CSE 4125: Distributed Database Systems
Chapter – 4
(Part – C)

Distributed Database Design

Topics to be discussed -

- Design of Derived Horizontal Fragmentation
- Design Vertical Fragmentation

The Design of
Derived Horizontal Fragmentation

Derived Horizontal Fragmentation

- The horizontal fragmentation of a relation cannot be based on a property of its own attributes, but is derived from the horizontal fragmentation of another relation.

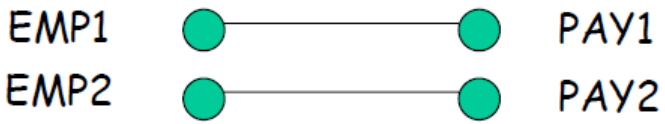
- Derived fragmentation is used to facilitate the join between fragments.

Distributed Join

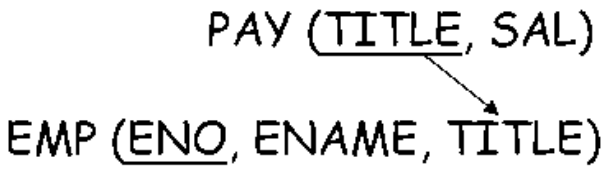
A distributed join is a join between horizontally fragmented relations.

Join Graph

- A distributed join is represented efficiently using **join graphs**.
- The join graph *G* of the distributed join *R JN S* is a *graph(N,E)*, where
  - ✓ nodes ***N:*** fragments of R and S.
  - ✓ non directed edges ***E:*** Join between fragments which are not intrinsically empty.

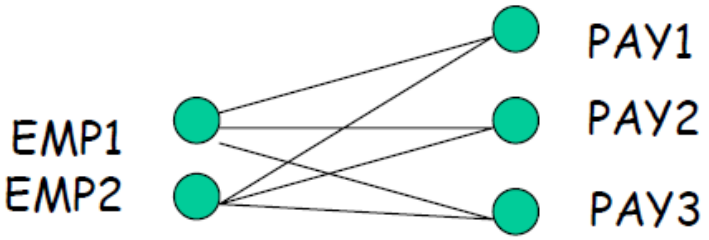


Example:

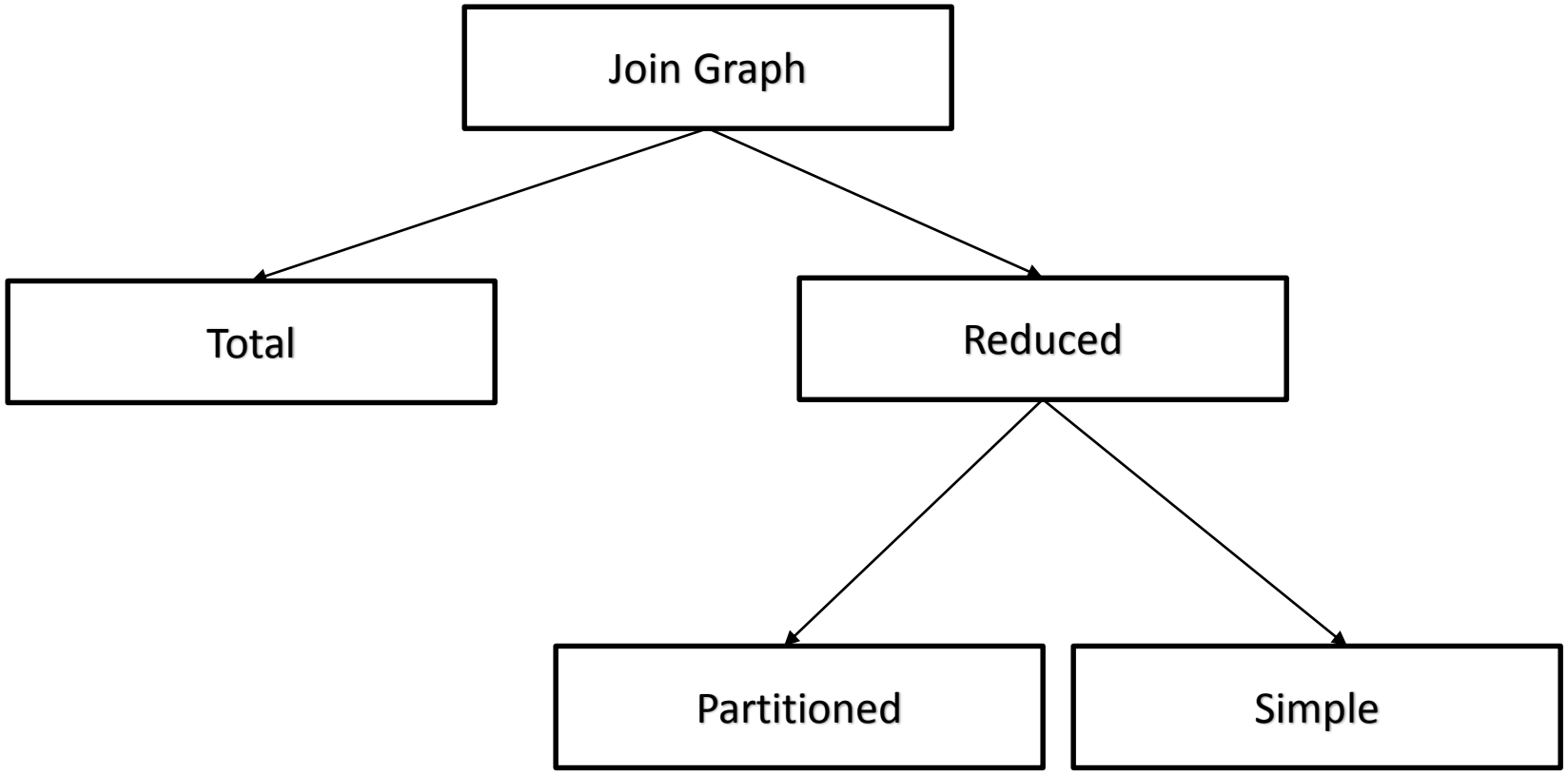


- ✓Divide ***EMP*** into ***EMP1*** and ***EMP2*** based on ***TITLE***
- ✓Divide ***PAY*** into ***PAY1, PAY2, PAY3*** based on ***SAL.***

To join ***EMP*** and ***PAY,*** we have one possible following scenario.

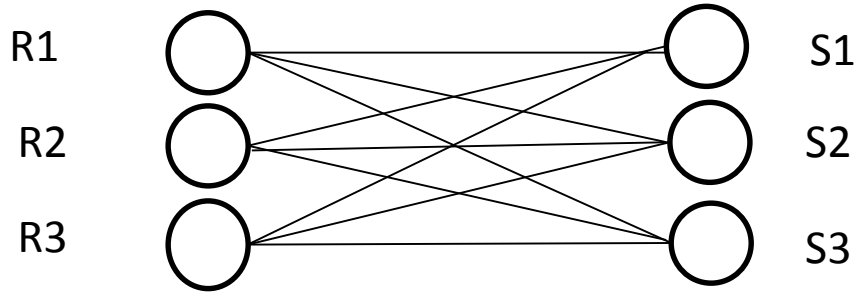


Types of Join Graph

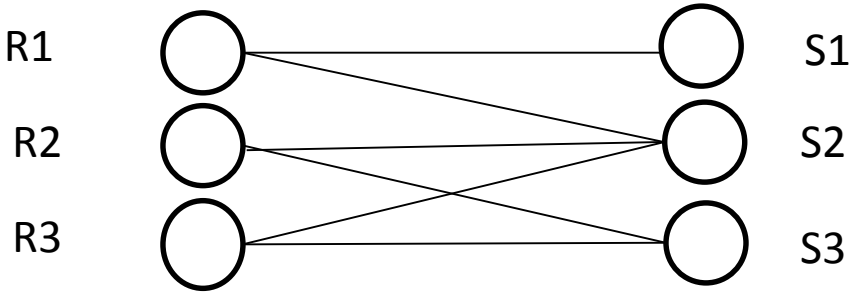


Types of Join Graph

Total Join Graph: when a join graph contains all possible edges between fragments of R and S.

