Desiderata	No Parallelism	Standard Model Parallelism	Pipelined Parallelism	Standard Task Parallelism	Hydra Shard Parallelism
In-Memory Model Training	✓	✓	✓	✓	✓
Arbitrary Single-Device Model Scaling					✓
Efficient Larger-than-Memory Model Training		✓	✓		✓
Efficient Training of Multiple In-Memory Models				✓	✓
Efficient Training of Multiple Larger-than-Memory Models					✓