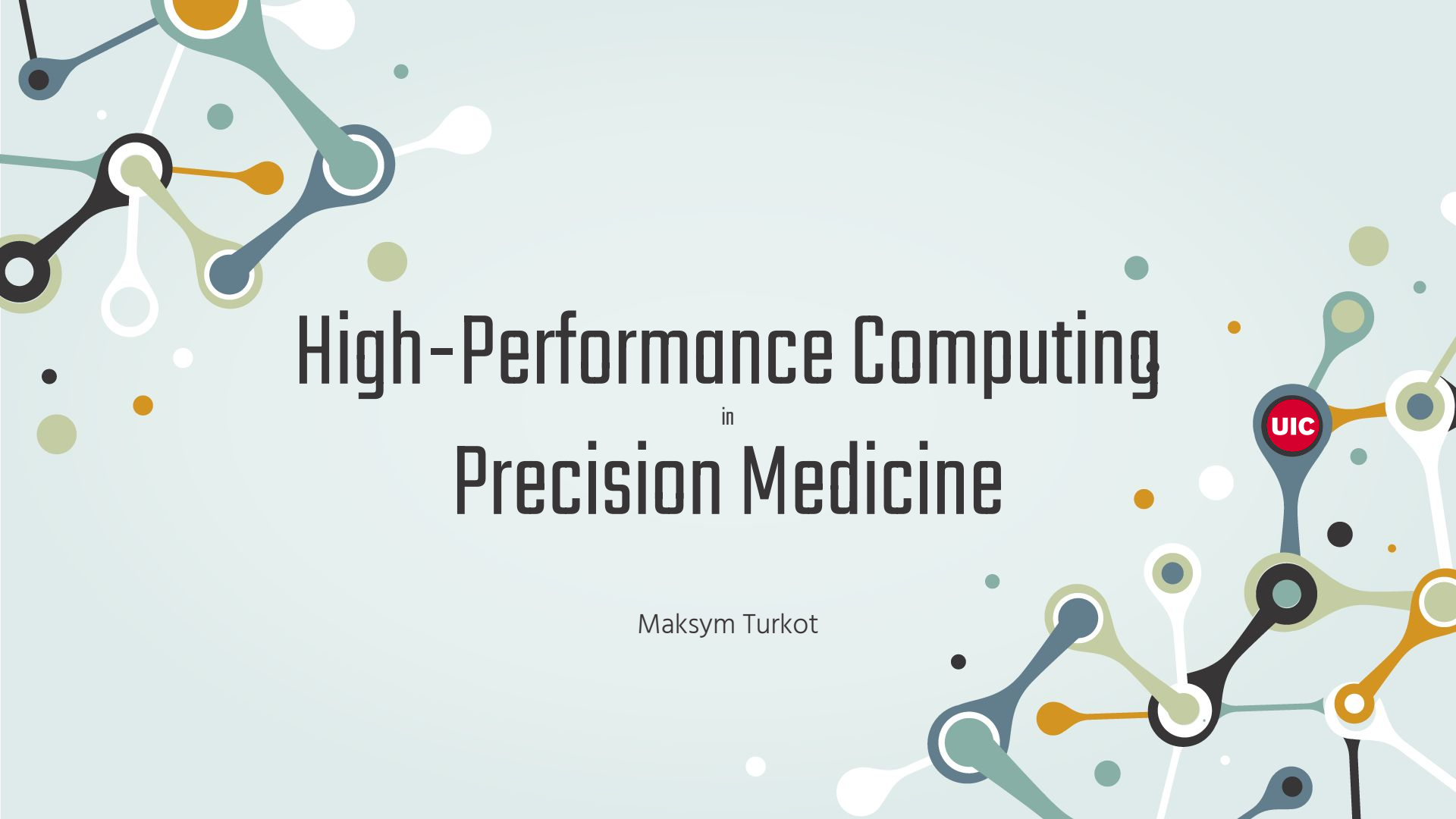
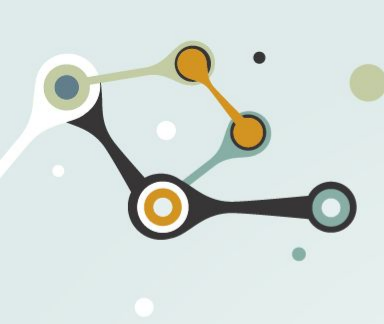


An abstract graphic featuring a network of interconnected nodes and lines, resembling a molecular structure or a data network. The nodes are represented by circles in various colors (blue, green, orange, black, white) and are connected by lines of different colors (blue, green, orange, black, white). The background is a light blue-grey color. The graphic is positioned on the left and right sides of the slide, framing the central text.

