

Architettura dei DBMS relazionali

Laboratorio di basi di dati
Dipartimento di Informatica
Università degli Studi di Milano



Schema di riferimento di un DBMS

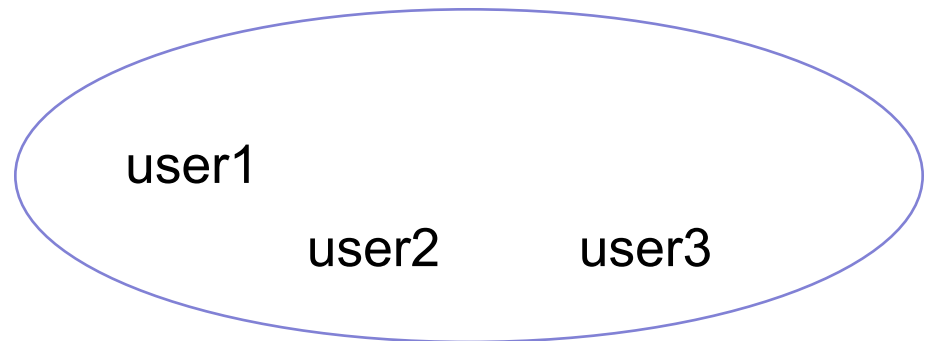
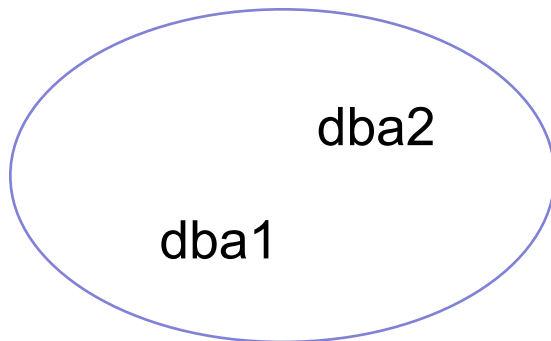
DB Administrator

