

Accelerating irregular bioinformatics algorithms on GPUs.

Muaaz Gul Awan

Application Performance Group

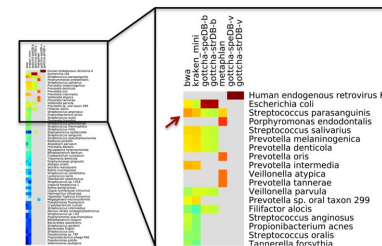
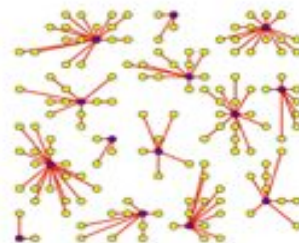
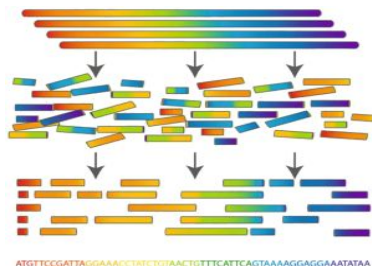


Outline

- Exabiome Project:
 - Common algorithmic motifs.
 - Irregular memory access patterns among bio-algorithms.
- ADEPT, a GPU accelerated sequence aligner:
 - Implementation on GPUs.
 - Results and Integration in exabiome software stack.
- GPU Local Assembly:
 - Local Assembly in MetaHipMer.
 - Challenges and implementation on GPUs.
 - Integration and results.

ExaBiome: Project Overview

Exascale algorithms & systems for previously intractable problems

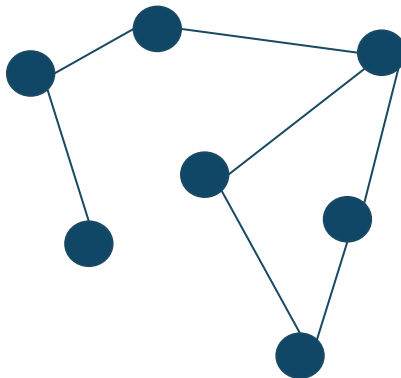


Problem Domain	Metagenome Assembly	Protein Clustering and Annotation	Comparative Analysis
Science need	Find species, genes and relative abundance in microbial communities	Improve understanding of tree of life for microbes; aid in identifying gene function	Track microbiome over time or space, changes in environment, climate, etc.
Computing techniques	hash tables, alignment, k-mer counts, graph walks	Hash tables tables, alignment, k-mer counts, sparse matrices, ML (clustering, GNNs)	hash tables, alignment, k-mer counts, ML (dimensionality reduction)

Motifs of Bio Data Analysis

Short Read Assembly:

- Dynamic Programming Algorithms.
- Distributed and local Graph Traversals.
- Distributed and local hash tables.



Long Read Assembly:

- Dynamic Programming Algorithms.
- Sparse matrix multiplication.
- Distributed hash tables.

Protein Similarity and Clustering:

- Dynamic Programming Algorithms.
- Sparse Matrix Multiplication.

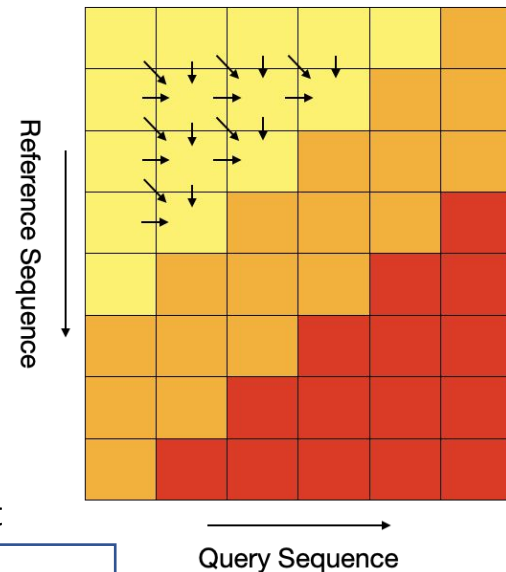
ATGCATG

ATGCATG

ATGCATG



kmers	Ext
TGCA	T
GCAT	G
CATG	C



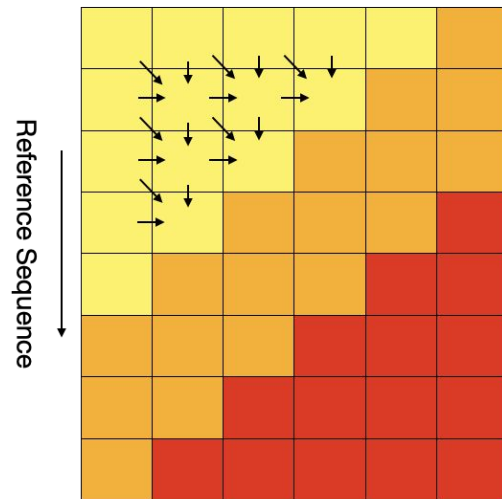
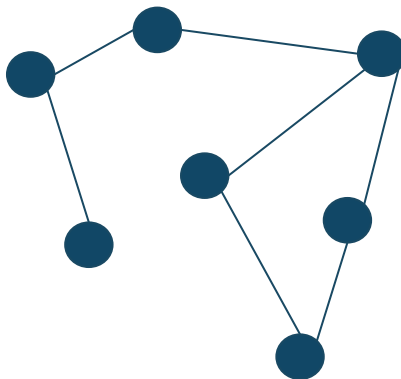
Challenges on GPUs

Ideal for GPUs:

- Localized and predictable memory access pattern.
- A lot of computations per each memory access.
- Equal amount of work can be distributed across threads.

What we have:

- Random or along diagonals memory access pattern.
- Integer only computations bound by memory bandwidth.
- Non deterministic amount of work.
- Varying or limited parallelism (DP and graph algorithms).



