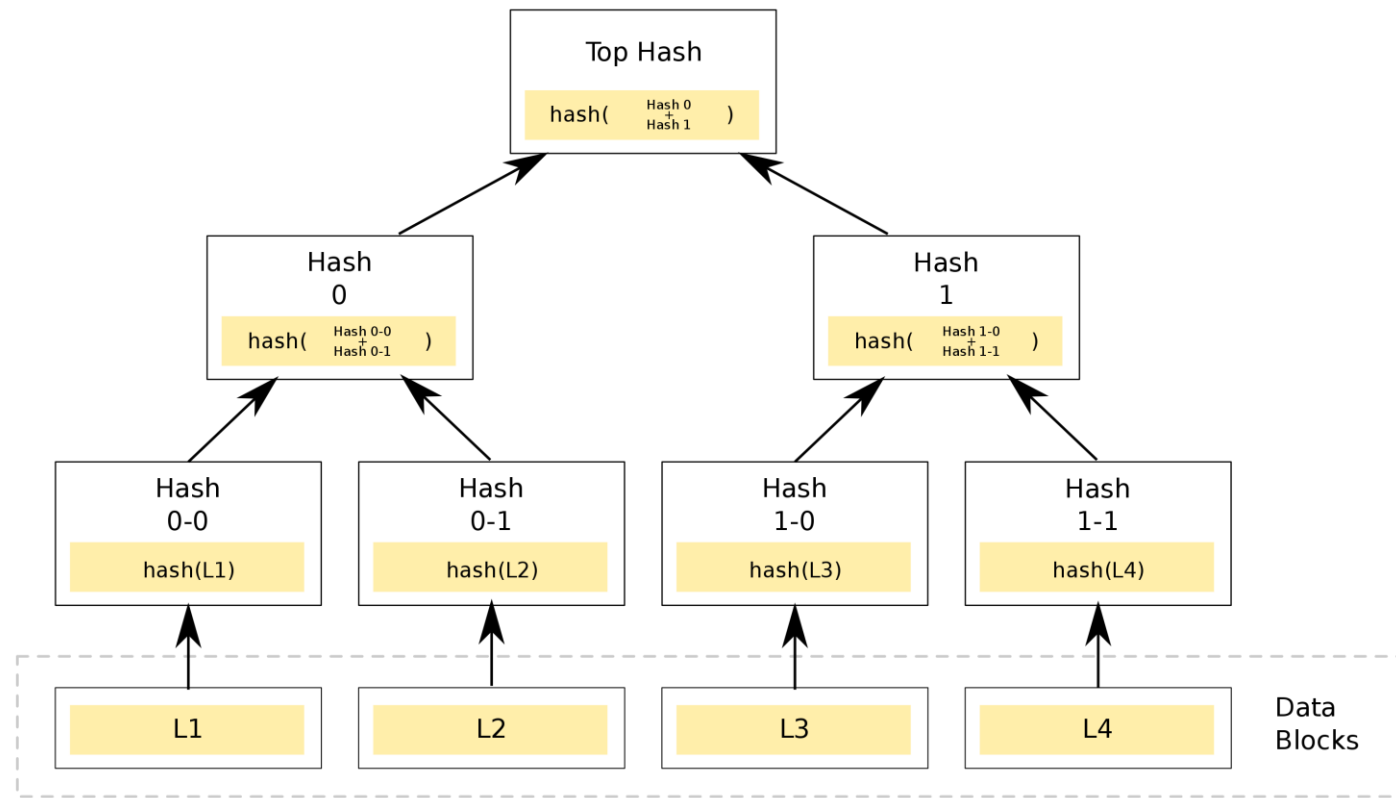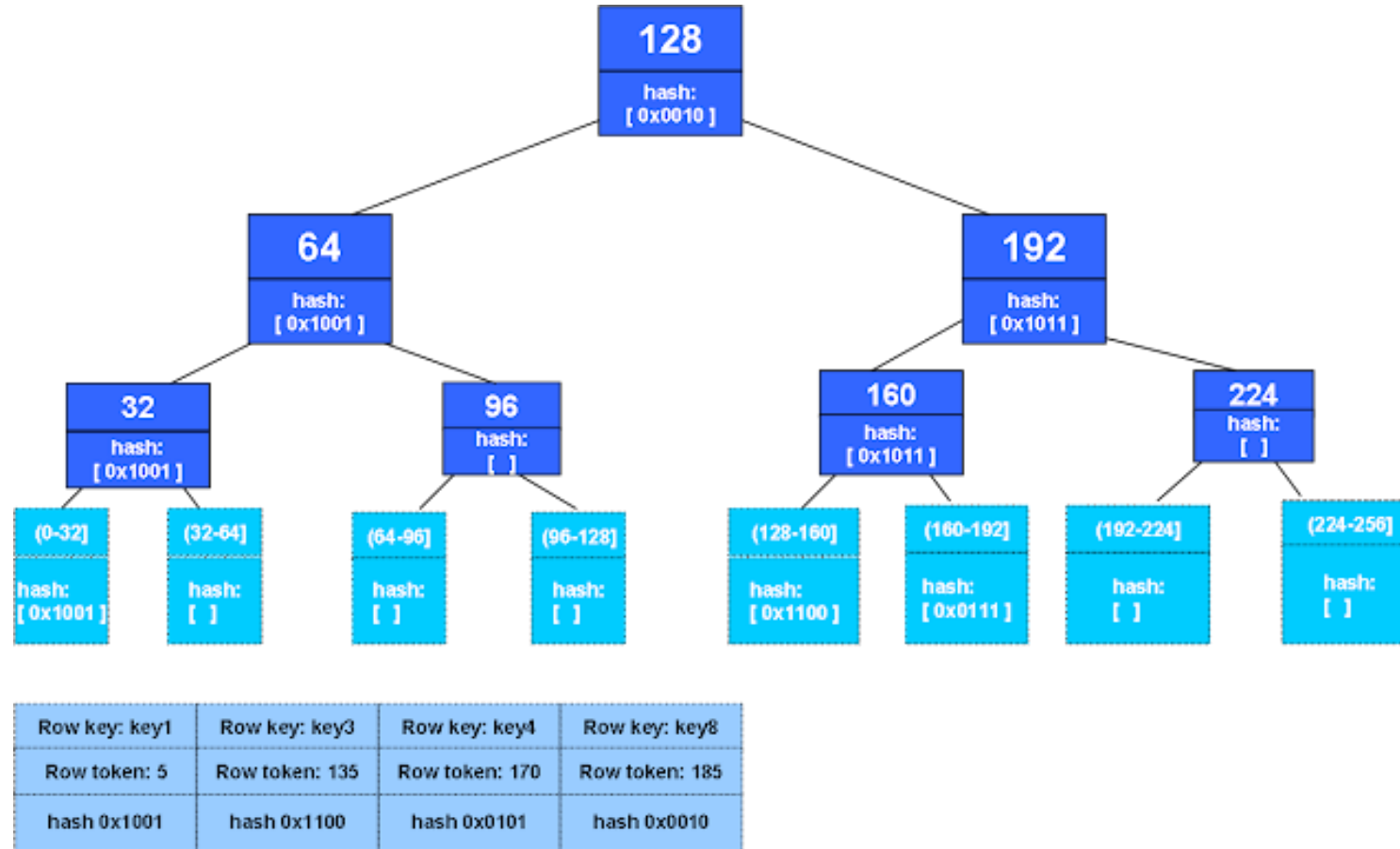# MerkleTrees