DB User

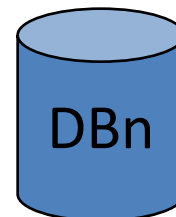
non expert
user

application
developers

occasional
user



DBMS



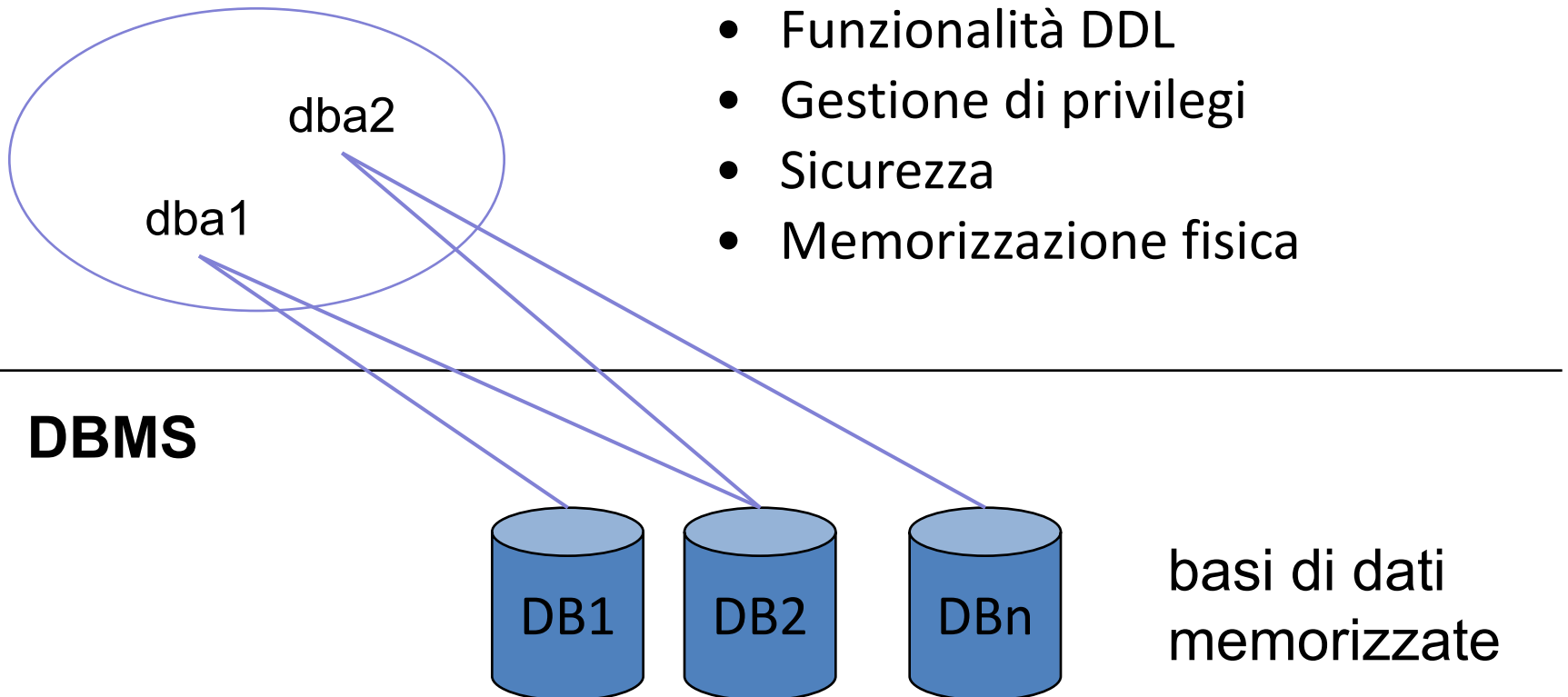
basi di dati
memorizzate

Schema di riferimento di un DBMS

DB Administrator

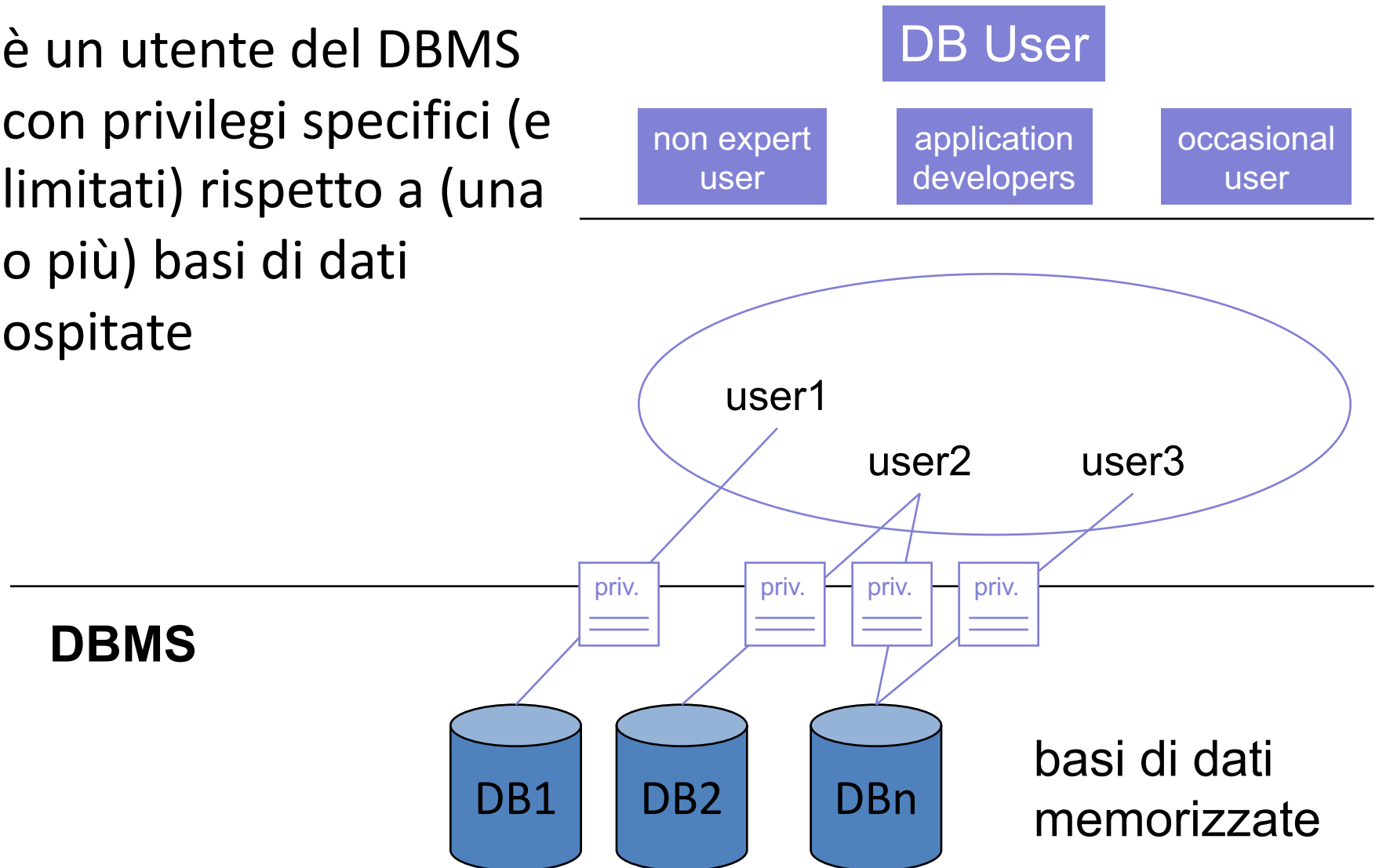
è un utente del DBMS con privilegi amministrativi rispetto a (una o più) basi di dati ospitate

