

Nouvelles approches pour la construction efficace de graphes pour l'utilisation des Graph Neural Networks au HL-LHC sur architectures CPU et GPU

jeudi 26 septembre 2024 09:30 (30 minutes)

La construction de graphes est une étape clé dans les méthodes basées sur les Graph Neural Network [1-4] qui sont développées pour la reconstruction de traces dans le détecteur ATLAS [5] au CERN [6] pour la phase de haute luminosité du LHC (HL-LHC) [7] qui doit démarrer en 2029. L'algorithme de construction calcule les connexions entre les hits dans le détecteur en se basant sur une carte des connexions possibles entre modules (la « module map ») et des critères de sélection géométriques. Le temps de calcul de cet algorithme est extrêmement contraint dans le contexte de reconstruction des traces durant HL-LHC. Jusqu'à récemment, cet algorithme avait un coût rédhibitoire (plusieurs centaines de secondes par événement). Nous présentons une première nouvelle approche dans la gestion des critères de sélection et de nouvelles optimisations de l'utilisation de la mémoire et d'accès aux données qui a permis de réduire considérablement le temps de calcul, permettant ainsi de passer à quelques secondes par événement sur 1 cœur CPU. Au HL-LHC nous considérons aussi la possibilité d'exploiter les architectures GPU qui deviennent de plus en plus compétitive en coût. Nous présentons une seconde approche qui redessine en profondeur l'algorithme dans une version hautement parallélisée d'accès la module map et qui a été portée intégralement sur GPU. Cette nouvelle approche permet de gagner deux ordres de grandeur supplémentaires et de réduire significativement le temps de calcul à quelques dizaines de millisecondes.

Les performances en termes de temps de calcul et de charge mémoire des différentes approches sur différentes architectures ainsi que la portabilité du code seront discutés.

Références:

- [1] Biscarat, Catherine et al. "Towards a realistic track reconstruction algorithm based on graph neural networks for the HL-LHC". In: EPJ Web Conf. 251 (2021), p. 03047. doi: 10.1051/epjconf/202125103047. URL: <https://doi.org/10.1051/epjconf/202125103047>.
- [2] Gage Dezoort, Peter W Battaglia, Catherine Biscarat, Jean-Roch Vlimant. Graph neural networks at the Large Hadron Collider. Nature Rev.Phys., 2023, 5 (5), pp.281-303. <10.1038/s42254-023-00569-0>. <hal-04096263>
- [3] Heberth Torres o.b.o ATLAS Collaboration « Physics Performance of the ATLAS GNN4ITk Track Reconstruction Chain » Proceedings of the CTD 2023, <https://cds.cern.ch/record/2882507>
- [4] Sylvain Caillou et al. « Novel fully-heterogeneous GNN designs for track reconstruction at the HL-LHC » Proceedings of the CHEP 2023, EPJ Web of Conferences 295, 09028 (2024) <https://doi.org/10.1051/epjconf/202429509028>
- [5] ATLAS experiment, url: <https://atlas.cern>
- [6] CERN, url: <https://home.cern/>
- [7] Apollinari et al., High-Luminosity Large Hadron Collider (HL-LHC), url: <https://cds.cern.ch/record/2284929>

Auteurs principaux: VALLIER, Alexis (L2I Toulouse, CNRS/IN2P3, UT3); Dr COLLARD, Christophe (L2I Toulouse, CNRS/IN2P3, UT3); TORRES, Heberth (L2I Toulouse, CNRS/IN2P3, UT3); STARK, Jan (L2I Toulouse, CNRS/IN2P3, UT3); CAILLOU, Sylvain (L2I Toulouse, CNRS/IN2P3, UT3)

Orateur: Dr COLLARD, Christophe (L2I Toulouse, CNRS/IN2P3, UT3)

Classification de Session: Le Calcul et l'IA