



Università di Pisa

Corso di Laurea in Ingegneria Informatica

Progetto del Corso di Basi di Dati

A.A. 2022-2023

Andrea Covelli, Samuele Riccomini

Sommario

1. Introduzione	4
2. Progettazione Concettuale.....	5
2.1. Dizionario delle entità	5
2.2. Dizionario delle relazioni.....	9
2.3. Regole di Vincolo	12
2.4. Regole di Derivazione.....	13
2.5. Descrizione E/R	14
3. Ristrutturazione E/R	17
3.1. Eliminazione delle Generalizzazioni	17
3.2. Eliminazione degli Attributi Multivalore	17
3.3. Accorpamento di Relazioni	18
3.4. Analisi delle Ridondanze	18
4. Operazioni sui dati	19
4.1. Tavola dei volumi.....	19
4.2. Operazioni.....	24
1) Registrazione utente.....	24
2) Pagamento fattura	26
- Introduzione della ridondanza	29
3) Inserimento di un file.....	31
4) Valutazione di un film da parte di un utente.....	34
5) Numero di fatture emesse in ciascun paese.....	35
6) Stampa stato Server	37
7) Numero di film disponibili per un abbonamento.....	38
- Introduzione della ridondanza	40
8) Numero di visualizzazioni di un film	42
- Introduzione della ridondanza	44
5. Funzionalità	47
5.1. Rating Assoluto	47
5.2. Scelta del server	51
5.3. Caching	53
6. Progettazione logica.....	56
6.1. Schema logico.....	56

7. Vincoli	57
7.1. Vincoli intrarelazionali di dominio.....	58
7.2. Vincoli intrarelazionali di n-upla	58
7.3. Vincoli interrelazionali di integrità referenziale.....	59
8. Analisi delle dipendenze funzionali e normalizzazione.....	61
9. Analytic functions.....	65
9.1. Classifica.....	65
9.2. Custom Analytic.....	66
10. Popolamento del database	69

1. Introduzione

Il seguente progetto consiste nella realizzazione di una Base di Dati che implementa un servizio di streaming e memorizzazione di contenuti chiamato *FilmSphere*.





La Base di Dati memorizza i dati degli utenti che si registrano quali e-mail, password e altri dati sensibili come nome, cognome e il metodo di pagamento per effettuare il pagamento e il rinnovo dell'abbonamento. Inoltre, sono presenti informazioni di carattere tecnico per la gestione dell'infrastruttura di *FilmSphere*, come i dati di ciascun server, i formati di ciascun file, la gestione delle connessioni degli utenti ai server e l'accesso ai vari film.

La seguente documentazione è stata suddivisa in sezioni, ognuna contenente una fase della progettazione della Base di Dati. Sono state inoltre introdotte 8 operazioni che operano sulla base di dati al fine di verificare il funzionamento della stessa.

Si rimanda per ulteriori approfondimenti a ciascuna sezione di questa documentazione in cui, accanto allo specifico argomento, sono state apposte delle note per aiutare il lettore nella comprensione della logica dietro alle scelte effettuate in fase di progettazione.

È stata introdotta una legenda in cui a ogni sezione dello schema E/R è stata assegnato un colore, al fine di migliorare la leggibilità e la chiarezza nella lettura.

Legenda

- Area Contenuti → 
- Area Clienti → 
- Area Streaming → 
- Area Formati → 

2. Progettazione Concettuale

2.1. Dizionario delle entità

La seguente sezione si occupa di elencare le entità che sono state inserite nella progettazione concettuale, eliminate/aggiunte in fase di ristrutturazione o eliminate in fase di normalizzazione.

Per rendere più completa la tabella abbiamo inserito i seguenti riferimenti:

(*) → indica un'entità/attributo aggiunta/o in fase di ristrutturazione.

(**) → indica un'entità/attributo eliminata/o in fase di ristrutturazione.

(#) → indica un attributo ridondante.

[attr₀, attr₁, ..., attr_n]AttrComposto → indica un attributo composto eliminato in fase di ristrutturazione e sostituito con il suo contenuto.

Entità	Attributi	Identificatori	Descrizione
Film	Titolo, Descrizione, AnnoDiProduzione, Genere**(1,N), Durata, NumVisualizzazioni#	Titolo	Film presente su <i>FilmSphere</i>
Genere*	NomeGenere*	NomeGenere	Genere a cui può appartenere un Film su <i>FilmSphere</i>
Attore**	/	Identificato da Artista	Specializzazione di Artista, attore partecipante ad un Film
Regista**	/	Identificato da Artista	Specializzazione di Artista, regista dirigente di un Film

Entità	Attributi	Identificatori	Descrizione
Artista	IDartista, NomeArtista, CognomeArtista, Popolarità	IDartista	Generalizzazione di Attore e Regista
Premio	IDpremio, NomePremio, Anno, Importanza, Attore*, Regista*	IDpremio	Premio vinto da un Film o da un Attore o da un Regista
Critico	NomeCritico, CognomeCritico, IDcritico	IDcritico	Critico cinematografico
Paese	NomePaese, InizioIP, FineIP, RegioneGeografica	NomePaese	Paese in cui è attivo <i>FilmSphere</i> , oppure produttore di un Film
Lingua	NomeLingua	NomeLingua	Lingua in cui è disponibile il doppiaggio o il sottotitolo di un Film
Utente	[Nome, Cognome, Paese, E-mail, Password] InformazioniPersonalì, CodiceUtente, IDultimaFattura#	CodiceUtente	Utente registrato al servizio <i>FilmSphere</i>
Fattura	DataEmissione, IDfattura	IDfattura	Fattura relativa alla sottoscrizione di un abbonamento di <i>FilmSphere</i>
Pagamento	DataPagamento, Importo, [N° Carta, CVV, Scadenza, Intestatario] CartaDiCredito**	IDfattura (esterna)	Rappresenta il pagamento di una fattura emessa

Entità	Attributi	Identificatori	Descrizione
CartaDiCredito*	N° Carta, Scadenza, CVV, Intestatario	N° Carta	Carta di credito usata da un Utente per effettuare i pagamenti
Connessione	IP, Inizio, Fine, [TipoDispositivo, RisoluzioneDispositivo] Dispositivo	IP, Inizio	Connessione con un dispositivo a <i>FilmSphere</i> , effettuata negli ultimi 30 giorni
Abbonamento	Tipologia, Tariffa, N° MaxOre, [N° Profili, N° Dispositivi, Download, Pubblicità] CapacitàSupplementari, NumeroFilm#	Tipologia	Abbonamento offerto da <i>FilmSphere</i>
Visualizzazione	IDvisualizzazione, InizioVisualizzazione, FineVisualizzazione(0,1)	IDvisualizzazione	Visualizzazione di un File da parte di un Utente durante una connessione, effettuata negli ultimi 30 giorni
Server	IDserver, Paese, SpazioArchiviazione, On/Off, CapacitàMassima	IDserver	Server appartenente alla CDN di <i>FilmSphere</i>

Entità	Attributi	Identificatori	Descrizione
File	IDfile, Dimensione, DataRilascio, Lunghezza	IDfile	File in cui è codificato un Film
Formato**	CodiceFormato**, BitRate**	CodiceFormato	Generalizzazione di FormatoAudio e FormatoVideo, indica il formato di codifica di un File
FormatoAudio	CodiceAudio*, BitRateAudio*, BitDepth	CodiceAudio	Specializzazione di Formato, rappresenta un possibile formato di codifica audio per un File
FormatoVideo	CodiceVideo*, BitRateVideo*, Risoluzione, RapportoD'Aspetto, FPS	CodiceVideo	Specializzazione di Formato, rappresenta un possibile formato di codifica video per un File

2.2. Dizionario delle relazioni

La seguente sezione si occupa di elencare le relazioni che sono state inserite nella progettazione concettuale, eliminate/aggiunte in fase di ristrutturazione o eliminate in fase di normalizzazione.

Per rendere più completa la tabella abbiamo inserito i seguenti riferimenti:

(*) → indica una relazione aggiunta in fase di ristrutturazione.

(**) → indica una relazione eliminata in fase di ristrutturazione.

Relazione	Descrizione	Componenti	Attributi
Partecipazione	Associa gli artisti ai film in cui hanno recitato	Film(1,N), Artista(0,N)	Ruolo
Direzione	Associa gli artisti ai film che hanno diretto	Film(1,N), Artista(0,N)	/
PremioAttore**	Associa un premio al relativo attore vincitore	Premio(0,N), Attore(0,N)	/
PremioRegista**	Associa un premio al relativo regista vincitore	Premio(0,N), Regista(0,N)	/
PremioFilm	Associa un premio al relativo film vincitore	Premio(1,1), Film(0,N)	/
PremioArtista*	Associa un premio al relativo artista vincitore	Premio(0,N), Artista(0,N)	/
Appartenenza*	Associazione tra film e genere	Film(1,N), Genere(1,N)	/
ValutazioneCritico	Valutazione, da parte di un critico, di un film	Critico(1,N), Film(0,N)	Testo, Data, Punteggio

Relazione	Descrizione	Componenti	Attributi
Produzione	Associazione tra film e il paese che l'ha prodotto	Film(1,1), Paese(0,N)	/
Codifica	Associa un file al corrispondente film che codifica	Film(1,N), File(1,1)	/
Sottotitolo	Sottotitolo, in una lingua, che può avere un file	Lingua(1,N), File(1,N)	/
Doppiaggio	Doppiaggio audio, in una lingua, che può avere un file	Lingua(0,N), File(1,N)	/
ValutazioneUtente	Valutazione, da parte di un utente, di un film	Utente(0,N), Film(0,N)	Stelle, Data, Commento
Intestazione	Associa una fattura emessa all'utente a cui è intestata	Fattura(1,1), Utente(1,N)	/
Transazione	Associa una fattura pagata al relativo pagamento	Pagamento(1,1), Fattura(0,1)	/
DettagliPagamento*	Associa al pagamento la carta di credito usata per effettuarlo	Pagamento(1,1), CartaDiCredito(0,N)	/
FormaPagamento	Associa una carta di credito ad un utente.	CartaDiCredito(1,1), Utente(1,1)	/
Accesso	Associa una connessione, al servizio FilmSphere, all'utente che la effettua	Utente(1,N), Connessione(1,1)	/

Relazione	Descrizione	Componenti	Attributi
Sottoscrizione	Associa una fattura all'abbonamento che sottoscrive	Fattura(1,1), Abbonamento(0,N)	/
Restrizione Abbonamento	Associa gli abbonamenti ai paesi nei quali non sono disponibili	Paese(0,N), Abbonamento(0,N)	/
Libreria	Associa ad un tipo di abbonamento i file a cui permette la visione	Abbonamento(1,N), File(0,N)	/
Fruizione	Associa ad una visualizzazione la connessione utilizzata	Connessione(0,N), Visualizzazione(1,1)	/
Collegamento	Associa ad una visualizzazione il server utilizzato	Visualizzazione(1,1), Server(0,N)	/
Visualizzazione File	Associa ad una visualizzazione il file visionato durante questa	Visualizzazione(1,1), File(0,N)	/
Memorizzazione	Associa ad ogni file i server che lo contengono	File(1,N), Server(0,N)	/
Restrizione File	Associa i file ai paesi in cui questi non sono disponibili	File(0,N), Paese(0,N)	/
Estensione**	Associa i file ai formati in cui essi sono codificati	File(1,1), Formato(0,N)	/
EstensioneVideo*	Associa ai file il formato video in cui sono codificati	File(1,1), FormatoVideo(0,N)	/
EstensioneAudio*	Associa ai file il formato audio in cui sono codificati	File(1,1), FormatoAudio(0,N)	/

2.3. Regole di Vincolo

Vincolo	Entità
Un artista deve aver recitato o diretto almeno un film	Artista
Un artista può essere presente nelle occorrenze di Partecipazione e di Direzione nel caso in cui sia un attore e il regista di un film, non necessariamente nello stesso	Artista, Film
Un premio con attributo <i>attore (regista)</i> a "true" deve essere assegnato ad un artista che compare nell'occorrenze di Partecipazione (Direzione) relative al film premiato	Artista, Premio, Film
Un premio dell'artista viene assegnato anche al film con cui ha vinto il premio	Premio, Artista, Film
Un paese può appartenere a più regioni geografiche se questo si trova al confine tra 2 o più regioni	Paese
Per tutti i nuovi utenti il primo mese, di qualsiasi abbonamento, è gratuito	Utente, Abbonamento
Ogni utente per poter usufruire del servizio <i>FilmSphere</i> deve avere intestata una fattura, non scaduta, che specifica il tipo di abbonamento che sta utilizzando	Abbonamento, Fattura, Utente
Non è possibile emettere una nuova fattura intestata ad un utente se questo non ha pagato la precedente	Utente, Fattura
<i>DataEmissione</i> di una nuova Fattura deve essere maggiore o uguale al valore di <i>DataEmissione</i> della Fattura precedente (intestata allo stesso utente) + 30 giorni	Fattura
<i>DataPagamento</i> in un Pagamento deve essere maggiore o uguale a <i>DataEmissione</i> della rispettiva Fattura	Pagamento, Fattura
Un file è disponibile agli utenti se compare nella libreria dell'abbonamento che hanno sottoscritto	Abbonamento, Utente, File
<i>IP</i> di una Connessione deve essere compreso tra <i>InizioIP</i> e <i>FineIP</i> di un Paese presente nella base di dati	Connessione, Paese
Il valore di <i>InizioVisualizzazione</i> di una Visualizzazione deve essere maggiore o uguale al valore di <i>Inizio</i> della rispettiva Connessione	Visualizzazione, Connessione
Il valore di <i>FineVisualizzazione</i> di una Visualizzazione deve essere minore o uguale al valore di <i>Fine</i> della rispettiva Connessione	Visualizzazione, Connessione

<i>InizioVisualizzazione</i> di una Visualizzazione deve essere sempre minore di <i>FineVisualizzazione</i> della stessa	Visualizzazione
Il valore di <i>DataRilascio</i> di un File deve essere maggiore o uguale al valore di <i>AnnoProduzione</i> del rispettivo Film	File, Film
Il valore di <i>RisoluzioneSchermo</i> di una Connessione deve essere maggiore o uguale al valore di <i>Risoluzione</i> del FormatoVideo in cui è codificato il file che viene visualizzato	File, FormatoVideo, Visualizzazione, Connessione

2.4. Regole di Derivazione

Vincolo	Entità
Il numero di visualizzazioni di un Film è uguale al numero di occorrenze in Visualizzazione con <i>InizioVisualizzazione</i> pari al mese precedente	Film
Il valore di <i>IDultimaFattura</i> di un Utente corrisponde a <i>IDfattura</i> dell'ultima fattura intestata a quell'utente	Utente
Il numero di film di un Abbonamento corrisponde al numero di occorrenze di Libreria, per quella tipologia di abbonamento, che fanno riferimento a film diversi	Abbonamento

2.5. Descrizione E/R

In questo paragrafo spiegheremo brevemente alcuni dettagli relativi ad attributi/entità/relazioni del diagramma E/R, il cui significato potrebbe risultare ancora poco chiaro.

