Analysis of Whole Genome Sequencing with MPI on HPC Architecture

Jarlier F [1, 2, 3], Fedy N [5], Sirotti L[5], Hupé P [1, 2, 3, 4]

(1) Institut Curie, Paris France, (2) INSERM U900, Paris France, (3) Mines ParisTech, Paris France, (4) CNRS UMR144, Paris France, (5) Université Paris V, Paris France

Motivation:

The Next Generation Sequencing (NGS) technology offers new insights in cancer research and personalized medicine. Due to its large scale approach, we can detect genetic alterations with an unprecedent accuracy. Due to the decreasing cost of sequencing, whole genome sequencing becomes more widely used in research project. In a near future, it will likely become a tool for daily clinical practice. The drawbacks of such breakthrough are the volume of generated data and also the complexity of the downstream analysis. For instance, a whole genome sequencing with a 40X coverage is 400GB and a variant analysis will last for a week with current pipelines. Actual software tools, used in pipelines, lacks of scalability.

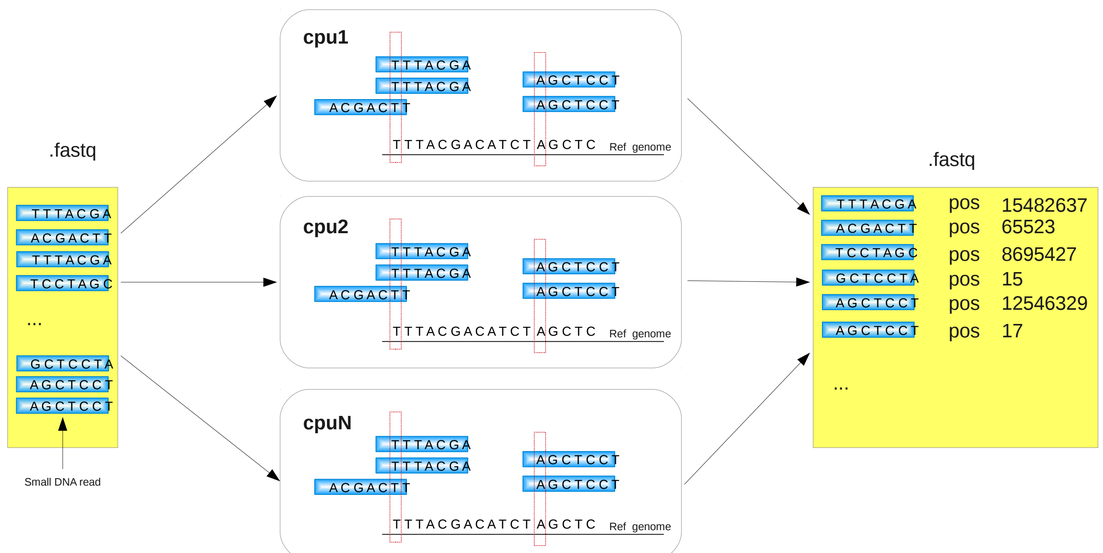
To tackle traditionnal bottlenecks, we have used a parallelization with message passing interface paradigm called MPI. This paradigm has many advantages : it transfers input-output (IO) file system latencies at network level ; MPI also provides many optimizations such as collective operations to optimize IO ; it provides communication between jobs ; it also avoids copy of data thanks to derived datatype.

The classical NGS pipeline consists of two steps. The first step is the alignment of small nucleotide sequences (called reads) on a referenced genome. The second step is the sorting of the alignment result according to chromosome and genomic position. These operations are essential but time consuming. Therefore we propose to optimize the two steps using MPI technology. In the first part we present an overview of the parallelized workflow that we have implemented and then we show the results we have obtained on whole human genome samples.

1) Description of the MPI workflow

Every NGS pipeline starts with the two following operations: the alignment and the sorting. The alignment consists in finding the position of small fragments of DNA produced by the sequencers. Number of reads at a particular position is called the coverage of a sample. The deeper the coverage the better the reconstructed sequence is. Nowadays a standard coverage is around 30X, 40X and even 100X. For instance, a whole genome sequencing with 100X coverage produces 1 TB of data.

After intensive study of the alignment and sorting algorithms we have noticed that a major part of bottlenecks are in the IO file access where all the latency is. To solve that problem we have decided to transfer IO constraints to the network level. Other optimizations have been implemented such as : collective operation, shared memory, indexing, derived datatype. It turns out that these optimizations have reduced drastically the time and the memory consumption.

