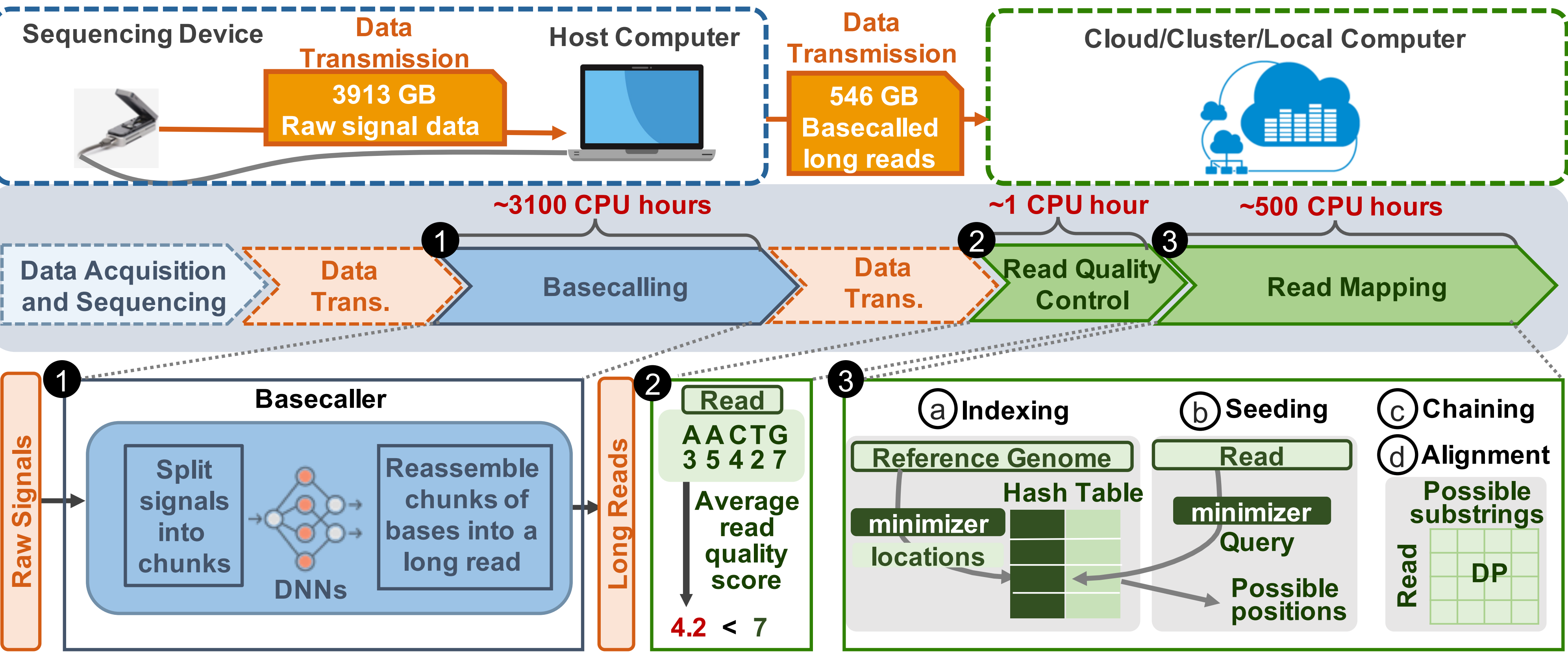


Haiyu Mao<sup>1</sup>, Mohammed Alser<sup>1</sup>, Mohammad Sadrosadati<sup>1</sup>, Can Firtina<sup>1</sup>,  
Akanksha Baranwal<sup>1</sup>, Damla Senol Cali<sup>2</sup>, Aditya Manglik<sup>1</sup>, Nour Almadhoun Alserr<sup>1</sup>, Onur Mutlu<sup>1</sup>