Abbonamento

Abbiamo ipotizzato che ogni abbonamento abbia una durata di 30 giorni.

Gli attributi *Download* e *Pubblicità* sono 2 booleani che indicano rispettivamente la possibilità di scaricare film per guardarli offline e la presenza di inserzioni durante la visione di un film.

L'attributo *N°MaxOre* indica il numero massimo di ore, disponibili in un abbonamento, dopo le quali non sarà più possibile visionare nessun film fino al rinnovo dell'abbonamento.

N° Dispositivi e *N° Profili* indicano, rispettivamente, il numero massimo di dispositivi su cui è possibile connettersi a *FilmSphere* e il numero massimo di profili che un utente può avere, nel caso in cui volesse dividere il proprio abbonamento con altre persone.

Infine, l'attributo *Tariffa* rappresenta il prezzo mensile dell'abbonamento.

Abbiamo ipotizzato che ogni abbonamento abbia associata una libreria di file di cui permette la visione; quindi, ogni utente ha accesso solo ai film che sono codificati nei file presenti nella libreria dell'abbonamento che ha sottoscritto.

Fattura

L'attributo *DataEmissione* indica il periodo di inizio della validità dell'abbonamento che sottoscrive tale fattura. Per i nuovi utenti, che hanno una sola fattura intestata, tale valore corrisponde al giorno in cui è avvenuta la registrazione, mentre per tutti gli altri utenti, corrisponde al giorno della scadenza della precedente fattura.

Abbiamo ipotizzato che il sistema emette una fattura al momento del pagamento: se questo avviene prima della data di scadenza dell'abbonamento sottoscritto dall'ultima fattura, cioè entro 30 giorni dall'emissione della fattura, allora il valore di *DataEmissione* della nuova fattura sarà pari a *DataEmissione* della precedente fattura + 30 giorni, altrimenti, se avviene dopo, corrisponderà al giorno del pagamento.

Visualizzazione

Ad ogni connessione sono associate una o più visualizzazioni di file, una per ogni file visto dall'utente durante quella connessione. Una visualizzazione corrisponde alla visione di un unico file, a cui è associato tramite la relazione *VisualizzazioneFile*, che si trova in un server a cui è necessario collegarsi per poterlo guardare.

Gli attributi *InizioVisualizzazione* e *FineVisualizzazione* indicano il timestamp di Inizio e Fine della visualizzazione: nel caso in cui il file sia stato visto per intero la differenza dei due attributi sarà pari in valore assoluto al valore dell'attributo *Lunghezza* di quel file, altrimenti sarà minore ma verrà comunque conteggiata come una Visualizzazione.

Indirizzo IP

Ogni paese possiede un intervallo di indirizzo IP, identificato dagli attributi *InizioIP* e *FineIP* dell'entità Paese, assegnato dal I.C.A.N.N. Quindi è possibile localizzare ogni utente che si collega, semplicemente guardando a quale paese appartiene l'intervallo che contiene l'indirizzo IP che identifica il dispositivo usato per la connessione.

Paese

L'attributo *RegioneGeografica* rappresenta un'area con superficie minore di un intero continente ma che può includere uno o più paesi differenti.

Server

L'attributo *CapacitàMassima* specifica il numero massimo di richieste che il server riesce a soddisfare contemporaneamente. Tale valore si misura in base al numero di dispositivi che sono collegati al server. Se il numero di collegamenti ad un server dovesse superare il suo valore di *CapacitàMassima* allora tale server viene considerato in sovraccarico.

L'attributo *SpazioArchiviazione* indica, invece, la dimensione complessiva della memoria che possiede il server. Ogni server è in grado di memorizzare un numero di file la cui somma dei valori di *Dimensione* è pari al valore di *SpazioArchiviazione* del server.

File

L'attributo *DataRilascio* serve per poter distinguere le varie versioni che col tempo possono uscire del solito film, le quali oltre ad avere un formato audio/video migliore, possono avere anche lunghezza differente.

FormaPagamento & CartaDiCredito

Abbiamo ipotizzato che ogni utente debba avere una sola carta di credito registrata sulla piattaforma, in quanto rappresenta l'unico metodo di pagamento consentito su *FilmSphere*.

Assumiamo inoltre che l'intestatario della carta di credito non corrisponda necessariamente all'utente a cui è associata.

3. Ristrutturazione E/R

3.1. Eliminazione delle Generalizzazioni

Artista

Artista è una generalizzazione totale e sovrapposta specializzata in Attore e Regista per distinguere tra le diverse professioni di un artista. Dal momento che le entità figlie non hanno attributi propri abbiamo deciso di accorparle nell'entità padre.

L'entità Artista guadagna la relazione PremioArtista, che la collega a Premio, con cardinalità (0,N). La relazione PremioArtista verrà discussa più nel dettaglio nella sezione Accorpamento di Entità e Relazioni.

Formato

Formato è una generalizzazione totale ed esclusiva specializzata in FormatoAudio e FormatoVideo per distinguere tra i diversi tipi di formati, ossia audio e video che possiede un file. Dal momento che entrambe le entità figlie hanno attributi propri e la generalizzazione è totale abbiamo deciso di accorpare l'entità padre in FormatoAudio e FormatoVideo.

Le entità FormatoAudio e FormatoVideo, guadagnano rispettivamente le relazioni EstensioneAudio e EstensioneVideo che le collegano a File. Queste relazioni hanno cardinalità (0,N) verso FormatoAudio e FormatoVideo e cardinalità (1,1) verso File. Inoltre, FormatoAudio ottiene gli attributi *CodiceAudio* e *BitRateAudio*, di cui il primo lo identifica univocamente. Mentre FormatoVideo ottiene gli attributi *CodiceVideo* e *BitRateVideo*, di cui il primo lo identifica univocamente.

3.2. Eliminazione degli Attributi Multivalore

Genere

Si considera che ogni film appartenga a 2 o più generi. L'attributo multivalore Genere è stato sostituito con l'entità Genere, con un unico attributo *NomeGenere*, che la identifica, e con la relazione Appartenenza, la quale ha cardinalità (1,N) sia verso Film che verso Genere.

3.3. Accorpamento di Relazioni

PremioArtista

Dopo aver eliminato la generalizzazione di Artista abbiamo deciso di accorpare le relazioni PremioAttore e PremioRegista, che servivano per distinguere i tipi di premi, cioè quelli agli attori da quelli ai registi, in un'unica relazione PremioArtista con cardinalità (0,N) sia verso Premio che verso Artista.

A seguito dell'accorpamento, per permettere la distinzione, abbiamo aggiunto all'entità Premio 2 attributi booleani *Attore* e *Regista*, i quali distinguono se un premio è assegnato ad un attore o a un regista oppure solo al film.

La nostra decisione di accorpamento è dovuta al minor numero di accessi che sono necessari per trovare un premio assegnato ad un attore o a un regista.

3.4. Analisi delle Ridondanze

L'analisi delle ridondanze è stata affrontata durante l'analisi delle singole operazioni, rimandiamo quindi al paragrafo Operazioni.

4. Operazioni sui dati

4.1. Tavola dei volumi

Nota: i volumi considerati derivano dal funzionamento di un anno della base di dati.

Concetto	Tipo	Volume	Nota
Film	E	1.000	Ipotesi
Artista	E	3.600	Ipotesi
Partecipazione	R	8.000	Assumendo che siano necessari, in media, 8 attori per fare un film
Direzione	R	1.200	Considerando che per un film possono esserci più registi
Premio	E	600	Ipotesi
PremioFilm	R	600	Assumendo che ogni premio faccia sempre riferimento ad un film anche se a vincerlo è un artista
PremioArtista	R	300	Ipotesi
Genere	E	27	Ipotesi
Appartenenza	R	2.000	Assumiamo che, in media, ogni film appartenga a 2 generi diversi
Paese	E	80	Ipotesi
Produzione	R	1.000	Assumiamo che ogni film sia prodotto da un solo paese

Concetto	Tipo	Volume	Note
Critico	E	30	Ipotesi
ValutazioneCritico	R	300	Ogni critico ha valutato, in media, 10 film
Lingua	E	20	Ipotesi
Doppiaggio	R	30.000	Assumiamo che ogni file abbia, in media, un doppiaggio in 5 lingue diverse
Sottotitoli	R	60.000	Assumiamo che ogni file abbia, in media, i sottotitoli in 10 lingue diverse
Codifica	R	6.000	Ogni file codifica un unico film
Utente	E	500.000	Ipotesi
ValutazioneUtente	R	25.000.000	Assumiamo che ogni utente ha, in media, recensito 50 film
Fattura	E	4.000.000	Considerando che per ogni utente viene emessa una fattura al mese, otteniamo che, in media, ogni utente ha intestate 8 fatture
Intestazione	R	4.000.000	Ogni fattura è intestata ad un unico utente
Abbonamento	E	5	Come da specifiche

Concetto	Tipo	Volume	Note
Sottoscrizione	R	4.000.000	Ogni fattura sottoscrive un unico abbonamento
Pagamento	E	3.500.000	Ipotizziamo che ogni giorno ciascun utente abbia intestata una fattura ancora da pagare
Transazione	R	3.500.000	Un pagamento fa riferimento ad un'unica fattura
CartaDiCredito	E	500.000	Il numero di carte di credito corrisponde al numero di utenti, poiché assumiamo che ognuno di essi abbia associata una carta di credito diversa
DettagliPagamento	R	3.500.000	Ogni pagamento deve possedere i dettagli della carta di credito con cui è stato effettuato
FormaPagamento	R	500.000	Assumiamo che ogni utente abbia una sola carta di credito salvata
RestrizioneAbbonamento	R	32	Ipotizziamo che il 20% dei paesi, in cui è attivo <i>FilmSphere</i> , non abbiano, in media, disponibili 2 tipi di abbonamenti

Concetto	Tipo	Volume	Note
Libreria	R	20.000	Assumiamo che, in media, ogni tipo di abbonamento permette di visionare 4.000 file
Connessione	E	15.000.000	Ipotizziamo che un utente effettui, in media, una connessione al giorno.
Accesso	R	15.000.000	Ogni connessione è relativa ad un unico utente
Server	E	160	Assumiamo che ogni paese posseda, in media, 2 server
Memorizzazione	R	120.000	Ipotizziamo che, in media, ogni file sia presente in 20 server
Visualizzazione	E	15.000.000	Assumiamo che, in media, durante una connessione si effettui una visualizzazione
Collegamento	R	15.000.000	Ogni visualizzazione viene effettuata collegandosi ad un unico server

Concetto	Tipo	Volume	Note
Fruizione	R	15.000.000	Ogni visualizzazione viene effettuata durante un'unica connessione
VisualizzazioneFile	R	15.000.000	Ogni visualizzazione fa riferimento ad un unico file
File	E	6.000	Assumiamo che, in media, ogni film abbia 6 file che lo codificano con formati audio/video diversi
FormatoAudio	E	8	Ipotesi
FormatoVideo	E	10	Ipotesi
EstensioneAudio	R	6.000	Ogni file possiede un'estensione audio
EstensioneVideo	R	6.000	Ogni file possiede un'estensione video
RestrizioneFile	R	4.000	Assumiamo che, in media, in ciascuno paese non siano disponibili 50 file diversi