High-Performance Computing in Precision Medicine

Maksym Turkot



OUTLINE

01

INTRODUCTION

Motivation and importance

02

BACKGROUND

Genomic analysis, data generation, read mapping

03

INDEXING & MOVEMENT

Indexing acceleration and data movement

04

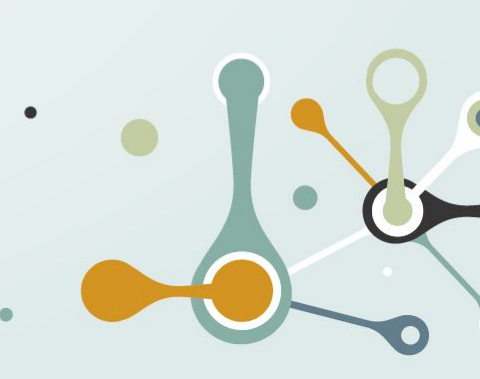
GPU ACCELERATION

NVIDIA Parabricks and benchmarks

05

CLOUD COMPUTING

Collaboration and performance



An abstract graphic design featuring a central orange circle with the white number '01'. This circle is connected by a teal line to a larger teal shape. To the left, a dark grey shape contains an orange circle, with a teal line connecting it to the central orange circle. Various other organic shapes in teal, orange, and white are scattered around the central elements.

01

INTRODUCTION

Introduction



- HPC in Precision Medicine
- Shifting Bottleneck in Genomic Research
- Acceleration Techniques
 - Algorithmic (indexing)
 - Hardware (GPUs)
- Cloud Computing in Genomic Workflows

