Merkle Root and Merkle Tree

The Merkle Root is one element of the Merkle Tree, a hash-based data structure – or hash tree – which is composed of leaves and branches as follows:

- Each leaf represents the hashed value of a transaction.
- Two leaves are then chained ("concatenated") and hashed to form a branch.
- Then two branches are "concatenated" and hashed to form another branch.

This process of re-hashing the branches is performed until the top of the tree – called the "root hash " – is reached.

By organizing the data following this structure, the Merkle trees take large amounts of data and make them more manageable to process. The use of the "Merkle Root" in the block header makes it possible to easily compare and verify large data sets of transactions without having to include the complete set of data of every transaction in the block header, while still providing a way to verify the entire blockchain on every transaction.

To compare two replicas of data sets (which can be of huge size), there is therefore no need to check all the data elements in both sets, but rather the difference between their two hash trees.

Furthermore, to compare two hash trees, one only needs to compare the root nodes of those trees. This results in much easier and more efficient data integrity and consistency verification.

If the two root nodes being compared are equal, then both the data and their recording order in the tree are valid. If not, then the trees contain inconsistent data records. In this case, tracing the source of inconsistency is also facilitated by the branched structure. To locate the origin of the difference between two trees, it suffices to go through them from top to bottom to find the nodes and then the leaves with different hashes.

The Merkle Tree structure significantly reduces the size needed to perform consistency and data verification, as well as data synchronization in peer-to-peer networks and distributed ledgers. File-sharing systems such as Google Drive and Dropbox are two applications that use Merkle Tree features: changes are detected by comparing the root, branch and leave nodes, and only data that need to be synchronized are transferred between the source and the destination.