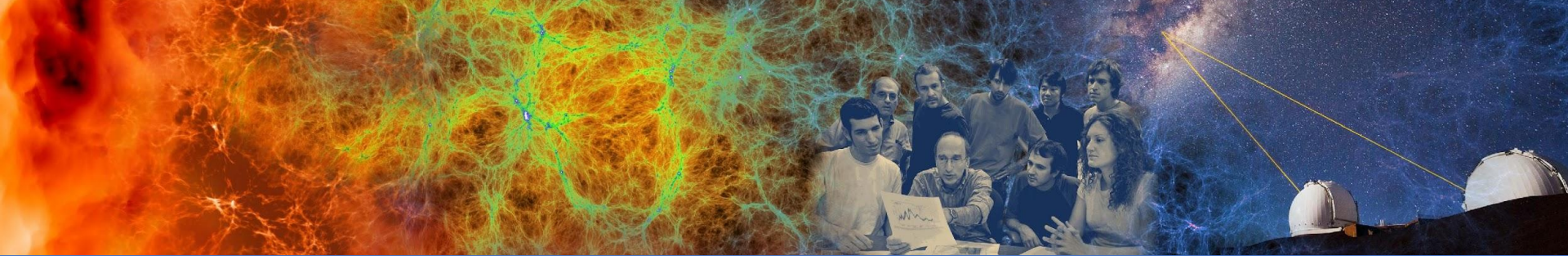
ATGCATG

ATGCATG

ATGCATG

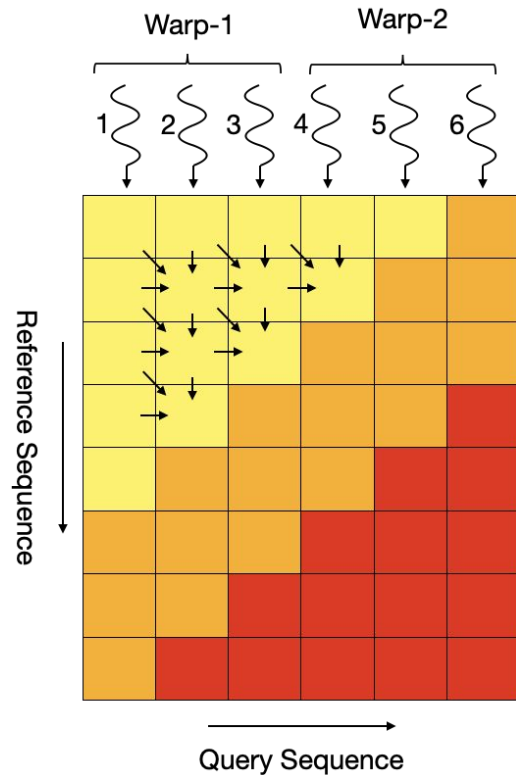


kmers	Ext	Query Sequence
TGCA	T	
GCAT	G	
CATG	C	

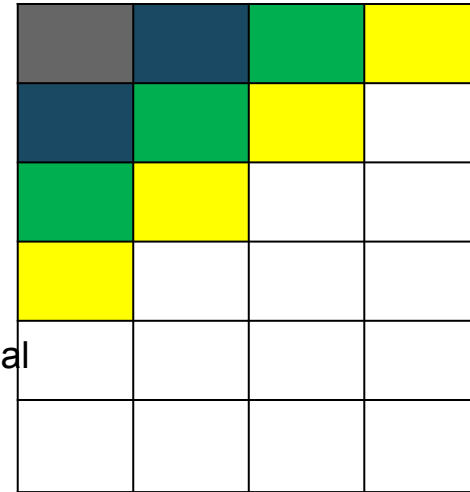


Sequence Alignment

*ADEPT: A GPU Accelerated Sequence Aligner

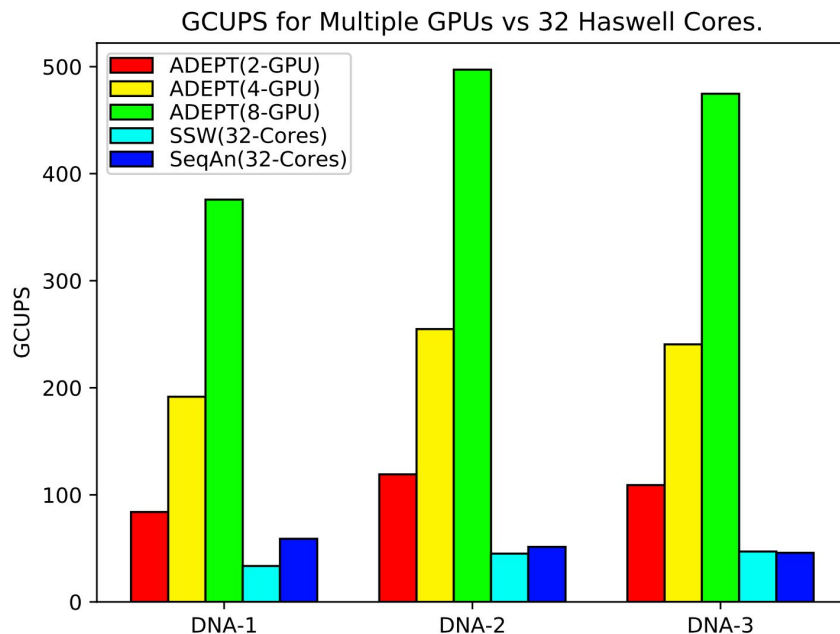


Diagonal Major Indexing
allows for coalesced
memory access in the global
memory



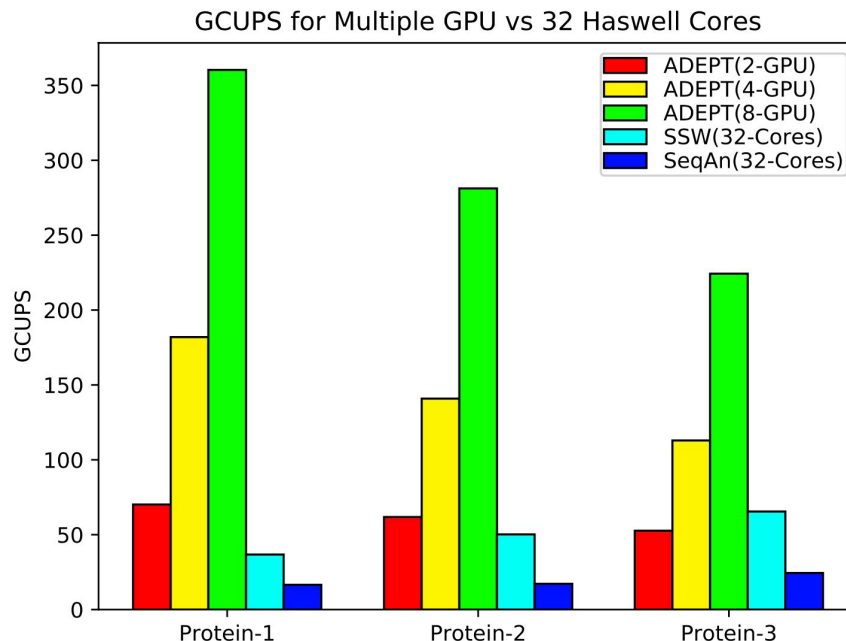
ADEPT's performance on DNA

Dataset	Query Set		Reference Set		Total Alignments
	Min. Size	Max. Size	Min. Size	Max. Size	
DNA-1	150	200	99	779	31,071,476
DNA-2	201	250	99	979	8,892,748
DNA-3	251	300	99	1,131	16,308,186



