

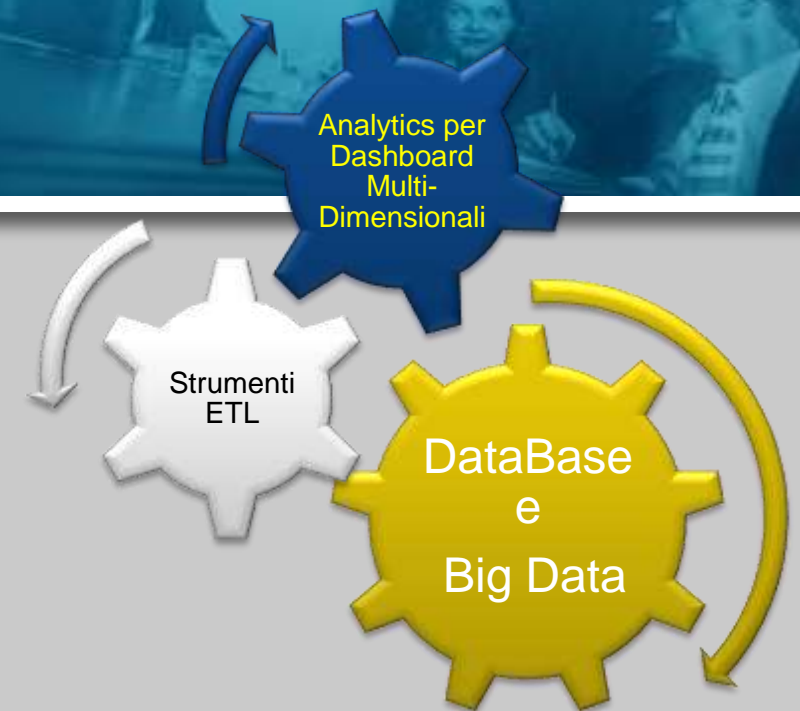
Best Solutions

Le migliori soluzioni tecnologiche in ambito informatico

**DESCRIZIONE EDWH
COMPETENZE, APPROCCIO**

Società consulenza che realizza progetti informatici su tematiche Business Intelligence, Data Warehouse e Big Data

Specializzata su architetture di gestione e controllo dei progetti di integrazione dati





Agenda

- Enterprise Data Warehouse
- Esperienza ed Approccio BS
- Architetture di Controllo





Enterprise Data Warehouse Componenti principali / Vantaggi

Un Enterprise Data Warehouse (EDWH) è una soluzione centralizzata per l'archiviazione, la gestione e l'analisi dei dati aziendali. Esso integra dati provenienti da varie fonti, fornendo un'unica fonte di verità che supporta il processo decisionale strategico.

Componenti Principali di un EDWH

1. Sorgenti di Dati (Data Sources)
2. Processi ETL (Extract, Transform, Load)
3. Data Warehouse
4. Data Marts
5. Strumenti di Business Intelligence (BI) e Analisi
6. Metadata e Data Governance

