Accelerating FM-index on the Intel KNL many-core processor

Jose M. Herruzo, Sonia Gonzalez, Oscar Plata
Dept. Computer Architecture, Univ. Malaga
jmherruzo@uma.es, sgn@uma.es, oplata@uma.es

UNIVERSIDAD DE MÁLAGA

Collaboration

UAB group: Analysis of the sequence alignment problem (based on FM-index) in bioinformatics.
UNIZAR group: Implementation of the FM-index algorithm on KNL.

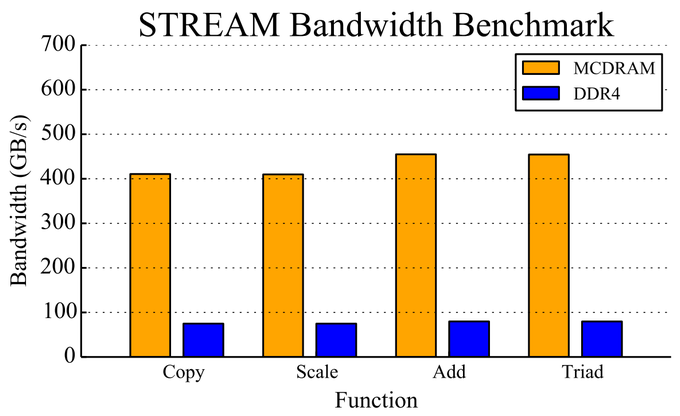
Motivation

FM-index:
Data structure based on the Burrows-Wheeler transform (BWT) that has been succesfully used for fast and memory-efficient alignment of sequences in bioinformatics.

Computational structure:
Alignments based on FM-index exhibits an irregular memory access pattern with low locality and low arithmetic intensity. Movement of data through the cache/memory hierarchy is the main performance bottleneck.

KNL processor:
Intel many-core processor with 64 Atom cores, L1/L2 cache hierarchy and 2D mesh interconnect. Each core supports 4 hardware threads and AVX-512 SIMD instructions.

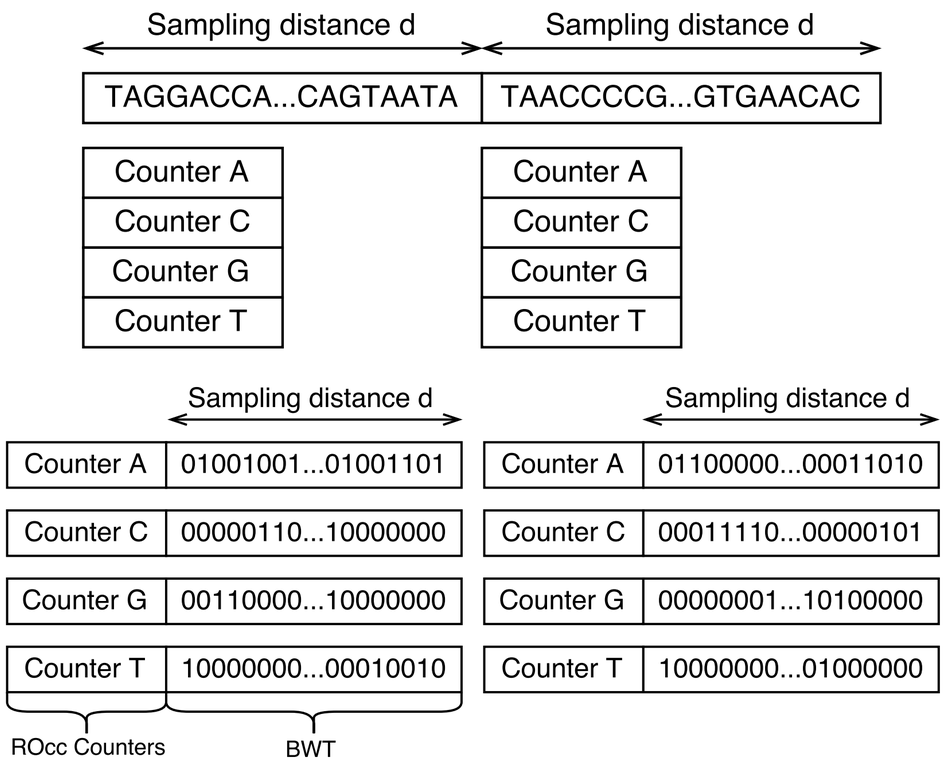
Asymmetric main memory:
192 GB DDR4 RAM, bandwidth: 90+ GB/s.
16 GB on-package MCDRAM, bandwidth: 400+ GB/s.



Accelerating FM-Index

Memory latency:
Hidden by manual and automatic (compiler) techniques: ILP, multi-threading, AVX-512 intrinsics, prefetching ...