BACKGROUND

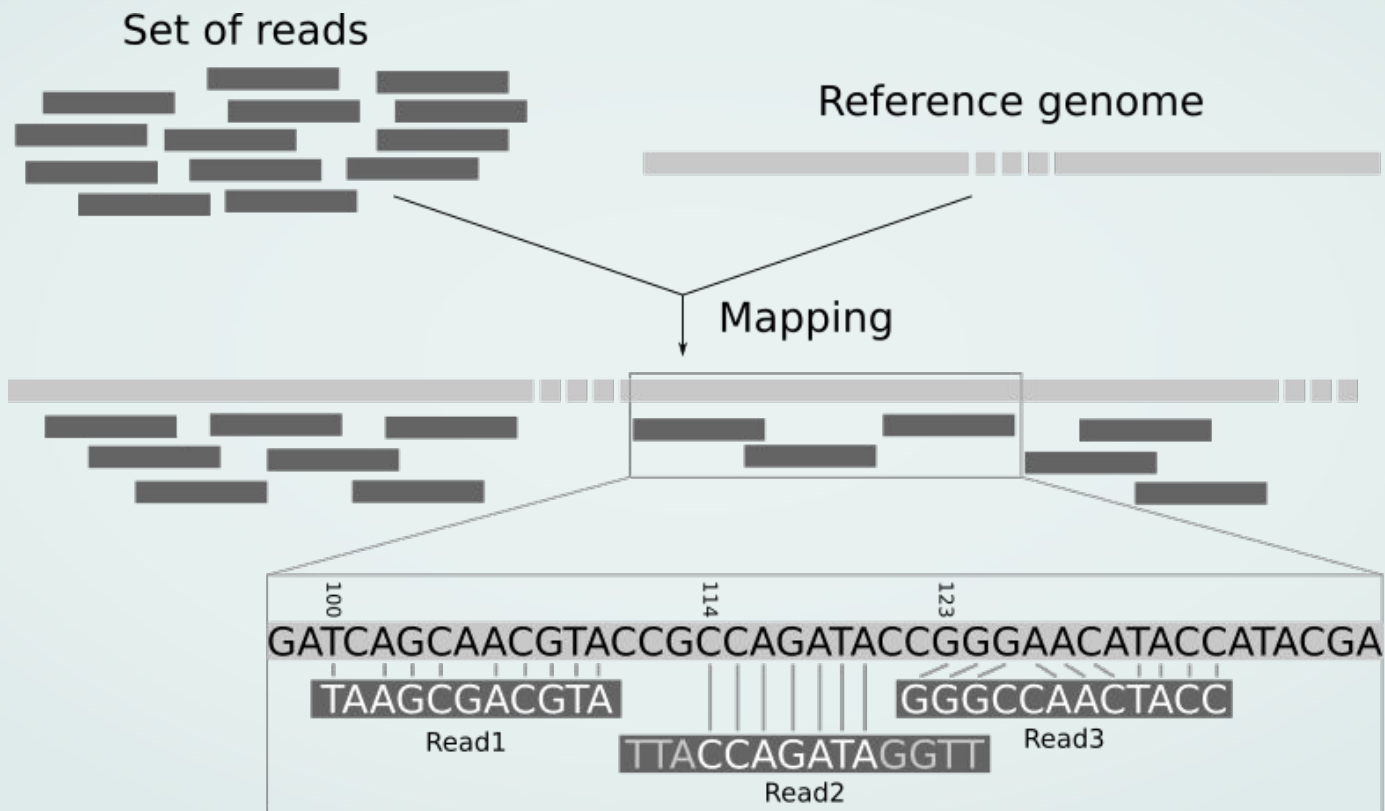


Obtaining Genomic Data

- Generation
 - Machine-based, sequencer
- Download
 - Prepared data
- Simulation
 - Mimic sequencer. Useful



Read Mapping

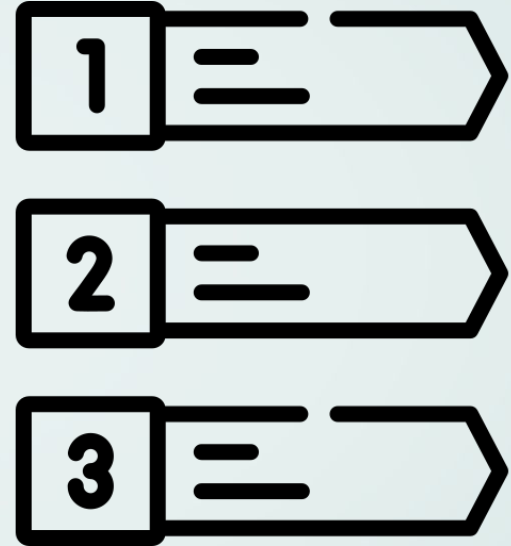


INDEXING & MOVEMENT



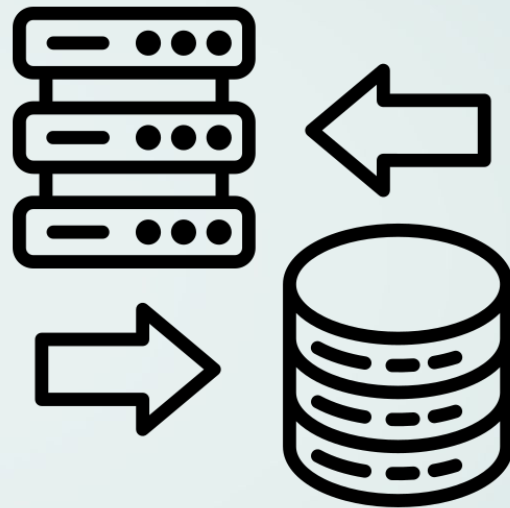
Indexing and Seeding Acceleration

- Indexing and Seeding
- Indexing Process
- Impact of Seed Characteristics
- Tradeoff and Improvements



Data Movement Reduction

- Memory-Intensive Tasks
- Processing-in-Memory (PIM)
- Alternative Approaches
- 3D ReRAM-based Memory



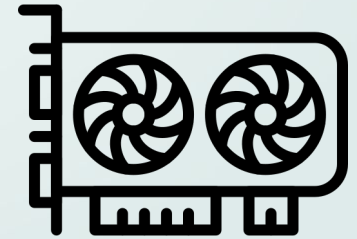
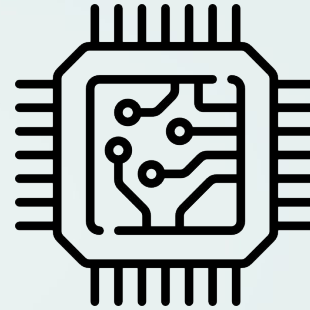
An abstract graphic design featuring organic, flowing shapes in orange, olive green, and dark grey. A central dark grey shape contains a circle with an orange border, which in turn contains a smaller olive green circle with the white number '04'. Other shapes include a large orange shape with an olive green circle, a white shape with a blue-green circle, and various smaller circles and teardrop shapes in the same color palette.