<sup>1</sup> ETH Zurich, <sup>2</sup> Bionano Genomics

## 1. Background and Motivation

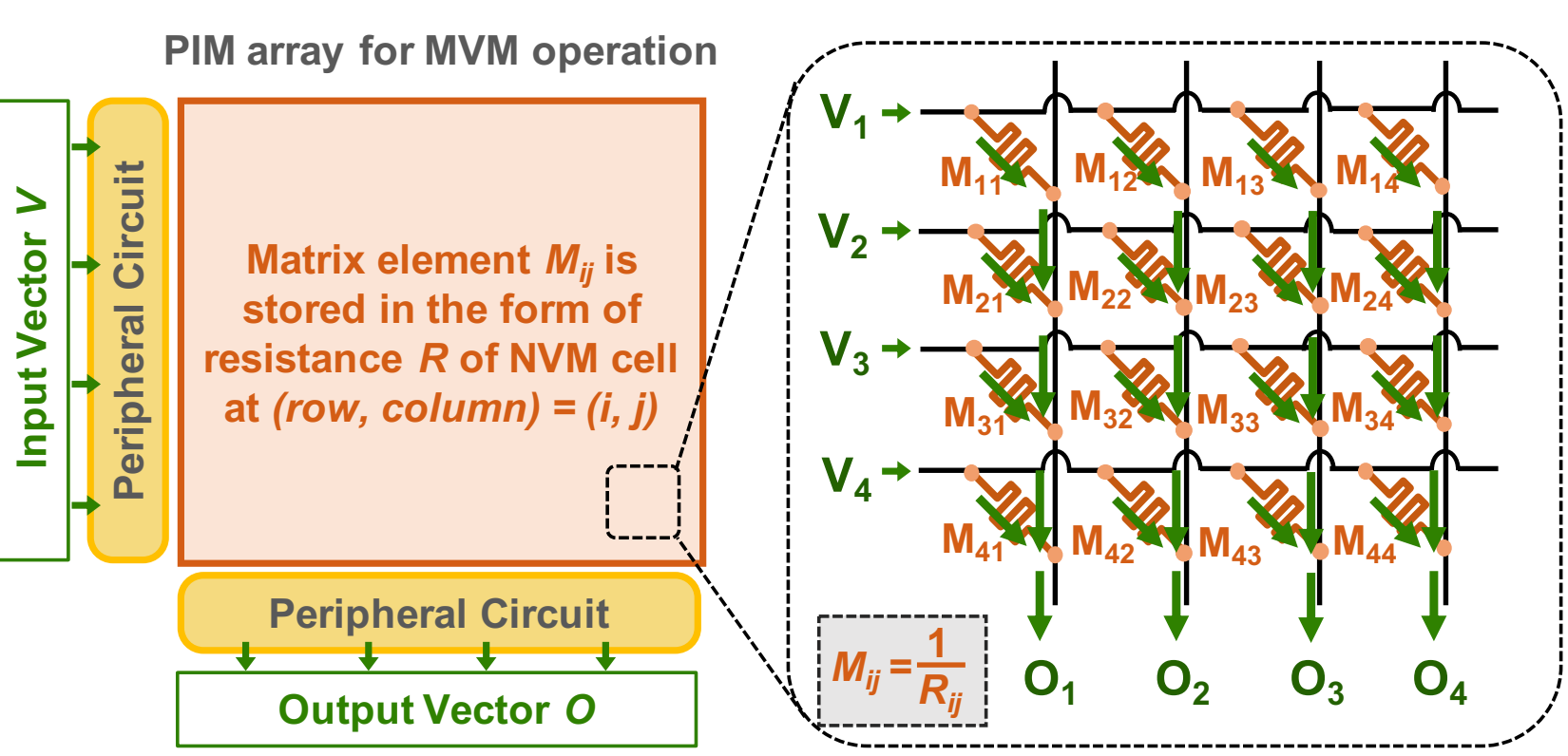
### ► The genome sequencing and analysis pipeline



