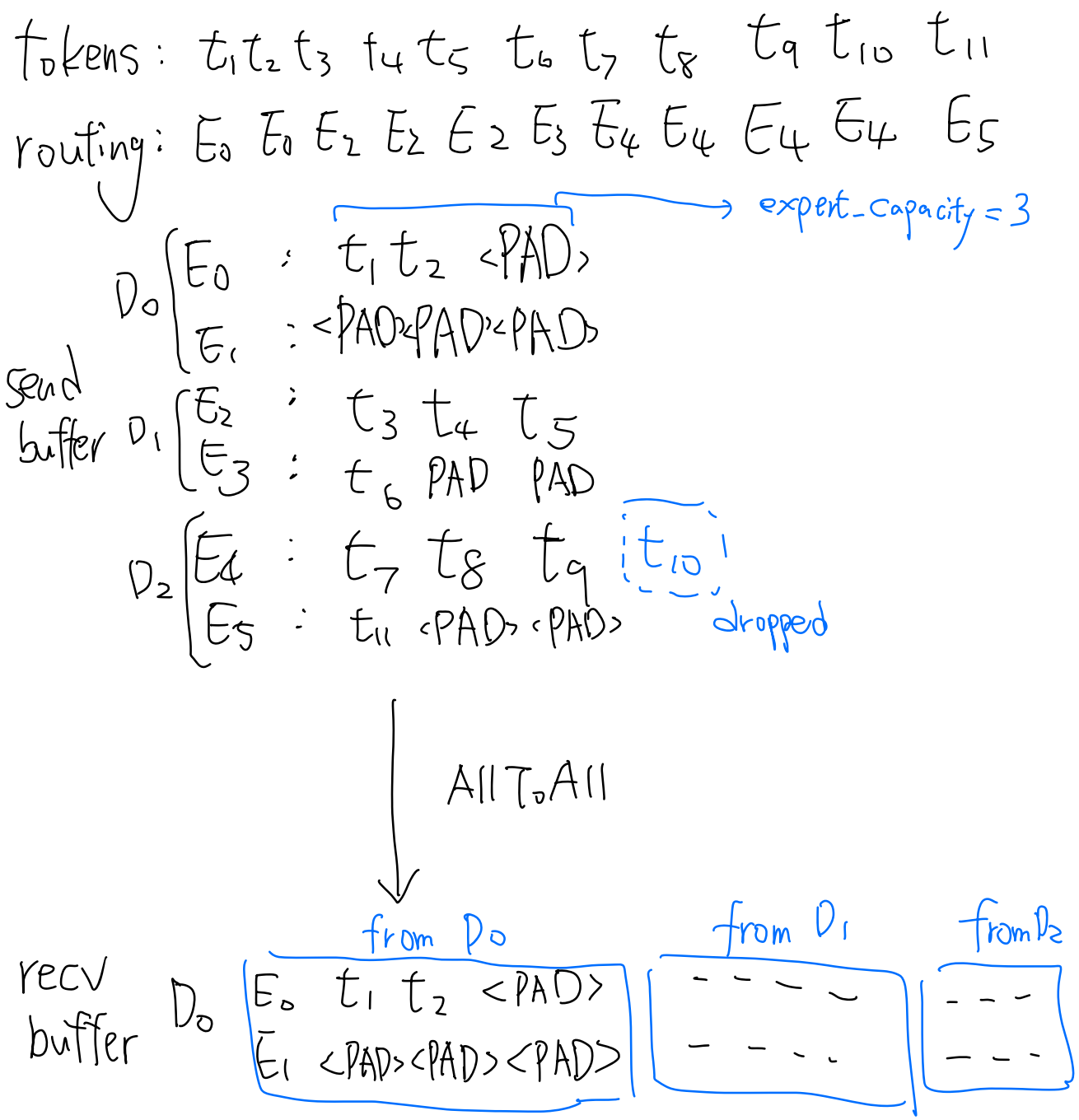
tokens: t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11}
routing: E_0 E_0 E_2 E_2 E_2 E_3 E_4 E_4 E_4 E_4 E_5
expert_capacity = 3
send buffer
D_0 E_0 : t_1 t_2 <PAD>
E_1 : <PAD> <PAD> <PAD>
D_1 E_2 : t_3 t_4 t_5
E_3 : t_6 PAD PAD
D_2 E_4 : t_7 t_8 t_9
E_5 : t_{11} <PAD> <PAD>
t_{10}
dropped
All To All
recv buffer
D_0
from D_0
E_0 t_1 t_2 <PAD>
E_1 <PAD> <PAD> <PAD>
from D_1
from D_2