

|   |         | Serial | Deterministic | Nondeterministic |                     |
|---|---------|--------|---------------|------------------|---------------------|
|   |         |        |               | $n = 5$          | $n = 10$            |
| Naive MatrixMul (f32, 512×32×512) @ 28 SM's [9.14 blocks/SM]  |         |        |               |                  |                     |
| DRAM reads  | $t = 4$ | 0      | 0 (0%)        | 0 (0%)           | 0 (0%)              |
|   | $t = 8$ |        | 0 (0%)        | 0 (0%)           |                     |
| DRAM writes   | $t = 4$ | 0      | 0 (0%)        | 0 (0%)           | 0 (0%)              |
|   | $t = 8$ |        | 0 (0%)        | 0 (0%)           |                     |
| L1D hit rate  | $t = 4$ | 96.1%  | 96.1% (0%)    | 96.1% (0%)       | 96.1% (0.1%)        |
|   | $t = 8$ |        | 96.1% (0%)    | 96.2% (0.1%)     | 96.1% (0%)          |
| L2D hit rate  | $t = 4$ | 100%   | 100% (0%)     | 100% (0%)        | 100% (0%)           |
|   | $t = 8$ |        | 100% (0%)     | 100% (0%)        | 100% (0%)           |
| Cycles  | $t = 4$ | 50.2K  | 50.2K (0%)    | 50.4K (0.3%)     | 50.3K (0.2%)        |
|   | $t = 8$ |        | 50.2K (0%)    | 50.4K (0.3%)     | 50.4K (0.3%)        |
| Exec time   | $t = 4$ | 85.5s  | 29.8s (2.9x)  | 27.0s (3.2x)     | <b>25.9s (3.3x)</b> |
|   | $t = 8$ |        | 24.2s (3.5x)  | 21.1s (4.1x)     | <b>19.6s (4.4x)</b> |
| Naive MatrixMul (f32, 512×32×512) @ 112 SM's [2.29 blocks/SM] |         |        |               |                  |                     |
| DRAM reads  | $t = 4$ | 0      | 0 (0%)        | 0 (0%)           | 0 (0%)              |
|   | $t = 8$ |        | 0 (0%)        | 0 (0%)           | 0 (0%)              |
| DRAM writes   | $t = 4$ | 0      | 0 (0%)        | 0 (0%)           | 0 (0%)              |
|   | $t = 8$ |        | 0 (0%)        | 0 (0%)           | 0 (0%)              |
| L1D hit rate  | $t = 4$ | 96.8%  | 96.8% (0%)    | 96.7% (0%)       | 96.7% (0%)          |
|   | $t = 8$ |        | 96.8% (0%)    | 96.7% (0.1%)     | 96.7% (0%)          |
| L2D hit rate  | $t = 4$ | 100%   | 100% (0%)     | 100% (0%)        | 100% (0%)           |
|   | $t = 8$ |        | 100% (0%)     | 100% (0%)        | 100% (0%)           |
| Cycles  | $t = 4$ | 20.2K  | 20.2K (0%)    | 20.2K (0.2%)     | 20.2K (0.3%)        |
|   | $t = 8$ |        | 20.2K (0%)    | 20.2K (0.3%)     | 20.2K (0.3%)        |
| Exec time   | $t = 4$ | 122.4s | 40.9s (3x)    | 37.7s (3.2x)     | <b>37.3s (3.3x)</b> |
|   | $t = 8$ |        | 29.2s (4.2x)  | 25.9s (4.7x)     | <b>25.1s (4.9x)</b> |