

The Grid File: An Adaptable, Symmetric Multikey File Structure

Authors: Nievergelt, Hinterberger, and Sevcik
Presented By: Neeraj Gupta

10/21/2009 EECS 584, Fall 2009 1

Contents

- Motivation
- Grid File Design
- Operations
 - Single Record Retrieval
 - Range Queries
 - Insertion
 - Deletion
- Conclusion

10/21/2009 EECS 584, Fall 2009 2

Motivation (B+ tree)

Query [1450-1600, c-g]
– Only 1 matching record

CAD, geo-data apps, interactive apps require efficient multi-D queries

10/21/2009 EECS 584, Fall 2009 3

Grid File is

- Adaptable** (w.r.t. insertions/deletions)
 - Efficient query handling
- Symmetric**
 - No distinction b/w primary, secondary keys
- Multi-key**
 - Reference records using subset of keys
- File Structure**
 - How data is stored in a file

10/21/2009 EECS 584, Fall 2009 4

Record Space and Grid Blocks

Divide record space into grid blocks

To modify partition?
Split/Merge intervals

10/21/2009 EECS 584, Fall 2009 5

Grid File Structure

- Allocates storage in units of fixed size
 - disk blocks/pages/buckets
- To map grid blocks to buckets?
 - Use grid directory

10/21/2009 EECS 584, Fall 2009 6

Grid Directory (k = 2)

Linear Scales (k 1-d arrays)

0	1000	1500	1750	1875	2000
1	2	3	4	5	

a f k p z

1 2 3 4

Grid Array (one k-d array)

Bucket

```
[... ]
[... ]
[1980, w, ... ]
[... ]
```

> Grid Directory (Grid Array + Linear Scales)
 > Many to one – low bucket occupancy
 > k = number of keys referencing records

10/21/2009 EECS 584, Fall 2009 7

Follows two principles

- Two-disk-access principle
 - Retrieve single record in at-most 2 disk accesses:
 - Accessing directory
 - Accessing bucket
- Efficient range queries w.r.t. all attributes

10/21/2009 EECS 584, Fall 2009 8

Single Record Retrieval

[1980, w, ... ,]

0	1000	1500	1750	1875	2000
1	2	3	4	5	

a f k p z

1 2 3 4

Grid Directory

Bucket

```
[... ]
[... ]
[1980, w, ... ]
[... ]
```

- Find [1980,w] in the bucket?

10/21/2009 EECS 584, Fall 2009 9

Range Query Example

[1450-1600, c-g, ... ,]

0	1000	1500	1750	1875	2000
1	2	3	4	5	

a f k p z

1 2 3 4

Buckets

Precision: (# matching records)/(# records retrieved)

10/21/2009 EECS 584, Fall 2009 10

Grid File Insertion

Grid Blocks

Buckets

A

Bucket A overflows !! Split it !

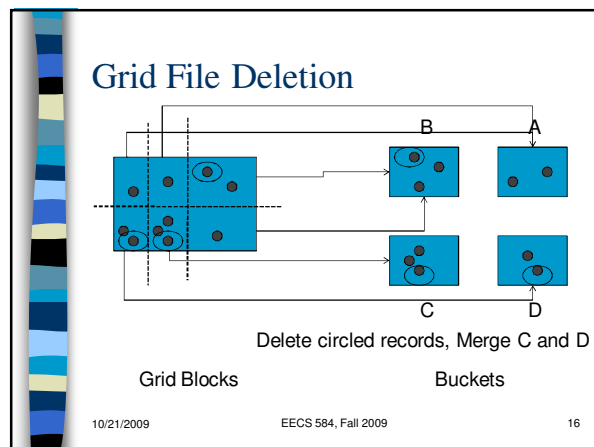
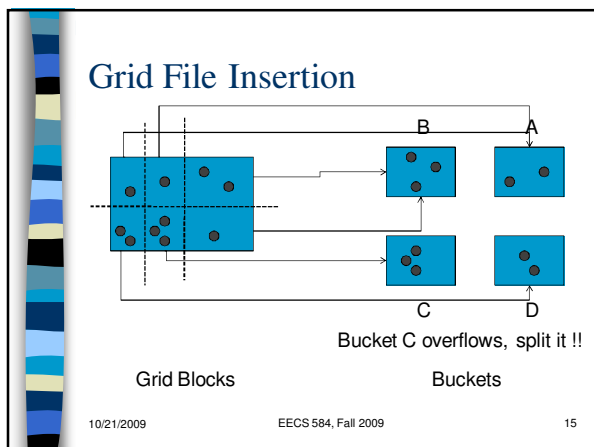
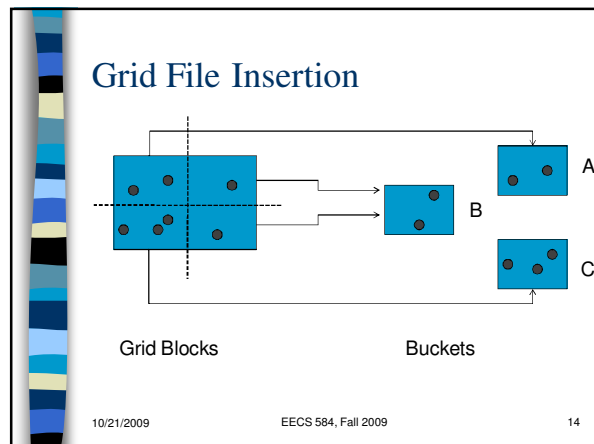
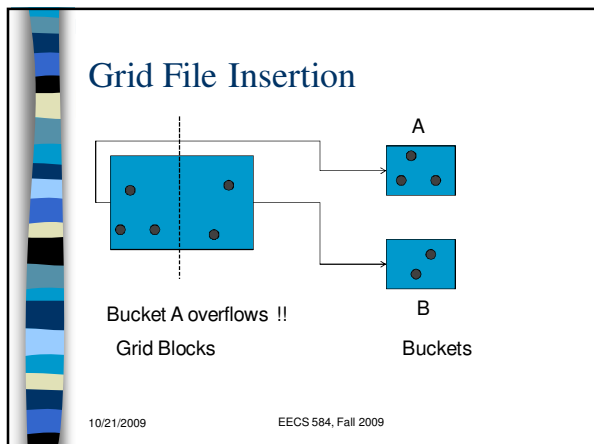
Overflow on 5 records

