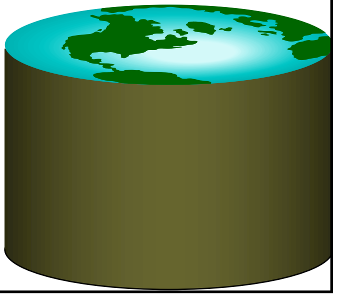
CSCD43: Database Systems Technology

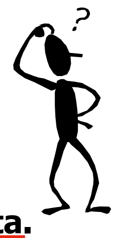
Winter 2020

Prof. Nick Koudas



1

What Is a Database *System*?



- **Database:**
  a very large, integrated collection of data.
- **Models a real-world *enterprise***
  - Entities (e.g., teams, games)
  - Relationships
    (e.g., The Raiders *are playing in* The Superbowl)
  - Also includes active components (e.g. "business logic")

- **A *Database Management System (DBMS)* is a software system designed to store, manage, and facilitate access to databases.**

2

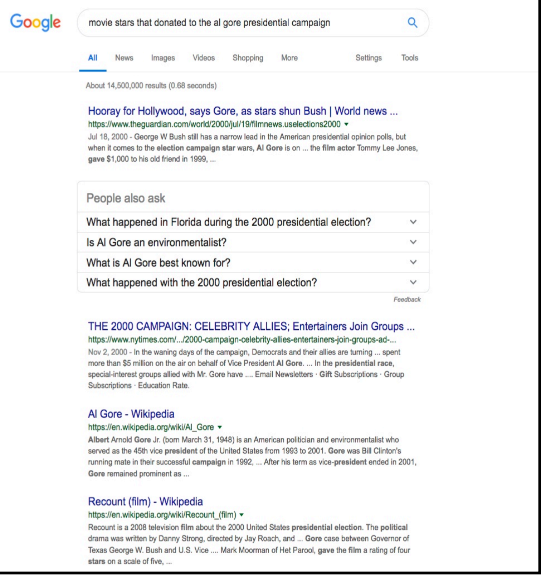
Is the WWW a DBMS?

- **Fairly sophisticated search available**
  - crawler *indexes* pages on the web
  - Keyword-based search for pages
- **But, currently**
  - data is mostly unstructured and untyped
  - search only:
    - can' t modify the data
    - can' t get summaries, complex combinations of data
  - few guarantees provided for freshness of data, consistency across data items (data quality), fault tolerance, …
  - Web sites (e.g. e-commerce) typically have a DBMS in the background to provide these functions.

3

"Search" vs. Query

- **What if you wanted to find out which movies stars donated to Al Gore' s presidential campaign?**
- **Try "movie stars that donated to al gore presidential campaign" in your favorite search engine.**



movie stars that donated to the al gore presidential campaign

About 14,500,000 results (0.68 seconds)

Hooray for Hollywood, says Gore, as stars shun Bush | World news ...
https://www.theguardian.com/world/2000/jul/19/filmnews.uselections2000
Jul 18, 2000 - George W Bush still has a narrow lead in the American presidential opinion polls, but when it comes to the **election campaign star** wars, **Al Gore** is on ... the **film actor** Tommy Lee Jones, **gave** $1,000 to his old friend in 1999, ...

People also ask
What happened in Florida during the 2000 presidential election?
Is Al Gore an environmentalist?
What is Al Gore best known for?
What happened with the 2000 presidential election?

THE 2000 CAMPAIGN: CELEBRITY ALLIES; Entertainers Join Groups ...
https://www.nytimes.com/.../2000-campaign-celebrity-allies-entertainers-join-groups-ad-...
Nov 2, 2000 - In the waning days of the campaign, Democrats and their allies are turning ... spent more than $5 million on the air on behalf of Vice President **Al Gore**. ... In the **presidential race**, special-interest groups allied with Mr. Gore have .... Email Newsletters · **Gift** Subscriptions · Group Subscriptions · Education Rate.

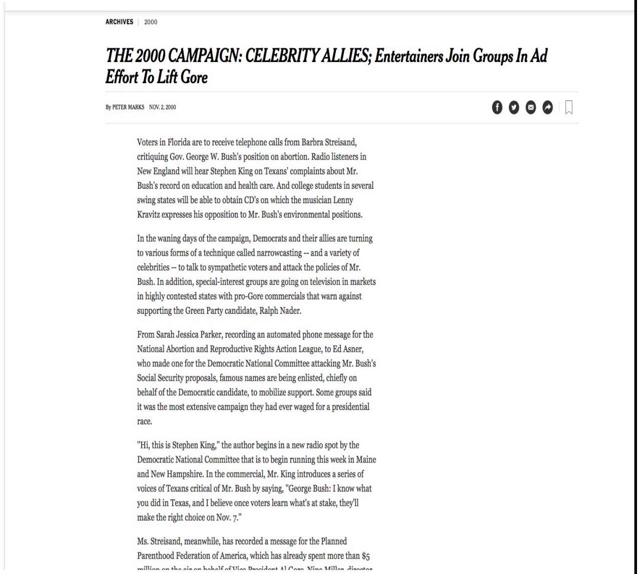
Al Gore - Wikipedia
https://en.wikipedia.org/wiki/Al_Gore
Albert Arnold **Gore** Jr. (born March 31, 1948) is an American politician and environmentalist who served as the 45th vice **president** of the United States from 1993 to 2001. **Gore** was Bill Clinton's running mate in their successful **campaign** in 1992, ... After his term as vice-**president** ended in 2001, **Gore** remained prominent as ...

Recount (film) - Wikipedia
https://en.wikipedia.org/wiki/Recount_(film)
Recount is a 2008 television **film** about the 2000 United States **presidential election**. The **political** drama was written by Danny Strong, directed by Jay Roach, and ... **Gore** case between Governor of Texas George W. Bush and U.S. Vice .... Mark Moorman of Het Parool, **gave** the **film** a rating of four **stars** on a scale of five, ...

4

"Search" vs. Query

- **"Search" can return only what's been "stored"**
- **E.g., best match Google, Bing top ten:**


ARCHIVES | 2000

THE 2000 CAMPAIGN: CELEBRITY ALLIES; Entertainers Join Groups In Ad Effort To Lift Gore

By PETER MARKS NOV. 2, 2000

Voters in Florida are to receive telephone calls from Barbra Streisand, critiquing Gov. George W. Bush's position on abortion. Radio listeners in New England will hear Stephen King on Texans' complaints about Mr. Bush's record on education and health care. And college students in several swing states will be able to obtain CD's on which the musician Lenny Kravitz expresses his opposition to Mr. Bush's environmental positions.

In the waning days of the campaign, Democrats and their allies are turning to various forms of a technique called narrowcasting -- and a variety of celebrities -- to talk to sympathetic voters and attack the policies of Mr. Bush. In addition, special-interest groups are going on television in markets in highly contested states with pro-Gore commercials that warn against supporting the Green Party candidate, Ralph Nader.

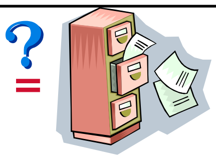
From Sarah Jessica Parker, recording an automated phone message for the National Abortion and Reproductive Rights Action League, to Ed Asner, who made one for the Democratic National Committee attacking Mr. Bush's Social Security proposals, famous names are being enlisted, chiefly on behalf of the Democratic candidate, to mobilize support. Some groups said it was the most extensive campaign they had ever waged for a presidential race.

"Hi, this is Stephen King," the author begins in a new radio spot by the Democratic National Committee that is to begin running this week in Maine and New Hampshire. In the commercial, Mr. King introduces a series of voices of Texans critical of Mr. Bush by saying, "George Bush: I know what you did in Texas, and I believe once voters learn what's at stake, they'll make the right choice on Nov. 7."

Ms. Streisand, meanwhile, has recorded a message for the Planned Parenthood Federation of America, which has already spent more than $5 million on the air on behalf of Vice President Al Gore. Nina Miller, director

5

Is a File System a DBMS?



- **Thought Experiment 1:**
  - You and your project partner are editing the same file.
  - You both save it at the same time.
  - Whose changes survive?

A) Yours B) Partner's C) Both D) Neither E) ???

- **Thought Experiment 2:**
  - You're updating a file.
  - The power goes out.
  - Which of your changes survive?

Q: How do you write programs over a subsystem when it promises you only "???" ?

A: Very, very carefully!!

A) All B) None C) All Since last save D) ???

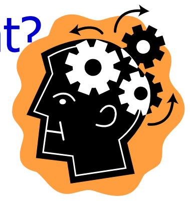
6

Why Study Databases??

- **The world is full of digital data, we better know**
- **how to manage them**
- **Need for DBMS has exploded in the last years**
  - Corporate: retail swipe/clickstreams, “customer relationship mgmt”, “supply chain mgmt”, “data warehouses”, etc.
  - Scientific: digital libraries, Human Genome project, NASA Mission to Planet Earth, physical sensors, grid physics network
- **DBMS encompasses much of CS in a practical discipline**
  - OS, languages, theory, AI, multimedia, logic
  - Yet traditional focus on real-world apps

7

What’s the intellectual content?



- **representing information**
  - data modeling
- **languages and systems for querying data**
  - complex queries with real *semantics**
  - over massive data sets
- **Query processing**
  - Generating query answers from declarative statements
  - Query optimization
  - Query execution
- ***concurrency control* for data manipulation**
  - controlling concurrent access
  - ensuring *transactional semantics*
- **reliable data storage**
  - maintain data semantics even if you pull the plug

* semantics: the meaning or relationship of meanings of a sign or set of signs

8

About the course: Workload

- **Projects with a “real world” focus:**
  - Modify the internals of a “real” open-source database system: PostgreSQL
    - Serious C system hacking
    - Measure the benefits of our changes
- **Exams – 1 Midterm & 1 Final**
- **Projects to be done in groups of up to 2**
  - Pick your partners ASAP

9

About the Course - Administrivia

- **http://www.cs.toronto.edu/~koudas**
  - Under teaching click on cscd43
- **Prof. Office Hours:**
  - Friday 11-12 in IC468 or by appointment
- **TAs: Yannis Xarchakos**
  - Office Hours: Tuesday after tutorial IC 400
  - Tutorials: Tuesday 9am-10am SW143
  - **Tutorials start Jan 21**

10

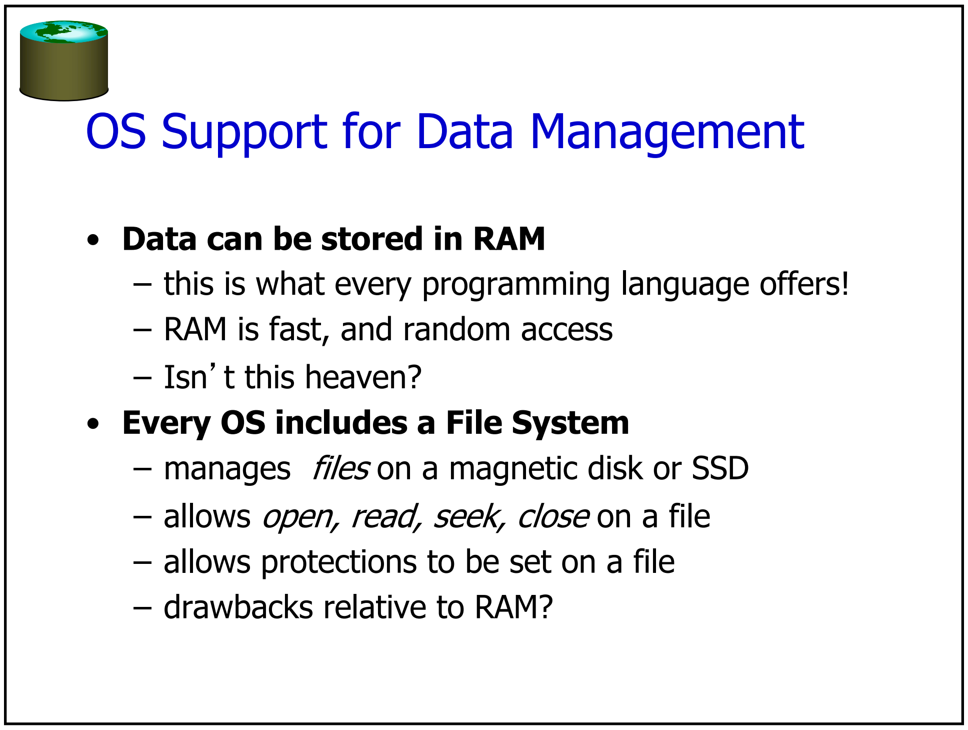
About the Course - Administrivia

- **Textbook**
  - Ramakrishnan et. al., Database Management Systems, 3rd edition
- **Cheating policy: zero tolerance**
- **Team Projects**
  - Three projects
  - Teams of up to 2
  - Submit; instructions on what exactly to submit will be provided with each project. Besides code, we will also be asking for performance evaluation for your algorithms compared with previous algorithms in the dbms.

11

- **A "free tasting" of things to come in this class:**
  - file systems & DBMSs
  - Query processing
  - Query Optimization
  - concurrent, fault-tolerant data management
  - DBMS architecture
- **Today's lecture is from Chapter 1 in R&G**

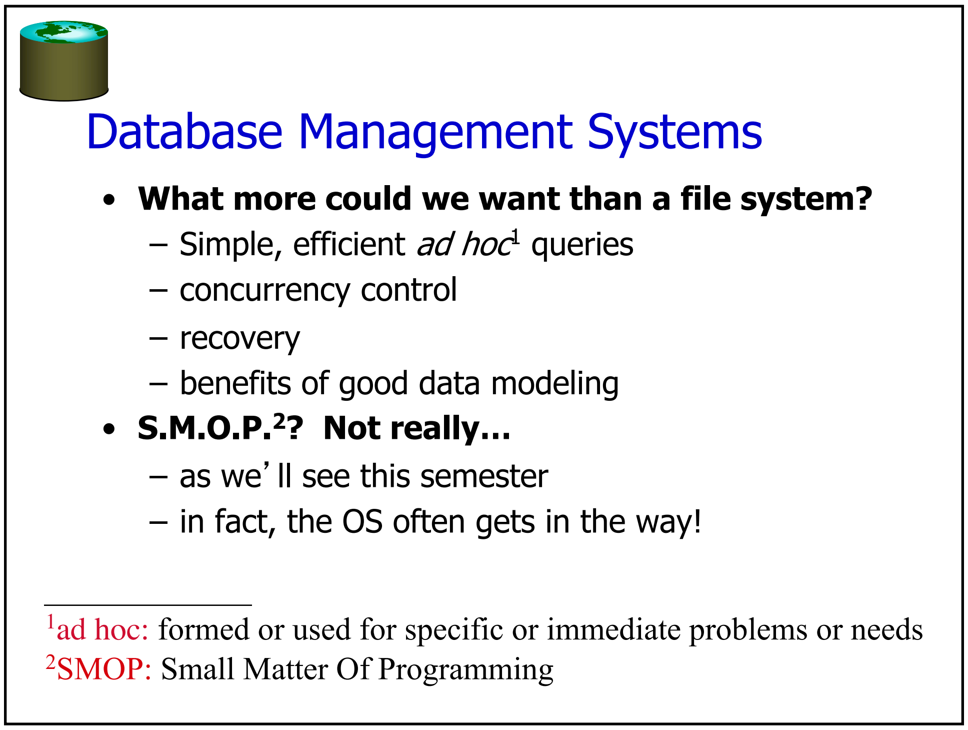
12



OS Support for Data Management

- **Data can be stored in RAM**
  - this is what every programming language offers!
  - RAM is fast, and random access
  - Isn't this heaven?
- **Every OS includes a File System**
  - manages *files* on a magnetic disk or SSD
  - allows *open, read, seek, close* on a file
  - allows protections to be set on a file
  - drawbacks relative to RAM?

