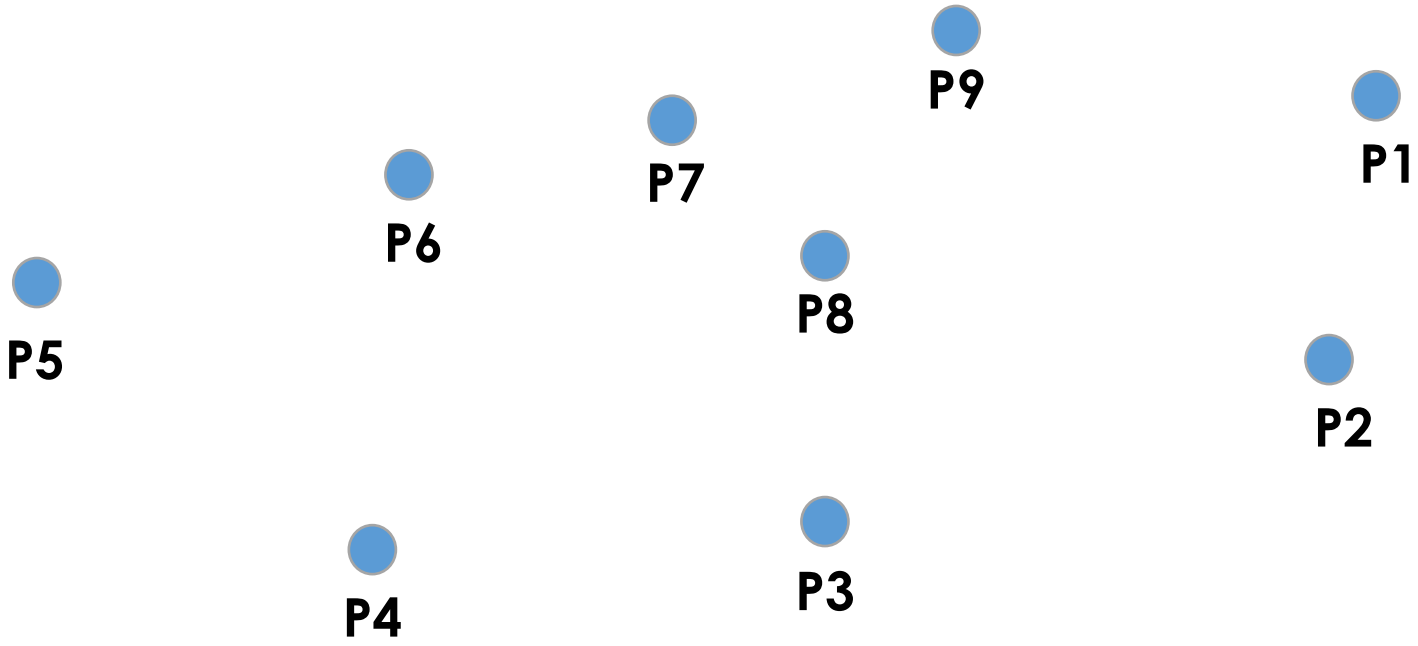
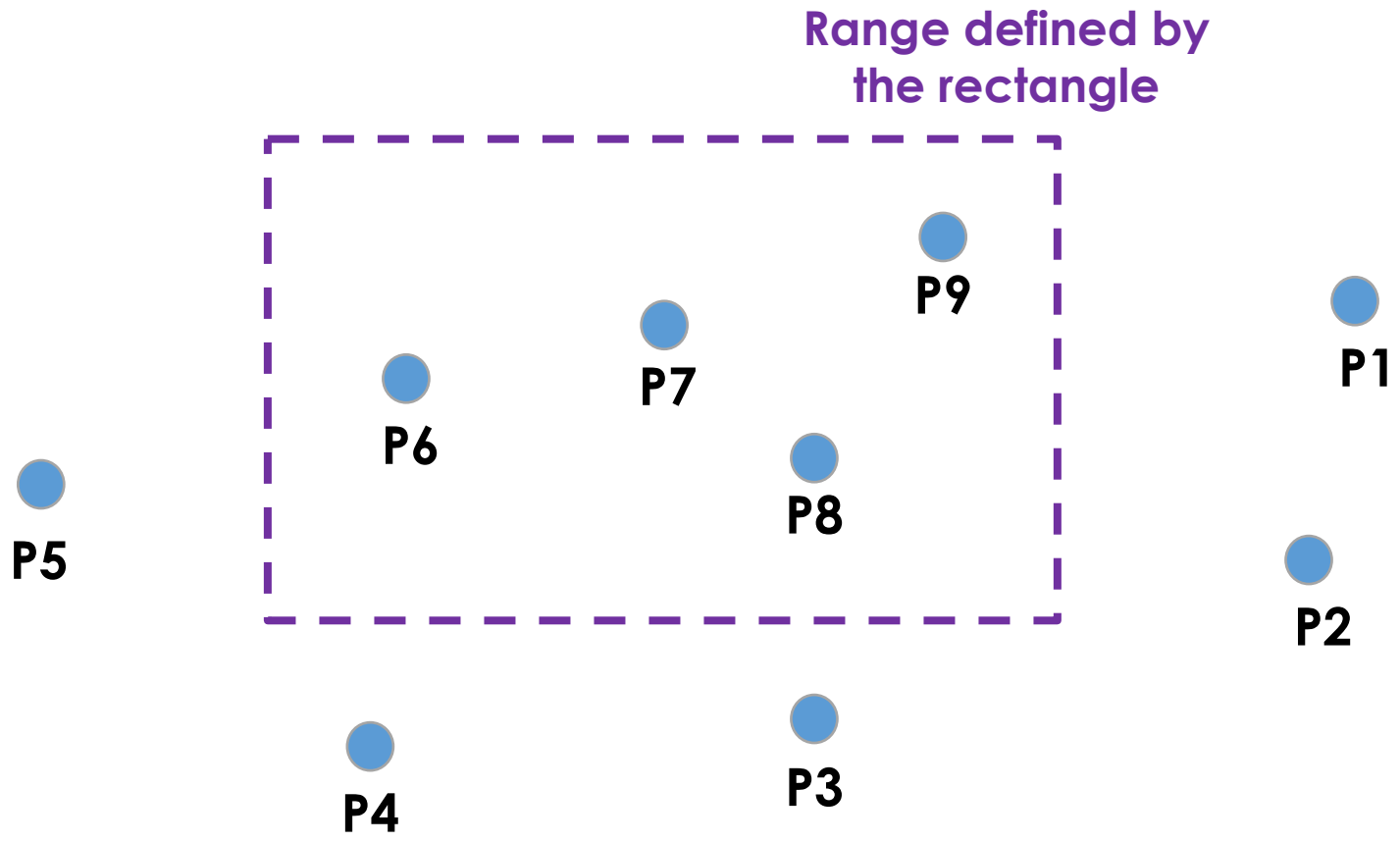
Spatial Data structures
For Point Data

J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Points in 2D space



Range Queries



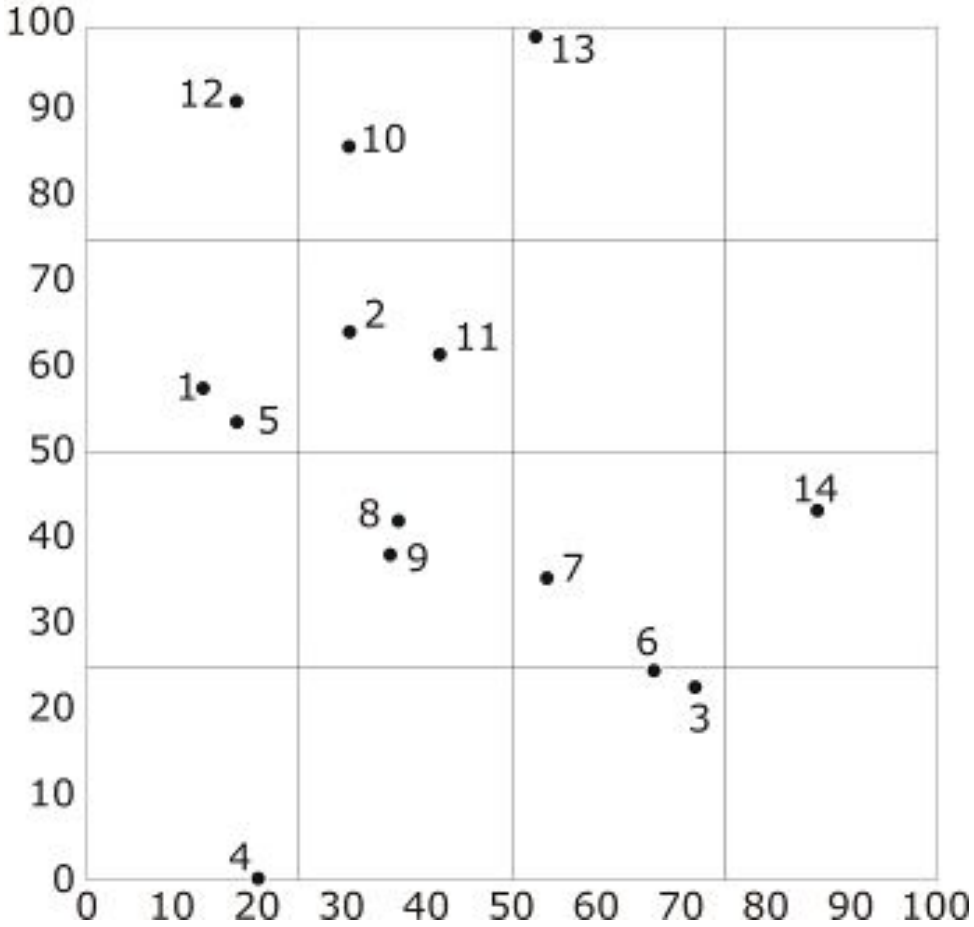
Retrieve all points inside the rectangle
E.g., return all restaurants in GK-1, New Delhi
Range can also be defined by a rectangle

How to store 2D points in a secondary memory ?

Data structure should support efficient query algorithms for range queries

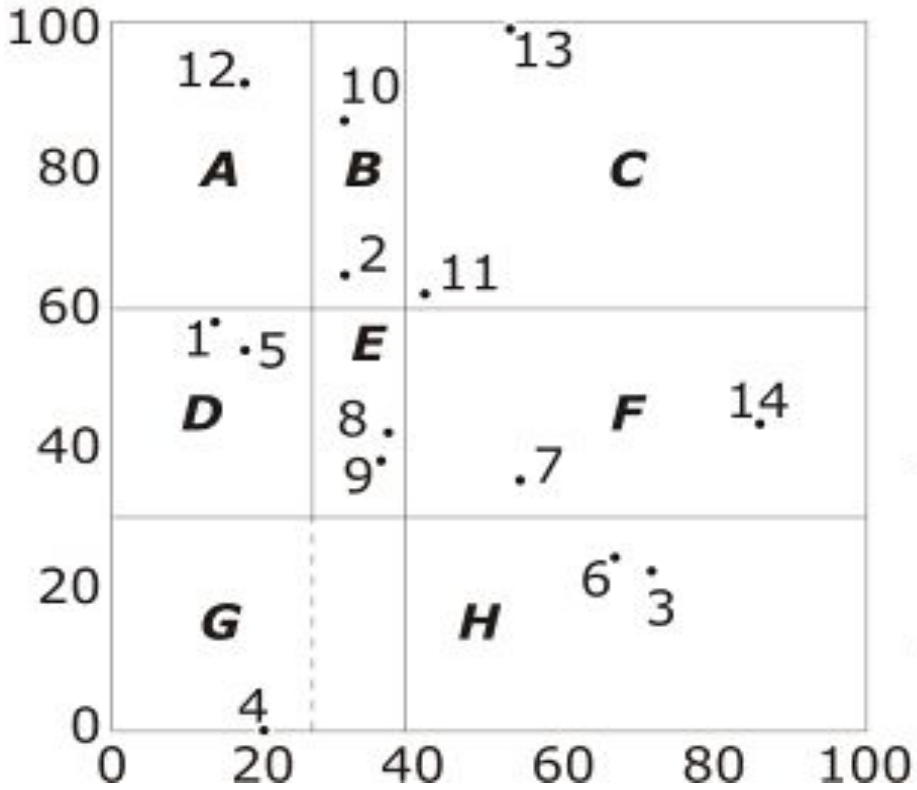
Uniform Grids over 2D space

- Divide space into cells by a grid
- Store data in each cell in distinct disk page
- A directory structure needed
- Efficient for find, insert, range and nearest neighbor
- **But wastes a lot of disk storage space**
- **Non-uniform data distribution over space ??**



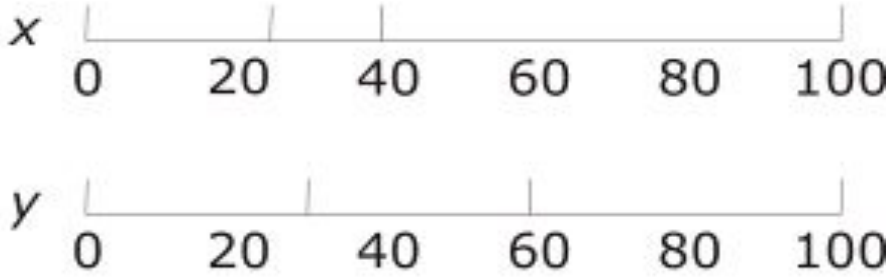
Refinement of basic idea into Grid Files

- Use non-uniform grids
- Linear scale store row and column boundaries
- Allow sharing of disk pages across grid cells