10/21/2009 EECS 584, Fall 2009

How to split ?

- Several possibilities
 - Dimension
 - Choose dimension according to fixed schedule, perhaps cyclically
 - May favor some attribute(s)
 - Location
 - Can be midpoint

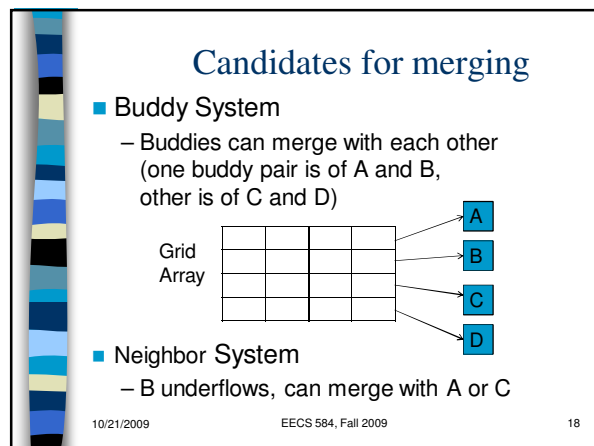
10/21/2009 EECS 584, Fall 2009 12



Bucket Merging

- Controlled by 3 decisions
 - What? (candidates for merging)
 - Which? (candidate has priority)
 - When? (trigger bucket occupancy merging)
 - Merging Threshold
 - % occupancy which resulting bucket shouldn't exceed when two buckets are merged
 - Up to 70% is good (shown experimentally)

10/21/2009 EECS 584, Fall 2009 17



Problem with Buddy System

- Might not guarantee high bucket occupancy for long deletions

A: [Block with 3 records]
 B: [Block with 3 records]
 C: [Block with 1 record]

Deletion from B → [Block with 2 records]

Cannot Merge C with B Cannot Merge B with C

10/21/2009 EECS 584, Fall 2009 19

Grid File Deletion

Grid Blocks Buckets

Delete circled records, Merge C and D

10/21/2009 EECS 584, Fall 2009 20

Grid File Deletion

Grid Blocks Buckets

60% bucket occupancy

10/21/2009 EECS 584, Fall 2009 21

Directory Merging

- Directory merging also possible
 - Shrinking file
 - Dynamic Weighting of attributes
 - Not using some attributes

10/21/2009 EECS 584, Fall 2009 22

Directory Merging Example

No queries between [a-k] and [0-1500]

0	1000	1500	1750	1875	2000
1	2	3	4	5	

a f k p z
1 2 3 4

Bucket

- [1876, w, ...]
- [1877, w, ...]
-
- [1900, w, ...]
-
- [2000, w, ...]

23

10/21/2009 EECS 584, Fall 2009

Directory Merging

No queries between [a-k] and [0-1500]

0-1500	1750	1875	2000
1	2	3	

a-k p z
1 2

Bucket

- [1876, w, ...]
- [1877, w, ...]
- [1900, w, ...]
-
- [2000, w, ...]

- Grid Directory trimmed on merging

24

10/21/2009 EECS 584, Fall 2009

Concurrent Access

[1450, w, ...] [1980, g, ...]

	0	1000	1500	1750	1875	2000
	1	2	3	4	5	
a						
f						
k						
p						
z						

1 2 3 4

[...]
[...]
[1980, g, ...]
[...]

[...]
[...]
[1450, w, ...]
[...]

No root node as in trees (bottleneck if present), allows concurrency

10/21/2009 EECS 584, Fall 2009 25

Assumptions

- No correlated attributes
- Small number of attributes, large domain

10/21/2009 EECS 584, Fall 2009 26

Conclusion

- Designed for efficient multi-D queries
 - 2 disk accesses for single record access
 - Efficient range queries
- Adaptable, Symmetric, Multi-key File Structure
- Splitting/Merging Policy configurable during insertions/deletions
- Allows simpler concurrency control protocols

10/21/2009 EECS 584, Fall 2009 27

Thanks !

10/21/2009 EECS 584, Fall 2009 28