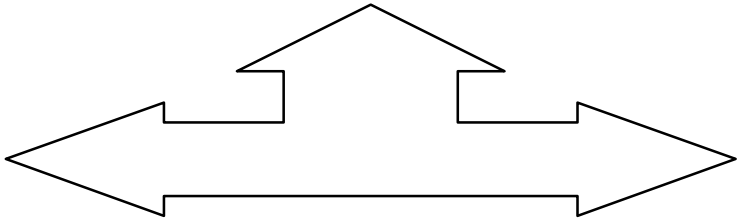
An Example of Fragmentation

JNO	JNAME	BUDGET	LOCATION
1	Instrumentation	1 500 000	London
2	CAD/CAM	1 650 000	New York
3	Development	2 000 000	Cleveland
4	Maintenance	950 000	Paris

User perception

New York fragment



London fragment

JNO	JNAME	BUDGET	LOCATION
2	CAD/CAM	1 650 000	New York
3	Development	2 000 000	Cleveland

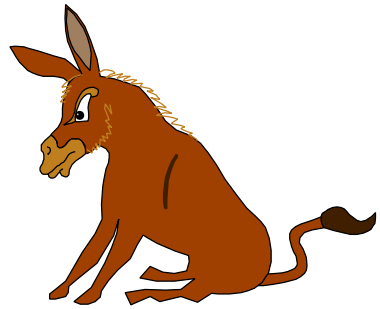
physical storage
New York

JNO	JNAME	BUDGET	LOCATI
1	Instrumentation	1 500 000	London
4	Maintenance	950 000	Paris

physical storage
London

An Example of Data Replication

JNO	JNAME	BUDGET	LOCATION
1	Instrumentation	1 500 000	London
2	CAD/CAM	1 650 000	New York
3	Development	2 000 000	Cleveland
4	Maintenance	950 000	Paris



User perception

New York fragment

JNO	JNAME	BUDGET	LOCATION
2	CAD/CAM	1 650 000	New York
3	Development	2 000 000	Cleveland

London fragment

JNO	JNAME	BUDGET	LOCATI
1	Instrumentation	1 500 000	London
4	Maintenance	950 000	Paris

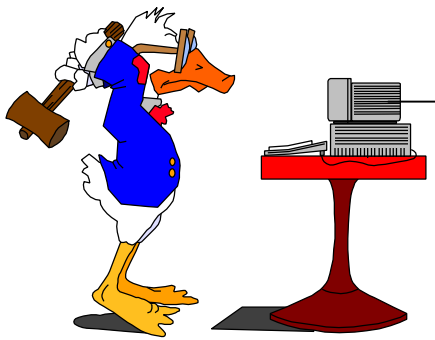
Replica of London fragment

JNO	JNAME	BUDGET	LOCATI
1	Instrumentation	1 500 000	London
4	Maintenance	950 000	Paris

Replica of New York fragment

JNO	JNAME	BUDGET	LOCATIO
2	CAD/CAM	1 650 000	New York
3	Development	2 000 000	Cleveland

DBMS independence – II



INGRES User

INGRES (SQL)

INGRES/ STAR

GATE WAY

ORACLE (SQL)

INGRES database

Distributed INGRES database

ORACLE database

History of distributed DB

- Concepts behind distributed DBMS were pioneered during the late 1970's in the IBM research project R*Star.

  IBM's subsequent delivery of distributed DBMS products has been part of a 10 year evolving technology known as DRDA (distributed relational data architecture).

- DRDA at this time is largely an approach for integrating data sets across the different versions of DB2 and was the breakthrough of distributed DB technology.

- The first well-publicised distributed DBMS product was INGRES/Star, announced in 1987. Oracle also announced distributed DBMS capabilities in 1987, but largely as a marketing ploy. The first Oracle product to reasonably support distributed database processing is Oracle 7, which has been in the market since 1993.

$\Rightarrow$ Constantly growing interests and use of distributed DBMS.

1st decade: Breakthrough

- Fail of database machines (Boral and DeWitt 1983)

  DB machines : An idea whose time has passed.

- From the 90th : Technological chances.

  Mainframe based servers achieved lesser performances than developed parallel machines (mainly by Teradata and Tandem).
  Shared nothing machine by Teradata built since 1978!

- Domination of the relational model in the 90th.

  From the appearance in the marketplace in 1983 to the domination. Relational queries are ideally suited to parallel execution.

$\Rightarrow$ Breakthrough of parallel database technology. Not so much older than its distributed brother.

2nd decade: Enthusiasm Highly parallel database systems

- MPP machines Intel, NCP, nCube machines up-to 100 processors based on conventional processors, memories and disks at low price.
- At least five running industrial products (Teradata, Tandem, Super Database, Persist by IBM, NCR).
- Success in OLTP against a very large database (near linear speed-up and scale-up in the Teradata and Tandem system).

$\Rightarrow$ Success and viability of highly parallel machines.

3rd decade: Problems

- Fail of Parallel Machines

  Too many parallel machine provider and fail of many machines (e.g. KSR, Transputers). Single processing components too less powerful.

- Parallel programming problems

  Difficulty of programming a parallel machines.

→ Shared virtual memories.
→ Performance problems.
Lack of parallel debuggers and monitoring tools.

- Performance problems

  Missing speed-up and scale-up for massive parallel machines. No efficient techniques for mixing complex batch and OLTP processing(multi-user databases).

⇒ Break down of enthusiasm.

4th decade: Consolidation

- Technology chances:

  New (virtual) parallel machines like cluster of workstations, PC's. Pile of PC's let reconsider the MPP aim.

- At least ten running industrial products (see before). Dominant in market-niches (e.g. great database server ALTAVISTA, OLTP processing insurance security, military ...), immense financial transactions.

- Programming advances:

  PVM, MPI, Linda and friends, RMI's.

- Application push: Data warehouse etc.

$\Rightarrow$ Consolidation and dominance in special applications.

Architecture of PDDBMS