COL 362/632
Introduction To Database Management Systems

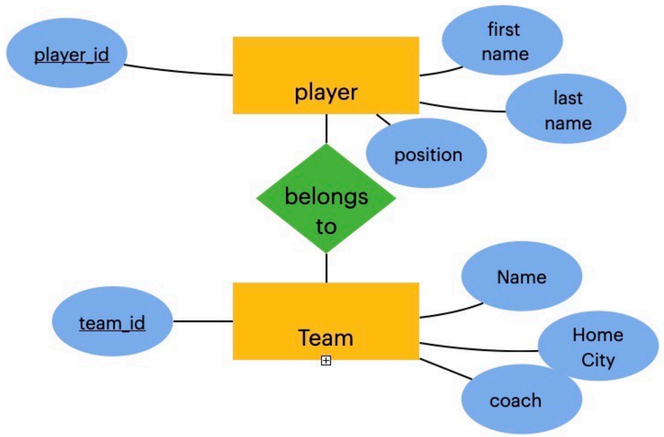
Database Internals — Overview

Kaustubh Beedkar

Computer Science & Engineering

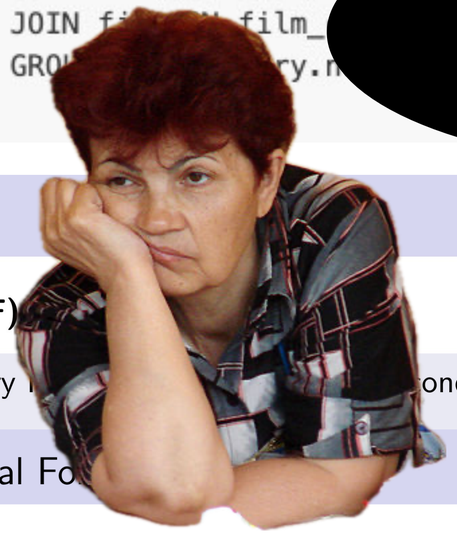
Indian Institute of Technology Delhi

Why Databases are cool




```
%%sql SELECT category.name, AVG(film.rental_duration)
      AS average_duration_category,
          (SELECT AVG(rental_duration)
          FROM film
          ) AS overall_average_duration
      FROM category
      JOIN film_category O
      JOIN fi  film_
      GRO  ry.n
```

Really?



Normal Forms

Boyce-Codd normal form (BCNF)

A relation R is in BCNF **iff** for every onditions hold

- $X \to Y$ is trivial FD
- X is a superkey

Player

PlayerId	Name	Position
123	MS Dhoni	7
193	Virat Kohli	3
131	Rohit Sharma	1
134	SK Yadav	5
154	David Warner	3

MI
DC

Normal Fo

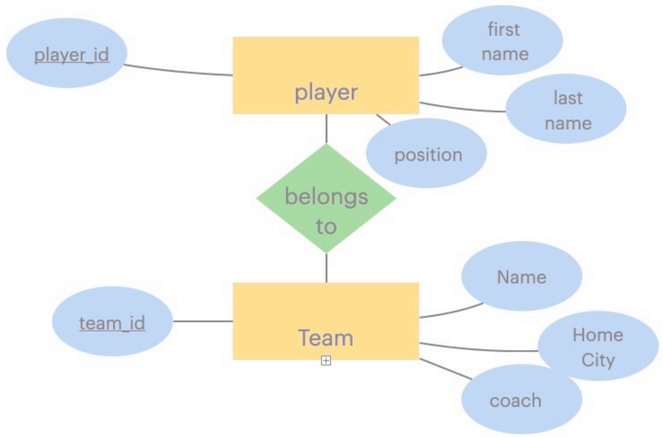
Third Normal Form (3NF)

A relation R is in 3NF if whenever $X \to Y$, one of the following is true

1. $X \to Y$ is trivial FD
2. X is superkey
3. $\forall a \in Y \setminus X$, a is prime attribute

Note: If R is in BCNF $\implies$ it is in 3NF

Why Databases are cool




```
%%sql SELECT category.name, AVG(film.rental_duration)
      AS average_duration_category,
          (SELECT AVG(rental_duration)
          FROM film
          ) AS overall_average_duration
      FROM category
      JOIN film_category
      JOIN film ON film_
      GROUP BY category.
```

This is ~~cool~~ **important** for

- User of the database
- Application developer

Not so exciting as a computer scientist!

