SCIENCE CHINA
Information Sciences

• NEWS & VIEWS •

January 2023, Vol. 66 117101:1–117101:2
https://doi.org/10.1007/s11432-022-3581-9

Hetu: a highly efficient automatic parallel distributed deep learning system

Xupeng MIAO[1], Xiaonan NIE[1], Hailin ZHANG[1], Tong ZHAO[1] & Bin CUI[1,2,3]*

[1]*School of Computer Science, Peking University, Beijing* 100871*, China;*
[2]*Center for Data Science, Peking University, Beijing* 100871*, China;*
[3]*National Engineering Laboratory for Big Data Analysis and Applications, Beijing* 100871*, China*

Received 2 March 2022/Revised 5 July 2022/Accepted 22 August 2022/Published online 1 December 2022

Citation Miao X P, Nie X N, Zhang H L, et al. Hetu: a highly efficient automatic parallel distributed deep learning system. Sci China Inf Sci, 2023, 66(1): 117101, https://doi.org/10.1007/s11432-022-3581-9

On July 18, 2021, the PKU-DAIR Lab[1)] (Data and Intelligence Research Lab at Peking University) openly released the source code of Hetu, a highly efficient and easy-to-use distributed deep learning (DL) framework. Hetu is the first distributed DL system developed by academic groups in Chinese universities, and takes into account both high availability in industry and innovation in academia. Through independent research and development, Hetu is completely decoupled from the existing DL systems and has unique characteristics. The public release of the Hetu system will help researchers and practitioners to carry out frontier MLSys (machine learning system) research and promote innovation and industrial upgrading.

Background. Machine learning, especially deep learning, has grown rapidly in recent years and has revolutionized the traditional big data computing systems. It brings the awareness of a new field focused on ML in practice — machine learning systems [1]. This field acts as a bridge between the domains of computer systems and artificial intelligence, considering the new challenges of ML with a lens shaped by traditional system research. Due to the explosive growth of diverse applications of ML in production, the continued growth in data volume, and the increasing complexity of large-scale learning systems, building practical large-scale ML systems for not only data scientists but also system engineers is increasingly challenging.

Over the past decade, advanced DL has driven the ML explosion and DL systems are becoming the cornerstone of artificial intelligence. Emerging DL models (e.g., feedforward neural network, convolutional neural network, recurrent neural network, graph neural network, Transformer, sparse embedding model, and mixture-of-experts) are playing crucial roles in real-world industrial tasks. Meanwhile, they have special computation and communication characteristics, suffering from distinct system challenges from different aspects, including data management, computation acceleration, hardware under-utilization, network communication, memory access, and execution scheduling. Therefore, the PKU-DAIR Lab launched a project to develop a novel and holistic artificial intelligence (AI) system to support various AI applications and deployments, regarding it as an important development trend for both AI and system communities.

Analysis of current DL frameworks. There are some well-known DL frameworks supporting distributed training ability. TensorFlow maps the nodes of a dataflow graph across many machines in a cluster (or multiple computational devices within a machine) and utilizes the parameter server architecture to manage the model state. Another representative system is PyTorch, which provides asynchronous dataflow execution and distributed data parallel module with the AllReduce primitive. Recently, PyTorch has been quite popular in academia due to its user-friendly imperative pythonic interfaces, which makes debugging easy, while TensorFlow has advantages in industrial scenarios since its rich ecosystems accelerate the development of ML services.

With the development of DL, distributed scalability is becoming the core competitiveness of DL systems. Although existing systems have provided some customized distributed interfaces, they still face severe challenges and obstacles.

- Functionality: Limited communication architectures, parallel strategies, and consistency protocols.
- Complexity: The system implementation of communication and computation is highly coupled and difficult to follow and optimize.
- Usability: The complicated deployment of distributed training paradigms requires human expert knowledge for better efficiency.

Moreover, they are also suffering from the efficiency and scalability bottlenecks for large-scale distributed training. These observations motivate us to break the current system abstraction, create a novel design to handle these concerns, and build a high-performance distributed DL system.

Hetu development process. In October 2018, the PKU-DAIR Lab proposed to set up a team in PKU to develop a

* Corresponding author (email: bin.cui@pku.edu.cn)
1) https://github.com/PKU-DAIR.

