The definition of a schedule:

Denote the computation time of node i in dependency graph G as p_i.
Denote the communication time between task i and task j when they are on different machines to be c_{ij}
Schedule S assigns each task i a machine M_i and a starting time t_i.
Schedule S is legal (possible) when:

(1) $t_i \geq 0$

(2) $t_i + p_i \leq t_j$ or $t_j + p_j \leq t_i$ when task i and task j are on the same machine

(3) $t_i + p_i \leq t_j$ when task i and task j are on the same machine, and there is an edge $i \rightarrow j$ in the dependency graph G

(4) $t_i + p_i + c_{i\rightarrow j} \leq t_j$ when task i and task j are on different machine, and there is an edge $i \rightarrow j$ in the dependency graph G

The scheduling problem for tensorflow graph is different from traditional scheduling problem. After the forward pass on the tensorflow graph G, the back propagation happens in G with reverse dependency.
According to the book "Neural Network and Deep Learning", at the same node in graph G, back propagation takes the same amount of time as the forward pass.
We can model the back propagation process with a graph G' with the same vertices and reverse edges. Denote the computation time of a node j in back propagation p'_j and the communication time between node j and node i in back propagation c'_{ji}
Let $C_0 = max\left(max_{j\in V(G)}\frac{p'_j}{p_j}, max_{i\rightarrow j\in E(G)}\frac{c'_{ji}}{c_{ij}}\right)$
We prove that by scheduling the nodes in back propagation on the same machine as the forward pass, the total makespan of 1 tensorflow iteration $\leq (C_0+1)makespan(S)$, where S is a forward iteration schedule. Furthermore, when the ratio of computational and communication time between forward pass and back propagation is constant in graph G, our algorithm gives the same approximation ratio for makespan as the schedule for forward pass S.

We prove below that:

- Given a schedule S for G on m processors, we can generate a schedule S' for G' on m processors where $makespan(S') \leq C_0 makespan(S)$.
- Given constant computation and communication ratio between forward and backward pass, the optimal schedule for G' takes C_0 times the amount of time for optimal schedule for G

Lemma 1: Given a schedule S for G on m processors, exists a schedule S' for G' on m processors with $makespan(S') \leq C_0 makespan(S)$
Denote the starting time of task i in schedule S as t_i
We know that $\forall i \rightarrow j \in E(G), c'_{j\rightarrow i} \leq C_0 c_{i\rightarrow j}$ and $\forall j \in V(G), p'_j \leq C_0 p_j$

Denote the makespan of schedule S as T.
We create the schedule S' as follows:

- put task i on the same machine as in S $\forall$ i
- give task i a starting time $t'_i = C_0(T - t_i - p_i)$

. We prove below that the schedule S' is legal:

(1) We know $T \leq t_i + p_i \forall i$
Therefore $t'_i = C_0(T - t_i - p_i) \geq 0 \forall i$

(2) Let i, j be 2 arbitrary tasks assigned to the same machine.
Since S is a legal schedule, we know that $t_i + p_i \leq t_j$ or $t_j + p_j \leq t_i$
$\therefore T - t_i - p_i \geq T - t_j - p_j + p_j$ or $T - t_j - p_j \geq T - t_i - p_i + p_i$
$\therefore C_0(T - t_i - p_i) \geq C_0(T - t_j - p_j) + C_0 p_j$ or $C_0(T - t_j - p_j) \geq C_0(T - t_i - p_i) + C_0 p_i$
$\therefore t'_i \geq t'_j + p'_j$ or $t'_j \geq t'_i + p'_i$

(3) Let $j \to i$ be an arbitrary edge in G' and i, j are assigned to the same machine in S
$i \to j \in E(G) <=> j \to i \in E(G')$
Since S is a legal schedule, we know that $i \to j \in E(G)$, $t_i + p_i \leq t_j$
$\therefore C_0(T - t_i - p_i) \geq C_0(T - t_j - p_j) + C_0 p_j$
$\therefore t'_i \geq t'_j + p'_j$ in G'

(4) Let $j \to i$ be an arbitrary edge in G' and i, j are assigned on different machines.
$i \to j \in E(G) <=> j \to i \in E(G')$
Since S is a legal schedule, we know that $i \to j \in E(G)$, $t_i + p_i + c_{i\to j} \leq t_j$
$\therefore C_0(T - t_i - p_i) - C_0 c_{i\to j} \geq C_0(T - t_j - p_j) + C_0 p_j$
$\therefore C_0(T - t_i - p_i) \geq C_0(T - t_j - p_j) + C_0 p_j + C_0 c_{i\to j}$
$\therefore t'_i \geq t'_j + p'_j + c'_{j\to i}$ in G'

S' satisfy all condition for a legal schedule and therefore is legal.
$makespan(S') = max_i(t'_i + p'_i) = max_i(C_0(T - t_i)) \leq C_0 T = C_0 makespan(S)$

Lemma 2: Given computation and communication time constant ratio between forward and backward pass, the makespan of optimal schedule for G' is C_0 times the makespan for optimal schedule for G

We know that $\forall j \in V(G'), \frac{p'_j}{p_j} = C_0$, and $\forall j \to i \in E(G'), \frac{c'_{ji}}{c_{ij}} = C_0$
Therefore $max\Big(max_{j\in V(G')}\frac{p_j}{p'_j}, max_{j\to i \in E(G')}\frac{c_{ij}}{c'_{ji}}\Big) = \frac{1}{C_0}$
Let OPT be the shortest makespan for any schedule on G with m processors. Let S^* be an optimal schedule for G
Let OPT' be the shortest makespan for any schedule on G with m processors. Let S'^* be an optimal schedule for G'
Since By **Lemma 1** we know that there is a schedule S' for G' where $makespan(S') \leq C_0 makespan(S^*)$, therefore $OPT' \leq makespan(S') \leq C_0 makespan(S^*) = C_0 OPT$
For the same reason $OPT \leq \frac{1}{C_0} OPT'$.
Therefore $OPT' = C_0 OPT$