Energy–Performance Trade-offs in Privacy-Preserving Federated Learning on SmartNIC-Enabled HPC Systems

Scientific Achievement

- Studied energy-performance trade-offs of privacy-preserving federated learning (FL) on high-performance computing (HPC) systems.
- Evaluated effects of network architecture and FL server placement.
- Showed InfiniBand improves runtime and energy efficiency over Ethernet.

Significance and Impact

- Provides system-level insights for deploying privacy-preserving FL efficiently on accelerator-rich HPC systems.
- Shows that communication architecture is the dominant driver of runtime and energy efficiency in FL workloads.
- Enables energy-aware design of distributed AI workflows for scientific computing, privacy-sensitive data analysis, and multi-institution collaborations.

Technical Approach

- Evaluated FL across three communication configurations.
- Combined node power telemetry with FL logs.
- Analyzed privacy overhead and scaling effects using transformer models (ALBERT, DistilBERT, BERT) under differential privacy settings.

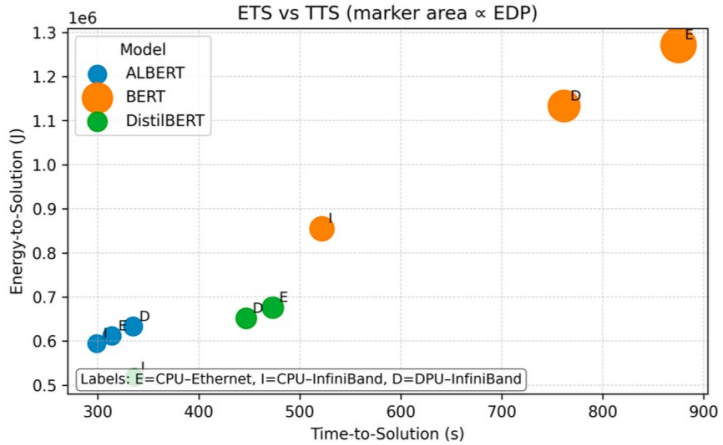

Figure 1: Energy-to-solution (ETS) versus time-to-solution (TTS) across models and configurations. Colors indicate transformer models and circle size indicate size of the model. Marker area is proportional to EDP.

PI(s)/Facility Lead(s): K. Kim (PI), O. Kotevska (ORNL's PI)
ASCR Program: AI for Science
ASCR PM: Xujing Davis
Publication(s) for this work: Kotevska, O., Duc, N., Hernandez, O. (2026). Energy–Performance Trade-offs in Privacy-Preserving Federated Learning on SmartNIC-Enabled HPC Systems. In IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW).