

INAF-Osservatorio Astronomico di Roma

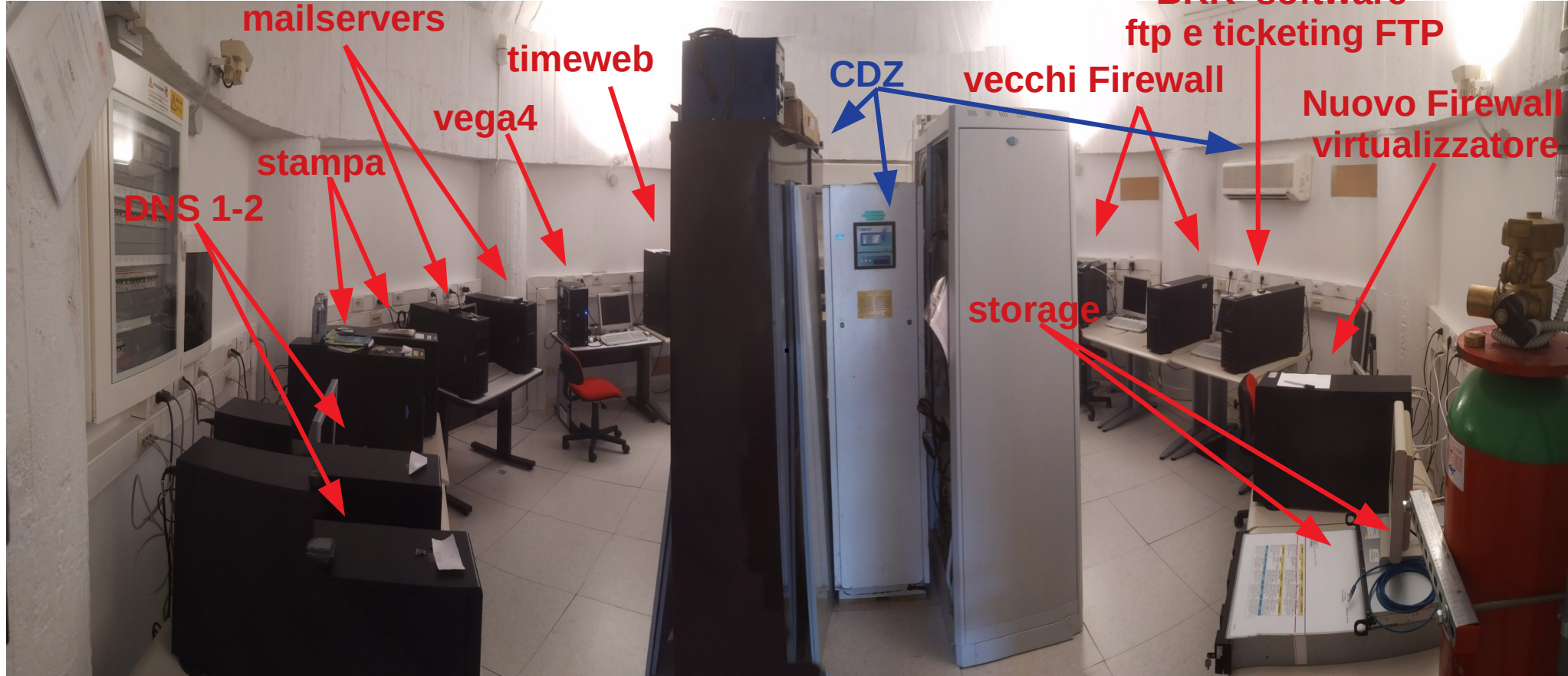


Esperienze e Know How Informatico acquisito dallo staff dell'OAR durante la collaborazione nei progetti scientifici nazionali ed internazionali

Author: Stefano Gallozzi

- Si Parlerà di:
- **Esperienze e Capacità nella gestione dei progetti di ricerca**
- **Soluzioni ai Problemi più importanti incontrati**
- **Know-How acquisito:**
 - ✓ **Virtualizzazione per servizi ridondati**
 - ✓ **Machine Learning ed uso GPUs**
 - ✓ **Trasferimento dati ed ottimizzazione rete**
 - ✓ **Archivi Astronomici di BigData**
 - ✓ **Scegliere il Database “giusto”**

Vecchio CED di Palazzo (x Servizi)