Normal Forms

Boyce-Codd normal form (BCNF)

A relation R is in BCNF **iff** for every FD $X \to Y$, at least o

- $X \to Y$ is trivial FD
- X is a superkey

Player

PlayerId	Name	Position	PlaysFor
123	MS Dhoni	7	CSK
193	Virat Kohli	3	RCB
131	Rohit Sharma	1	MI
134	SK Yadav	5	MI
154	David Warner	3	DC

Normal Forms

Third Normal Form (3NF)

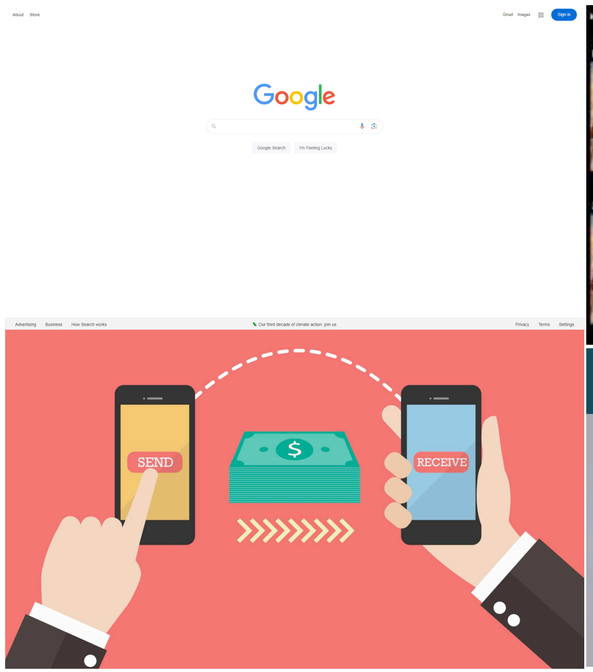
A relation R is in 3NF if whenever $X \to Y$, one of the following is true

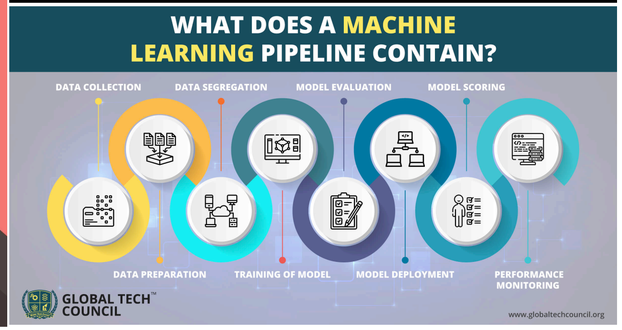
1. $X \to Y$ is trivial FD
2. X is superkey
3. $\forall a \in Y \setminus X$, a is prime attribute