- Funzionalità DDL
- Gestione di privilegi
- Sicurezza
- Memorizzazione fisica



Schema di riferimento di un DBMS

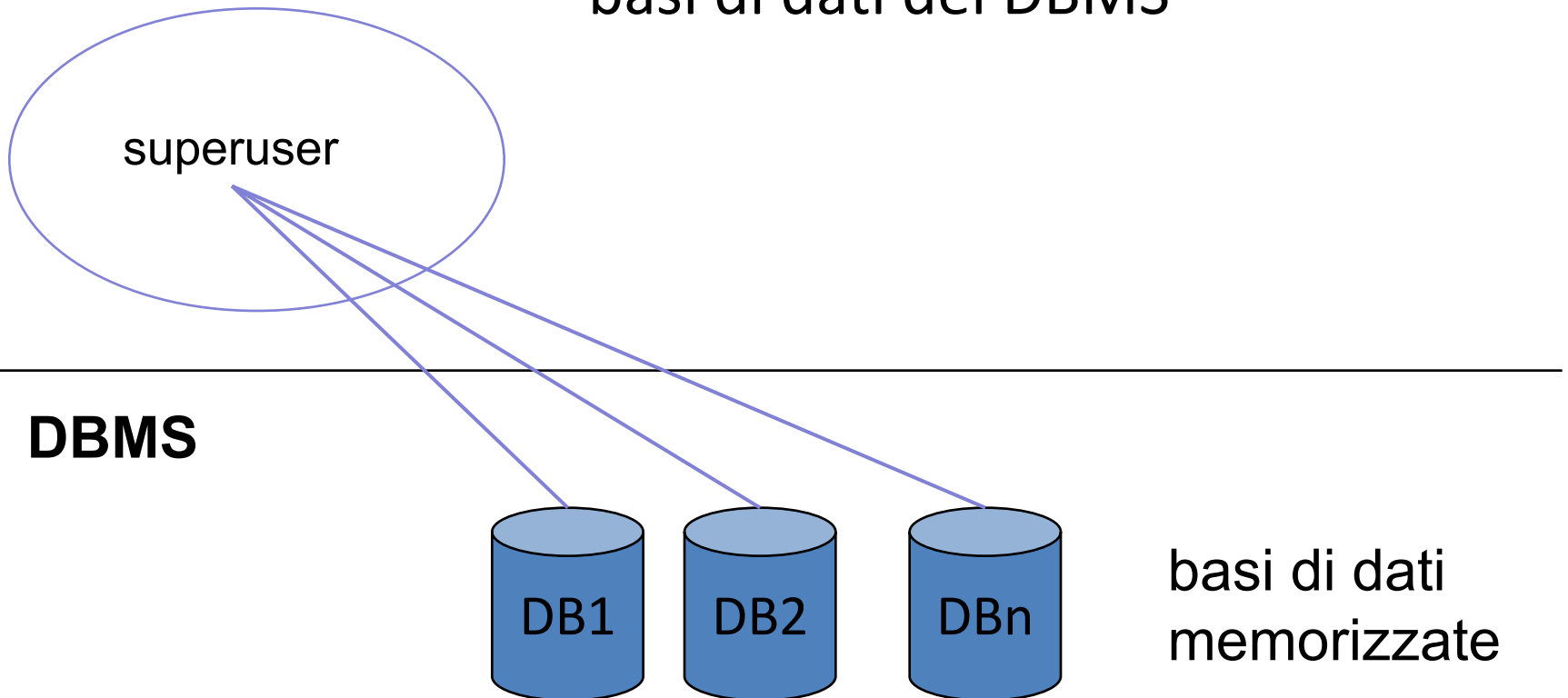
è un utente del DBMS
con privilegi specifici (e
limitati) rispetto a (una
o più) basi di dati
ospitate