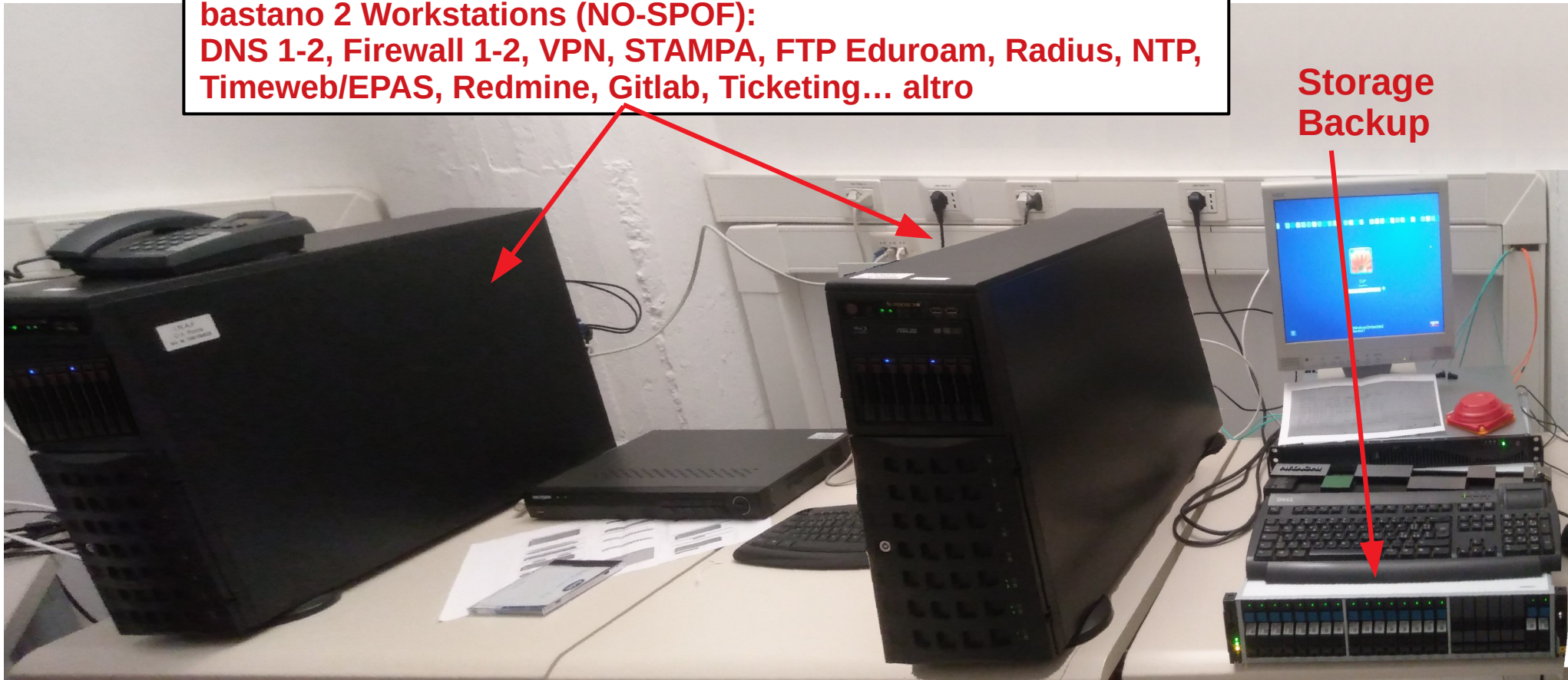


Nuovo CED di Palazzo (x Servizi)

SERVIZI VIRTUALIZZATI E RIDONDANTI

bastano 2 Workstations (NO-SPOF):

DNS 1-2, Firewall 1-2, VPN, STAMPA, FTP Eduroam, Radius, NTP, Timeweb/EPAS, Redmine, Gitlab, Ticketing... altro



**Storage
Backup**

CED Ricerca

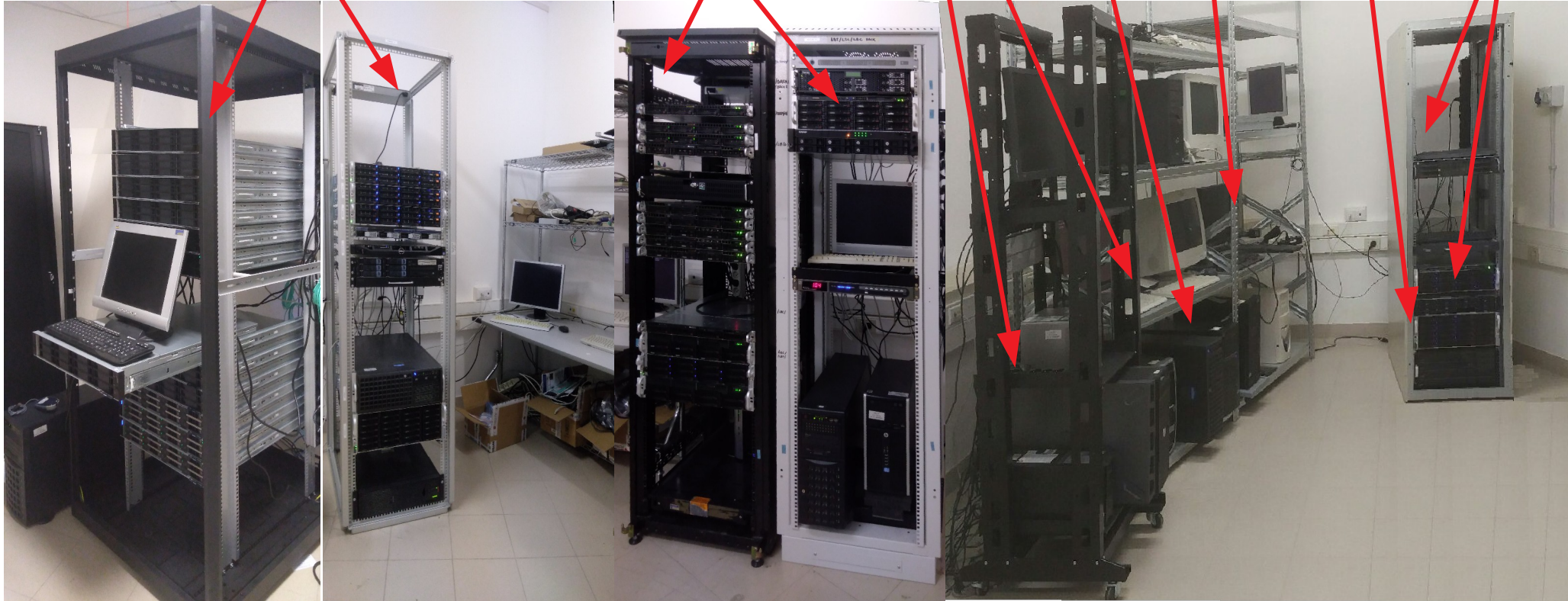
ASTRI Miniarray & CTA

EXTRA GALATTICI - JWST
LBT & LSC + EUCLID

DIVA

KEPLER
STELLARI

EST/IBIS
SOLARI
GRAVITOWN



CED Ricerca

ASTRI Miniarray & CTA


EXTRA GALATTICI - JWST DIVA

KEPLER EST/IBIS SOLARI

LBT & LSC + EUCLID TEORICI

STELLARI

GRAVITOWN



Gallozzi S.
Fiordoliva F.
Lucarelli F.
Lombardi S.
Lamastra A.
Visconti F.
Mastropietro M.
Saturni F.G.
Bigongiari C.
Antonelli L.A.



Paris D.
Gallozzi S.
Castellano M.
Merlin E.
Santini P.
Bonchi A.
Zappacosta L.



Valiante R.
Schneider R.
Menci N.



Israrel G.
Rodriguez G.
Nicastro F.

Pulone L.
Limongi M.
D'Alessio F.
Leoni R.
Ventura P.

Giorgi F.
Viavattene G.
Ermolli I.
Piazzesi R.
Biazzo K.
Brocato E.

CED Ricerca

**Cold Storage
(TAPE LIBRARY)**



**High Performance Computing
N. 2 Rack KLN GALILEO (ex-Cineca)**



Hardware & Software Facilities for high Throughput / Performance Computing

Boards:

- CARMA development-KIT
- NVIDIA K20 GPU
- NVIDIA RTX 3050
- n.2 NVIDIA Jetson TX1
- n.2 NVIDIA RTX 3090 Ti



Softwares:

Esperienza maturata nella scrittura di software dedicato a tasks di classificazione e regressione, principalmente nel campo della fisica delle Altissime Energie.