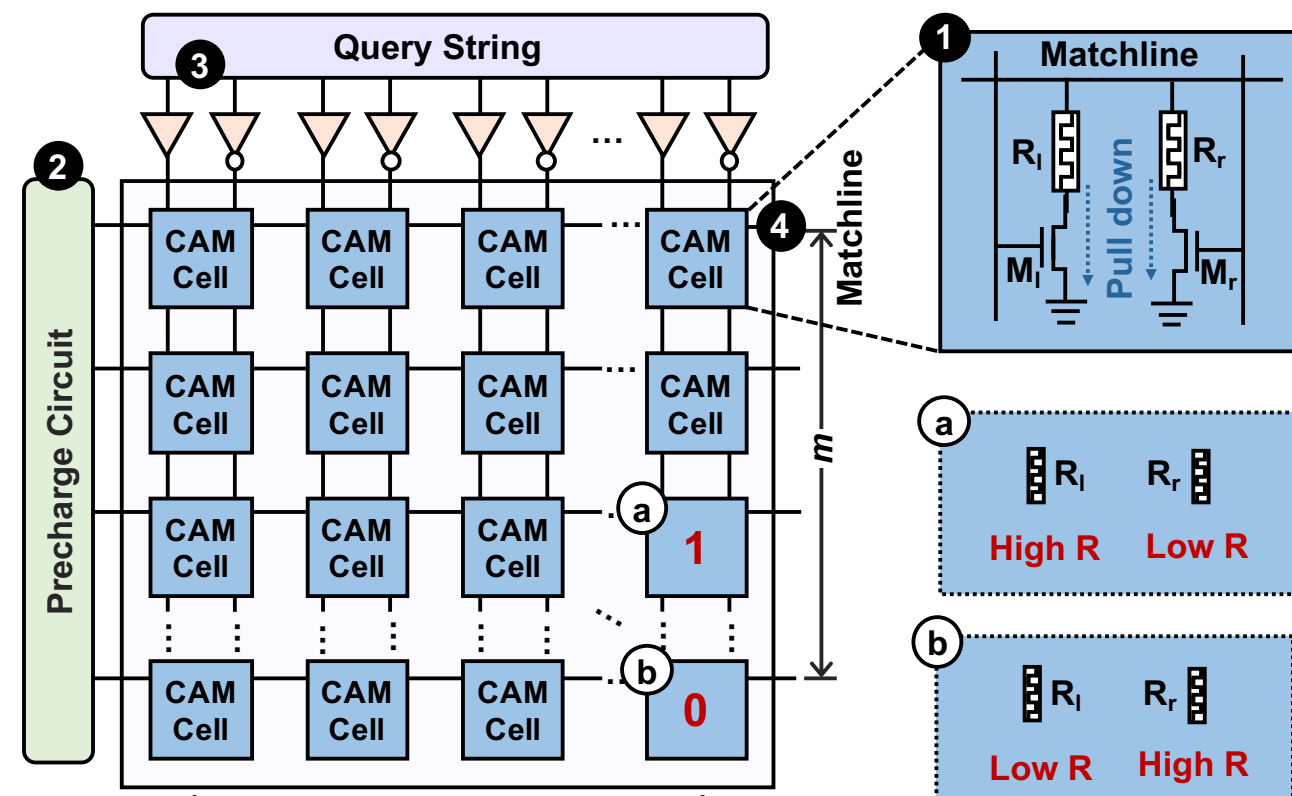
- The basecalling step ① and the read mapping step ③ are the **two most time-consuming steps** in the genome analysis pipeline
- The read quality control step ② is a highly-recommended but optional step to reduce the workload of read mapping by eliminating **unnecessary computation**
- The basecalling step and the read mapping step are **executed separately**

## 2. The state-of-the-art Solution

### ► NVM-based PIM accelerators



(a) NVM-based PIM array designed for Matrix-Vector-Multiplication (MVM) operation