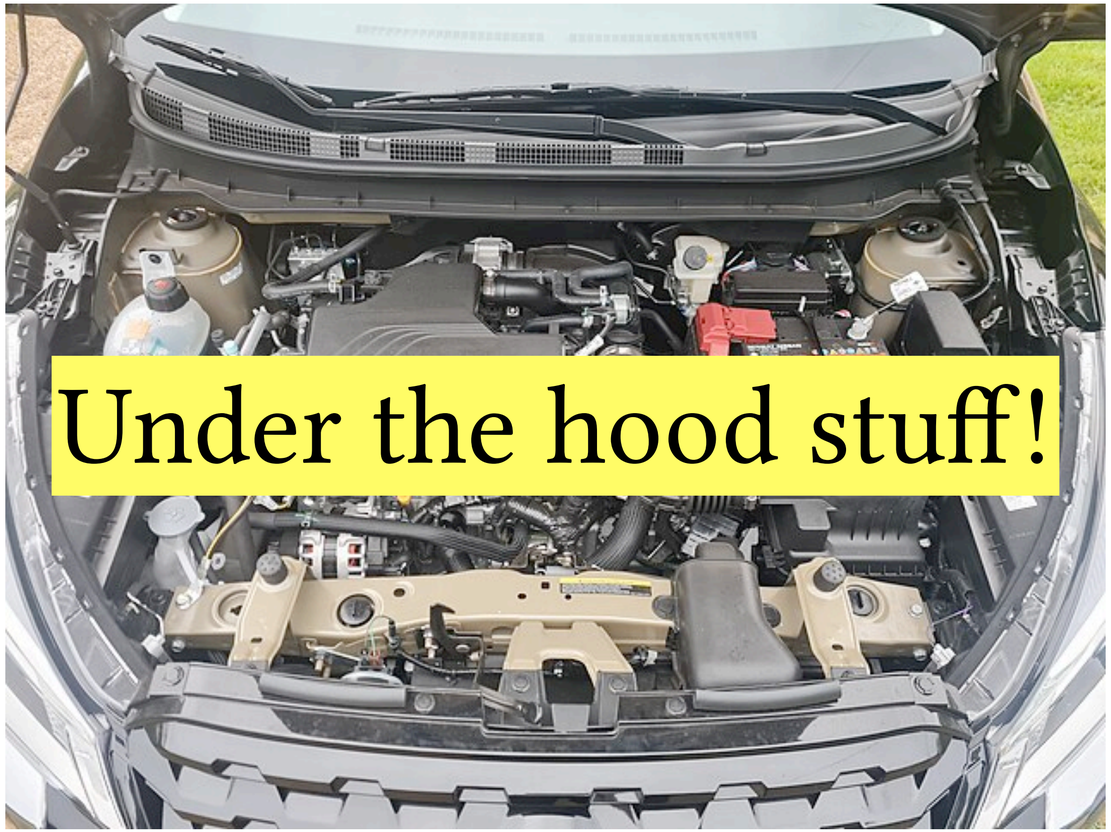
Note: If R is in BCNF $\implies$ it is in 3NF

