Perspectives on HPC Hardware/Software

Lucian Anton
Cray UK

NEMO HPC working group meeting
7th February 2018

Disclaimer: competence without comprehension

Outline

- Overview of the current HPC hardware
- Perspectives for the next 5 years
- Impact on applications

Overview

- **Since mid 2000 the processor frequency scaling has ended**
  - Fast clock and ILP have driven the performance
  - The size of combinatorial logic needed to handle efficiently multiple instructions in flight does not scale well with pipeline width.
- **OO superscalar architecture was topped up with multicores, hardware threads, vector floating point units and several levels of memory caches**
- **These kind processors are in high demand from laptops to data centers**
  - Flexible at the cost of relatively higher power consumption
    - But simpler architecture struggle to deliver performance for algorithms with complex data dependencies
  - Currently the HPC market is dominated by Intel Xeon processors
  - and accelerators

Next 5 years (I)

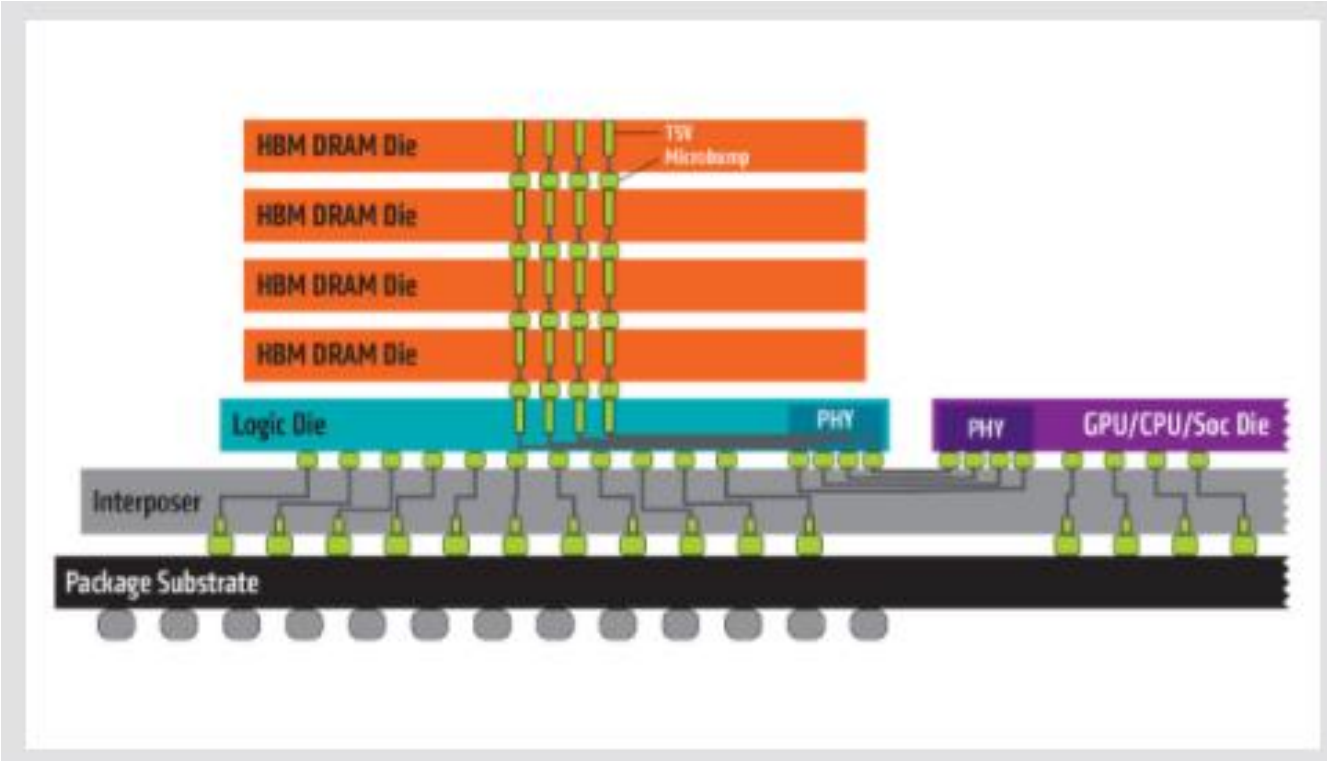
- **Number of cores keep increasing (Gordon's law still working 14nm → 7-5 nm)**
  - Cache coherence needs a different approach (directory based protocol) or coherency domains
- **Vector processing unit width length to reach 512 bits + more complex instructions set**
  - The instruction set AVX-512 was extended by Intel
  - New instruction set, SVE, proposed by ARM
- **High bandwidth memory will be present in many server class processors, … probable**
- **More vendor choice:**
  - Intel: Skylake …
  - AMD: Naples, Rome, …
  - Cavium: ARM ThunderX{2,3}
  - IBM Power 9?
- **For a Thunder X2 vs Broadwell comparison see:**
  - http://www.goingarm.com/slides/2017/SC17/GoingArm_SC17_Bristol_Isambard.pdf

Next 5 years (I)

- **HPC interconnects will offer more bandwidth**
  - Needed as the number of cores/node keeps growing
  - More intelligent data routing: adaptive, levels of priority, congestion avoidance mechanisms
- **Storage: SSD and NVRAM will allow for higher bandwidth, better network balance to support more complex workflows, faster data transfers.**
- **Hopefully many of these enhancements will be controllable via the current communication or storages APIs or perhaps automatically optimized**
  - Developers life is hard enough as it is

Applications: HBM (I)

- High bandwidth memory has the potential to increase significantly the performance of many scientific or engineering HPC applications.
- ~500-1000 MB/s
- ~16-32 GB/processor
- expensive



- http://www.dcs.warwick.ac.uk/pmbs/pmbs/PMBS/papers/paper7.pdf

Applications: HBM (II)

- **HBM can be used as addressable memory or as one more level of cache**
  - Cache is simpler, but …
    - because it is directly mapped: 1 node cache trashing will slow down the whole computation.
- **Faster compute inside node will increase the relative cost of load imbalance and communication**
  - Data partition and communication patterns might need reanalysis

Applications: Vector instructions

- **AVX-512 and SVE aim to increase the amount of code vectorization**
  - Variable vector length, gather-scatter memory access pattern, line predication, fast transcendental functions
  - Most probable the width of vector register will stop at 512 bits for a while
  - Hopefully vectorization will be handled well by the compilers, but one might need to do it by hand in the code critical sections.
    - Having 2 vector instruction sets is one more ache for developers

Applications: Cache coherency costs

- **As the number of cores increases the cache coherence protocol had to adapt**
  - >~60 cores/socket directory based cache coherence.
  - implies more operations across the internal network, higher latency
  - Hence, as ever, data locality is important for performance.

Conclusions

- **I expect that NEMO's performance will grow automatically on the future processors with help from the new hardware and supporting software (compiler, libraries).**
- **However, I also expect that going beyond the average will require code changes, probable both: algorithms and implementation.**
  - Having processors from 3 different vendors will most probable increase work load regarding performance portability.
  - But basic rules and performance modelling should work the same
    - Vectorisation and cache optimization; memory bound vs compute bound, etc…