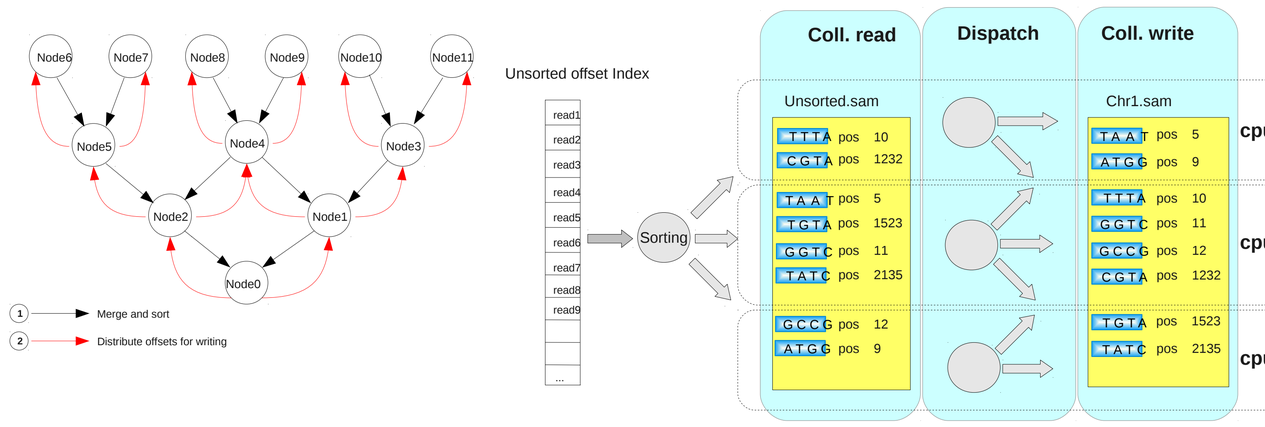


Step 1 :
- MPI Collective read
- Data organization

Step 2 :
- Alignment (BWA MEM)
- Shared memory

Step 3 :
- MPI Collective write

Overview of the alignment algorithm



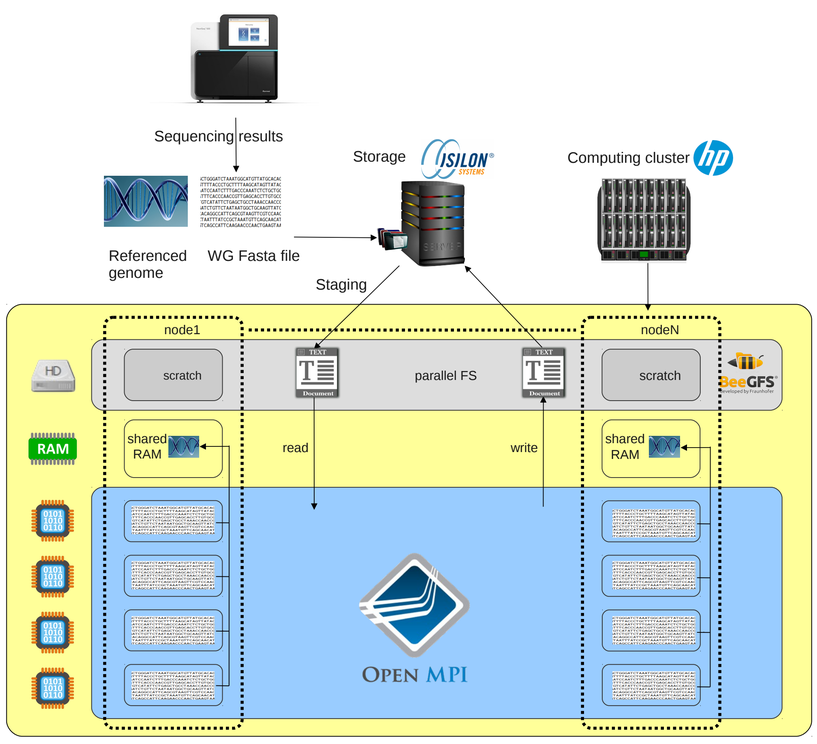
Step 4 :
- Collective read
- Indexing
- Doubling recursion sort of genomic position

