Stream Joins on Many Cores

Laboratory, Semester Project, or Master Thesis

As a consequence of *Moore's law* and the idea of *multi-core architectures*, core counts in modern computing systems are increasing rapidly (tens of cores have already become commodity). This trend further and further increases the pressure on software makers to find efficient algorithms that can utilize such high degrees of hardware parallelism.

As a side project to the *Avalanche* project in the Systems Group at ETH, we have developed a new database algorithm that can solve one particular database task in a massively parallel way. More specifically, our operator uses a *pipelining model* to orchestrate available CPU cores and can use available parallelism to run *window-based stream joins* at very high throughput rates.

After we demonstrated the potential of our technique for a particular workload type, we would like to deeper explore how our approach will react to different real-world situations (*e.g.*, load fluctuations, dynamic CPU allocation, heterogeneous compute resources, virtualized environments). To this end, we would like to have a **visualization and analysis framework for our system** that helps us understand its characteristics and extend its scope to further workload types. Possibly that could also lead an inclusion of further compute resources, for instance *graphics processors*, into our current framework.

The interested student would get an insight into modern computing platforms and the challenges around multi-core processing. From the student we would expect some familiarity with the C programming language (our current code is written in C), though the framework itself could also be written in a different language (*e.g.*, Java). The work would be done within a very vibrant and active research team; we would encourage the student to participate in this team and contribute his own fresh ideas.

We would define the exact scope of the work depending on whether you want to join us in the context of a *laboratory*, *semester project*, or *Master thesis*.

Contact: Jens Teubner, `jens.teubner@inf.ethz.ch`