

# Sistem Basis Data Terdistribusi

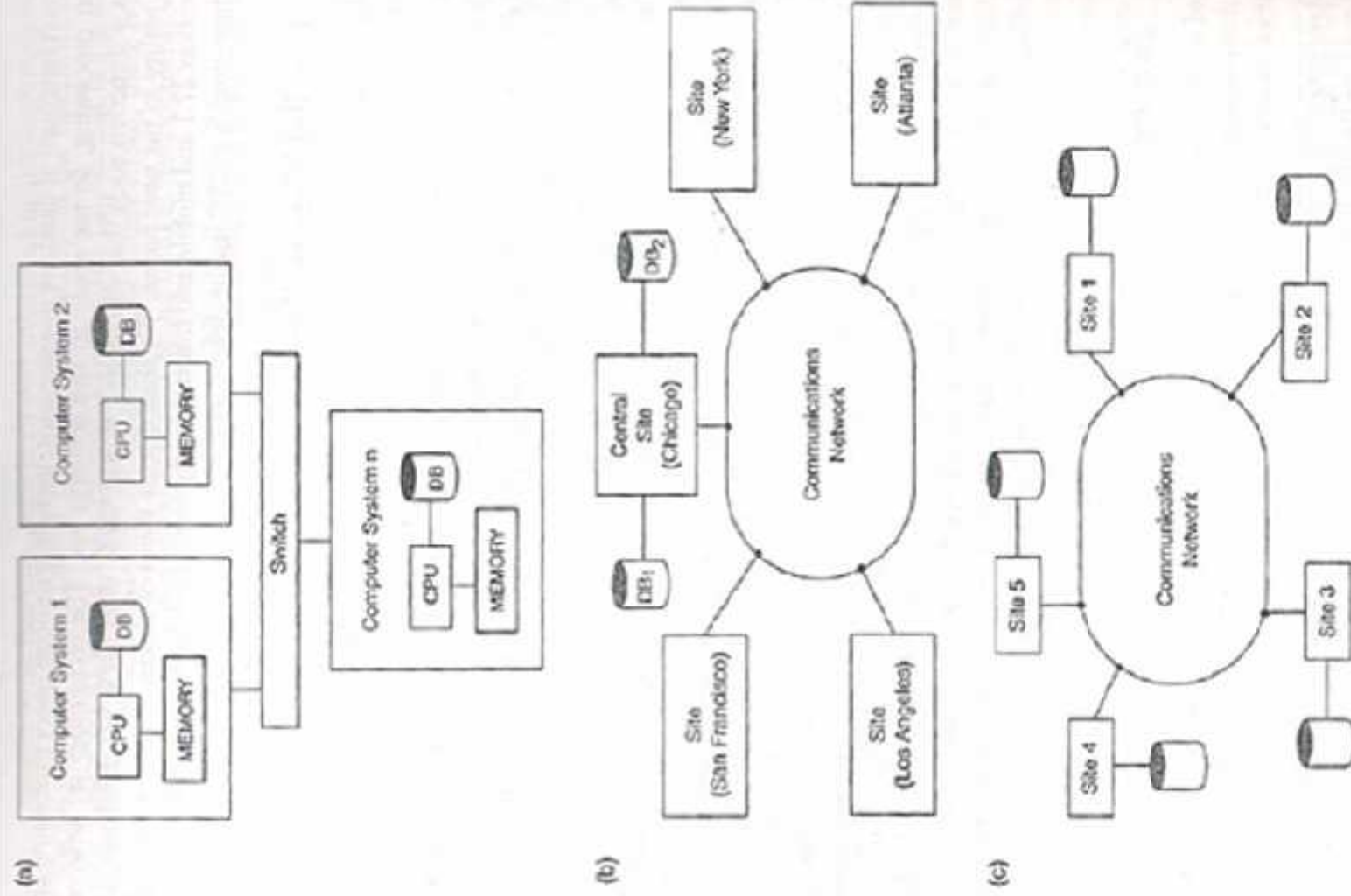
Sistem Basis Data Lanjut

Prepared by: MT. Wilson

Sumber: *Fundamentals of Database Systems, Third Edition ch.24*, Elmasri

# Konsep Basis Data Terdistribusi

- Sistem Komputasi Terdistribusi adalah sejumlah elemen proses yang terkoneksi melalui jaringan komputer dan saling bekerjasama dalam melakukan suatu tugas
- Basis Data Teridistribusi adalah kumpulan basis-basis data yang saling berhubungan secara logika dan tersebar pada sebuah jaringan komputer
- Sistem Manajemen Basis Data adalah sebuah sistem software yang mengelola basis data terdistribusi



**Figure 24.1** Some different database system architectures. (a) Shared nothing architecture. (b) A networked architecture with a centralized database at one of the sites. (c) A truly distributed database architecture.