Step 5 :
- 3 phases collective read, dispatch and write

Overview of the sorting algorithm

2) Description of the IT architecture

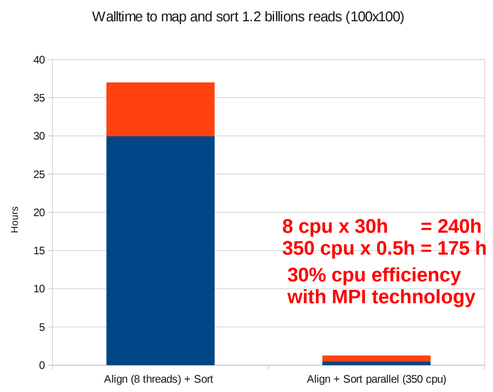
For fast reading and writing, collective operations are mandatory but these optimizations can only be achieved upon a custom distributed file system. At Institut Curie we use BeeGFS. We have also tested our solution on Lustre with the same performances.



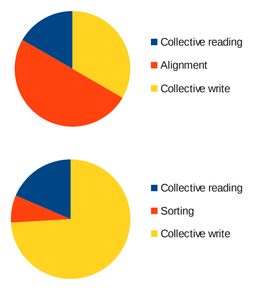
Cluster architecture at Institut Curie

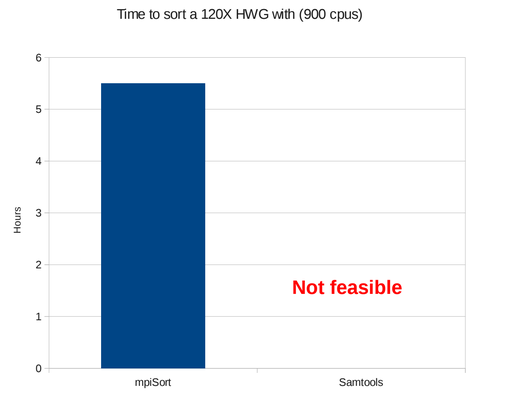
3) Results

In this section we present results we have obtained on both the Insitut Curie cluster and TGCC cluster (CEA, Bruyères-Le-Châtel, France).

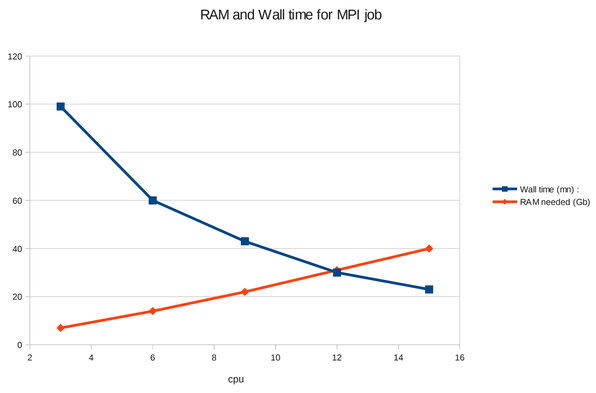


Wall time for aligning and sorting a 40X human whole genome.
1.2 billion reads occupy 400GB on the disk. The efficiency is due to transfert of IO bounds to network bounds.





Wall time for sorting a 120X human whole genome.
This represents 6 billion reads and occupy 1.2TB on the disk.



Influence of the number of cpus for parallel alignement of 50 millions reads file. The memory usage increase linearly.

Influence of the number of cpus for parallel alignement over 5, 50 and 100 Millions of reads.

3) Conclusion

We have developed a MPI framework able to tackle current bottleneck of whole genome sequencing analysis pipeline. In this study we have decided to use parallelization for the alignment and the sorting of NGS data. We have drastically reduced the alignement time and even allows the sorting of reads that was not feasible before. From our results MPI technology is an efficient candidate and performs very well on the cluster architecture we have tested.

The cpu, memory usage, the network and file systems are crucial for a good scalability. MPI adresses efficiently these different aspects as we have shown here. MPI among other optimization technics will definitely help bioinformatics developpers to cross the barrier of the Big Data.

References : www.open-mpi.org ; Fast and accurate short read alignment with Burrows–Wheeler transform (Li H. et al. 2009) ; The Sequence Alignment/Map format and SAMtools (Li H et al. 2009) ;

Acknowledgements : We thank François Prud'homme and Maxime Chevailliot from the Institut Curie IT department for their help in setting up the cluster architecture. We thank Eric Viara for his advices on code optimizations. We thank Claude Scarpelli for providing us with TGCC cluster access. Finally we thank Alban Lermine and Nicolas Servant for NGS support and analysis validation.