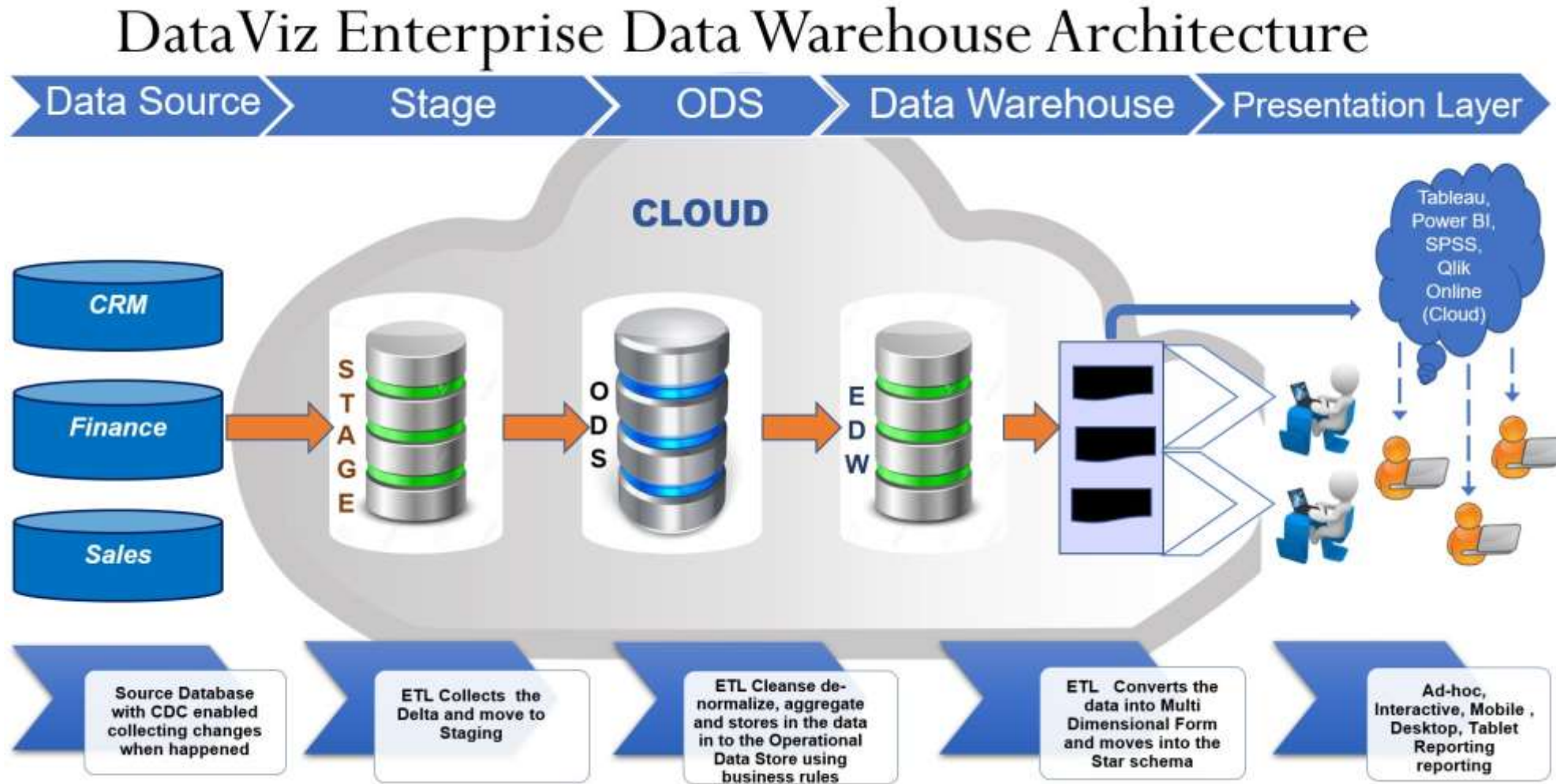
Vantaggi di un EDWH

1. Unica Fonte di Verità
2. Supporto alle Decisioni
3. Miglioramento dell'Efficienza
4. Analisi Avanzate
5. Scalabilità





Enterprise Data Warehouse Schema Tipico





Enterprise Data Warehouse Concetti Chiave

- 1. Data Integration:** Unificare dati da diverse fonti in un'unica piattaforma.
- 2. Data Quality:** Assicurare che i dati siano accurati, completi, consistenti, tempestivi, validi e integri.
- 3. Scalabilità:** Progettare l'EDWH per poter gestire un aumento del volume dei dati e delle richieste.
- 4. Data Governance:** Implementare politiche e procedure per la gestione dei dati, inclusi i diritti di accesso e la sicurezza dei dati.
- 5. Performance Optimization:** Utilizzare tecniche come l'indicizzazione, la partizionamento e la materializzazione delle viste per migliorare le prestazioni delle query.





Enterprise Data Warehouse Project Management

1. Raccolta dei Requisiti

- Business Requirements: Identificare le esigenze di business e gli obiettivi specifici che l'EDWH deve supportare.
- Data Requirements: Determinare quali dati sono necessari e da quali fonti verranno estratti.

2. Progettazione

- Architettura: Definire l'architettura dell'EDWH, che include il modello logico e fisico dei dati.
- Modello di Dati: Progettare il modello di dati, spesso utilizzando un approccio star schema o snowflake schema.
- ETL (Extract, Transform, Load): Progettare i processi ETL per estrarre i dati dalle sorgenti, trasformarli secondo le esigenze e caricarli nel data warehouse.

3. Sviluppo

- Implementazione ETL: Sviluppare e implementare i processi ETL.
- Database Development: Creare il database del data warehouse, implementando il modello di dati progettato.
- Data Integration: Integrare dati da varie fonti per creare una visione unificata.

4. Testing

- Unit Testing: Testare singoli componenti del sistema.
- System Testing: Testare il sistema nel suo complesso per garantire che tutti i componenti funzionino correttamente insieme.
- User Acceptance Testing (UAT): Consentire agli utenti finali di testare il sistema per assicurarsi che soddisfi i requisiti di business.

5. Deployment

- Data Migration: Migrare i dati dai sistemi legacy al nuovo EDWH.
- Go-Live: Rilasciare il sistema in produzione.

