PHYSICAL DATABASE DESIGN AND PERFORMANCE

MSBA 7024 / MACC 7020
Database Design and Management

OBJECTIVES

Definition of terms

Describe the physical database design process

Choose storage formats for attributes

Know when and how to use denormalization

Select appropriate file organizations

Describe three types of file organization

Describe indexes and their appropriate use

Translate a database model into efficient structures

PHYSICAL DATABASE DESIGN

Purpose–translate the logical description of data into the *technical specifications* for storing and retrieving data

Goal–create a design for storing data that will provide *adequate performance* and insure *database integrity*, *security*, and *recoverability*

PHYSICAL DESIGN PROCESS

Inputs

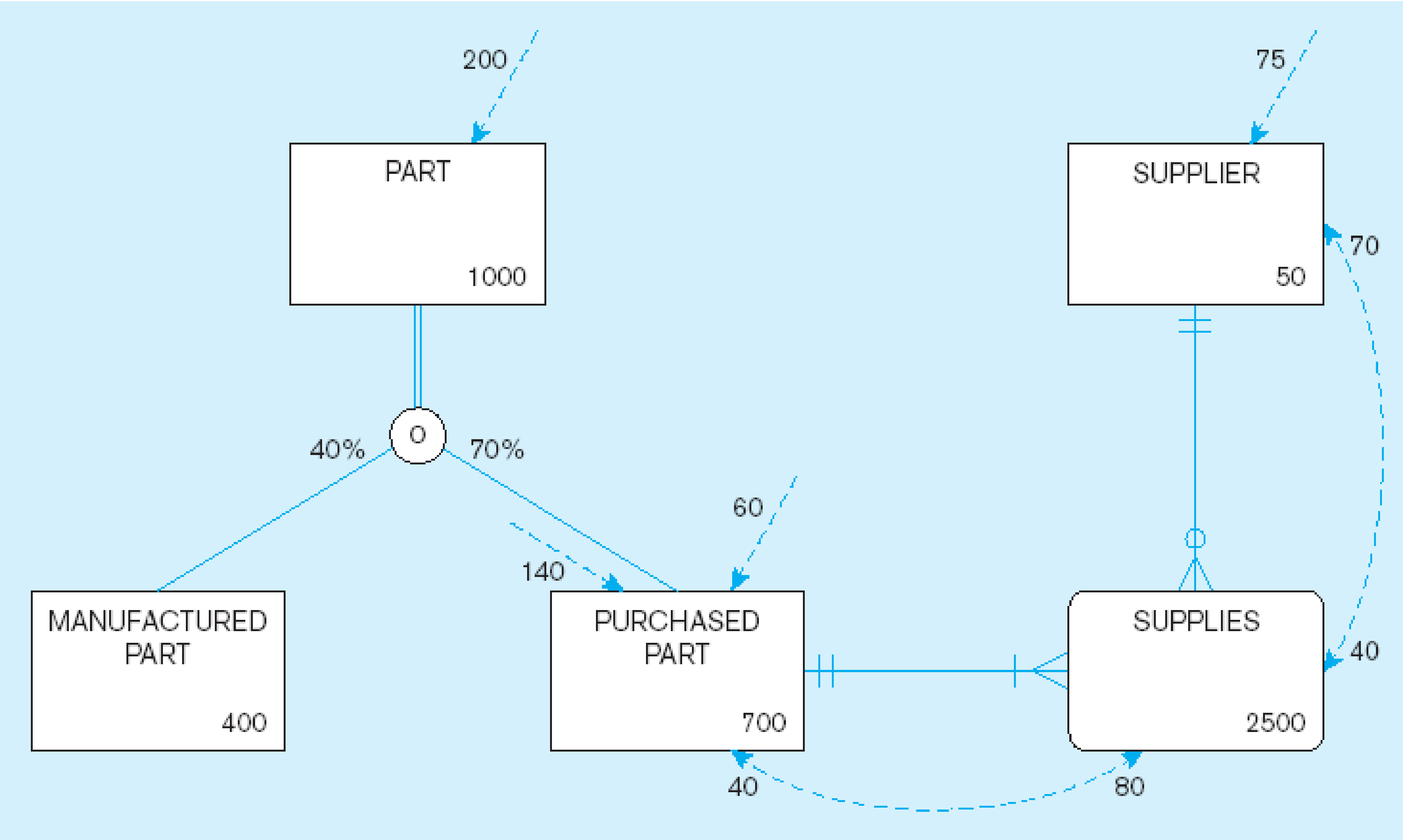
- Normalized relations
- Volume estimates
- Attribute definitions
- Response time expectations
- Data security needs
- Backup/recovery needs
- Integrity expectations
- DBMS technology used

Leads to

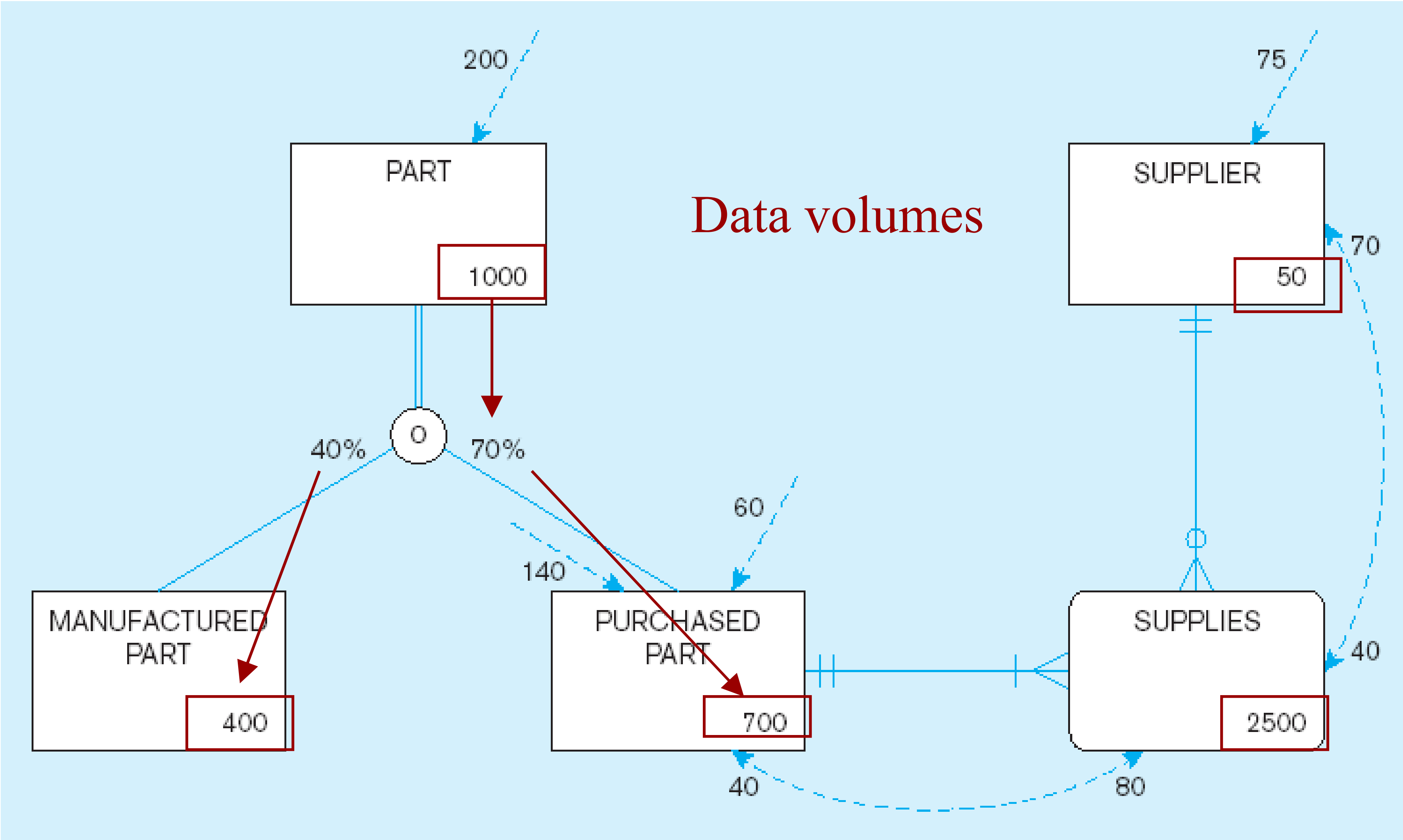
Decisions

- Attribute data types
- Physical record descriptions (doesn't always match logical design)
- File organizations
- Indexes and database architectures
- Query optimization

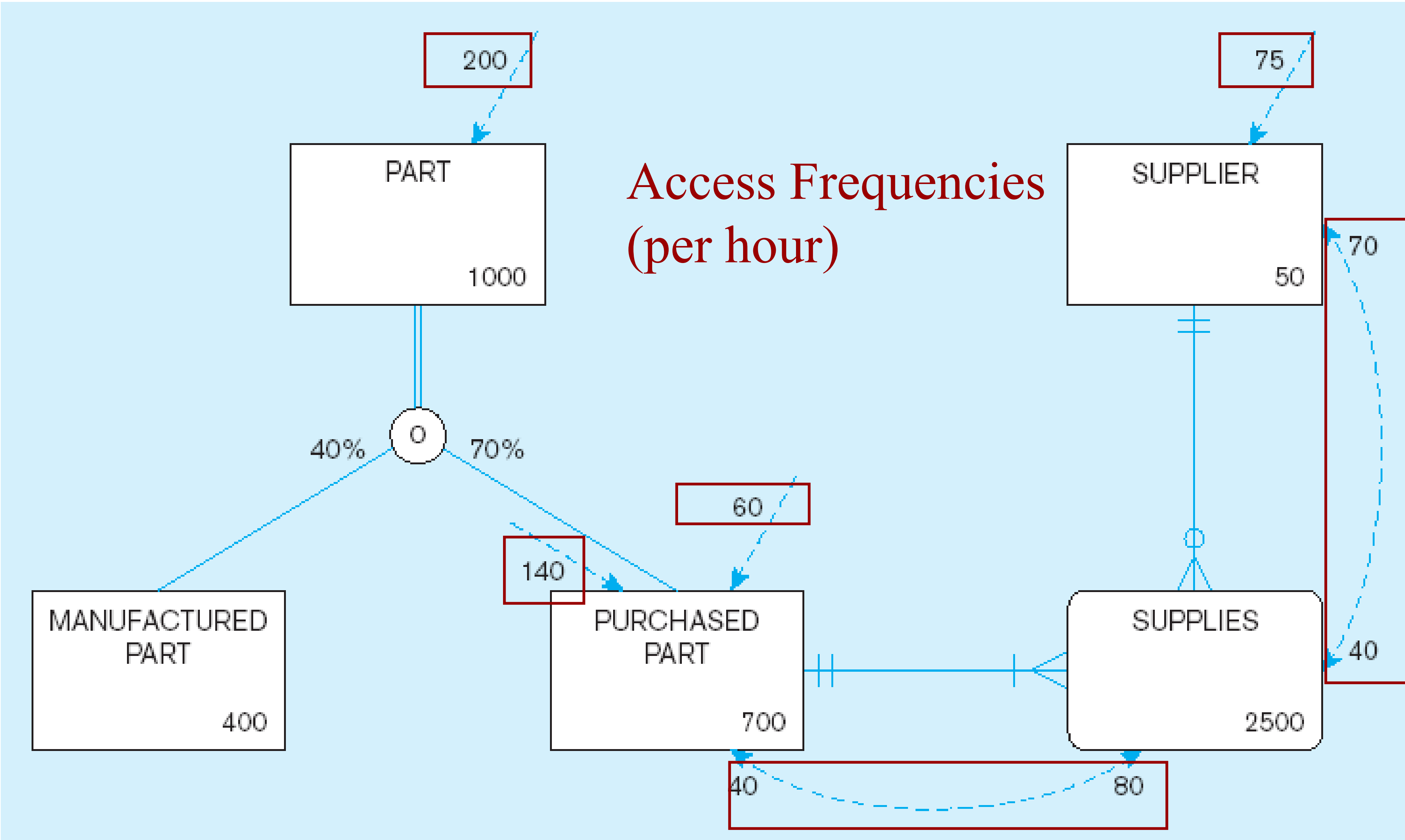
Composite usage map (Pine Valley Furniture Company)



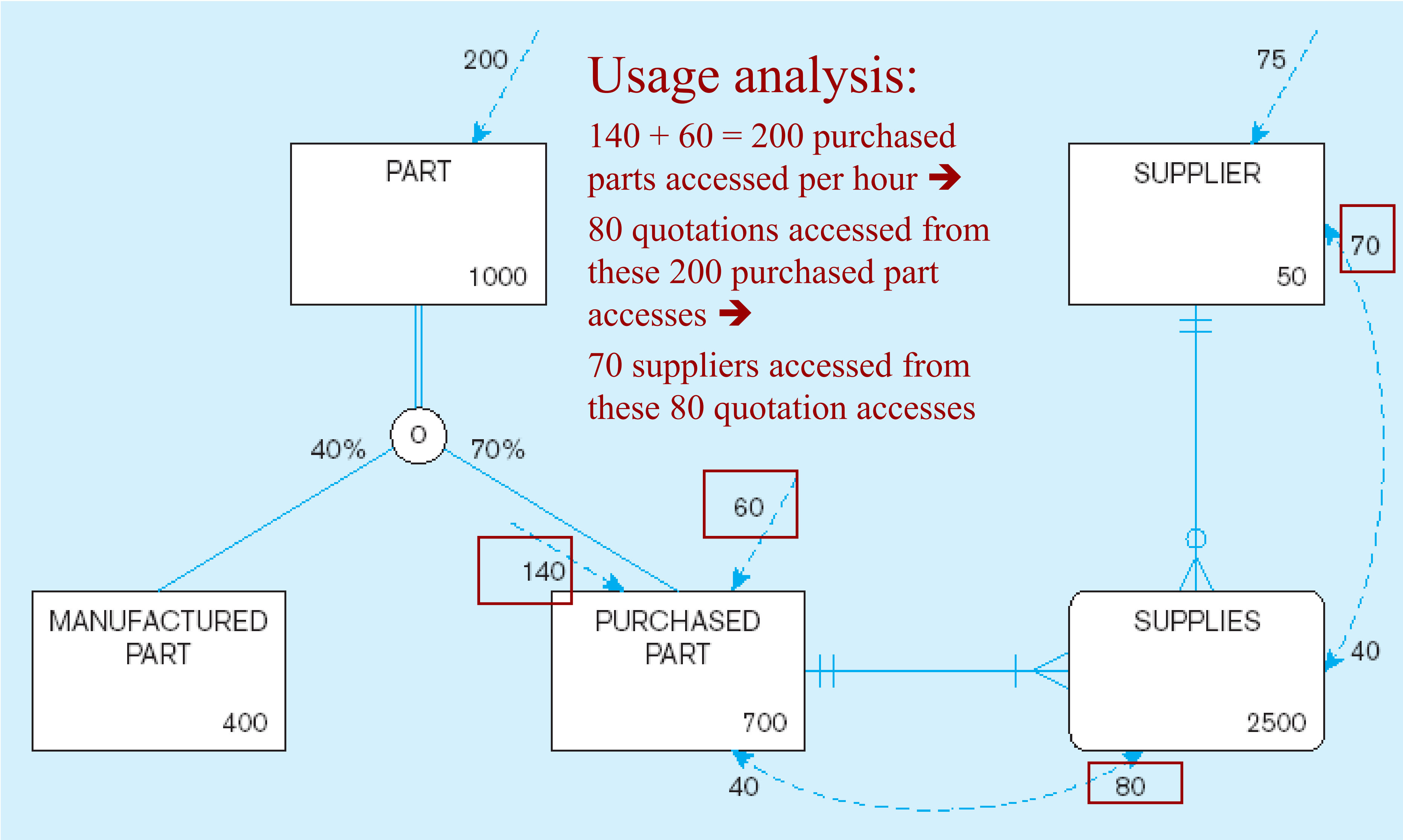
Composite usage map (Pine Valley Furniture Company) (cont.)