Why Databases are cool

- Database **technology** is cool!
- Powering the digital world

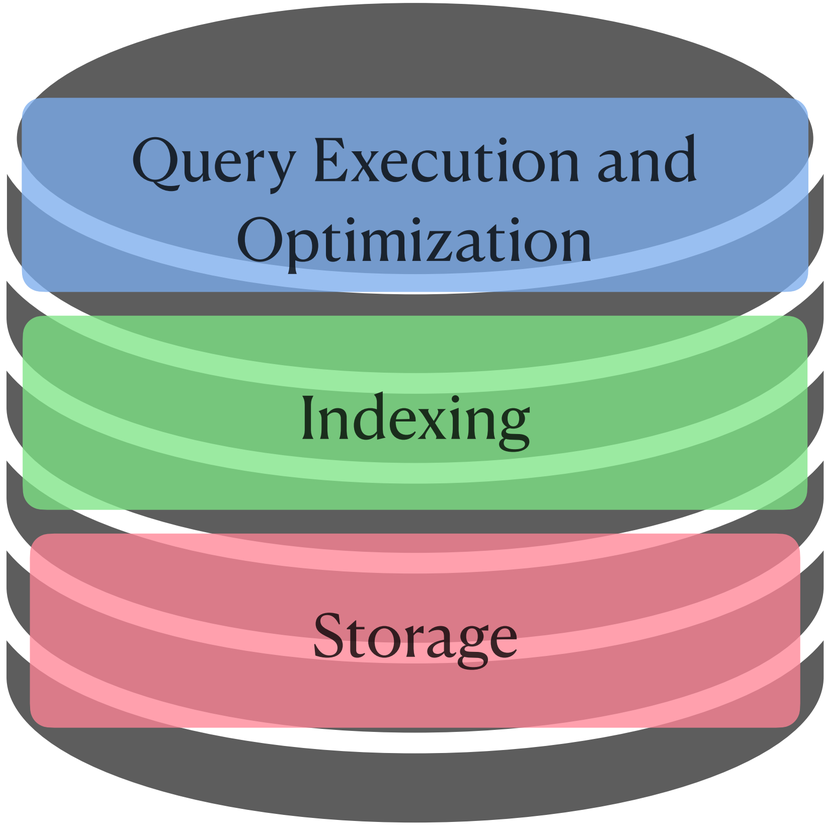






DB internals

- Three fundamental layers



Algorithms

Data Structures

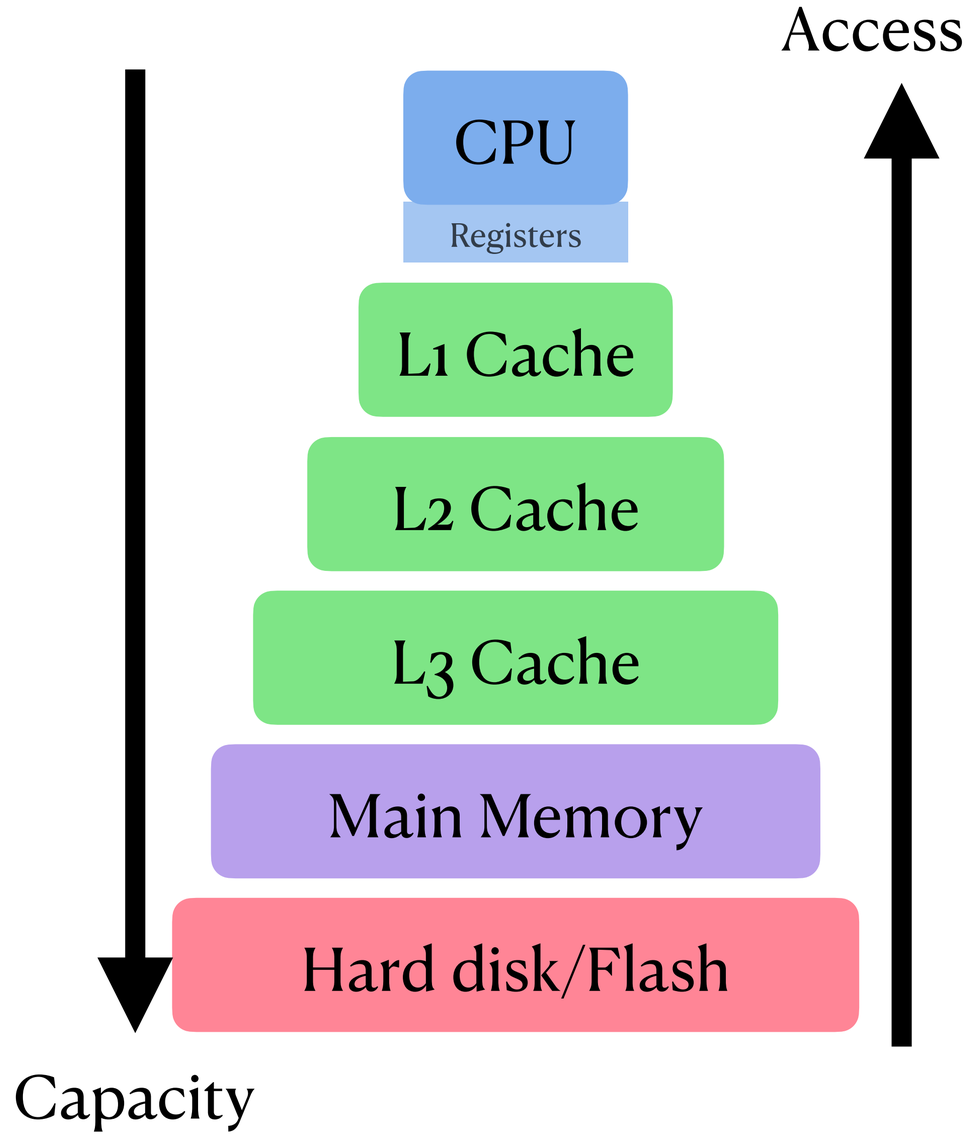
Data Layout

Storage

Imagine you have to store a trillion rows of data and retrieve them efficiently. How would you do it?

Challenges of Storage

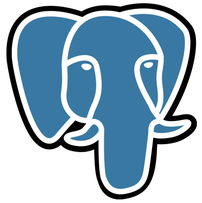
- Storing huge amounts of data while keeping it **durable**, **fault-tolerant**, and **efficient**
- Dealing with disk vs. memory trade-offs.



