

Ciclo di vita dimensionale

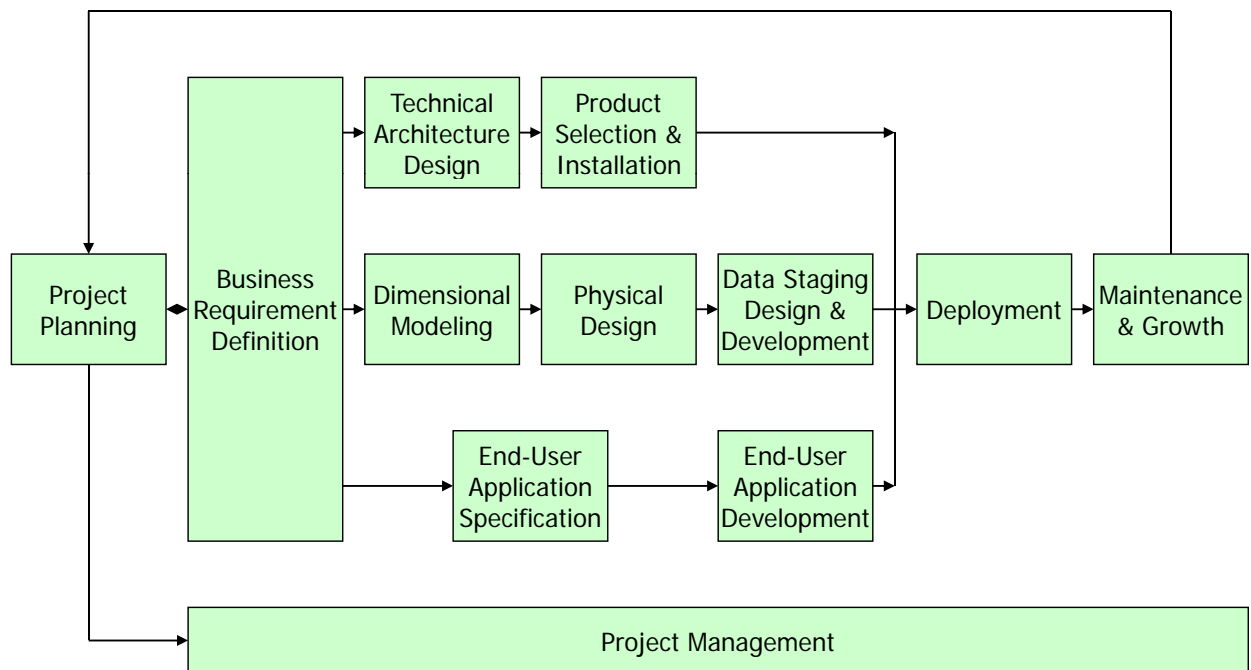
Luca Cabibbo
marzo 2010

Ciclo di vita dimensionale

Viene ora descritta una metodologia completa di progettazione e realizzazione di data warehouse

- la metodologia è descritta dal **ciclo di vita dimensionale (Business Dimensional Lifecycle)** dei data warehouse
- fornisce il contesto di riferimento per la progettazione e realizzazione di data warehouse dimensionali

Fasi nel ciclo di vita dimensionale



3

Ciclo di vita dimensionale

Luca Cabibbo

Pianificazione e gestione del progetto

Il ciclo di vita di un data warehouse ha inizio con la pianificazione, e per tutta la sua durata richiede una attività di gestione

- la pianificazione riguarda la definizione del progetto di data warehouse e dei suoi confini
 - verifica di “readiness” dell’organizzazione
 - analisi preliminare dei requisiti e dei confini del progetto
 - valutazioni economiche
 - identificazione delle risorse
 - formazione dei gruppi di lavoro
 - definizione del piano di sviluppo (iterativo)
- gestione del progetto
 - monitoraggio e controllo
 - gestione dei cambiamenti