13



Database Management Systems

- **What more could we want than a file system?**
  - Simple, efficient *ad hoc*[1] queries
  - concurrency control
  - recovery
  - benefits of good data modeling
- **S.M.O.P.[2]? Not really…**
  - as we'll see this semester
  - in fact, the OS often gets in the way!

[1] ad hoc: formed or used for specific or immediate problems or needs
[2] SMOP: Small Matter Of Programming

14

Describing Data: Data Models

- **A *data model* is a collection of concepts for describing data.**
- **A *schema* is a description of a particular collection of data, using a given data model.**
- **The *relational model of data* is the most widely used model today.**
  - Main concept: *relation*, basically a table with rows and columns.
  - Every relation has a *schema*, which describes the columns, or fields.

15

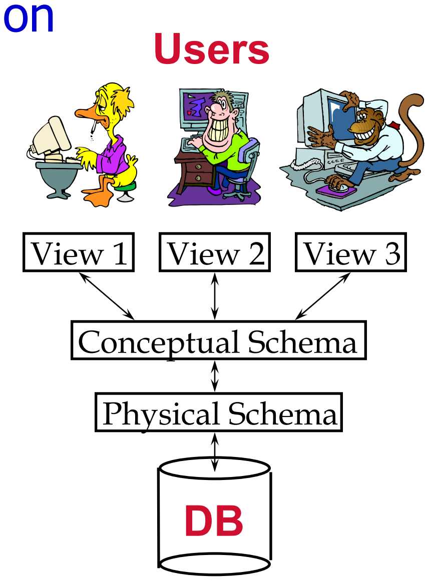
Example: University Database

- **Conceptual schema:**
  - *Students(sid: string, name: string, login: string, age: integer, gpa:real)*
  - *Courses(cid: string, cname:string, credits:integer)*
  - *Enrolled(sid:string, cid:string, grade:string)*
- **Physical schema:**
  - Relations stored as unordered files.
  - Index on first column of Students.
- **External Schema (View):**
  - *Course_info(cid:string,enrollment:integer)*

16

Levels of Abstraction

- Views describe how users see the data.
- Conceptual schema defines logical structure
- Physical schema describes the files and indexes used.
- (sometimes called the ANSI/SPARC model)



17

Data Independence

- **Applications insulated from how data is structured and stored.**
- **Logical data independence: Protection from changes in *logical* structure of data.**
- **Physical data independence: Protection from changes in *physical* structure of data.**
- **Q: Why is this particularly important for DBMS?**

18

Concurrency Control

- **Concurrent execution of user programs: key to good DBMS performance.**
  - Disk accesses frequent, pretty slow
  - Keep the CPU working on several programs concurrently.
- **Interleaving actions of different programs: trouble!**
  - e.g., account-transfer & print statement at same time
- **DBMS ensures such problems don't arise.**
  - Users/programmers can pretend they are using a single-user system. (called "Isolation")
  - Thank goodness! Don't have to program "very, very carefully".

19

Transaction: An Execution of a DB Program

- **Key concept is a transaction: an atomic sequence of database actions (reads/writes).**
- **Each transaction, executed completely, must take the DB between consistent states.**
- **Users can specify simple integrity constraints on the data. The DBMS enforces these.**
  - Beyond this, the DBMS does not understand the semantics of the data.
  - Ensuring that a single transaction (run alone) preserves consistency is ultimately the user's responsibility!

