

name

- Encoder-Decoder DeepSpeed 128 Experts
 - Encoder-Decoder MegaBlocks 128 Experts
- Encoder-Decoder Sliced 128 Experts