Schema di riferimento di un DBMS

DB Administrator

L'utente **superuser** è un DB Administrator con privilegi di amministrazione rispetto a tutte le basi di dati del DBMS



Utenti di un DBMS

- Un DBMS possiede almeno un utente superuser (inizialmente è l'unico utente del DBMS)
- Una base di dati possiede almeno un utente administrator (il proprietario – owner – della base di dati)
- I privilegi di un utente possono variare a seconda della base di dati

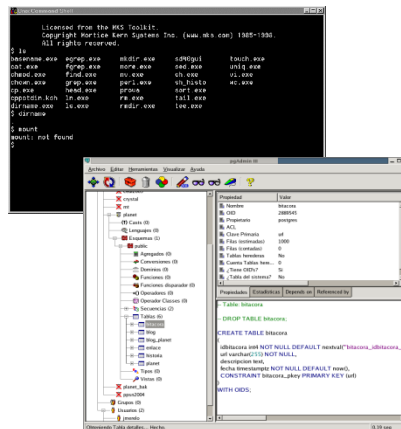
Architettura client/server dei DBMS

LOCAL CLIENT

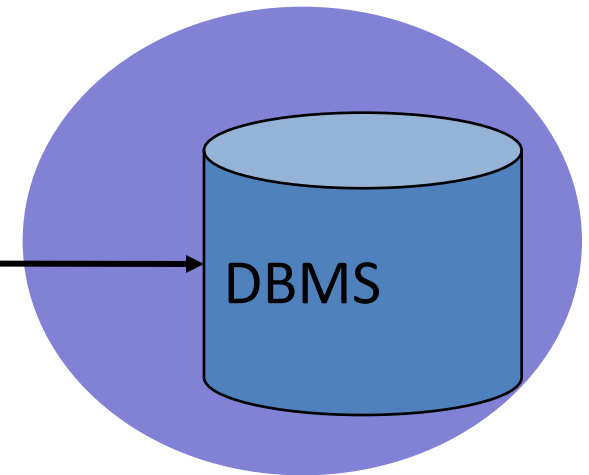
- è un utente del DBMS (DB User/DB Administrator) che si collega mediante pipe
- l'utente è *logged* sullo stesso elaboratore che ospita il DBMS

