



$$\begin{pmatrix} \mathbf{u}_{\text{energy}} \end{pmatrix} \left[ \begin{pmatrix} \text{key} \\ \mathbf{K} \end{pmatrix} \begin{pmatrix} U_{\text{key}} \end{pmatrix} + \begin{pmatrix} \text{query} \\ \mathbf{x}_t \\ X_t \end{pmatrix} \begin{pmatrix} U_{\text{query}} \end{pmatrix} \right] = \begin{pmatrix} \mathbf{e}_t \end{pmatrix}$$

$$\begin{pmatrix} \sigma(\mathbf{e}_t) \end{pmatrix} \begin{pmatrix} \text{value} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} \alpha_t \end{pmatrix} \begin{pmatrix} \text{value} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} \gamma_t \end{pmatrix}$$