

Танцуем от  
полезности  
инструмента для  
студентов в их  
проектах.

Значит ускорение на  
GPU:

<https://github.com/citiususc/veryfasttree>

## Phylogenomic tools with GPU acceleration. List more github repos order by date of last release

### 🔗 Sources

How is Phylogenetic Analysis performed using GPU? | SeiMaxim  
seimaxim · 1

BEAGLE 3 High-performance Computational Library...  
hal · 2

beagle-dev/beagle-lib: general purpose library for evaluating the ...  
github · 3

View 5 more

### ✳ Answer

Here is an updated list of GitHub repositories related to phylogenomic tools with GPU acceleration, ordered by the date of their last release:

#### 1. NVIDIA Parabricks

- **Last Release:** March 19, 2024
- **Description:** A comprehensive tool for multi-omics analysis, supporting DNA, RNA, and methylation data with significant speed improvements through GPU acceleration.
- **GitHub Repository:** [NVIDIA Clara Parabricks Workflows](#).

#### 2. VeryFastTree

- **Last Release:** July 2023
- **Description:** An efficient tool for phylogenetic tree inference designed to handle massive datasets, leveraging parallelization and vectorization to speed up computations.
- **GitHub Repository:** [VeryFastTree](#).

#### 3. BEAGLE

- **Last Release:** Version 3.1.0 (exact date unspecified)
- **Description:** A high-performance library that supports Bayesian and Maximum Likelihood phylogenetics, optimized for GPU acceleration to enhance likelihood computations.
- **GitHub Repository:** [BEAGLE Library](#).

#### 4. IQ-TREE

- **Last Release:** January 4, 2024
- **Description:** A software tool for phylogenetics that uses GPU acceleration to improve the efficiency of tree inference.
- **GitHub Repository:** [IQ-TREE](#).

#### 5. RAXML-NG

nickjcroucher/  
gubbins

Rapid phylogenetic analysis of large samples of recombining bacterial whole genome sequences using Gubbins

At 20 Contributors · 6 Issues · 166 Stars · 49 Forks

citiususc/  
veryfasttree

Fast phylogenetic tree inference for massive genomic datasets

At 3 Contributors · 108 Stars · 5 Forks

ALPHA: Automated Local Phylogenomic Analyses

Launch

View More

No Videos Found

PRO Generate Image

v4.0 (July 2023):

Introduction of new thread levels for improved parallelization.

Enhanced performance through new parallel regions (e.g., ML Lengths, ML splits, LogLk, etc.).

Threads used in tree creation: Top hits, TopHitNJSearch, FastNJSearch, and ExhaustiveNJSearch (-slow).

Implementation of a **faster tree partitioning** approach with significant speed improvements.

Tree partitioning limited to NNI, SPR, and upProfiles computations for memory conservation:

Parallel tree traversal implemented for remaining parts.

Replacement of disk storage for profiles with **Disk Computing**.

Shared and reused Top upProfiles among threads for memory efficiency and accelerated sequential parts.

Improved non-deterministic mode with removal of mutex usage.

Optimized performance by parallelizing non-deterministic parts in deterministic mode:

Also implemented non-deterministic parts in deterministic mode for improved performance.

Deterministic mode now outperforms non-deterministic mode in terms of speed.

Tree partitioning method logging now hidden by default.

Support for **Fastq format and libBZ2 compression**.

Support for reading trees from NEXUS block trees.

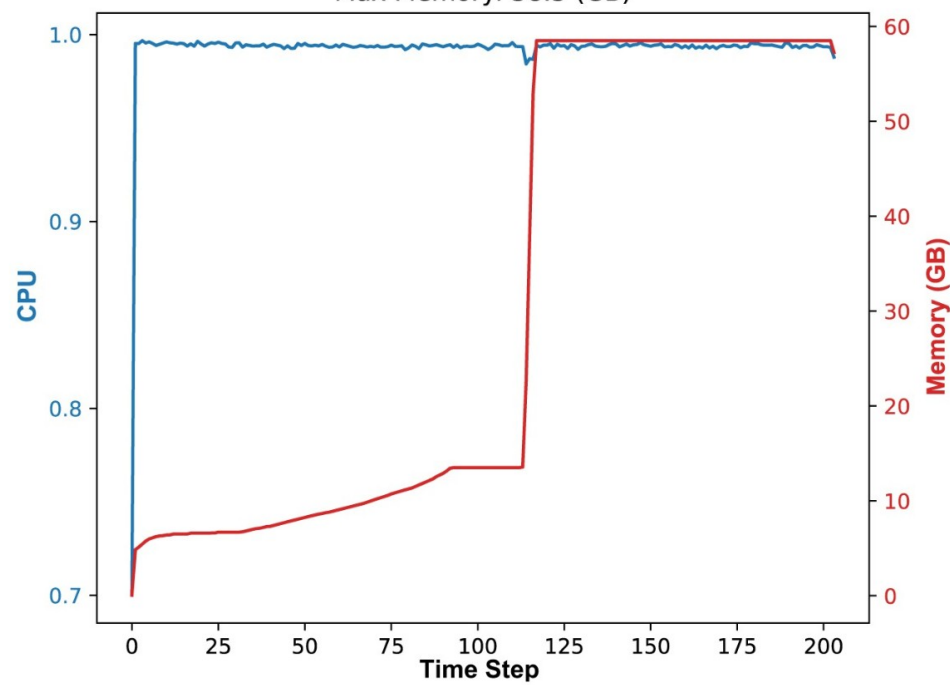
Nvidia **CUDA GPU computing support (experimental)**