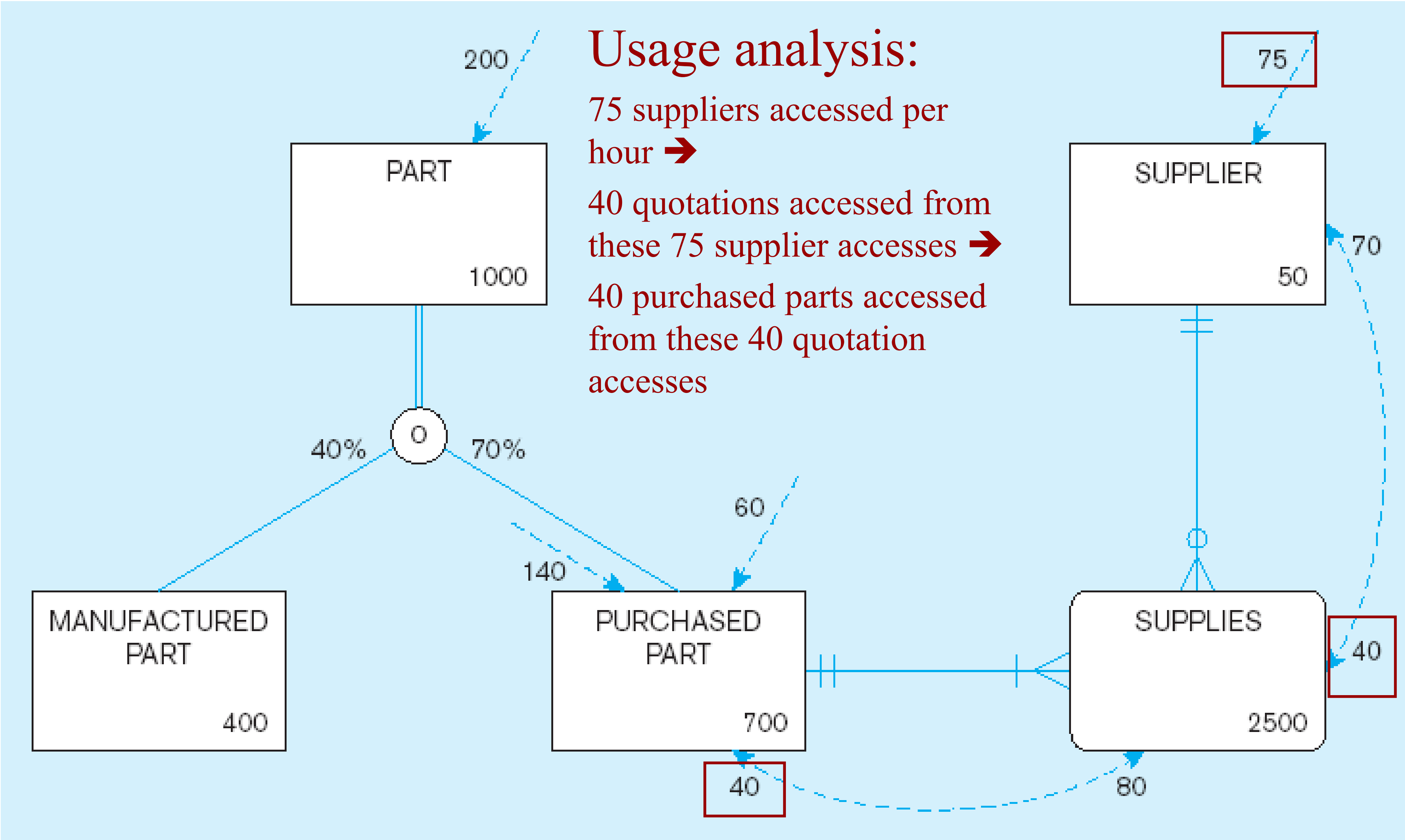
Composite usage map
(Pine Valley Furniture Company) (cont.)



Composite usage map
(Pine Valley Furniture Company) (cont.)



Composite usage map (Pine Valley Furniture Company) (cont.)



Usage analysis:

75 suppliers accessed per hour ➔

40 quotations accessed from these 75 supplier accesses ➔

40 purchased parts accessed from these 40 quotation accesses

DESIGNING FIELDS

Field: smallest unit of data in database

Field design

- Choosing data type
- Coding, compression, encryption
- Controlling data integrity

CHOOSING DATA TYPES

CHAR – fixed-length character string (e.g. CHAR(5))

VARCHAR2 – variable-length character string (e.g. VARCHAR2(30))

NUMBER – positive/negative number

(e.g., NUMBER(5), NUMBER(5,2))

INTEGER – positive/negative whole number

DATE – actual date (and time)

BLOB – binary large object (good for graphics, sound clips, etc.)

Example code look-up table
(Pine Valley Furniture Company)

PRODUCT File

Product_No	Description	Finish	...
B100	Chair	C	
B120	Desk	A	
M128	Table	C	
T100	Bookcase	B	
...	...	...	

FINISH Look-up Table

Code	Value
A	Birch
B	Maple
C	Oak

Code saves space, but costs an additional lookup to obtain actual value

FIELD DATA INTEGRITY

Default value–assumed value if no explicit value

Range control–allowable value limitations (constraints or validation rules)

Null value control–allowing or prohibiting empty fields

Referential integrity–range control (and null value allowances) for foreign-key to primary-key match-ups

Sarbanes-Oxley Act (SOX) legislates importance of financial data integrity

DENORMALIZATION

Transforming ***normalized*** relations into ***unnormalized*** physical record specifications

Benefits:

- Can improve performance (speed) by reducing number of table lookups (i.e. *reduce number of necessary join queries*)
- Many data mining/analytics tools require one single table as input

Costs (due to data duplication)

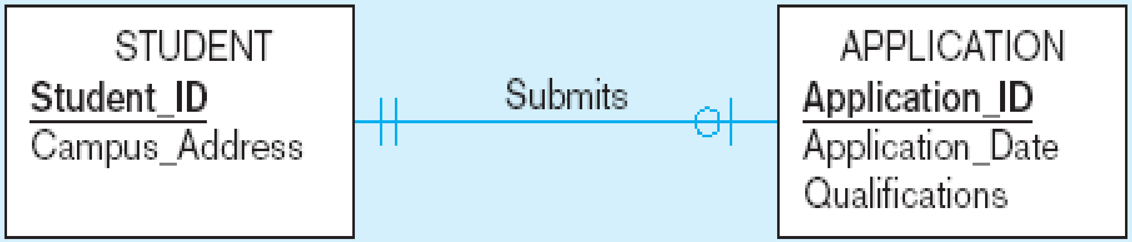
- Wasted storage space
- Data integrity/consistency threats

DENORMALIZATION

Common denormalization opportunities

- One-to-one relationship
- Many-to-many relationship with attributes
- Reference data (1:N relationship where 1-side has data not used in any other relationship)

A possible denormalization situation: two entities with one-to-one relationship



Normalized relations:

STUDENT

Student_ID	Campus_Address

APPLICATION

Application_ID	Application_Date	Qualifications	Student_ID

Denormalized relation:

STUDENT

Student_ID	Campus_Address	Application_Date	Qualifications

and Application_Date and Qualifications may be null

A possible denormalization situation: a many-to-many relationship with nonkey attributes

Normalized relations:

VENDOR

Vendor_ID	Address	Contact_Name

ITEM

Item_ID	Description

PRICE QUOTE

Vendor_ID	Item_ID	Price

Extra table access required

Denormalized relations:

VENDOR

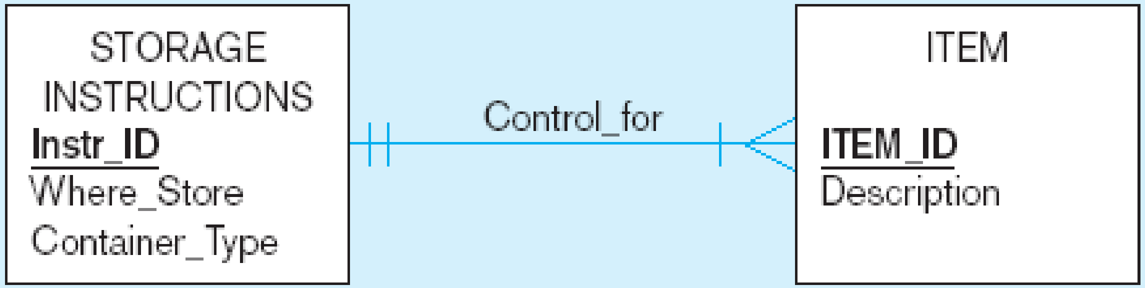
Vendor_ID	Address	Contact_Name

ITEM QUOTE

Vendor_ID	Item_ID	Description	Price

Data duplication

A possible denormalization situation: reference data



Normalized relations:

STORAGE

Instr_ID	Where_Store	Container_Type

ITEM

Item_ID	Description	Instr_ID

Extra table access required

Denormalized relation:

ITEM

Item_ID	Description	Where_Store	Container_Type

Data duplication

PARTITIONING

Horizontal Partitioning: Distributing the rows of a table into several separate files

- Useful for situations where different users need access to different rows
- Three types: Key Range Partitioning, Hash Partitioning, or Composite Partitioning

Vertical Partitioning: Distributing the columns of a table into several separate relations

- Useful for situations where different users need access to different columns
- The primary key must be repeated in each file

Combinations of Horizontal and Vertical

Partitions often correspond with User Schemas (user views)

PARTITIONING (CONT.)

Advantages of Partitioning:

- Efficiency: Records used together are grouped together
- Local optimization: Each partition can be optimized for performance
- Security, recovery
- Load balancing: Partitions stored on different disks, reduces contention
- Take advantage of parallel processing capability

Disadvantages of Partitioning:

- Inconsistent access speed: Slow retrievals across partitions
- Complexity: Non-transparent partitioning
- Extra space or update time: Duplicate data; access from multiple partitions

DATA REPLICATION

Purposely storing the same data in multiple locations of the database

Improves performance by allowing multiple users to access the same data at the same time with minimum contention

Sacrifices data integrity due to data duplication

Best for data that is not updated often

DESIGNING PHYSICAL FILES

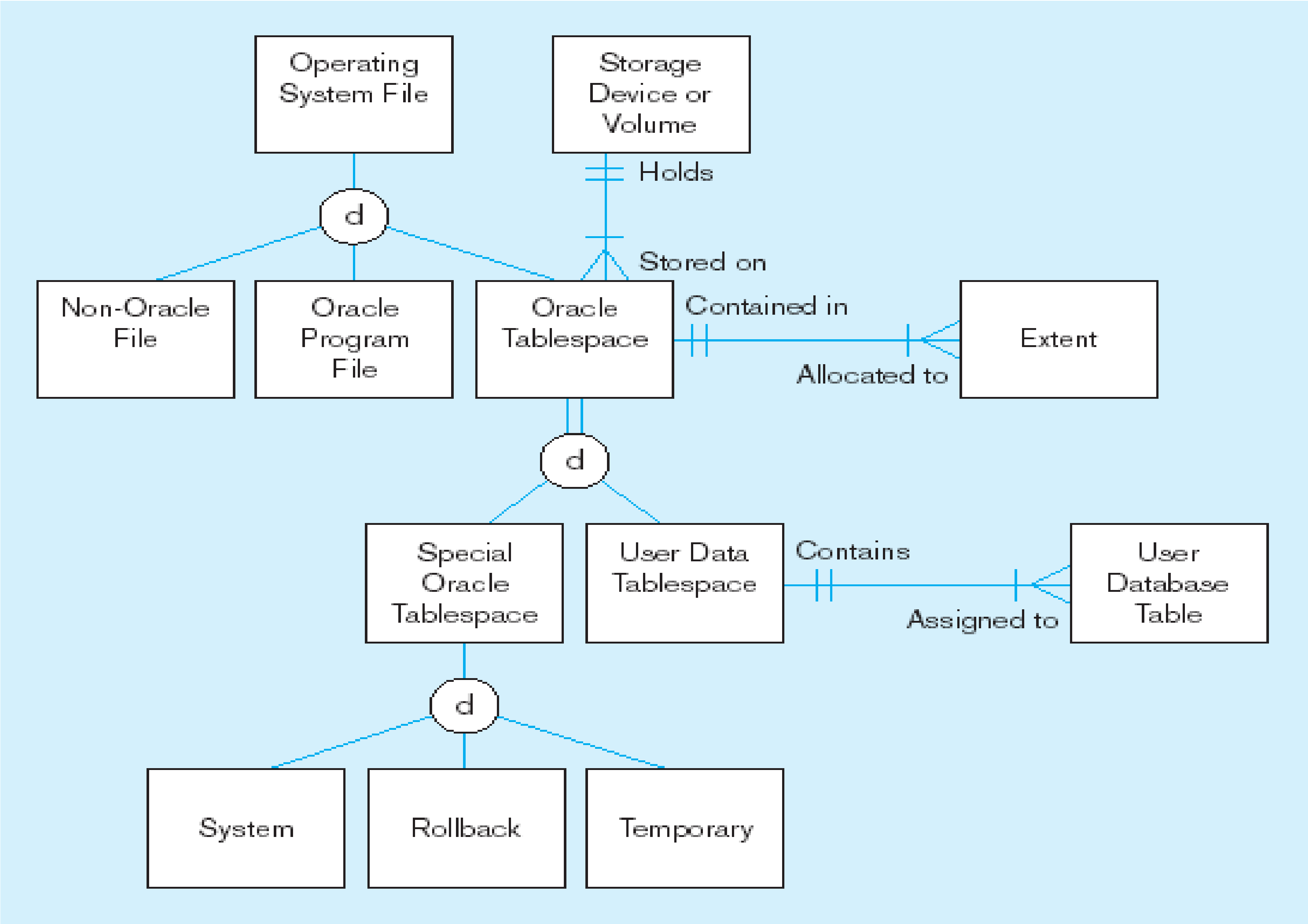
Physical File:

- A named portion of secondary memory allocated for the purpose of storing physical records
- Tablespace–named set of disk storage elements in which physical files for database tables can be stored
- Extent–contiguous section of disk space

Constructs to link two pieces of data:

- Sequential storage
- Pointers–field of data that can be used to locate related fields or records

Physical file terminology in an Oracle environment



PHYSICAL RECORDS

Physical Record: A group of fields stored in adjacent memory locations and retrieved together as a unit

Page: The amount of data read or written in one I/O operation

Blocking Factor: The number of physical records per page

FILE ORGANIZATIONS

Technique for physically arranging records of a file on secondary storage

Factors for selecting file organization:

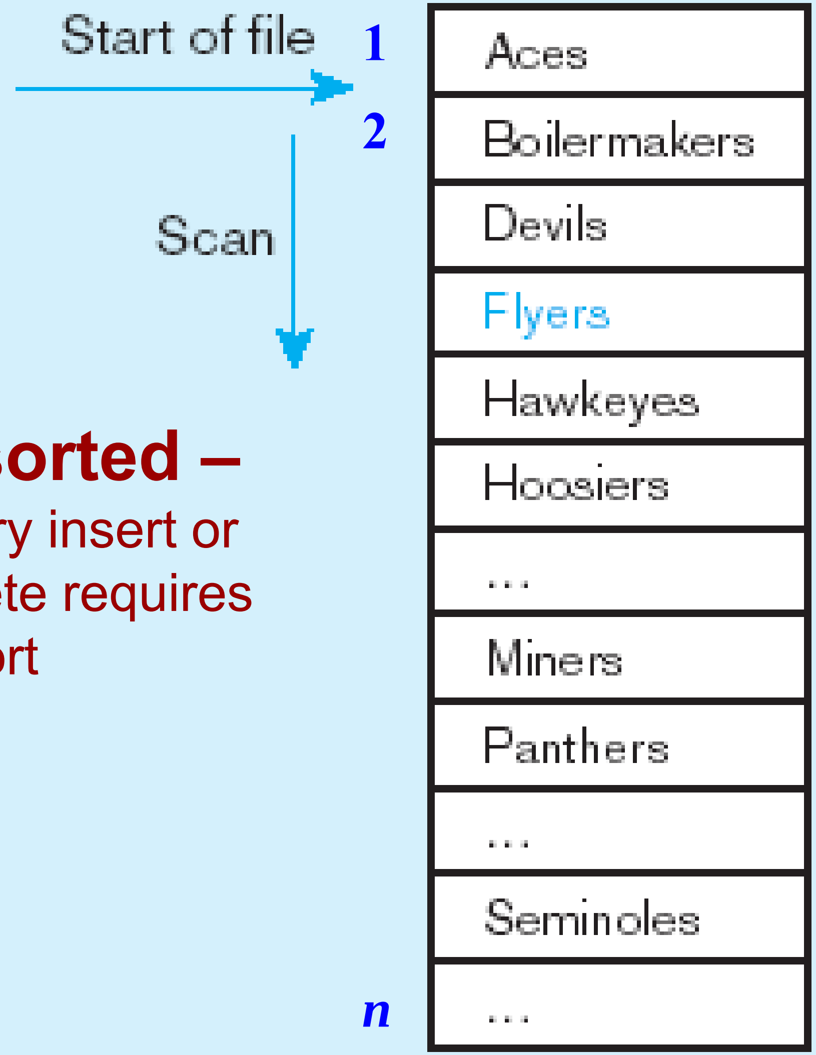
- Fast data retrieval and throughput
- Efficient storage space utilization
- Protection from failure and data loss
- Minimizing need for reorganization
- Accommodating growth
- Security from unauthorized use

Types of file organizations

- Sequential
- Indexed
- Hashed

Sequential file organization

Records of the file are stored in sequence by the primary key field values



If sorted – every insert or delete requires resort

INDEXED FILE ORGANIZATIONS

Index – a separate table that contains organization of records for quick retrieval

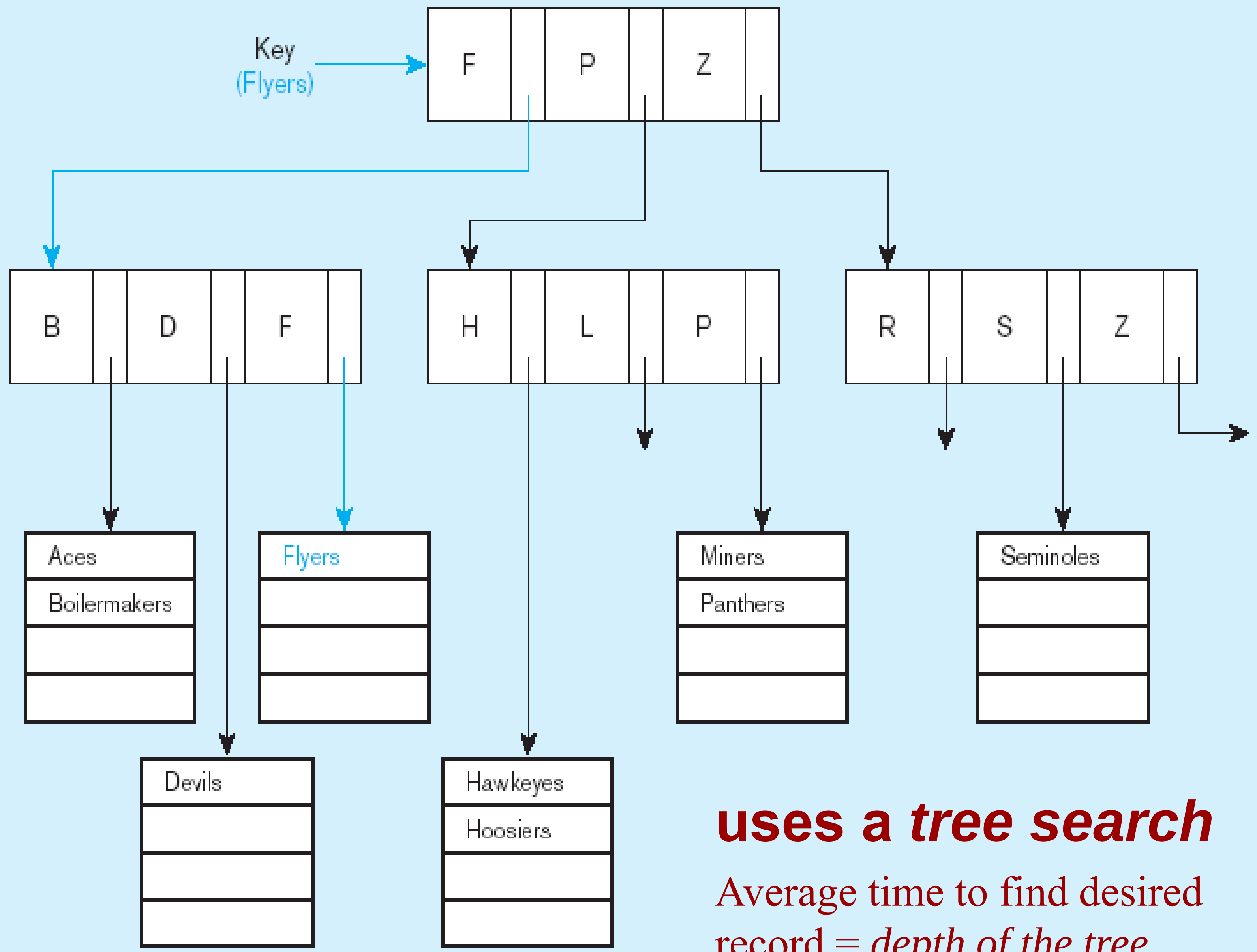
Primary keys are automatically indexed

Oracle/MySQL has a CREATE INDEX operation, and MS ACCESS allows indexes to be created for most field types

