



HAL
open science

Processing-in-Memory to speed up NGS analysis

Dominique Lavenier

► **To cite this version:**

Dominique Lavenier. Processing-in-Memory to speed up NGS analysis. SFA²F 2022 - Sequencing to Function: Analysis and Application for the Future, Jun 2022, Santa Fe, United States. hal-03817360

HAL Id: hal-03817360

<https://hal.science/hal-03817360>

Submitted on 18 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Processing-in-Memory to speed up NGS analysis

Dominique Lavenier
GenScale group
Univ Rennes, IRISA/CNRS, Inria, France

All current computing platforms are designed following the von Neumann architecture principles, originated in the 1940s, that separate computing units from memory. Processing-in-memory (PIM) consist of processing capabilities tightly coupled with the main memory. Contrary to bringing all data into a centralized processor, which is far away from the data storage, in-memory computing process the data directly where it resides, suppressing most data movement, and, thereby greatly improving the performance of massive data application by orders of magnitude.

NGS data analysis completely falls in these application domains where PIM can strongly accelerate the main time-consuming software in genomic and metagenomic areas. More specifically, mapping algorithms, intensive sequence comparison algorithms or bank searching, for example, can highly benefit of the parallel nature of the PIM concept.

New memory components based on PIM principles have been developed by the UPMEM company, a young startup created in 2015. The company has designed an innovative DRAM Processing Unit (DPU), a RISC processor integrated directly in the memory chip, on the DRAM die. An UPMEM PIM server counts no less than 2560 DPUs for 160 GB of PIM memory and 256 GB of legacy memory.

First experiments on the UPMEM PIM server have demonstrated that an average speed-up of X20 can generally be obtained on various time-consuming tasks of NGS pipelines.

The talk will detail the concept of Processing-in-Memory, which category of NGS algorithms can benefit of the PIM architecture, the way they can be parallelized on such structure, and what are the expected gains compared to a standard multicore platform.