Introduced parallel compilation.

Incorporation of changes from FastTree-2.1.11.

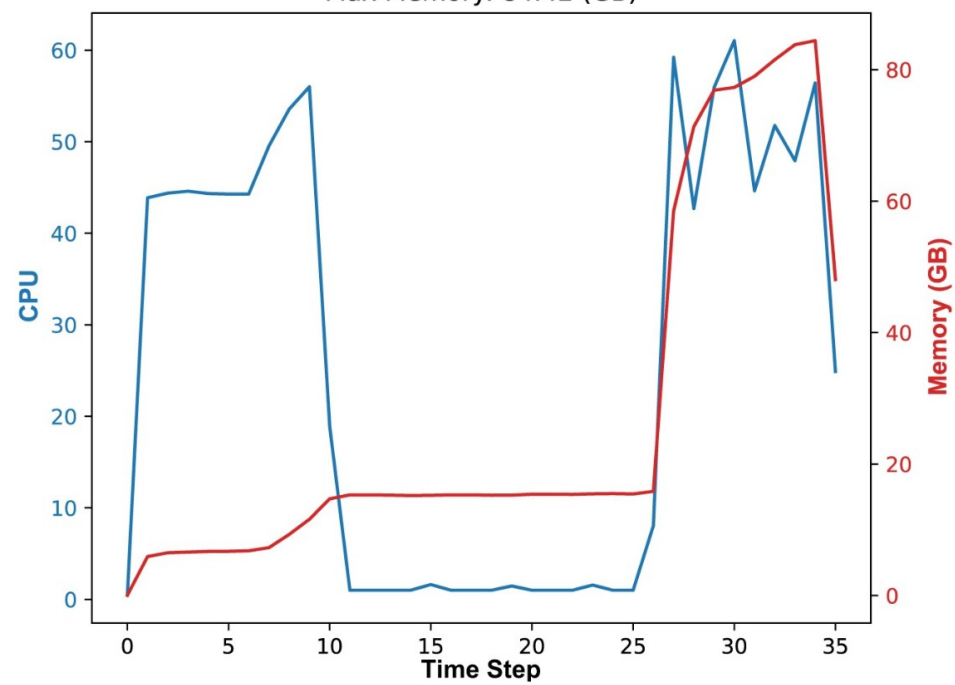
Clang Support

Max Memory: 58.5 (GB)



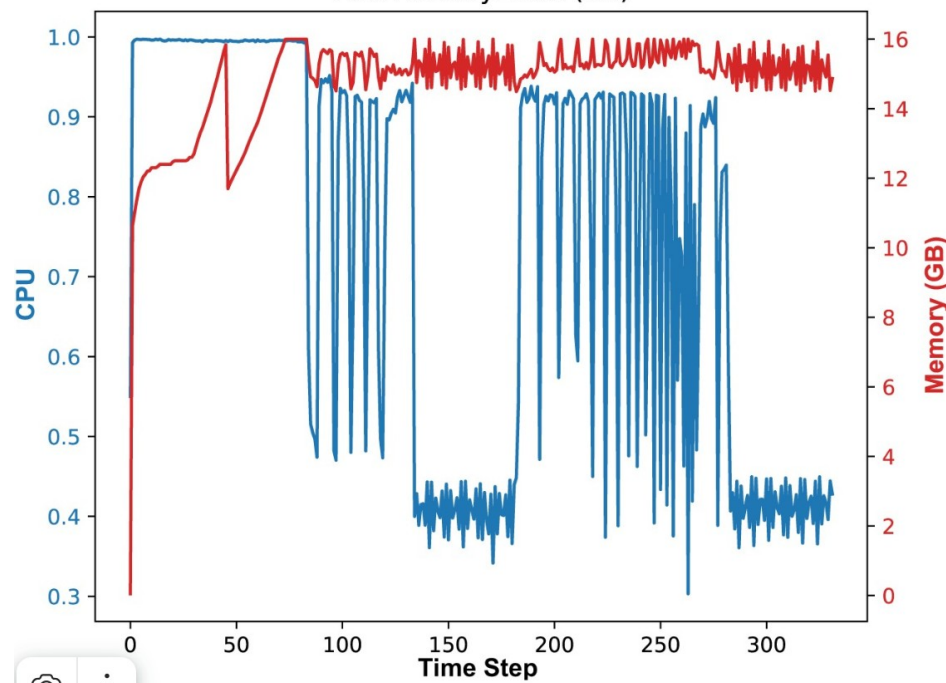
**A** VFT4 (1 thread)

