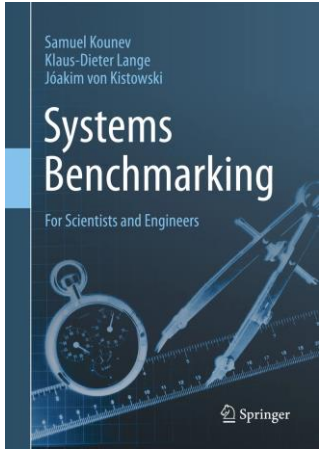
The Vision of OpenBehch

Wei Wang

ECNU / X-lab / ONES Group

2022-11

Why Benchmark ?



TOP500

From Wikipedia, the free encyclopedia

The **TOP500** project ranks and details the 500 most powerful non-distributed computer systems in the world. The project was started in 1993 and publishes an updated list of the supercomputers twice a year. The first of these updates always coincides with the International Supercomputing Conference in June, and the second is presented at the ACM/IEEE Supercomputing Conference in November. The project aims to provide a reliable basis for tracking and detecting trends in high-performance computing and bases rankings on HPL,[1] a portable implementation of the high-performance LINPACK benchmark written in Fortran for distributed-memory computers.

Currently the latest TOP500 list is the 59th, published in June 2022. Since June 2022, USA's Frontier is the world's most powerful supercomputer, reaching 1102 petaFlops (1.102 exaFlops) on the LINPACK benchmarks.[2] The United States has by far the highest share of total computing power on the list (nearly 50%),[3] while China currently leads the list in number of systems with 173 supercomputers, with the USA not far behind in second place.

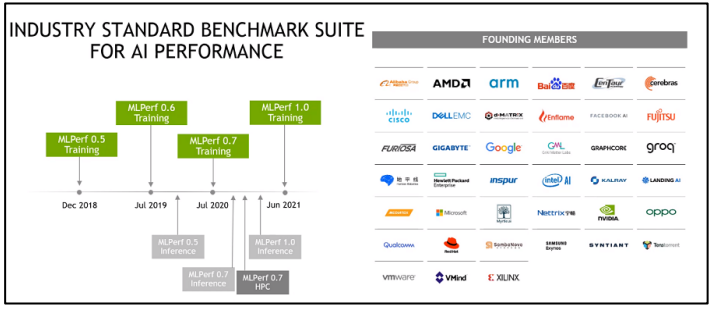
TOP500
TOP 500
The List.

Key people	Erich Strohmaier Jack Dongarra Horst Simon Martin Meuer
Established	24 June 1993
Website	top500.org

The TOP500 list is compiled by Jack Dongarra of the University of Tennessee, Knoxville, Erich Strohmaier and Horst Simon of the National Energy Research Scientific Computing Center (NERSC) and Lawrence Berkeley National Laboratory (LBNL), and, until his death in 2014, Hans Meuer of the University of Mannheim, Germany.

The TOP500 project lists also Green500 and HPCG benchmark list.

