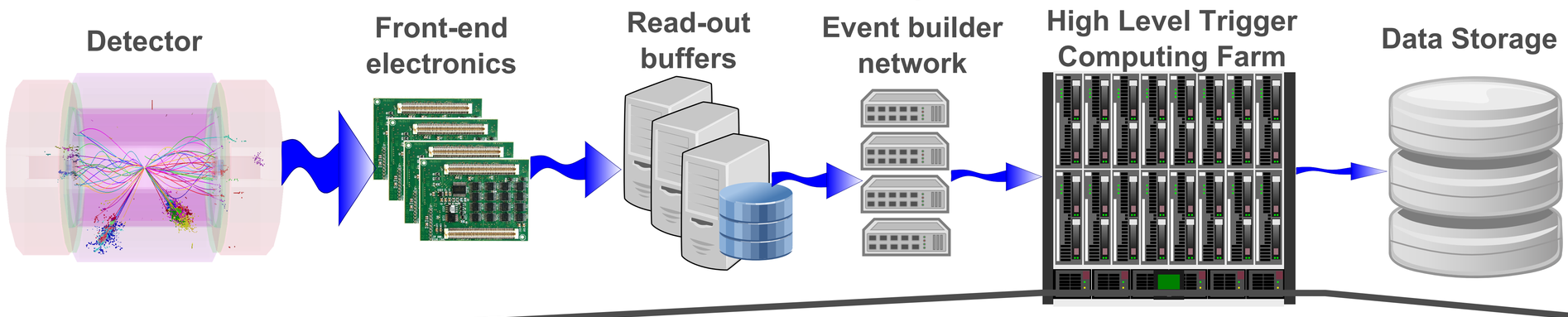
Many-core processors in High Energy Physics

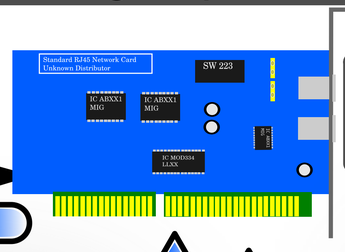
by Przemyslaw Karpinski and Aram Santogidis

May 2014

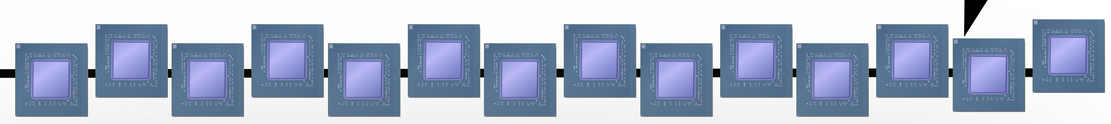


A Data Acquisition System (DAQ) is required to perform the initial selection of data coming from a detector. More sophisticated processing is performed in the High Level Trigger (HLT).

Network Interface Card



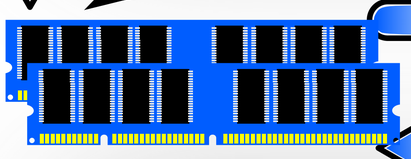
Many-core architecture



Network-Host-Coprocessor Data Transfer

The event data sets arrive from the network to be stored either to the Host machine's RAM or directly to the Coprocessor. We investigate rDMA technologies to optimize the data transfer. Two prominent technologies are **iWARP** for the host and **CCL** for the Coprocessor. The Coprocessor Communication Link (CCL) is a proxy driver that allows inter-node (i.e., between nodes) rDMA by virtualizing direct access to a hardware InfiniBand Host Channel Adapter (HCA) from Xeon Phi.

Host Memory

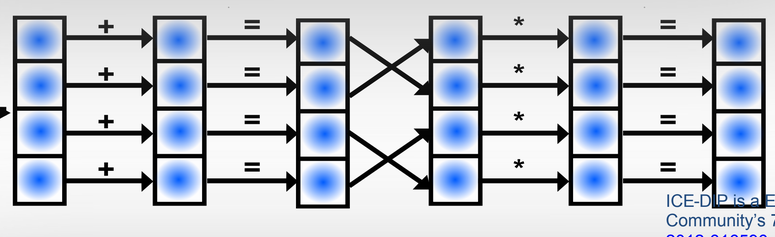


PCI Express Bus

r D M A

Problem definition for a parallel environment

Parallelization of a computational task brings additional issues: task scheduling, process synchronization, data non-locality. New solutions need to consider that current software frameworks are not adapted for parallization, reconstruction algorithms need to be redesigned to fully utilize hardware capabilities, and users need tools that are intuitive and easy to use. Xeon Phi together with **MPSS** (Intel Manycore Platform Software Stack) are offering a possible solution to all these problems.



ICE-DIP is a European Industrial Doctorate project funded by the European Community's 7th Framework programme Marie Curie Actions under grant PITN-GA-2012-316596

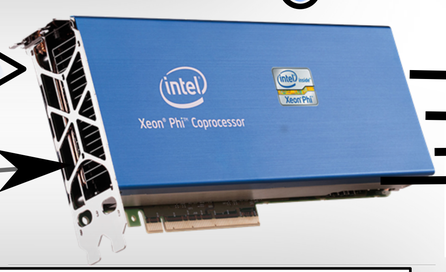
Host-Coprocessor Data Transfer

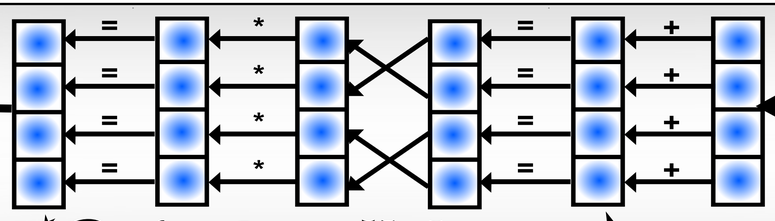
There are a number of data transfer mechanisms enabling bi-directonal communication between the Host and the Coprocessor over the PCIe bus. (**MYO**, **COI**, **SCIF**, **NetDev**).We investigate the performance of the available options in respect to the workloads running at CERN DAQ systems. Additional research will be done for custom data transfer mechanisms tailored to the needs of specific applications related to HEP.

Using Vector instructions (SIMD)

Vector instructions allow to perform the same operations on multiple sets of data, at the same time. But in Iterative algorithms, data dependencies can cancel out possible speedup gains. Performing vector operations on scattered data sets may also cause memory access stalls and further increase the execution time. High SIMD utilisation is a key for increasing performance in parallel applications.

Xeon Phi™ Coprocessor





Data Parallelism

Profiling Performance and Power

In order to measure improvements either in computation performance or power consumption, we utilize the PMUs present in modern architectures in order to compute and compare against metrics. (e.g. bandwidth utilization, access latency, performance per Watt etc.). We use industry standard tools enabling access to the hardware performance counters and develop our own custom utilities for performance monitoring. We also benchmark HEP related sample software in order to identify potential improvements.

Programming models & Tools

New architectures are designed so specific programming models execute with better efficiency than others. Comparison of both classical and experimental software technologies, like: **pthreads**, **MPI**, **openMP**, **TBB**, **Cilk++**, **OpenCl** and others is being conducted. Diversity of problems present at CERN require diversity of tools and knowledge how to pick the right tools.

ICE-DIP is a European Industrial Doctorate project funded by the European Community's 7th Framework programme Marie Curie Actions under grant PITN-GA-2012-316596