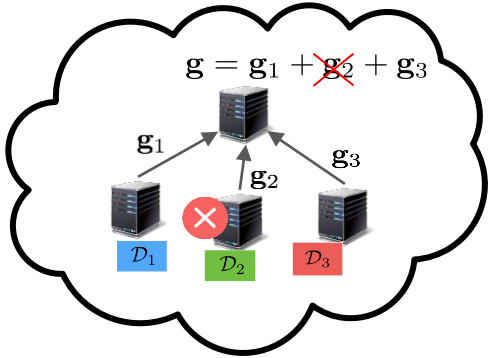
Large-Scale Distributed Learning

1. Distributed Learing in Cloud



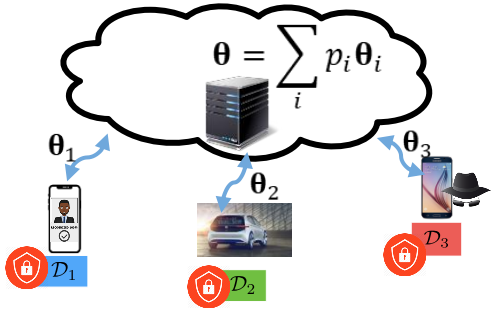
Challenges

- Communication bottleneck
- Straggling worker nodes

Proposal: CodedReduce [1]

- Novel ***tree topology*** for inter-cluster parallelization
- Smart algrebraic ***coding over gradients*** for straggler resiliency
- Gains of up to **30x** over prior benchmarks over AWS

2. Federated Learning



Challenges

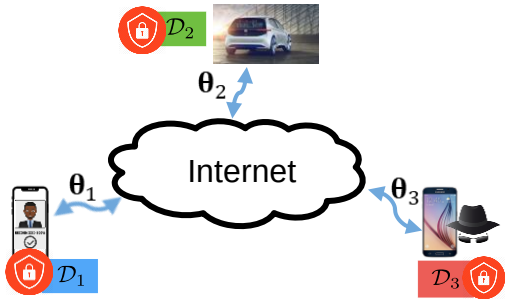
- Data privacy
- Data heterogeneity
- Model inversion attacks
- Byzantine attacks

Proposal: DiverseFL [2]

- Provides robustness against Byzantine attacks, particularly with ***non-IID*** client data
- Novel ***per client criteria*** for finding Byzantines
- ***TEE-asisted*** secure aggregation at server
- Outperforms prior SOTA for Byzantine mitigation, such as FLTrust (NDSS'21)

(More details in next 2 slides)

3. Decentralized Serverless Learning



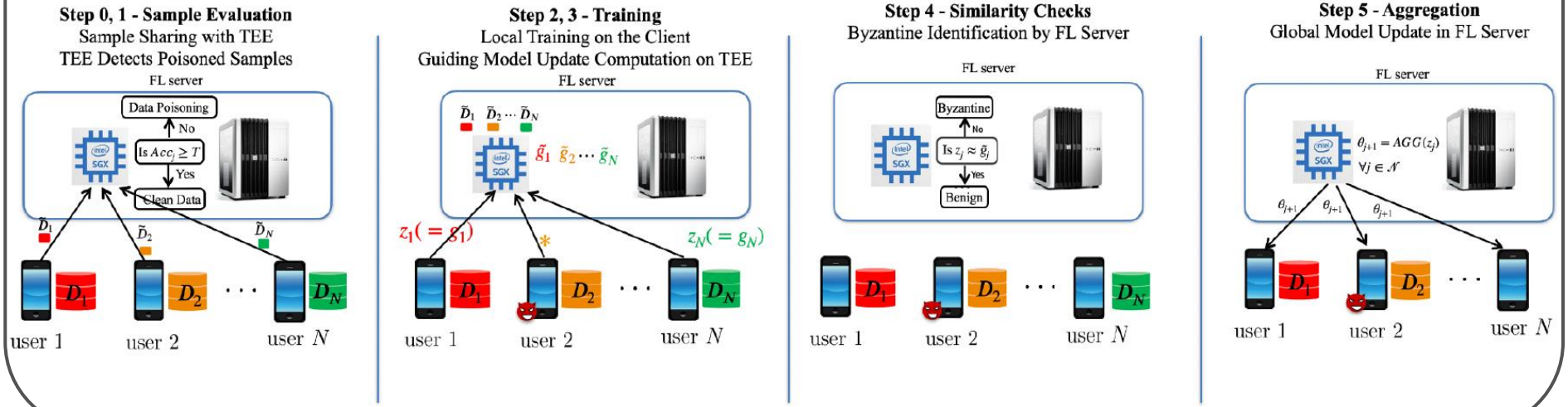
Challenges

- Lack of central coordinator
- Topology design for decentralized training
- Security against Byzantine attacks

Proposal: Basil [3]