Acquisite competenze nell'uso di due frameworks open source:

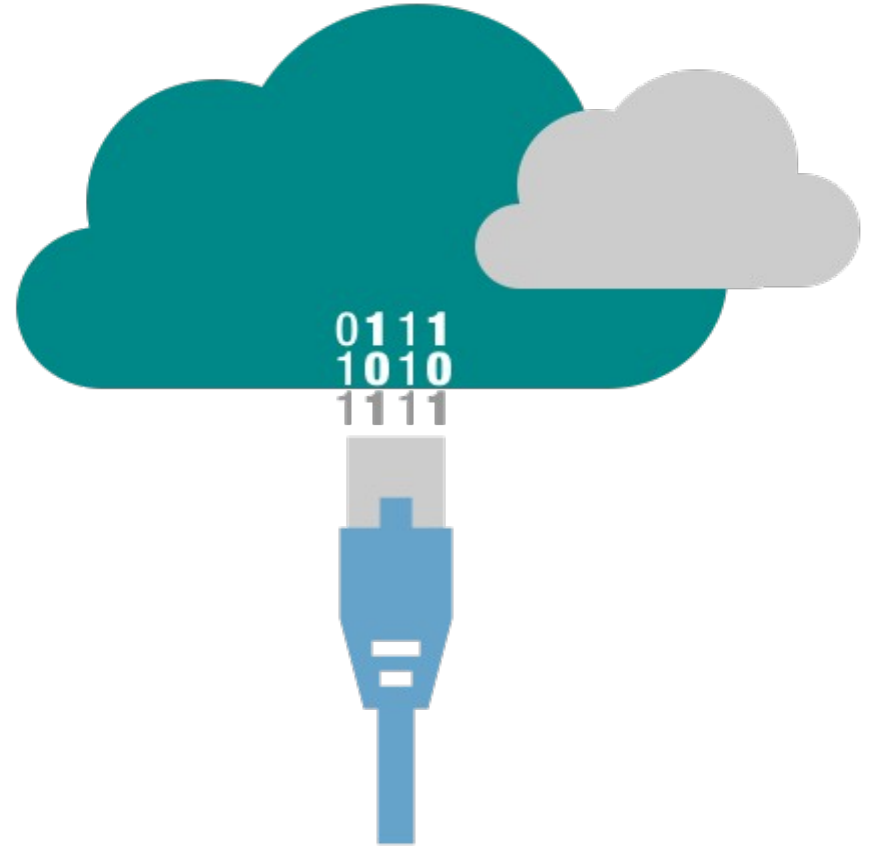
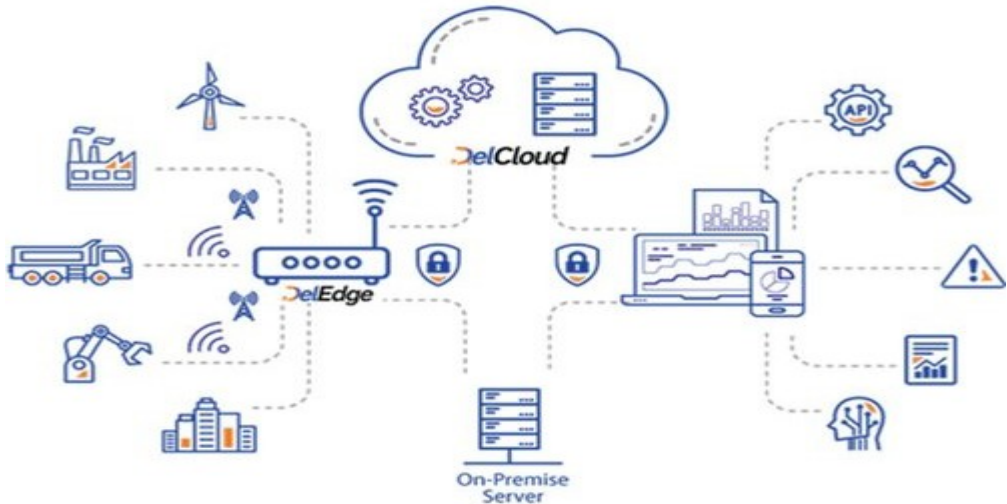
- Tensorflow
- PyTorch

Anche se i benchmarks sono complessi, in media, si riesce ad incrementare la velocità di espletamento dei tasks di classificazione con le tecniche del DEEP LEARNING (reti neurali di autoapprendimento): con un auto-apprendimento di poco più di 5 minuti si riescono a processare quasi 200k immagini su una RTX 3050.

■ In fase di acquisto

Trasferimento Dati ed Ottimizzazione Risorse

Grandi Progetti come LBT ed ASTRI necessitano di trasferimento dati continuo ed efficiente
NON BASTA ATTACCARE IL CAVO DI RETE PER TRASFERIRE DATI



Trasferimento Dati ed Ottimizzazione Risorse

- colli di bottiglia e performance rete:
 - agire point-to-point a ritroso
- modificare l'MTU su ogni switch della tratta (1500 → 9000)
- scegliere filesystem più adatto
- scegliere dischi e/o raid più adatti
- scegliere i sftw di trasferimento (rsync/scp/sftp/ftp/gridFTP,aria2)
- ottimizzare kernel e servers (apache)
- mettere regole ACL dirette su routers se possibile bypassare i firewall (P2P)



(da 5-8 volte più veloce con configurazioni ottimizzate)

con connessione ~10Gbps si riesce a saturare l'IO di scrittura dei dischi

Big Data ed Archivi Astronomici

Se stimiamo la quantità di dati da archiviare e rendere accessibile ai vari “utenti”:

{ PRESA DATI **LBT** ~ 15-30 GByte/Notte
PRESA DATI **ASTRI/MA** ~ 0.5-5.0 TByte/Notte
PRESA DATI **CTA** ~ 5.0-50 TByte/Notte

→ si incontra un technological challenge da superare
a medio/lungo termine che comprende anche l'infrastruttura ICT collegata (espandibilità)