DBMS SERVER

- è un software in esecuzione su un elaboratore server
- è generalmente implementato come servizio di sistema
- può essere configurato per accettare connessioni locali (pipe) e/o connessioni remote (TCP socket)



pipe di comunicazione



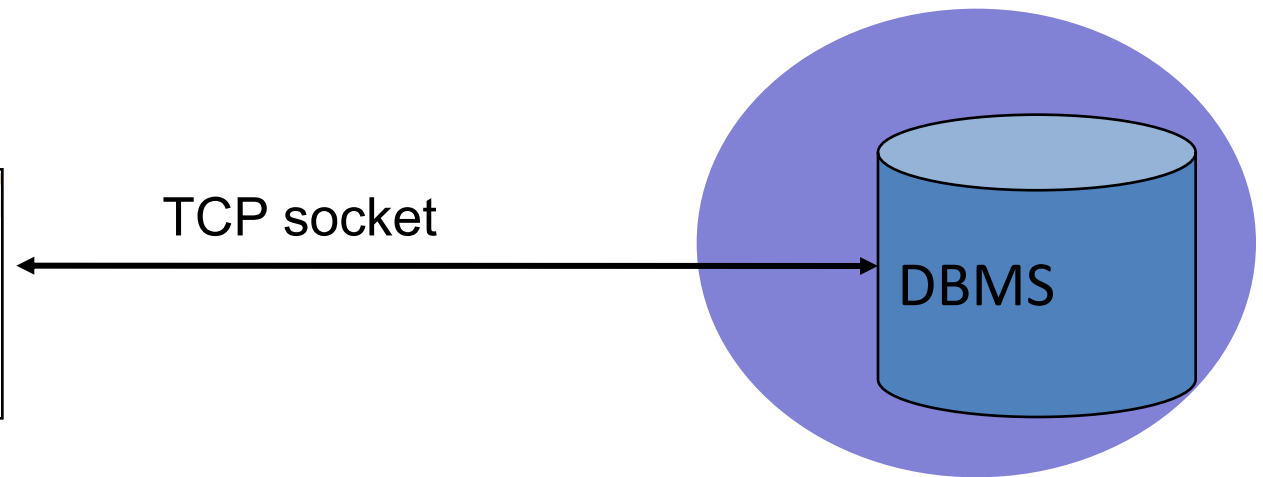
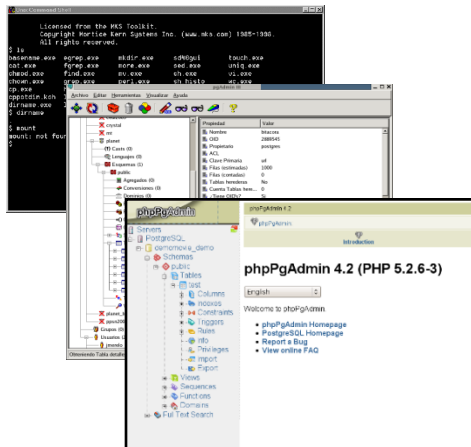
Architettura client/server dei DBMS

REMOTE CLIENT

- è un utente del DBMS (DB User/DB Administrator) che si collega mediante TCP socket
- l'utente è logged su un elaboratore diverso da quello che ospita il DBMS

