Protea: Client Resource Profiling Engine for Federated Systems

Wanru Zhao[†], Xinchi Qiu[†], Javier Fernandez-Marques[◇]
Pedro P. B. de Gusmão[†], Nicholas D. Lane[†]
[†]University of Cambridge [◇]Samsung AI

Federated Learning (FL) [1, 2] has emerged as a prospective solution that facilitates the training of a high-performing centralised model without compromising the privacy of users. While successful, FL research is currently limited by the difficulties of establishing a realistic large-scale FL system at the early stages of experimentation. Simulating such scenarios has become the go-to approach.

Simulation can be highly beneficial if it allows enough degrees of freedom to study different scenarios at the system and algorithmic levels. Scale and heterogeneity [3] are the two main concerns for realistic large-scale FL system simuation. A typical procedure to launch an FL simulation requires specifying how many clients operate in each round and the resources each client needs. However, few frameworks provide this interface [4, 5, 6, 7].

To facilitate efficient scalable FL simulation of heterogeneous clients, we design and implement `Protea`, a flexible and lightweight client profiling component within federated systems using the FL framework Flower [8]. It allows automatically collecting system-level statistics and estimating the resources needed for each client, thus running the simulation in a resource-aware fashion. The results show that our design successfully increases parallelism for 1.66 $\times$ faster wall-clock time and 2.6$\times$ better GPU utilisation, which enables large-scale experiments on heterogeneous clients.

References

[1] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[2] Keith Bonawitz, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloe Kiddon, Jakub Konečnỳ, Stefano Mazzocchi, H Brendan McMahan, et al. Towards federated learning at scale: System design. *arXiv preprint arXiv:1902.01046*, 2019.

[3] Peter Kairouz, H Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Keith Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, et al. Advances and open problems in federated learning. *arXiv preprint arXiv:1912.04977*, 2019.

[4] Chaoyang He, Songze Li, Jinhyun So, Mi Zhang, Hongyi Wang, Xiaoyang Wang, Praneeth Vepakomma, Abhishek Singh, Hang Qiu, Li Shen, Peilin Zhao, Yan Kang, Yang Liu, Ramesh Raskar, Qiang Yang, Murali Annavaram, and Salman Avestimehr. Fedml: A research library and benchmark for federated machine learning. *Advances in Neural Information Processing Systems, Best Paper Award at Federate Learning Workshop*, 2020.

[5] Dimitrios Dimitriadis, Mirian Hipolito Garcia, Daniel Madrigal, Andre Manoel, and Robert Sim. Flute: A scalable, extensible framework for high-performance federated learning simulations, March 2022.

[6] G Anthony Reina, Alexey Gruzdev, Patrick Foley, Olga Perepelkina, Mansi Sharma, Igor Davidyuk, Ilya Trushkin, Maksim Radionov, Aleksandr Mokrov, Dmitry Agapov, Jason Martin, Brandon Edwards, Micah J. Sheller, Sarthak Pati, Prakash Narayana Moorthy, Shih han Wang, Prashant Shah, and Spyridon Bakas. Openfl: An open-source framework for federated learning, 2021.

[7] Fan Lai, Yinwei Dai, Xiangfeng Zhu, Harsha V Madhyastha, and Mosharaf Chowdhury. Fedscale: Benchmarking model and system performance of federated learning. In *Proceedings of the First Workshop on Systems Challenges in Reliable and Secure Federated Learning*, pages 1–3, 2021.

[8] Daniel J. Beutel, Taner Topal, Akhil Mathur, Xinchi Qiu, Javier Fernandez-Marques, Yan Gao, Lorenzo Sani, Kwing Hei Li, Titouan Parcollet, Pedro Porto Buarque de Gusmão, and Nicholas D. Lane. Flower: A friendly federated learning research framework. *CoRR*, abs/2007.14390, 2020.