Indexing approaches:

- B-tree index
- Bitmap index
- Hash Index
- Join Index

B-tree index

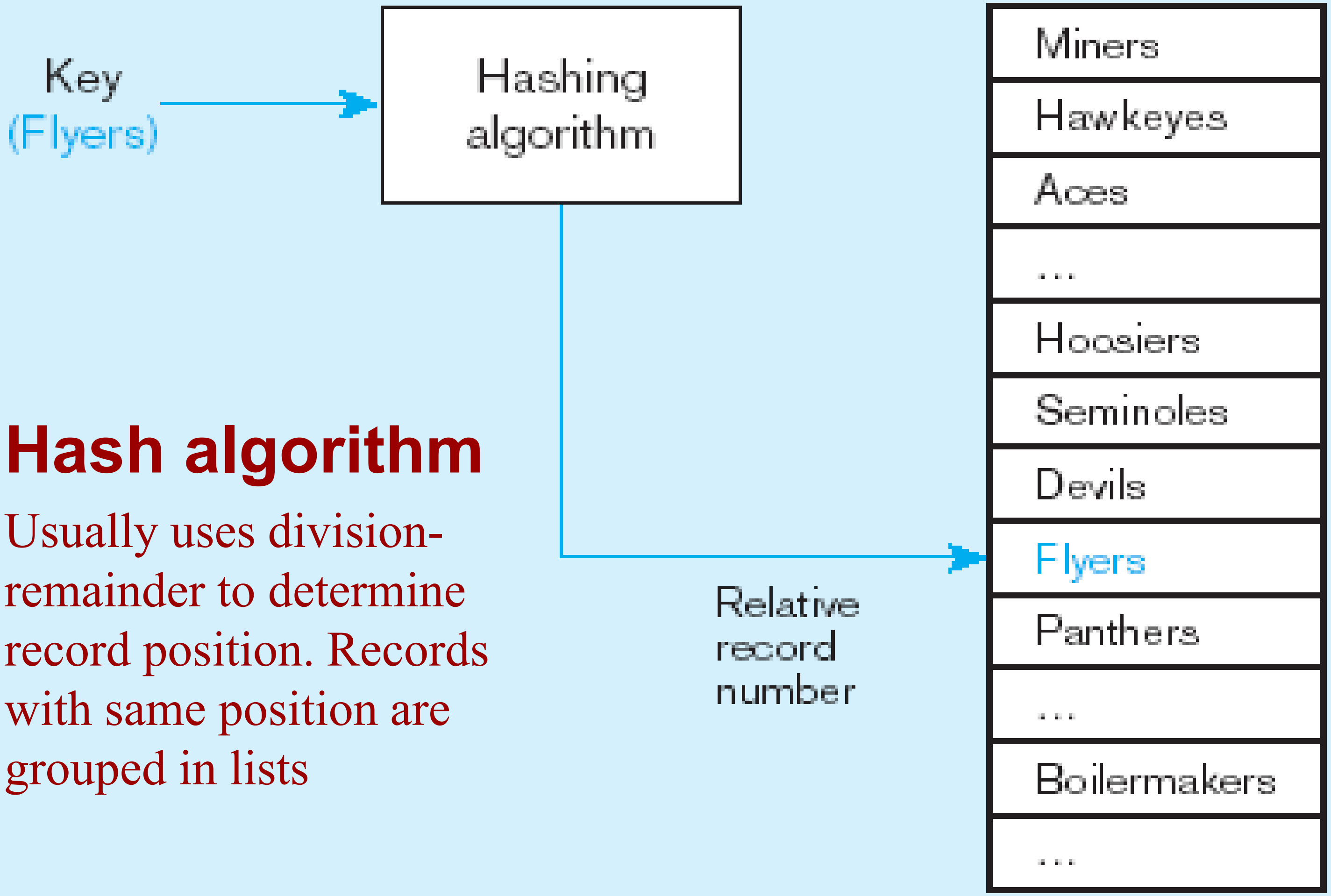


Leaves of the tree are all at same level → consistent access time

uses a *tree search*

Average time to find desired record = *depth of the tree*

Hashed file or index organization



Hash algorithm

Usually uses division-remainder to determine record position. Records with same position are grouped in lists

Bitmap index organization

Bitmap saves on space requirements

Rows - possible values of the attribute

Columns - table rows

Bit indicates whether the attribute of a row has the values

Price	Product Table Row Numbers									
	1	2	3	4	5	6	7	8	9	10
100	0	0	1	0	1	0	0	0	0	0
200	1	0	0	0	0	0	0	0	0	0
300	0	1	0	0	0	0	1	0	0	1
400	0	0	0	1	0	1	0	1	1	0

Products 3 and 5 have Price $100
Product 1 has Price $200
Products 2, 7, and 10 have Price $300
Products 4, 6, 8, and 9 have Price $400

Join Indexes–speeds up join operations

Customer

RowID	Cust#	CustName	City	State
10001	C2027	Hadley	Dayton	Ohio
10002	C1026	Baines	Columbus	Ohio
10003	C0042	Ruskin	Columbus	Ohio
10004	C3861	Davies	Toledo	Ohio
. . .				

Store

RowID	Store#	City	Size	Manager
20001	S4266	Dayton	K2	E2166
20002	S2654	Columbus	K3	E0245
20003	S3789	Dayton	K4	E3330
20004	S1941	Toledo	K1	E0874
. . .				

Join Index

CustRowID	StoreRowID	Common Value*
10001	20001	Dayton
10001	20003	Dayton
10002	20002	Columbus
10003	20002	Columbus
10004	20004	Toledo
. . .		

Order

RowID	Order#	Order Date	Cust#(FK)
30001	O5532	10/01/2001	C3861
30002	O3478	10/01/2001	C1062
30003	O8734	10/02/2001	C1062
30004	O9845	10/02/2001	C2027
. . .			

Customer

RowID	Cust#(PK)	CustName	City	State
10001	C2027	Hadley	Dayton	Ohio
10002	C1026	Baines	Columbus	Ohio
10003	C0042	Ruskin	Columbus	Ohio
10004	C3861	Davies	Toledo	Ohio
. . .				

Join Index

CustRowID	OrderRowID	Cust#
10001	30004	C2027
10002	30002	C1062
10002	30003	C1062
10004	30001	C3861
. . .		