- *Fast* and *computationally efficient* Byzantine-resilient decentralized training
- *Logical ring* topology for sequential training for low power edge devices
- Novel *performance-based approach* for Byzantine mitigation
- Superior performance with *non-IID data*, where prior SOTA fails completely

DiverseFL Overview:

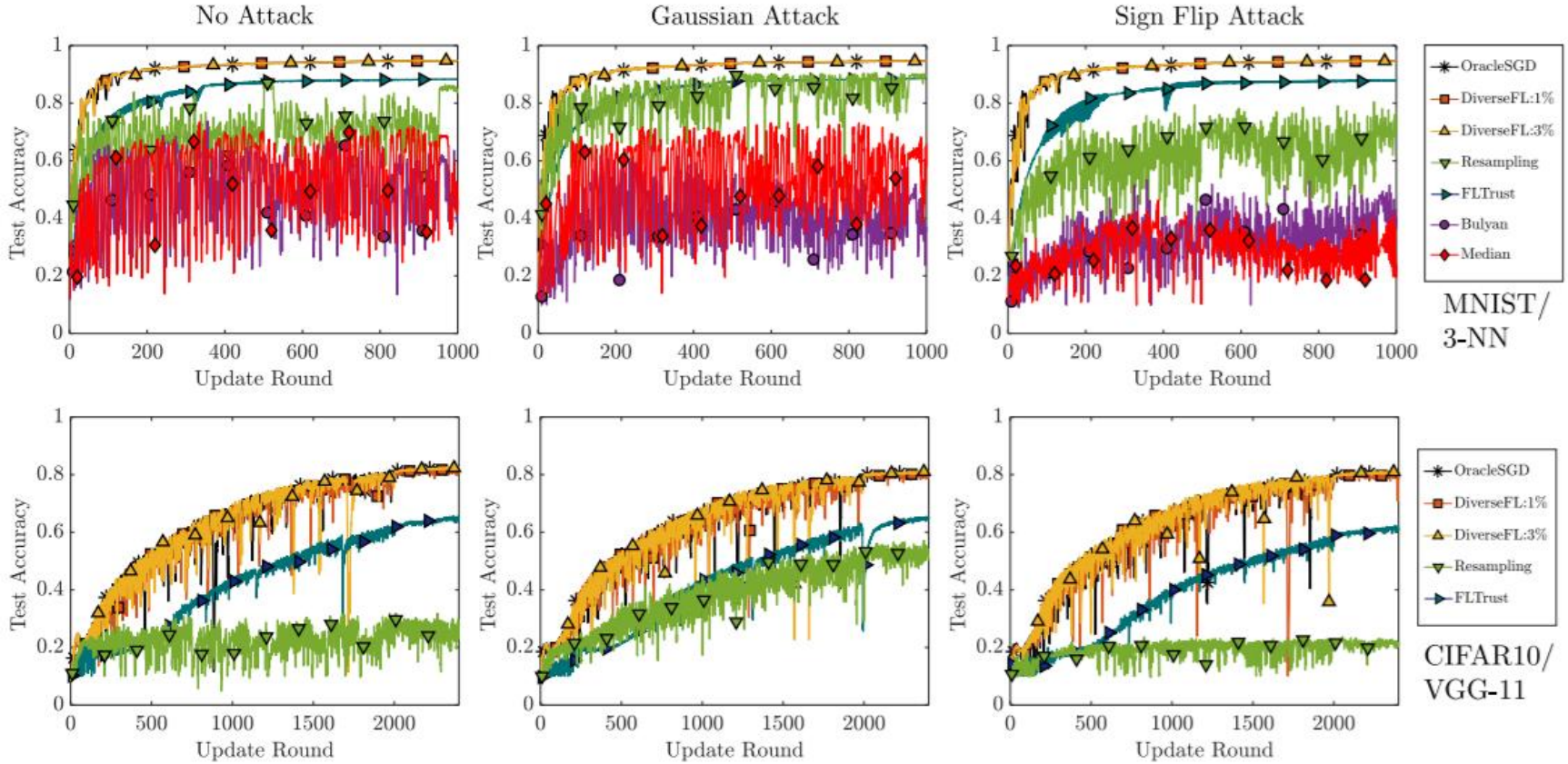


DiverseFL Convergence: For communication round i, and optimal global model θ^*, current model θ^i satisfies the following with high probability:

$$\left\|\theta^i - \theta^*\right\| \leq (1-\rho)^i \left\|\theta^0 - \theta^*\right\| + \frac{\alpha(2+\epsilon_2)(4\Gamma_1+\beta)}{\rho}$$

DiverseFL Performance in Practice

Setting:
Total 23 clients; 5 Byzantine nodes; non-IID data distribution



DiverseFL achieves **stable**, **near optimal** performance across multiple attacks and **datasets**