Jack Dongarra

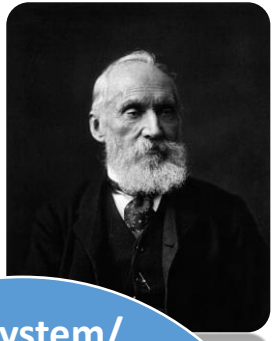


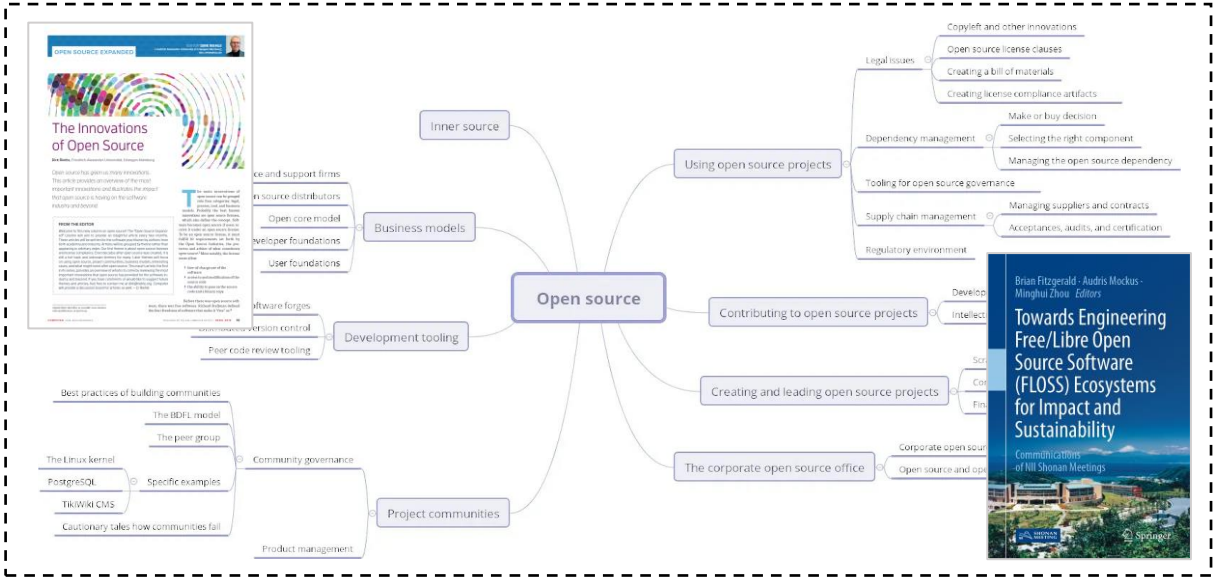
MLPerf
AI Benchmarks

Why OSS ecosystem benchmark ?

“When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind.”

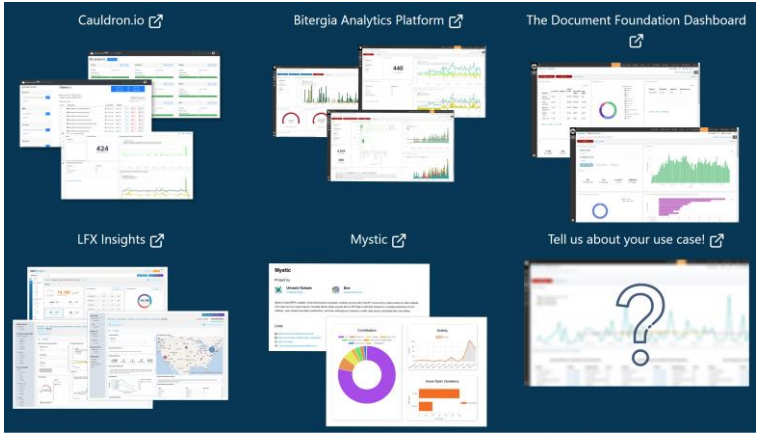
— **Lord Kelvin**





Open Source Expanded (IEEE Computer magazine column)

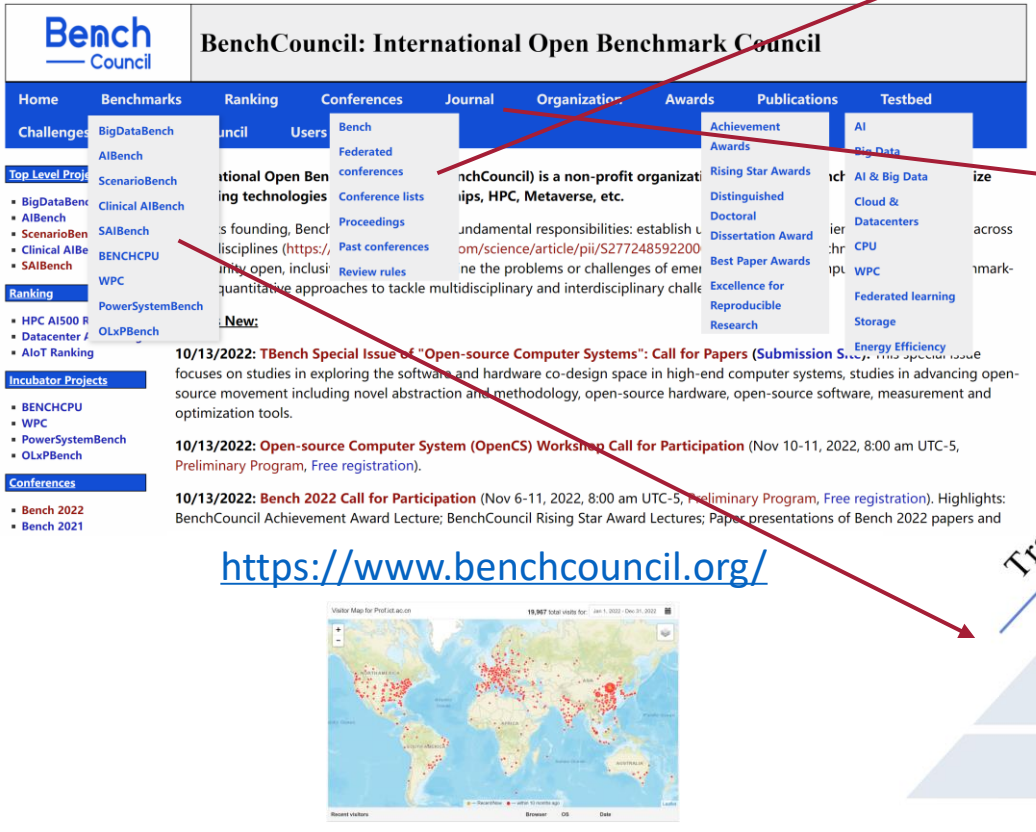
Why OSS ecosystem benchmark ?