	File Organization		
Factor	**Sequential**	**Indexed**	**Hashed**
Storage Space	No wasted space	No wasted space for data, but extra space for index	Extra space may be needed to allow for addition and deletion of records after initial set of records is loaded
Sequential Retrieval on Primary Key	Very fast	Moderately fast	Impractical, unless use hash index
Random Retrieval on Primary Key	Impractical	Moderately fast	Very fast
Multiple Key Retrieval	Possible, but requires scanning whole file	Very fast with multiple indexes	Not possible, unless use hash index
Deleting Records	Can create wasted space or require reorganizing	If space can be dynamically allocated, this is easy, but requires maintenance of indexes	Very easy
Adding New Records	Requires rewriting file	If space can be dynamically allocated, this is easy, but requires maintenance of indexes	Very easy, except multiple keys with same address require extra work
Updating Records	Usually requires rewriting file	Easy, but requires maintenance of indexes	Very easy

RULES FOR USING INDEXES

1. Use on larger tables
2. Index the primary key of each table
3. Index search fields (fields frequently in WHERE clause)
4. Fields in SQL ORDER BY and GROUP BY commands
5. When there are >100 values but not when there are <30 values

RULES FOR USING INDEXES (CONT.)

6. Avoid use of indexes for fields with long values; perhaps compress values first
7. DBMS may have limit on number of indexes per table and number of bytes per indexed field(s)
8. Null values will not be referenced from an index
9. Use indexes heavily for non-volatile databases; limit the use of indexes for volatile databases

   Why? Because modifications (e.g., inserts, deletes) require updates to occur in index files

CLUSTERING FILES

In some relational DBMSs, related records from different tables can be stored together in the same disk area

Useful for improving performance of join operations

Primary key records of the main table are stored adjacent to associated foreign key records of the dependent table

e.g. Oracle has a CREATE CLUSTER command

RAID

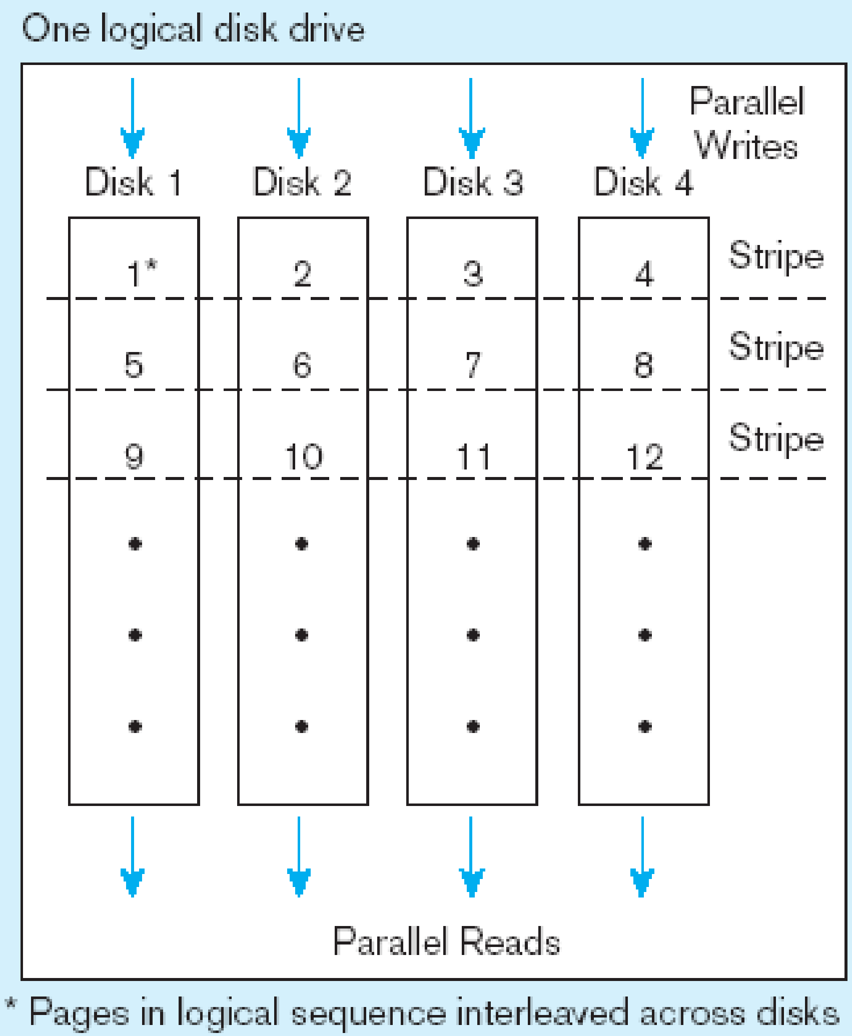
Redundant Array of Inexpensive Disks

A set of disk drives that appear to the user to be a single disk drive

Allows parallel access to data (improves access speed)

Pages are arranged in **stripes**

RAID with four disks and striping



Here, pages 1-4 can be read/written simultaneously

RAID TYPES

Raid 0

Maximized parallelism

No redundancy

No error correction

no fault-tolerance

Raid 1

Redundant data – fault tolerant

Most common form

Raid 2

No redundancy

One record spans across data disks

Error correction in multiple disks– reconstruct damaged data

Raid 3

Error correction in one disk

Record spans multiple data disks (more than RAID2)

Not good for multi-user environments,

Raid 4

Error correction in one disk

Multiple records per stripe

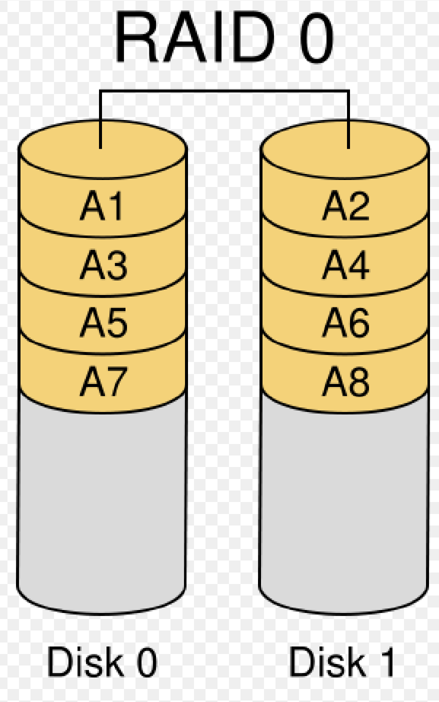
Parallelism, but slow updates due to error correction contention

