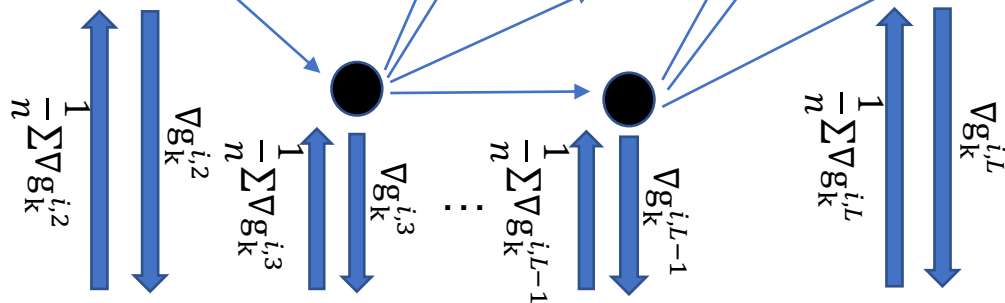


Back-propagation updates the model parameters of each layer

Input data  $\hat{x}_j$   
coming from  
a mini-batch  
at node i

Layer 1 ← Layer 2 ← Layer 3 ← ... ← Layer L-1 ← Output Layer

Forward pass  
produces output  $\hat{y}_j$   
and calculates loss



Communication Backend