6. Manutenzione e Ottimizzazione

- Monitoraggio: Monitorare le prestazioni del data warehouse e l'integrità dei dati.
- Ottimizzazione: Ottimizzare le query e i processi ETL per migliorare le prestazioni.
- Manutenzione: Aggiornare il sistema e aggiungere nuove funzionalità in base alle esigenze aziendali.





Agenda

- Enterprise Data Warehouse
- Esperienza ed Approccio BS
- Architetture di Controllo





Esperienza ed Approccio BS Progetti in ambito BI/DW e Big Data

Problem solving
Assessment
Stime
Analisi funzionali



Disegno ER
Modellazione e
gestione Data Base



Sviluppo e
test Procedure



Realizzazione
dashboard
Analytics
Multidimensionali



Manutenzione
• Presidio
• Correttiva
• Evolutiva



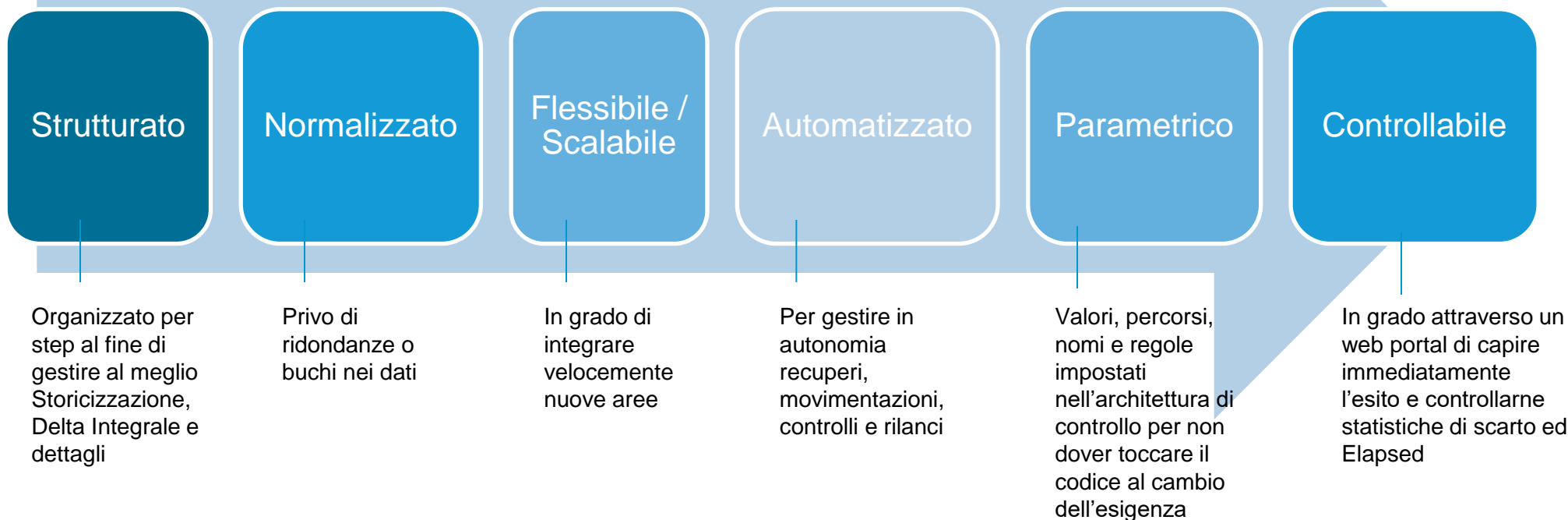
Architetture Data Warehouse di gestione e controllo





L'esperienza ventennale sulla realizzazione di progetti di EDWH

Metodologia e caratteristiche dei nostri Progetti EDW





Lo sviluppo logico del DW

Staging Area

- Presa in carico dei parametri
- gestione
- movimentazione
- archiviazione
- controlli formali dati (key, date, etc.)

Di ogni procedura alimentante



EDW liv1

- Normalizzazione «verticale»
- Gestione Slowly Changing Dimensions (SCDs) in Data Warehouses e Storizzazione dato

Di ogni procedura alimentante



EDW liv2

- Integrazione di tutti i flussi alimentanti in un unico EDW
- in Terza Forma Normale
- privo di ridondanze
- e collegato tramite surrogate key la dove necessario

