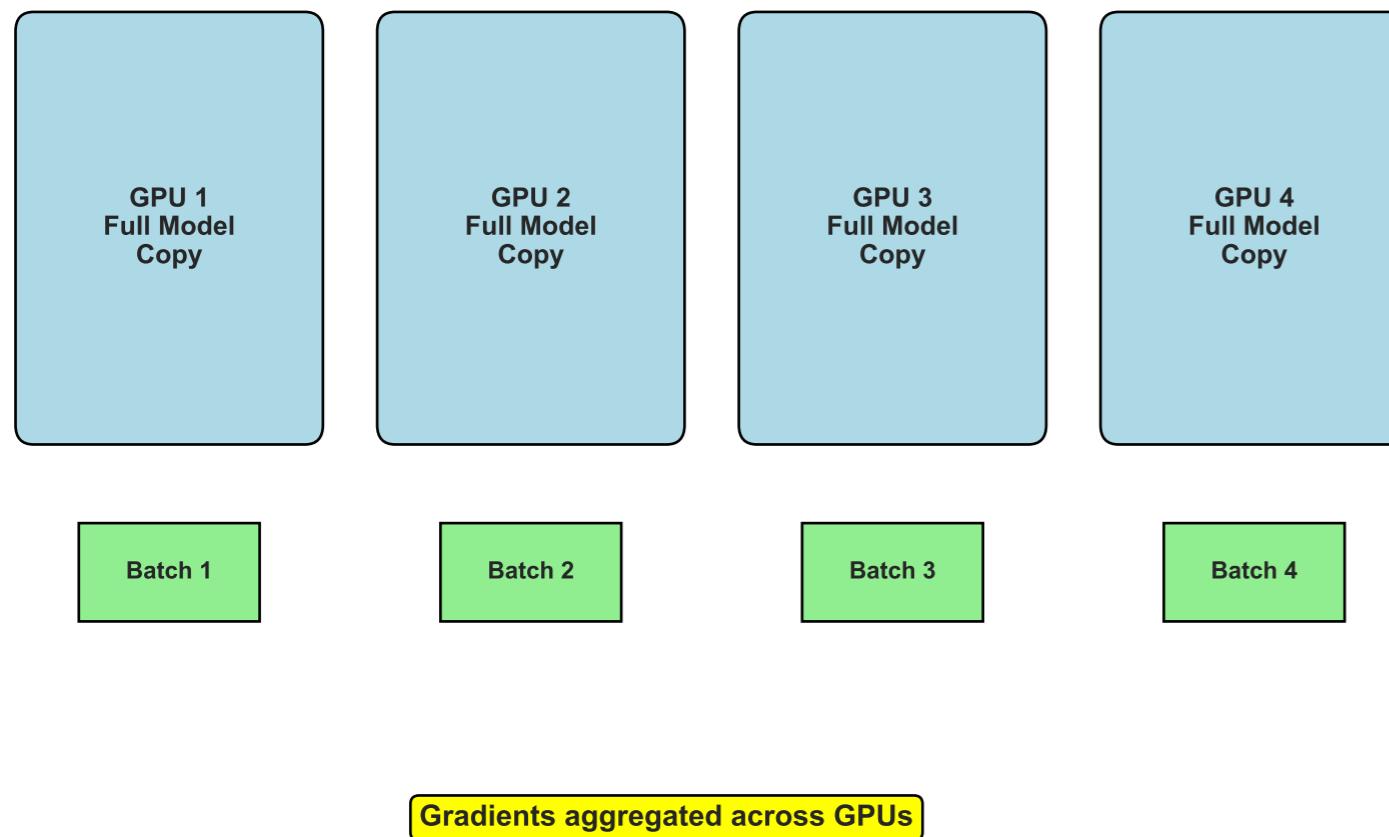
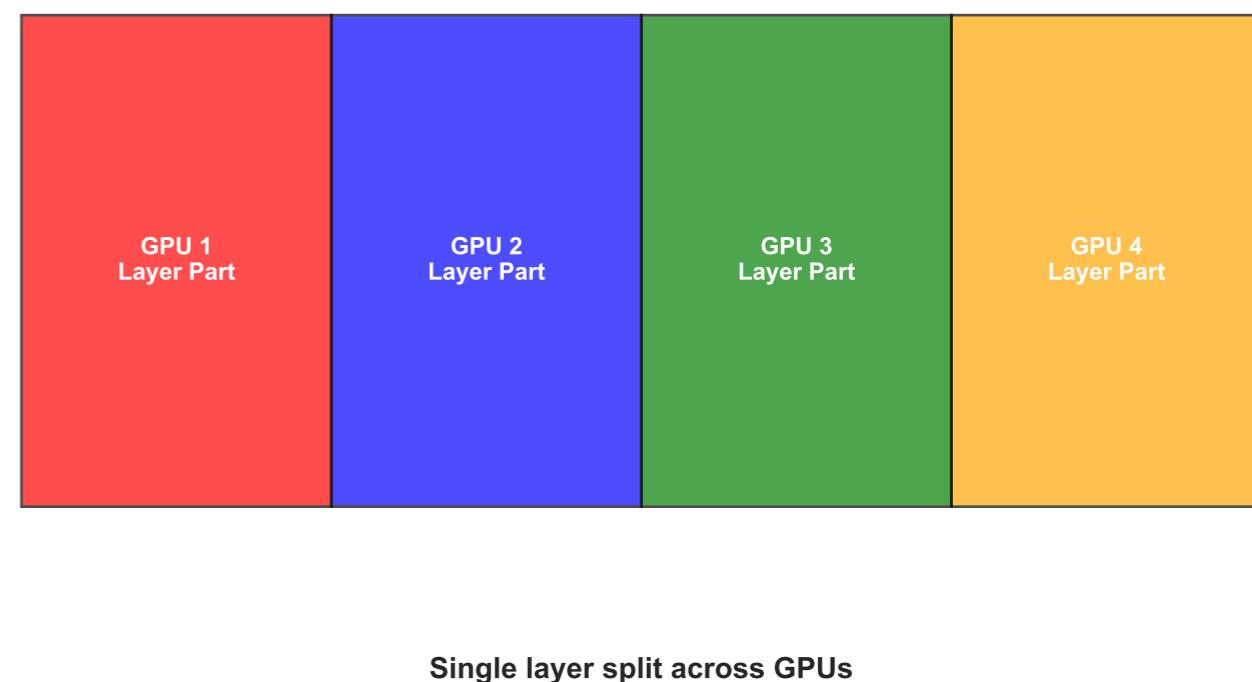