Scale-up _vs_ Scale-out centralizzato vs distribuito

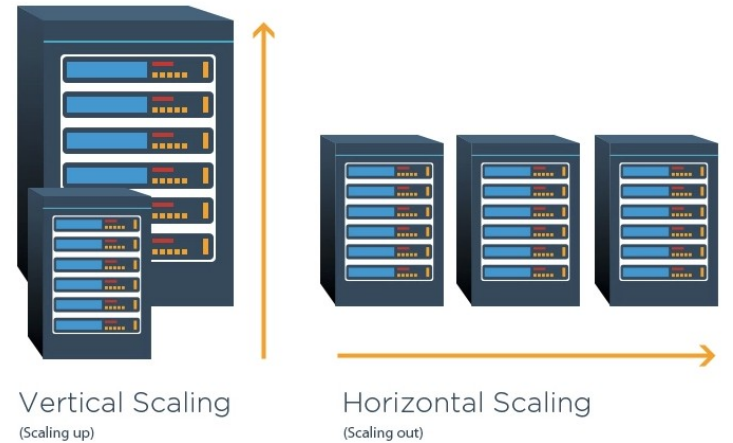
Ad un certo punto ci si scontrerà sempre su come gestire lo “scaling” come espandere un sistema in uso

SCALE-UP ~ sostituire l'hardware con altro di migliori performances (maggiore potenza/capacità)

SCALE-OUT ~ aggiungere nodi di calcolo/storage con le stesse performances (stessa potenza/capacità)

Quando un progetto prevede scalabilità di una qualunque risorsa nel corso degli anni l'unica scelta possibile è lo SCALE-OUT

- si mantengono le prestazioni originarie
- si aggiunge ridondanza
- si rende il sistema no-SPOF
- costa enormemente di meno



Scale Up



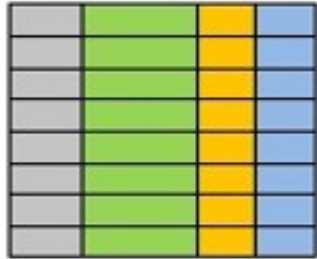
Scale Out



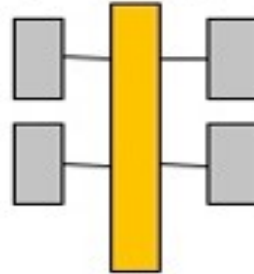
Databases: le scelte possibili

Fixed-Schema Relational DBs

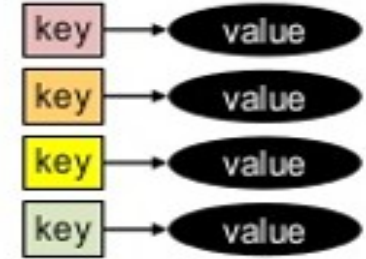
Relational



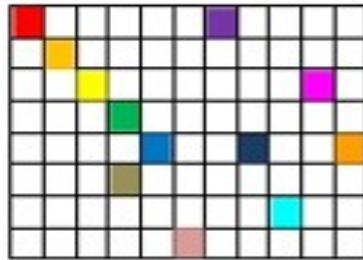
Analytical (OLAP)



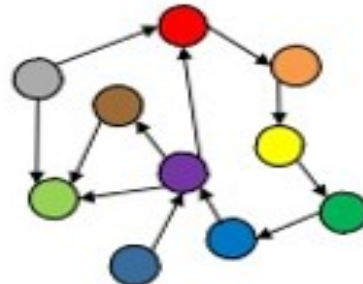
Key-Value



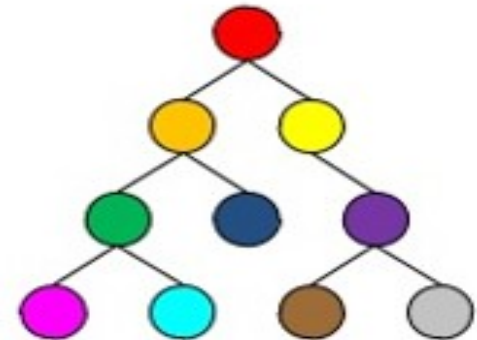
Column-Family



Graph



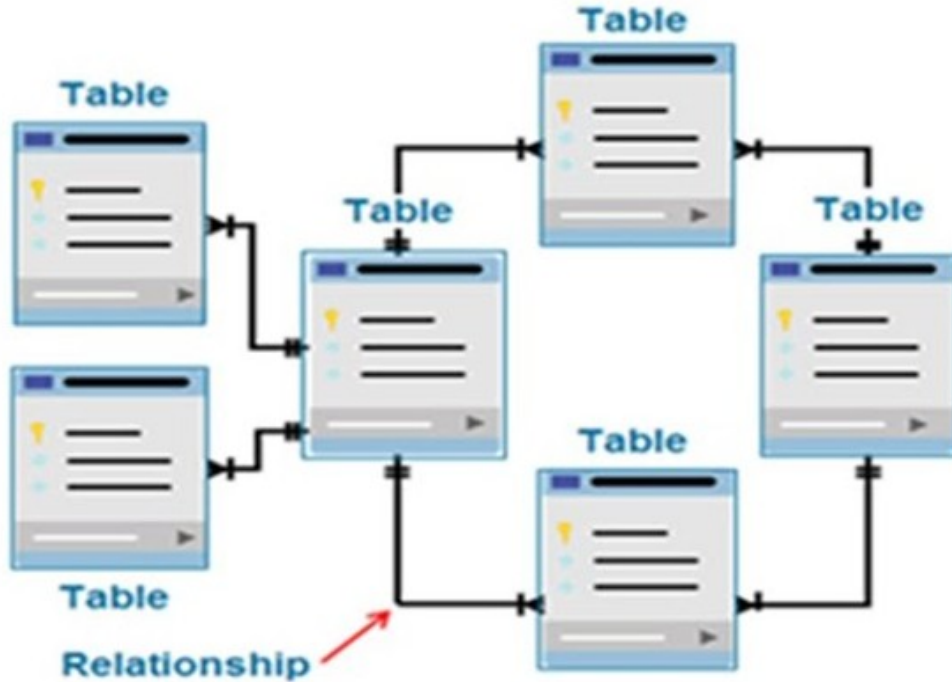
Document



Schemaless – Not Only Relational

Relational Dbs (A.C.I.D. transactions)

