Impact of leading-chunk preservation Perplexity 10² Transformer Attamba C8 **L8** Attamba C8 L16 6×10^{1} Attamba C8 **L**1 WK2 Attamba C64 L64 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