Storage

Key Concepts

- Buffer management
- Storage structures
  - Heap files: unstructured simple storage
  - B-Trees: used in most relational databases
  - LSM Trees: log-structured merged trees, used in No-SQL databases
  - Column Stores: used in analytical databases Snowflake, Google BigQuery



Storage

Key Concepts

- Data layout

Player

PlayerId	Name	Position	PlaysFor
123	MS Dhoni	7	CSK
193	Virat Kohli	3	RCB
131	Rohit Sharma	1	MI
134	SK Yadav	5	MI
154	David Warner	3	DC

?

2-Dimentional

1-Dimentional

Player

PlayerId	Name	Position	PlaysFor
123	MS Dhoni	7	CSK
193	Virat Kohli	3	RCB
131	Rohit Sharma	1	MI
134	SK Yadav	5	MI
154	David Warner	3	DC

Row layout

...123 | MS Dhoni | 7 | CSK || 193 | Virat Kohli | 3 | RCB || 131| Rohit Sharma | 1 | MI || ...

Storage

Key Concepts

- Data layout

Player

PlayerId	Name	Position	PlaysFor
123	MS Dhoni	7	CSK
193	Virat Kohli	3	RCB
131	Rohit Sharma	1	MI
134	SK Yadav	5	MI
154	David Warner	3	DC

?

2-Dimentional

1-Dimentional

Player

PlayerId	Name	Position	PlaysFor
123	MS Dhoni	7	CSK
193	Virat Kohli	3	RCB
131	Rohit Sharma	1	MI
134	SK Yadav	5	MI
154	David Warner	3	DC

Column layout

123 | 193 | 131 | 134 | 154 | | MS Dhoni | Virat Kolhi | Rohit Sharma | SK Yadav | David..

