The brain on low power scalable architectures: efficient simulation of cortical slow waves and asynchronous states.

Andrea Biagioni[1] for the INFN WaveScalES and ExaNeSt team
[1] INFN, Sezione di Roma, Italy

E-mail: andrea.biagioni@roma1.infn.it

Abstract. The brain cycles among entirely different states. The unconscious deep-sleep state and the asynchronous awake state represent the extremes of the range. The deep-sleep is believed to play an essential role for memorization of experiences and action plans, and for bringing back the system toward its optimal energetic and computational working point. During deep-sleep, the brain expresses slow oscillations which appear to be synchronized at a macroscopic scale in space and time as travelling waves (slow-wave activity, SWA). On the contrary, the processes observed during wakefulness include asynchronous irregular activity (AW, Asynchronous Awake).

Such non-homogenous behavior poses different challenges to simulation engines. Efficient brain simulation is a scientific grand-challenge, a parallel/distributed coding challenge and a source of requirements and suggestions for future computing architectures. Indeed, the human brain includes about 10^{15} synapses and 10^{11} neurons activated at a mean rate of several Hz. Full brain simulation poses Exascale challenges even if simulated at the highest abstraction level.

The WaveScalES experiment in the Human Brain Project (HBP) has the goal of matching experimental measures and simulations of slow waves during deep-sleep and anesthesia and the transition to other brain states. The focus is the development of dedicated large-scale parallel/distributed simulation technologies.

The ExaNeSt project designs an ARM-based, low-power HPC architecture scalable to million of cores, developing a dedicated scalable interconnect system, and SWA/AW simulations are included among the driving benchmarks.

At the joint between both project is the INFN proprietary Distributed and Plastic Spiking Neural Networks (DPSNN) simulation engine. DPSNN can be configured to stress either the networking or the computation features available on the execution platforms. The simulation stresses the networking component when the neural net – composed by a relatively low number of neurons, each one projecting thousands of synapses – is distributed on a large number of hardware cores. For growing number of neurons per core the computation starts to be the dominating component. In this talk, I will report about preliminary performance results obtained on an ARM-based HPC prototype developed in the framework of the ExaNeSt project. Furthermore, a comparison is given of instantaneous power, total energy consumption, execution time and energetic cost per synaptic event of SWA/AW DPSNN simulations when executed either on an ARM- or INTEL-based server platform.