# Keuntungan Basis Data Terdistribusi

- Manajemen data terdistribusi dengan tingkat transparansi yang berbeda
- Keandalan dan ketersediaan
- Peningkatan performa
- Ekspansi yang lebih mudah

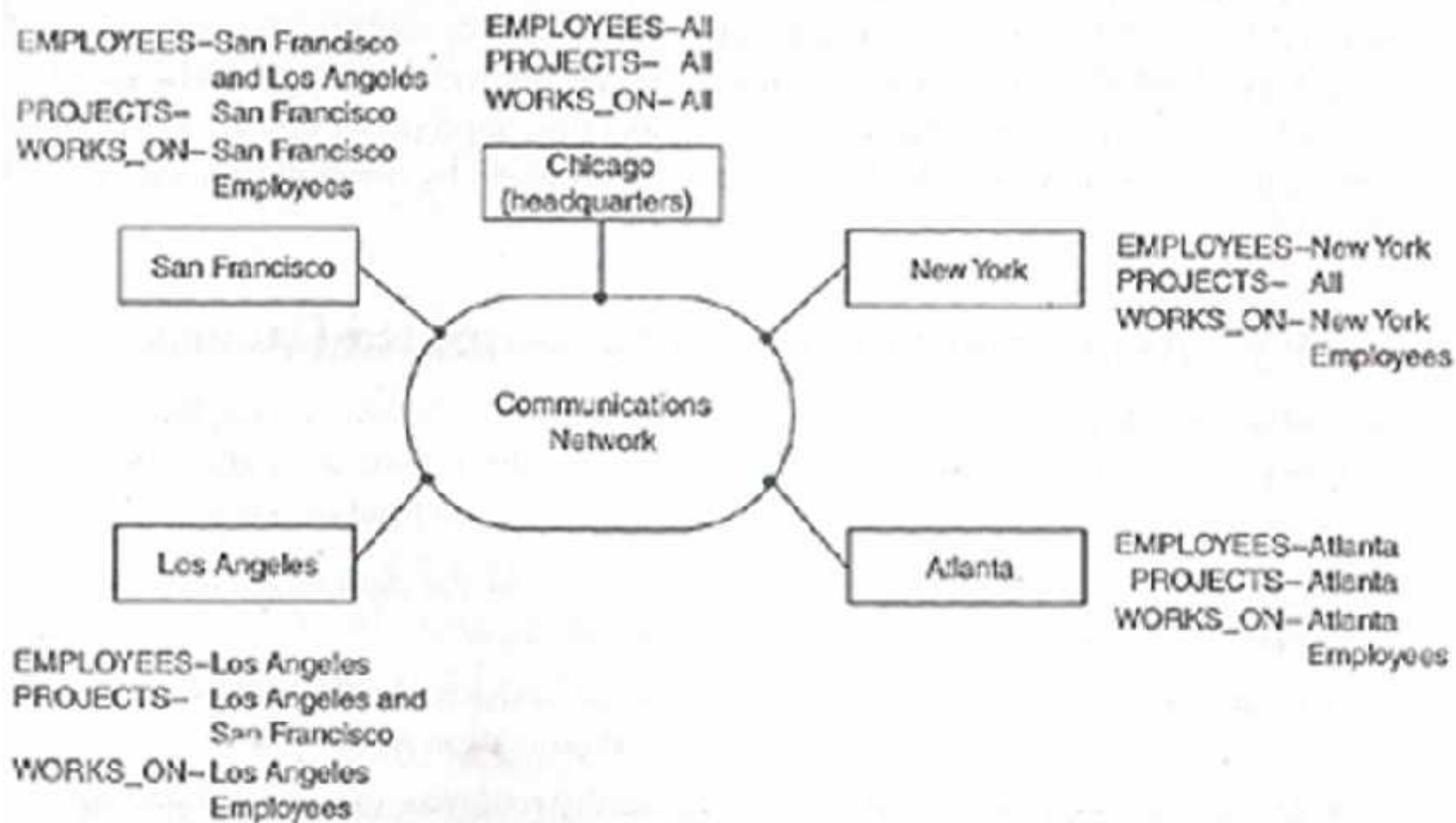


Figure 24.2 Data distribution and replication among distributed databases

# Fungsi Tambahan Basis Data Terdistribusi

- Keeping track of data
- Proses query yang terdistribusi
- Manajemen transaksi yang terdistribusi
- Manajemen replikasi data
- Pemulihan basis data terdistribusi
- Keamanan
- Manajemen direktori (katalog) terdistribusi

# Perbedaan DDBMS dengan Sistem Sentral: Hardware

- Ada banyak komputer yang disebut **sites** atau **nodes**
- Site-site ini terhubung oleh **jaringan komunikasi** untuk mengirim data dan perintah-perintah di antara site-site tersebut

# Teknik Memecah Basis Data

- Memecah basis data menjadi unit-unit logika yang disebut **fragmen-fragmen** yang bisa *diassign* untuk disimpan di berbagai site
- Penggunaan **replikasi data** yang memungkinkan data tertentu untuk disimpan di lebih dari satu site
- Proses **alokasi** fragmen –atau replika fragmen- untuk penyimpanan di berbagai site



# Teknik Memecah Basis Data (2)

- Teknik-teknik di atas digunakan dalam proses perancangan basis data terdistribusi
- Informasi yang berhubungan dengan fragmentasi, alokasi dan replikasi data disimpan dalam sebuah **global directory** yang diakses oleh aplikasi DDBS

# Fragmentasi Data

- Fragmentasi Horizontal
  - Membagi sebuah relasi secara horizontal dengan mengelompokkan baris-baris untuk membuat subset dari tuples
- Fragmentasi Vertical
  - Membagi sebuah relasi secara vertikal berdasarkan kolom
- Fragmentasi Hibrid

# Replikasi dan Alokasi Data

- Berguna untuk meningkatkan ketersediaan data
- Kasus paling ekstrim: replikasi keseluruhan basis data pada setiap site dalam sistem terdistribusi → **fully replicated distributed database**
  - Keuntungan: sistem terus beroperasi selama paling tidak satu site tetap berjalan
  - Kelemahan: operasi update menjadi lambat
- Kasus paling ekstrim lainnya: **tidak ada replikasi**