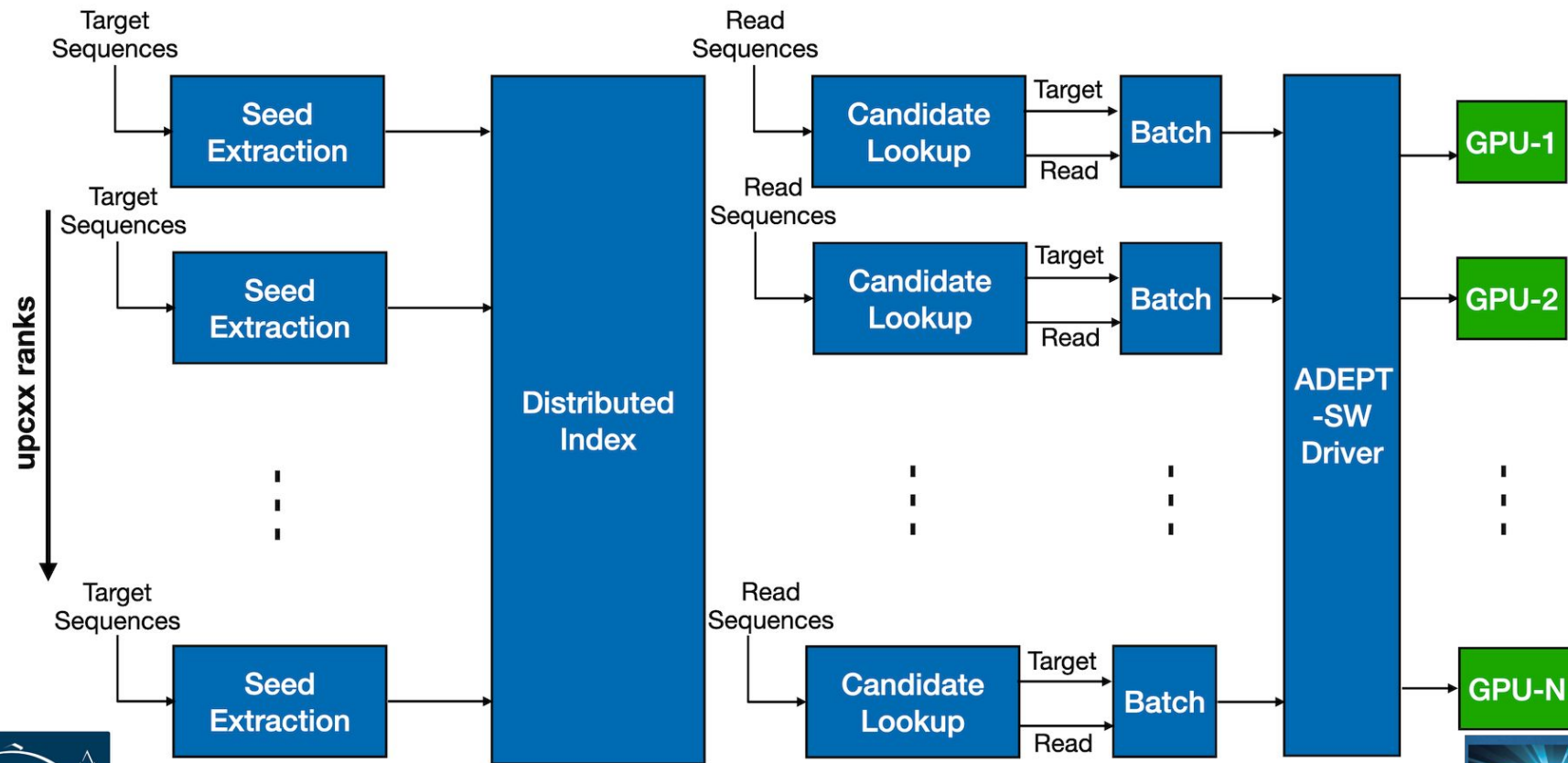
ADEPT's performance on Protein

Dataset	Query Set		Reference Set		Total Alignments
	Min. Size	Max. Size	Min. Size	Max. Size	
Protein-1	20	200	200	1,664	31,846,093
Protein-2	20	400	400	1,664	38,610,219
Protein-3	20	600	600	1,664	12,148,680

