Hardware-accelerated detection and classification of pathogens of epidemic significance

Leonid Yavits
EnICS Labs
Bar Ilan University
Primary Research Directions

- SoC/RISC-V Architecture
- Hardware Security
- Low Power Digital/Analog
- Image Sensors
- Hardware accelerators
- Quantum Computing
- Bio-convergence
- and many others...
Motivation: Viral pandemic

• Imagine that in 2030, the world is struck by a new viral pandemic
  • The virus mutates quickly, creating many hundreds of viral lineages
• Existing diagnostic tools are inadequate for variant discovery
  • PCR requires specifically designed per-lineage primers
• Identifying the lineage of a viral sample is practical only through
  • DNA sequencing and computational pangenomic analysis
• This is a problem of a very large scale
  • 100M are tested daily, of which 10M test positive, of which 1M is sequenced
  • Each one of these 1M daily sequenced genomes needs to be analyzed in a very wide context
Motivation: Sepsis

• An operation is successful but a patient whose immune system is weakened by operation-related treatments develops blood infection
• Sepsis causes 20–50% of hospital deaths in the United States alone
• Early microbial pathogen identification is crucial
  • In over 30% of cases the underlying pathogen is never even identified
• As a result, antibiotic treatment is typically not pathogen-targeted
  • Clinical decision-making is based on epidemiological information rather than individual patient data
  • Clinicians often continue antimicrobial treatment despite negative microbiologic testing

Motivation: AMR

- **Antimicrobial resistance** is an escalating global crisis
  - The spread of drug resistance is outpacing the development of new antimicrobials
  - Infections due to AMR bacteria are more difficult and expensive to treat
  - WHO has declared that AMR is one of the top 10 global public health threats facing humanity
  - AMR already causes at least 700,000 deaths a year and is expected to cause over 10 million deaths annually by 2050
Accurate pathogen detection and classification is a **time-critical** problem.

It is also a problem of **a very large scale**, especially for viral pathogens with pandemic potential.

Traditional methods of classification are either **prohibitively slow** or **inaccurate**.

- Including the computational ones.
And the solution:
Real-time hardware-accelerated pathogen classifier

• Input: Unaligned DNA reads of a metagenomic sample

• Output:
  • Detection of specific pathogens or pathogen lineages
  • Pathogen profiling

• Requirements
  • Real time operation (for example, at the speed of a sequencer)
  • High accuracy, sensitivity and precision
  • Portable implementation (low power, small form factor)
Hamming Distance tolerant Content Addressable Memory (HD-CAM) based pathogen identifier

CAM allows (in addition to read and write):
To compare a query pattern with the contents of the entire memory in a single cycle.
Edit Distance tolerant (EDAM) based pathogen classifier

- DNA reads are prone to edits
- EDAM tolerates Edit rather than Hamming distance
- Highly sensitive to high error and low quality (ambiguous) DNA reads
- Difficult to scale
  - Solution: high-density resistive memory
Please consider contributing to and participating in AACBB

In conjunction with ISCA
Orlando Florida June 18
in parallel with ETH PiM tutorial

https://aacbb-workshop.github.io/