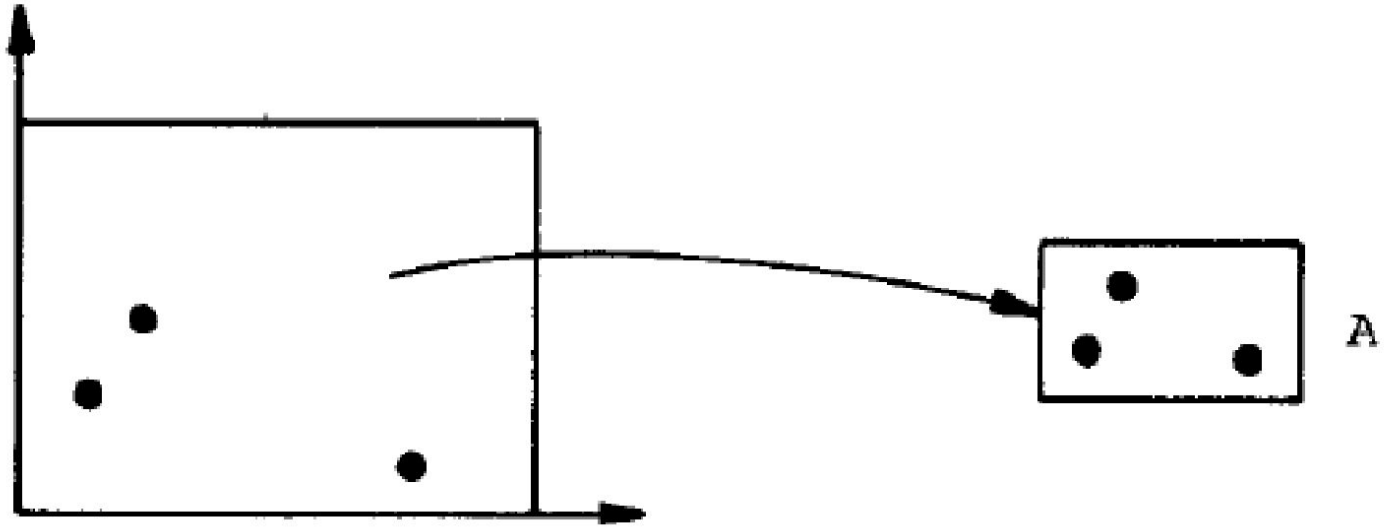
A	B	C
D	E	F
G	G	H

Grid directory



Grid Files (insertion example)

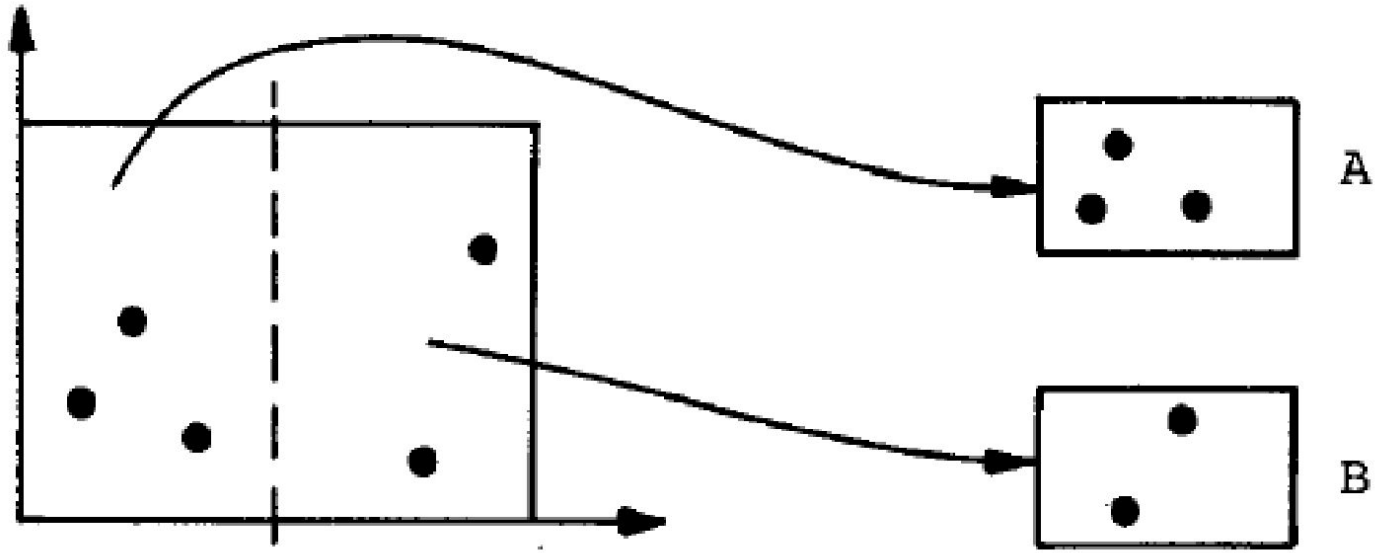
- Capacity of bucket = 3



J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (insertion example)

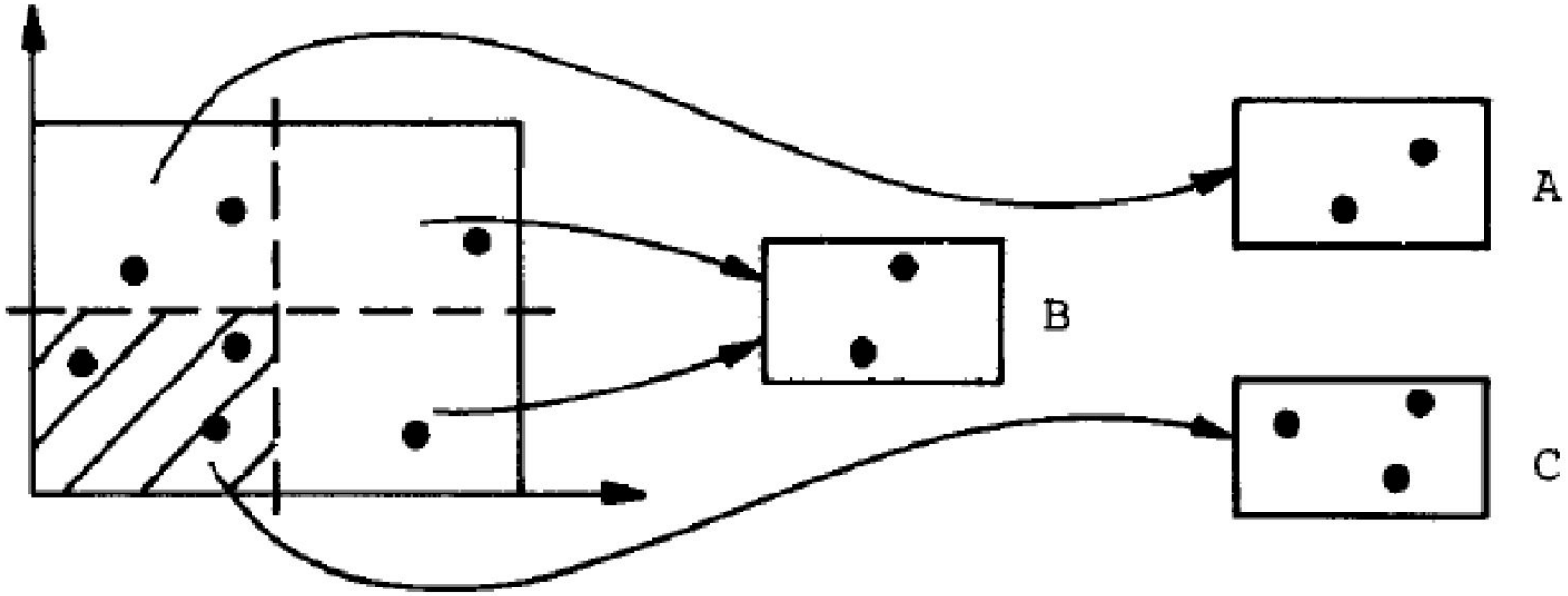
- When the bucket overflows we split it.
- A new bucket is made.
- Records that lie in one half of the space are moved to the new bucket.



J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (insertion example)

- Bucket A overflows again.

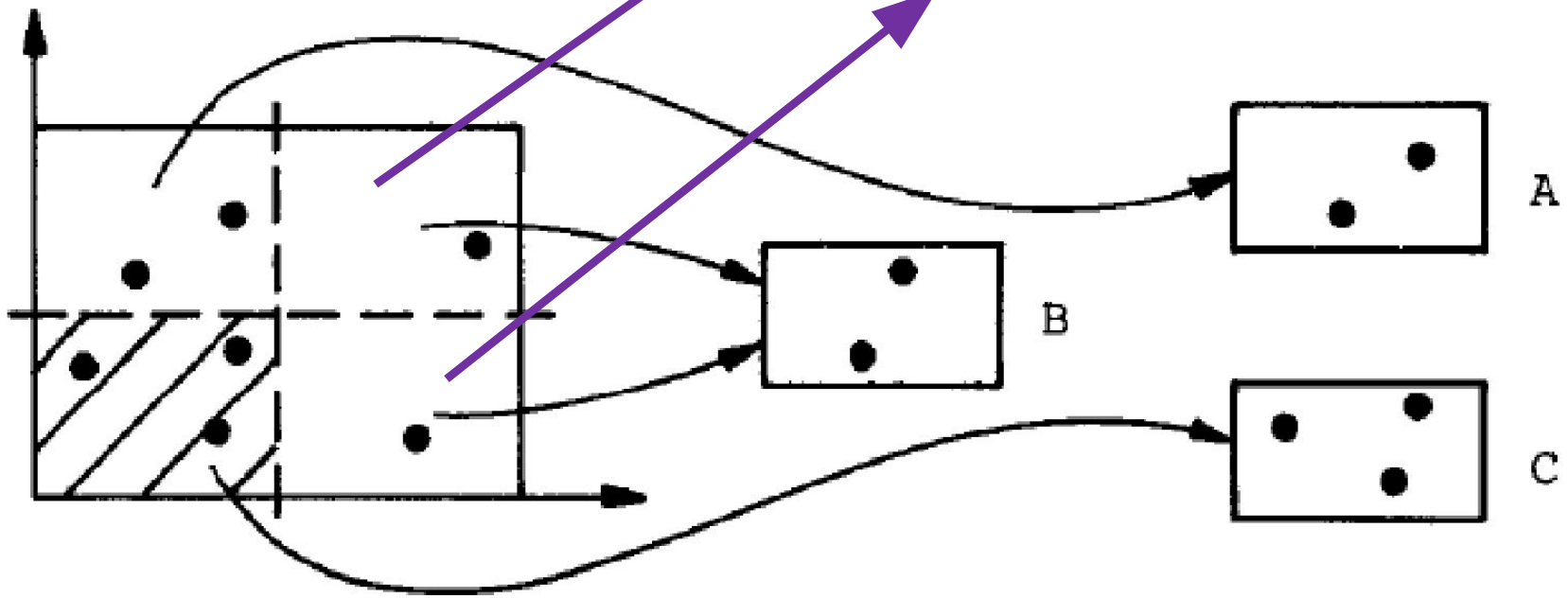


J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (insertion example)

- Bucket A overflows again.

Very Imp: Splitting of A is full horizontal split, i.e., region of B is also split. But B was not overflowing, so both buckets still point to B only

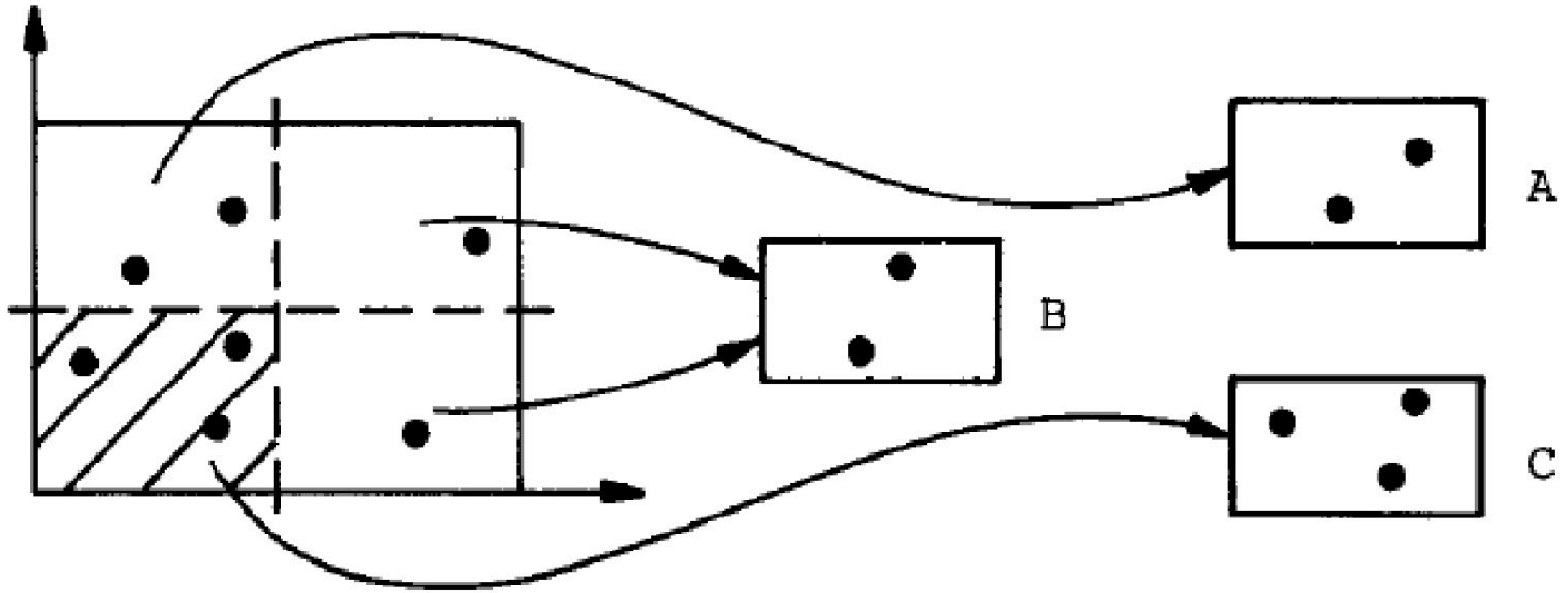


J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (insertion example)

- Bucket A overflows again.

In Grid files, data space which are the buckets is different from the geographic spread of the data.