20

Scheduling Concurrent Transactions

- **DBMS ensures that execution of {T1, ... , Tn} is equivalent to some *serial* execution T1' ... Tn'.**
  - Before reading/writing an object, a transaction requests a lock on the object, and waits till the DBMS gives it the lock. All locks are held until the end of the transaction. (Strict 2PL locking protocol.)
  - Idea: If an action of Ti (say, writing X) affects Tj (which perhaps reads X),
    … say Ti obtains the lock on X first
    … so Tj is forced to wait until Ti completes.
    This effectively orders the transactions.
  - What if
    … Tj already has a lock on Y
    … and Ti later requests a lock on Y?
    (Deadlock!) Ti or Tj is aborted and restarted!

21

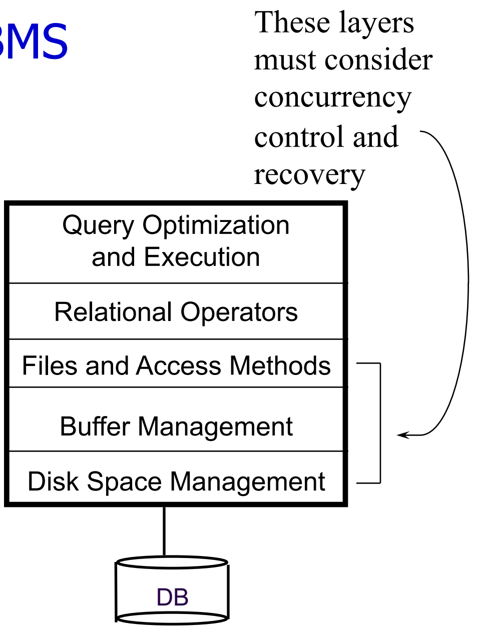
Ensuring Transaction Properites

- **DBMS ensures *atomicity* (all-or-nothing property) even if system crashes in the middle of a Xact.**
- **DBMS ensures *durability* of *committed* Xacts even if system crashes.**
- **Idea: Keep a *log* (history) of all actions carried out by the DBMS while executing a set of Xacts:**
  - Before a change is made to the database, the corresponding log entry is forced to a safe location. (*WAL protocol*; OS support for this is often inadequate.)
  - After a crash, the effects of partially executed transactions are *undone* using the log. Effects of committed transactions are *redone* using the log.
  - trickier than it sounds!

22

Structure of a DBMS

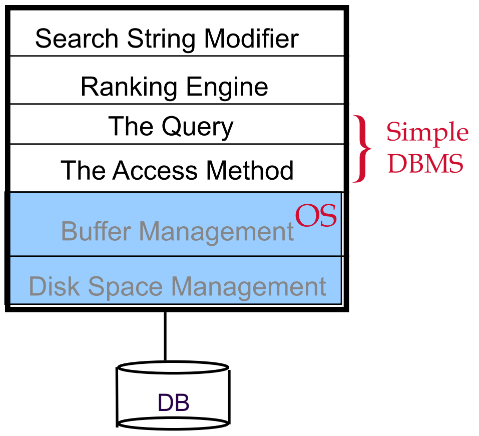
- **A typical DBMS has a layered architecture.**
- **The figure does not show the concurrency control and recovery components.**
- **Each system has its own variations.**
- **You will see the "real deal" in PostgreSQL.**
  - It's a pretty full-featured example


These layers must consider concurrency control and recovery
Query Optimization and Execution
Relational Operators
Files and Access Methods
Buffer Management
Disk Space Management
DB

24

FYI: A text search engine

- **Less "system" than DBMS**
  - Uses OS files for storage
  - Just one access method
  - One hardwired query
    - regardless of search string
- **Typically no concurrency or recovery management**
  - Read-mostly
  - Batch-loaded, periodically
  - No updates to recover
  - OS a reasonable choice
- **Smarts: text tricks**
  - Search string modifier (e.g. "stemming" and synonyms)
  - Ranking Engine (sorting the output, e.g. by word or document popularity)
  - no semantics: WYGIWIGY



There may be time to talk about some of these text tricks in this class, but it won't be a focus.

25

Advantages of a DBMS

- **Data independence**
- **Efficient data access**
- **Data integrity & security**
- **Data administration**
- **Concurrent access, crash recovery**
- **Reduced application development time**
- **So why not use them always?**
  - Expensive/complicated to set up & maintain
  - This cost & complexity must be offset by need
  - General-purpose, not suited for special-purpose tasks (e.g. text search!)