Atomicità, Consistenza, Isolazione e Durabilità

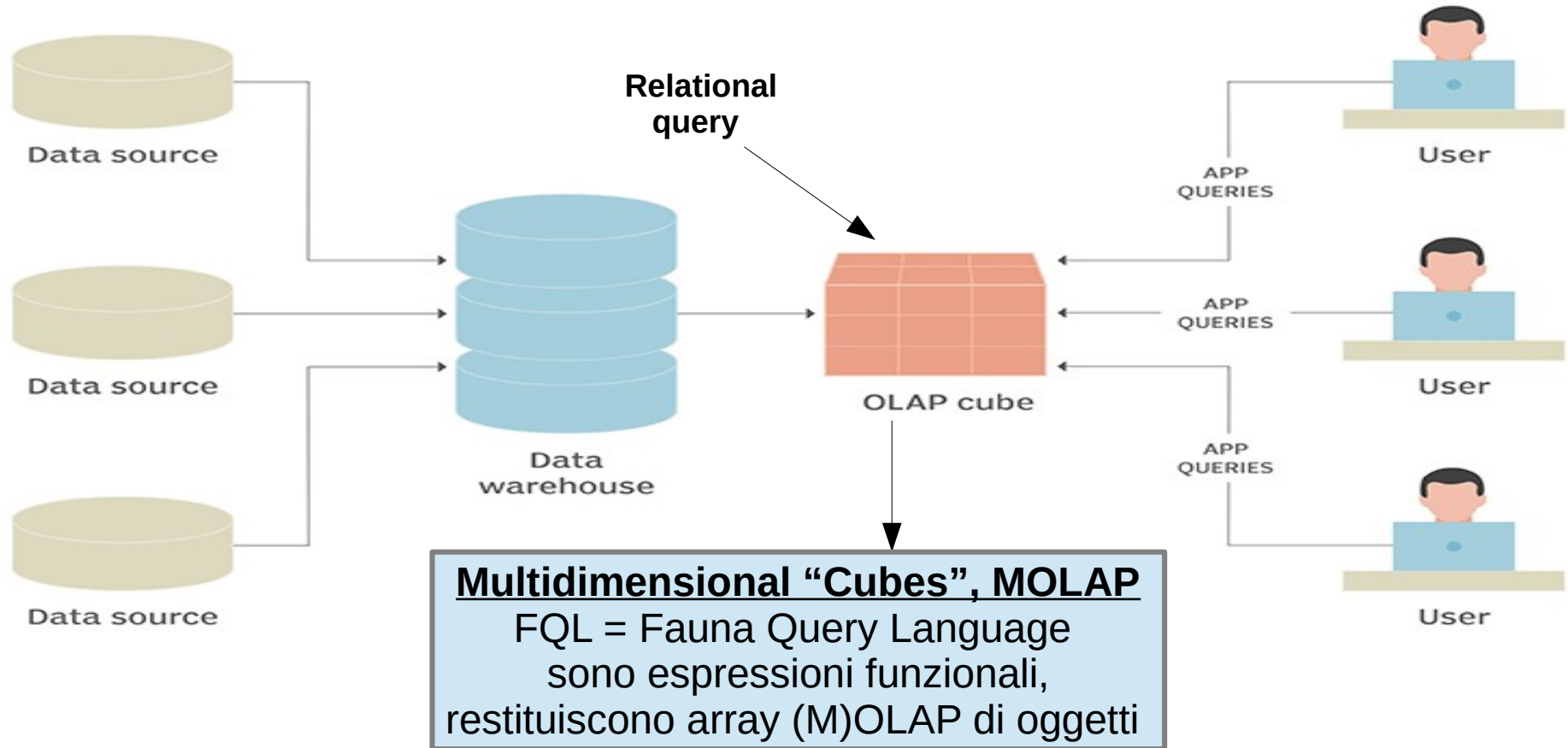


RDBMS

Relational Databases

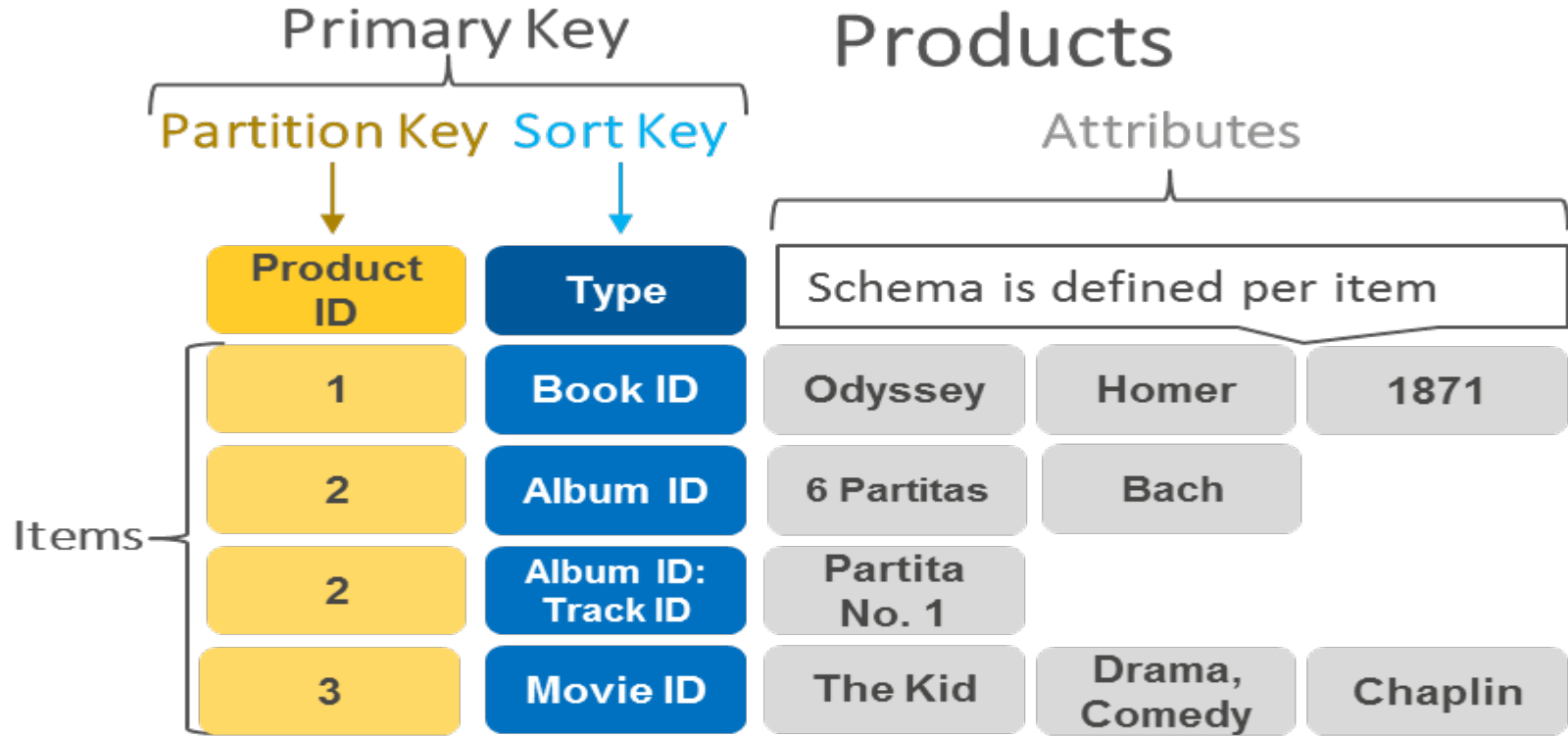
diverse tabelle messe tra loro in relazione attraverso chiavi e/o tabelle intermedie

