

# Scalable Federated Learning for Scientific Foundation Models on Leadership-Class Systems

## Scientific Achievement

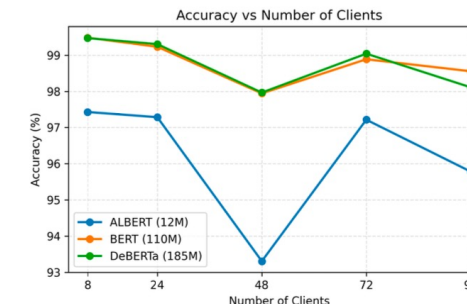
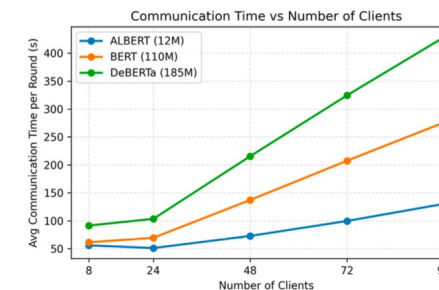
- Provided the first system-level study of federated fine-tuning of foundation models on an exascale supercomputer.
- Showed that model accuracy remains stable under heterogeneity, dropout, and differential privacy (DP) noise.
- Identified that communication and synchronization, not learning, are the main bottlenecks at scale.

## Significance and Impact

- Establishes a practical baseline for deploying federated learning (FL) on leadership-class HPC systems.
- Demonstrates that privacy-preserving FL is feasible at large scale without accuracy loss.
- Informs future system design by highlighting communication as the dominant scaling constraint.

## Technical Approach

- Conducted large-scale experiments with up to 96 clients on an exascale system using real workloads.
- Evaluated system behavior across client scale, model size, heterogeneity, dropout, and DP settings.
- Introduced system-aware strategies to reduce runtime overhead.



(a) Communication time per round vs. number of clients for ALBERT, BERT, and DeBERTa. (b) Test accuracy vs. number of clients for the three model sizes.

Figure 1: Effect of model size on FL scalability. Each curve averages over FedAvg, FedProx, and FedOpt algorithms and all data distributions. Communication cost scales roughly linearly with client count and proportionally with model size, while accuracy remains high for all three transformers.

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., Nguyen, T., da Silva, R., Engelmann, C., Balaprakash, P. (2026, April). Scalable Federated Learning for Scientific Foundation Models on Leadership-Class Systems. In The 6th Workshop on Machine Learning and Systems (EuroMLSys).