4.2. Operazioni

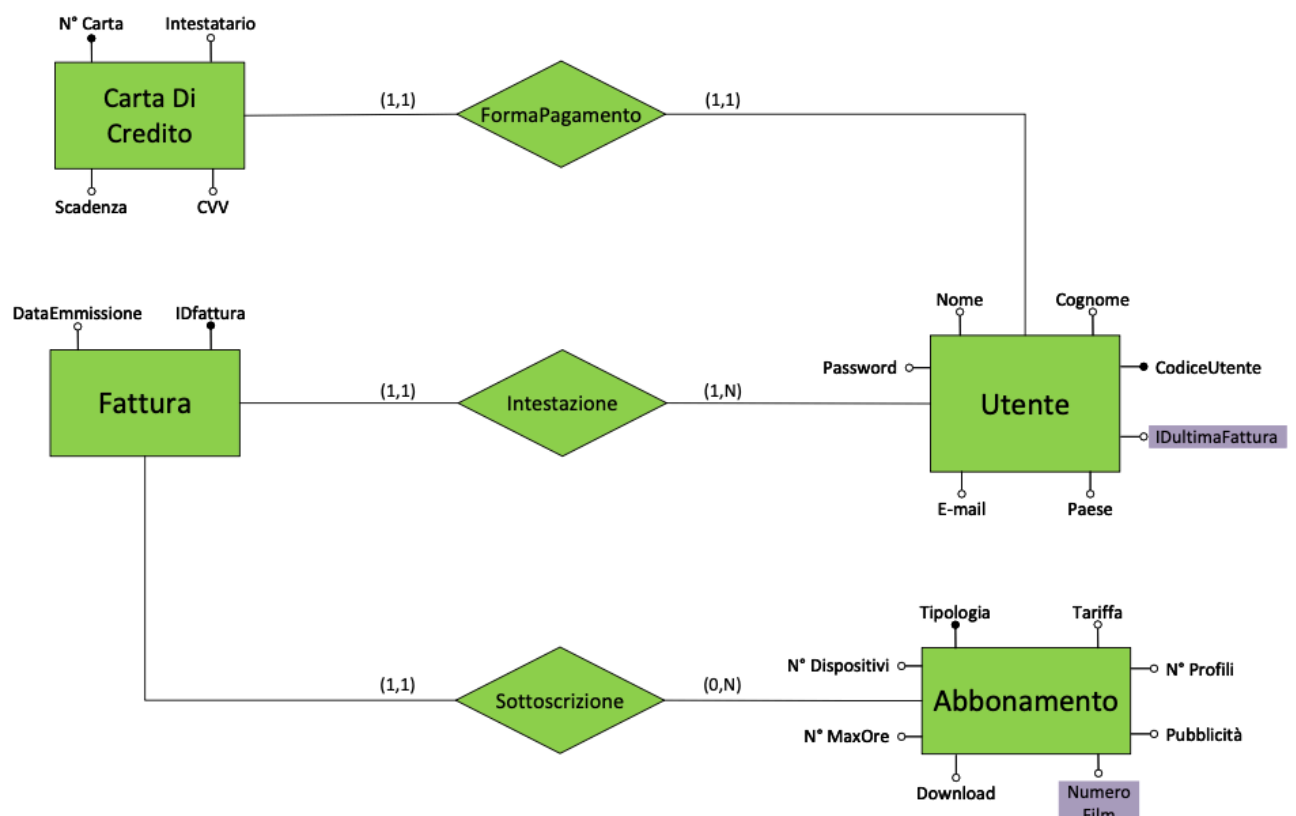
Nel seguente paragrafo verranno analizzate una serie di operazioni, utilizzando come volumi quelli ipotizzati nella sezione precedente. Il numero di accessi, invece, dipende dal tipo di accesso: se viene effettuato per chiave allora è puntuale e possiamo stimare il numero di accessi, altrimenti è necessario eseguire una scansione completa.

Per ogni operazione abbiamo stimato la frequenza giornaliera con cui questa viene ripetuta, che ci ha permesso di calcolare il numero di accessi giornalieri di ciascuna operazione. Durante il calcolo di quest'ultimi abbiamo raddoppiato il numero di ogni accesso in scrittura in quanto consideriamo loro più pesanti rispetto agli accessi in lettura.

1) Registrazione utente

Descrizione: la seguente operazione effettua la registrazione di un nuovo utente su *FilmSphere*. Il primo periodo di utilizzo è gratuito quindi non è richiesto nessun pagamento iniziale al momento della registrazione. Verrà però emessa una fattura intestata al nuovo utente che dovrà pagare per poter continuare a usufruire del servizio nel periodo successivo.

Porzione interessata di diagramma E/R



Input	Output	Frequenza
Nome, Cognome, E-mail, Password, Paese, Tipologia, N° Carta, Scadenza, Intestatario, CVV	/	500 volte al giorno

Volumi coinvolti

Nome	Tipo	Volume
Utente	E	500.000
CartaDiCredito	E	500.000
FormaPagamento	R	500.000
Fattura	E	4.000.000
Intestazione	R	4.000.000
Sottoscrizione	R	4.000.000
Abbonamento	E	5

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Utente	E	500.000	R	Si verifica che l'indirizzo e-mail inserito non sia già presente
2	Utente	E	1	W	Si crea una nuova istanza di utente
3	CartaDiCredito	E	1	W	Si crea una nuova istanza di CartaDiCredito
4	FormaPagamento	R	1	W	Associamo il nuovo utente alla nuova carta di credito

5	Fattura	E	1	W	Creiamo una nuova istanza di fattura
6	Sottoscrizione	R	1	W	Associamo la fattura alla tipologia di abbonamento che il nuovo utente vuole sottoscrivere
7	Intestazione	R	1	W	Associamo la nuova fattura al nuovo utente

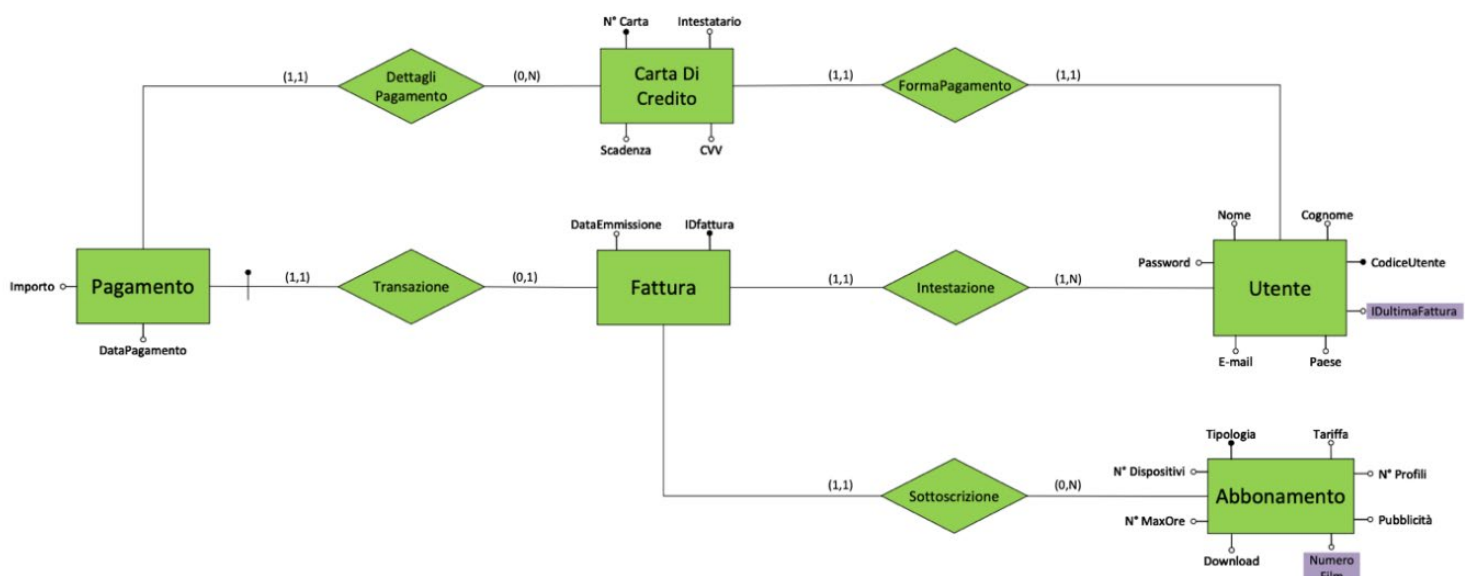
$$o^T = 500.000 + 12 = 500.012$$

$$n^T = 500 \cdot 500.012 = 250.060.000$$

2) Pagamento fattura

Descrizione: un utente effettua il pagamento di una fattura non ancora pagata, la quale sarà l'ultima a lui intestata, poiché il sistema non emette nuove fatture intestate a un utente se prima questo non è in pari con pagamenti. L'operazione comporta anche l'emissione automatica di una nuova fattura intestata a quell'utente, la quale sottoscrive la stessa tipologia di abbonamento della precedente.

Porzione interessata di diagramma E/R



Nota sulla frequenza: ipotizzando che tutti gli abbonamenti abbiano una tariffa mensile, otteniamo che ogni giorno $\frac{4.000.000}{365} \cong 11.000$ utenti devono effettuare il pagamento dell'abbonamento per il mese successivo.

Input	Output	Frequenza
CodiceUtente	/	11.000 volte al giorno

Volumi coinvolti

Nome	Tipo	Volume
Utente	E	500.000
Fattura	E	4.000.000
Intestazione	R	4.000.000
Transazione	R	3.500.000
Pagamento	E	3.500.000
FormaPagamento	R	500.000
CartaDiCredito	E	500.000
DettagliPagamento	R	3.500.000
Sottoscrizione	R	4.000.000
Abbonamento	E	5

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Intestazione	R	4.000.000	R	Si cercano le fatture intestate all'utente. Ogni utente ha, in media, 8 fatture intestate.
2	Fattura	R	8	R	Cerchiamo tra le 8 fatture trovate quella con la data di emissione più recente
3	FormaPagamento	R	1	R	Si cerca la carta di credito dell'utente da usare per il pagamento

4	Pagamento	E	1	W	Si effettua il pagamento di quella fattura creando quindi una nuova istanza di Pagamento
5	DettagliPagamento	R	1	W	Si associa la carta di credito dell'utente al nuovo pagamento
6	Transazione	R	1	W	Si associa il pagamento alla relativa fattura
7	Sottoscrizione	R	1	R	Si cerca la tipologia di abbonamento che sottoscriveva la fattura appena pagata
8	Fattura	E	1	W	Emissione automatica di una nuova fattura
9	Sottoscrizione	R	1	W	Si associa la nuova fattura alla stessa tipologia di abbonamento della precedente
10	Intestazione	E	1	W	Intestazione della nuova fattura all'utente

$$o^T = 4.000.012 + 12 = 4.000.024$$

$$n^T = 11.000 \cdot 4.000.024 = 44.000.300.000$$

- Introduzione della ridondanza

Descrizione: inseriamo la ridondanza IDultimaFattura come attributo dell'entità Utente, la quale tiene conto di quale sia l'ultima fattura intestata all'utente. Risulta ridondante in quanto tale valore potrebbe essere ricavato cercando la fattura più recente tra quelle intestate all'utente. Verifichiamo dunque se il numero di accessi dell'operazione n°2 si riduce, per trarre infine la conclusione più conveniente.

Lettura della ridondanza

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Utente	E	1	R	Leggiamo il valore dell'attributo IDultimaFattura
2	FormaDiPagamento	R	1	R	Si cerca la carta di credito, con la chiave di Utente, da usare per il pagamento
3	Pagamento	E	1	W	Si effettua il pagamento di quella fattura
4	DettagliPagamento	R	1	W	Si associa la carta di credito dell'utente al nuovo pagamento
5	Transazione	R	1	W	Si associa il pagamento alla relativa fattura
6	Sottoscrizione	R	1	R	Si cerca la tipologia di abbonamento che sottoscriveva la fattura appena pagata
7	Fattura	E	1	W	Emissione automatica di una nuova fattura
8	Sottoscrizione	R	1	W	Si associa la nuova fattura alla stessa tipologia di abbonamento della precedente

9	Intestazione	E	1	W	Intestazione della nuova fattura all'utente
---	--------------	---	---	---	---

$$o_{RID}^T = 3 + 12 = 15$$

$$n_{RID}^T = 11.000 \cdot 15 = 165.000$$

$$\Delta = n^T - n_{RID}^T = 44.000.300.000 - 165.000 = 44.000.135.000$$

Scrittura della ridondanza

Descrizione: il tipo di aggiornamento previsto per la ridondanza è l'Immediate Refresh: essa viene aggiornata ad ogni inserimento nell'entità Fattura.

Nota sulla frequenza: l'operazione d'aggiornamento avverrà ogni volta che l'utente effettua un pagamento (11.000 volte al giorno) ma anche ogni volta che si registra un nuovo utente (500 volte al giorno), in quanto in entrambi i casi viene emessa una nuova fattura.

Input	Output	Frequenza
IDfattura	/	11.500 volte al giorno

Volumi coinvolti (mantenimento della ridondanza)

Nome	Tipo	Volume
Fattura	E	4.000.000
Intestazione	R	4.000.000
Utente	E	500.000

Tavola degli accessi (mantenimento della ridondanza)

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Intestazione	R	1	R	Leggiamo a quale utente è intestata la nuova fattura emessa

2	Utente	E	1	W	Aggiorniamo il valore di <i>IDultimaFattura</i> con il valore di <i>IDfattura</i>
---	--------	---	---	---	---

$$o^A = 1 + 2 = 3$$

$$n^A = 11.500 \cdot 3 = 34.500$$

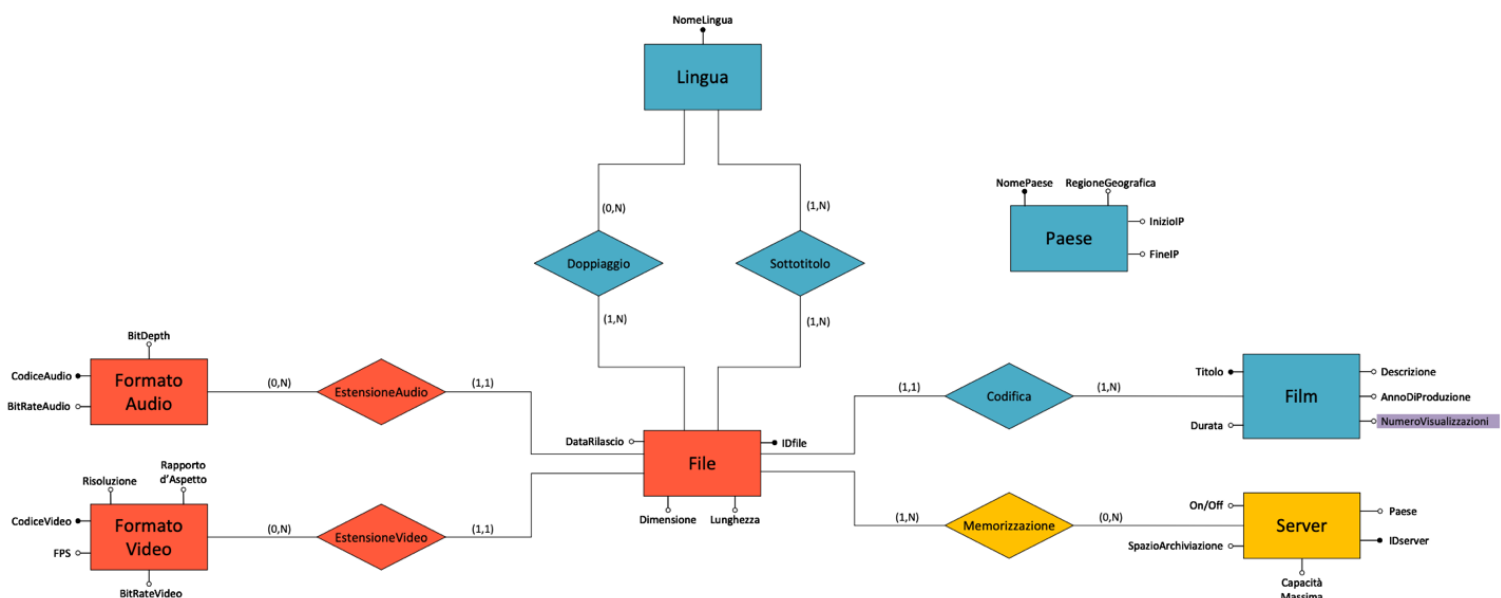
$$\Delta = 44.000.135.000$$

$$n^A \ll \Delta \Rightarrow \text{conviene inserire la ridonanza!}$$

3) Inserimento di un file

Descrizione: si aggiunge un nuovo file alla base di dati, il quale codifica un film: con doppiaggio e sottotitolo in una lingua specifica, codificato in un determinato formato audio e video, e lo inseriamo nei server localizzati in una regione geografica specifica. I file aggiunti possono codificare un film già presente nella base di dati ma con un formato video e/o audio differente, oppure un nuovo film, che sarà reso disponibile agli utenti in un secondo momento. In questa operazione ipotizziamo che il file faccia riferimento ad un film già presente nella base di dati in quanto è l'opzione più comune.