OnLineAnalyticalProcessing (OLAP) DBs



Si estraggono selettivamente al volo alcuni set di dati per vedere dei trend a più alto livello

Key-Value DBs



I dati immagazzinati in formato chiave-valore, che possono essere oggetti semplici o articolati

Columns DBs

Column View

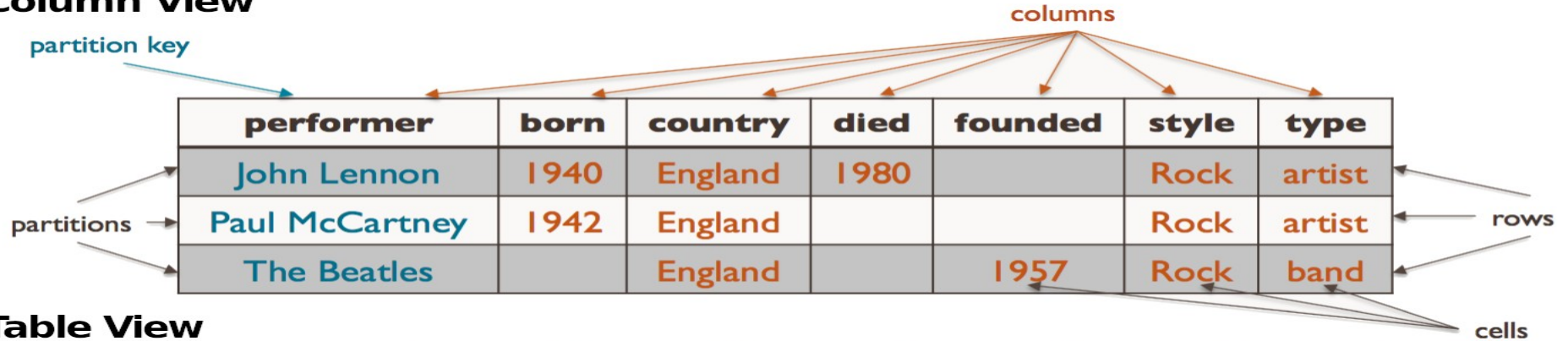
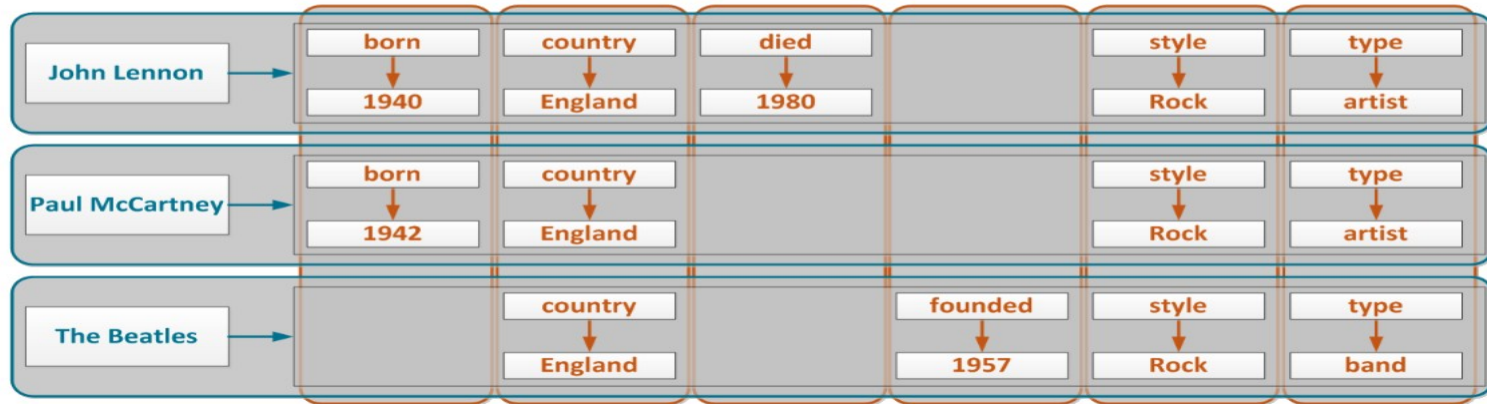
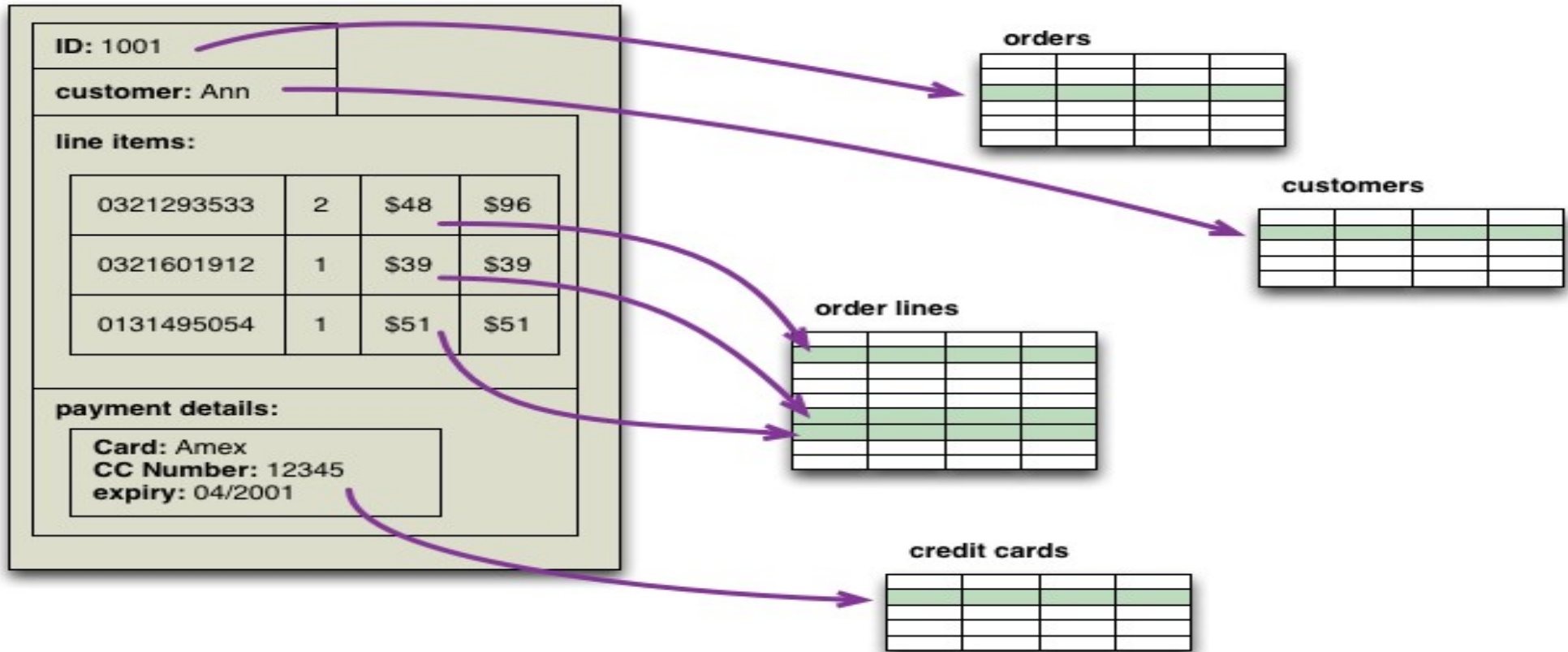