Storage

Beyond Databases

- Cloud storage
- Blockchain
- High performance logging

Amazon S3

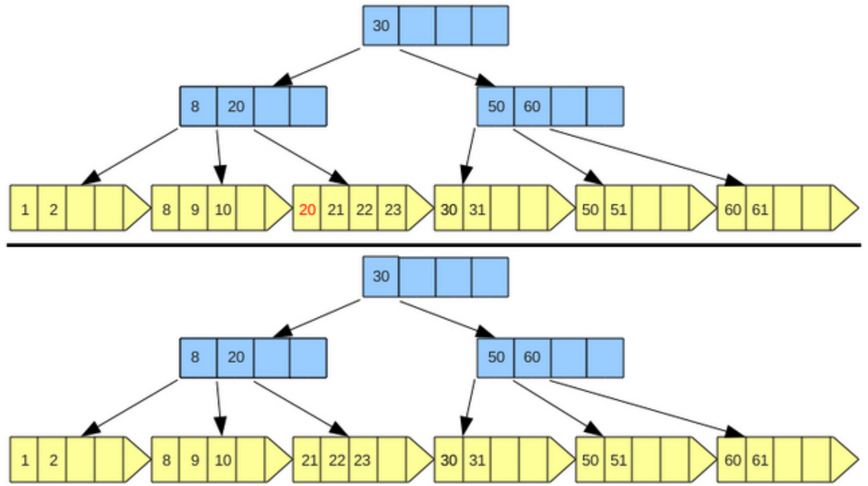
APACHE kafka®

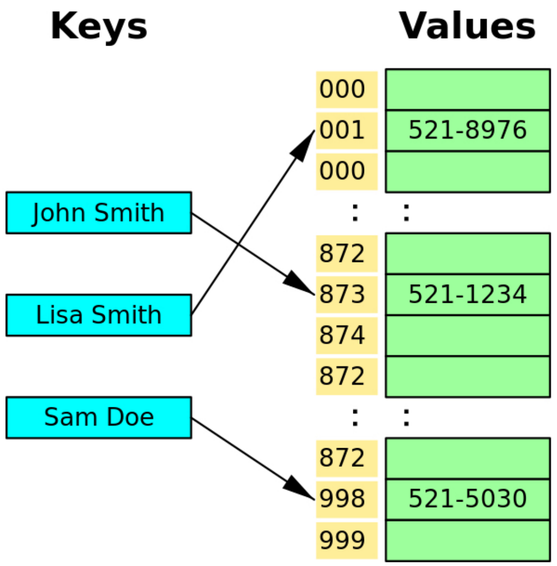
RocksDB

Indexing

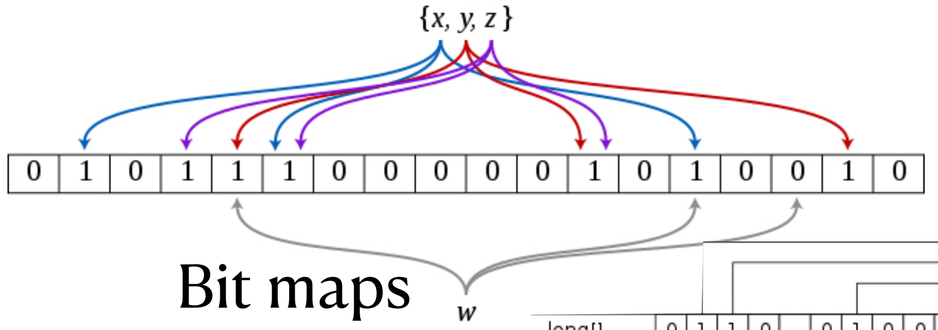
- Without indexing, searching data would require **full table scans**– painfully slow!
- Indexing **makes lookups fast** by keeping a structured map of data locations.