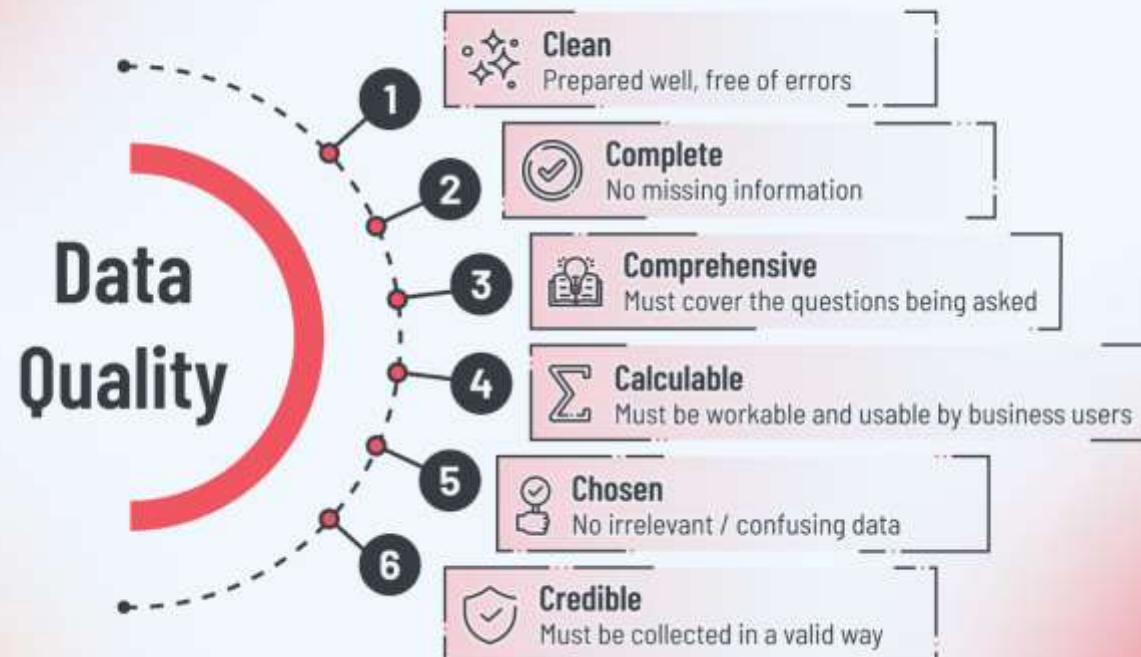


Per Procedura Alimentante si intendono i flussi di: Banking, Factoring, Commun, CSQ, Collection



Concetti Chiave della Data Quality

- **Accuratezza:** Misura quanto i dati rappresentano correttamente la realtà.
- **Completezza:** Indica se tutti i dati necessari sono presenti.
- **Consistenza:** Riguarda la coerenza dei dati tra diverse fonti e sistemi.
- **Tempestività:** Rileva se i dati sono aggiornati e disponibili al momento giusto.
- **Validità:** Verifica che i dati rispettino le regole aziendali e di formato.
- **Integrità:** Assicura che le relazioni tra i dati siano corrette e integre.