Table View



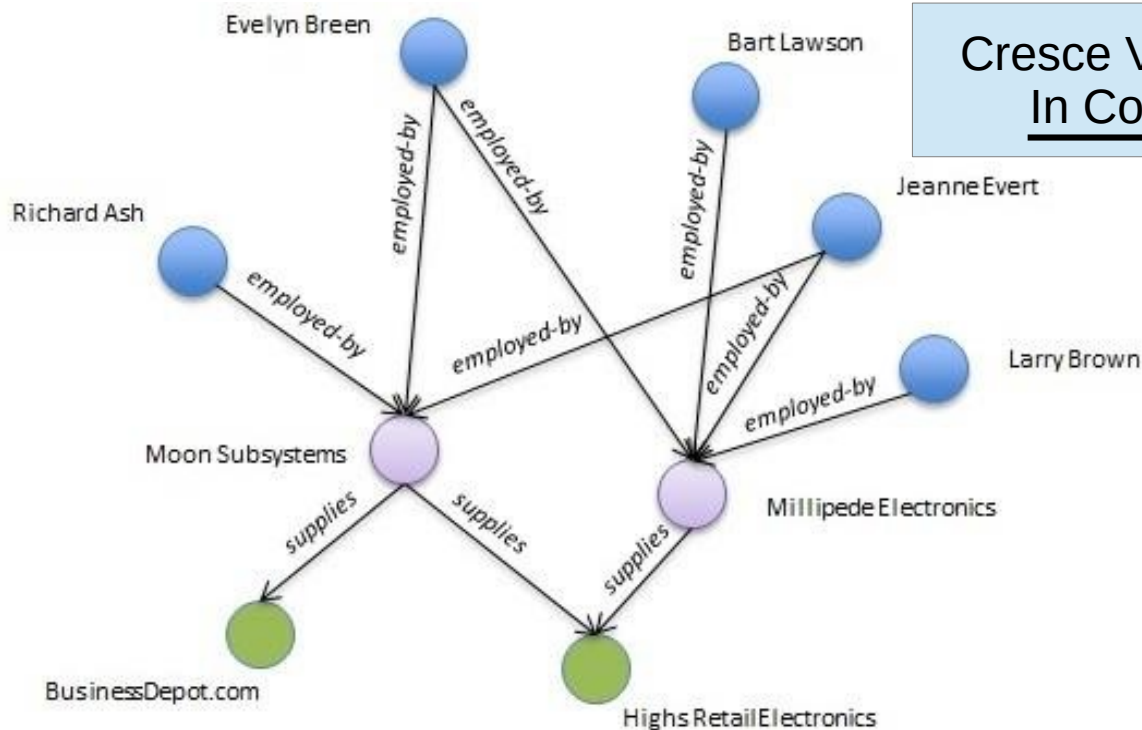
Grandi tabelle di dati (anche nulli) con le colonne trattate separatamente ed aggregate al bisogno

Document DBs



Tutti dati tra loro coerenti messi insieme in "documenti" raggruppati in collezioni di "documenti"

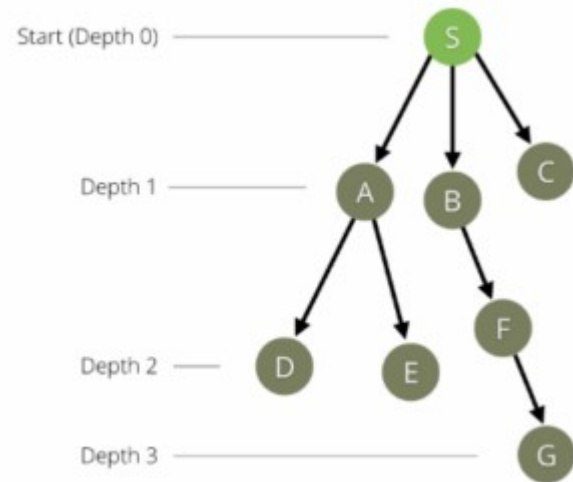
Graph DBs



Cresce Velocemente
In Complessità



DAG, Directed Acyclic Graph
(per workload Management Sys)



Strutture di “nodi” legati assieme da frecce che sono le “relazioni” e le “operazioni” tra nodi