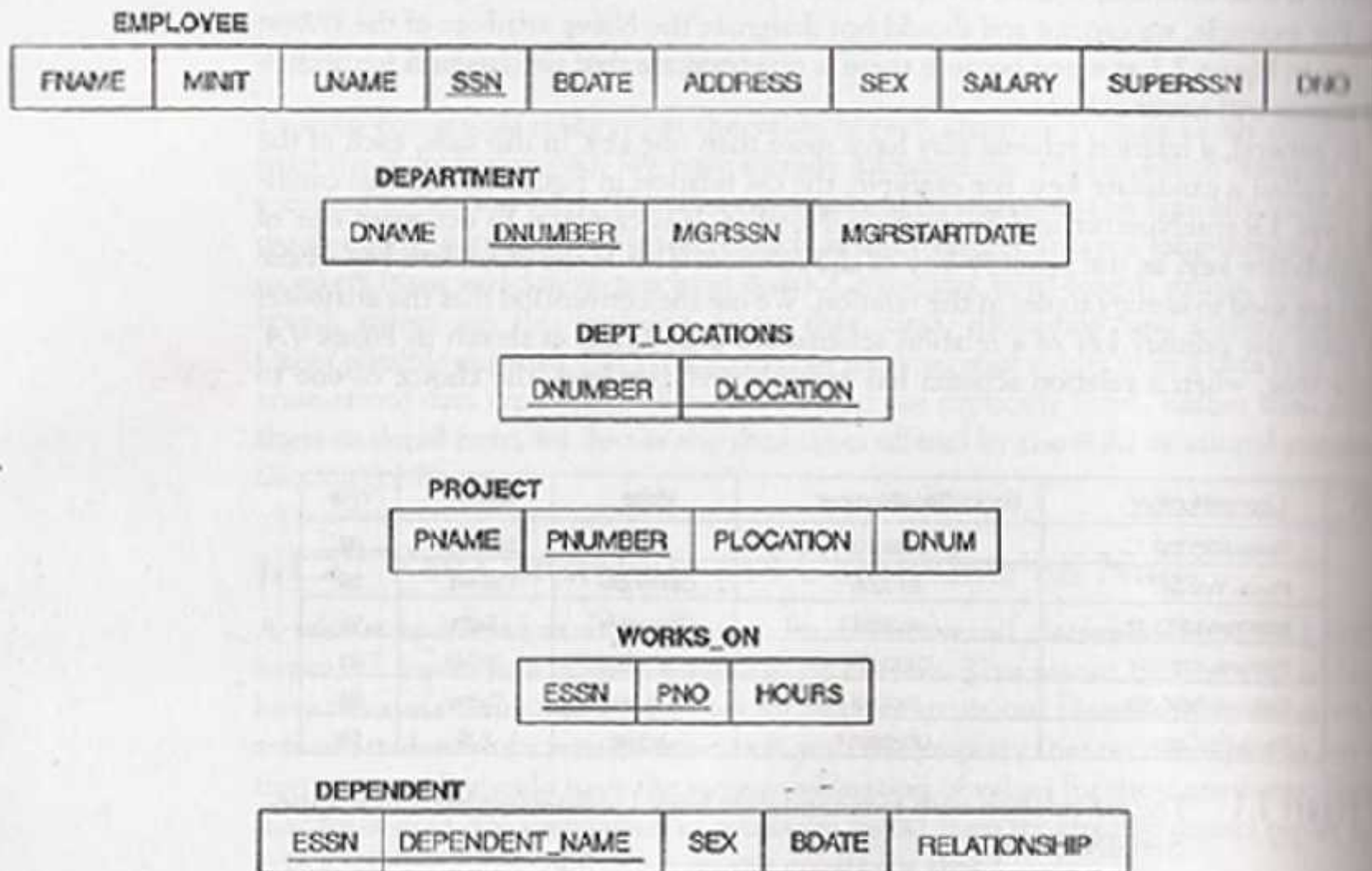
# Replikasi dan Alokasi Data (2)

- Kasus khusus replikasi partial untuk pekerja yang mobile –seperti sales rep, financial planners dan claim adjustors- mereka membawa basis data replikasi dalam laptop atau PDA dan melakukan sinkronisasi secara periodik dengan server basis data

# Replikasi dan Alokasi Data (3)

- Masing-masing fragmen –atau salinan dari sebuah fragmen- harus *diassign* ke site tertentu dalam sistem terdistribusi. Proses ini disebut **distribusi data** (atau **alokasi data**)

# Contoh Fragmentasi, Alokasi dan Replikasi



**Figure 7.5** Schema diagram for the *COMPANY* relational database schema; the primary keys are underlined.

EMPLOYEE	FNAME	MINT	LNAME	SSN	BOATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	1955-01-09	721 Fardem, Houston, TX	M	30000	333445555	5	
Franklin	T	Wong	333445555	1955-12-01	528 Voss, Houston, TX	M	40000	888667777	5	
Alisa	J	Zeigler	999887777	1960-07-19	3321 Coase, Spring, TX	F	25000	987654321	4	
Jennifer	S	Wallace	987654321	1941-06-03	2518 Barry, Dallas, TX	F	40000	888667777	4	
Farrell	K	Narayan	666884444	1935-09-15	975 Firo Oak, Humble, TX	M	30000	333445555	5	
Joyce	A	English	453453453	1972-07-31	5031 Rice, Houston, TX	F	25000	333445555	5	
Ahmed	V	Jobor	007087987	1956-03-29	980 Dallas, Houston, TX	M	25000	987654321	4	
James	E	Darg	888667777	1937-11-10	450 Stone, Houston, TX	M	55000	null	1	