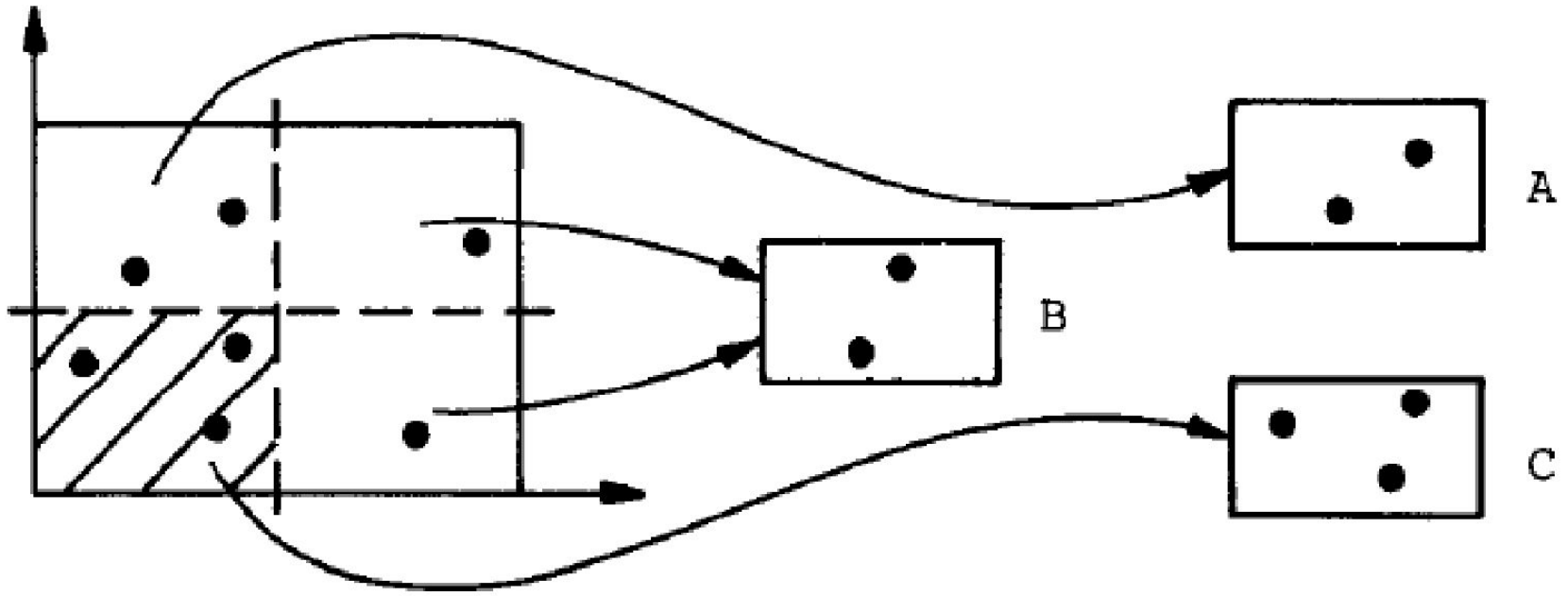


J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (insertion example)

- Bucket A overflows again.

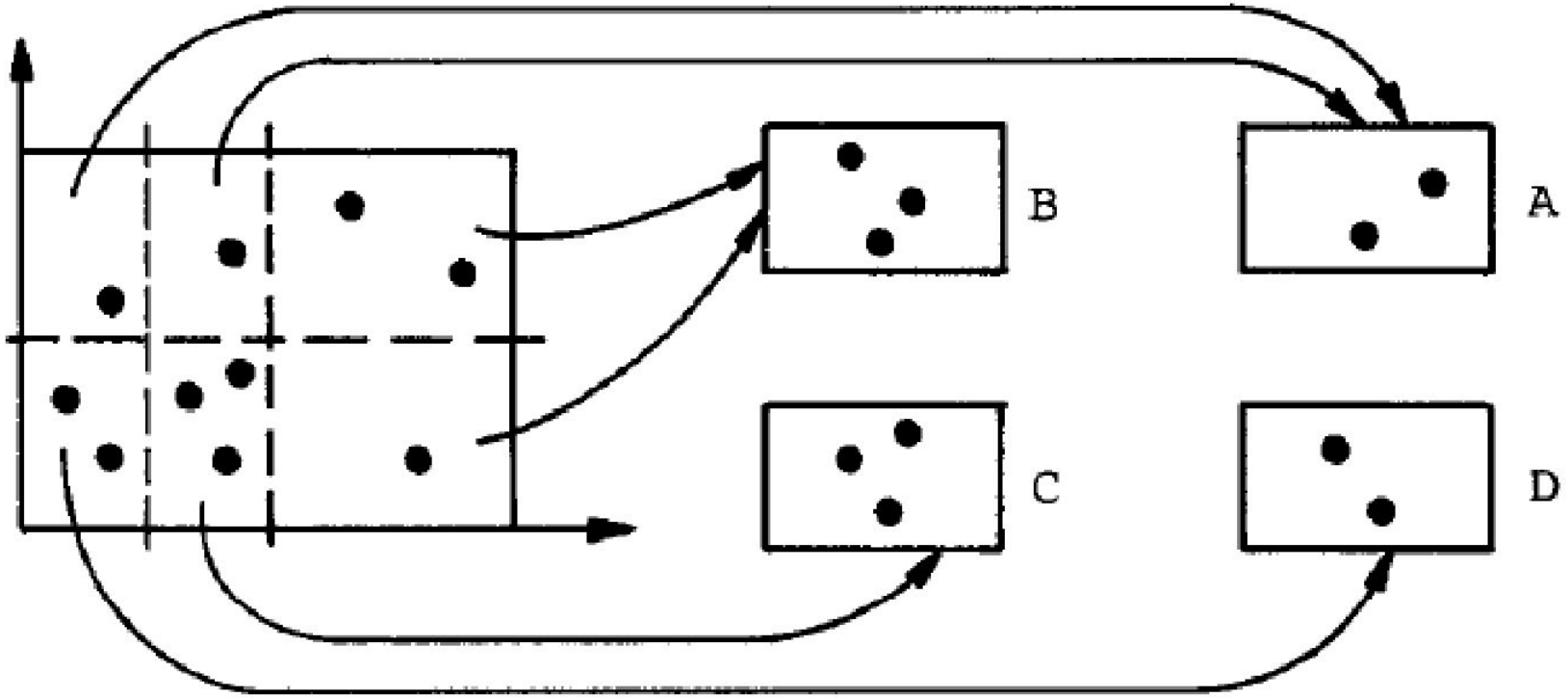
Splits in any dimension are made through and trough out. This makes the task of maintain linear scales easy



J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (insertion example)

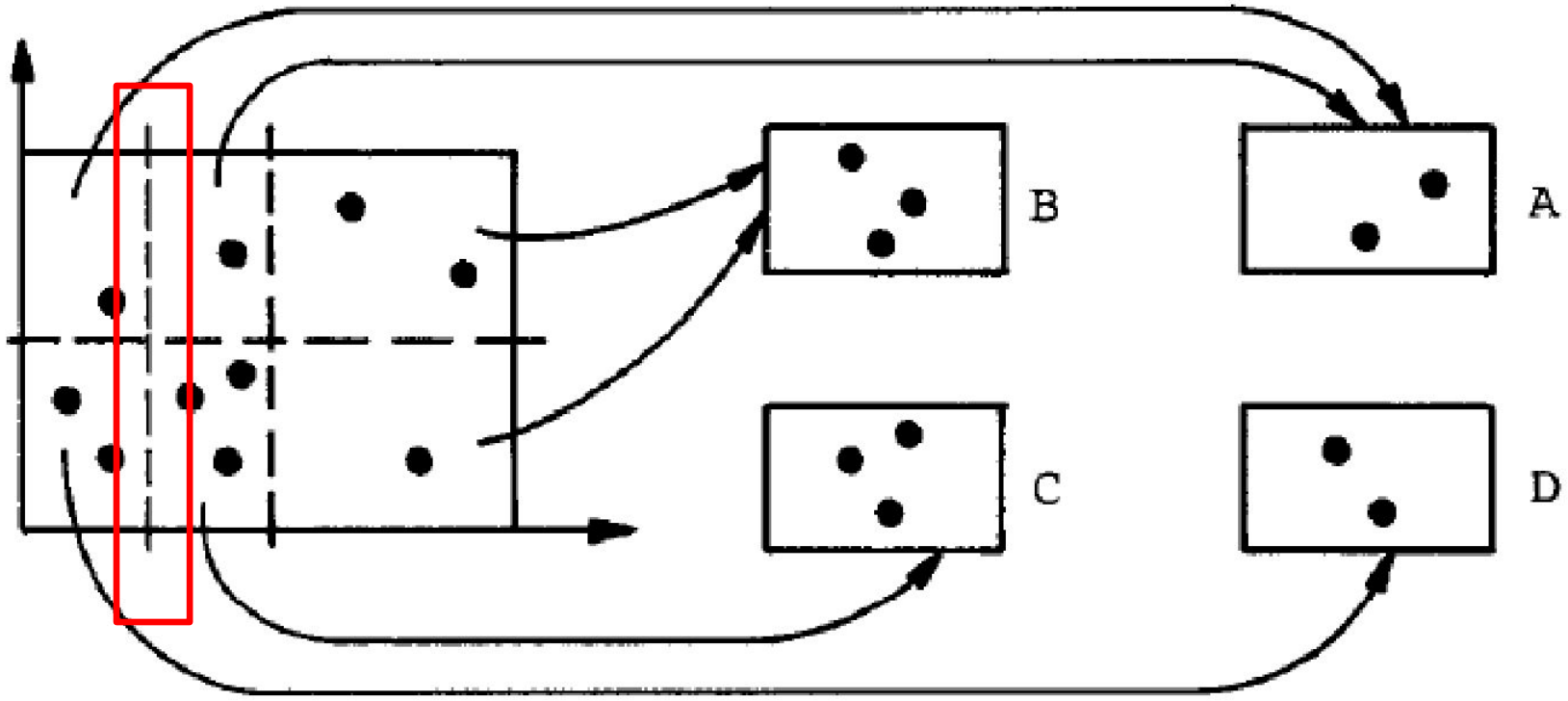
- One more split.



J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (insertion example)

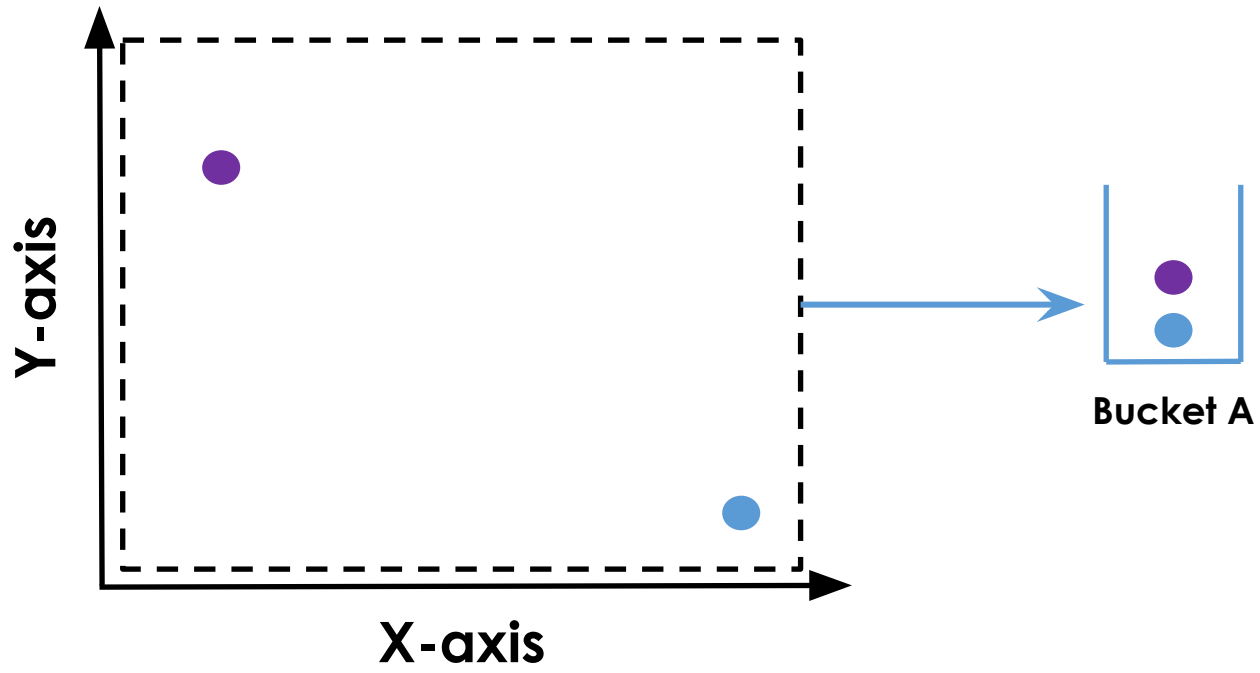
- One more split.
- Note that splits in any dimension are made through and trough.



J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

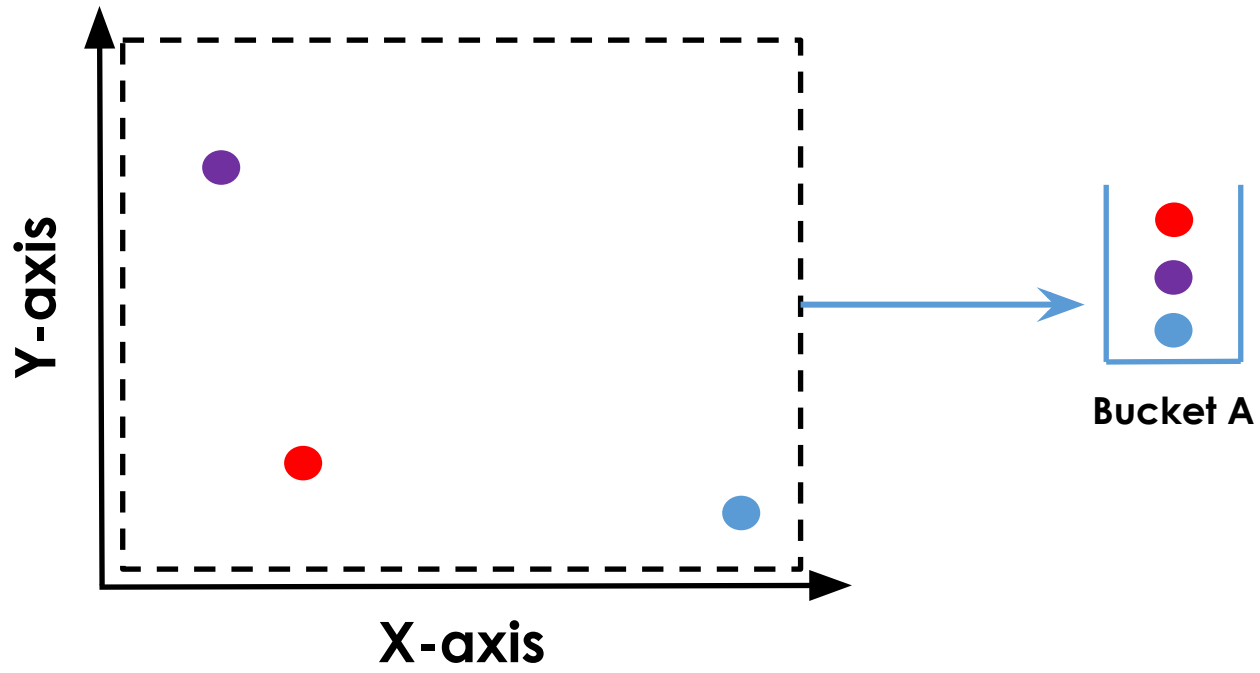
Grid Files Inserting Points (Another example)

