Performance for Model: 8B, Seq Len: 8192 on H100 176 0.39 0.42 0.48 0.55 0.56 0.56 8.0 160 0.39 0.42 0.48 0.55 0.56 0.56 -0.7Host Memory (GB) 12 128 144 0.39 0.42 0.48 0.54 0.56 0.56 - 0.6 0.39 0.42 0.48 0.55 0.56 0.56 - 0.5 0.39 0.42 0.48 0.55 0.56 0.56 0.4 96 0.39 0.42 0.48 0.55 0.56 0.56 -0.380 0.50 0.56 0.52 0.53 0.55 0.56 40 78 30 50 60 70 Device Memory (GB)