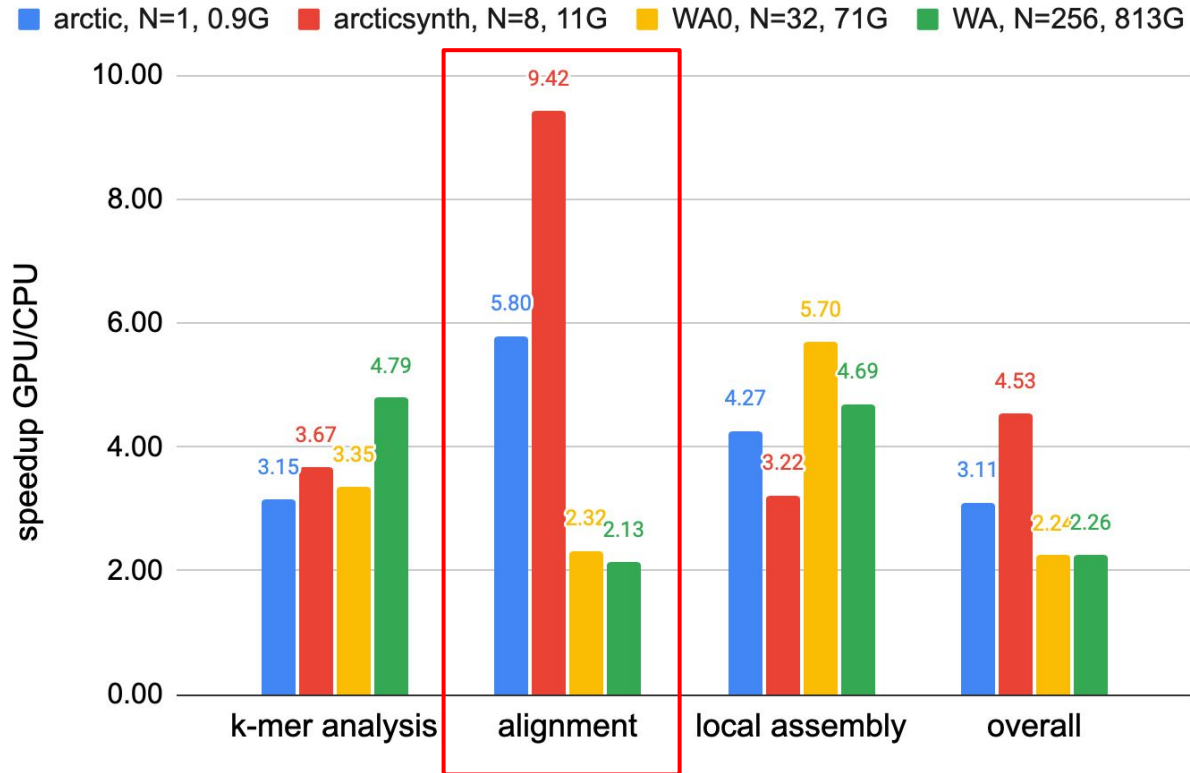


*Fox, Naomi K., Steven E. Brenner, and John-Marc Chandonia. "SCOPe: Structural Classification of Proteins—extended, integrating SCOP and ASTRAL data and classification of new structures." *Nucleic acids research* 42.D1 (2014): D304-D309.

ADEPT in MetaHipMer

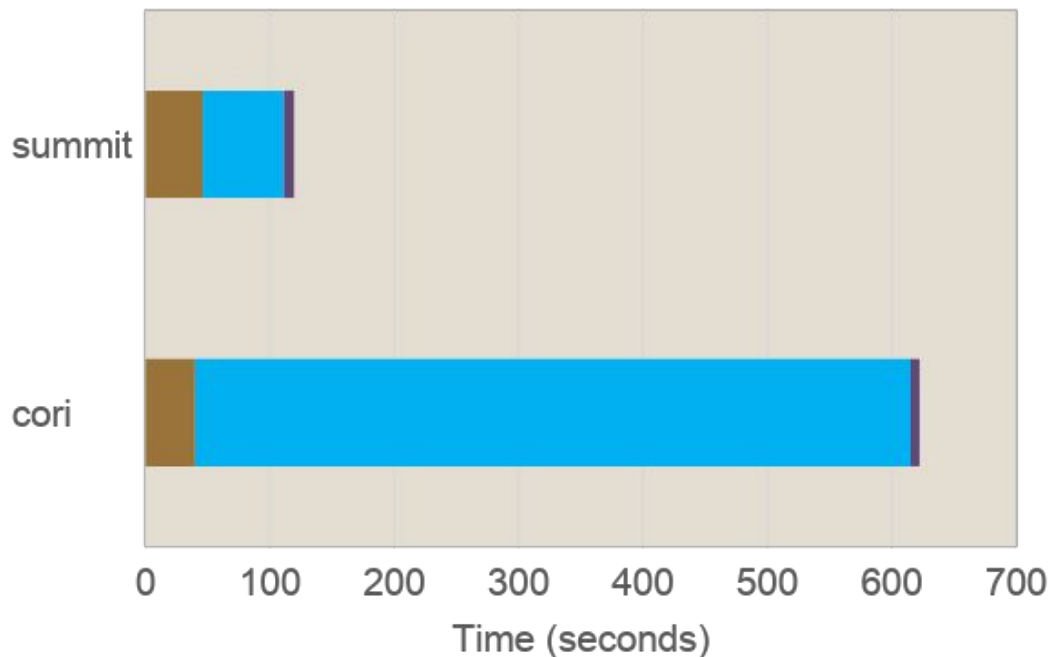


MetaHipMer Comparison



ADEPT in PASTIS

GPU-enabled PASTIS



■ sparse ops ■ alignment ■ other (I/O, wait, etc.)

