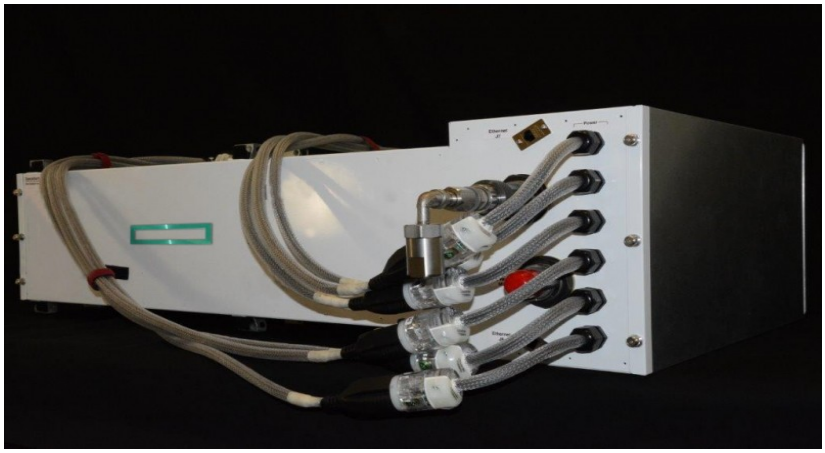

Figure 2. Spaceborne computer (look from the side)