How would you synchronize 2 databases?

- Both have millions of rows
- Crappy ingestion resulted in many rows missing
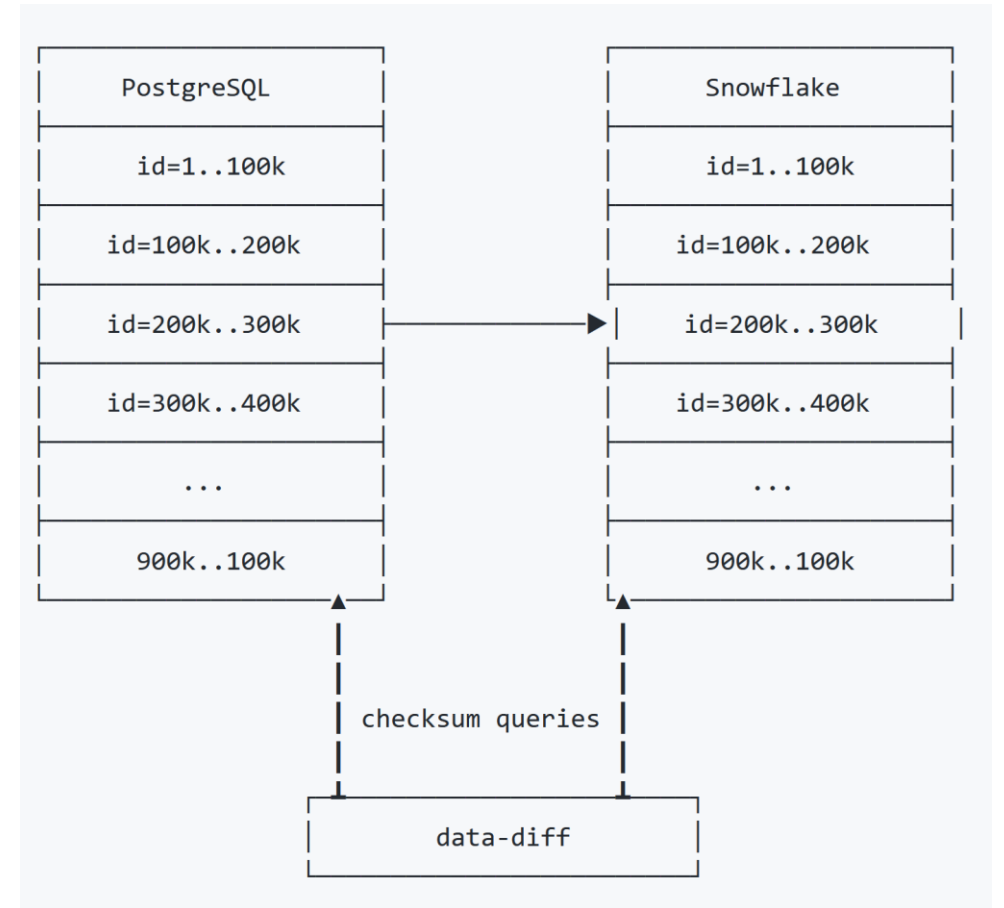- Crappy internet makes batch dump to costly

MerkleTree example

MerkleTree Cassandra

Data-Diff

Poor-mans merkletree

1. min/max on id column
2. divide into X buckets
3. hash each bucket
4. compare hashes
5. repeat

MerkleTree, hackaton

- https://github.com/NielsZeilemaker/merkle_tree_tutorial