- Assume Bucket size = 3



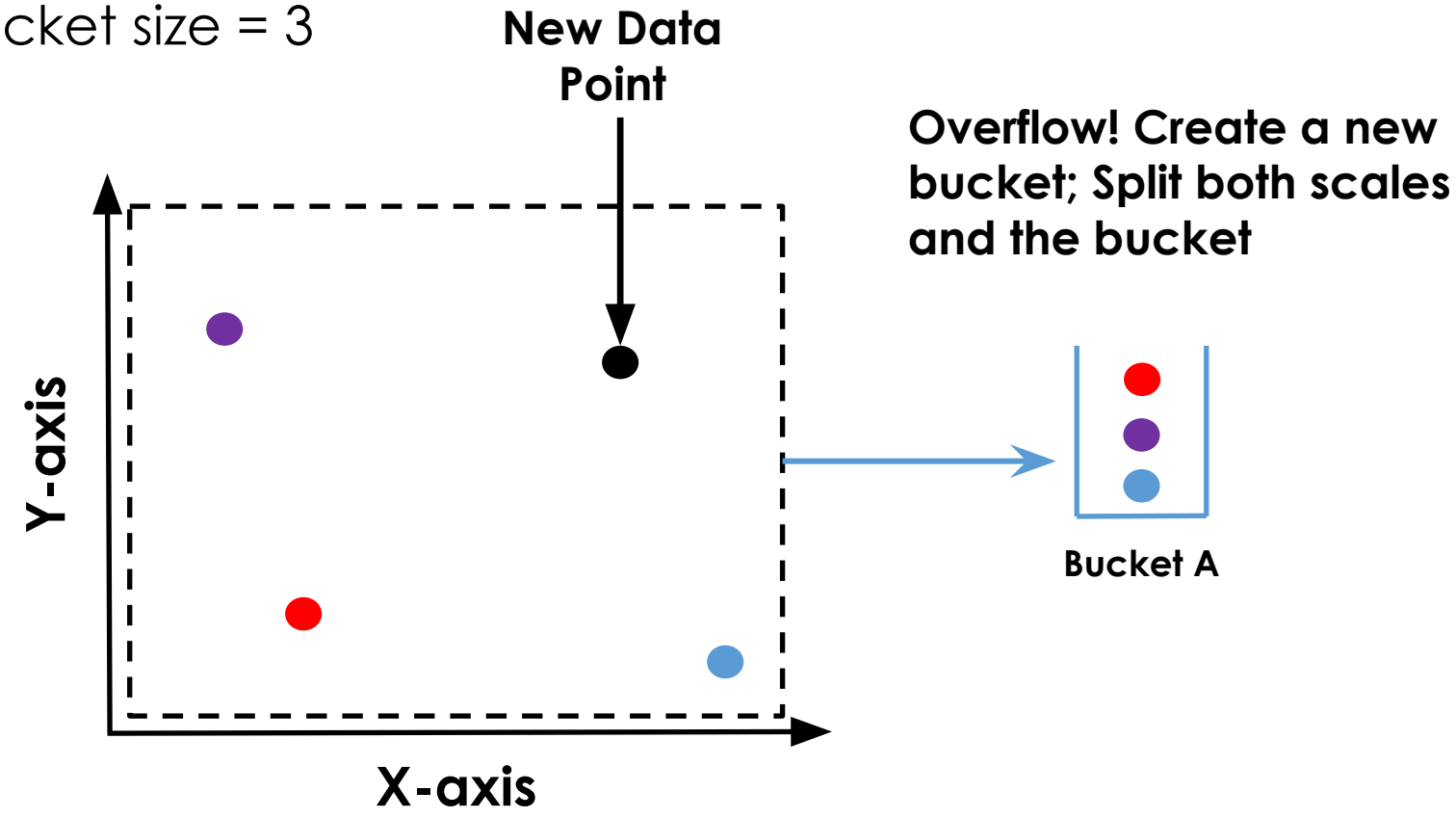
Grid Files Inserting Points (Another example)

- Assume Bucket size = 3



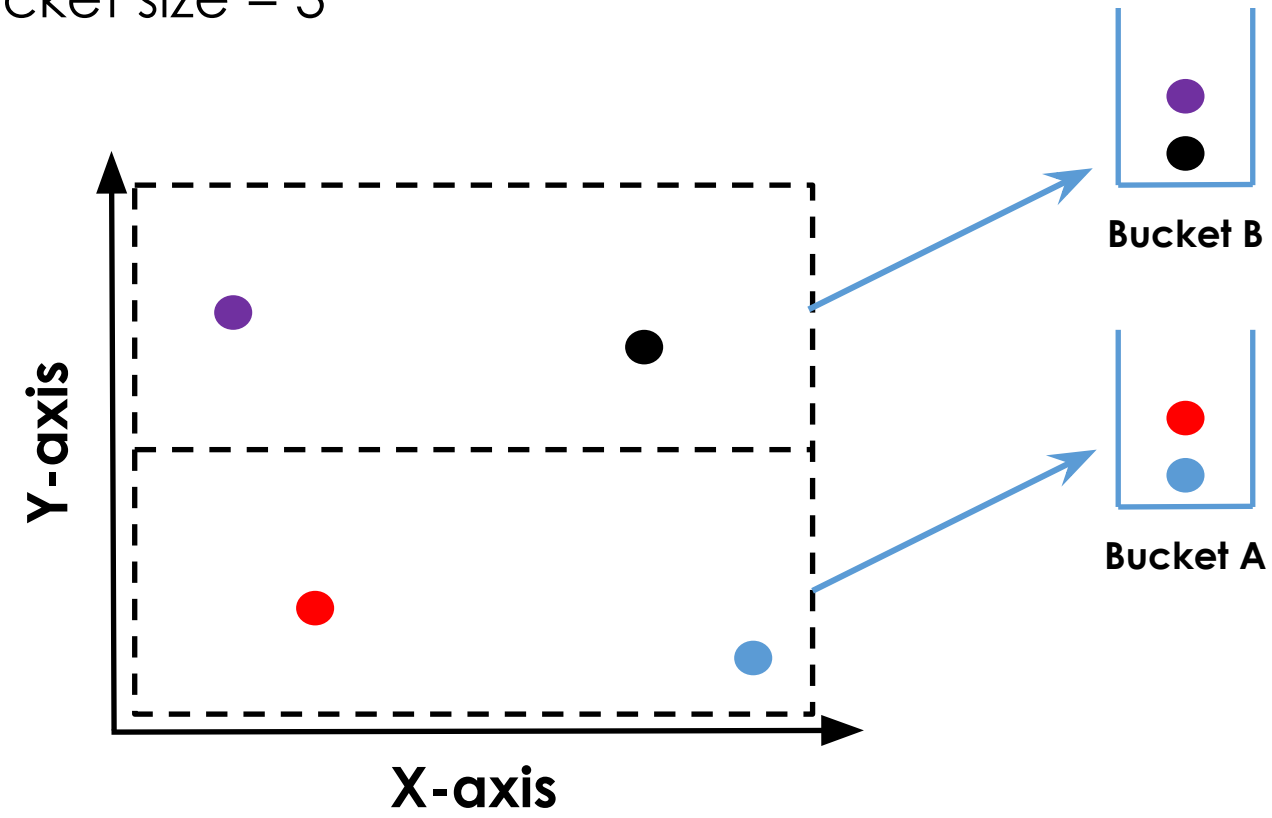
Grid Files Inserting Points (Another example)

- Assume Bucket size = 3



Resulting Grid File

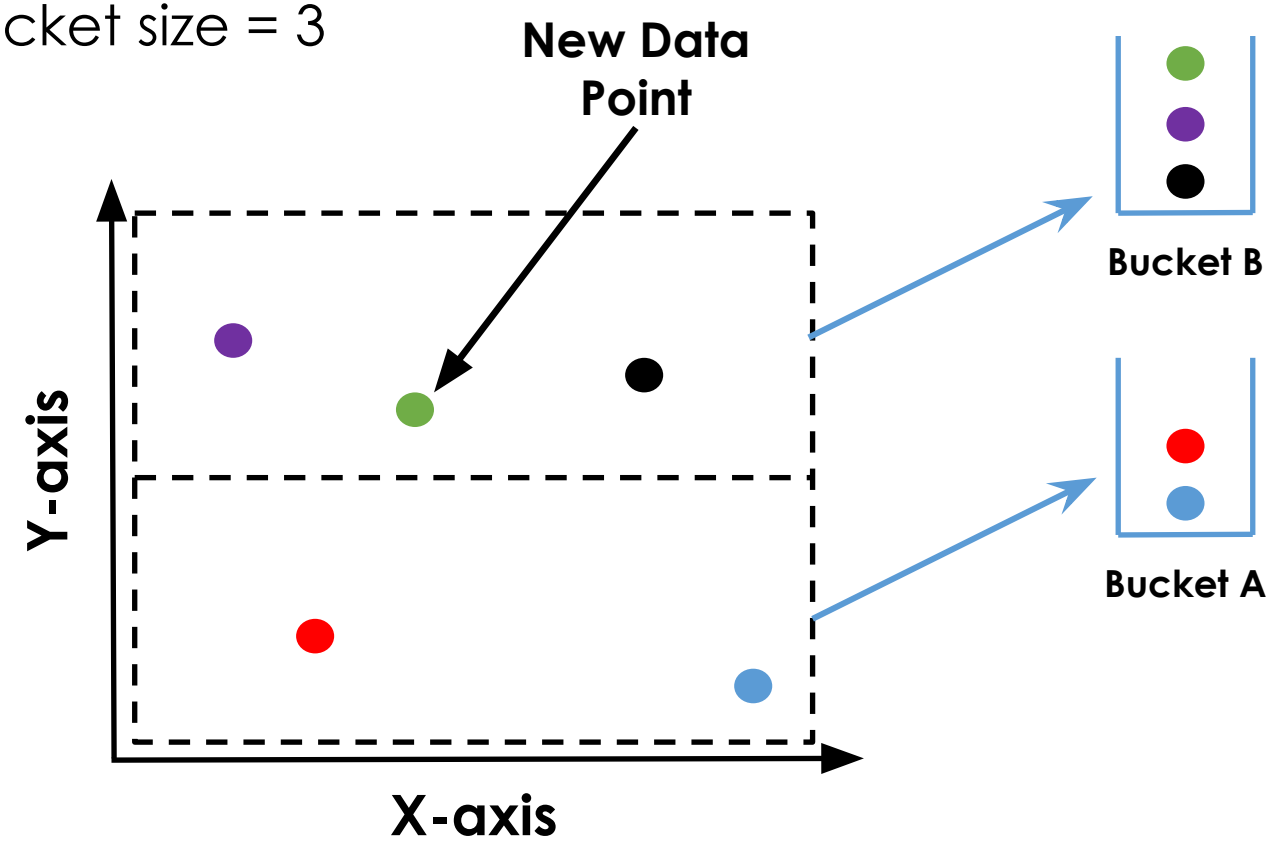
- Assume Bucket size = 3



Both Directory (x & y scales) and the Bucket are split

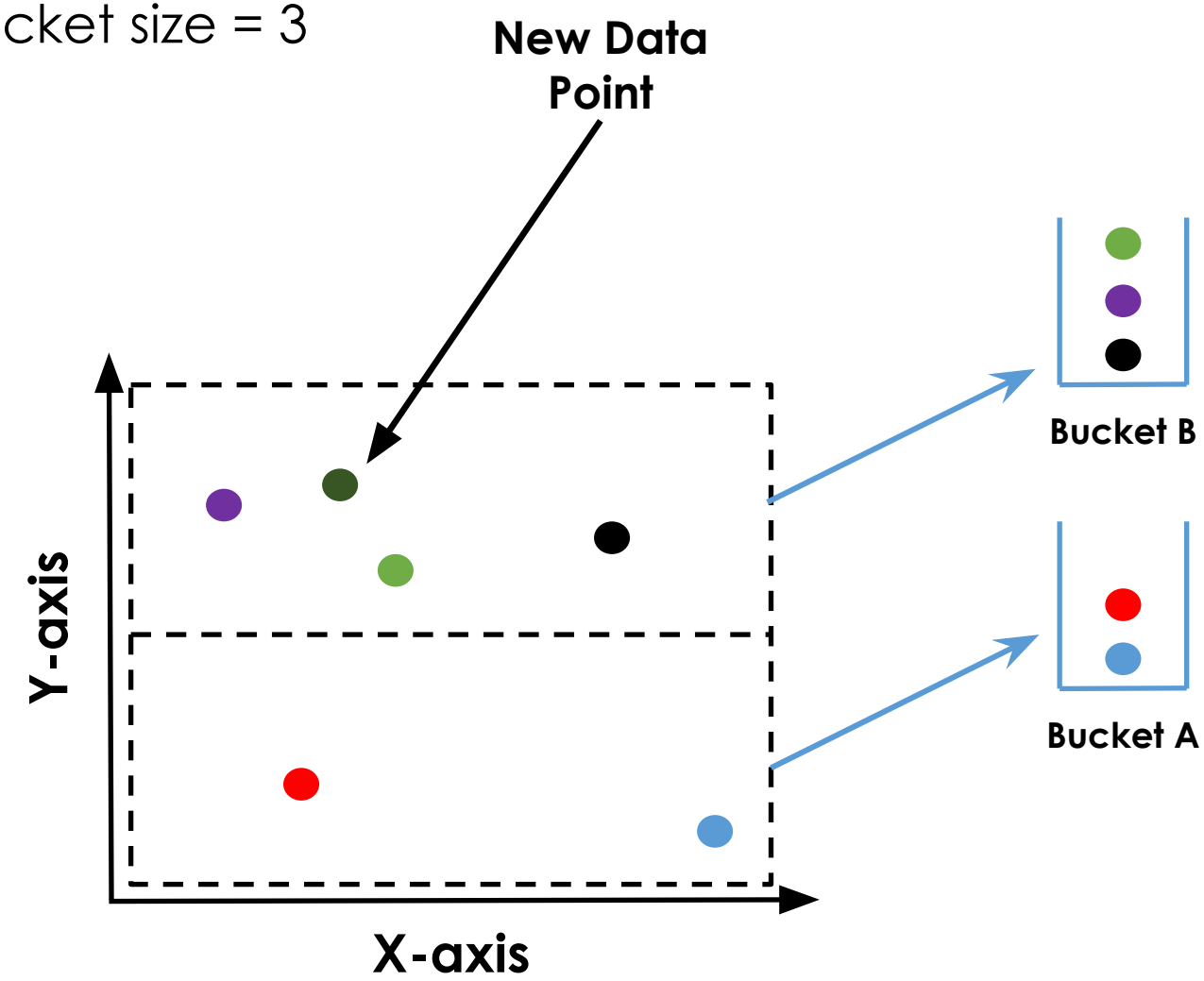
Grid Files Inserting Points (Another example)