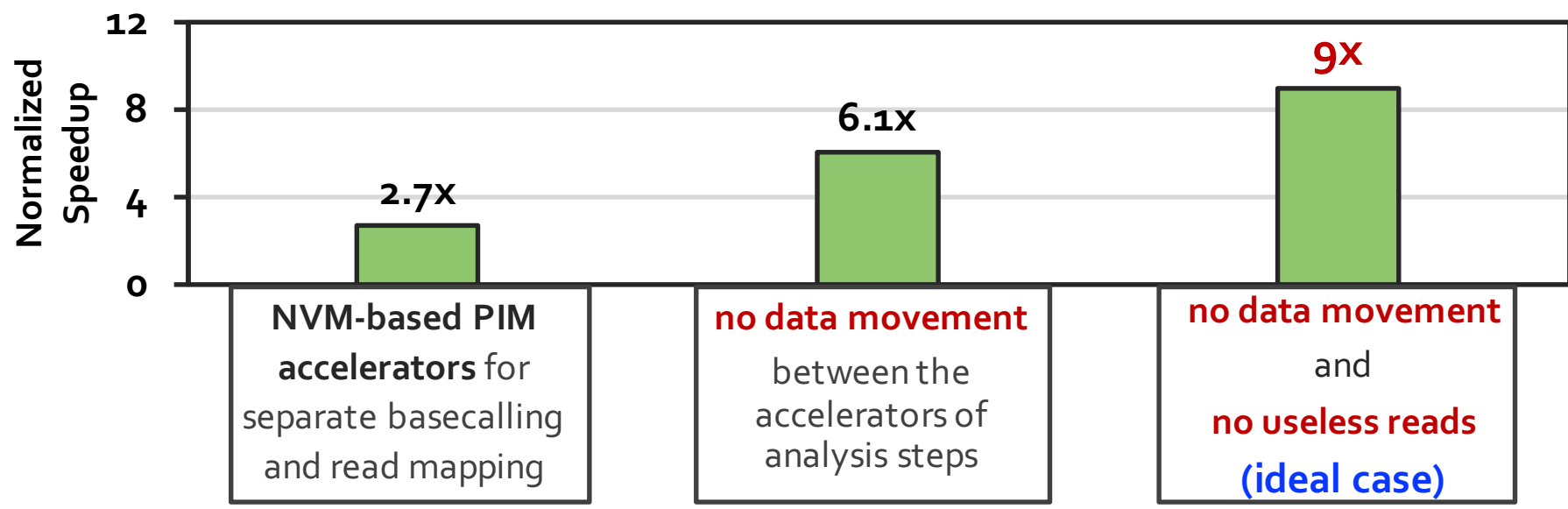


(b) NVM-based CAM array designed for string matching operation

- State-of-the-art works accelerate the basecalling and read mapping steps separately
- No prior system supports multiple key steps of the genome analysis pipeline

- Large data movement between multiple genome analysis steps
- A lot of wasted computation done on data that is later discovered to be useless