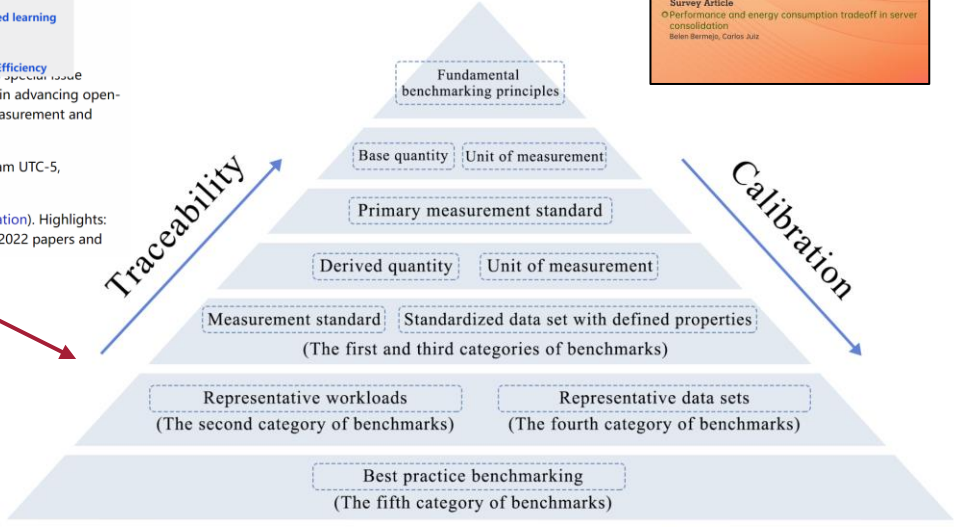
CHAOSS

GRIMOIRELAB

How ? Bench Council



https://www.benchcouncil.org/

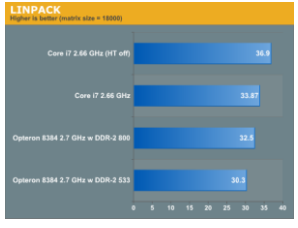


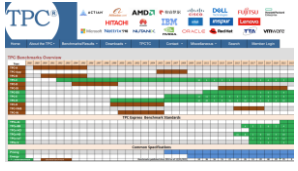
Jianfeng Zhan, Call for establishing benchmark science and engineering, BenchCouncil Transactions on Benchmarks, Standards and Evaluations (TBench), 2021

Benchmark hierarchy

First category

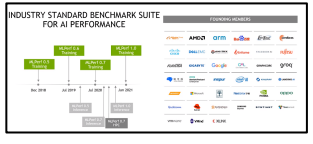
The first category is a **measurement standard**. In the computer discipline, the Linpack benchmark is of this category, which is widely used to report the performance of a high-performance computer.

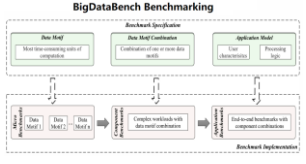




Second category

The second one is the **representative workloads** that run on the systems under measurement. The application benchmarks or synthetic benchmarks in the computer discipline are of this category.