- Assume Bucket size = 3



Grid Files Inserting Points (Another example)

- Assume Bucket size = 3

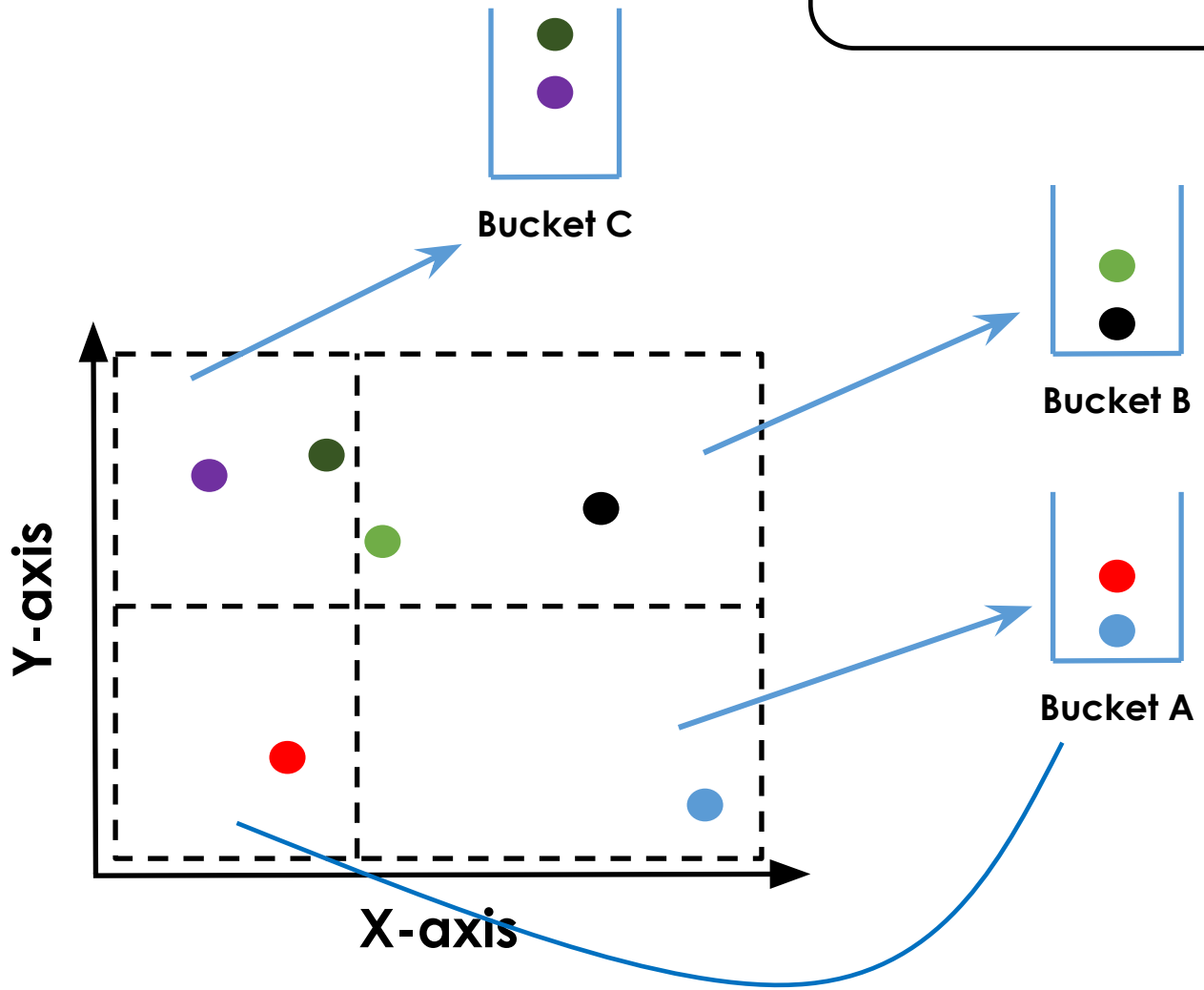


Overflow! Create a new bucket; Split both scales and the bucket.

Resulting Grid File

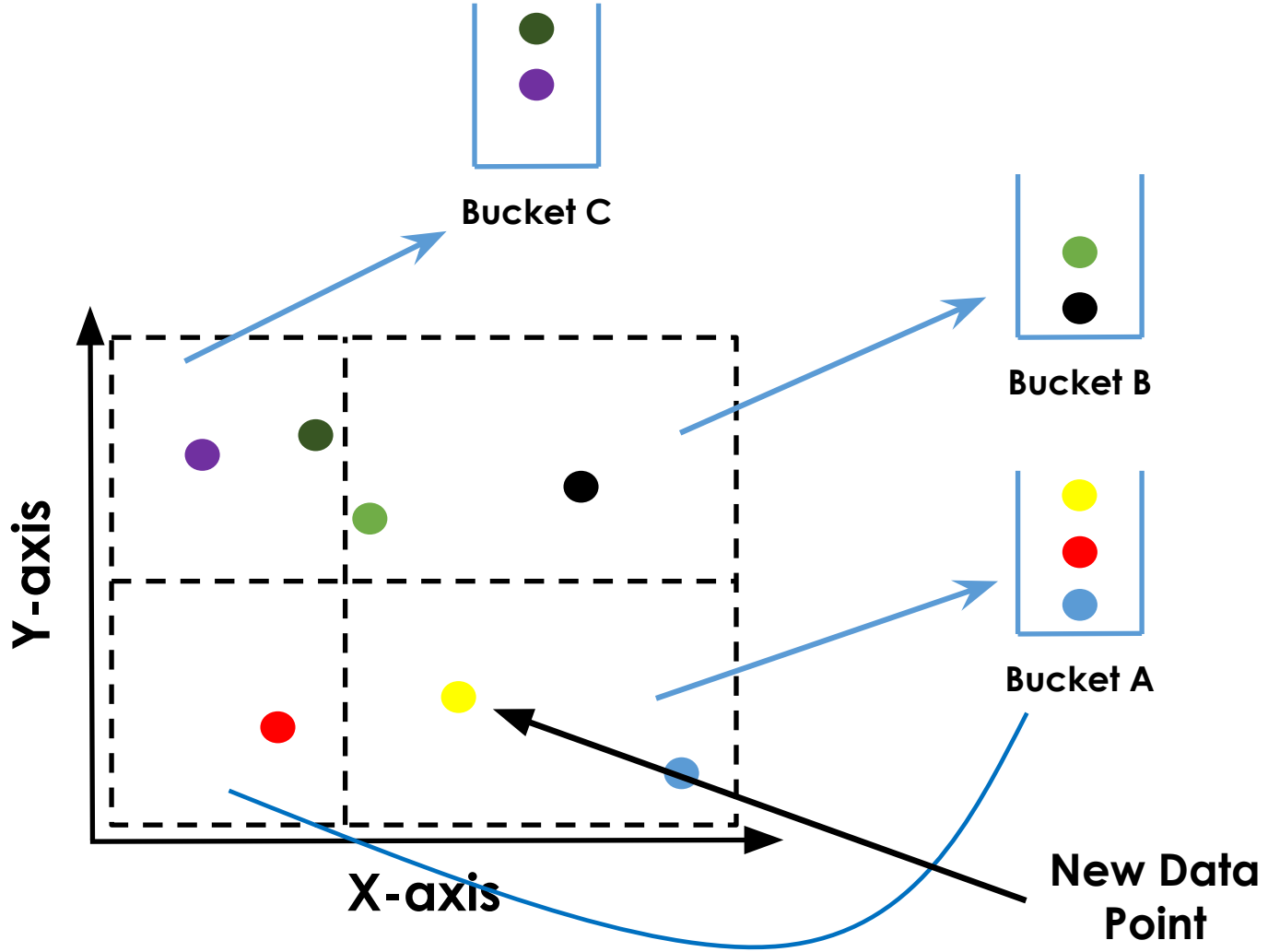
- Assume Bucket size = 3

Both Directory (x & y scales) and the Bucket are split



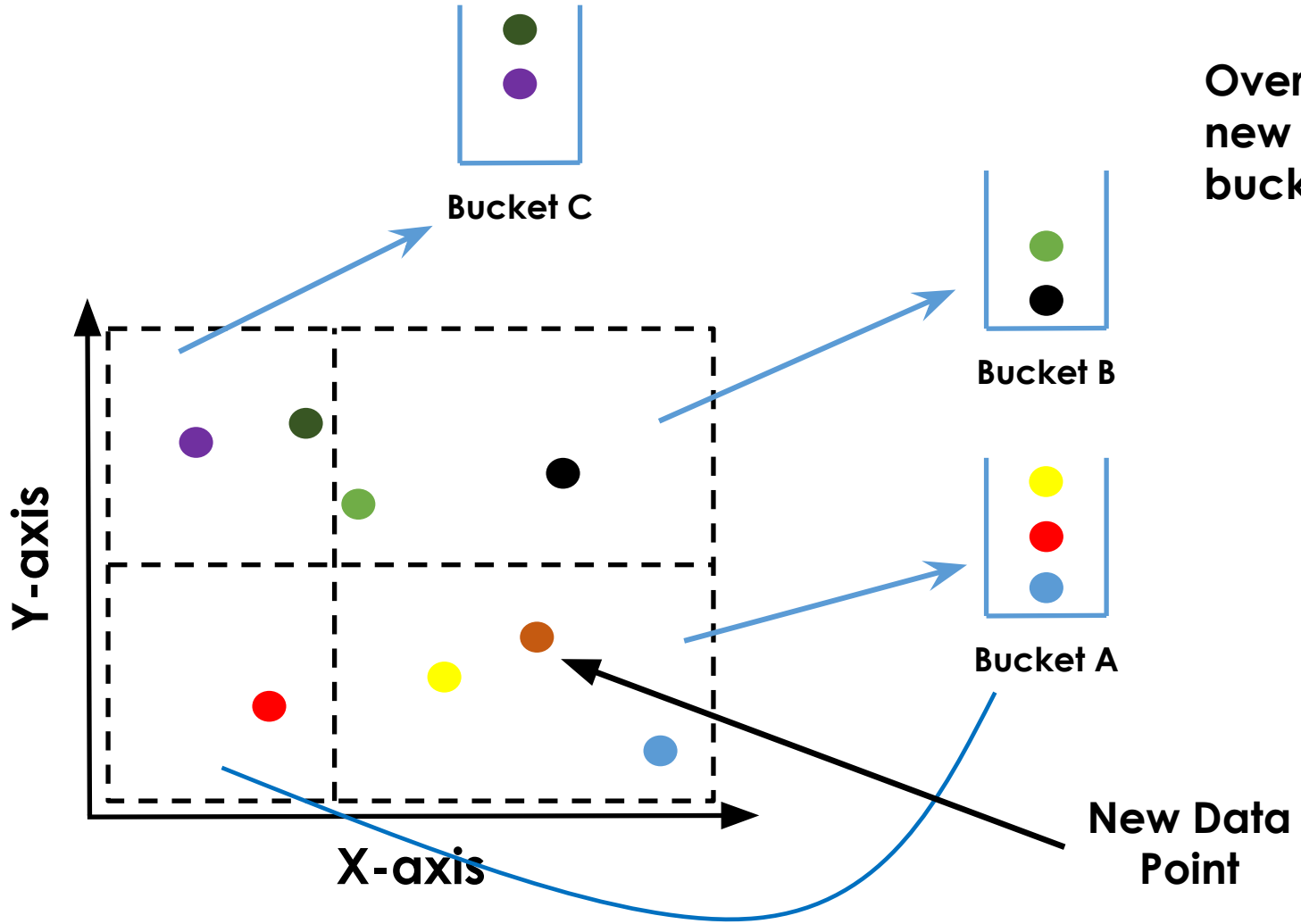
Grid Files Inserting Points (Another example)

- Assume Bucket size = 3



Grid Files Inserting Points (Another example)

