CPU Usage (GB) vs. parameter count (millions) google/gemma-3-270m-it on CPU 2.25 Qwen/Qwen2.5-0.5B-Instruct on CPU × meta-llama/Llama-3.2-1B-Instruct on CPU * 2.00 1.75 CPU Usage (GB) 1.50 1.25 1.00 0.75 0.50 300 500 400 600 700 800 900 1000 Parameter count (millions)