Esperienza ed Approccio BS Tecnologie BI/DW

ORACLE®

IBM

DB2

TERADATA

Microsoft®
SQL Server

VERTICA

Greenplum



Data Base

IBM

Datastage

informatica

Microsoft®
SQL Server
INTEGRATION SERVICES

talend
Data Integration

ORACLE®
DATA INTEGRATOR



ETL

SAP BusinessObjects®

QlikView

Power BI

Microsoft®
SQL Server
ANALYSIS SERVICES

JASPER SOFT

apphera

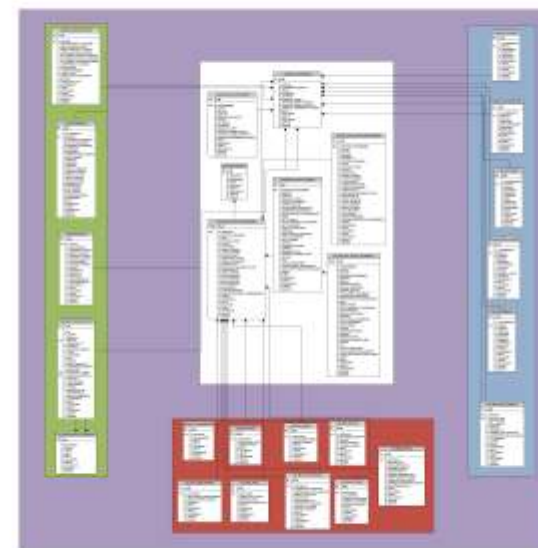
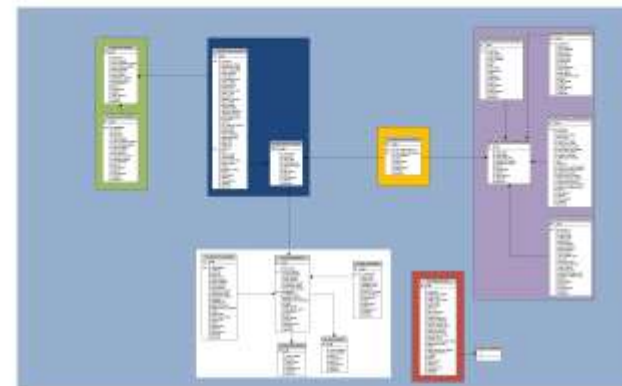
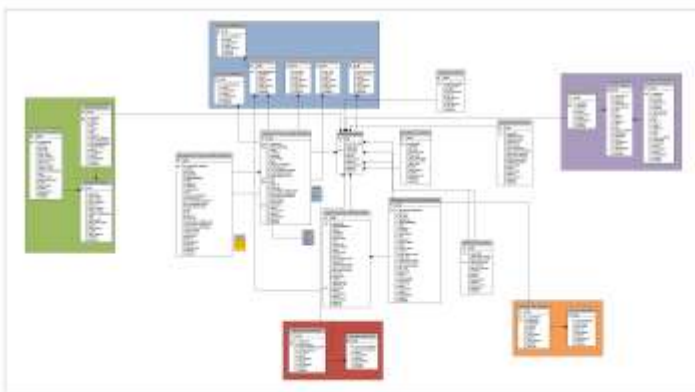
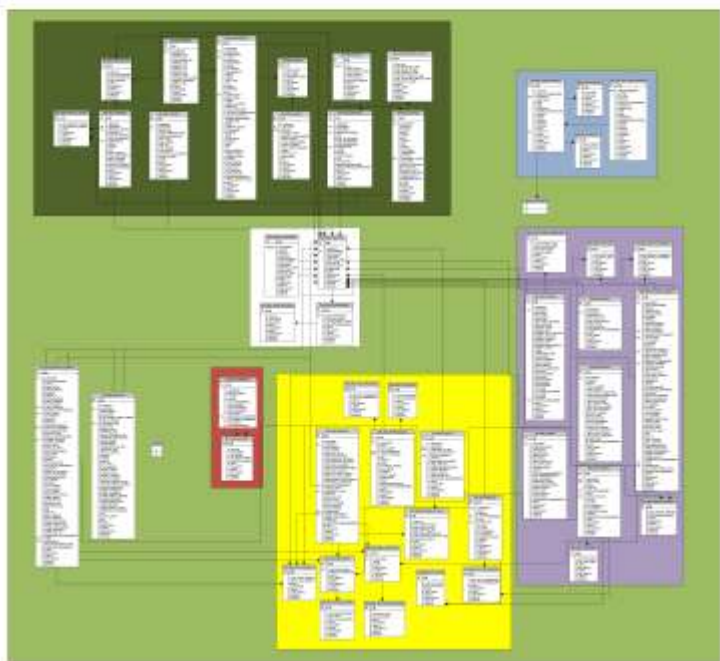


ANALYTICS



Esperienza ed Approccio BS modelli ER

In terza forma normale





Agenda

- Enterprise Data Warehouse
- Esperienza ed Approccio BS
- Architetture di Controllo





Architetture di controllo cosa sono

Le architetture di controllo per un Enterprise Data Warehouse (EDWH) sono fondamentali per garantire la qualità, l'integrità e la sicurezza dei dati all'interno del data warehouse. Queste architetture comprendono una serie di processi, strumenti e best practice che assicurano che i dati siano gestiti e monitorati in modo efficace. Ecco una panoramica delle componenti principali delle architetture di controllo per un EDWH