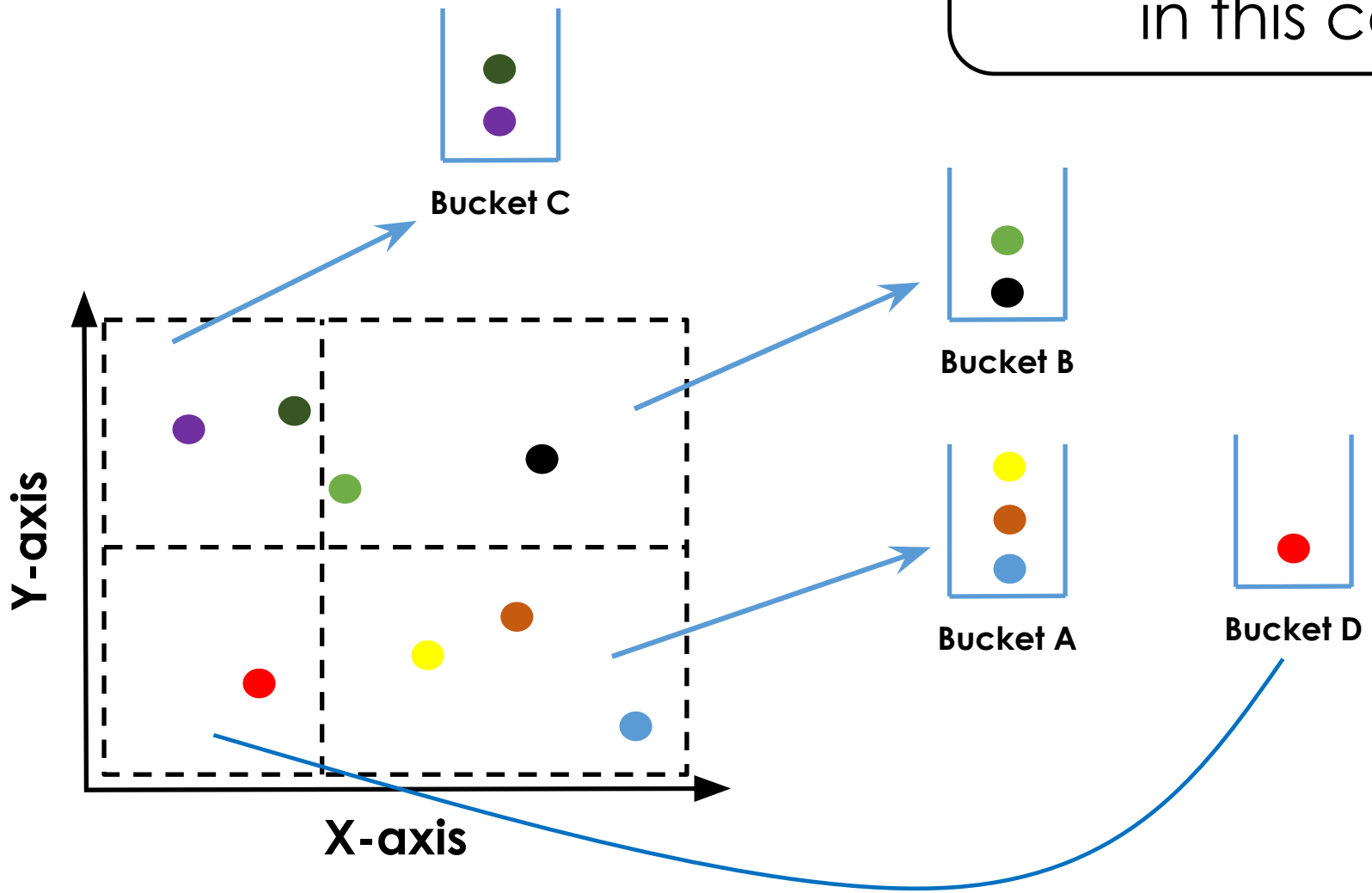
- Assume Bucket size = 3



Resulting Grid File

- Assume Bucket size = 3

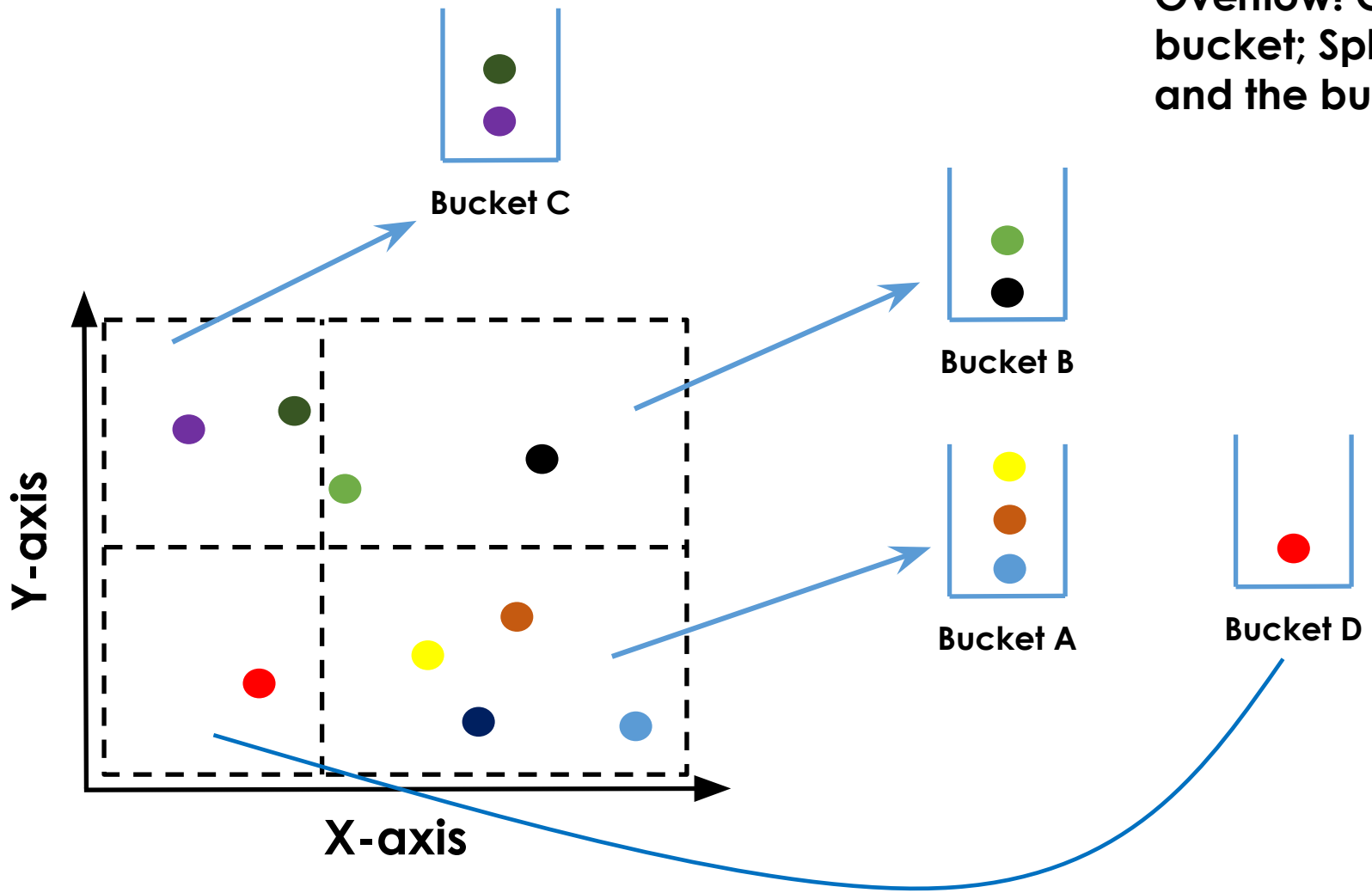
Only Bucket was split. Directory was not split in this case.



Grid Files Inserting Points (Another example)

- Assume Bucket size = 3

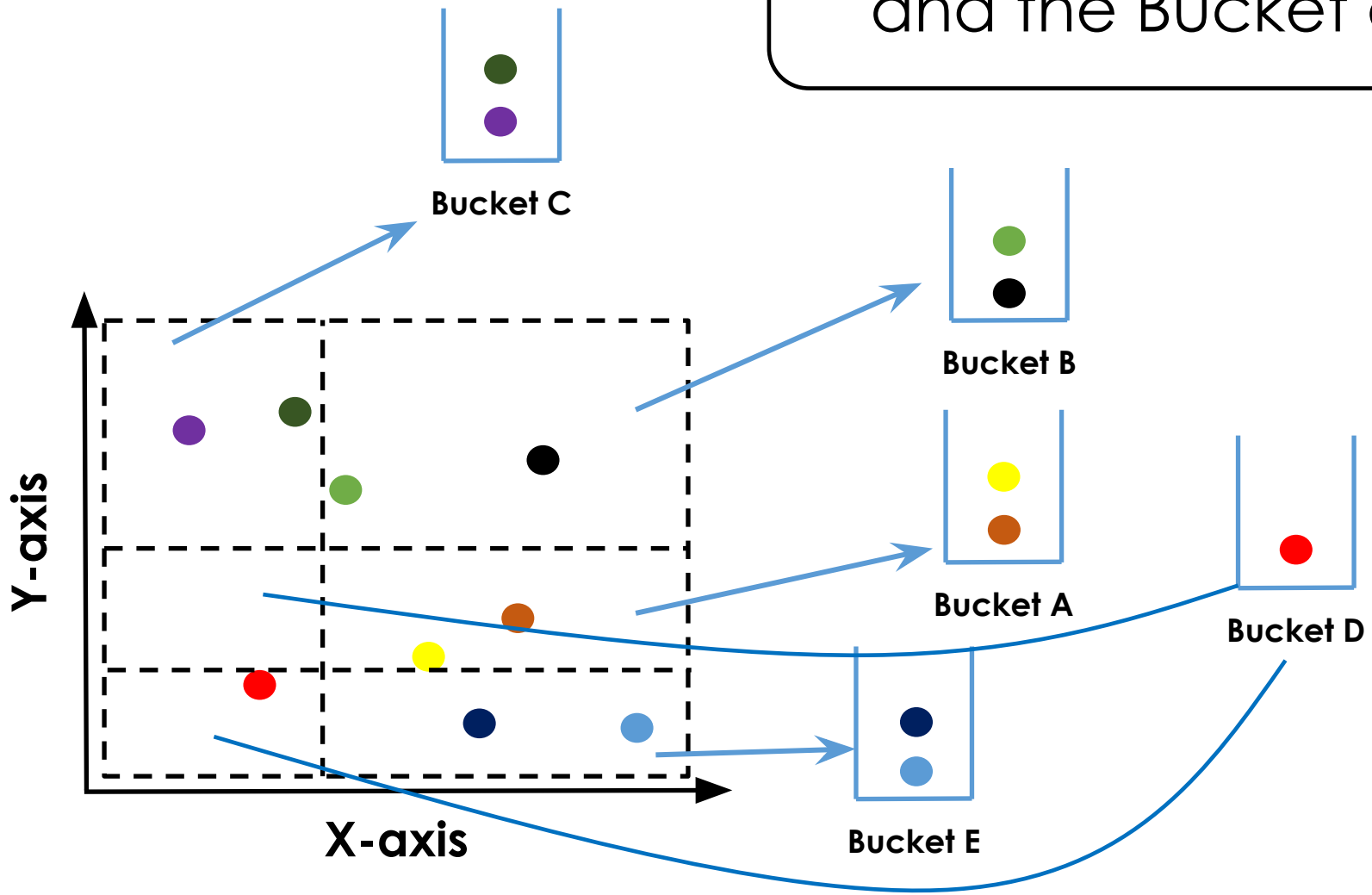
Overflow! Create a new bucket; Split both scales and the bucket.



Grid Files (Another example)

- Assume Bucket size = 3

Both Directory (x & y scales) and the Bucket are split



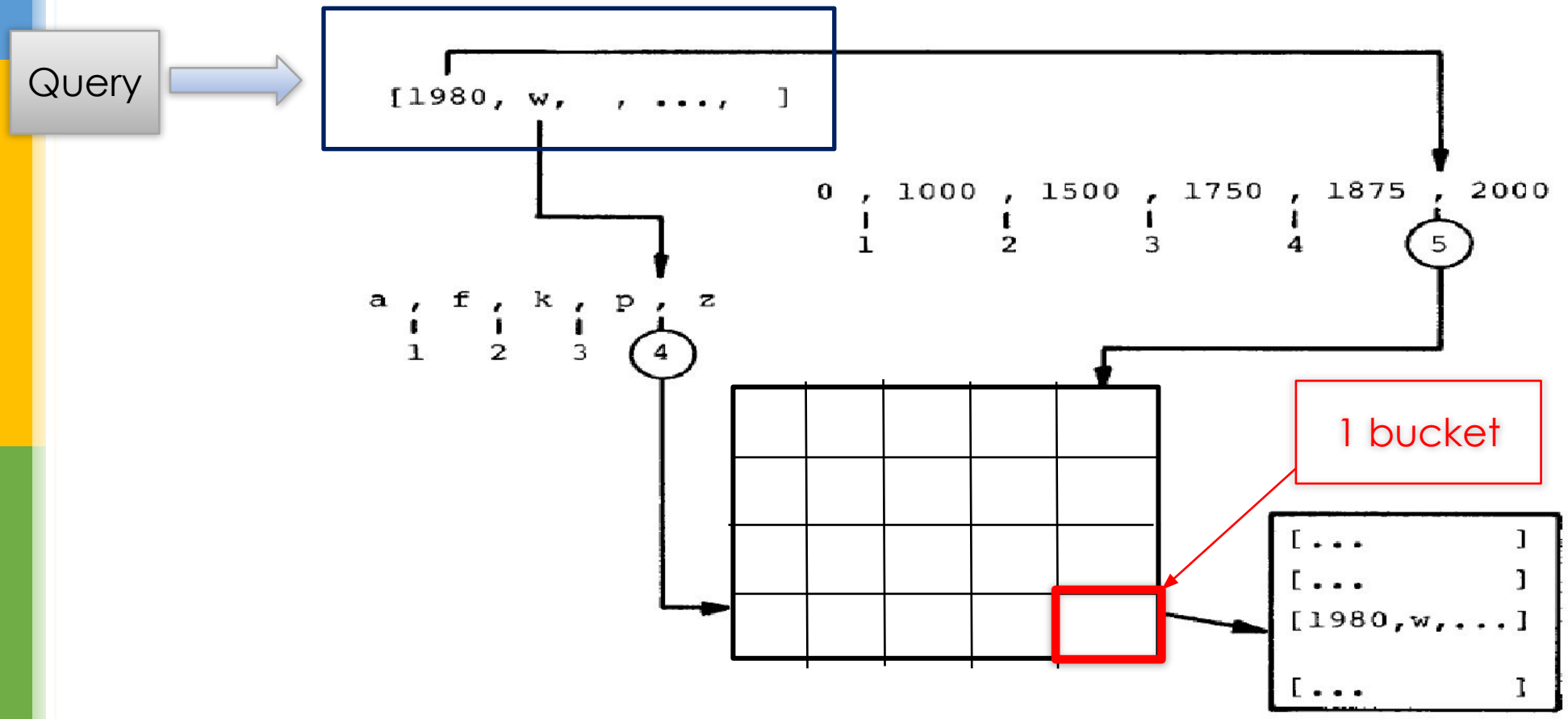
Grid Files (Splitting Policies)

- **Splits:**
  - Can happen during insertion.
  - Overflow of a bucket corresponding to a grid partition leads to a split.
  - Can also happen if bucket containing records from several grid partition fills up.
  - Splitting dimension can be changed alternatively.
  - Splitting point may not always be the middle point, other algorithms are also possible.

