Accelerating Distributed ML Training via Selective Synchronization

Sahil Tyagi and Martin Swany
{**styagi**, **swany**}**@iu.edu**

Research Problem

- Distributed Data-parallel (DDP) training algorithms suffer different trade-offs
- *BSP* or *Bulk-Synchronous-Parallel* has high synchronization overhead
- *Federated Averaging* or *FedAvg* has low accuracy for both IID/Non-IID data from infrequent updates
- *Stale-Synchronous Parallel* or *SSP* also has low accuracy despite its bounded-asynchronous approach

Contributions

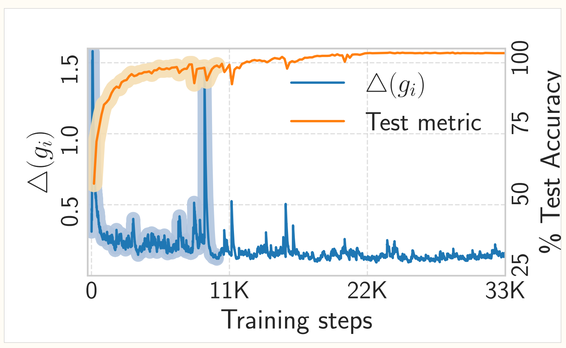
- **SelSync**, a semi-synchronous training technique that switches between synchronous and local-SGD based on importance of updates
- Propose a new data-partitioning scheme for semi-synchronous training.
- Show that local and global updates diverge more in parameter aggregation vs. gradient averaging.
- Develop randomized data-injection for Non-IID data training

Balancing parallel and statistical aspect of DDP Training

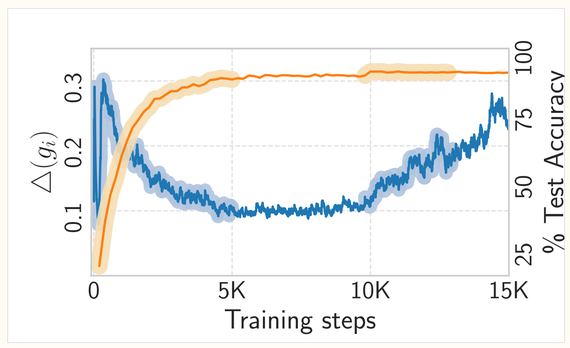
- Unlike traditional distributed computing, DDP also has a statistical aspect due to stochastic nature of SGD
- DDP has sensitive or critical regions in training
- Gradients get smaller and saturate, but magnitude and trajectory varies
- Crucial updates can be detected as:

$$\triangle(g_i) = \left| \frac{\mathbb{E}[||\nabla \mathscr{F}_{(i)}||^2] - \mathbb{E}[||\nabla \mathscr{F}_{(i-1)}||^2]}{\mathbb{E}[||\nabla \mathscr{F}_{(i-1)}||^2]} \right|$$

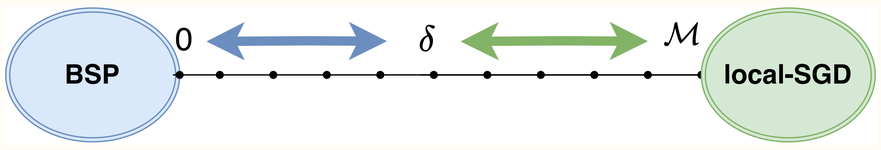
ResNet101



VGG11

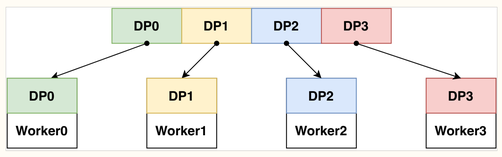


- **SelSync** update rule: For threshold δ, use BSP if $\triangle(g_i) \geq \delta$; else local-SGD

