COMP283-Lecture 2
Applied Database Management

Introduction	
DB Design	Considerations
Built in data types	MS-SQL, Oracle, MySQL
DB Structure	DB Files
	Pages & Extents

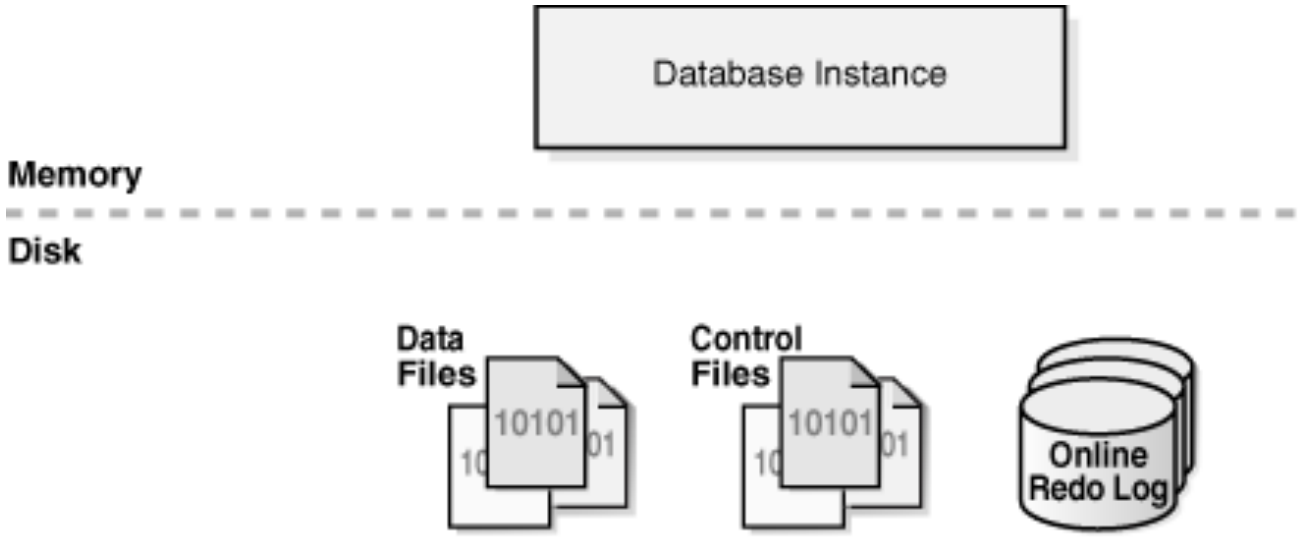
COMP283-Lecture 2

Main Database Files:

MS-SQL	Oracle	MySQL
Primary Data (MDB)	Tablespaces	FRM
Transacton Logs (LDB)	Transaction Logs	Transaction Logs
Secondary Data (NDB)	Redo logs	MYD
TempDB	Control Files	MYI

COMP283-Lecture 2

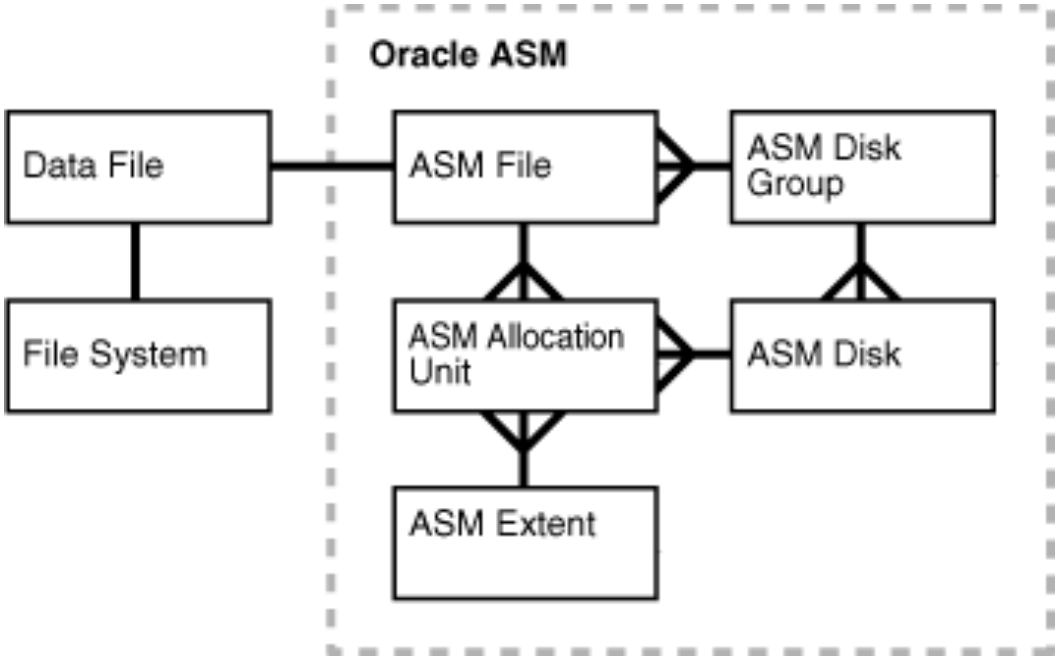
Main Database Files:



- In Oracle, a Database Instance is a set of memory structures that manage database files - associated with a single database.
- Oracle can be configured to use “raw” disks. (raw??)

COMP283-Lecture 2

Main Database Files:



- Oracle ASM: Automatic Storage Management
- What are some advantages of using ASM vs standard filesystem?

COMP283-Lecture 2
DB Design: Assessing file store requirements

- Assess current and expected storage capacity
  - Disk Space
  - Disk throughput
  - Location storage requirements

COMP283-Lecture 2
DB Design: Integer Data Types

	MS-SQL	Oracle	MySQL
bit	0 or 1 1 byte		0 or 1 up to 8 bytes
tinyint	0..255 1 byte		-127..127 or 0..255 1 byte
smallint	-2^{15}.. $2^{15}-1$ 2 bytes		-2^{15} .. 2^{15} or 0..$2^{16}-1$ 2 bytes
mediumint			$-2^{23}..2^{23}$ or 0..$2^{24}-1$ 3 bytes
Int(eger)	-2^{31}.. $2^{31}-1$ 4 bytes	-2^{37} .. 2^{37} or 0.. $2^{38}-1$ Up to 5 bytes	$-2^{31}..2^{31}$ or 0..$2^{32}-1$ 4 bytes
bigint	-2^{63}.. $2^{63}-1$ 8 bytes		$-2^{63}..2^{63}-1$ or 0..$2^{64}-1$ 8 bytes

COMP283-Lecture 2

DB Design: Real Number Data Types

	MS-SQL	Oracle	MySQL
numeric/ decimal	$-2^{38}+1$ to $2^{38}-1$ 5 to 17 bytes		$-10^{64}+1$ to $10^{64}-1$ or 0 to $10^{65}-1$ Variable bytes
number		$-2^{38}+1$ to $2^{38}-1$ Variable bytes	
smallmoney	-214748.3648 to 214748.3647 4 bytes		
money	-922337203685477.5808 to 922337203685477.5807 8 Bytes		
real	- 3.40E+38 to -1.18E-38, 0 and 1.18E-38 to 3.40E+38 4 bytes		–1.79308E38 to –2.23308E38, 0, and 2.23308E38 to 1.79308E38 4 or 8 bytes
float(n)	- 1.79E+308 to -2.23E-308, 0 and 2.23E-308 to 1.79E+308 4 or 8 bytes		–3.438E38 to –1.1838E38, 0, and 1.1838E38 to 3.438E38 4 bytes
double			–1.79308E38 to –2.23308E38, 0, and 2.23308E38 to 1.79308E38 4 or 8 bytes

COMP283-Lecture 2
DB Design: Character Data Types

	MS-SQL	Oracle	MySQL
CHAR(n)	n = 1 to 8000 bytes	n=1 to 2000 bytes	n=1 to 255 bytes
VARCHAR(n)	n = 1 to 8000 bytes or max (2^{31}-1 bytes)		n=1 to 65535 bytes (+ 1 or 2 length bytes)
VARCHAR2(n)		n=1 to 4000 bytes	
NCHAR(n)	n=1 to 4000 1 to 8000 bytes	n=1 to 2000 bytes	n=1 to 255 bytes
NVARCHAR(n)	n = 1 to 4000 (3 to 8002 bytes) or max (2^{31}-1 bytes)		
NVARCHAR2(n)		n=1 to 4000 bytes	

COMP283-Lecture 2

DB Design: Large Object Data Types

	MS-SQL	Oracle	MySQL
Binary(n)	1 to 8000 bytes		
Varbinary(n)	1 to 8000 bytes or max (2^{31}-1 bytes)		
blob		8 to 128 TBytes	65535 Bytes