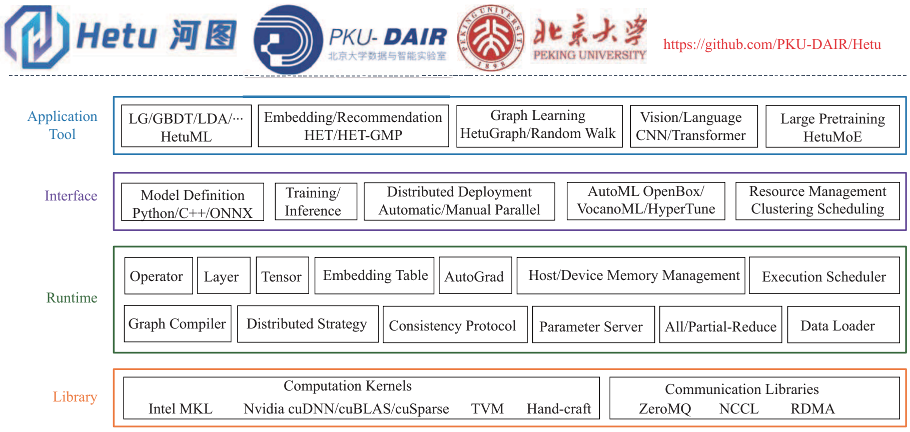

Figure 1 Overview of Hetu.

DL system for high-dimensional and large-scale deep learning scenarios. The team comprises more than 15 graduate and undergraduate students, led by a senior Ph.D. student who has participated in developing an industrial distributed ML system [2]. This research is supported by the National Key R&D Program of China (No. 2018YFB1004403). On July 18, 2021, version 1.0 of this DL system was publicly released as open-source, and the system was named Hetu (a mysterious symbol in Chinese, representing ancient Chinese philosophical thoughts).

Hetu design and features. Hetu is designed to make ML development more generic, efficient, flexible, scalable, reliable, deployable, and easy-to-use, particularly for DL. Overall, the main design principle of Hetu concentrates on both performance and usability, leading to a simple and performant system. Currently, Hetu and its ecosystems contain four layers as illustrated in Figure 1, i.e., application tools, user interfaces, Hetu runtime, and execution libraries.

- First, Hetu supports various AI applications with novel accelerations, such as HetuML for over ten traditional ML problems (e.g., linear/tree/topic models), HET [3] and HET-GMP [4] for large embedding models, HetuGraph for graph learning models, HetuMoE for large pretraining models, and vision/language Transformer models.
- Second, Hetu provides a perspicuous interface to satisfy specified requirements from different users. For fresh users without background knowledge, we provide automatic parallel deployment for efficient and easy-to-use distributed training and AutoML services for convenient development.
- Third, Hetu runtime follows the dataflow graph abstraction and makes enormous efforts to enrich the functionalities and enable the innovations, e.g., host/device memory management, compilation optimization, flexible synchronization [5], and hybrid communication architecture.
- Finally, the underlying libraries for computation kernels and communication primitives support various hardware (e.g., GPU) with hand-crafted optimizations.

Automatic parallelization. Hetu allows developers to build DL models in the logical graph as if they are using a single device, then it automatically generates a physical graph describing the tensor distribution and computation parallelization over multiple devices in distributed environments. Hetu adopts a unified intermediate representation covering almost all distributed training functionalities in existing DL systems and many advanced features to meet new challenges. Based on these attributes, the graph transformation process manages to find an efficient semantically equivalent parallel execution plan through the cost estimation mechanisms. It can also be started from some hints (e.g., how to determine the device placement, and how to distribute tensors, how to partition operators) to utilize the expert experiences.

Achievements and applications. Hetu has accomplished remarkable achievements after being open-sourced. In December 2021, Hetu achieved the highest outstanding award in the CCF Big Data & Computing Intelligence Contest 2021 and the first prize in the China Software Open-source Innovation Competition held by CCF ChinaSoft 2021. In January 2022, Hetu was awarded the Top-10 Open-source Events by the 2021 Synced Machine Intelligence Awards. Some key innovative technologies of Hetu have been successfully applied in real industrial applications, such as large-scale advertising recommendation in Tencent, TMALL 11-11 commodity recommendation and City Brain transportation forecasting in Alibaba Cloud, and AutoML in Kuaishou.

Access methods. Hetu is published under the MulanPSL 2.0 open source license. Its source code is available. Detailed information can be found on the websites[2)3)].

References

1 Abadi M, Barham P, Chen J, et al. TensorFlow: a system for large-scale machine learning. In: Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation, 2016. 265–283
2 Zhang Z P, Cui B, Shao Y X, et al. PS2: parameter server on spark. In: Proceedings of the 2019 International Conference on Management of Data, 2019. 376–388
3 Miao X P, Zhang H L, Shi Y N, et al. HET: scaling out huge embedding model training via cache-enabled distributed framework. In: Proceedings of the VLDB Endowment, 2021. 312–320
4 Miao X P, Shi Y N, Zhang H L, et al. HET-GMP: a graph-based system approach to scaling large embedding model training. In: Proceedings of the 2022 International Conference on Management of Data, 2022. 470–480
5 Miao X P, Nie X N, Shao Y X, et al. Heterogeneity-aware distributed machine learning training via partial reduce. In: Proceedings of the 2021 International Conference on Management of Data, 2021. 2262–2270

2) https://www.gitlink.org.cn/PKU-DAIR/Hetu.
3) https://github.com/PKU-DAIR/Hetu.