Dataset size: 5 million protein sequences

9.9 billion candidate alignments

~853 Million alignments

Protein similarity graph

▷ 5 Million nodes

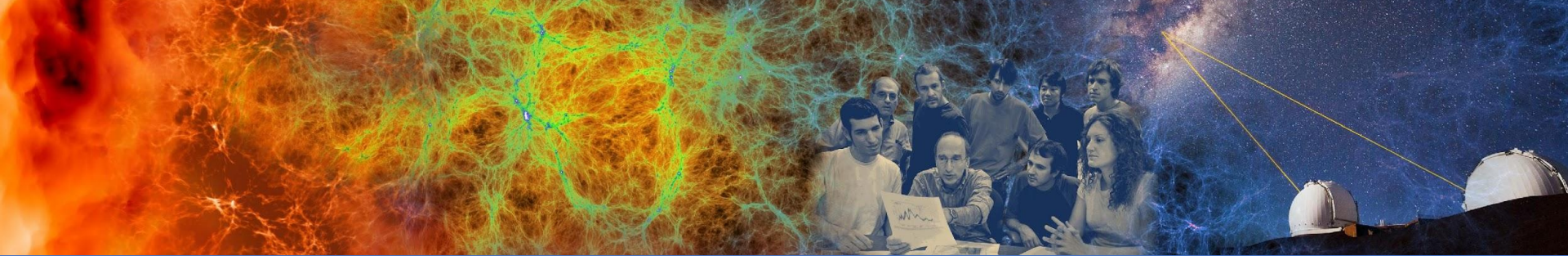
▷ 64.6 million edges

100 nodes of NERSC Cori and Summit

▷ ADEPT

▷ SeqAn

Alignment: 5.2x speedup



Local Assembly



BERKELEY LAB



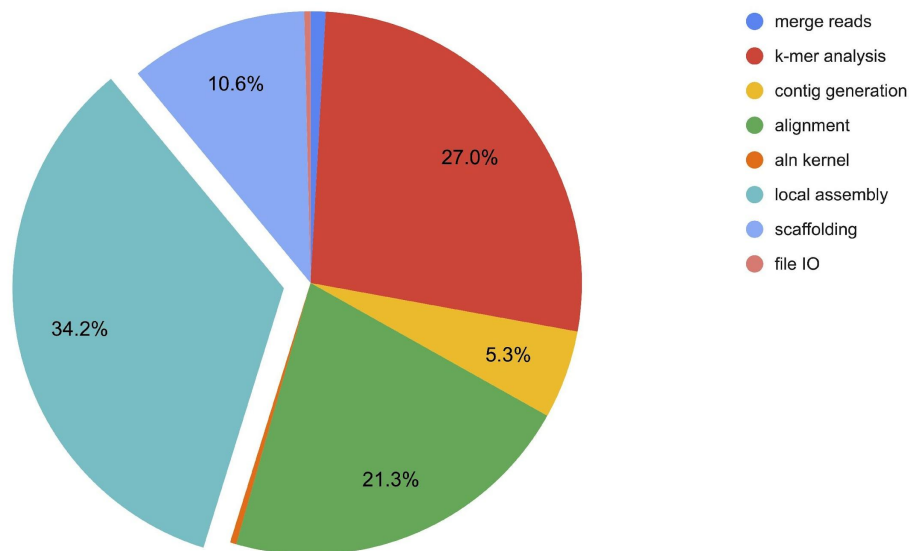
U.S. DEPARTMENT OF
ENERGY

Office of
Science

MetaHipMer Time Breakdown

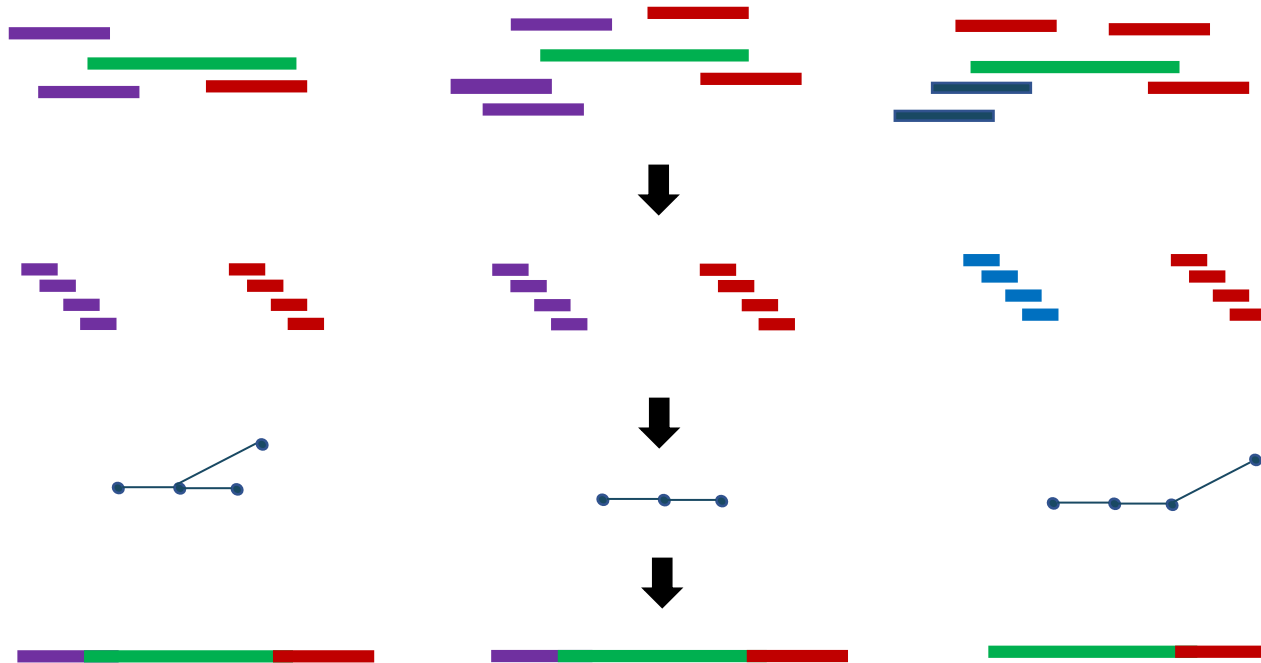
- 64 nodes of Summit system using the marine communities dataset*
- Total time 2128 seconds.
- Local assembly module takes the most amount of time.

MetaHipMer profile (CPU local assembly)



