Impact of leading-chunk preservation Perplexity 102 Transformer 29.5 Attamba C8 **L8** 34.0 Attamba C8 L16 33.9 Attamba C8 37.2 WK2 Attamba C64 L64 35.7 Attamba C64 L16 Attamba C64 L1 4×10^{1} Attamba C128 L128 Attamba C128 L16 Attamba C128 L1 3×10^{1} 10^4 Global Step