

[Inicio](#) > SORS: Graph Neural Networks on CPUs: Enabling Affordable and Distributed Training and Inference

[SORS: Graph Neural Networks on CPUs: Enabling Affordable and Distributed Training and Inference](#)

Objectives

[Download here the slides of the presentation](#)

Abstract: Graph Neural Networks (GNNs) have gained significant popularity in computer vision and natural language processing (NLP) for their ability to model complex relationships and dependencies among entities in data. However, the high cost of GPUs and TPUs has made it difficult to deploy GNNs on a large scale. On the other hand, CPUs are widely available and more affordable, making them an attractive alternative.

In this talk, we will discuss the potential of using CPUs for GNN training and inference. We will explore different techniques for optimizing GNNs on CPUs, including parallelization and vectorization. Furthermore, we will discuss the potential advantages of using multi-CPU for GNN training, including lower costs, improved scalability, and faster computation times.

We will also examine several applications of GNNs in computer vision and NLP, highlighting their potential for real-world solutions. Finally, we will discuss future research directions in this field, including the development of new techniques for optimizing GNNs on CPUs and the integration of GNNs with other machine learning models.



Short Bio: Mahdi Jelodari

holds a PhD in computer science from the University of Manchester, UK. He is currently a ML/IoT consultant based in London, leading several R&D projects in the area of Healthcare. Mahdi is a 2x entrepreneur and patent holder in the area of compilers and machine learning. After executing several successful pilots with customers, He is now looking into cost optimising AI models specifically Graph based neural nets at scale. He is currently a Severo Ochoa Mobility Grant holder at BSC working with Computer Architecture For Parallel Paradigms group.

Speakers

Speaker: Mahdi Jelodari, PhD is patent holder in the area of high-level synthesis and programming language design for code acceleration on FPGAs.

Host: Osman Unsal, Computer Architecture For Parallel Paradigms Group Manager, CS, BSC

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 7 Sep 2024 - 13:52): <https://www.bsc.es/es/research-and-development/research-seminars/sors-graph-neural-networks-cpus-enabling-affordable-and-distributed-training-and-inference>