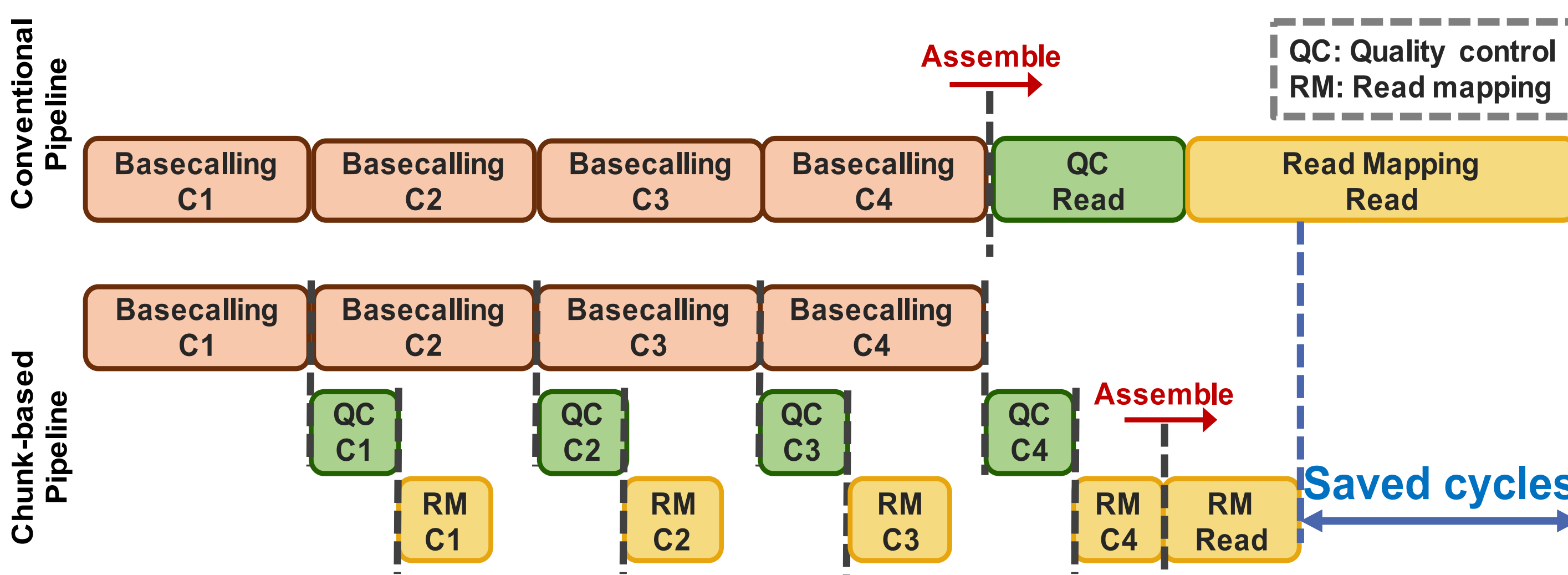
## 3. Potential Benefits



Goal:  
Efficiently accelerate the *entire* genome analysis pipeline while *minimizing data movement and useless computation*

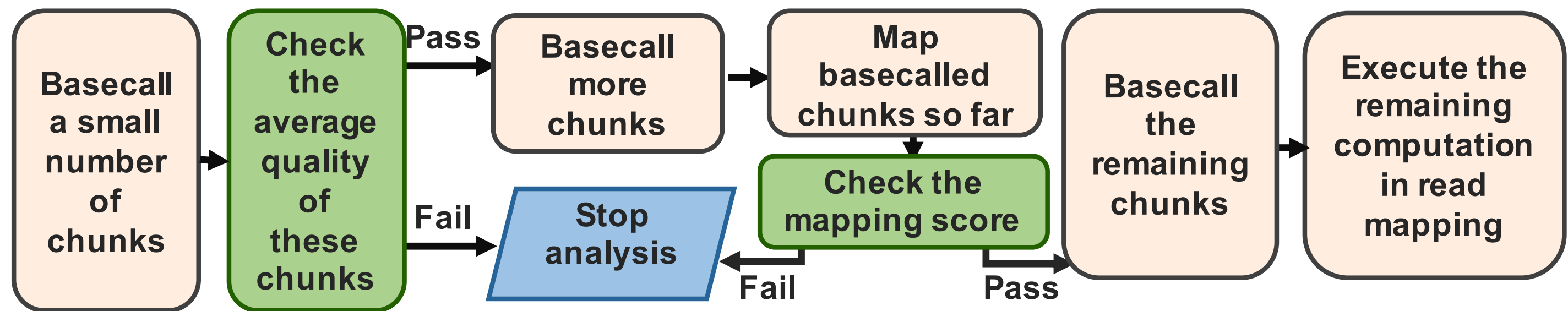
## 4.1. GenPIP: Chunk-based Pipeline

A read consists of four chunks: C1, C2, C3, C4



- Enables **fine-grained pipelining** of genome analysis steps
- Processes reads **at chunk granularity** (i.e., a subsequence; 300 bases)