DBMS SERVER

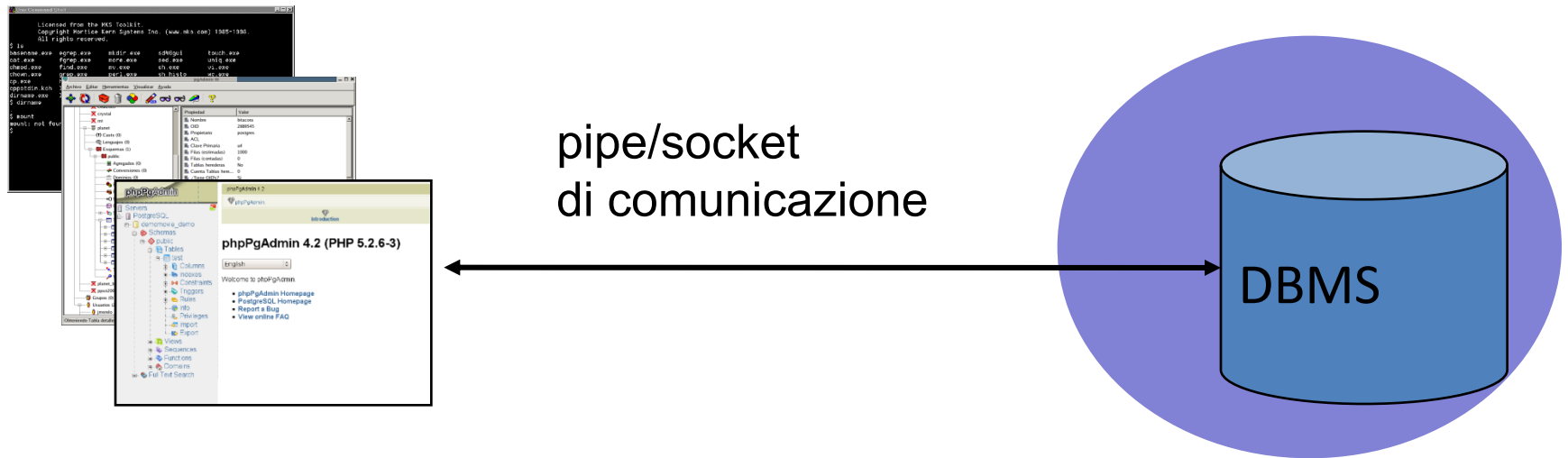
- è un software in esecuzione su un elaboratore server
- è generalmente implementato come servizio di sistema
- può essere configurato per accettare connessioni locali (pipe) e/o connessioni remote (TCP socket)



Architettura client/server dei DBMS

Gli utenti del DBMS (DB Administrator, DB User) utilizzano un software client per comunicare con il server

- client a riga di comando
- client con interfaccia grafica (GUI)
- client con interfaccia web



Utente iniziatore del servizio

- Il DBMS viene avviato da un utente iniziatore che è proprietario del servizio di sistema (è auspicabile che tale utente non sia amministratore del sistema)
- L'utente iniziatore del servizio è un utente di sistema e **NON** è un utente del DBMS
 - Questo significa che l'utente iniziatore non ha accesso alle basi di dati presenti sul DBMS

Scegliere il DBMS

Laboratorio di basi di dati
Dipartimento di Informatica
Università degli Studi di Milano



Scegliere il DMBS

- Il mercato offre un'ampia gamma di soluzioni
- Non esiste un prodotto migliore in assoluto
- La scelta deve risultare da una valutazione costi/benefici
 - anche in termini di complessità della soluzione rispetto all'ambito applicativo
- Tecnologia privilegiata al corso: PostgreSQL

DBMS Commerciali

- DBMS commerciali
 - MS SQL Server
 - Oracle
 - DB2
 - MySQL
- DBMS open-source like
 - MySQL
 - MariaDB
 - PostgreSQL
 - HSQLDB

MS SQL Server ⁽¹⁾

- DBMS relazionale sviluppato da Microsoft
- Sistemi operativi: MS Windows, Windows Server, Linux
- Cenni storici
 - 1989: prima versione per OS/2, base code acquistato da Sybase SQL Server (competizione con Oracle e IBM)
 - 2000: SQL Server 2000, con performance migliori, IDE tools lato client, e diversi sistemi complementari (e.g., supporto per il data mining e OLAP)
 - 2005: SQL Server 2005, con supporto nativo per XML e database mirroring
 - 2008: SQL Server 2008, con full-text search
 - 2014: SQL Server 2014, con supporto per Hekaton (in-memory database)
 - 2022: SQL Server 2022

MS SQL Server (2)

Diverse edizioni offerte



Express/web

Free, entry-level database for small web and mobile apps

Feature highlights

- Up to 16 cores of CPU
- Up to 64 GBs of memory
- In-memory OLTP and Columnstore
- End-to-end encryption with secure enclaves
- Full T-SQL surface area
- Support for Linux and Windows containers
- UTF-8 character encoding
- Data classification and auditing



Standard

Full featured database with Big Data Clusters for mid-tier applications and data marts