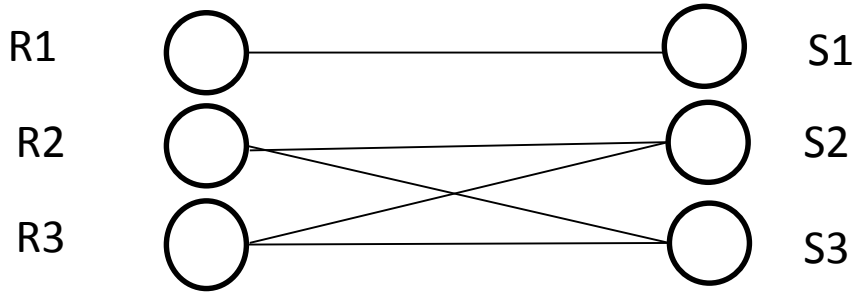


Reduced Join Graph: when some of the edges between fragments of R and S are missing.

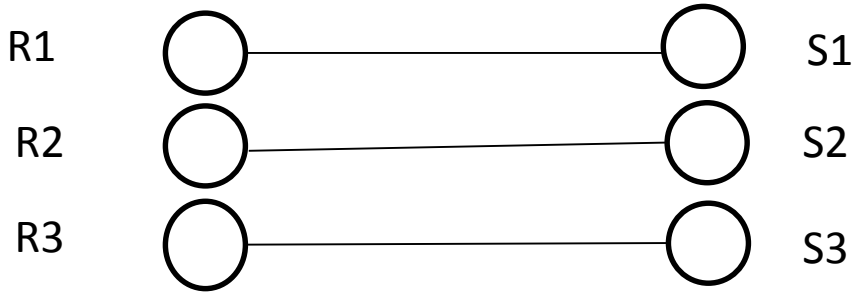


Types of Join Graph

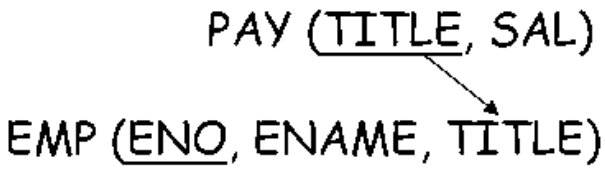
Partitioned Join Graph: when a reduced join graph is composed of two or more subgraphs without edges between them.



Simple Join Graph: when a reduced graph is partitioned and each subgraph has just one edge.



Example:



EMP

ENO	ENAME	TITLE
1	A	Dev
2	B	CAD
3	C	Main
4	D	Dev

PAY

TITLE	SAL
Dev	10 K
CAD	20 K
Main	30 K

After JOIN

ENO	ENAME	E.TITLE	P.TITLE	SAL
1	A	Dev	Dev	10 K
4	D	Dev	Dev	10 K
2	B	CAD	CAD	20 K
3	C	Main	Main	30 K