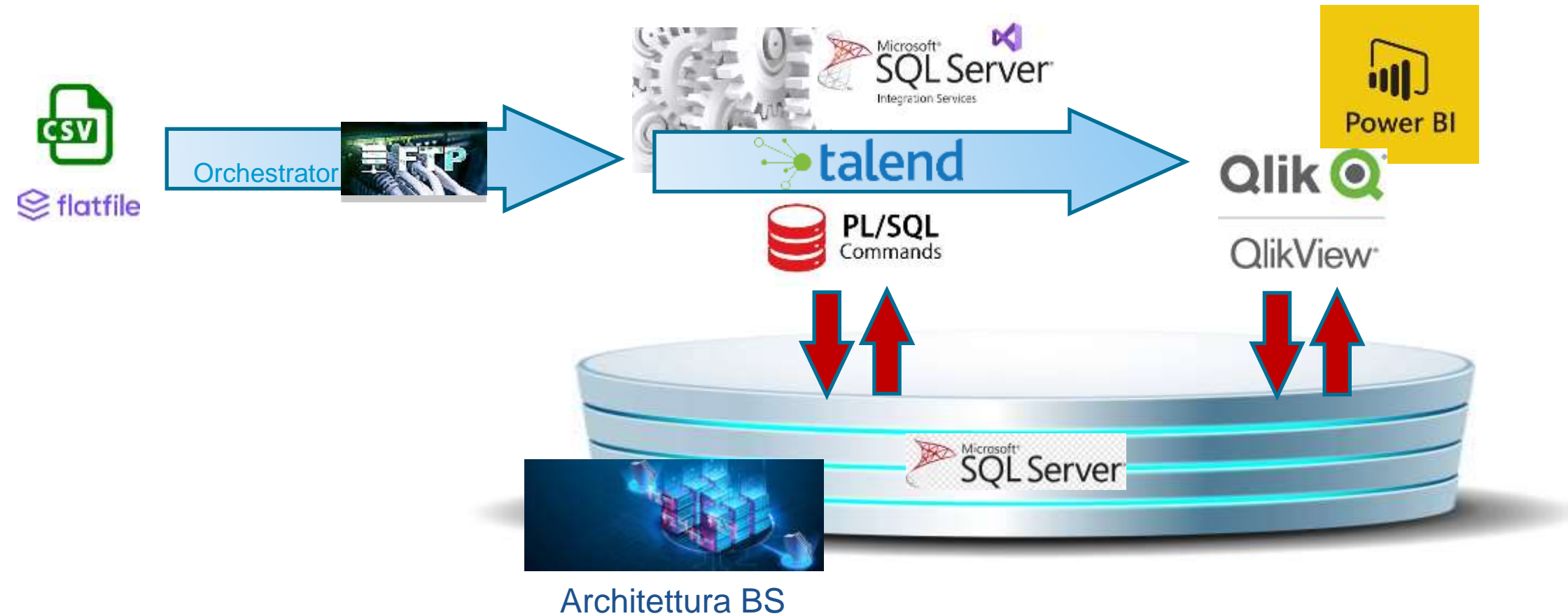
Componenti delle Architetture di Controllo EDWH

- 1. Data Governance**
- 2. Quality Control**
- 3. Security and Access Control**
- 4. Data Integration and ETL Control**
 - 1. ETL Process Monitoring:** Monitorare e gestire i processi di estrazione, trasformazione e caricamento dei dati.
 - 2. Error Handling:** Implementare meccanismi per gestire gli errori e le eccezioni nei processi ETL.
 - 3. Audit Trails:** Mantenere tracce di audit dettagliate per monitorare le modifiche ai dati e i processi ETL.
- 5. Performance Monitoring**
- 6. Data Lineage and Metadata Management**
 - 1. Data Lineage:** Tracciare l'origine e la trasformazione dei dati attraverso il sistema.
 - 2. Metadata Management:** Gestire i metadati per garantire che gli utenti abbiano accesso a informazioni contestuali sui dati.



Architetture di controllo Schema tipo (Banca Sistema)

Esempio di architettura





Architetture di controllo Esigenza 1/2

Un processo male organizzato e non automatizzato **COSTA**, sia in termini di progettazione e sviluppo sia di manutenzione



- Più il progetto diventa esteso e complesso, più si aggiungono e si sviluppano flussi al suo interno, più diventano determinanti alcuni fattori per assicurarne il successo e semplificarne la manutenzione:
 - Gestione parametrica completa e flessibile
 - Per migliorare i passaggi tra ambienti di Sviluppo, Test e Produzione (es. Cambio di utenze e password)
 - Per intervenire il meno possibile sullo sviluppo a fronte di modifiche strutturali (Es. Cambio di path o nomi dei file)
 - Riduzione tempi di analisi e intervento
 - A fronte di Anomalie eventuali
 - A fronte di Nuove “Richieste Utente”
 - Schedulazione e controllo flussi
 - Gestendone ad alto livello le dipendenze funzionali



Architetture di controllo Esigenza 2/2

- Uno degli aspetti più critici dei sistemi di Business Intelligence, che spesso causa il fallimento di interi progetti, è costituito dall'**affidabilità del dato**: l'utente che prende le decisioni di business deve *fidarsi* dei dati mostrati su analisi e report



A tale scopo, diventano elementi chiave:

- ✓ il **controllo** dei processi di costruzione dell'informazione
- ✓ la **rapidità** di diagnosi ed intervento a fronte di anomalie di funzionamento e segnalazioni da parte degli utenti.