Feature highlights

- Up to 24 cores of CPU
- Up to 128 GBs of memory
- SQL Server 2019 Big Data Clusters
- Data virtualization through PolyBase
- Enhanced in-memory performance
- Automatic intelligent database tuning
- Azure Data Studio with notebook support

+ Express/web features



Enterprise

Mission-critical performance and intelligence for tier 1 databases

Feature highlights

- Unlimited cores of CPU
- Unlimited memory
- Industry-leading performance with unmatched scalability
- Unlimited virtualization benefits
- Petabyte scale data warehousing
- Business critical HA on Windows and Linux
- Low TCO with free DR replicas
- Access to Power BI Report Server

+ Standard features

+ Express/web features



Developer

Free to use with all the features of Enterprise Edition specifically for dev/test in non-production environments

<https://www.microsoft.com/it-it/sql-server/sql-server-2019-comparison>

Oracle ⁽¹⁾

- ORDBMS (DBMS relazionale e ad oggetti) sviluppato da Oracle Corporation
- Compatibile con diversi sistemi operativi
- Cenni storici
 - 1979: Oracle V2, supporto per funzionalità base di SQL
 - 1985: Oracle V5, supporto per interazioni client-server
 - 1992: Oracle V7, supporto per stored procedure e trigger
 - 1999: Oracle 8i, orientato alle interazioni tramite Internet
 - 2003: Oracle 10g, supporto per grid computing
 - 2013: Oracle 12c, supporto per cloud computing
 - 2023: Oracle 23c

Oracle ⁽²⁾

- Disponibile in diverse edizioni, fra cui:
 - Express Edition: versione entry level e limitata a 4GB di dati utente, 1GB di memoria, 1 CPU
 - Standard Edition One
 - Standard Edition: funzionalità di SE One + supporto per macchine più grandi e clustering di servizi
 - Enterprise Edition: tutte le funzionalità di Oracle Database + supporto per applicazioni di online transaction processing (OLTP) ad alto volume, query-intensive data warehouses, applicazioni Internet onerose
 - Personal Edition: quasi tutte le funzionalità di EE, limitato a single-user development

DB2 ⁽¹⁾

- DBMS relazionale sviluppato da IBM
- Compatibile con diversi sistemi operativi (fino al 1990 solo mainframe IBM)
- Cenni storici
 - 1983: prima versione di DB2, basata su un prodotto precedente di IBM, QBE
 - 2006: DB2 9 che supporta XML
 - 2007: DB2 9.5 orientato ad ambienti distribuiti
 - 2009: DB2 9.7, supporto alla frammentazione dei dati
 - 2013: DB2 10.5, migliorie di performance e funzionalità mobile
 - DB2 13, ulteriori migliorie (e.g., supporto per AI)

DB2 ⁽²⁾

- Disponibile in diverse edizioni, fra cui:
 - Workgroup Server Edition (WSE): limiti su processore and memoria
 - Enterprise Server Edition (ESE): funzionalità di WSE + materialized query tables, no limiti su processori, memorie, o dimensioni dei dati
 - Advanced Enterprise Server Edition (AESE): funzionalità di ESE + column organized tables, in-memory database, data compression, workload management, replication, ...
 - Advanced Workgroup Server Edition (AWSE): simile ad AESE, ma con limiti su processori e memoria
 - Developer Edition: supporto per un unico application developer

MySQL ⁽¹⁾

- DBMS relazionale
 - sviluppato in origine da MySQL AB, ora parte di Oracle
 - disponibile con licenza GNU General Public License (alcune edizioni a pagamento per scopi commerciali)
- Compatibile con diversi sistemi operativi, molto diffuso per lo sviluppo di Web Application
- Cenni storici
 - Sviluppato a partire dal 1995 da Michael Widenius e David Axmark
 - 1998: prima versione per Windows
 - 2005: MySQL V5, supporto a viste e trigger
 - 2008: MySQL V5.1, supporto a frammentazione
 - Versione corrente: MySQL 8