Porzione interessata di diagramma E/R



Nota sulla frequenza: ipotizziamo che *FilmSphere* possedeva al momento del lancio 1.000 file e che nel corso di un anno ne siano stati aggiunti 5.000. Otteniamo quindi che, in media, sono stati aggiunti $\frac{5.000}{365} \cong 13$ file ogni giorno.

Input	Output	Frequenza
Titolo, Lunghezza, Dimensione, DataRilascio, CodiceAudio, CodiceVideo, RegioneGeografica, NomeLingua	/	13 volte al giorno

Volumi coinvolti

Nome	Tipo	Volume
File	E	6.000
Codifica	R	6.000
Film	E	1.000
Lingua	E	20
Paese	E	80
Doppiaggio	R	30.000
Sottotitolo	R	60.000
FormatoAudio	E	8
FormatoVideo	E	10
EstensioneAudio	R	6.000
EstensioneVideo	R	6.000
Server	E	160
Memorizzazione	R	120.000

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	File	E	1	W	Si crea una nuova istanza di file
2	Codifica	R	1	W	Si associa il file al film specificato in input

3	Doppiaggio	R	1	W	Creiamo una nuova occorrenza di doppiaggio con il valore di Lingua passato in input
4	EstensioneAudio	R	1	W	Associamo il file alla sua estensione audio
5	EstensioneVideo	R	1	W	Associamo il file alla sua estensione video
6	Paese	E	80	R	Cerchiamo i paesi appartenenti alla regione geografica passato in input
7	Server	E	160	R	Cerchiamo il valore di <i>IDserver</i> dei server localizzati nei paesi appartenenti alla regione geografica specificata. Ipotizziamo che, in media, in ogni regione geografica ci siano 16 server
8	Memorizzazione	R	16	W	Inseriamo il nuovo file nei 16 server trovati

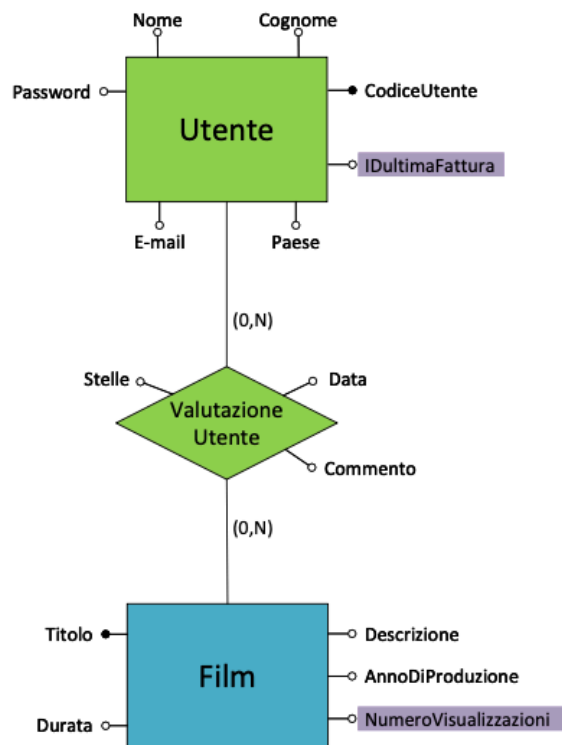
$$o^T = 240 + 42 = 282$$

$$n^T = 13 \cdot 282 = 3.666$$

4) Valutazione di un film da parte di un utente

Descrizione: un utente aggiunge una recensione ad un film.

Porzione interessata di diagramma E/R



Nota sulla frequenza: in media otteniamo che l'operazione si ripete

$$\frac{25.000.000}{365} \cong 68.500 \text{ volte al giorno.}$$

Input	Output	Frequenza
CodiceUtente, Titolo, Stelle, Data, Commento	/	68.500 volte al giorno

Volumi coinvolti

Nome	Tipo	Volume
Utente	E	500.000
ValutazioneUtente	R	25.000.000
Film	E	1.000

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	ValutazioneUtente	R	1	W	Si crea una nuova occorrenza con <i>CodiceUtente</i> e <i>Titolo</i> passati in input

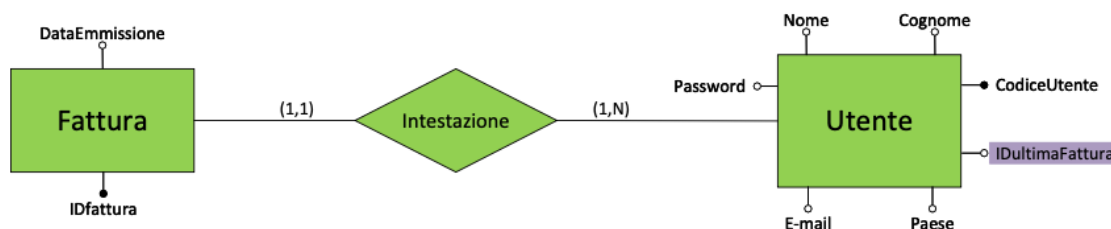
$$o^T = 2$$

$$n^T = 2 \cdot 68.500 = 137.000$$

5) Numero di fatture emesse in ciascun paese

Descrizione: a fine statistico viene calcolato il numero di fatture emesse nel mese precedente in ciascuno dei paesi presenti nella base di dati una volta al mese alla fine del mese.

Porzione interessata di diagramma E/R



Input	Output	Frequenza
DataMesePrecedente	Paese, N° di fatture	1 volta al mese

Volumi coinvolti

Nome	Tipo	Volume
Fattura	E	4.000.000
Intestazione	R	4.000.000
Utente	E	500.000

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Fattura	E	4.000.000	R	Si cercano tutte le fatture emesse nel mese precedente
2	Intestazione	R	300.000	R	Avendo la chiave di Fattura effettuo un accesso puntuale a Intestazione. Si stima di trovare, in media, 300.000 fatture emesse in quel mese (4.000.000/12)
3	Utente	E	300.000	R	Avendo associato ad ogni fattura il suo intestatario, raggruppo loro per <i>Paese</i> . Contando quanti utenti appartengono a ciascun gruppo ricavo il numero di fatture emesse per ogni paese nel mese precedente

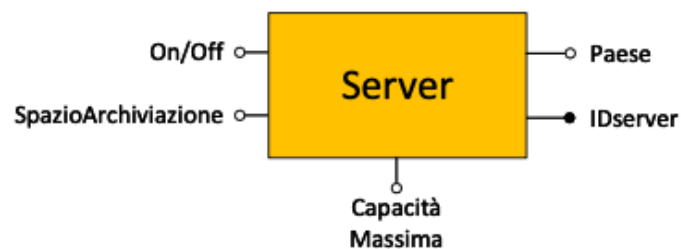
$$o^T = 4.000.000 + 600.000 = 4.600.000$$

$$n^T = \frac{1}{30} \cdot 4.600.000 = 153.334$$

6) Stampa stato Server

Descrizione: viene visualizzato lo stato di tutti i server, cioè, il valore di tutte le informazioni che i server possiedono nella base di dati al momento della richiesta.

Porzione interessata di diagramma E/R



Nota sulla frequenza: ipotizziamo che ogni minuto venga visualizzato lo stato di ciascun server.

Input	Output	Frequenza
/	IDserver, Paese, SpazioArchiviazione, On/Off, CapacitàMassima	1.440 volte al giorno

Volumi coinvolti

Nome	Tipo	Volume
Server	E	160

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Server	R	160	R	Si leggono tutti i dati relativi ad ogni server

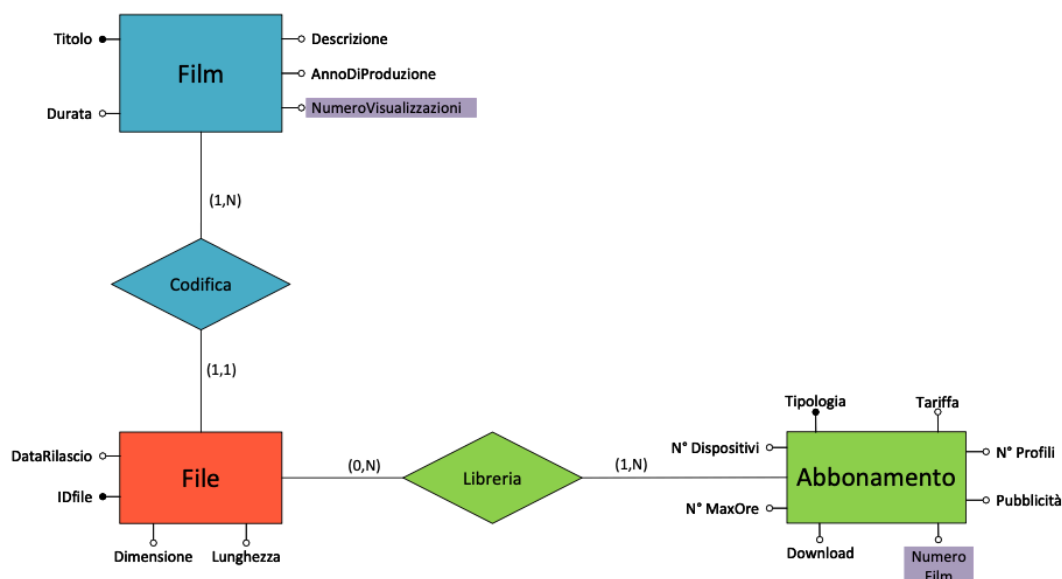
$$o^T = 160$$

$$n^T = 1.440 \cdot 160 = 230.400$$

7) Numero di film disponibili per un abbonamento

Descrizione: l'operazione mostra all'utente il numero di film previsti da un abbonamento (tale valore rappresenta una stima, in quanto non prende in considerazione i file che non sono disponibili in ciascun paese; quindi, il valore corretto sarà leggermente inferiore rispetto a quello calcolato dall'operazione).

Porzione interessata di diagramma E/R



Nota sulla frequenza: l'operazione viene eseguita quando:

- Ogni volta che un nuovo utente si registra sulla base di dati, quindi 500 volte al giorno
- Ogni volta che un utente decide di cambiare abbonamento. Ipotizziamo che ogni mese il 15% degli utenti (75.000) cambia abbonamento, quindi ciò avverrà, in media, $\frac{75.000}{30} = 2.500$ volte al giorno

In totale otteniamo che l'operazione si ripete $2.500 + 500 = 3.000$ volte al giorno.

Input	Output	Frequenza
Tipologia	N° Film	3.000 al giorno

Volumi coinvolti

Nome	Tipo	Volume
Abbonamento	E	5
Libreria	R	20.000
File	E	6.000
Film	E	1.000
Codifica	R	6.000

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Libreria	R	20.000	R	Si cercano tutti i file associati ad esso, che abbiamo ipotizzato essere in media 4.000
2	Codifica	R	4.000	R	Con la chiave di File associamo a ogni file il film che codifica e contiamo a quanti film diversi otteniamo

$$o^T = 20.000 + 4.000 = 24.000$$

$$n^T = 3.000 \cdot 24.000 = 72.000.000$$

- Introduzione della ridondanza

Descrizione: inseriamo la ridondanza NumeroFilm come attributo dell'entità Abbonamento la quale tiene conto del numero di film disponibili per ogni abbonamento. Risulta ridondante in quanto può essere calcolata come mostrato nelle nell'operazione sopra. Verifichiamo se il numero di accessi dell'operazione n°7 vengono ridotti per trarre infine la conclusione più conveniente.

Lettura della ridondanza

Tavola degli accessi (con ridondanza)

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Abbonamento	E	1	R	Leggiamo il valore di <i>NumeroFilm</i> da restituire come output

$$o_{RID}^T = 1$$

$$n_{RID}^T = 3.000 \cdot 1 = 3.000$$

$$\Delta = n^T - n_{RID}^T = 72.000.000 - 3.000 = 71.997.000$$

Scrittura della ridondanza

Descrizione: il tipo di aggiornamento previsto per tale ridondanza è il Deferred Refresh con cadenza mensile: assumendo che all'inizio del mese vengono resi disponibili agli utenti nuovi film allora, una volta al mese viene ricalcolato il numero di film disponibili per ciascun abbonamento e aggiornata la ridondanza.