J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (Querying example)

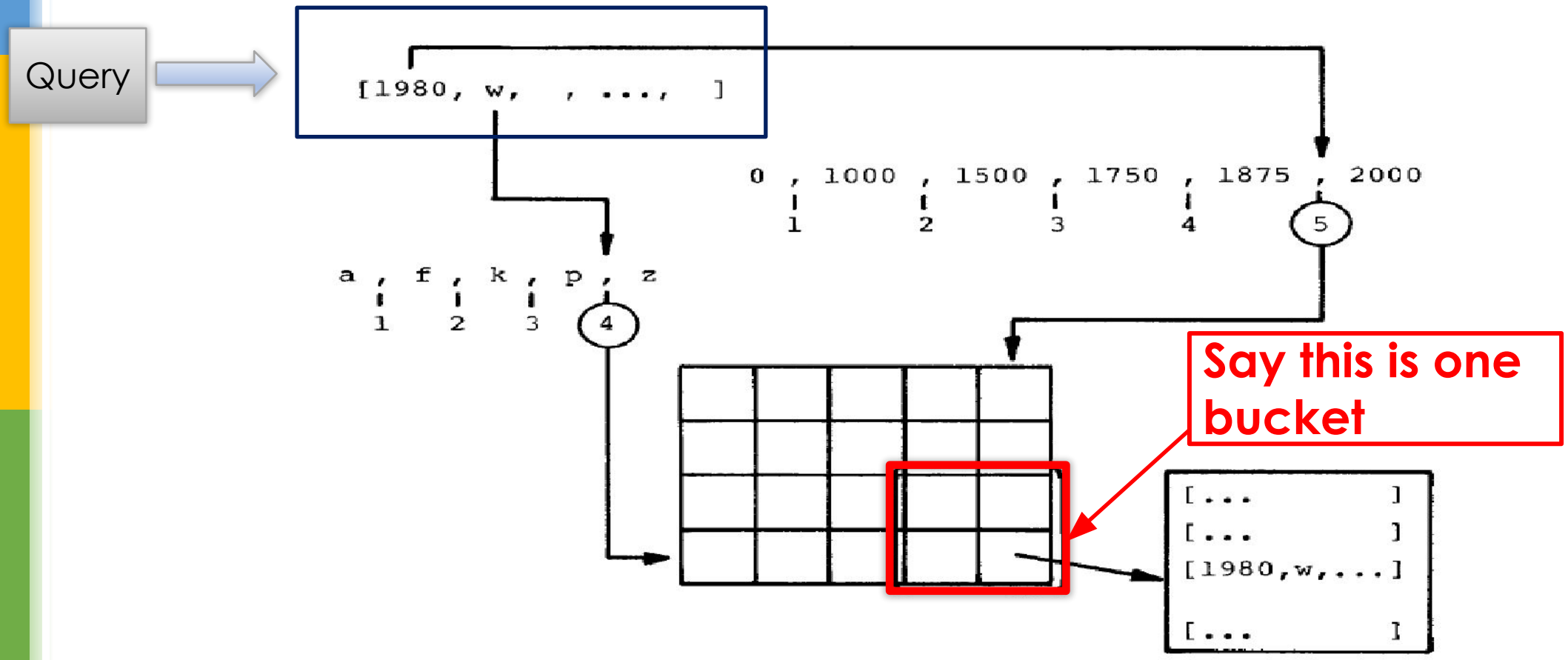
- X-partitions (0,1000,1500,1750,1875,2000)
- Y-partitions (a, f, k, p, z).



J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

Grid Files (Querying example)

- X-partitions (0,1000,1500,1750,1875,2000)
- Y-partitions (a, f, k, p, z).



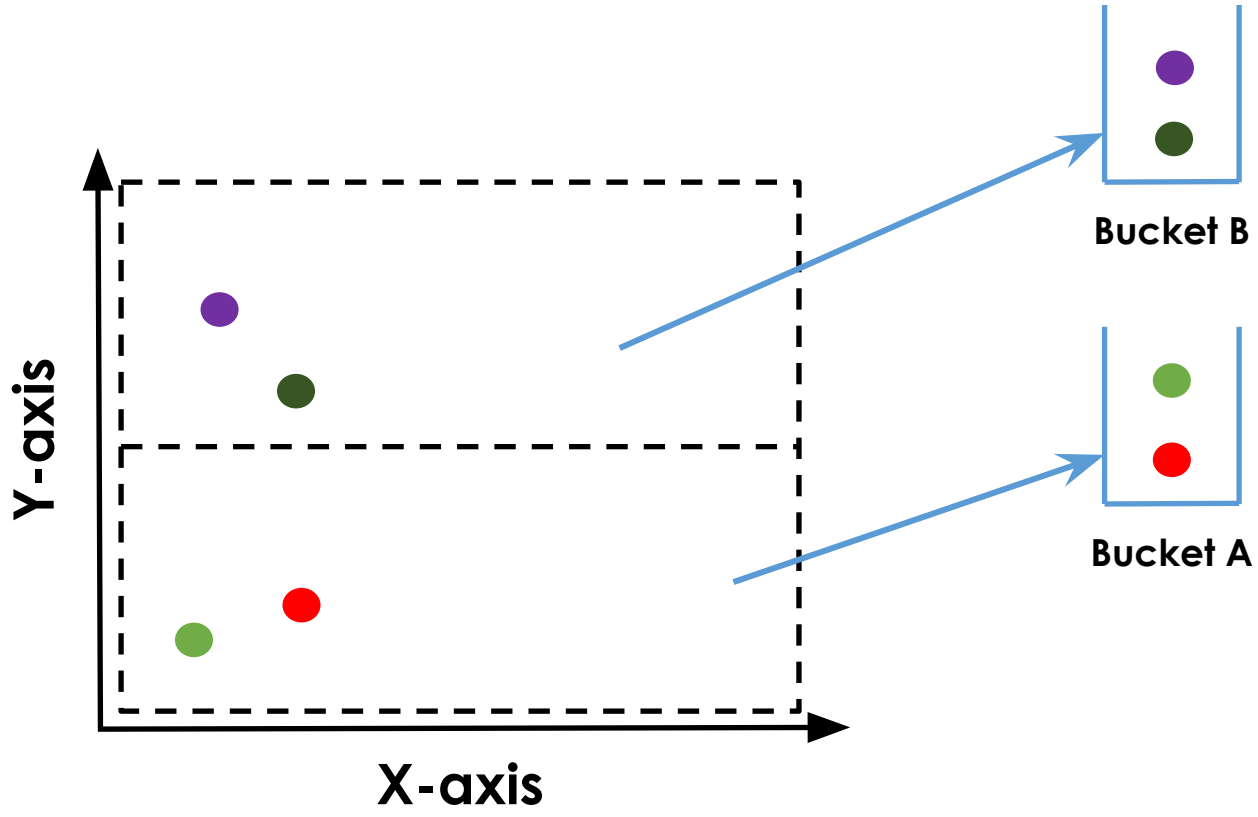
J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

More Examples on Grid Files (Inserting Points)

J. Nievergelt and H. Hinterberger. The Grid File: An Adaptable, Symmetric Multikey File Structure. ACM Transactions on Database Systems, Vol. 9, No. 1, March 1994

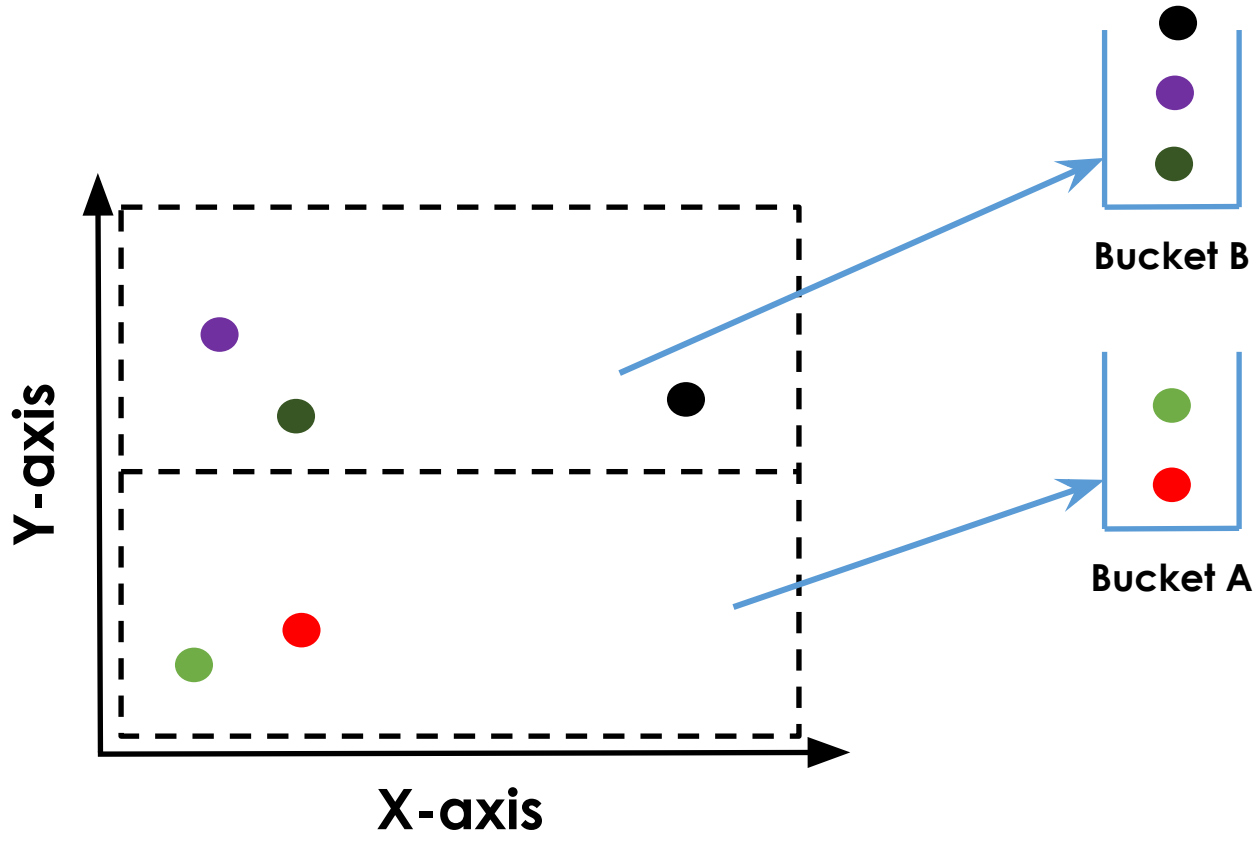
Grid Files (Another example)

- Assume Bucket size = 3



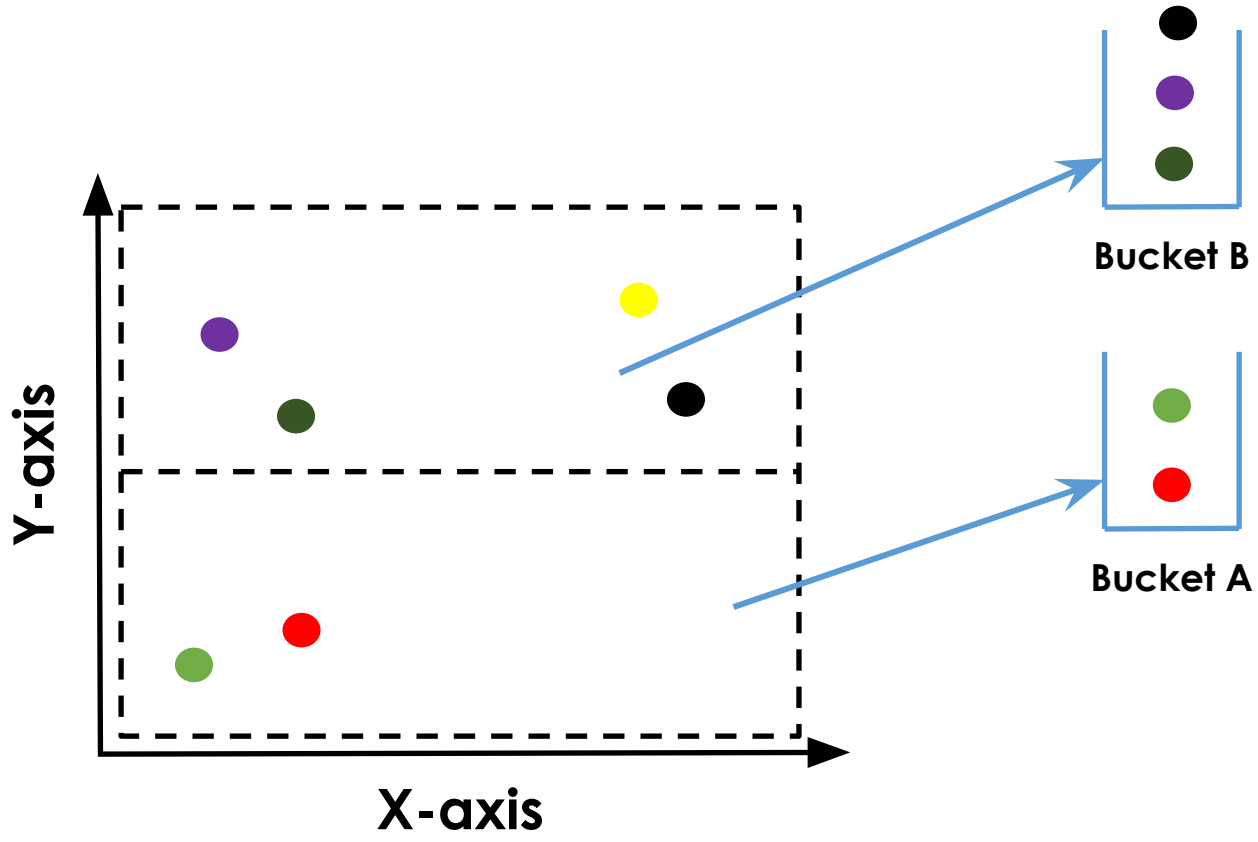
Grid Files (Another example)

- Assume Bucket size = 3



Grid Files (Another example)