Trees



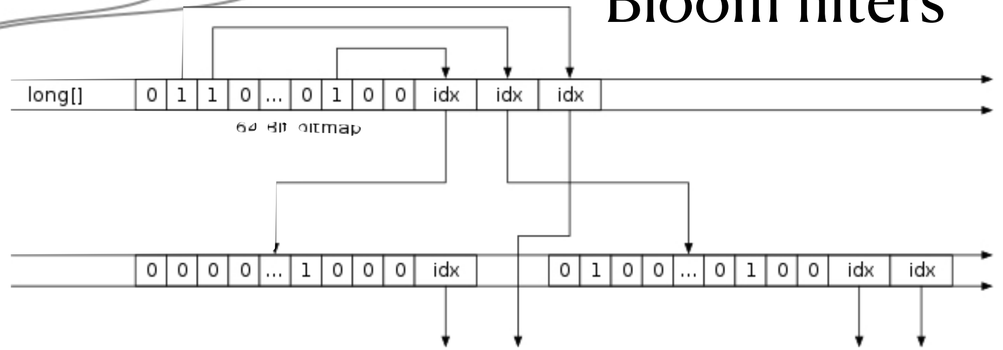


Hash Tables



Bit maps

Bloom filters

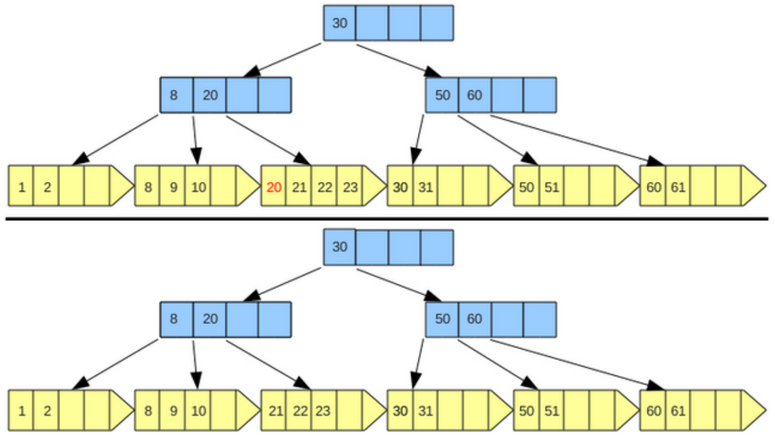


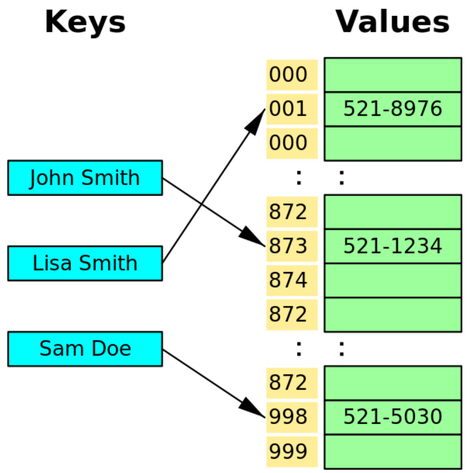
Indexing

Types of Indexes

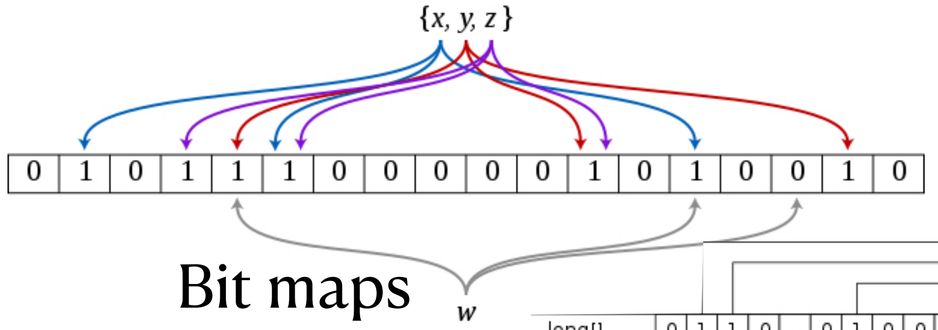
- B-Trees: Backbone of traditional indexing (used in almost all databases)
- Hash Indexes: Used for key-value stores
- LSM Trees: Used in write-heavy No-SQL systems
- Learned Indexes (?)

Trees



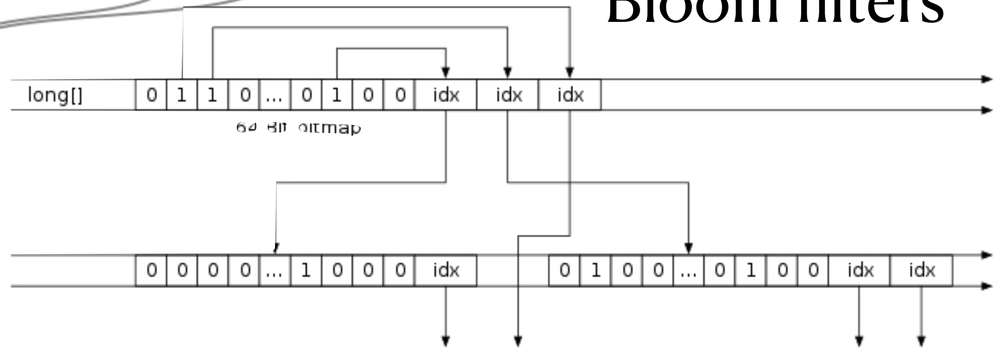


Hash Tables



Bit maps

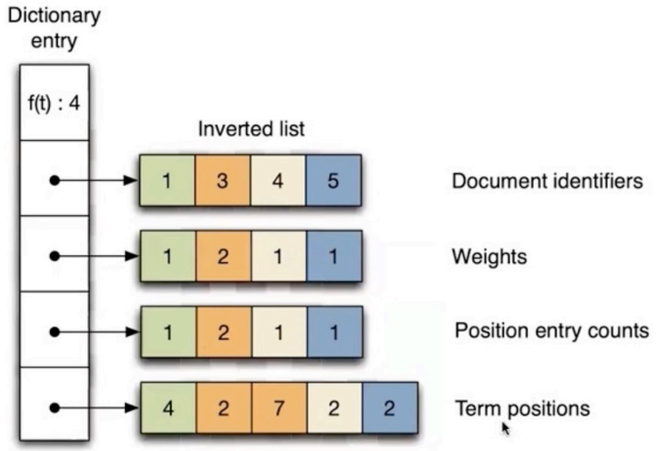
Bloom filters



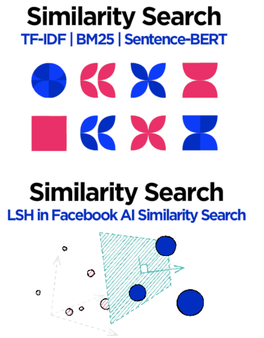
Indexing

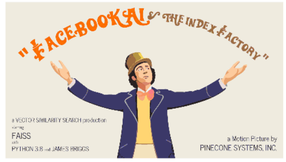
Beyond Databases

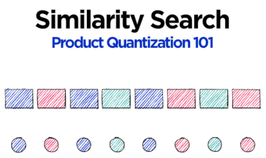
- Search engines — inverted index
- AI Retrieval — vector databases