Divide ***EMP*** into ***EMP1*** and ***EMP2*** based on ***TITLE***

EMP1

ENO	ENAME	TITLE
1	A	Dev
4	D	Dev

EMP2

ENO	ENAME	TITLE
2	B	CAD
3	C	Main

Divide ***PAY*** into ***PAY1, PAY2, PAY3*** based on ***SAL.***

PAY1

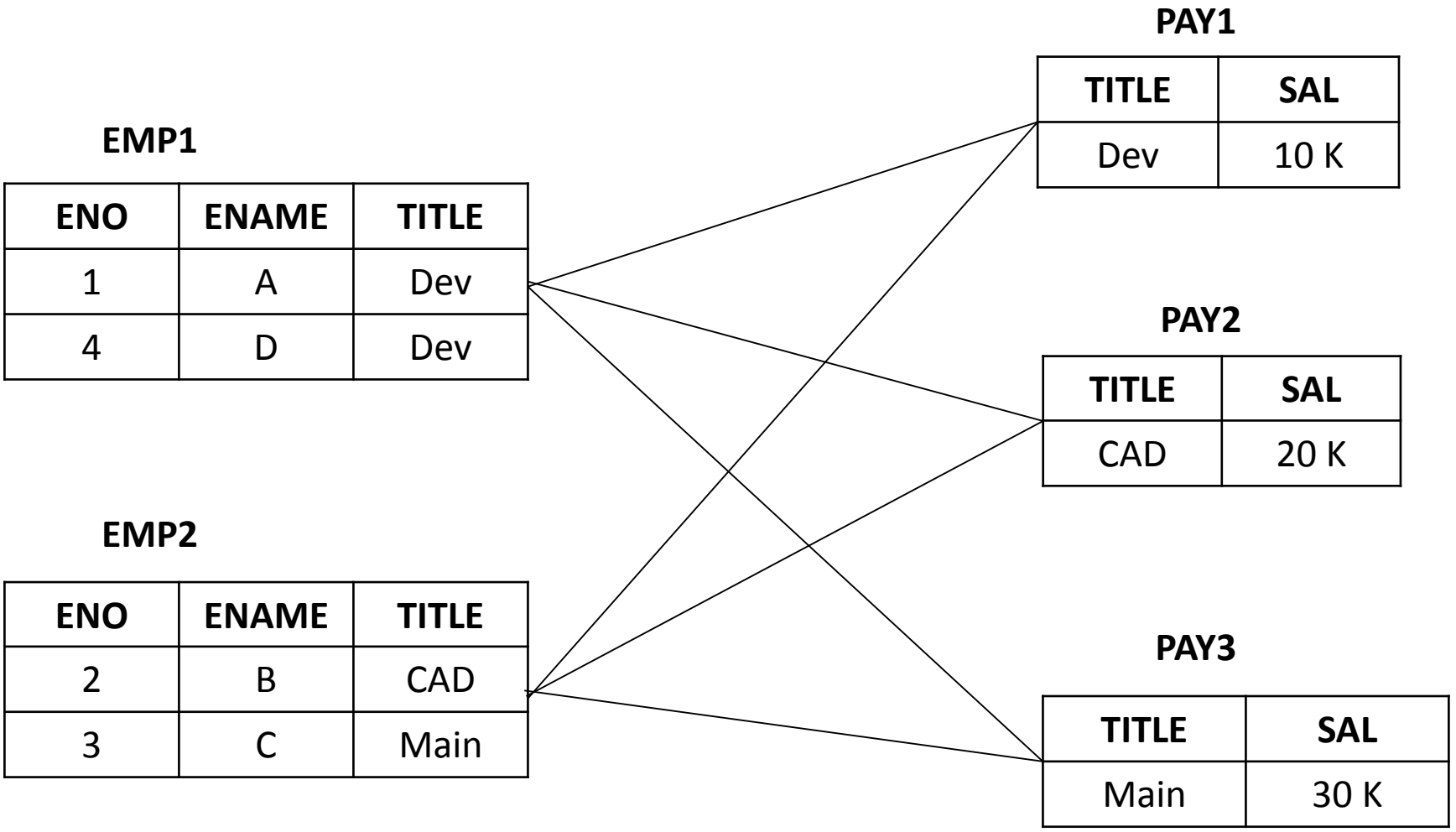
TITLE	SAL
Dev	10 K

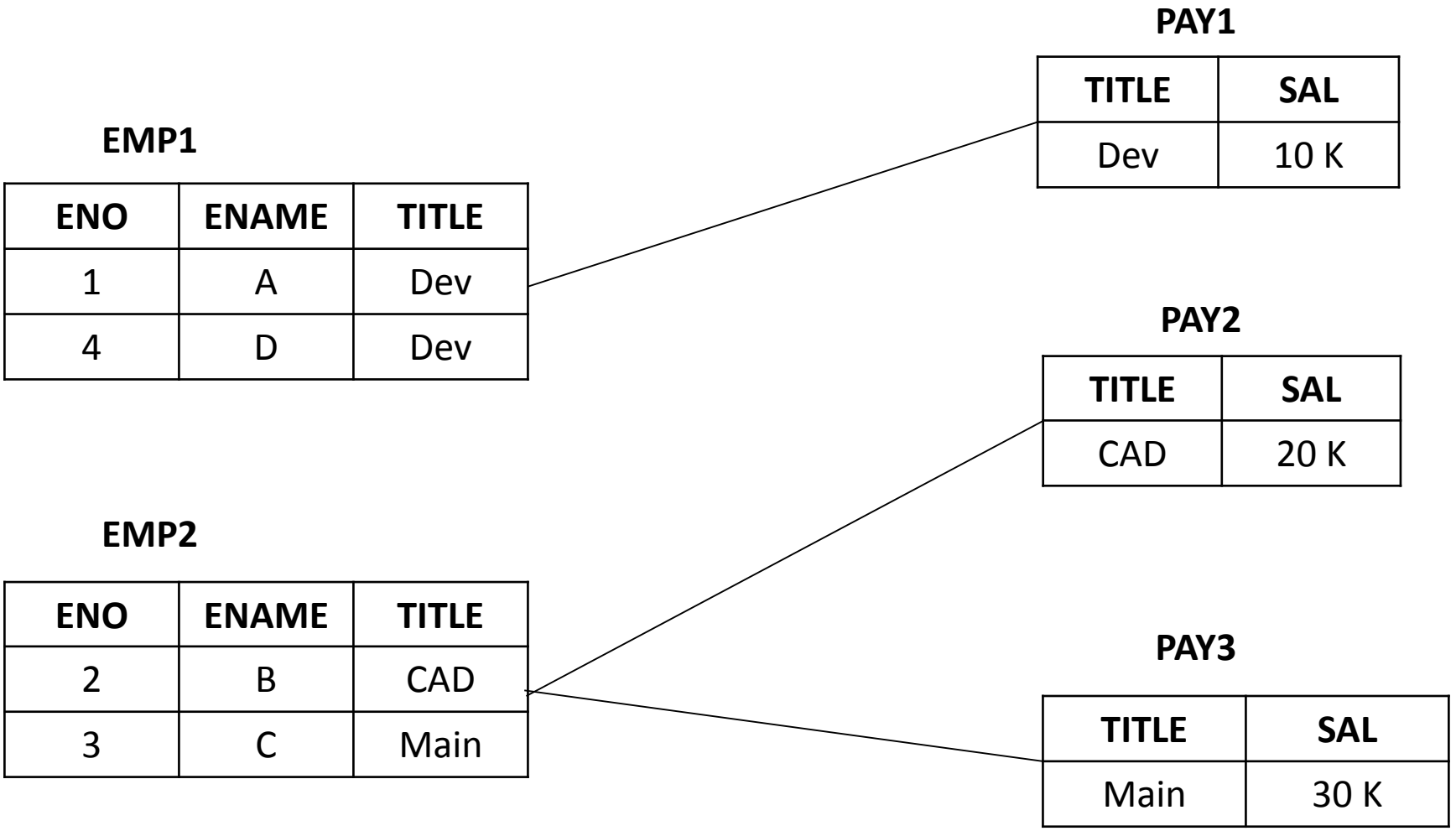
PAY2

TITLE	SAL
CAD	20 K

PAY3

TITLE	SAL
Main	30 K


EMP1
ENO
ENAME
TITLE
1
A
Dev
4
D
Dev
EMP2
ENO
ENAME
TITLE
2
B
CAD
3
C
Main
PAY1
TITLE
SAL
Dev
10 K
PAY2
TITLE
SAL
CAD
20 K
PAY3
TITLE
SAL
Main
30 K


PAY1
TITLE
SAL
Dev
10 K
EMP1
ENO
ENAME
TITLE
1
A
Dev
4
D
Dev
PAY2
TITLE
SAL
CAD
20 K
EMP2
ENO
ENAME
TITLE
2
B
CAD
3
C
Main
PAY3
TITLE
SAL
Main
30 K