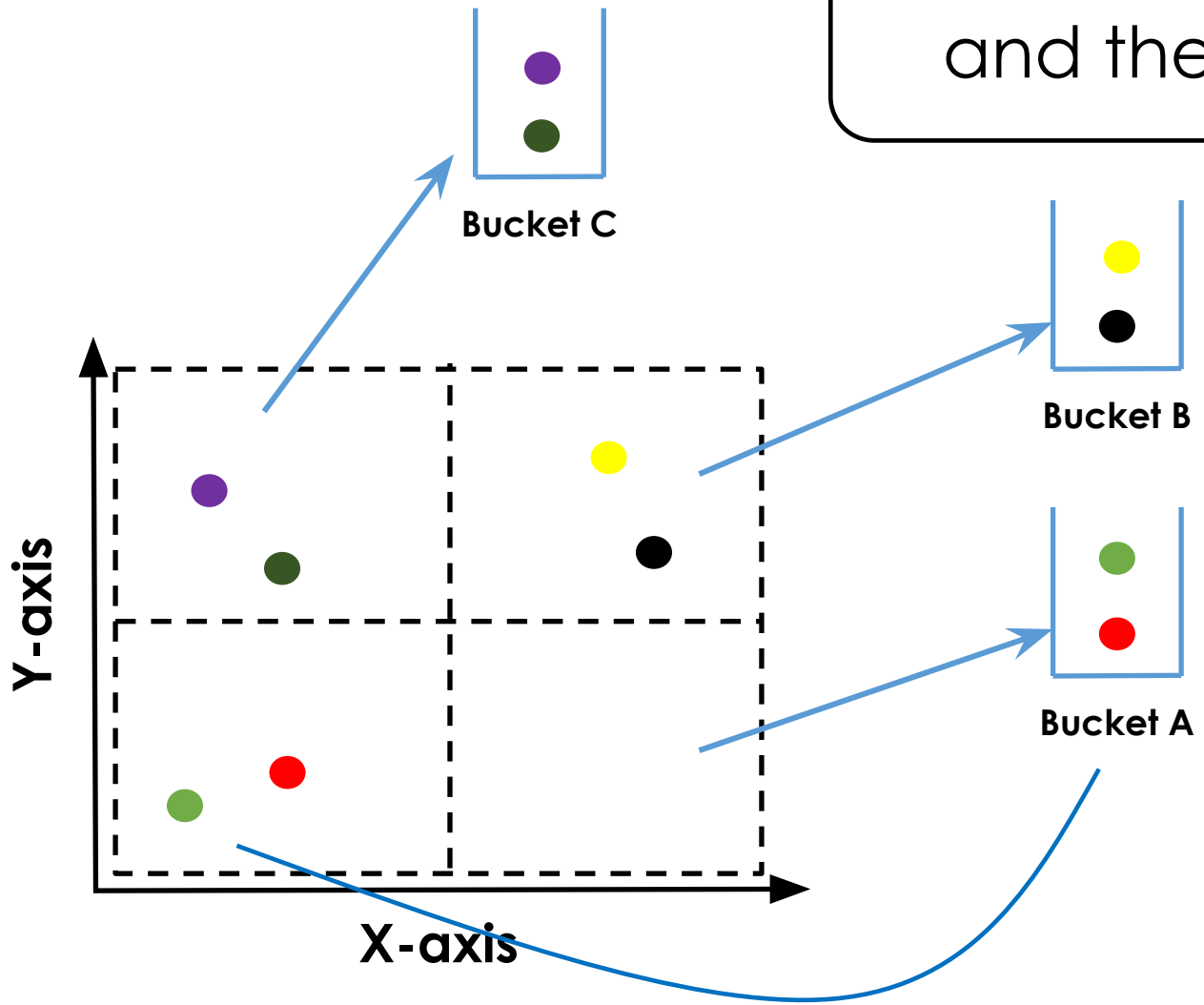
- Assume Bucket size = 3



Grid Files (Another example)

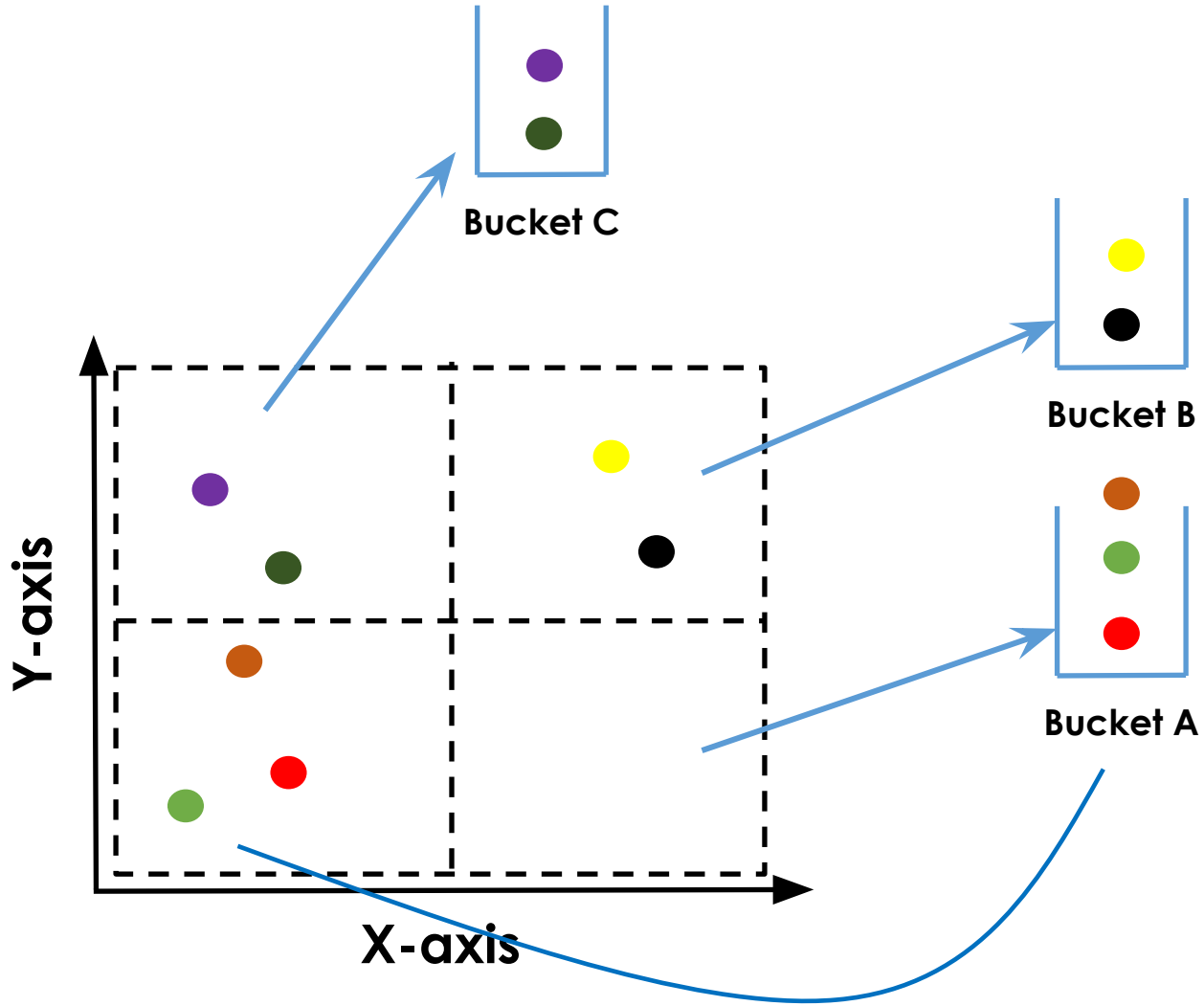
- Assume Bucket size = 3

Both Directory (x & y scales) and the Bucket are split



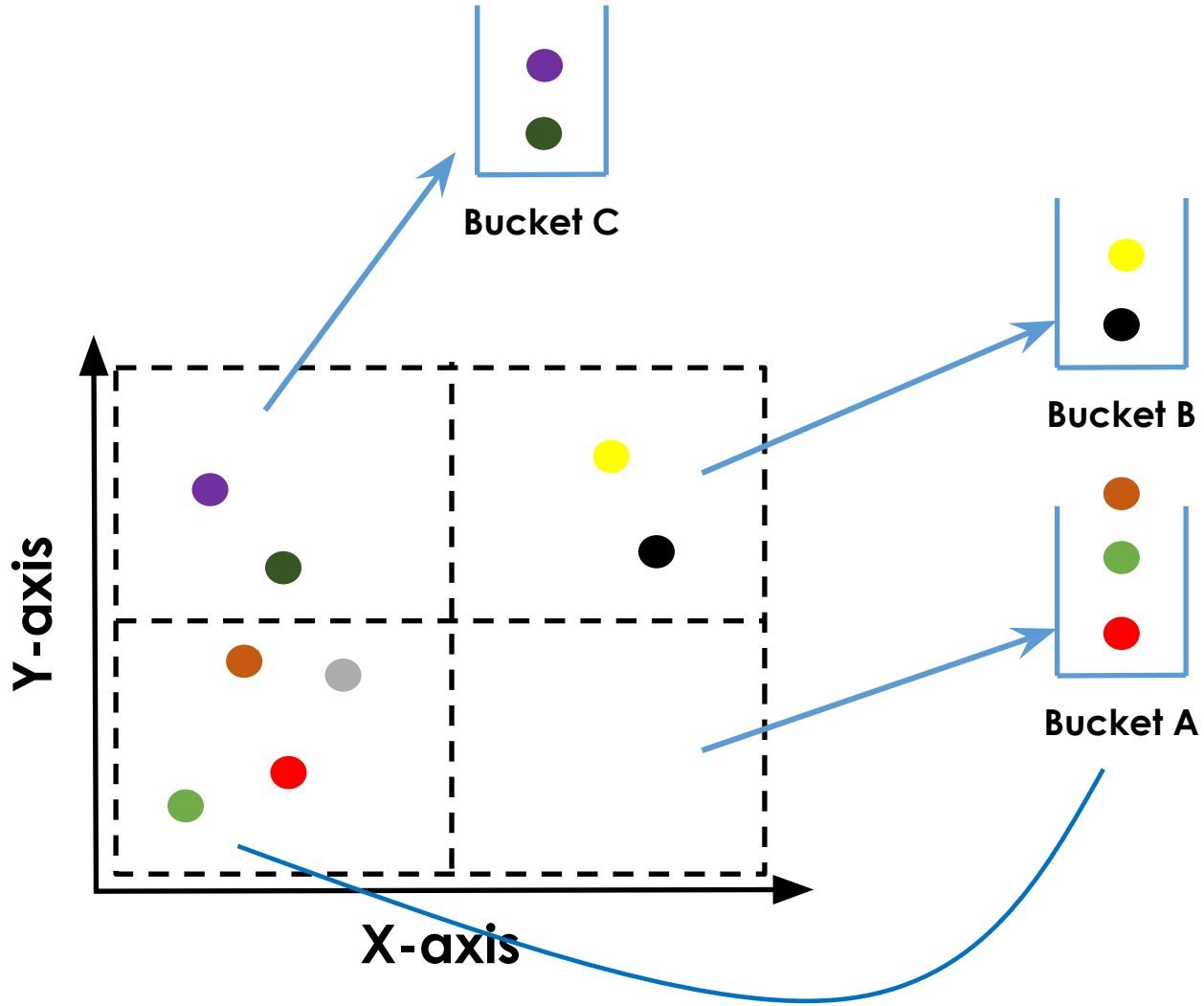
Grid Files (Another example)

- Assume Bucket size = 3



Grid Files (Another example)

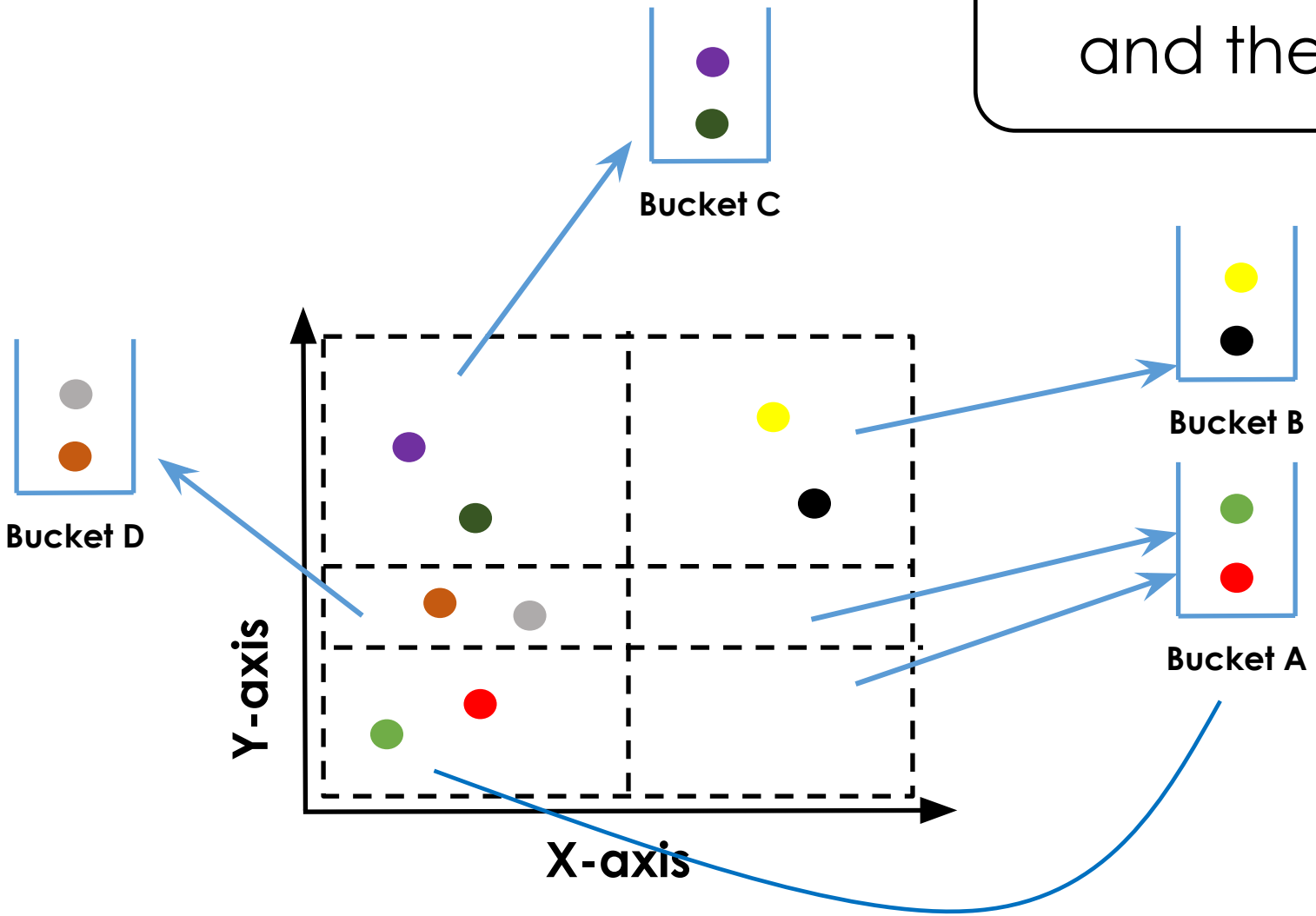
- Assume Bucket size = 3



Grid Files (Another example)

- Assume Bucket size = 3

Both Directory (x & y scales) and the Bucket are split



- Assume Bucket size = 3

In general, there can be multiple options while splitting. For e.g., cell C1 could be pointing to bucket D as well. Also it is ok for multiple cells to point to the same bucket.