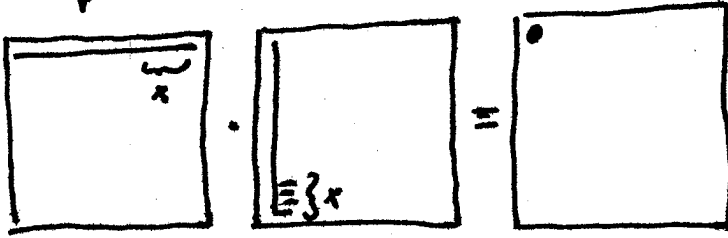


Why blocking?

assume: cache size $\ll n$
 cache line = 8 doubles
 only 1 cache

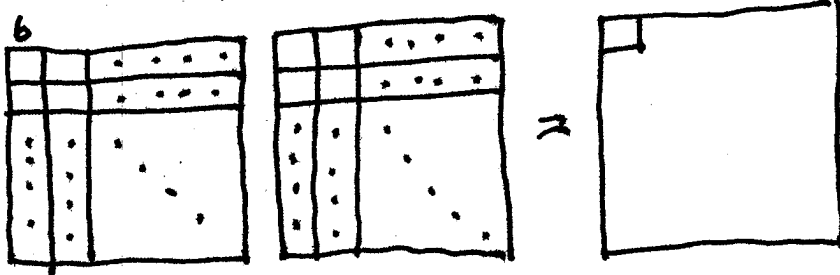
CIT = cache miss

1.) Triple loop MIMD



- entry: $\frac{n}{8} + 4$ CITs (compulsory)
 afterwards: x is in cache
 - entry: no reuse, so again $\frac{n}{8} + 4$ CITs
- \Rightarrow total = $(\frac{n}{8} + 4)n^2 = \frac{9}{8}n^3$ CITs

2.) blocked MIMD



choose: $b \geq 8$
 (cache line)
 and $81b$
 and $3b^2 \leq c$
 $c =$ cache size

- block: $\frac{nb}{8} + \frac{4b}{8} = \frac{4b}{4}$ CITs
 - block: same
- \Rightarrow total = $\frac{4b}{4} \cdot \left(\frac{n}{b}\right)^2 = \frac{4}{4b}n^3$
- choose $b = \sqrt{\frac{c}{3}} \Rightarrow \frac{\sqrt{3}}{4\sqrt{c}}n^3$ CITs
- gain: $\approx 2.5\sqrt{c}$

- Explains much of triple loops poor performance
- Blocking achieves both: data reuse
 neighbor use
 i.e. exhibits spatial and temporal locality