Architetture di controllo Overview funzionalità



Gestione Work flow

- **Schedulazione**
 - Batch
 - On-Demand
 - Real Time
- **Dipendenze funzionali**
 - Anche di natura tecnologica differente



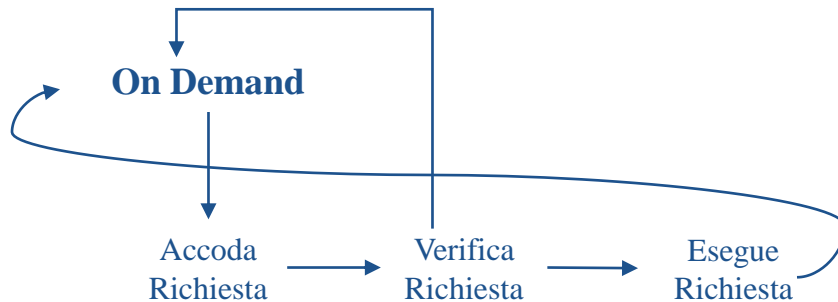
Controllo Esecuzione

- **Verifica parametri e richieste di lancio**
- **Movimentazione automatica sorgenti e target**
 - Controllo
 - Recupero
 - Movimentazione
 - Archiviazione



Monitoraggio Processi

- **Controllo degli esiti dei flussi (pre-post esecuzione)**
- **Gestione dei dettagli di elaborazione**
- **Gestione scarti ed errori**
- **Segnalazione immediata e gestione eventuale punto di arresto**
- **Notifica e-mail differenziata**
- **Storicizzazione dati e dettagli di esecuzione**
- **Unica interfaccia utente (Console) windows client / server**





Web Control

| | | | | |
|----------------------------|-----------------------------|---------------------|-------------------------|--------------------------------|
| Ultima esecuzione corretta | Seguibile Data U. esc. cor. | Ultima esecuzione | Stato Ultima esecuzione | Data Ultima Esc. corretta att. |
| 14/02/2023 06:00:18 | 2023-01-31 | 16/02/2023 06:00:18 | OK | 00:31:27 |

IPVAC (M1)

| | | | |
|----------------------------|-----------------------------|---------------------|-------------------------|
| Ultima esecuzione corretta | Seguibile Data U. esc. cor. | Ultima esecuzione | Stato Ultima esecuzione |
| 16/02/2023 06:00:18 | 2023-01-31 | 16/02/2023 06:00:18 | OK |

| Area | Ultima esecuzione corretta | ESECUTIVAZIONE AREA IN corso | Stato esecuzione | Stato |
|-------------|----------------------------|------------------------------|------------------|-------|
| BANCA | 16/02/2023 06:00:18 | 16/02/2023 06:00:18 | OK | OK |
| COLLETTORI | 16/02/2023 06:00:18 | 16/02/2023 06:00:18 | OK | OK |
| COMUNE | 16/02/2023 06:00:18 | 16/02/2023 06:00:18 | OK | OK |
| CROSS | 16/02/2023 06:00:18 | 16/02/2023 06:00:18 | OK | OK |
| FACTORIZING | 16/02/2023 06:00:18 | 16/02/2023 06:00:18 | OK | OK |

MEMI

| | | | |
|----------------------------|-----------------------------|---------------------|-------------------------|
| Ultima esecuzione corretta | Seguibile Data U. esc. cor. | Ultima esecuzione | Stato Ultima esecuzione |
| 16/02/2023 06:00:18 | 2023-01-31 | 16/02/2023 06:00:18 | OK |

Not Found File

Filtri:

| | |
|-------------------------------|-------------------------------------|
| EDWH: | Area: |
| <input type="checkbox"/> EDW1 | <input type="checkbox"/> BANKING |
| <input type="checkbox"/> EDW2 | <input type="checkbox"/> COLLECTION |
| <input type="checkbox"/> STG | <input type="checkbox"/> COMUNE |
| | <input type="checkbox"/> CQS |
| | <input type="checkbox"/> CROSS |
| | <input type="checkbox"/> FACTORING |
| | <input type="checkbox"/> FINALE |

From: 01/02/2023 To: 16/02/2023

| Row Number | Message Area | Area | Request | State | N. File Annot. | N. File Applic. | Nome File Messaggio |
|------------|--------------|---------|----------|-----------|----------------|-----------------|-----------------------------------|
| 1 | STG | BAIBD10 | 20230212 | NOT FOUND | 23 | 23 | SCBA0740_COSP.ac, SCBA0740_LIQ.ac |
| 2 | STG | BAIBD10 | 20230212 | NOT FOUND | 05 | 23 | SCBA0740_COSP.ac, SCBA0740_LIQ.ac |

