SemBoost: accelerating semantic intelligence applications on massively parallel architectures

Paolo Burgio, Micaela Verucchi, Marko Bertogna
HiPeRT Lab, Università di Modena e Reggio Emilia, Italy

Tools for Semantic Analysis process structured text
- E.g., web pages
- To extract semantic information
- Almost in real-time (few milliseconds)

Still cannot exploit massively parallel architectures

Technological Development

SemBoost project
Application ➔ Search in a network
Using GP-GPUs
Data structure ➔ Network of concepts

Effectively achieving performance through parallelization requires:
1. Adequate application partitioning
2. Ad-hoc data structure

Traditional/"Naive" representations are:
- Pointer-based
- Sparse in memory
- Hard to move across different memory spaces

CompactArr (CA) is an ad-hoc data structure explicitly designed for host+accelerator systems
1. Compact
2. Cache-friendly
3. Easy to move across different memory spaces

Comes from LightKer project @UNIMORE

Cache-friendly
- Speedup also on CPUs
- 2-3X against pointer-based implementation

Reduced memory occupation
- <2MB against 26MB
- ~15X reduction

CompactArr
Enables adoption of GPU accelerators
- 20% speedup vs. pure CPU

Contact: Marko Bertogna
E-mail: marko.bertogna@unimore.it
TETRACOM contribution: 30000 €
Duration: 4/1/2015-2/1/2016