4

Ciclo di vita dimensionale

Luca Cabibbo

Raccolta e analisi dei requisiti

Comprensione dettagliata dei requisiti del data warehouse

- requisiti raccolti dai responsabili dei processi
 - obiettivi dell'organizzazione
 - processi dell'organizzazione
 - metriche di prestazione dei processi
 - dimensioni
 - modalità correnti e desiderate di analisi
- requisiti raccolti dai responsabili del sistema informativo
 - sorgenti di dati
 - modalità correnti di analisi

Progettazione del data warehouse

La progettazione del data warehouse avviene su tre tracce parallele

- progettazione dei dati
 - progettazione dimensionale (logica) e fisica del data warehouse
 - progettazione e realizzazione dell'area di preparazione dei dati
- progettazione tecnologica
 - progettazione dell'architettura del data warehouse
 - scelta e installazione di prodotti commerciali
- progettazione delle applicazioni
 - specifica delle applicazioni
 - sviluppo delle applicazioni

Progettazione dimensionale

La progettazione dimensionale è la progettazione logica dei dati del data warehouse, basata sull'architettura a bus

- progettazione dei data mart
- progettazione di un insieme di dimensioni conformi
- progettazione degli schemi dimensionali
- analisi delle sorgenti informative
 - comprensione delle sorgenti informative disponibili e delle loro qualità
 - progettazione preliminare del mapping dei dati dalle sorgenti informative al data warehouse
- piano preliminare delle aggregazioni

Lo schema logico del data warehouse è la pietra angolare della progettazione dell'intero data warehouse

Progettazione fisica

Definizione delle strutture fisiche adeguate a una rappresentazione efficiente dello schema logico del data warehouse

- definizione dello schema dei dati
- progettazione preliminare delle strategie di
 - indicizzazione
 - partizionamento dei dati

Progetto e sviluppo della preparazione dei dati

Il processo di preparazione dei dati

- ha tre passi principali
 - estrazione
 - trasformazione
 - caricamento
- deve gestire la qualità dei dati presenti nelle sorgenti informative
- deve essere realizzata in due forme
 - preparazione al caricamento iniziale del data warehouse
 - preparazione ai caricamenti incrementali

Progettazione tecnologica

Progettazione dell'architettura globale del data warehouse (ovvero, del sistema di data warehousing)

- richiede l'integrazione di numerosi sistemi e tecnologie
 - i sistemi (legacy) su cui risiedono le sorgenti informative, su cui viene parzialmente eseguita la preparazione dei dati
 - i sistemi (solitamente aperti e basati su tecnologia relazionale) su cui risiedono i server di presentazione
 - i sistemi client

Selezione e installazione dei prodotti

- scelta degli ambienti hardware e software

Progettazione delle applicazioni

Specifica delle applicazioni utente

- non tutti gli utenti useranno strumenti ad hoc di interrogazione
- è quindi necessario sviluppare un insieme di applicazioni per gli utenti finali
 - ad esempio, per ciascun rapporto bisogna progettare la struttura del rapporto, i parametri e i calcoli eseguiti
- basata sullo schema logico (dimensionale) dei dati

Sviluppo delle applicazioni

- solitamente basato sull'uso di
 - strumenti avanzati di accesso ai dati
 - metadati

Installazione e avviamento

Le tracce di progettazione dei dati, tecnologica e delle applicazioni convergono in questa fase

- il data warehouse viene costruito e reso accessibile agli utenti finali
 - installazione delle applicazioni
 - formazione degli utenti
 - supporto agli utenti

Manutenzione e crescita

La manutenzione comprende

- supporto e formazione continui
- verifica di qualità del data warehouse
 - soddisfazione degli utenti
 - misurazione delle prestazioni e dell'uso
 - ottimizzazione

Un data warehouse realizzato con il ciclo di vita dimensionale è pronto a evolvere e crescere

- l'evoluzione di un data warehouse è indice di successo, non di fallimento
- solitamente l'evoluzione riguarda il miglioramento dei data mart esistenti e la realizzazione di nuovi data mart