## throughput Mixtral-8x7b-v0.1 6 $\delta$ $\omega$ $\delta$ Throughput (Higher is Better) 8 6 <sup>10</sup> <sub>20</sub> <sub>30</sub> $B_{atch}$ Size40 50 60