04

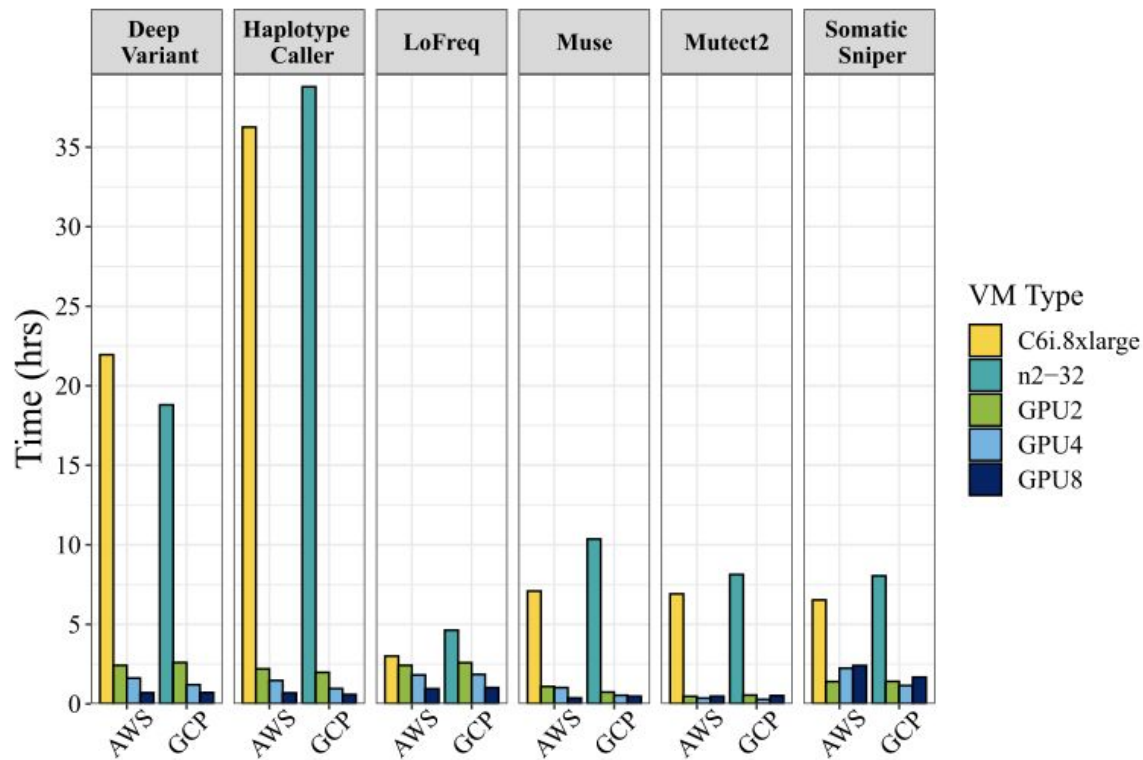
GPU ACCELERATION

GPU Acceleration

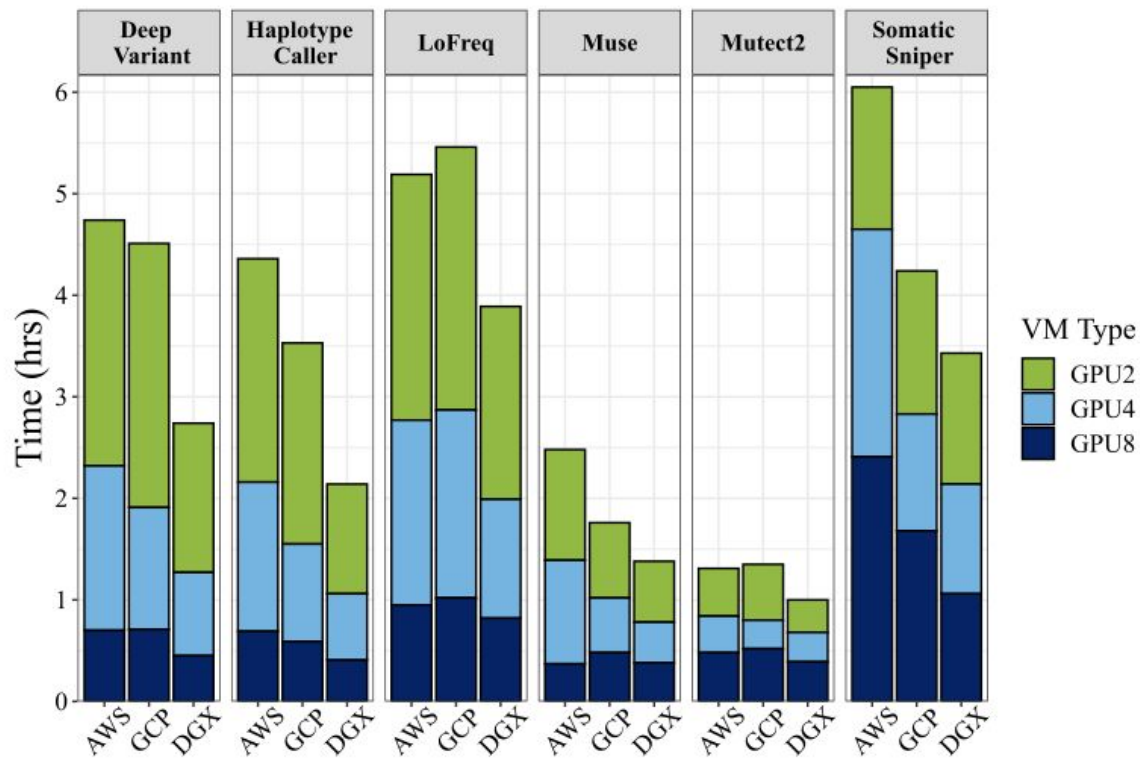
- Data Analysis Bottleneck
- NVIDIA Parabricks
- Varying CPU performance
- Massive GPU acceleration



Runtimes of All Cloud-Based Analyses



GPU Runtimes Across Platforms





05

CLOUD COMPUTING

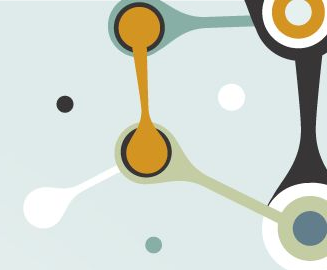
Cloud Computing

- An Invaluable Tool
- Reproducibility & Accessibility
- Simplified setup
- Collaboration
- Advantage in genome work
- Scaling





References



- [1] M. Alser, Z. Bing¹ ol, D. S. Cali, J. Kim, S. Ghose, C. Alkan, and O. Mutlu, “Accelerating genome analysis: A primer on an ongoing journey,” *IEEE Micro*, vol. 40, no. 5, pp. 65–75, 2020.
- [2] A. Taylor-Weiner, F. Aguet, N. J. Haradhvala, S. Gosai, S. Anand, J. Kim, K. Ardlie, E. M. Van Allen, and G. Getz, “Scaling computational genomics to millions of individuals with GPUs,” *Genome Biology*, vol. 20, pp. 1–5, 2019.
- [3] K. A. O’Connell, Z. B. Yosufzai, R. A. Campbell, C. J. Lobb, H. T. Engelken, L. M. Gorrell, T. B. Carlson, J. J. Catana, D. Mikdadi, V. R. Bonazzi et al., “Accelerating genomic workflows using NVIDIA Parabricks,” *BMC Bioinformatics*, vol. 24, no. 1, p. 221, 2023.
- [4] B. Langmead and A. Nellore, “Cloud computing for genomic data analysis and collaboration,” *Nature Reviews Genetics*, vol. 19, no. 4, pp. 208–219, 2018.
- [5] M. Alser, J. Lindegger, C. Firtina, N. Almadhoun, H. Mao, G. Singh, J. Gomez-Luna, and O. Mutlu, “From molecules to genomic variations: Accelerating genome analysis via intelligent algorithms and architectures,” *Computational and Structural Biotechnology Journal*, vol. 20, pp. 4579–4599, 2022.

Thank You!

Question time

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