Input	Output	Frequenza
/	/	1 volta al mese

Volumi coinvolti (mantenimento della ridondanza)

Nome	Tipo	Volume
Abbonamento	E	5
Libreria	R	20.000
File	E	6.000
Film	E	1.000
Codifica	R	6.000

Tavola degli accessi (mantenimento della ridondanza)

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Libreria	R	20.000	R	Raggruppiamo ogni occorrenza in base alla tipologia di abbonamento
2	Codifica	E	6.000	R	Associamo ad ogni file il titolo del film che codifica e contiamo quanti film diversi abbiamo per ogni abbonamento
3	Abbonamento	E	5	W	Aggiorniamo la ridondanza di ogni tipologia di abbonamento

$$o^A = 20.000 + 6.000 + 10 = 26.010$$

$$n^A = \frac{1}{30} \cdot 26.010 = 867$$

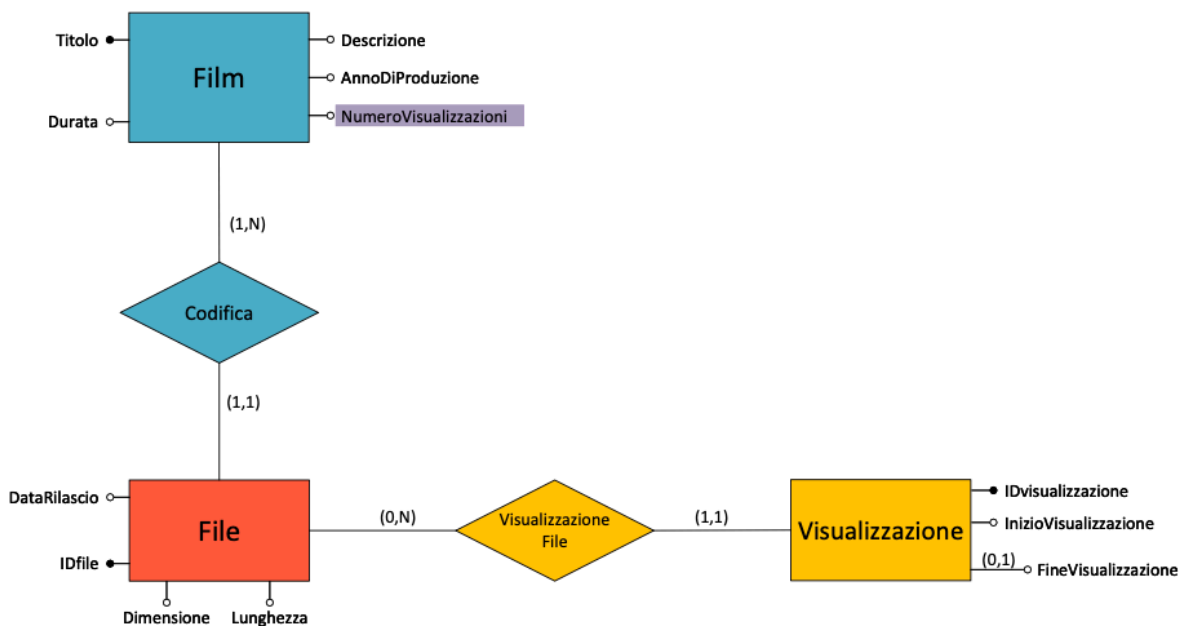
$$\Delta = 72.997.000$$

$$n^A \ll \Delta \Rightarrow \text{conviene inserire la ridondanza!}$$

8) Numero di visualizzazioni di un film

Descrizione: mediante la seguente operazione quando un utente clicca su un film gli viene mostrato il numero di visualizzazioni che ha avuto quel film nel mese precedente.

Porzione interessata di diagramma E/R



Nota sulla frequenza: considerando che, in media, ogni giorno tutti gli utenti sono attivi e guardano un solo film, ipotizziamo che ognuno visualizzi le informazioni di circa 5 film prima di trovare quello da vedere. Otteniamo quindi che l'operazione si ripete, in media, $500.000 \cdot 5 = 2.500.000$ volte al giorno.

Input	Output	Frequenza
Titolo	N° Visualizzazioni	2.500.000 volte al giorno

Volumi coinvolti

Nome	Tipo	Volume
Film	E	1.000
Codifica	R	6.000
File	E	6.000
VisualizzazioneFile	R	15.000.000
Visualizzazione	E	15.000.000

Tavola degli accessi

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Codifica	R	6.000	R	Si controlla a quali file è associato il film passato in input. Si trovano in media 6 file per ogni film
2	VisualizzazioneFile	R	15.000.000	R	Associamo ad ogni file la relativa visualizzazione. Si stima di trovare in media $\frac{15.000.000}{6.000} \cdot 6 = 15.000$ visualizzazioni per ogni film
3	Visualizzazioni	E	15.000	R	Selezioniamo solo le visualizzazioni che sono state effettuate nel mese precedente e contiamole

$$o^T = 15.021.000$$

$$n^T = 2.500.000 \cdot 15.021.000 = 37.552.500.000.000$$

- Introduzione della ridondanza

Descrizione: inseriamo la ridondanza NumeroVisualizzazioni come attributo dell'entità Film, la quale tiene conto del numero di visualizzazioni che quel film ha ricevuto il mese precedente. Verifichiamo se il numero di accessi dell'operazione n°8 vengono ridotti per trarre infine la conclusione più conveniente.

Lettura della ridondanza

Tavola degli accessi (con ridondanza)

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	Film	E	1	R	Si legge il valore della ridondanza del film dato in input

$$o_{RID}^T = 1$$

$$n_{RID}^T = 2.500.000 \cdot 1 = 2.500.000$$

$$\Delta = n^T - n_{RID}^T =$$

$$37.552.500.000.000 - 2.500.000 = 37.552.497.500.000$$

Scrittura della ridondanza

Descrizione: il tipo di aggiornamento previsto per tale ridondanza è il Deferred Refresh con cadenza mensile: poiché la ridondanza fa riferimento al mese precedente, è sufficiente aggiornarla una volta al mese, alla fine del mese, così che tutte le occorrenze di Visualizzazione facciano riferimento al periodo di nostro interesse.

Input	Output	Frequenza
/	/	1 volta al mese

Volumi coinvolti (mantenimento della ridondanza)

Nome	Tipo	Volume
Film	E	1.000
Codifica	R	6.000
File	E	6.000
VisualizzazioneFile	R	15.000.000
Visualizzazione	E	15.000.000

Tavola degli accessi (mantenimento della ridondanza)

N°	Concetto	Costrutto	Accessi	Tipo	Note
1	VisualizzazioneFile	R	15.000.000	R	Leggiamo tutte le visualizzazioni effettuate negli ultimi 30 giorni. Ipotizziamo che circa 2.000 file diversi vengano visualizzati ogni mese
2	Codifica	R	2.000	R	Con la chiave di File associamo ad ogni file, il titolo del corrispondente film e raggruppando per film si contano le righe di ogni gruppo
3	Film	E	1.000	W	Si incrementa il valore della ridondanza in ciascun film pari al numero di visualizzazioni giornaliere che ha ricevuto

$$o^A = 15.000.000 + 2.000 + 2.000 = 15.004.000$$

$$n^A = 15.004.000 \cdot \frac{1}{30} \cong 500.000$$

$$\Delta = 37.552.497.500.000$$

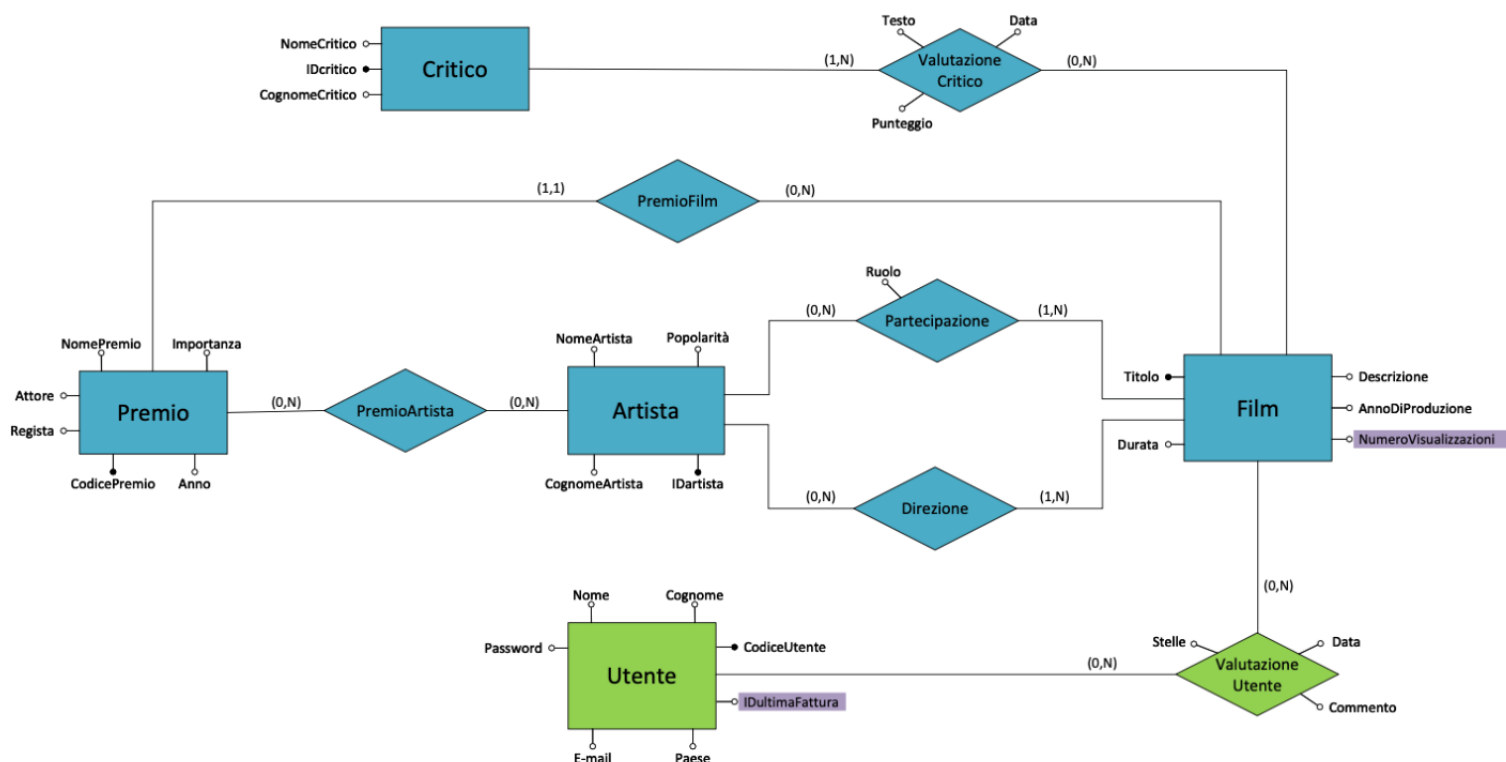
$$n^A \ll \Delta \Rightarrow \text{conviene inserire la ridonanza!}$$

5. Funzionalità

5.1. Rating Assoluto

Descrizione: il Rating Assoluto consiste nell'assegnare un punteggio su una scala da 1 a 10 a ciascun film presente in *FilmSphere*. Il punteggio è basato su dei criteri elencati e spiegati all'interno della successiva sezione Metrica.

Porzione interessata di diagramma E/R



Metrica

La metrica adottata per calcolare il Rating assoluto di un film di *FilmSphere* è la seguente:

$$Rating_{abs} = \mu_{rating-utenti} + \mu_{rating-critici} + \delta_{regista} + \delta_{attori} + \lambda_{premi-film} + \lambda_{premi-attori-film} + \gamma$$

Il valore di $Rating_{abs}$ è compreso tra 1 e 10 (estremi inclusi).

Glossario

- $\mu \triangleq$ media
- $\delta \triangleq$ coefficiente di popolarità (premi vinti in assoluto)

- $\lambda \triangleq$ coefficiente dei premi vinti (relativi al film)
- $\gamma \triangleq$ coefficiente di successo (numero di visualizzazioni)

Calcolo dei singoli fattori

$$1. \mu_{rating-utenti} = \beta_1 \mu_1$$

Il seguente fattore *media* consiste nel calcolare la media del rating assegnato dagli utenti di tutta *FilmSphere* per quel film. Il calcolo è abbastanza semplice dato che la media considerata è una media nella versione semplice e non in quella ponderata.

Il valore della media μ_1 è compreso tra 1 e 5 (¹).

Questo valore viene poi moltiplicato per un *peso* di valore $\beta_1 = 0.4$

$$2. \mu_{rating-critici} = \beta_2 \mu_2$$

Il seguente fattore *media* consiste nel calcolare la media del rating assegnato dai critici che hanno recensito almeno un film di *FilmSphere*. Il calcolo è abbastanza semplice dato che la media considerata è una media nella versione semplice e non in quella ponderata.

Il valore della media μ_2 è compreso tra 1 e 5.

Questo valore viene poi moltiplicato per un *peso* di valore $\beta_2 = 0.3$ (²)

$$3. \delta_{regista} = \beta_3 \delta_1$$

Il seguente fattore *popolarità* consiste nel calcolare la popolarità di un regista sulla base di tutti i premi che ha vinto nel corso della sua carriera, comprendendo quindi sia il film in questione che tutti gli altri film di *FilmSphere*.

