1 Problem statement (global)

1.1 Version 2

Input:

- Set $\mathcal{R} = (R_1, R_2, ..., R_m)$ representing available resources, where each resource $R_k \in \mathcal{R}$ has capacity c_k.
- Value $c^{(GPU)}$ representing GPU capacity.
- Set of tasks $\mathcal{T} = (T_1, T_2, ..., T_n)$.
- Values $p_{i,k}$, $a_{i,k}$ and $b_{i,k}$ representing processing time, slope and intercept of task T_i on resource R_k. If task T_i cannot be executed on resource R_k, then $p_{i,k} = \infty$.
- Values $p_{i,k}^{(GPU)}$, $a_{i,k}^{(GPU)}$ and $b_{i,k}^{(GPU)}$ representing processing time, slope and intercept of task T_i on resource R_k if the task is offloaded to GPU. If task T_i cannot be executed on resource R_k, or cannot be offloaded to GPU, then $p_{i,k}^{(GPU)} = \infty$.
- Value h representing major frame length.

ILP model:

$$\min \sum_{W_j \in \mathcal{W}} \left(\sum_{T_i \in \mathcal{T}} \sum_{R_k \in \mathcal{R}} (A_{i,j,k} + A_{i,j,k}^{(GPU)}) + \max_{\substack{T_i \in \mathcal{T} \\ R_k \in \mathcal{R}}} \{B_{i,j,k}, B_{i,j,k}^{(GPU)}\} \right) \cdot \frac{1}{h} \tag{1}$$

$$x_{i,j,k} = 1 \implies A_{i,j,k} = 0.6 \cdot a_{i,k} \cdot l_j \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{2}$$

$$x_{i,j,k}^{(GPU)} = 1 \implies A_{i,j,k}^{(GPU)} = 0.6 \cdot a_{i,k}^{(GPU)} \cdot l_j \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{3}$$

$$x_{i,j,k} = 1 \implies B_{i,j,k} = 0.6 \cdot b_{i,k} \cdot l_j \quad \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{4}$$

$$x_{i,j,k}^{(GPU)} = 1 \implies B_{i,j,k}^{(GPU)} = 0.6 \cdot b_{i,k}^{(GPU)} \cdot l_j \quad \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{5}$$

$$\sum_{W_j \in \mathcal{W}} \sum_{R_k \in \mathcal{R}} (x_{i,j,k} + x_{i,j,k}^{(GPU)}) = 1 \quad \forall T_i \in \mathcal{T} \tag{6}$$

$$\sum_{W_j \in \mathcal{W}} l_j \leq h \tag{7}$$

$$l_j \geq \frac{x_{i,j,k} \cdot p_{i,k} + x_{i,j,k}^{(GPU)} \cdot p_{i,k}^{(GPU)}}{0.6} \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{8}$$

$$\sum_{T_i \in \mathcal{T}} (x_{i,j,k} + x_{i,j,k}^{(GPU)}) \leq c_k \quad \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{9}$$

$$\sum_{T_i \in \mathcal{T}} \sum_{R_k \in \mathcal{R}} x_{i,j,k}^{(GPU)} \leq c^{(GPU)} \quad \forall W_j \in \mathcal{W} \tag{10}$$

$$x_{i,j,k} = 0 \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} : p_{i,k} = \infty \tag{11}$$

$$x_{i,j,k}^{(GPU)} = 0 \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} : p_{i,k}^{(GPU)} = \infty \tag{12}$$

$$x_{i,j,k} \in \{0,1\}, x_{i,j,k}^{(GPU)} \in \{0,1\} \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{13}$$

$$A_{i,j,k} \in \mathbb{R}_0^+, A_{i,j,k}^{(GPU)} \in \mathbb{R}_0^+ \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{14}$$

$$B_{i,j,k} \in \mathbb{R}_0^+, B_{i,j,k}^{(GPU)} \in \mathbb{R}_0^+ \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{15}$$

$$l_j \in \mathbb{N} \quad \forall W_j \in \mathcal{W} \tag{16}$$

1.2 Version 1

Input:

- Set $\mathcal{R} = (R_1, R_2, ..., R_m)$ representing available resources, where each resource $R_k \in \mathcal{R}$ has capacity c_k.
- Value $c^{(GPU)}$ representing GPU capacity.
- Set of tasks $\mathcal{T} = (T_1, T_2, ..., T_n)$.
- Values $p_{i,k}$ and $E_{i,k}$ representing processing time and energy consumption of task T_i on resource R_k. If task T_i cannot be executed on resource R_k, then $p_{i,k} = \infty$.
- Values $p_{i,k}^{(GPU)}$ and $E_{i,k}^{(GPU)}$ representing processing time and energy consumption of task T_i on resource R_k if the task is offloaded to GPU. If task T_i cannot be executed on resource R_k, or cannot be offloaded to GPU, then $p_{i,k}^{(GPU)} = \infty$.
- Value h representing major frame length.

Note:

$$E_{i,k} = \left(a_{i,k} + \frac{b_{i,k}}{c_k} \right) \cdot p_{i,k} \tag{17}$$

Where $a_{i,k}$ is the slope and $b_{i,k}$ is the intercept of the approximation function.

ILP model:

$$\min \sum_{T_i \in \mathcal{T}} \sum_{W_j \in \mathcal{W}} \sum_{R_k \in \mathcal{R}} (a_{i,j,k} \cdot E_{i,k} + a_{i,j,k}^{(GPU)} \cdot E_{i,k}^{(GPU)}) \quad \text{subject to:} \tag{18}$$

$$\sum_{W_j \in \mathcal{W}} \sum_{R_k \in \mathcal{R}} (a_{i,j,k} + a_{i,j,k}^{(GPU)}) = 1 \quad \forall T_i \in \mathcal{T} \tag{19}$$

$$\sum_{W_j \in \mathcal{W}} l_j \leq h \tag{20}$$

$$l_j \geq \frac{a_{i,j,k} \cdot p_{i,k} + a_{i,j,k}^{(GPU)} \cdot p_{i,k}^{(GPU)}}{0.6} \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{21}$$

$$\sum_{T_i \in \mathcal{T}} (a_{i,j,k} + a_{i,j,k}^{(GPU)}) \leq c_k \quad \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{22}$$

$$\sum_{T_i \in \mathcal{T}} \sum_{R_k \in \mathcal{R}} a_{i,j,k}^{(GPU)} \leq c^{(GPU)} \quad \forall W_j \in \mathcal{W} \tag{23}$$

$$a_{i,j,k} = 0 \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} : p_{i,k} = \infty \tag{24}$$

$$a_{i,j,k}^{(GPU)} = 0 \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} : p_{i,k}^{(GPU)} = \infty \tag{25}$$

$$a_{i,j,k} \in \{0,1\}, a_{i,j,k}^{(GPU)} \in \{0,1\} \quad \forall T_i \in \mathcal{T}, \forall W_j \in \mathcal{W}, \forall R_k \in \mathcal{R} \tag{26}$$

$$l_j \in \mathbb{N} \quad \forall W_j \in \mathcal{W} \tag{27}$$