DEPT.	LOCATIONS	DNUMBER	DLOCATION
		1	Houston
		4	Stafford
		5	Dallas
		5	Springfield
		5	Houston

DEPARTMENT	DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research		5	333445555	1988-06-22
Administration		4	987654321	1956-01-01
Headquarters		1	888667777	1981-06-19

WORKS ON	ESSN	PNO	HOURS
	123456789	1	32.5
	123456789	2	7.5
	666884444	3	40.0
	453453453	1	20.0
	453453453	2	20.0
	333445555	2	10.0
	333445555	3	10.0
	333445555	10	10.0
	333445555	20	10.0
	999887777	30	30.0
	999887777	10	10.0
	987654321	10	35.0
	987654321	20	5.0
	987654321	30	20.0
	987654321	25	15.0
	888667777	20	null

PROJECT	PNAME	PNUMBER	PLOCATION	PNUM
	ProductX	1	Dallas	5
	ProductY	2	Springfield	5
	ProductZ	3	Houston	5
	Computerization	10	Stafford	4
	Reorganization	20	Houston	1
	Newproducts	30	Stafford	4

DEPENDENT	ESSN	DEPENDENT_NAME	SEX	BOATE	RELATIONSHIP
	333445555	Alice	F	1960-04-05	DAUGHTER
	333445555	Theodore	M	1955-10-25	SON
	333445555	Jay	F	1958-05-03	SPOUSE
	987654321	Ahmed	M	1942-02-28	SPOUSE
	123456789	Michael	M	1968-01-04	SON
	123456789	Alice	F	1968-12-30	DAUGHTER
	123456789	Elizabeth	F	1967-09-05	SPOUSE

Figure 1.6 One possible relational database state corresponding to the COMPANY schema.



(a)

EMP05	FNAME	MINIT	LNAME	SSN	SALARY	SUPERSSN	DNO
	John	B	Smith	123456789	30000	300445555	5
	Prakash	T	Wong	300445555	40000	800665555	5
	Ramesh	H	Nayyar	800665555	35000	300445555	5
	Joyce	A	English	453453453	25000	300445555	5

DEP5	DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
	Research	5	300445555	1980-05-22

DEP5_LOCS	DNUMBER	LOCATION
	5	Bellare
	5	Sugarland
	5	Houston

WORKS_ON5	ESSN	PNO	HOURS
	123456789	1	32.5
	123456789	2	7.5
	066884444	3	90.0
	453453453	1	20.0
	453453453	2	20.0
	300445555	2	10.0
	300445555	3	10.0
	300445555	10	10.0
	300445555	20	10.0

PROJ55	PNAME	PNUMBER	PLOCATION	DNUM
	Product X	1	Bellare	5
	Product Y	2	Sugarland	5
	Product Z	3	Houston	5

Data at Site 2

(b)

EMP04	FNAME	MINIT	LNAME	SSN	SALARY	SUPERSSN	DNO
	Alice	J	Zelazny	999887777	25000	987654321	4
	Jennifer	G	Walls	057654321	40000	800665555	4
	Abirad	V	Jacob	987654321	25000	987654321	4

DEP4	DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
	Administration	4	987654321	1995-01-01

DEP4_LOCS	DNUMBER	LOCATION
	4	Stafford

WORKS_ON4	ESSN	PNO	HOURS
	300445555	10	10.0
	999887777	30	30.0
	999887777	10	10.0
	987654321	10	20.0
	987654321	30	5.0
	987654321	30	20.0
	987654321	20	15.0

PROJ04	PNAME	PNUMBER	PLOCATION	DNUM
	Computerization	10	Stafford	4
	Newberetts	30	Stafford	4

Data at Site 3

Figure 24.3 Allocation of fragments to sites. (a) Relation fragments at site 2 corresponding to department 5. (b) Relation fragments at site 3 corresponding to department 4.