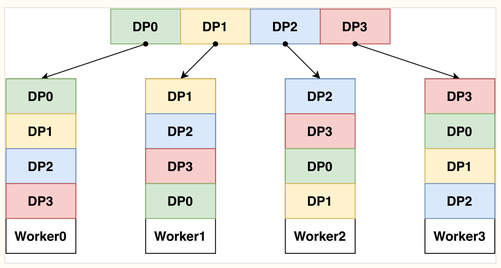


SelSync Data-Partitioning Scheme

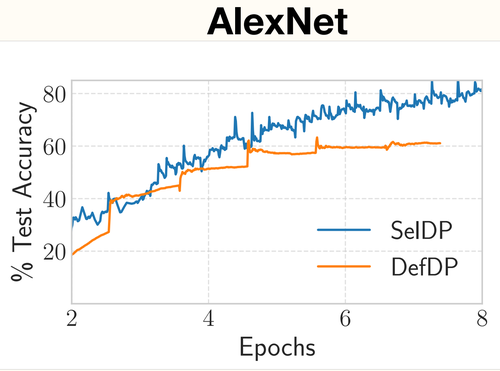
Default-DP



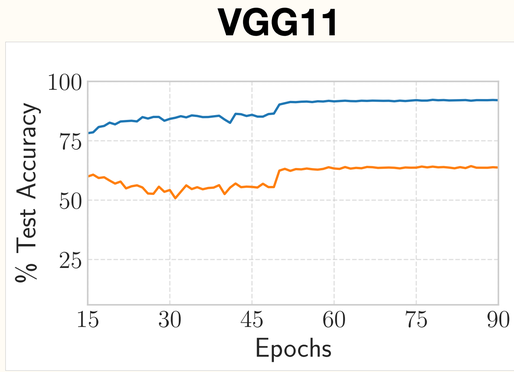
SelSync-DP



AlexNet



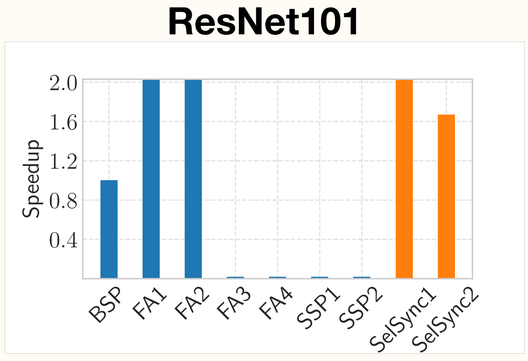
VGG11



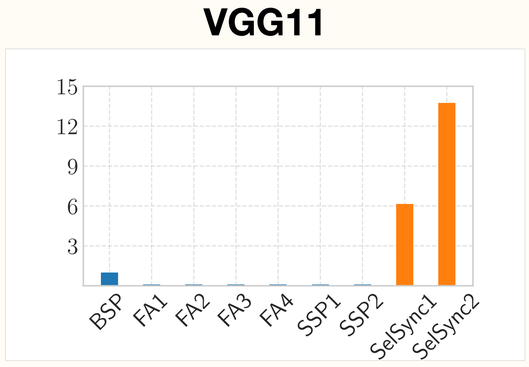
Evaluation over IID Data

- Tested on **ResNet101**, **VGG11**, **AlexNet** and **Transformer** to achieve at-least BSP-level accuracy/PPL: 98.5%, 90.9%, 85.15% and 90.0 respectively
- Compared with ***BSP***, ***FedAvg*** and ***SSP;*** FedAvg config.: (1,.25), (1,.125), (.5,.25), (.5,.125); SSP config.: staleness threshold 100 and 200; SelSync config.: δ=0.3 and 0.5

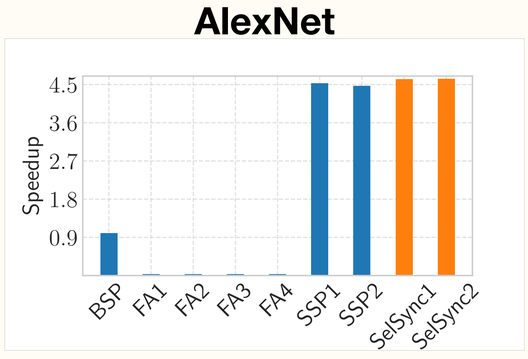
ResNet101



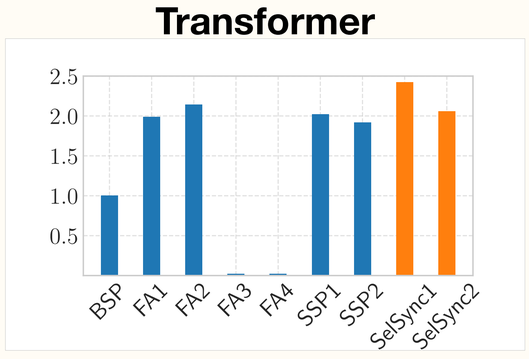
VGG11



AlexNet



Transformer



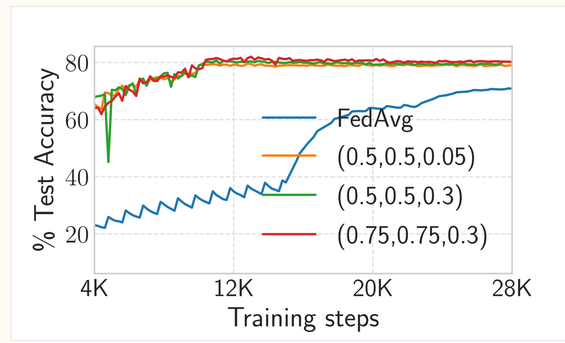
Evaluation over Non-IID and Unbalanced Data

- **Data-injection**: A random subset of workers communicate and share partial training samples to improve overall data distribution
- Configured as (α, β) where α is the *fraction of workers chosen* and β is the *fraction of batch-size to be shared*; (0.5, 0.5) implies half the workers share half the training samples at each iteration
- Batch-size is a sensitive hyperparameter so cumulatively should remain fixed

$$b' = \frac{b}{(1 + \alpha\beta N)}$$

ResNet101

VGG11



Conclusion

- SelSync attains similar model accuracy as BSP while speed-up training
- Threshold δ decides the **Local-To-Synchronous Step Ratio (LSSR)** and synchronizes critical updates only and training locally otherwise
- Sweet spot between fully synchronous and asynchronous training
- Data-injection in semi-synchronous training improves convergence