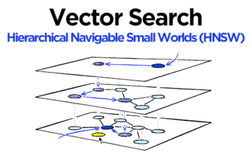


B. Cambazoglu and R. Baeza-Yates, "Scalability Challenges in Web Search Engines"









Query Optimization and Execution

Making declarative work imperatively

```
Select id, first_name, last_name
from actor left join film on ...
where exist (
...  )
```

?

```
HashTable<Integer, Record>
hashTable = new HashTable...
for(int i = 0; i < table1.numRows;
i++) {
hashTable.put(table.getField(i));
...
} for(int i;i < table2.numRows; i++)
{
...
...
}
```

You describe what you want

The database figures out the **best way** to execute your query

Query Optimization and Execution

Making declarative work imperatively

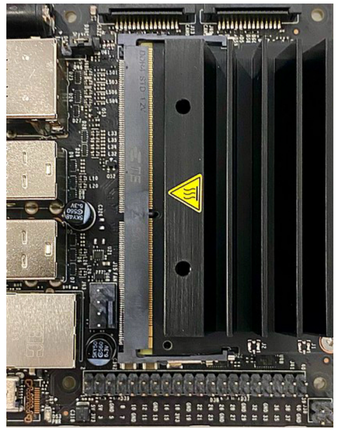
Key steps in query execution

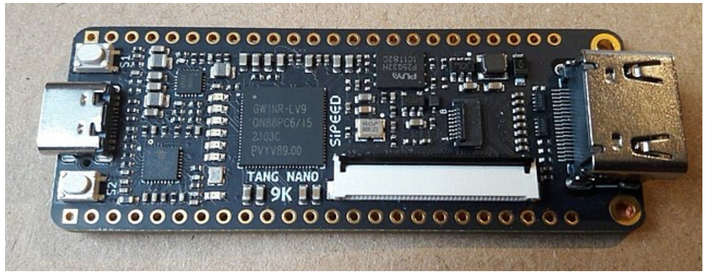
1. Parsing and validation
2. Query optimization
    a. Generate multiple possible execution plans
    b. Use cost-based optimization to chose the best one
3. Execution: run the plan using techniques like parallel execution and caching

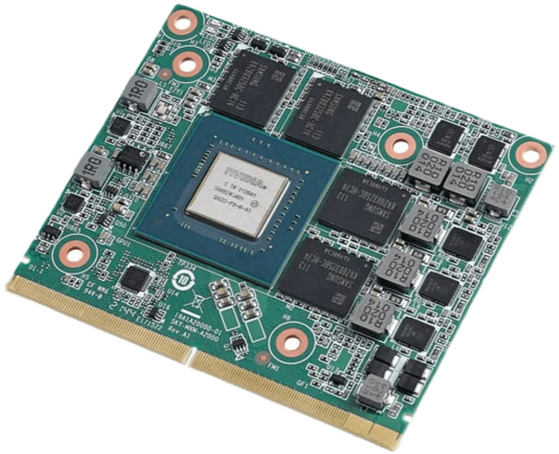
Query Optimization and Execution

Algorithms

1. Join
2. Scans
3. Sorting
4. ....







Aware of underlying compute and storage hardware

Why care about DB internals?

- Database internals are at the core of the biggest technological revolutions
- Databases are not just about SQL

- **AI & Machine Learning:** Data pipelines rely on indexing and query optimization.
- **Cloud Infrastructure:** Distributed databases like Spanner run on database internals.
- **Streaming & Real-time Analytics:** Kafka, Apache Flink use database-style event processing.
- **Graph Analytics:** Used in social networks, fraud detection, and knowledge graph

Every major tech company – Google, Amazon, Facebook – is, at its core, a database company