

Device 1 (rank = 1)

x' = split(x)[rank] w' = split(w)[rank] z' = matmul(x', w') z = all\_reduce(z', rank)

Device 2 (rank = 2)

x' = split(x)[rank] w' = split(w)[rank] z' = matmul(x', w') z = all\_reduce(z', rank)