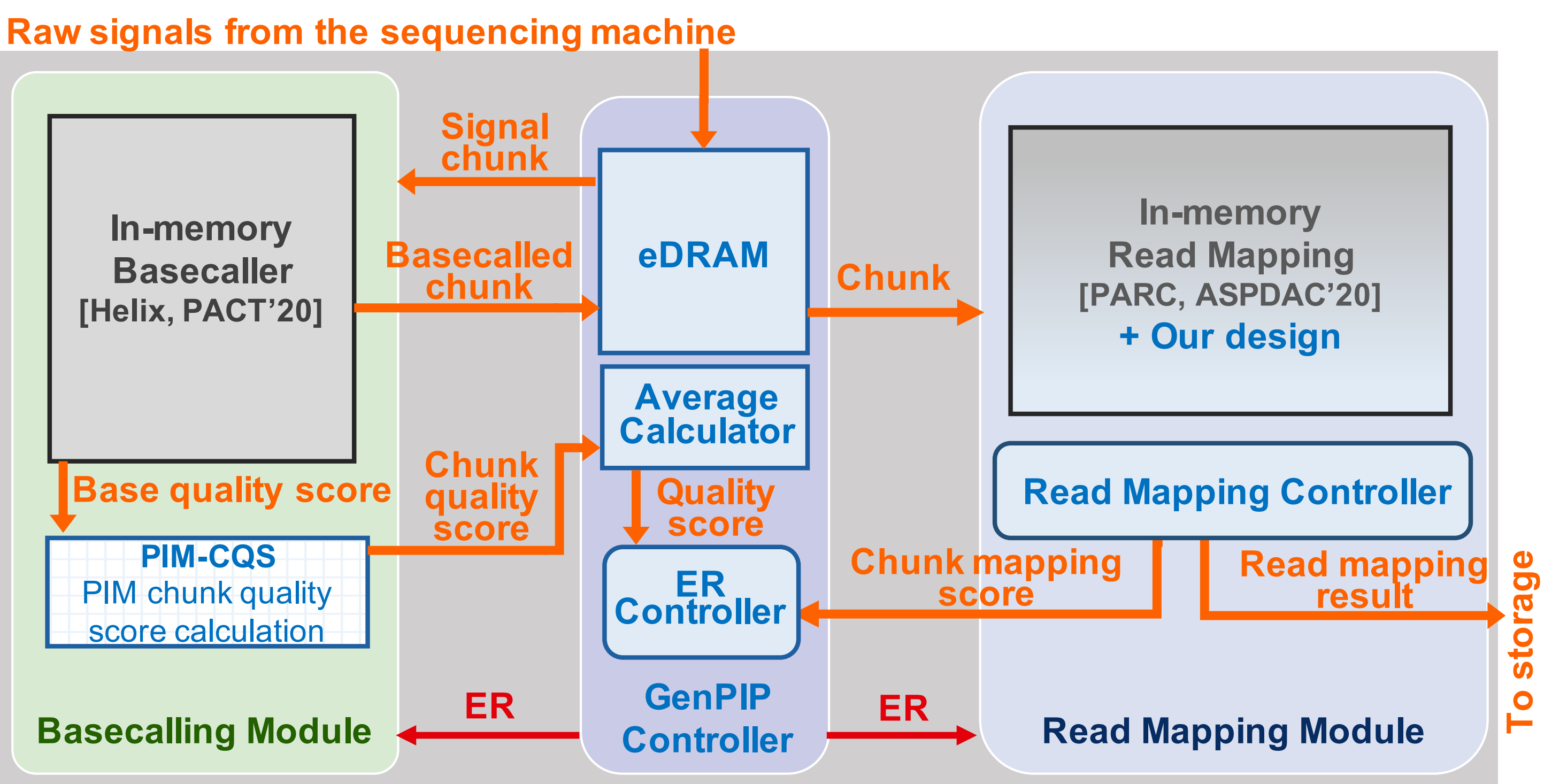
## 4.2. GenPIP: Early Rejection



- Stops the execution on **useless reads** as early as possible by using a small number of chunks to **predict the usefulness of a read**