*<https://gold.jgi.doe.gov/biosamples?id=Gb0192059>

Local Assembly



Reads aligning to ends of contigs are obtained from the alignment stage

Reads are broken down into kmers to construct De Bruijn graphs

De Bruijn graphs are traversed to extend contigs

Building kmer hashtable

contig

ATGCATGCATGCA

reads

ATGCATG
ATGCATG
ATGCATG



kmers

Ext

TGCA	T
ATGC	A
GCAT	G
CATG	C

DNA Walks

ATGCATGCA

ATGCATGCA

ATGCATGCAT

ATGCATGCATG

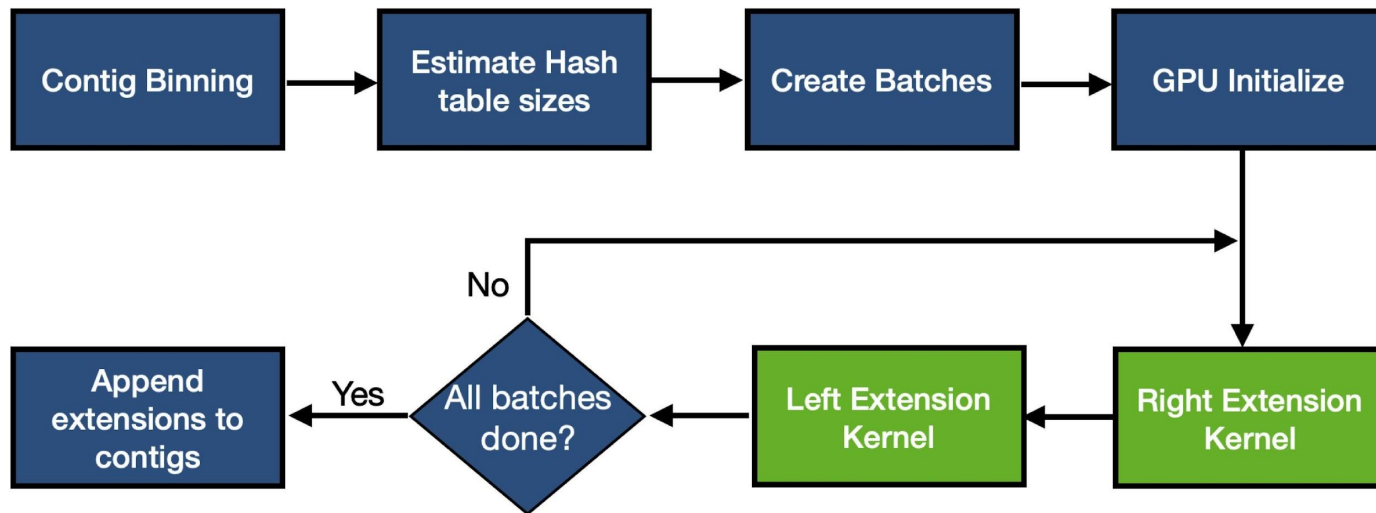
ATGCATGCATGC

kmers	Ext
TGCA	T
GCAT	G
CATG	C

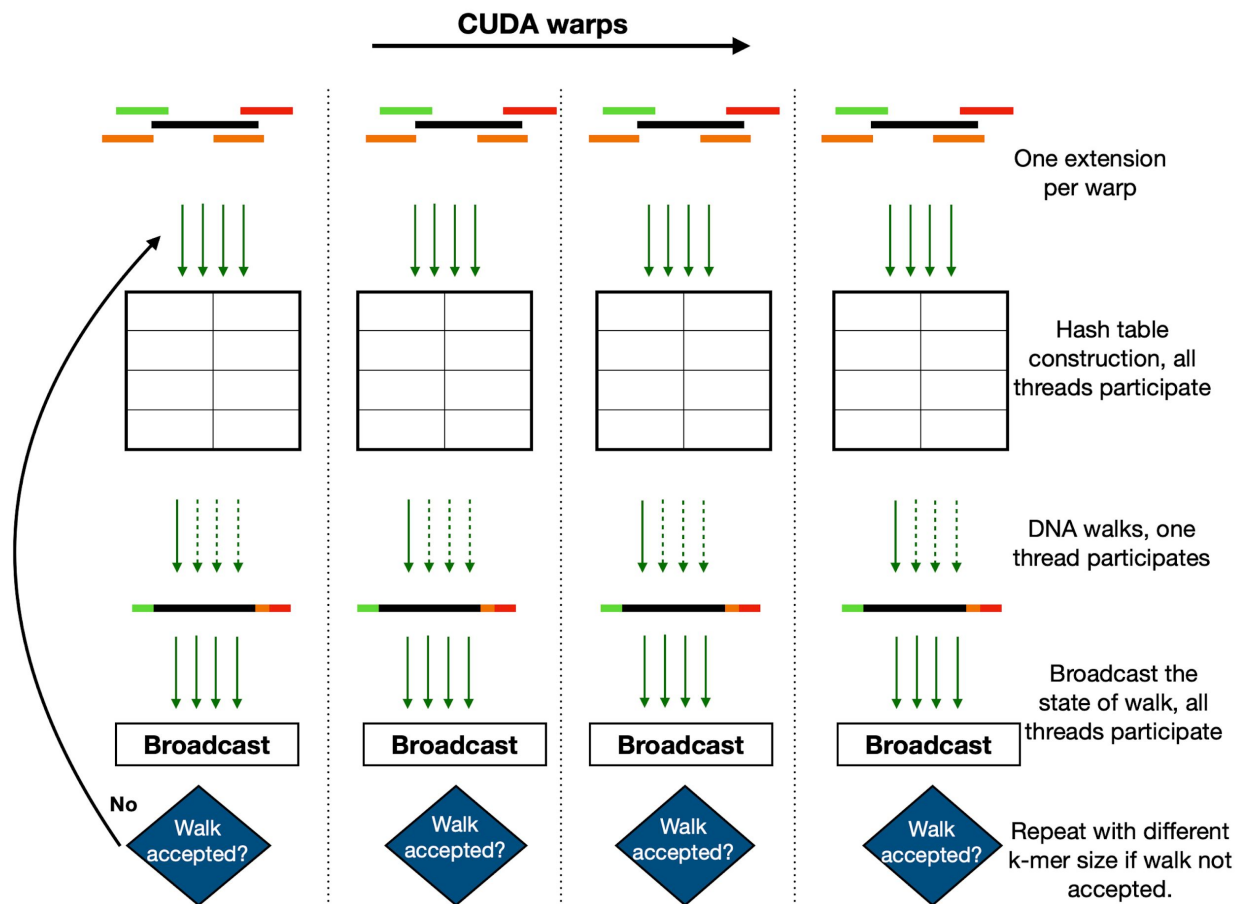
Implementation Challenges on GPU

- Not a typical GPU problem
- No dynamic memory allocation on GPUs and no support for STL containers.
- Length of walks is non-deterministic.
- Static memory allocation can be used.
- Accurate memory usage needs to be known before kernel launch time.

GPU Local Assembly



Local Assembly Kernel



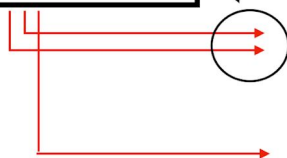
Local Assembly Kernel

Hash collision, resolved through linear probing

Threads

1 2 3 4

ATGCTATGGA



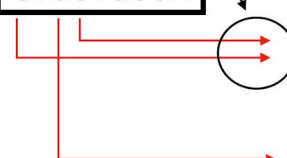
k-mers	Extensions
ATGC	T
GCTA	T

Thread collision, resolved through exclusive access

Threads

1 2 3 4

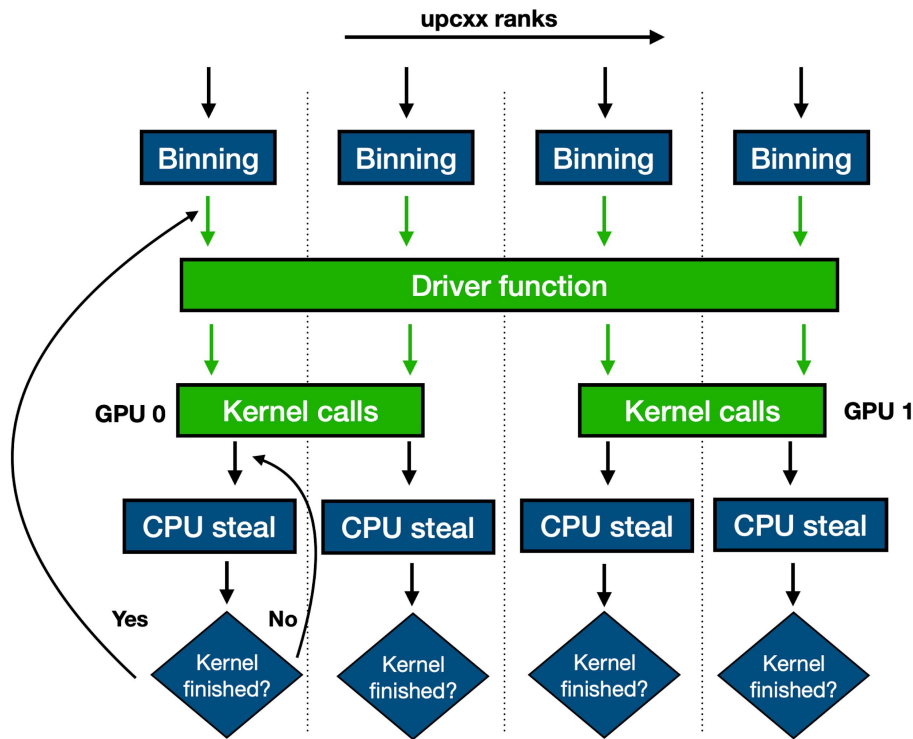
CTGCTGCGA



k-mers	Extensions
CTGC	T
GCTG	C

CUDA's atomic CAS paired with *match_any_sync* was used for mimicking an atomic region.

