Analyzer Report {Loop Blocking}

This report generated by Convolutional Neural Network Inference Analyzer (CNN-IA) to summarize the analysis needed to reach the optimal loop blocking for mlp_fc3_batch16 using restricted schedule space.

Memory Architecture:

	L0	L1	L2
Capacity	16	16384	536870912
Access cost	0.05	3.84	200.0
Static cost	0.0	0.0	0.0
Parallel count	256	1	1
Parallel mode	1	0	0
Parallel cost	2.0	0.0	0.0

Precision : 16
Minimum utilization : 0.0%
Outputs can be buffered by MAC : 0
Replication to improve utilization : True

Glossary:

- Memory Levels : (L0, L1, L2)
  The smallest index the nearest to CPU.

- Loop Notations : (FX, FY, OX, OY, OC, IC, ON)

FX : FILTER WIDTH
FY : FILTER HEIGHT
OX : OUTPUT WIDTH
OY : OUTPUT HEIGHT
OC : OUTPUT CHANNEL
IC : INPUT CHANNEL
ON : BATCH

Map Configuration

Loop Blocking (factors):

	L0	L1	L2
FX	3.0	1.0	1.0
FY	1.0	1.0	1.0
OX	1.0	1.0	1.0
OY	1.0	1.0	1.0
OC	3.0	5.0	5.0
IC	1.0	1.0	500.0
ON	1.0	16.0	1.0

Loop Partitioning (units):

	L0	L1	L2
FX	1.0	1.0	1.0
FY	3.0	1.0	1.0
OX	1.0	1.0	1.0
OY	13.0	1.0	1.0
OC	4.0	1.0	1.0
IC	1.0	1.0	1.0
ON	1.0	1.0	1.0

Loop Ordering (from the innermost):

	L0	L1	L2
FX	0.0	6.0	6.0
FY	1.0	6.0	6.0
OX	6.0	6.0	6.0
OY	2.0	6.0	6.0
OC	3.0	1.0	1.0
IC	6.0	6.0	0.0
ON	6.0	0.0	6.0

(Hinted schedule configurations are in green)

Schedule

MEM - L2:
for (OC, 5b, 1p)
for (IC, 500b, 1p)

MEM - L1:
for (OC, 5b, 1p)
for (ON, 16b, 1p)

MEM - L0:
for (OC, 3b, 4p)
for (OY, 1b, 13p)
for (FY, 1b, 3p)
for (FX, 3b, 1p)
spatially unrolled loops: (FX)(FY)(OY)(OC)

Cost

MEM	ENERGY (PJ)
L0	399800.0
L0-PARA	20050000.0
L1	16592640.0
L2	33800000.0
TOTAL	70842440.0