Per assegnare un valore al seguente fattore abbiamo deciso di usare delle fasce di popolarità elencate nella tabella seguente:

Numero di premi vinti dal regista	Valore del fattore δ_1
0-19	1
20-39	2
40-59	3
60-79	4
80-99+	5

¹ Non è necessario avere una tabella di corrispondenza come quelle presenti sotto perché le valutazioni, sia degli utenti che dei critici, sono già comprese tra 1 e 5

² Il peso assegnato al rating medio dei critici e degli utenti è lo stesso per dare uguale importanza ai due fattori di valutazione.

Il fattore δ_1 viene moltiplicato per un peso assegnato $\beta_3 = 0.2$

$$4. \delta_{attore} = \beta_4 \delta_2$$

Il seguente fattore *popolarità* consiste nel calcolare la popolarità di un attore sulla base di tutti i premi che ha vinto nel corso della sua carriera, comprendendo quindi sia il film in questione che tutti gli altri film di *FilmSphere*.

Per assegnare un valore al seguente fattore abbiamo deciso di usare delle fasce di *popolarità* elencate nella tabella seguente:

Numero di premi vinti dall'attore	Valore del fattore δ_2
0-9	1
10-19	2
20-29	3
30-39	4
40-49+	5

Il fattore δ_2 viene moltiplicato per un peso assegnato $\beta_4 = 0.3$ ⁽³⁾

$$5. \lambda_{premi-film} = \beta_5 \lambda_1$$

Il seguente fattore *premi vinti* consiste nell'assegnare un punteggio ai premi vinti dal film, inerenti unicamente al film.

Per assegnare un valore al seguente fattore abbiamo deciso di usare delle fasce *numero-premi* elencate nella tabella seguente:

Numero di premi vinti dal film	Valore del fattore δ_2
0-4	1
5-9	2
10-14	3
15-19	4
20-24+	5

Il fattore λ_1 viene moltiplicato per un peso assegnato $\beta_5 = 0.3$

³ Si è deciso di assegnare un valore maggiore (0.3 contro 0.2) alla popolarità dell'attore rispetto a quella del regista perché abbiamo ritenuto che il punteggio di un film sia più influenzato dagli attori presenti che dal regista del film in questione.

$$6. \lambda_{\text{premi-attore-film}} = \beta_6 \lambda_2$$

Il seguente fattore *premi vinti* consiste nell'assegnare un punteggio ai premi vinti da ciascun attore in quel film.

Per assegnare un valore al seguente fattore abbiamo deciso di usare delle fasce, numero-*premi-attore* vinti per quel film, elencate nella tabella seguente:

Numero di premi vinti dal film	Valore del fattore δ_2
0-4	1
5-9	2
10-14	3
15-19	4
20-24+	5

Il fattore λ_2 viene moltiplicato per un peso assegnato $\beta_6 = 0.2$

$$7. \gamma = \beta_7 \gamma_1$$

Il seguente fattore *numero visualizzazioni* consiste nell'assegnare un punteggio al film sulla base del numero di visualizzazioni registrate dal film nell'ultimo mese. Quest'ultimo fattore, così come il rating, vengono aggiornati una volta al mese.

Per assegnare un valore al seguente fattore abbiamo deciso di usare delle fasce *numero-visualizzazioni* elencate nella tabella seguente:

Numero di visualizzazioni	Valore del fattore δ_2
0-99.999	1
100.000-299.999	2
300.000-599.999	3
600.000-999.999	4
1.000.000+	5

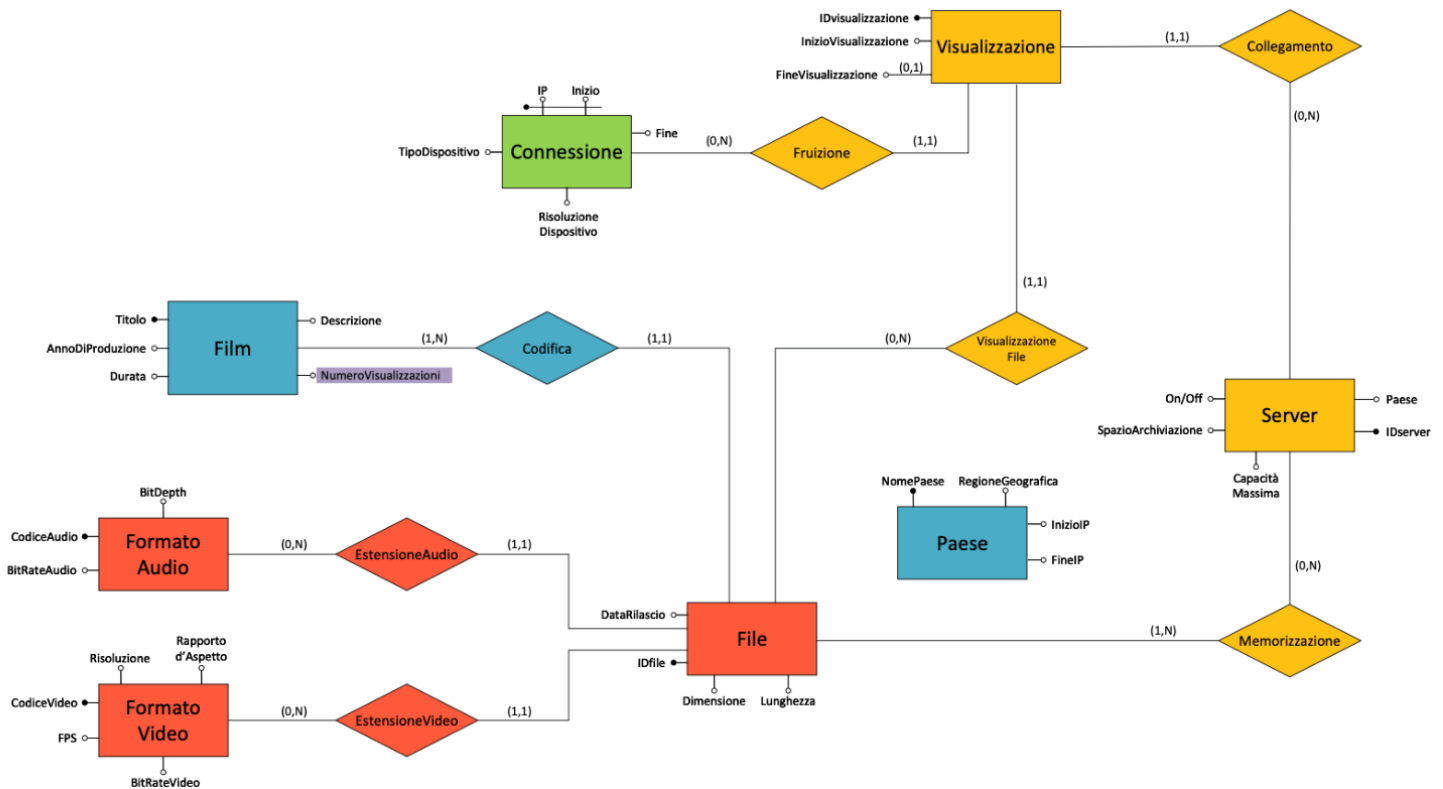
Il fattore γ_1 viene moltiplicato per un peso assegnato $\beta_7 = 0.3$

Nota: i vari pesi β sono stati scelti in modo da assegnare un "peso" diverso a ciascun fattore presente in $Rating_{abs}$.

5.2. Scelta del server

Descrizione: quando un utente richiede di visualizzare un film al servizio, la funzionalità Scelta del Server, sceglie il server della CDN più vicino all'utente che ha il file richiesto, per ridurre i tempi di buffering e migliorare l'esperienza di visione.

Porzione interessata di diagramma E/R



Metrica

Per realizzare la seguente operazione abbiamo adottato come metrica il Paese in cui si trova l'utente, rintracciabile tramite indirizzo IP. All'interno del paese dal quale l'utente si collega, cerchiamo il primo server che possiede il file richiesto e che non sia in sovraccarico. Se non troviamo un server con tali caratteristiche all'interno del paese, espandiamo la ricerca ai paesi della stessa regione geografica

Se infine si presentasse il caso in cui tutti i server della regione geografica sono occupati o non hanno il file richiesto, la ricerca del server verrebbe ripetuta dall'inizio ma su un altro file, che codifica il solito film.

Nel caso in cui non venga trovato nessun server all'interno della regione geografica in cui si trova l'utente, dovremmo espandere la ricerca ad un'altra regione geografica. Ipotizziamo però che questo evento sia molto raro e che la ricerca si concluda in uno dei casi precedenti.

Input	Output	Frequenza
Titolo, IP	IDserver	500.000 volte al giorno

Nota sulla frequenza: la scelta del server avviene ogni volta che un utente vuole visualizzare un file e abbiamo ipotizzato che ciò avviene, in media, per ogni utente una volta giorno, di conseguenza, la frequenza giornaliera è pari al numero degli utenti.

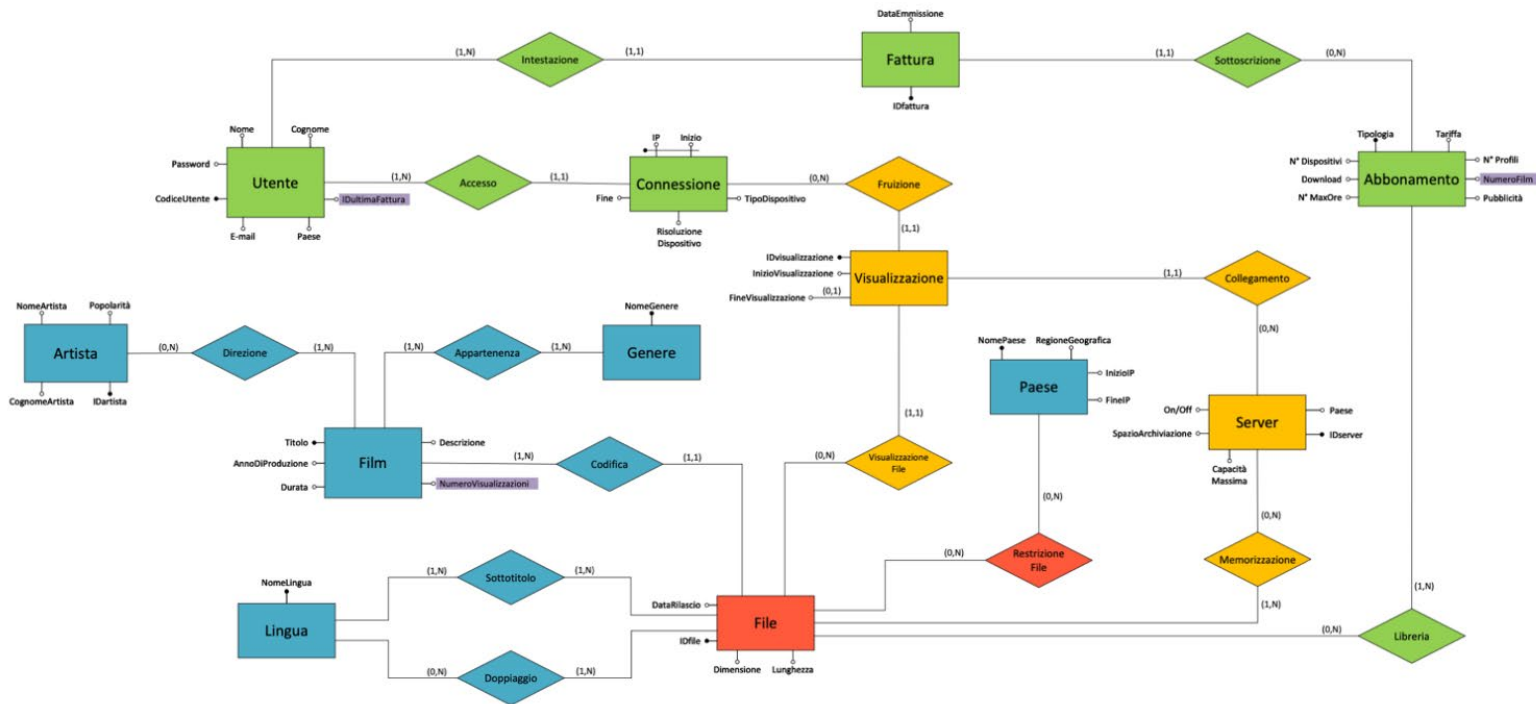
Procedura

1. Grazie al valore di *IP* risaliamo al Paese e alla regione geografica in cui si trova l'utente cercando l'intervallo di appartenenza di tale valore nell'entità Paese. Inoltre, in quanto chiave di Connessione, ricaviamo anche i dati relativi al dispositivo, tra cui la *Risoluzione*, la quale ci servirà per trovare un file, che codifica il film richiesto, con un formato video supportato dal dispositivo.
2. Con il valore di *Titolo*, e il valore di *Risoluzione*, cerchiamo il file, che codifica il film richiesto, avente la massima risoluzione supportata dal dispositivo.
3. Cerchiamo tra le occorrenze di Memorizzazione in quali server è presente quel file.
4. Se sono stati trovati dei server leggiamo, con il valore della chiave *IDserver*, in quale paese si trovano e verifichiamo se almeno uno corrisponde al paese in cui si trova l'utente. Se tale ricerca ha successo l'operazione termina, restituiamo *IDserver* di quel server e saltiamo al punto 7, altrimenti proseguiamo.
5. Si ripete la ricerca con un altro Paese appartenente alla stessa regione geografica fino a trovare un server non in sovraccarico contenente il file richiesto.
6. Nel caso in cui non venisse trovato nessun server ripartiamo dal punto 2 prendendo un nuovo file con risoluzione minore rispetto al precedente.
7. Prima di terminare l'operazione crea una nuova istanza di visualizzazione riferita al file selezionato, al server trovato e alla connessione in atto.

5.3. Caching

Descrizione: la funzionalità di Caching consiste nell'allocare i file dei film che potrebbero interessare all'utente nei server a lui più vicini, con il fine di garantire una migliore fruizione dei contenuti.

