Assignments 8
LSM / Mongo DB

eva.knirsch@kiu.edu.ge

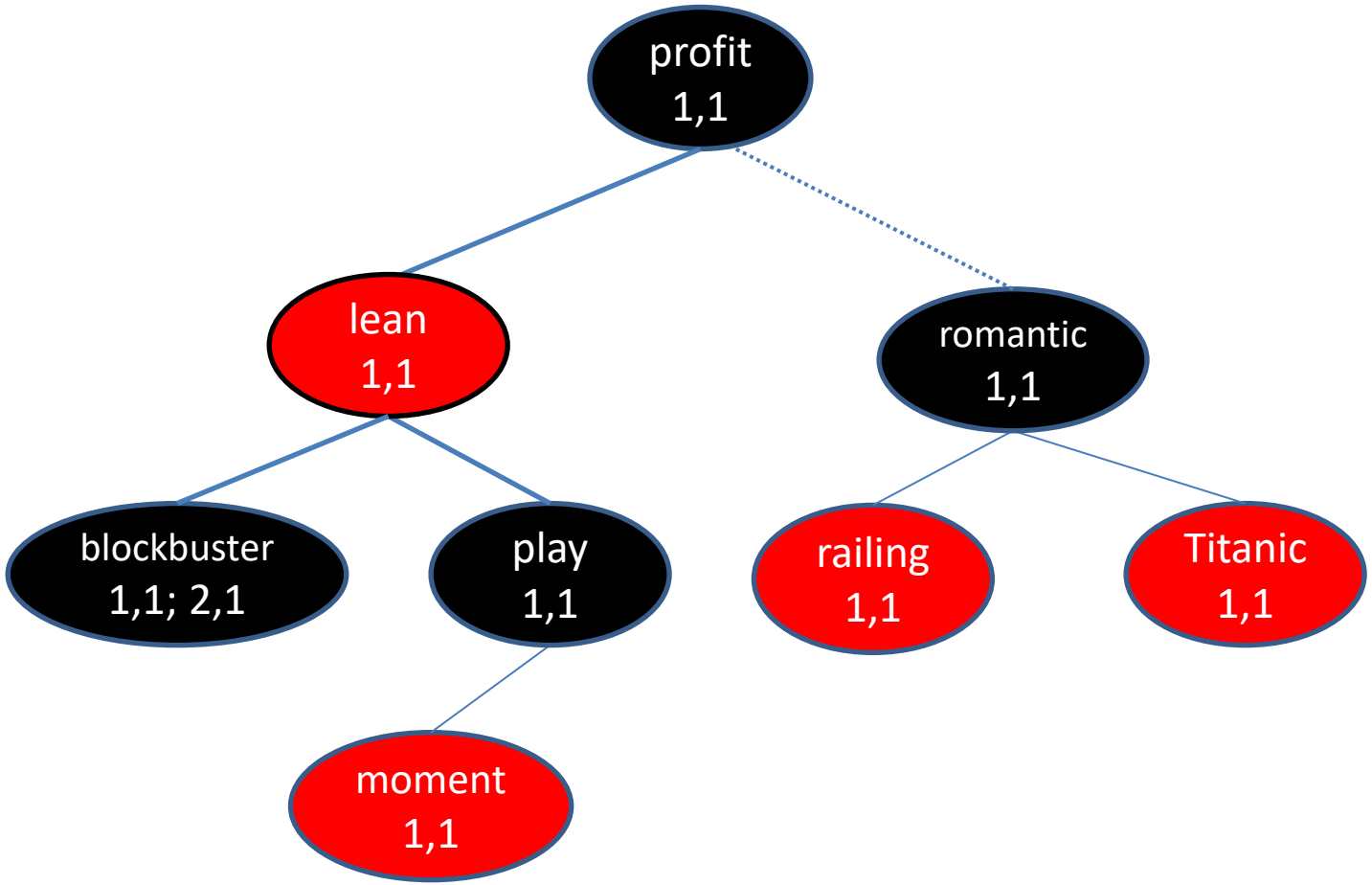
Insert :