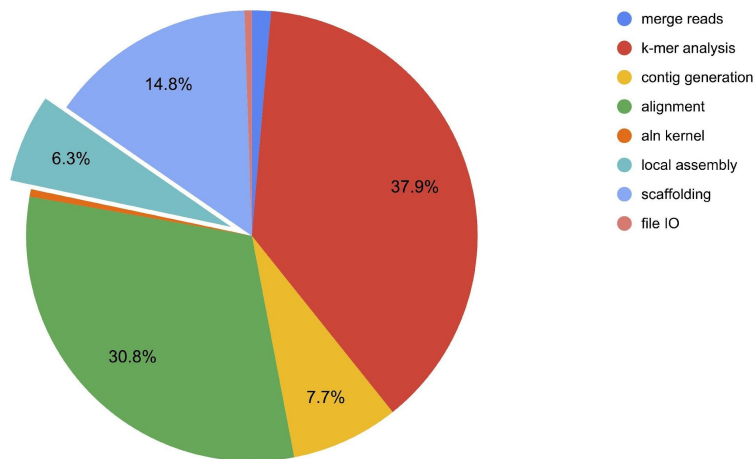
GPU Accelerated Local Assembly



- Driver function packs the data and makes calls to GPU kernels.
- Driver is called from a separate CPU thread which runs in background.
- This allows for CPU stealing.
- NVIDIA's MPS handles multiple kernel launches from different ranks.

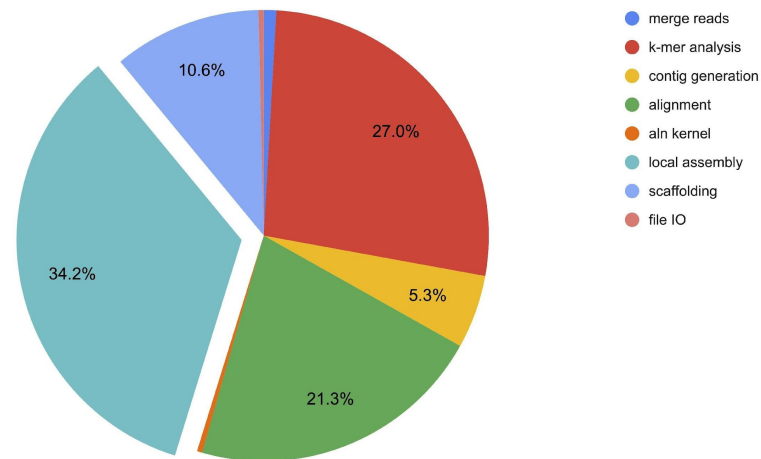
GPU Accelerated Local Assembly

MetaHipMer profile (GPU local assembly)



With GPU local assembly

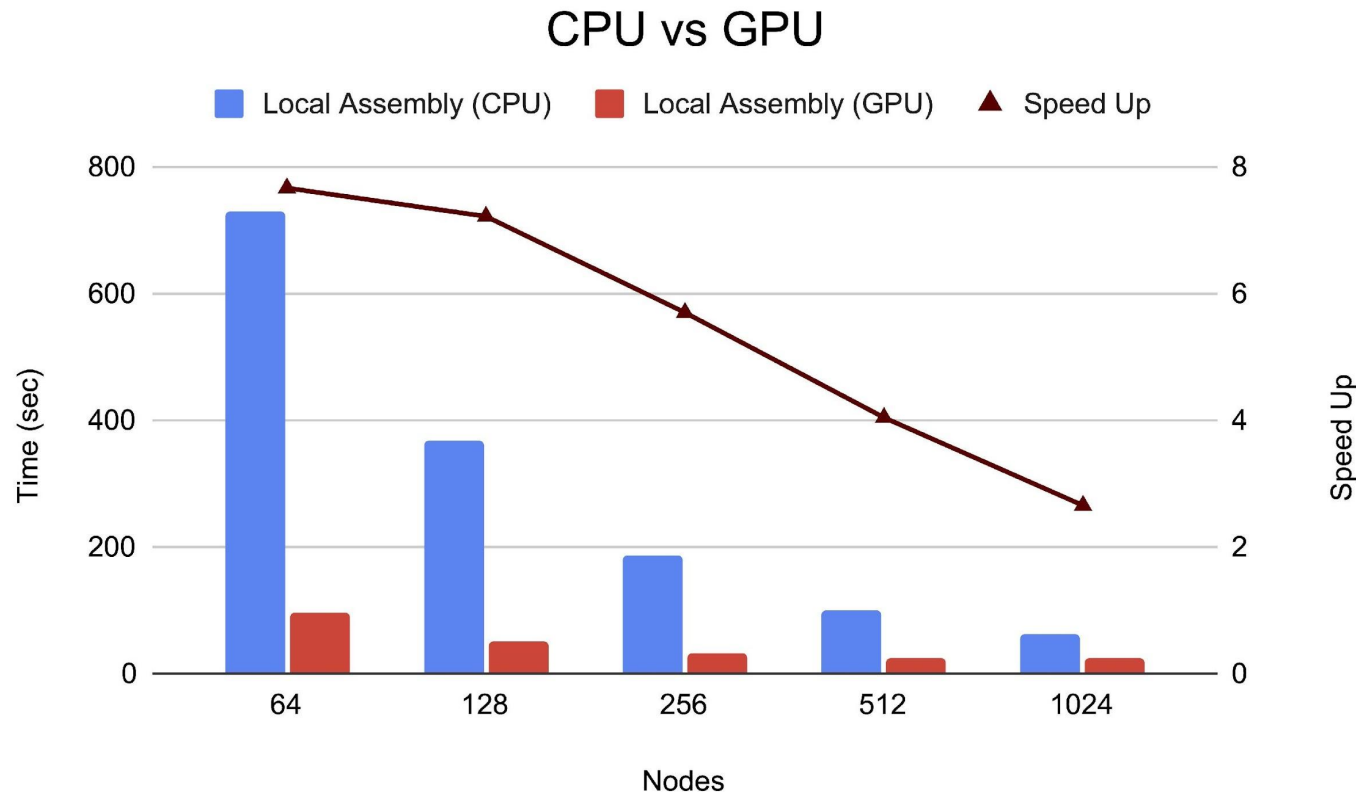
MetaHipMer profile (CPU local assembly)