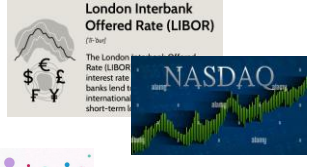


Third category

The third is a **standardized data set** that represents real-world **data science problem**, with defined properties and **ground truth**. This benchmark is often used to measure against different **algorithms**. The state-of-the-art algorithm implementation plus the data set constitutes the benchmark of the second category.

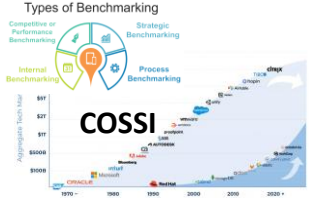
Fourth category

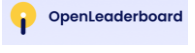
The fourth is a **representative data set**, used as a reference. For example, a financial benchmark is an index (statistical measure), calculated from a representative set of underlying data, is used as a reference for financial instruments or contracts.



Fifth category

The fifth is **the industry best practices** in different domains. Benchmarking is the continuous process of searching the industry best practices that lead to superior performance and measuring products, services, and processes against them.





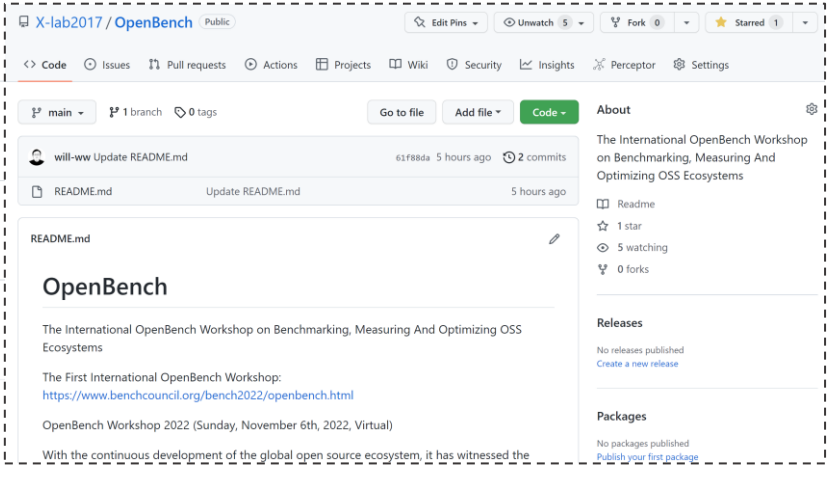
Goals of OpenBench

- **It aims to promote the sustainable growth and evolution of the open source software ecosystem by establishing benchmarks, standards and evaluations of international consensus.**
- **The forum invites researchers from industry, academia, foundations, associations and government organizations to jointly explore the frontier direction of open source software ecological benchmarks, , including but not limit to the topics around:**
  - OSS metrology and measurement methods
  - OSS measurement data standards,
  - Definition and classification of OSS ecological benchmarks
  - Standard datasets and analysis tasks
  - Indicator system and benchmarking system
  - Application scenarios and cases of open source ecological benchmarking
  - So much more …