COMP283-Lecture 2
DB Design: Date & Time Data Types

	MS-SQL	Oracle	MySQL
smalldatetime	Jan 1, 1900 to Jun 6, 2079 (4 bytes)		
datetime	Jan 1, 0001 to Dec 31, 9999 (8 bytes)		Jan 1, 1000, 00:00:00 to Dec 31, 9999, 23:59:59
Date	Jan 1, 0001 to Dec 31, 9999 (3 bytes)	Jan 1, 4712, 00:00:00 BC to Dec 31,9999 23:59:59	Jan 1, 1000 to Dec 31, 9999
time	00:00:00.0000000 to 23:59:59.9999999 (5 bytes)		-838:59:59 to 838:59:59 (as time difference)
datetimeoffset	Jan 1, 0001, 00:00:00 to Dec 31, 9999 23:59:59.9999999 (10 bytes)		
datetime2	Jan 1, 0001, 00:00:00 to Dec 31, 9999 23:59:59.9999999 (varies)		

COMP283-Lecture 2

DB Design: Data and Transaction Log Files in MS-SQL

- The primary data files (.MDF).
- Secondary data files (.NDF).
  - Filegroups.
- Log Data Files (.LDF).

DB Design: Pages & Extents in MS-SQL

- A Page is the smallest unit of data that can be stored in a data file.
- An Extent is a group of 8 contiguous pages.
  - Uniform Extents
  - Mixed Extents

COMP283-Lecture 2

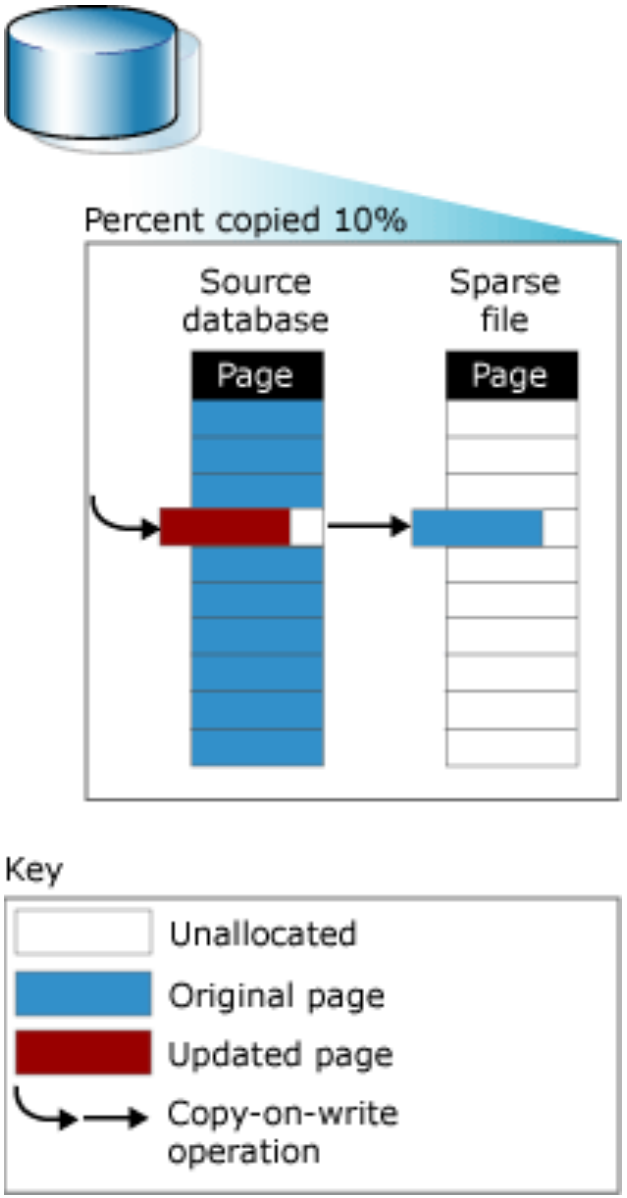
DB Design: Database Snapshot files in MS-SQL

- A database snapshot provides a read-only, static view of a source database as it existed at snapshot creation, minus any uncommitted transactions.
- Database snapshots are dependent on the source database and must be on the same server instance as the database.
- If that database becomes unavailable for any reason, all of its database snapshots also become unavailable.
- Useful - Can revert a database using snapshots.
- But why the limitations?

COMP283-Lecture 2

DB Design: Database Snapshot files in MS-SQL

- Each database snapshot is a Sparse file.
  - Initially an empty file allocated no space.
- Each page of the file is only "filled out" when the original data in that page of the database changes.
  - known as "Copy on Write"
- But why?



DB Design: Database Capacity

- Define a horizon.
- Design the database server with the right storage of current and expected use.
- Design the database with the right sized data files for the current data and expected future use.

COMP283-Lecture 2

Conclusion

- Main database file types.
- Introduced factors to be considered for DB design
- Identified built in data types and the differences between different DBMS.
- Talked about Sparse Files and Copy on Write