DB2 – WS 2017/2018

Hashorganisation – 01.11.2017

Dr. Karsten Tolle

Hash(funktion)

- Anwendungen:
  - Passwörter/Kryptographie
  - Prüfsummen
  - **Datenbanken**

- Hashfunktion $h:K \rightarrow S$ ist eine Funktion, wobei die Mächtigkeit von S kleiner oder gleich von K ist ($|K| \geq |S|$) .

Hash in Java

```
Hashtable<Integer,String> myTable= new Hashtable<Integer,String>();
myTable.put(1, "John");

HashMap<Integer,String> myMap= new HashMap<Integer,String>();
myMap.put(1, “John");
```

Key, Value

```
HashSet<String> mySet= new HashSet<String>();
mySet.add ("First");
```

Key

HashSet()	Constructs a new, empty set; the backing HashMap instance has default initial capacity (16) and load factor (0.75).

https://www.quora.com/What-is-the-difference-between-HashSet-HashMap-and-hash-table-How-do-they-behave-in-a-multi-threaded-environment

Hash in Datenbanken als

- **Indexstruktur**
- Partitionierung
- Join-Implementierung

HASH Partitions

Partition Determined by a Hash from a Given Expression

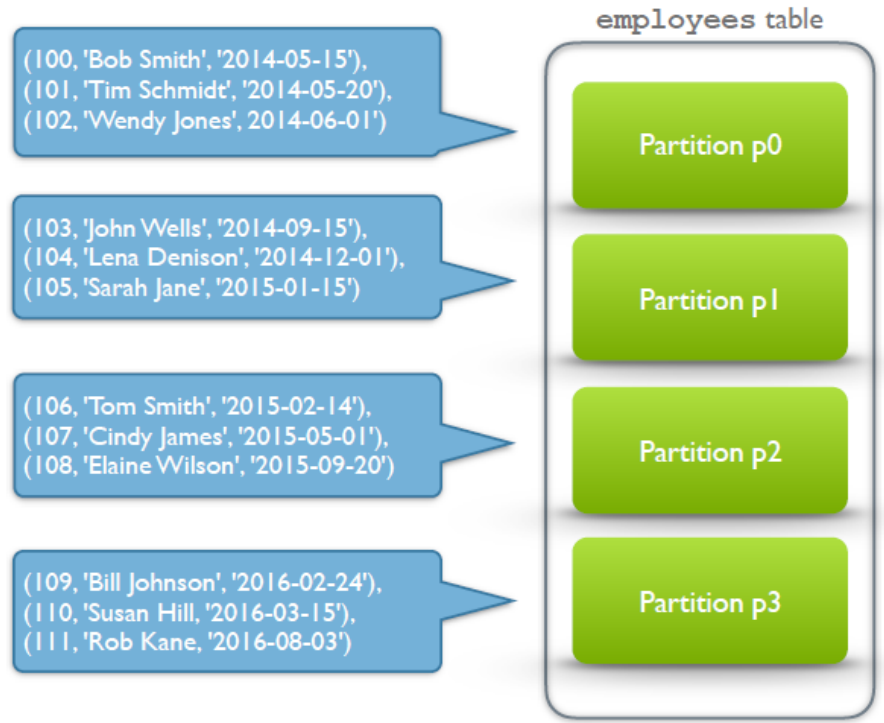
```
CREATE TABLE employees (
  id INT AUTO_INCREMENT KEY,
  name VARCHAR(20), dept_id INT,
  join_date DATE)
ENGINE = MyISAM
PARTITION BY HASH(id) PARTITIONS 4;
```

