



2025 1학기 컴퓨터소프트웨어학부 CSE융합세미나



장소

ITBT관 911호

날짜 / 시간

2025.05.28 16:00~18:00

Upcoming Talk



이승원

Master / 삼성전자 SAIT

Processing In-Memory and
Near-Memory
(Accelerating Single Large AI
with Memory Centric Computing)

In this talk, we explore current trends affecting supercomputers (TOP500) and highlight case studies of large jobs at Samsung Advanced Institute of Technology (SAIT).

For Molecular Dynamics (MD) and Density Functional Theory (DFT) simulations, we use GPUs instead of CPUs for improved simulation speed up to 300 times.

In addition, we optimized a large language model (BERT) and we can overlap between computation and communications. Using a hyper-parameter auto-tuning method, we can enlarge the batch size and achieve the best time-to-result of BERT training on MLPerf Training v1.1 and v2.0 when we use 1024 NVIDIA® A100 GPUs.
(<https://github.com/SAITPublic/>).

Lastly, we also propose an OpenACC extension for Processing-in-memory (PIM) and some results of Processing-near-memory(PNM).



한양대학교
HANYANG UNIVERSITY