Porzione interessata di diagramma E/R



Metrica

La metrica adottata per scegliere di quali film deve essere effettuato il caching si basa sui seguenti criteri:

1. Si considerano le lingue audio e dei sottotitoli in cui l'utente ha visualizzato gli ultimi 10 film, i film i cui file non hanno almeno una di quelle lingue sia per l'audio che per i sottotitoli gli viene assegnata una probabilità pari a 0

$$p_{film} = \begin{cases} \alpha & \text{se il film rispetta la condizione 1.} \\ 0 & \text{altrimenti} \end{cases}$$

2. Ai film che rispettano la condizione sopra citata viene assegnata una probabilità non nulla, determinata dai seguenti fattori:

$$\text{Genere } p_{genere} = \begin{cases} 0.10 & \text{se il film } \in \text{ al genere di uno degli ultimi 10 film vis.} \\ 0.01 & \text{altrimenti} \end{cases}$$

$$\text{Visualizzazioni } p_{vis.} = \begin{cases} 0.50 & \text{se il n° di vis. del film} \geq 500.000 \\ 0.10 & \text{altrimenti} \end{cases}$$

$$\text{Regista } p_{regista} = \begin{cases} 0.15 & \text{se il regista è lo stesso di almeno uno degli ultimi 10 film} \\ 0.05 & \text{altrimenti} \end{cases}$$

$$\text{Durata } p_{durata} = \begin{cases} 0.05 & \text{se la durata} \leq \text{durata media degli ultimi 10 film} \\ 0.01 & \text{altrimenti} \end{cases}$$

$$\alpha = p_{genere} + p_{vis.} + p_{regista} + p_{durata}$$

La probabilità α ottenuta indica la probabilità che un film venga visualizzato dall'utente.

Si può osservare che la probabilità massima che un film può raggiungere è pari allo 0.80 (lo 80%). Nessun film può dunque raggiungere la probabilità del 100%. Questa decisione è stata presa considerando che, generalmente, l'utente richiede lo 80% delle volte un contenuto simile a quello che ha già visualizzato e, un restante 20%, per film con caratteristiche diverse dagli ultimi 10 film visualizzati.

Come ulteriore osservazione si osserva che al n° di visualizzazioni è stato assegnato il *peso* maggiore. Il motivo di questa scelta risiede nella scelta di porre maggior risalto alle preferenze degli utenti in generale, pur mantenendo un restante 50% di variabilità su altri fattori.

Input	Output	Frequenza
CodiceUtente	/	100.000 volte al giorno

Nota sulla frequenza: poiché durante l'operazione viene effettuato il caching di 5 film supponiamo che per i 5 giorni successivi non sia necessario effettuare nuovamente il caching per quell'utente, di conseguenza, ripetiamo l'operazione 1 volta ogni 5 giorni per ciascun utente ottenendo una frequenza giornaliera di $\frac{500.000}{5} = 100.000$ volte.

Procedura

1. Con il valore di *CodiceUtente* leggiamo le occorrenze di Accesso per trovare i valori di chiave di Connessione.
2. Associamo ad ogni connessione le visualizzazioni effettuate.
3. Prendiamo in considerazione solamente le ultime 10 visualizzazioni e da quelle ricaviamo i file e di conseguenza gli ultimi 10 film visualizzati dall'utente attraverso la relazione Codifica.
4. Di ciascun film trovato leggiamo la relazione Direzione per associarvi il regista.
5. Con la chiave di utente leggiamo il valore di IDultimaFattura e dalla fattura corrispondente risaliamo, attraverso la relazione Sottoscrizione, al tipo di abbonamento che attualmente l'utente ha sottoscritto.
6. Con la chiave di abbonamento accediamo alle occorrenze di Libreria e ricaviamo la chiave dei file a cui l'utente ha accesso con il proprio abbonamento.
7. Con il valore dell'attributo *Paese* dell'entità Utente accediamo per chiave alla relazione RestrizioneFile così da ottenere la chiave dei file ristretti nel paese dell'utente e tali file vengono rimossi da quelli trovati al punto precedente.
8. Ogni file rimasto viene associato al relativo film e per ciascuno viene effettuato il calcolo della probabilità illustrato nella sezione Metrica.
9. Si cercano i server presenti nel paese dell'utente e vengono aggiunti ad uno di essi i primi 5 film con probabilità α maggiore.

6. Progettazione logica

6.1. Schema logico

Nella traduzione dal modello concettuale E/R ristrutturato al modello logico sono state adottate le seguenti convenzioni:

- i. Le sezioni delle relazioni sono state colorate secondo gli stessi colori usati nel modello E/R.
- ii. Le chiavi di ciascuna relazione sono state sottolineate.

Area contenuti

Film(Titolo, Descrizione, AnnoDiProduzione, Durata,
NumeroVisualizzazioni, NomePaese)

Genere(NomeGenere)

Appartenenza(Titolo, NomeGenere)

Partecipazione(Titolo, IDArtista, Ruolo)

Direzione(Titolo, IDArtista)

Artista(IDArtista, NomeArtista, CognomeArtista, Popolarità)

PremioArtista(CodicePremio, IDArtista)

Premio(CodicePremio, NomePremio, Importanza, Anno, Regista, Attore, Titolo)

ValutazioneCritico(IDcritico, Titolo, Testo, Data, Punteggio)

Critico(IDcritico, NomeCritico, CognomeCritico)

Sottotitolo(NomeLingua, IDfile)

Doppiaggio(NomeLingua, IDfile)

Lingua(NomeLingua)

Paese(NomePaese, InizioIP, FineIP, RegioneGeografica)

Area formati

File(IDfile, Dimensione, DataRilascio, Lunghezza, Titolo, CodiceAudio, CodiceVideo)

FormatoAudio(CodiceAudio, BitRateAudio, BitDepth)

FormatoVideo(CodiceVideo, FPS, Risoluzione, BitRateVideo, RapportoD'Aspetto)

RestrizioneFile(NomePaese, IDfile)

Area clienti

Utente(CodiceUtente, Nome, Cognome, E-mail, Paese, Password, IDultimaFattura,
N° Carta)

ValutazioneUtente(CodiceUtente, Titolo, Stelle, Data, Commento)

Fattura(IDfattura, DataEmissione, Tipologia, CodiceUtente)

Pagamento(IDfattura, DataPagamento, Importo, N° Carta)

CartaDiCredito(N° Carta, DataScadenza, Intestatario, CVV)

Connessione(IP, Inizio, RisoluzioneDispositivo, TipoDispositivo, Fine, CodiceUtente)

Abbonamento(Tipologia, Tariffa, N° Profili, Pubblicità, N° Dispositivi, Download,
N° MaxOre, NumeroFilm)

RestrizioneAbbonamento(NomePaese, Tipologia)

Libreria(IDfile, Tipologia)

Area streaming

Visualizzazione(IDvisualizzazione, InizioVisualizzazione, FineVisualizzazione, IP,
Inizio, IDserver, IDfile)

Server(IDserver, SpazioArchiviazione, Paese, On/Off, CapacitàMassima)

Memorizzazione(IDfile, IDserver)

7. Vincoli

7.1. Vincoli intrarelazionali di dominio

Nota:

- i. Gli attributi di tipo DATE hanno tutti come vincolo che la data deve essere maggiore o uguale a quella di lancio di *FilmSphere*
- ii. Gli attributi di tipo INT devono assumere tutti un valore ≥ 0

Attributo	Min	Max
Film(Anno)	1900	YEAR(CURRENT_DATE)
Film(Durata) <<in minuti>>	5	360
CartaDiCredito(N°carta)	0000 0000 0000 0000	9999 9999 9999 9999
CartaDiCredito(CVV)	100	999
CartaDiCredito(MeseScadenza)	1	12
CartaDiCredito(AnnoScadenza)	2022	-
Connessione(IP)	0	255 255 255 255
Fattura(N°carta)	0000 0000 0000 0000	9999 9999 9999 9999
Paese(InizioIP)	0	255 255 255 255
Paese(FineIP)	0	255 255 255 255
Premio(Importanza)	0	100
ValutazioneUtente(Stelle)	1	5
ValutazioneCritico(Punteggio)	1	5
Artista(Popolarità)	1	10
FormatoVideo(FPS)	22	60
Abbonamento(N° Profili)	1	10
Abbonamento(N° Dispositivi)	1	5

7.2. Vincoli intrarelazionali di n-upla

- i. Premio(Film) *IS NOT NULL*
- ii. Paese(InizioIP) deve essere minore di Paese(FineIP)

- iii. Il valore di CartaDiCredito(Mese) e CartaDiCredito(Anno) deve essere maggiore della data in cui viene effettuato l'inserimento della carta
- iv. Connessione(Inizio) deve essere minore di Connessione(Fine)
- v. Visualizzazione(InizioV) deve essere minore di Visualizzazione(FineV)

7.3. Vincoli interrelazionali di integrità referenziale

Attributo	Vincolo di integrità ref. con
Film(Paese)	Paese(NomePaese)
File(Titolo)	Film(Titolo)
Direzione(Titolo)	Film(Titolo)
Direzione(IDartista)	Artista(IDartista)
Partecipazione(Titolo)	Partecipazione(Titolo)
Partecipazione(IDartista)	Partecipazione(IDartista)
PremioArtista(IDartista)	Artista(IDartista)
PremioArtista(CodicePremio)	Premio(CodicePremio)
PremioFilm(CodicePremio)	Premio(CodicePremio)
ValutazioneCritico(IDcritico)	Critico(IDcritico)
ValutazioneCritico(Titolo)	Film(Titolo)
ValutazioneUtente(CodiceUtente)	Utente(CodiceUtente)
ValutazioneUtente(Titolo)	Film(Titolo)
Connessione(CodiceUtente)	Utente(CodiceUtente)
Utente(Paese)	Paese(NomePaese)

Attributo	Vincolo di integrità ref. con
Utente(N° Carta)	CartaDiCredito(N° Carta)
Utente(CodiceUtente)	Connessione(CodiceUtente)
Pagamento(N° Carta)	CartaDiCredito(N° Carta)
Fattura(CodiceUtente)	Utente(CodiceUtente)
Fattura(Tipologia)	Abbonamento(Tipologia)
Visualizzazione(IDserver)	Server(IDserver)
Visualizzazione(IDfile)	File(IDfile)
Visualizzazione(IP, Inizio)	Connessione(IP, Inizio)
Memorizzazione(IDserver)	Server(IDserver)
Memorizzazione(IDfile)	File(IDfile)
Libreria(Tipologia)	Abbonamento(Tipologia)
Libreria(IDfile)	File(IDfile)
RestrizioneFile(NomePaese)	Paese(NomePaese)
RestrizioneFile(IDfile)	File(IDfile)
RestrizioneAbbonamento(NomePaese)	Paese(NomePaese)
RestrizioneAbbonamento(Tipologia)	Abbonamento(Tipologia)
Doppiaggio(NomeLingua)	Lingua(NomeLingua)
Doppiaggio(IDfile)	File(IDfile)
Sottotitolo(NomeLingua)	Lingua(NomeLingua)
Sottotitolo(IDfile)	File(IDfile)

8. Analisi delle dipendenze funzionali e normalizzazione

La seguente sezione è dedicata a verificare che le relazioni prodotte nella progettazione concettuale siano in forma BCNF.

Nota: alle relazioni che soddisfano la BCNF vi è stato apposto un segno di spunta ✓

Area contenuti

- ✓ Film(Titolo, Descrizione, AnnoDiProduzione, Durata, NumeroVisualizzazioni, NomePaese)
 - Titolo → Descrizione, AnnoDiProduzione, Durata, NumeroVisualizzazioni, NomePaese
- ✓ Genere(NomeGenere)
 - La relazione contiene un unico attributo che è chiave
- ✓ Appartenenza(Titolo, NomeGenere)
 - La relazione contiene 1 chiave
- ✓ Partecipazione(Titolo, IDArtista, Ruolo)
 - Titolo, IDArtista → Ruolo
- ✓ Direzione(Titolo, IDArtista)
 - La relazione contiene 1 chiave
- ✓ Artista(IDArtista, NomeArtista, CognomeArtista, Popolarità)
 - IDArtista → NomeArtista, CognomeArtista, Popolarità
- ✓ PremioArtista(CodicePremio, IDArtista)
 - La relazione contiene 1 chiave
- ✓ PremioFilm(CodicePremio, Titolo)
 - La relazione contiene 1 chiave
- ✓ Premio(CodicePremio, NomePremio, Importanza, Anno, Regista, Attore)
 - CodicePremio → NomePremio, Importanza, Anno, Regista, Attore
- ✓ ValutazioneCritico(IDcritico, Titolo, Testo, Data, Punteggio)
 - IDcritico, Titolo → Testo, Data, Punteggio
- ✓ Critico(IDcritico, NomeCritico, CognomeCritico)
 - IDcritico → NomeCritico, CognomeCritico

- ✓ Sottotitolo(NomeLingua, IDfile)
 - La relazione contiene 1 chiave
- ✓ Doppiaggio(NomeLingua, IDfile)
 - La relazione contiene 1 chiave
- ✓ Lingua(NomeLingua)
 - La relazione contiene un unico attributo che è chiave
- ✓ Paese(NomePaese, RegioneGeografica, InizioIP, FineIP)
 - NomePaese → InizioIP, FineIP, RegioneGeografica

Area formati

- ✓ File(IDfile, Dimensione, DataRilascio, Lunghezza, Titolo)
 - IDfile → Dimensione, DataRilascio, Lunghezza, Titolo
- ✓ EstensioneAudio(CodiceAudio, IDfile)
 - La relazione contiene 1 chiave
- ✓ EstensioneVideo(CodiceVideo, IDfile)
 - La relazione contiene 1 chiave
- ✓ FormatoAudio(CodiceAudio, BitRateAudio, BitDepth)
 - CodiceAudio → BitRateAudio, BitDepth
- ✓ FormatoVideo(CodiceVideo, FPS, Risoluzione, BitRateVideo, RapportoD'Aspetto)
 - CodiceVideo → FPS, Risoluzione, BitRateVideo, RapportoD'Aspetto
- ✓ RestrizioneFile(NomePaese, IDfile)
 - La relazione contiene 2 chiavi