# Model Parallelism: Training at Scale

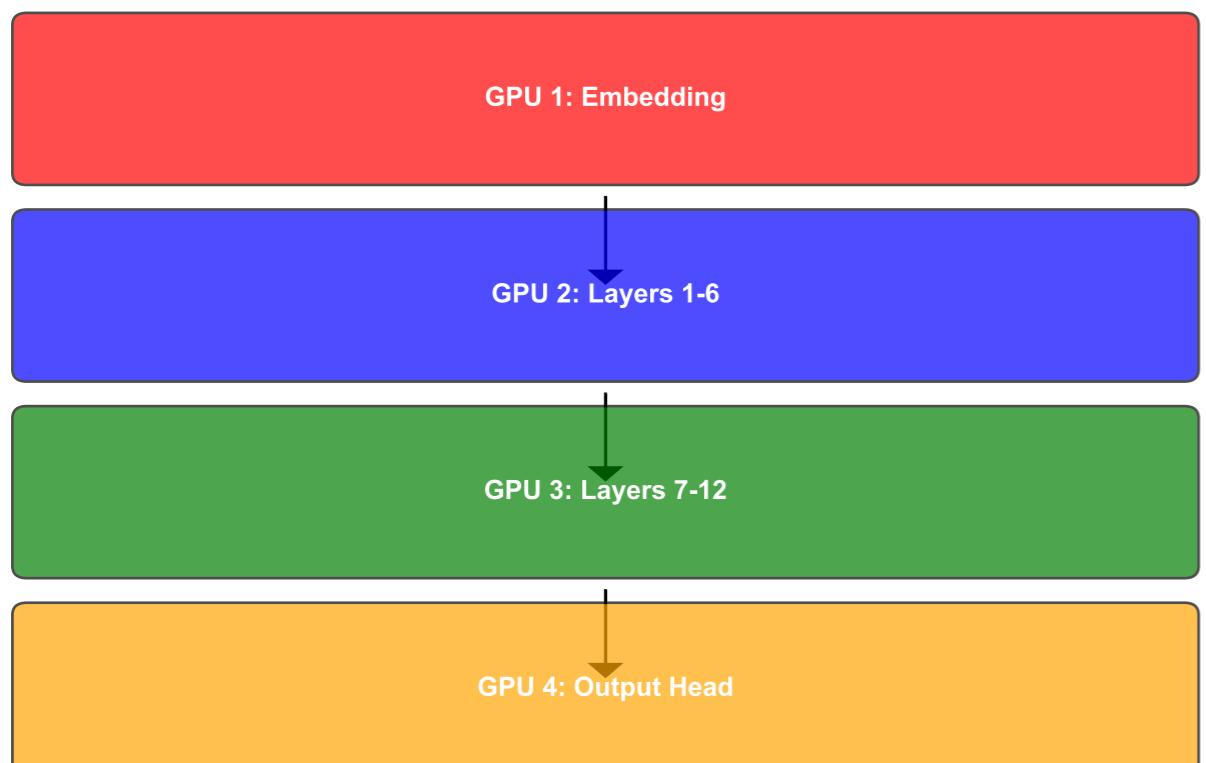
## Data Parallelism



## Tensor Parallelism



## Pipeline Parallelism



## Parallelism Strategy Comparison

