

## B-Trees

## B-Trees

Generalizes multilevel index.

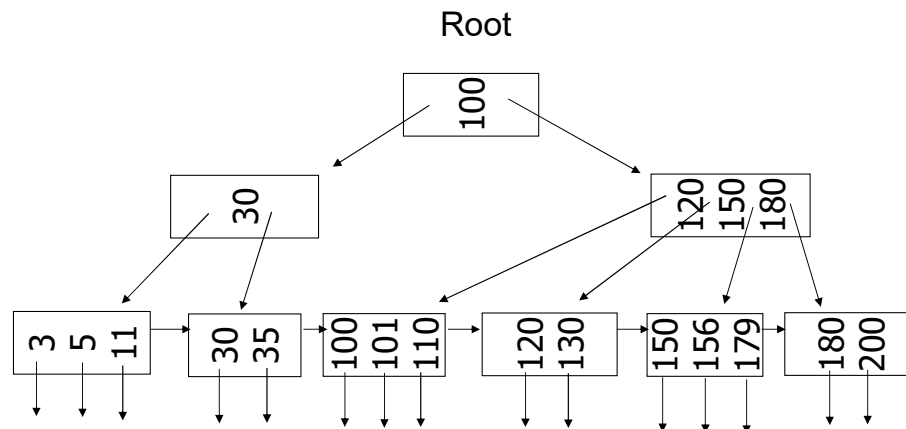
Number of levels varies with size of data file, but is often 3.

Different variants, we start with B+-trees.

Useful for primary, secondary indexes, primary keys, nonkeys.

Each node in the tree represents a block.

## B+Tree Example



## Nodes of B+ Tree

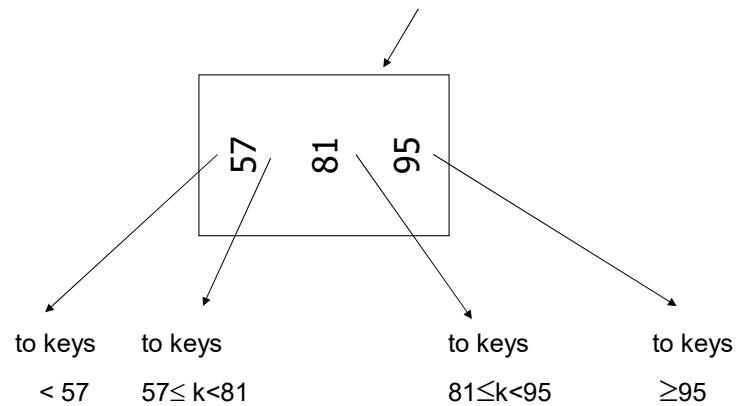
### Leaves

- One pointer to next leaf.
- keypointer pairs for records of data file.
- At least half of these (round up) occupied.

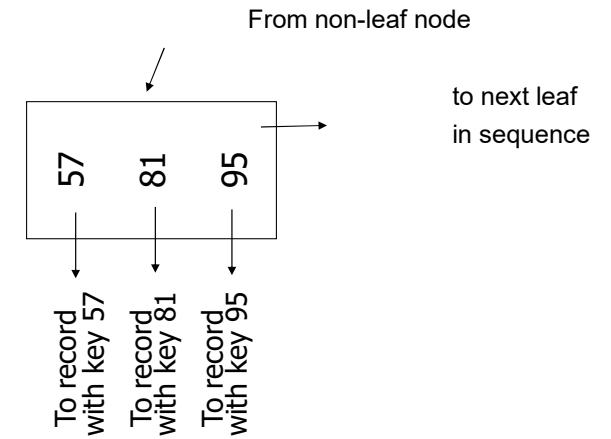
### Interior Nodes

- k keys form the divisions among k+1 subtrees.
- Key i is least key reachable from (i + 1)st child.

## Sample non-leaf



## Sample Leaf Node



## Don't want nodes to be too empty

Trees have an order that determines the maximal number of keys in a node

Use in a tree of order  $n$  at least

Non-leaf:  $\lceil (n+1)/2 \rceil$  pointers to children

Leaf:  $\lfloor (n+1)/2 \rfloor$  pointers to records

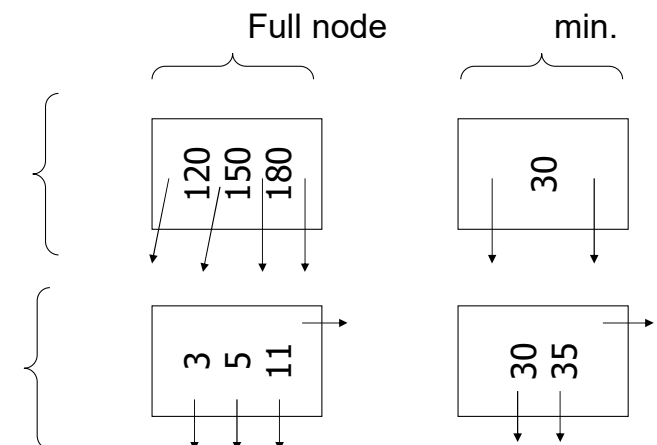
Root is a special Case

$n=3$

node

Non-leaf

Leaf



## B+ Tree rules (Tree of order n)

- (1) All leaves at same lowest level  
(balanced tree)
- (2) Pointers in leaves point to records  
except for "sequence pointer"
- (3) Number of pointers/keys for B+ tree (except for  
sequence pointers)

	Max ptrs	Max keys	Min ptrs→data	Min keys
Non-leaf (non-root)	n+1	n	$\lceil (n+1)/2 \rceil$	$\lceil (n+1)/2 \rceil - 1$
Leaf (non-root)	n	n	$\lfloor (n+1)/2 \rfloor$	$\lfloor (n+1)/2 \rfloor$
Root	n+1	n	1 (if leaf)	1

Prof. Dr. Justus Klingemann

## Lookup

### Lookup in B+ Tree

- Start at root.
- Until you reach a leaf, follow the pointer that could lead to the key you want.
- Search that leaf (and leaves to the right if duplicates are possible).

Prof. Dr. Justus Klingemann

## B+ Tree Insertion

Search for the key being inserted.

If there is room for another key-pointer pair at that leaf, insert there.

If no room, split leaf.

- Split of leaf results in insert of key-pointer pair at level above.
  - key is **copied** to level above
- Thus, recursive splitting all the way up the tree is possible.
  - split of non-leaf results in **moving** one key to level above
- Convention: If the number of keys in the two nodes resulting from the split is uneven, put one more key in the left node.  
Otherwise: both nodes get the same number of keys

Prof. Dr. Justus Klingemann