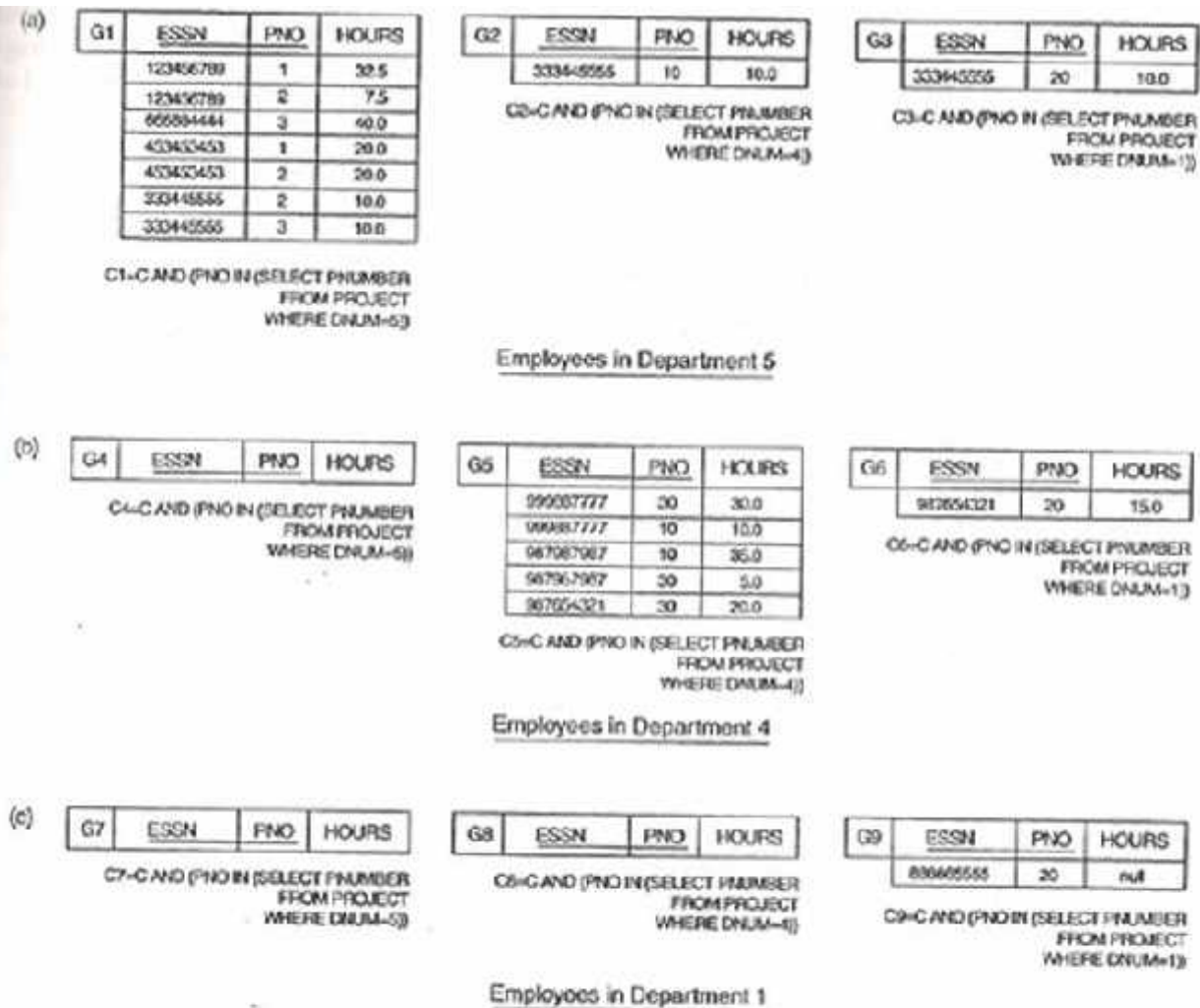


Figure 24.4 Complete and disjoint fragments of the `WORKS_ON` relation. (a) Fragments of `WORKS_ON` for employees working in department 5 ( $\sigma_{(C=[ESSN \text{ IN } (SELECT SSN \text{ FROM EMPLOYEE WHERE DNO=5)])}$ ). (b) Fragments of `WORKS_ON` for employees working in department 4 ( $\sigma_{(C=[ESSN \text{ IN } (SELECT SSN \text{ FROM EMPLOYEE WHERE DNO=4)])}$ ). (c) Fragments of `WORKS_ON` for employees working in department 1 ( $\sigma_{(C=[ESSN \text{ IN } (SELECT SSN \text{ FROM EMPLOYEE WHERE DNO=1)])}$ ).