With CPU local assembly

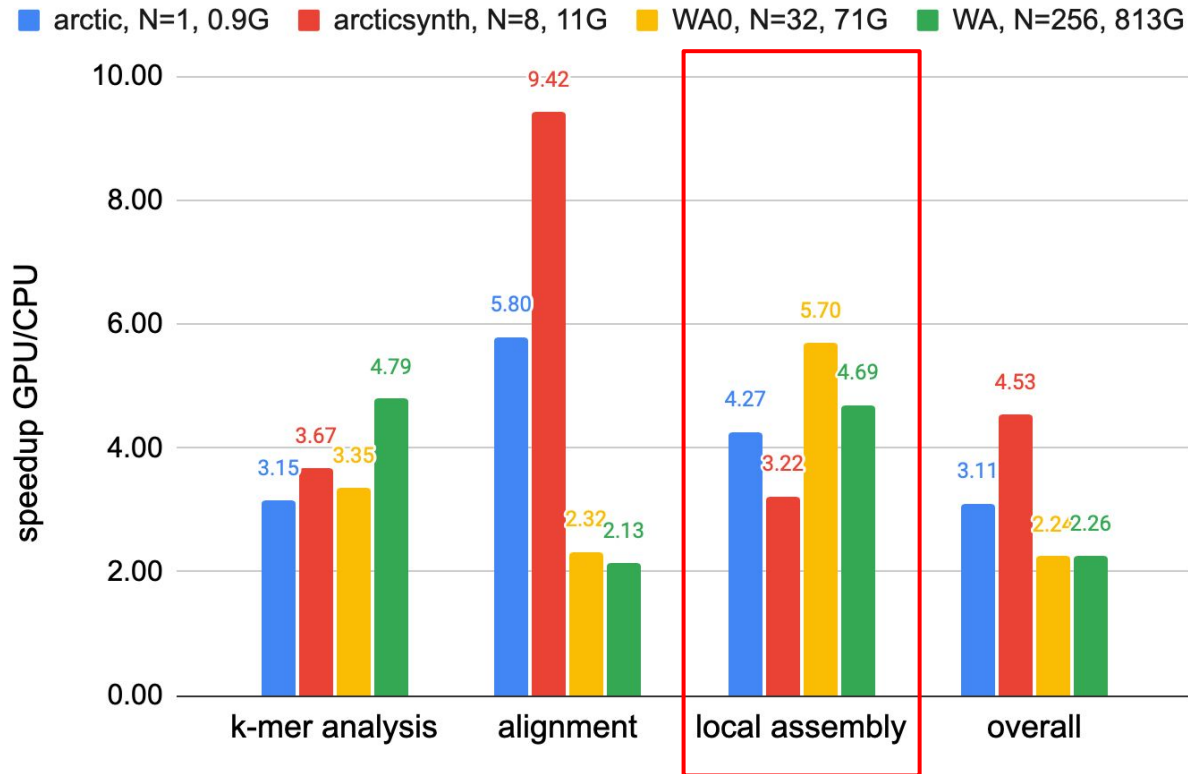
Awan, Muaaz Gul, et al. "Accelerating large scale de novo metagenome assembly using GPUs." *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 2021.

GPU Accelerated Local Assembly



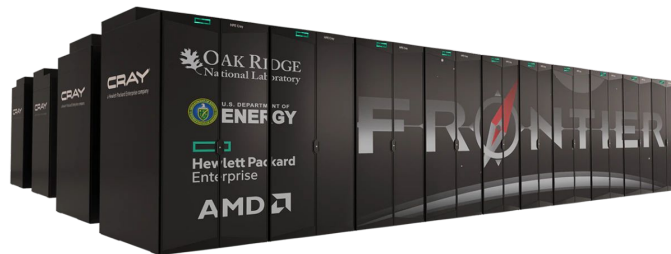
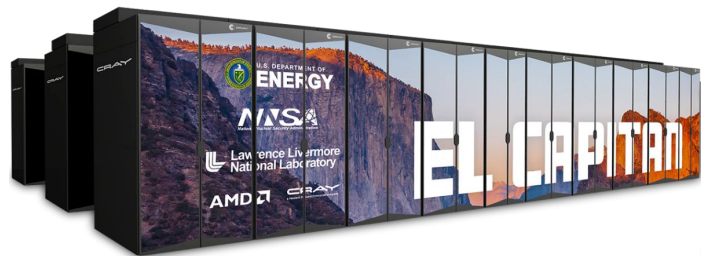
Awan, Muaaz Gul, et al. "Accelerating large scale de novo metagenome assembly using GPUs." *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 2021.

MetaHipMer Comparison



Porting to non-NVIDIA devices

- Planned and current exascale supercomputers do not have NVIDIA devices.
- HIP and SYCL are possible choices for AMD and Intel GPUs respectively.
- Some efforts within DOE are focussed on supporting multiple backends and runtimes.



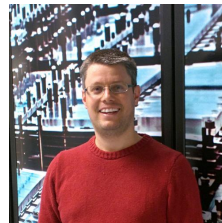
Acknowledgements



Rob Egan



Steven Hofmeyr



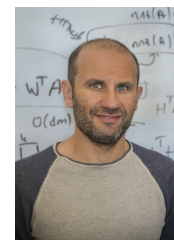
Jack Deslippe



Oguz
Selvitopi



Aydin Buluc



Leonid Olikar



Katherine Yelick



LEADERSHIP
COMPUTING
FACILITY

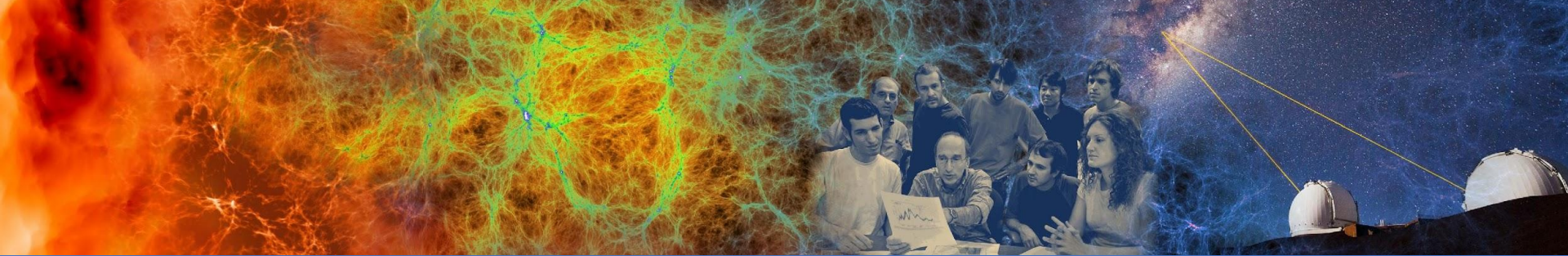


EXASCALE COMPUTING PROJECT



U.S. DEPARTMENT OF
ENERGY

Office of
Science



Thank you!