Area clienti

- ✓ Utente(CodiceUtente, Nome, Cognome, E-mail, Paese, Password, IDultimaFattura, N° Carta)
 - CodiceUtente → Nome, Cognome, E-mail, Paese, Password, N° Carta, IDultimaFattura
 - E-mail → Nome, Cognome, CodiceUtente, Paese, Password, N° Carta, IDultimaFattura

Si hanno 2 dipendenze funzionali che sono entrambe chiave della relazione Utente
- ✓ ValutazioneUtente(CodiceUtente, Titolo, Stelle, Data, Commento)
 - CodiceUtente, Titolo → Stelle, Data, Commento

Si considera che un utente possa scrivere al massimo una recensione per ciascun film
- ✓ Fattura(CodiceFattura, DataEmissione, Tipologia, CodiceUtente)
 - CodiceFattura → DataEmissione, Tipologia, CodiceUtente
- ✓ Pagamento(IDfattura, DataPagamento, Importo, N° Carta)
 - IDfattura, DataPagamento → Importo, N° Carta
- ✓ CartaDiCredito(N° Carta, DataScadenza, Intestatario, CVV)
 - N° Carta → DataScadenza, Intestatario, CVV
- ✓ Connessione(IP, Inizio, RisoluzioneDispositivo, TipoDispositivo, Fine, CodiceUtente)
 - IP, Inizio → RisoluzioneDispositivo, TipoDispositivo, Fine, CodiceUtente
- ✓ Abbonamento(Tipologia, Tariffa, N° Profili, Pubblicità, N° Dispositivi, Download, N° MaxOre, NumeroFilm)
 - Tipologia → Tariffa, N° Profili, Pubblicità, N° Dispositivi, Download, N° MaxOre, NumeroFilm
- ✓ RestrizioneAbbonamento(NomePaese, Tipologia)
 - La relazione contiene 1 chiave
- ✓ Libreria(IDfile, Tipologia)
 - La relazione contiene 1 chiave

Area streaming

- ✓ Visualizzazione(IDvisualizzazione, InizioV, FineV, IP, Inizio, IDserver, IDfile)
 - IDvisualizzazione → InizioV, FineV, IP, Inizio, IDserver, IDfile
- ✓ Server(IDserver, SpazioArchiviazione, Paese, On/Off, CapacitàMassima)
 - IDserver → SpazioArchiviazione, Paese, On/Off, CapacitàMassima
- ✓ Memorizzazione(IDfile, IDserver)
 - La relazione contiene 1 chiave

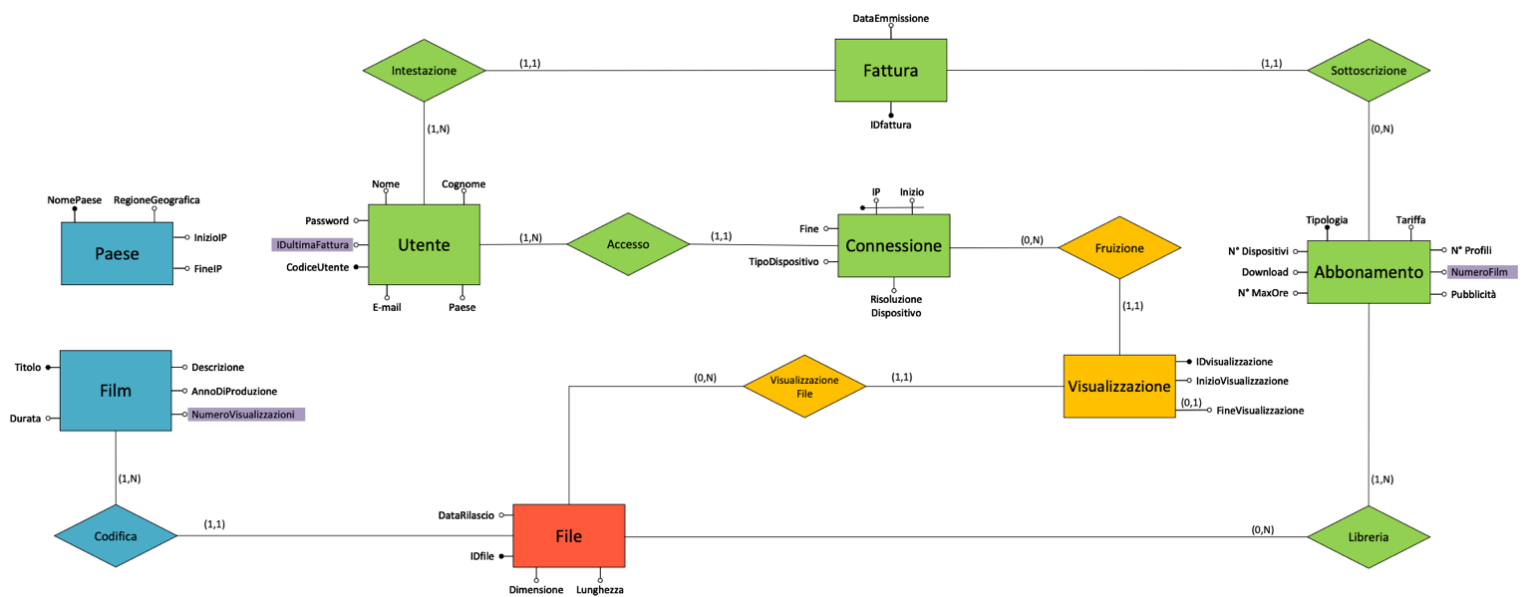
Dopo un'analisi di tutte le relazioni abbiamo verificato che tutte le relazioni prodotte sono normalizzate in BCNF.

9. Analytic functions

9.1. Classifica

Descrizione: questa analytic si occupa di stilare una classifica dei film più visti a seconda del Paese, e della tipologia di abbonamento.

Porzione interessata di diagramma E/R



Input	Output	Frequenza
NomePaese, Tipologia	Classifica	1 volta al mese

Procedura

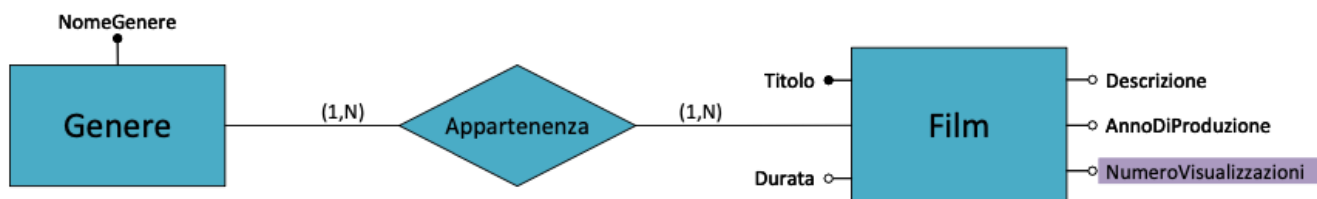
1. Leggiamo ogni occorrenza presente nella relazione *VisualizzazioneFile* e otteniamo la coppia di chiavi *IDfile* e *IDvisualizzazione*.
2. Associamo ad ogni file il film che codifica accedendo per chiave alla relazione *Codifica*.
3. Associamo ad ogni visualizzazione il *CodiceUtente* accedendo prima alla relazione *Fruizione*, ottenendo la chiave di *Connessione*, e successivamente alla relazione *Accesso* ottenendo la chiave di *Utente*.
4. Con il valore di *CodiceUtente* accediamo direttamente all'entità *Utente* per leggere i valori dell'attributo *Paese* e l'attributo *IDultimaFattura* così da ottenere la tipologia di abbonamento che ha sottoscritto accedendo alla relazione *Sottoscrizione*.

5. Selezioniamo le occorrenze che hanno il valore di Paese e Tipologia uguali ai valori passati in input.
6. Raggruppiamo per film e contiamo le righe di ogni gruppo.
7. Ordiniamo in ordine decrescente per ottenere la classifica con in testa il film con il maggior numero di visualizzazioni.

9.2. Custom Analytic

Descrizione: la custom analytic Generi più richiesti consiste nel trovare i 10 generi più richiesti dagli utenti nell'ultimo mese nella piattaforma *FilmSphere*. Per scegliere quali sono i generi più richiesti abbiamo deciso di adottare una metrica basata su dei criteri elencati e spiegati all'interno della successiva sezione Metrica.

Porzione interessata di diagramma E/R



Input	Output	Frequenza
/	GeneriMigliori	1 volta al mese

Metrica

La metrica adottata è basata sulla Popolarità_{genere} cioè su un valore, compreso tra 1 e 10. Sulla base di questo valore, calcolato per ciascun film avente almeno 100.000 visualizzazioni nell'ultimo mese, si basa la custom analytic. Il valore Popolarità_{genere} viene calcolato nel seguente modo:

$$\text{Popolarità}_{\text{genere}} = \chi_{\text{rating}} + \psi_{n^{\circ}\text{visual}}.$$

Il valore di Popolarità_{genere} è compreso tra 1 e 10 (estremi inclusi).

Glossario

- $\chi_{\text{rating}} \triangleq$ coefficiente legato al rating
- $\psi_{n^{\circ}\text{visual}} \triangleq$ coefficiente relativo al n° di visualizzazioni dell'ultimo mese

Calcolo dei singoli fattori

$$1. \chi_{rating} = \beta_1 \frac{\sum_i^n rating_i}{n}$$

Il seguente fattore *rating consiste* nell'effettuare la media dei rating dei film appartenenti al genere preso in considerazione e moltiplicarlo per un *peso* di valore $\beta_1 = 1.50$, dove n = numero di film per quel genere

- Il valore di $\chi_{rating} \in [0.50, 2.50]$

$$2. \psi_{n^{\circ}visual.} = \beta_2 \frac{\sum_i^n coeff_{num.visual.,i}}{n}$$

Il seguente fattore consiste nell'assegnare un punteggio al singolo film in base al n° di visualizzazioni registrate dal film, nell'ultimo giorno del mese.

- Il valore di $\psi_{n^{\circ}visual.} \in [1.50, 7.50]$

Questo valore viene poi moltiplicato per un *peso* di valore $\beta_2 = 0.50$ perché si è deciso di dare una maggiore rilevanza al rating di un film piuttosto che al n° di visualizzazioni: il rating fornisce una valutazione qualitativa di un film, permette di valutare l'apprezzamento a lungo termine, ma ha come svantaggio il fatto che non tutti gli utenti rilasciano una recensione, mentre il n° di visualizzazioni permette di avere una valutazione quantitativa che riflette la popolarità di un film, ma è influenzato da fattori, come la durata di visualizzazione, che non necessariamente indicano la qualità del contenuto.

- Il valore di $coeff_{num.visual.}$ viene assegnato secondo la seguente tabella:

N° di visualizzazioni	Valore assegnato a $coeff_{num.visual.}$
100.000-199.999	1.5
200.000-499.999	2.5
500.000-699.999	3.5
700.000-999.999	4
1.000.000+	5

Nota: i vari pesi β sono stati scelti in modo da assegnare un'importanza diversa a ciascun fattore presente in $Popolarità_{genere}$.

Procedura

1. Leggiamo tutte le occorrenze della relazione *Appartenenza* e ottiene la coppia di chiavi *Genere, Titolo*.
2. Per ogni genere presente in *FilmSphere* vengono selezionati tutti i film che hanno registrato un numero di visualizzazioni nell'ultimo mese pari o superiore a 100.000.
3. Ottenuta la chiave dei film interessati, inerenti a quel genere, viene calcolato il termine χ_{rating} .
4. Viene ripetuto lo stesso passaggio per il termine $\psi_{n^{\circ}visual}$.
5. Viene effettuata la somma dei 2 valori appena ottenuti per ogni genere.
6. Si selezionano, sulla base del termine $Popolarità_{genere}$ in ordine decrescente, i 10 generi col valore maggiore.

10. Popolamento del database

La seguente sezione si occupa di popolare il database con dei dati all'interno di ciascuna tabella presente. I dati inseriti permettono di riprodurre in scala i volumi sopra citati, riducendo dunque i volumi effettivi nell'implementazione fisica senza compromettere l'efficienza del database. Ciascuna tabella è stata popolata secondo una scala diversa, a seconda della dimensione dei volumi di partenza nella *Tavola dei volumi*.

Concetto	VolumeTeorico	VolumeDatabase
Film	1.000	100
Genere	27	27
Appartenenza	2.000	200
Artista	3.600	36
Partecipazione	8.000	600
Direzione	1.200	100
Premio	600	60
PremioFilm	600	20
PremioArtista	300	40
Paese	80	80

Concetto	VolumeTeorico	VolumeDatabase
Critico	30	30
ValutazioneCritico	300	60
Lingua	20	20
Doppiaggio	30.000	300
Sottotitolo	60.000	600
Utente	500.000	1.000
ValutazioneUtente	25.000.000	3.000
Fattura	4.000.000	2.500
Abbonamento	5	5
Pagamento	3.500.000	2.500
CartaDiCredito	500.000	1.000
RestrizioneAbbonamento	32	32
Libreria	20.000	1.500
Connessione	15.000.000	5.000

Concetto	VolumeTeorico	VolumeDatabase
Server	160	160
Memorizzazione	120.000	2.000
Visualizzazione	15.000.000	5.000
File	6.000	400
RestrizioneFile	4.000	20
FormatoAudio	8	8
FormatoVideo	10	10
EstensioneAudio	6.000	400
EstensioneVideo	6.000	400