Partitions contain Same Number of Rows

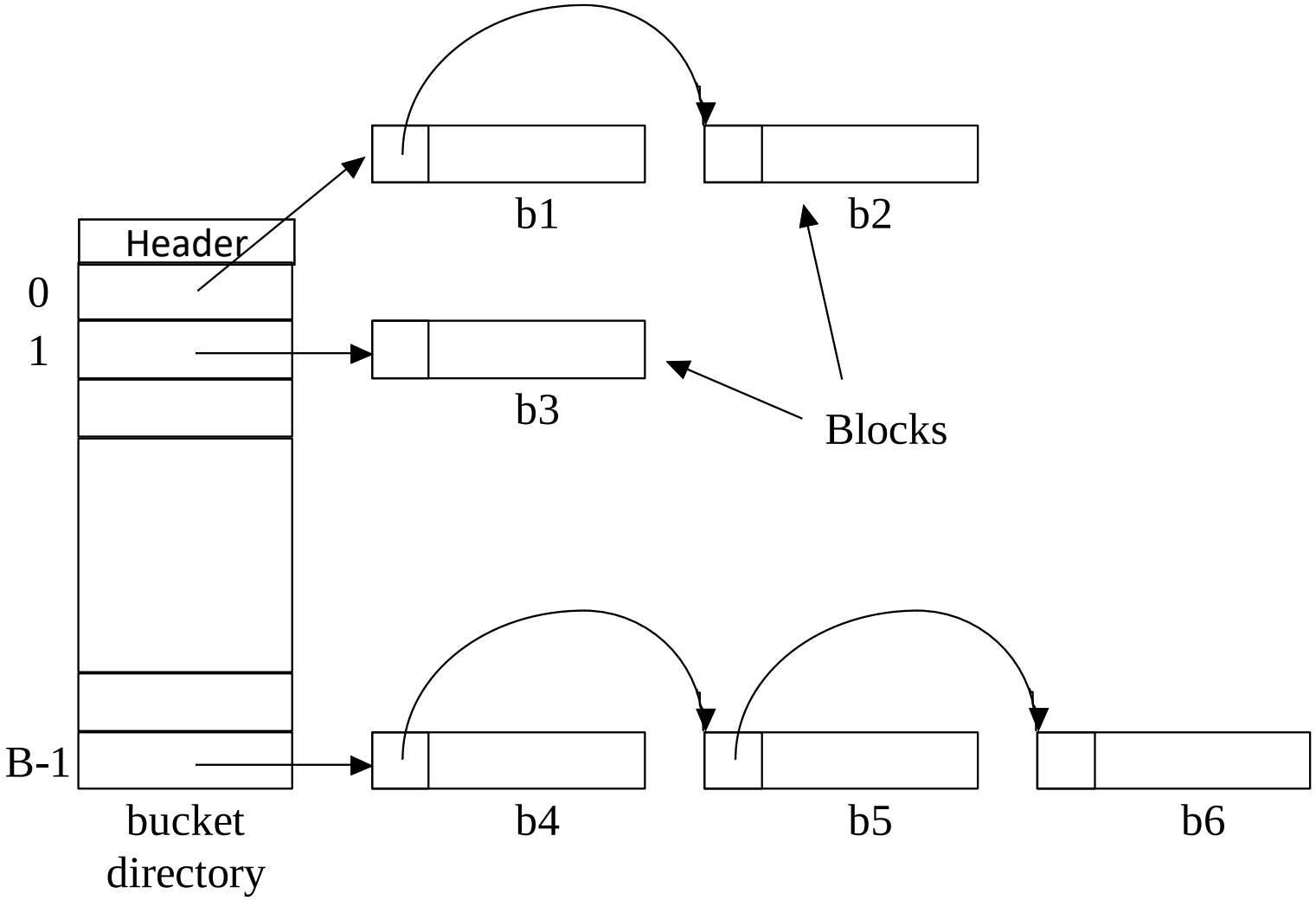
Use **COALESCE** to Reduce Number of Partitions

```
ALTER TABLE employees
COALESCE PARTITION 1;
```

Distribution of Rows



Hashorganisation – Indexstruktur über Schlüssel



Beispiel (wir gehen von unpinned Datensätzen aus!)

PNR	Name	Gehalt
1	Albert	500
12	Marie	1000
3	Anna	2000
24	Anna	200
5	Edsger	800
6	Einstein	900
15	Anna	500
16	Curie	700
21	Turing	1000

Integer; char(10); Integer

Hashfunktion: h(PNR) := PNR mod 3

Blockgröße: 32 Byte
Integer: 1 Byte
char(n): n Byte
Verweis auf Block: **x Byte**

Fragen zu einer gegebenen Situation

- Wie viele Blöcke werden für das Bucket-Directory benötigt?
- Wie viele Datensätze passen in einen Block?
- Beim Zeichnen auf Einfüge-Reihenfolge achten.