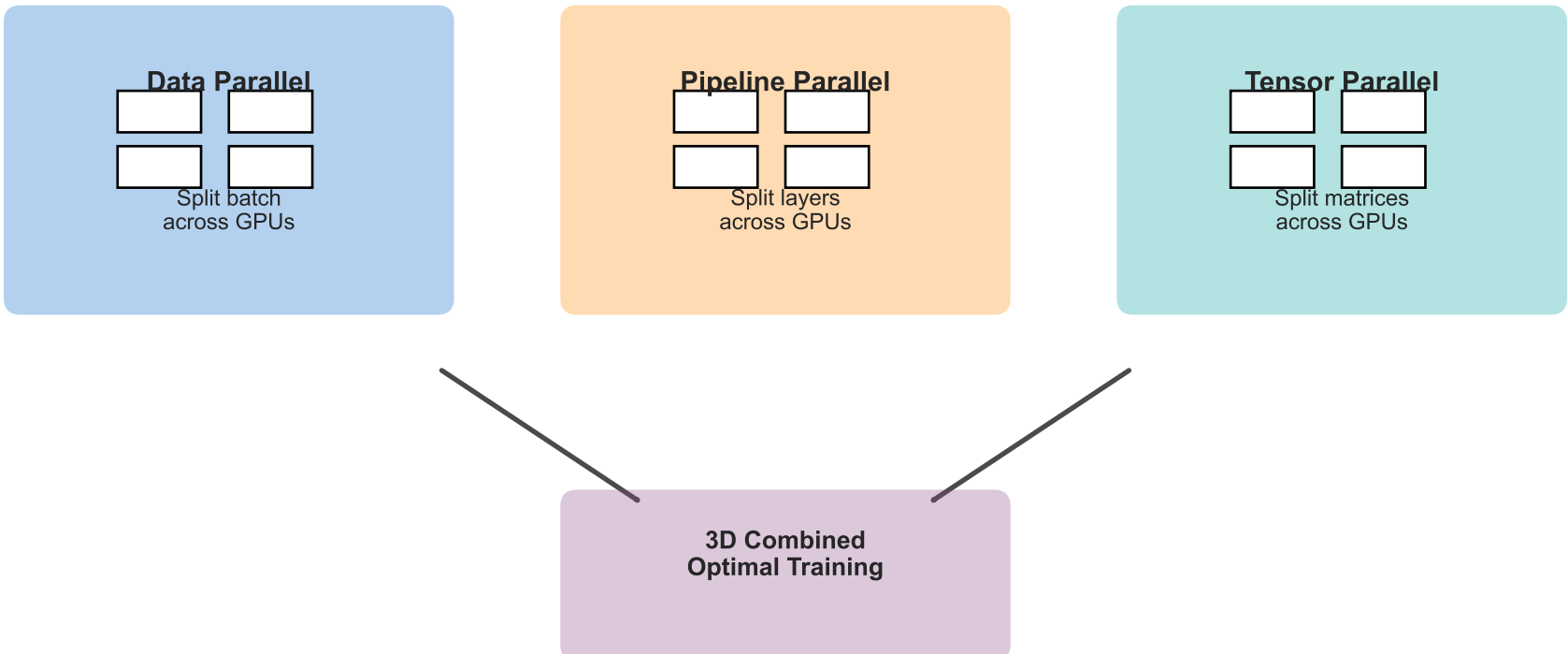3D Parallelism for Large Model Training