Max Memory: 84.42 (GB)



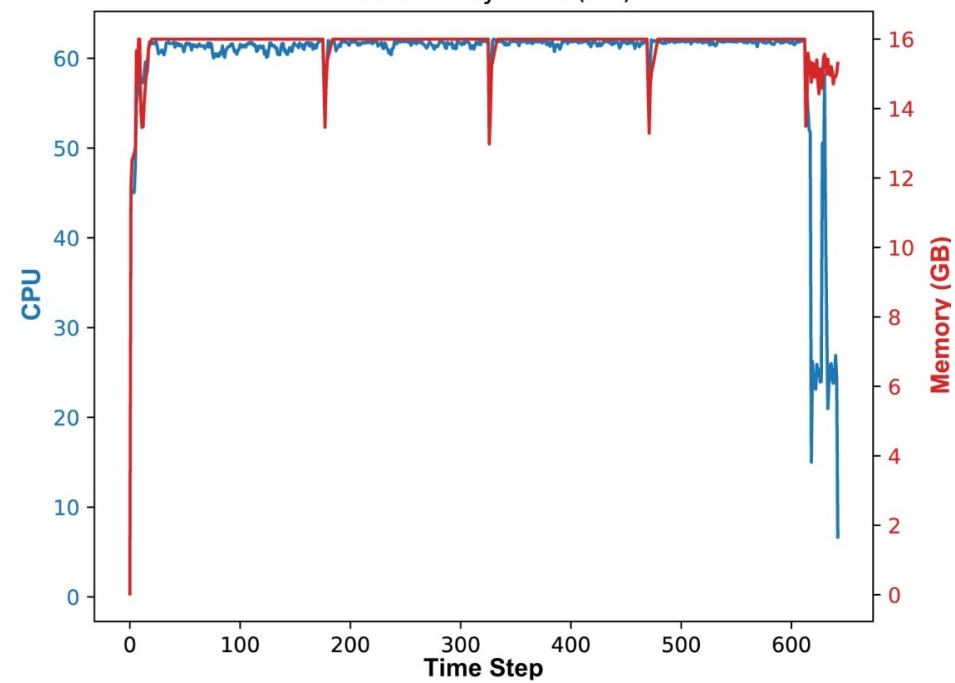
**B** VFT4 (64 threads)

Max Memory: 16.0 (GB)



**C** VFT4 with disk-computing enabled (1 thread)

Max Memory: 16.0 (GB)



**D** VFT4 with disk-computing enabled (64 threads)

VeryFastTree является самым быстрым инструментом на современном этапе для оценки филогении ML. Например, он способен обрабатывать массивные наборы данных, содержащие **1 миллион таксонов, всего за 36 часов**, что в несколько раз быстрее, чем другие инструменты.

VeryFastTree позволяет обрабатывать наборы данных, которые в противном случае были бы трудноразрешимыми или требовали чрезмерно большого времени вычислений.

Новой характеристикой VeryFastTree является то, что мы называем **дисковыми** вычислениями, которые позволяют обрабатывать чрезвычайно большие наборы данных **на недорогих серверах** с ограниченными ресурсами памяти.

Efficient phylogenetic tree inference for massive taxonomic datasets: harnessing the power of a server to analyze 1 million taxa

César Piñeiro, Juan C Pichel

GigaScience, Volume 13, 2024, giae055, <https://doi.org/10.1093/gigascience/giae055> Published: 08 August 2024

### **Доступность данных**

Наборы данных, подтверждающие результаты этой статьи, были получены следующим образом:

Большой набор данных был получен с веб-сайта инструмента FastTree-2 [17].

Очень большой набор данных был получен из базы данных Greengenes [27].

Сверхбольшой набор данных был получен из хранилища Kim Lab for Computational Evolutionary Biology (Университет Пенсильвании) [28].

Набор данных для моделирования выравнивания белков на 5000 последовательностей был получен с веб-сайта FastTree-2 tool [17].

Вспомогательные данные и архивная копия кода доступны через репозиторий GigaScience, GigaDB [29].