{necessary: 2,1}

{film: 2,2}

{actor:2,3}

{life:2,1}

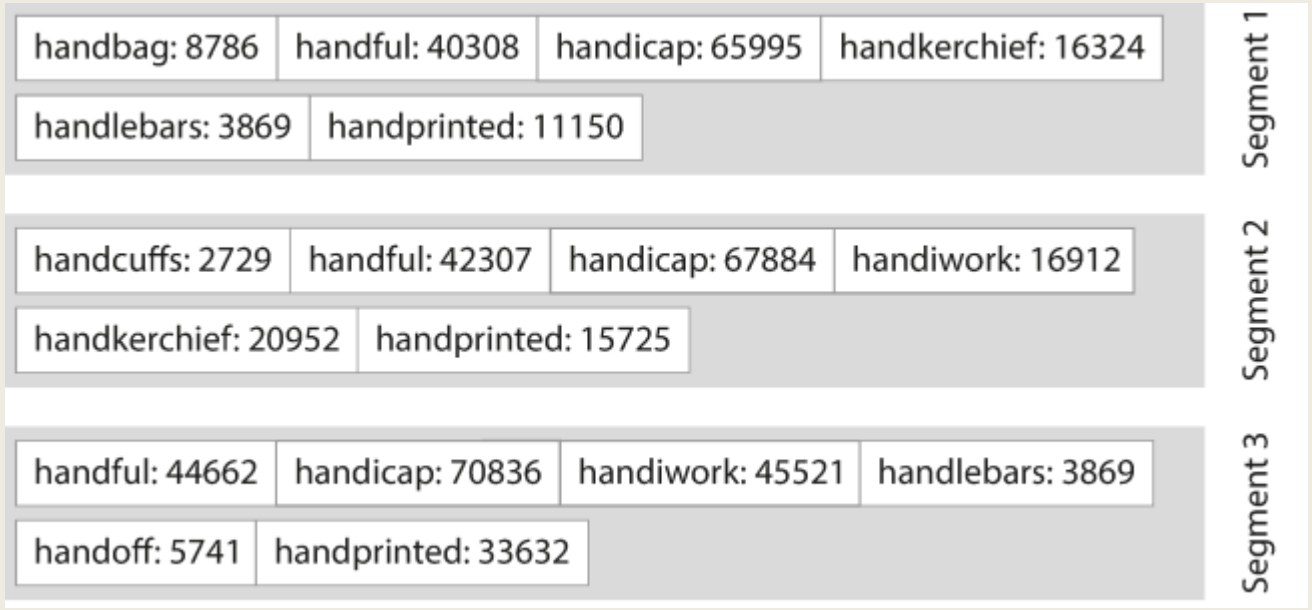


Document each step.

eva.knirsch@kiu.edu.ge

LSM-Storage: 3. Compaction / Merge

- Compaction means to merge segments on disk and discard duplicate keys with old values.



Segment 1 is oldest, segment 3 is most recent.

If there are duplicate keys in different segments, which value is the current one and which can be discarded?

What does the merged segment look like?

Example taken out of [Kl], p. 76

eva.knirsch@kiu.edu.ge

LSM-Storage: Delete Operations

Assume that the key-value pair handprinted: 33632 needs to be deleted.

The deletion process in LSM works with a tombstone as value for the key to be deleted.

1. Explain: What is a tombstone process?
2. Why is a tombstone process necessary in LSM storage? Give a concrete example.
3. When can the tombstone key-value pair be finally be deleted itself?

eva.knirsch@kiu.edu.ge

Sparse Index

LSM works with sparse indexes of SSTables. These indexes are held in memory. They are used to speed up read operations

1. What is a sparse index? Give an example with data and sparse index.

2. Why is it not necessary to hold a full index in memory for LSM?

3. How can a sparse index speed up a search request? Give a concrete example.

eva.knirsch@kiu.edu.ge

Mongo DB

eva.knirsch@kiu.edu.ge

Install Mongo DB locally

1.Download MongoDB Community Server
https://www.mongodb.com/try/download/community
Installs the server and the GUI client, including MongoDB Compass shell

2.Shell download – optional – command line shell
https://www.mongodb.com/try/download/tools
MongoDB Shell

3.Add path to the **bin** directory of the shell as environment variable.

4.Test: enter on command line.
mongosh -help

5. If you want to download additional mongodb tools:
https://www.mongodb.com/try/download/database-tools

knirsch@htw-berlin.de

Document Databases

Terminology

RDBMS	Key-Value	Document
database	cluster	database
table	bucket	collection
tuple	key-value	document
PK / rowID	key	objectID

Data Import

1. Create db lesson in MongoDB
2. Create teacher and student collection in MongoDB.
3. Export your teacher and student tables in .json format out of Postgres lesson db
4. Import into MongoDB. You can use Compass GUI to do this.
5. Each tuple becomes a document.
6. What does MongoDB add to each document?
7. What index does MongoDB create on each collection automatically?

knirsch@htw-berlin.de

Mongo DB Syntax

1. Commands start with: **db.**
2. Followed by the **name of the collection** (to which the operation refers).
3. Then followed by the **method()**

Query commands refer to one collection

Examples:

db.<collectionname>.countDocuments()	Count documents of a collection
db.<collectionname>.find({})	Display all documents of a collection

db.collection.find(query, projection, options)

find() returns a cursor to the documents that match the specified query criteria. Per default, it returns the whole document. If you only want to get certain fields back, you need to add a corresponding projection.

knirsch@htw-berlin.de

Basic Queries

Counting the number of documents:

```
db.<collectionname>.count()                 #depreciated
db.<collectionname>.countDocuments()        #new
```

Show all documents:

```
db.<collectionname>.find({})
```

Pretty output format: (line breaks, one value per line):

```
db.<collectionname>.find({}).pretty()
```

Return documents, sorted by name, limit number of returned documents:

```
db.teacher.find({}).sort({name:-1}).limit(5)
```

Return only documents that match a query condition, return only specific fields (projection):

```
db.teacher.find({gender: "m", },{name:1, gender:1})
```

Return documents that match a query condition, return all fields except DoB:

```
db.teacher.find({education: "Master", },{dob:0})
```


knirsch@htw-berlin.de

Basic Queries

Get some practice:

1. How many male teachers are in teacher collection?
2. Return documents of teachers that live in postalcode 4600. Do not show object_id.
3. Return all documents of male teachers in postalcode 4600.
4. Return the documents of teachers Krawinkel and Kaiser (or condition on any other two teachers)
5. Teacher Krawinkel teaches EN, DE and FR. Add the subjects to the document. (updateOne())
6. Teacher Alt teaches MA, PH and CH. Add the subjects to the doucment

knirsch@htw-berlin.de