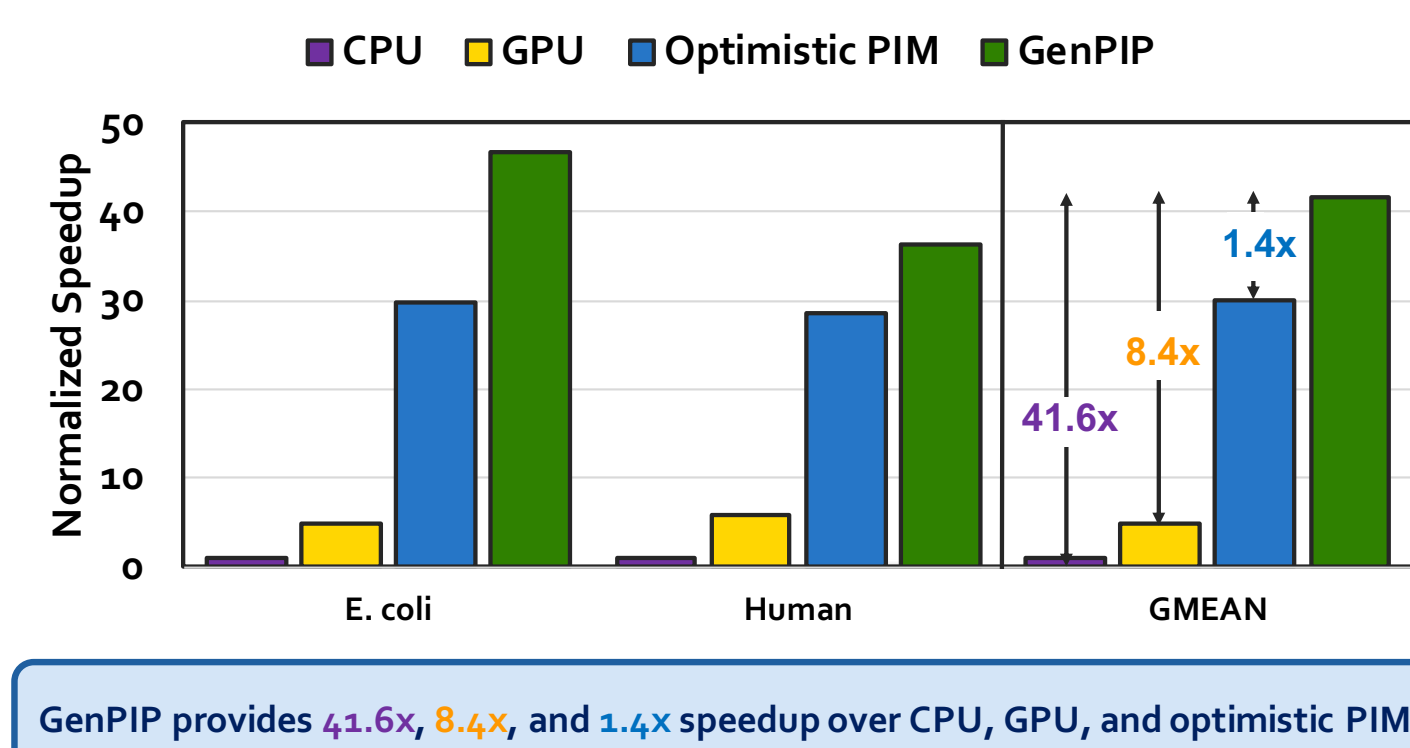
## 5. GenPIP Implementation

### ► NVM-based PIM implementation



## 6. Results

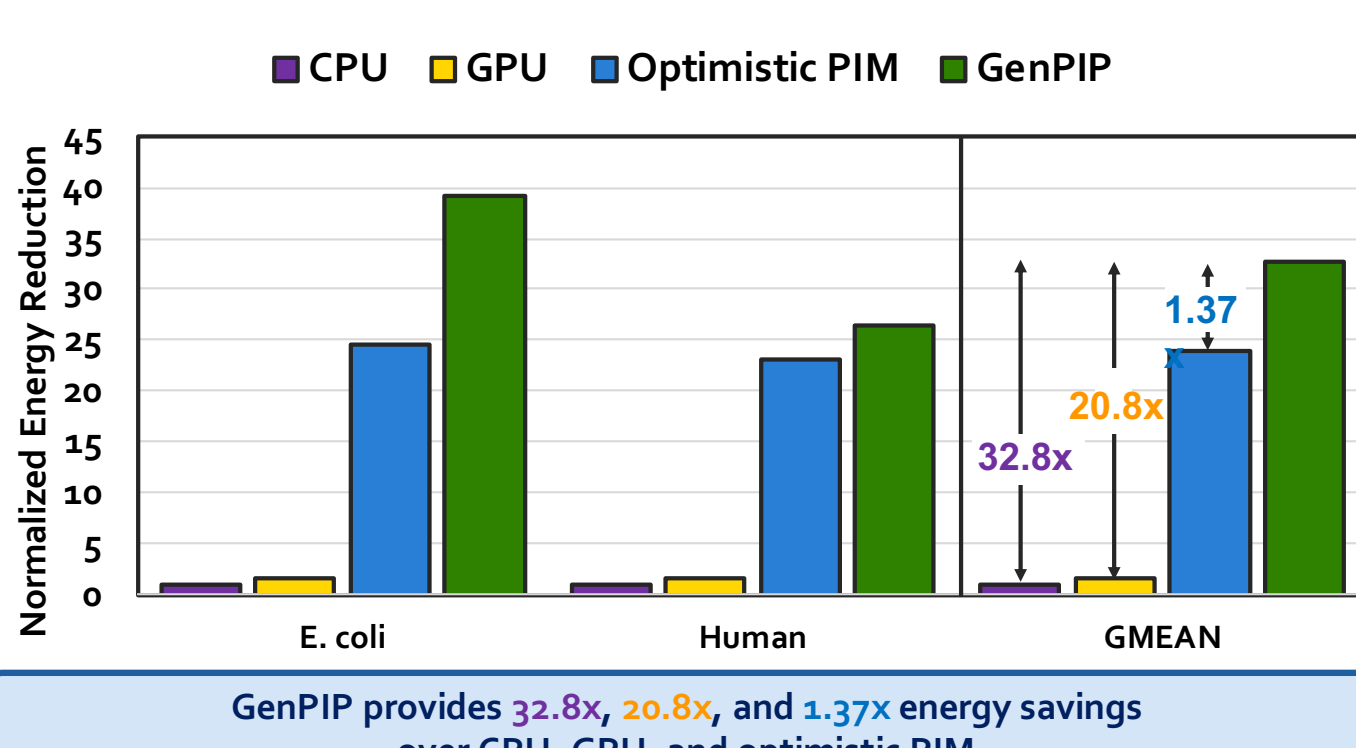
### ► Speedup



GenPIP provides 41.6x, 8.4x, and 1.4x speedup over CPU, GPU, and optimistic PIM

Both CP and ER are critical to the speedup

### ► Energy Efficiency



GenPIP provides 32.8x, 20.8x, and 1.37x energy savings over CPU, GPU, and optimistic PIM

ER is especially critical to the energy efficiency

## ACKNOWLEDGMENTS

We thank the SAFARI Research Group members for valuable feedback and the stimulating intellectual environment they provide. We acknowledge the generous gifts provided by our industrial partners, including Google, Huawei, Intel, Microsoft, and VMware. This research was partially supported by the Semiconductor Research Corporation and the ETH Future Computing Laboratory.



Talk Video



Full Paper



Contact Me