

GPU Memory Usage: 485B Model Training (8192 GPUs, A800-SXM4-80GB)

