

## LEGEND



← Tensor  
distribution



← Tensor



← Gradient  
distribution

AR: AllReduce

T : Transpose

Processor grid:  
2D Cartesian



Split in y &  
Replicate in x:



2D distribution:  
(for all A-shape data)



Split in x &  
Partial sum in y:



Split in x &  
Replicate in y:



Split in y &  
Partial sum in x:

