



UNIVERSITÀ DI PISA

Corso di Laurea in Ingegneria Informatica

## **Progetto di Basi di Dati**

*Guidetti Samuele, Marianelli Alessio*

# Indice

1. Introduzione , p. 4
  - 1.1 Abbonamenti , p. 4
2. Progettazione Concettuale , p. 5
  - 2.1 Dizionario delle Entità p. 5
    - 2.1.1 Entità aggiunte nella ristrutturazione , p. 8
    - 2.1.2 Commenti Aggiuntivi , p. 8
  - 2.2 Dizionario delle Relazioni , p. 9
    - 2.2.1 Relazioni aggiunte nella ristrutturazione , p. 13
3. Ristrutturazione , p. 13
  - 3.1 Eliminazione delle Generalizzazioni , p.13
  - 3.2 Attributi Composti , p.14
4. Tavole dei volumi , p. 14
  - 4.1 Area Connessioni , p. 14
  - 4.2 Area Utenti , p. 15
  - 4.3 Area File , p. 16
  - 4.4 Area Film , p. 17
5. Analisi delle Operazioni sui Dati , p. 17
  - 5.1 Raccomandazione Contenuti , p. 17
  - 5.2 Caching , p. 22
  - 5.3 Inserimento Rating , p. 26
    - 5.3.1 Calcolo Rating , p. 26
    - 5.3.2 Operazione , p. 27
  - 5.4 Richiesta Riproduzione Film , p. 29
  - 5.5 Attivazione Abbonamento , p. 37
  - 5.6 Inserimento Recensione , p. 39
  - 5.7 Recensioni Film , p. 40
  - 5.8 Formati Disponibili , p. 41
6. Progettazione Logica , p. 43
  - 6.1 Ridenominazioni , p. 43
  - 6.2 Traduzione in Schema Logico , p. 43
  - 6.3 Vincoli di Integrità Referenziale , p. 45
  - 6.4 Vincoli Intrarelazionali , p.47

- 6.4.1 Vincoli di Dominio , p.47
  - 6.4.2 Vincoli di Tupla , p.47
- 6.5 Vincoli Interrelazionali , p. 47
- 6.6 Dipendenze Funzionali , p. 48
- 7. Analytics , p. 51
  - 7.1 Classifiche , p. 51
  - 7.2 Custom Analytics – Espansione Catalogo , p. 54
  - 7.3 Bilanciamento Carico , p. 57
- 8. Implementazione , p. 58
  - 8.1 Trigger , p. 58
    - 8.1.1 Premiazione Precoce Film , p. 58
    - 8.1.2 Intervallo Data Film , p. 58
    - 8.1.3 Max Tre Preferenze Genere , p. 58
    - 8.1.4 Max Tre Preferenze Attore , p. 59
    - 8.1.5 Max Tre Preferenze Regista , p. 59
    - 8.1.6 Coerenza Anno E Premio , p. 59
    - 8.1.7 Controlli di Coerenza Connessione , p. 59
    - 8.1.8 Pagamento Valido , p. 59
    - 8.1.9 Coerenza Recensione , p.60
    - 8.1.10 Server Pieno , p. 60
    - 8.1.11 Obbligo Regia , p. 60
    - 8.1.12 Obbligo Interpretazione , p. 60
    - 8.1.13 Aggiunta Rating , p.60
    - 8.1.14 Accesso In Connessione , p.60
    - 8.1.15 Aggiornamento Carico Attuale Fine Riproduzione , p. 60
    - 8.1.16 Coerenza Riproduzioni , p. 61
  - 8.2 Aggiorna Rating , p. 61
  - 8.3 Popolamento , p. 61

# 1. Introduzione

Il progetto consiste nella realizzazione di una Base di Dati per un servizio di streaming di film chiamato “*FilmSphere*”. L’obiettivo è quello di applicare tutte le metodologie apprese, al fine di realizzare una Base di Dati che rappresenti al meglio la realtà che si sta interpretando.

Nella Sezione introduttiva successiva verranno spiegate le scelte che abbiamo fatto in relazione agli Abbonamenti.

## 1.1 Abbonamenti

Come da specifiche abbiamo realizzato 5 “Tier” di abbonamento ognuno con delle caratteristiche a sé e dei prezzi mensili propri. Di seguito l’elenco:

Tipo	MaxDispositivi	OreMaxMensili	RisoluzioneMax	Download	Prezzo Mensile (€)
Basic	1	30	1080p	No	2.99
Premium	1	60	2k, 1440p	No	5.99
Pro	2	100	4k, 2160p	No	7.99
Deluxe	2	130	4k, 2160p	Si	9.99
Ultimate	3	illimitate	8k, 4320p	Si	11.99

**Ulteriori chiarimenti:** I prezzi presentati sopra sono soggetti a possibili variazioni nel tempo. Pertanto, sono presenti nel Database al solo scopo di mostrare (se necessario) il prezzo mensile standard di quel momento all’Utente. I pagamenti effettuati da ogni Utente, possono essere soggetti a variazioni dovute a fattori come sconti per abbonamenti di 3, 6 o 12 mesi oppure altri sconti dovuti ad offerte erogate da FilmSphere (si pensi ad esempio a degli sconti natalizi). Questo per sottolineare che non c’è una correlazione diretta tra le quantità pagate nelle fatture e i prezzi mensili sopra riportati.

## 2. Progettazione Concettuale

### 2.1 Dizionario delle entità

Di seguito è riportato il dizionario delle entità relativo allo schema ER non ristrutturato. Per comodità e chiarezza abbiamo ritenuto opportuno evidenziare le diverse sezioni dello schema ER con colori diversi:

- Area Utenti
- Area Connessioni
- Area File
- Area Film

Gli attributi sottolineati sono le chiavi primarie delle Entità dove compaiono.

Entità	Attributi	Descrizione
Film	<u>IDFilm</u> , Titolo, Descrizione, Uscita, Rating	Contenuto offerto dal servizio FilmSphere. Con Uscita si intende la data di uscita del Film. Per il Rating, controllare l'apposita Sezione.
Attore	<u>IDAttore</u> , Nome, Cognome, DataNascita, Notorietà	Attore che fa un'interpretazione nel film. Specializzazione di VIP. Notorietà è un valore tra 1 e 100, deciso arbitrariamente dai gestori di FilmSphere.
Genere	<u>Nome</u>	Genere cinematografico (thriller, horror, commedia ecc.)
Regista	<u>IDRegista</u> , Nome, Cognome, DataNascita, Notorietà	Regista del film. Specializzazione di VIP. Notorietà è un valore tra 1 e 100, deciso arbitrariamente dai gestori di FilmSphere.
PremioVip	<u>NomePremioVip</u> , <u>Descrizione</u> , TipoVincitore, Peso	Premio assegnato ad un VIP. Nome e Descrizione formano la chiave di ogni premio (es. Oscar, miglior regista). Peso è un valore tra 1 e 100 deciso arbitrariamente dai gestori di FilmSphere.
VIP	<u>IDVip</u> , Nome, Cognome, DataNascita, Notorietà	Generalizzazione di Attore e Regista.

PremioFilm	<u>NomePremioFilm</u> , <u>Descrizione</u> , Peso	Premio assegnato ad un film.
Utente	<u>IDUtente</u> , E-mail, Nome, Cognome, Password.	Cliente della piattaforma FilmSphere (Generalizzazione di Critico e STD User). Ogni Utente ha un'Email diversa.
Critico	<u>IDcritico</u> , Nome, Cognome, Password, E-mail	Utente al quale è riconosciuto il titolo di Critico da FilmSphere. Specializzazione di Utente.
STD User	<u>IDSTDUser</u> , Nome, Cognome, Password, E-mail	Cliente generico della piattaforma FilmSphere. Specializzazione di Utente.
Riproduzione	<u>Inizio</u> , Fine	Riproduzione di un file da parte di un Utente.
Connessione	<u>IDConnessione</u> , IP, Inizio, Fine, Posizione	Connessione di un utente ad un server tramite un dispositivo registrato. Posizione è un attributo che contiene: Latitudine, Longitudine
Server	<u>IDServer</u> , BandaMax Capacità, Posizione	Server della rete di distribuzione CDN. Posizione è un attributo che contiene: Latitudine, Longitudine. CapacitàMax indica la memoria interna del server destinata all'archiviazione dei File[Gb]. BandaMax è la capacità massima del server nel sostenere Riproduzioni[Mbps].
Dispositivo	<u>IDdispositivo</u> , Tipo, MaxRisoluzione, HDR	Dispositivo dal quale l'utente ha fatto accesso almeno una volta.
Paese	<u>Nome</u>	Paese nel mondo. Ogni paese può essere soggetto a restrizioni sui formati.

File	<u>IDFile</u> , Dimensione, Durata, CostoRiproduzione, Risoluzione	File all'interno dei quali troviamo i Film presenti su FilmSphere. CostoRiproduzione è l'incidenza del File sul carico di un Server quando ne viene richiesta la riproduzione.
Formato	<u>NomeFormato</u> , DataRilascio, Bitrate, RapportoDaspetto, CompressioneVideo, QualitàVideo CompressioneAudio, QualitàAudio	Formato di un file. I formati sono di vario tipo ed ognuno possiede caratteristiche diverse.
Lingua	<u>Nome</u>	Lingue disponibili per Doppiaggio e Sottotitoli
Abbonamento	<u>Tipo</u> , RisoluzioneMax, DownloadFile, OreMaxMensili, MaxDispositivi, PrezzoMensile	Abbonamento sottoscritto da un Utente. Gli abbonamenti hanno diversi livelli e vantaggi sempre migliori a seconda della tariffa pagata.
Fattura	DataEmissione, <u>Codice</u> , Emittente, Importo, Scadenza, DuarataAbbonamento	Una fattura emessa dalla piattaforma FilmSphere nel momento in cui un Utente acquista un Abbonamento. (La presenza di <i>Scadenza</i> verrà rivalutata in fase di progettazione Logica)
Carta di Credito	<u>Numero</u> , MeseScadenza, AnnoScadenza, CVV, NomeIntestatario, CognomeIntestatario	Carta di Credito salvata come metodo di pagamento su FilmSphere da un Utente. Il database salva anche le carte scadute, in modo da poter sempre risalire ai dati di chi ha pagato una data fattura.

### 2.1.1 Entità aggiunte nella ristrutturazione

Entità	Attributi	Descrizione
VincitaVIP	Anno	Reificazione della relazione VincitaVIP

### 2.1.2 Commenti aggiuntivi

In questa sezione si trovano spiegazioni ulteriori per certe entità che abbiamo pensato avessero bisogno di una descrizione più completa.

**Dispositivo:** Mantiene tutti i dispositivi che sono stati utilizzati almeno una volta da un utente FilmSphere. La loro rimozione (per dispositivi inutilizzati da tempo ad esempio) deve avvenire solo dopo un accurato controllo delle connessioni e delle riproduzioni più vecchie in modo che non si presentino casi dove abbiamo connessioni che non hanno corrispondenza con dispositivi a causa di una prematura eliminazione.

**Connessione:** Con connessione si intende l'apertura di FilmSphere tramite App, sito o altri mezzi su cui è possibile ottenere accesso al Servizio. Non necessariamente una Connessione identifica una Riproduzione. Si può ad esempio avviare una connessione per sfogliare semplicemente il catalogo. Per semplicità ipotizziamo che quando un Utente si connette ci viene fornita automaticamente la sua Posizione grazie al suo IP.

**Accesso:** Accesso indica la presenza di un dispositivo nel quale un Utente ha fatto Login a FilmSphere. In Accesso sono presenti quei dispositivi nei quali l'account dell'Utente è presente al momento attuale. A seguito di un Logout il dispositivo viene eliminato da Accesso, pur rimanendo nel Database tra i dispositivi registrati all'interno di Dispositivo. L'entità Accesso è fondamentale per garantire il rispetto del limite massimo di dispositivi imposto dai piani di Abbonamento.

**Critico:** Sono Utenti che hanno ricevuto un riconoscimento in quanto personaggi pubblici oppure in quanto particolarmente attivi nella sezione recensioni di FilmSphere. Le loro recensioni possiedono un peso maggiore.



## 2.2 Dizionario delle relazioni

Di seguito è riportato il dizionario delle relazioni relativo allo schema ER non ristrutturato.

Relazione	Attributi	Descrizione	Entità coinvolte
VincitaVIP	Anno	Relazione che indica la vincita di un premio da parte di un vip. Un vip vince un premio in relazione alla partecipazione ad un film.	VIP (0,N), PremioVIP (1,N) Film (0,N)  Un VIP <u>può</u> vincere un PremioVIP. Un Film <u>può</u> far vincere un PremioVIP ad un VIP. Un PremioVIP <u>è</u> stato vinto da <u>almeno un VIP</u> che partecipava ad un Film
VincitaFilm	Anno	Relazione che indica la vincita di un premio da parte di un film.	Film (0,N) PremioFilm (1,N)
GenereFilm		Appartenenza di un film a più generi cinematografici.	Genere(1,N) Film(1,N)
Interpretazione		Recitazione di un attore in un film	Attore(1,N) Film(1,N)
Regia		Indica la regia di un regista per un film	Regista (1,N) Film(1,1)
Recensione	Voto, Commento, DataRecensione	Valutazione da parte di un Utente di un film. Un Utente ne può fare max una al giorno per Film.	Film(0,N) Utente(0,N)

PreferenzaGenere		Corrispondenza di un utente con le sue scelte di generi preferiti (max 3)	Utente(0,N) Genere(0,N)
PreferenzaAttore		Corrispondenza di un utente con le sue scelte di attori preferiti (max 3)	Utente(0,N) Attore(0,N)
PreferenzaRegista		Corrispondenza di un utente con le sue scelte di registi preferiti (max 3)	Utente(0,N) Regista(0,N)
Dispositivo Connessione		Indica una connessione di un utente tramite un dispositivo dal quale l'utente ha fatto accesso.	Dispositivo(1,N) Connessione(1,1)
Richiesta Riproduzione		Richiesta di riproduzione di un file (Film) da parte di un utente durante la connessione ad un server	Connessione(0,N) Riproduzione(1,1)
Connesso		Identifica una connessione ai servizi di FilmSphere da parte di un Utente tramite un Dispositivo Registrato.	Utente(0,N) Connessione(1,1)
Riproduzione Server		Relazione che identifica quale Server sta erogando il file all'utente che l'ha richiesto.	Riproduzione(1,1) Server(0,N)

POP		Indica la presenza di un file su di un server. Il server è “punto di presenza” del file. Quando appena installato, un server può non contenere File.	File(1,N) Server(0,N)
FileRiproduzione		Corrispondenza tra Riproduzione e il File che viene riprodotto	Riproduzione(1,1) File(0,N)
Produzione		Corrispondenza tra un film ed il suo Paese di produzione	Film(1,1) Paese(0,N)
Disponibilità		Indica la Presenza su FilmSphere di un Film sotto forma di File riproducibile	Film(0,N) File(1,1)
Restrizione		Indica quali Formati sono vietati in un determinato Paese.	Paese(0,N) Formato(0,N)
Doppiaggio		Doppiaggio di file in una lingua.	File(1,N) Lingua(0,N)
Sottotitolaggio		Sottotitolaggio di un file in una lingua. Un file può non avere sottotitoli.	File(0,N) Lingua(0,N)
Codifica		Codifica di un File in un Formato. Se nuovo, un Formato può non codificare nessun file.	Formato(0,N) File(1,1)

Intestazione		Corrispondenza tra una fattura emessa da FilmSphere e l'utente intestatario che deve pagarla.	Fattura(1,1) Utente(1,N)
Accesso		Identifica la presenza dell'account di un utente su di un dispositivo. Il numero di dispositivi a cui un utente ha accesso è determinato dall'abbonamento sottoscritto.	Utente(0,N) Dispositivo(0,1)
Sottoscrizione	DataInizio, DataFine	Indica l'abbonamento vigente per l'utente che l'ha sottoscritto.	Abbonamento (0,N) Utente(0,N)
Pagamento	Data	Relazione che identifica il pagamento di una fattura tramite una carta.	Fattura(0,1) Carta di Credito(0,N)
Possesso		Indica Il possesso (attuale) di una carta da parte di un utente.	Utente(1,1) Carta di Credito(0,N)
Bene		Bene mette in relazione la fattura con il bene acquistato, ovvero l'abbonamento	Abbonamento(0,N) Fattura(1,1)

### 2.2.1 Relazioni aggiunte nella ristrutturazione

Relazione	Attributi	Descrizione	Entità coinvolte
VIP-VincitaVIP		Relazione nata dalla reificazione di VincitaVIP	VIP(0,N) VincitaVIP(1,1)
Vincita VIP-Premio VIP		Relazione nata dalla reificazione di VincitaVIP	VincitaVIP(1,1) Film(0,N)
VincitaVIP-Film		Relazione nata dalla reificazione di VincitaVIP	PremioVIP(1,N) VincitaVIP(1,1)

## 3. Ristrutturazione

### 3.1 Eliminazione delle Generalizzazioni

Nel nostro diagramma ER iniziale erano presenti 2 generalizzazioni:

- 1) STD User e Critico sono specializzazioni di Utente dove un Utente è Critico oppure STD User. STD User e Critico compongono la totalità degli Utenti. La Generalizzazione è Totale ed Esclusiva. Nella ristrutturazione abbiamo optato per un accorpamento delle figlie nel genitore con l'inserimento dell'attributo Critico all'interno di Utente per garantire la distinzione tra STD User e Critico.
- 2) Attore e Regista sono specializzazioni di VIP dove un VIP può essere sia Regista che Attore. Per inciso, un Regista può partecipare al suo stesso Film come Attore. Attore e Regista compongono la totalità dei VIP. La Generalizzazione è Totale ed Sovrapposta. Anche qui la nostra scelta è stata di accorpare le figlie nel genitore. La scelta è stata guidata dal fatto che i loro attributi erano identici, oltre al fatto che la generalizzazione è Sovrapposta.

## 3.2 Attributi composti

Nel nostro schema abbiamo un solo attributo composto ovvero *Posizione* che compare in Server e Connessione. L'attributo rappresenta una coppia di numeri Latitudine e Longitudine che servono per identificare la posizione sul Globo dei Server e dell'Utente durante la Connessione a FilmSphere in modo da facilitare le operazioni relative al CDN. Invece di proseguire con la scomposizione dei due, abbiamo preferito mantenere Posizione ed utilizzare il dato di tipo POINT di MySql che ci permette di mantenere attributi costituiti da coppie di numeri.

## 4. Tavole dei Volumi

### Premessa

Le stime sotto riportate si basano su un arco temporale di un anno. Anche le frequenze riportate all'interno dell'analisi delle operazioni si basano sullo stesso concetto.

### 4.1 Area Connessioni

Concetto	Tipo	Volume	Note
Dispositivo	Entità	2.161.500	Stima basata su una suddivisione degli Utenti in base agli abbonamenti: 65%; 9%; 10%; 11%; 5%. Inoltre, il 10% degli utenti che in un anno ha aggiunto un nuovo dispositivo
Connesso	Relazione	208 Milioni	Cardinalità (1,1) con Connessione
Connessione	Entità	208 Milioni	Stima raggiunta considerando che un Utente si connetta a FilmSphere in media 4 volte a settimana per una media di un milione di utenti attivi in un anno.
Dispositivo Connessione	Relazione	208 Milioni	Cardinalità (1,1) con Connessione
Richiesta Riproduzione	Relazione	156 Milioni	Cardinalità (1,1) con Riproduzione

Riproduzione	Entità	156 Milioni	Stima raggiunta considerando una media di 3 Riproduzioni a settimana per Utente e una media di un milione di utenti attivi nell'anno di attività.
Riproduzione Server	Relazione	156 Milioni	Cardinalità (1,1) con Riproduzione
Server	Entità	35	Stima Plausibile
File Riproduzione	Relazione	156 Milioni	Cardinalità (1,1) con Riproduzione
POP	Relazione	33600	Considerando che un File in media sta su 7 Server.

## 4.2 Area Utenti

Concetto	Tipo	Volume	Note
Utente	Entità	1.5 Milioni	Stima plausibile
PreferenzaGenere	Relazione	2.5 Milioni	Stima plausibile
PreferenzaAttore	Relazione	1.7 Milioni	Stima plausibile
PreferenzaRegista	Relazione	2.1 Milioni	Stima plausibile
Possesso	Relazione	1.5 Milioni	Cardinalità (1,1) con Utente
Carta di Credito	Entità	1.8 Milioni	Considerato che: Più utenti possono condividere la stessa Carta e che Una fetta di utenti ha avuto più di una carta, si è fatto conto che le carte siano poco più numerose degli utenti
Abbonamento	Entità	5	Ci sono 5 tipi di Abbonamento (vedi sez. Abbonamenti).
Sottoscrizione	Relazione	5.4 Milioni	Non tutte le Fatture vengono pagate subito. Stimiamo che in un dato momento il 90% delle fatture presenti siano state Pagate e che quindi il corrispettivo abbonamento sia stato associato all'Utente. Nel restante 10% sono contenute anche le Fatture non pagate in tempo (scadute).

Fattura	Entità	6 Milioni	Le Fatture indicano Promesse di pagamento o Pagamenti di Abbonamenti. Abbiamo stimato che ogni Utente in media compra un abbonamento ogni 3 mesi, considerando un anno di attività di FilmSphere.
Bene	Relazione	6 Milioni	Da cardinalità (1,1) con Fattura
Intestazione	Relazione	6 Milioni	Da cardinalità (1,1) con Fattura
Pagamento	Entità	5.4 Milioni	Un'istanza di Pagamento indica una Fattura pagata e quindi una Sottoscrizione effettuata con successo.
Accesso	Relazione	1,965 Milioni	Stima superiore basata sul fatto che ogni Utente faccia accesso sul numero massimo di dispositivi che gli sono concessi e che ogni dispositivo rimane con l'account anche se l'Utente non ha rinnovato l'abbonamento (vedi Sez. Abbonamenti)

## 4.3 Area File

Concetto	Tipo	Volume	Note
File	Entità	4800	Considerando una media di 4 File per Film.
Sottotitolaggio	Relazione	19200	In media ogni file ha 4 sottotitolaggi
Doppiaggio	Relazione	24000	In media ogni file ha 5 doppiaggi
Lingua	Entità	20	Stima Plausibile.
Codifica	Relazione	4800	Cardinalità (1,1) con File
Formato	Entità	85	Si stima che siano presenti circa 85 Formati per i File di FilmSphere
Restrizione	Relazione	39	Si considera che in media un paese ogni 5 abbia restrizioni per un formato.
Paese	Entità	195	Paesi del Mondo.
Disponibilità	Relazione	4800	Cardinalità (1,1) con File
PaeseConnessione	Relazione	208 Milioni	Cardinalità (1,1) con Connessione



## 4.4 Area Film

Concetto	Tipo	Volume	Note
Film	Entità	1200	Stima plausibile
VincitaFilm	Relazione	1400	Ogni premio è stato vinto una volta l'anno. La stima viene fatta risalire a 30 anni fa, meno qualche film che potrebbe non essere su FilmSphere
PremioFilm	Entità	50	Stima plausibile
Genere	Entità	30	Dato frutto di una ricerca web
GenereFilm	Relazione	3600	In media ogni film ha 3 generi.
Vip	Entità	3200	Stima plausibile
Interpretazione	Relazione	24000	Ogni film in media ha 20 attori.
Regia	Relazione	1200	Ogni film ha un regista.
Premio Vip	Entità	35	Stima plausibile
VincitaVip	Entità	990	Come VincitaFilm
VincitaVip PremioVIP	Relazione	990	Da cardinalità (1,1) vincitaVip
Vip VincitaVip	Relazione	990	Da cardinalità (1,1) vincitaVip
VincitaVip Film	Relazione	990	Da cardinalità (1,1) vincitaVip
Recensione	Relazione	780 mila	Un film in media ha 650 recensioni
Produzione	Relazione	1200	Cardinalità (1,1) film

## 5. Analisi Delle Operazioni Sui Dati

In questa sezione vengono affrontate le Operazioni sui dati. Al fine del calcolo degli accessi, si faccia attenzione alle scritture poiché un accesso in scrittura **costa il doppio** di uno in lettura.

### 5.1 Raccomandazione Contenuti

**FREQUENZA:** 500mila volta al giorno (ogni volta che un Utente si connette).

**DESCRIZIONE:** Questa operazione esamina le visualizzazioni dei film degli ultimi 75 giorni e conta quante volte sono stati visti i film per ogni regista, genere e attore. Se uno di questi (regista, genere o attore) corrisponde alle preferenze dell'utente, riceve un aumento di importanza. Successivamente, identifica i migliori 5 per ogni categoria. Infine, attribuisce a ciascun film un valore basato sulla quantità di parametri corrispondenti presenti. Più questo valore è elevato maggiore è la possibilità che l'Utente gradisca il Film.

**INPUT:** IDUtente

**OUTPUT:** Classifica di film che potrebbero interessare all'utente

Parte schema E-R interessata

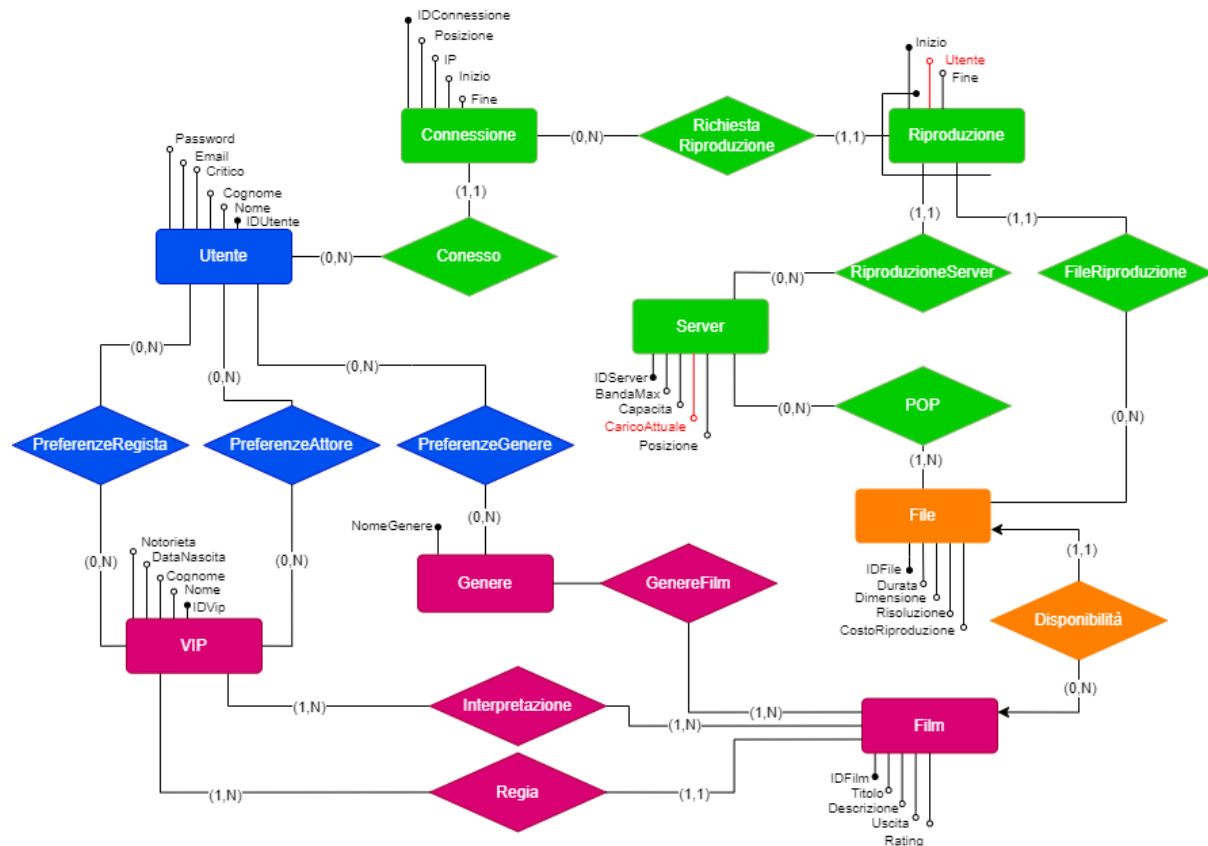


Tavola dei volumi

Concetto	Tipo	Volume	Note
GenreFilm	Relazione	3600	In media ogni film ha 3 generi.
Interpretazione	Relazione	24000	Ogni film in media ha 20 attori.
Regia	Relazione	1200	Ogni film ha un regista.
PreferenzaGenere	Relazione	2.5 Milioni	Stima plausibile
PreferenzaAttore	Relazione	1.7 Milioni	Stima plausibile
PreferenzaRegista	Relazione	2.1 Milioni	Stima plausibile
Connesso	Relazione	208 Milioni	Cardinalità (1,1) con Connessione
Riproduzione	Entità	156 Milioni	Stima raggiunta considerando una media di 3 Riproduzioni a settimana per Utente e una media di un milione di utenti attivi nell'anno di attività.
File Riproduzione	Relazione	156 Milioni	Cardinalità (1,1) con Riproduzione
Disponibilità	Relazione	4800	Cardinalità (1,1) con File

## Tavola degli accessi

Senza *Utente* in Riproduzione

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	Connesso	Relazione	R	208 Milioni	Full scan per ottenere le chiavi di Connessione in modo da ottenere tutte le connessioni effettuate dall'Utente.
2	Riproduzione	Entità	R	156 Milioni	Full scan di Riproduzione per trovare tutte le riproduzioni dell'Utente negli ultimi 75 giorni.
3	FileRiproduzione	Relazione	R	33	Stimando che un Utente in media faccia 3 riproduzioni alla settimana per (poco meno di) 11 settimane. Accedo per ottenere le chiavi di File.
4	Disponibilità	Relazione	R	33	Ad ogni File corrisponde un Film. Accedo per ottenere gli IDFilm.
5	GenereFilm	Relazione	R	3600	Full scan di GenereFilm, #(N,N) per trovare i Generi di ogni Film.
6	Interpretazione	Relazione	R	24000	Full scan di Interpretazione per ottenere gli attori di ogni Film.
7	Regia	Relazione	R	33	Ogni Film avrà un solo regista.

8	PreferenzaGenere	Relazione	R	2.5 Milioni	Full scan per trovare tutte le preferenze dell'Utente.
9	PreferenzaAttore	Relazione	R	1.7 Milioni	Full scan per trovare tutte le preferenze dell'Utente.
10	PreferenzaRegista	Relazione	R	2.1 Milioni	Full scan per trovare tutte le preferenze dell'Utente.
ACCESSI TOTALE GIORNALIERI		185.163.849.500.000			

### Tavola degli accessi

Con *Utente* in Riproduzione

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	Riproduzione	Entità	R	156 Milioni	Full scan di Riproduzione per trovare tutte le riproduzioni dell'Utente negli ultimi 75 giorni.
2	FileRiproduzione	Relazione	R	33	Stimando che un Utente in media faccia 3 riproduzioni alla settimana per (poco meno di) 11 settimane. Accedo per ottenere le chiavi di File.
3	Disponibilità	Relazione	R	33	Ad ogni File corrisponde un Film. Accedo per ottenere gli IDFilm.

4	GenereFilm	Relazione	R	3600	Full scan di GenereFilm, #(N,N) per trovare i Generi di ogni Film.
5	Interpretazione	Relazione	R	24000	Full scan di Interpretazione per ottenere gli attori di ogni Film.
6	Regia	Relazione	R	33	Ogni Film avrà un solo regista.
7	PreferenzaGenere	Relazione	R	2.5 Milioni	Full scan per trovare tutte le preferenze dell'Utente.
8	PreferenzaAttore	Relazione	R	1.7 Milioni	Full scan per trovare tutte le preferenze dell'Utente.
9	PreferenzaRegista	Relazione	R	2.1 Milioni	Full scan per trovare tutte le preferenze dell'Utente.
ACCESSI TOTALE GIORNALIERI		81.163.849.500.000			

## Gestione Ridondanza Raccomandazione Contenuti

Dopo aver valutato le due possibilità, abbiamo ovviamente deciso di procedere con l'inserimento della ridondanza all'interno di Riproduzione. L'ulteriore vantaggio di questa aggiunta è che *Utente* non è un attributo che deve essere aggiornato. L'unica gestione che è richiesta è quella che prevede il suo inserimento (o, raramente, la sua eliminazione) al momento di una richiesta di riproduzione assieme agli altri attributi della tabella Riproduzione.

## 5.2 Caching

**FREQUENZA:** Una volta ogni due settimane per ogni Utente. Circa 71000 volte al giorno.

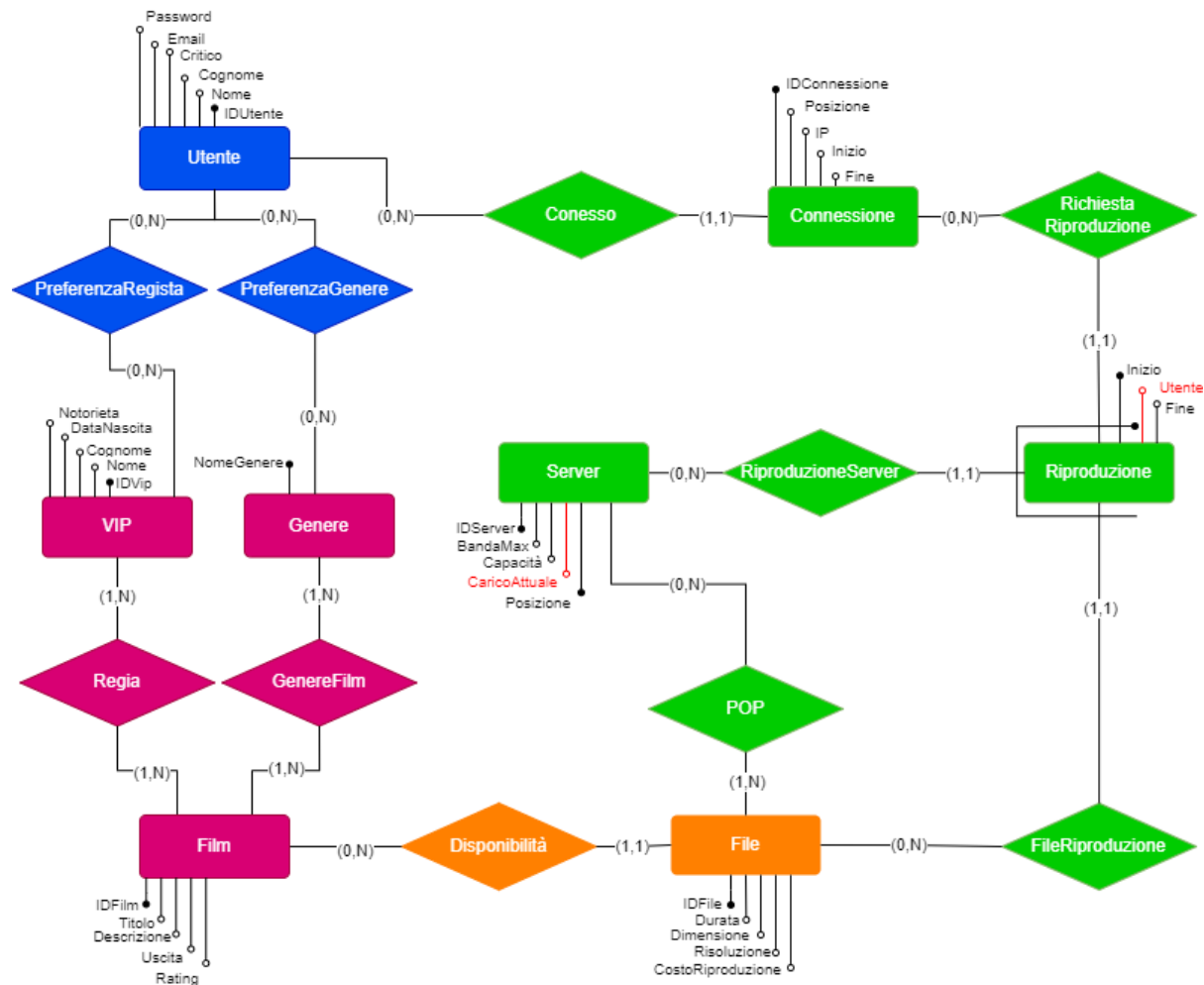
**DESCRIZIONE:** Questa operazione si basa su due idee principali: la prima è che un Utente (in media) ha buone probabilità di guardare di nuovo lo stesso film dopo un certo lasso di tempo; la seconda è che c'è una buona probabilità, che Utenti delle stesse regioni guardino film simili nello stesso periodo.

Inizia considerando le connessioni dell'utente avvenute nelle ultime due settimane e determina la posizione (latitudine, longitudine) a frequenza maggiore. Successivamente, ottiene i film che potrebbero piacere all'utente, basandosi sulle riproduzioni dell'ultimo mese e sulle sue Preferenze e li classifica tramite un punteggio. Sceglie poi i primi 3 film, ordinati in base al valore numerico ottenuto dalla funzione. Infine, carica questi film (o meglio, i file corrispondenti) sul server che si trova più vicino alla posizione più frequente, assicurandosi che non sia pieno (se è pieno, passa al successivo). SE un Utente non si è connesso nelle ultime due settimane, allora l'operazione restituisce un messaggio di errore. SE un Utente non ha effettuato Riproduzioni nell'ultimo mese, non inserisco nulla.

**INPUT:** IDUtente

**OUTPUT:** lista di inserimenti eseguiti in POP

## Parte Schema E-R interessata



## Tavola dei volumi

Concetto	Tipo	Volume	Note
Connesso	Relazione	208 Milioni	Cardinalità (1,1) con Connessione
Connessione	Entità	208 Milioni	Stima raggiunta considerando che un Utente si connetta a FilmSphere in media 4 volte a settimana per una media di un milione di utenti attivi in un anno.
Server	Entità	35	Stima Plausibile
POP	Relazione	33600	Considerando che un File in media sta su 7 Server.
Riproduzione	Entità	156 Milioni	Stima raggiunta considerando una media di 3 Riproduzioni a settimana per Utente e una media di un milione di utenti attivi nell'anno di attività.
FileRiproduzione	Relazione	156 Milioni	Cardinalità (1,1) con Riproduzione

File	Entità	4800	Considerando una media di 4 File per Film.
Disponibilità	Relazione	4800	Cardinalità (1,1) con File
GenereFilm	Relazione	3600	In media ogni film ha 3 generi.
Regia	Relazione	1200	Ogni film ha un regista.
PreferenzaGenere	Relazione	2.5 Milioni	Stima plausibile
PreferenzaRegista	Relazione	2.1 Milioni	Stima plausibile

Tavola degli accessi

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	Connesso	Relazione	R	208 Milioni	Full scan di Connesso per trovare le connessioni dell'Utente.
2	Connessione	Entità	R	208	L'Utente medio in un anno ha effettuato circa 208 connessioni (4 alla settimana). Accediamo per ottenere Inizio e Posizione in modo da prendere le connessioni delle ultime due settimane e calcolare la posizione più frequente.
3	Riproduzione	Entità	R	156 milioni	Full scan di Riproduzione per ottenere i dati sulle riproduzioni dell'Utente nell'ultimo mese.
4	FileRiproduzione	Relazione	R	12	In media un Utente fa 12 riproduzioni alla settimana, perciò, in un mese sono 12. Accedo per ottenere gli IDFile corrispondenti.



5	Disponibilità	Relazione	R	12	Accedo a disponibilità tramite IDFile per ottenere gli IDFilm corrispondenti.
6	Regia	Relazione	R	12	Accessi puntuali per ottenere gli IDVip dei Registri di ogni Film.
7	PreferenzaRegista	Relazione	R	12	Accedo puntualmente per controllare se l'Utente ha messo come sua preferenza uno o più dei registi dei Film che ho trovato.
8	GenereFilm	Relazione	R	3600	Full scan per trovare i Generi dei Film trovati.
9	PreferenzaGenere	Relazione	R	36	Ogni Film ha in media 3 Generi. Accedo puntualmente tramite IDUtente e NomeGenere per controllare se l'Utente ha inserito come preferenze i generi trovati.
10	Server	Entità	R	35	Full scan di Server per leggere la Posizione e la Capacità di ogni server.
11	POP	Relazione	R	33600	Full scan di POP per valutare quali Server possono ospitare altri File e quali invece hanno raggiunto la Capacità massima.

12	File	Entità	R	4800	Full scan di File per ottenere la dimensione dei file già presenti su ogni Server e anche quella dei file da inserire, entrambe necessarie per il calcolo della capacità attuale dei Server.
13	POP	Relazione	W	12	Scrittura in POP dei File in corrispondenza dei Server liberi più vicini all'Utente.
TOTALE ACCESSI GIORNALIERI		25.847.006.921.000			

## 5.3 Inserimento Rating

### 5.3.1 Calcolo Rating

Prima di proseguire definiamo al meglio l'attributo Rating di Film e come viene gestito. Rating è semplicemente una metrica di FilmSphere che attribuisce ad un film un punteggio, sulla base di alcune caratteristiche come: i premi che ha vinto, la notorietà degli attori e dei registi e i voti delle recensioni.

Vediamo l'operazione nel dettaglio.

Definiamo: voto medio degli Utenti ( $R_u$ ), voto medio dei Critici ( $R_c$ ), peso dei premi vinti dal Film ( $P_m$ ), notorietà degli Attori ( $N_a$ ) e notorietà del Regista ( $N_r$ ). Il criterio scelto per la Notorietà degli attori è quello di effettuare la media delle Notorietà, mentre per i premi si prende il premio col Peso più elevato. L'operazione effettuata è la seguente:

$$R_f = \left\lfloor \frac{P_m \cdot 0.1 + (N_a + 1.5 \cdot N_r) \cdot 0.04 + (R_u + 2 \cdot R_c) \cdot 0.3}{3} \right\rfloor$$

Il rating così ottenuto sarà un valore compreso tra 1 e 10.

Rating non è un valore che necessita una precisione elevatissima, per questo abbiamo deciso di aggiornarlo tramite un Event alla mezzanotte di ogni giorno. L'Event in questione verrà discusso nella sezione di *Implementazione*.

La funzione Inserimento Rating non fa altro che chiamare a sua volta Calcolo Rating ed eseguire l'update sulla base di ciò che Calcolo Rating fornisce in output. Quando un Film è nuovo oppure è stato appena inserito, è probabile che questo non abbia recensioni o premi e per questo il Rating avrebbe un valore non valido. Per evitare questo scenario, abbiamo scelto di attribuire ai valori sopracitati un valore iniziale standard che mantenga il Film in un Rating medio, in modo da non scoraggiare gli Utenti e anzi, invogliarli a lasciare recensioni a seconda dei loro gusti.

### 5.3.2 Operazione

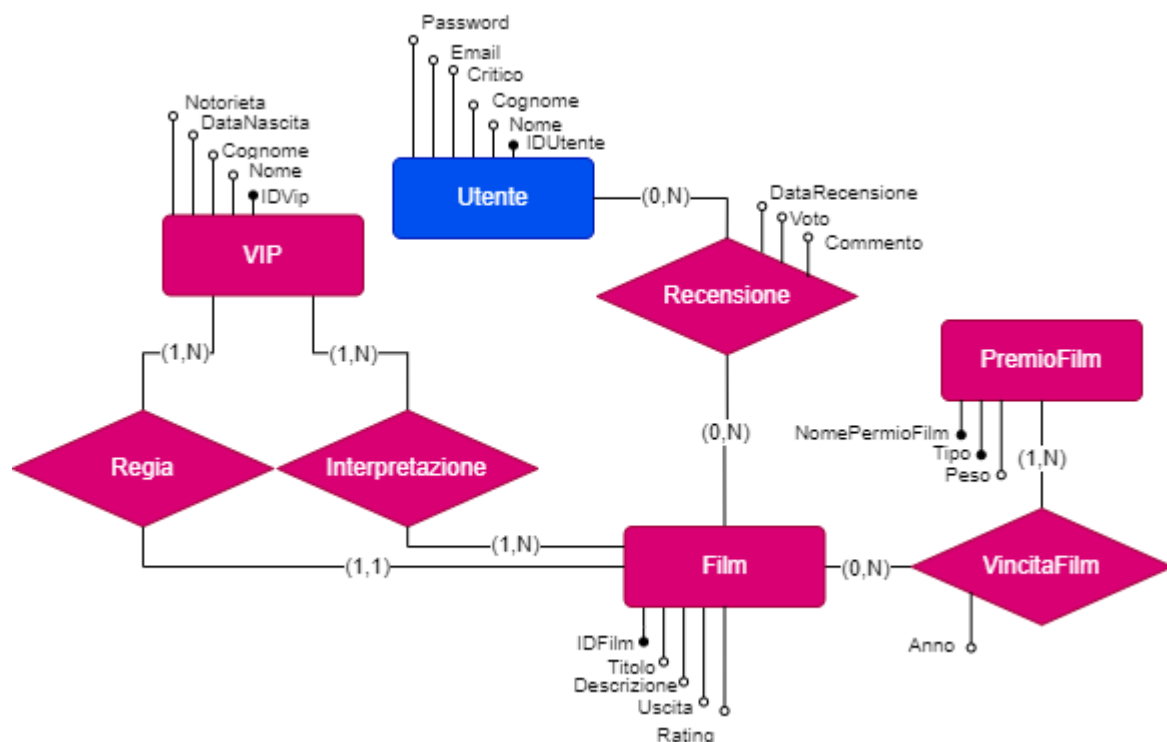
**FREQUENZA:** una volta al giorno per Film quindi 1200

**DESCRIZIONE:** Come descritto sopra, l'operazione si occupa di calcolare ed inserire il Rating di ogni Film.

**INPUT:** IDFilm

**OUTPUT:** Rating

### Parte Schema E-R interessata



## Tavola dei volumi

Concetto	Tipo	Volume	Note
Film	Entità	1200	Stima plausibile
VincitaFilm	Relazione	1400	Ogni premio è stato vinto una volta l'anno. La stima viene fatta risalire a 30 anni fa, meno qualche film che potrebbe non essere su FilmSphere
PremioFilm	Entità	50	Stima plausibile
Vip	Entità	3200	Stima plausibile
Interpretazione	Relazione	24000	Ogni film in media ha 20 attori.
Regia	Relazione	1200	Ogni film ha un regista.
Recensione	Relazione	780 mila	Un film in media ha 650 recensioni
Utente	Entità	1.5 Milioni	Stima plausibile

## Tavola degli accessi

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	VincitaFilm	Relazione	R	1400	Full Scan di VincitaFilm per trovare tutte le Vincite relative al Film.
2	PremioFilm	Entità	R	1	Accesso in lettura per leggere i premi vinti dal Film e il relativo peso. In media ogni Film ha vinto un Premio (1400/1200)
3	Recensione	Relazione	R	780mila	Full Scan per trovare tutte le recensioni relative al Film e gli utenti che le hanno effettuate.
4	Utente	Entità	R	650	Accedo tramite gli IDUtente delle recensioni per distinguere tra critici e STD Users.

5	Interpretazione	Relazione	R	24000	Full scan per trovare gli attori che interpretano un ruolo all'interno del Film.
6	VIP	Entità	R	20	Accedo puntualmente tramite gli IDVip per trovare le notorietà degli attori. Ogni Film ha in media 20 attori.
7	Regia	Relazione	R	1	Accedo puntualmente a Regia per ottenere l'IDVip del Regista.
8	VIP	Entità	R	1	Accedo puntualmente per leggere la Notorietà del Regista.
9	Film	Entità	W	1	Accedo a Film per inserire il nuovo valore di Rating.
ACCESSI TOTALI GIORNALIERI		967.290.000			

## 5.4 Richiesta Riproduzione Film

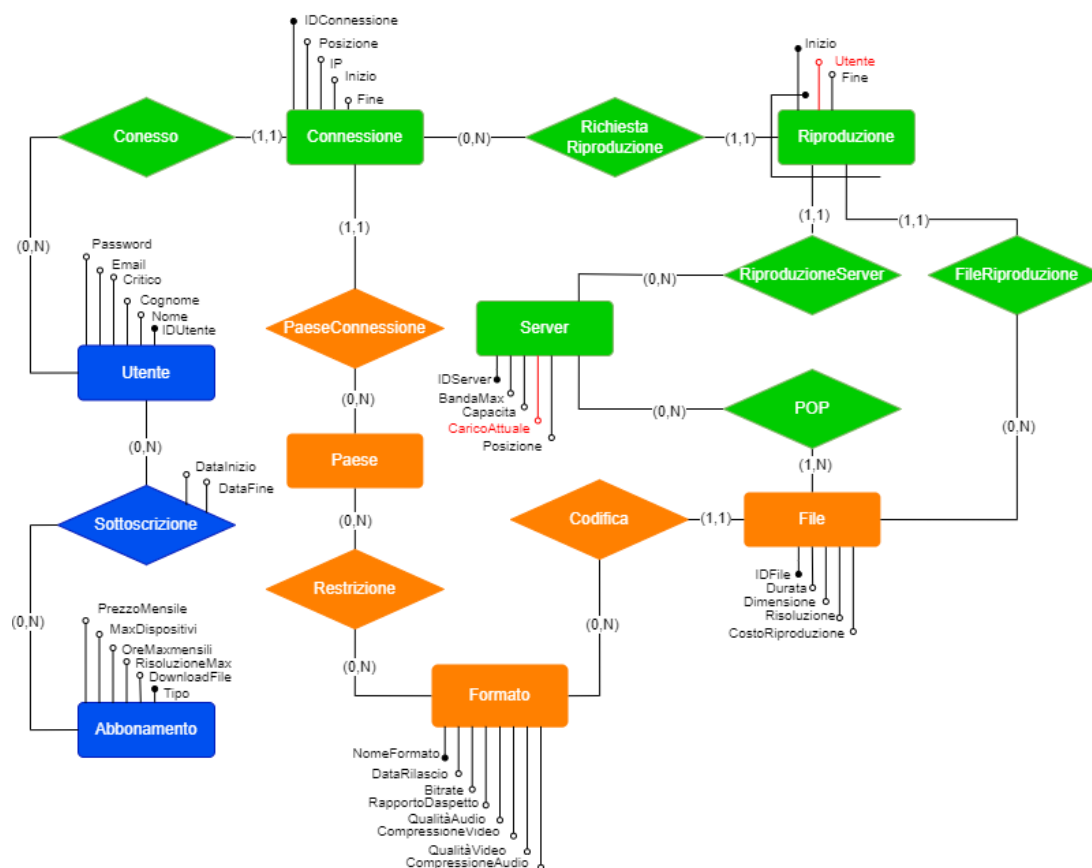
**FREQUENZA:** 400mila volte al giorno

**DESCRIZIONE:** Questa operazione richiede in input una connessione e un file. Per prima cosa, controlla i limiti dettati dall'Abbonamento attivo dell'Utente che ha stabilito la connessione. In particolare, controlla il limite di ore mensili e la risoluzione massima consentita. Se il controllo va a buon fine, procede al controllo delle restrizioni presenti nel paese da dove l'Utente si connette. Successivamente, individua il server più vicino all'utente contenente il file. Se il server non è sovraccarico o offline, avvia una nuova riproduzione del file su quel server, creando un nuovo record nella tabella 'Riproduzione'. In caso contrario, restituisce un messaggio di errore.

**INPUT:** IDConnessione, IDFile

**OUTPUT:** Booleano: TRUE, successo ed inserimento in Riproduzione; FALSE, operazione fallita.

## Parte Schema E-R interessata



## Tavola dei volumi

Concetto	Tipo	Volume	Note
Connessione	Entità	208 Milioni	Stima raggiunta considerando che un Utente si connetta a FilmSphere in media 4 volte a settimana per una media di un milione di utenti attivi in un anno.
Richiesta Riproduzione	Relazione	156 Milioni	Cardinalità (1,1) con Riproduzione
Riproduzione	Entità	156 Milioni	Stima raggiunta considerando una media di 3 Riproduzioni a settimana per Utente e una media di un milione di utenti attivi nell'anno di attività.
Riproduzione Server	Relazione	156 Milioni	Cardinalità (1,1) con Riproduzione

Server	Entità	35	Stima Plausibile
FileRiproduzione	Relazione	156 Milioni	Cardinalità (1,1) con Riproduzione
POP	Relazione	33600	Considerando che un File in media sta su 7 Server.
File	Entità	4800	Considerando una media di 4 File per Film.
Utente	Entità	1.5 Milioni	Stima plausibile
Abbonamento	Entità	5	Ci sono 5 tipi di Abbonamento (vedi sez. Abbonamenti).
Sottoscrizione	Relazione	5.4 Milioni	Non tutte le Fatture vengono pagate subito. Stimiamo che in un dato momento il 90% delle fatture presenti siano state Pagate e che quindi il corrispettivo abbonamento sia stato associato all'Utente. Nel restante 10% sono contenute anche le Fatture non pagate in tempo (scadute).

## Gestione Ridondanza Richiesta Riproduzione Film

L'attributo CaricoAttuale in Server è una ridondanza in quanto sarebbe ottenibile attraverso il calcolo delle Riproduzioni attualmente in corso. Data la situazione, abbiamo analizzato l'operazione in questione valutandola con e senza la ridondanza.

### Senza CaricoAttuale

#### Tavola degli accessi

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	Connesso	Relazione	R	1	Accedo puntualmente a Connesso per leggere l'IDUtente in modo da controllare quante ore l'Utente ha già utilizzato per il mese corrente.
2	Sottoscrizione	Relazione	R	5.4 Milioni	Full Scan di Sottoscrizione per trovare l'abbonamento attivo dell'Utente

3	Abbonamento	Entità	R	1	Accesso puntuale per la lettura delle caratteristiche dell'Abbonamento attivo (OreMaxmensili e RisoluzioneMax)
4	Riproduzione	Entità	R	156 Milioni	Full Scan di Riproduzione per calcolare le ore di Riproduzione totali utilizzate nel mese corrente dall'Utente. Controlliamo anche se l'Utente ha già una riproduzione in corso per questa connessione.
5	File	Entità	R	1	Accedo puntualmente a File per ottenere CostoRiproduzione, Risoluzione e Durata del file da riprodurre.
6	Codifica	Relazione	R	1	Accedo a Codifica tramite IDFile per ottenere il Formato corrispondente.
7	PaeseConnessione	Relazione	R	1	Accedo puntualmente a PaeseConnessione tramite IDConnessione per ottenere il Paese da cui l'Utente si sta connettendo.
8	Restrizione	Relazione	R	1	Accedo puntualmente tramite Paese e Formato, per controllare se il formato del File è ristretto nel paese di connessione dell'Utente.



9	POP	Relazione	R	33600	Full Scan di POP: ho solo la chiave di File. Devo trovare tutti i server su cui si trova il File.
10	Server	Entità	R	7	Accedo a Server per leggere BandaMax e Posizione. In media un File sta su 7 Server
11	RiproduzioneServer	Relazione	R	156 Milioni	Full Scan RiproduzioneServer Ho solo la chiave di Server. Mi servono tutte le riproduzioni (attive) dei Server che mi interessano.
12	Riproduzione	Entità	R	31.200.000	Numero di Accessi dato dal fatto che ogni Server ha circa 4.457.142 riproduzioni e ho bisogno dei dati di 7 Server.
13	FileRiproduzione	Relazione	R	20000	Considerando una media di circa 20000 Riproduzioni attive in un momento generico. Ogni Riproduzione identifica un File, ma più copie dello stesso File possono essere riprodotte contemporaneamente.

14	File	Entità	R	1440	Stimando che le riproduzioni considerate siano occupate dal 30% dei File presenti su FS, abbiamo 1440 File. Quindi accedo per prendere i CostoRiproduzione di ognuno e calcolare così il Carico di ogni Server.
15	Connessione	Entità	R	1	Accedo puntualmente a Connessione tramite IDConnessione per ottenere la Posizione dell'Utente e calcolare il Server più vicino che è POP del File.
16	Riproduzione	Entità	W	1	Una volta trovato il Server più vicino, inserisco una nuova istanza di Riproduzione.
TOTALE ACCESSI GIORNALIERI		139.462.022.400.000			

## Con CaricoAttuale

### Tavola degli accessi

N	Concetto	Tipo concetto	Tipo Accesso	Accessi	Note
1	Connesso	Relazione	R	1	Accedo puntualmente a Connesso per leggere l'IDUtente in modo da controllare quante ore l'Utente ha già utilizzato per il mese corrente.
2	Sottoscrizione	Relazione	R	5.4 Milioni	Full scan di Sottoscrizione per trovare l'abbonamento attivo dell'Utente

3	Abbonamento	Entità	R	1	Accedo puntualmente per la lettura delle caratteristiche dell'Abbonamento attivo (OreMaxmensili e RisoluzioneMax)
4	Riproduzione	Entità	R	156 Milioni	Lettura di tutte le riproduzioni mai fatte da un Utente tramite Full Scan per calcolare le ore di Riproduzione totali utilizzate nel mese corrente. Controlliamo anche se l'Utente ha già una riproduzione in corso per questa connessione
5	File	Entità	R	1	Accesso puntale a File per ottenere CostoRiproduzione,Risoluzione e Durata del file.
6	Codifica	Relazione	R	1	Accedo a Codifica tramite IDFile per ottenere il Formato corrispondente.
7	PaeseConnessione	Relazione	R	1	Accedo puntualmente a PaeseConnessione tramite IDConnessione per ottenere il Paese da cui l'Utente si sta connettendo.
8	Restrizione	Relazione	R	1	Accedo puntualmente tramite Paese e Formato, per controllare se il formato del File è ristretto nel paese di connessione dell'Utente.
9	POP	Relazione	R	33600	Full scan di POP per trovare il Server che è POP del File richiesto.
10	Server	Entità	R	7	Accedo a Server per leggere Posizione e CaricoAttuale. In media un File sta su 7 Server
11	Connessione	Entità	R	1	Accedo puntualmente a Connessione tramite IDConnessione per ottenere la Posizione dell'Utente e calcolare il Server più vicino che contiene il File.
12	Riproduzione	Entità	W	1	Una volta trovato il Server più vicino, inserisco una nuova istanza di Riproduzione.

13	Server	Entità	W	1	Dopo aver aggiunto la nuova riproduzione è necessario aggiornare CaricoAttuale, sommando il CostoRiproduzione al vecchio valore. Dopodiché lo aggiorniamo.
TOTALE ACCESSI GIORNALIERI		64.573.447.200.000			

La nostra scelta è stata ovviamente quella di mantenere la ridondanza CaricoAttuale poiché semplifica di molto il processo e lo snellisce. Di seguito i calcoli per la valutazione del risparmio che questa ridondanza ha portato.

T := Operazione che stiamo considerando.

A := Operazione per aggiornamento ridondanza.

Frequenza giornaliera:  $f^T = g^A = 400k$ ;

Operazioni Elementari (senza ridondanza):  $o^T = 348.655.056$ ;

OE giornaliera (senza ridondanza):  $n^T = f^T * o^T = 139.462.021.200.000$ ;

OE (con ridondanza):  $o^T_{RID} = 161.433.618$

OE giornaliera (con ridondanza):  $n^T_{RID} = f^T * o^T_{RID} = 64.573.447.200.000$ ;

OE risparmiate:  $\Delta = n^T - n^T_{RID} = 74.888.574.000.000$ ;

OE per aggiornamento:  $o^A = 2+2$ ; (il +2 è dato dal fatto che consideriamo che ad ogni inizio di riproduzione corrisponde anche una fine, la quale prevederà una scrittura che aggiorni CaricoAttuale.)

OE giornaliera:  $n^A = g^A * o^A = 1.600.000$ ;

Valutazione finale:

$$n^A + n^T_{RID} < n^T \implies n^A + n^T_{RID} - n^T_{RID} < n^T - n^T_{RID} \implies$$

$$n^A < \Delta \implies 1.600.000 < 74.888.574.000.000$$

## 5.5 Attivazione abbonamento

**FREQUENZA:** Considerando una crescita semi costante della User Base durante l'anno di attività, si stimano circa 2000 nuovi Utenti ogni giorno a cui si aggiunge una percentuale media di rinnovi dello 0.5% giornaliero per un totale di circa **10000** attivazioni di abbonamenti giornalieri.

### **DESCRIZIONE:**

Premessa:

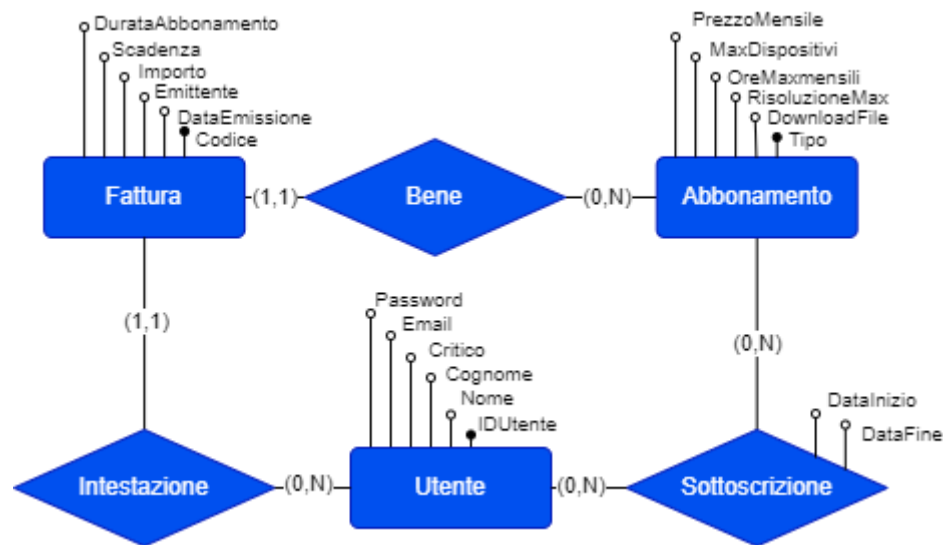
Si considera una Fattura pagata quando essa ha un corrispettivo nella tabella "Pagamento", e quando ciò accade l'Utente ha diritto a ricevere il suo abbonamento.

L'Operazione va in esecuzione dopo che è stato inserito un record all'interno di Pagamento. Dopo che è stato registrato il pagamento di una fattura, memorizzo il tipo di abbonamento e la sua durata. Successivamente, verifico se l'utente ha un abbonamento "attivo". Se non trovo abbonamenti attivi, inserisco un nuovo record nella tabella 'Sottoscrizione'. In questo nuovo record, l'attributo 'DataInizio' viene impostato come il giorno in cui si verifica l'evento, mentre 'DataFine' è calcolato come giorno dell'evento più la durata dell'abbonamento. Se l'utente ha già uno o più abbonamenti attivi, individuo quello che scade per ultimo e utilizzo quella data come punto di partenza per il nuovo abbonamento. La nuova 'DataInizio' sarà quindi la data di scadenza più lontana rispetto alla data attuale più un giorno, e aggiungendo la durata dell'abbonamento pagato, ottengo la 'DataFine' del nuovo abbonamento.

**INPUT:** Codice (Fattura), IDUtente

**OUTPUT:** Inserimento in Sottoscrizione

## Parte dello schema E-R interessata



## Tavola dei volumi

Concetto	Tipo	Volume	Note
Utente	Entità	1.5 Milioni	Stima plausibile
Abbonamento	Entità	5	Ci sono 5 tipi di Abbonamento (vedi sez. Abbonamenti).
Sottoscrizione	Relazione	5.4 Milioni	Non tutte le Fatture vengono pagate subito. Stimiamo che in un dato momento il 90% delle fatture presenti siano state Pagate e che quindi il corrispettivo abbonamento sia stato associato all'Utente. Nel restante 10% sono contenute anche le Fatture non pagate in tempo (scadute).
Fattura	Entità	6 Milioni	Le Fatture indicano Promesse di pagamento o Pagamenti di Abbonamenti. Abbiamo stimato che ogni Utente in media compra un abbonamento ogni 3 mesi, considerando un anno di attività di FilmSphere.
Bene	Relazione	6 Milioni	Da cardinalità (1,1) con Fattura
Intestazione	Relazione	6 Milioni	Da cardinalità (1,1) con Fattura

## Tavola degli accessi

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	Sottoscrizione	Relazione	R	5.4 Milioni	Full scan per ottenere i dati su tutti gli abbonamenti attivi dell'Utente.
2	Fattura	Entità	R	1	Accesso puntuale per ottenere Durata dell'abbonamento.
3	Bene	Relazione	R	1	Una Fattura identifica un Tipo di Abbonamento. Accedo per leggere il Tipo da inserire nel nuovo record.
4	Sottoscrizione	Relazione	W	1	Inserisco il nuovo record in Sottoscrizione, con le modalità corrette.
TOTALE ACCESSI GIORNALIERI		54.000.040.000			

## 5.6 Inserimento Recensione

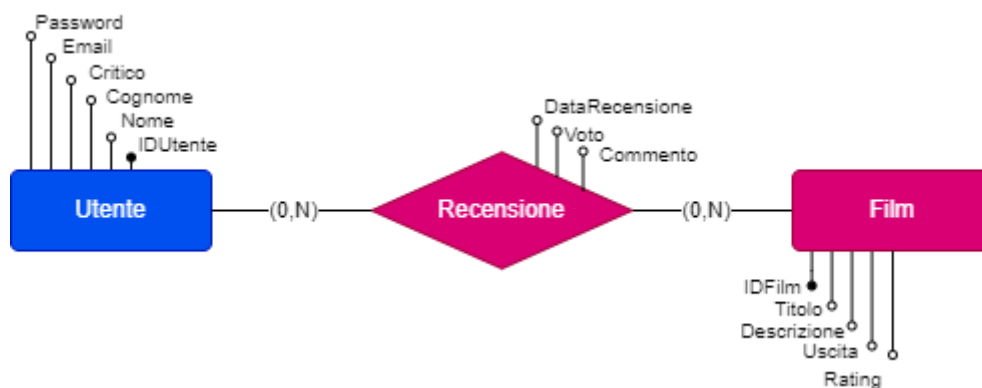
**FREQUENZA:** Si stima che in media gli Utenti lascino 1800 recensioni al giorno (stima al ribasso; media esatta: 2136).

**DESCRIZIONE:** L'operazione aggiunge un nuovo record nella tabella 'Recensioni', inserendo tutti i dati forniti come input. Un utente può inserire al massimo una recensione al giorno per un dato Film.

**INPUT:** IDUtente, IDFilm, Voto, Commento

**OUTPUT:** Inserimento in Recensione

## Parete schema E-R interessata



## Tavola dei volumi

Concetto	Tipo	Volume	Note
Film	Entità	1200	Stima plausibile
Recensione	Relazione	780 mila	Un film in media ha 650 recensioni
Utente	Entità	1.5 Milioni	Stima plausibile

## Tavola degli accessi

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	Recensione	Relazione	W	1	Inserisco una nuova recensione con i valori in input
TOTALE ACCESSI GIORNALIERI		3600			

## 5.7 Recensioni Film

**Frequenza:** Considerando che in media 500 mila persone si connettono a FilmShere e 1/6 di esse guardano le recensioni di un film, questo ci porta ad una frequenza giornaliera approssimativa di 84 mila.

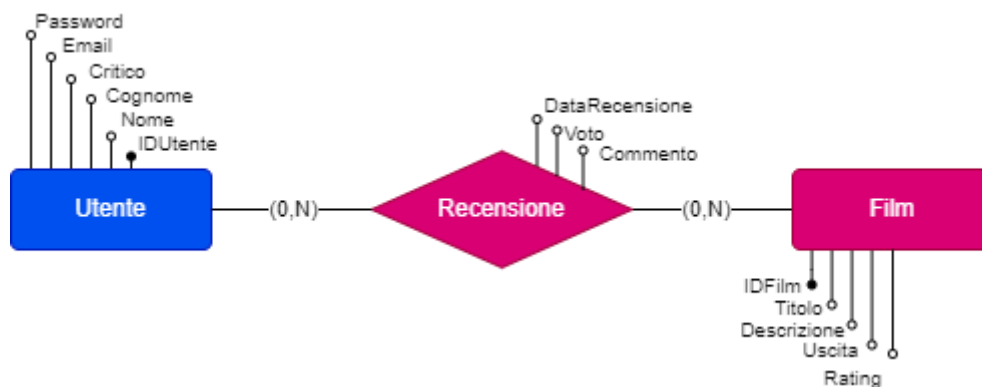
**Descrizione:** Prende un IDFilm e trova tutte le recensioni

**Input:** IDFilm

**Output:** Tutte le recensioni di quel film.



## Parete schema E-R interessata



## Tavola dei volumi

Concetto	Tipo	Volume	Note
Film	Entità	1200	Stima plausibile
Recensione	Relazione	780 mila	Un film in media ha 650 recensioni
Utente	Entità	1.5 Milioni	Stima plausibile

## Tavola degli accessi

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	Recensione	Relazione	R	780 mila	Accedo in lettura per leggere tutte le recensioni e trovo quelle dove ho l' <b>IDFilm</b> in input
TOTALE ACCESSI GIORNALIERI		65.520.000.000			

## 5.8 Formati Disponibili

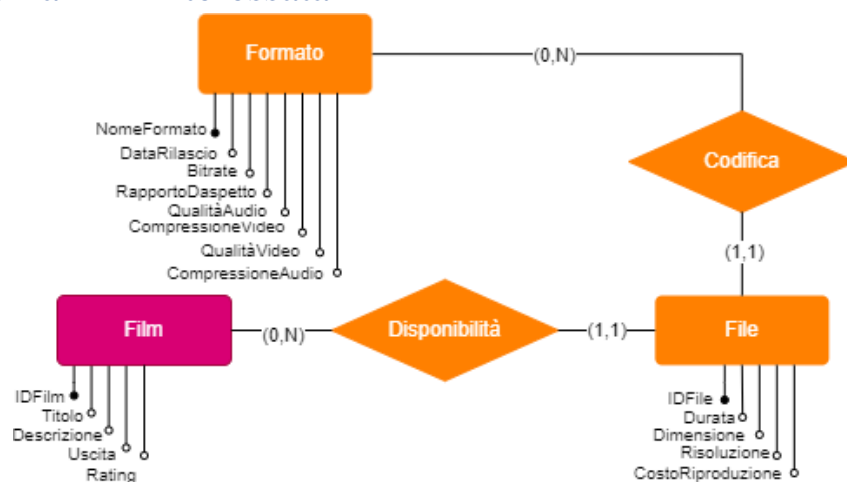
**Frequenza:** Ogni volta che un utente richiede di vedere un film. In media 400mila volte al giorno

**Descrizione:** Dato un film restituisce tutti i formati in cui il film è disponibile su FilmSphere

**Input:** IDFilm

**Output:** Tutti i NomeFormato disponibili per quel film

## Parte schema E-R interessata



## Tavola dei volumi

Concetto	Tipo	Volume	Note
Film	Entità	1200	Stima plausibile
File	Entità	4800	Considerando una media di 4 File per Film.
Disponibilità	Relazione	4800	Cardinalità (1,1) con File

## Tavola degli accessi

N	Concetto	Tipo concetto	Tipo accesso	Accessi	Note
1	Disponibilità	Relazione	R	4800	Full Scan perché un Film identifica potenzialmente più File.
2	Codifica	Relazione	R	4	In media un Film ha 4 File. Avendo #(1,1) con Codifica, vi accedo 4 volte.
3	Formato	Entità	R	4	Accedo tramite IDFile per trovare i NomeFormato che mi servono.
TOTALE ACCESSI GIORNALIERI		1.923.200.000			

## 6. Progettazione Logica

### 6.1 Ridenominazioni

Nel passaggio allo schema logico sono stati effettuati dei cambiamenti per quanto riguarda i nomi di alcuni attributi. Di seguito l'elenco dei cambiamenti:

Genere: Nome → NomeGenere

PremioVip: Descrizione → DescrizionePremioVip

PremioFilm: Descrizione → DescrizionePremioFilm

Riproduzione: Inizio → InizioRiproduzione, Fine -> FineRiproduzione

Connessione: Inizio → InizioConnessione, Fine -> FineConnessione

Paese: Nome → NomePaese

Lingua: Nome → NomeLingua

Pagamento: Data → DataPagamento

### 6.2 Traduzione in Schema Logico

Di seguito la traduzione dello schema concettuale in schema logico:

**Utente** (IDUtente, Nome, Cognome, Critico, Email, Password, CartaIntestata)

**Sottoscrizione** (TipoAbbonamento, Utente, DataInizio, DataFine)

**Abbonamento** (Tipo, DownloadFile, RisoluzioneMax, OreMaxMensili, MaxDispositivi, PrezzoMensile)

**Fattura** (Codice, DataEmissione, Emittente, Importo, DurataAbbonamento, Intestatario, Abbonamento)

**CartaDiCredito** (Numero, MeseScadenza, AnnoScadenza, CVV, NomeIntestatario, CognomeIntestatario)

**Pagamento** (NumeroCarta, CodiceFattura, DataPagamento)

**PreferenzaGenere** (Utente, Genere)

**PreferenzaAttore** (Utente, Attore)

**PreferenzaRegista** (Utente, Regista)

**Accesso** (Dispositivo, Utente)

**Genere** (NomeGenere)

**Vip** (IDVip, Nome, Cognome, DataNascita, Notorieta)

**GenereFilm** (Film, Genere)

**Inerpretazione** (Film, VIP)

**VinicitaVip** (PremioVip, CategoriaVinicitaVip, Film, Vip, Anno)

**PremioVip** (NomePremioVip, DescrizionePremioVip, TipoVincitore, Peso)

**VincitaFilm** (PremioFilm, CategoriaVincitaFilm, Film, Anno)

**PremioFilm** (NomePremioFilm, DecrizionePremioFilm, Peso)

**Recensione** (DataRecensione, Utente, Film, Voto, Commento)

**Film** (IDFilm, Titolo, Descrizione, Uscita, NomePaeseProduzione,  
Regista, Rating)

**Connessione** (IDConnessione, IP, Paese, Posizione, InizioConnessione,  
FineConnessione, Utente, DispositivoConnessione)

**Dispositivo** (IDdispsitivo, HDR, MaxRisoluzione, Tipo)

**Server** (IDServer, BandaMax, Capacità, Posizione, CaricoAttuale)

**Riproduzione** (FileRiprodotta, Server, Connessione, InizioRiproduzione,  
FineRiproduzione, Utente)

**Pop** (File, Server)

**File** (IDFile, Durata, Dimensione, Risoluzione, CostoRiproduzione, Formato,  
Film)

**Lingua** (NomeLingua)

**Sottotitolaggio** (File, LinguaSottotitolaggio)

**Doppiaggio** (File, LinguaDoppiaggio)

**Formato** (NomeFormato, DataRilascio, Bitrate, RapportoDaspetto,  
QualitaAudio, CompressioneVideo, QualitàVideo,  
CompressioneAudio)

**Restrizione** (Paese, Formato)

**Paese** (NomePaese)

### 6.3 Vincoli di Integrità Referenziale

Tra l'attributo *Utente* di **Sottoscrizione** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *TipoAbbonamento* di **Sottoscrizione** e l'attributo *Tipo* di **Abbonamento**

Tra l'attributo *Intestatario* di **Fattura** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *Abbonamento* di **Fattura** e l'attributo *Tipo* di **Abbonamento**

Tra l'attributo *NumeroCarta* di **Pagamento** e l'attributo *Numero* di **Carta di credito**

Tra l'attributo *CodiceFattura* di **Pagamento** e l'attributo *Codice* di **Fattura**

Tra l'attributo *CartaIntestata* di **Utente** e l'attributo *Numero* di **Carta di credito**

Tra l'attributo *Utente* di **PreferenzaGenere** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *Genere* di **PreferenzaGenere** e l'attributo *NomeGenere* di **Genere**

Tra l'attributo *Utente* di **PreferenzaAttore** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *Attore* di **PreferenzaAttore** e l'attributo *IDVip* di **Vip**

Tra l'attributo *Utente* di **PreferenzaRegista** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *Regista* di **PreferenzaRegista** e l'attributo *IDVip* di **Vip**

Tra l'attributo *Dispositivo* di **Accesso** e l'attributo *IDDispositivo* di **Dispositivo**

Tra l'attributo *Utente* di **Accesso** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *Genere* di **GenereFilm** l'attributo *NomeGenere* di **Genere**

Tra l'attributo *Film* di **GenereFilm** e l'attributo *IDFilm* di **Film**

Tra l'attributo *VIP* di **Inerpretazione** e l'attributo *IDVip* di **Vip**

Tra l'attributo *Film* di **Interpretazione** e l'attributo *IDFilm* di **Film**

Tra l'attributo *PremioVip* di **VinicitaVip** e l'attributo *NomePremioVip* di **PremioVip**

Tra l'attributo *CategoriaVincitaVip* di **VinicitaVip** e l'attributo *DescrizionePremioVip* di **PremioVip**

Tra l'attributo *Film* di **VinicitaVip** e l'attributo *IDFilm* di **Film**

Tra l'attributo *Vip* di **VincitaVip** e l'attributo *IDVip* di **Vip**

Tra l'attributo *PremioFilm* di **VincitaFilm** e l'attributo *NomePremioFilm* di **PremioFilm**

Tra l'attributo *CategoriaVincitaFilm* di **VincitaFilm** e l'attributo *DescrizionePremiFilm* di **PremioFilm**

Tra l'attributo *Film* di **VincitaFilm** e l'attributo *IDFilm* di **Film**

Tra l'attributo *Utente* di **Recensione** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *Film* di **Recensione** e l'attributo *IDFilm* di **Film**

Tra l'attributo *Regista* di **Film** e l'attributo *IDVip* di **Vip**

Tra l'attributo *NomePaeseProduzione* di **Film** e l'attributo *NomePaese* di **Paese**

Tra l'attributo *Utente* di **Connessione** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *Paese* di **Connessione** e l'attributo *NomePaese* di **Paese**

Tra l'attributo *DispositivoConnessione* di **Connessione** e l'attributo *IDDispositivo* di **Dispositivo**

Tra l'attributo *FileRiprodotta* di **Riproduzione** e l'attributo *IDFile* di **File**

Tra l'attributo *Server* di **Riproduzione** e l'attributo *IDServer* di **Server**

Tra l'attributo *Connessione* di **Riproduzione** e l'attributo *IDConnessione* di **Connessione**

Tra l'attributo *Utente* di **Riproduzione** e l'attributo *IDUtente* di **Utente**

Tra l'attributo *File* di **Pop** e l'attributo *IDFile* di **File**

Tra l'attributo *Server* di **Pop** e l'attributo *IDServer* di **Server**

Tra l'attributo *Film* di **File** e l'attributo *IDFilm* di **Film**

Tra l'attributo *Formato* di **File** e l'attributo *NomeFormato* di **Formato**

Tra l'attributo *File* di **Sottotitolaggio** e l'attributo *IDFile* di **File**

Tra l'attributo *LinguaSottotitolaggio* di **Sottotitolaggio** e l'attributo *NomeLingua* di **Lingua**

Tra l'attributo *File* di **Doppiaggio** e l'attributo *IDFile* di **File**

Tra l'attributo *LinguaDoppiaggio* di **Doppiaggio** e l'attributo *NomeLingua* di **Lingua**

Tra l'attributo *Paese* di **Restrizione** e l'attributo *NomePaese* di **Paese**

Tra l'attributo *Formato* di **Restrizione** e l'attributo *NomeFormato* di **Formato**

## 6.4 Vincoli Intrarelazionali

### 6.4.1 Vincoli di Dominio

- *CVV* di **CartaDiCredito** deve essere un numero a 3 cifre con valore minimo “000” e valore massimo “999”.
- *Notorietà* di **Vip** deve essere un numero intero tra 1 e 100.
- *TipoVincitore* di **PremioVip** è una stringa di caratteri che assume il valore di “Attore” oppure di “Regista”.
- *Peso* di **PremioFilm** e *Peso* di **PremioVip** devono essere un numero intero tra 1 e 100.
- *Voto* di **Recensione** deve essere un numero intero compreso tra 1 e 10.
- *Uscita* di **Film** deve essere compresa tra “1895-03-19” e `CURRENT_DATE`.
- *Tipo* di **Dispositivo** è una stringa di caratteri che assume i valori di: “TV”, “Telefono”, “Tablet”, “Computer”.

### 6.4.2 Vincoli di Tupla

Abbiamo cercato di ridurre al minimo il numero di valori NULL presenti. Per questo gli unici attributi che possono essere NULL sono:

- *Fine* di **Connessione**
- *Fine* di **Riproduzione**
- *OreMaxMensili* di **Abbonamento**

Abbiamo poi individuato i seguenti vincoli:

- *Inizio* di **Connessione** e *Inizio* di **Riproduzione** devono sempre essere minori di **Connessione** e *Fine* di **Riproduzione** (presi a coppie).
- *DataInizio* di **Sottoscrizione** deve sempre essere minore di *DataFine* di **Sottoscrizione**.

## 6.5 Vincoli Interrelazionali

- La data di un **Pagamento** deve essere maggiore o uguale della data di emissione della **Fattura** corrispondente.
- Il **Pagamento** di una **Fattura** non può essere effettuato dopo la scadenza della stessa.

- La data di una **Recensione** deve essere maggiore dell'anno di uscita del Film a cui si riferisce.
- Una carta scaduta non può effettuare pagamenti, perciò il mese di scadenza di **CartaDiCredito** e l'anno di scadenza di **CartaDiCredito** devono essere maggiori o uguali del mese e dell'anno di **Pagamento** dove la carta compare.
- Un film può vincere premi solo dopo la sua uscita. Perciò la data di **VincitaFilm** deve essere maggiore dell'uscita di **Film**.
- Un Vip può vincere premi relativi ad un Film solo dopo l'uscita di quest'ultimo. Perciò, la data di **VincitaVip** deve essere maggiore dell'uscita di **Film**.
- Un **Regista** per vincere un premio alla regia deve aver diretto almeno un **Film** che compare su FilmSphere.
- Un **Attore** per vincere un premio per l'**Interpretazione** deve aver partecipato ad almeno un **Film** in qualità di **Attore**.
- Un **Utente** per effettuare una Connessione deve avere un abbonamento attivo in Sottoscrizione.
- Un **Dispositivo** può essere utilizzato al massimo per una **Connessione** alla volta.
- Un **Utente**, in un dato istante, può avere un numero massimo di Connessioni pari al numero dispositivi consentiti dal suo **Abbonamento** attivo.
- Il **Server** su cui avviene la **Riproduzione** deve essere **POP** del **File** riprodotto.
- Il **Formato** del file della **Riproduzione** deve essere consentito dal **Paese** di **Connessione** dell'utente.
- La **Risoluzione** del file riprodotto in **Riproduzione** deve rientrare nella soglia consentita dall' **Abbonamento** dell'**Utente**.
- Una **Riproduzione** deve iniziare dopo la rispettiva **Connessione**, e terminare prima (o contemporaneamente) della stessa.
- La somma complessiva mensile delle ore di **Riproduzione** di un **Utente** deve essere inferiore al numero di ore disponibili per il suo **Abbonamento** attuale.

## 6.6 Dipendenze Funzionali

### **Utente:**

IDUtente → (Nome, Cognome, Critico, Email, Password, CartaIntestata);

Email → (Nome, Cognome, Critico, IDUtente, Password, CartaIntestata);



Per come abbiamo strutturato FilmSphere, ogni Utente ha una mail univoca e che quindi lo identifica. Abbiamo preferito però scegliere come chiave primaria IDUtente. **Utente** rimane comunque in forma BCNF.

#### **Sottoscrizione:**

(TipoAbbonamento, Utente, DataInizio) → (DataFine);

#### **Abbonamento:**

Tipo → (DownloadFile, RisoluzioneMax, OreMaxMensili, MaxDispositivi, PrezzoMensile);

#### **Fattura**

Codice → (DataEmissione, Emittente, Importo, DurataAbbonamento, Intestatario, Abbonamento, *Scadenza*);

DataEmissione → (Scadenza);

Una volta emessa una Fattura la sua Scadenza viene automaticamente posta a 30 giorni dopo. Questo fa sì che data una DataEmissione si riesce sempre a ricavare la Scadenza. Piuttosto che una decomposizione della relazione abbiamo optato per la più semplice rimozione di Scadenza in quanto appunto facilmente ottenibile da DataEmissione.

#### **Carta di credito**

Numero → (MeseScadenza, AnnoScadenza, CVV, NomeIntestatario, CognomeIntestatario);

#### **Pagamento**

(NumeroCarta, CodiceFattura) → (DataPagamento);

#### **Vip**

IDVip → (Nome, Cognome, DataNascita, Notorieta);

#### **VincitaVip**

(PremioVip, CategoriaVincitaVip, Film, Vip) → (Anno);

#### **PremioVip**

(NomePremioVip, DescrizionePremioVip) → (TipoVincitore, Peso);

### VincitaFilm

(PremioFilm, CategoriaVincitaFilm, Film) → (Anno);

### PremioFilm

(NomePremioFilm, DescrizionePremioFilm) → (Peso);

### Recensione

(DataRecensione, Utente, Film) → (Voto, Commento);

### Film

IDFilm → (Titolo, Descrizione, Anno, NomePaeseProduzione,  
Regista, Rating);

### Connessione

IDConnessione → (Paese, IP, Posizione, InizioConnessione,  
FineConnessione, Utente, DispositivoConnessione);

### Dispositivo

IDdispositivo → (HDR, MaxRisoluzione, Tipo);

### Server

IDServer → (BandaMax, Capacita, Posizione, CaricoAttuale);

### Riproduzione

(FileRiprodotta, Server, Connessione, InizioRiproduzione)  
→ (FineRiproduzione, Utente);

### File

IDFile → (Durata, Dimensione, Risoluzione, CostoRiproduzione,  
Formato, Film);

### Formato

NomeFormato → (DataRilascio, Bitrate, RapportoDassetto,  
QualitàAudio, CompressioneVideo, QualitàVideo,  
CompressioneAudio);

## 7. Analytics

### 7.1 Classifiche

La funzionalità Classifiche è uno strumento che permette di controllare quali Film e quali Formati sono i più apprezzati dagli Utenti, sulla base delle Riproduzioni effettuate.

Classifiche permette inoltre di distinguere le riproduzioni in base all'Abbonamento e al Paese, ma facciamo chiarezza:

Classifiche richiede un Paese e un tipo di Abbonamento. In base a cosa viene inserito, possiamo avere 4 diversi scenari:

#### Paese e Abbonamento NOT NULL

In questo caso le classifiche stilate faranno riferimento al singolo paese inserito e ci interesseremo solo delle riproduzioni che sono state effettuate da Utenti che, al momento della riproduzione, avevano come abbonamento attivo l'abbonamento richiesto.

#### Paese NULL e Abbonamento NOT NULL

In questo caso le classifiche stilate faranno riferimento alle riproduzioni fatte in ogni paese dove è possibile utilizzare FilmSphere ma ci interesseremo solo delle riproduzioni che sono state effettuate da Utenti che, al momento della riproduzione, avevano come abbonamento attivo l'abbonamento richiesto.

#### Paese NOT NULL e Abbonamento NULL

In questo caso le classifiche stilate faranno riferimento al singolo paese inserito e non faremo distinzione sulla base dell'Abbonamento degli Utenti, trovando quindi le statistiche generali di quel Paese.

#### Paese e Abbonamento NULL

In questo caso le classifiche stilate faranno riferimento alle riproduzioni fatte in ogni paese dove è possibile utilizzare FilmSphere senza tenere di conto delle caratteristiche dell'abbonamento. Questo tipo di esecuzione permette di avere delle statistiche “globali”.

```

drop procedure if exists Classifiche;
delimiter ;;
create procedure Classifiche(in _Paese varchar(15), in _Abbonamento
varchar(10))
begin
    drop temporary table if exists Base;
    create temporary table Base(
        Count smallint unsigned auto_increment,
        Formato varchar(24) not null,
        IDFilm smallint unsigned not null,
        TipoAbbonamento varchar(10) not null,
        primary key(Count)
    ) engine = InnoDB default charset = UTF8MB4;

```

La tabella “Base” immagazzina i dati necessari per stilare entrambe le classifiche, ovvero: il formato del file, l’identificativo del film, il tipo di abbonamento che aveva l’utente al tempo della riproduzione e un valore “count” che funziona da chiave in modo da mantenere i duplicati.

```

insert into Base(formato, idFilm, tipoAbbonamento) (
    with AbbonamentoAllaConnessione as (
        select s.TipoAbbonamento, c.idconnessione, c.paese
        from connessione as c inner join sottoscrizione as s on s.utente =
            c.utente
        where c.inizioConnessione between s.DataInizio and s.DataFine
    )
)

```

AbbonamentoAllaConnessione trova quale era l’abbonamento attivo dell’Utente al momento della Connessione e da quale paese si stava connettendo.

```

select f.formato, f.film, a.tipoAbbonamento
from AbbonamentoAllaConnessione as a
    inner join riproduzione as r on r.Connessione = a.idConnessione
    inner join file as f on r.FileRiprodotta = f.IDFile
where a.paese = _Paese or _Paese is null
);

```

Trovo le informazioni sul File che stava venendo riprodotto e in particolare mi interessa al Film corrispondente e al formato del File, i cui risultati vengono inseriti nella tabella “Base”. Essa riporterà quindi le informazioni di quale Film stava guardando l’Utente, in che Formato era il File da cui lo stava guardando e il tipo di abbonamento che l’Utente aveva al momento della riproduzione. Ovviamente il tutto è filtrato in base al paese di connessione dell’Utente, che deve essere lo stesso di quello richiesto in ingresso. Se non è richiesto un Paese in particolare, allora la classifica sarà una classifica globale.

```

with NumeroRipFilm as (
    select distinct b.idFilm, count(*) over (partition by b.idfilm)
    as numeroR
    from Base as b
    where b.tipoAbbonamento = _Abbonamento or _Abbonamento is null
)

```

In NumeroRipFilm raccogliamo per ogni film il numero di volte che è stato riprodotto. Se è stato richiesto un particolare abbonamento allora prendiamo solo le riproduzioni che sono state effettuate da Utenti che, al momento della riproduzione, avevano l'abbonamento corrispondente.

```

select dense_rank() over(order by n.numeroR desc) as RankFilm, f.titolo,
    n.numeroR as NumeroRiproduzioni
from film as f inner join NumeroRipFilm as n on n.idfilm = f.IDFilm;

```

Ordiniamo i Film per numero di riproduzioni.

```

with NumeroRipFormato as (
    select distinct b.formato, count(*) over (partition by b.formato)
    as numero
    from Base as b
    where b.tipoAbbonamento = _Abbonamento or _Abbonamento is null
)

```

In NumeroRipFormato raccogliamo per ogni formato il numero di volte in cui esso compare come Formato del file riprodotto. Se è stato richiesto un particolare abbonamento allora prendiamo solo le riproduzioni che sono state effettuate da Utenti che, al momento della Riproduzione, avevano l'abbonamento corrispondente.

```

select dense_rank() over(order by n.numero desc) as RankFormato,
    n.formato, n.numero as NumeroRiproduzioni
from NumeroRipFormato as n;

drop temporary table if exists Base;
end;;
delimiter ;

```

Ordiniamo i Formati per numero di riproduzioni, e distruggiamo la tabella “Base”.

## 7.2 Custom Analytics - Espansione Catalogo

La funzionalità *Espansione Catalogo*, si occupa di creare combinazioni tra i Generi, i Registri e gli Attori più amati dal pubblico di FilmSphere, basandosi sulla popolarità di ognuno di essi.

Essa lavora coi dati delle Riproduzioni degli ultimi 3 mesi e con le Preferenze selezionate da ogni Utente, dando priorità a queste ultime.

La riteniamo un buono strumento di valutazione per i gestori di FilmSphere che permette di capire la direzione dei gusti degli spettatori, ed agire di conseguenza durante la valutazione dei nuovi contenuti da inserire.

Attraverso la generazione di combinazioni, infatti, può fornire indicazioni circa il tipo di contenuti che il pubblico desidererebbe vedere aggiunti su FS, perché già esistenti nell'industria, o addirittura contenuti non ancora realizzati.

Ovviamente le combinazioni sono una lama a doppio taglio e possono fornire risultati impossibili da realizzare o aggiungere a FS a causa diritti di acquisizione da parte di altri servizi o impossibilità nel realizzare nuovi Film con determinati Registri o Attori.

Di seguito il codice con commento:

```
drop procedure if exists Espansione_Catalogo;
delimiter ;;
create procedure Espansione_Catalogo()
begin
    declare numLimit smallint unsigned default 10;
    with base as (
        select f.film
            from riproduzione as r
            inner join file as f on f.IDFile = r.FileRiprodotta
            where timestampdiff(month,r.fineRiproduzione, current_timestamp())<3
    ),
```

Prima CTE dentro la quale raccolgo tutti i Film riprodotti negli ultimi 3 mesi con ripetizioni.

```

RegistiPreferiti as (
    select distinct v.Nome as nomeRegista, v.Cognome as cognomeRegista,
    count(*) over(partition by f.regista) + (
        select count(p.Regista) * 3
        from preferenzaRegista as p
        where p.regista = f.regista ) as preferibilitàRegista
    from base as b inner join film as f on b.film = f.idfilm inner join vip
    as v on v.IDVip = f.Regista
),

```

Contiamo quante volte un regista compare alla regia nei film riprodotti negli ultimi 3 mesi.

A questo valore sommiamo il numero di volte in cui quel regista compare tra le preferenze degli Utenti, moltiplicandolo però per 3, così da dare maggiore importanza alle preferenze degli Utenti.

```

AttoriPreferiti as (
    select distinct v.nome as nomeAttore, v.cognome as CognomeAttore, count(*)
    over(partition by i.vip) + (
        select count(p.attore) * 2
        from preferenzaattore as p where p.attore=i.vip) as preferibilitàAttore
    from base as b inner join interpretazione as i on i.Film = b.film
    inner join vip as v on i.vip = v.idVip
),

```

Stesso concetto di RegistiPreferiti con la differenza che qui le comparse in preferenza vengono moltiplicate per 2.

```

GeneriPreferiti as (
    select distinct g.genere, count(*) over(partition by g.genere) + (
        select count(p.genere) * 1.5
        from preferenzagenere as p
        where p.genere = g.genere) as preferibilitàGenere

    from base as b inner join genereFilm as g on g.film = b.Film
)

```

Stesso concetto di RegistiPreferiti con la differenza che qui le comparse in preferenza vengono moltiplicate per 1.5.

```

select rank() over(order by rankGenere, rankRegista, rankAttore)
as RankGlobale, generi.genere, registi.nomeRegista, registi.cognomeRegista,
attori.nomeAttore, attori.cognomeAttore
from (

```

```

        select g.genere, dense_rank() over(order by g.preferibilitàGenere desc)
            as RankGenere
        from GeneriPreferiti as g
        limit numLimit
    ) as Generi

    cross join (
        select r.nomeRegista, r.cognomeRegista,
            Dense_rank() over(order by r.preferibilitàRegista desc) as
            RankRegista
        from RegistriPreferiti as r
        order by rankRegista
        limit numLimit
    ) as Registri

    cross join (
        select a.nomeAttore, a.cognomeAttore, dense_rank() over(order by
            a.preferibilitàAttore desc) as rankAttore
        from AttoriPreferiti as a
        order by rankAttore
        limit numLimit
    ) as Attori

    where generi.rankGenere <= registri.rankRegista and
        generi.rankGenere <= attori.rankAttore and
        registri.rankRegista <= attori.rankAttore;
end;;
delimiter ;

```

Realizziamo un cross join tra le tre CTE realizzate precedentemente e ora rankate sulla base della preferibilità di ogni elemento per ogni categoria (Genere, Registri, Attori).

Come condizione sui risultati, (se i rank non sono a parimerito) imponiamo che:

- Il Genere abbia il rank più alto tra i tre;
- Il Regista abbia un rank minore o uguale del Genere;
- L'Attore abbia il rank più basso tra i tre;

Queste condizioni ci garantiscono combinazioni che danno priorità al genere, garantendo che i generi con rank elevati compaiano in più combinazioni. Abbiamo scelto di prioritizzare genere in quanto pensiamo che sia la componente di scelta più rilevante quando si parla di Film.

Infine, i risultati vengono ordinati per "RankGlobale", dando sempre priorità al Genere.



## 7.3 Bilanciamento Carico

La funzione *Bilanciamento Carico* si occupa di fornire suggerimenti per gli spostamenti di file da un Server all'altro. In particolare, dato un Server, controlla le riproduzioni degli ultimi 6 mesi ed effettua un controllo su quali File sono stati i più riprodotti. Dopodiché, trova i server più vicini che non sono POP dei vari File e suggerisce una serie di spostamenti, dando priorità ai File più richiesti.

```
drop procedure if exists Bilanciamento_Carico;
delimiter ;;
create procedure Bilanciamento_Carico(in _IDServer varchar(3))
begin
    with fileRealmenteRiprodotti as (
        select r.FileRiprodotta, count(*) as numeroRiproduzioni,
               Max(r.InizioRiproduzione) as ultimaRichiesta
        from riproduzione as r
        where timestampdiff(month, r.fineRiproduzione, current_timestamp()) <
              6 and r.Server = _IDServer
        group by r.FileRiprodotta
    ),
```

Troviamo i File riprodotti sul server richiesto negli ultimi 6 mesi e contiamo quante volte sono stati riprodotti. Teniamo anche il timestamp dell'ultima riproduzione per ogni file in modo da avere un ulteriore criterio di ordinamento se le riproduzioni sono a parimerito.

```
RankRiproduzioniFile as (
    select p.File, ifnull(f.numeroRiproduzioni, 0) as richieste,
           f.ultimaRichiesta
    from pop as p left outer join fileRealmenteRiprodotti as f on p.file =
           f.fileRiprodotta
    where p.server = _IDServer
    order by richieste desc
)
```

Controlliamo se i File riprodotti sono ancora presenti nel Server e li ordiniamo per numero di Riproduzioni.

```
select r.file, dense_rank() over(order by r.richieste desc,
                                   r.ultimaRichiesta desc) as piùRichiesto, (
    select distinct first_value(s.IDServer) over(order by
    ST_Distance_Sphere (s.posizione, (
        select s1.posizione
        from server as s1
        where s1.idserver = _IDServer))),s.IDServer)
```

```

        from server as s
        where s.IDServer not in
              (select p.server from pop as p where p.file = r.file)
        ) as ServerConsigliato
        from RankRiproduzioniFile as r;
end;;
delimiter ;;

```

Troviamo il server più vicino che non contiene il file ed esso sarà il Server verso il quale suggeriamo lo spostamento. Se il file è stato rimosso dal server sul quale effettuiamo la chiamata, esso comparirà come server consigliato. Infine, ordiniamo i file per richieste di riproduzione e vi accostiamo il server consigliato per lo spostamento.

## 8. Implementazione

### 8.1 Trigger

Di seguito alcuni dei principali Trigger necessari per il corretto funzionamento del Database.

#### 8.1.1 Premiazione Precoce Film

- **Tabella:** vincitaFilm
- **Tipo:** *before insert*
- **Funzione:** Controlla che un film non vinca un premio prima della sua uscita

#### 8.1.2 Intervallo Data Film

- **Tabella:** Film
- **Tipo:** *before insert*
- **Funzione:** Va a controllare che la data di uscita del film sia tra il 19/03/1895 e la data dell'inserimento

#### 8.1.3 Max Tre Preferenze Genere

- **Tabella:** PreferenzaGenere
- **Tipo:** *before insert*
- **Funzione:** Controlla se un utente sta esprimendo più di tre preferenze di genere

#### 8.1.4 Max Tre Preferenze Attore

- **Tabella:** PreferenzaAttore

- **Tipo:** *before insert*
- **Funzione:** Controlla se un utente sta esprimendo più di tre preferenze di attori

#### 8.1.5 Max Tre Preferenze Regista

- **Tabella:** PreferenzaRegista
- **Tipo:** *before insert*
- **Funzione:** Controlla se un utente sta esprimendo più di tre preferenze di registi

#### 8.1.6 Coerenza Anno E Premio

- **Tabella:** VincitaVip
- **Tipo:** *before insert*
- **Funzione:** Controlla se un vip ha vinto un premio in un anno antecedente alla uscita del film. Inoltre, se il premio è per un regista controlla che il regista in questione abbia diretto almeno un film presente su FS. Fa lo stesso controllo se il vincitore è un Attore.

#### 8.1.7 Controlli Di Coerenza Connessione

- **Tabella:** Connessione
- **Tipo:** *before insert*
- **Funzione:** Si assicura che una connessione rispetti certi criteri di coerenza: che una connessione avvenga in un periodo in cui un utente ha un abbonamento attivo; che lo stesso dispositivo non sia in uso da più di una connessione allo stesso tempo; che il numero di connessione di un certo utente in un periodo sia minore uguale al numero massimo di dispositivi che quell'utente può avere in base al suo abbonamento (dato che non posso fare due connessioni in contemporanea con lo stesso dispositivo);

#### 8.1.8 Pagamento Valido

- **Tabella:** Pagamento
- **Tipo:** *before insert*
- **Funzione:** Controllo che una fattura venga pagata dopo che è stata emessa, che lo stia facendo con una carta non scaduta e che ciò avvenga prima della sua scadenza (30 giorni dall'emissione)

### 8.1.9 Coerenza Recensione

- **Tabella:** Recensione
- **Tipo:** *before insert*
- **Funzione:** Controlla se la recensione avviene prima dell'uscita del film

### 8.1.10 Server Pieno

- **Tabella:** pop
- **Tipo:** *before insert*
- **Funzione:** controlla se il server su cui andiamo a inserire può ospitare il file che ci vogliamo inserire

### 8.1.11 Obbligo Regia

- **Tabella:** PreferenzaRegista
- **Tipo:** *before insert*
- **Funzione:** Controllo che il VIP espresso come preferenza sia un regista

### 8.1.12 Obbligo Interpretazione

- **Tabella:** PreferenzaAttore
- **Tipo:** *before insert*
- **Funzione:** Controllo che il VIP espresso come preferenza sia un attore

### 8.1.13 Aggiunta Rating

- **Tabella:** Film
- **Tipo:** *before insert*
- **Funzione:** Mette un valore standard al rating di un film appena inserito

### 8.1.14 Accesso In Connessione

- **Tabella:** Connessione
- **Tipo:** *before insert*
- **Funzione:** Controllo che un utente non si stia connettendo con un dispositivo a lui non associato

### 8.1.15 Aggiornamento Carico Attuale Fine Riproduzione

- **Tabella:** riproduzione
- **Tipo:** *after update*
- **Funzione:** Tolgo da carico attuale il peso della riproduzione di quel film a server che sosteneva la riproduzione

### 8.1.16 Coerenza Riproduzioni

- **Tabella:** riproduzione
- **Tipo:** *before insert*
- **Funzione:** Mi assicuro che le riproduzioni siano in regola: inizino e terminino dentro una connessione, con controllo per connessione non terminata; il file che si vuole riprodurre si trovi sul server a cui ci si è collegati; il paese di connessione non blocchi la visualizzazione del formato del file; e che l'utente con questa visualizzazione non superi il numero massimo di ore concesse dal suo abbonamento

## 8.2 Aggiorna Rating

É un Event che aggiorna l'attributo rating di film

- **Periodicità:** un giorno
- **Inizio:** alla mezzanotte del giorno di creazione
- **Funzione:** ha lo scopo di aggiornare il rating con una frequenza di un giorno. Ciò perché non ci serve un valore aggiornatissimo e ci aspettiamo che il valore non abbia grandi variazioni in lassi di tempo minori di un giorno

## 8.3 Popolamento

Il nostro database è stato popolato tramite due metodi:

- Il primo ha previsto una raccolta dati tramite generazione su ChatGPT o ricerche sul web. I dati raccolti sono stati poi utilizzati in degli script Python per la generazione delle insert into statiche. Questo metodo ha interessato la maggior parte delle tabelle.
- Alcune tabelle hanno richiesto una maggiore attenzione in quanto strettamente legate al Tempo. Per questo la loro generazione è stata effettuata tramite apposite procedure in sql in modo da avere dati corretti.