To design an optimal solution, we apply a general technique that is known as *Deferred Merge Embedding* (DME) by researchers in the VLSI community. It's a greedy algorithm that works as follows. Let v denote the root, with v' and v'' its two children. Let d' denote the maximum root-to-leaf distance over all leaves that are descendants of v', and let d'' denote the maximum root-to-leaf distance over all leaves that are descendants of v''. Now:

- If $d' > d''$, we add $d' - d''$ to the length of the v-to-v'' edge and add nothing to the length of the v-to-v' edge.
- If $d'' > d'$, we add $d'' - d'$ to the length of the v-to-v' edge and add nothing to the length of the v-to-v'' edge.
- $d' = d''$ we add nothing to the length of either edge below v.

We now apply this procedure recursively to the subtrees rooted at each of v' and v''.

Let T be the complete binary tree in the problem. We first develop two basic facts about the optimal solution, and then use these in an "exchange" argument to prove that the DME algorithm is optimal.

(i) Let w be an internal node in T, and let e', e'' be the two edges directly below w. If a solution adds non-zero length to both e' and e'', then it is not optimal.

Proof. Suppose that $\delta' > 0$ and $\delta'' > 0$ are added to $\ell_{e'}$ and $\ell_{e''}$ respectively. Let $\delta = \min(\delta', \delta'')$. Then the solution which adds $\delta' - \delta$ and $\delta'' - \delta$ to the lengths of these edges must also have zero skew, and uses less total length. ■

(ii) Let w be a node in T that is neither the root nor a leaf. If a solution increases the length of every path from w to a leaf below w, then the solution is not optimal.

Proof. Suppose that $x_1, \ldots, x_k$ are the leaves below w. Consider edges e in the subtree below w with the following property: the solution increases the length of e, and it does not increase the length of any edge on the path from w to e. Let F be the set of all such edges; we observe two facts about F. First, for each leaf x_i, the first edge on the w-x_i path whose length has been increased must belong to F, (and no other edge on this path can belong to F); thus there is exactly one edge from F on every w-x_i path. Second, $|F| \geq 2$, since a path in the left subtree below w shares no edges with a path in the right subtree below w, and yet each contains an edge of F.

Let e_w be the edge entering w from its parent (recall that w is not the root). Let δ be the minimum amount of length added to any of the edges in F. If we subtract δ from the length added to each edge in F, and add δ to the edge above w, the length of all root-to-leaf paths remains the same, and so the tree remains zero-skew. But we have subtracted $|F|\delta \geq 2\delta$ from the total length of the tree, and added only δ, so we get a zero skew tree with less total length. ■

We can now prove a somewhat stronger fact than what is asked for.

[1] ex179.790.171

(iii) The DME algorithm produces the unique optimal solution.

Proof. Consider any other solution, and let v be any node of T at which the solution does not add lengths in the way that DME would. We use the notation from the problem, and assume without loss of generality that $d' \geq d''$. Suppose the solution adds δ' to the edge (v, v') and δ'' to the edge (v, v'').

If $\delta'' - \delta' = d' - d''$, then it must be that $\delta' > 0$ or else the solution would do the same thing as DME; in this case, by (i) it is not optimal. If $\delta'' - \delta' < d' - d''$, then the solution will still have to increase the length of the path from v'' to each of its leaves in order to make the tree zero-skew; so by (ii) it is not optimal. Similarly, if $\delta'' - \delta' > d' - d''$, then the solution will still have to increase the length of the path from v' to each of its leaves in order to make the tree zero-skew; so by (ii) it is not optimal. ■