The First International OpenBench Workshop

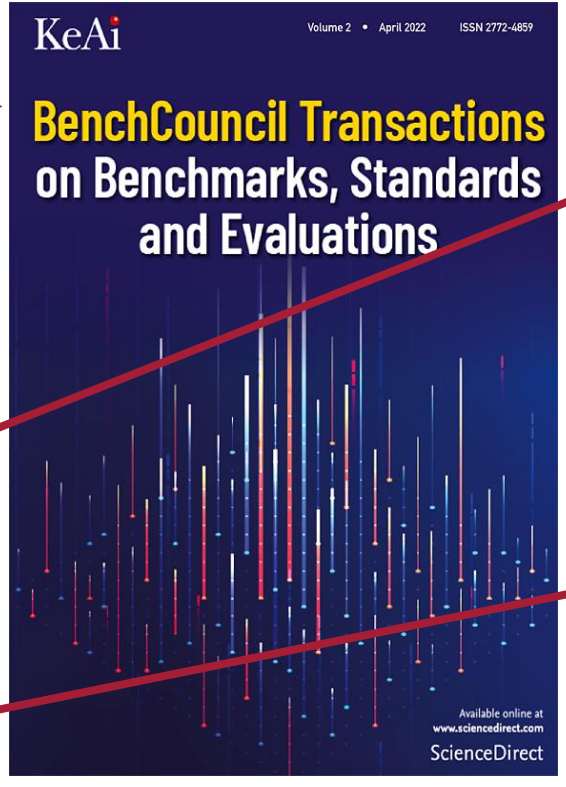
Time	Event
UTC-5 (Washington, New York), 07:50-08:00 UTC (London), 12:50-13:00 UTC+8 (Beijing), 20:50-21:00	**Opening remarks**
UTC-5 (Washington, New York), 08:00-08:20 UTC (London), 13:00-13:20 UTC+8 (Beijing), 21:00-21:20	**Evolution of CHAOSS: How we got where we are from our beginnings** **Sean P. Goggins**, Professor, University of Missouri
UTC-5 (Washington, New York), 08:20-08:40 UTC (London), 13:20-13:40 UTC+8 (Beijing), 21:20-21:40	**Data-driven of Open Source & InnerSource community build evaluation** **Willem Jiang**, Open Source Expert of Huawei, Mentor of Huawei InnerSource Foundation,
UTC-5 (Washington, New York), 08:40-09:00 UTC (London), 13:40-14:00 UTC+8 (Beijing), 21:40-22:00	**Open Source Security: Challenges, Solutions, and Opportunities** **Yang LIU**, Professor, Nanyang Technological University
UTC-5 (Washington, New York), 09:00-09:20 UTC (London), 14:00-14:20 UTC+8 (Beijing), 22:00-22:20	**Explore the Use of Data Heuristics in OSPO work** **Richard Sikang Bian**, Head of Open-Source Program Office (OSPO), Technical Strategy Expert of Ant Group
UTC-5 (Washington, New York), 09:20-09:40 UTC (London), 14:20-14:40 UTC+8 (Beijing), 22:20-22:40	**Commercial Participation in OSS: Models, Collaborations, and Turnover** **Yuxia Zhang**, Assistant professor at school of computer science & technology, Beijing Institute of Technology (BIT)
UTC-5 (Washington, New York), 09:40-10:00 UTC (London), 14:40-15:00 UTC+8 (Beijing), 22:40-23:00	**Model-informed Automatic Quality Assessment for Open Datasets** **Jidong Tian & Yaohui Jin**, Shanghai Jiao Tong University
UTC-5 (Washington, New York), 10:00-10:20 UTC (London), 15:00-15:20 UTC+8 (Beijing), 23:00-23:20	**Community Responsiveness Studies with Apache DevLake** **Chenhui Zhang**, Data Science Lead, Merico DevInsight
UTC-5 (Washington, New York), 10:20-10:40 UTC (London), 15:20-15:40 UTC+8 (Beijing), 23:20-23:40	**Measuring qualitative data with predictability** **Samantha Venia Logan**, Expert in Online Community Management and full-stack marketing
UTC-5 (Washington, New York), 10:40-11:00 UTC (London), 15:40-16:00 UTC+8 (Beijing), 23:40-24:00	**Round table (discussion) & Closing Remarks**

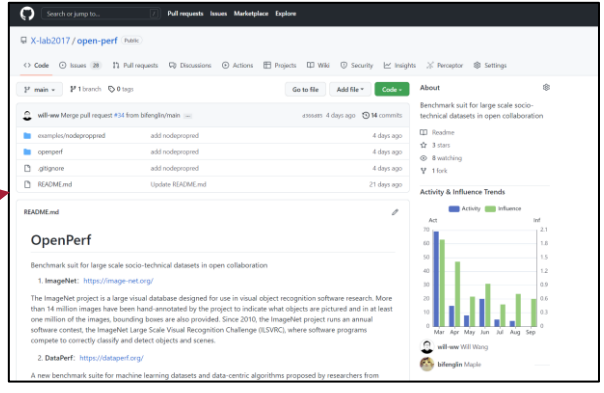
https://github.com/X-lab2017/OpenBench



Future plans (2023)

- Call for **OpenBench program committee** (PC) members
- TBench Special Issue of "**Benchmarking Open Source Software Ecosystems for Impact and Sustainability**": Call for Papers
  - Position Papers
  - Full-Length Articles
  - Review Papers
  - Short Communications
- **OpenBench 2023**
  - Topics
  - Call for papers
  - Call for Benchmarks, Standards, data sets, et al.
- **OpenPerf** project/repo





https://github.com/X-lab2017/open-perf

Thank you !