EDWH Details

Filtri:

| | | |
|-------------------------------|-------------------------------------|----------------------------------|
| EDWH: | Area: | Stato: |
| <input type="checkbox"/> EDW1 | <input type="checkbox"/> BANKING | <input type="checkbox"/> FAILED |
| <input type="checkbox"/> EDW2 | <input type="checkbox"/> COLLECTION | <input type="checkbox"/> RO |
| <input type="checkbox"/> STG | <input type="checkbox"/> COMUNE | <input type="checkbox"/> NOT RUN |
| | <input type="checkbox"/> CQS | <input type="checkbox"/> OK |
| | <input type="checkbox"/> CROSS | |
| | <input type="checkbox"/> FACTORING | |
| | <input type="checkbox"/> FINALE | |

Order by:

| |
|--------------------------------------|
| <input type="checkbox"/> Data |
| <input type="checkbox"/> Area |
| <input type="checkbox"/> Data e Area |
| <input type="checkbox"/> Area e Data |

Statistiche riassuntive:

| | | | |
|---------------|-----------------|---------------------|----------------------|
| Media Elapsed | Tot. right left | Tot. right caricata | Tot. right rifiutata |
| 00:00:27 | 89556275 | 70198072 | 48099 |

| giugno | | febbraio 2023 | | marzo | | |
|--------|-----|---------------|-----|-------|-----|-----|
| set | dom | lan | mar | ven | sab | dom |
| 01 | 02 | 03 | 04 | 05 | 06 | 07 |
| 08 | 09 | 10 | 11 | 12 | 13 | 14 |
| 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 29 | 30 | 31 | | | | |

Una settimana | recenti relativi al giorno: 2023-02-15

| Area | Area | Execuzione | Request | EDWH | Stato | right left | right caricata | right rifiutata | elapsed (sec) |
|------|-------------|---------------------|-----------|---------------------|-------|------------|----------------|-----------------|---------------|
| BAIB | COLLETTORI | 16/02/2023 06:00:18 | 11:01:003 | W_CRS_2001_CQS_BAIB | OK | 118 | 0 | 118 | 05 |
| BAIB | COMUNE | 16/02/2023 06:00:18 | 11:01:003 | W_CRS_2001_CQS_BAIB | OK | 0 | 0 | 0 | 0 |
| BAIB | FACTORIZING | 16/02/2023 06:00:18 | 11:01:003 | W_CRS_2001_CQS_BAIB | OK | 0 | 0 | 0 | 0 |

EDWH Errori

Filtri:

| | | |
|-------------------------------|-------------------------------------|----------------------------------|
| EDWH: | Area: | Stato: |
| <input type="checkbox"/> EDW1 | <input type="checkbox"/> BANKING | <input type="checkbox"/> FAILED |
| <input type="checkbox"/> EDW2 | <input type="checkbox"/> COLLECTION | <input type="checkbox"/> RO |
| <input type="checkbox"/> STG | <input type="checkbox"/> COMUNE | <input type="checkbox"/> NOT RUN |
| | <input type="checkbox"/> CQS | <input type="checkbox"/> OK |
| | <input type="checkbox"/> CROSS | |
| | <input type="checkbox"/> FACTORING | |
| | <input type="checkbox"/> FINALE | |

Order by:

| |
|--------------------------------------|
| <input type="checkbox"/> Data |
| <input type="checkbox"/> Area |
| <input type="checkbox"/> Data e Area |
| <input type="checkbox"/> Area e Data |

Statistiche riassuntive:

| | | |
|-----------------|---------------------|----------------------|
| Tot. right left | Tot. right caricata | Tot. right rifiutata |
| 15910406 | 12107554 | 6742 |

From: 01/02/2023 To: 16/02/2023

| Row Number | Message Area | Area | Execuzione | EDWH | Stato | N. Errori | Tipi Errori | right left | right caricata | right rifiutata | elapsed (sec) |
|------------|--------------|---------|---------------------|---------------------|--------|-----------|------------------------------------|------------|----------------|-----------------|---------------|
| 1 | EDW1 | BAIBD10 | 11/02/2023 21:11:27 | W_CRS_2001_CQS_EDW1 | FAILED | 4 | Timeout Error Caricamento Error | 4,074,398 | 343,418 | 6,742 | 212 |
| 2 | STG | CQS | 11/02/2023 21:04:44 | W_CRS_1004_CQS_STG | OK | 4 | File not found | 204,276 | 204,276 | 0 | 74 |
| 3 | STG | CQS | 11/02/2023 21:06:07 | W_CRS_1004_CQS_STG | OK | 4 | File not found | 204,204 | 204,204 | 0 | 102 |