Scalabilità dei DBs

Relational DBMS

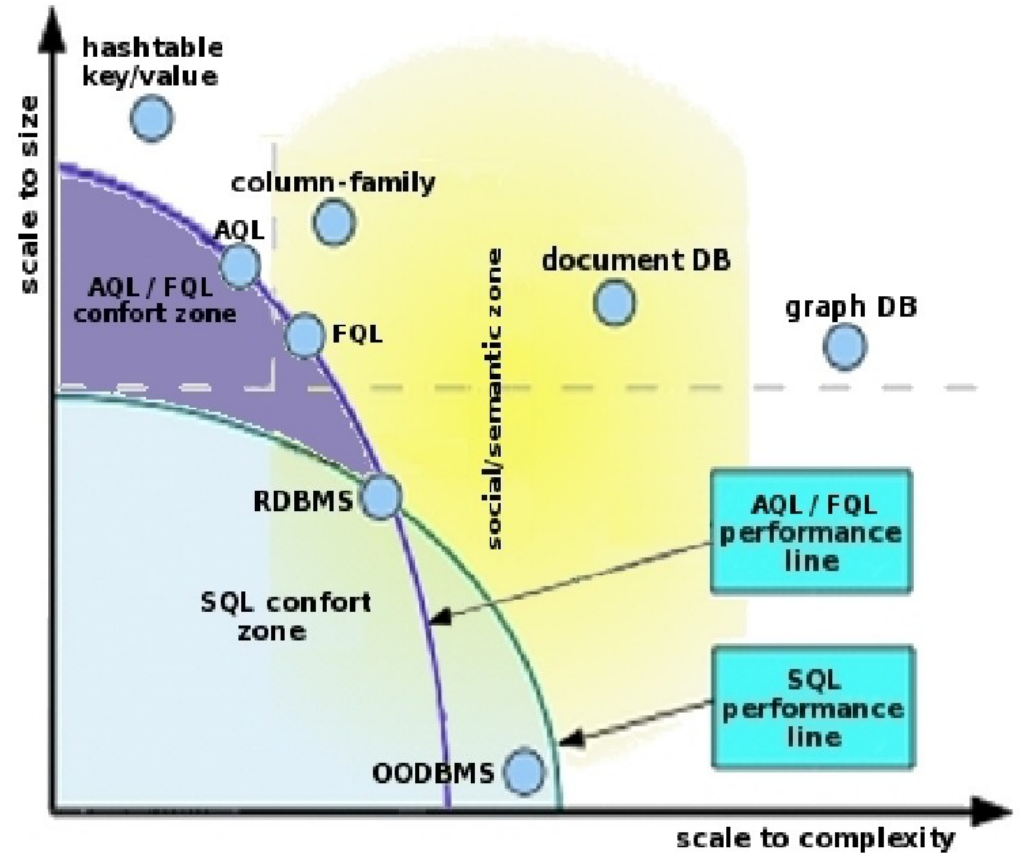
VS

NO-SQL Approach

Balance if needed a fixed-schema Relational DB or a versatile approach of a schema-less NO-SQL

BUT
keep in mind the

C.A.P. Theorem



Distributed – Social / Semantic DBs the CAP Theorem

Consistency

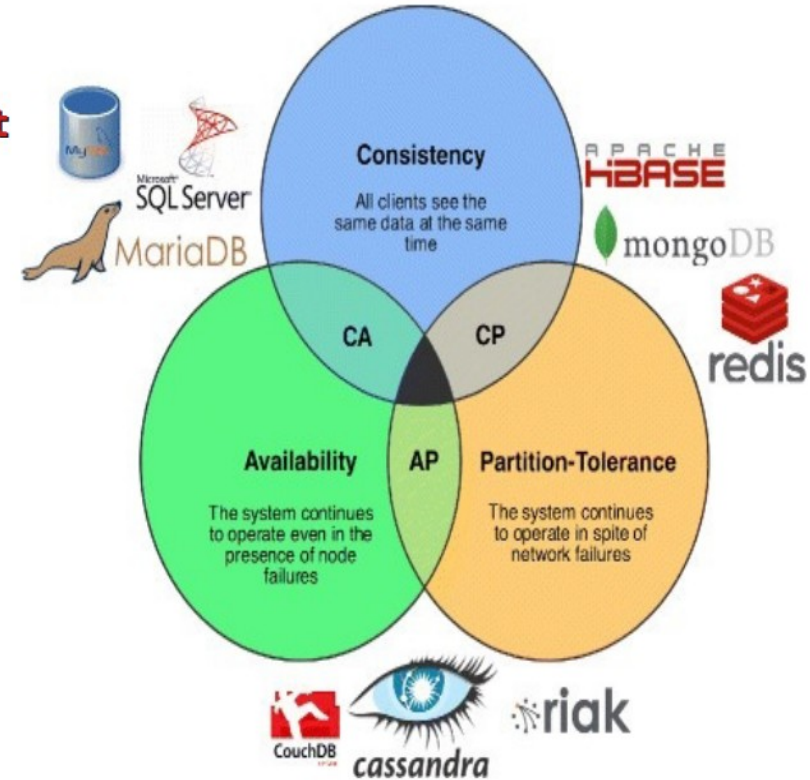
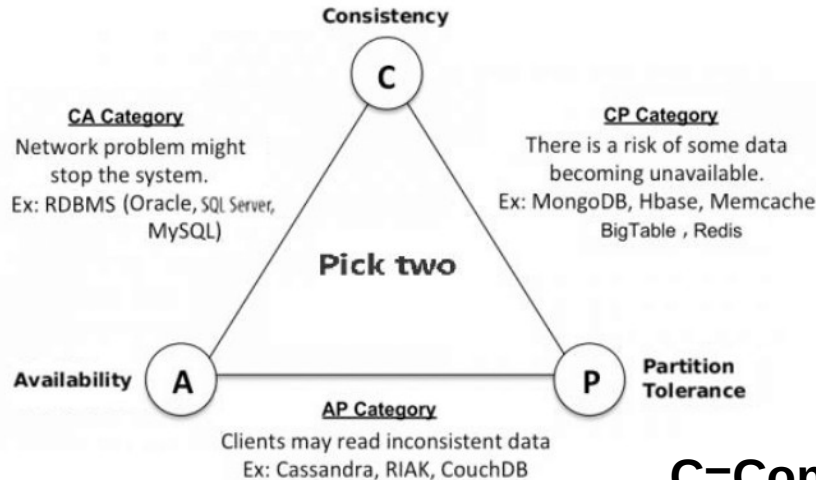
Every read receives the most recent write or an error.

Availability

Every request receives a (non-error) response, without the guarantee that it contains the most recent write.

Partition tolerance

The system continues to operate despite messages being dropped by the network between nodes.



C=Consistency is guaranteed by WORM paradigm!

DOMANDE?

