



Device 1 (rank = 1)

$x' = \text{split}(x, \text{sharding_ratios}[\text{rank}])$

$z' = \text{matmul}(x', w)$

$z = \text{het_all_gather}(z', \text{rank})$

Device 2 (rank = 2)

$x' = \text{split}(x, \text{sharding_ratios}[\text{rank}])$

$z' = \text{matmul}(x', w)$

$z = \text{het_all_gather}(z', \text{rank})$