Memory bandwidth:
Adapting FM-index data structure to fit MCDRAM available bandwidth (reducing data movement).

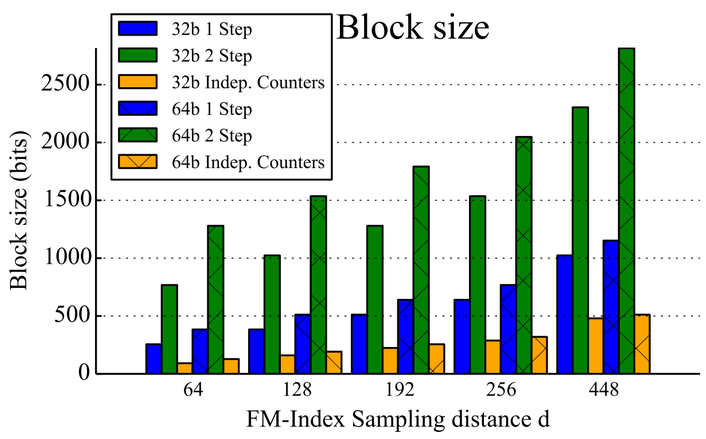


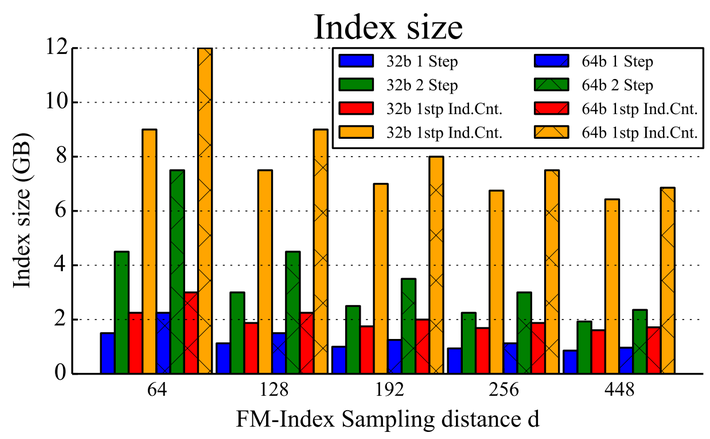
Data required to search a symbol in a 2-step FM-index is loaded in a single memory block. The computational cost is reduced but the total memory footprint is increased.

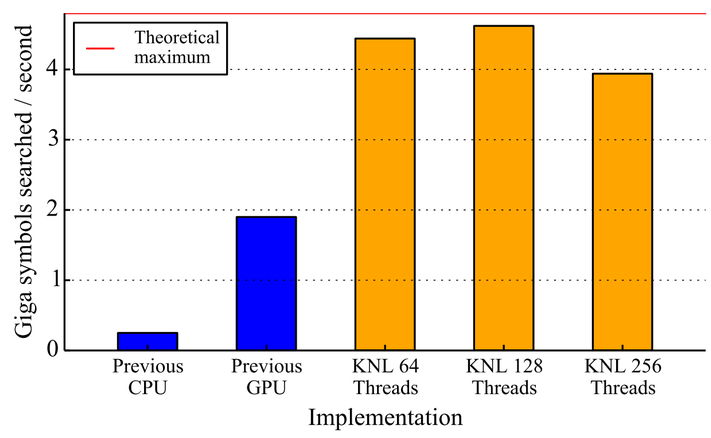
```
INITIALIZE Reg[0], Reg[1]
NextSymb[0] ← ENCODE_SYMB(Reg[0],Seqs,Index)
NextSymb[1] ← ENCODE_SYMB(Reg[1],Seqs,Index)
PREFETCH(Reg[0],Index)
PREFETCH(Reg[1],Index)
Comment: multi-threaded: loop in each thread
while (FinishedSeqs < Nsequences/Nthreads) do
   for (k=0; k<2; k++) do
      Comment: LF vectorized using AVX-512 intrinsics
      LF(Reg[k], Index)
      if (Any Sequence from Reg[k] finished) then
         FinishedSeqs++
         UPDATE_RESULTS(Reg[k],Results)
         UPDATE_REGISTER(Reg[k],Seqs,Ndone,ThrId)
      end if
      NextSymb[k] ← ENCODE_SYMB(Reg[k],Seqs,Index)
      PREFETCH(Reg[k],Index)
   end for
end while
```

Parallel 2-step FM-index algorithm exploting the high bandwidth of MCDRAM and the AVX-512 SIMD support.

Experimental Results







A. Chacon, *et.al.*, *Boosting the FM-index on the GPU: ...*, IEEE/ACM TCBB (2015)

SyeC Supercomputación y eCiencia