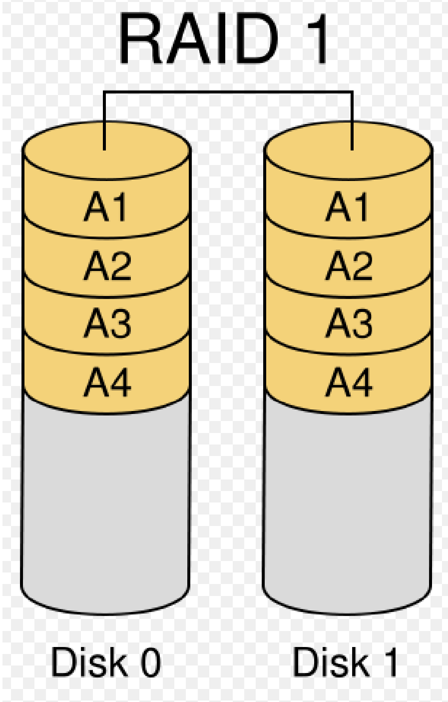
Raid 5

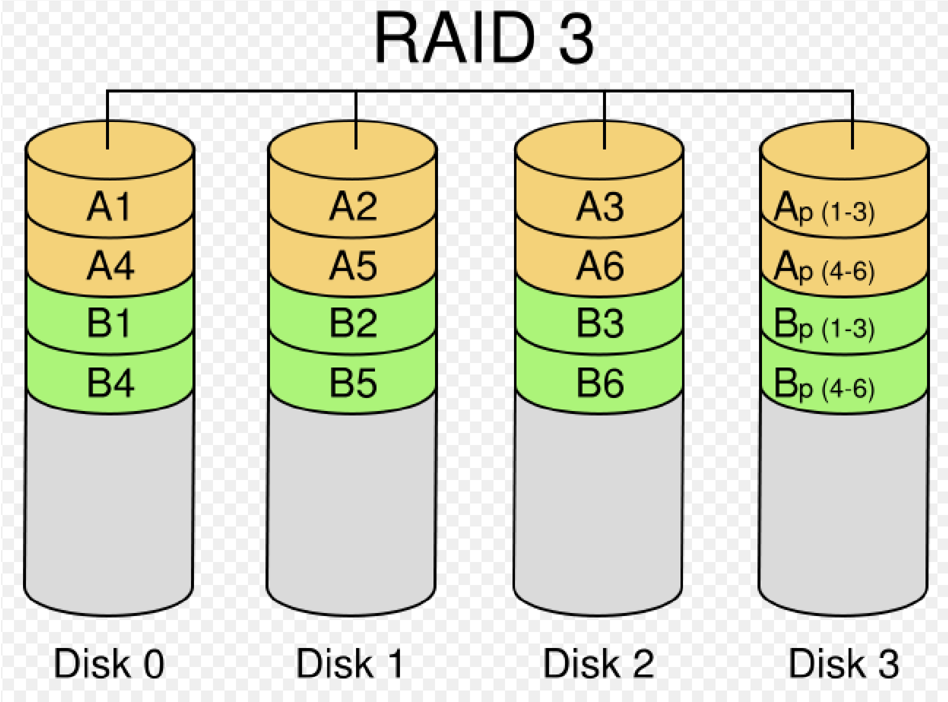
Rotating parity array

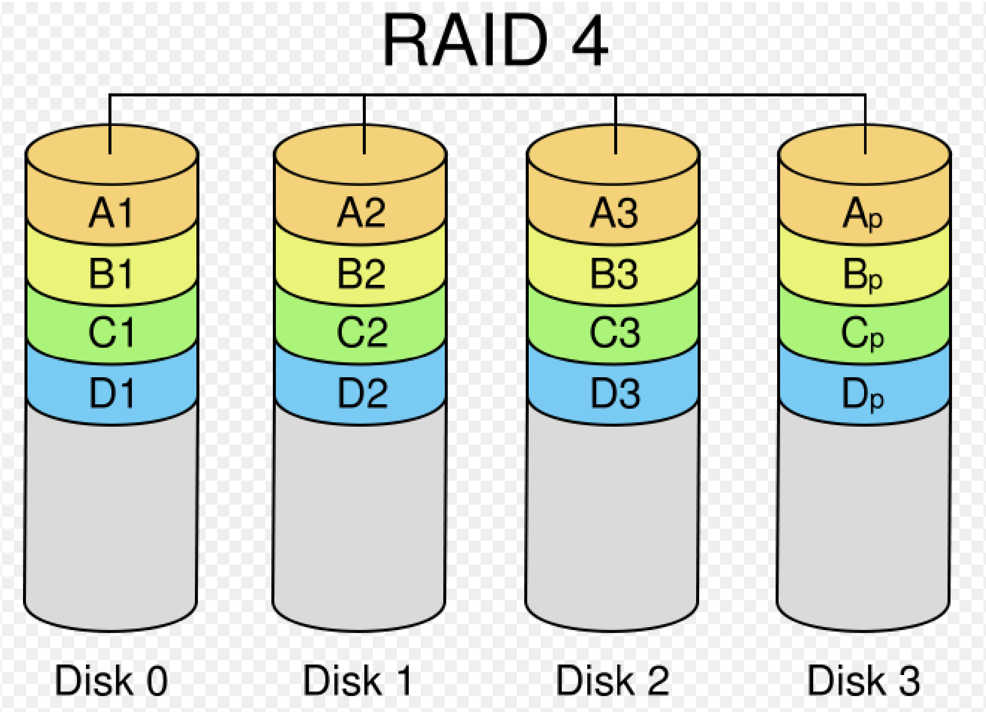
Error correction takes place in same disks as data storage

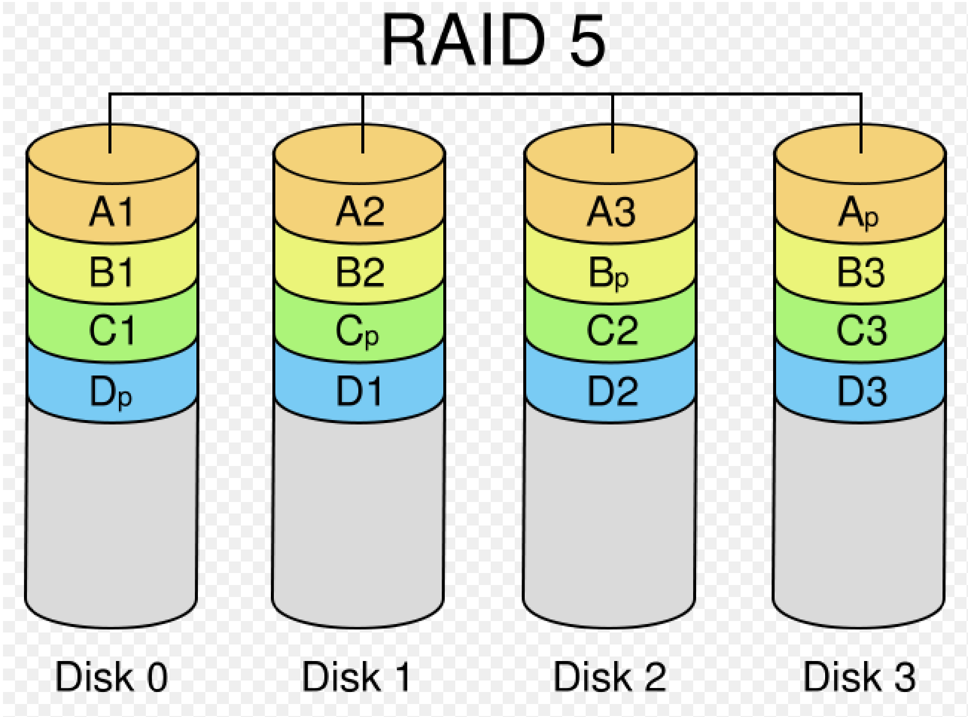
Parallelism, better performance than Raid4


RAID 0
A1
A3
A5
A7
A2
A4
A6
A8
Disk 0
Disk 1


RAID 1
A1
A2
A3
A4
A1
A2
A3
A4
Disk 0
Disk 1


RAID 3
A1
A4
B1
B4
A2
A5
B2
B5
A3
A6
B3
B6
Ap (1-3)
Ap (4-6)
Bp (1-3)
Bp (4-6)
Disk 0
Disk 1
Disk 2
Disk 3


RAID 4
A1
B1
C1
D1
A2
B2
C2
D2
A3
B3
C3
D3
Ap
Bp
Cp
Dp
Disk 0
Disk 1
Disk 2
Disk 3


RAID 5
A1
B1
C1
Dp
A2
B2
Cp
D1
A3
Bp
C2
D2
Ap
B3
C3
D3
Disk 0
Disk 1
Disk 2
Disk 3