26

Databases make these folks happy ...

- **DBMS vendors, programmers**
  - Oracle, IBM, MS, Spark, main memory DB vendors
- **End users in many fields**
  - Business, education, science, …
- **DB application programmers**
  - Build enterprise applications on top of DBMSs
  - Build web services that run off DBMSs
- **Database administrators (DBAs)**
  - Design logical/physical schemas
  - Handle security and authorization
  - Data availability, crash recovery
  - Database tuning as needs evolve

…must understand how a DBMS works

27

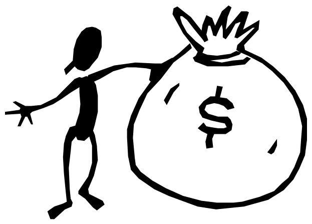
Summary (part 1)

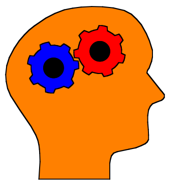
- **DBMS used to maintain, query large datasets.**
  - can manipulate data and exploit *semantics*
- **Other benefits include:**
  - recovery from system crashes,
  - concurrent access,
  - quick application development,
  - data integrity and security.
- **Levels of abstraction provide data independence**
- **In this course we will explore:**
  1) How to be a sophisticated user of DBMS technology
  2) What goes on inside the DBMS

28

Summary, cont.

- DBAs, DB developers the bedrock of the information economy





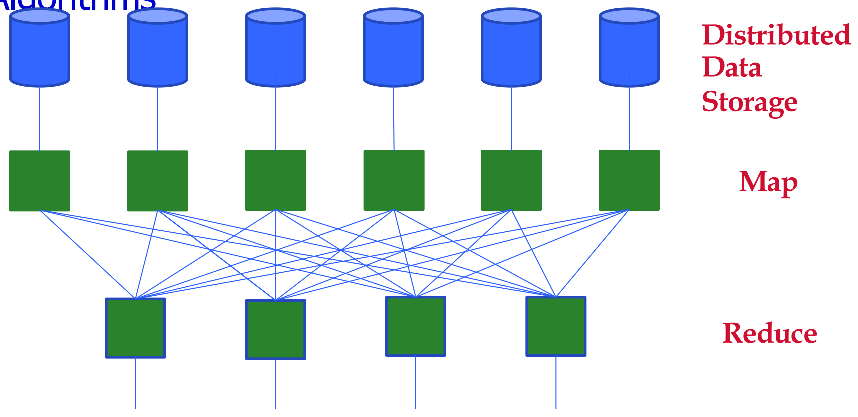
- DBMS R&D represents a broad, fundamental branch of the science of computation

29

1. THE MAPREDUCE ABSTRACTION

30

The Map Reduce Abstraction for Distributed Algorithms



31

2. SPARK

32

Intro to Spark

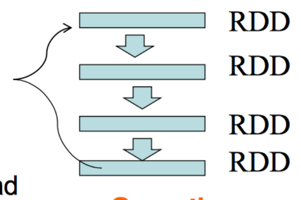
- Spark is really a different implementation of the MapReduce programming model
- What makes Spark different is that it operates on Main Memory
- Spark: we write programs in terms of operations on resilient distributed datasets (RDDs).
- RDD (simple view): a collection of elements partitioned across the nodes of a cluster that can be operated on in parallel.
- RDD (complex view): RDD is an interface for data transformation, RDD refers to the data stored either in persisted store (HDFS) or in cache (memory, memory+disk, disk only) or in another RDD

33

RDDs in Spark

RDD: Resilient Distributed Datasets

- **Like a big list:**
  - Collections of objects spread across a cluster, stored in RAM or on Disk
- **Built through parallel transformations**
- **Automatically rebuilt on failure**



Operations

- **Transformations (e.g. map, filter, groupBy)**
- **Make sure input/output match**

34