



$$\text{softmax} \left( \frac{\mathbf{Q} \times \mathbf{K}^T}{\sqrt{d}} \right) \times \mathbf{V} \quad \text{FP Matrix Mult-Adds}$$