MySQL ⁽²⁾

- Disponibile in diverse edizioni, fra cui:
 - MySQL Standard Edition
 - MySQL Enterprise Edition: funzionalità standard +altre (e.g., sicurezza, monitor di performance)
 - MySQL Cluster Carrier Grade Edition (CGE): funzionalità enterprise +altre (e.g., storage engine addizionali, partizionamento automatico dei dati (sharding) con bilanciamento del carico)
 - MySQL Community Edition: licenza GNU

MariaDB

- "Cugino" molto stretto di MySQL
- DBMS nato nel 2009 come fork di MySQL
 - in seguito all'acquisizione da parte di Oracle
 - idea: garantire SW free e open source con licenza GNU
- Il versioning ha seguito il versioning di MySQL fino alla versione 5.5, successivamente nasce un versioning separato
 - nel 2012 si passa dalla versione 5.5 alla 10
 - versione corrente: 11

PostgreSQL

- ORDBMS (relazionale e ad oggetti) rilasciato con licenza PostgreSQL
 - gratuito ed open-source
- Compatibile con diversi sistemi operativi
- Cenni storici
 - Nasce da un progetto condotto all'università della California@Berkley (Ingres project)
 - 1988: viene rilasciato il primo prototipo
 - 1995: primo rilascio di Postgres95
 - 2000: supporto per chiavi esterne e sintassi SQL92 per i join
 - 2010-oggi (v16): miglioramenti a gestione di indici, trigger e viste, viste materializzate, meccanismi di lock avanzati, supporto JSON, ...

II DBMS PostgreSQL

Laboratorio di basi di dati
Dipartimento di Informatica
Università degli Studi di Milano



PostgreSQL

- Disponibile Open Source con licenza BSD
- Disponibile per Windows / Linux / Mac OSX / Solaris / ...
- Estensione proprietaria a SQL procedurale (PLpgSQL)
- Dispone di interfaccia grafica
 - phpPgAdmin (web-based)
 - pgAdmin IV
 - DataGrip
- Dispone di interfaccia a linea di comando
 - psql

PostgreSQL



Installazione di PostgreSQL

- Possibilità offerte per l'installazione:
 - Compilazione dei sorgenti
 - Utilizzo di una distribuzione binaria
- PostgreSQL può essere installato senza possedere i privilegi di root (in una porzione del FileSystem accessibile con privilegi RW all'utente che esegue l'installazione)
- Installare la versione stabile più recente

Configurazione di PostgreSQL

- Le principali impostazioni di configurazione del DBMS sono contenute nel file **postgresql.conf** memorizzato nella cartella di installazione di PostgreSQL
- In particolare, in postgresql.conf si trovano impostazioni su
 - connessioni (e.g., porta di ascolto e numero massimo di connessioni concorrenti)
 - utilizzo di memoria
 - log di operazioni/errori
 - gestione dei lock (uso di transazioni)

Template di PostgreSQL

- Dopo l'installazione, su PostgreSQL sono presenti tre database (**template0**, **template1**, **postgres**) che sono i modelli a partire dai quali vengono generati nuovi database
 - template0 è un database di sistema non accessibile agli utenti
 - template1 è il database contenente il modello predefinito
 - postgres è una copia di template1
- Utilizzando template1, è possibile creare un proprio template come modello per la generazione di nuovi database

Connessione mediante psql

- Connessione al DBMS mediante il client a riga di comando psql:
 - E' necessario specificare il nome dell'utente, l'indirizzo del server e il database a cui connettersi:

psql -U <nomeutente> -h <indirizzoserver> <nomedb>

Esempio: `psql -U postgres -h localhost template1`

Comandi di psql

- Help in linea
 - \?[comando]: help sui comandi del client
 - \h[comando]: help su sql
- Principali comandi di psql:

\l	lista dei database presenti sul DBMS
\c <nome_db>	connessione al database specificato
\dt	elenco tabelle del database a cui si è connessi
\dn	elenco degli schemi del database a cui si è connessi
\du	lista utenti del DBMS
\d <nome_obj>	descrizione della struttura dell'oggetto specificato
\q	chiusura della connessione e uscita da psql